VTT SYMPOSIUM 171

COMADEM '97

10th International Congress and Exhibition on Condition Monitoring and Diagnostic Engineering Management

Vol. 1

TECHNICAL RESEARCH CENTRE OF FINLAND
ESPOO 1997

VTT SYMPOSIUM 171

UDC 658.58:061.3
Keywords:
condition monitoring, diagnostic engineering, meetings, expositions

COMADEM '97
10th International Congress and Exhibition on Condition Monitoring and Diagnostic Engineering Management
Vol. 1

Espoo, Finland, 9 - 11 June, 1997

Edited by

Erkki Jantunen
Kenneth Holmberg
VTT Manufacturing Technology, Finland

Raj B.K.N. Rao
COMADEM International, UK

Organized by

Technical Research Centre of Finland
Helsinki University of Technology
Finnish Maintenance Society
Scandinavian Center for Maintenance Management
COMADEM International

TECHNICAL RESEARCH CENTRE OF FINLAND
ESPOO 1997

ISBN 951-38-4562-1
ISSN 0357-9387

JULKAISIJA – UTGIVARE – PUBLISHER

Valtion teknillinen tutkimuskeskus (VTT), Vuorimiehentie 5, PL 2000, 02044 VTT
puh. vaihde (09) 4561, faksi (09) 456 4374

Statens tekniska forskningscentral (VTT), Bergsmansvägen 5, PB 2000, 02044 VTT
tel. växel (09) 4561, fax (09) 456 4374

Technical Research Centre of Finland (VTT), Vuorimiehentie 5, P.O.Box 2000, FIN–02044 VTT, Finland
phone internat. + 358 9 4561, fax + 358 9 456 4374

VTT Valmistustekniikka, Käyttötekniikka, Metallimiehenkuja 6, PL 1702, 02044 VTT
puh. vaihde (09) 4561, faksi (09) 460 627

VTT Tillverkningsteknik, Driftsäkerhetsteknik, Metallmansgränden 6, PB 1702, 02044 VTT
tel. växel (09) 4561, fax (09) 460 627

VTT Manufacturing Technology, Operational Reliability,
Metallimiehenkuja 6, P.O.Box 1702, FIN–02044 VTT, Finland
phone internat. + 358 9 4561, fax + 358 9 460 627

Technical editing Kerttu Tirronen

VTT OFFSETPAINO, ESPOO 1997

DISCLAIMER

Portions of this document may be illegible electronic image products. Images are produced from the best available original document.

PREFACE

COMADEM 97 is the tenth in the series of international conferences on condition monitoring and diagnostic engineering management. The large number of research contributions, altogether over one hundred papers, at the conference and their scientific and technical level is an indication of how much the importance and economical influence of condition monitoring and diagnostic engineering management is recognized world wide. The conference proceedings contain papers on a wide variety of subjects, such as machinery and plant monitoring, instrumentation and control, advances in diagnostic engineering, expert systems, applications of neural networks, modern maintenance management, environmental pollution monitoring and industrial case studies. The proceedings are published in two volumes, VTT Symposium 171 and 172.

The local organizers, the Technical Research Centre of Finland, Helsinki University of Technology, the Finnish Maintenance Society and Scandinavian Center for Maintenance Management regard it as a privilege and honour to be able to offer this international, interesting and high-level programme as a contribution to the condition monitoring and diagnostic engineering management society. The organizers would like to thank all the authors for their efforts and co-operation in preparing the excellent papers that have been selected for this conference. The organizers also wish to express their gratitude to the International Advisory Board and National Programme Committee for their expert knowledge, energy and co-operation in the evaluation process of the papers and for setting up the scientific programme.

Financial support from the European Commission, the Academy of Finland and the support from the City of Espoo are gratefully acknowledged.

Erkki Jantunen
Conference Manager

CONTENTS

Expert systems

Electrical motors

ADVANCED SOLUTIONS FOR OPERATIONAL RELIABILITY IMPROVEMENTS

Kenneth Holmberg, Research Professor
VTT Manufacturing Technology
P.O. Box 1704, Espoo, Finland

ABSTRACT

A great number of new technical tools are today developed for improved operational reliability of machines and industrial equipment. Examples of such techniques and tools recently developed at the Technical Research Centre of Finland (VTT) are: metallographic approach for steam-piping lifetime estimation, an expert system AURORA for corrosion prediction and material selection, an automatic image-processing-based on-line wear particle analysis system, microsensors for condition monitoring, a condition monitoring and expert system, CEPDIA, for the diagnosis of centrifugal pumps, a machine tool analysis and diagnostic expert system, non-leakage magnetic fluid seals with extended lifetime and diamond-like surface coatings on components with decreased friction and wear properties. A hyperbook-supported holistic approach to problem solving in maintenance and reliability engineering has been developed to help the user achieve a holistic understanding of the problem and its relationships, to navigate among the several technical tools and methods available, and to find those suitable for his application.

1 INTRODUCTION

The operational reliability of machines and instruments is becoming a crucial factor of competitiveness in applications where safety and availability are important. Automation and integrated production have resulted in larger and more complex systems, which are more vulnerable and susceptible to diverse consequential effects because of breakdowns. Observing the significant cost of downtime and maintenance activities, and the economical benefit to be gained from

higher process availability, industry has increasingly shown great interest in operational reliability and maintenance management techniques.

In some sectors of industry the cost of maintaining plant assets is comparable with, or even higher than that of operating the process, in the range of 30–50% of the operating budget, when process energy and depreciation are ignored [1,2]. In general, the maintenance in industry accounts for 8% of the turnover, 16% of the employment costs and 10% of the total national product in Northern European countries.

There is today a great variety of different techniques and technologies available for the improvement of reliability and maintainability throughout the lifetime of the product. These include technical fields, such as management technologies, material science, design calculations, monitoring techniques and methods for information processing [3,4,5].

This paper presents some general technological trends in the field of reliability, availability and maintenance, a holistic general approach and a number of examples of different advanced technical solutions for the improvement of operational reliability.

2 TECHNOLOGICAL TRENDS

Advanced methods for reliability control were developed early in the fields of nuclear energy, aviation and the space industry for obvious safety reasons. Mainly economic but also environmental aspects have brought a demand for similar approaches in the conventional power industry and the process industry, but recently also in flexible manufacturing systems in the metal industry and in automated production in other industrial areas. Design philosophy, manufacturing strategy, lifetime and requirements on maintenance and safety are very different when comparing a nuclear power plant and a consumer product like, for example, a vacuum cleaner. Thus, the methods for reliability control used in a nuclear power plant cannot directly be used for a vacuum cleaner. The acquired knowledge can be utilized but new methods suitable for consumer products need to be developed.

Today, new technological aids for improved availability and reliability are at the development stage and - to some extent - already on the market. New condition monitoring techniques, such as on-line wear particle analysis, can be used as complementary techniques for more qualified condition information, or in some cases even to replace the vibration monitoring techniques dominant today. Advanced knowledge-based software, such as expert systems, fuzzy logic and neural networks and hybrid systems that combine the strength of various approaches, seem to offer a possibility to perform automatic diagnosis where previously it has been difficult, because of the amount and complexity of data or for other reasons. Today, telecommunications offer a new and interesting concept of remote diagnosis of machinery, such as diesel engines, located far away from qualified maintenance personnel.

In complex production lines, such as flexible production systems (FMS) in the metal industry, good availability can be achieved only by systematic and statistics-based failure analysis methods. The integration of the maintenance software to the process control software becomes a key issue.

Statistical methods for analysis of the reliability of electronic components is well developed, but the failure and lifetime control of mechanical components is not well understood. Software faults are today mainly controlled by modelling approaches, and the control of human error certainly needs more attention. New hybrid method approaches are under development to combine and take advantage of both the traditional bottom-up reliability analysis methods and the top-down safety analysis methods [6].

General trends in the technologies utilized for improved availability and reliability are towards

(a) intelligent monitoring supported by multi- and microsensor techniques and failure modelling,
(b) automatic diagnosis based on on-line monitoring and knowledge-based information systems,
(c) system reliability control, including statistical and system analysis methods and
(d) advanced software integration and user interface solutions.

3 A HOLISTIC SYSTEMATIC APPROACH

Because of the great variety of different techniques based on expert knowledge in several fields of technology involved, there is a need to approach the reliability and maintainability problems from a general holistic point of view starting from the customer's problem and ending with the satisfied user. VTT has developed a systematic approach with the aim improving the synergistic interactions between the different fields of expertise by showing a logical and comprehensive structure, where each expert can find his place and see the connections to experts from other fields who can all work with the same aim of satisfying the customer, as shown in Figure 1.

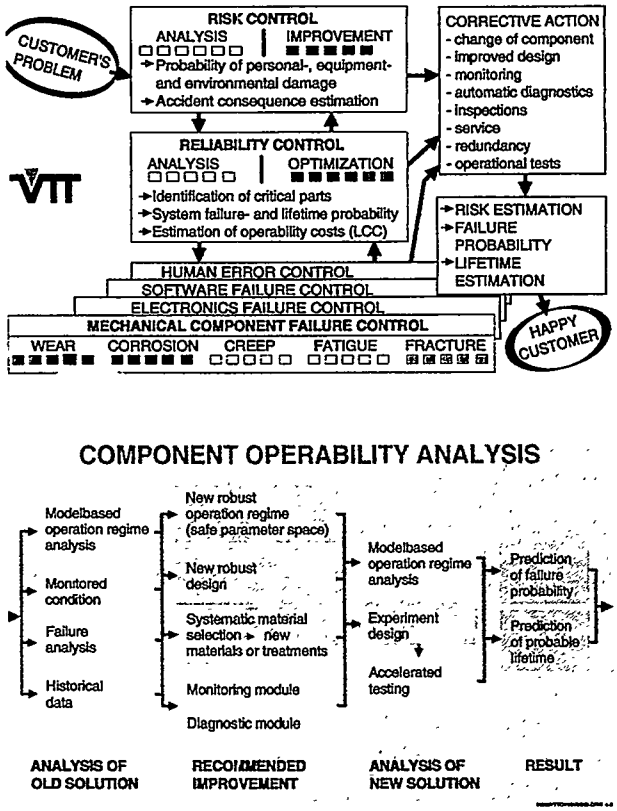


Fig. 1. A holistic systematic approach to the improvement of the operational reliability, safety and availability of products and industrial production systems.

The probability of personnel, equipment and environmental damage is analyzed and the accident consequences estimated by the systematic methods of risk control. Critical parts are identified, the probability of system failure and lifetime calculated, and operability costs estimated by statistically based techniques of reliability control. When the critical parts of the production system that need improvement have been identified, the right techniques and tools for improvement actions are found in the fields of mechanical component failure control, electronics failure control, software failure control or human error control.

When a critical function is identified, such as for example the wear endurance life of a certain component, a component operability analysis is carried out, including analysis of the old solution, a robust lifetime design approach to recommended improvements, analysis of the new solution, and as a result the improvement actions with estimated improved failure probability and probable lifetime.

The recommended measures to be taken can be a change of component, redundancy, improved design, extended monitoring, automatic diagnosis, inspections, operational tests or service instructions. The output of the holistic approach is recommendations for improvements together with estimations on their effects on the risks, the probability of failure and the lifetime.

4 FAILURE PREDICTION AND LIFETIME ESTIMATION

The material degradation by wear, corrosion, creep, fracture and fatigue are often unavoidable mechanisms that cause failure and limit the lifetime of machinery and structures. Even if it is often not possible to completely eliminate these degradation mechanisms, they can be controlled, their effects can be limited and their progression can be measured or estimated for condition assessment and functional lifetime prediction.

Steam pipings in power plants that operate at elevated temperatures degrade by creep, fatigue or combined mechanisms. This degradation results in increasing strain and decreasing residual strain to fracture. The metallographic techniques have been developed to estimate the minimum (safe) residual lifetime for the piping by replica-based

microstructure inspection and assessment of the degree of grain boundary creep cavitation [7,8,9].

The close relation between cavitation, strain and safe residual creep life can explain the success of cavitation measurements as indications of the creep condition of a high-temperature steam piping. It is possible to predict the residual lifetime by using the relation between cavitation and strain, as shown in Figure 2. In such a prediction it is not necessary to assume that cavitation is the most important life-determining factor in the total damage. However, this relation between damage and life is usually based on experience and depends on, e.g. material and component type, as the effect and details of loading histories vary from one plant to another.

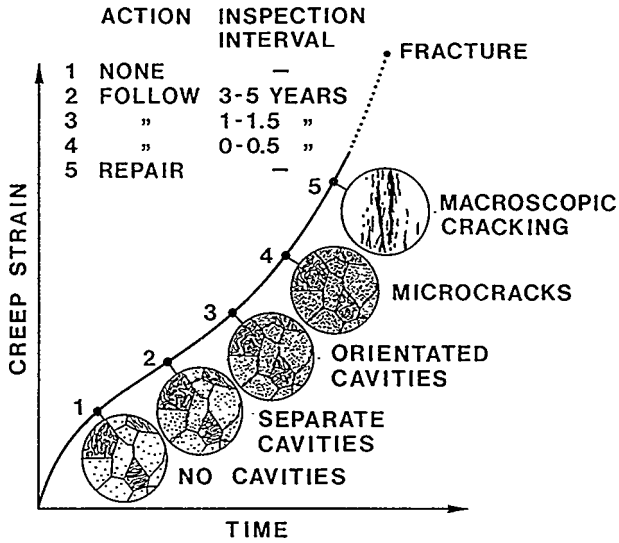


Fig. 2. Assessment rules for creep damage and residual lifetime prediction of a hot steam piping.

Material degradation by corrosion is a serious problem in several industrial environments, e.g. in the process industry. An expert system module, AURORA-STACTOR, for predicting the possibility of corrosion attack of stainless steels in aqueous environments was developed as an example of applying various approaches of using expert systems and semi-empirical models of corrosion to support appropriate material selection from a corrosion point of view for a certain environment, to predict the corrosion risk and for failure analysis [10,11,12].

The AURORA family consists of separate prototype modules dealing with various aspects of corrosion in different alloys. The operation of the systems is based on evaluating the corrosion risk of an alloy environment combination characterized by the user. The detailed relation of the metal surface with the environment is characterized through questions posed to the user. All the main variables influencing the probability of corrosion are considered. The prototype modules include, e.g. the prediction of corrosion risk and failure analysis of stainless steels and copper-based alloys, as well as the risk prediction of bimetallic corrosion.

In the finalized module, AURORA-STACOR, the result of the analysis is given as "probabilities" the propobility of uniform corrosion, pitting/underdeposit corrosion, crevice corrosion and stress corrosion cracking separately and all together. The environmental conditions specified by the user and the advice on how to reduce the possibility of the particular form of corrosion are also shown by the system (Figure 3).

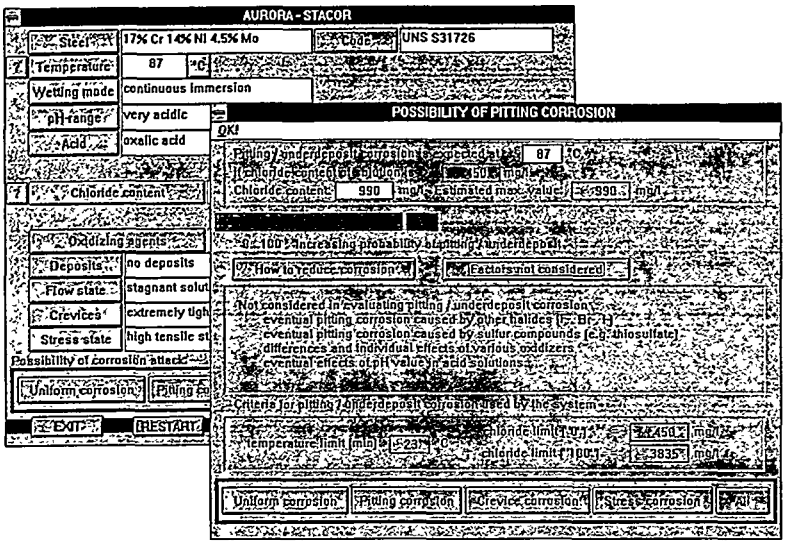


Fig. 3. Display examples from the corrosion expert system AURORA-STACOR showing the specified environmental conditions, the possibility of pitting corrosion and advice on how to reduce corrosion.

5 CONDITION MONITORING

For machines with a critical position in the production system, real-time condition monitoring is of great importance. Oil analysis methods provide complementary information on the condition of machines compared to the widely used vibration monitoring methods. In some applications the oil analysis is more sensitive to indicating a starting failure and can be a substitute to vibration monitoring [13]. The drawback with oil analysis methods has been the lack of reliable and convenient on-line measurement equipment that would provide real-time condition data.

A new automatic on-line wear particle monitoring and diagnosis system has been developed, Figure 4. It consists of a novel optic wear particle sensor, pattern recognition for particle identification, fuzzy- logic-based reasoning for wear mechanism determination and an expert system for the machine condition diagnosis [14,15,16].

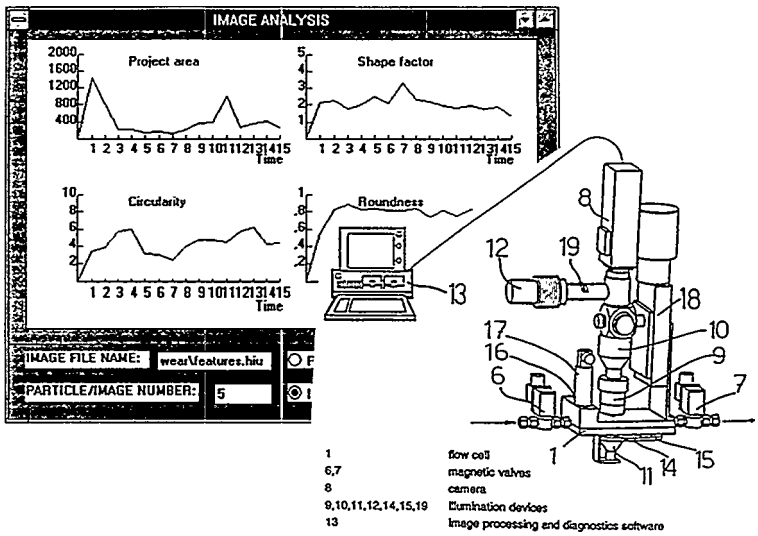


Fig. 4. Schematic illustration of the on-line wear particle monitoring instrument with the image analysis interface of the integrated expert system.

The particle analysis method consists of four main sub-analyses. The wear origin is determined according to the particle colour. The particles are classified by image-processing software into five categories based on their shapes. The surface texture of the particle is

analyzed according to its smoothness, the existence of grooves, pits or holes, and the twist of the particle. The number and size of the particles are connected to the wear mechanism deduction process. The wear mechanisms are divided into the following five basic types: adhesive, abrasive, cutting, fatigue and corrosive wear.

The conclusions of the analysis are based on wear severity determination, which is carried out separately for each pre-classified particle group. Severity determination is a result of the logical classification operations with material, wear degree, wear mechanism and wear factor parameters. The conclusion includes information on the current wear stage and proposals concerning desired actions.

MicroElectroMechanical Systems (MEMS) is a new and rapidly developing technology that offers interesting perspectives for the condition monitoring and diagnostics of machinery. The MEMS are a spin off from the technology developed to fabricate integrated circuits on silicon chips. While integrated circuits are designed to exploit the electrical properties, MEMS take advantage of the electrical, optical, thermal and mechanical properties of silicon [17].

Although MEMS can be made small mechanical devices or actuators, they have been most commonly used for sensors. Already today there are MEMS sensors available for the measurement of e.g. temperature, pressure, force, acceleration, flow and rotational frequency.

For condition monitoring purposes the possibility of producing cheap and very tiny, less than 1 mm^2 small integrated sensors, e.g. for measuring accelerations and temperature is most interesting. It would give the possibility to largely integrate a multitude of sensing elements into almost any part of a component or machinery. This would of course have a remarkable effect on the whole concept and business of condition monitoring.

VTT has designed and fabricated an integrated MEMS sensor, of $12 \times 22 \times 5$ mm^3 size for the spectrometric measurements of the composition of gases, shown in Figure 5 [18], and is investigating the possibilities of developing MEMS sensors for wider use in condition monitoring.

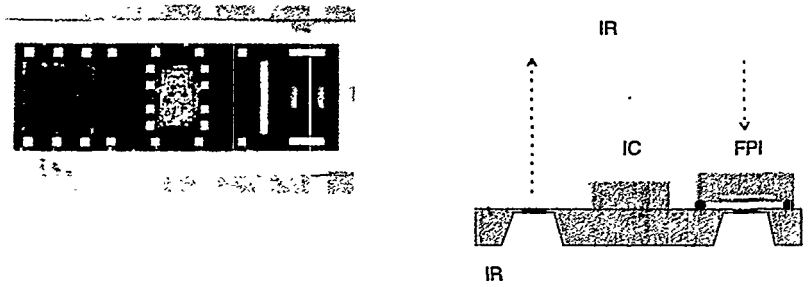


Fig. 5. The spectrometric MEMS sensor designed and fabricated by VTT Electronics.

6 DIAGNOSTICS

Failures and inefficient functioning due to uncontrolled machinery condition is a problem today in the process industry and in energy production. In Finland 10% of the total electrical energy used and 20% of the industrial energy consumption goes to pumping. However, when 63 pumps in several different factories were analyzed it turned out that their average efficiency in pumping was less than 40% and more than 10% of the pumps were pumping at an efficiency less than 10% [19]. From this it is obvious that considerable economic savings can be achieved with improved pumping condition monitoring and diagnostics.

Based on laboratory testing an expert system, CEPDIA, was developed for the diagnosis of centrifugal pump condition and for the definition of pump energy-savings potential [20]. Separate diagnose are made for the different components of the pump: the pump, the mechanical seal, the stuffing box, the bearings and the electrical motor. The diagnosis is based on condition monitoring measurements from sensors and information on maintenance actions carried out on the pump and its components (Figure 6). Although the diagnosis system is capable of handling monitored information from a great number of different transducers (up to 20), it can also base the diagnosis on a very limited amount of information.

A number of databases are included in the system for handling and saving the measured data, technical information and maintenance actions carried out on the pumps. The diagnosis can also be based on

vibration signature analysis, which is quite effective in defining which fault is the actual cause of the malfunction of the pump or the components. CEPDIA can be used for the calculation of the efficiency of the electrical motor and the pump. This information is used for economic evaluation of the pump configuration, whether the configuration is acceptable or if it is economically justified to make some changes to the system in order to improve the economics of the pumping.

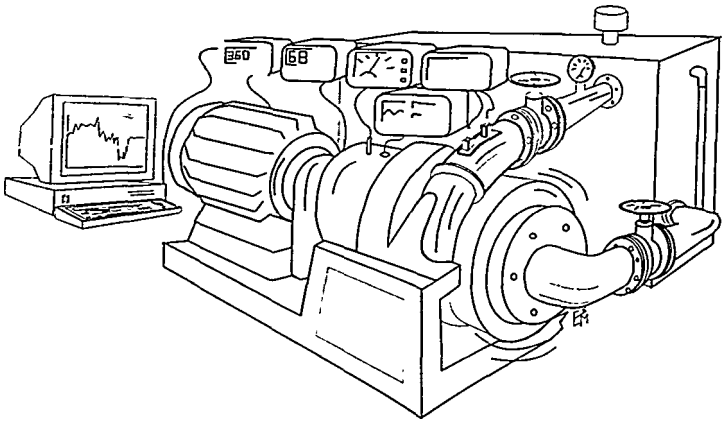


Fig. 6. The CEPDIA expert system performs diagnosis of centrifugal pums based on monitored data, technical pump information and history data from maintenance actions.

An increasing number of Flexible Manufacturing Systems (FMS) have been installed in Europe during the past few years. A general experience is that the availability of the installed FMS is not as high as was originally expected, and especially the unmanned use in three shifts has not been successful. To improve the availability of FMS the following systems were developed in a Finnish-UK EUREKA project: automated data collection for utilization and failure data, automated analysis of historical problems, machine tool condition monitoring, diagnostic expert systems, multimedia information systems, quality support including methodology, statistical process control, and ball bar diagnostics, and maintenance planning. All these systems were tested at four pilot sites covering large and medium-sized companies, and from users to machine manufacturers [21].

The VTT work focused on the improvement of the machine tool utilization rate by an advanced condition monitoring system using modern sensor, signal-processing and diagnostic techniques [22,23]. A comprehensive cutting test procedure was carried out with different tools and measurement arrangements. The recorded signal information was processed in several ways, both in time and frequency domain. The effectiveness of the best sensors and analysis methods was verified with a developed program module in the prediction of the remaining lifetime of a tool in use. The results from the statistical analysis show that vibration, sound and acoustic emission measurements are more reliable for tool wear measurement than the most common methods, such as power consumption, current, and force measurements used in commercially available systems. Even better results are obtained based on the use of FFT methods, especially when at least two signals are available for frequency analysis. The relations between the analyzed signals and tool wear form a basis for the diagnosis rules that are used in a diagnostic expert system, schematically shown in Figure 7.

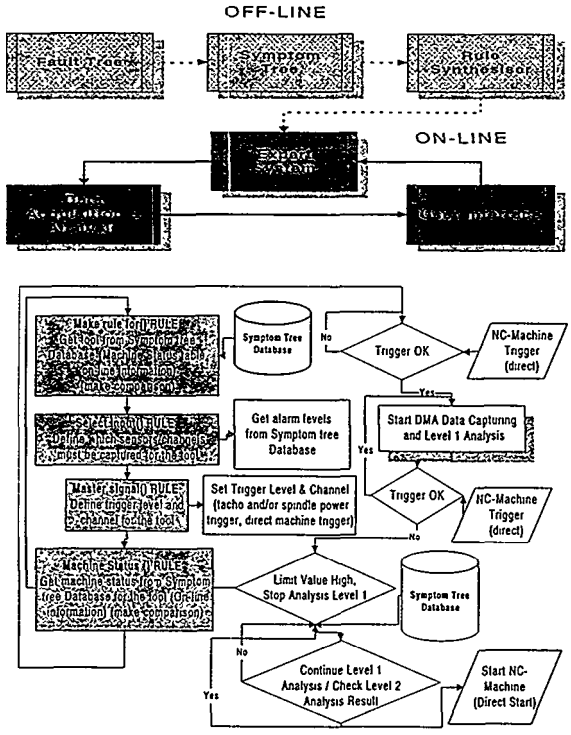


Fig. 7. Principles of the expert system rule generation and the data acquisition system.

7 ADVANCED COMPONENT DESIGN

One important and very basic way of improving the operational reliability of a system is the improvement of the reliability of each component in the system. Of course, there is the biggest need for reliability improvement of those components that have turned out to be or can be expected by calculations to be the most unreliable components in the system. When history data is available there is good knowledge of which these components are. On the other hand, at the design phase, the critical parts in the system have to be located by the available calculation methods. For this reason it is important to develop experimental test methods that result in appropriate reliability data, such as the failure probability as a function of age and the probable functional lifetime, as shown in Figure 1.

Improved reliability of components can be achieved by new design concepts or solutions or by improved material properties. Components often fail by defects in their surface such as wear and corrosion. To inhibit this there has been a tremendous development during the last few decades to find new surface treatment or surface coating techniques [24]. Today the problem of the component designer is more to find the optimal and most suitable surface modification technique of all those available and the optimal surface modification parameters for that specific technique to suit his application.

One of the most interesting recently developed new component coatings for improved performance and reliability is diamond and diamond-like hard carbon coatings. A very thin, only about 1 μm thick, diamond-like hard carbon layer deposited on the component surface (see Figure 8a) can improve the wear resistance by more than one order of magnitude, increase the lifetime by even two orders of magnitude and decrease the friction by one order of magnitude down to friction coefficient values as low as $\mu = 0.02$. This makes the surface of the component in completely dry conditions as slippery as if it was covered by oil [25,26].

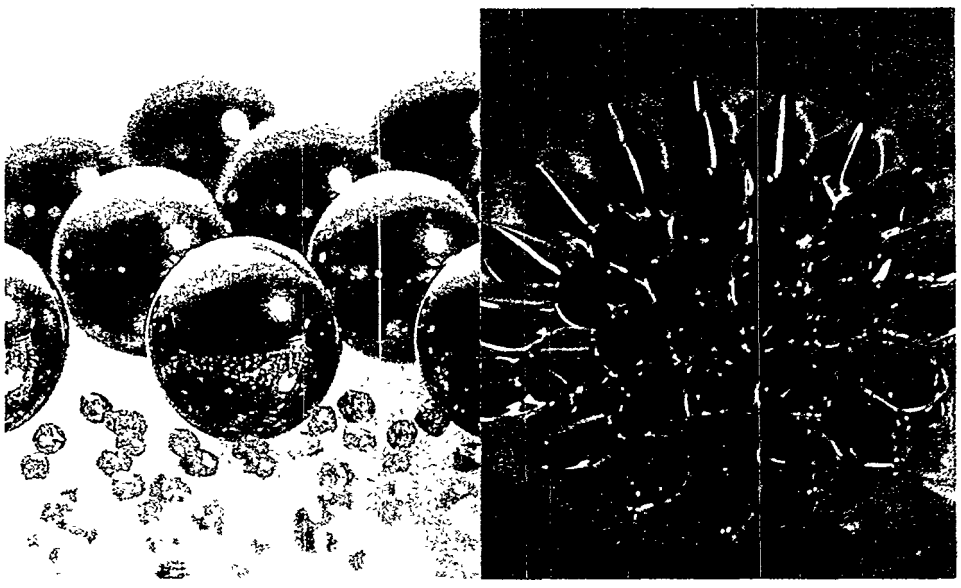


Fig. 8. a) Components coated by a hard diamond or diamond-like carbon coating results in improved reliability, increased lifetime and decreased friction. b) The use of magnetic fluids in seals makes them completely leakage, free and increases the lifetime considerably.

In many industrial systems seals are a critical component and there are several examples of disastrous accidents because of the failure of this component, the callenger accident perhaps being the most famous one. New and more reliable completely leakage, free seals have been developed by using magnetic fluids, as shown in Figure 8b [27]. Magnetic fluids or ferrofluids consist of magnetic particles suspended in the carrier fluid, which, can be e.g. a lubricating oil. The suspended ferrous particles are so small in size, about 5-10 nm in diameter, that they have no abrasive effects but they make the fluid act like a ferromagnetic material in the presence of an external magnetic field [28,29]. A sealing design with a magnet around the seal keeping the ferrofluid by a concentrated magnetic field just in the seal-shaft gap is completely leakage free, has no wear because of no solid contacts between the moving parts and has a coefficient of friction less than one order of magnitude lower than ordinary seals deigned for the same purpose.

8 A HYPERBOOK SUPPORTED TOOL FOR FINDING SOLUTIONS

We have seen in the presentation above that numerous different technologies are available today for solving problems of operational reliability in industry, in energy production and in communications. These include technical tools developed in areas such as material science, sensor development, signal processing, diagnostic engineering, information technology, telecommunications, risk management, reliability and safety engineering. The know-how of how to use these tools is, however, both in industry and in research organizations today spread amongst experts found in different companies or different units in a larger organization.

VTT has developed a holistic approach widely covering the different technologies, and technical tools available for solving maintenance and reliability problems. The PC-based hyperbook application OPERA helps the user achieve a holistic understanding of the problem, to navigate among the several technical tools and methods available, to understand the possibilities and limitations of each tool and method, to help choose the most appropriate tool, to see the synergistic possibilities that a suitable combination of tools can offer, and to find out where the needed expertise can be found [30]. The OPERA approach is based on the concept presented in section 3 and shown in Figure 1.

9 CONCLUSIONS

There is an increasing need for improved operational reliability of machines and equipment in industry, communications and energy production. Several new technical tools are developed today for this purpose and some of them are already on the market. Examples of such techniques and tools developed at VTT are presented in this paper:

- a metallographic replica technique for steam - piping lifetime estimation,
- an expert system, AURORA, for corrosion prediction and material selection,
- an automatic image-processing-based on-line wear particle analysis system,
- microsensors for condition monitoring,

- a condition monitoring and expert system, CEPDIA, for the diagnosis of centrifugal pumps,
- a machine tool analysis and diagnostic expert system,
- non-leakage magnetic fluid seals with extended lifetime and
- diamond-like surface coatings on components with decreased friction and wear properties.

A hyperbook, supported holistic approach to problem solving in maintenance and reliability engineering has been developed to help the user achieve a holistic understanding of the problem and its relationships, to navigate among the several technical tools and methods available and to find those suitable for his application.

REFERENCES

1. Dekker, R. "Applications of maintenance optimization models: a review and analysis", Reliability Engineering and Systems Safety, 51 (1996), pp. 229–240.
2. van Rijn, C. F. H. & Scholten, P. "Integral management of productyion assets", Maintenance, 11 (1996) 3, pp. 3–13.
3. Holmberg, K. & Folkesson, A. (Eds.) "Operational reliability and systematic maintenance". London: Elsevier Applied Science, 1991. 331 p.
4. Holmberg, K., Enwald, P. & Priha, I. "Maintenance and reliability - advanced technologies and technological trends", Proc. of 7th Int. Congress on Condition Monitoring and Diagnostic Engineering Management – COMADEM'94, New Delhi, India, 26–29 September 1994. 10 p.
5. Holmberg, K. & Ronkainen, H. "Käyttövarmuuden tekniikat / Operational reliability", VTT Research Notes 1557 (in Finnish), Espoo, Finland: VTT, 1994. 96 p.
6. Suokas, J. & Rouhiainen, V. "Quality management of safety and risk analysis", Amsterdam: Elsevier, 1993. 291 p.
7. Auerkari, Pertti & Vierros, Paula, "Relation between creep damage, strain and residual life in power plant steam pipes". Proc. of 3rd Int. Conference on Creep and Fracture of Engineering Materials and Structures, Swansea, UK, 21–24 April, 1987, The Institute of Metals. Pp. 905–914.
8. Auerkari, P., Salonen, J. & Borggreen, K. "Guidelines for evaluating in-service creep damage", VTT Manufacturing

Technology, NT Technical Report 302, Espoo, Finland, 1995. 15 p.

9. Auerkari, P., Salonen, J., McNiven, U. &. Jovanovic, A. "Classification of damage and prediction of its development in high temperature components of the power industry", Penny (Ed.) Risk, Economy and Safety, Failure Minimisation and Analysis, Rotterdam: Balkema, 1996. Pp. 269–280.
10. Hakkarainen, T., Carpén, L. & Hakkarainen, T. "The microcomputer as an corrosion expert", Proc. of the 1st Int. Symp. on Corrosion Prevention in the Process Industries, Amsterdam, The Netherlands, 8–11 Nov. 1988. Pp. 349–357.
11. Hakkarainen, T. "A model for prediction of possibility of localized corrosion attack of stainless steel", CORROSION/96, Houston, USA: NACE International, Paper 363, 1996. 7 p.
12. Hakkarainen, T. "Transferring corrosion knowledge to computers: Models for predicting corrosion", CORROSION/97, Houston, USA: NACE International, Paper 97396, 1997. 8 p.
13. Hunt, T. "Handbook of wear debris analysis and particle detection in liquids", Elsevier Applied Science, London, UK, 1993. 488 p.
14. Enwald, P., Liu, G. Q. & Holmberg, K. "An expert system for machine part condition assessment", IMECO XIII World Congress., Torino, Italy, 5-9 September, 1994. 6 p.
15. Enwald, P. & Liu, G. Q. "Systematic feature analysis of wear particles – a base for the automation of wear monitoring", Proc. Int. Conf. on Condition Monitoring´94, Swansea, UK, 21–24.3.1994. 11 p.
16. Hämäläinen, J. & Enwald, J. "Inspection of wear particles in oils by using a fuzzy classifier", Int. Conf. on Automatic Visual Inspection, Optics for Productivity in Manufacturing, Frankfurt, Germany, 20–24 June 1994. 9 p.
17. Paura, G. "MEMS sensors branch out", Mechanical Engineering, Oct. 1996, pp. 65–68.
18. Heimala, P., Hokkanen, A., Keinänen, K., Keränen, K., Tenhunen, J. & Lehto, A. "Mikroanturisysteemien tutkimus-ohjelma 1994–96 / Micro sensors systems development programme 1994–96", VTT Research Notes 1825, Espoo, Finland: VTT, 1997. 47 p. (in Finnish).
19. Jantunen, E. "Maintenance and downtime costs of centrifugal pumps in Finnish industry", Condition Monitor (1992)82, pp. 5–8.

20. Jantunen, E. & Ollila, J. "Diagnosis of centrifugal pumps", Euromaintenance´96 Int. Conf., Copenhagen, Denmark, 21–23.5.1996, 10 p.
21. Milne, R. & Malinen, P. "FMS MAINT: Maintenance for the future", Euromaintenance´96 Int. Conf., Copenhagen, Denmark, 21–23 May 1996. 11 p.
22. Jantunen, E., Jokinen, H. & Holmberg, K. "Monitoring of tool wear", Proc. of 8th Int. Congr. on Condition Monitoring and Diagnostic Engineering Management COMADEM´95, Kingston, Canada, 26–28 June 1995, pp. 143–150.
23. Jantunen, E. & Jokinen, H. "Reduction of data needed in an expert system for condition monitoring of FMS using regression analysis techniques" Proc. of 9th Int. Congr. on Condition Monitoring and Diagnostic Engineering Management COMADEM´96, Sheffield, UK, 16–18 August, 1996. 10 p.
24. Holmberg, K. & Matthews, A. "Coatings tribology – properties, techniques and applications in surface engineering", Elsevier Tribology Series 28, Amsterdam, 1994. 442 p.
25. Holmberg, K., Koskinen, J., Ronkainen, J., Vihersalo, J., Hirvonen, J-P. & Likonen, J. "Tribological characteristics of hydrogenerated and hydrogen-free diamond-like carbon coatings", Diamond Films and Technology 4(1994)2, 113–129.
26. Holmberg, K., Matthews, A. & Ronkainen, H. "Coatings tribology – contact mechanisms and surface design", World Tribology Conference, London, UK, 8–12.9.1997, 20 p. (invited presentation, to be published).
27. Säynätjoki, M. & Holmberg, K. "Magnetic fluids in sealings and lubrication – a state of art review", Synthetic Lubrication 10(1993)2, 119–132.
28. Mäkelä, K. "Electro- and magnetorheological fluids, Part I: Fundamentals" Finnish Journal of Tribology 15(1996)1, 41–47.
29. Mäkelä, K. "Electro- and magnetorheological fluids, Part I: Material aspects" Finnish Journal of Tribology 15(1996)2, 3–15
30. Holmberg, K., Vuori, M., Toola, A. & Säynätjoki, M "A hyperbook supported holistic approach to problem solving in maintenance and reliability engineering", Proc. of 8th Int. Congr. on Condition Monitoring and Diagnostic Engineering Management COMADEM´95, Kingston, Canada, 26–28 June 1995, 195-202.

THE MAINTENANCE WHEEL, OR, WHY DID MY GRANDFATHER'S STUFF LAST SO LONG

Kenneth Culverson,
Reliability Engineering Consultant
Alameda, CA, USA

ABSTRACT

In the beginning, people used simple single cell tools. Often hand made, each tool had a very utilitarian function which was easy to understand. The feel, sound, smell and sight of the tool at work attested to its quality. These same senses told the user when things were not right, and appropriate action was taken to attempt to reduce the degradation of function, or restore the tool to its proper form.

As tools grew more complex, more parts, often more functions were added and it was not always as easy for the average user to determine when the tool was not working right. This led to the growth of the "master" who had a "feel" for the tool and became the local expert on keeping everything right, or fixing the tool if necessary.

1 THE MASTER MECHANIC

Condition monitoring is not just an interesting academic pursuit. Condition monitoring is being applied ever more widely in industry all the time. The issues and ideas that we discuss at this congress and other meetings like this one have a direct effect on the way maintenance is performed in society at large. As anyone who has attempted to institute a condition based maintenance system can tell you, there is a large learning experience that must be gone through by the organization for the system to succeed.

Today, in industry, we have to deal with the perceived gulf between the "knowledge" workers, and the "hands-on" workers. There are prejudices toward the opposite group in most organizations. Worse, the attempt to

institute a condition based maintenance program is usually imposed on the craftsmen by the "knowledge" worker, i.e. the engineers. As any master machinist or electrician can tell you, the engineers do not know anything about keeping a plant running perfectly. Besides, these old timers have done an excellent job keeping things running in the past, and will continue to do so regardless of what that young engineer thinks. This is the world I work in every day. I not only have to develop the condition monitoring program, I have to sell it to the master craftsmen.

2 EARLY CONDITION MONITORING

In the early days of the industrial revolution, most factories were small affairs employing a few people in the manufacture of a single product or a small line of allied products. The tools usually had one function. For example, a lathe was used to turn wood and later metal parts out of roughly square stock. The power for the lathe was simple. The earliest lathes used foot power and a simple cord and pulley power transmission apparatus. The operator of the lathe knew when the power was applied evenly to the turning head stock by the feel of the piece against the tool, and the speed of the cut. When water power and later steam power were used, the power transmission was changed from a cord to a belt but the method of determining the efficiency of the drive did not change. Human senses provided the tool to monitor the condition of the drive.

Likewise, the condition of the tool was gauged by the operator by the amount of "chatter" (vibration) the tool made doing the work and the heat generated by the cut. Another condition that was important, because the tools had to last for a long time, was lubrication. The lubricant had to have the right viscosity, and be clean. The sometimes crude bearing surfaces did not tolerate any debris in the lubricating film.

This craftsman most likely had a forge nearby to dress tools, manufacture fittings and make the fasteners used in the final product. As anyone who has ever done any work on a forge can tell you, getting the temperature right is important not only in keeping the work malleable but also to decrease the time and effort required to do the job. A good "smithy" could break the slag formed on the top of the forge, add just enough air to achieve the proper look of the coals, and proceed to heat and work the metal in hand with ease. This required knowing the condition of the forge and the material to select the best heat for the job.

If this craftsman were like my grandfather, he would assemble the various parts he had patiently made of wood, leather and metal into a useful article which would be expected to last. The tool case my grandfather made over one hundred years ago is still in use, and contains many of the tools he made, all still serviceable.

2.1 The Rise Of Specialization

As the pace of the industrial revolution picked up, groups of people came together to establish firms which would allow them to be more productive and increase their wealth. This growth from individual craftsmen to small factories to the huge factories that rose in the last century had several benefits for the financial backers of the enterprises and for the populace in general. More goods were available at prices many could afford. The down side of this effort, from the perspective of condition monitoring, was the loss of the general knowledge of a tool or products condition by the general laborer.

One reason for this loss of tool knowledge was the drive to produce goods as cheaply as possible, and as quickly as possible. This led to increased fragmentation in the work place. Even the early attempts at automation made it possible to employ workers with lesser skills to operate, or tend, these new machines. A small number of very experienced workers could be retained to maintain the equipment. While these "master mechanics" brought skills and experience to the job, the scope of many factories made it difficult, at best, to know the personality of each machine. Since the mechanics did not operate the machinery, they did not have the operational feel of each machine and could only guess what the operator actually did during a work shift. Anyone who has talked to these "old-timers" know that they have a very low regard for most operators, who "break everything they touch" and worse. Still, these mechanics brought a wealth of knowledge of the machines in general and knew the proper feel of a healthy machine, and the symptoms of a sick machine.

Even today, if a plant is fortunate enough to keep a long term group of mechanics, you will still see the old "master mechanics" feeling bearing surfaces for abnormal heat; feeling and smelling the lubricant for loss of lubricant properties or contamination; tapping on a structure to determine its sturdiness; or listening to a motor end through a 12" screwdriver to

hear vibratory events in the motor. This early form of condition monitoring and diagnostics is alive and well in many firms worldwide.

2.2 The Effect Of Aviation On Maintenance

In the beginnings of aviation, the aircraft were usually home designed and home built. If they did not kill their builders on the first take-off, they were maintained by their designer/builder. These early aircraft had the minimum number of components required to achieve flight. Everything was critical, and failure of any part usually ended with tragic results. This complexity of function, which required the fine interaction of so many different components, soon became more than one man or group of men could master. Increased specialization of crafts accelerated as one group maintained the structure, another, the engine, another group the controls, etc. As the aircraft became larger and more complex, people wanted to ride on the aircraft and the airline industry was born. Of course, there were risks associated with leaving the earth's surface to travel and the requirements for safe flight and availability led to the concept of time (or hour) based maintenance on everything. Much of this type of maintenance was the result of crashes and the resulting investigations that usually found that if someone had checked some part before the flight, the crash would not have happened. Besides, it was difficult to carry the mechanic along on every flight.

Along with the beginnings of aviation came the beginning of the automobile. Speed was the key, speed greater than anything people were used to became the norm. These new technological advances liberated us from the "old ways' and the concepts of aviation maintenance took hold in other areas as well. After all, if it was good enough for the modern technological marvels people were seeing, then it must be good for the more mundane tasks. Along with the new speed of traveling, was a rise of a new class of worker who dealt with paper rather than working with tools. There was a further loss of the old "feel" of how things should behave. The increasing complexity and cost of tools made it more difficult for the average person to do jobs that had been considered normal just a couple of generations before. The same forces were at work in factories. Maintenance was a cost, better avoided if at all possible. But the machinery had to be kept busy with the lowest cost. What was the best to keep the machinery working? Time based maintenance became the method of choice if any maintenance choice at all was made, and the concept of condition monitoring began to fade.

This led to those industries who embraced preventive, time based maintenance as the cure for all their problems, the progressive ones, and those industries that decided either through economic pressures or ignorance to let things run until they failed and then try to fix them. Many of the basic industries took the latter approach with great economic and human loss.

3 CONDITION MONITORING RESEARCH

All the above is not to say that condition monitoring was completely dead, or that nothing was being done in the field to expand the capabilities and knowledge of how machines worked. Basic research was conducted in universities worldwide on the principles which would lead to the integrated technologies used today. The great depression which descended on the developed nations in the 1920's and 1930's led to massive public works projects including funding much of the basic research we consider to be the roots of condition monitoring today.

3.1 Basic Physics And Engineering

The underlaying principles used by condition monitoring in an industrial setting are based on some fundamental laws of motion and heat. In the second quarter of this century, considerable research was conducted into new ways of analyzing these laws and determining new ways of analyzing the phenomena. Advances in the mathematics of vibration laid the groundwork for future work. The chemical engineering schools made great strides in researching the properties of lubricants and elastomeric materials used in power transmission. Mechanical engineers searched for better and more efficient ways of transmitting power, developing new techniques and solving problems caused by the new techniques. Electrical engineers searched for new ways to manipulate data through the use of analog computers. Physicists went back to elementary thermodynamics to explore new ways of measuring temperature.

3.2 Advances In Computing

While all this research provided a sound foundation for the application of condition monitoring as a practical field of endeavor, it was hampered by

the time it took to process any data obtained. While this was not a real concern in the school environment at the time, the lack of adequate computing power kept these techniques confined to academia.

The development of digital computers during World War II was just what this basic research had required to go beyond the laboratory. Here, at last, was a method of crunching numbers well beyond anything that had been considered before. Of course, these computers were large, expensive and in limited supply. The government and a few large firms acquired the early computers and with the aid of some brilliant programmers started using computers to perform extremely fancy mathematical calculations to a precision not before possible. Still, unless you worked for one of the government or corporate research laboratories, all you could do was dream about the possibilities. Then you went back to the pad and paper, your log tables and slide rule and brute force went through your problem.

3.3 Solid State Electronics To The Rescue

The invention of the transistor, and all the solid state components which followed will have to be marked as one of the seminal points in history. Now, for just a few hundred dollars, one could obtain a desktop calculator which could do routine mathematics with just the push of a few buttons. You still needed the log tables to handle more complicated calculations, but there was hope on the horizon.

4 THE MODERN ERA

Precisely when the modern era of condition monitoring began could start a lively discussion that would go on forever. From my perspective, the modern era of condition monitoring began when machinery reliability became an important issue in the design and maintenance of systems. Aviation is generally accepted as the leader in looking at basic reliability of systems as a method of determining what maintenance is required to achieve the basic reliability designed into the equipment.

Aviation had grown and flourished since its founding at the turn of the century. Following World War II, the new techniques used to manufacture military aircraft during the war were turned to the

commercial sector, and for the first time, air travel became a realistic, if somewhat expensive, way for the average person to travel. Along with new processes for manufacturing airplanes, came greater redundancy. Greater redundancy provided greater reliability and safety for the flying public. Maintenance remained time based, however. The introduction of turboprop and pure turbojet engines decreased travel time even further, and airplanes started flying higher to take advantage of the operational efficiencies inherent in the new engines. Still maintenance was time based, there was just more of it.

4.1 The Economics Of Time Based Maintenance

The turbojet airplane brought some new, unforseen maintenance consequences. The loss of two De Haviland Comets was determined to be the result of stress caused by the pressure cycling during flight. The basic structure of an aircraft had not been the focus of most maintenance programs, but now this was added to the maintenance requirements. More time was spent in maintenance and less time spent flying.

In the 1960's, Boeing developed the B-747 aircraft. When the prospective owners and operators of this aircraft looked at the size and weight of the aircraft, it was felt that something had to be done to simplify maintenance or the aircraft would be uneconomical.

There exists an algorithm which uses the aircraft weight to determine the number of hours of maintenance will be required per year. The number was staggering for the B-747 so the airline industry, led by United Airlines and the Boeing Company researched failures and maintenance to see if something could be done to alleviate the problem. Thus was born the first Maintenance Steering Group which ultimately developed Reliability Centered Maintenance. For the first time, maintenance was dictated by the function and criticality of that function. Maintenance tasks were keyed to operational factors such as flight hours or landings and takeoffs, rather than the calendar.

This process was applied to all later jumbo jets, and to several jets in service with the airlines. The U.S. Department of Defense adopted the techniques for aviation and the nuclear submarine fleet, as did the nuclear power industry.

Condition monitoring began in earnest in these industries. The use of oil analysis to determine the condition of engines, Thermographic analysis of composite parts, traditional non destructive inspections were expanded in use to determine the condition of fasteners and airframe components.

Once the non-aviation sector began looking at the ways reliability could be used in maintenance, the growth of condition monitoring was ready to take off, with vibration analysis taking much of the lead. Oil analysis was healthy and in use as well.

4.2 Technology Issues

Those people who used some of the early vibration data collectors, or oil analysis equipment or Thermographic equipment can testify to the difficulty of taking data, analyzing data, and making the data presentable to those people who had to make decisions.

The microprocessor changed the way the world does business, and the way we do condition monitoring, as well. Finally, it was possible to pack all the electronics required to acquire a signal, process that signal and display the results in a very small box. Portability was the real key to making condition monitoring widely acceptable. For all the research done in various laboratories and universities, it was the microchip which fueled the growth we see today in condition monitoring.

5 THE MASTER MECHANIC

In many ways, we have come full circle in maintenance. The use of portable equipment which allows us to see inside a machine and determine its condition is giving rise to a new generation of "Master Mechanics". While they may not have all the years of experience the previous generation had, they are able to make diagnoses of their equipment with the same degree of expertise their predecessors did. In fact, the diagnoses made by the new generation of "Master Mechanic" are probably more accurate and timely than that of their predecessors.

It is interesting to watch the attitudes change of these new mechanics who have taken to the technology. They gain a missionary aura fairly rapidly, attempting to convert all those around them to condition

monitoring. What is perhaps even more interesting is the response of those older more experienced mechanics when they finally start to use the new technologies. It is like they have gained new powers and are usually the first to find new ways to use the technologies.

This is probably the reason that many companies who have adopted the condition monitoring philosophy have returned the ownership of machinery to those who maintain it, and in some cases, to those who operate it.

There have been great advances made in the world of condition monitoring during all of our lifetimes. But I feel it is important to remember that what we are now doing is not completely new. Yes, the tools are new, and the analysis is more precise, but we should all remember the early craftsman who practiced condition monitoring by instinct. These were truly the pioneers in our industry.

I feel that if my grandfather were around in these times, he would embrace the technologies and see them as an extension of the skills he mastered so many years ago.

TPM EXPERIENCE - EUREKA MAINE PROJECT EU1190: THE FAST TRACK TO WORLD CLASS MANUFACTURING

Peter Willmott B Tech (Hons), CE, FIEE.
Managing Partner
WCS International
Bath, England

ABSTRACT

You have restructured your manufacturing operation in order to bring down labour costs. You have refocused the company around its core activities. You have introduced Just-in-Time to reduce stocks. You have sought to build a Total Quality Culture. What else can you do to improve your competitiveness?

Total Productive Maintenance (TPM) can slash unit costs in manufacturing and process industries by ensuring that plant and equipment are used to their maximum effectiveness. In the last few years, a number of large European companies have implemented TPM. For many, the process has been hugely successful.

The time has now come to review the European experience of TPM, good and bad, in order to develop reliable models for its widespread adoption. TPM offers Europe an opportunity to increase its competitivity in world markets. Two acknowledged experts - Willmott Consulting Services International (WCS), based in the United Kingdom and the Swedish Institute of Management for Innovation and Technology (IMIT) - have joined forces for this EUREKA project, which falls under MAINE, the EUREKA umbrella for maintenance techniques in manufacturing. This paper will review the progress to date including:

- Why and How participating companies are implementing TPM
- Their Key Success Factors and pitfalls to avoid
- How to benchmark, audit and review your TPM Progress both from the hard quantifiable aspects as well as the soft cultural issues.

1.0 Introduction

1.1 Project Summary

- The project will Research, Quantify and Qualify the TPM Experiences of a number of European TPM-practising Companies and Industries.
- These experiences (both good and bad) will be Interpreted and Developed
- into a Methodology which can be Tested and Transferred to both Non-Practising & Practising Companies, especially Small, Medium Enterprises and Key Suppliers.
- The Project Deliverables will also Define the Parameters, Steps and Measurement Criteria for Developing a European TPM Benchmark Standard.

1.2 Project Objectives

To develop TPM Technology Transfer "Models"

- From Large Companies to Small Medium Enterprises
- From One Sector to Another
- From Supplier to Manufacturer to Customer
- From One Nation to Another

LEADING TO .....

EUROPEAN TPM BENCHMARK STANDARDS

1.3 Benefits of Participation

As a Participant in the "TPM Experience" Project, you will be influencing and contributing to the Development of the latest European TPM "Body of Knowledge" This will include:

- The Development of Practical Help Guides and Checklists to Successful and Speedier TPM Planning and Implementation

- Access to Latest TPM Developments. For example: TPM for Design and TPM in the office.

- The Use and Trialing of exacting TPM Self Assessment Criteria (Hard Measures and Cultural/Soft Measures) against which to Audit, Benchmark and Monitor your Progress.

- Networking with similar (and dis-similar) industrial companies within Europe.
- Providing you with Closer and Practical Links with your Supplier and Customer Base.
- A Single Point Access to TPM "Best Practice" (and Pitfalls to avoid)

NAMELY: A HIGH QUALITY, NO/LOW COST DYNAMIC TPM INFORMATION NETWORK

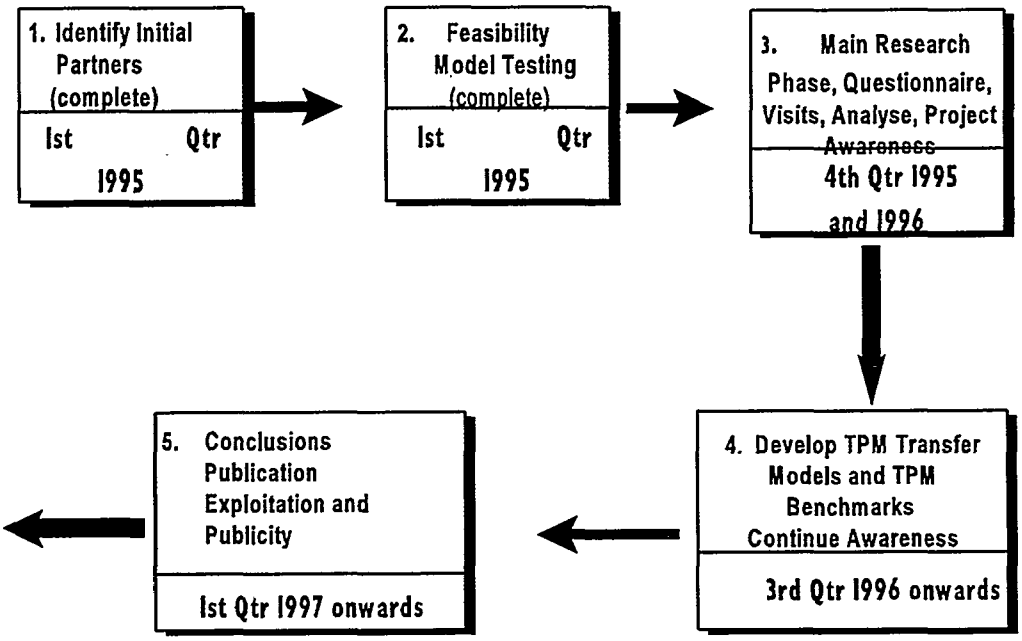
1.4 Industrial Participants

Abbey Corrugated
Advanced Films Ltd
Amcor Packaging
AE Goetze Automotive
Alcan Rolled Products UK
Ashley and Rock Ltd
Aylesford Newsprint Ltd
BFF Non-Wovens
BICC Cables
Biwater Industries
Bonar Teich Flexibles Ltd
Bourns Electronics (Ireland)
BP Exploration (M)
Bridon International
British Aerospace (M)
British Steel
Caradon Ltd (M)
Cassidy Brothers PLC
Cleveland Potash
Clyde Bonding Company
Coils UK Ltd
Courage Brewery (M)
Croda Colloids Ltd
Croxton & Garry
Devro
Dutton Engineering
Edward Hall Ltd
EEV Ltd
Esselte (M)
Euro Gas Turbines
Exxon Chemicals (M)
Fine Organics
GEC Plessey Semics
Gorenje Inova
Guinness Brewing
Hoechst Trespathan
Howegarden Ltd
H & R Johnson
Hy Q International Ltd
Ibstock Brick
KNP - BT Packaging (M)
Leaf UK Ltd
Kesslers International Ltd
LIP (Equipment & Services)
Lyme Green Machining
McDougalls
Nova Aerospace
Mitsubishi Electric UK Ltd
Norton Chemical Process Prdts
Pilkington Glass
Pirelli
Pressfab Sections Ltd
Rank Hovis (M)
Rhone-Poulenc Chemicals
RHP Group (M)
Rolls Royce Cars
Rolls Royce Associates
Rover
Royal Mail
RR Donnelly
Sandos Ringaskiddy Ltd
Schade UK
Schlumberger Industries
Securistyle
Sellotape GB Ltd
Snack Factory
Solvay Interox
Sony Manufacturing Co Ltd
SPS Technologies
Stevens & Bullivant
Strix Ltd
Synthomer Ltd
3M (M)
Tekniska Hogskolan I Lulea
Thorn Lighting
TRW Steering Systems
Tuberex
Tucker Fasteners Ltd
UES Steels
Unilever Int Marketing
United Distillers
Warner Lambert
William Lee
Yardley
ZDB

1.5 Non-Industry Participants

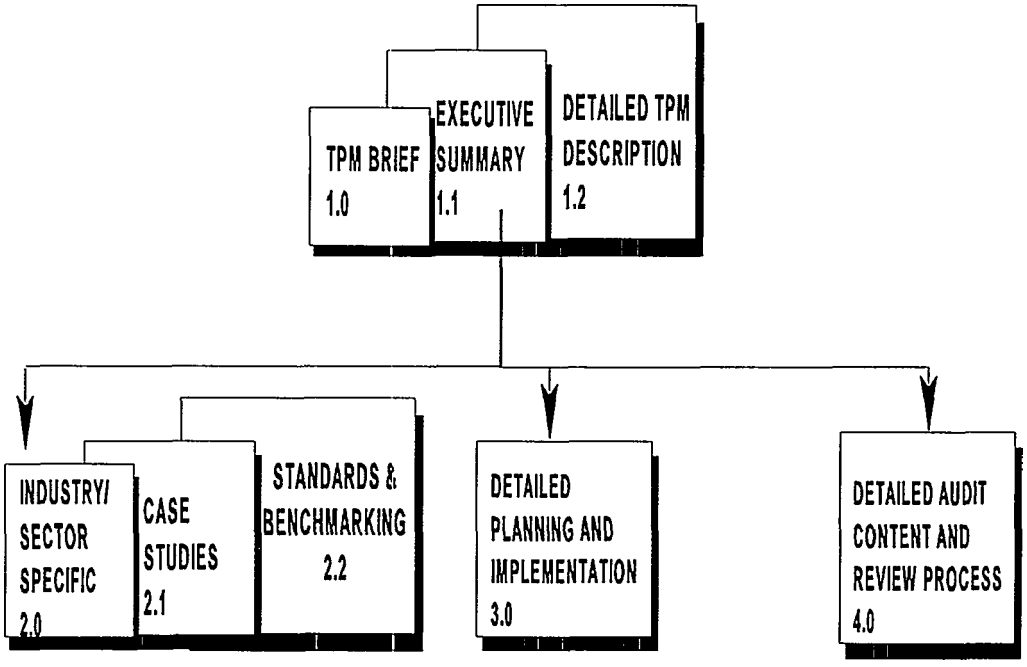
- WCS International, Project Leaders, UK
- Findlay Publications, Awareness Partner UK
- F J Systems, UK Software Partner UK
- DTI (EUREKA-MAINE Agent), UK
- Chalmers University, IMIT, Sweden
- AEEU, (Elec/Mech Trade Union), UK
- ASSYNT
- Conference Communication, UK
- EPIC Interactive, UK
- European Union
- FORBAIRT (EUREKA-MAINE Agent), Ireland
- Informa, Austria
- IRISH MANAGEMENT INSTITUTE, Ireland
- IWT Belgium
- MCP, UK
- NCSI
- Rogaland, Norway
- Ryn Consult, Holland
- STATUS Meetings, UK
- ZDB, Czech Republic

1.6 Project Stages



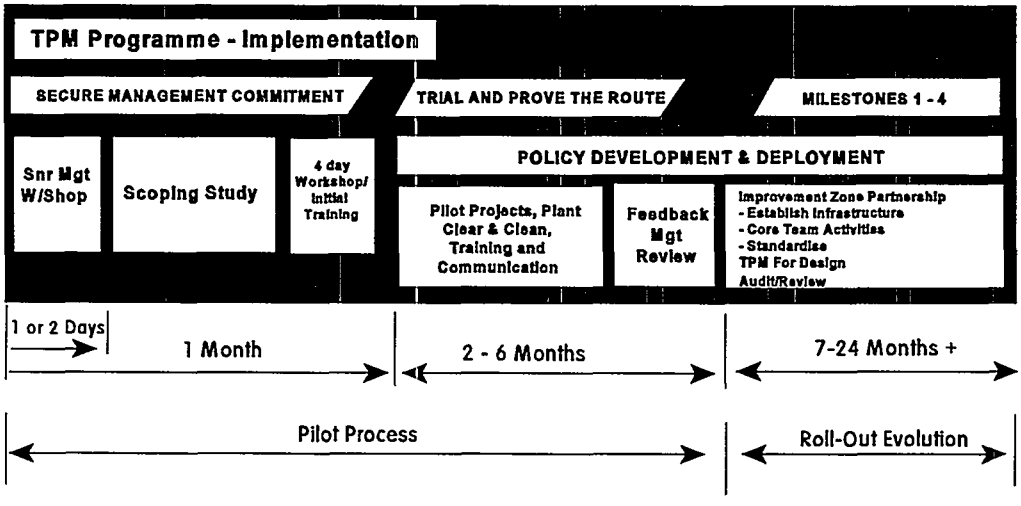
2.0 Framework for Dissemination

2.1 Structure and Content



2.2 TPM Implementation Process

Approach



2.3 Auditing and Measurement of TPM Progress

One common denominator of World Class Manufacturing Companies is their determination to set tough improvement targets and then make sure they have a robust measurement system in place to monitor their progress. Common sense you might say, but rarely done on practice.

If TPM is to deliver the business benefits and cultural change indicators it is capable of, then we must have an Auditing and Measurement process in place which can provide answers to the following eight questions:

- Where are we now? (the reference point)
- Where do we want to get to? (the business drivers and vision)
- How do we get there? (the route map)
- How long should it take us? (the journey time)
- What are the measures? (both 'hard' and 'soft')
- What are the milestones? (intermediate checkpoints)
- What are the targets? (specific measurement goals)
- How do we compare with others? (Benchmarking)

In essence, we need a top- down delivery system with a bottom- up measurement process covering key performance indicators (KPI's) for:

- Productivity
- Quality
- Cost
- Delivery
- Safety
- Morale

Some typical but not exhaustive KPI's for TPM are shown below:

P
- Overall Equipment Effectiveness
- Productivity (Output/Person/Day)
- Number of Failures
- Number of Minor Stops

Q
- Number of Defects
- Number of Customer Claims

C
- Maintenance Cost as % of Sales

D
- Inventory turn (Days of Sales/WIP)
- Lead Times
- Order Completion

S
- Number of Accidents

M
- Number of Suggestions
- Dust
- Temperature
- Noise

Also shown below are some of the achievements of five Japanese plants who are TPM Award Winners. This level of performance sets the World Class Standards to which European Manufacturing Plants must aspire if they are to be truly competitive.

KPI PLANT	BOOK COVERS	CLOTHING	CHEMICALS	AIR CONDITIONERS	LIGHTING SYSTEMS
BREAKDOWNS PER MONTH	18 5 3	707 2 ½ 15	200 4 10	250 6 5	387 4 33
PRODUCTIVITY	10 0 2 125	100 2 ½ 120	100 4 150	100 6 200	100 4 247
OEE%	55 3 75	71 3 85	100 4 160	65 6 88	71 4 88
INVESTMENT $m	1.9 5 3.4	1.0 per $ 2 ½ 10.0	1.0 per $ 4 3.0	1.0 per $ 6 4.5	0.75 4 3.5

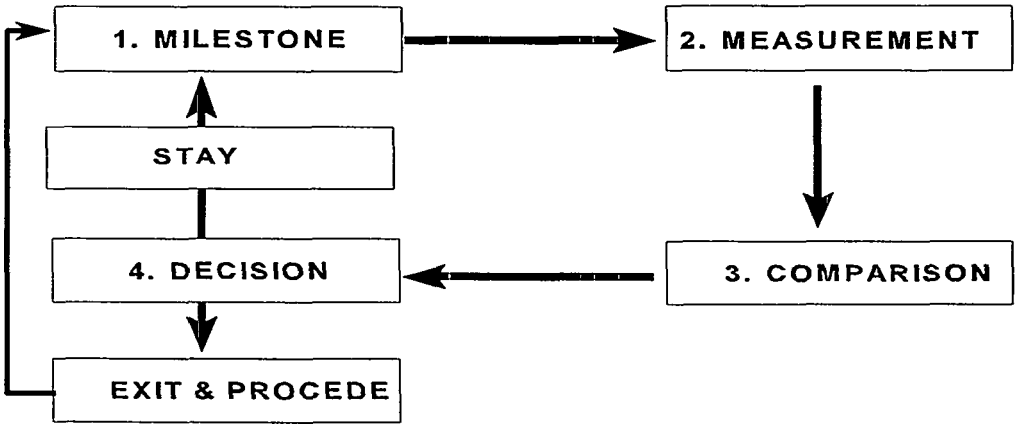
KEY:

Reference Point
Time (Yrs)
Achievement

The messages which are now beginning to emerge from our "TPM Experience" Industrial Partners are that we need three sets of KPI's covering:

- Direct Cost of Maintenance. (Value for money)
- Indirect/Lost Opportunity Costs (Benefit Driven)
- Cultural 'Soft' Measures/Personal Development Indicators (Benefit Driven)

Also, if we are going to measure something, then we need clear milestones and exit criteria, where the concept is basically straight forward and can be illustrated as:



Also, the other essential requirements are that any audit and measurement process must be:

- Simple to carry out (but not simplistic)
- Relevant, manageable and economic
- Capable of demonstration
- Actionable
- A Step - by - step approach
- Capable of recognising and measuring personal development and cultural differences

Through the "TPM Experience" Project, we are developing four milestones:

Milestone 1 - **Introduction** including Awareness, Pilots, Plant Clear and Clean and Getting Everyone Involved

Milestone 2 - **Refining Best Practice and Standardising** across all areas and shifts to get the Basics Right

Milestone 3 - Where we are **Building Capability** of our People, Equipment and Processes towards innovation

Milestone 4 - **Striving for Zero Losses** and Achieving the Vision which was set at the launch of TPM

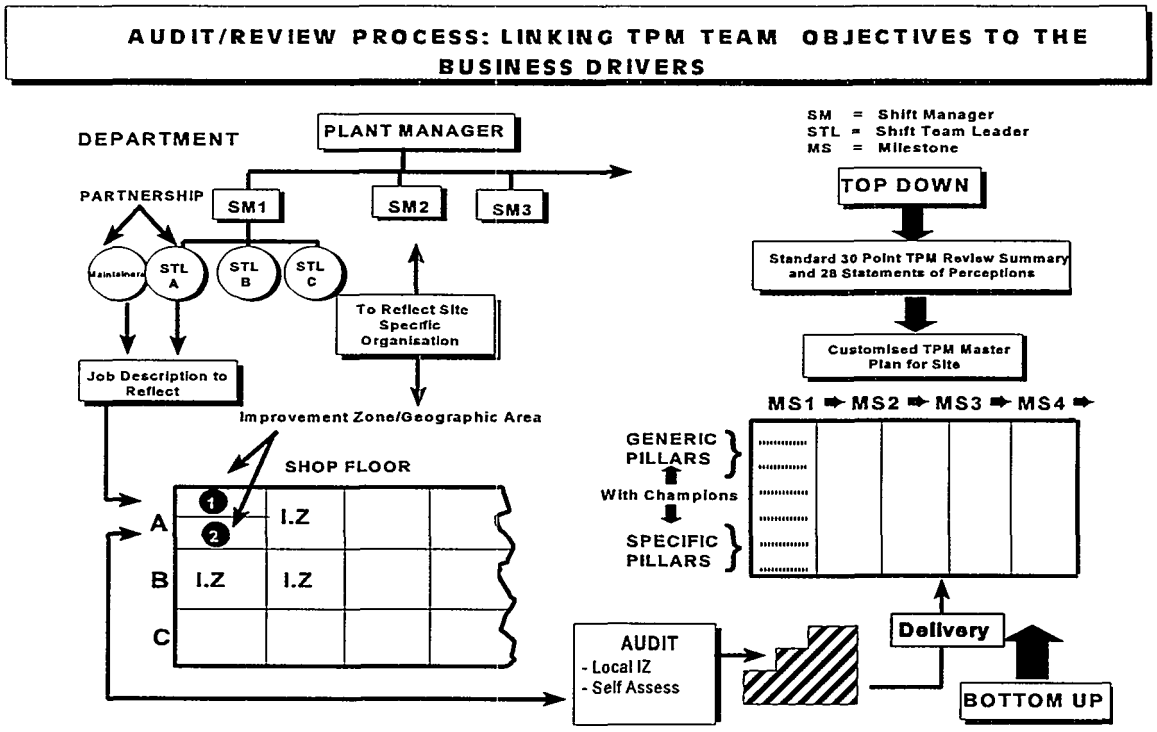
Broadly speaking, each milestone will usually take around 12 months to achieve.

The Top Down Management Review process is being developed under seven headings of:

- Business Management /Strategy
- Infrastructure and Technology
- Systems in Place and Working
- Objective Feedback Communication
- Training and Skills Development
- Cultural Indicators
- Motivation

In addition there is a 28 statement evaluation to measure individual's opinions, whether they are managers, operators or maintainers, and their readiness to accept the challenges and cultural changes implicit in the TPM process. (See Works Management January 1996).

In summary therefore we can link all of the above into the Audit Review process shown below where achievement of Milestone 2 is equivalent to Autonomous Maintenance step 4.



The Bottom Up Audit process, (which also checks that the top-down TPM Management Review Process is actually being delivered) is a series of structured checklists linked to the five pillars of TPM of :

- Increasing the overall Equipment Effectiveness (OEE)
- Improving the existing Planned Maintenance Systems
- Autonomous Maintenance (Self Determined Asset Care)
- Increased Skills (hand/operational, teamworking and problem solving skills)
- Maintenance Prevention through Early Equipment Management

and the four milestones detailed above in order to **Pursue Ideal Conditions**.

Only by linking the TPM Team Objectives to the Business Drivers will we know that TPM is becoming an essential part of "the way we do things here" process rather than simply a project with a start and a finish.

3.0 Vehicles for Dissemination

Vehicle	Seriously Considering TPM	About to Launch TPM	Already using TPM
• Written word (Publications)	3	3	3
• Audio/Visual (Video, CD-ROM)	3	3	3
• Manuals	3	3	3
• Conference/Seminars	3	2	1
• Workshops	2	3	2
• Industrial & Professional Bodies and Agencies	3	1	1
• Business Links	2	3	2
•European TPM Practioneer's Network	1	1	3
• Inter Plant Visits	3	3	3
• Study Tours (countries)	1	2	3
• Web Site	3	3	3

1 = Minimal Value 2 = Worthwhile 3 = High Value

4.0 Summary

The latest project situation as at June 1997 will be presented at the end of this paper.

EFFECTIVENESS OF VIBRATION-BASED MONITORING SYSTEMS IN PAPER MILLS: TWO CASE STUDIES

Basim Al-Najjar
Lund University and Växjö University, Sweden

Uday Kumar
Luleå University of Technology, Sweden

ABSTRACT

The effectiveness of vibration-based monitoring, (VBM), system may be measured in terms of reduced stoppage rates and increased use of potential bearing life. Analysis of data from roller bearings at two paper mills suggests that greater bearing lives can be safely achieved by increasing the accuracy of the vibration data. This paper relates bearing failures to the observed vibration spectra and their development over the bearings' lives. The literature was found mainly to confirm the author's analysis. A systematic approach which describes the performance and objectives of the studies is discussed. Explanations of the mechanisms behind some frequent modes of early failure and ways to avoid them are suggested. It is shown theoretically, and partly confirmed by the analysis of unfortunately incomplete data from two paper mills over many years, that accurate prediction of remaining bearing life requires (a) enough vibration measurements, (b) numerate records of operating conditions, (c) better discrimination between frequencies in the spectrum and (d) correlation of (b) and (c). This is because life prediction depends on precise knowledge about primary, harmonic and side band frequency amplitudes and their development over time. Much money could be saved because some policies utilise as little as half of the bearing's potential life.

1 INTRODUCTION

Often, defects which began in one part of a bearing cause damage to other parts after a while. Thus, the bearing condition cannot be assessed reliably without considering the significant damage as recorded by its vibration history. Vibration frequencies and their multiples, generated by bearing defects whose amplitudes exceed predetermined levels, can be utilised to

indicate bearing condition, [1, 2, 3, 4, 5, 6]. They can used for identification of failure causes, defect developing mechanisms and failure modes of most types of faults in rotating and reciprocating machines. Stiff competition made companies within the process and chemical industries, such as paper mills and refineries, intolerant of unplanned and not warning maintenance. VBM systems provide earlier indications of the changes in the machine state. These can also be used in detecting deviations in the product quality before they show on quality control charts [7]. In spite of using VBM programs, manufacturing machinery still experiences failures and unplanned-but-before-failure replacements, (UPBFR), which arise when a defect develops without being detected early by the system or maintenance staff due to personnel or system error, or some unexpected failure cause [8]. The diagnosis and prognosis problem can be stated in this way; we can detect many faults but we cannot say with a reliable degree of certainty how serious they are. The seriousness of the amplitude of a particular vibration frequency or frequency band is usually evaluated subjectively. The usual rule is; replace the bearing as soon as the amplitude of some of the bearing defect frequencies exceeds the normal level. With more accurate data and better records, analysis strongly suggests that more of the potential life of the bearings can be used. It was impossible to present all the data collected because editorial space limitations.

2 CASE STUDIES

The study analyses vibration measurements collected during 2 to 4 years from four types of roller bearings and used by two different paper machines and companies. The paper mill companies, A and B, have 4 and 3 paper mill machines respectively, of different ages. Vibration is the main parameter used for monitoring the machines. The vibration measurements are collected using Microlog and Presim^2 software in A, and CSI with the Master Trend Program from Computational Systems Inc. in B. The measurements are all in mm/s, RMS. The interval between measurements varied between 2 to 5 months in A and about 3 weeks in B. The paper machines, PM10 at A and PM01 at B were selected for investigation because their databases include more replacements of identical bearings than the other machine databases. The systematic approach adopted for the analysis is summarised at Fig.1.

2.1 Case study A

Measurements of ten identical replaced spherical roller bearings of type 23228ck/SKF which are usually used at the driven side of the leading roller of drying cylinders in PM10 formed the data. The bearings were selected because they belong to the most troublesome area in the machine. There were not enough replaced bearings from other types to be included. Table 1 gives a summary of the information regarding these ten bearings.

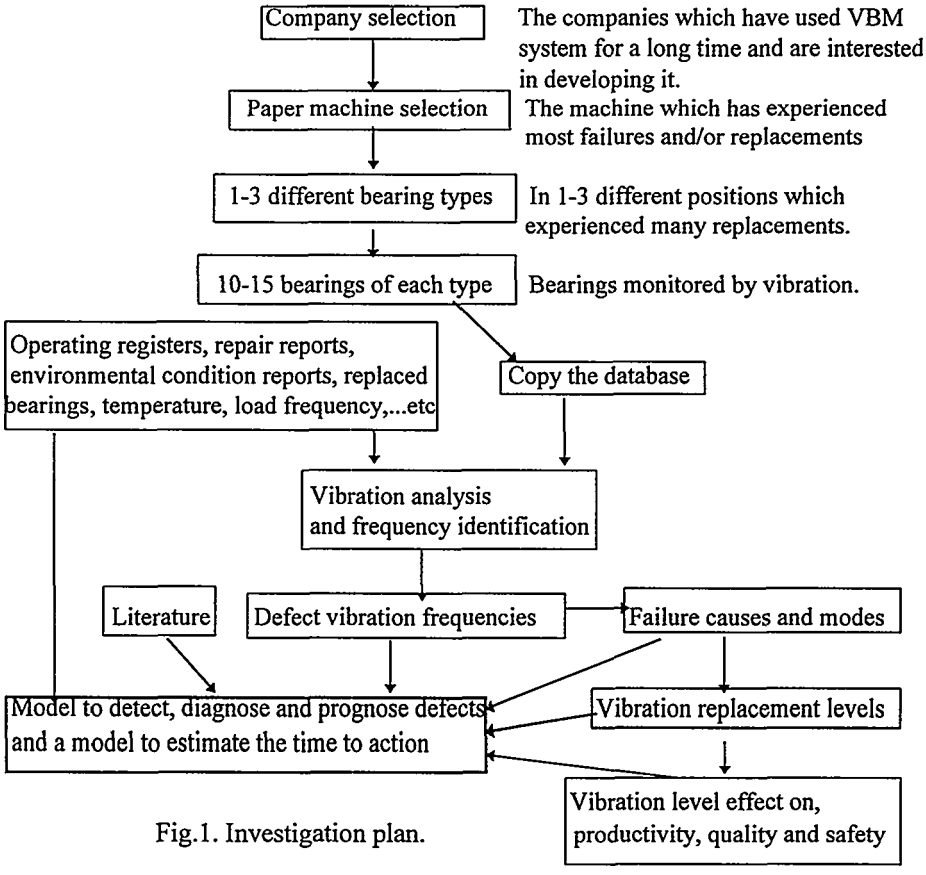


Fig.1. Investigation plan.

2.1.1 Analysis and results

The higher machine stiffness in the vertical direction makes the vibration amplitude lower than at the horizontal direction. Bearing frequencies of amplitude equal to or less than the noise level (0.05 mm/s) experienced at the machine, are considered insignificant and those equal to or larger than 0.15 mm/s significant. In general, the change in the vibration amplitude can be considered significant if it exceeds double the original level [1,8]. Frequencies which acquire amplitudes between 0.05 mm/s and 0.15 mm/s

are considered significant if there are other detectable bearing defect frequencies in the spectrum. This means that the significance of defect vibration frequency amplitude is weighted by its severity and by the whole bearing damage severity. Let X and i*X denote the machine speed and its ith multiple respectively. The existence of; 2*X only, 1*X and 2*X simultaneously, many multiples of X, and many multiples of X and especially 2*X in each vibration spectrum, are denoted by 2X, 1.2X, iX and i.2X, respectively. Where BPFI, BPFO, BSF & FTF are defect frequencies for inner race, outer race, rolling elements and cage, respectively [10].

Roller No	Date of installation	Date of replacement	Direction	Low level BPFI BPFO BSF FTF 1X 2X 3X .	High level BPFI BPFO BSF FTF 1X 2X 3X .	Replacement vibration level, is the overall level, rms, in mm/s vertical	horizontal	axial
41	880331	941027	ver. axi.	I O B F iX - - - - -	- - - - 2X I O B F 2X	2.68	5.27	2.68
56	July/83	940929	ver. axi.	I O B F iX I O B F iX	- - - - - - - - - -	1.21	5.13	2.03
59	July/83	950228	ver. axi.	- - - - - - O - - iX	I - - - - I - - F -	0.86	0.87	1.66
64	July/83	941124	ver. axi.	I O B F iX - - - - iX	- - - - 2X I - B F 1.2X	0.97	2.26	2.54
65	July/83	940721	ver. axi.	- O B F iX - O B - iX	I - - - 2X I - - F 2X	0.93	1.29	1.17
72	July/83	941208	ver. axi.	I O B F - - - - - -	- - - - iX I OB F iX	1.58	6.55	2.85
75	July/83	941226	ver. axi.	- - B - iX - - - - -	I O - F - I OB F 2X	2.94	2.66	2.94
80	July/83	941027	ver. axi.	- O - F iX - - - - -	I - B - 2X I OB F iX	2.07	6.64	4.35
85	July/83	950228	ver. axi.	I O B F iX - - B F iX	- - - - - I O - - 2X	0.72	1.53	1.32
95	July/83	950227	ver. axi.	- O B F iX - - - - -	I - - - 2X I O - F i.2X	0.91	1.31	3.14

Table 1. Replaced bearings information, ver. and axi. are vertical and axial respectively

The vibration levels of BPFI, BPFO, BSF, FTF, X and their higher multiples are given in the vertical and axial directions, and are classified into high and low. Undetectable frequencies are denoted by (-) and the detectable frequencies which acquired high or low levels have been denoted by I, O, B, F and jX, corresponding to BPFI, BPFO, BSF, FTF and X, respectively, for j=2, 1.2, i, i.2. Vibration levels larger than 0.3 mm/s are considered high and those less than 0.3 low. Low vibration levels are considered high if other bearing defect frequencies are also detectable in the spectrum. The analysis of the vibration spectra history of these ten

bearings reveals that it is possible to indicate changes in the bearing condition at an early stage when using bearing defect frequencies, their higher multiples and the combined frequencies. This use can be improved if the following conditions are met: (1) Resolution line should be<1 Hz. Its exact value depends on machine speed. (2) The interval between measurements should not be too long to avoid missing defect initiation and development. (3) Vibration measurements after maintenance actions, especially renewals, are most important to control the action quality, identify defects and their causes and follow their development easily. (4) The use of the frequencies representing the sum or difference of some of the bearing defect frequencies (BPFI, BPFO, BSF, FTF) or their multiples, e.g. (BPFI+2*FTF) or (BPFO+2*BSF), is reliable for detecting defects especially when it is impossible to recognise bearing defect frequencies and their multiples. (5) For an effective diagnosis, variations in bearing defect frequencies should be considered.

Bearings at the leading rollers of the drying cylinders suffer high ambient temperature which reduces the lubricant viscosity and leads to a reduced thickness of the oil film and severe wear due to metal-to-metal contact. This caused more frequent replacements of the bearings and is noticed at all leading roller bearing positions. A combination of this phenomenon and high axial vibration level was (probably) behind the failure involving looseness of the tapered clamping sleeve in one of the leading rollers. This is why the company considers changes in either the machine construction or in the lubricant system to be necessary. Several multiples of the machine speed can be found in almost all the spectra indicating: (1) Waviness in the inner and/or outer races, (2) Rotation of the outer ring in the bearing housing or of the inner ring on the shaft, and/or (3) Bearing misalignment due to faulty installation of the bearing or bent shaft. The third reason is the most probable due to existence of high vibration levels at 2*X in both radial and axial directions in almost all the examined spectra, when no coupling misalignment is recorded. The "bent" shaft is probably hogged be excessive thermal expansion which results in misalignment in the bearing. This is a design fault [10]

2.2 Case study B

In the database of the VBM program at company B, three types of spherical roller bearing, which are usually used in the drying cylinders, were selected for deeper analysis due to their large number of replacements. These bearings are 23052 cck/SKF at the driven side (DnS) and 23060/ HA3C4V33/SKF, (was replaced by 23060cck/C4S3V33/ SKF)

at the driving side (DS), which have the same defect frequencies. 49 bearings of each type are considered in this study. The replacements are divided into 5 generations. 56 bearings, 20 at DnS and 36 at DS, have run since their installation in Aug.1977, i.e. 221 months. There is only one bearing position which has experienced 4 replacements, two have experienced 3, 16 have 2 replacements and the rest have one replacement during the period Aug.'77-Jan.'96. They started keeping the vibration measurements in November-1992. The measurements taken straight after installation are available only for the bearings the drying cylinders number (11, 12, 15 and 32)/DnS, and 27/DS because they were installed during the last 3 years. The vibration measurements cover frequency range 0-300 Hz, with 400 resolution lines. The comparison between the first two replacement generations at both DS and DnS is given in Table 2. Notice that the maximum and minimum life lengths reveal the extreme limits of these groups. The number of replaced and unreplaced bearings, replacement date and position are stated in Table 3. The replacements which are performed at DnS is 220% more than that performed at DS.

	DnS, 1st generation	DnS, 2nd generation	DS, 1st generation	DS, 2nd generation
Number of replacements, s	29	13	13	6
Number of bearings, n	49	29	49	13
Total life length, months	3165	892	1413	264
Average life length, months	109.14	68.6	108.7	44
Sample standard deviation, months	45.5	50.6	47.6	44
Maximum life length, months	207	164	180	108
Minimum life length, months	28	2	5	6
The number of bearings, k, which have life length < 60 months	2	6	1	4
(k / s) 100%	7%	46%	8%	67%

Table 2. Comparison between two bearing replacement generations at PM01/DS & DnS

Bearing type	Side	Bearings n	Replace-ments, s	Unrepla-ced, m	m/n	s/n	s/(n-m)	Time interval
23052 cck	DnS	49	44	20	0.41	0.90	1.52	Aug.'77-an.'96
23060 cck	DS	49	20	36	0.73	0.41	1.54	Aug.'77-an.'96
Total		98	64	56	0.57	0.65	1.53	

Table 3. Bearing replacements at drying cylinders/ PM01.

The number of replacements at DS and DnS during the same period for the first two generations are plotted against bearing life length, see Figs 2, 3, 4 and 5. In the plots, the time intervals are of 12 months. These 4 groups of replacements are, however, distinguishable from each other. Each group

clearly exhibits several modes, each of which probably represents a different damage syndrome, but the records are not clear on this.

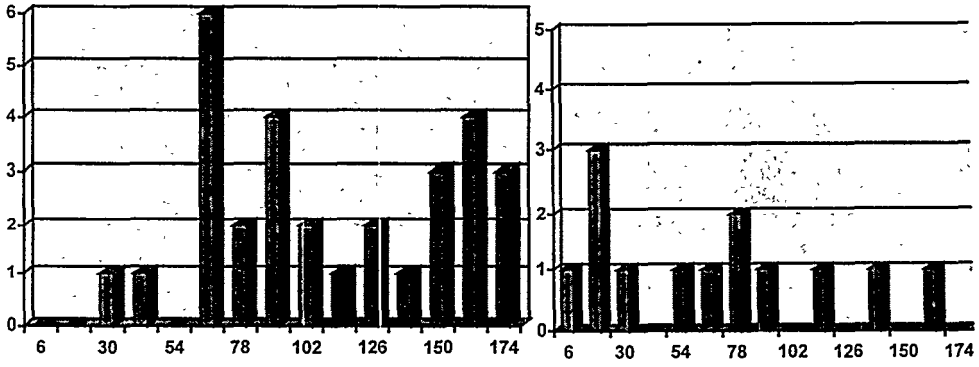


Fig.2 & 3. Replacements at DnS/1st and 2nd generation (the life in months).

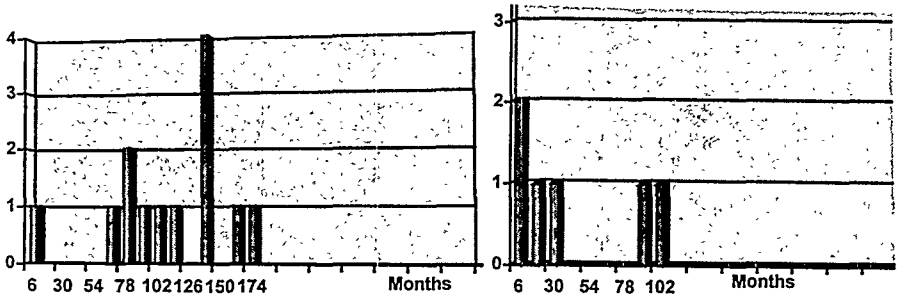


Fig.4 & 5. Replacements at DS/ 1st and 2nd generation (the life in months).

2.2.1 Bearing life length and spectra analysis

The 1st bearing generation was installed at Aug.'77 and the installation of the 2nd bearing generation was at different opportunities, i.e. failures or assessed conditions. The replaced bearing condition may vary between the slightly damaged and the completely deteriorated because the assessment is usually subjective. In many cases, the quality of bearing installation could not be evaluated because there were no vibration measurements straight after installation. Some bearings have run without replacement for over 221 months at DS and DnS. The maximum is about 207 and 180 months for the 2nd generations at DnS and DS, respectively.From Table 2:

1. The average life of the 1st generation/DnS is longer by about 59% than that for the 2nd generation/ DnS.
2. The average life of the 2nd generation/DS is only about 40% of the 1st generation/DS due to one or more of the following reasons: (a) Faulty installation. (b) A design fault, which was passive during first replacement generation period, or induced faulty construction due to

some constructional changes in the machine. (c) Changes in operating conditions such as rotational speed, loading and temperature.
(d) Misuse of bearings, e.g. overloading, high felt tension. (e) Faulty service, e.g. excess of grease, unsuitable lubricant, pollution in the lubricant. Defects initiated in DS bearings due to one or more of these reasons may develop faster because of, e.g. higher loading and/or temperature [10].

3. The average life of the 2nd generation/DnS is longer than that experienced at the 2nd generation/DS by about 56%.
4. Maximum lives of the 1st generations/DnS and DS are longer by about 26% and 67% respectively than 2nd generation/DnS and DS.
5. The lowest and highest sample standard deviations, 44 & 50.6 months, occurred in the 2nd generations at DS and DnS respectively.

Assume that the bearings of life lengths less than or equal to 60 months are those exposed to some of the above mentioned causes of rapid deterioration. Faulty installation can seriously decrease bearing life and make the replacements much more frequent, (about 3 to 5 times) during the final stage of operation[9]. Numbers of such bearing are 6 and 4, i.e. 46% and 67%, in the 2nd generations/DnS and DS, respectively. It is possible to recognise from the histograms that there exist four modes which may be two failure modes, probably abrasive wear and sub-surface fatigue, plus two replacement policies, probably replacement when the vibration level first deviates from normal and replacement at higher levels, i.e. failures or UPBFR. Lack of information regarding censored and full failure data and the replaced bearing condition, made statistical reliability analysis impossible.

The vibration spectra of 13 bearings installed at drying cylinders at both DnS and DS were selected for deeper analysis. These bearings were selected because there were vibration measurement records covering most of their lives, including measurements taken straight after installation. The analysis results can be summarised by:

(1) Using 400 resolution lines for a range of 300 Hz made the diagnosis easier, but, the assessment of badly damaged bearings possibly generating frequencies above 300 Hz became difficult. (2) Some of the bearings had detectable vibration frequencies in the measurements straight after the installation. Later measurements revealed defect developments. (3) In many spectra it was not difficult to identify multiples of FTF. This probably occurred because the cage and the rollers were strongly squeezed due to thermal hogging [10]. (4) In many cases, the overall vibration level increased appreciably when several harmonics of the machine speed and bearing defect frequencies became detectable [10]. This phenomenon is

noticed in many spectra which have high vibration levels in both radial and axial directions. (5) Variations in bearing defect frequencies were detectable in almost all the cases. (6) When the frequencies BPFO and BPFI or their multiples are detectable it is possible to find a side band, whose frequencies are modulated by 1*X or higher [10]. (7) In some cases, the vibration levels of several defect frequencies were higher than the levels at which these bearings were replaced. (8) At almost all the analysed spectra, no multiple of the machine speed was identified when it was 219 RPM. The reason for this may be that the machine speed was not correctly recorded.

3 GENERAL COMMENTS AND CONCLUSIONS

The study results were presented and discussed with the maintenance staff including the analysts and technicians responsible for vibration monitoring. The objectives were to establish changes in the measurement, analysis, diagnosis and replacement policies based on these results.

The importance of analysing variations in bearing life length may be considered from different aspects, e.g. operational safety, product quality, maintenance cost, production losses, which may be divided into; economical, technical and organisational categories. A high sample standard deviation in bearing lives translates, in general, into higher proportion of failure cycles, which increases maintenance costs [11], unless the modes can be separated by better data discrimination and records.

Collected data and analysis results lead technically to the following:
(1) There is no clear vibration level-based replacement policy. (2) Insufficient on-renewal vibration spectra made the identification of failure causes less certain. (3) The 1st bearing generations survived longer than 2nd bearing generations in spite of improvements in bearing manufacturing due to TQM. No data concerning bearing replacements from company A were received. The speed of PM10 as measured varies from 489 to 547 RPM. This made following bearing defect frequencies impossible when using Palogram in Prism^2 because they are functions of RPM. At company A, 400 resolution lines are used for 800 Hz range, which made it sometimes impossible to differentiate between close frequencies. Deficiencies in data coverage and quality prevented the identification of all the actual vibration levels at failures and other renewals, which would be necessary for statistical analysis and optimisation. In order to improve the condition monitoring system, we

suggest: (1) Apply the study results. (2) Investigate closely the operating conditions of the bearings of shorter life in order to identify and eliminate the real causes. (3) Establish a clear and written vibration level-based replacement policy based on economics [11]. (4) Operating conditions data are useful to discover reasons behind vibration level variations. (5) A record describing bearing damage is important to correlate vibration history with damage found on renewal [11]. (6) Analysis of complete data (vibration history, working conditions, damage on renewal) for long-lasting bearings may lead to improvement of other bearings' lives, establishment of deterioration models and better economics.

ACKNOWLEDGEMENT

The study is sponsored by Stora Hylte AB and Stiftelse Svenskt Underhållsteknik Centrum (UTC/FoU, project 9571).Authors are grateful to the maintenance staff at the companies and to Professor David Sherwin.

REFERENCES

[1] Collacott, R.A. Mechanical Fault Diagnosis and condition monitoring. Chapman and Hall, London., 1977.

[2] Tranter, J. The fundamentals of, and the application of computers to, condition monitoring and predictive maintenance. Inter. Machinery Monitoring and diagnostics Conference 1 Las Vegas, USA 1989.

[3] Lai, M.S. and Reif, Z. Prediction of ball bearing failures. Ist International Machinery Maintenance Diagnosis Conference and exhibit 1989.

[4] Bloch, H.P. and Geitner, F.K. Machinery failure, analysis and troubleshooting. Gulf Publishing Company, London, 1994.

[5] Al-Najjar, B. A new vibration based approach for monitoring rolling element bearing. Nordic Conf. on Vehicle and Machine Vibration 1994, Sweden.

[6] Al-Najjar, B. On the effectiveness of vibration-based programs. ISSN 1400-1942, ISRN HV/MASDA/SE/R/--9581--SE, Växjö, Sweden, April (1996I).

[7] Al-Najjar, B. TQMain; A common database for total Quality maintenance, An approach for continuous reduction in costs of quality products. Journal of Quality in Maintenance Engineering, 2-20, Vol 2, Number 3, (1996III).

[8] Lipovszky, György, S. K. and Varga, G. Vibration testing of machines and their maintenance. Elsevier, Amsterdam, 1990.

[9] Barkov, A. and Barkova, N. Condition assessment and life prediction of rolling element bearings-Part 2. Sound and Vibration, 27-31, September 1995.

[10] Al-Najjar, B. Improvement in effectiveness of vibration-based monitoring systems in paper mills, (submitted for publication)

[11] Al-Najjar, B. Criteria for the selection of a cost-effective maintenance policy. ISSN 1400-1942, ISRN HV/MASDA/SE/R/--9585--SE, Växjö, Sweden, May (1996II).

INFLUENCE OF RANDOM OUTER LOAD CHANGE AND RANDOM GEARING FAULTS TO VIBRATION DIAGNOSTIC SIGNAL GENERATED BY GEARBOX SYSTEMS

Walter Bartelmus
Technical University of Wrocław Faculty of Mining Machinery
System Division
Wrocław Poland

ABSTRACT
The paper deals with the investigations of influence of an outer load to diagnostic signals generated by a meshing of a gear. The investigations are done by computer simulations. Different outer load patterns and different patterns of gearing conditions described by error functions of a gearing are investigated. Trials of elimination of influence of outer load patterns to diagnostic signals are presented.

1 INTRODUCTION

Mathematical modelling and computer simulation gives new possibilities of investigations of vibration diagnostic signals generated by a gearing of gearbox systems for aiding diagnostic inference. In the paper [1] a mathematical model for gear systems is presented. The model, after further development, was used for investigations, using computer simulations, of unstability of gear system. It was published in [2]. The results shone in [2] were compared to the results obtained by measurements, taken by strain gauges attached to a pinion of a gear. The results of measurements are presented by Rettig H., in [3]. Signals taken from strain gauges gives direct evaluation of gearing condition co-operation, so does computer simulation. These are very expensive measurements which are done for reference, scope of the experiments presented in [3] was very limited. Good consistence of results obtained by computer simulations, paper [2], and measurements, paper [3], gives the base for replacement further measurements by computer simulations. The results presented in [2] were also compared to measurements presented in paper [4]. Simulations may be done for any conditions

which may change the vibration diagnostic signal. The gear conditions under which investigations were done by Rettig were limited to change of operation caused by change of rotation speed of gears. The gear conditions may be described, as stated in [4], by four groups of factors: design, production technology, operation, change of gear condition. The very important problem in diagnostic of gearboxes is influence of outer load random change (operation factor) to diagnostic signal. In the proposed paper this problem will be considered. For elimination of influence of the random outer load change normalisation of the signals is taken. The influence of different reference value for normalisation will be presented. Very interesting problem in evaluation of gear condition is visualisation of diagnostic signal at the condition of broken or a partly broken gear tooth (condition change factor). This problem is presented in paper [5]. The development of condition change of gearing caused by random change of errors of the teeth of a gear to the diagnostic signal is developed in presented paper. The increased errors are caused by wear and pitting. During the computer simulations any choice of gearing errors may be taken and relation between gearing errors and visualised signal may be obtained. The aim of the paper is to present the possibility of increasing the expert knowledge on the signals generated by a gearing of gearbox systems at different conditions, caused by random teeth errors and by outer load random change, on the basis of computer simulations.

2 RANDOM LOAD AND RANDOM ERRORS DESCRIPTION

Inter teeth load of a gearing can be considered at acceleration of the system and at the normal operation. The inter teeth load at acceleration of the system depends of an electric motor characteristic and a system inertia. The acceleration load mode is given by Ma [Nm] moment which may be constant Fig.26. or variable. So the acceleration load mode is described as Ma(const/var). Outer load of a gear system at the normal operation is described by Ml(w;r;pw;rp;kr) where: w - coefficient of load changeability (0 -1); r - coefficient of random load change scope; pw - coefficient of load shape (0 -1); rp - coefficient of random load shape scope; kr - coefficient of load periodic changeability. In Fig.1. the load mode Ml(1;0;0.9;0;8) is given; where Tl[s] period of the load change, Mr - reference of outer load [Nm]. The coefficient of load periodic changeability is ratio of Tl/T=kr, where T - period of meshing of a gearing. In Fig.2. the load

mode Ml(1;0.3;0.9;0.3;8) is given. It means that load peaks for r=0.3 changes 30% and the place of the load peaks changes 30% for rp=0.3.

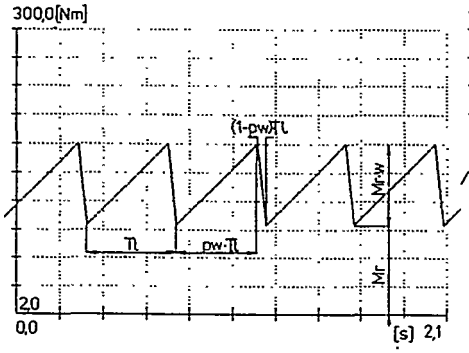


Fig.1. Load function for Ml(1;0;0.9;0;8)

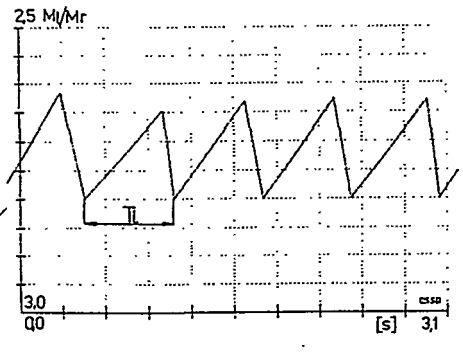


Fig.2. Load function for Ml(1;0.3;0.9;0.3;8)

A value of Mr peak is given: Mr (peak) =Mr + [1 - r*(1 - li)*Mr;
* - multiplication sign; li - random value (0 - 1).
Place of peak is given by: pw(random) = [1 - r*(1 - li)*pw]
The errors in meshing are described by the error mode E(a;e1;r;ra); where a,e1 parameters of the error mode which describe place of an error peak (0 - 1) and e1 max. value of error [μm]. The parameters r and ra (0 - 1) describe the scopes of a random changeability and are given by:
e(random) = [1 - r*(1 - li)]*e1.
The place of a random value of error is given
a(random) = [1 - ra*(1 - li)*a.

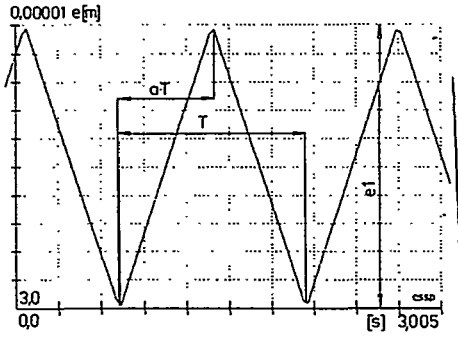


Fig.3. Error function for E(0.5;10;0;0)

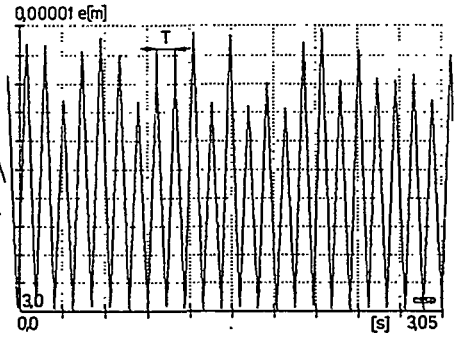


Fig.4. Error function for E(0.5;10;0.3;0)

In Fig.3. the error mode E(0.5;10;0;0) is given, and in Fig.4. the error mode E(0.5;10;0.3;0) is given; where T - meshing period.

3 NORMALISATION OF DIAGNOSTIC SIGNAL

Inter teeth forces can be presented as a ratio Kd=F(t)/Fr; where F(t) - current inter teeth force Fr - rated inter teeth force. Kd is called the dynamic coefficient. The current force is normalised by the rated inter teeth force. For assessment of the meshing condition change current inter teeth forces, caused by change of a mesh condition, ought to by measured. The change of a mesh condition can be descried by the error mode E(a;e1;r;ra). The given error mode generate specific F(t). When outer load of the gearbox is constant, F(t) gives the signal of the mesh condition change. But to the signal F(t) has influence an outer load given by the outer load mode Ml(w;r;pw;rp;kr). One of the most impotent problem for condition monitoring is reduction of influence of the outer load mode to the diagnostic signal. For the reduction the diagnostic signal is normalised in the form of ratios Kd1=F(t)/F1(t); where F1(t) - measured current force on an input shaft of a gearbox. One can make other ratios: Kd2 =F(t)/Fl(t); where F1(t) outer load force given by the outer load mode Ml(w;r;pw;rp;kr); Kd3 =F(t)/F2(t); where F2(t) measured current force on an output shaft of a gearbox; Kd4=F(t)/Fs, where Fs force counted from power consumption by an electric engine. The functions for Kd; Kd1;Kd2;Kd3;Kd4 are given in Figures; Fig.5., Fig.6., Fig.7., Fig.8., Fig.9..

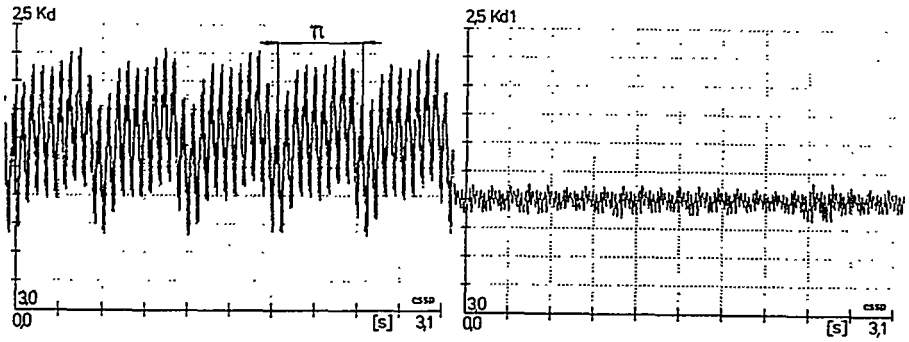
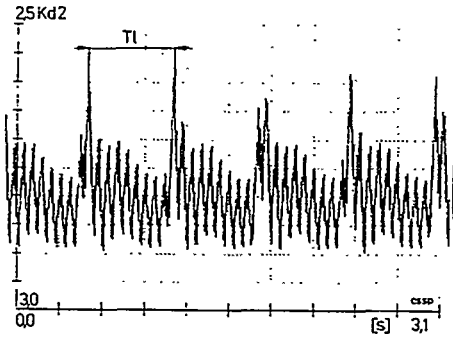


Fig.5. Function of Kd for Ml(1;0;0.9;0;8) and E(0.5;10;0;0)

Fig.6. Function of Kd1 for Ml(1;0;0.9;0;8) and E(0.5;10;0;0)

As it is seen from Fig.5. to Fig.9. the lowest influence of the load mode to the signal is seen in Fig.6. for Kd1. Periodic influence of the outer load to other functions is seen in all Fig.5 to Fig.9 but Fig.6. As an unsuitable functions for the diagnostic assessment of the mesh condition are function given in Fig.7. and Fig.8..



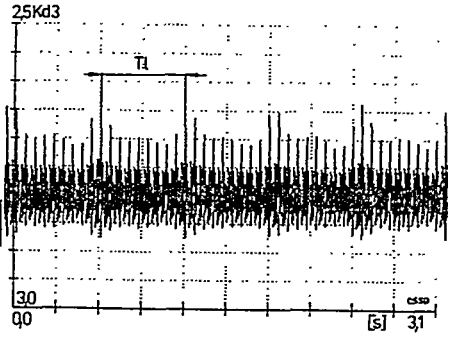
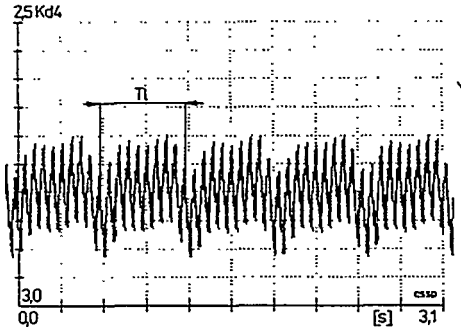


Fig.7. Function of Kd2 for Ml(1;0;0.9;0;8) and E(0.5;10;0;0)

Fig.8. Function of Kd3 for Ml(1;0;0.9;0;8) and E(0.5;10;0;0)



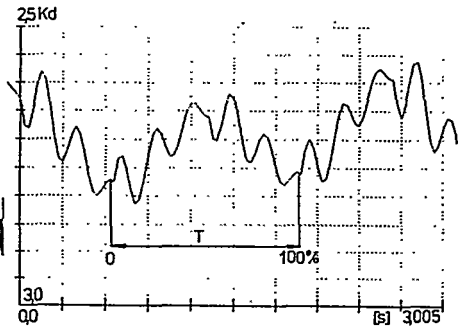


Fig.9. Function of Kd4 for Ml(1;0;0.9;0;8) and E(0.5;10;0;0)

Fig.10. Zoom of Kd function for Ml(1;0;0.9;0;8) and E(0.5;10;0;0)

The zoom given in Fig.10. is equivalent to the results given in the paper [3].

4 INFLUENCE OF RANDOM LOAD AND RANDOM GEARING FAULTS

As it was mentioned the random load and the random faults of gearing are described by the load modes Ml(w;r;pw;rp;kr), Ma(const/var) and the error mode E(a;e1;r;ra). In Fig.11. the error

function for the error mode E(0.5;10;1;0) is given. This error mode describes the condition of a very deteriorated gear mesh. The Kd function for the error mode is given in Fig.12. This result is taken for the load mode Ml(0;0;0;0;0). it means that the outer load is content end equals to the rated load.

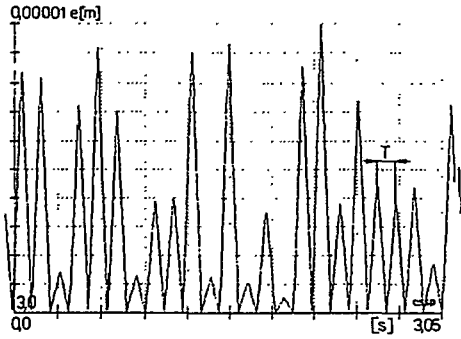


Fig.11. Error function for E(0.5;10;1;0)

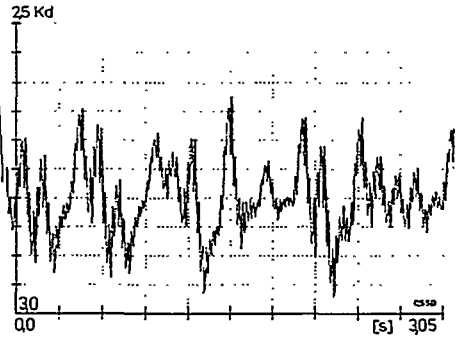


Fig.12. Kd function for E(0.5;10;1;0) and Ml(0;0;0;0;0)

The function of Kd1 under mentioned state of load Ml(0;0;0;0;0) and condition of the mesh E(0.5;10;1;0) is given in Fig.13.

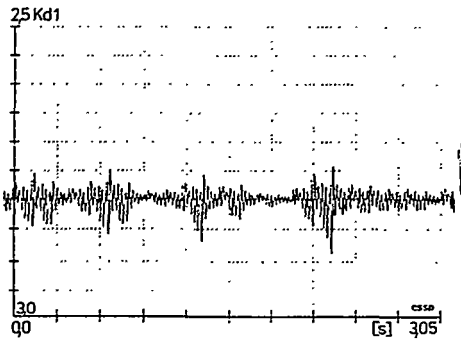


Fig.13.Function of Kd1 for Ml(0;0;0;0;0) and E(0.5;10;1;0)

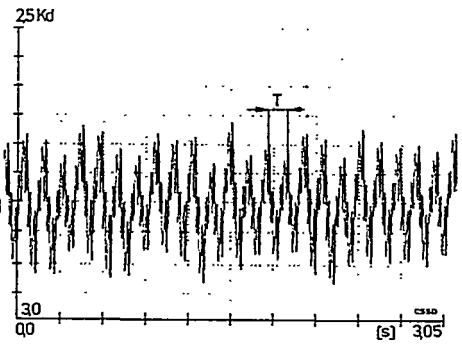


Fig.14. Function of Kd for Ml(0;0;0;0;0) E(0.5;10;0.3;0)

For a medium deteriorated gearing mesh its condition may be described by E(0.5;10;0.3;0) look Fig.4., function Kd under this condition and the load mode Ml(0;0;0;0;0) is given in Fig.14. The

function of Kd1 is given in Fig.15. If you compare Fig.13. to Fig15 and Fig.12. to Fig.14. the influence of r for error mode is seen.

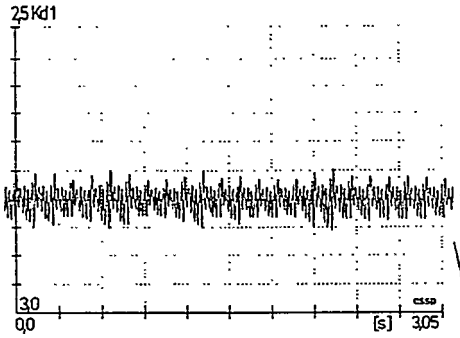


Fig.15.Function of Kd1 for M1(0;0;0;0;0) E(0.5;10;0.3;0)

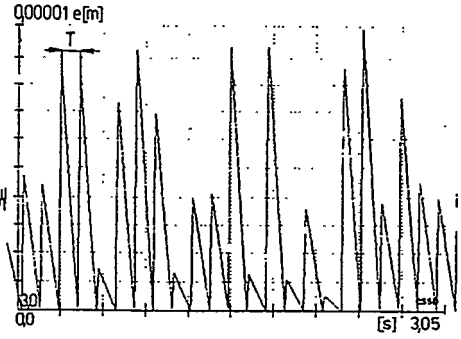


Fig.16. Error function for E(0.1;10;1;0)

If one change the error mode to the form E(0.1;10;1;0) the error function is as it is given in Fig.16. By this error function one may describe the condition of co-operation of a mesh when the gearing is very deteriorated and a supporting bearing is also in a bad condition.

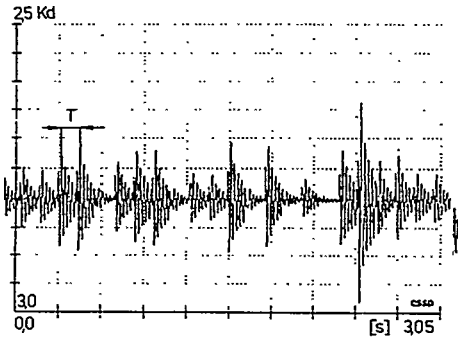


Fig.17. Function of Kd for M1(1;1,0.9,0.3;8) E(0.1;10;1;0)

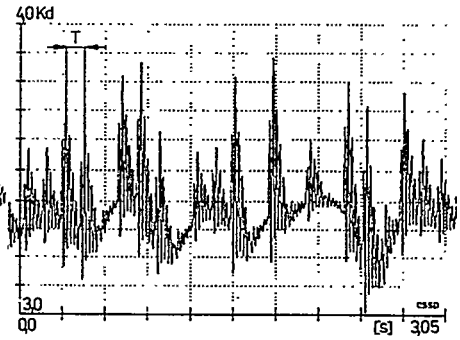


Fig.18. Function of Kd1 for M1(1;1,0.9,0.3;8) E(0.1;10;1;0)

Fig.17. and Fig.18 give the results of computer simulations under the outer load given by M1(1;1,0.9,0.3;8) end the gearing condition given by error mode E(0.1;10;1;0). It easy to see the influence of error mode parameters , compare Fig.13. to Fig.17. or Fig.12. to Fig.18. Further

computer investigations of the diagnostic signals shows good properties of evaluation of the gearing condition by investigation the signal during acceleration of a gearbox when state of load is described by Ma(const/var). During computer simulations the load mode was given by Ma(const). It means a constant electric motor moment [Nm] which drives the gear systems during the rotation acceleration of the system, Fig.26.. The system change rotation from 0 to rated rotation [rpm].

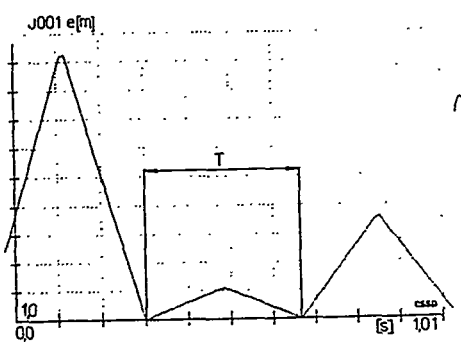


Fig.19. Error mode function for E(0.5;10;1;0.1) and Ma(const)

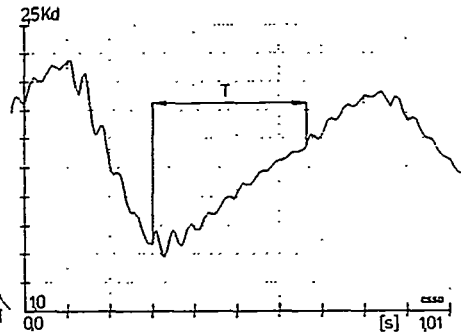


Fig.20. Function of Kd for E(0.5;10;1;0.1) and Ma(const)

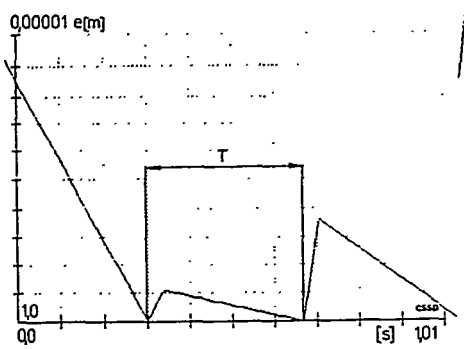


Fig.21. Error mode function for E(0.1;10;1;0;0) and Ma(const)

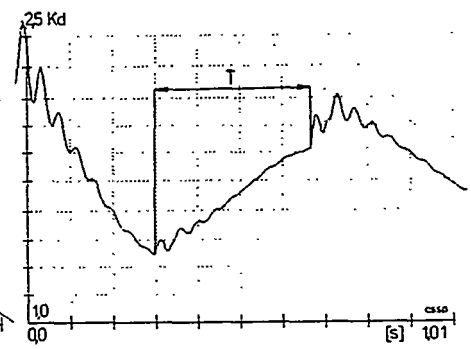


Fig.22. Function of Kd for E(0.1;10;1;0;0) and Ma(const)

Fig.19. and Fig.20. together with Fig.21. and Fig.22 show relation between the gearing condition given by the error modes E(0.5;10;1;0.1) and E(0.1;10;1;0.1). The change of a - parameter from a = 0.5 to 0.1 may be evaluated, look Fig.20. and Fig.22. More

enhanced relation between the condition of the gearing given by error mode E(0.1;10;1;0) and diagnostic signal given by Kd is shown in Fig.23. and Fig.24. These results of simulations where taken for load mode Ma(const). The load mode under which the most of simulations were done is given in Fig.25. for which the load mode is given Ml(1;1;0.9,0.3;8). The Function of an electric motor driven moment Ms [Nm] is given in the Fig.26.

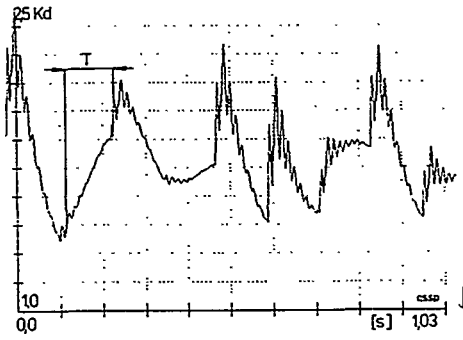


Fig.23. Error mode function for E(0.1;10;1;0)

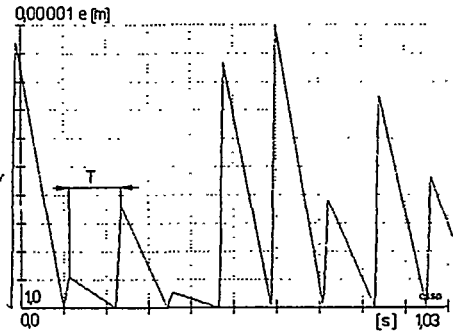


Fig.24. Function of Kd for Ml(1;1;0.9,0.3;8) E(0.1;10;1;0)

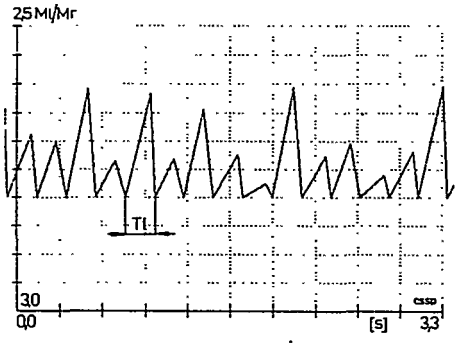


Fig.25. Function of load mode for Ml(1;1;0.9,0.3;8)

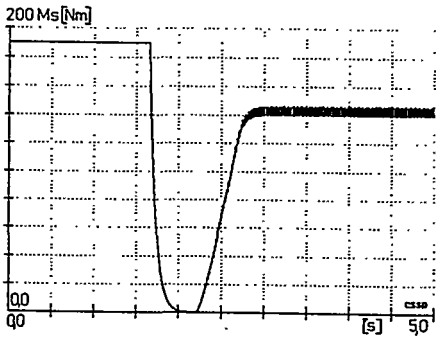


Fig.26. Function of electric motor moment for Ma(Ms=const) and Ml(1;1;0.9,0.3;8)

In the Fig.26 one may see the motor moment during the acceleration of the system Ms=2*Mr, Mr=95Nm, the electric motor moment Ms under condition of free rotation when Ms=0 and under condition when the system is loaded by the outer load Ml(1;1;0.9,0.3;8).

5 CONCLUSIONS

The paper presents relation between the random error mode of gearing and diagnostic signals. The assessment of the gearing condition may be done by examine time functions obtained during rotation acceleration of a gearbox system. Compare Fig.19. to Fig.20 and Fig.21. to Fig.22. or Fig.23. to Fig.24. The relation between the random error mode and diagnostic signal of inter teeth forces during quasi study rotation of the gearing is also given. Compare Fig.16. to Fig.17. and Fig.18. or Fig11 to Fig.12. and Fig.13. or Fig.4. to Fig.14. and Fig.15.. By proposed normalisation one may eliminate the influence of outer load to the diagnostic signal. The normalisation has to be done in the form Kd1=F(t)/F1(t), where F(t) current inter teeth force, F1(t) current force measured at an input shaft of a gearbox. Finely you ought to mind that measurements of inter teeth forces are difficult so we have to measure accelerations of the gearbox walls but there is a good relation between inter teeth forces and the accelerations of gear wheels support The relation will be given during presentation of the paper.

REFERENCES

[1] Bartelmus W., Computer Simulation of Vibration Generated by Meshing of Toothed Wheel for Aiding Diagnostic of Gearboxes, Conference proceedings Condition Monitoring '94 Swansea, UK PINERIDGE PRESS Swansea, UK

[2] Bartelmus W., Diagnostic Symptoms of Unstability of Gear Systems Investigated by Computer Simulation. Proceedings of 9th International Congress COMADEM 96 Sheffield July 1996

[3] Rettig H.; Innere Dynamische Zusatzkrafte bei Zahanadgetrieben Ant. Antriebstechnik 16(1977) no. 11 s 655-663

[4] Bartelmus W., Vibration Condition Monitoring of Gearboxes, Machine Vibration (1992) 1: 178-189 Springer-Verlag London Limited

[5] Bartelmus W., Visualisation of vibration signal generated by gearing obtained by computer simulation. Proceedings of XIV IMEKO World Congress (in print)

[6] Siwicki I., Manual for CCCP WARSAW 1992

DIFFICULTIES IN CONDITION MONITORING OF SLOWLY ROTATING ROLLING ELEMENT BEARINGS

Henry Mikkonen
ABB Service Oy, Helsinki Finland

ABSTRACT

The condition monitoring of slowly rotating machinery has recently become a considerable problem in the process industry. These machines are often critical for the process, which means that an unexpected breakdown can cause huge expenses. The present equipment allows to monitor the condition of the machines that rotate relatively quickly very effectively, which is why the mills are unwilling to shut the process down for the sake of the periodic inspection of the slowly rotating machinery. However, making these inspections is often the only reliable method to make sure that the machine will work until the next shut-down.

In this paper it the principles of why monitoring of slowly rotating bearings is difficult is discussed. Also some guidelines and recommendations on how to measure them are given.

1. INTRODUCTION

It has been calculated that in Finland, slowly rotating machines cause annually hundreds of millions of US$ of maintenance costs. Utilising right methods for different rotating conditions these costs can dramatically be reduced.

International standard on condition monitoring ISO 2372 [1] covers the machines with rotating speeds from 10 to 200 rev/s. The condition monitoring of machines outside this area is difficult, firstly, because there is no reliable information or recommendations about acceptable

vibration levels available. Secondly, the present vibration monitoring systems are mainly designed for monitoring machines that work at the standardised speeds only.

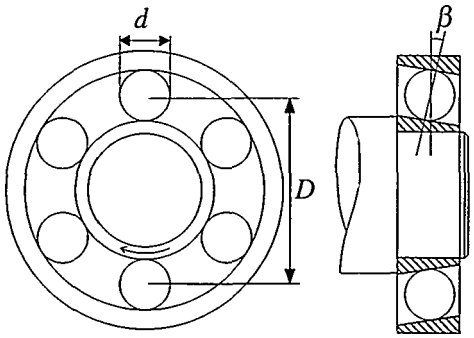
2. VIBRATION OF A ROLLING ELEMENT BEARING

A rotating rolling element bearing generates forced vibrations. These vibrations are generated for example because of elastic deformations, surface roughness and inaccuracies in manufacturing. They are discussed extensively in Ref. [2]. These vibrations are mainly weak but sometimes, especially in a heavily loaded condition, a ball or roller passing frequency of the outer ring (see fig. 1) can be seen in vibration measurements, even if the bearing is in faultless condition. This vibration results from deformation of the construction, which occurs when the rolling element moves into and out of the loaded zone.

A faulty bearing can vibrate at following frequencies [3]:

- random, ultrasonic frequencies,
- natural frequencies of bearing components (and other natural frequencies),
- rotational defect frequencies and
- sum and difference frequencies.

Normally the existence of these bearing characteristic frequencies is considered a sign of a defect in the bearing. The most common way to detect bearing faults is by studying velocity or acceleration spectra in an attempt to find the vibration components that appear at bearing defect frequencies and especially at their harmonics. These frequencies can be calculated according to following formulas [4, 5].



N = number of rolling elements
n = rotating frequency of the shaft

Ball Passing Frequency of the Outer ring:

$$BPFO = \frac{N}{2}(1 - \frac{d}{D}\cos\beta) \cdot n$$

Ball Passing Frequency of the Inner ring:

$$BPFI = \frac{N}{2}(1 + \frac{d}{D}\cos\beta) \cdot n$$

Ball Spinning Frequency:

$$BSF = \frac{D}{2d}\left[1 - \left(\frac{d}{D}\cos\beta\right)^2\right] \cdot n$$

Fundamental Train Frequency:

$$FTF = \frac{1}{2}(1 - \frac{d}{D}\cos\beta) \cdot n$$

Figure 1. Formulas for calculating bearing defect frequencies, if the outer ring is fixed.

3. EXPERIMENTAL

The results discussed here are based both on laboratory experiments using 'full scale' industrial bearings (shaft diameter 110-120 mm) and on real world measurements from several plants.

3.1 Appearance of faults in measurements

A defect in the bearing generates a shock pulse, when either the rolling element hits the discontinuity on the race or the defect on the rolling element hits the race. The shock wave travels through the boundaries, interacts with other vibration signals present and some of it reflects back from the surface.

As the vibration from a defected bearing is studied further, it becomes evident why the collection of bearing faults especially in slowly rotating machinery is difficult. This can be seen in Fig. 2. where a time-domain signal measured from an outer-race-defected bearing with a single pulse enlarged is shown.

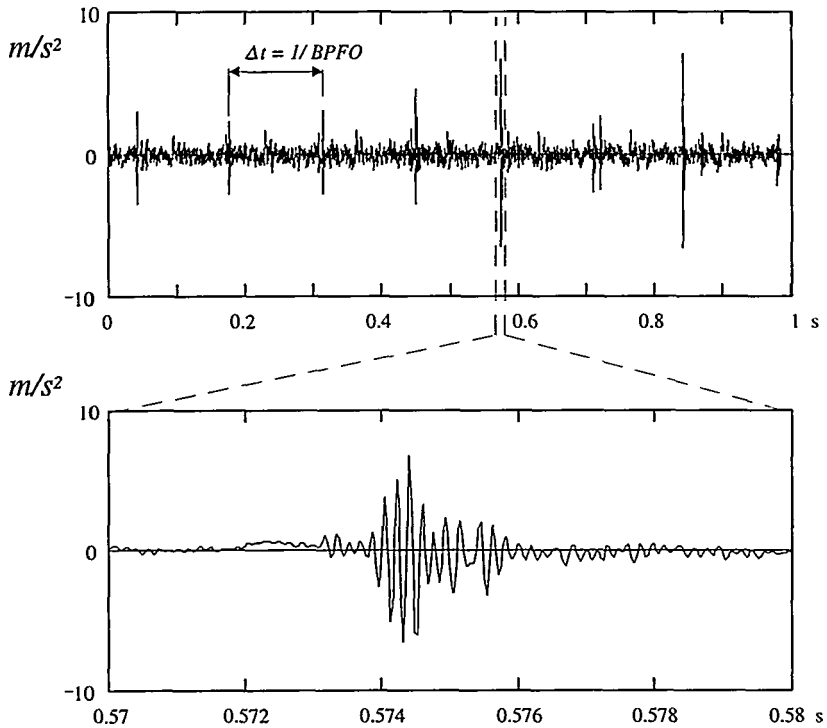


Figure 2. A time-domain signal measured from an outer-race-defected bearing. Frequency range 5 Hz...10 kHz. BPFO = 7.47 Hz [6].

It can clearly be seen that even when the frequency between shock pulses is low, as in this case, the frequency exited by the pulses is high. This vibration of high frequency could appear for example at the natural frequency of a bearing component, bearing housing or other parts of the construction. Also accelerometer mounting resonance or accelerometer resonance itself could be exited by the defect pulses.

Actually the main point is to understand that the frequencies that bearing defects generate are high even if calculated bearing defect frequencies are low. Another example of this is seen in Fig. 3. It displays how filtering affects the shape of the signal. In Fig. 3a) there is a wide band acceleration signal measured in a frequency band from 5 Hz to 10 kHz. The bearing is defected on its outer race. In Fig. 3b) there is a signal measured from the same bearing, but the frequency ranges now between 5 and 650 Hz. The pulses generated by the defect

are now much more difficult to notice than in Fig. 3a). In Fig. 3c) there is again a signal measured from the same condition, but now with the frequency band from 500 Hz to 10 kHz.

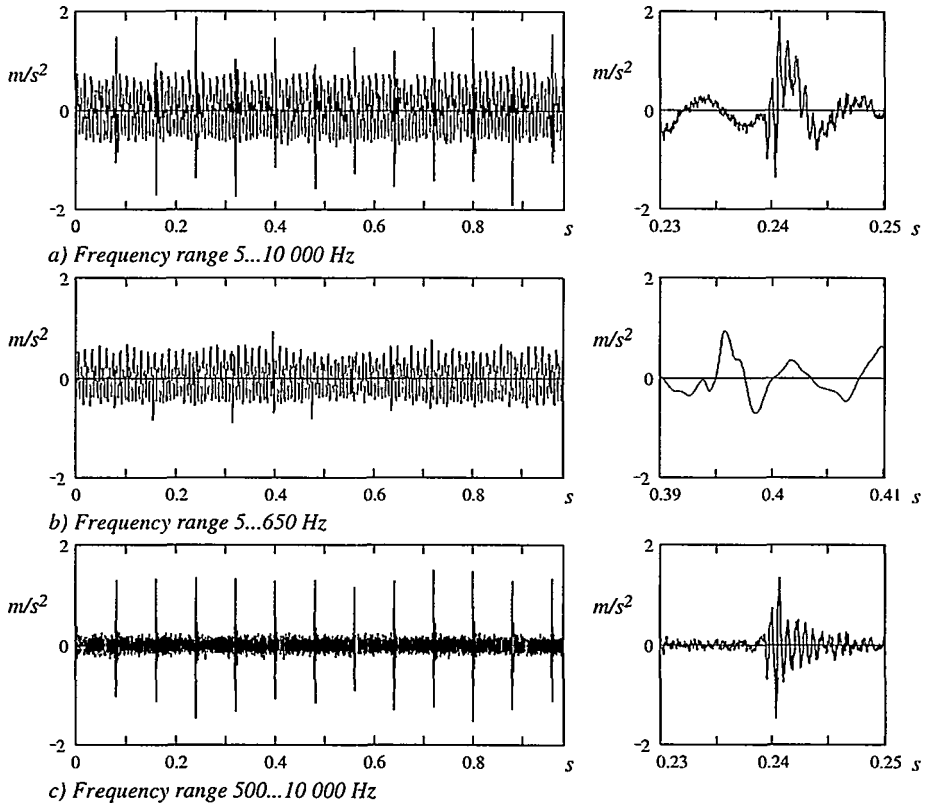


Figure 3. An example of how filtering of data affects the shape of the signal. Time-domain signals measured from an outer-race-defected bearing. BPFO= 12.6 Hz [6].

It is evident that if the cutoff frequency is too low, as in Figure 3b), the information from the defect is filtered out. This is actually exactly what an anti-aliasing filter will do when spectra are measured. For example if vibration is measured according to ISO 2372, the cutoff frequency is 1 kHz, which often filters the most important information out. The trial runs indicated that the upper limit of the frequency range should be at least 2 kHz, if the bearing fault is to be detected at an early stage. The time duration should be selected so that that a time-domain signal contains at least 5 to 10 individual

pulses. If the frequency spectrum is used for the analysis, it is better to either have a signal measured during a longer time or to use averaging. In practice, the normal velocity spectrum has proved not to be an effective tool to monitor slowly rotating rolling element bearings.

The best results were achieved when accelerometer was stud mounted in the bearing housing in the load zone so that the direction of the sensor was the direction of the load. When a measurement was made with these optimal parameters, the result could not be significantly improved through signal processing. If one or more of the parameters differed from the optimal, the improving effect of the signal processing methods was remarkable. This is noteworthy, because in industry, it is rarely possible to make the measurement optimally.

3.2 Difficulties in fault detection as rotating speed is decreased

As proved above it, is important to measure high frequencies, even though the calculated bearing defect frequencies are low. On the other hand, a low defect frequency leads to a long measurement time, for a sufficient number of individual pulses must be measured before the fault can be recognised. Both these conditions will lead to a huge number of samples to be taken. If for example the upper limit of the frequency range is 10 kHz, the sample rate has to be at least 25600 samples/s. If the defect frequency is low enough it is evident that the amount of data cannot be handled by conventional techniques.

Shortly, the difficulties in detecting faults in slowly rotating rolling element bearings are as follows:

1. If the time interval between the defect pulses is long and the pulses are damped rapidly, the effect of the pulses on the frequency spectrum is minor.
2. The time interval between the pulses is not precisely constant due to, for example, the sliding of rolling elements or variation in rotating speed.
3. The power of the defect pulses will weaken as the rotating speed is decreased.
4. Because the sample rate has to be high, the number of samples becomes impossible to handle when rotating speed is low enough.

The first two of the problems can be avoided by using the time-domain signal instead of the frequency spectrum. Problem 3 can be diminished if the signal is processed so that the vibrations generated by the bearing defect are highlighted. The best signal processing method proved to be the filtering of acceleration signal by using a proper high-pass filter. The most suitable frequency range according to tests and real-world measurements is from 500 Hz to 10 kHz; it is wide enough to be universal. An interesting new approach to condition monitoring and signal processing is the use of higher order derivatives of acceleration as presented in Ref. [7]. The test results using differentiated acceleration signals were satisfactory.

If the rotating speed is so slow that it is not possible to measure the time-domain signal long enough to record multiple fault pulses, the only way to make analysis is to use special techniques. Envelope analysis is a way of compressing data so that when it is used, the measurement can be made at high frequency range, but the analysis can be performed at low frequencies. Thus defect pulses will not be filtered out by an anti-aliasing filter and the measurement time can be long. Enveloping proved to be an extremely good way to solve problem 4.

A point of concern here is that as cutoff frequency is high, there is a possibility of different resonances that can alter the level of vibrations measured. This makes it impossible to make analysis on bearing condition just relying on overall levels.

4. CONCLUSION

The measurement parameters and signal processing methods have to be selected carefully in order that proper results will be achieved. The bearing defect frequency determines whether the analysis should be made in the time or frequency domain. The following points concern the detection of failures in relation to the rotating speed:

1. When the defect frequencies are high, the failure can be clearly seen in the frequency spectrum, for the time interval between the pulses is short compared with the time that it takes until the pulse is damped. In this case, the fault increases the rms-value

and the frequency components in the frequency spectrum significantly. The number of points to be measured can therefore be relatively small. As the rotating speed is high, the pulses are powerful, which also makes the detection of failures easier. The detection of failures can be made easier, for example by filtering of data, as when slowly rotating machines are handled. However this kind of processing is usually not necessary, because the failure can be detected without it.

2. When speeds are lowered, the time interval between the pulses will be longer and the pulses will not be as powerful as before. Thus the effect on the spectrum components is minor. Further, the upper limit of the frequency range has to be high, which makes the low defect frequencies difficult to be detected. In addition, frequency spectrum is difficult to use, because the intervals between the pulses are not precisely constant due to the sliding of rolling elements. In this case, the pulses can clearly be seen in the time domain, especially if the signal is processed as suggested earlier.
3. If the rotating speed of the bearing is very slow, the pulses can be highlighted as mentioned earlier. However, the detection of bearing faults is impossible with the conventional methods, because single pulses occur so rarely that the number of samples to be measured would be excessive. Envelope analysis allows to measure even very slow bearings effectively.

Even though envelope analysis is a reliable way of monitoring slow machinery it is also effective method of detecting faulty bearings no matter what the rotating speed is. In ABB Service for example at least one envelope measurement is carried out from all bearings under mesurement program to make sure faults are detected in an early stage.

The recommendations presented in this paper are valid only as far as bearing faults that generate frequent impacts are concerned. Failures of this kind are the most common ones [1], thus the observations cover at least 80% of the bearing faults. Some bearing faults do not generate any impacts. The detection of them, especially when the rotating speeds are low, is difficult by using an accelerometer. The only method that seems to be reliable is the use of fixed sensors, which measure the relative displacement of the shaft.

REFERENCES

1. ISO 2372 Mechanical vibration of machines with operating speeds from 10 to 200 rev/s - Basis for specifying evaluation standards. International Standard, 1974. 9 p.
2. Harris, T. A. Rolling Bearing Analysis. Third Edition. New York 1991. John Wiley & Sons, Inc. 1013 p.
3. Berry, J. E. How to Track Rolling Element Bearing Health with Vibration Signature Analysis. Sound and Vibration, 25(1991)11, p. 24 - 35.
4. Palmgren, A. Ball and Roller Bearing Engineering. Copyright 1945 SKF Industries, Inc. Printed by S. H. Burbak & Co. Inc., Philadelphia.
5. Springer, C. W. The Effects of Internal Geometry Changes on the Frequencies Generated by Deep Groove Ball Bearings. Tappi Journal, 73(1990)5, p. 101 - 105.
6. Mikkonen, H. Condition Monitoring of Slowly Rotating Rolling Bearings (in Finnish). Licenciate thesis. University of Oulu. Department of Mechanical Engineering. Oulu 1995. 90 p.
7. Lahdelma, S. On the Higher Order Derivatives in the Laws of Motion and their Application to an Active Force Generator and to Condition Monitoring. Academic dissertation. University of Oulu. Department of Mechanical Engineering. Report 101. Oulu 1995. 88 p.

DIAGNOSIS OF MISALIGNMENT IN FLEXIBLE COUPLINGS

Maurice F. White and Mladen Jecmenica
Department of Marine Engineering
Norwegian University of Science and Technology
Trondheim, Norway

ABSTRACT

Angular or parallel misalignment of shafts with flexible couplings is a common source of vibration in industrial rotating machinery. In this paper a description is given of the dynamic forces and moments produced by a misaligned flexible coupling. Emphasis is placed on double couplings with spacers such as those commonly used for high performance machinery. Results are presented for radial misalignment of membrane and gear couplings, and include examples of the predicted vibration state of a misaligned shaft-bearing system in the time and frequency domains.

1. ANALYSIS OF GEAR COUPLINGS

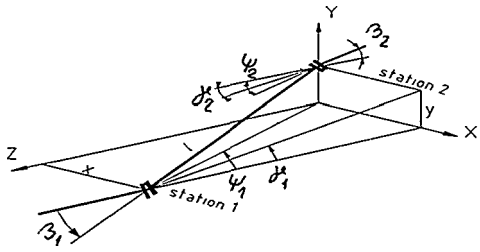
Gear couplings are often used in rotor systems for power transmission in turbines, centrifugal compressors, pumps and other similar high speed machines. Gear tooth clearances and tooth profiles allow misalignment between shafts, and gear couplings can transmit high torque loads while accommodating misalignment.

When a gear coupling is misaligned , it has been suggested in [1] & [2] that primarily three moments are developed :

M_t - due to shaft torque being turned through a misalignment angle,
M_c - due to contact points on gear teeth being displaced axially,
M_f - due to friction between mating teeth.

These moments and the dynamic response they can produce at the bearing locations are described below. A double coupling is represented schematically in Fig. 1, where stations 1 and 2 are locations of the two gear wheels separated by a spacer sleeve of length ℓ. In general both angular and parallel misalignment can occur between input and driven shafts, and

the misalignment angle β does not necessarily lie in the X-Z or Y-Z planes. It is therefore necessary to define forces and moments about the X and Y axes for stations 1 and 2. To simpify the analysis, an angular misalignment of β in the X-Z plane at station 1 is considered here. For the coordinate system in Fig. 1 the moments and rotations are defined as positive in the clockwise direction looking along each axis, where γ and ψ are misalignment angles in the vertical Y-Z and horizontal X-Z planes.



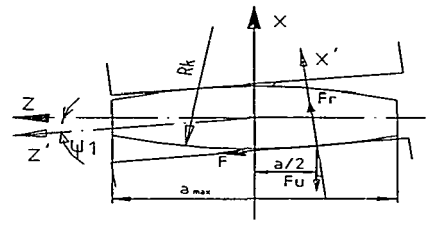


Figure 1 Coordinate system Figure 2 Gear tooth contact

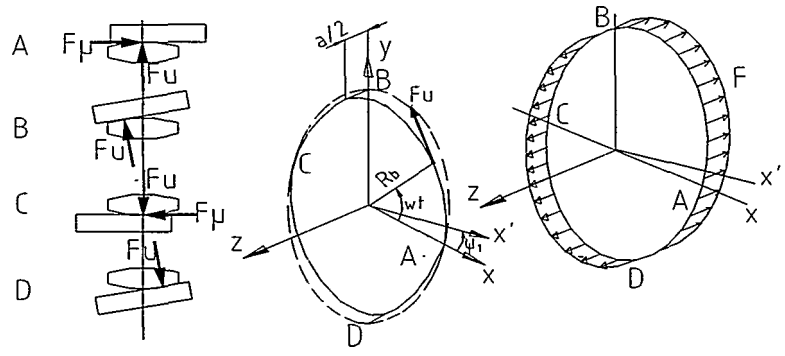
1.1 Moment due to shaft torque turning through misalignment angle

The angular velocity and acceleration of the driven shaft, from [3] is :-

$$\omega_1 = \frac{\omega \cos\beta}{1-\cos^2\omega t \sin^2\beta}; \quad \varepsilon_1 = \frac{d\omega_1}{dt} = \frac{-\omega^2 \cos\beta \ \sin^2\beta \ \sin 2\omega t}{\left(1-\cos^2\omega t \sin^2\beta\right)^2} t \qquad (1.1.1)$$

From Newtons law, the shaft dynamic torque is proportional with acceleration, and hence the moment M_t produced by turning the shaft torque, T, through the misalignment angle, $\beta = \psi_1$ at station 1, has a frequency of two times shaft speed:- $M_t = T \sin\beta \cos 2\omega t$ (1.1.2)

This moment acts about an axis normal to the plane of misalignment.



a) Gear tooth b) Contact point c) Friction force

Figure 3 Forces on a gear tooth due to contact point and friction

1.2 Moment due to axial location of contact points on gear teeth

From Figs. 3a and 3b, the contact point between the gear hub teeth and sleeve teeth will move in the axial direction due to the misalignment and the gear tooth crowning. At point B, the transmitted tooth force F_u (due to shaft torque T) is displaced a distance a/2 in the Z-direction. The moment that resists misalignment due to axial distance between points of contact on a gear tooth is given by eq. 1.2.1.

$$M_c = \frac{T}{Rb} \cdot \frac{a}{2} = F_u \cdot \frac{a}{2} \qquad \text{where} \qquad \frac{a}{2} = R_k \sin\beta = \frac{R_b}{tg\alpha} \sin\beta \qquad (1.2.1)$$

For misalignment in the X-Z plane and clockwise shaft rotation (see Figs. 2 and 3), this moment will vary during one revolution. For a fixed radius of tooth crowning R_k, pitch radius R_b and contact angle α (e.g. 20°), the motion of the contact point between a gear tooth and the coupling sleeve is sketched in Fig. 3 a & b, and may be described by f_c.a/2, where

$$f_c = \frac{\omega t}{\pi/2} \qquad for \quad -\pi/2 \geq \omega t \geq \pi/2$$

$$f_c = 1 - \frac{\omega t - \pi/2}{\pi/2} \qquad for \quad \pi/2 \geq \omega t \geq 3\pi/2 \qquad (1.2.2)$$

At the same time the tooth force components in the X and Y directions, due to transmitted shaft torque, vary as $F_{ux}=F_u \cos \omega t$, $F_{uy}=F_u \sin \omega t$. For a single tooth contact, this produces moments about the X and Y axes of M_{cx} and M_{cy} respectively as indicated in Figs. 4a and 4b :-

$$M_{cx} = -F_u \frac{a_x}{2} f_c \cos\omega t \qquad M_{cy} = -F_u \frac{a_y}{2} f_c \sin\omega t \qquad (1.2.3)$$

If misalignment is not in the X-Z plane, then a/2 must be decomposed in X and Y components. Fourier decomposition of M_c gives predominantly second order contributions, with some higher order harmonics from M_{cx}.

1.3 Moment due to friction between mating teeth

The moment produced by the misaligned gear coupling due to the friction between the mating teeth is $M_f = T \cdot \mu$ (1.3.1)

Refer again to Figs. 1, 2 & 3, with misalignment in the X-Z plane. Starting from the X-axis with $\omega t=0$, the gear tooth contact point will initially be centred on the Y-axis. As ωt increases from 0 to $\pi/2$ (and also from $3\pi/2$ to 2π), tooth frictional force $F=\mu F_u$ is in the negative Z direction (opposing motion). The force becomes positive from $\omega t=\pi/2$ to $3\pi/2$ for angular misalignment in the X-Z plane as given by :-

$$f_\mu = \frac{\cos\omega t}{|\cos\omega t|} \qquad \text{and in Y-Z plane} \qquad f_\mu = \frac{\sin\omega t}{|\sin\omega t|} \qquad (1.3.2)$$

The moment arms about the X and Y axes vary as $R_b \sin \omega t$, $R_b \cos \omega t$. This produces moments about X and Y axes of M_{fx} and M_{fy} as shown in Figs. 4a and 4b, where for one tooth contact :-

$$M_x = F_u \mu R_b f_\mu \sin \omega t \qquad M_y = F_u \mu R_b f_\mu \cos \omega t \qquad (1.3.3)$$

When contributions from all gear teeth are summated, assuming symmetric tooth geometry and load distribution, then $\Sigma M_{fx}=0$. In practice, machining tolerances, transmission errors, wear, non-linearity etc., may give some cyclic variation and produce higher order frequency components. ΣM_{fy} consists mainly of a static component with a small dynamic moment at tooth passing frequency. For couplings with a large number of teeth the moment will be approximately constant with $M_{fy} \cong \mu T$.

1.4 Total moment produced by the misaligned gear coupling

To find the frequency content and magnitude of these moments the three components M_t, M_c and M_f must be summated during one revolution. This leads to general expressions for moments on the X and Y axes at station 1:-

$$M_{x1} = T \sin \psi_1 \cos 2\omega t - F_u \frac{a_{1x}}{2} f_c \cos \omega t - F_u \mu R_b f_\mu \sin \omega t$$

$$M_{y1} = T \sin \gamma_1 \sin 2\omega t - F_u \frac{a_{1y}}{2} f_c \sin \omega t + F_u \mu R_b f_\mu \cos \omega t \qquad (1.4.1)$$

These moments are periodic and consist of combinations of sine and cosine functions. The amplitudes of the various frequency components can thus be found using Fourier analysis. Similar expressions may be obtained for station 2. Forces generated at stations 1 and 2 are given by:-

$$F_{x1} = \frac{-M_{y1} \mp M_{y2}}{l} = -F_{x2} \qquad F_{y1} = \frac{M_{x1} \pm M_{x2}}{l} = -F_{y2} \qquad (1.4.2)$$

The signs depend on the point of intersection of the shaft centrelines.

For intersection between stations 1 and 2 use $+M_{y2}$, $-M_{x2}$
For intersection outside stations 1 and 2 use $-M_{y2}$, $+M_{x2}$
For intersection at station 1 use $M_{y2}=M_{x2}=0$

Static and dynamic force components also occur in the axial direction, and may be found by decomposition of F_x and F_y through the misalignment angle i.e. $F_{1,axial} = F_x \sin\gamma_1 + F_y \sin\psi_1$ etc.

1.5 Example for Gear Coupling

Gear coupling HCCE 010R (SKF), 0.8° rated misalignment angle in X-Z plane at station 1, distance between gear wheels is 168 mm, speed 1500 rpm, transmitted torque T=127.0 [Nm], and friction coefficient μ=0.15.

Moments about X and Y-axes due to angular misalignment are shown in Figs. 4a & 4b for forces from one gear tooth during one shaft revolution.

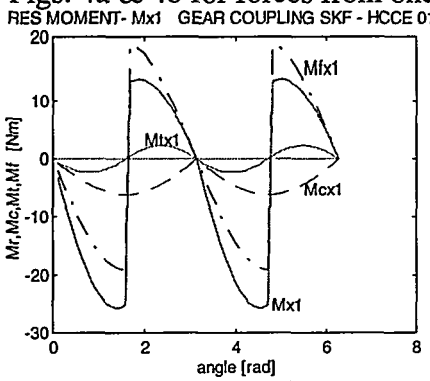


Figure 4a. Gear coupling, M_x

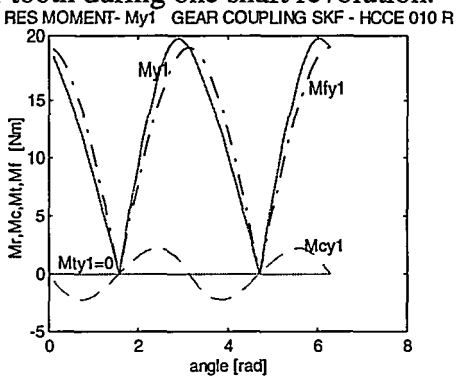


Figure 4b. Gear coupling, M_y

2. ANALYSIS OF MEMBRANE COUPLING

This type of coupling is torsionally stiff but allows flexure. The maximum allowable angular misalignment depends on the number of bolts, the membrane characteristics torque and speed etc. When a membrane coupling is misalignment two moments are developed :-

M_t - produced by torque and misalignment angle $\quad M_t = T \sin\beta \cos 2\omega t$

M_s - due to bending stiffness of the membranes $\quad M_s = k_\Theta \beta$

The angular stiffness k_θ of the coupling is asymmetric and varies between maximum and minimum values during one revolution of the shaft depending on angular location of the connecting bolts, and the geometry :-

$$k_\Theta = \frac{k_{\max} + k_{\min}}{2} + \left(\frac{k_{\max} - k_{\min}}{2}\right) \sin m\omega t \qquad (2.1)$$

Misalignment angle β may not only lie in the X-Z or Y-Z planes, and the moment must therefore be defined about the X and Y-axes [4],[5]. For angular misalignment at station 1 it can be shown that:-

$$M_{x1} = T \sin\psi_1 \cos 2\omega t - k_\Theta \gamma_1 \quad \text{and} \quad M_{y1} = T \sin\gamma_1 \sin 2\omega t - k_\Theta \psi_1 \qquad (2.2)$$

Similar expressions may be obtained for station 2, and the forces can be defined in the same way as for the gear coupling above.

2.1 Example of a membrane coupling

Membrane coupling AN4200 GR102 (ARPEX), 0.4° rated misalignment angle in X-Z plane at station 1 only, distance between membranes is 140 mm, speed 1500 rpm and transmitted torque T=127.0 [Nm] and angular stiffness k_θ = 12 Nm/deg.

Moments about X and Y-axes due to angular misalignment of the coupling at station 1 are shown in Figs. 5a and 5b for one shaft revolution.

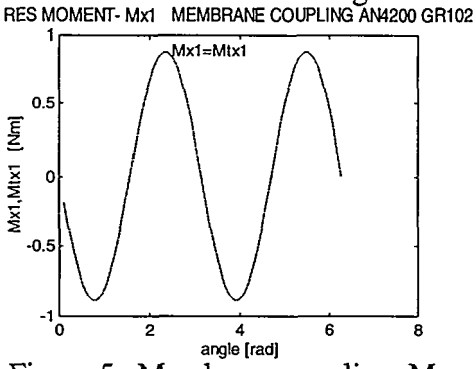


Figure 5a Membrane coupling, M_x

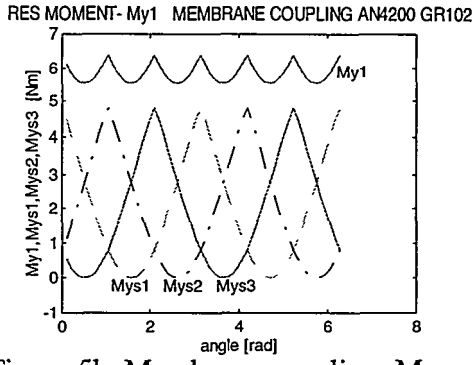


Figure 5b. Membrane coupling, M_y

3. PREDICTED VIBRATION RESULTS AND DISCUSSION

A shaft-bearing system with flexible coupling was modelled to investigate and diagnose the effects of misalignment. Both shafts were supported on circular journal bearings with the following linear dynamic coefficients:-

Stiffness coefficients [N/m]		Damping coefficients [Ns/m]	
Kxx	.56570E+07	Cxx	.33259E+06
Kxy	.26062E+08	Cxy	-.43668E+05
Kyx	-.23883E+08	Cyx	-.11343E+05
Kyy	.39961E+07	Cyy	.30398E+06

The equation of motion of the complete rotor system can be written as :

$$[M]\{\ddot{q}\}+[D]\{\dot{q}\}+[K]\{q\}=\{Q\} \tag{3.1}$$

The force vector {Q} includes the dynamic forces and moments generated by misalignment of the flexible coupling, and an unbalance value of G16 at the centre of each shaft. The equation may be solved to predict the vibration in system.

As usual, vibrations of the rotor system are complex and include several frequency components (see Fig. 6), such as 1x from rotating unbalance forces and 2x from shaft torque turning moment, M_t. The friction and contact moments, M_f and M_c, in the gear coupling and stiffness moment, M_s, in the membrane coupling (see Figs. 2 & 3) were described with Fourier series. Frequency components up to eight times the shaft speed have been included (i.e. 1x, 2x, 3x, ..., 8x; where x is shaft speed).

Components with frequency 4x, 6x, 8x, usually have low values (2-10 % of second order 2x component). Figs. 6 a, b&c show predicted shaft vibration, orbits and spectra at the drive shaft bearing located closest to the coupling at station 1. Only the dynamic part of the result is presented as

this is typical of what is measured for diagnostic purposes. In the simulation the forces have been applied some time previously, and the resulting response is slowly converging into a stable orbit. Fig. 6a is for a coupling with very little misalignment, and shows mainly 1x unbalance response. Fig. 6b is the membrane coupling with a horizontal misalignment of $\beta = 0.4°$ with T = 127 Nm and $k_\theta = 500$ Nm/deg. Fig. 6c is the gear coupling with $\beta = 1°$ (horizontal), T = 127 Nm, $R_k = 20\ R_{ko}$. Increased 2x and higher order frequencies is clearly seen for the more severe cases of misalignment.

Figs. 7 and 8 respectively show results for the gear and membrane couplings for a range of conditions. Effects of misalignment on shaft static equilibrium position is also indicated in these figures. Static equilibrium position due to gravity load of the rotor is at X=4.5, Y=-1.0 μm. For the bearing described above, increasing M_c or M_s, will move the shaft in the +Y direction. This is also the case for M_f , resulting from the very strong cross coupling stiffness in the bearing, even though M_f produces a bearing force in the -X direction. Moment M_t has no static component.

4. CONCLUSIONS

1. For both gear and membrane couplings, misalignment causes bearing forces at second harmonic of shaft speed due to the transmitted shaft torque being turned through the misalignment angle. The moment produced acts about an axis perpendicular to the plane of misalignment, and the amplitude is dependent on shaft torque and misalignment angle.

2. In misaligned gear couplings, tooth contact forces due to contact point effects and sliding friction can produce additional static and periodic moments about coordinate axes perpendicular to the shaft line. Dynamic forces produced at bearings will be predominantly at twice shaft speed.

3. Gear coupling faults such as tooth damage, poor lubrication, lockup, transmission errors, or poor manufacturing quality, may produce asymmetric distribution of tooth forces. In the worst case, all the torque could be transmitted by one tooth. For diagnosis, such faults could be characterised by an increase in higher order frequency components.

4. Membrane couplings having asymmetric bending stiffness can generate high order harmonics at frequencies related to the number of connecting bolts, in addition to second order due to transmitted torque.

5. Synchronous forces do not appear to be generated by the misalignment itself. However, the moments produced are periodic functions of shaft speed, and inherent unbalance or runout in the coupling

will usually be present. These first order frequency effects may be enhanced by any eccentricity caused by the misalignment forces.

6. Shaft position in a journal bearing clearance will be influenced by the static moments from gear tooth friction or membrane stiffness in the misaligned coupling. However, bearing direct and cross coupling stiffness and damping characteristics complicate the response, and a thorough knowledge of bearing behaviour is necessary before such data can be used for coupling diagnostics.

Acknowledgments

This work was partially supported by EU contract BE95-2015 Modiarot.

References

1. E. L. Pernleitner, Ausrichten von maschinensatzwellen, VDI-Verlag GmbH, Dusseldorf 1995.
2. D. L. Dewell, L. D. Mitchell, Detection of a misaligned disc coupling using spectrum analysis, UKM Paper No: 4160/10
3. T. Bevan, Theory of Machines, Longmans,1965
4. A. S. Sekhar, A. Sreenivasa Rao, Crack Versus Misalignment in Rotor-Coupling Bearing System, Machine Vibration (1996), 5: 179-188.
5. A.S. Sekhar, B.S. Prabhu, Effect of coupling misalignment on vibrations of rotating machinery, J. of Sound and Vibration (1995), 185(4), 655-671.

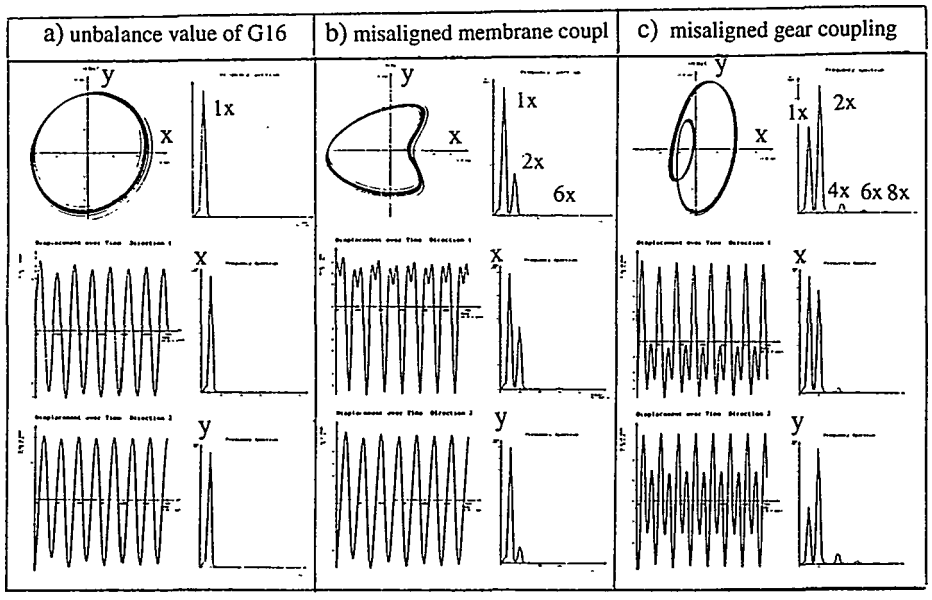


Figure 6 Time and frequency plots for examples of misalignment

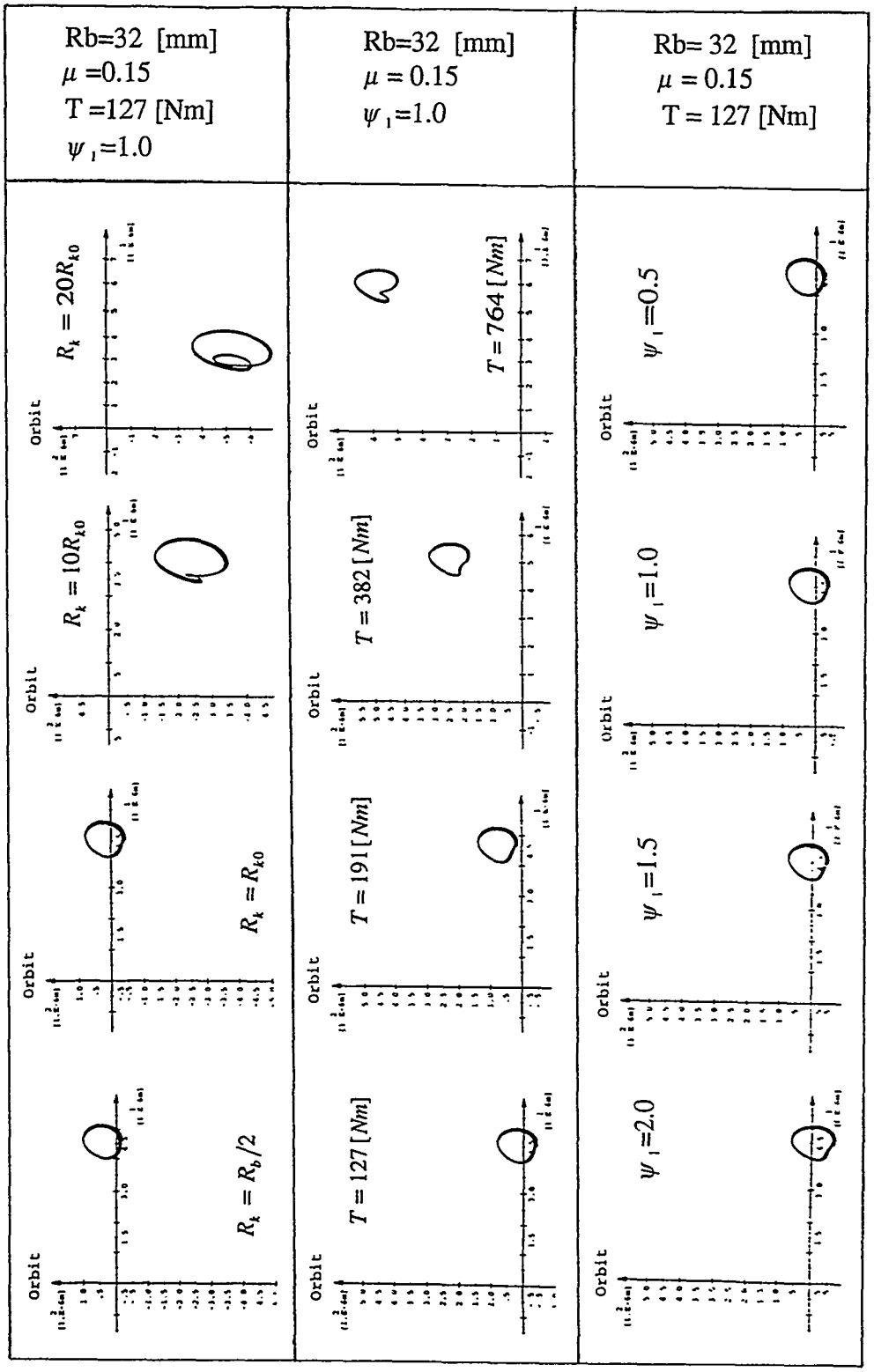


Figure 7 Orbit plots for a misaligned gear coupling

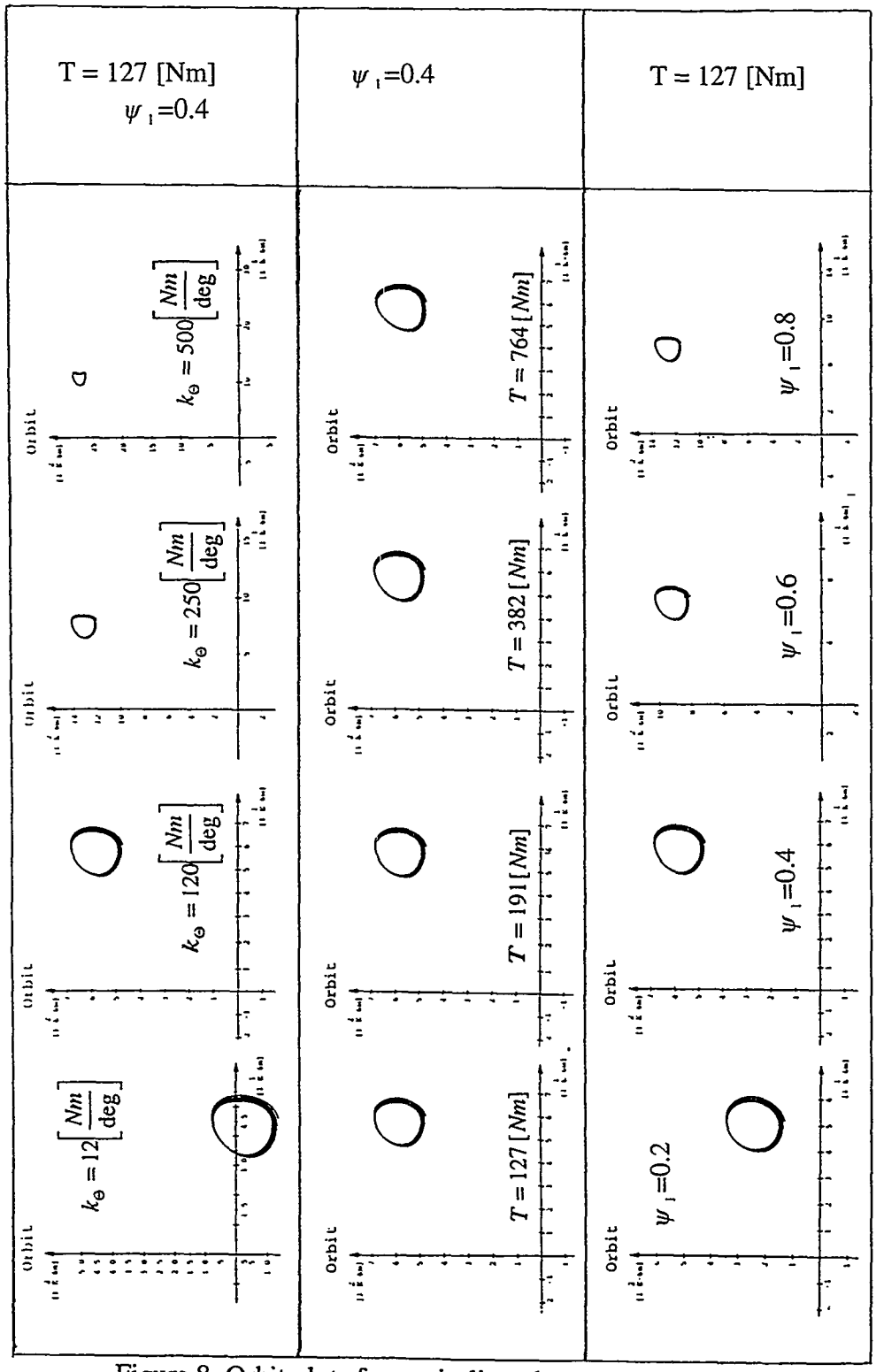


Figure 8 Orbit plots for a misaligned membrane coupling

VIBRATION MONITORING OF AIRCRAFT RECIPROCATING ENGINES
PART 1: DYNAMIC CHARACTERISATION OF THE ENGINE

Júlio Montalvão e Silva, Professor of Mechanical Engineering
Instituto Superior Técnico, Universidade Técnica de Lisboa
Lisboa, Portugal (email: D569@beta.ist.utl.pt)

Nuno Neves Nunes, Assistant Professor of Mechanical Engineering
Escola Superior de Tecnologia, Instituto Politécnico de Setúbal
Setúbal, Portugal (email: nnunes@bocage.ips.pt)

ABSTRACT
Aircraft engines operation and maintenance follow very demanding rules as any failure, when in flight, will put at risk the lives of crew and passengers. Regular inspections and different levels of maintenance actions at specified operation time intervals, are implemented according to strict and well defined preventive maintenance planning. Despite all these careful procedures, accidents occur and it is a fact that some of them are due to engine failure.
Condition monitoring of aircraft engines is therefore a welcome additional preventive procedure as it allows increasing the safety margins of operation. Automatically controlled systems with different levels of sophistication, have been used with reported success. However, small aircraft with reciprocating engines tend to be somewhat neglected.
The present paper describes a case study (currently under implementation) of aircraft engines condition monitoring involving a pilot training fleet of FAP (Portuguese Airforce) constituted by 23 engines for the 16 *EPSILON* aeroplanes equipped with 6-cylinder *Lycoming* reciprocating engines. The condition monitoring techniques under implementation are based on vibration analysis. An adequate implementation of these techniques implies thorough knowledge of the aircraft engine dynamic characteristics. Thus, the description of this case-study is divided in two parts: Part 1,. which constitutes the contents of this paper, presents and discusses all the theoretical and experimental analyses that were carried out in order to characterise the vibration behaviour of the engine. Crankshaft-propeller system, reciprocations masses, torsion vibration absorber, gears, camshafts,

fuel pump, oil pump and magnets, among other important engine parts, were carefully analysed and their vibration properties were quantified. The thermodynamic behaviour of the engine was also taken into consideration. As a result, the engine fundamental vibration properties were defined so that they could be used to interpret the data acquired by on-operation condition monitoring procedures. Part 2 is presented in a following paper and describes the techniques and procedures implemented to monitor the condition of the engines of the FAP *EPSILON* fleet. Results are presented and discussed.

1. INTRODUCTION

During the first forty years in the history of aviation, internal gas combustion engines were the most used class of aircraft motor-propulsion.

At the end of the Second World War the turbo-jet was introduced. Although this motor became the means of propulsion *par excellence*, the0 alternative internal combustion engines, what we might call conventional, continue to be widely used in small civilian and military aircraft. This is due to the low acquisition and maintenance costs of these engines that, functioning at low rotation, do not have high demands in the characteristics and quality of materials relative to the turbo-jets.

Throughout the years there have been several types of these engines differing according the positioning of the cylinders. Today, the most common engine has its cylinders opposing one another (or in V) and can be found in the engines of small aircraft with four, six or eight cylinders.

The operation and maintenance norms of these engines obey very demanding patterns. Any kind of anomaly or breakdown occurring in flight could put at risk the lives of its crew and the loss of the aircraft. For security measures the engine is obligatorily inspected at regular intervals of function, with different levels of inspection being foreseen for differing functional hours.

The evolution of technology is leading to the increasing use of automatic control systems that allow for forecasting the state of functioning of equipment. One of the control processes being used in aircraft engines is the analysis of engine vibration levels [1].

By doing this analysis, the functioning state of the engine may be forecast, this being a very useful tool in the field of maintenance.

The final objective of the work in progress is the creation of an expert system that could be installed to, through the acquisition of continuous

information, inform the pilot about the functioning state of the engine by giving anticipated reports of eventual anomalies in function.
With this in mind an exhaustive study of the engine was initiated with view to its dynamic characterization as described in this paper. In parallel, an experimental analysis of engine behaviour of the FAP (Portuguese Air Force) fleet of aircraft was undertaken, in order to establish a database. The procedures, adopted methodologies and obtained results up to this date will be presented in the second part of this report [2].

2. ENGINE DYNAMICS CHARACTERIZATION

The Lycoming engine, model AEIO-540-L1B5D, was used for the analysis.

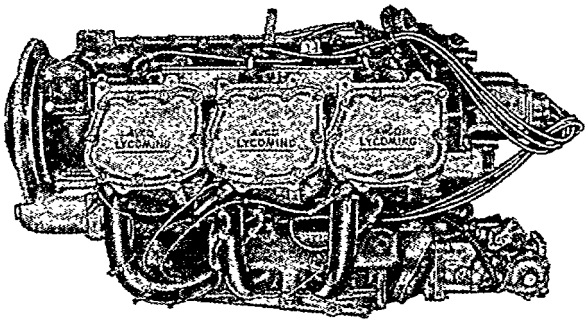


Fig. 2.1 - Six cylinder *Lycoming* Engine

Although engine characterization involves various aspects, this paper will only concern itself with those considered relevant to the dynamic characterization that influence the analysis of vibrations, having as its objective the definition of the typical engine frequencies.

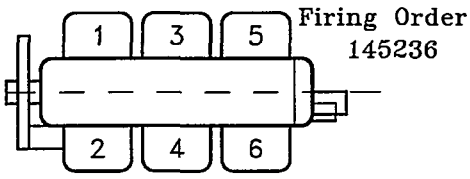


Fig. 2.2 - Numbering of the engine cylinders

The engine under study has no roller-bearing. The crankshaft is supported by sliding blocks, just like the came shaft.
The numbering of the engine cylinders is indicated in fig. 2.2. After the construction of the phase diagram for an engine with these characteristics the following would be obtained: cylinder position and firing sequence with the following major and minor orders (see fig. 2.3).

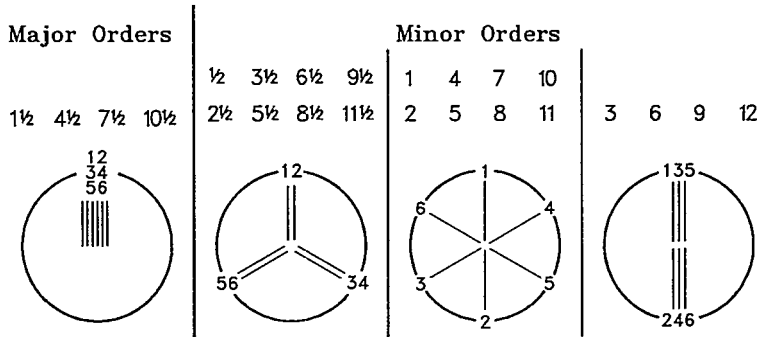


Fig. 2.3 - Diagram of engine phases

From the phase diagram it can be seen that the harmonics which are not in equilibrium (Major Orders) are: 1½, 4½, 7 ½ and 10 ½.

In the engines with alternative firing, the crankshaft is subjected to a discrete periodic disturbance, applied through the connecting rods resulting from the firing which occurred in the cylinders. These forces create in the crankshaft not only bending but torsion. It is therefore fundamental to know the natural frequencies associated with the vibration mode in bending and torsion [3], [4].

On the other hand, the analysis of disturbance frequency allows for the determination of the predominant excitation frequencies that, in the case of an engine with a four-stroke cycle, are usually multiples of the half harmonic of the engine's rotational velocity.

Among other projects already analysed with these objectives we quote as an example the one by *Sobieraj* [5] that conceived a discrete crankshaft/propeller model of a reciprocating aircraft engine, with the objective of determining the dynamic properties of the system using the finite elements method. In his work he characterizes the elements and proposes a type of mesh depending on the model under study. *Dzygadlo* [6] also used the model described above to determine the dynamic properties of a crankshaft/propeller system. With this model it was possible to trace the transference functions of the system and to establish the relations $\ddot{\theta}/M$ in frequency function and the determination of the natural frequencies of torsion of the system.

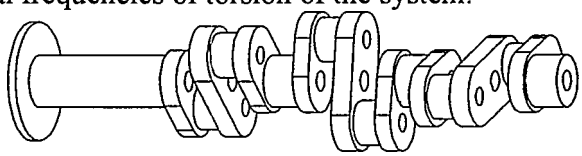


Fig. 2.4 - Engine crankshaft under analysis

In fig. 2.4 the crankshaft under analysis is represented. The associated components of the crankshaft (propeller, equivalent masses of the

piston connecting rod assembly and hind gear train) were taken into consideration and, because of having mass and inertia on the same large scale as the crankshaft, could not be ignored.
As a first approximation, a theoretical study was made of the crankshaft/propeller assembly as represented in fig. 2.5.

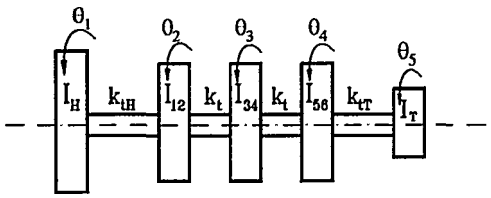


Fig. 2.5 - Schematic model of the system to be studied

I_H- Moment of Inertia of the propeller
I_{12}, I_{34}, I_{56}- Moment of Inertia of the cranks and equivalent piston connecting rod assembly
I_T- Moment of Inertia of the hind gear train
k_{ti}- Coefficient of Stifness to the Torsion of the shaft
θ_i- Degree of freedom of the system

The equation that establishes equilibrium is given by:

$$[I_n]\{\ddot{\theta}\}+[k_t]\{\theta\}=\{0\} \tag{2.1}$$

where: $[I_n]$ represents the matrix of inertia of the system
$[k_t]$ is the matrix of torsional stifness
$\{\ddot{\theta}\}$ and $\{\theta\}$ are the vectors of acceleration and angular positions

The moments of inertia of the various components were calculated or experimentally determined, giving rise to the values listed in Table 2.1.

Table 2.1 - Moments of Inertia of the system under study

Set	Component	I[$kg\,m^2$]	Crankshaft	I[$kg\,m^2$]	Var.	I[$kg\,m^2$]
Propeller	Propeller	2.26700	Bearing	0.00101	I_H	**2.26801**
Cylinders 1,2	Equiv.Piston/Rod	0.00429	Bearing+Cranks	0.13201	I_{12}	**0.13630**
Cylinders 3,4	Equiv.Piston/Rod	0.00429	Bearing+Cranks	0.13093	I_{34}	**0.13522**
Cylinders 4,5	Equiv.Piston/Rod	0.00429	Bearing+Cranks	0.13093	I_{56}	**0.13522**
Gear Train	Engrenagens	0.00264	Bearing	0.00024	I_T	**0.00288**

The torsional stifness of the crankshaft was calculated considering that only its supports contribute to it.

Table 2.2 - Torsional Stifness of the Supports

Bearing	Var.	Stifness [Nm]
Propeller	k_{tH}	827095
Others	k_{tO}	3544692

Eq. (2.1) may be writen

$$\left[[k_t]-\omega^2[I_n]\right]\vec{\theta}=\vec{0} \tag{2.2}$$

Dividing Eq.(2.2) by $\lambda=\frac{1}{\omega^2}$, premultiplying by $[k_t]^{-1}$ gives

$$\left[\lambda[I]-[D]\right]\vec{\theta}=\vec{0} \tag{2.3}$$

where $[I]$ is the matrix identity

and $[D]=[k_t]^{-1}[I_n]$ is the matrix dynamic of the system

Eq. (2.3) represents mathematically a problem of eigen values whose solution are the values indicated in Table 2.3, being the vibration mode schematically represented in fig. 2.6.

Table 2.3 - Natural Frequencies (Eigen Values)

	Eigen Value $1/\omega^2$	Torsion Frequency [*Hz*]	Vibration Modes				
			θ_1	θ_2	θ_3	θ_4	θ_5
Rigid	-0.811582	≈ 0	0.062017	0.062018	0.062018	0.062018	0.062018
1°	**4.79362×10^{-7}**	**230**	**0.074996**	**-0.354012**	**-0.425718**	**-0.463468**	**-0.464254**
2°	3.44416×10^{-8}	858	0.008170	-0.642291	-0.076998	0.573768	0.587606
3°	1.26233×10^{-8}	1418	0.001972	-0.426388	0.772465	-0.368292	-0.393580
4°	7.93866×10^{-10}	5649	4.03E-09	-0.000014	0.000657	-0.030317	1.398804

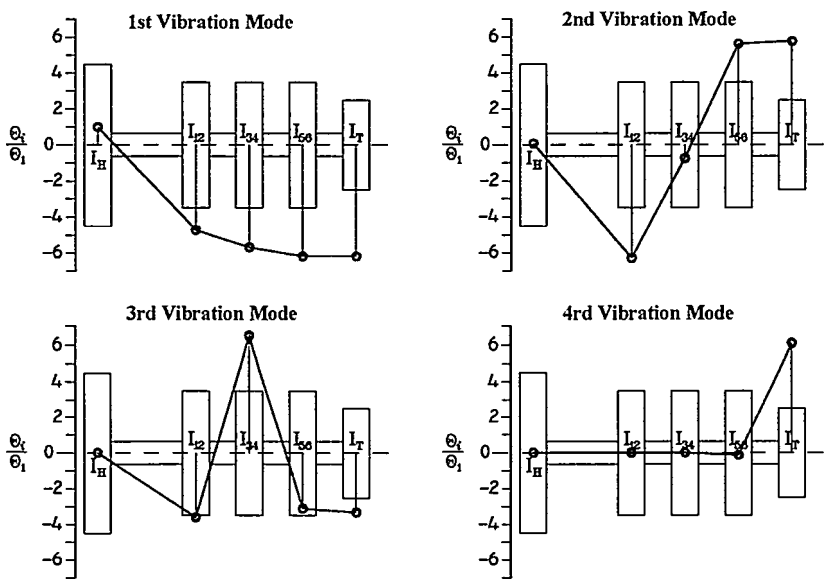


Fig. 2.6 - Torsional vibration modes of the crankshaft

As a second approximation, an experimental verification of the vibration mode of the previously analysed system was undertaken. For this, modal analysis techniques were applied.

With the view to simulating the "*free free*" boundary condition, the crankshaft was suspended by two cables in a portico. Additional masses were placed corresponding to the inertias of the propeller and the hind gear train, likewise the equivalent masses of the piston connecting rod assembly. Fig. 2.7 illustrates the apparatus used in the experimental model.

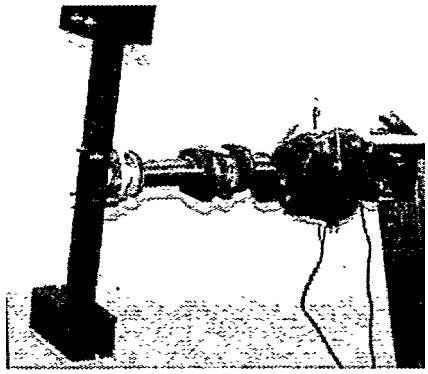


Fig. 2.7 - Test model

The process used was based on the application of a disturbance to one of the cranks. The applied disturbance through a vibrator was transferred to the crankshaft through the application of a force and of a moment in relation to the crankshaft axis. With transducers placed in T-form on blocks located on the cranks of the crankshaft the response of the system is measured. The frequency response of the system was thus determined and the natural torsional frequencies were identified. The measurement chain was mounted as shown in fig. 2.8.

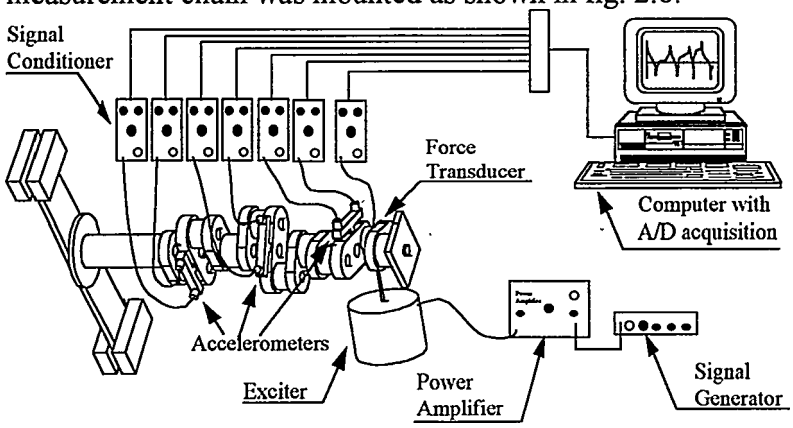


Fig. 2.8 - Measurement chain for the determination of torsional frequencies

The objective of using T-blocks was to rapidly identify the type of vibrational movement (dislocation or rotation) [7]. Considering the block in fig. 2.9, the dislocation of point **P** can be given by:

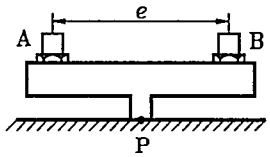


Fig. 2.9 - T-block to measure rotations

$$x = \frac{x_A + x_B}{2} \qquad (2.4)$$

The rotation of point **P** can be obtained by:

$$\theta = \frac{x_A - x_B}{2e} \qquad (2.5)$$

After the assembly of the system the process of testing took place. Fig. 2.10 represents the spectral analysis of the function $\ddot{x}/F$.

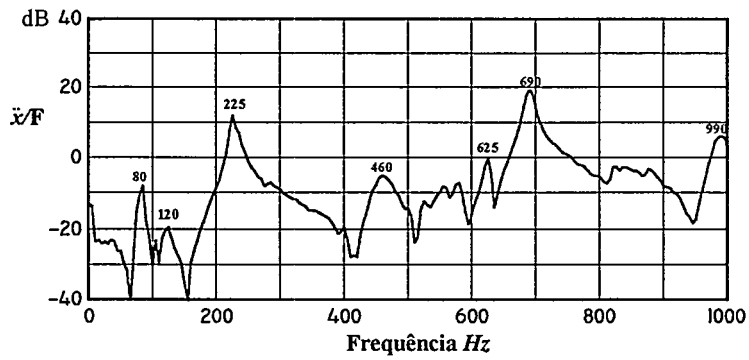


Fig. 2.10 - Frequency Response of Function $\ddot{x}/F$ (range 0 a 1000 *Hz*)

As you can see from the spectrum in fig. 2.10 there is a group of natural frequencies of the system. With each of these frequencies there is an associated mode of vibration. As the objective was the experimental identification of the modes of torsional vibration it was therefore necessary to analyse the following frequencies more closely: 80; 106; 117; 145; 227; 460; 625; 690 and 990 *Hz*.

To identify the vibration mode simple harmonic disturbances were applied to the system at the frequencies listed above. The time responses of the accelerometers were registered in files. The technique consisted in determining the phases of the block's accelerometers and from that to conclude the fundamental nature of the movement of the pair of cranks of the crankshaft under analysis, identifying whether treated to a dislocation or a rotation.

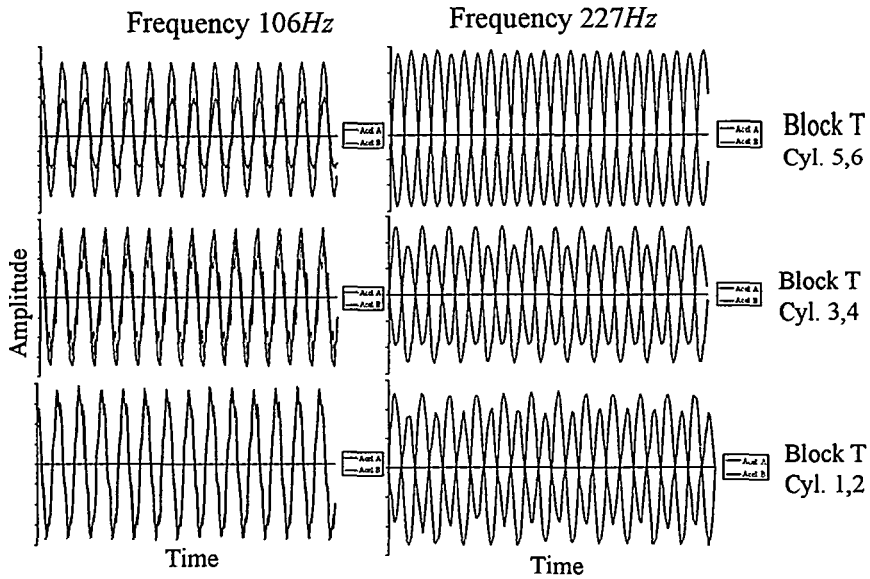


Fig. 2.11 - Response of the T-block accelerometers to an excitation of 106 *Hz* and 227 *Hz*

From observation of the signal phases it can be concluded that for a frequency of 106 Hz the system suffers a dislocation (bending), while for a frequency of 227 Hz the system suffers a rotation (torsion).

Table 2.4 presents the determined natural frequencies and the identification of the associated mode of vibration.

Table 2.4 - Crankshaft natural frequencies - Experimental Analysis

Frequency [*Hz*]	Vibration Mode
80	Bending
106	Bending
117	Bending
145	Bending
227	**Torsion**
460	Bending
625	Bending
685	**Torsion**
990	Bending

It can be seen from the table data that the first natural frequency of torsion is close to the one calculated analytically. The same is not true for the second frequency which is significantly lower than the one calculated earlier. This fact is easily justifiable by the simplifications used in the definition of the simplified theoretical analysis.

In the engine under study the existence of a counter-weight was seen with the finality of decreasing the level of torsional vibrations to the natural torsional frequency of the crankshaft/ propeller system functioning like a dynamic absorber in the crankshaft (see fig. 2.12). This is placed in a way so as to have an oscillation frequency equal to the natural torsion frequency of the crankshaft. Its

pendular movement was studied so as to oscillate in phase opposition with the torsional vibrations of the crankshaft at the torsion frequency. Once the system is a little damped, it is possible in this way to decrease the torsional oscillations that could otherwise originate high levels of vibrations with unpredictable consequences.

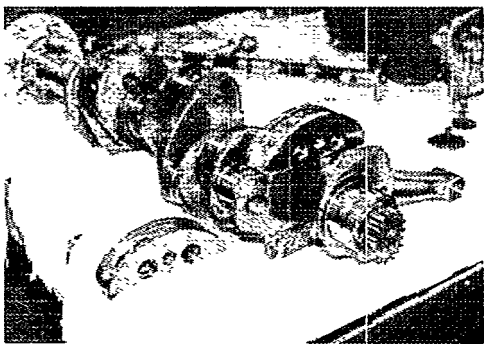


Fig. 2.12 - Dynamic counter-weight destined to torsional vibrations

In the hind part of the engine (see fig. 2.13) there is a gear train that through a carriage moved by the crankshaft transmits movement to the following components: fuel pump, box of magnetos; came shaft; oil pump and vacuum pump. In the front part of the engine there is a conical gear that is responsible for the movement of the propeller's variable step pump. It is important to know the geometry and relations of transmission of the gears in order to determine the typical frequencies one expects to find in the spectrum of frequencies [8].

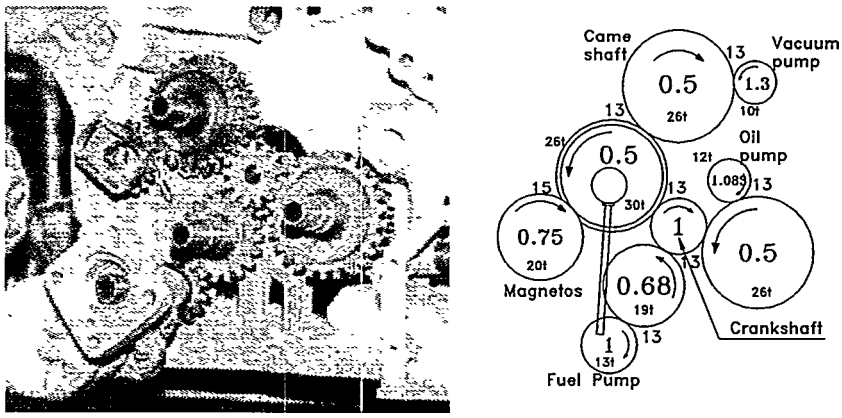


Fig. 2.13 - Hind Gear Train

The values presented are normalized at the rotation frequency of the crankshaft representing harmonics of rotation velocity. It can be seen that the harmonics which require special attention are: the 13th (fuel pump; came shaft; oil pump and vacuum pump) and the 15th (magnetos).

The information relative to the engine gears can be resumed as in Table 2.5.

Table 2.5 - Relations of Transmission of the Engine Gears

Component	Number of Teeth	Normal Velocity	Gear Frequency
Crankshaft	13	1	13
Came Shaft (Hind)	26	0.5	13
Magnetos	20	0.75	15
Fuel Pump	13	1	13
Oil Pump	12	1.083	13
Vacuum Pump	10	1.3	13
Iddle Gear (Magnetos)	26 and 30	0.5	13 and 15
Iddle Gear (Oil Pump)	26	0.5	13
Iddle Gear (Fuel Pump)	19	0.684	13
Came Shaft (Anterior)	34	0.5	17
Iddle Gear	19 and 18	0.895	17 and 16
Variable Step Pump	17	0.947	16

3. TYPICAL ENGINE VIBRATION SIGNATURES

Fig. 3.1, 3.2 and 3.3 show typical spectra of frequencies of the engine in various functional regimes, obtained in accordance with the methodologies described in the 2nd part of this report.

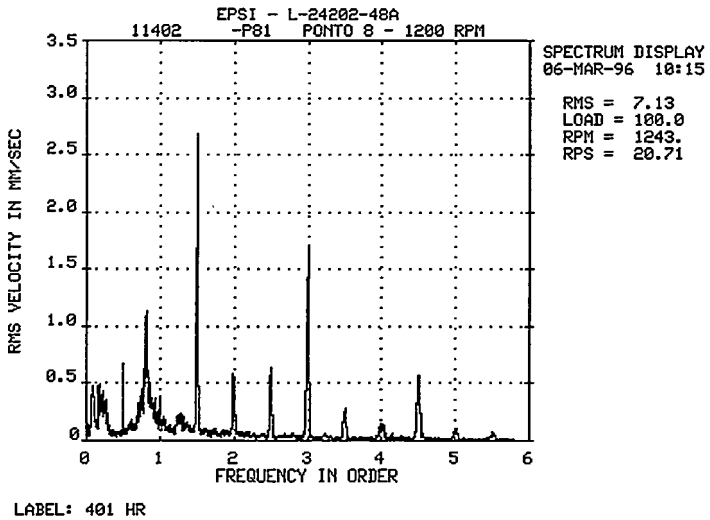


Fig. 3.1 - Typical Spectrum at 1200 *rpm*

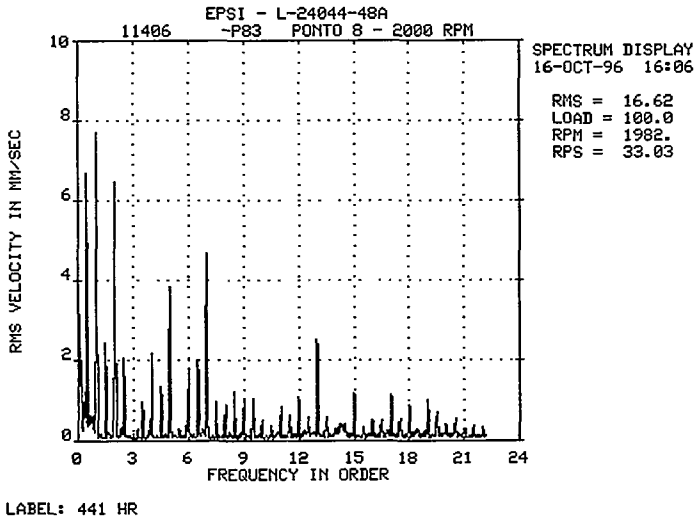


Fig. 3.2 - Typical Spectrum at 2000 *rpm*

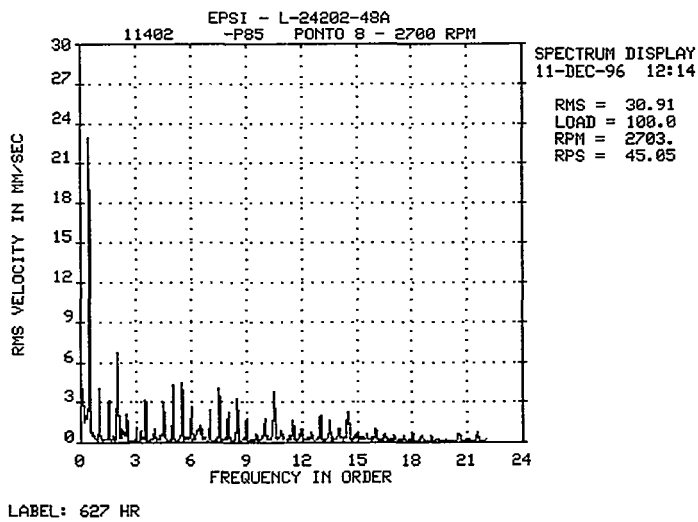


Fig. 3.3 - Typical Spectrum at 2700 *rpm*

From observation of the spectra it can be seen that there is a group of typical frequencies that characterize the functioning of the engine. The identification of the components of the collected spectra is not yet complete. However, there is a group of frequencies that have already been identified in this paper and are presented in Table 3.1.

Table 3.1 - Identification of the components of the spectra of the engine

Frq[*Hz*]	Ord	Description	Sympton
Regime de funcionamento 1200 rpm			
2	--	Structural natural frequency	Engine support
8	--	Structural natural frequency	Engine support
20	1	Rotational velocity	Propeller disequilibrium
30	1½	Engine typical	
90	4½	Engine typical	
Regime de funcionamento 2000 rpm			
16.7	½	Explosion in cylinders	Difference in cylinders explosions
33.3	1	Rotational velocity	Propeller unbalanced
66.6	2	Blades	Missalignement
166.7	5	Engine typical	
233.3	7	Natural torsion frequency	Dynamic absorber (counter-weights)
433.3	13	Hind gear mesh frequency	Oil pump, fuel pump, magnetos, etc ...
500	15	Magnetos mesh frequency	Magnetos
Regime de funcionamento 2700 rpm			
22.5	½	Thermodyamics process	Difference in cylinders explosions

4. FINAL NOTES

The techniques used and the results obtained from the determination of the characteristic frequencies were described for both structural and for engine functioning. The levels of engine vibration were registered for various functioning regimes.

Although a group of typical frequencies was identified, it was shown that a much wider and better knowledge of the engine is necessary in order to identify the components of the spectrum that reveals itself to be complex in the case of reciprocating engines.

REFERENCES

1- Kroes, M.J. and Wild, T.W. - "*Aircraft Powerplants*", McGraw-Hill, 7ªEd, USA 1994

2- Silva, J.M. and Nunes, N.N., "*Vibration Monitoring of Aircraft Reciprocating Engines- Part 2: Contribuition of the Definition of the Monitoring Parameter*", COMADEM 97

3- Nestorides, E.J., B.I.C.E.R.A. "*A Handbook on Torsional Vibration*", Cambridge University Press, 1958

4- Den Hartog, J.P. - "*Mechanical Vibrations*", Dover Publications, 4ªEd, USA 1985

5- Sobieraj, W. - "*A Discret Dynamic Model of the Cranckshaft-Airscrew Assembly of Aircraft Piston Engine for the Purpose of Vibration Analysis by the Method of Finite Elements*", Journal of Technical Physics, 1988

6- Dzygadlo, Z. and Sobieraj, W. - "*Numerical Analysis of Forced Vibrations of a Cranckshaft-Airscrew System*", Journal of Technical Physics, 1989

7- Ewins, D.J., Modal Testing: Theory and Practice, Research Studies Press, Ltd. 1984

8- Mitchell, J.S. - "*An Introduction to Machinery Analisys and Monitoring*", Ed. Pennwell, 2ªEd., USA 1993

VIBRATION MONITORING OF AIRCRAFT RECIPROCATING ENGINES
PART 2: CONTRIBUTION OF THE DEFINITION OF THE MONITORING PARAMETER

Júlio Montalvão e Silva, Professor of Mechanical Engineering
Instituto Superior Técnico, Universidade Técnica de Lisboa
Lisboa, Portugal (email: D569@beta.ist.utl.pt)

Nuno Neves Nunes, Assistant Professor of Mechanical Engineering
Escola Superior de Tecnologia, Instituto Politécnico de Setúbal
Setúbal, Portugal (email: nnunes@bocage.ips.pt)

ABSTRACT
The present paper constitutes the second part of a case-study of aircraft engine condition monitoring, involving 23 6-cyclinder *Lycoming* reciprocating engines that equip an *EPSILON* pilot training fleet used by FAP (Portuguese Airforce). The monitoring techniques are based on on-operation vibration analysis. Part 1 of this case-study constitutes a preceding paper and presented the theoretical and experimental engine vibration characterisation. The objective of this work is to define monitoring procedures that will allow for increased safety and longer operational life. For this purpose, the vibration responses of the 23 engines were regularly measured under normal operation conditions, after an initial bench-test behaviour characterisation. Vibration signatures were analysed using FFT techniques and results were correlated with the "a priori" known engine dynamic properties. Consideration of parameters such as engine speed, propeller balance, time of operation after last major overhaul, and many others, together with statistical data analysis allowed average values to be obtained and, through a weighting criterion, defining what was *called "Degree of Operationality"*. The corresponding monitoring vibration data analysis and discussion is presented in this paper. An example of fatigue failure of a crankshaft, during actual flight, is presented and it is shown that vibration signatures indicated a trend in the spectra that could be related to the developing anomaly.

1. INTRODUCTION
After the dynamic characterization of the engine made in the first part of this work, a systematic collection of the vibration levels data of the

engines installed in the EPSILON aircrafts that make up the 101 Fleet of the Portuguese Air Force was then made.
The data was later registered in a database and was subjected to statistical analysis so as to establish normal values and to make possible the eventual detection of anomalies in the engines.
Based on the collected data it was possible to identify some anomalies that after interventions were corrected led to the reduction in vibration levels. These actions improved the engine's performance avoiding unnecessary wear of the components and made possible the reduction in the number of breakdowns caused by excessive vibration.
In this paper will be presented several cases of anomalies with the respective corrective measures.

2. UNBALANCED PROPELLER

The unbalance is characterized by a high amplitude at the rotation velocity. The measurements taken showed the existence of very different values at the rotation velocity. The spectra in Fig. 2.1 show two different cases where in b) a unbalance situation can be seen.

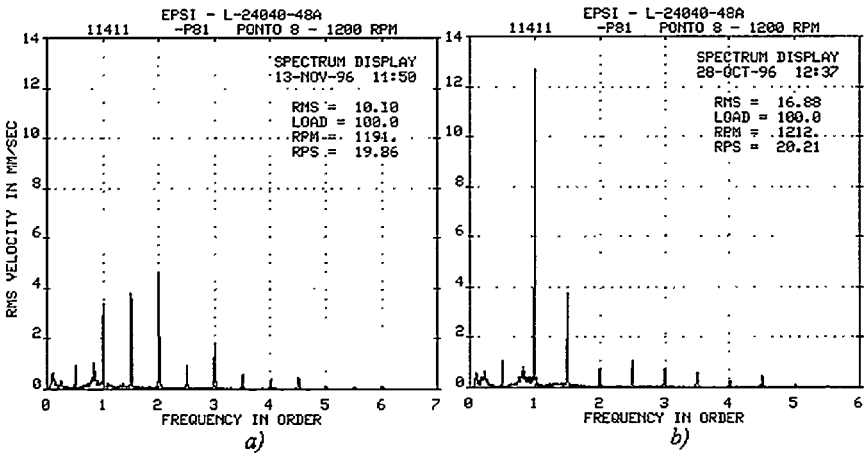


Fig. 2.1 - Example of unbalanced propeller - 1200 rpm

The aircraft fleet was subjected to a balancing procedure. It was decided to subject all engines that presented rotation velocity values higher than 5 *mm/s* (RMS) to this procedure.

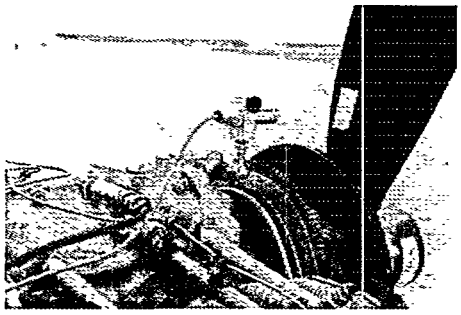


Fig. 2.2 - Balancing Apparatus

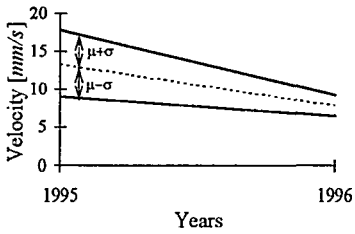


Fig. 2.3 - Evolution of vibration level at 1200 *rpm*

The method used was to balance the crankshaft/propeller assembly. The measurement of the unbalance was taken with the engine functioning at 1200 rpm. Thus it was possible to determine the equilibrium mass and its position angle. The apparatus in Fig. 2.2 was used.
These actions significantly improved the vibration level of the fleet. The reduction in values over a year period is as presented in Fig. 2.3.

3. PROBLEM IN THE LUBRIFICATION SYSTEM

The 13th harmonic represents the frequency of the engine's hind gear train. The average level observed for the group of engines measured is 1.8 *mm/s* at 2000 rpm. On one of the engines measured a sudden increase of this value was detected as being around 4 *mm/s* as was the increase of other harmonic components, namely the 15th and the 17th. Fig. 3.1 represents two spectra where a) refers to an engine that is considered good, and b) refers to an engine where the anomaly was detected.

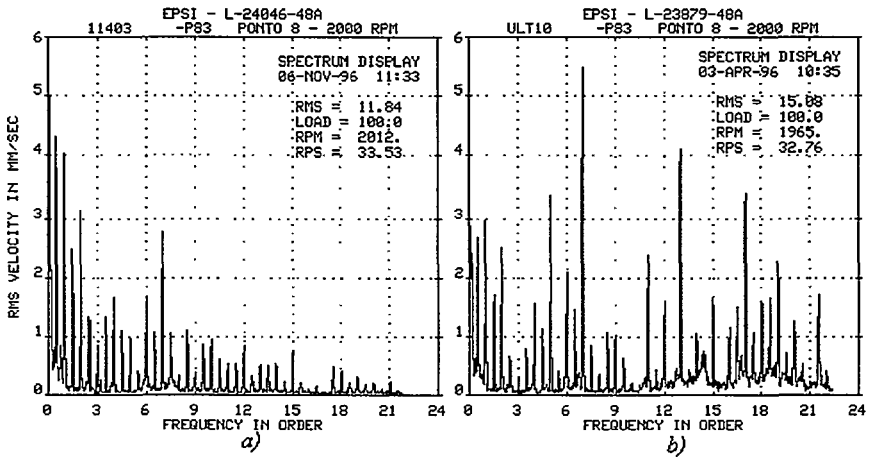


Fig. 3.1 - Anomaly in the gear train - Spectra

This motor was subjected to closer attention and after an oil analysis the presence of iron particles in the oil was detected. Later the presence of rubber coming from the magnetos was verified which had got lodged in the oil housing alongside the suction duct of the oil tube, leading to an overload of the oil pump that functions coupled to the gear train.

4. DISEQUILIBRIUM IN THE FIRING CYLINDERS

One aircraft, at 2700 rpm, had a high component of the ½ harmonic of the engine. The average value for the group of engines measured is 25 *mm/s*. In the considered problematic engine values higher than 40 *mm/s* were measured. Fig. 4.1 presents the spectra of a good engine *a)* and the spectra of an engine with an anomaly *b)*.

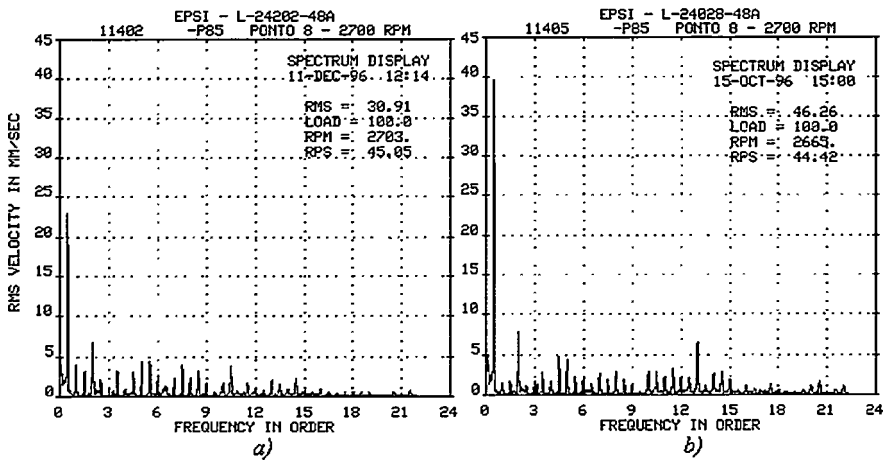


Fig. 4.1 - Disequilibrium in the firing cylinders - Spectra

After the rate of compression of the cylinders, the fuel feed system and the ignition rate (magnetos) were verified, it was detected through a thermography test that one of the cylinders had a working temperature of around 150°C lower than that of the other cylinders indicating a default in the combustion process.

5. FRACTURE OF THE CRANKSHAFT IN FLIGHT

One of the objectives of the current project is to create a process that permits the categorization of the engines by registered vibration levels. The algorithm that is still being developed takes into consideration the functioning hours of the engine and the way it is used which is obtained through a mission profile.

Based on this data a value known as the *"Degree of Operationality"* is obtained. Engines have thus been categorized by the *"Degree of Operationality"*.

In October 1996 an accident occurred with an aircraft of the fleet which fractured the crankshaft in flight. It was verified that the origin of

the fracture was a fissure that propagated until reaching its critical length. This aircraft made several flights with the process of propagation of the crack taking place until collapsing.

After the *"Degree of Operationality"* was consulted it was verified that the engine in question was classified in the last place with the lowest degree.

From observing the spectra of this engine it was verified that the component of the 7th Harmonic at 2000 rpm had a high value. Fig. 5.1 represents two spectra at 2000 rotations, *a)* corresponding to a good engine, and *b)* corresponding to an engine with problems in the crankshaft.

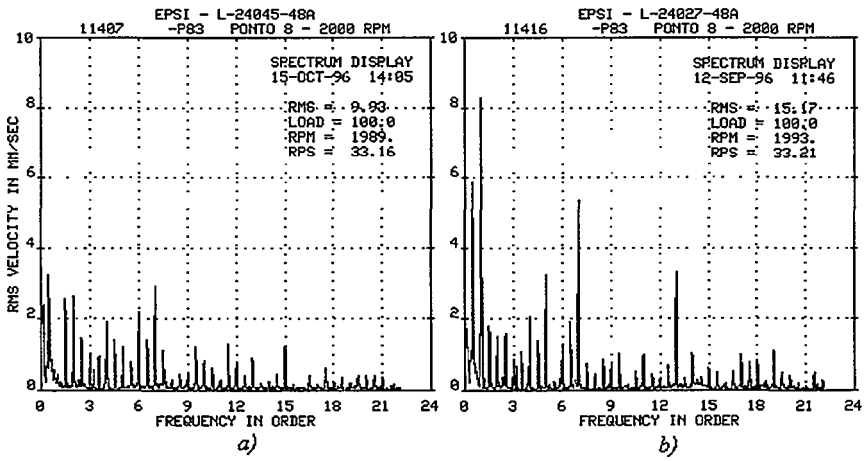


Fig. 5.1 - The fracture in the crankshaft - Spectrum

The 7th harmonic at 2000 rpm corresponds to approximately 233 Hz, i.e., close to the 1st natural frequency in torsion. Admittedly, the dynamic properties of the crankshaft have altered, thus preventing the dynamic absorber (counter-weight) in the crankshaft from carrying out its function.

6. FINAL NOTES

After the dynamic characterization of the engine made in the first part of this paper, in which the characteristic components of the spectra and of the first registers made in the engines were identified and where the typical spectra as well as their normal values were established, it was verified that it was possible to identify a group of problems that came up. Corrective measures were also taken according to the registered values and the improvement of the fleet's engines performance was achieved.

The collected data has been useful to analyse and understand anomalies in engines and to determine the respective causes.

SEAL WATER QUALITY CONTROL AND SEAL CONDITION MONITORING SYSTEM SAFEMATIC LCS (LCS=LEAKAGE CONTROL SYSTEM)

Esa Salovaara, Development Manager, Safematic Oy, Muurame, Finland
Eero Hagrén, Managing Director, Exéns Development Oy, Kajaani, Finland

ABSTRACT

The LCS system has been developed to monitor seal water quality and the condition of seals located in the most demanding environments.

System sensors continuously analyse water from seal units with the central unit providing information for process control and detection of seal failures. This information can be sent to a millwide control system or to a local PC. On request graphical displays of the data are available from the system enabling overall trends of conditions to be observed.

By using the LCS system a seal leakage can be detected earlier than by manual (human) methods, which allows more time to plan required seal changes.

The greatest benefit of LCS is when it is applied to a 'closed seal water system', which often involves corrosive process fluids. Detection of such fluids entering the seal water can be quickly and easily achieved using LCS. As a result the amount of waste water generated by a sealing system is minimised and closing of water circulations can be made in a controlled way.

1 INTRODUCTION

It has become increasingly important for industry to develop processes that are environmentally friendly. The development of the LCS system allows seal water to be monitored based on the individual requirements of each sealing point.

During normal operation the seal water flowing via seals does not require cleaning, but occasionally process liquid can leak into the sealing water.

Mechanical seals can operate faultlessly for up to 5 or 7 years and in large mills there can be as many as 300 seals. An undetected leakage is quite common and as a result requires that all of the seal water must be cleaned. Altogether when each of 300 pumps uses 4-5 l of seal water per minute, the annual water consumption is 600 000 m3 - 750 000 m3. The price of water, being in average USD 0.47 per m3 it means that USD 280 000 - 350 000 is spent on water annually.

2 USING THE LCS SYSTEM

The LCS control system has been designed for the pulp industry to maximise pump runnability, to indicate any seal leakage and to minimise the amount of waste water produced by a mill.

Modern environmental standards are easier to be fulfilled by using LCS-system both for realtime and historical data collection and analysis. All measured and collected data can be checked as a millwide trend or in details sensor by sensor.

LCS helps to detect irregular leakage of process fluid in seal water at a very early stage thus making the replacement of seals predictable. Only the amount of seal water that really needs to be cleaned, shall be sent to the waste water plant. This feature is even more important when using a closed water circulation. In such an environment a single leaking seal can contaminate all of the water in a system and cause malfunctions in process equipment.

By analysing the collected data from the LCS system it is possible to observe whether there are simultaneous process factors (such as pressure spikes) that have contributed to seal degradation.

3 TECHNOLOGY

The LCS system consists of a central unit and a number of sensor boxes, which are connected between the sealing unit and a sewer / recirculation pipe (figure 1). The central unit can be freely located to minimise cabling cost and a maximum of 50 sensors can be connected to a single central unit. Several central units can be connected in the same bus thus enabling applications covering big mills and large areas (figure 2).

Additionally each sensor box has an optical indication of the status of the device such as on/off and ok/alarm.

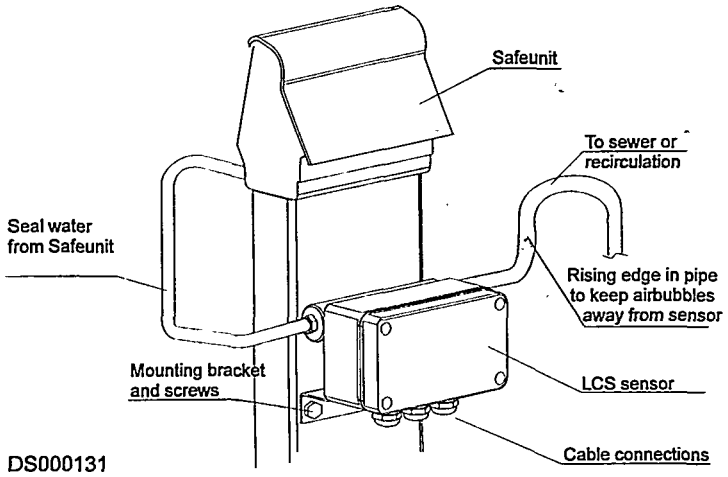


FIGURE 1

3.1 CABLING

A sensor bus cable connects each sensor serially to the central unit. Data transmission and power supply are provided through the sensor bus hence reducing cabling costs in comparison to conventional "sensor-to-central unit" devices. Adding extra sensors is easy and inexpensive, because the extra sensor only requires connection to the last sensor in the system.

LCS-SYSTEM
System structure and connections

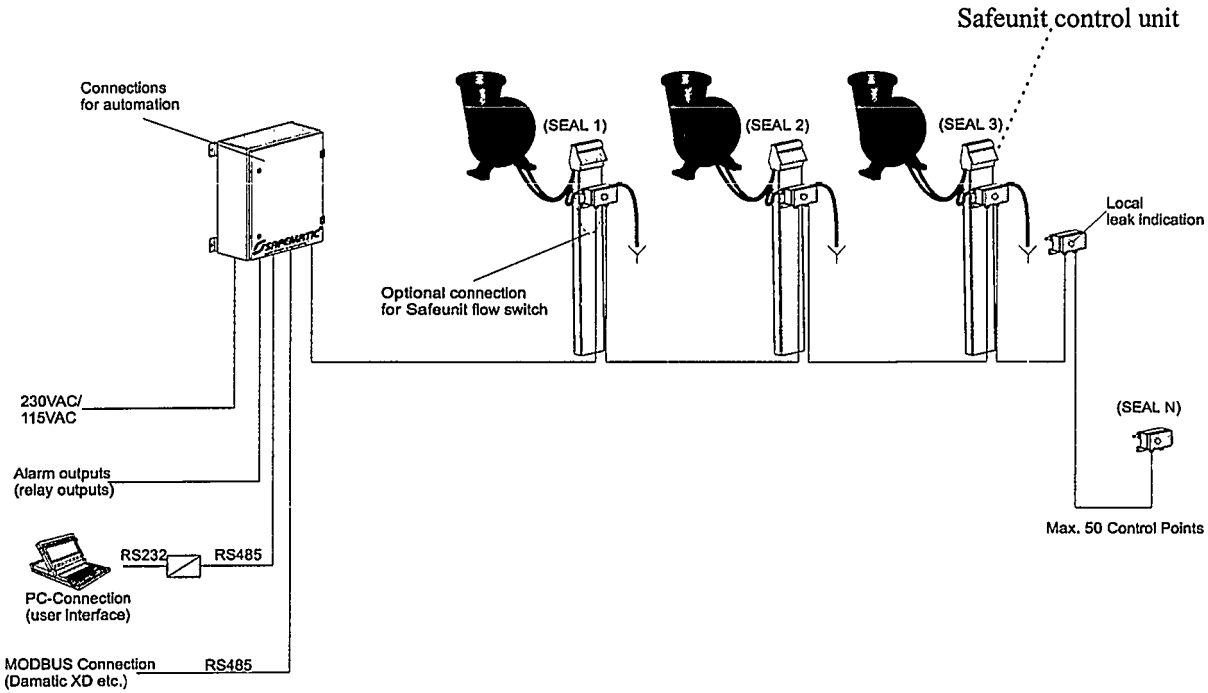


FIGURE 2

Each sensor sends measurement and diagnostic information to a computer on request and all sensors can be controlled and calibrated via the sensor bus from a single PC (figure 3).

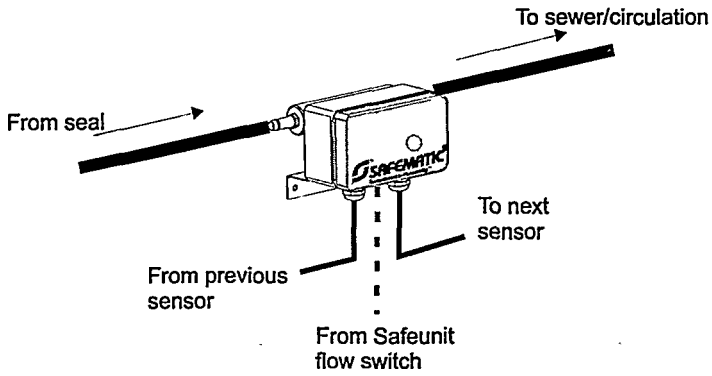


FIGURE 3

3.2 CENTRAL UNIT

The central unit controls each sensor and receives data from each sample point. If any of the sensors fails to respond, the central unit alerts the user to the problem.
Sensor power supply is derived from the central unit.
The data collected can be stored for a set period and analysed to examine system conditions. A leakage can thus be seen and analyzed from the earliest stage and correct measures for maintenance operations made.

Each of the measuring sensors can also been configured with respect to a reference sensor. The reference sensor provides values that describe the quality of the original seal water. The measuring sensors examine the seal water to determine whether contamination has occurred during process operation.
Should the original seal water characteristics change the measurement sensors are automatically reconfigured hence eliminating possible false alarms.

3.3 MILL WIDE CONNECTIONS

LCS System is connected to other system via central unit.A PC can be connected to the LCS system on a permanent basis or only during reception / transmission of data. The PC connection to the LCS is achieved using a RS 485 serial bus. All sensors are configured via this single PC connection.
The LCS system can also be connected to the millwide process control system using a serial bus connection (MODBUS). Monitoring personnel can access updated information concerning seal condition, measurement data and alarm conditions. Alarm levels and diagnostic information can be controlled / manipulated using MODBUS. The same information and rights to modify can also be transmitted to the PC network of the maintenance department as well.
LCS system has relay outputs for controlling selected mill connections according to the measurement results and set alarm levels. These alarm outputs can be grouped to different relay outputs according to the priority of each pump involved.

3.4 ADDITIONAL FEATURES

Each sensor unit has the option to monitor flow rate when the flow switch is connected (see figure 3). Should the flow rate of the seal water decrease too low an alarm message is generated.
All information concerning the sensor units is transferred to a single computer, which allows minimising the number of connections required to the process control unit. Each sensor does not require separate cable connection to the control room, which is a major advantage to the mill production personnel.

3.5 THE OPERATION PRINCIPLE OF THE SENSOR

The original basis for designing the LCS system was to monitor pulp mill seal leakages. The operation of the sensor unit is effectively to measure the conductivity of water. In principle this is achieved using four electrodes, the two of which measure the voltage drop produced by the liquid. Conductivity can be calculated using both set and measured voltage values.

The system has been designed to compensate for any contamination which occurs on the electrodes. Should the problem continue over the limit the sensor alarms and the operation is disabled.
Water temperature also effects conductivity. The temperature of the seal water is measured on each sensor to help calculate real conductivity. It is recommended that the seal water should not exceed +60 °C. From this information it is possible to ascertain whether there is sufficient cooling water be supplied to the seals.

At this first stage LCS is able to indicate leakages of most demanding process liquids like black and white liquors using the conductivity principle. In future the sensors can be developed to operate on different measurement principles (e.g. optical measurements). The LCS system allows easy integration of such intelligent measurement sources.

4 THE OPERATION OF THE SYSTEM IN PRACTICE

A trial system incorporating three sensors has been installed at a pulp mill liquor cooking department since May 1996. The purpose is to measure the seal water from circulation pumps. When required the system produces graphical diagrams showing the conductivity and temperature of any of three pumps for example. From this information it is possible to observe, if a seal has started to leak periodically (as seen in figure 4) and replacement of the seal can be made as a preventative measure. Comparisons of process equipment and process parameters can be made to determine, whether there are any contributing factors to a seal malfunction.
The LCS system can be controlled also from a remote location via a modem and PC.

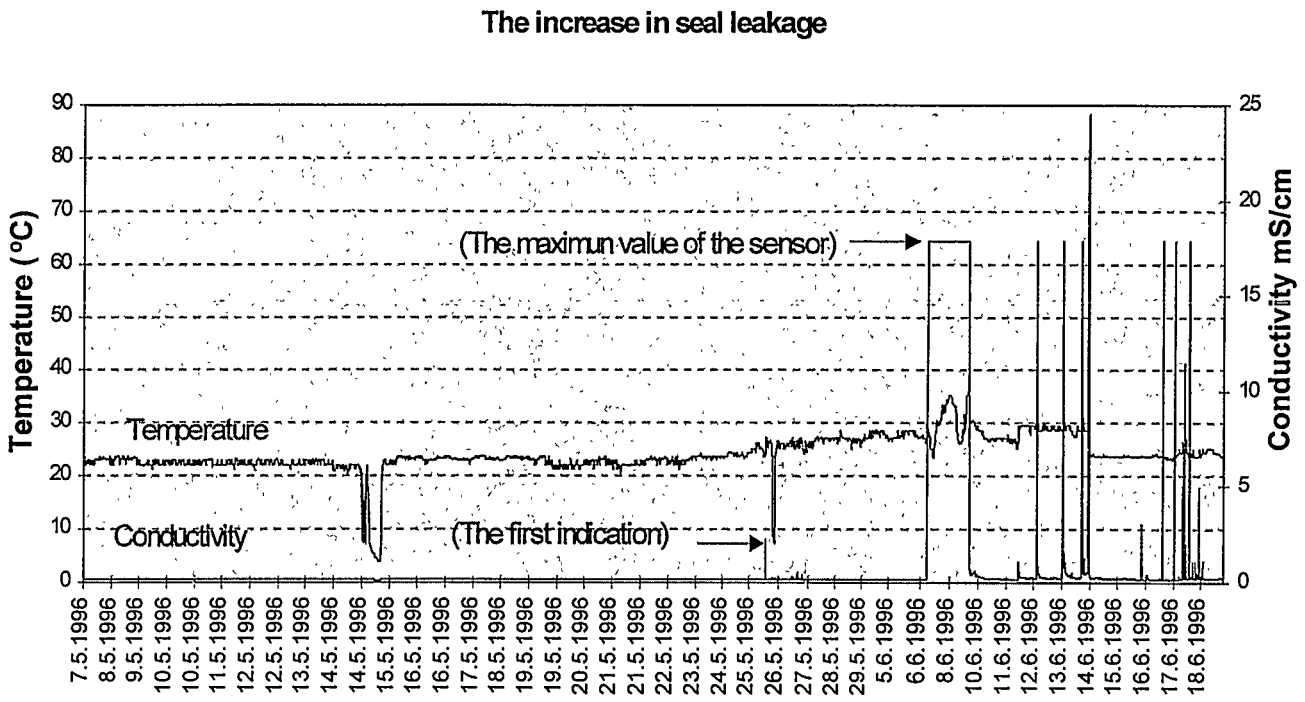

The increase in seal leakage
Temperature (°C)
Conductivity mS/cm
(The maximun value of the sensor)
Temperature
Conductivity
(The first indication)
90
80
70
60
50
40
30
20
10
0
25
20
15
10
5
0
7.5.1996
8.5.1996
9.5.1996
10.5.1996
10.5.1996
11.5.1996
12.5.1996
12.5.1996
13.5.1996
14.5.1996
14.5.1996
15.5.1996
16.5.1996
16.5.1996
17.5.1996
18.5.1996
19.5.1996
19.5.1996
20.5.1996
21.5.1996
21.5.1996
22.5.1996
23.5.1996
24.5.1996
25.5.1996
26.5.1996
27.5.1996
28.5.1996
29.5.1996
5.6.1996
6.6.1996
8.6.1996
10.6.1996
11.6.1996
12.6.1996
13.6.1996
14.6.1996
15.6.1996
16.6.1996
17.6.1996
18.6.1996

FIGURE 4

CHANGE DETECTION OF SENSOR SIGNALS; APPLICATION TO BINARY DISTILLATION PROCESS

Petri Uusitalo and Pekka Häyhä, Research Engineer
Department of Mathematics (DM) and Automation and Control Institute (ACI), Tampere University of Technology (TUT), Tampere, Finland

ABSTRACT

In this paper statistical change detection methods in fault detection and process monitoring are discussed. The goal is to apply statistical methods for abrupt change detection in measurements and residual signals. Both of these signal types are very common in industrial processes. With the help of the change detection methods it is possible to get more reliable information about the behavior of the signals. In practice, change detection methods alone give rarely good enough solution for any fault diagnosis problems, but they are often useful tools in total solutions. The main contribution of this paper is to research the reliability and the usability of the change detection algorithms in real process environment with noisy measurements. Comparison between different algorithms is realized with the help of the real-time tests. The theoretical background of the algorithms is also presented in the paper. Application environment is a pilot binary distillation process at the Automation and Control Institute of Tampere University of Technology.

1 INTRODUCTION

Modern industrial plants are complex systems that contain several machines and a large number of measurements. The Operator gets more and more information (i.e. raw measurements) from the process. In many case the operator can not exploit this information efficiently enough. The role of the operator is changed radically.

Earlier the number of the operators was much larger than nowadays and they were familiar with the characteristics of the process. Today the operator is more the supervisor that controls the operation of the whole industrial plant. They have not time to monitor each measurement or individual machine. For these reasons information should be in very compact and in clear form. One efficient way to package information into the compact form is to use change detection algorithms: only changes in signal lines is detected and advised to supervisors or higher level systems.

For these reasons the use of the diagnosis and condition monitoring systems are continuously increased. The purpose of the fault diagnosis system is to detect failures in process and inform the operator. In this kind of application is possible to use change detection algorithms for many purposes and in many different ways. For example, it is possible to apply statistical change detection algorithms to residual signals that are generated with diagnosis models. Alarm limits to detect abnormal operation of the process are often generated according to modeling uncertainty. The determining of the limit values can be difficult, especially when measurements are noisy. With the aid of change detection methods it may be possible to use stricter alarm limits, because the alarm threshold can be defined more sophisticated.

2 APPLICATION ENVIRONMENT

Application and test environment is the pilot binary distillation process at the Automation and Control Institute of Tampere University of Technology. The control of the distillation column has been performed by Valmet DSS 2000 automation system. In the control room there are two pentium PCs. The user interface has been made with SCADA (Supervisory Control and Data Acquisition) program for Windows NT environment. This kind of open automation system is a good environment to test and develop these algorithms, because much more engineering and software tools are available than in conventional automation systems with closed architecture [2].

The change detection algorithms are integrated to the automation system. For example the operation of some process measurements is monitored by these algorithms. We have also a fault diagnosis system for the temperature measurements of the column. In this application changes in residual signals are monitored [2].

3 THEORETICAL BACKGROUD

3.1 Base of The Used Algorithms

The starting point of the algorithms is independent random variables $(y_k)_k$. The probability density function (pdf) $p_\theta(y)$ of the random variables depends on one scalar parameter θ. Before the unknown change time t_0, parameter θ is constant and equal to θ_0 and after change it is equal to θ_1 ($\theta_0 \neq \theta_1$). The problem is to detect the change in the parameter on-line as soon as possible. In next sections some basic algorithms are introduced briefly [1].

Common base for the used algorithms is the logarithm of likelihood ratio (LLR):

$$s(y_i) = \ln \frac{p_{\theta 1}(y_i)}{p_{\theta 0}(y_i)}. \tag{3.1}$$

Suppose probability density functions $p_{\theta 1}$ and $p_{\theta 0}$ have expectations $E_{\theta 1}$ and $E_{\theta 0}$ then $E_{\theta 0}(s) < 0$ and $E_{\theta 1}(s) > 0$ i.e. if change happens in the parameter θ the sign of the mean value of the log-likelihood ratio changes.

3.2 Limit Check Algorithm and Shewhart Control Chart

Limit check algorithm is based on the idea: take sequences of N samples long from signal and test every sequence under hypotheses H: $\theta=\theta_0$ and G: $\theta=\theta_1$. Testing is repeated until hypothes G is true for parameter θ i.e. change is detected. The sum of LLR's, $S_j^k = \sum_{i=j}^{k} s_i(y_i)$, is used in the testing.

The optimal decision rule d and the stopping rule (set-point of alarm time) t_a for this algorithm is

$$d = \begin{cases} 0 \text{ if } & S_1^N < h; H \\ 1 \text{ if } & S_1^N \geq h; G \end{cases} \text{ and } t_a = N \min\{k : d_k = 1\}, \tag{3.2}$$

where d_k is a decision rule for a sample number k, size of the sample is N and h is chosen threshold. Sum equation S_1^N is called decision function.

Stewhart control chart is basically the same as limit checking algorithm. It is only formulated and reparametrized to be easy to calculate with a computer. The algorithm is: the alarm is set the first time when

$$|\bar{y}(k) - \mu_0| \geq \kappa \frac{\sigma}{\sqrt{N}}, \tag{3.3}$$

where $\bar{y}(k)$ is mean of k first samples of sequence under testing, σ is estimated variance of data, μ_0 is estimated value θ_0 of parameter θ and κ is tuning parameter for alarm threshold.

3.3 Weighted Average Algorithms

Geometric moving average control chart (GMA) is based on two ideas. The first is the use of LLR and the second is the exponential weighting of observations. In this case decision function can be formulated as

$$g_k = \sum_{i=0}^{\infty} \gamma_i \ln \frac{p_{\theta 1}(y_{k-1})}{P_{\theta 0}(y_{k-1})} = \sum_{i=0}^{\infty} \gamma_i s_{k-i}, \qquad \gamma_i = \alpha(1-\alpha)^i, \quad 0 < \alpha \leq 1 \quad , \tag{3.4}$$

where α is called a forgetting factor. The alarm time is

$$t_a = \min\{k: g_k \geq h\}. \tag{3.5}$$

Finite moving average control chart (FMA) is same method as GMA. The only difference is the choice of weights γ_i. In FMA the weights are coefficients of chosen causal filter.

Filtered derivative algorithm (FDA) uses the same ideas as the GMA. The differences are: the weights in equation (3.4) are set to ones and instead of g_k is used derivate ∇g_k. The changes are detected with ∇g_k and stopping rule

$$\nabla g_k = g_k - g_{k-1} \text{ and } t_a = \min\{k: \sum_{i=0}^{N-1} 1_{\{\nabla g_{k-1} \geq h\}} \geq \eta\}, \tag{3.6}$$

where η is the number of alarms allowed before stopping time, when $\nabla g_{k-1} \geq h$ which is chosen threshold. Derivate ∇g_k is usually formed by using the weighting of several differences of g_k. In our study we used integrating filter with constant weights and Bartlett -windowed filter.

3.4 Cumulative Sum of Differences

For *cumulative sum of differences* CUSUM algorithms decision function, decision rule and stopping time are derived from basic forms:

$$S_j^k = \sum_{i=1}^{k} \ln \frac{p_{\theta_1}(y_i)}{p_{\theta_0}(y_i)},\ g_k = \max_{1 \le j \le k} S_j^k \text{ and } t_a = \min\{k : g_k \ge h\}. \quad (3.7)$$

Typically cumulative sum of difference S_j^k decreases before the change and increases after the change. Because the detection of the change is difficult if S_j^k is very small, CUSUM algorithm can be restarted if S_j^k decreases under the chosen threshold ε . These algorithms are called *repeated sequential probability ratio CUSUM algorithms* (SPRT). For SPRT the decision rule and the stopping time can be stated as

$$d = \begin{cases} 0, \text{ if } S_1^T \le -\varepsilon \\ 1, \text{ if } S_1^T \ge h \end{cases} \text{ and } T = T_{-\varepsilon,h} = \min\{k : (S_1^K \ge h) \cup (S_1^K \le -\varepsilon)\} \quad (3.8)$$

Where ε and h are chosen thresholds. If d=0 algorithm is restarted and if d=1 alarm is set. If $\varepsilon = -\infty$, the algorithm is called open ended.

When the magnitude of the change in the mean value ν of the signal is known a priori, but the sign of change is not known, it is reasonable to use *two sided CUSUM algorithm.* This algorithm uses two CUSUM algorithms together: one to detect the rise of the mean and one to detect the decrease of the mean. In this case the decision functions upwards and downwards and stopping time can be formulated to recursive form:

$$g_k^+ = \left(g_{k-1}^+ + y_k - \mu_0 - \frac{\nu}{2}\right)^+,\ g_k^- = \left(g_{k-1}^- - y_k + \mu_0 - \frac{\nu}{2}\right)^+ \text{ and } t_a = \min\{k : (g_k^+ \ge \bar{h}) \cup (g_k^- \ge \bar{h})\}, \quad (3.9)$$

where $(x)^+$ denotes sup(0,x).

Open ended reverse time repeated sequential probability ratio CUSUM algorithm (OERTRSPV) is obtained, if decision function in equation (3.8) is rewritten. The new decision rule can be stated as

$$d=1, \text{ if } \sum_{i=j_0}^{k}(y_i-\mu_0-\frac{\nu}{2})\geq \bar{h}, \tag{3.10}$$

where $\bar{h}$ is not constant threshold anymore and j_0 is restart time. $\bar{h}$ is a straight line with slope $\omega\tan(\alpha)$, where ω is a coefficient correcting the scale between horizontal and vertical axes and α is the angle between the line and horizontal. If equation (3.10) is reparametrized, choosing $\tan(\alpha)=\frac{\nu}{2\omega}$ and $d=\bar{h}/\tan(\alpha)$, decision threshold forms V-shaped mask due to cumulative sum S_j^k. Now decision rule can be rewritten as

$$d=1, \text{ if } \sum_{i=j_0}^{k}[y_i-\mu_0-\omega\tan(\alpha)]\geq d\tan(\alpha). \tag{3.11}$$

This is an equation for upper side control of S_j^k, but it can be easily used as two sided algorithm using the symmetry of the adaptive threshold. We used in this algorithm in two sided form.

χ^2 *-weighting of CUSUM* is a good choice for the change detection algorithm when the value of the parameter θ is not known after the change. In this case it is assumed that the cumulative distribution of $F(\theta)=e^{LLR}$ contains all possible values of the unknown parameter in chosen interval. In this interval it is possible to estimate change time by weighting $F(\theta)$. χ^2-weighting of $F(\theta)$ leads to CUSUM algorithm which has the decision function

$$g_k=\max_{1\leq j\leq k}\left[\ln\cosh(b\tilde{S}_j^k)-\frac{b^2}{2}(k-j+1)\right], \tag{3.12}$$

where

$$b=\frac{\upsilon}{\sigma},\ \tilde{S}_j^k=\frac{1}{\sigma}\sum_{i=j}^{k}(y_i-\mu_0) \text{ and } t_a=\min\{k: g_k\geq h\} \tag{3.13}$$

are signal to noise ratio, cumulative sum and stopping time.

3.5 Generalised Likelihood Ratio -Algorithm

Generalised likelihood ratio (GRL) uses the same basic idea as the previous algorithm.

However in this case the weighting is replaced by the maximum estimate of the exponent of LLR. Now the observations from time j to time k can be expressed as

$$S_j^k(\theta_1) = \sum_{i=j}^{k} \ln \frac{p_{\theta 1}(y_i)}{p_{\theta 0}(y_i)}. \tag{3.14}$$

This equation contains two independent variables: change time t_a and parameter θ_1 after the change. The decision function is obtained by using double maximisation. Decision function and stopping time are:

$$g_k = \max_{1 \le j \le k} \sup_{\theta_1 \{|\theta_1 - \theta_0| \ge v_m > 0} S_j^k(\theta_1) \text{ and } t_a = \min\{k: g_k \ge h\}, \tag{3.15}$$

where v_m is a known minimum jump.

4 PROCESS EXPERIMENTS

4.1 Background

The first task was to ensure the behavioral characteristics of the algorithms by simulation tests. In the simulation tests data was theoretically correct: normally distributed and changes occurred within one sample period. In the process experiments two datasets with different characteristics were used: the flow measurement of the reflux and the position signal of the control valve. The flow measurement was normally distributed, and it contained many amplitude changes with arbitrary gain and direction. It also included a number of changes in variance and strong disturbances. The position signal of the valve was a typical actuator signal: changes were sharp and variance was constant. In the practical implementations of the change detection system standard digital signal processing methods should be used for proper choice of frequency content of the sensor signals [3]. This enables change detection algorithms to work efficiently and get best possible results. In the tests with experimental data the valve signal was raw measurement and the flow measurement was decimated, because the algorithms didn't work properly without it.

In the tuning of the algorithms it was used reference data from the measurements. First the wanted level of changes of mean was fixed and the corresponding change times were defined. Then the algorithms were tuned to find the given change times as accurately as possible. The amount of false alarms was fixed beforehand. The obtained tuning was used in the actual tests. The performance of the tuned algorithms applied to detect real changes was collected as results. Used performance indexes for the behaviors of the algorithms were (see Table 1.): (1.) mean time between false alarms [samples], (2.) mean delay of detection [samples], (3.) probability of false detection [%], and (4.) probability of nondetection [%].

4.2 Results and comments

The tests with simulated signals indicated that the best choices for the change detection algorithm should be selected from the set of the weighted average type algorithms (3.3). On the other hand the differences between the algorithms were quite small. All algorithms gave good results with the simulated data. The numerical results of the process experiments are presented in the Table 1. With experimental data the differences between different the algorithms were considerably larger than in with simulated data. In the test 1 the position signal for the control valve was used. This case was fairly easy. Test 2 was " the worst case test". It was made with the flow measurement signal. This test was fairly demanding for the algorithms, that can be seen also in the results in the table 1.

Table 1. The results of the detection tests.

	Test 1:				Test 2:			
	1.	2.	3.	4.	1.	2.	3.	4.
1. Limit Check	21433	23.6	3.3	6.5	563	94	32	12.5
2. Shewhart	25330	33.6	1.5	6.5	891	93	22.7	12.5
3. GMA	10716	5.9	6.5	6.5	10687	15	2.5	18.8
4. FMA	5358	4.8	12.2	6.5	5344	19	4.9	18.8
5. FMAI	∞	9.1	0.0	3.2	5344	20	4.3	6.3
6. FMAT	7144	13.8	9.1	3.2	10687	27	2.4	14.6
7. SPRT	21433	37.0	3.9	19.4	254	13	63.6	50.0
8. 2-S.CUSUM	7144	5.1	9.7	6.5	2671	12	8.0	4.2
9. OERTRSPV	10717	49	6.7	9.7	972	34	21.6	16.7
10. χ^2–CUSUM	10717	22.3	6.9	12.9	*			
11. GRL	10717	4.4	6.9	12.9	356	7.8	44.8	22.9

* No results, problems with tuning of the algorithm

Table 2. The summary from the algorithms

	Tuning Parameters	Tunability	Comments
1. Limit Check	2	Easy	Slow, very robust
2. Shewhart	2	Easy	Slow, very robust
3. GMA	3	Quite easy	Fast, fairly robust, fairly reliable
4. FMA	3	Difficult	Fast, fairly reliable
5. FMAI	3	Quite easy	Very reliable, fairly robust
6. FMAT	3	Quite easy	Average
7. SPRT	2	Normal	Unreliable, not work in all case
8. 2-S. CUSUM	2	Normal	Fast, fairly robust, work well
9. OERTRSPV	3	Very diff.	Slow, not very good results
10. χ^2-CUSUM	2	Difficult	Very unreliable, many problems
11. GRL	2	Normal	Very fast, slightly unreliable

All of the tested algorithms operated very good, if the environmental conditions were good, data well prepared and the tuning was optimal. Table 2 presents the characteristics of the algorithms, like robustness, fastness, reliability and ease of tuning in the worst case situation. *Robustness* means that an algorithm is unsensitive for changes in tuning parameters. *Fastness* describes how quickly an algorithm alarms after a change has occurred. *Reliability* means the tolerance of the algorithm against disturbances. With optimally tuned algorithms it is possible to obtain excellent results in the test situation, but in practical applications the robustness of the algorithms guarantees good results in constantly changing environment.

The behavior of the detection algorithm is presented in Fig. 1. Subplot on the left shows the example of the flow measurement signal where occurs a change of 8 [l/h] in the mean of the signal between samples 144 and 149. Subplot on the right shows the response of GRL algorithm to the flow measurement.

Experiments with algorithms that were adapted to detect changes in the variance gave same kind of results than previous ones. The main difference was the behavior of the algorithms after a change: usually algorithms needed a new tuning. The another point was that the detrending of the measurements improved results considerably.

In the future our plan is to combine the change detection algorithms and the knowledge-based methods. The system will consist of a two level. At the first level the changes of the signals are detected. At the second level the reasons of the changes are searched by using fuzzy logic. The decision-making is based on the knowledge-based rules .

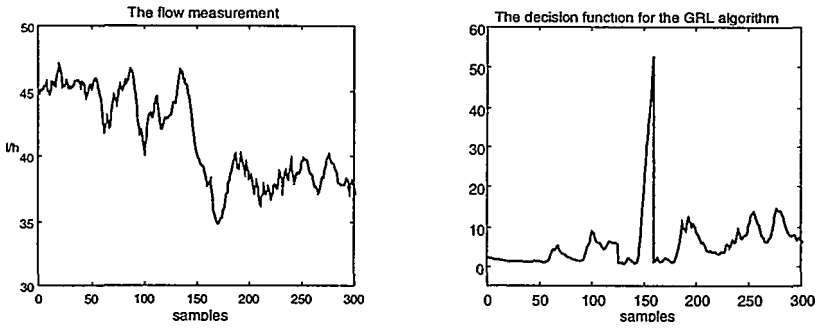


Figure 1. The behavior of the GRL algorithm; The change of the mean in the flow measurement.

5 CONCLUSIONS

The presented and tested change detection algorithms gave promising results. Important issues in the operation of the algorithms were the preprocessing of the signal with the aid of standard digital signal processing, robustness and tunability of the algorithms. Statistical change detection methods are one possibility to find essential information from measurement signals. Signals can be monitored automatically and operators get alarms only when real changes occur in signals i.e. something interesting happens in the process. It seems that it is possible to apply this kind of algorithms to fault diagnosis and process monitoring more widely.

REFERENCES:

1. Basseville, M., Nikiforov, I. *Detection of Abrupt Changes, Theory and Applications.* Prentice Hall Information and System Sciences Series. Englewood Cliffs, New Jersey 07632. 1993.

2. Lautala, P., Majanne, Y., Häyhä P., Airikka P., *Fault diagnosis in open automation environment; application to binary distillation process.* 13 Th IFAC World Congress, San Francisco, USA, June 30 – July 5, IFAC, pp. 91 - 96. 1996.

3. Mitra, S. & Kaiser, J. *Handbook of Digital Signal Processing.* John Wiley & Sons, Inc. New York. 1993.

APPLICATION OF THE GAS-PATH-ANALYSIS (GPA) FOR THE NON-STATIONARY OPERATION OF A JET-ENGINE

Reinhart Lunderstädt, Professor and Rainer Junk, Assistant Professor,
Institute of Automation, University of German Armed Forces,
Hamburg, Germany

ABSTRACT

The Gas-Path-Analysis (GPA) is a powerful tool for modern engine condition monitoring. On the basis of stationary modeling and diagnosis a non-stationary procedure is developed. It can be used during start-up and other non-stationary operations of jet engines, especially in the military field. The theoretical background are nonlinear parameter estimation algorithms and neural networks for the calculation of the non-stationary reference base lines.

1 INTRODUCTION

The high requirements both in the civil and military field of air traffic as well as in the operation of power plants, which may be characterized by the keywords of "safety, reliability and economy", have led to the development of methods of engine condition monitoring at an early stage.

The first automated diagnostic systems were introduced in the early sixties. These systems were designed to supervise mainly the life cycle of important engine components by counting load cycles and temperature strains. System theory regarding diagnostic procedures for jet engines and gas turbines was first considered in the early seventies. These considerations were based on **steady-state** (thermodynamic) models for engine dynamics, that had been deduced from working procedure data processing. Basic achievements in that field have been contributed by URBAN [1]. Further progress was made by ROESNICK [2] and later on by FIEDLER and LUNDERSTÄDT [3] who developed diagnostic software for online operations in many fields of application.

Due to measurement noise and also to systematic measurement errors, the diagnosis comprises the use of estimation procedures. Referring hereto, algorithms for state evaluation were developed and checked on their reliability. Especially for systematic sensor errors a new theory was developed by LUNDERSTÄDT, HILLEMANN and JUNK [4,5], which could meanwhile be proven on two- and three-shaft jet engines and on an appropriate gas turbine with different working points.

The application of the thermodynamic diagnosis, known as Gas-Path-Analysis (GPA), is restricted on stationary models and thereby on a **stationary operation**. Especially in the case of military jet engines stationary working points are not common. Because of this the goal of the paper to be presented is the development of a procedure which leads to a **non-stationary use** of the GPA as well as to a non-stationary modeling of the engine. As an example, the jet engine LARZAC which is used in the French/German training airplane ALPHA JET, is taken for explanation.

The engine consists of five modules: low pressure and high pressure compressor, combustion chamber and high pressure and low pressure turbine. These modules are diagnosed separately by state variables which are the mass flows and the efficiencies of the different modules. The states cannot be measured directly; they must be derived from temperatures and pressures at the in- and outlet of each module. In addition each shaft speed and the fuel consumption are needed.

2 SAMPLING DIAGNOSIS SYSTEM

2.1 Structure of the Diagnosis System

The modeling of the non-stationary engine is done by the presented kind of diagnosis with an approach which requires no exact knowledge of the process. Furthermore the statement is adjusted to the real engine by measurement data. The basis of modeling of the engine is the fact that each (dependent) measurement quantity Y_j can be described as a function of the independent measurement quantities which are in the case of the LARZAC engine:

- ambient atmospheric pressure p_0,
- ambient atmospheric temperature T_0 and
- fuel flow $\dot{m}_F$.

If the inputs p_0, T_0 and $\dot{m}_F$ of the engine are constant then (because of the physics of the engine) after a sufficiently long time the dependent measurement quantities will be constant, too. These final values were the basis of the modeling for the stationary diagnosis used till now. Due to the fact, that these relations are generally valid and assuming that all thermodynamical compensation processes have ended and the system is balanced, they also are valid for each working point. In the case of military engines normally one cannot wait for the stationary states. In addition control processes in the engine cause changes in the fuel flow so that the "stationary" states are only reached internally. Basing on the thermodynamics of the engine the time dependence of the "real" measurement quantities is modeled with two time-delay systems of first order connected in series. The time constants of the two time systems are unknown and have to be determined from the measurement data. By this it is achieved that an optimal adaptation of the model to the real engine is done without a detailed knowledge of the parameters of the system.

The procedure of the state diagnosis consists of two steps. In the first step the mathematical diagnosis model is worked out with reference measurement data. In the second step the state diagnosis is done using the model of the first step. Both steps are characterized by the same procedure of data acquisition. The following steps of further processing then are different.

2.2 Measurement Data Acquisition

The aim of the measurement is to find out the functional relation between the dependent and the independent measurement quantities and the size of the time constants - already mentioned -, which describe the dynamics of the system. The functions are extremely nonlinear. A linearization of these functions in the neighborhood of a fixed performance point simplifies the relation for the engines internal quantities by

$$Y^*(p_0, T_0, \dot{m}_F) = Y_A^* + c_{pA} \cdot \Delta p_0 + c_{TA} \cdot \Delta T_0 + c_{\dot{m}_F A} \cdot \Delta \dot{m}_F \ , \qquad (2.1)$$

with

$$\Delta p_0 = \frac{p_0 - p_{0A}}{p_{0A}}, \Delta T_0 = \frac{T_0 - T_{0A}}{T_{0A}}, \Delta \dot{m}_F = \frac{\dot{m}_F - \dot{m}_{FA}}{\dot{m}_{FA}} \quad .$$

The vector $\underline{Y}_A^*$ represents the quantities of the depended measurement data of the working point. The constants c_i $(c_{pA}, c_{TA}, c_{\dot{m}_F A})$, which are

partial derivatives, are to be determined by reference measurement data. Eq. (2.1) is fulfilled sufficiently well if there are only small changes in the independent quantities Δp_0, ΔT_0 and $\Delta \dot{m}_F$.

The stationary final states of the system can be formed out with the unknown parameters c_i if the input quantities are given. Nevertheless, to determine the parameters of the engine it is also necessary to consider the dynamics of the engine, because the measurement is done during a non-stationary operation of the engine. Therefore it is necessary to calculate the adjusting measurement quantities for a given time behavior of the stationary final value Y^*. For this it is at first necessary to determine the output value x_A of a series connection of two time delays of first order in general if the input value x_E is given. The choice of this kind of mathematical model for the transition behavior follows from the knowledge of the heat transfer and heat conduction which take place during the non-stationary operation of the engine. The determination of the time constants is only possible if the input and output values are known and an appropriate statement for the functional relation between these values is given. In order to avoid the explicit determination of the time functions of the inputs and the outputs the continuous system is changed into a sampling system.

The transition behavior of any linear system in the time domain is determined by the relation

$$\underline{x}(t) = \underline{\Phi}(t - t_k) \cdot \underline{x}(t_k) + \int_{t_k}^{t} \underline{\Phi}(t - \tau) \cdot \underline{b} \cdot x_E(\tau) \cdot d\tau \, , \; t \geq t_k \quad , \tag{2.2}$$

with

$\underline{x}(t)$	:	state vector
$\underline{\Phi}(t)$	:	transition matrix
$\underline{x}(t_k)$	:	state vector for time $t = t_k$ (initial condition),
$\underline{b}$	:	control vector,
$x_E(t)$	:	input of the system.

The state vector $\underline{x}$ consists of system immanent physical quantities including the output of the system. The dynamic behavior is determined by the transition matrix $\underline{\Phi}(t)$ which is the solution of the homogeneous differential equation of the system. The control vector $\underline{b}$ describes the effect of the inputs on the different states of the system. In the case of a time delay system of second order, the state variables

correspond to the outputs of the two subsystems. The transition from the continuous to the discrete system is realized in a way that Eq. (2.2) is evaluated for definite points of time t. For

$$t = t_{k+1} = t_k + T \quad ; \qquad \text{T: sampling time} \tag{2.3}$$

Eq. (2.2) goes over in

$$\underline{x}(t_{k+1}) = \underline{\Phi}(T) \cdot \underline{x}(t_k) + \int_0^T \underline{\Phi}(T-\tau) \cdot \underline{b} \cdot x_E(\tau) \cdot d\tau \quad . \tag{2.4}$$

With Eq. (2.4) the formula for the recursive application and by this for the online realization of evaluation algorithm is given. The state variables of the latest sampling point are used as starting values for the next sampling step so that Eq. (2.4) is evaluated recursively. Thereby it has to be noticed that the function of the input $x_E(\tau)$ in the integral of Eq. (2.4) is the function $x_E(t)$ in the interval $[t_k, t_{k+1}]$. The evaluation of the sampling equation is not directly possible because the time behavior of the input is not known between the two sampling points. Only the values of the sampling points themselves can be used. For a sufficiently small sampling time T it is nevertheless possible to make a linear interpolation. After the determination of the transition matrix $\underline{\Phi}(t)$ the integrals in Eq. (2.4) can be exploited so that the output can be calculated as

$$x_{A,k+1} = f\left(T, x_{E,k}, x_{E,k+1}, T_I, T_{II}\right) \quad . \tag{2.5}$$

The inputs are measured and the sampling time is preset. So it is possible to determine from Eq. (2.5) the dynamic parameters of the system, that means the time constants T_I and T_{II}, if the outputs are measured, too. After the influence of the time delay of second order on the dynamic behavior of each dependent measurement value j is known, the output of the time delay system can be compared with the real measurement value. This comparison is necessary for the numerical determination of the unknown system parameters c_{pA}, c_{TA}, $c_{\dot{m}_F A}$, T_I and T_{II}. For each dependent measurement quantity x_A such a set of unknown parameters exists. The explicit evaluation follows thereby using

$$\begin{aligned} x_{E,k} &= Y_j^*\left(p_{0,k}, T_{0,k}, \dot{m}_{F,k}\right) \\ &= Y_A^* + c_{pA} \cdot \Delta p_{0,k} + c_{TA} \cdot \Delta T_{0,k} + c_{\dot{m}_F A} \cdot \Delta \dot{m}_{F,k} \end{aligned} \tag{2.6}$$

for the input.

For the determination of the system parameters a nonlinear parameter estimation procedure is used which estimates the parameters on the basis of **all** measurement values which are gathered till the actual evaluation time. In Fig. 1 the procedure for the algorithm is outlined which connects the independent and the dependent measurement quantities. By a suitable choice of the unknown system parameters, the dependent measurement quantity is sufficiently well approximated. From Fig. 1 it can be seen that the influence of the inputs of the engine is certainly different on the measurement quantity T_7 (temperature behind the low pressure turbine), which is chosen as an example. By this it can be noticed without any signal analysis that the influence of the fuel consumption will be dominant because the time behavior of T_7 and $\dot{m}_F$ is similar. But also the ambient air pressure and the ambient air temperature are included in the measurement signal; their influence is nevertheless low. This results in small parameters c_{pA} and c_{TA} (in comparison to $c_{\dot{m}_F A}$).

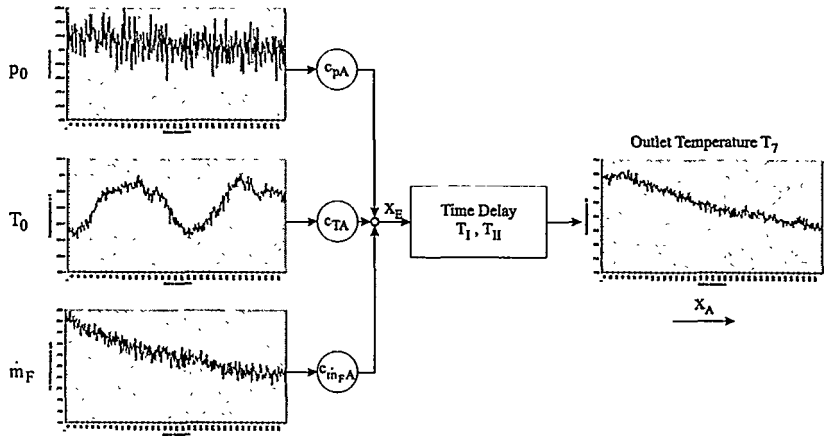


Fig. 1: Influence of the independent measurement quantities on the dependent values for the engine outlet temperature T_7 as an example.

Fig. 2 gives an overview on the behavior of the estimated and the measured outlet temperature after the parameter estimation is finished. The correlation coefficient reached for this example is $r = 0.972$. That means that a high correspondence between both quantities is achieved.

The stationary temperature can now be determined on the basis of the independent measurement quantities and with the help of the esti-

mated coefficients c_{pA}, c_{TA} and $c_{\dot{m}_F A}$. In Fig. 3 the stationary temperature is depicted as function of the measured fuel consumption.

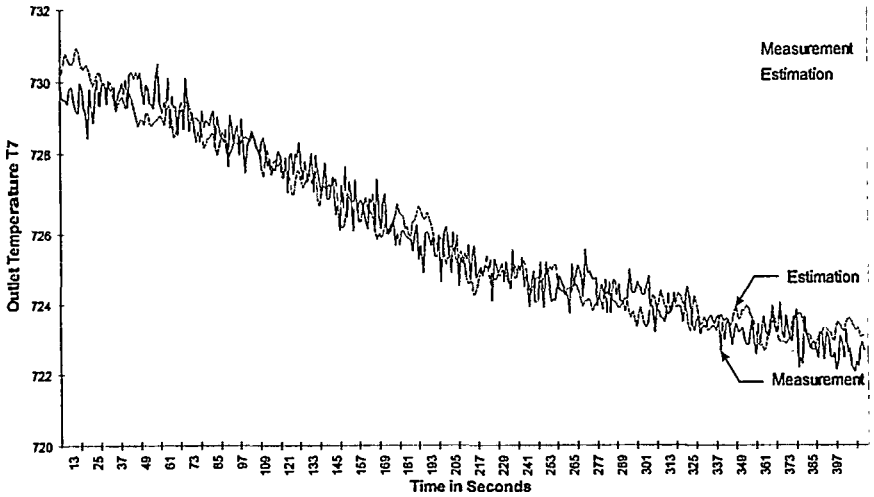


Fig. 2: Estimated and measured outlet temperature T_7.

It can be noticed that a definite association of the dependent measurement quantity is possible using the time model of Fig. 1. By this a high model quality is reached. Deviations from the linearity result from the coincidental changes in p_0 and T_0.

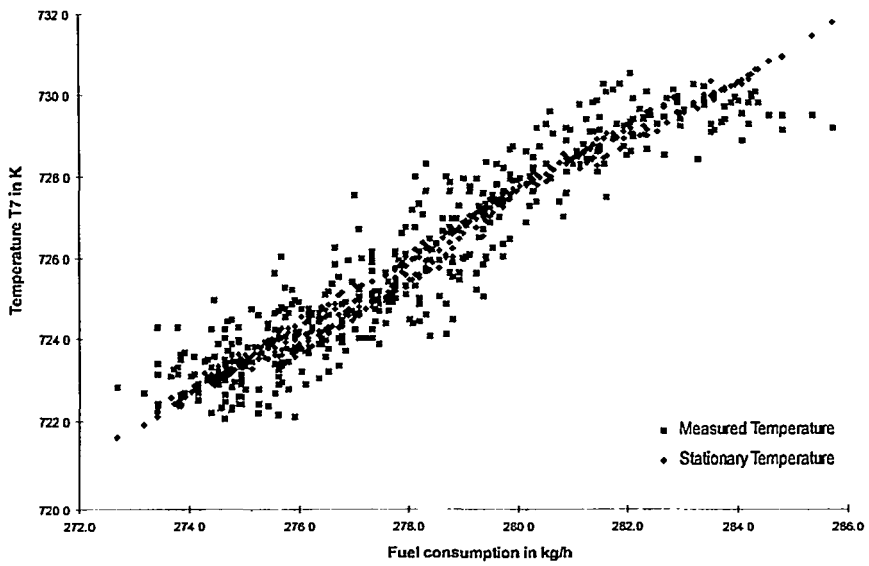


Fig. 3: Stationary outlet temperature T_7 as a function of the $\dot{m}_F$.

2.3 Modeling

As in the case of the stationary diagnosis the fault free system is described by a reference model. In the stationary case **one** reference measurement vector is needed for each working point. For the non-stationary diagnosis these few reference points are not sufficient. For the whole operating range a field of characteristics must be established from which the reference measurement vectors can be taken also under varied input quantities.

For this each dependent measurement quantity is handled by a neural network fulfilling this requirement. During the modeling the reference measurement data are stored in their nets and they are recalled from there during the diagnosis. For the quantitative diagnosis the diagnosis matrix of the specific working point is addionally necessary. It establishes the connection between the deviations of the measurement quantities at the reference point and the real change of the state variables. The derivation of the diagnosis matrices is done in the same way as in the stationary case [3]. Five diagnosis matrices are provided for the operating range between 40 % and 100 % thrust level. In the case where a working point is situated between two reference points, a matrix is used which is derived by linear interpolation.

2.4 Online Diagnosis

Also during the diagnosis procedure at first the dynamic parameters of the engine are determined to find out the stationary final values of the dependent measurement quantities. Additional measurement time for the diagnosis is not necessary because the measurement data, which are required for calculating the engine parameters, can be used again. Therefore the actual measurement time is restricted to a few minutes. After this the diagnosis is done on the basis of the measured data. For this at first the nominal values of the fault free system are ascertained. In dependence of the independent measurement quantities belonging to the diagnosis run of the engine, the nominal values are calculated with the help of the neural networks. For this the independent measurement quantities are applied to the nets and the reference values are calculated in the recall phase. Because the neural nets are built up internal by simple mathematical functions this calculation is done very fast and without time lag. Subsequently the deviations of the actual measurement values from the nominal values of the measurement quantities are calculated for each sampling point k. In order to use these deviations for a quantitative statement for the state variables, they have to be multiplied with the diagnosis matrix that repre-

sents the specific working point. Because of the modeling only a restricted number of working points is at disposal. The appropriate diagnosis matrix has to be calculated actually by interpolation between two working points in dependence of the fuel consumption.

3 RESULTS

In order to test the diagnosis algorithm, one module of the LARZAC engine was changed. The high pressure turbine was selected because there was one engine available that had severe damages in the sealings of the angular clearance. After a three-minute-diagnosis run the dynamic parameters of the system were determined. The calculation of the states of the engine on the basis of the acquired measurement data took about thirty seconds. The output of the diagnosis program can be seen in Fig. 4.

The diagnosis records a small increase of the mass flow and a high decrease of the efficiency of the high pressure turbine (states x_5 and x_6). The small changes of all other states are less than 1 % and permit no further condensation. Because of the high change in the efficiency, it is easy to determine that the high pressure turbine is defect. The example shows impressively that with a small measurement time a diagnosis of the engine can be carried out. Thereby it is not necessary to have special test stand conditions. Plans for the future comprise data acquisition inflight, provided all sensors are available. Additional costs for fuel will not come up because the start-up of the engine, and the stopping, respectively can be used for data acquisition. The diagnosis procedure yet used in practice reduces costs in view of fuel consumption, apparatus expense and man power. The price of the system is very low because it is for the most part only software oriented.

4 OUTLOOK

The modeling and the diagnosis outlined base on the assumption of fault free measurement data. But, the theory of filtering and compensation of fault affected measurement data [4, 5] can be extended from the stationary to the non stationary case considered in the paper. The presentation will include very new results for the non stationary fault affected measurement data handling. These results originate from simulations and actual test stand data for the LARZAC and the RB 199 jet engine of trinational TORNADO military airplane.

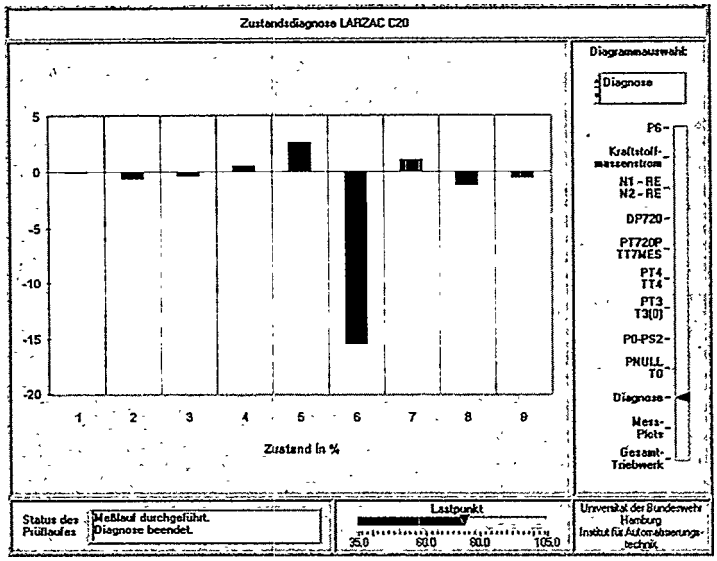


Fig. 4: Output of the diagnosis results for a LARZAC engine with defect high pressure turbine.

5 REFERENCES

1. Urban, L. A.: Gas Path Analysis - A tool for engine condition monitoring. 33rd Annual International Air Safety Seminar, Christchurch, N.Z., 1980.
2. Roesnik, M.: Eine systemtheoretische Lösung des Fehlerdiagnoseproblems am Beispiel eines Flugtriebwerkes. Dissertation, Fachbereich Maschinenbau, UniBwH, Hamburg, 1984.
3. Lunderstädt, R.; Fiedler, K.: Zur systemtheoretischen Diagnose von Strahltriebwerken. Automatisierungstechnik (at) 33 (1985), S. 272-279, 313-317.
4. Lunderstädt, R.; Hillemann, Th.: Sensor Fault Detection for Gas Turbines with Knowledge Based Methods. Preprints of 12th IFAC Congress, Sydney, 1993.
5. Junk, R.; Lunderstädt, R.: Model Optimization of a Gas Turbine with Fault Affected Reference Measurements by Neural Networks for State and Sensor Diagnosis. Preprints of 5th European Propulsion Forum, Pisa, April 5-7, 1995.

FAULT DETECTION AND ANALYSIS IN ROTATING MACHINERY USING HIGHER ORDER TIME FREQUENCY METHODS

S. K. Lee * **and P. R. White** , E-mail *: skl@isvr.soton.ac.uk
Signal Processing and Control Group, ISVR(Institute of Sound and Vibration Research), University of Southampton, Highfiled Street Southampton SO17 1BJ, England

ABSTRACT

Impulsive acoustic and vibration signals within rotating machinery are often induced by irregular impacting. Thus the detection of these impulses can be useful for fault diagnosis. Recently there is an increasing trend towards the use of higher order statistics for fault detection within mechanical systems based on the observation that impulsive signals tend to increase the kurtosis values. We show that the fourth order Wigner Moment Spectrum, called the Wigner Trispectrum, has superior detection performance to the second order Wigner distribution for typical impulsive signals found in a condition monitoring application. These methods are also applied to data sets measured within a car engine and industrial gearbox.

1 INTRODUCTION

Acoustic and vibration data from rotating machinery have been used for fault detection for a long time. These data usually consist of tonal signals, periodic impulses due to the fault and broadband noise. The impulsive components are often hidden by the competing signal components. Recently a pre-processor, dubbed the two-stage ALE (Adaptive Line Enhancer) [2], has been proposed for enhancing the impulsive signals prior to analysis with a bilinear time-frequency method. However, at low Signal to Noise Ratios (SNRs) significant residual noise can remain even after enhancement. This residual noise is generally dominated by the broadband components which can serve

to cause significant problems in the bilinear time-frequency analysis. We seek to explore the use of alternative time-frequency tools for this last analysis stage. Specifically we seek methods which are inherently more robust in the presence of random (Gaussian) noise. Higher Order Spectral (HOS) analysis has been considered for the detection of impulsive signals in Gaussian noise [3]. Time-frequency distributions using HOS have been developed and applied to the impulsive signals [3]. This method is based on the higher order extensions to the WVD. In this paper WHOMS (Wigner Higher Order Moment Spectra) are discussed in detail, with emphasis on the cross-terms and ambiguity function. The detection performance of the WHOMS is compared with that of the WVD via Receiver Operating Characteristic (ROC) curves[4]. It is also shown, using measured data sets, that this approach can be successfully applied to the detection of faults in a car engine and to data from an industrial gearbox. Finally it is shown that useful quantitative information can be extracted from this analysis about the timing of the impulses and their spectral characteristics.

2 HIGHER ORDER WIGNER-VILLE MOMENT DISTRIBUTION

The n+1th order WHOMS is defined as [3]

$$W_{n+1}(\tau, f_1, .., f_n) = \int_{\tau_1} .. \int_{\tau_n} R^i_{n+1}(t, \tau_1, .., \tau_n) \prod_{i=1}^{n} e^{-2j\pi f_i \tau_i} d\tau_i \tag{1}$$

where

$$R^i_{n+1}(t, \tau_1, ..., \tau_n) = s^*(t-\alpha) \prod_{i=1}^{n} s(t+\tau_i-\alpha) \tag{2}$$

In order to extend most of the properties of the WVD, the lag centring condition has to be imposed, leading to

$$\alpha = \frac{1}{n+1} \sum_{i=1}^{n} \tau_i \tag{3}$$

and thus WHOMS is rewritten as

$$\begin{aligned} W_{n+1}(t, f_1, .., f_n) = \int_{\tau_1} .. \int_{\tau_n} s^*(t - \frac{1}{n+1} \sum_{m=1}^{n} \tau_m) \\ \prod_{i=1}^{n} s(t + \frac{n}{n+1}\tau_i - \frac{1}{n+1} \sum_{k=1, k\neq i}^{n} \tau_k) e^{-2\pi j f_i \tau_i} d\tau_i \end{aligned} \tag{4}$$

Therefore we can obtain the Wigner Bispectrum (WB) and the Wigner Trispectrum (WT) using equation (4) as follows;

(a) WVD (n=1)

$$W_2(t, f) = \int_{\tau} s^*(t - \tau/2) s(t + \tau/2) e^{-2\pi j f \tau} d\tau \tag{5}$$

(b) WB (n=2)

$$W_3(t,f_1,f_2)=\iint s^*(t-\frac{\tau_1+\tau_2}{3})s(t+\frac{2\tau_1-\tau_2}{3}) \\ s(t+\frac{2\tau_2-\tau_1}{3})e^{-2\pi j(f_1\tau_1+f_2\tau_2)}\,d\tau_1\,d\tau_2 \tag{6}$$

(c) Symmetric WB (n=3)

$$W_4(t,f_1,f_2,f_3)=\iiint s^*(t-\tau)s(t-\tau+\tau_1)s(t-\tau+\tau_2) \\ s(t-\tau+\tau_3)e^{-2\pi j(f_1\tau_1+f_2\tau_2+f_3\tau_3)}d\tau_1 d\tau_2 d\tau_3 \tag{7}$$

where

$$\tau=\frac{\tau_1+\tau_2+\tau_3}{4} \tag{8}$$

3 SLICE WIGNER HIGHER ORDER MOMENT SPECTRA

Whilst the WVD has significant resolution advantages over other time-frequency methods, its application is dogged by problems associated with cross/interference terms. If we consider a signal x(t) which consists of two components $x_1(t)$ and $x_2(t)$, such that :

$$x(t) = x_1(t) + x_2(t) \tag{9}$$

Consider a specific forn of the two components,

$$x_1(t)=s(t-t_1)e^{-2\pi j\alpha t} \tag{10.a}$$

$$x_2(t)=s(t-t_2)e^{-2\pi j\beta t} \tag{10.b}$$

where s(t) is a prototype signal, t_1 and t_2 are time shifts and α and β are frequency shifts. In general, the n+1th order Wigner distribution for a two component signal is the sum of 2^{n+1} distributions, of which two are auto terms whilst 2^{n+1}-2 are cross terms. See [6] for the detail analysis on cross terms of WHOMS.

The full WHOMS is a function of n+1 variables, time and n frequency variables. This presents problems when attempting to display results, since the resulting plots exist in n+2 dimensional space. Further, there is also a heavy computational burden as one is required to perform n dimensional FFTs for each time segment. For these reasons it is common to consider a subset of the WHOMS called the principal slice. The principal slice consists of the plane defined by $f_1 = f_2 = f$; all values of t are considered. This slice includes both auto-terms and cross-terms, however the number of cross-terms is significantly reduced. Figure 1 shows a pictorial explanation of the principal slice for the WHOMS using a signal x(t) which consists linear sum of $x_1(t) = A_0 \cdot \exp(j\omega_1 t)$ and $x_2(t) = A_0 \cdot \exp(j\omega_2 t)$. In this figure the black circles are auto-terms and the circles represent the cross-terms, the "P"

denotes the principal slice. Figure 1(a) shows the WVD for a signal x(t) at arbitrary time t_0. According to reference [6] the cross-terms are real and oscillate along the time axis. Figure 1(b) shows the WB for a signal x(t) at arbitrary time t_0 there are 6 cross-terms. However, if we use the principal slice, these are reduced to only two cross-terms. But these two cross terms are complex. In particular even auto-terms in the WB are complex in nature; consequently it is common to consider only the magnitude of the WB. The SWB (Sliced WB) is obtained by setting $f_1 = f_2 = f$ in the definition of the WB. Figure 1(c) shows the WT for x(t) at arbitrary time t_0. Similarly to the SWB the SWT (Slice WT) is obtained by setting $f_1 = f_2 = -f_3 = f$ in the definition of the WT. From reference [6] we see there are 14 cross-terms in the WT but, in this example (assuming $\omega_1 \neq \omega_2$), if we use the SWT these reduce to only two. These two cross terms in the SWT are real and oscillate between the two auto-terms. The auto-terms in the WVD and the SWT are sinusoidal with amplitudes A_0^2 and A_0^4 respectively. If Gaussian random noise n(t) with SNR $(\sigma_0^2 / A_0^2) < 1$ is added to x(t), where σ_0^2 is noise variance, then the relative magnitudes of the SWT for signal x(t) to that for noise n(t) are larger and clearer than the relative magnitudes of the WVD. Figure 2 illustrates this result. Figures 2(a),(b) and (c) show the WVD, SWB and SWT respectively, for x(t) with added Gaussian noise (σ_0^2=0.5). Therefore it is identified that the SWB and SWT for deterministic signals are less sensitive to Gaussian random noise than the WVD. However both the SWB and the SWT for the signal x(t) still suffer from cross terms. These cross terms can be smoothed via the ambiguity function and an exponential kernel[1,6].

4 DETECTION PERFORMANCE OF THE WVD AND THE SWT

In this section we examine the various distributions for a simple synthetic transient signal. The model used is a complex sinusoid modulated by a Gaussian pulse, with additive Gaussian noise. Specifically

$$x(t) = \left(\frac{\gamma}{\pi}\right)^{1/4} e^{-\gamma(t-t_1)^2} e^{j\omega_0 t} + n(t) \tag{11}$$

The parameter γ controls the width of the Gaussian pulse. For this signal the WVD and SWT are computed and are depicted in Figures 3(a) and (b). The kurtosis of this signal is 5.2 and the SNR is -3.4dB. This demonstrates that the SWT yields a result in which the signal

component is more visible than in the WVD. This implies that the SWT may be a more efficient detector of impulsive signals and hence be able to detect faults at an earlier stage within a condition monitoring application. Next we formalise these ideas via a more critical assessment of the performance of WVD and WT as detectors.

The detection performance of the WVD and WT are compared by estimating the ROC curves for the synthetic Gaussian pulse data set described previously. The WVD and WT for this signal x(t) are shown in Figures 3(a) and (b). The estimated ROC curves for this data are computed using 1000 Monté Carlo simulations. This ROC curve is shown in Figure 3(c). From this figure, the superior detection performance of the WT to the WVD is evident.

As a second example the model (11) is modified so that the kurtosis is less than 3, and the SNR is large. These results are shown in Figure 4. Both the WVD and WT produce acceptable distributions, as shown in Figures 4(a) and (b). Further the ROC curves, shown in Figure 4(c), are both ideal curves in that they show that, for this data set, it is possible to select a threshold such that the probability of detection is 1 and probability of false alarm is zero (to within estimation errors).

5 APPLICATION OF WHOMS TO FAULT DETECTION IN ROTATING MACHINERY

The impulsive signals generated by faults in a rotating machine are generally immersed in a variety of harmonic and broadband signals. If such raw signals are analysed using the WVD (or SWT) the plethora of cross-terms created makes interpretation of the time-frequency plot extremely difficult. To overcome this problem we pre-process our signals, prior to the time-frequency analysis, to reduce tonal signals and to eliminate some of the broadband noise using methods described in [2], referred to as the two-stage ALE (Adaptive Line Enhancer). In this section we compare the use of the WVD and SWT for use in the time-frequency analysis of the pre-processed signals.

The lower trace in figure 5(a) shows the time series of an acoustic signal radiated from an automotive engine in which the spark plug is loose[2]. The data which has been pre-processed using the two-stage ALE [2] is shown in temporal axis of Figure 5(b). In the pre-processor output the impulsive signal has clearly been enhanced. Figure 5(a) shows the WVD of the raw data. The WVD of the enhanced data

shown in Figure 5(b). In Figure 5(a) the cross-terms effectively mask the impulsive signal. In Figure 5(b) the impulsive signal is readily identifiable and we estimate that the impulsive signal occurred at a crankshaft angle of 530°, however, it is difficult from this plot to identify the dominant frequencies in the impulse. If we apply the SWT to the enhanced signal, the impulse can be clarified by utilising an exponential kernel [1], to create Figure 5(c). This allows us once again to conclude that the impulse occurs at a crank angle of 530°, but we can now further identify that the two frequency components, at 85th order and 95th order (1 order = 14.7 Hz), appear to be the dominant terms.

The problem of fault detection in industrial gears has received wide attention [5]. The lower trace in figure 6(a) shows a time averaged data of vibration signal measured from faulty gear. In order to remove the regular signal, *i.e.* the periodic signals at the fundamental and harmonics of the tooth meshing frequency, the two-stage ALE is again applied, the resulting (residual) signal is shown in figure 6(b). Figure 6(a) shows the WVD of the raw signal. Again the masking effects of the cross-terms serve to hide the impulsive signal. However, the WVD of the pre-processed signal, Figure 6(b), reveals some information about the signal structure but is still unsatisfactory. If we apply the SWT to the residual signal, it can be further enhanced by smoothing with an exponential kernel, the results are shown in figure 6(c). From these results we surmise that that the gear fault occurs at a shaft angle of 210° and is centred on a frequency corresponding to the 48 order of shaft rotating speed. From these results it appears that the SWT may offer a more effective detector for measured data sets. To check this hypothesis we compute the ROC curves based on the measured data sets.

To study the performance of these algorithms as detectors, we compute the ROC curves for the WVD and SWT exploiting measured data. This test used the pre-processed impulsive signal in the region of a shaft angle of 530° and 210° shown in Figure 5(b) and 6(b) respectively. These impulsive signals were then added to 1000 realisations of Gaussian white noise in order to allow estimation of the ROC curves, which are plotted in Figure 5(d) and 6(d) respectively. From Figure 5(d) and 6(d) we see that on this measured data set the SWT outperforms the WVD as a detection tool.

6 CONCLUSIONS

This paper has demonstrated that the WHOMS (Wigner Higher-Order Moment Spectra) have potentially significant advantages over the WVD in a condition monitoring environment. In order to exploit WHOMS it is necessary to work in some reduced domain, in this case the slice domain proved effective. The SWT has the advantage over the SWB of being easier reduce cross-terms by smoothing. It is also a more natural tool since the SWT utilises the fourth order moments of a signal, which have previously been considered for condition monitoring. It has been demonstrated that visually the results from the SWT are more pleasing than those of the WVD. Moreover rigorous analysis, based on ROC curves, has demonstrated that its performance is better than WVD for detection purposes.

ACKNOWLEDGEMENTS

The authors sincerely thank Dr. P. D. McFadden for supplying his measured data found in reference [5].

REFERENCES

1. J. K Hammond and P. R. White, 1996, *Journal of Sound and Vibration* **190**(3), 419-447. The Analysis of Nonstationary Signals using Time-Frequency Methods.
2. S. K. Lee and P.R White, 1996, *IEE Colloquium on Modelling and Signal Processing for Fault Diagnosis*, Ref. No.1996/260,1/6-6/6. Fault Diagnosis of Rotating Machinery Using a Two-stage Adaptive Line Enhancer.
3. J.R. Fonollosa and C.L Nikias, 1993, *IEEE Trans. Signal Processing*, **41**(1), 245-266. Wigner-Higher-Order Moment Spectra Definition, Properties, Computation and Application to Transient Signal Analysis.
4. H.L. Van Trees, 1968, ***Detection, Estimation and Modulation Theory*** John Wiley & Sons New York
5. J. Wang and P. D. Mcfadden, 1993, *Mechanical Systems and Signal Processing* **17**(3),193-203, Early Detection of Gear Failure by Vibration Analysis -I and II Calculation of the Time-Frequency Distribution.
6. S.K Lee and P. R. White, 1997, *ASME 16th biennial conference* submitted paper ,"Identification of fault signals in rotating machinery using higher order time-frequency analysis" .

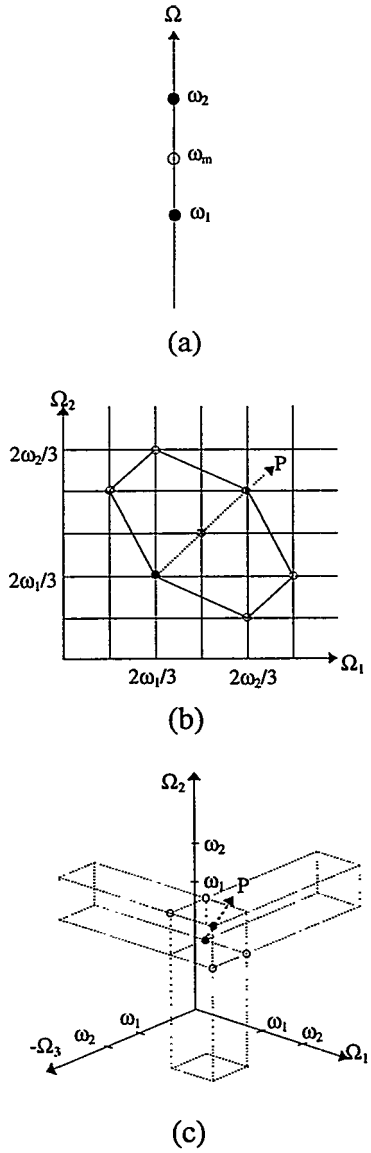


Figure 1 Pictorial explanation for the WHOMS (a) WVD at time t_0 (b) WB at time t_0 (c) WT at time t_0 : "P" designates the principal slice, circles are the cross-terms and black circles are auto-terms.

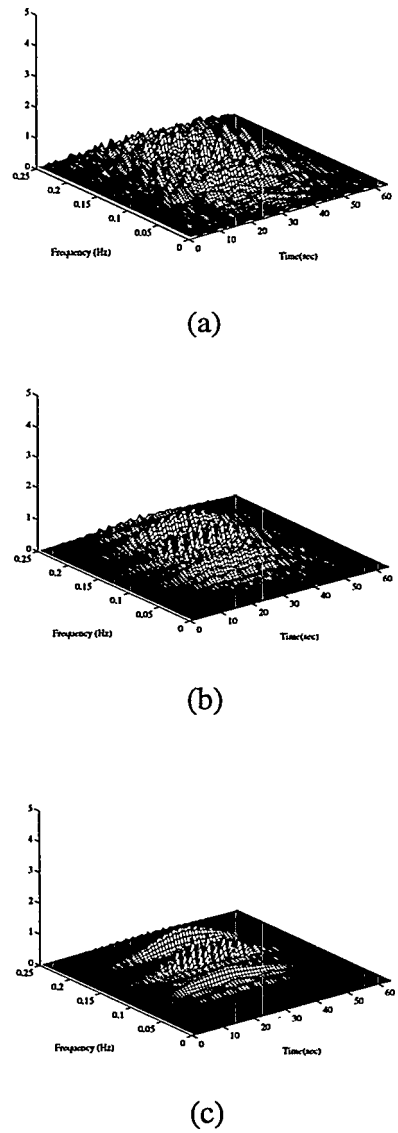


Figure 2 A comparison of the effect of Gaussian noise added to signal x(t) (a) WVD (b) SWB (c) SWT

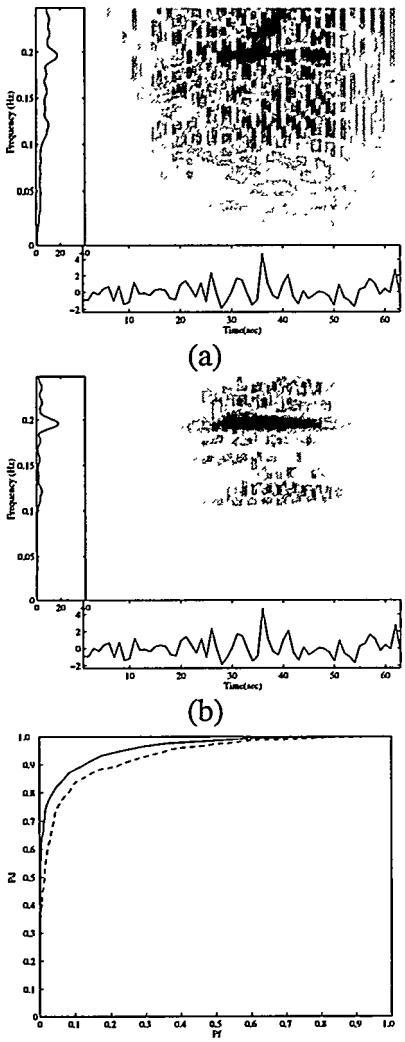


Figure 3 Comparison between WVD and SWT single complex signal modulated by a Gaussian window plus Gaussian noise: (a) WVD (b) SWT (c) ROC curve

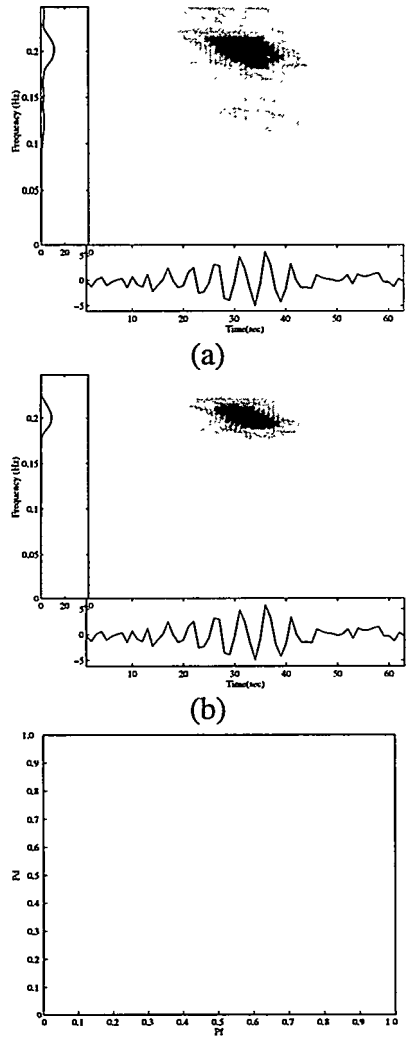


Figure 4 Comparison between WVD and SWT single complex signal modulated by a Gaussian window plus Gaussian noise: (a) WVD (b) SWT (c) ROC curve

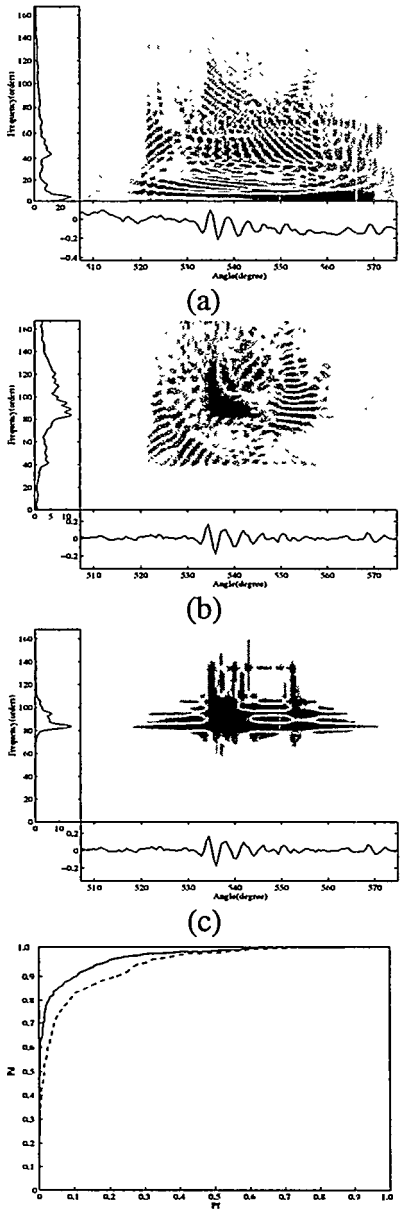


Figure 5 (a) WVD of raw data (b) WVD of pre-processed data(c) Smoothed SWT of pre-processed data (d) ROC curves for WVD (Dotted line) and SWT (Solid line) computed for measured engine noise data.

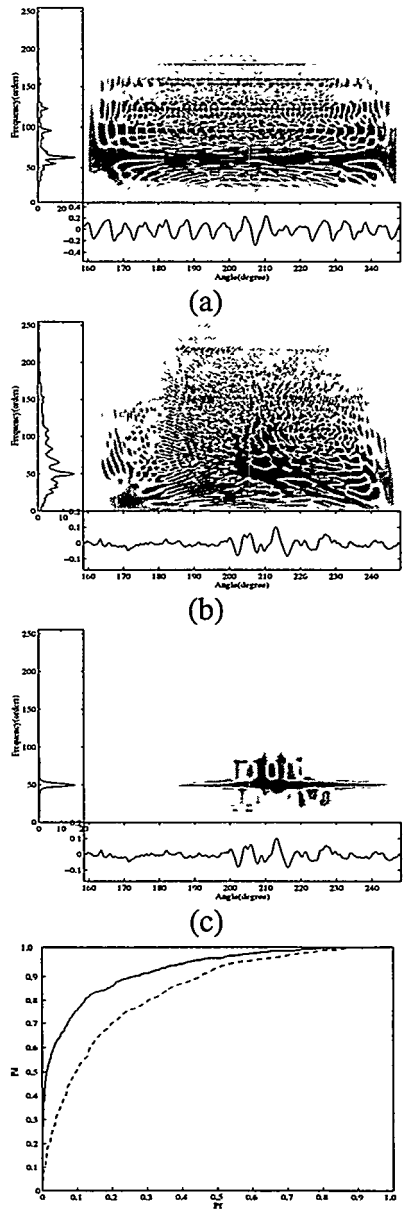


Figure 6 (a) WVD of raw data (b) WVD of pre-processed data(c) Smoothed SWT of pre-processed data(d) ROC curves for WVD (Dotted line) and SWT (Solid line) computed for measured gear box data.

FLAME IMAGE MONITORING AND ANALYSIS IN COMBUSTION MANAGEMENT

Mr. D. Popovic, Operation Manager
CEZ, a.s. Elektrarna Detmarovice, 73571 Detmarovice, Czech Republic

Mr. A.V.Andersin, Product Manager
Mr. A.J.Huttunen, Manager
Mr. J. J. Nihtinen, Product Manager
Imatran Voima Oy, IVO Technology Centre, Rajatorpantie 8, 01019 IVO, Vantaa, Finland

ABSTRACT

Modern power plants are amongst the most highly automated and centrally controlled production facilities. Increased energy costs have placed demands on power producers for increased burning efficiency, high equipment availability, low maintenance and safe operation. Environmental constraints have imposed limits on emissions, requiring producers to pay further attention to combustion characteristics in their plants.

When NO_x emissions are reduced with new low-NO_x burners and in-furnace modifications in old pulverised fuel boilers, many changes in the firing conditions may occur. Depending on coal quality and the original furnace design, low-NO_x burners, overfire air, low-excess-air firing and other primary modifications in various combinations may cause flame instability, increased slagging, increased minimum load and other difficulties in controlling the burning process.

In most pulverised fuel combustion systems, the common factor in tackling these operational problems is improving the quality of burning. Combustion management is, therefore, one of the most critical components of power generation. Management and control of burners are, however, only as good as the information provided. Flame data is one of the most critical information sources, but

conventional flame monitoring systems are basic in operation and serve primarily as safety devices. These systems are essentially no more than flame detectors, indicating the presence or absence of flame.

If more accurate burner or burner level-specific information is available, it is possible to make more sophisticated decisions concerning combustion operating efficiency.

To find and solve these problems quicker, a new type of burner management system for pulverised fuel and oil-fired boilers was developed by Imatran Voima Oy. The DIMAC combustion management system monitors and analyses individually each burner or burner level. There are special software for wall and corner fired boilers. The DIMAC system is comprised of two functional subsystems: flame monitoring and flame analysis. The DIMAC system have been installed in 11 power plants.

The DIMAC enables the power plant operators to minimise NO_x emissions and optimise the burning efficiency with varying coal qualities and boiler loads at the same time so that slagging, unburnt carbon in fly ash and flame stability stay in acceptable limits. It also quarantees that burners operate in good safety conditions in each burner level. The DIMAC system monitors perpendicularly each individual burner and evaluates flame parameters. Real-time flame monitoring and analysis allows the operator to directly see the effect of changing fuel distribution on flame pattern and flame stability.

Based on data from the DIMAC references the system can improve boiler efficiency by 0.2 - 0.5 per cent unit as a result of more efficient control of the burning process. At the same time, the NO_x formation can be reduced by 10 - 20%. As a result of the enhanced boiler efficiency and availability, DIMAC offers a pay-back time of as short as 1 year.

This burner-specific flame monitoring and analysis system is introduced in the following. The experiences and the results about commissioning and operation of the DIMAC system in Detmarovice Power Plant in Czech Republic is also described in this paper.

1 SYSTEM AND FUNCTIONAL DESCRIPTION

The DIMAC combustion management system is developed to monitor and analyse combustion-specifically at the pulverised fuel and oil-fired power plants. There are specific software for wall and tangentially fired burning systems.

The DIMAC system is comprised of two functional subsystems: the flame monitoring system and the optional flame analysis system (Figure 1).

The DIMAC system monitors each burner or burner level individually. DIMAC semi-conductor type cameras are set perpendicularly to the burners to achieve the most informative view of the flame. DIMAC views the flames through the boiler wall, and the flame images are fed as video signals from the cameras to the B&W or colour monitors and also to the flame image analysing boards in the instrument cabinet of the analysing system. To install the camera, a 40 - 100 mm viewport is required in the boiler wall.

The analysing unit is responsible for all computational functions regarding the ignition and combustion. The flame image is analysed for burner type-specific flame parameters using specific algorithms. There is a digital image processing board (the RIP 9600) for each camera signal so that true on-line, real-time flame processing can be achieved. The RIP 9600 card is designed specifically to control and rapidly process picture information acquired from a flame camera. The system analyses each image twenty-five times a second and, after averaging, produces flame parameters every second, providing truly continuous monitoring of the combustion quality.

The analysing cards evaluate burning parameters from flame pictures for specified window area. Windows for the flame areas are set during the commissioning of the system. These windows, which are user adjustable, are used for the evaluation of flame parameters, as well as for setting of alarm limits.

The management system collects and stores burner parameters from the analysing unit. It stores historical data, indicates alarms and drives user interfaces, displays and other peripheral devices. The main components of the management system and user interface are a track

ball, a multisync colour monitor and an industrial PC/AT with digitiser and I/O boards.

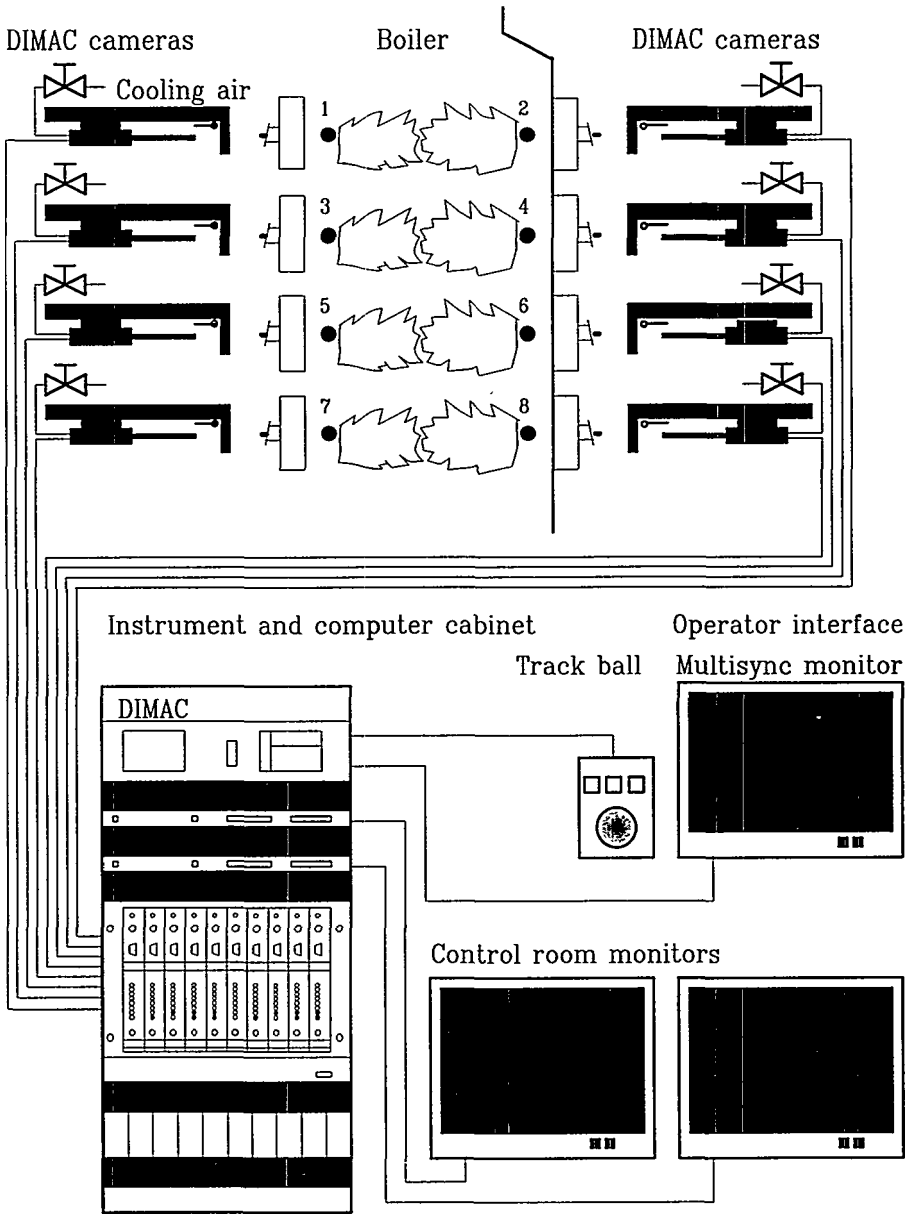


Figure 1. The DIMAC system is composed of two functional systems. The monitoring system consist of cameras, video multiviewers and control room monitors. The analysis system includes analysing unit, industrial PC and multisync monitor.

2 FLAME PARAMETERS

2.1 Parameters in front and opposite wall fired boilers

The calculated flame parameters in front wall and opposite wall fired systems are:

1. ignition point
2. stability of ignition point
3. average intensity of the flame
4. total intensity of the flame.

Ignition point parameter shows how close the ignition point is to the burner mouth. Stability of ignition point shows how stable the flame is. Stability of the ignition point is calculated as mathematical variance of the ignition point.

Air distribution is normally adjusted at the time of start-up for a particular set point. With DIMAC, indications of flame shape, stability and intensity distribution provide real-time information on all combinations of firing rates, as well as mill and burner combinations. This information allows adjustment of air distribution in the furnace for actual existing firing conditions.

Improved information on flame stability reduces the possibility of loss of flame. Allowing reduced margin for loss of flame enables minimisation of the requirement for supplementary fuel firing. Information on individual burner stability allows supplemental firing to be adjusted on an individual basis. This results in a greatly reduced consumption of more costly secondary fuels.

DIMAC provides the operator with information on changes in the location of the ignition point and a direct numeric and graphic presentation of flame stability. Operators can observe trends that would lead to loss of ignition in the future.

Operators can also adjust the burners and fuel system to prevent loss of ignition, rather than responding to the loss of flame. Loss-of-flame control becomes pro-active rather than reactive. Operators have information on the margin of safety to loss of ignition and can adjust the system to control this margin as the circumstances warrant.

2.2 Parameters in tangentially fired boilers

In tangential burning systems, potential problems and required information are different from those of the wall burning systems. In tangential fired boilers, the burner or burner level-specific information can be used to minimise slagging of the burner area, to get more information about fuel distribution, to minimise excess air in various boiler loads and to control ignition of the fuel flow.

In tangential burning systems, the evaluated parameters for conventional type jet burners are:

1. position of ignition point on fuel stream
2. stability of ignition location
3. height of fuel stream
4. upper flashpoint in combustion window
5. lower flashpoint in combustion window

The ignition point on fuel stream indicates how the fireball is located to the boiler cross section. Air flows can be balanced so that the fireball is in the middle of the cross section to balance heat transfer in the furnaces. This parameter also gives an alarm, if ignition point is too close to the burner: Burner malfunctions or incorrect burner tunings (fuel/air ratio or air velocity) can be identified.

In certain burners, flames try to reach upper and lower air registers around burners. In such case, parameters four and five can trigger alarms, if needed.

In new type low-NO_x burners it is also important to know the intensity distribution in the flame. The opening angle of the flame is also calculated. This parameter helps the operator to fine tune secondary air flows and the ratio of secondary and tertiary air flows so that the flame shape is correct. Intensity distribution and opening angle of the flame are calculated in corner fired low-NO_x burner installations.

3 EXPERIENCES OF COMMISSIONING AND OPERATION OF DIMAC AT DETMAROVICE POWER PLANT

3.1 System description

The DIMAC system has been installed at the same time with the Low-NO_x combustion system in CEZ Detmarovice Power Plant to units 1,2 and 4 in 1994-1997. Installation at unit 3 are to be completed in 1997 and commissioning of this last unit will executed in the end of 1997. The delivery for each unit includes the following subsystems:

- **IVO RI-JET low-NO_x burners (16 pcs) and air box modifications.** The burners are situated in the corners of the furnace (tangential firing) in four levels. Boilers are operated with two or three burner levels in use. Each burner level is fed by one coal mill. Mills are equipped with static classifiers.

- **Over fire air (OFA) nozzles (4 pcs) and OFA ducts.** Air to the nozzles is taken from the windboxes after Ljungström air heaters. OFA nozzles can be tilted in vertical direction (actuators) and in horizontal direction (manual).

- **OFA automation and control.** The OFA control system regulates the OFA flow rate based on a floating set-point that is derived from boiler load. Purpose of the control is to keep constant air factor (stoichiometric ratio) at the burner zone. Air flow rate is regulated by four control dampers.

- **DIMAC Combustion Management System.** DIMAC system includes eight flame monitoring cameras, monitors and a analysing system. Two cameras are installed on each burner level so that on each level coal flames can be monitored from the "cold" side as well as from the "hot" side.

A separate unit (SAIA) was used for the automation and control system in boilers 1 and 2. The I&C system in boilers 3 and 4 has been/will be modernised by Siemens and the OFA automation is included in this main system.

3.2 Detmarovice power plant

There are four similar units at Detmarovice power plant. Below is presented the process data of the units:

• Power output	200 MWe
• Steam flow rate	max. 650 t/h
• Live steam	540°C, 165 bar
• Reheat steam	540°C
• O2 at FG exit	about 4%
• NOx emissions	820-1150 mg/Nm^3 (before low-NO_x)
	300-400 mg/Nm^3 (after low-NO_x)

There have been carried out several modifications of existing control system to reach operation conditions according to frequency control system of the Western Europe (UCPTE). Nowadays turbine keeps the power (secondary control mode) and in addition balance the frequency (primary control mode). This control mode means that we have to change power by ± 5 MWe (± 50 mHz) and follow by power drift 3 MWe per minute at loads from 130 MWe to 200 MWe; below this we operate in hand control mode. Boiler have to balance mentioned changes by combustion system in pressure range ± 10 bars.

We have also spread the operation range of the units so that we nowadays operates at loads from 50 to 100 %.

The units have experienced operational difficulties related to pulverised coal fineness:

-design value (1974): 30% particles >90μm
-operation value (1997): 40-50% particles >90μm

Above mentioned reasons have placed demand on very carefully combustion management and more detailed information on combustion.

3.3 Commissioning of the system

During the commissioning DIMAC system was used to monitor changes in flames when burner settings, air flow rates etc. were adjusted. With DIMAC system it was possible to get immediate, on-

line information about flames and combustion that significantly helps the system commissioning.

During commissioning DIMAC system was used to

- monitor flame shape
- monitor flame intensity (brightness) with certain restrictions
- monitor flame stability (ignition point, "black areas" in flames)
- monitor differences in flames on different burner levels
- monitor differences in flames from "cold" and "hot" sides

Main results achieved by executing the low-NO_x modification:

- NOx emission was reduced to level 300-400 mg/Nm^3 (reduction 50-70%)
- O_2 in flue gases after the boiler decreased to 3,9%
- Unburnt carbon in fly ash and in bottom slag was less than the guaranteed 3.5%
- somewhat higher UBC values were observed in unit 2 but later some adjustments were done with positive results
- boiler efficiency was increased due to lower flue gas exit temperatures
- no changes HP steam parameters
- no changes in boiler capacity
- no remarkable effect on boiler dynamics; very important due to UCPTE operation (frequency control)
- reheat steam temperatures were decreased at lower loads due to changes in heat distribution
- safer operation at low loads

Reheat steam temperatures will be higher in unit 4 due to the modifications made at reheater heat transfer sections. These modifications will also be made in unit 3 at the same time with low-NO_x modification (-97) and in units 1 and 2 in the next major overhaul.

3.4.DIMAC System operation experiences

In daily power plant operation DIMAC system is used to

- monitoring of flame shape, stability etc. followed by corrective actions to optimise combustion
- monitoring of normal deterioration of the combustion system and possible mechanical faults
- monitoring of process disturbances and mill malfunctions

With DIMAC we have reach very convenient operation environment for the operators. They can immediately recognise problems with flame stability and control combustion process accordingly.

REFERENCES

1. Lawn, C.J., Principles of Combustion Engineering for Boilers, Academic Press, pp 50-58 (1987)

2. Hirvonen, J.,et.al., Image Processing in Combustion Control. International Journal of Pattern Recognition and Artificial Intelligence Vol.10, No 2, 1996.

RESEARCH ON WEAR DIAGNOSTIC SOFTWARE FOR TRIBO-SYSTEM

X. P. Yan , Professor , Director of Research Group
Department of Marine Mechanical Engineering ,
Wuhan Transportation University , Wuhan 430063 , P. R. China

ABSTRACT

Oil monitoring is referred to the wear particle or debris analysis and physical test of lubricant's properties. It is necessary to improve the data processing method and to increase the analysis speed. A wear diagnostic software based on oil monitoring is introduced in this paper. The deceloped software contains three subsystems which are respectively called oil monitoring & diagnostic pattern, oil monitoring & diagnostic program and oil monitoring & diagnostic knowledge base. The structure and main menus of the software are outlined. The principle of data fusion and mutlisensors management will be adopted in the developed software and an expert system based on oil monitoring will be studied in the future.

1 INTRODUCTION

When one surface moves over another, there is always some resistance to movement, and the resisting force is called friction. In order to reduce the friction, a lubricate is used to improve the smoothness of movement of one surface over another. So lubrication is a primary purpose of the lubricants, and the lubricants will also be there to help cool the tribo-system. Furthermore, they can clean and remove wear particle or debris from critical regions on the surface of parts. It is this multipurpose characteristic of lubricants which enables them to be such a good condition monitoring medium, and it is not a remote medium, it is very close to the real mechanism of wear and deterioration. It may be sampled at some distance from the source of trouble, but the evidence has been retained, virtually without any loss in intensity[1]. This is why oil monitoring can be used to diagnose the wear state in a

machine under running condition. Oil monitoring is referred to the wear particle or debris analysis and physical test of lubricant's properties. Oil monitoring is a powerful technique to diagnose wear extent, cause of wear, failure of moving parts and tendency of wear by analyzing concentration, size distribution and types of debris. Also oil monitoring is used to detect the used lubricant's properties such as viscosity, water, flash point etc. by the standard testing method. All data obtained and expressed in some way to display the wear condition in the monitored machine. Traditionally, data is processed by manual, that is a time consuming and tedious work for analysis, and influences the practice applications, particularly in industries. It is very important to increase the analysis speed and make the data processing automatically. For this reason, a wear diagnostic software has been developed. This software can be used to deal with the data from the different oil analyzer, for instance, data from chemical and physical test; data from spectrometric oil analysis; data from ferrographic analysis; data from particle counting and data from infrared spectrum analysis etc. It can help the operator to write the diagnostic report and get the diagnosis results fast than before. Its application in industry showed that this software is a useful tool for oil monitoring.

2 DESIGN OF WEAR DIAGNOSTIC SOFTWARE

2.1 Structure

Data such as numbers, texts, graphs and images is included in oil monitoring. So it is not easy to deal with all this data by manual, and also difficult exchange the data between the analysis Lab and machine management department, especially, when these are hundreds of machines should be monitored in an enterprise. Only is computer used to aid the data processing, the oil monitoring can be effectively made its applications[2]. Figure 1 is the structure of the wear diagnostic software which has been developed by the author's research group. The software contains three subsystems which are respectively called oil monitoring & diagnosis pattern, oil monitoring & diagnostic programs and oil monitoring & diagnostic knowledge base. There are input interfere, database and output (both tables and reports)for each subsystem. For example, the first subsystem called oil monitoring & diagnostic pattern has the functions of generating failure statistical reports and data querying reports. The second subsystem called oil monitoring &

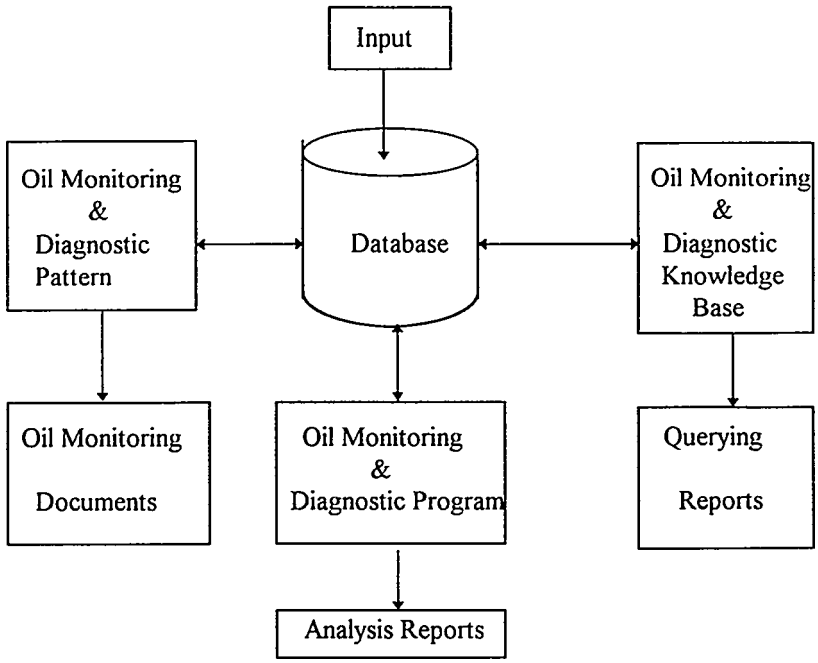


Figure 1. Structure of wear diagnostic software

diagnostic program has the functions of processing data and outputting the diagnosis result, failure analysis tendency curves and wear particle image. The third subsystem called oil monitoring & diagnostic knowledge base contains information such as standards, instruments, lubricants and also images of wear particle. This subsystem has the function of providing various knowledge to analyst, when he is doing the analysis.

2.2 Database

There are many different kinds of data in the wear diagnostic software and the data is stored in the different database. The databases are as follow.

2.2.1 Database in oil monitoring & diagnostic pattern

1) Management department (or organization) database
2) The monitoring instruments database
3) The monitored machines database

4) The moving pairs of monitored machine database
5) The lubricating system of monitored machine database
6) The tables used in oil monitoring database
7) Sampling requirements database

2.2.2 Database in oil monitoring & diagnostic program

1) Sampling database
2) Spectrographic analysis database
3) Wear particle concentration database
4) Wear particle types database
5) Wear particle image database
6) Lubricant physical testing database
7) Infrared spectra analysis database
8) Particle counting database
9) Maintenance records database

2.2.3 Database in oil monitoring & diagnostic knowledge base

1) Oil monitoring literature database
2) Specifications of oil monitor database
3) Ferrographic analysis standard database
4) Spectrographic oil analysis standard database
5) Lubricants chemical and physical testing standard database
6) Specifications of lubricants database
7) Cleanliness standard database
8) Wear mode and mechanism database
9) Failure analysis database

3 FUNCTIONS OF WEAR DIAGNOSTIC SOFTWARE

This wear diagnostic software is a computer aided system for oil monitoring. It can be applied to process the various data such as numbers, texts graphs and images from oil monitoring. The software is used to help the analysts to do oil monitoring effectively and easily. As an example, the menus of the monitoring & diagnostic program, which is the main subsystem in this software are shown in Figure 2. The functions of data management, information processing and fault diagnosis are included in this subsystem. The functions of those three modules are discussed as follow.

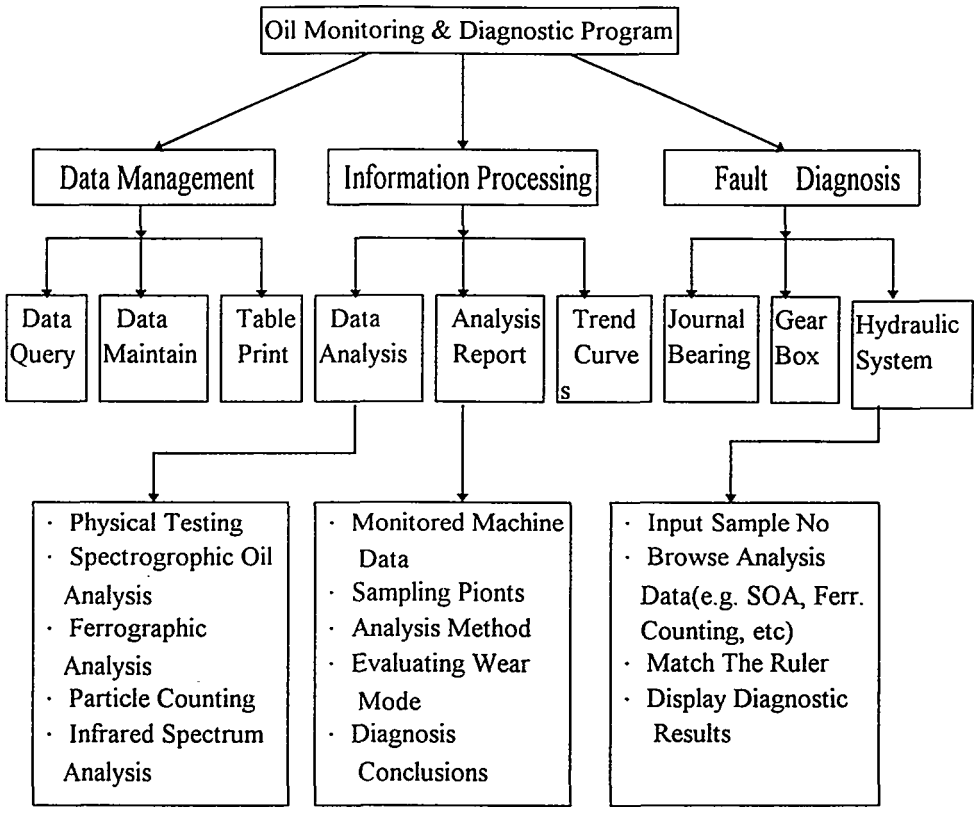


Figure 2. Schematic diagram of the menu used in oil monitoring & diagnostic program

In data management module, "Data Maintain" menu has the option which is used to add, modify, delete and edit etc. "Data Query" menu can be used to search for which record the analyst wants. "Table Print" menu is used to output the tables containing the interesting data from the specific databases.

In information processing module, if "Data Analysis" menu is selected, the data from five oil monitoring methods are automatically processed and several mathematics patterns such as liner regression, grey predict and fuzzy identification are applied in this software. The analysts can effectively finish their works with the help of this subsystem in the routine of oil monitoring.For example, when a specific analysis method is selected in the menu of "Data Analysis", only the parameter of "Sampling number" is input by the analyst, the analysis is automatically

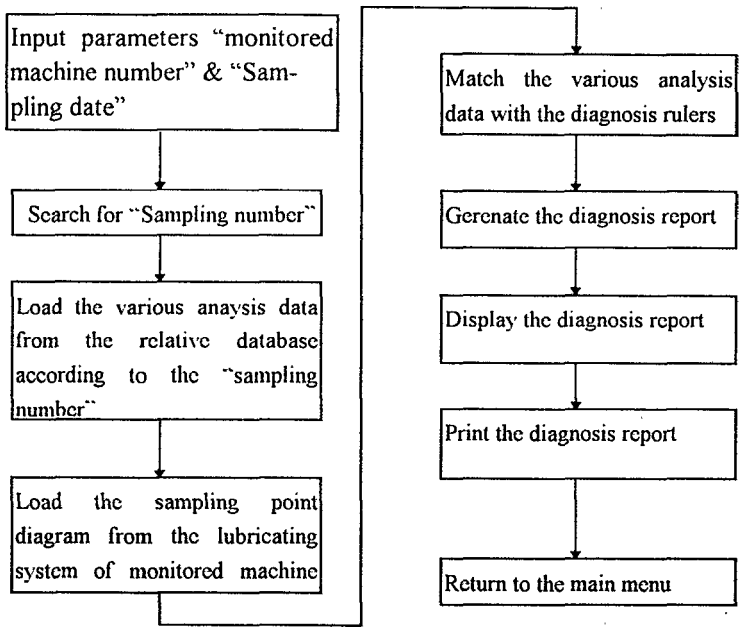


Figure 3. Flow diagram of analysis report generating procedure

completed and the analysis result will be output in form of diagnostic table. The flow diagram of "Analysis Report" is shown in Figure 3 which illustrates the report generating procedures. In this module, after the parameters such as "monitored machine number" and "Sampling Date" are input, the data relative to those parameters is located for from the different database e.g. ferrographic analysis data from Wear Particle Concentration database and Wear Particle Types database, Spectrometric Oil Analysis database etc. The sampling point diagram is loaded from the Lubricating System of Monitored Machine database.

In fault diagnosis module, the failures generated in three typical tribo-systems can be determined. When this module is selected, the analyst inputs the sampling number and then the diagnosis goes on in the way of man-machine dialog. The diagnosis rulers expressed in the method of "If-Then" are edited in the software, and the rules are summarized from the successful monitoring examples by the designer of this software.

4 DISCUSSION

Oil monitoring has grown considerably in last ten years, but there are still many problems to be solved in the future, especially, on-line oil monitoring, automated identification of wear particles and intelligent diagnosis system based on oil monitoring are the important and interesting research areas[3]. The wear diagnostic software developed in the author's laboratory is a computer-aided diagnostic system, and not an expert system. When an expert system is designed, the rules (or knowledge) are more important than the structure of the expert system itself. At the present, it is necessary to collect the data from various machine monitored and to summarize the successful experiences from the different fields in industry. In some way, the wear diagnostic software is useful to help the analyst to arrange the data and acquire the knowledge. This software is adopted by a iron & steel company to help the analyst to do the analysis in the routine of oil monitoring. The application of this software noted that the analysis speed is increased. Although this software is useful, it will be modified to meet the needs for the different fields in industry, and the principle of data fusion and mutilsensors management will be applied[4], an expert system based on oil monitoring will be resulted from this software in the future.

5 CONCLUSION

A wear diagnostic software is developed and it is powerful tool for the analyst to deal with the data from oil monitoring. This software is used to help both the managers to understand the oil monitoring and the analysts to do the accurate and fast analysis in the routine of oil monitoring. The principle of data fusion and mutilsensors management will be adopted in the software and an expert system will be resulted from this software in the advancing research.

REFERENCES

1. Hunt, T. M. "Oil Debris Monitoring", Handbook of Condition Monitoring, First edition ,eds Rao, B. K. N. Elsevier Advanced Technology, 1996, 325-340

2. Yan, X. P., Xie, Y. B. & Dong, L. Z. "The Design of A Computer-aided Diagnostic System Based on Oil Monitoring", Proceedings of the Fifth International Condition Monitoring Conference, eds J. Zhu, Y. Y. Zhang, H. L. Xiao and M. H. Jones, Defense Press, 1997,304-308

3. Yan, X. P., Xie, Y. B. & Xiao, H. L. "Development of Oil Monitoring Techniques and its Tendency", China Mechanical Engineering(in Chinese), 8(1), 1997

4. Manyika, J. & Durrant-Whyte, H. "Data Fusion and Sensor Management: A Decentralized Information- Theoretic Approach", Ellis Horwood Limited, 1994

MACHINE AND LUBRICANT CONDITION MONITORING FOR EXTENDED EQUIPMENT LIFETIMES AND PREDICTIVE MAINTENANCE

Malte Lukas and Daniel P. Anderson
Spectro Incorporated
Littleton, Massachusetts, USA

ABSTRACT

Predictive maintenance has gained wide acceptance as a cost cutting strategy in modern industry. Condition monitoring by lubricant analysis is one of the basic tools of a predictive maintenance program along with vibration monitoring, performance monitoring and thermography.

In today's modern power generation, manufacturing, refinery, transportation, mining, and military operations, the cost of equipment maintenance, service, and lubricants are ever increasing. Parts, labor, equipment downtime and lubricant prices and disposal costs are a primary concern in a well run maintenance management program. Machine condition monitoring based on oil analysis has become a prerequisite in most maintenance programs. Few operations can afford not to implement a program if they wish to remain competitive, and in some cases, profitable.

This paper describes a comprehensive Machine Condition Monitoring Program based on oil analysis. Actual operational condition monitoring programs will be used to review basic components and analytical requirements. Case histories will be cited as examples of cost savings, reduced equipment downtime and increased efficiencies of maintenance programs through a well managed oil analysis program.

1.0 OIL ANALYSIS

An effective predictive maintenance program based on condition monitoring through oil analysis must determine both machine condition and lubricant condition. Lubricating oil may be used as a diagnostic medium which

carries wear debris away from the wearing surfaces. Analysis of the wear debris can, therefore, provide important information about the condition of the internal parts of a machine or engine. On the other hand, the condition of the lubricant itself is important to know. Does the lubricant meet specification? Is the viscosity correct? Is the oil contaminated with water, particulates or chemical compounds?

Condition monitoring by oil analysis can be broken down into two main categories: Debris Monitoring and Lubricant Condition Monitoring. Debris monitoring measures the trace quantities of wear particles carried away from the wearing surfaces by the lubricant. Lubricant condition monitoring determines if the lubricant itself is fit for service based on physical and chemical tests. These techniques, when combined with statistical trending and data based management techniques, provide a complete program of machine condition monitoring by oil analysis.

1.1 Debris Monitoring

Debris monitoring pertains primarily to the detection, and sometimes also the analysis, of metallic wear particles. The most common techniques and devices applied to this category of condition monitoring include atomic emission spectroscopy (AES), atomic absorption spectroscopy (AAS), X-ray fluorescence spectroscopy (XRF), ferrography, magnetic plugs, magnetic chip detectors, and microscopic examination of filter debris.

Debris monitoring is the backbone of oil analysis condition monitoring programs. It is effective in the sense that tests can be applied to determine that a system is nearing, or has reached, a failure mode. Further damage can thus be contained or avoided through immediate shutdown and repair.

Spectroscopy is the most widely applied technique for debris monitoring. It provides a quantitative, multi-elemental analysis of wear debris in lubricating oil. The elemental concentration of as many as 20 elements are reported in parts per million (ppm). Wear metals such as iron, aluminum, chromium, copper, tin, lead, silver, titanium and nickel are detectable, as well as lubricant additives such as calcium, barium, zinc, phosphorus, magnesium, boron and molybdenum. Certain contaminants such as silicon, sodium and potassium are also routinely detected. Trends are used to determine the mechanical health of a system. Concentration trends are established through routine monitoring to indicate if a continuing wear condition exists, the rate of wear, and as a consequence, the immediacy of the wear problem.

There are several types of spectrometers used for debris monitoring. These include, rotating disk arc emission (RDE), atomic absorption (AAS), X-ray fluorescence (XRF) and inductively coupled plasma (ICP) emission spectrometers. Each has its own advantages and disadvantages.

1.2 Lubricant Condition Monitoring

The second part of an oil analysis program is lubricant condition monitoring. Through periodic sampling of the lubricant, the laboratory can determine the effectiveness and remaining life of the lubricant based on degradation and/or contamination analysis.

Many oil analysis laboratories perform one or more ASTM approved tests to determine oil condition. These are both physical and chemical tests. Some typical tests for oil condition are:

- Viscosity
- TBN (total base number)
- TAN (total acid number)
- Water content (Karl Fischer)
- Total solids content

In recent years Fourier Transform-Infrared (FT-IR) spectroscopy has been applied to used lubricating oil analysis. FT-IR analysis is quick and inexpensive to perform and is an excellent screening tool. ASTM tests can be performed in cases where a definitive answer is required. FT-IR analysis quantifies the presence of various types of chemical bonds. Differences between the infrared spectra of the sample and the unused lubricant indicate chemical changes in the used oil as well as various types of contamination. A thirty second analysis provides information on oil contaminants include water, blowby products (soot), coolant chemicals (ethylene glycol) and unburned fuel and degradation based on nitration, oxidation, and sulfation.

2.0 TURNKEY USED OIL ANALYSIS LABORATORIES

Configuration and instrumentation of a laboratory will vary based on the machines being monitored and the sample work load. A full-service laboratory is shown in Figure 1, but the basic minimum components consist of an emission spectrometer, a Fourier Transform-Infrared spectrometer (FT-IR) and a viscometer. Each instrument sends its results to a data based laboratory information management system for data storage, evaluation and reporting.

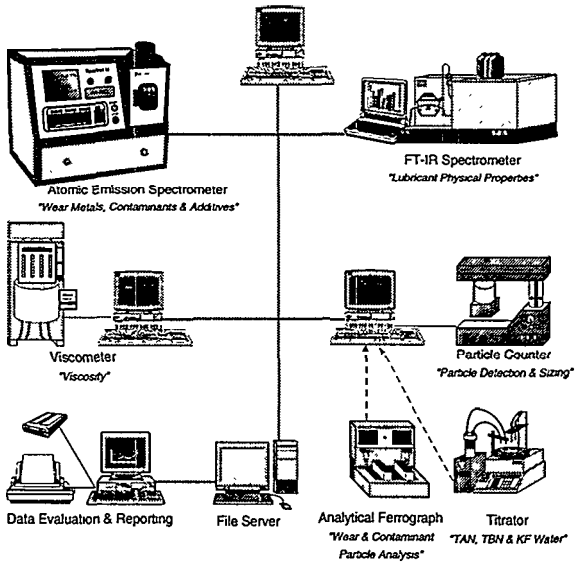


Figure 1, Full-service turnkey used oil analysis laboratory

3.0 CASE HISTORIES

The following are a series of actual predictive maintenance examples based on oil analysis case histories. They show the effectiveness and versatility of well managed and properly applied condition monitoring programs based on oil analysis.

3.1 Cost Avoidance by Detecting Lubricant Mix-Up

A serious recurring problem in maintenance procedures is the use of an incorrect lubricant. A condition monitoring program can readily identify such problems through the analysis of the lubricant additive package and lubricant physical property analysis.

The most common occurrence of lubricant mix-ups occur when an oil system is "topped off" to replace the oil that has been lost due to use or leakage. Usually a small amount of incorrect oil in a large closed loop system presents few immediate problems. This is, however, not the case in certain diesel engines as illustrated by this example.

Table 1 is a summary of the last four oil analyses for a medium speed diesel engine from a locomotive. Only the most significant analytical data is shown.

Table 1, Spectrometric Results for an EMD Medium Speed Diesel Locomotive

Date	Fe	Cu	Ag	Mg	P	Zn
9/30	19	10	0	0	0	3
12/23	21	10	0	0	9	3
3/23	27	13	2	107	75	90
6/11	25	30	10	220	110	123

The data clearly shows that after the first two samples, an incorrect oil was used to top-off the reservoir. The three additive metals magnesium (Mg), Phosphorus (P), and zinc (Zn) appear in the third analysis and increase in the fourth, a clear indication that the oil formulation has changed. In this type of engine, an incorrect oil which contains a zinc based additive package can result in severe wear problems. Several components such as bearings and wrist pins have silver coatings which corrode and wear in the presence of zinc. The early stages of the corrosive action cause by the zinc additive is indicated by the increase in the iron, copper and silver wear metals. A recommendation based on the analysis was made to drain and flush the system and to observe correct top-off oil requirements. Without oil analysis, the wear problem could have resulted in a bearing failure and a major overhaul costing over $150,000.

3.2 Contamination Example on a Pump Turbine

Pump turbines are used in many parts of the world to generate electrical power. Water is pumped to an elevated reservoir at night when power is relatively inexpensive. During peak power requirement periods, the water is allowed to flow downhill to turn a turbine which is coupled to a generator. These are reliable systems. However, condition monitoring based on oil analysis can be very effective at predicting a possible failure in the very early stages of the problem and prior to secondary damage or catastrophic failure.

A pump storage system of an electric utility was part of a condition monitoring program when the laboratory detected an increase in "coarse" wear particles in the upper guide bearing assembly of the turbine. Although the normal analysis using an emission spectrometer was acceptable, the laboratory requested more frequent sampling based on the data for iron and babbit metals obtained with a large particle detection system option

to the emission spectrometer. The original and the next three analyses using the standard emission spectrometer and the large particle detection (Rotrode Filter Spectroscopy, RFS) technique are shown in Figure 2. Although the normal emission spectrometric analysis does not show a trend, the RFS analysis definitely does.

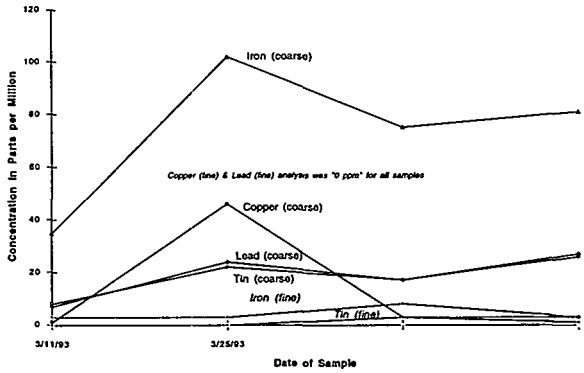


Figure 2, Pump Turbine Guide Bearing Wear Trend

Ferrographic analysis on the last two samples verified the presence of large cutting wear particles, Figure 3, causing the laboratory to issue an ALERT. However, the presence of spheres on the ferrogram was the eventual indicator which lead to the source of the wear problem.

Figure 3, Ferrogram Showing Cutting Wear and Weld Beads

Tilting pad bearings such as those used on the turbine do not generate spheres in a wear mode. Weld beads were suspect and it was eventually verified that the turbine had not been protected during overhead construction work. Weld debris including weld beads, and not a defective component, were the cause of the wear trend. Although the wear was not critical, the oil was cleaned as a precaution and more frequent oil analysis monitoring was recommended. The wear trend if undetected by oil analysis may or may not have lead to a catastrophic failure. The thought of failure is not a pleasant one, especially in view that such a failure can require an expensive multi-million dollar overhaul.

3.3 Slinger Ring Problem in Journal Bearing

The previous case history dealt with oil analysis from plain journal bearings common to many large industrial motors and turbines. A variation of this bearing, found in high horsepower motors is the ring oiled sleeve. The rings (slingers) sit on the shaft and rotate with it, splashing oil on top of the shaft. In forced lubrication systems, these slinger rings become redundant. However, in the event of an emergency shutdown, the forced feed lubrication is lost, and these slingers act as a safety device during coast down to prevent lubricant starvation.

Slingers are generally made of clock brass or general duty bearing bronze. They are considered "soft" material compared to the carbon steel shaft. On occasion, the slinger rings can stick, causing them to wear abnormally against the rotating shaft. As this happens, they release a large amount of nonferrous copper alloy wear particles into the circulating oil.

Abnormal slinger wear was suspected in an extruder main drive motor by the predictive maintenance engineer at a large petrochemical complex. The 6,000 HP motor has a force-feed lubrication system which started to be monitored regularly through oil analysis. The oil analysis data detected serious wear from the slingers, Figure 4, and it was confirmed by ferrographic analysis.

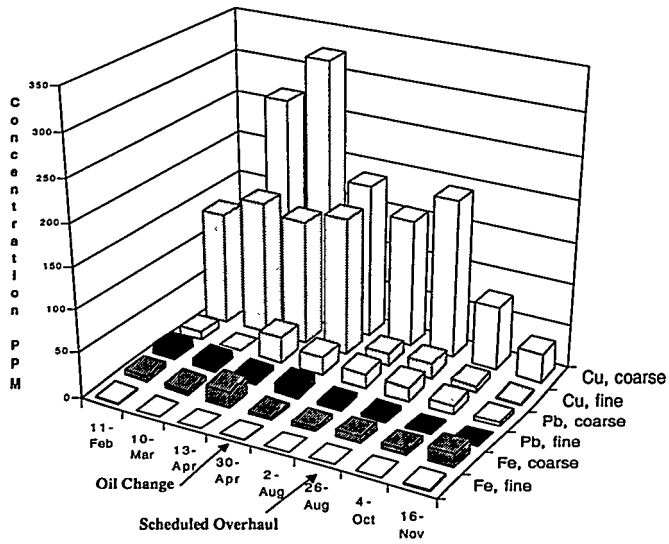


Figure 4, Wear Trend in Extruder Drive Motor

The lubrication system was placed on a monthly sampling cycle. As the wear trends increased, the oil analysis laboratory recommended an oil change based on the April 30 analysis. It was assumed that the concentration of debris within the system would be reduced, thereby minimizing the risk of damage to other components in the lubrication cycle until a detailed inspection of the bearing could be carried out at the next scheduled maintenance overhaul.

During the scheduled overhaul in October, the slingers were found to have substantial wear. The debris from the wear contributed substantially to scoring of the babbit liner on the sleeve section, Figure 5. A decision was made to reinstall the bearing, change the oil and continue monitoring on a regular monthly basis. Replacement slingers were machined in-house during the scheduled overhaul. An unscheduled shutdown and resulting disruption of the process line was avoided. The condition monitoring engineers also approached the motor manufacturer regarding design changes to reduce abnormal slinger wear.

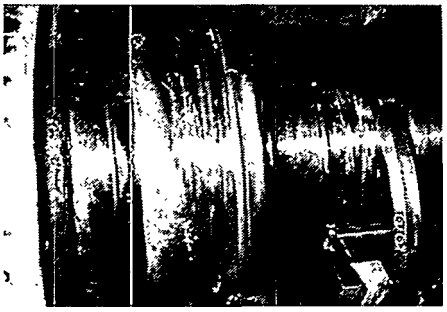


Figure 5, Slinger Ring Assembly Showing Abnormal Wear

3.4 Cost Avoidance by Reducing Unexpected Bearing Failures

The Southern California Edison Mohave Generating Station implemented a full condition monitoring program in the early 1990's. At that time they set up and placed into operation their own on-site oil analysis laboratory. Confirmed cost benefits were realized within the first year of operation.

Bearing failures and replacement costs were targeted as an immediate application for oil analysis. Over a four year period, they were able to decrease the cost of general use roller and ball bearings by more than a factor of four. The costs savings were documented as bearing expenditures

which were reduced as follows:

Year	Bearing Costs
1992	$415,000
1993	$264,000
1994	$209,000
1995	$100,000

Condition monitoring also reduced the replacement cost of mill and PA Fan motor bearings from $320,000 to $30,000.

The Mohave Generating Station has 40 Dynacore centrifuge rotors in operation to remove water from coal slurry prior to burning. In 1991 they experienced 35 feed pipe bearing failures. Today, failures have been almost completely eliminated as a result of their oil analysis program. The actual failure rate history over since the implementation of oil analysis has been as follows:

Year	No. Bearing Failures
1991	35 failures
1992	17 failures
1993	4 failures
1994	3 failures
1995	1 failure

4.0 CONCLUSION

It is never too late to implement a machine condition monitoring program. The benefits of the program can be realized in a very short period of time. Figure 6 is a typical summary of the types of problems that will be encountered in most instances.

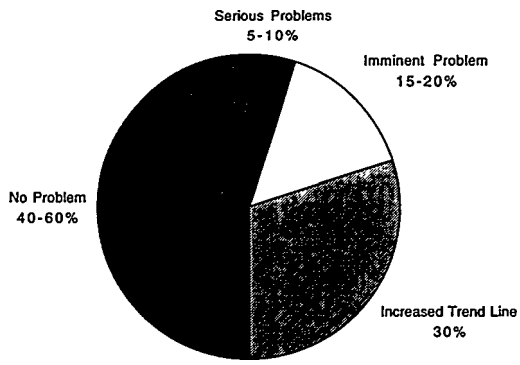


Figure 6, Summary of Typical Problems in a Maintenance Program

A number of serious or critical problems will be identified almost immediately. These will require immediate attention to avoid secondary damage, unexpected downtime or a major overhaul. A surprising number of imminent problems will also be identified. These are the future unplanned failures and should be scheduled for action and/or repair during the next scheduled maintenance shutdown.

The objectives of a predictive maintenance program based on condition monitoring through oil analysis is to identify potential failures in their early stages when repairs can still be initiated and costly secondary damage is avoided. A second objective is to monitor the quality of lubricants and to reduce lubricant usage through extended oil change intervals. The net benefits are reduce maintenance costs, increase equipment availability and life, reduce lubricant usage and improve safety. They can be summarized as follows:

1. *REDUCE MAINTENANCE COSTS*

2. *INCREASE EQUIPMENT AVAILABILITY*

3. *REDUCE LUBRICANT USAGE*

4. *IMPROVE SAFETY .*

It is almost impossible in today's competitive environment to operate without some kind of predictive maintenance program. Condition monitoring based on oil analysis is a proven technique which leads to more efficient use of equipment and maintenance savings. Some basic principles that must be followed in implementing such a program to fully realize its benefits are:

1. *Well Defined Purpose*
2. *Appropriate Tests*
3. *Careful and Timely Sampling*
4. *Commitment to Act on the Information*

CONTAMINANT MONITORING OF HYDRAULIC SYSTEMS - THE NEED FOR RELIABLE DATA

M J DAY, Senior Manager
Pall Europe Ltd, Portsmouth, UK
&
J RINKINEN, Associate Professor,
Tampere University of Technology,
Tampere, Finland

ABSTRACT

The need for both reliable operation of hydraulic and lubrication systems and long component lives has focused users to the benefits of controlling the contamination in the hydraulic fluid. Maximum operating (target) levels are being implemented as part of a condition based maintenance regime. If these are exceeded, maintenance effort is directed to correcting the rise in contamination level, and so make optimum use of resources as maintenance effort is only affected when it is necessary to do so.

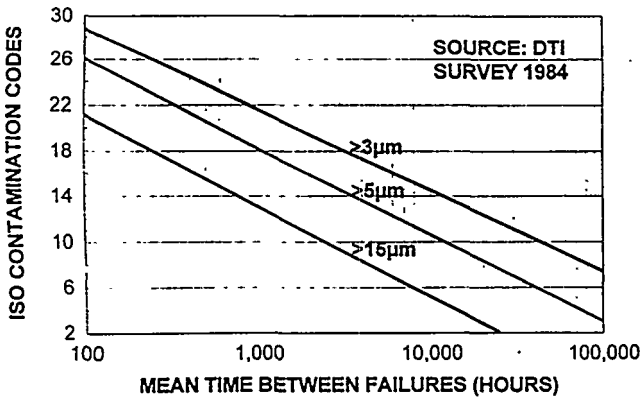
Fundamental to this aspect of condition based monitoring is the provision of accurate and reliable data in the shortest possible time. This way, corrective actions can be implemented immediately so minimising the damage to components. On-line monitoring devices are a way of achieving this and are seeing increased use, but some are affected by the condition of the fluid. Hence, there is a potential for giving incorrect data which will waste time and effort by initiating unnecessary corrective actions. A more disturbing aspect is the effect on the user of continual errors. The most likely effect would be a loss of confidence in the technique or even complete rejection of it and hence the potential benefits will be lost.

This paper explains how contaminant monitoring techniques are applied to ensure that the potential benefits of operating with clean fluids is realised. It examines the sources of error and shows how the user can interrogate the data and satisfy himself of its authenticity.

1. INTRODUCTION

The presence of dirt in hydraulic and lube systems has long been acknowledged as the major cause of failure, but was rarely quantified on a wide scale. It was not until the UK's Department of Trade three year research programme into the effects of contamination on components and systems that the magnitude of the effect was quantified [1]. The Field Studies project [2] gave a relationship between the dirt level as represented by the ISO 4406 code [3] and the reliability of systems [Figure 1]. Put quite simply, the better the cleanliness of the hydraulic fluid, the higher the level of reliability. Although these clean fluids were produced by so called wear control filters, generally 6 or 3 μm, the report stated that it was also a function of how they were performing in the system concerned and how they were maintained.

FIGURE 1 RELATIONSHIP BETWEEN DIRT LEVEL AND RELIABILITY



Since the study, matters have improved considerably as all sectors of industry realise the benefits of operating with cleaner fluids. However, just fitting a finer filter to achieve a reduction in contamination level can only be considered as a 'fix'. What is required is a practice of Total Cleanliness Control (TCC™) [4] which embraces all engineering aspects; the design of components and systems for cleanliness control; cleaner machining coolants; filtered wash fluids; improved assembly and build techniques, and finally better maintenance procedures. It will need some investment in both capital and training, but a study by the UK's Institute of Mechanical Engineers *Tribology in Action* programme indicated that for every £1000 invested, a return of £40,000 could be realised [5].

Fundamental to TCC™ practice is the setting of realistic standards of fluid cleanliness and, of course, the specification of filters to achieve them. Once a target cleanliness level has been selected, then the system has to be regularly monitored to see that it is being maintained.

If it exceeded, maintenance activity is initiated to correct the situation and restore it back to the original design level. Hence, the availability of accurate and consistent data is essential to ensure that time and, hence money, is not wasted through false alarms.

2. PHILOSOPHY OF CLEANLINESS MONITORING

The aim of TCC™ is to minimise the introduction of dirt into the process, whether it be a build-up of machining debris at the sub-component level, to the control of wear debris during system operation. To do this, the source of the dirt has to be investigated, a means of control established and specifications detailed. Procedures or instructions will have to be documented showing how to monitor the process and what to do when the target cleanliness level is exceeded.

An illustration of this is seen when cleanliness monitoring by particle counting was integrated into the management of the hydraulic systems of one of the Tunnel Boring Machines which operated under the English Channel. Here, huge financial penalties of £500,000 per day would have been imposed in the event of project overrun. At the design stage, maximum contamination or Action Levels were placed on the particle counts of the system oil beyond which corrective actions had to be taken. This is the foundation of a condition based maintenance regime.

Table 1 Action Levels for Tunnel Boring Machine

TARGET ISO CODES 2 / 5/15 µm	COMMENTS	ACTIONS
17/15/12	*Normal running level*	No action
19/17/14	*Action level:* Filters not controlling	Check all filters for blockage, replace where necessary
20/18/15	*Immediate attention level:* Regenerative wear	Stop system, replace filters where required, flush system at low pressure.
15/13/10	*Flushing level*	Flushing complete, operate system.

Monitoring was effected at the component stage where the 'normal running level' of ISO 17/15/12 was selected as the cleanliness level for delivered components (clean by 1991 standards).

The 'flushing level' was selected for commissioning both at the manufacturing plant and subsequently at the tunnel site following strip down and rebuild.

3. CONSEQUENCES OF INCORRECT DATA

With the type of methodology where actions are generated whenever the target levels are exceeded, it is essential that the data supplied is accurate and consistent. Imagine what could happen if, in the above example, the laboratory supplying the information sent out incorrect data at, say, 23/19/13; all tunnelling would have stopped for maintenance, only to find at the end that it was a false alarm. Hence, time and effort would be wasted, and with project overrun charges at about £500,000 per day, this could have proved to be financially disastrous.

After a few instances of receiving incorrect data, users usually send 'nominally' identical samples to other laboratories in an attempt to obtain the correct level. If the results differ, a round of meetings of interested parties is usually convened, samples are exchanged between the laboratories, and even third party experts are called in. All of this absorbs a lot of time which could be more profitably employed elsewhere. A more significant and serious aspect is the potential loss in confidence in contaminant monitoring by users as a result of continual 'false alarms'. If this occurs there will be an erosion of confidence throughout the hydraulics industry, monitoring will be abandoned and users will not realise the benefits to be gained.

Hence, the provision of valid contaminant monitoring data is essential.

4. MONITORING METHODS

It is beyond the scope of this paper to discuss the options for contaminant monitoring of each step in the various process and information on the types can be found discussed elsewhere [6, 7]. The choice of instrument and how it is used depends on a number of factors.

4.1 ***On-Line*** instruments are connected directly to the process concerned (coolant line, wash plant or system pipework) and analyse a small proportion of the main flow. The benefit of this method is that the results are available very quickly, provided that the sampling point has been flushed adequately (see section 5.3), and hence corrective actions can be implemented almost immediately. Also, this mode is potentially

the most accurate as ingression of external contaminant is eliminated. However, direct connection to the system can also be problematic. The use of a small subsidiary flow can present difficulties in obtaining a representative sample. Also, connection to the system is often by opaque hoses or pipes and hence the sample fluid is not seen so that problems with the fluid (air, water in oil, oil in water) which could interfere with the operation of the instrument will be missed.

The instruments which operate in this mode are: Automatic optical particle counters (APC's); Mesh Blockage instruments; Image Analyses (under development); Magnetic Detectors

4.2 Off-Line analysis involves the extraction of fluid samples from the system and collection in sample bottles for subsequent analysis. This can either be in-house or using a specialised external laboratory. Here the major disadvantage is the cycle time before the results are available, which can be considerable if external laboratories are used. Hence, delays in implementing corrective actions will be incurred particularly if repeat samples have to be analysed.

Off-line samples can be greatly affected by errors because of the number of stages in the process, from sampling to analysis, and the errors become more significant as the level of cleanliness improves. The advantage of off-line methods are the variety of instrument types that are available, many of them relatively inexpensive. Also some will give information on the types of contaminant which is necessary when the source of any increase has to be investigated.

The candidate instruments for cleanliness monitoring are: Microscopic methods - Counting, Patch testing or Image analysis; Automatic Particle Counters; Mesh Blockage instruments and Magnetic Debris instruments.

5 FACTORS INVOLVED IN OBTAINING RELIABLE DATA

The major cause of differences in cleanliness data, unfortunately, is a result of the people involved with the operation.
Procedures exist for taking valid samples [8] and for analysing and reporting data [9, 10], but it is the departure from these procedures, or the use of un-validated procedures, that give rise to variations in data.

The factors involved are briefly discussed below and most are applicable to all monitoring techniques. APC's, by their nature, demand

special consideration as their accuracy of repeatability, speed of analysis, and ability to work directly on a system, can actually work against them. Put simply, APC's can count incorrectly with the same degree of repeatability as they can count correctly.

5.1 Sample Location - The reason for sampling has to be establish first as this will ultimately decide where samples will taken. For coolant, wash systems and systems where knowledge of the cleanliness level of the fluid entering into a component is required, then the sampling point must be located in the line leading to the component. On the other hand, if information on the general cleanliness level of hydraulic or lube systems is required, then the sample can be taken from any of the major flow lines provided that good mixing conditions exist and conditions are stable. However, it must be remembered that in hydraulic and lube systems, the contamination level will vary around the system as dirt will be added through ingression and generation, and will be removed by filtration. Also, the particle generation rate of the system will vary depending on the pressure and duty cycle. Hence, when monitoring regularly, it is essential to sample from the same part of the system or process, when conditions have stabilised and the process is up to its normal running temperature.

5.2 Cleaning of Sample Bottles - ISO 3722 [11] - If samples are extracted from the system and collected in sample bottles for off-line analysis, then only bottles which have been cleaned and verified by ISO standards must be used. Modern hydraulic systems which feature highly effective filters have fluid cleanliness levels which approach the cleanliness levels of the bottles themselves [12, 13], so the potential for error is great. Table 2 overleaf compares the cleanliness levels of sample bottles with some system levels which are currently being experienced.

The use of uncleaned or so called "sterilised" bottles cannot be tolerated when sampling from modern hydraulic systems.

5.3 Method of Sampling - This can often be the greatest major source of variability, again because the contribution of extraneous contaminant can completely obscure the results. The level of variability can, potentially, be much greater with on-line particle counting as recent research shows. In many systems studied the actual cleanliness is substantially cleaner than originally thought.

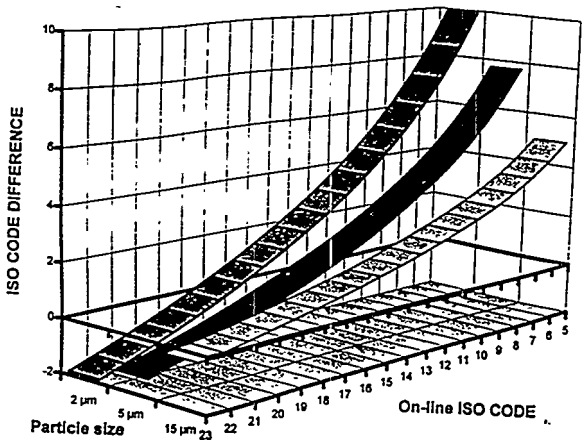
Table 2 - Cleanliness Levels of Sample Bottles

SAMPLE	PARTICLE COUNTS PER 100 ml			
	> 2μm	> 5μm	> 15μm	> 25μm
250 ml Sample Bottles				
Cleaned to ISO 3722	285	150	26	8
As received (glass)	5,900	3,480	1,090	425
As received (plastic)	10,150	6,615	2,410	1,115
System Levels				
Servo controlled fatigue rig	4,200	2,268	640	283
Plastic injection moulding machines	8,160	2,900	784	192
Agricultural tractors (roll-off cleanliness)	-	14,200	804	308
Paper Machine (hydraulic) (on-line analysis)	1,250	325	15	0

This is because most of our previous experience was generated using sample bottles analysis which effectively limited the threshold of cleanliness to about ISO 13/11/9. On-line studies of hydraulic systems by Tampere University of Technology have shown that, for many systems, the true cleanliness level can only be determined by on-line particle counting. This is seen in Figure 2 which shows the differences in ISO codes between off-line and on-line APC data as a function of the fluid cleanliness.. At the 'dirtier' levels (ISO 19/17/14) there is little difference in the two methods, but as the fluid gets cleaner then there is a substantial increase through both inadequate flushing of the sample point and the addition of extraneous contamination.

In this on-line study, the researchers found that they had to flush for significant periods - sometimes up to 15 to 20 minutes to remove the debris in the connection points, in order to obtain valid data. A similar situation was found by Svedberg of KTH Stockholm. The problem here is usually for one of having insufficient flowrate to adequately flush fittings and connection hoses as the APC can only work on low flowrates (20 - 100 mL/min). When sampling for off-line analysis this situation should not occur as sufficient flowrate should be available i.e. 1-5 L/min.

FIGURE 2 DIFFERENCES BETWEEN ON-LINE AND OFF-LINE COUNTS



Sampling for off-line analysis is detailed in ISO 4021 and most of the recommendations are valid for on-line analysis. The points to remember are:

- always use approved sample valves or connection points
- site the valve in areas which see the majority of the flowrate and where flow is turbulent
- use an appropriate flowrate (e.g. 1-5 L/min) and connect the monitor to the sample line if necessary
- flush as long as it is necessary to obtain representative data before counting. If on-line, count for a minimum of two samples to establish the trend. Accept if they are within 20% each other.

Never take samples from dead legs, quiescent areas or reservoir drain valves. If a sampling or connection point does not exist - fit one in an appropriate position.

5.4 Off-line Analysis Procedures - In view of the comments expressed about system cleanliness levels in section 5.2, all sources of extraneous contamination must be eliminated, especially those from the analysis procedures. Samples should be prepared and counted under the cleanest environmental conditions possible; laminar flow cabinets will ensure the correct environment. Before preparing the sample for counting, it essential that the sample container is examined as it may give an indication to potential problems. The external surfaces should be cleaned and the sample examined for:

- The presence of large settled contaminant, as this may cause blockage of the APC sensor.
- Darkening of the fluid, which indicates the presence of finely divided wear debris or oxidation products and will give rise to problems in counting.

- A cloudy appearance can indicate the presence of water in oils and a darkening of an otherwise clear oil can indicate the presence of 'silt' (sub 5μm particles).

The contents of the sample must always be agitated/shaken to redistribute the particles so that they are then representative of their condition in the system from which it was taken. The sample must be counted in accordance with approved procedures as standards are written by experts who have been through the learning process and know the pitfalls. In-house procedures are often prepared by inexperienced personnel. It is advisable to perform an exploratory count on the sample, either using by APC or a microscope, on a reduced sample volume to establish the likelihood of any problems with the analysis.

With microscope based counting methods, it is easy to see if there are any problems, but it is always necessary to make sure that the contaminant is evenly distributed over the surface of the membrane. The volume filtered prior to counting should not be so high that coincident particles result, and the operator has to ensure that a sufficient number of randomly selected areas are counted so that the count is representative. APC's, because they are effectively "blind", demand special consideration and this is discussed in the next section.

6. PARTICLE COUNTING USING APC'S

Unfortunately, the instruments that are most frequently giving difference in particle count data currently are APC's, and the benefits of the technique sometimes are not fully realised. It is not intended to examine all the sources of error, as these are well documented [14].
Once again, it is the inexperience of the user and is inexcusable if erroneous data is issued by independent laboratories as they should have the capabilities and resources to understand the operating principles. They are summarised below:

* ***Coincidence*** - the presence of large numbers of small particles (say, less than 5μm and sometimes called "silt") can cause over counting at larger sizes. This is the most commonly reported error, and is illustrated in Table 3.

Table 3 - Coincidence Errors in APC Data

SYSTEM	MODE	PARTICLE COUNTS PER 100 mL > 2μm	> 5μm	> 15μm	> 25μm	ISO CODES 5/15μm
Agricultural Tractor						
- as reported	Bottles	2,030,000	823,000	1,480	416	22/11
- 16 x dilution	Bottles	857,000	19,100	1,800	256	15/11
- microscope	Bottles	-	13,200	1,320	324	14/11
Aircraft Hydraulics						
- as reported	On-Line	-	364,600	1,010	90	19/11
- 2 x dilution	Bottles	-	20,520	760	208	15/11
- microscope	Bottles	-	12,500	780	390	14/11

* ***Saturation*** - APC's have a limit on the number of particles that can be counted and this results in undercounting at the smaller sizes and over counting at the larger ones through coincidence errors.

* ***Calibration*** - Although the standard for the hydraulics industry is ISO 4402 using AC Fine Test Dust [15] there is another method which is sometimes used, monosized latex spheres, and these give different results when identical samples are analysed.

* ***"Other fluids"*** - APC's work on be the basis of light interaction, that is, the difference in the refractive indices of the particle and the fluid. The presence of other interfaces, e.g. water in oil, immiscible liquids, air, etc, will be recorded as particles. Fortunately, most can be detected visually if the sample fluid is contained in a clear bottle, but difficulty will be experienced if the APC is working directly on a system with closed piping.
* ***Flow rate*** - this will affect the voltage signal generated by the particle and, hence, its perceived size. This should be slight provided that the changes in flow rate are moderate. The major problem is with on-line work where it affects the volume analysed and hence the count level.

When analysing using APC's or other automatic monitors perform one or two exploratory counts before the main analysis.

7. CHECKING PARTICLE COUNT DATA

It is surprising how much particle count data, especially that from APC's, leaves the laboratory without being checked for validity. The

reason for this is the belief that electronic instruments are never wrong, insufficient training and commercial pressures. For this reason, the draft ISO standard for APC's contains a checking procedure. The data is checked either by performing a 1:1 dilution on the final diluted sample and observing that the counts reduce by half, or by comparing the APC data with a microscopic count at 5 µm on the membrane that should have been prepared earlier. If the counts differ by more than 40%, then errors are likely.

Another good guide for users is to review the range differences between the ISO codes at > 5 and > 15µm. For samples collected in bottles these should be in the order of 3 to 4 for 'normal' samples. If it is greater (5 to 6), it can point to coincidence errors, and a 4 to 5 code difference is usual if the APC calibration method is using Latex spheres. For clean samples (ISO 13 at > 5µm), a 2 code difference is common because sampling errors inflate the counts at the higher sizes. Unfortunately, these guidelines have been developed using sample bottles and a period of learning is now required for on-line analysis.

8. CONCLUSIONS

The use of cleanliness monitoring techniques in the management of fluid system is increasing rapidly as component manufacturers, system builder and operators realise the benefits of a condition based monitoring policy. This way the processes will be kept at their design values ensuring relatively trouble free operation, leading to improvements in productivity, product quality, profitability and customer satisfaction.
The type of monitor chosen will depend on individual requirements, but it is essential that the user becomes familiar with the basic principles so that he is aware of its advantages and disadvantages.
The sampling position must be carefully selected and the sampling flowrate must be sufficient to flush out any residual contamination in the lines and that caused by connection. Sample collection for on-line analysis should not proceed until this is affected.

All data must be reviewed for validity and checked if necessary. The use of the 'patch test' is recommended for checking on-line monitoring data if the validity is in doubt.

12 REFERENCES

1 DTI, 'The UK fluid power contamination control programme", Department Trade and Industry, N.E.L., East Kilbride, UK., February 1984.

2 DTI, 'Contamination control in fluid power systems, Vol 1 Field Studies, 'Department Trade and Industry, N.E.L., Glasgow, UK., February 1984.

3 ISO 4406, 'Hydraulic fluid power - Fluids - Method for coding level of contamination by solid particles", International Standards Organisation, Geneva, Switzerland, 1987.

4 M J Day, 'Condition monitoring of fluid systems - The complete approach", Presented at the Conf on profitable condition monitoring, Harrogate, UK, December 1996, Published by MEP, Bury St Edmonds, UK, ISBN 1860580122.

5 I Mech E, Research material compiled by IMechE 'Tribology Action' Committee, I Mech E, London UK, 1992.

6 M J Day, 'Condition Monitoring of hydraulic systems", Chapter 8, Handbook of Condition Monitoring; Elseviere Sciences, Oxford, UK, 1997, ISBN 1856/72341.

7 T.M HUNT, 'Handbook of wear debris analysis and particle detection in liquids"; published by Chapman & Hall, London 1993, ISBN 1-85166-962-

8. ISO 4021, 'Hydraulic Fluid Power - Particle contamination analysis - Extraction of fluid samples from lines of an operating hydraulic system" International Standards Org, Geneva, Switzerland, 1992

9. ISO 4407 , 'Hydraulic fluid power - Fluid contamination - Determination of particulate contamination by the counting method using a microscope", International Standards Org, Geneva, Switzerland, 1992.

10 ISO/FDIS 11500, 'Hydraulic Fluid Power- Determination of particulate contamination using automatic counters operating on the light extinction principle, " International Standards Org., Geneva, Switzerland

11 ISO 3722, 'Hydraulic fluid power - Fluid sample containers - Qualifying and controlling cleaning methods", International Standards Org, Geneva, Switzerland,1978.

12. G C SVEDBERG, 'Using automatic particle counters on-line - Extremely low contamination levels can be detected", Presented at The Fourth Scandinavian International Conference on Fluid Power, pp 759 to 774, Tampere Univ. of Tech., Tampere, Finland, ISBN 951-722-374-9.

13. J RINKINEN,T KIISO, "Using portable particle counter in oil system contamination control" Third Scandinavian Conf, Univ. Lindkoping, Lindkoping, Sweden, May 1993.

14 M J DAY, 'The use and application of automatic particle counters", Presented at the 3rd International Conf on Condition Monitoring, Cambridge, London, UK, PP131-152, Elswere Applied Science, UK, 1990, ISBN 1-85166-633-8.

RULER ANALYSES TO DETERMINE THE ANTIOXIDANT, TOTAL ACID NUMBER, AND TOTAL BASE NUMBER VALUES OF USED DIESEL ENGINE OILS TO ENHANCE CONDITION MONITORING PROGRAMS

Robert E. Kauffman, Research Chemist,
University of Dayton Research Institute
Dayton, USA and
Johan Ameye,
Socomer Scientific Instruments, Brussels, Belgium

• ABSTRACT

This paper describes the development of rapid antioxidant, total acid number (TAN), and total base number (TBN) techniques based on voltammetry. The voltammetric techniques use 0.1- 0.5 milliliter (mL) of oil and 5 mL of acetone or ethanol solution, require less than 2 minutes for each analysis, and are performed in vials using inexpensive, portable instruments (**RULER™**). In contrast to currently used TAN (ASTM Method 664) and TBN (ASTM Method 2896 or 4739) techniques, the voltammetric techniques do not involve titrations, and consequently, do not rely on operator dependent endpoints. Once prepared, the oil/analysis solution mixture can be reanalyzed to verify the voltammetric test results. The current ASTM methods require resampling followed by a new titration for duplicate analyses.

This paper also presents results for the voltammetric and ASTM method (TAN and TBN) analyses of various new, laboratory stressed, and authentic used oils obtained from different types of normally and abnormally operating diesel engines. The results presented herein demonstrate that combined knowledge of the antioxidant/TAN/TBN measurements improves the capability of the analyst or equipment operator to evaluate the remaining useful life of the tested fluid as well as the condition of the operating equipment. Interrelationships among the different voltammetric analyses and wear debris analyses of the used oils are also discussed.

1 INTRODUCTION

Previous research has shown that voltammetry is well suited for determining the antioxidant concentrations of a wide variety of lubricating oils and greases [1-3]. In particular, the antioxidants present in automotive and diesel engine oils [1] have been successfully analyzed using voltammetry. Antioxidants are used to increase the thermal-oxidative stabilities of the various type lubricants. Once the antioxidants become depleted during extended use in operating equipment, the basestock of the lubricant begins to oxidize at a more rapid rate, causing accelerated increases in organic acids, viscosity, color, etc. indicating the useful life of the oil has ended.

The antioxidant level at which the useful life of the stressed oil ends is dependent on the efficiency of the antioxidant, the severity of equipment operating conditions, the thermal-oxidative stability of the lubricant's basestock, and so on. Consequently, techniques which monitor the condition of the used oil are needed to improve the oil monitoring capabilities of voltammetry.

Total acid number (TAN) measurements have been standardized by ASTM and used by condition monitoring programs to quantitate the organic acid accumulation of automotive and diesel engine oils. The standardized TAN measurements use alcoholic solutions containing strong bases (usually potassium hydroxide) to titrate solvent (toluene/alcohol) diluted oil samples to color (ASTM Method D-974) or potentiometric (ASTM Method D-664) endpoints. The standardized TAN methods are time consuming, operator dependent, produce large quantities of waste, and are best suited for laboratory use.

In addition to organic acid accumulation resulting from oxidation, automotive and diesel engine oils experience mineral (sulfur and nitrogen containing) acid accumulation resulting from the combustion gases of the fuels. Consequently, diesel engine oils are formulated with mineral acid neutralizing compounds, e.g., calcium carbonates, over based detergents, etc to inhibit acidic corrosion of the lubricated, metal surfaces. Analytical techniques, termed total base number (TBN) measurements, are used to quantitate the lubricant's capacity to neutralize mineral acids.
TBN measurements have been standardized by ASTM and used for several decades to quantitate the mineral acid neutralizing capabilities of automotive and diesel engine lubricating oils.

Two standardized TBN measurements (ASTM Methods D-664 and D-974) use an alcoholic solution containing a strong acid (usually hydrochloric acid) to titrate solvent (toluene/alcohol) diluted oil samples to a color or potentiometric endpoint. A third TBN method (ASTM Method D-2896) uses an acetic acid - perchloric acid solution titration of the solvent diluted oil sample. The standardized TBN methods are time consuming, are operator dependent, produce large quantities of waste, and are best suited for the laboratory.

Therefore, research was performed to develop rapid voltammetric techniques capable of determining the TAN and TBN values of diesel engine oils at remote sites as well as in the laboratory. The voltammetric techniques were performed with a **RULER™** instument and do not involve titration allowing for repeat analyses of each oil sample (less than 1 milliliter) prepared in a 6 milliliter vial. The developed voltammetric methods require less than one minute to perform (less than 10 seconds for each repeat analysis) using the same voltammetric instrument designed for antioxidant analysis [1-3].

Voltammetric antioxidant, TAN and TBN analyses of new and used diesel engine oils obtained from different type engines and tests were then performed to evaluate the capabilities of the combined voltammetric measurements for use by condition monitoring programs. The interrelationships among the different voltammetric analyses and wear debris analyses of the used oils were also investigated.

2 EXPERIMENTAL

2.1 Lubricating Oils

The new and used lubricating oils analyzed in this paper were obtained from normally and abnormally operating equipment as well as laboratory stressing tests and diesel engine test stands. Diesel engine oils containing ZDDP (zinc dialkyl/aromatic dithiophosphates) or non-ZDDP (e.g., calcium phenate) antioxidant/antiwear additives were obtained and analyzed.

2.2 ASTM TAN and TBN Methods

The ASTM TAN and TBN results reported herein were performed by industrial oil analysis laboratories from which the used lubricants and fluids were obtained.

2.3 Voltammetric Additive Analyses

The voltammetric method was performed with a commercially available voltammograph (**RULER™** instrument) equipped with a digital readout, two outputs (stripchart and RS-232) , and a three electrode sensing system (glassy carbon working electrode, a platinum wire reference electrode, and a platinum wire auxiliary electrode). A fresh oil typical of the application (100% standard) and the solvent system (0% standard) were used to calibrate the voltammetric instrument for % remaining antioxidant determinations.

The oil samples (100 - 500 μL) were diluted with acetone or ethanol (5 mL) containing a dissolved electrolyte and +325 mesh sand (1g). For hindered and overbased (calcium phenates) phenol-type antioxidants, the solvent system was water/ethanol and the electrolyte was potassium hydroxide. For ZDDP type additives, the solvent system was a water/acetone solution and the electrolyte was lithium perchlorate. When the oil/solvent/sand mixture was shaken, the insoluble oil coated the sand, and upon standing, the agglomerated particulates quickly settled out to produce a clear solution for analysis. The voltage of the auxiliary electrode was scanned from 0.0 to 1.0 V at a rate of 0.5 V/second. The height (or area when intergration available) of the peak(s) produced by the voltammetric method for the new (100%) and used oil were then used to calculate the percent remaining additives of the used samples.

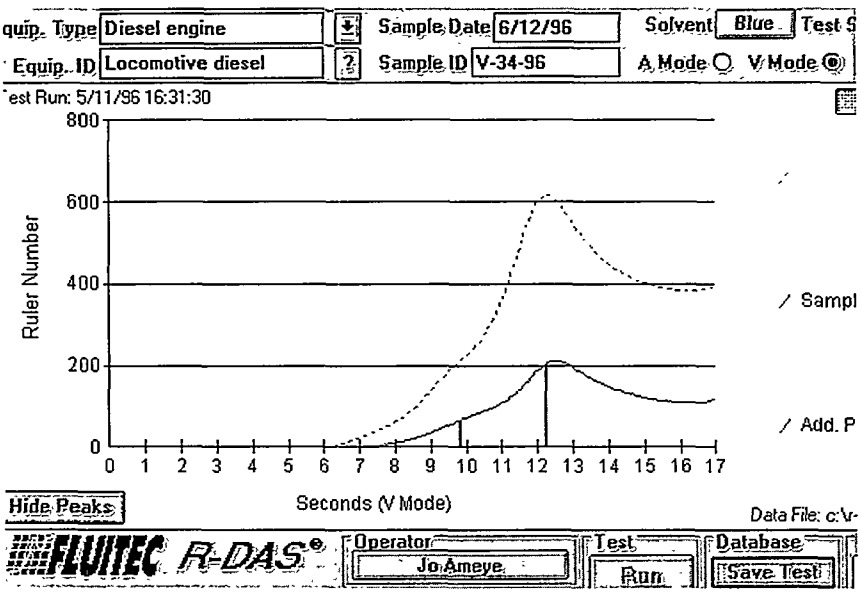


Figure 1 : RULER™ Voltammogram from Diesel Engine Oil

2.4 Voltammetric TAN Method

The voltammetric TAN method involves adding 0.1-0.3 milliliter of oil to a vial containing 3 milliliters of an ethanolic solution and shaking the closed vial for 20 seconds. The ethanolic solution contains an acid reactive compound (0.02N) such as sodium phenate which produces a voltammetric peak. If acids are present, they react with the sodium phenate to produce a sodium salt of the acid and phenol reducing the height of the sodium phenate peak. The voltammetric instrument measures the height of the sodium phenate peak before and after addition of the oil sample. An alcoholic solution containing 0.2N of hydrochloric acid is used to calibrate the sodium phenate peak height so that the amount of reduction in the peak height can be used to quantitate the amount of acid present in the oil sample. (**figure 2**)

The voltammetric TAN method can distinguish between strong and weak acids by varying the shaking time. Strong acids (hydrochloric acid) react immediately requiring only 2 seconds of shaking whereas organic acids react slowly requiring 20 seconds for complete reaction. Linearity plots were established with both strong (hydrochloric) and organic (oleic) acid solutions for TAN values of 0.1-10.0 mg of KOH/g of oil.

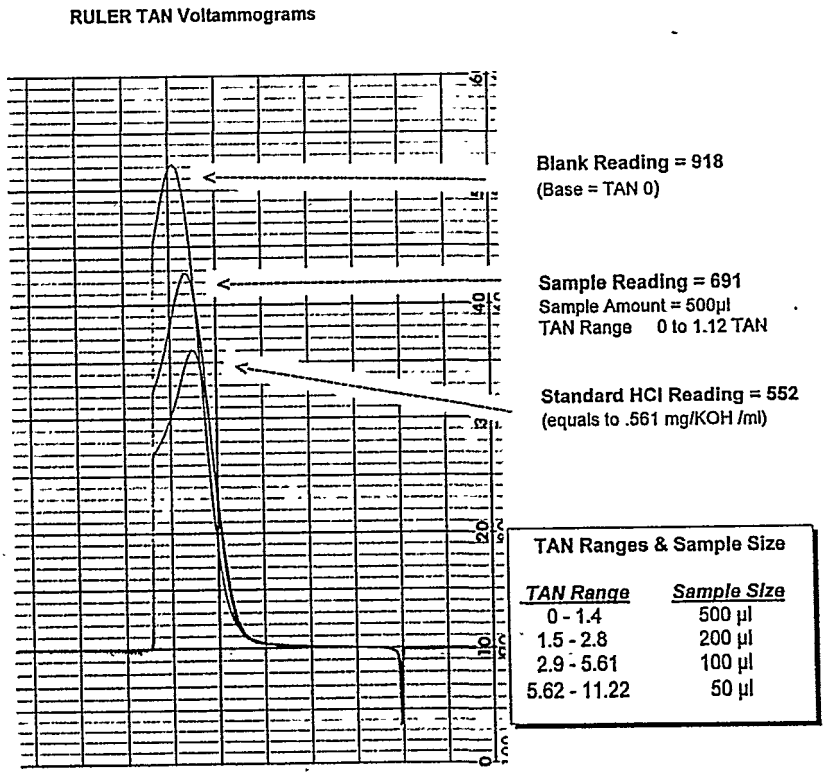


fig 2 : RULER TAN Voltammogram

2.5 Voltammetric TBN Method

The voltammetric TBN method involves adding 0.1-0.3 milliliters of oil to a vial containing 3 milliliters of an ethanolic solution and shaking the closed vial for 20 seconds. The ethanolic solution contains hydrochloric acid (0.05N). The reacted acidic ethanolic solution/oil mixture is then poured into a second vial containing a cuprous oxide or sulfide powder and the closed vial is shaken an additional 20 seconds.
During the second shaking period, any hydrochloric acid not neutralized by the oil sample reacts with the cuprous oxide or sulfide powder to produce cuprous chloride which produces a voltammetric peak. (**figure 3**)
The voltammetric instrument measures the peak height of the cuprous chloride to quantify the amount of acid not neutralized by the oil sample. An alcoholic solution containing 0.1N of hydrochloric acid is used to calibrate the cuprous chloride peak height so that the oil sample's TBN can be calculated. Linearity plots were established for TBN values of 1 to 20 mg of HCl/g of oil.

RULER TBN Voltammograms

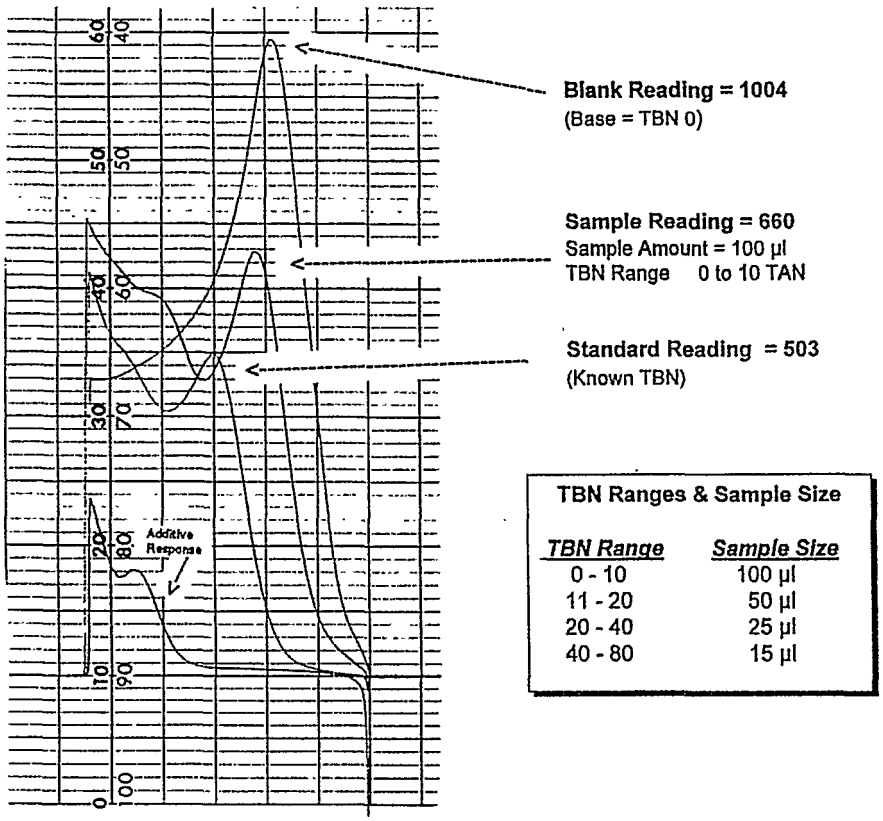


fig 3 : RULER TBN Voltammogram

3 RESULTS AND DISCUSSION

3.1 Diesel Engine Oils - ZDDP

To evaluate the voltammetric (RULER™) additive/TAN/TBN potential for analyzing diesel engine oils containing the ZDDP antioxidant/antiwear additive, used engine oils were obtained from an extended diesel engine test (test usually terminated at 48 hours). The voltammetric additive analyses and the ASTM TAN/ TBN measurements were made by the laboratory personnel performing the engine test. The voltammetric TAN/TBN measurements were performed by the authors.

The resulting voltammetric and ASTM measurements of the new and used diesel engine oils were plotted versus engine operating time as shown in **Figure 4**.
The results in **Figure 4** show that the trending capabilities of the voltammetric and ASTM TAN measurements are in good agreement with the voltammetric values being slightly lower during the useful life of the oil (TAN values stable). The lower voltammetric TAN values are attributed to the fact that the voltammetric test detects the acidic additives (ZDDP, detergents, etc.) to a lesser extent than the ASTM method.

The results in **Figure 4** also show that the voltammetric and ASTM TBN measurements are also in good agreement. The difference between the TBN tests after 48 hours is attributed to the fact that the voltammetric test is unaffected by the organic acids produced by oxidation while the ASTM method is buffered by the produced organic acids resulting in superficially high TBN values (TBN value should decrease, not increase, with use).

In addition to the TAN/TBN measurements, the results in **Figure 4** show that the voltammetric additive (ZDDP) analyses are the only analyses capable of determining the remaining useful life (RUL) of the oil being tested in the engine test. The oxidative RUL of the oil ends (TAN increases) when the ZDDP depletes to approximately 15% of its original concentration. A dramatic increase in the Fe concentration occurred (not plotted) at a ZDDP level of 10%, indicating the initiation of accelerated wear in the diesel engine correlated with the voltammetric ZDDP analyses.

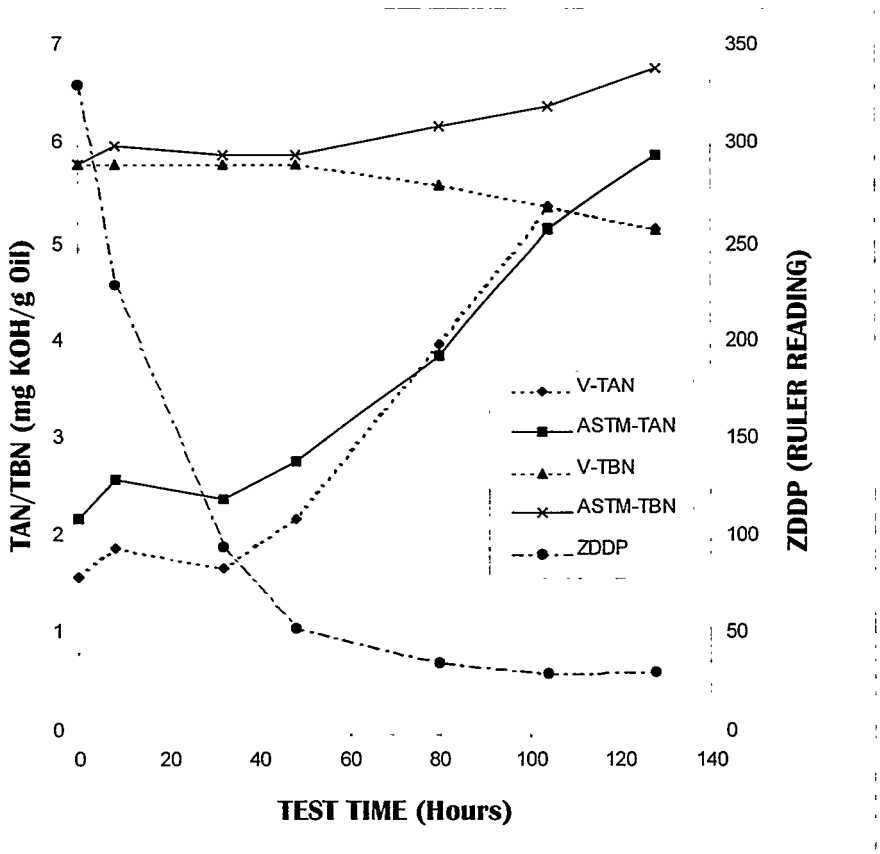


Figure 4. Plots of voltammetric [ZDDP (RULER™), V-TAN and V-TBN] and ASTM analyses versus engine test time for used diesel engine oils obtained from an extended diesel engine test.

In addition to the used oils obtained from the diesel engine test, used oils obtained from truck fleets were also analyzed by the voltammetric additive/TAN/TBN tests. A wide range of diesel engine oils (all contained ZDDP) were used in the fleet studies. Although the types of additives detected by the voltammetric analyses were similar in the different new oils, the concentrations of the additives varied with manufacturer. Consequently, the voltammetric analyses could be used for qualitative assurance testing of new oils (verify additive concentrations of incoming batches) as well as monitoring used oils. The results from the truck fleet evaluations were in good agreement with the engine test stand results. For all of the different diesel engine oils and diesel engines monitored, whenever the ZDDP level of the used oil decreased to below 20% of the original concentration, the TAN and iron concentrations of the used oil would be increased significantly indicating the used oil was at the end of its useful life.

Therefore, the results in **Figure 4** and from the fleet studies for used diesel engine oils demonstrate that the combined voltammetric additive/TAN/TBN measurements provide the user with on-site capabilities to monitor both the RUL and condition of used diesel engine oils with minimal time and cost requirements. The results in **Figure 4** also indicate that the voltammetric TBN values may become more accurate than the ASTM TBN values for used diesel engine oils with increased TAN levels. If necessary, the accuracy of voltammetric TAN results could be improved by taking into account the acidic (ZDDP, etc.) additive concentrations detected by the voltammetric additive analyses.

3.2 Diesel Engine Oils - Non-ZDDP

To further demonstrate the enhanced oil condition monitoring capabilities of the combined voltammetric antioxidant, TAN, and TBN measurements, 30 used oils obtained from normally (scheduled maintenance) and abnormally (unscheduled maintenance due to high wear metal, pentane insolubles, etc.) operating locomotive diesel engines were analyzed. These oils rely on calcium phenate and other non-ZDDP additives for antioxidant/antiwear protection. In every case, the used oil samples obtained from normally operating engines had greater than 60% of the original antioxidant remaining (calcium phenate), TAN values below 0.5 mg of KOH/g of oil, and TBN values (originally 17) above 13 mg of KOH/g of oil sample. In contrast to the normally operating equipment, the used oil samples obtained from abnormally operating equipment had less than 5% of the original antioxidant remaining, TAN values up to 1.5 mg of KOH/g of oil, and TBN values below 11 mg of KOH/g of oil. Consequently, the antioxidant values indicate, and the TAN values confirm that the oils were used past the ends of their useful (oxidatively stable) lives resulting in accelerated oxidation.

4 SUMMARY

This paper has described the development of rapid TAN and TBN tests based on voltammetry. This paper also presents results for the voltammetric and ASTM method (TAN and TBN) analyses of various new, laboratory stressed, and authentic used oils obtained from different types of normally and abnormally operating diesel engines. Interrelationships among the different voltammetric analyses and wear debris analyses of the used oils were demonstrated. The results presented herein demonstrate that combined knowledge of the antioxidant/TAN/TBN (if applicable) measurements improves the capability of the analyst or equipment operator to evaluate the remaining useful life/condition of the tested fluid as well as the condition of the operating equipment. The combined additive/degradation measurements have the following capabilities:

(1) Quality assurance of incoming batches of lubricants - check type and concentration of additives, compare antioxidant packages of different manufacturers, etc.

(2) Determine correct oil change intervals - determine the depletion rate of the antioxidants and the antioxidant concentration at which the fluid degradation rate dramatically increases to predict correct oil change intervals for each piece of operating equipment.

(3) Detect abnormally operating equipment - once the normal depletion rates of the antioxidants have been established, accelerated antioxidant depletion rates can be used to detect abnormally operating equipment.

(4) Monitor non-oxidative degradation mechanisms such as hydrolysis, overheating, and acid accumulation from combustion gases.

Voltammetric antioxidant and TAN measurements of other type lubricants such as phosphate ester hydraulic fluids, transmission fluids, aircraft engine oils, and metal working fluids will be reported in future papers.

REFERENCES

1. Kauffman, R. E., "Remaining Useful Life Measurements of Diesel Engine Oils, Automotive Engine Oils, Hydraulic Fluids, and Greases Using Cyclic Voltammetric Methods," Lub. Eng. 51(3), pp 223-229 (1995).

2. Kauffman, R. E., "On-Line and Off-Line Techniques for Monitoring the Thermal and Oxidative Degradation of Aircraft Turbine Engine Oils - Part 1: Laboratory Evaluations," Lub. Eng., 51(11), pp 914-920 (1995).

3. CRC Handbook of Lubrication and Tribology, "Volume III Monitoring, Materials, Synthetic Lubricants, and Applications," Rapid Determination of Remaining Useful Lubricant Life, pp 89-100, Editor E. Booser, CRC Press, Ann Arbor, MI (1994).

ACOUSTIC MULTIVARIATE CONDITION MONITORING - AMCM

Per-Einar Rosenhave, Research Engineer, Vestfold College,
Maritime Dep., Tønsberg, Norway.
E-mail: Per-Einar.Rosenhave@hive.no

ABSTRACT

In Norway, Vestfold College, Maritime Department presents new opportunities for non-invasive, on- or off-line acoustic monitoring of rotating machinery such as off-shore pumps and diesel engines. New developments within acoustic sensor technology coupled with chemometric data analysis of complex signals now allow condition monitoring of hitherto unavailable flexibility and diagnostic specificity. Chemometrics paired with existing knowledge yields a new and powerful tool for condition monitoring. By the use of multivariate techniques and acoustics it is possible to quantify wear and tear as well as predict the performance of working components in complex machinery. This paper describes the AMCM method and one result of a feasibility study conducted onboard the LPG/C "Norgas Mariner" owned by Norwegian Gas Carriers as (NGC), Oslo.

1. CHEMOMETRICS

Chemometrics is a method which has been developed by chemists over the past ten to fifteen years. The method is highly applicable for almost any set of data. Chemometrics has properties for the reduction of dimensions, thus enabling the user to handle a vast amount of variables simultaneously. Multivariate data analysis forms the basis of all chemometrics. The combination of mathematical, statistical and data analytical methods makes it feasible to extract problem related information from vast and incomprehensible sets of

data and present the results in a simple and easily understood manner. For the purpose of prediction, empirical regression models may be developed. By use of multivariate techniques, such as Principal Component Analysis (PCA) and Partial Least Squares Regression (PLS-R or just PLS), models are developed based on the existing correlation between all variables simultaneously.

Empirical modelling is an extremely powerful alternative compared with the more commonly used theories of model development used in modern science and technology. Chemometric methods do not start from known relations in the data set. They reveal the hidden empirical relations and structures merely from one set of already observed input and output data. System knowledge could be an asset, but is not a prerequisite.

1.1 PARTIAL LEAST SQUARES REGRESSION

The method of partial least squares regression goes further than principal component analysis. PLS includes two matrixes, a Y-matrix, which consists of the dependent variables and an X-matrix which contains the independent variables. The idea is to let the dependent variables Y interact with the independent variables X, while establishing the PLS-components. A PLS-component is the Y-guided analogue of a principal component in an X to Y regression sense. In this way it is possible to establish a regression model for the relation between X and Y. Such a model can then be used for future prediction. If the model is calibrated in the right way, and there is indeed a connection between X and Y in our data, we are able to find a new Y from a given X, or predict Y from X. If there is a correlation between the "noise" emitted from a specific machine and the condition of the machine with PLS, then we can use the "noise" as variables in the X-matrix and predict the condition of the machine as the variable in Y. At first PLS can be regarded as an "individual" PCA performed on the X-matrix and one performed on the Y-matrix, but with the *very important* difference that in the case of a PLS, the two models are *not* independent. The structure in the Y-matrix is used as a guideline while decomposing the X-matrix to determine the model. With PLS the idea is not to describe as much as possible of the variation in X, but *to seek that part of X that is most relevant for the*

description of Y. Two types of PLS exist: the PLS1 and the PLS2. The difference between them is that a PLS1 models only one Y-variable, while a PLS2 can model several Y-variables.

2. AMCM

Acoustic Multivariate Condition Monitoring, or AMCM for short, is a novel combination of the best from proven sensor technology and new sophisticated data analysis. This combination allows a complex and powerful analysis of machine condition, that does not require extensive personnel training. AMCM can be specifically designed to allow engineers access to monitoring of results and machine condition diagnosis displays.

The use of well-proven sensor technology assures flexible and reliable signals over a wide range of acoustic levels and frequencies; both AE sensors as well as accelerometers may be used as sources either alone or in combination with ordinary process- data and signals. However, it is the application of powerful chemometric data analysis methods that distinguishes AMCM from both the simpler AE and frequency domain vibration approaches. Indeed, AMCM makes use of only the best information from both these successful techniques, and adds the new powerful feature of multivariate calibration data analysis to the task. This allows empirical calibration of virtually any type of instruments and/or signals. Thus, AMCM is not relying on specific sensors or one specific method of analysis. Rather AMCM is composed of a set of flexible options for the optimal combination of sensors and their accompanying pre-processing methods, which act as an alternative, or complementary, input for signals and for multivariate instrument and signal calibrations.

AMCM is a new generic approach, which may use any relevant acoustic signal to characterize machine operation. AMCM implies that passive and/or active measuring techniques based on acoustics, process data, signal conditioning combined with multivariate analysis are used to detect and predict changes that occur in various engine systems. Such changes can occur due to normal wear, sudden damage or by variation in operating parameters. The changes can be acoustically detected, processed and modeled. The resulting models can be used for example for prediction and predication of the

percentual deviation of wear from the nominal condition, detection of a sudden damage or for optimized process control. In the near future we foresee an engine control room where the engineer can press a button and instantly see a comprehensible display of relevant data. Such information could be, for example: "The fuel valve performance in cylinder five is reduced to 70% of its potential" or, "Cylinder blow-by has increased by 28%" or, "There is a leakage in the second stage pressure valve of the cargo compressor." All this can be achieved by use of non-intrusive clamp-on sensors and multivariate analysis. Multivariate calibration is, of course, essential. By careful calibration it will be possible to extract the desired condition and performance data of various components from one sensor. AMCM is based on the age-old fact that it is possible to hear vibrational changes and emissions due to wear, tear, damage and changes in operational parameters.
AMCM will give increased functionality, early warning and will also satisfy the demand for a simple, robust and cost-effective measuring equipment. AMCM is patent pending.

2.1 COLLECTING ACOUSTIC SIGNATURES

The different acoustic signatures, which form the base for the subsequent analysis, originate from different physical processes within the machinery itself and represent energy released as a consequence of the machinery being run at a specific load, being in a particular condition or the like.

This paper describes how wear and tear and sudden damage of a sea-water pump impeller can be detected and quantified by the use of AMCM. The "sound" or the acoustic signature of the pump was collected by an accelerometer mounted in different positions either on the pump housing or on the connecting pipelines. See fig 1. The analogue signal from the accelerometer was amplified and low-pass filtered in a signal amplifying unit close to the signal source, before being digitalized by the use of an analogue to digital converter (ADC). This digital signal could then either be processed directly or stored for later processing and analysis. The amplification and low-pass filtering is necessary in order to gain the best signal to noise ratio if aliasing- or folding- effects are to be avoided.

Figure 1 shows the main sea-water pump used during this experiment. Sea-water enters the pump from the right hand side of Figure 1. Hopefully test-point two and three is visible.
Test-point number six is located on the back of the pump-housing, between location two and three. Test-point one is located on the outlet side of the non-return valve, while test-point number four is located on the right hand side of the middle piece between the pump-housing and the electrical motor.

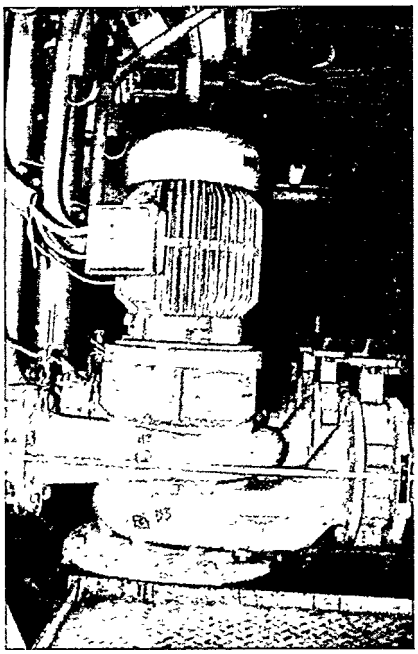


Figure 1: The sea-water pump with some of the test-points shown. Flow exit and test-point no. one is to the left. Test-point no. six is behind and between location two and three.

Figure 2: The acoustic signal is picked up by a simple uni-axial accelerometer mounted directly on the component to be monitored.

Figure 2 illustrates how simple it is to attach an accelerometer to a component. The attachment is provided by the use of cementing studs glued onto a chosen location on the component in question. Thus one and the same accelerometer can be used for all measurements at all locations, securing reproducibility at all times.

2.2 PRE-ANALYSIS SIGNAL PROCESSING

After digitalization, see Figure 3, the acoustic signatures are usually transformed from the time domain to any another domain more suitable for the subsequent multivariate analysis. The choice of the "correct" transformation method to use is "problem dependent", but transformations like the well known Fast Fourier Transform (FFT) or the newer Angular Measuring Technique (AMT) could be used.

Given that FFT was chosen, different techniques could be used in order to achieve the best signal to noise ratio and keep as much of the information from the time-domain as possible.

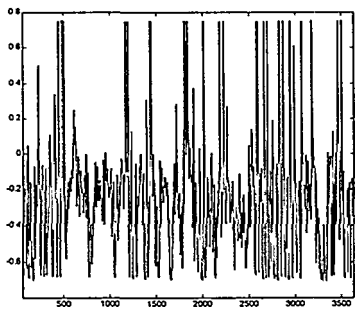


Figure 3: The raw-signal, that is the signal in time-domain after digitalization.

In Figure 4 the resulting power spectrum is shown as a result of an estimation of the power spectrum by use of the Welch periodogram method. During transformation the data sequence was divided into equal length data sections. The data sections were overlapped by 75% and the periodograms were computed from these data sections after detrending and a subsequent Hanning-windowing. The resulting signal contained frequencies between 0 and 44 kHz.

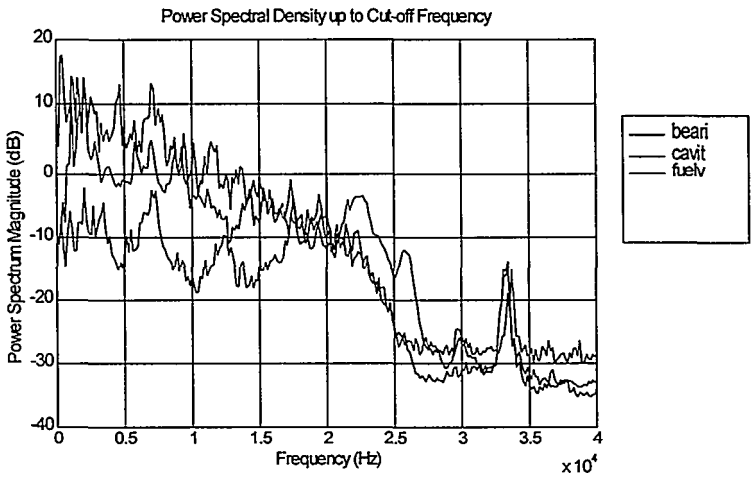


Figure 4: Power Spectral Density of three different raw-signals. The over-all signal bandwidth is 40 kHz.

3. ACOUSTIC CHARACTERIZING WITH A PLS1-DISCRIM

The purpose of the described test was to establish whether or not it was possible to discriminate acoustically between the different induced damages, a worn impeller and a new impeller in a working pump.

In order to accomplish this, a PLS1-discrim X-matrix was compiled consisting of collected data from four principally different conditions: the signature of the old worn impeller, two induced damages "a" and "b", and finally, the signature of the pump with a brand new impeller mounted. The two induced damages were introduced by removing two small pieces of metal from the inlet section of the impeller. In the case of damage "a", a piece of 25 g was removed, while in case "b", another 15 g was removed. In total the removed mass adds up to only 0.13% of the total impeller mass of 30 kg.

All measurements were done with a new wearring mounted in the pump and the acoustical signature from the four different conditions were recorded at seven different locations on the pump housing and on the connecting pipes. The X-matrix therefore consists of 28 objects and 256 frequency variables. The objects have been labeled w1 to w7, a1 to a7, b1 to b7, and n1 to n7 respectively, representing the worn impeller, the two induced damages "a" and "b" and finally the new impeller.

The PLS1-discrim Y-matrix contains one response variable specially designed to quantify the wear and tear and the two induced damages. Establishing such a variable can be a difficult task. As an example, the removal of a small piece of metal from the impeller inlet will not only manifest itself as a damage to the impeller. It will also introduce new acoustic signatures due to the effects of unbalance, changes in the flow pattern, turbulence or other factors. One possible way of establishing such a response variable would be to use the impeller mass or use the weight reduction relative to the mass of a new impeller. Another quantification method, and the one chosen, was to give all instances where the new impeller was used the factor zero. The worn impeller was given a factor one, and the two damages were given a factor three and a four respectively. This semi-quantification index includes both the induced damages and normal wear. The

PLS1-discrim Y-matrix is then made up of 28 objects and a single response variable in the range of 0 to 4.
A cross validated PLS1-discrim model was developed using centered and scaled data-matrixes. Using two PLS1 components, the model was able to explain 75% of the variance in the Y-space using 66% of the variance in the X-space.

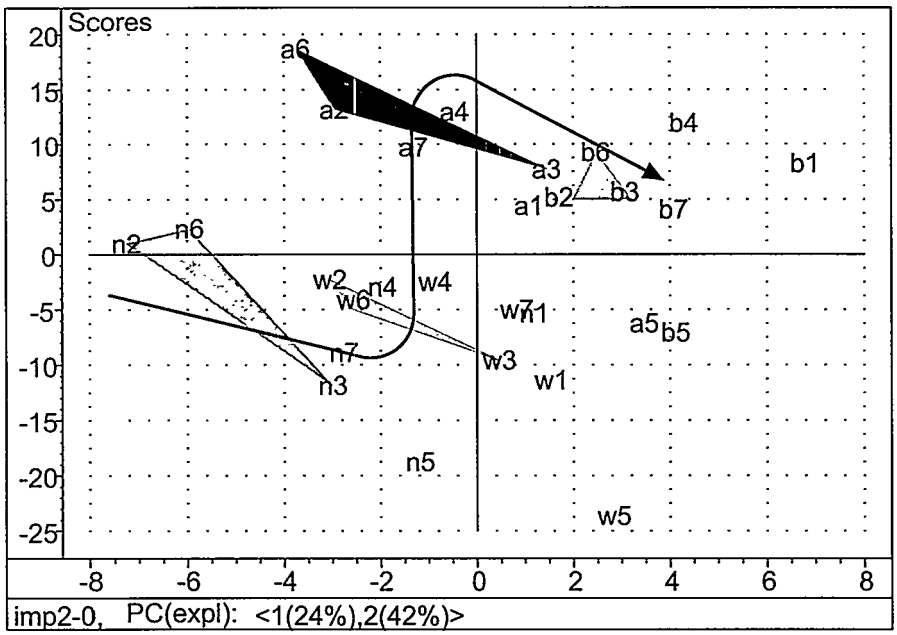


Figure 5: Wear appears as a movement to the right, while a damage will move the locations up in the t1t2 plot

Figure 5 shows the PLS1-discrim t1t2 plot with an arrow. The figure shows seven different sensor positions and the four cases of wear and damage.
The signature, with the sensor in position one and with the new impeller installed, is labeled n1. The signature from a worn impeller picked up at the same position has been labeled w1. Signatures from the two induced damages taken at the same sensor position has been labeled a1 and b1 respectively.

The arrow represents the evolution of the acoustic spectra due to the four different situations that have been examined. The same trend can be found by following the four signatures of wear and damage from sensor position five, namely n5 → w5 → a5 → b5. When a new

impeller is gradually worn, its signature will move mainly to the right in the plot. Introducing a damage by breaking off a part of the impeller makes the new spectra move upwards in the plot. This sudden change in position indicates that a new situation has emerged, which is very unlike the plot representation caused by normal wear. Removing yet another part of the impeller does not represent a new situation. However, it merely manifests itself as "yet another damage" moving to the right on the "damage level". From this plot it can be seen that development in wear or damage will move the positions of the spectra to the right along t1, while a change caused by a sudden damage will cause the spectra to move upwards in the plot along t2. This plot demonstrates, with all clarity, that the combination of acoustics and chemometrics in wear and damage monitoring has arrived!

4. CONCLUSIONS

The AMCM approach proposed in this paper is based on the combination of using acoustic signatures and different sensor technology coupled with "chemometric" data analysis of complex signals. The approach has been illustrated for normal impeller wear in combination with induced impeller damage. It was found that a PLS1-discrim performed on the X-matrix with a designed Y- response-matrix was able to discriminate clearly between the four experimental conditions; a new impeller, a worn impeller, an induced damage and secondary induced damage to the impeller.

PLS1-discrim represents a very effective technique, which will establish a response-optimized regression between X and Y. In this case the X-matrix contains Fourier-power-spectra based upon the individual acoustic signatures. The response variables corresponding to the acoustic signatures are contained in the Y-matrix. In this case these response variables express quantifications of wear and sudden damage. When the model has been developed and validated, these response variables are no longer needed. They can be discarded and a newly collected acoustic signature can be fed directly into the model resulting in the production of new process and maintenance data, which can predict failure.

5. BIBLIOGRAPHY

Rosenhave P. E., *Acoustic Chemometric Condition Monitoring of Marine Components - A Feasibility Study*, Vestfold College, Norway, 1996.

Esbensen K. H.; Schönkopf S.;Midtgaard T., *Multivariate Analysis - in practice*, Camo AS, 1994, ISBN 82-993330-0-8.

Geladi, P.; Kowlaski, B. R., *Partial Least Squares Regression: A Tutorial*,Anal. Chim. Acta, 1986, 185, 1-17.

Martens H,; Næs T.; *Multivariate Calibration*, J. Wiley & Sons, 1994, ISBN 0-471-93047-4.

Randall R. B., *Frequency Analysis*, 1987, Brüel & Kjær, ISBN 87-87355-07-8.

Trampe Broch J., *Mechanical Vibration and Measurements*, 1984, Brüel & Kjær, ISBN 87-87355-36-1.

6. ACKNOWLEDGMENTS

The author gratefully acknowledge the participation of Norwegian Gas Carriers (NGC), Oslo, in this work. Their tolerance and support has been invaluable.

PREDICTIVE MAINTENANCE AND INSPECTION THROUGH AIRBORNE ULTRASOUND TECHNOLOGY

Alan Bandes
UE Systems, Inc.
Elmsford, NY 10523-2536 USA

ABSTRACT

Airborne ultrasound can be considered an ideal integrating technology in that these instruments can stand alone to detect a variety of potential problems or they can be used to support vibration and infrared inspection programs. Usually portable, these instruments detect leaks in both pressurized gas systems or vacuum systems and related equipment such as tanks, pipes, heat exchangers, valves and steam traps. Additional applications include inspection of high voltage apparatus for corona, arcing and tracking. They are used to trend bearing failure as well as to detect conditions such as lack of lubrication and rubbing. A brief overview of the technology its applications and suggested inspection techniques are explained.

BACKGROUND AND INTRODUCTION

In many companies and institutions, there has been an almost compelling need to find areas of cost reduction. For many this has meant the cutting of budgets which unfortunately included a cutting of personnel. In many departments, this has resulted in many doing more with less. This has also impacted on many maintenance and production departments around the world. In order to maintain an efficiently running operation, there has been a trend towards various forms of Predictive Maintenance. Along with this trend has come a need for effective technologies and instruments that enable personnel to become more and more productive in terms of equipment inspection and problem solving. Opportunities in predictive maintenance have lead to improvements in production, reduced maintenance costs, reduced energy consumption, efficient use of personnel and increased profitability.

ultrasound. Of the various technologies, one of the least understood is airborne ultrasound. This technology can be considered an integrated technology since it can be used with both infrared and vibration inspections as well as stand alone to perform a multiplicity of inspection activities. Instruments based on this technology apply to a wide range of plant operations and yet are simple enough to be used with little training. This makes them ideally suited for such programs as TPM (Total Productive Maintenance) in which they may be used by both maintenance and production personnel.

Light weight and portable, these instruments may be used to inspect potential problems in practically every type of equipment and system in a plant. Some typical applications include:

> leak detection in pressure and vacuum systems such as, boilers, heat exchangers, condensers, chillers, distillation columns, vacuum furnaces, specialty gas systems, bearing inspection, steam trap inspection, detection of valve blow-by, detection of cavitation in pumps, detection of corona in switch gear, compressor valve analysis, even the integrity of seals and gaskets in tanks, pipe systems and large walk-in boxes.

What makes airborne ultrasound so effective? All operating equipment and most leakage problems produce a broad range of sound. The high frequency ultrasonic components of these sounds are extremely short wave in nature. A short wave signal tends to be fairly directional. It is therefore easy to isolate these signals from background plant noises and to detect their exact location. In addition, as subtle changes begin to occur in mechanical equipment, the subtle, directional nature of ultrasound allows these potential warning signals to be detected early, before actual failure.

Airborne ultrasound instruments, often referred to as "ultrasonic translators", provide information two ways: qualitative through the ability to "hear" ultrasounds through a noise isolating headphone and quantitative via incremental readings on a meter.

Although the ability to gauge intensity and view sonic patterns is important, it is equally important to be able to "hear" the ultrasounds produced by various equipment. That is precisely what makes these instruments so popular. They allow inspectors to confirm a diagnosis on the spot by being able to clearly discriminate among various equipment sounds. This is accomplished in most ultrasonic

translators by an electronic process called "heterodyning" that accurately converts the ultrasounds sensed by the instrument into the audible range where users can hear and recognize them through headphones.

The high frequency, short wave characteristic of ultrasound enables users to accurately pinpoint the location of a leak or of a particular sound in a machine.

Most of the sounds sensed by humans range between 20 Hertz and 20 kilohertz (20 cycles per second to 20,000 cycles per second). The average human threshold is actually 16.5 kHz. These frequencies tend to be relatively gross when compared with the sound waves sensed by ultrasonic translators. Low frequency sounds in the audible range are approximately 1.9 cm (3/4") up to 17 m (56') in length, whereas ultrasounds sensed by ultrasonic translators are only 0.3 cm (1/8") up to 1.6 cm (5/8") long. Since ultrasound wave lengths are magnitudes smaller, the "ultrasonic environment" is much more conducive to locating and isolating the source of problems in loud plant environments.

The basic advantages of ultrasound and ultrasonic instruments are:

1. they are directional and can be easily located
2. They provide early warning of impending mechanical failure
3. instruments are easier to use in a loud, noisy environment
4. They support and enhance other PDM technologies or can stand on their own in a maintenance program

INSTRUMENTATION

Airborne Ultrasound translators are relatively simple to use. They consist of a basic hand held unit with headphones, a meter, a sensitivity adjustment, and (most often) interchangeable modules that are used in either a scanning mode or a contact mode. Some instruments have the ability to adjust the frequency response from between 20 to 100 kHz. An ultrasonic transmitter called a tone generator is often included.

Many of these features are useful in helping a user adapt to a specific test situation. As an example, should a ultrasound source be too difficult to locate due to an intense signal, a downward adjustment of the sensitivity, will help a user focus in on the exact site. In another instance, should a low level leak occur in a water valve, the frequency tuning can be adjusted to help a user hear the trickle of the water leak.

The interchangeable modules allow users to adjust for different types of inspection problems. The scanning mode is used to detect ultrasounds that travel in the atmosphere such as a pressure leak or a corona discharge, while the contact mode is used to detect ultrasounds generated within a casing such as in a bearing, pump, valve or steam trap housing.

APPLICATIONS

Generically, applications for ultrasonic translators fall under three basic categories: mechanical inspection, leak detection and electrical inspection.

MECHANICAL INSPECTION

Operating mechanical equipment produces a "normal" sound signature while operating effectively. As components begin to fail a change in the original sonic signature occurs. This change can be noted on a meter or a recording device such as an oscilloscope, portable PC, portable vibration analyzer or chart recorder. The sound quality will be heard through headphones.

The key to mechanical inspection relies on a consistency factor. Variables should be kept to a minimum. To accomplish this, whether trouble shooting or trending equipment, a test point should be established. This test point can then be used for comparison with other test points on similar equipment or compared with itself over time.

As an example, for bearing inspection, in order to determine whether a bearing is in a good or failed mode, touch the bearing housing using the contact probe at one point, usually the grease fitting, and adjust the sensitivity to get a specific meter reading. Compare this reading at the same sensitivity setting on a similar reference point on a bearing operating under the same conditions. The meter reading and the sound quality should be similar. This same reading can then be used to trend each bearing over time to determine lack of lubrication or failure mode.

According to NASA research[1], ultrasonic monitoring of bearings provides the earliest warning of bearing failure. They noted that an increase in amplitude of a monitored ultrasonic frequency of 12-50 times over baseline will indicate the initial stages of bearing failure. This change is detected long before it is indicated by changes in vibration or temperature.

With the above described procedure, problems such as cavitation in pumps, compressor valve leakage, faulty gears, excessive rubbing, poor connections, etc., can be determined in their early stages before break down.

Should an existing vibration program already exist for bearing analysis, it should be understood that an ultrasonic bearing monitoring program works very well with on-going vibration programs.

Ultrasound translators can be used to aid a diagnosis. The high frequency, short wave characteristic of ultrasound allows the signal to be isolated so that a user can hear and determine if, in fact, a bearing has been correctly diagnosed as failed.

At times there can be false signals generated by equipment connected to a particular bearing. By adjusting the sensitivity, the frequency and listening to the sound, it can be determined whether it is the bearing or a rotor or something else that is the root of the problem. The ability to hear what is going on can prove very important.

Ultrasound detectors work well on slow speed bearings. In some extreme cases, just being able to hear some movement of a bearing through a well greased casing could provide information about potential failure. The sound might not have enough energy to stimulate classic vibration accelerometers, but will be heard via ultrasonic translators, especially those with frequency tuning.

ULTRASONIC INSPECTION AND VIBRATION ANALYZERS

A recent development in the field of ultrasonic inspection has been to connect these instruments with portable vibration analyzers. Historically vibration instruments of this type have been limited to receiving only the lower frequency "vibration" signals produced by operating equipment. By receiving the heterodyned ultrasonic signal, the capacity of these instruments has been greatly enhanced.

As an example, one inspector , Mr. Steve Seeber of Mid Atlantic Infrared recently reported on some case histories[2] describing his use of combining ultrasonics with vibration. In one instance a boiler feedwater pump had a history of failures. His vibration spectra displayed no bearing fault yet there was a distinctive buzzing noise that suggested a problem developing. By feeding the ultrasonic information into the vibration spectra the calculated fault frequencies for this bearing corresponded precisely to the peak frequencies showing the demodulated ultrasound spectrum. Once the bearing was replaced, the audible buzz that was previously heard disappeared.

Ultrasound is also effective in detecting problems related to impact and rubbing. Mr. Seeber also related an experience in which a gear box was producing an unusual noise. His checks with vibration and sound revealed a spectrum with numerous harmonics of the input shaft speed but the sound was uniformly distributed around the gearbox. It was assumed that the noise was a looseness around the input shaft or clutch. An ultrasonic inspection demonstrated a strong signal at the drive side input shaft bearing. When the bearing was replaced, the noise was eliminated.

LEAK DETECTION

The category of leak detection covers a wide area of plant operations. It can be looked on as a way of keeping a system running more efficiently. Some plants include it as part of an energy conservation program while others refer to it as fugitive emissions. No matter what, leaks can cost money, effect product quality and reek havoc with the environment.

Leakage can occur in liquid or gas systems. The greatest advantage of ultrasonic detection is that it can be used in a variety of leak situations since it is sound sensitive and not "gas" specific.

The reason ultrasound is so versatile is that it detects the sound of a leak. When a fluid (liquid or gas) leaks, it moves from the high pressure side of a leak through the leak site to the low pressure side where it expands rapidly and produces a turbulent flow. This turbulence has strong ultrasonic components. The intensity of the ultrasonic signal falls off rapidly from the source. For this reason the exact spot of a leak can be located.

The method of generalized gas leak detection is quite easy. All one does is scan an area, listening for a distinct rushing sound. With continued adjustments of the sensitivity, the leak area is scanned until the loudest point is heard.

Some instruments include a rubber focusing probe which narrows the area of reception so that a small emission can be pinpointed. The rubber focusing probe is also an excellent tool for confirming the location of a leak by pressing it against the surface of the suspected area to determine if the sound of the leak remains consistent. If it decreases in volume, the leak is elsewhere.

Vacuum leaks may be located in the same manner. The only difference is that the turbulence will occur within the vacuum chamber. For this reason, the intensity of the sound will be less than that of a pressurized leak. Even though it is most effective with low-mid to gross leaks, the ease of ultrasound detection makes it the instrument of choice for most vacuum leak problems.

Liquid leaks are usually relegated to determining leakage through valves and steam traps although there have been some reports of success in locating water leaks from pressurized pipes buried underground. The key to determining whether a product can be checked for leakage is its ability to produce some turbulence as it leaks.

Valves are usually checked for leakage with the contact probe. The downstream side is used to determine leakage. This is accomplished by first touching the upstream side and adjusting the sensitivity to read about 50% of scale. The downstream side is then touched and the sound intensity is compared. If the signal is lower than upstream, the valve is considered closed. If it is louder than upstream and is accompanied by a typical rushing sound, it is considered to be leaking.

Steam traps are also inspected easily with ultrasonic translators. It is important to determine exactly how a particular trap is supposed to operate. This can be accomplished by consulting with steam trap suppliers. In some instances, manufacturers of ultrasonic translators supply video cassette training tapes that show exactly how each type of trap can be inspected.

The method is quite simple. A steam trap is touched with the contact probe. By listening to the trap operation and by observing the meter, trap condition can be interpreted.

The speed and simplicity of this type of test allows every trap in a plant to be routinely inspected.

Leaking tubes in heat exchangers and condensers as well as boiler casing leaks are detectable with ultrasonic translators.

In most power plants, the problem of condenser in-leakage is a major cause of concern. Most often condenser fittings are routinely inspected utilizing the leak detection method previously described.

Should a leak be suspected in a condenser tube bundle, it is possible to locate the leak by putting a condenser at partial load and opening up a water box of a

suspected tube bundle. After the tube sheet is cleared of debris, the tube sheet is scanned.

Heat exchangers may be tested in a similar fashion. The header is removed and the shell side is either placed under vacuum or is pressurized.

There will be some instances where it is difficult or too time consuming to inspect under pressure or vacuum. In this instance a test unique to ultrasound is incorporated. This method uses an ultrasonic transmitter called a "tone generator". The "tone generators" are placed in the various access ports or fittings to produce an intense, uniform ultrasound within the shell side. Since the sound waves are high frequency, short wave, they will tend to deflect off the solid surface of solid, intact tubes but will penetrate through the leak site of a tube. By scanning the tube sheet with the ultrasonic scanner an operator listens for a distinct high frequency signal indicating the leaking tube.

The preferred method is pressure or vacuum, but the tone generator method is a good back-up for difficult situations.

ELECTRIC INSPECTION

Electrical problems are also detected with ultrasonic translators. When arcing, tracking or corona discharges occur, they ionize air molecules producing ultrasound. Loose connections, buss bars, junction boxes, etc., can be listened to for the high frequency sounds of an arc. This will usually be heard as a buzzing or frying sound in the headphones.

Another area of inspection for ultrasonic detectors is switchgear and overhead high voltage lines for location of corona or tracking problems. Although infrared has often been used to locate electrical problems, it has been found that these instruments are often "blind" to corona and tracking in high voltage systems (13 KV and up). Ultrasonic detectors "hear" the sound of corona and enable users to locate them quite quickly. For this reason, many inspectors now use ultrasonic translators to support their infrared electric monitoring programs. In fact, those inspectors that use both technologies often relate that they prefer to screen enclosed switchgear with ultrasound instruments to detect the possibility of corona, arcing or tracking by scanning around door seals and air vents.

One of the most impressive new developments in ultrasonic scanning is the ultrasonic wave form concentrator, a parabolic reflector that is used to detect

leakage at a distance. These accessories are useful in situations where it is difficult to reach a leak area, such as in a high ceiling, or where the leak might prove dangerous, as with high voltage leaks or hazardous gases.

WHAT TO LOOK FOR IN AN ULTRASONIC TRANSLATOR

Before purchasing equipment of this nature, investigate the reliability of the company. As with any system or instrument you purchase, you want to be sure of receiving good after-sales service. In addition, check to see what type of training support you will receive. Are there instruction manuals? Are they easily understood? Are any video or audio cassette programs included or available to assist new staff learn how to use the instrument in various applications?

Technically, it is advisable to have instruments that are sensitive enough to detect the type of problem typically encountered in a plant. A wide dynamic range in an instrument opens up the spectrum of testing which may include small leaks on one end and gross mechanical problems on the other.

Since sound quality is an important aspect of this type of diagnosis, make sure the instrument heterodynes the ultrasonic signal. This will insure users are getting an accurate reproduction of the ultrasonic signal which provides signal clarity and ease of interpretation of the sound quality they receive in the headphones. Heterodyning is an important characteristic for effectively connecting with vibration analyzers and for computerized sound analysis programs.

Noise attenuating headphones with good sound quality are essential. If the sound quality is not clear, then, it will be difficult to understand what is being sensed. It is advisable to get the over-the-ear headphones that will block out ambient plant sounds during your inspections.

Surprisingly some instruments do not have RF shielding. Without proper shielding, stray electronic signals will interfere with test results. In some instances, radio programs have been reported being heard which totally confused operators. So be sure that your instrument has RF shielding.

Since every plant is different, there might be special accessories needed to assist in some testing situations. Check for the availability of accessories to meet the needs of special tests. As an example, for compressor valve analysis, a magnetically mounted probe with oscilloscope interface might be important to you.

If you are going to inspect a variety of equipment or have fluids of different viscosities, it would be useful to have the ability to change frequencies. This is a feature offered by some instruments.

For leak detection of potentially explosive or flammable gases, it is advisable to use equipment rated intrinsically safe.

CONCLUSION

Airborne ultrasound instruments are becoming an important part of Predictive Maintenance, Fugitive Emissions and Energy Conservation programs. Their versatility, ease of use and portability enable managers to effectively plan and implement inspection procedures. By locating leaks, detecting high voltage electrical emissions and sensing early warning of mechanical failure, these instruments contribute to cost reductions, improved system efficiencies and reduced downtime. For optimum effectiveness, it is recommended that all major technologies, infrared, vibration and ultrasound, be used as part of a comprehensive inspection program.

REFERENCES

1. B72-10494, " A System for Early Warning of Bearing Failure" NASA Tech Brief, Marshall Space Flight Center (August 1972)

2. Seeber, S. "Ultrasound Testing Enhances Vibration Analysis", P/PM Technology, Volume 8, Issue 5, October, 1995

DIESEL ENGINE COOLANT ANALYSIS, NEW APPLICATION FOR ESTABLISHED INSTRUMENTATION

Daniel P. Anderson, Malte Lukas and Brian K. Lynch
Spectro Incorporated
Littleton, MA, U.S.A.

ABSTRACT

Rotating disk electrode (RDE) arc emission spectrometers are used in many commercial, industrial and military laboratories throughout the world to analyze millions of oil and fuel samples each year. In fact, RDE spectrometers have been used exclusively for oil and fuel analysis for so long that it has nearly been forgotten by most practitioners that when RDE spectrometers were first introduced more than 40 years ago, they were routinely used for aqueous samples as well. This paper reviews early methods of aqueous sample analysis using RDE technology.

This paper also describes recent work to calibrate an RDE spectrometer for both water samples and for engine coolant samples which are a mixture of approximately 50% water and 50% ethylene or propylene glycol. Limits of detection determined for aqueous standards are comparable to limits of detection for oil standards. Repeatability of aqueous samples is comparable to the repeatability achieved for oil samples.

A comparison of results for coolant samples measured by both inductively coupled plasma (ICP) and rotating disk electrode (RDE) spectrometers is presented. Not surprisingly, RDE results are significantly higher for samples containing particles larger than a few micrometers.

Although limits of detection for aqueous samples are not as low as can be achieved using the more modern ICP spectrometric method or the more cumbersome atomic absorption (AA) method, this paper suggests that RDE spectrometers may be appropriate for certain types of aqueous samples in situations where the more sensitive ICP or AA spectrometers and the laboratory environment and skilled personnel needed for them to operate are not conveniently available.

1 INTRODUCTION

When an internal combustion engine burns fuel, heat is created at temperatures as high as 4000°F (2200°C). This heat must be removed by some form of cooling. The two most common ways to dissipate heat are by air cooling or liquid cooling. This paper discusses only the liquid cooling systems used in most modern engines.

A liquid cooling system contains the following components; radiator, fan, thermostat, water pump, engine water jacket and the cooling liquid. This paper will concentrate on the liquid contained inside the cooling system.

The analysis of used coolant samples has been a successful technique for scientific preventive maintenance. It is applicable to any closed loop cooling system, but is applied primarily to diesel and gasoline engines because they are the most likely component to suffer from a poorly operating cooling system. Overheating causes oil deterioration, oxidation, reduced lubricity and damage to all oil wetted components. The longevity of liquid cooled transmission and hydraulic system components are also dependent on a properly operating cooling system. A properly maintained cooling system not only prevents overheating but also maintains a constant engine temperature. Improperly maintained engine temperatures can result in the type of problems shown in Table 1.

Table 1 - Problems due to Improper Engine Temperatures

High Temperature Problems	Low Temperature Problems
• Pre-ignition • Detonation/Knock • Lubrication Failure • Burnt Pistons & Valves	• Unnecessary Wear • Poor Fuel Economy • Accumulation of Water & Sludge

2 WHY COOLANT ANALYSIS ?

Most people don't give much thought to the condition of their engine coolant system until it is too late. More than 40% of all diesel engine maintenance problems can be attributed to poor cooling system maintenance [1]. Poorly operating cooling systems cause the engine to run hotter which in turn causes the lubricant to oxidize and loose it's lubricity thus causing abnormal wear in all oil wetted areas. The following is a list of reason why to do coolant analysis.

- Protect against gel formation
- Protect against corrosion and rust

- Protect against over/under concentration of SCA's
- Extend drain intervals
- Protect your engine
- Environmental/disposal concerns

Another factor that recently gained world-wide attention is the impact of used coolant disposal on the environment. Ethylene glycol is extremely hazardous if ingested by humans or animals. Because of this, most large users of coolants operate the coolants longer in order to reduce the need for disposal. Others have started recycling and reconditioning the used coolants. Because of this, the need for coolant analysis has increased dramatically over the past few years. Disposal of used coolants can be difficult and expensive and must be done in accordance with local, state or federal laws.

The following is the Cummins Engine Company recommended cooling system maintenance intervals: [2]

- Replace coolant filter at every oil change.
- Top off the cooling system at filter changes.
- Test/replenish SCA package at filter change.
- Test the coolant twice a year.
- Replace coolant every two years or 240,000 miles (6,000 hours).

The following is the Caterpillar Inc. (CAT SOS coolant analysis) recommended cooling system maintenance intervals: [3]

- Every 250 hours check glycol level, freeze and boil protection, SCA concentration, pH, and conductivity.
- Every 1000 hours or a minimum of twice a year check the same as above, plus identify metal corrosion, contaminant levels and built-up impurities.

3 COOLANT ANALYSIS

To be effective, a used coolant analysis program should determine both the coolant condition and the presence of any contaminants or debris. The coolant fluid can be used as a diagnostic medium as the coolant carries not only heat away from the engine parts but also carries fine debris from the interior surfaces of the cooling system. Analysis of the wear debris can provide important information about the condition of the internal parts of the cooling system. However, the condition of the coolant itself is important to know. Does the coolant meet specification? Is the SCA package correct? Is the coolant contaminated with solids, metal particulate or chemical degradation products?

In a modern condition monitoring program based on coolant analysis, a coolant sample is taken from a piece of equipment at periodic sampling intervals and sent to the laboratory for analysis. Based on the analysis, a diagnostic report is made and a recommendation is sent to the personnel responsible for the equipment. The report may show that everything is normal, warn of a possible problem or make a specific maintenance recommendation. The entire process, from sample taking to the diagnostic report, should take less than 24 hours. In a modern coolant analysis program, the data generated and collected by the laboratory is also used to provide periodic maintenance summaries. These reports can be statistical in nature and provide an insight to management personnel on the effectiveness of the program, efficiency of the maintenance department, repair status of equipment, recurring problems, and even information on the performance of coolants.

Condition monitoring by coolant analysis can be broken down into two main categories: Debris Monitoring and Coolant Condition Monitoring. Debris Monitoring spectrochemically measures the trace elements carried away from the cooling system by the coolant. Coolant Condition Monitoring determines if the coolant itself is fit for service based on physical and chemical tests. These two techniques, when combined with statistical trending and data-based management, provide a complete program of condition monitoring by coolant analysis.

3.1 Debris Monitoring for Coolants

Debris monitoring pertains primarily to the detection of metallic wear particles, corrosion products, degradation products and contaminants. Spectroscopy is the most widely applied technique for debris monitoring. Commercial labs in the USA have been using either ICP or AA spectrometers for coolant analysis. Table 2 lists elements routinely detected and quantified for coolant analysis.

Table 2 - Elements Routinely Detected and Quantified Coolant Analysis

Wear Metals	Contaminants	Additives
Iron	Silicon	Potassium
Zinc	Magnesium	Silicon
Lead	Calcium	Boron
Copper		Sodium
Aluminum		Molybdenum
Magnesium		Phosphorus

Table 3 lists the typical elements which are routinely analyzed and provides examples to their origin in a diesel engine cooling system.

Table 3 - Sources of Various Elements in Used Coolants

WEAR METALS
Iron (Fe) - Liners, water pump, cylinder block, cylinder head.
Zinc (Zn) - Brass from components.
Lead (Pb) - Solder in radiator, oil cooler, after cooler, heater core.
Copper (Cu) - Radiator, oil cooler, after collar, heater core.
Aluminum (Al) - Radiator tanks, coolant elbows, piping, spacer plates, thermostat housing.
Magnesium (Mg) - Cast alloys.

CONTAMINANTS
Silicon (Spectro Incorporated) - Dirt.
Magnesium (Mg) - Hard water scaling problem.
Calcium (Ca)- Hard water scaling problem.

ADDITIVES
Potassium (K) - Buffer.
Silicon (Si) -Anti-foaming agent, anti-corrosion for aluminum.
Boron (B) - pH buffer, anti corrosion for ferrous metals.
Molybdenum (Mo) - Anti-cavitation, silicate.
Phosphorus (P) - pH buffer, anti corrosion for ferrous metals.

The RDE/AES technique has been in use for over 50 years for the analysis of a variety of samples. Water samples were routinely analyzed by RDE spectrometers [4]. Low limits of detection were achieved by first concentrating the sample by evaporating most of the water by putting the water sample in an oven.

With the introduction of the ICP/AES technique in the late 1970's, water samples were no longer run on RDE spectrometers because of the significantly lower limits of detection as well as superior precision offered by ICP spectrometers. Concentration of water samples by evaporation was no longer necessary, at least not for routine samples, because the limits of detection were so very much lower.

For lubricating oil and fuel analysis, the RDE technique continued to be a preferred method due to its simplicity of operation and reliability. Sample introduction is simple. Both AA and ICP spectrometers require that the oil sample first be diluted, usually with kerosene, so that the sample can be nebulized to form an aerosol. The RDE technique also has the ability to more efficiently analyze the larger particulate in the used sample. RDE spectrometers lend themselves to

on-site analysis in less than optimum working environments whereas AA or ICP spectrometers required a laboratory environment as well as more highly skilled personnel for their successful operation.

Recently, several RDE spectrometers for commercial customers were calibrated with coolant standards so that these spectrometers could be used not only for oil analysis but also for coolant analysis.

The limits of detection (LOD) that can be achieved on a RDE instrument are not as low as can be achieved using an inductively coupled plasma (ICP) spectrometer, but for coolant analysis low limits of detection are not very important since the results are rounded off to the nearest ppm for reporting purposes and sub-ppm levels are of no practical consequence.. Table 4 compares the LOD's for ICP and RDE spectrometers and shows that the RDE spectrometers are appropriate for coolant samples.

Table 4 - Limits of Detection

	RDE, Oil	ICP, Oil	ICP, Water	RDE, Water
Na	0.04	0.20	0.015	0.03
Mo	0.46	0.08	0.005	0.37
Mg	0.01	0.20	0.0002	0.01
P	3.00	0.50	0.008	3.20
B	0.07	0.01	0.002	0.13
Ca	0.03	0.03	0.0001	0.03
Cu	0.04	0.02	0.002	0.06
Al	0.35	0.02	0.001	0.43
Pb	0.80	0.10	0.02	0.75
Fe	0.23	0.02	0.002	0.21
Si	0.20	0.10	0.006	0.33
Zn	0.10	0.02	0.002	0.18

It may be of interest to note that the LOD's for oil are not substantially different than those obtained for water (or for water/glycol mixtures).

LOD's are not by any means the only measure of spectrometric performance. In the case of coolant analysis the purpose is to detect abnormal circumstances. LOD's have little to do with the quality of coolant analysis furthermore, real measurements are many times LOD so that LOD's much less than 1 ppm are not important.

The ICP technique measures only the most finely divided material, i.e., the material cannot be in the form of large particles. This is because the sample introduction system of an ICP spectrometer

includes a spray chamber, the purpose of which is to remove the largest aerosol droplets generated by the nebulizer. The size of the aerosol droplets which reach the plasma torch where emission takes place are on the order of 1 or 2 μm. The RDE spectrometer, on the other hand, is able to detect somewhat larger particles, up to and beyond 10 μm in size.

Table 5 and Table 6 compare results obtained with an ICP spectrometer to results obtained on an RDE spectrometer on used coolant samples. Table 5 is a comparison of used coolant samples that were relatively clean, (no particulate could be visibly detected). Table 6 is a comparison of used coolant samples where particulate could be visibly detected. The data clearly show that if particles are present the results for the wear metals and contaminants are substantially higher when the sample is run on the RDE spectrometer. It may be concluded from this data that neither instrument is entirely quantitative for these samples because there will typically be some fraction of larger particles that is not completely measured by either an ICP or RDE spectrometer. Nevertheless, the purpose of the analysis is served by indicating which coolant systems are in distress. It may be argued that since the RDE spectrometer is more responsive to large particles it is more capable of indicating abnormal coolant conditions than an ICP spectrometer.

Table 5 - ICP/RDE Comparison of "Clean" Coolant Samples

INST	Al	Cu	Fe	Pb	Zn	Mg	Ca	Si	Mo	P	B	Na
RDE	1.1	0.4	0	0	0	0.5	0	47.6	18.6	590	389	2069
ICP	0	0	0	0	0	0	0	41	19	406	248	1078
RDE	14.8	6	7.8	8.7	2	0	0	37.6	0	122	263	1546
ICP	6.7	0	3.8	3.2	0	0	0	26	0	90.2	217	914
RDE	1.3	0.4	0	0	0	0.1	0	34.8	0.2	511	649	2541
ICP	0	0	0	0	0	0	0	28	0	315	421	1309
RDE	0.5	0	0	0	0	0	0	3.2	13.5	77.5	615	2409
ICP	0	0	0	0	0	0	0	3.7	18	60	480	1283
RDE	0.4	0.7	5.7	7.1	1.3	0.5	0.8	40	58.9	984	635	2835
ICP	0	0	2.7	2.8	0	0	0	26	44	585	345	1385
RDE	0.8	0.7	1	0	0	0	0	3.7	0	445	325	1364
ICP	0	0	0	0	0	0	0	3.6	0	335	213	852
RDE	0.3	0.9	0.6	6	1.9	0.4	12.1	61	1.2	122	840	5972
ICP	0	0	0	2.6	0	0	3.6	43	0	74	515	2650

Table 6 - ICP/RDE Comparison of Coolant Samples Where Particles Were Detected

INST	AL	CU	FE	PB	ZN	MG	CA	SI	MO	P	B	NA
RDE	0.2	39	0	39	8.7	0.9	0	53.6	0	1677	630	5440
ICP	0	12	0	17	2.5	0	0	34	0	1163	321	2733
RDE	11	23	60	2.8	4.3	1.3	0	34.2	2.5	61.8	806	3365
ICP	6	5.6	37	0	0	0	0	23	4	53.7	672	2134
RDE	0.5	47	0	0	0	0.2	0	59	23.3	583	1415	6471
ICP	0	8.7	0	0	0	0	0	30	15	271	748	2552
RDE	2	1	1	4.9	2.8	1.3	5.8	42.4	68.5	326	1088	5013
ICP	0	0	0	2.2	0	0	0	29	58	203	678	2668
RDE	0.9	16	75	25	40	0.9	4.2	12.4	0.2	12.1	1368	5118
ICP	0	2.3	28	7.5	19	0	0	5.2	0	12.5	870	2495
RDE	1.1	1.7	24	2.5	2.5	0.4	0.1	24.5	66.3	709	385	1652
ICP	0	0	7.5	0	0	0	0	19	61	522	232	959
RDE	9.4	12	114	49	6.9	10	66.4	29.4	0.6	948	1424	6567
ICP	0	0	25	8.9	0	2.2	12	12	0	561	746	2904
RDE	0	28	130	70	14	3.7	1.2	39	7	327	322	1671
ICP	0	0	47	18	6.7	0	0	32	8.7	222	211	939
RDE	5.7	44	1.5	3.4	28	27	269	150	260	2110	751	3585
ICP	0	12	0	0	7.6	7.4	44	83	193	1315	334	1564

3.1.2 Coolant Physical Property Monitoring

The second part of an effective coolant analysis program is coolant condition monitoring. Through periodic sampling of the coolant, the laboratory can determine the effectiveness and remaining life of the coolant based on additive degradation and contamination analysis.

ASTM (American Society for Testing and Materials) tests are mostly written for quality control and quality assurance requirements of new and sometimes used coolants. Therefore, ASTM procedures are often modified to reduce analysis times in the interest of economics. The number and type of tests that are performed on a used coolant sample vary. Table 7 below summarizes the physical property tests performed by the typical used coolant analysis laboratory.

Table 7 - Typical Physical Property Tests for Used Coolant Analysis

- pH, ASTM 1287
- Reserve Alkalinity, ASTM D1121

- Percent EG/PG
- Freeze/Boiling Point
- Nitrite, ppm NO2
- SCA Levels
- Total Dissolved Solids (TDS)
- Appearance

pH - measures the acidity or alkalinity of the used coolant sample. This can be measured by performing ASTM D1287, which is very precise. The use of an inexpensive pH meter can provide quick and accurate results in a non-laboratory environment. Most pH meters have the capability to measure the conductivity of the coolant which also can be used to determine the percentage of TDS. Most major engine manufacturers recommend coolant pH levels between 8.5 to 10.5. If pH levels fall below 8.0 rapid nitrite depletion will occur. Coolant pH levels above 11.5 will corrode aluminum and promote scaling[2].

Reserve Alkalinity - measures the amount of alkaline inhibitors present in the used coolant. This gives an indication of the coolant's ability to provide corrosion protection. This test can be performed by ASTM D1121. If the buffering agents are not at correct levels, corrosion and rapid additive depletion will occur due to a reduction in pH values. The result will be cylinder liner pitting.

Percent Antifreeze EG/PG - This test uses a refractometer to quantify the amount of EG/PG in a coolant sample. Most major engine manufacturers recommend coolants composed of 50/50 water/glycol solution to provide satisfactory freeze and boil point protection. An operating range of 40 to 60 % antifreeze is acceptable, however, the use of antifreeze in concentrations over 65% may cause SCA drop-out, water pump seal damage and engine overheating.

Freeze/Boiling Point - Once the percent of EG or PG is determined, the freeze and boiling point can be calculated by using charts provided by the antifreeze manufacture. Freeze point can also be measured by performing ASTM D1177. If the percent antifreeze is not known, an inexpensive hydrometer can be used to measure the density of the coolant and then calculate the freeze point. The amount of freeze protection required should be based on the lowest expected temperature in your region.

Nitrites/SCA Package Analysis - The analysis of the primary corrosion inhibitor, nitrite, can be performed by various tests. The most accurate method is by ion chromatography. For the most accurate results this test must be performed in a laboratory. The

second method is the colormetric analysis with nitrite test sticks. This is a very quick and easy way to measure the concentration of NO_2. Fleetguard offers a test stick that will measure nitrites (NO_2), molybdates (MoO_4) and freeze point protection colormetricly on the same test stick. As with all elements contained in the SCA package, the concentrations must remain within 10% of the new coolant SCA additive level.

Total Dissolved Solids (TDS) - This is a measurement of the dissolved solids in the coolant. The dissolved solids are composed of the basic inhibitor chemicals, silicates, active SCAs, spent SCAs, contaminants and water hardness compounds[2]. The higher the dissolved solids the higher the conductance. The percentage TDS can be quantified with a conductivity meter. This meter will also be capable of measuring the pH of the coolant. Cummins recommends a maximum of 5% TDS, higher levels may cause water pumps seal failure[2].

Appearance - This quick and easy test records the overall condition, color and visible contamination of the coolant. Color is important because most manufacturers identify coolants by color. It is important to notify maintenance personnel if oil or large particles are present in the coolant sample.

4 CONCLUSION

RDE spectrometers, when appropriately calibrated, may be used as part of a condition monitoring program based on coolant analysis. The same instrument as used for used oil analysis can be modified and calibrated to also effectively analyze coolants. The added capability provides the laboratory with an supplementary tool to increase its capabilities and effectiveness,

References

[1] Professional Services Industries, Inc. TAI/Faber Division, Coolant Analysis
[2] Cummins Engine Company, Inc., Service Bulletin #3666
[3] Caterpillar Inc., SOS Homepage World Wide Web
[4] ASTM, E-2 SM 11-18, 1971 "Suggested Method for Spectrochemical Analysis of Water by the Rotating Disk technique Using an Optical Emission Spectrometer", submitted by J.F. Kopp

FAULT DETECTION AND ISOLATION IN PROCESSES INVOLVING INDUCTION MACHINES

Caj Zell, Alexander Medvedev

Control Engineering Group, Luleå University of Technology
S-971 87 Luleå, SWEDEN
Caj.Zell@sm.luth.se

ABSTRACT

A model-based technique for fault detection and isolation in electro-mechanical systems comprising induction machines is introduced. Two coupled state observers, one for the induction machine and another for the mechanical load, are used to detect and recognize fault-specific behaviors (fault signatures) from the real-time measurements of the rotor angular velocity and terminal voltages and currents. Practical applicability of the method is verified in full-scale experiments with a conveyor belt drive at SSAB, Luleå Works.

1 INTRODUCTION

The induction machine has become a popular drive in recent years and has gradually evolved into a viable alternative to the DC-motor in electro-mechanical applications.

In many cases, the induction machines used in industrial installations are over-dimensioned and often run in a semi-functional condition. Furthermore, due to an excessive torque delivered by the induction machine, incipient faults symptoms in the driven mechanical part are less pronounced and can easily be missed by the maintenance personnel. As a result, the induction machine energy consumption rises and the unattended, for a longer time, electrical and mechanical faults might eventually lead to a complete system break-down.

Thus, from the energy-saving and process availability points of view, there is an industrial need for real-time condition monitoring covering not only the induction machine itself, but as

well the specific process in which the induction motor is being used. The latter is hereafter referred to as the "mechanical process". To this end, the model-based Fault Detection and Isolation (FDI) methods can be taken advantage of in order to extract the fault-specific information from the real-time signals available in the electro-mechanical systems.

A natural and appealing idea is to use the induction machine as a "sensor" for the mechanical process, yielding information on the mechanical load through the rotor angular velocity and terminal voltages and currents. Accordingly, it is assumed that abnormal behavior and faults related to the mechanical process will result in discernible deviations in the load-torque acting on the rotor of the induction machine.

In order to obtain information about the health of the mechanical process, a reliable estimate of the load torque could be in place. For this purpose, a load-torque observer is derived based on the mechanical equation governing dynamical behavior of the rotor axis. In this equation, the driving torque acting on the rotor axis is present and has therefore to be evaluated.

The driving torque is a nonlinear function of the stator and rotor magnetic fluxes of the induction motor which can be estimated by the observer techniques. However, the magnetic fluxes are subject to fast fluctuations and therefore high sampling rates have been traditionally required. In this paper, a method of reducing the computational burden of implementing flux observers in general and for induction machines fed directly from the power net in particular, is described.

Due to the presence of rotating parts, a fault occurring in the mechanical process is likely to give rise to a periodic component in the load torque. The load torque is therefore modeled as the sum of the outputs of an integrator and a finite number of marginally stable dynamical systems. This model is incorporated in the load-torque observer.

The fault signatures are established by analyzing different faults and the effects they will have on the states in the load-torque observer. The decision of whether a fault has occurred or not is made by comparing the distances (vector norms) between the current state and the fault signatures and using the threshold technique. The results are illustrated by experiments performed with a 37 kW induction motor in a test rig.

The work described here has been done as a part of the project "Intelligent Alarm Management". The project is a collaboration between SSAB Luleå Works, MEFOS and Luleå University of Technology and is financed by the Swedish Na-

tional Board for Industrial and Technical Development and SSAB.

The Blast Furnace 2 at SSAB, Luleå, is subject to a large number of alarms representing different types of process faults and disturbances. In a specific situation, a fault can result in a massive amount of alarms causing the operator to experience an information overflow where appropriate action is hard to determine. In this case, a suggestion from a monitoring system can be of crucial importance in helping the operator to make the right decisions or providing support for decisions already under consideration. In an attempt to cope with these problems, the research project "Intelligent Alarm Management" has been launched in the spring of 1993.

The type of electro-mechanical systems considered in this paper assumes on the induction machine as electrical drive. As the induction machine is widely used in the industry, this covers a wide area of prospective applications. Still, the results can also fairly easily be generalized to include other drives.

2 PROCESS MODEL

The dynamics of the induction machine is nonlinear and time-invariant. The nonlinearity is due to the coupling of the electrical and mechanical subsystems by the rotor angular velocity and the time-invariance is preserved as long as the variations in the process parameters can be neglected. However, the equations describing the electrical subsystem, when viewed explicitly, with the rotor angular velocity as a process parameter, are linear and time-varying.

The state equations describing the induction motor's dynamics can be formulated as

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + Bu(t) \\ J\frac{d\omega}{dt} &= \tau_e(x) - \tau_l(\omega, t) \end{aligned} \tag{1}$$

where τ_e- electric torque, τ_l - load torque, and

$$x = \begin{pmatrix} \Psi_s \\ \Psi_r \end{pmatrix}, \quad x \in C^2, \quad u \in C, \quad \omega \in R, \quad l = L_s L_r - L_m^2$$

$$A = -\frac{1}{l}\begin{pmatrix} R_s L_r & -R_s L_m \\ -R_r L_m & R_r L_s - j\omega(t) l \end{pmatrix}, \qquad B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

The notation used in the induction machine model is summarized in the Nomenclature section at the end of the paper .

Logically, the induction machine can be viewed as consisting of two interacting parts, the electrical subsystem and the mechanical subsystem (Fig. 1).

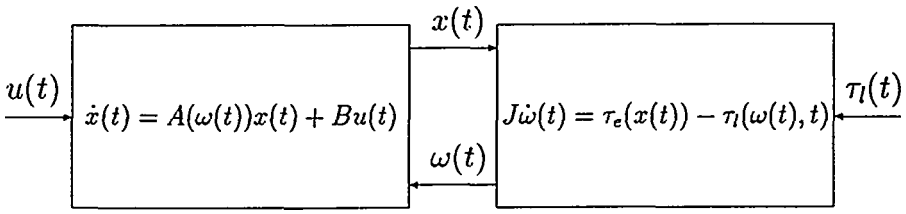


Fig. 1: Decomposition of the induction machine

2.1 The electrical subsystem

In the equations governing the magnetic fluxes (1), the rotor angular velocity ω can be considered to be a time-varying parameter. Since, it becomes possible to treat the electrical subsystem as a time-variant linear system.

$$\begin{aligned} \dot{x}_e(t) &= A(t)x_e(t) + Bu(t) \\ y(t) &= Cx_e(t) \end{aligned} \tag{2}$$

where

$$x_e = \begin{pmatrix} \Psi_s \\ \Psi_r \end{pmatrix} \quad A = -\frac{1}{l}\begin{pmatrix} R_s L_r & -R_s L_m \\ -R_r L_m & R_r L_s - j\omega(t) l \end{pmatrix}$$

$$B = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad C = \frac{1}{l}\begin{pmatrix} L_r \\ -L_m \end{pmatrix}^T \quad y = i_s \quad x \in \mathcal{C}^2 \quad u, y \in \mathcal{C}$$

The continuous-time flux model is apparently linear, time-variant, of second order and with state variables in the complex plane. The system matrix time-dependence poses certain problems when it comes to discretization of the induction machine mathematical model (2).

- The first difficulty concerns derivation of the transition matrix. A closed-form expression is not readily available and therefore alternative methods of discretization have to be considered. One way is to use an approximation, assuming, for example, that the rotor angular velocity is constant between the sampling instants. In this case, the Jordan factorization of the system matrix is given by the following expression

$$A = PA_jP^{-1} \tag{3}$$

where

$$A_j = \begin{pmatrix} \frac{-R_sL_r-R_rL_s+j\omega l+\alpha}{2l} & 0 \\ 0 & \frac{-R_sL_r-R_rL_s+j\omega l-\alpha}{2l} \end{pmatrix}$$

$$P = \begin{pmatrix} \frac{R_rL_s-R_sL_r-j\omega l+\alpha}{2L_mR_r} & \frac{R_rL_s-R_sL_r-j\omega l-\alpha}{2L_mR_r} \\ 1 & 1 \end{pmatrix}$$

$$P^{-1} = \begin{pmatrix} \frac{L_mR_r}{\alpha} & \frac{-R_rL_s+R_sL_r+j\omega l+\alpha}{2\alpha} \\ -\frac{L_mR_r}{\alpha} & -\frac{-R_rL_s+R_sL_r+j\omega l-\alpha}{2\alpha} \end{pmatrix}$$

and

$$\alpha = \sqrt{(R_sL_r - R_rL_s + j\omega l)^2 + 4R_rL_m^2R_s} \tag{4}$$

- The second difficulty concerns the choice of sampling interval. The frequency properties of the system are not easily established which fact implies problems in choosing the sampling frequency. In this work, an induction machine directly fed from the power net $(50 - 60Hz)$ is investigated and experience has shown that sampling time $h = 0.001s$ is a reasonable choice.

2.2 The mechanical subsystem

The mechanical subsystem is governed by the Newton's second law for rotational bodies. This is the second equation in (1) which describes the dynamics of the rotor axis. The electric torque acting on the rotor axis is a nonlinear function of the flux components [1] and is given by

$$\tau_e(x(t)) = -\frac{3L_m}{2L}(\Psi_{sD}\Psi_{rQ} - \Psi_{sQ}\Psi_{rD}) \tag{5}$$

Since almost all parts of the mechanical process participate in one or another kind of rotational movement, the load torque τ_e, can be assumed to be periodic and is therefore modeled, for instance, as the sum of a finite number of weighted sinusoids.

$$\tau_l(t) = a_0 + \sum_{i=1}^{n} a_i sin(\omega_i t) \tag{6}$$

This corresponds to a bank of marginally stable linear autonomous systems where the sum of the outputs yields the load torque signal (6). Now, the dynamics of the induction machine rotor and the mechanical process can be combined in a model for the mechanical system as a whole. The resulting flowchart is depicted in Fig. 2

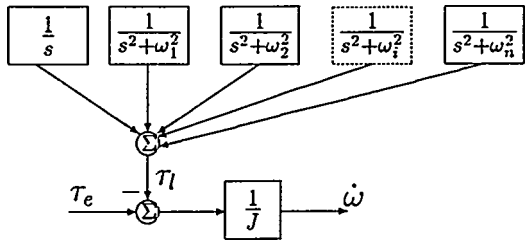


Fig. 2: The mechanical system

The mechanical system can be put in state-space form as

$$\begin{aligned} \dot{x}_m(t) &= Ax_m(t) + Bu(t) \\ y(t) &= Cx_m(t) \end{aligned} \tag{7}$$

where

$$u = \tau_e \qquad y = \omega_r$$

$$A_m = \begin{pmatrix} 0 & -\frac{1}{J} & -\frac{1}{J} & 0 & -\frac{1}{J} & 0 & \cdots & -\frac{1}{J} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -\omega_1^2 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & -\omega_2^2 & 0 & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & -\omega_n^2 & 0 \end{pmatrix}$$

3 OBSERVERS

An observer-based approach is used to solve the fault detection problem. Two observers are used: one to estimate the driving electric torque acting on the rotor axis and another to estimate the load torque.

3.1 The Flux Observer

The flux estimation problem has been investigated in depth in [2], [3]. A short summary is given here.

The discretization of the electrical subsystem is significantly simplified through the Jordan factorization (3). Assuming the rotor velocity to be constant between the sampling instants, the transition matrix can be found as

$$e^{Ah} = Pe^{\Lambda_j h}P^{-1} \tag{8}$$

Now, a discrete-time flux observer can be formulated as

$$\begin{gathered}\hat{x}(t+h) = Pe^{A_j h}P^{-1}x(t)+ \\ \int_0^h Pe^{A_j(h-\tau)}P^{-1}Bu(\tau)d\tau + K(y(t)-\hat{y}(t)) = \\ Pe^{A_j h}P^{-1}x(t)+ \\ PA_j^{-1}(e^{A_j h}-I)P^{-1}Bu(t) + K(y(t)-\hat{y}(t))\end{gathered} \tag{9}$$

Here, K is chosen so that the observer is stable for all constant rotor velocities up to the synchronous velocity. This design approach is practically motivated and suffices in this particular application.

3.1.1 Sinusoidal Input Voltage

In case certain input signal characteristics, such as amplitude and frequency, are known a priori, the observers performance can significantly be enhanced. In practice, many drives utilize sinusoidal inputs of constant amplitude and frequency. In such a case, the input signal can be expressed as

$$u(t) = c\exp(j\phi_s(t)) \tag{10}$$

where c is a real constant, ϕ_s is the stator voltage phase angle, and $j = \sqrt{-1}$. Now, evaluating (9) with $u(t)$ given by (10) one gets,

$$\begin{aligned}\hat{x}(t+h) &= \Phi(h)\hat{x}(t) + \int_0^h \Phi(h-\tau)Bu(\tau)d\tau + K\left(y(t)-\hat{y}(t)\right) \\ &= \Phi(h)\hat{x}(t) + \Gamma_s u(t) + K(y(t)-\hat{y}(t))\end{aligned} \tag{11}$$

where

$$\begin{aligned}\Gamma_s &= P(j\omega_s hI - A_j)^{-1}(\exp(j\omega_s hI) - \exp(A_j h))P^{-1}B \\ \omega_s(t) &= \dot{\phi}_s(t)\end{aligned}$$

In most practical cases, the stator voltage frequency ω_s can be assumed to be known a priori.

Note that the equation above reduces to (9) when $\omega_s = 0$, i.e. for constant input voltage during the sampling intervals. Furthermore, the implementation of (11) does not require any further calculations compared to the case when the assumption of the input voltage being piecewise constant is made. Generalizations to more complex periodic inputs can be made by analyzing the frequency components of the input and performing the calculations for each harmonic.

3.2 The Load Torque Observer

The load torque observer is based on the model (7). Since the system is LTI, the discretization procedure is simple and a discrete-time observer can be designed using pole placement.

4 FDI IN ELECTRO-MECHANICAL SYSTEMS

The proposed method is illustrated on a number of faults occurring in an AC-motor driven conveyor belt. A conveyor belt includes a large number of rotating parts and therefore the mechanically induced fault modes show periodic behavior. Thus, the load torque can be spectrally decomposed and thereupon seen as a sum of sinusoidal contributions of different magnitudes and frequencies. Examples of faults quite likely to occur are slipping in the transmission belts, broken bearings and sliding of the conveyor belt.

The method proposed in this paper is based on the assumption that the magnitudes of the sinusoidal contributions constitute a unique fault signature for each investigated fault.

The load torque observer will contain information on the sought magnitudes in the state estimates. For example, the magnitude of the sinusoid with frequency ω_1, can be computed as

$$a_1 = \sqrt{\hat{x}_3^2 + (\hat{x}_4/\omega_1)^2} \tag{12}$$

where $\hat{x}_3$ and $\hat{x}_4$ are the estimates, produced by the load torque observer, of x_{m3} and x_{m4} in (7) respectively.

The decision on whether a fault has occurred or not is then taken using threshold technique by evaluating a vector norm of the discrepancy between the current estimated magnitudes and the "fault signature magnitudes". An in-depth discussion of relevant vector norms and robustness towards modeling errors etc. is challenging but outside the scope of this work.

The fault signatures can be obtained by fault modeling or simply provoking faults on purpose while logging sensory data.

5 EXPERIMENTS

The experiments have been conducted on a conveyor belt which transports iron ore to the blast furnace. In this process, slipping in the transmission belts is a disturbing problem. From

practical experiments, by artificially provoking slipping, it is found that the resulting load torque can be modeled as the sum of a constant and one sinusoid, with reasonable accuracy. This representation is then incorporated in the load torque observer and an algorithm for monitoring this particular fault can be implemented. An alarm is triggered whenever the norm of the deviation from a fault signature is sufficiently small.

In Fig. 3, the 2-norm of the deviation of the estimated load torque components and the fault signature components is depicted. Slipping of transmission belts is provoked at approximately $t = 20s$. It can be seen that the fault takes approximately two seconds to detect and the this is due to time it takes for the load torque observer to converge.

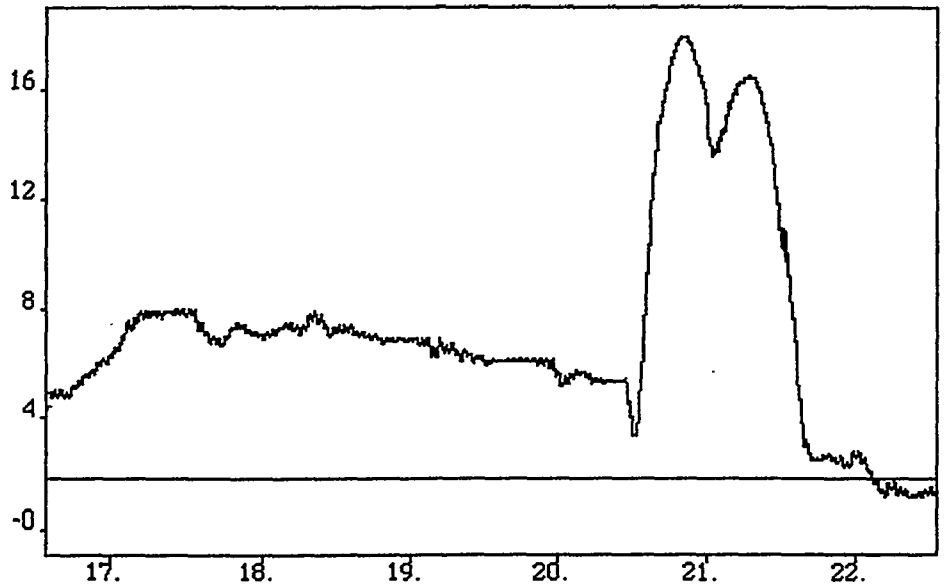


Figure 3: The vector norm (2-norm) of the deviation of the estimated load torque and fault signature components and the threshold (experimental data).

6 CONCLUSIONS

In this paper, a model-based method for detecting and isolating faults in electro-mechanical systems is proposed. Particular attention is paid to the case of an induction motor driving a mechanical process. The method utilizes sensor data from the induction machine to extract information on the operating conditions of the mechanical process.

A time-varying observer is used to reconstruct the magnetic flux in the induction machine, and thereby evaluate the driv-

ing electrical torque. The faults that occur in the mechanical system result in periodic fluctuations in the load torque on the motor's rotor axis. A load torque observer coupled with the flux observer contains a model of the spectral contents of the load torque. The spectral information is compared to the spectral composition related to specific faults, i.e. the fault signatures. The threshold technique is utilized to decide if the deviations are sufficiently small to activate an alarm.

References

[1] P. Vas, *Vector Control of AC Machines.* Oxford Science Publications, 1990.

[2] C. Zell, "Flux estimation in induction machines," licentiate thesis 1996:02l, Luleå Institute of Technology, 1996.

[3] C. Zell and A. Medvedev, "Sampled-data flux estimation in induction machines," in *IEEE Conference on Control Applications*, (Detroit, MI, USA), pp. 1102–1107, 1996.

Nomenclature

Ψ_{sD} - direct–axis stator flux
Ψ_{sQ} - quadrature–axis stator flux
Ψ_{rd} - direct–axis rotor flux
Ψ_{rq} - quadrature–axis rotor flux
i_{sD} - direct–axis stator current
i_{sQ} - quadrature–axis stator current
u_{sD} - direct–axis stator voltage
u_{sQ} - quadrature–axis stator voltage
L_s - self inductance of a stator phase winding
L_r - self inductance of a rotor phase winding
L_m - mutual inductance between rotor and stator
R_s - stator resistance
R_r - rotor resistance
ω - rotor speed

ENHANCED ENVELOPE ANALYSIS OF BEARING SIGNALS USING DIGITAL TECHNIQUES

R.B.Randall, Associate Professor, Y.Gao, Research Fellow, and D.Ho, Research Student
School of Mechanical and Manufacturing Engineering,
The University of New South Wales,
Sydney 2052, Australia

ABSTRACT

Envelope analysis has been shown to be a very powerful technique for detecting and diagnosing faults in rolling element bearings. However, it is most common to use analogue bandpass filters, rectifiers, and smoothing circuits to generate the envelope signal even when the latter is to be analysed in an FFT analyser. This paper shows that all these processes are better carried out digitally in the FFT analyser (or using a digital signal processing package) either by using a real-time zoom processor, or by inverse transforming a spectrum block corresponding to the desired passband. Not only is this more efficient, by eliminating the need for extra hardware, but it gives more flexibility in choosing the demodulation band, and a much sharper filtration. It is also shown that when the signal/noise ratio is greater than unity, considerable enhancement results from analysing the squared envelope signal. The signal/noise ratio can be improved by optimal choice of the demodulation band (where the spectral change is greatest) and in the case of contamination by gear noise, by using self adaptive noise cancellation to remove it. The paper describes the digital processing techniques and illustrates them using simulated and practical examples.

1 INTRODUCTION

Faults in rolling element bearings give rise to a series of high frequency bursts where the information about the fault is contained in the spacing of the bursts rather than in their frequency content, which is usually a mixture of all resonance frequencies excited. Figure 1(a),(b) illustrates that even when the bursts are completely regular, the repetition frequency information is contained in the spacing of the harmonics, but that the actual "ballpass frequencies" and their low harmonics are virtually not present. In the actual situation, in particular with low speed machines, the burst spacing varies randomly to a certain extent, due to varying speed, slip and load angle, and then there is no information to be gained from the raw spectrum (Fig.1(d),(e)). If the envelope of the bearing signal is analysed, however, (by amplitude demodulating the bearing signal), the repetition frequencies are found clearly, even in the case of a typical amount of random fluctuation (Fig.1(c),(f)).

For this reason, envelope analysis has become one of the most powerful techniques used in the diagnostics of rolling element bearings, but in most cases the envelope signal is obtained by adding analogue hardware, even when the envelope signal is to be analysed using an FFT analyser. The envelope signal can be formed by bandpass or highpass filtering the original signal to remove masking signals and enhance the bearing dominated components. This signal is then rectified and smoothed, typically using an RC smoothing circuit. This smoothing operation is unnecessary if the envelope analysis is to be carried out in an FFT analyser, as the antialiasing filters remove high frequency ripple. This paper shows how all these functions can be carried out by digital techniques in an FFT analyser or using a signal analysis package such as Matlab [1].

2 AMPLITUDE DEMODULATION

The signal could of course be demodulated by rectification as in the analogue case, by taking the square root of the square of each value, but this is only efficient if the signal is not to be bandpass filtered. In the latter case there are two ways in which the bandpass filtration and amplitude demodulation can be carried out, both giving a reduction in

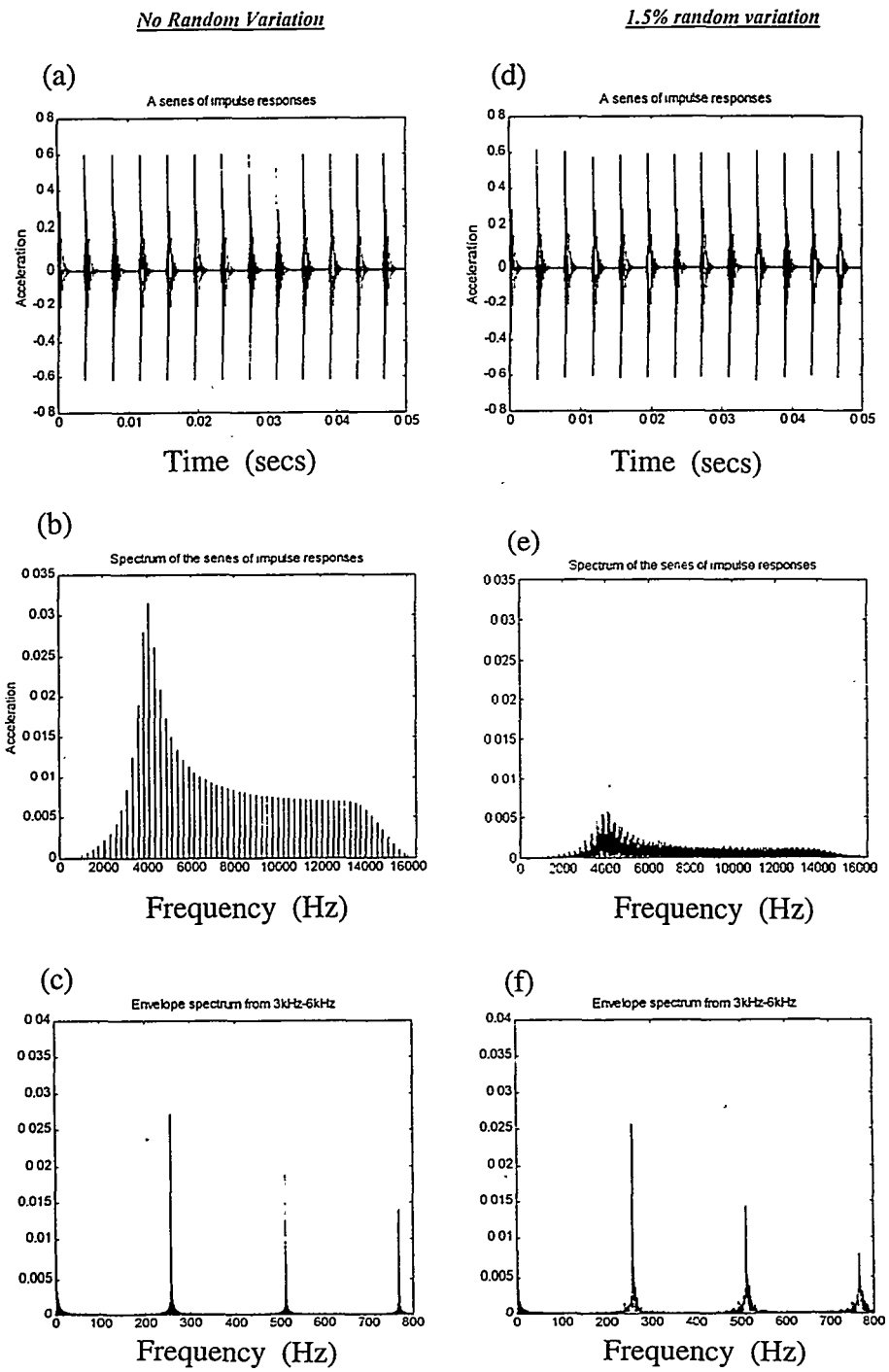


Figure 1. Simulated outer race bearing fault signal (acceleration) with and without random spacing fluctuation.
(a), (b), (c) No random variation. (d), (e), (f) 1.5% random variation.
(a), (d) Time signals. (b), (e) Raw spectra. (c), (f) Envelope spectra.

sampling frequency, and thus of the record size for a given time record. Both make use of the fact that a signal with a one-sided spectrum, a so-called analytic signal, is complex, but its imaginary part is the Hilbert transform of the real part [2]. An analytic signal can also be expressed in the form $A(t)e^{j\phi(t)}$ where $A(t)$ is an amplitude modulation function, and $e^{j\phi(t)}$ is a rotating unit vector with a modulated phase function, representing fluctuations around a carrier frequency in the centre of the band. Thus $\phi(t)$ can be more specifically represented as $2\pi f_c t + \phi_m(t)$, where $\phi_m(t)$ is the phase modulation of the carrier frequency f_c. If the bandpass filtered analytic signal is frequency shifted by an amount f_k, it becomes $A(t)e^{j(\phi(t)-2\pi f_k t)}$, which still has the same amplitude function $A(t)$ and thus the same envelope.

In FFT analysers with a real-time zoom function, the signal is first frequency shifted in the time domain (by multiplication by $e^{-j2\pi f_k t}$), and then lowpass filtered around the new zero frequency corresponding to the zoom band [2]. Figure 2 illustrates how this results in a complex time signal whose magnitude is the required envelope signal of the signal bandpassed in the zoom band.

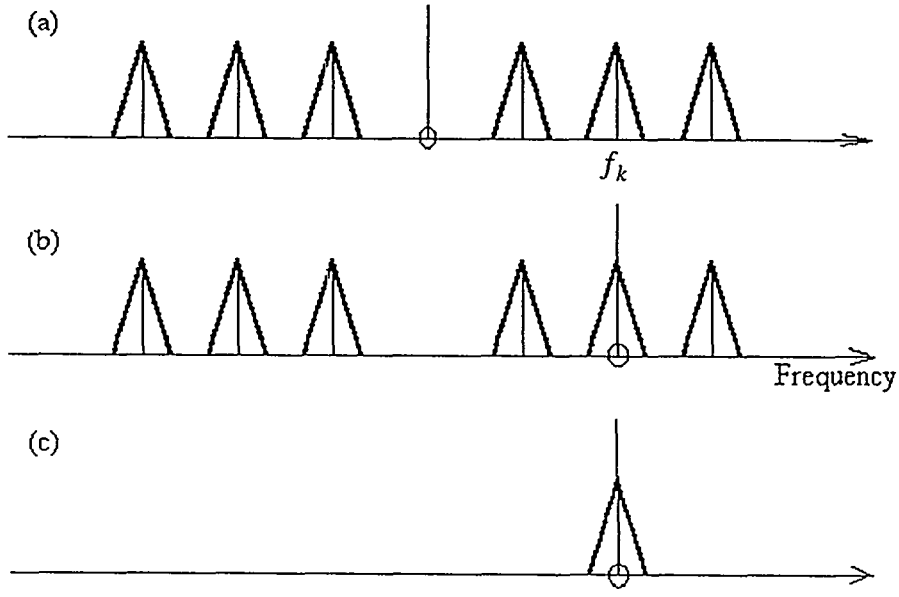


Figure 2. Extraction of a frequency band by a zoom processor
(a) Spectrum of original signal. (b) Spectrum after frequency shift by f_k (multiplication by $e^{-j2\pi f_k t}$). (c) Spectrum after lowpass filtration

Note that the lowpass filtration is done by filters of antialiasing quality (typ. 120 dB/octave) thus allowing bearing signals to be separated from adjacent masking components which can be much larger. This is not normally the case with analogue envelopers. After the lowpass filtration, the sampling rate is reduced accordingly, so that the frequency range of the final envelope analysis is appropriate to the frequency content of the envelope signal.

Where a digital signal processing package is being used, basically the same result can be achieved by a different procedure. An initial large FFT transform is normally required to encompass the high resonance frequencies to be demodulated, at the same time as providing sufficient resolution in the spectrum to resolve the modulation sidebands (ie a length equal to several periods of the modulation frequencies, the lowest typically being cage speed, or 40% of shaft speed). As illustrated in Figure 3, the band to be demodulated is then extracted into a smaller buffer and inverse transformed to a complex time signal. Even though the bandpass filtering and frequency shift have been done in the frequency domain, the finally resulting complex signal and its envelope are very similar to what is produced by the zoom technique. The separation in the frequency domain is perhaps even better, depending on what window function is used for the initial transform, but even the common Hanning weighting gives complete separation over a wide dynamic range in a small number of lines. In Figure 3(b) the situation is illustrated where the centre of the demodulation band is shifted to zero frequency (as for the zoom technique). Though this is necessary for phase demodulation, it has been shown above that the envelope signal is unaffected by the amount of frequency shift, and therefore the procedures of Figure 3(c) & (d) may be used, these being much simpler. It can be shown that forming the (squared) envelope of the complex time signal corresponds in the frequency domain to a convolution of the spectrum band with its complex conjugate. Using FFT procedures, the convolutions are circular, and to avoid any chance of wraparound effects, it is best to zero pad the spectral segment up to double its size, as in Figure 3(d). This means that the transform size and sampling frequency are double what they otherwise would be, though it does not change the resolution of the final envelope analysis. In most cases, however, a resonance peak is demodulated, and the concentration of the spectral information in the centre of the band means that wraparound effects are negligible and the procedure of Figure 3(c) can be used.

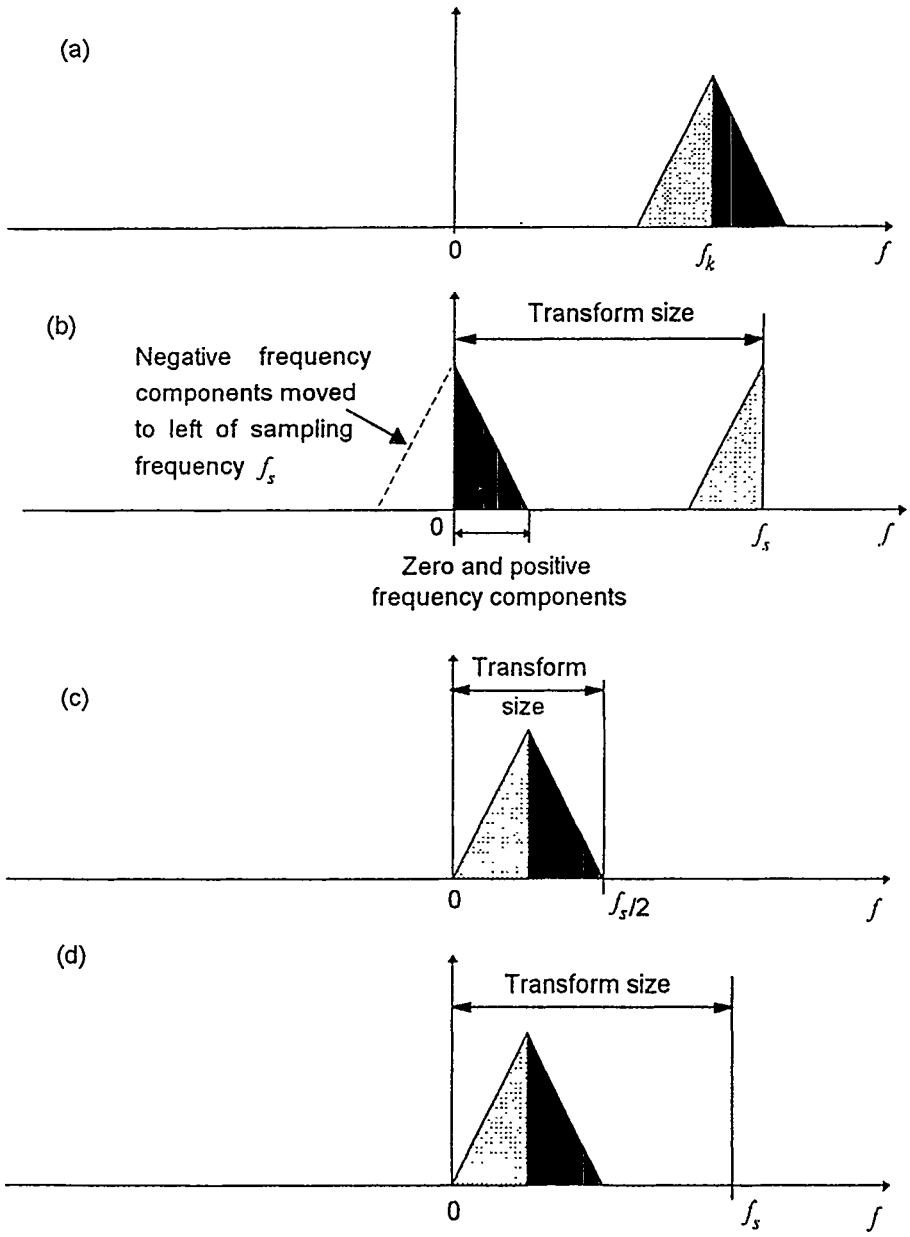


Figure 3. Block shift procedure for selection of frequency band for inverse transformation. (a) Original spectrum of one-sided bandpass section (b) Frequency shift by f_k (centre of passband).
(c) Frequency shift by amount corresponding to lower limit of passband (half size transform).
(d) Frequency shift by amount corresponding to lower limit of passband (full size transform).

3. MASKING

In a study published elsewhere [3], the masking effects of extraneous random and discrete frequency signals were studied in detail. One thing which emerged from this study was that the signal/noise ratio of the squared envelope signal was approximately the square of the signal/noise ratio (bearing/extraneous) of the original signal. Thus, if the latter can be made greater than unity, an advantage will be gained by analysing the squared envelope signal. One way of improving the signal/noise ratio of the signal to be demodulated is to select that frequency band where the greatest spectral change has occurred. This shows the advantages of working digitally, as there is complete freedom in choosing the band to be filtered out and demodulated, and of course it is very simple to square the envelope signal or not according to circumstance.

Where the masking signal is dominated by gear noise in the same band so that it cannot be removed by simple bandpass filtration, it was shown in [4] that use can be made of self adaptive noise cancellation to remove the gear noise and leave a much cleaner bearing signal.

3.1 Self Adaptive Noise Cancellation

Self adaptive noise cancellation is a further development of adaptive noise cancellation, for which it is necessary to have two signals, one (the primary signal) containing two components to be separated, and the other (the reference signal) coherent with one component only. An adaptive filter is caused to act on the reference signal, adapting its linear transfer function in such a way as to make the filter output most similar to the coherent component in the primary signal. The other component can then be obtained by subtraction. Self adaptive noise cancellation can be used where there is only a primary signal available, but where the two components have different characteristics, in particular where one is made up of discrete frequency components and the other random. The degree of randomness in normal bearing signals satisfies this requirement with respect to gear signals, which are locked to shaft speeds. As illustrated in Figure 4, the reference signal can then be made a delayed version of the primary signal, so that the gear component is still correlated, but the random bearing component uncorrelated.

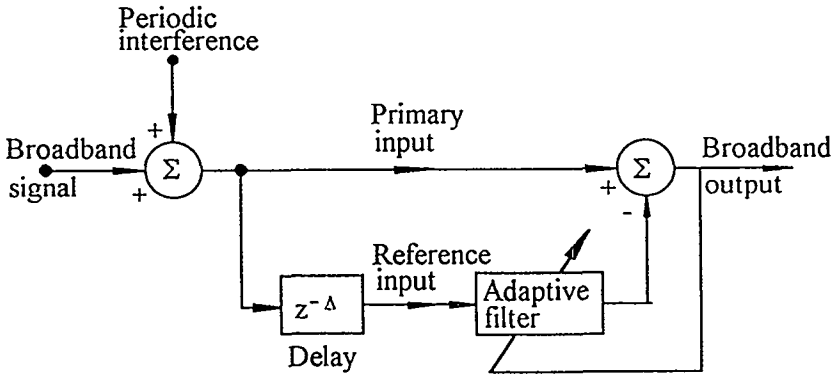


Figure 4. Principle of self adaptive noise cancellation

4. PRACTICAL EXAMPLES

The first example is from a paper mill bearing with an outer race fault. Even though the shaft speed is less than 2 Hz, and BPFO (ballpass frequency, outer race) only 15.4 Hz, the largest spectral change occurred at 5.4 kHz, but appeared completely noise-like. Demodulating this band where the spectral change from the fault was approx. 15 dB gave the results illustrated in Figure 5. Figure 5(a) shows the amplitude spectrum of the envelope signal obtained by the block shift method of Figure 3, and thus representing the best result that could be expected from an analogue envelope device. Figure 5(b) shows the enhancement gained by presenting the amplitude squared spectrum of the squared envelope signal, everything else being the same.

The second example is from a helicopter gearbox, where both inner race and outer race faults were present in a planet bearing. Because of the intimate mixing of the gear and bearing signals from the planet gear (as measured externally on the casing) there was a maximum spectral increase of only 6 dB, in the high frequency range from 13-16 kHz. Bandpass filtration in this range produced a result where both the BPFO and BPFI (ballpass frequency, inner race) could be detected, but were heavily masked by gear components (Figure 6(a)). Applying self adaptive noise cancellation to this band improved the signal/noise ratio such that the envelope signal could be squared,

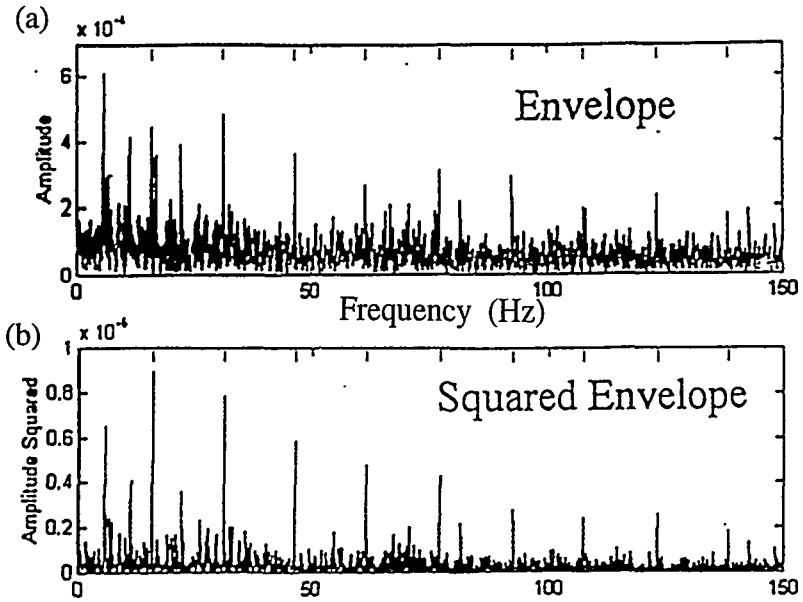


Figure 5. Example of the enhancement given by squaring the envelope signal. Outer race fault in a paper machine. (a) Amplitude spectrum of envelope signal, by demodulation of optimum passband (5.2-5.6 kHz). (b) Amplitude squared spectrum of squared envelope signal.
Tickmarks on upper scale show harmonics of BPFO, enhanced by this process.

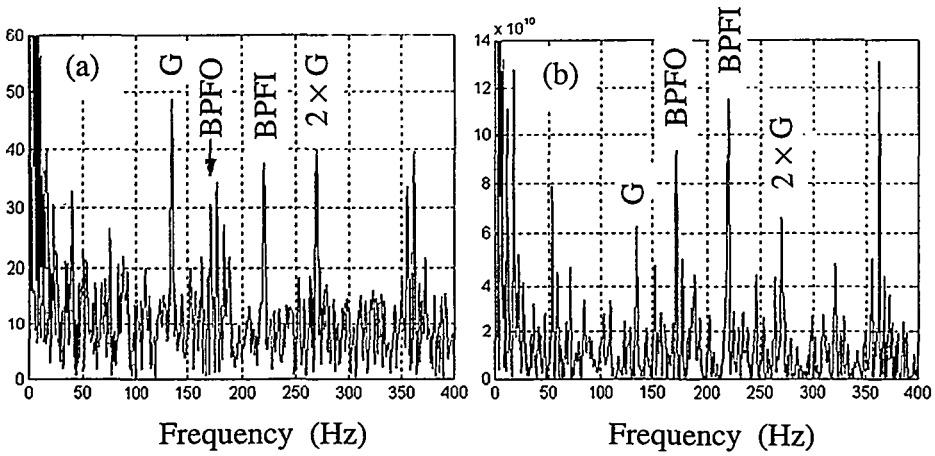


Figure 6. Use of band selection and Self Adaptive Noise Cancellation (SANC) to enhance the envelope analysis of inner and outer race bearing faults in a helicopter gearbox planet bearing.
(a) Envelope analysis after bandpass filtering in the range 13-16 kHz.
(b) Analysis of squared envelope after SANC.
G = gear related frequency, BPFO = ballpass frequency, outer race
BPFI = ballpass frequency, inner race

producing the greatly enhanced result of Figure 6(b), where the bearing components dominate over the gear effects.

CONCLUSION

The paper demonstrates that all functions associated with the generation of envelope signals for bearing diagnostics are more efficiently carried out by digital techniques in an FFT analyser or signal processing package, thus avoiding the need for supplementary hardware. The extra flexibility and more selective filtration associated with the digital processing mean that the signal/noise ratio of bearing to extraneous signal can often be made greater than unity, whereupon considerable enhancement is given by squaring the envelope signal, a function which is also easily carried out digitally. Where the masking signal is from gears, or other discrete frequency components locked to shaft speed, use can additionally be made of self adaptive noise cancellation to further improve the signal/noise ratio.

REFERENCES

[1] Matlab (Windows version 4.2b), The MathWorks Inc., Natick, MA, USA, 1994.
[2] R.B.Randall, (1987), *Frequency Analysis.* Bruel & Kjaer, Copenhagen.
[3] R.B.Randall and Y.Gao, "Masking effects in digital envelope analysis of faulty bearing signals", Vibrations in Rotating Machines, I.Mech.E., Oxford, September, 1996, pp 351-359.
[4] R.B. Randall and Y.J. Li, "Diagnostics of planetary gear bearings in the presence of gear vibrations". 2nd Int. Conf. on Gearbox Noise, Vib. and Diag., I.Mech.E., London, Nov. 1995, pp 73-80.

NEXT GENERATION SMART ACCELEROMETERS AND INTERFACE STANDARDS FOR THE CONDITION MONITORING INDUSTRY

Thurston L. Brooks, Director of Engineering
Brett D. Gidge, Electronic Design Engineer
Wilcoxon Research, Inc.
Gaithersburg, MD 20878, USA

ABSTRACT

Piezoelectric accelerometers are the primary sensor technology used in condition monitoring systems. These sensors have evolved from the first generation "charge mode" types of the 1960's, to the second generation internally amplified designs widely used today. The introduction of "smart sensors" began with third-generation vibration transducers. These smart sensors use mixed-mode analog and digital to communicate with the condition monitoring equipment. Efforts to create a mixed-mode interface standard, IEEE P1451.3, are underway. The standard will provide a plug-and-play connection for standardizing transducer module formats and protocols. The emerging fusion of condition monitoring sensors and modern control-bus systems will require the further development of smarter, processor-based, fourth-generation sensors. These sensors will be characterized by full digital transmission with bi-directional command and data communication capabilities.

1 INTRODUCTION

Piezoelectric accelerometers are the most common vibration sensor technology used in condition monitoring systems. These sensors have evolved from the first-generation, unamplified "charge mode" sensors of the 1960's, to the second-generation, internally-amplified designs widely used today. With the advent of application specific integrated circuits (ASIC) and hybrid technology, third-generation vibration

sensors began to appear on the scene in the late 1980's.[1] Third-generation devices use mixed-mode analog and digital transmission through the use of on-board electronics. These two-wire, mixed-mode systems contain digital data sheets that provide specifics of the sensor, as well as, analog functions that include the raw vibration signal, calibrated verification signal, and a health signal.

Development of fourth-generation vibration sensors, those with full digital communications and on-board intelligence, has not happened as quickly as many had envisioned. The development of smart sensors for condition monitoring applications has lagged behind the development of smart pressure, temperature, flow, and other sensory modalities primarily due to the shear magnitude of the data to be processed and transmitted. Consider that existing smart-sensor network technology such as HART, FieldBus, Profibus, DeviceNet, CAN, etc. all have maximum data rates far below that necessary for vibration monitoring (FieldBus H2 and Profibus-DP being recent new exceptions). For raw vibration data to be transmitted over a digital bus the required rate is a minimum of 320k buad.[2] It is interesting to note that none of the current sensor network/bus standards, including the FieldBus Foundation standard, has defined data formats for vibration. A simple fact is that a single vibration sensor would bog down most existing bus standards, much less the hundreds, and in some cases thousands, of sensors found in typical condition based maintenance (CBM) monitoring applications. Boeing Aircraft has estimated their needs for a vibration sensor bus at a minimum 1M bits per second (bps) [1]. They conclude that adding in command structure, overhead, and error correction would allow only 2 sensors to share a 1Mbps bus. Interestingly, even with these handicaps, Boeing believes the advantages of smart sensors far outweigh the existing centralized, analog sensors [1].

This paper highlights the developments that have occurred, or are anticipated to occur, in smart sensors intended specifically for vibration

---

[1]Third-generation smart accelerometers are already in use in a few military applications (e.g., the Wilcoxon FN8 developed in 1988).

[2]Desired bandwidth 1-10kHz with 16bits A/D and Nyquist sampling at 2 times (required for FFT calculation). Note: Bus capacity is usually calculated assuming 16 bits for the A/D conversion, but it is interesting that piezoelectric accelerometers typically have 140-160dB of dynamic range, thus technically one would require an A/D conversion of 23-27 bits to capture its full resolution.

monitoring. Although many of the issues discussed here apply to other sensory modes (acoustics, pressure, temperature, etc.) this paper primarily deals with the unique aspects of smart vibration sensors.

Table 1. Why Smart Sensors?

• Simplifies change-out and replacement of transducers through in situ normalizing of sensor parameters
• A/D conversion at the sensor will result in higher-quality, EMI immune data
• Distributed intelligence allows specialized and/or user-defined processing by location
• Total system costs will decrease thru the use of low-cost cabling and reduced interface complexity (no more signal conditioners, high cost cables, A/D conversion systems with differential instrumentation amps, etc.)
• The centralized-processor's decision, analysis and control will focus on the system-level and optimization will be enterprise-wide
• Localized processing can identify regional problems faster than current systems using round-robin poling of sensors

2 THIRD-GENERATION SMART VIBRATION SENSORS

Third-generation vibration transducers introduced low-level mixed-mode smart sensors capable of simple unidirectional communication. The third generation sensors have the ability to transmit manufacturer and user-defined information which is preprogrammed into its memory during fabrication. Some designs perform a health validation by transmitting the data through the sensing element and the analog electronics. Once the data transmission from memory is complete, the sensor immediately returns to a second-generation mode of operation where it continues to output an analog signal proportional to the vibration input. Miniaturization, along with increased performance and reliability, of IC (integrated circuit) chip design had to precede the development of a mixed mode smart vibration transducer. Also of importance was the need to keep a two wire system platform in order to mate with the existing hardware interfaces.

An actual smart technology application: Wilcoxon Research (WR) first demonstrated third-generation technology in a military application in the late 80's. To fit inside a hermetically sealed TO-8 package, the mixed mode electronics had to be implemented in an ASIC (application specific integrated circuit) design as shown in Figure 1.

The digital electronics, using a current detecting operational amplifier scheme, were triggered by a 2mA drop in the current source for 11ms.

Programmable read only memory (PROM) chips were used to store an autotest sequence and sensor identification code. Figure 2 shows the digital output sequence for the sensor used in this application.

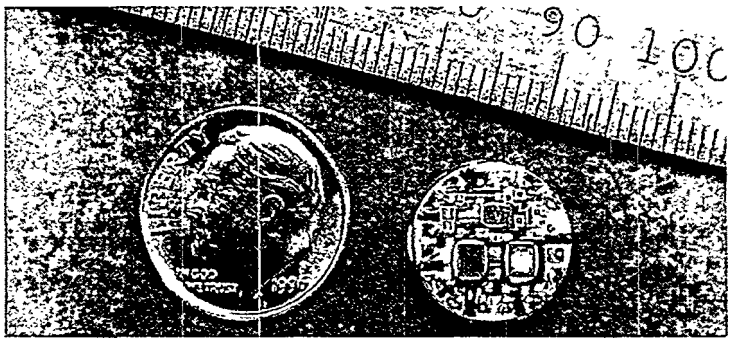


Figure 1. The FN8 Smart Electronics.

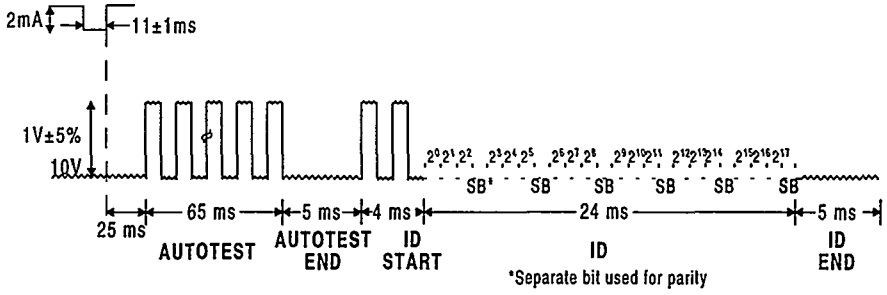


Figure 2. Output Sequence

The autotest, consisting of a 65ms string of 0's and 1's, was primarily used by the military to verify operation of the piezoelectric sensing element. The application only required the digital output of the sensor identification code but more data could have been programmed. Later development of the WR Model 200 IQ™ included the sensor type, manufacturer, serial number and standard sensitivity. A block diagram of this smart industrial vibration sensor is illustrated in Figure 3.

Standards of Tomorrow -- IEEE-P1451.3: Due to the effort involved in interfacing smart sensors to measurement and control systems, the lack of a standard could expose a great deal of risk to both transducer manufacturers and customers. NIST (National Institute of Standards and Technologies) is leading an effort to develop an IEEE interface standard, P1451.3, for mixed-mode smart transducers. The proposed standard will cover the transducer module which includes the transducer electronic data sheet (TEDS), the communication protocol, and the interface module.

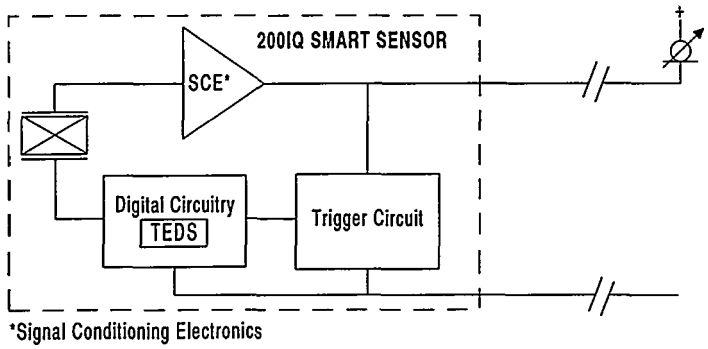


Figure 3. Block Diagram of 200IQ Smart Sensor

P1451.3 Purpose: The goal is to simplify the development of the smart transducer interface to measurement and control systems. The standard will provide a plug-and-play type connection setup for versatility and standardize transducer module formats and protocols.

P1451.3 Objectives: Objectives of P1451.3 include the following: (1) Create a mixed mode smart transducer interface standard which will be compatible with the future P1451 network capable smart transducer interface and based on existing analog interfaces; (2) Create a communication protocol for the digital electronics through existing analog connections; (3) Specify TEDS formats; and (4) Provide a technology to incorporate P1451.2 and P1451.3 capabilities inside a single transducer package.

P1451.3 TEDS Format: The TEDS simply provides a way of describing the transducer to the control equipment. The communication protocol which will be defined will provide the means for accessing the TEDS. Table 2 highlights the proposed categories for the TEDS format.

Table 2. TEDS Categories

Identification Parameters	Device Parameters	Calibration Parameters	Application Parameters
• Manufacturer name	• Sensor type	• Last date calibrated	• Channel ID
• Model number	• Sensitivity	• Cal curve fit function	• Channel grouping
• Serial number	• Bandwidth	• Cal transfer function	• Location/orientation
• Revision number	• Units		
• Date code	• Accuracy		

P1451.3 Communication Protocol: Several communication protocols for triggering the digital circuitry have been discussed as potential mixed-mode schemes. The temporary dropping of the power supply current

has already been implemented with current sensing amplifiers. The current drop could also enable a local IQ chip in the sensor. The chip would then disconnect the analog electronics through an analog switch and then wait to receive a digital code that would contain task instructions. This protocol technique would require a sensor with slightly more sophistication as shown in Figure 4.

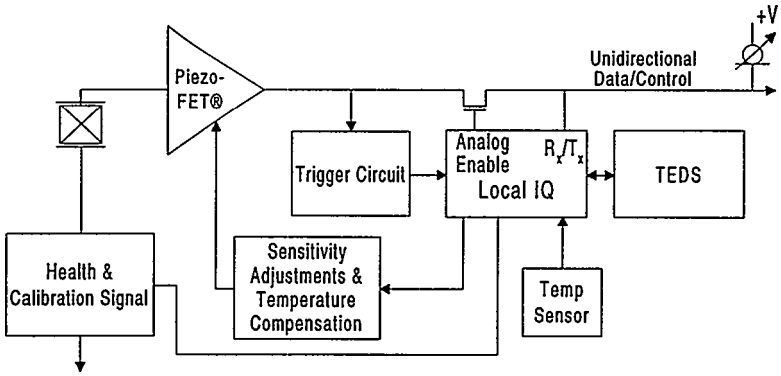


Figure 4. Advanced Third-Generation Schematic

As illustrated, the use of a centralized control chip would allow for additional features in the transducer. Digitally controlled gain adjustments have already been successfully applied in other sensor applications, most notably in a smart pressure transducer [5].

Other concepts for implementing mixed-mode smart sensors have been proposed. One such method would use power supply polarity to enable/disable the analog & digital electronics [2]. Under "normal" operation, the connection of the power supply would only forward bias the diode integrated in the analog circuitry. Reversing the transducer leads would then forward bias the other diode and enable the digital electronics while simultaneously blocking the analog output. Additional protocols for communicating with the sensor are currently being investigated.

P1451.3 Interfacing: The proposed IEEE standard hardware interfacing will be based on current analog interfaces such as those for voltage modulation and 4-20mA current loop.

3 FOURTH-GENERATION SMART VIBRATION SENSORS

Fourth-generation smart vibration transducers will be characterized by a number of attributes: 1) bi-directional command and data

communication, 2) all digital transmission, 3) local digital processing, 4) preprogrammed decision algorithms, 5) user-defined algorithms, 6) internal self-verification/ diagnosis, 7) compensation algorithms, and 8) onboard data/command storage. Figure 5 highlights our view of a fourth-generation "smart" sensor for vibration monitoring.

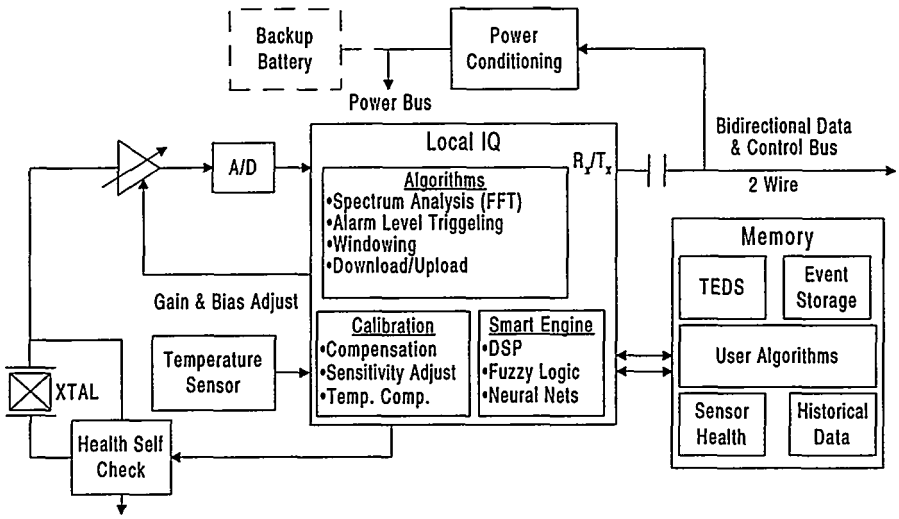


Figure 5. Fourth-Generation Smart Sensors

Bi-directional command and data communication: In contrast to third-generation smart sensors which have unidirectional control and data communication, fourth-generation smart sensor capabilities will allow the sensor to send control commands to the decision support processor, as well as, accept commands. In addition, data flow will be bi-directional, the user can download to the sensor, as well as, upload from the sensor. Hence, a particular mounting point can maintain location specific data, even when the sensor is replaced, by downloading the old sensor's site specific data before replacing.

All digital communication network: Another feature of a fourth-generation smart sensor is that all communications are performed digitally. One particular benefit is error immune transmission that results from techniques such as parity, CRCs, or check sums with retransmission of missing or corrupted data. Hence, EMI concerns are greatly reduced. Cable runs using regeneration techniques such as repeaters will enable transmission over extremely long distances without corruption. Boeing Aircraft reports that 300 meter sensor cable runs are required for 777 testing, resulting in the need for expensive signal conditioning equipment [1]. Fourth-generation smart vibration transducer networks are expected to use 2-wire interfaces and a daisy-

chain topology. This topology minimizes cabling cost per unit length and simultaneously minimizes total cable usage (length) in a given application. Two-wire networks have been identified by a number of user groups (e.g., NASA [3], IEEE P1451.3 [4], Boeing [1], etc.) as the desired solution for sensor nets.

Local Digital Processing: It has only been recently that significant processing power has become available for a low cost. For example, 50Mips DSPs are now available for $8 (US). Combine this with low cost sigma-delta A/D converters ($5 US) and you have a revolution about to take off. Does this mean that centralized CBM processors will disappear, and all processing will occur at the smart sensors? The answer is unequivocally, no. The processing power of distributed sensors will actually enhance CBM capabilities. With hundreds of individual smart sensor DSPs each calculating their own FFT, higher order FFTs could be calculated in the same time that current systems take to calculate one FFT. This would lead to more powerful and sophisticated algorithms involving phase and complete vibrational state analysis of machinery vibration. Subtle changes in machine state that currently go unnoticed will be recognized as significant indicators of machinery health. This higher order analysis can only be performed by a central processor that integrates all of the sensor states into a cohesive whole.

Preprogrammed decision algorithms: The algorithms that can be embedded in a smart transducer vary from the simplistic to the highly sophisticated. Alarm level triggering based on absolute levels would be an example of simple decision making. More sophisticated types of alarm level triggering are priority levels, delta change, windowing, band-alarming, etc. Even more sophisticated concepts such as neural nets and fuzzy logic could be used within the sensor to aid in localized decision making. Historical data comparisons such as trending of data could be easily performed by an intelligent sensor, as well. Interestingly, the storage requirements for trending are minimal, since spectral data is a very compact representation of considerable real-time data.

User-defined algorithm capabilities: This level of functionality would allow each sensor's computational power to be tailored to the specific needs of the customer. For example, after an accelerometer has been in place for a few months, the user may decide that the sensor's amplitude range is too low during machine startup and shutdown (resulting in distortion), but perfect for normal operation. The sensor could be commanded to lower the gain during startup/shutdown and then increased as a function of machine stability and speed for maximum

resolution during normal operation. The concept of extensible sensor object models would allow local smart sensors to be reconfigured for new tasks as required (see also "Onboard Data/Command Storage").

Internal self-verification/diagnosis: Sensor data will also become more reliable in fourth-generation sensors through the ability to constantly monitor its own health. These capabilities can be built into both software and hardware to ensure sensor integrity. Instances occur where CBM systems are unaware that a sensor has failed because of a faulty sensor mimicking a healthy machine. In addition to self-verification, another useful smart sensor function would be self-diagnostic capability. Once an error has been detected, the ability to diagnose the problem and localize the fault can ensure that the problem is fixed quickly. Also, when a problem is suspected by the user, the capacity to command all sensors to verify and diagnose, can locate hidden problems.

Compensation algorithms: A smart sensor can monitor temperature, age, signal amplitude, etc. and compensate directly for local conditions. For example, piezoelectric crystal sensitivity changes with age. Smart sensors could automatically compensate for this drift, saving any costs for recalibration. Another compensation algorithm, direct compensation of sensor nonlinearities (i.e., calibration), could be implemented through the use of lookup tables to linearize the output to a high degree of accuracy. All instrumentation systems are affected by temperature, but these effects can be readily removed by a smart sensor before the data is even processed. Yet another compensation technique involves rescaling of the input amplitude to the amplifier to prevent washover distortion from aliasing the data.

Onboard data/command storage: One feature of having onboard storage in the sensor is the ability to implement lookup tables to adjust and/or compensate for sensor environmental deviations. For example, if once every fifteen seconds a large transient occurs due to another machine's operation, the sensor can create a lookup table that compensates for the transient deviation, thus avoiding false triggers. There are other important advantages of having onboard storage. Consider that most CBM systems are typically set by the users to round-robin pole the sensors once a day, with once-an-hour poling being the exception rather than the rule. This means that if the random or unexpected events occur, the likelihood of catching an event is small. Dedicated sensor processors would allow the CBM Manager to record all significant events for subsequent analysis. This form of event storage would be akin to an airplane "blackbox" that could be interrogated after an unexpected

accident. Another feature of onboard data/command storage is that it enables download and upload capability. This means that the sensor represents an object to the CBM system, an object that has all of the associated benefits of object oriented programming -- reuse and portability, type casting, information hiding, specification (and re-specification) of allowed operations and domain values, and machine/application independencies.

4 CONCLUSION

The realization and implementation of fourth-generation CBM sensors will be ultimately decided by the market--customers will decide based on cost, size, interface utility, functionality, and most importantly benefit. As processing and decision support are incorporated into the sensor package (at low-cost through the use of ASICs) and if the data can be accessed in real time without simplification, fourth-generation CBM smart sensors will become a reality.

REFERENCES

1. Eccles, L., "A Smart Sensor Bus for Data Acquisition", Presented at the NIST IEEE P1451.3 Meeting, March 14, 1997.
2. Brooks, T. & Gidge, B., "Smart Industrial Piezoelectric Accels", IEEE NIST Workshop on Communication Interface for High-Speed Mixed-Mode Analog Smart Transducers, June, 1996.
3. Brooks, T. & Wilcox, B., "Summary of Workshop on the Application of VLSI for Robotic Sensing", NASA Workshop held Mar., 1983, Proceedings of the Computers in Engineering 1984, American Society of Mechanical Engineers, ed. W. A. Gruver, 1984.
4. Chen, S. & Kang, L., "Proposed IEEE P1451.3 Mixed-Mode Smart Transducer Interface," Sensors Expo, Philadelphia, Oct., 1996.
5. Bryzek, J., "The Evolution of Smart Sensor and Transducer Design," Sensors, Vol. 10, No. 10, Oct., 1993.

THE EFFECTS OF SYSTEM CHARACTERISTICS ON THE DIAGNOSIS OF BEARING DEFECTS

Wenyi Wang, Lecturer
Gippsland School of Engineering, Monash University, VIC 3842, Australia
Bruce Kuhnell, Professor
Mechanical Engineering, Monash University, Clayton, VIC 3168, Australia
Michael Harrap, Lecturer
Mechanical Engineering, Australian Defence Force Academy, ACT 2600, Australia

ABSTRACT

The system characteristics of a bearing structure, such as system damping factor and bearing load distribution, play a significant role in the defect diagnosis of the rolling element bearings. In this paper, the effects of these characteristics on various bearing diagnostic techniques are discussed. In particular, the envelope autocorrelation and envelope spectral techniques are compared in different circumstances. It is found that the envelope autocorrelation is superior to the envelope spectrum for cases which involve a sharp load distribution and heavy damping (only qualitative measures for this paper). The envelope autocorrelation and spectral techniques are comparable for moderate damping. For extremely light damping conditions, the envelope cepstrum is also examined. Both synthetic and experimental analysis results are presented.

1 INTRODUCTION

Defect diagnosis of rolling element bearings is often made possible by identifying the repetitive nature of the vibration impulse train produced by faulty bearings. The amplitude and the transient character of each impulse is heavily influenced by the dynamic characteristics of the bearing structure, such as system damping factor, vibration propagation path and the bearing's load distribution.

It is readily shown [4] that the damping factor and the vibration propagation path influence the duration of attenuation of the impulses

generated by a bearing defect. Sharp impulses (heavy damping) in the time domain correspond to widely spread harmonic components in the frequency domain. Consequently, in cases where this duration is comparable to the repetition period of the defect impulse train, it will be difficult to diagnose the bearing defect using the envelope spectral technique. When the defect is located in the inner race or one of the rolling elements of the bearing, the characteristics of the load zone is of great importance in the defect detection. The shape of the load zone is likely to influence the bandwidth of the modulation waveform [1]. A sharp load distribution tends to produce slowly attenuated modulation sidebands (a large modulation bandwidth) in the envelope spectrum.

In this paper, the effects of these bearing system features on the fault diagnosis of bearings are investigated. In particular, the envelope autocorrelation and envelope spectral techniques are compared in different circumstances. Both synthetic and experimental analysis results show that, for conditions where the loading zone is sharp and the damping is light or heavy in the measuring direction, the envelope autocorrelation technique is superior to the envelope spectral technique. These two techniques are comparable for moderate damping conditions. The envelope cepstrum also proves to be an appropriate tool for the extreme light damping condition.

2 EXPRESSION OF THE SYSTEM CHARACTERISTICS IN A BEARING SIGNAL MODEL

In this section, system damping and loading characteristics are expressed in a bearing signal model. The vibration signal generated by a bearing fault can be described by combining Mcfadden's [1] and Braun's [2] models. This signal plus other vibration sources and noise will form the measured vibration signal. It is as follows

$$x(t) = \left\{ m(t) \sum_k \exp[-(t-kT)/\alpha] \cdot \sin[2\pi f^*(t-kT)] \cdot U(t-kT) \right\} \otimes h(t) + n(t) \tag{1}$$

where $\otimes$ denotes the convolution operation, T is the characteristic defect period (ie. the reciprocal of the defect frequency $1/f$), and f^* the dominant structure resonant frequency exited by the bearing defect. α denotes the time constant for the exponential decay of the resonant oscillations, which is determined by system damping. $U(t)$ is a unit step function and $n(t)$ the vibration produced by other machine components

(narrow band) plus broadband noise. $m(t)$ represents the amplitude modulating function [1,3] determined by the defect location. $m(t)$ is uniform for outer race defects, and has a waveform similar to a half-wave sinusoid pulse train with the shaft rotation period for inner race defects and cage rotation period for roller/ball defects respectively. In practice, the actual measured signal will convolve the impulse response $h(t)$ of the vibration propagation path (dependent on the machine structure) with the bearing fault induced vibration.

In the above equation, $n(t)$ may be significantly attenuated using the high frequency resonance technique at the chosen resonant frequency f^*. Therefore, the envelope signal will be mainly determined by the defect period T, the time constant α, propagation influence $h(t)$ and the amplitude modulation waveform $m(t)$. The effects of α and $m(t)$ on the spectrum of $x(t)$ will be substantial. For simplicity, it is assumed in the following discussion, that the mechanical systems have a unity gain propagation path, ie. $h(t) = \delta(t)$.

3 ANALYSIS OF SYSTEM DAMPING EFFECTS ON BEARING DIAGNOSIS

In this section, we will focus on the effects of the time constant α on the diagnosis of bearing faults. For a bearing system, α is basically determined by the system structural damping factor. When there exists a bearing fault, heavy damping means a small time constant α, which leads to sharp fault impulses in the time domain. The spectrum representation of such an exponentially decayed impulse is broadband, the bandwidth B is normally determined by α. The repetition of the individual impulses makes it possible to distinguish the defect frequency f in the frequency domain. In contrast, light damping produces a slowly attenuating impulsive series and thus a narrow band spectrum. Figure 1 serves to illustrate these points.

With the same defect frequency f, the broader the impulse band the more harmonic components. Generally, with $B >> f$, the bigger the time constant α (the longer pulse attenuation duration) the higher the signal energy (the easier the fault detection). In the case where the defect frequency f is high and comparable to the impulse bandwidth B, the defect frequency components will be truncated by the bandwidth B.

However, in cases where α is very big, the tails of the previous impulse

may overlap the next impulse. The interaction of these impulses will blur the defect frequency components in the envelope spectrum. In this circumstance, the envelope spectrum may only reveal the modulation frequency for the inner race or ball fault case. The autocorrelation function will show the corresponding defect period clearly. This is due to the truncation of the high frequency harmonics caused by the narrow impulse bandwidth. In the other extreme of the system damping, very small α makes the bearing impulses quickly attenuate. Therefore, the corresponding spectrum will be broadband and exhibit many harmonics of the defect frequency. The autocorrelation function however will have sharp lag impulses indicating the defect period.

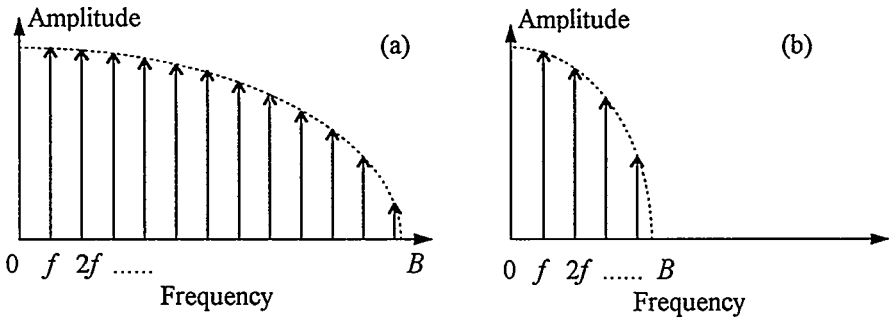


Figure 1. Schematic spectrum representation of exponentially decayed impulse series (the dashed line represents the spectrum of a single impulse). **(a)** Heavy damping (small α); **(b)** Light damping (large α).

4 ANALYSIS OF LOAD ZONE EFFECT ON THE DIAGNOSIS OF BEARING DEFECTS

According to the Stribeck equation [1], the shape of a bearing's loading zone is mainly determined by the distribution angle, the load intensity and the load distribution factor ε. They are functions of loading, the fit between the bearing and its housing and the type of bearings. When the outer race is fixed, the load distribution of the bearing mainly affects the modulation function $m(t)$ and thus the sideband attenuation in the spectrum. A sharp load distribution tends to produce slowly attenuated sidebands around each harmonic component. Figure 2 shows the schematic spectral representations with different load distributions. When heavy damping and sharp load distribution coexist and the defect frequency is smaller than the modulation bandwidth, as shown in Fig.2b, it is very difficult to use the spectral technique. This is due to the great number of spectral lines and the intertwinement between sidebands of various harmonic components. In this case, the autocorrelation technique will show an advantage over the spectral technique.

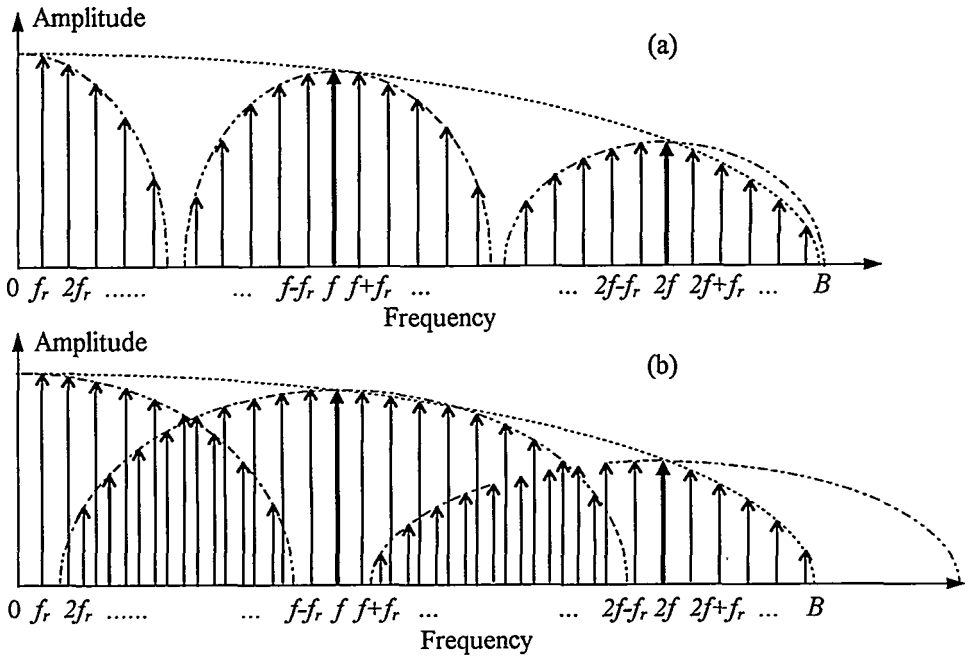
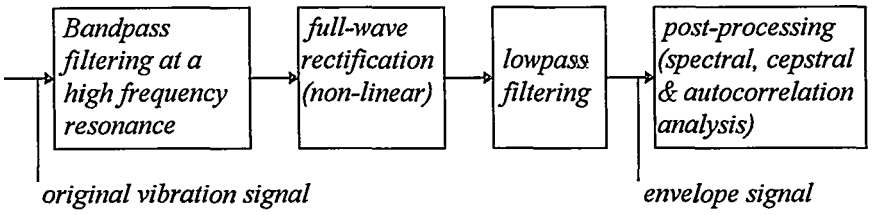


Figure 2. Schematic spectrum representation indicating the effect of load distribution, where f_r is the modulation (shaft or cage) frequency, f the defect frequency (the double dashed chain line represents the envelope of the load distribution spectrum) and B, the same for (a) and (b), the impulse bandwidth. **(a)** Flat load distribution (large ε); **(b)** Sharp load distribution (small ε).

5 ANALYSIS RESULTS USING SYNTHETIC DATA

The synthetic signals used in this section are based on the model in Equation (1). The following diagram shows the analysis procedure.



For the signal used to produce the results shown in Figure 3, the inner race defect frequency is 242.72Hz, shaft rotation frequency is 29.15Hz, and a moderate system damping factor (α=25) a normal load distribution factor (ε=0.5) are used. Comparing the envelope spectrum (Fig.3b) and envelope autocorrelation (Fig.3c) shows that the spectrum presents a better display in identifying both the fundamental defect frequency and the cage frequency (modulation sidebands). Notice that each spike

in the autocorrelation function corresponds to an integer multiple of the inner race defect period (4.12ms). Usually a bearing doesn't have an integer multiple relationship between the defect frequency and the modulation frequency (asynchronisation). Consequently, the modulation period is not directly shown in the autocorrelation [3].

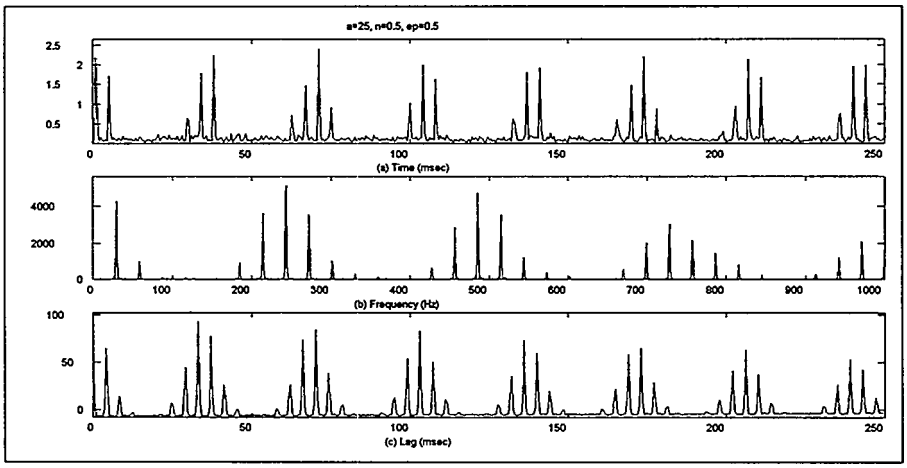


Figure 3. The envelope analysis for a synthetic vibration signal induced by an inner race fault (load distribution factor [1] $\varepsilon = 0.5$, time constant α=25). **(a)** The time signal; **(b)** Envelope spectrum; **(c)** Envelope autocorrelation function. f_i= 242.72Hz, (T_i= 4.12ms), f_r= 29.15Hz (T_r= 34.31ms).

In the case where the impulses attenuate more slowly (ie. larger α), the corresponding spectral band tends to be narrower and fewer harmonic components are shown in the spectrum [4]. In the autocorrelation function, the lag impulses become widened. As the α increases to 250, the attenuation duration of impulses (>>5ms) will be longer than the defect repetition period (4.12ms for f_i=242.72Hz). This means that individual impulses are overlapped together. Therefore, both the envelope spectrum and envelope autocorrelation function are likely to fail in identifying the defect components (see Fig.4b and Fig.4c). The explanation is that the long tailed impulses are always associated with narrow spectral bandwidth. Consequently, the spectral peaks (defect frequency components) produced by the repetitive impulses will be cut off by this bandwidth. Therefore, the defect frequency components may be truncated in cases where the narrow impulse bandwidth is accompanied by a high impulse repetition rate.

In this case, the power cepstral analysis of the envelope signal could result in a good solution (see Fig.4d) due to its ability to deconvolve the source vibration and the effects of system propagation paths [4]. In

fact, the autocorrelation function shown in Fig.4c is dominated by the two largest peaks in the power spectrum (Fig.4b) which correspond to the modulation sidebands of the shaft frequency (less important than the defect frequency for detection purpose). The logarithmic process of the power cepstrum (Fig.4d), on the other hand, emphasises the whole series of harmonics in the power spectrum even though some have very small amplitudes.

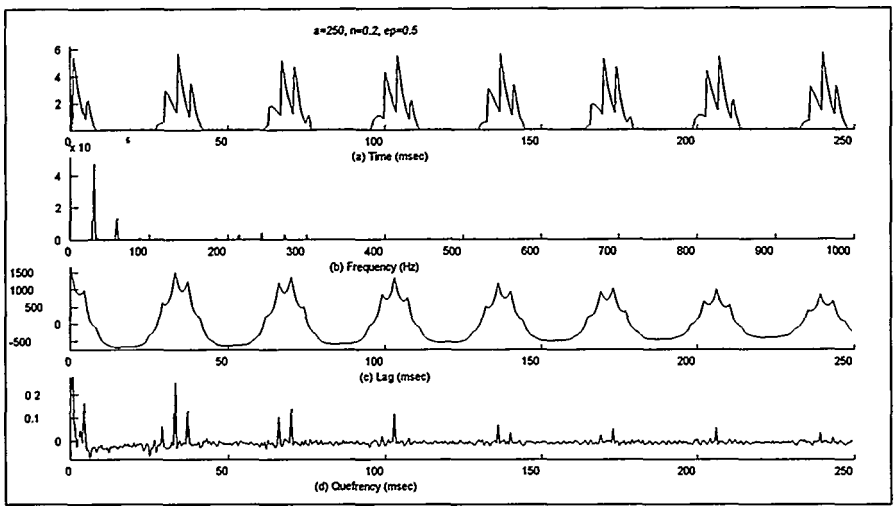


Figure 4. The envelope analysis for a synthetic vibration signal induced by an inner race fault (load distribution factor [1] $\varepsilon = 0.5$, α=250). **(a)** Time signal; **(b)** Envelope spectrum; **(c)** Envelope autocorrelation; **(d)** envelope cepstrum. f_i= 242.72Hz, (T_i= 4.12ms), f_r= 29.15Hz (T_r= 34.31ms).

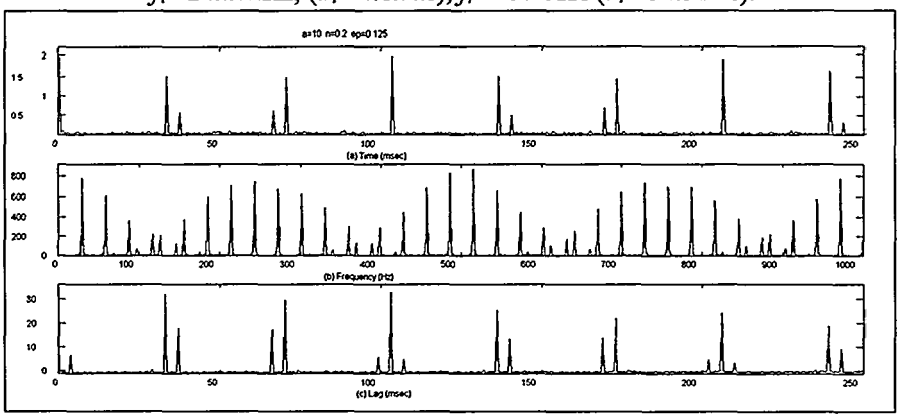


Figure 5. The envelope analysis for a synthetic vibration signal induced by an inner race fault (load distribution factor [1] $\varepsilon = 0.125$, time constant α=10). **(a)** Time signal; **(b)** Envelope spectrum; **(c)** Envelope autocorrelation function. f_i= 242.72Hz, (T_i= 4.12ms), f_r= 29.15Hz (T_r= 34.31ms).

At an earlier stage of fault development, the impulses induced by the

fault may have a very short transient. The corresponding spectrum tends to be widened. Moreover, as mentioned in the previous sections, when the bearing's load zone is narrow and sharp, the modulation sidebands will attenuate slowly and interfere with the harmonic components. An example of this case is given in Fig.5 where α is 10 and ε is 0.125. The envelope spectrum (Fig.5b) shows the interaction among the spectral harmonics and their slowly attenuated sidebands. The envelope autocorrelation function (Fig.5c), on the other hand, produces neat and narrow spikes in the lag domain which correspond to integer multiples of the inner race defect period. In this case, any broadband noise disturbance will be much more destructive to the spectrum than to the autocorrelation. This is because the noise autocorrelation components are concentrated at the zero-lag region [3].

6 EXPERIMENTAL ANALYSIS

Experiments were performed on a machine fault demonstration rig with NSK EN202 test bearings. During the experiments, the effects of load distribution and system damping on the detectability of the envelope analysis are the major concerns.

The first test was performed to create a case with light damping and a relatively high defect frequency (198.52Hz). A ball fault (unknown size) was implanted in a test bearing using welding sparks. During the test, a radial load of 200N was applied to the test bearing and the shaft was run at 50Hz. The vibration signal is shown in Fig.6a. Notice that the non-intermittent time signal has a similar appearance to that produced by an inner race fault. This suggests that the fault may even produce strong impulses when it strikes the cage. The envelope spectrum (Fig.6b) indicates the cage frequency clearly but shows no evidence of the ball defect frequency (two peaks around 200Hz are 190Hz and 209Hz respectively). However, the envelope autocorrelation function (Fig.6c), explicitly exhibits the basic ball defect period (5.04ms) and its integer multiples as well as the cage modulation effect. The envelope power cepstrum (Fig.6d) also gives a good indication of both the ball defect components and the effects of modulation. There are similarities between this case and that of synthetic signals with very large time constant (α=250) - see Figure 4. This confirms that the test rig for the measurement direction was lightly damped (the mode damping ratio ζ=0.01) and the impulse transients overlapped each other. When an enlargement of the time signal around 0.07~ 0.1 second

is shown (Fig.6e) the interference of the individual impulses is immediately visible.

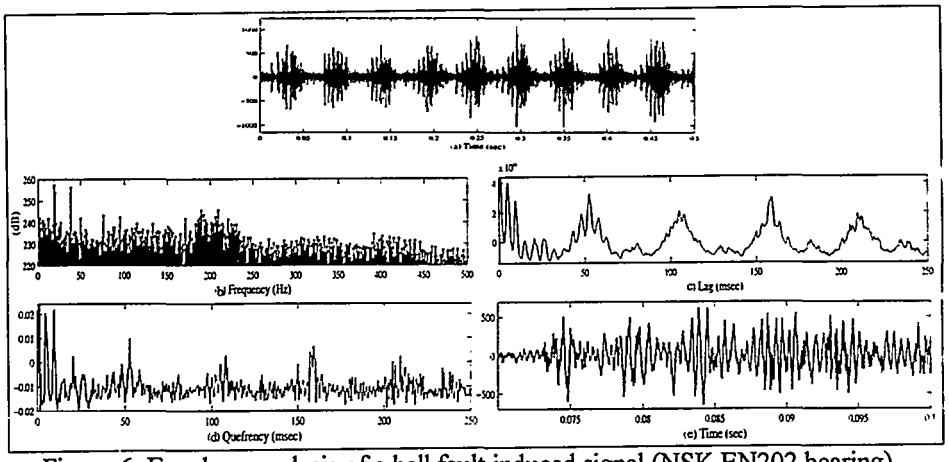


Figure 6. Envelope analysis of a ball fault induced signal (NSK EN202 bearing) with a light system damping. The defect components are: f_b=198.52Hz (T_b=5.037ms), f_c=19.06Hz (T_c=52.47ms). **(a)** Original signal, **(b)** envelope spectrum, **(c)** envelope autocorrelation, **(d)** envelope cepstrum, **(e)** segment enlargement of the time signal shown in (a).

The second test aimed at a case with heavier system damping (the mode damping ratio ζ=0.02), sharp load distribution (by increasing load intensity) and a low defect frequency (69.31Hz). To simulate a spalling damage, an inner race fault (about 1.0mm in diameter) was etched using nitric acid. During the test, the radial load of the test bearing was 500N and the shaft speed 14Hz. With the vibration signal shown in Fig.7a, the envelope spectrum (Fig.7b) almost loses its ability to detect the fault. However, the envelope autocorrelation function (Fig.7c) shows some integer multiples of the defect period. The peaks at 72ms, 144ms and 216ms correspond to 5, 10 and 15 multiples of the inner race defect period (T_i=14.4ms) respectively. It is worthwhile to point out that, because of the adjacency between T_r (71.4ms) and the 5th multiple of T_i (72ms) for this particular type of bearing, the envelope spectrum in Fig.7b seems to be dominated by the shaft frequency components (14Hz, ...). The inner race defect frequency components (69.31Hz, ...) are not immediately distinguishable. In fact, the inner race defect frequency components still exist in the spectrum (indicated by arrows) but are embedded in the slowly attenuated modulation components. This phenomenon is caused by a sharp modulation waveform. The sharp modulation waveform also results in amplitude attenuation of some integer multiples of the ball defect period in Fig.7c, such as 1st, 6th, 11th and 16th multiples marked by arrows.

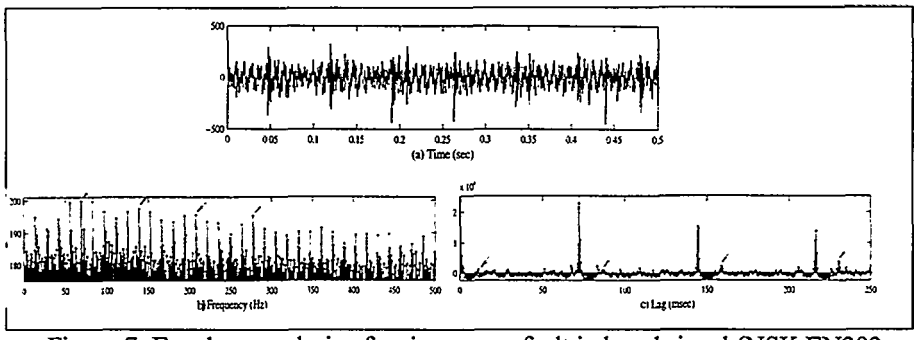


Figure 7. Envelope analysis of an inner race fault induced signal (NSK EN202 bearing) with a sharp load distribution. The defect components are: f_i=69.31Hz (T_i=14.43ms), f_r=14Hz (T_r=71.43ms). **(a)** Original signal, **(b)** envelope spectrum, **(c)** envelope autocorrelation (15 averages).

7 CONCLUSIONS

The effects of system damping and load distribution on various bearing diagnostic techniques have been examined in the preceding sections. The concluding remarks can be drawn from the above results:

i) System damping affects the diagnosis of inner race and ball defects substantially. For lightly damped systems, the envelope cepstrum is a good choice; for heavily damped systems, the envelope autocorrelation is superior to the envelope spectrum.

ii) The load distribution influences the attenuation rate of sidebands around each harmonic component. When light damping and sharp load distribution coexist, the autocorrelation technique shows an advantage over the spectral technique because of the severe intertwinement between sidebands and harmonics in the spectrum.

REFERENCES

(1) Mcfadden, P.D. and Smith, J.D. "Model for the Vibration Produced by a Single Point Defect in a Rolling Element Bearing". Journal of Sound and Vibration 96, 1984, p69-82.

(2) Braun, S. and Datner, B. "Analysis of Roller/Ball Bearing Vibrations". Journal of Mechanical Design 101, 1979, p118-125.

(3) Wang, W.Y. and Harrap, M. "Condition Monitoring of Ball Bearings Using an Envelope Autocorrelation Technique". Journal of Machine Vibration 5, 1996, p34-44.

(4) Randall, R.B. Frequency Analysis, Bruel and Kjaer, Denmark, 1987, pp.165-184, 271-305.

SENSOR AND ACTUATOR OPTIMIZATION BASED ON CONFINED VIBRATIONS

Daryoush Allaei[1] and Yin-Tsan Shih
R & D Department, QRDC, Inc., PO Box 562, Excelsior, MN 55331-0562, USA

ABSTRACT

The main purpose of this paper is to report the preliminary results of a study to determine the influence of confined vibrations on the optimum locations of sensors and actuators. In order to effectively incorporate confined vibrations in the design stage of sensor/actuator arrays used in smart structures, a new design methodology is also introduced. A specified performance index based on the observability and controllability grammians of sensors and actuators, respectively, is utilized to evaluate the influence of confined vibrations on plate-type structural components. It is clearly demonstrated that modal confinement has a significant impact on the optimization of sensors and actuators. The proposed optimization procedure and design methodology is proven to be effective in addressing some of the important issues surrounding this subject.

1 INTRODUCTION

One of the challenging research topics in active noise and vibration control is the performance optimization of sensor/actuator arrays given the constraints in real engineering applications (i.e., their limited number, restricted locations, power consumption, and cost). Sensor/actuator arrays are usually optimized based on their overall efficiency, required control effort (input energy), temporal characteristics (i.e., settling time, rise time, etc.), frequency characteristics (i.e., bandwidth, cut-off frequency, etc.), spatial distribution over the structure, rate of energy dissipation, and the robustness of the control system. Several combinations of these measures have been reported as the optimal sensor/actuator design criteria. However, the influence of the spatial distribution of vibration energy, namely vibration confinement, has not been previously investigated. The latter topic is the main focus of the work presented in this paper.

Based on a variety of criteria, several sensor/actuator optimization schemes have been reported in the open literature [1-12]. Schulz and Heimbold [1] presented an optimization criterion based on the maximum dissipation energy through the control action under velocity feedback control law. Kondoh et. al. [2] proposed the determination of the sensor/actuator positioning and feedback gain based on the minimization of the quadratic cost functional in the standard optimal control problem. Skelton et al. [3-5] have proposed two methods: Modal Cost Analysis (MCA) and Component Cost Analysis (CCA). They argued that the number of sensors, actuators, and system components can be reduced by removing those with the smallest modal or component cost.

An effective sensor/actuator optimization criteria is based on the measures of controllability and observability. Various measures of controllability and observability have been proposed [6-12]. Hac et al. [12] proposed an observability and controllability measure based on the controllability and observability grammians. The method is not computationally intensive since closed form solutions of the grammians can be found for flexible structures. The latter measure is utilized in the work discussed in this paper.

---

[1] This author can also be reached at da-qrdc@msn.com or http://www.mntech.org/mcdc/qrdc.htm

It is known that the control system may fail due to one or more uncontrollable or unobservable modes. For example, if a sensor or an actuator is placed at or near the nodes of a normal mode of a structure, the observability and controllability of that mode could be very low. This problem may become more severe when the modes of a structure have a large number of nodal points, nodal lines, or nodal surfaces (in the case of confinement). Based on this argument, the presence of confined vibrational energy may be detrimental to the performance of sensors or actuators, and may eventually result in the failure of the vibration/noise control systems. The former effect is due to the fact that confined vibrations result in the conversion of nodal points to nodal lines and nodal lines to nodal surfaces. In other words, the areas over which the modes of vibration have very low observability and/or controllability will significantly increase due to confinement. Figures 2a and 2b show the unconfined and confined modes of a plate-type structural component. Note that the nodal lines of the unconfined plate (with normal modes shown in Figure 2a) are converted to a nodal surface when vibrations of the plate are confined as shown in Figure 2b. Such regions with low vibration levels can be major troublesome areas when placing sensors or actuators.

Fundamentals of mode localization in disordered periodic structures have been extensively investigated [16-19] since the mid 1970's. However, their applications to real engineering structures and non-periodic systems have surfaced only since the early 1990's. Confined vibrations in structures is a more general form of mode localization [16-19]. Confinement could be the result of localized modes, or could be incorporated in the total vibration response of a structure without localizing its individual modes. However, mode localization may not necessary result in the confinement of the total vibration response of a structure when subjected to certain loading configurations. In this paper, confinement is defined based on the total vibration response of a structure, and it can be induced via mode localization or application of external force. Since the focus of this paper is not how to induce confinement, no more discussion will be presented on the latter subject. For more detailed description on how these two types of confinement can be induced, the authors recommend the references [18-19] listed in this paper.

In summary, confined vibration may be detrimental to vibration control systems if it occurs in an uncontrolled manner and without the knowledge of the control system designers. On the other hand, one may take advantage of the presence of the confined vinrations to optimize the sensor/actuator system by focusing on the areas where vibrations are confined. This concept may have the potential to decrease the required number of sensors and actuators, increase their effectiveness, and improve the overall performance of the vibration control system by a significant factor. However, the current sensor and actuator optimization techniques may be unacceptable for the realization of the beneficial or detrimental effects of the occurrence of mode localization and/or vibration confinement in structures. A new methodology and design tool are required to incorporate (when advantageous) or to prevent (when disadvantageous) the occurrence of confined vibrations in the optimization of sensors and actuators, and thereby significantly improve the performance of noise and vibration control systems.

The purpose of this paper is to present the preliminary results of a study [13] to determine the influence of the confined vibrations on the sensor and actuator optimization. To the best knowledge of the authors, this is the first study in which the effect of the confinement on the performance of sensor/actuator is investigated. Also, a new design methodology to incorporate the confined modes in the design stage of the sensor/actuator sets of a structure is introduced in this paper.

To demonstrate the influence of the confinement on the optimization of the sensor and actuator locations, Hac's optimization method [12] is applied to two plate structures whose only difference is the degree of confinement. A Performance Index (PI) based on the controllability and observability grammians is used as the

optimization criteria to optimally locate a sensor and/or actuator array. The results presented in this paper can also be demonstrated by employing other optimization schemes. The PI method is selected because it can clearly show the dependency of the modal confinement to the numerical values of PI.

The remaining sections of this paper are as follows. The main formulas describing the PI method is presented in Section 2.1 followed by a brief description of the optimization method in Section 2.2. A brief discussion on the anticipated impact of confinement on the numerical values of PI is provided in Section 2.3. A demonstrative application problem is introduced in Section 3.1. The numerical results and analysis are presented in Section 3.2. A methodology to incorporate the concept of Vibration Control by Confinement (VCC) in the optimization of sensor/actuator array is proposed in Section 3.3. Summary and conclusions are presented in Section 4 followed by a brief list of references.

2 FORMULATION, OPTIMIZATION SCHEME, AND CONFINED VIBRATIONS

The key PI-formulas utilized in this work are presented in Section 2.1. The reader is referred to the paper by Hac [12] for detailed formulations. An optimization criteria based on PI is presented in Section 2.2. Finally, a brief introduction on the basics of confined vibration and its potential influence on the sensor/actuator optimization is discussed in Section 2.3. For a detailed discussion on this subject, the reader is referred to articles and reports written by the first author [13, 16-19].

2.1 Formulation

The derivation presented in this section is based on the assumption that the dynamics of the structure can be represented by a linear second order system. The formulation is applicable to cases when a structure is subjected to either persistent or transient disturbances.

The PI criterion relies on the observability and controllability in the context of location of sensors and actuators. The PI formulation proposed in reference [12] is based on the argument that the size of the controllability and observability grammian and their individual eigenvalues must be as large as possible. These two conditions can be characterized by the sum and product of their eigenvalues as shown in Eq. (1).

$$PI = \left(\sum_{j=1}^{2md} \lambda_j \right) \sqrt[2md]{\prod_{j=1}^{2md} (\lambda_j)} \quad (1)$$

Where λ_j is the jth eigenvalue of the controllability or observability grammian. The same criterion is applicable for both displacement and velocity sensors and actuators. The PI defined by Eq. (1) guarantees a minimum control force and energy. In the case when both displacement and velocity sensors are used, it can be shown [12] that the observability grammian is the sum of two individual grammians.

In the case of an actuator, for a system with small damping and well spaced natural frequencies, the eigenvalues of the controllability grammian can be approximated [12] as shown below.

$$\lambda_{2i-1} = \lambda_{2i} = \beta_i / 4\zeta_i \omega_i \quad (2)$$

$$\beta_i = \sum_{q=1}^{m} [\phi_i(x_q)]^2 \quad (3)$$

where i is the number of modes, and x_q is the location of actuator q. Eq (3) shows how the eigenvalues of controllability grammian are related to the magnitude of the mode shapes at the location of the actuators. The eigenvalues of the controllability grammians are used to determine the performance index defined in this section. It should be noted that if the actuators are placed at the vibration nodal points, the eigenvalues of controllability grammian will be significantly low. In the case of confinement, in which nodal segments and surfaces may be present,

the magnitudes of the eigenvalues of the controllability grammian can become severely low. The latter is an indication that the actuators are ineffective.

Assuming that r sensors (displacement or velocity type) are used to measure the vibration response of the structure, the observability grammian can be formulated in a closed form solution [12] for both displacement and velocity sensors. For structures with small damping and well separated natural frequencies, the eigenvalues of the observability grammian can be approximated for displacement and velocity sensors as shown by Eqs. (4) and (5), respectively.

$$\lambda_{2i-1} = \lambda_{2i} = c_{di} / 4\zeta_i\omega_i \quad , \quad c_{di} = \sum_{k=1}^{r}\left(\frac{\Phi_i(x_k)}{\omega_i}\right)^2 \tag{4}$$

$$\lambda_{2i-1} = \lambda_{2i} = c_{vii} / 4\zeta_i\omega_i \quad , \quad c_{vi} = \sum_{k=1}^{r}\left(\Phi_i(x_k)\right)^2 \tag{5}$$

As it can be observed from Eqs (4) and (5), the eigenvalues of the observability are strongly related to the magnitude of the mode shapes determined at the sensor locations. As it was previously pointed out, it is expected that when confinement is induced, the eigenvalues of the observability grammian will be significantly lower than when the modes are extended throughout the structure (i.e., the structure has normal or unconfined modes). The latter is due to the fact that the confined structure will have nodal regions in contrast to nodal points.

2.2 Optimization Approach

As defined by Eq. (1), the Performance Index, PI, is a function of the eigenvalues, λ_j, of the controllability or observability grammians. As it was previously discussed, eigenvalues of the controllability and observability grammians are in turn functions of the locations of sensors and actuators. Therefore, sensor/actuator locations can be selected as design variables in the optimization problem. The natural frequencies, the damping factor, and the mode shapes of the unconfined and confined structures are assumed to be independent of the locations of the sensors and actuators. The optimization problem is solved by finding the best location for sensors and actuators so that (-PI) has a minimum value at (in other words, PI is maximized). Where x_a^* and x_s^* are optimal locations of actuators and sensors, respectively.

$$\min_{x_a} \; -PI(x_a) \rightarrow x_a^* \quad , \quad \min_{x_s} \; -PI(x_s) \rightarrow x_s^* \tag{6}$$

$$x_a \min \le x_a \le x_a \max \; , \; x_s \min \le x_s \le x_s \max \tag{7}$$

Eqs. (6)-(7) represent a nonlinear constrained optimization problem. The interior penalty function method [14] is utilized to account for the constraints. In this method, a new function is constructed by adding a penalty term to the objective function, -PI. The penalty term is chosen such that its value will be small at points away from the constraint boundaries and will tend to infinity as the constraint boundaries are approached. That is the penalty term forces the optimization program to find the optimal locations of sensors and actuators away from the constraint boundaries. By adding the penalty term to the objective function, the constrained minimization is transferred to an unconstrained minimization problem. Any unconstrained minimization method can be used to obtain the optimum locations of the sensors and actuators. A general purpose unconstrained optimization technique, the metric variable method [14], is adopted for this study. This method was first proposed by Davidon [20] and was then extended by Fletcher and Powell in 1963 [15]. The metric variable method generates a sequence X_i ($X = x_a$ or x_s) converging to a local optimal solution X^*. This optimization technique can be though of as a quasi-Newton method and also as a conjugate gradient approach. It is very powerful and converges quadratically due to its similarity to the conjugate gradient method. The method is also very stable and

continues to progress towards the minimum even while minimizing highly distorted and eccentric functions.

2.3 Confined Vibrations

Mode localization implies the confinement of the modal response to a limited segment of a structure, as opposed to a conventional modal response which extends throughout a structure. Vibration confinement, however, suggests that the total vibration response of a structure is confined to a limited region. The latter can be caused by either inducing mode localization or applying external forces whose magnitude may depend on the displacement field and its derivatives. More details on the basic concept of mode localization and an extensive review of the state-of-art of the Vibration Control by Confinement (VCC) can be found in reference [16-19].

The occurrence of mode localization and confinement of vibrational energy in structures is now widely accepted because its presence and effects have been demonstrated since 1958 in a variety of structures [18-19, 21-29]. Until the early 1990's [21-22], however, it was believed that mode localization occurs only in nearly periodic/cyclic lattices and structures. Furthermore, the cause of the occurrence of the phenomenon in such structures was reported to be due to unavoidable small material tolerances, defects caused during manufacturing processes, and/or assembly variations that lead to small differences between the individual components of nearly periodic structures. Thus, the phenomenon was often associated with its random behavior, uncontrollable features, skewing of the prediction models, and catastrophic failures that it can cause. In early the 1990's, the research community began questioning the validity of these reports. It has been demonstrated [13, 16-19, 23-29] that the phenomenon is more general than previously thought. First, it is not limited to nearly periodic structures, rather mode localization can occur in any structure should the design parameters fall in certain ranges. Second, as it was previously thought, the occurrence of the phenomenon can be random due to the randomness in the variations of the design parameters. However, confinement of vibrations can be passively and/or actively induced in a controlled manner in both modal domain and global dynamic response of structures. Third, on one hand, displacement and stress concentrations due to localized vibrations can be potentially harmful to structures and, on the other hand, confinement of vibrations can be advantageous if used to enhance the performance of the passive/active vibration/noise control systems. The focus of the present work is to demonstrate that confinement of vibrational response has a significant influence on the sensor/actuator optimization. Furthermore, a systematic procedure is proposed to account for such influences when integrating sensor/actuator arrays in active vibration control systems and smart structures. Finally, it will be shown that the performance of sensor/actuator arrays can be improved by inducing confinement.

There are several ways [16-19] to induce confinement in a structure. Variations in material and/or structural parameters can cause the occurrence of confinement whose severity depends on the its sensitivity to the variation of such design parameters. In terms of material properties, variation in elastic constants and mass density can be selected as the design parameters to either avoid or induce confinement. In terms of structural parameters, boundary conditions and coupling constants are among the parameters that can cause confinement. A collection of methods to induce confinement (and mode localization) in structures has been reported in references [16-19]. In the present work, one of the simplest means of inducing modal confinement is selected because the focus of the study is to investigate the impact of confinement of the sensor/actuator optimization rather than how to implement confinement.

In this work, confined modes are induced by adding a rib stiffener to the main plate structure. The dimensions of the rib stiffener is adjusted to induce confined

vibrations in the plate structure as shown in Figure 1. System and confinement specifications are listed in Table 1. The natural frequencies of the main plate structure undergo a relatively small shift due to the additional stiffness. In other words, the two structures, unconfined and confined plates, have parallel vibration characteristics.

3. NUMERICAL RESULTS, ANSLYSIS, AND DESING

3.1 Structural Specifications

In order to demonstrate the impact of confinement on the optimization of sensors and actuators, a plate-type structure is selected as an example. The plate displayed in Figure 1 is supported by metal skirts at its four boundaries. The metal skirts are used to simulate pinned boundary conditions. To induce confinement, a ring stiffener is placed at 40% of the plate length as shown in the figure. The plate without the rib does not exhibit confinement (i.e. its natural modes or total vibration response extend throughout the plate). The specifications of the plate structure are listed in Table 1.

Table 1 Specifications of the plate and its rib stiffener

Specification	Plate	Ring Stiffener
Material	Stainless Steel	
Young's modulus: E	19.03×10^{10} N/ m^2	
Density: ρ	7746.7 Kg/m^3	
Poisson's Ratio: μ	0.3	
Thickness, Width, and Length: h, w, L	1.0, 300, 450 mm	6.0, 20, 300 mm
Location of the center of the Rib	180 mm (~40%) measured from left edge	
Theoretical Modal Damping: ζ	0.01 for all modes	

3.2 Results and Analysis

In this paper, we limit the presentation of our results to the case of one or two sensors due to the page limitation. However, based on the formulations presented in the previous sections and our observation, the discussion and conclusions presented in this section can be extended to other cases involving multiple displacement sensors, velocity sensors, and actuators.

The lowest four modes of the unconfined and confined plates are shown in Figures 2a and 2b, respectively. It should be noted that in the case of confined plate (Figure 2b), modes 1 and 3 are confined to right while modes 2 and 4 are confined to the left side of the plate. The lowest four modes are chosen as the controlled modes.

In the case of the plate with unconfined modes, the numerical values of normalized PI versus the location of a displacement sensor are displayed in Figure 3a. The values of PI are normalized with respect to the global maximum. The PI curve is symmetric with respect to the center of the plate. Furthermore, its values drop to zero when the sensor is placed at any of the nodal points. The latter observation confirms the fact that that PI is strongly sensitive to the magnitude of modes at the sensor locations.

Figure 3b shows the PI values versus the location of a displacement sensor attached to the plate-rib structure whose lowest four modes are confined. Again, the values of PI are normalized with respect to the global maximum. Note that contrary to Figure 3a, the PI curve shown in Figure 3b is not symmetric with respect to the center of the plate. The latter is due to the fact that confined mode shapes are not symmetric with respect to the center of the plate. Furthermore, the left portion of the PI curve shown in Figure 3b has near zero values. In other words, from an observability point of view, the left side of the plate is converted to a nodal segment in contrast to nodal lines. It should be pointed out that when the sensor is located on the right or left side of the plate-rib structure, one of the

modes is not observable. In the case of the unconfined plate, all modes are observable unless the sensor is placed at one of the nodal lines.

Next, the optimum location of the sensor is evaluated for the unconfined and confined plates. Based on the objective function defined in Section 2.2, the optimal location of a sensor, for example, is at the point with the local maximum PI. The results of three cases calculated by the optimization program for a single displacement sensor are shown in Table 2. Table 3 lists the optimal points when two displacement sensors are located on the plate.

Table 2 Optimal position of one sensor based on PI maximization

		Unconfined		Confined	
Case	Initial Position (x, y), [m]	Optimal Position (x, y), [m]	Normalized PI at Optimal Position	Optimal Position (x, y), [m]	Normalized PI at Optimal Position
1	(0.06, 0.06)	(0.093, 0.082)	0.436	(0.0757, 0.082)	0.414
2	(0.15, 0.105)	(0.180, 0.121)	1.0	(0.124, 0.126)	0.212
3	(0.24, 0.21)	(0.269, 0.217)	0.814	(0.250, 0.217)	0.390

Table 3 Optimal position of two sensors based on PI maximization

		Unconfined		Confined	
Case	Initial Position [m] (x1, y1) (x2, y2)	Optimal Position [m] (x1, y1) (x2, y2)	PI at Optimal Pos. [$\times 10^{-12}$]	Optimal Position [m] (x1, y1) (x2, y2)	PI at Optimal Pos. [$\times 10^{-13}$]
1	(0.06,0.06); (0.15,0.105)	(0.13,0.08); (0.19,0.13)	4.85	(0.08,0.16); (0.07,0.10)	3.00
2	(0.06,0.06); (0.24,0.21)	(0.09,0.08); (0.18,0.16)	5.03	(0.08,0.08); (0.30,0.22)	7.07
3	(0.15,0.105); (0.24,0.21)	(0.18,0.16); (0.27,0.21)	5.78	(0.10,0.13); (0.26,0.22)	5.34

Note that the PI values given in Table 3 are not normalized because the global optimal position was not evaluated. In order to determine the global optimal location and its corresponding PI for these two sensors, an iterative process would need be carried out. The global maximum is found for each location of a fixed sensor/actuator while the other is swept at various positions on the plate. The optimum of these global optimal positions will then be the final global optimum location. This is a very time consuming process. Due to the limited scope of this project, the final global optimal was not determined.

Based on above results and comparison, it is clear that vibration confinement has a significant effect on the optimal location(s) of the sensors. Laboratory tests (not presented here) were conducted to verify the above theoretical results.

3.3 A Proposed Design Methodology

In the previous section, the impact of confinement on the optimal position of one and two sensors was presented. It was shown that when the natural modes of the structure are confined, there will be a larger region where the structure is not observable (nodal line or nodal surface). The objective function of the optimal control algorithm (in terms of location of the sensors) should be refined in order to incorporate the effect of the confined modes. Without considering the modal confinement of the structure, the optimal design of the sensor/actuator system may not be achieved when vibration response of a structure is confined. On the other hand, one may take advantage of the presence of the modal confinement to optimize the sensor/actuator system by focusing on the areas where vibrations are confined. As it was pointed out, this concept may have the potential to decrease the required number of sensors and actuators, increase their effectiveness, and improve the overall performance of the vibration control system. Therefore, a new design methodology is proposed by integrating some of the above observations in the optimization process. The following steps are recommended in order to account for modal confinement in the optimal sensor/actuator design problem.

1. Identify the vibration control requirement based on the safety and/or performance of critical regions of the structure. For example, the critical

sections may include areas at which the structure is highly sensitive to the external disturbances, sensitive equipment is installed, or radiated noise must be a minimum.

2. Identify the confinement requirements and configuration. Such requirements may include modal characteristics, energy distribution, frequency range of interest, and severity of confinement.
3. Choose the most suitable means of inducing confinement or to avoid its random occurrence. The structure should be confined so that excess vibrational energy is kept away from the identified critical areas. The control effort could then be concentrated in the limited critical areas.
4. Identify sensor and actuator requirements. This includes the number, type, power requirement, location constraints, frequency range of interest, and the number of controlled modes.
5. Utilize an optimization program to find the optimal location and number of the sensors and actuators. Sensors and actuators can be placed in the confined areas to obtain high performance.
6. Check the results against the design criteria. If the results are not satisfactory, then go back to step 2. Modify the confinement requirements and configuration and then repeat the steps 3 to 6 until a satisfactory design is achieved.

The above proposed procedure integrates two techniques: optimal sensor and actuator techniques used in active vibration control and passive and/or active version of Vibration Control by Confinement (PVCC or AVCC). Since the area to be controlled is reduced, the number of sensors/actuators is expected to be reduced while the overall power consumption by the control system should also be significantly decreased.

4. SUMMARY AND CONCLUSIONS

It was shown that the influence of energy confinement on the optimal position of sensors/actuators can be significant. A performance index based on a measure of the observability and controllability gramians was used to evaluate the efficiency of the sensors and actuators. It was demonstrated that when natural modes were confined, the PI curve had a significantly different shape, and the optimal sensor positions were considerably different than those of the unconfined cases. This study has shown the necessity to consider the modal or global confinement in the design stage of a control system.

A new design methodology was also proposed which incorporates the effects of vibration confinement into the optimal locations of sensors and actuators. Based on the proposed methodology, the energy is confined to non-critical areas to which concentrated control effort can be applied to in order to gain an improved control performance.

REFERENCES

1. Schulz, G. and Heimbold, G., "Dislocated Actuator/ Sensor Positioning and Feedback Design for Flexible Structures," J.of Guidance, V. 6(5), pp. 361-367, 1983.
2. Kondoh, S., Yatomi, C., and Inoue, K., "The Positioning of Sensors and Actuators in the Vibration Control of Flexible Systems," JSME International Journal, Series III, Vol. 33, No. 2, pp. 145-152, 1990.
3. Skelton, R. E. and Hughes, P. C., "Modal Cost Analysis for Linear Matrix-Second-Order Systems," J. of Dyn. Sys., Meas., and Control, Vol. 102, pp. 151-158, 1980.
4. Skelton, R. E., "Cost Decomposition of Linear Systems with Application to Model Reduction," International Journal of Control, Vol. 32, No. 6, pp. 1031-1055, 1980.
5. Yousuff, A. and Skelton, R. E., "Controller reduction by Component Cost Analysis," IEEE Transactions on Automatic Control, Vol. AC-29, No. 6, pp.520-530, 1984.

6. Moore, Bruce C., "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction," IEEE Transactions on Automatic Control, Vol. AC-26, No. 1, pp. 17-32, 1981.
7. Viswanathan, C. N., Longman, R. W., and Likins, P. W., "A Degree of Controllability Definition : Fundamental Concepts and Application to Modal Systems," Journal of Guidance, Vol. 7, No. 2, 1984, pp. 222-230.
8. Viswanathan, C. N., Longman, R. W., and Likins, P. W., "A Definition of the Degree of Controllability - A Criterion for Actuator Placement," 2nd VPI&SU/AIAA Symp. on Dynamics and Control of Large Flexible Spacecraft, pp. 369-384, 1979.
9. Lindberg, R. E. and Longman, R. W., "The Influence of Time an Normalization on Actuator Placement by Degree of Controllability," Proceeding of the AAS/AIAA Astrodynamics Specialist conference, Lake Placid, NY, Aug. 1983.
10. Lindberg, R. E. and Longman, R. W., "Optimization of Actuator Placement via Degree of Controllability Criteria Including Spillover Considerations," AIAA/AAS Paper 82-1435, San Diego, CA, Aug. 1982
11. Hamdan, A.M.A. and Nayfeh, A.H., "Measures of Modal Controllability and Observability for First- and Second-Order Liner Systems," J. of G. V12(3), 1989.
12. Hac, A. and Liu, L., "Sensor And Actuator Location in Motion Control of Flexible Structures," Journal of Sound and Vibration Vol. 167, No. 2, 1993, pp. 239-261
13. Allaei, D., "Application of Localized Vibration and Smart Materials in Controlling the Dynamic Response of Structures, Part V: Sensor and Actuator Optimization, " DARPA semi-annual report, DARPA Cont. No. DAAH01-94-C-R001, May 1996.
14. Rao, S. S., "Optimization -- Theory & Applications," 2nd edition, John W&S, 1986.
15. Fletcher, R. and Powell, M.J.D., "A rapidly convergent descent method for minimization ,"Computer Journal, Vol. 6, No. 2, pp.163-168, 1963.
16. Allaei, D., "Application of Localized Vibration and Smart Materials in Controlling the Dynamic Response of Structures, Part III, " DARPA report, 1994.
17. Allaei, D., "Application of Localized Vibration and Smart Materials in Controlling the Dynamic Response of Structures, Part IV, DARPA report, 1995.
18. Allaei, D., "Vibration Control by Confinement: An Overview", 4th International Congress on Sound and Vibration, St. Petersburg, Russia, June 1996.
19. Allaei, D., "Passive and Active Vibration Control by Confinement: An Overview", Society of Engineering Science, 33rd Annual Technical Meeting, October 1996.
20. Davidon, W.C., "Variable Metric Method of Minimization", Argonne National Laboratory Report No. ANL-5990, 1959.
21. Hodge, C.H., "Confinement of Vibration by Structural Irregularities," Journal of sound and Vibration, Vol. 28, No. 3, pp. 411-424, 1982.
22. Levine-West, M.B. and Salama, M.A., "Mode Localization Experiments on a Ribbed Antenna," AIAA J., 31(10), 1993.
23. Tarnowski, D. J., Allaei, D., Landin, D., Chen, C., and Liu, S. "An innovative noise and vibration suppression technique based on the combination of vibration control by confinement and layered damping approaches," presented at the 3rd Joint Meeting of the Acoustical Society of America and Acoustical Society of Japan, Hawaii, 1996.
24. Yigit, A. and Choura, S., "Vibration Confinement in Flexible Structures via Alternation of Mode Shapes by Using Feedback," J. of S. & Vib. 179(4), 1995.
25. Shelley, F.J. and Clark, W.W., "Active Mode Localization in Distributed Parameter Systems with Consideration of Limited Actuator Placement," 15th ASME Biennial Conf. On Vibration & Noise, 1995.
26. Allaei, D., "Performance Comparison Between Vibration Control by Confinement and Conventional Techniques," to be presented Symposium on Active and Hybrid Structural Control, ASME 16th Biennial Conf. on Mech. Vib. & Noise, 1997.
27. Allaei, D. and Shih, Y. T., "Influence of Confined Vibration on Sensor and Actuator Optimization," SAE paper 97NV48, 1997.
28. Allaei, D. and Shih, Y. T., "A comparison between Passive Vibration Control by Confinement and Current Passive Techniques," SAE paper 97NV61, 1997.
29. Clark, W.W., Sun, F., and Tarnowski, D.J., "On the Use of Vibration Confinement to Enhance Conventional Vibration Control," submitted to J. of Vib. and Control, 1997.

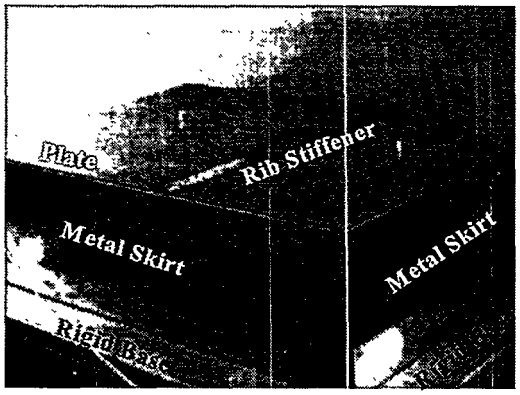
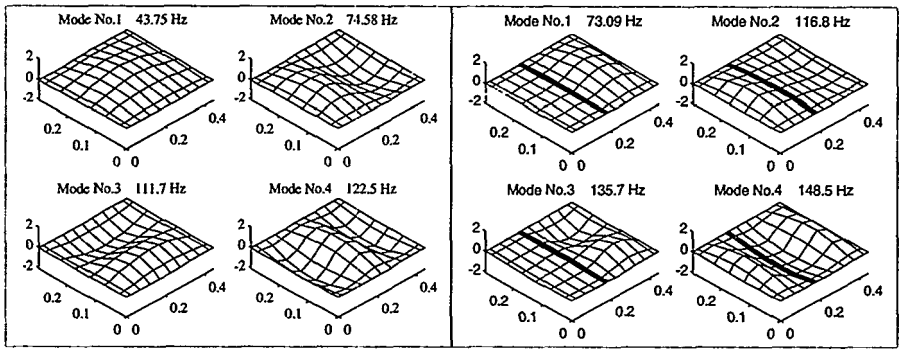
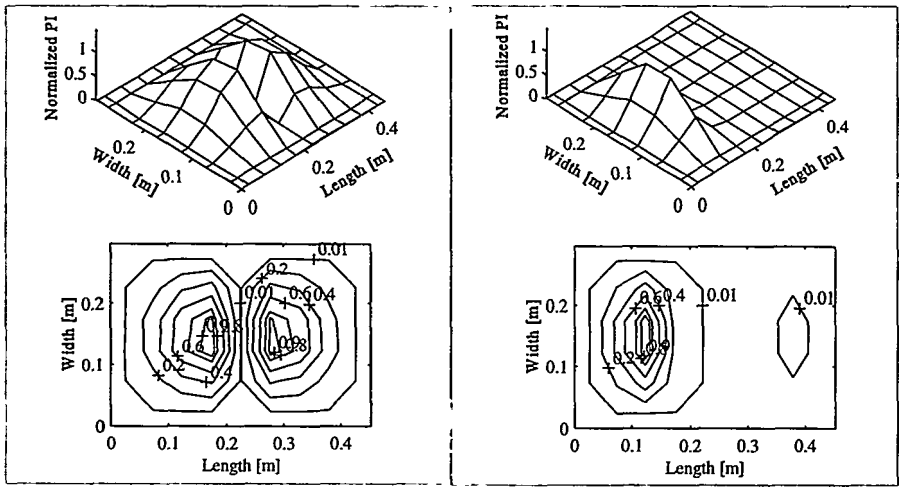


Figure 1 Photo of the plate-rib structure used in this study



(a) Unconfined plate (b) confined plate-rib

Figure 2 The lowest four modes of the SSSS plate structure



(a) Unconfined plate (b) Confined plate-rib

Figure 3 3D and contour views of the PI for the plate and plate-rib structures

ROLLING ELEMENT AND FLUID FILM BEARING DIAGNOSTICS USING ENVELOPING METHODS

Anton Azovtsev, Alexei Barkov
VibroAcoustical Systems and Technologies, Inc.
VAST, Inc., 22 Rozenshteina st., St. Petersburg, 198095, Russia
http:// www.inteltek.com

ABSTRACT

Diagnostics of rotating machinery using the envelope spectrum of high frequency vibration exited by friction forces is becoming widely used by many companies and experts. In recent years, automatic diagnostic systems based on this method have been developed for rolling element bearings. These systems can provide detailed diagnostics and condition prediction of a bearing by a single vibration measurement. Automatic diagnostic systems for simultaneous diagnostics of a gears and rolling element bearings in a gear transmissions have also been developed. But until recently, there was no diagnostic systems that used envelope spectra of random vibration for detection and identification of defects in fluid film bearings.

The main problems in the use of standard enveloping methods developed for the diagnostics of rolling element bearings in the diagnostics of fluid film bearings are discussed in this paper. Methods for solution of these problems are analyzed together with the practical results achieved in this field.

1. PECULIARITIES OF ROLLING ELEMENT BEARINGS DIAGNOSTICS

The method of rolling element bearings diagnostics by the spectrum of high frequency vibration envelope is based on the analysis of characteristics in the formation of friction forces in good and defective bearings as well as in the features of shock pulses that appear in the

interaction of rolling surfaces with cavities, spalls, or cracks in the bearing elements [1].

The friction forces depend on the rolling friction coefficient and the load on rolling elements. In good bearings, the friction is uniform in time, i.e. it does not depend on the rotation angle of the rotating race or of the cage.

In rolling element bearings with defects of installation including misalignment of races and non-uniform radial tension, loads on the rolling elements increase and, more importantly, these loads become dependent on the rotation angle of the rotating race and the cage. As a result, the friction forces, together with the random vibration excited by them, become amplitude modulated.

In bearings with non-uniform wear of inner and outer races and rolling elements, the friction coefficient in turn depends on the rotation angle of rotating race and cage which results in similar amplitude modulation of the friction forces and the resultant high frequency vibration.

Finally, shock pulses in bearings with cavities and cracks on rolling surfaces and races produce vibration as well. On the resonant frequencies of rolling elements and races, this vibration is actually attenuated self-oscillations that should not be considered as random vibration. At other frequencies, shock pulses excite random, fast attenuated vibration that is also modulated in amplitude.

As a result, all defects of bearing installation, wear, and cavities can be detected by the spectrum analysis of the high frequency vibration envelope. Figure 1 gives examples of envelope spectra from a good rolling element bearing (a), a bearing with an installation defect (b), with wear of a rolling surface (c), and with a crack on a rolling surface(d).

The apparent simplicity of defect detection and identification by the envelope spectrum of a rolling element bearing vibration can not be fully realized in practice. There are two main characteristics of the friction forces and the resultant random vibration formation that are the reasons for this.

The first reason is connected with the characteristics of the loads applied to the bearings in the real machines. In addition to the normal load, the rotating load from the shaft wobbling of an unbalanced rotor may be applied to the bearing. This additional load may also depend on the rotation angle of the shaft which significantly complicates the problem of defect identification. For example, a bearing may be exposed to shock loads due to defects of a gear transmission.

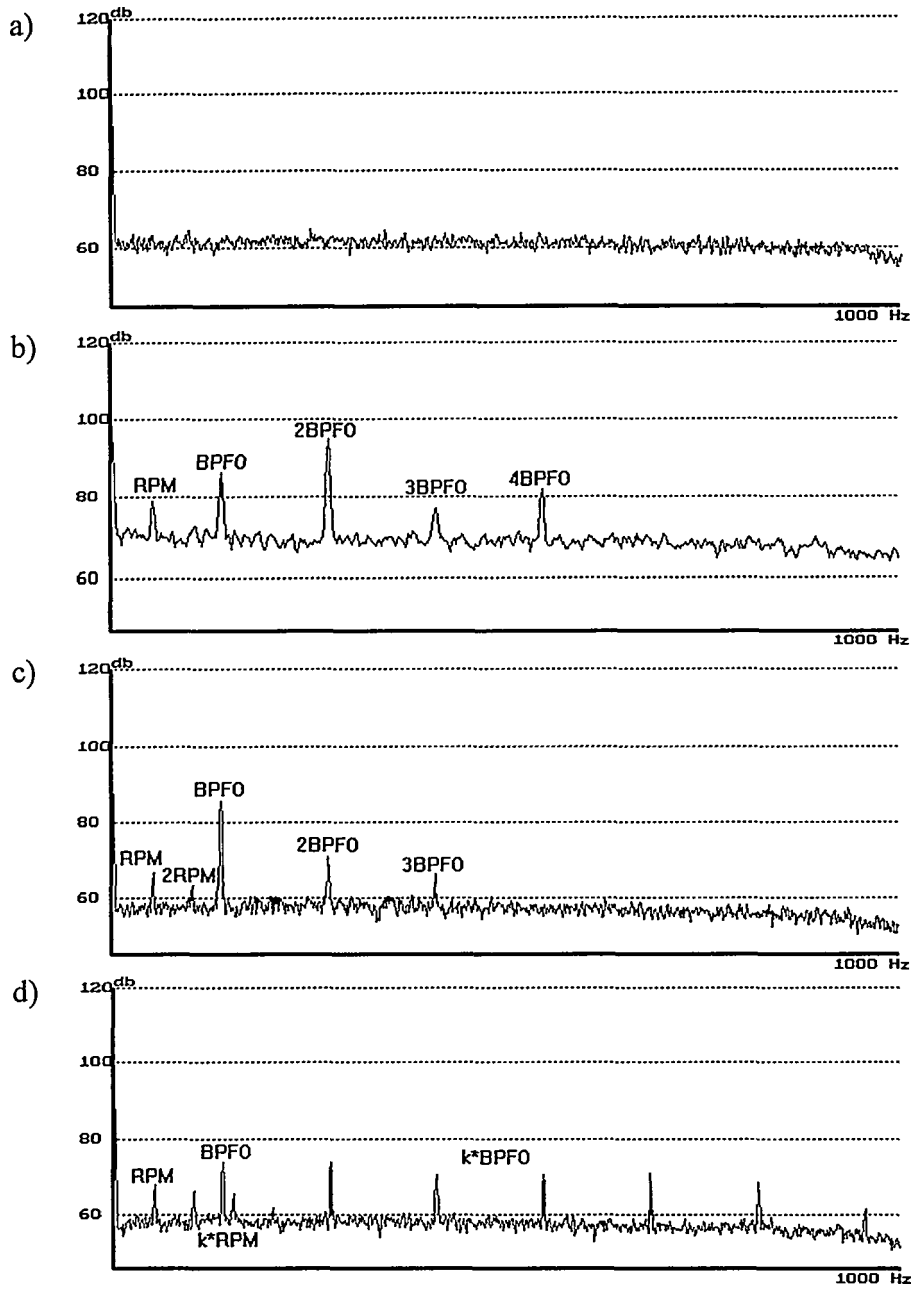


Figure 1. Examples of envelope spectra of rolling element bearings high frequency vibration.

a) A good rolling element bearing.
b) Bearing with a misalignment of an outer race.
c) Bearing with a wear of a outer race.
d) Bearing with a crack on outer race.

The gearing defects, for example, are detected by measuring the results of this shock load on the rolling element bearings [2]. Figure 2 presents envelope spectra of a rolling element bearing from a gearbox with one defective gear. Here, you can see envelope spectra measured on the bearing of the shaft with defective gearing and the adjacent shaft. The defects of gearing and cracks on the bearing surfaces can be distinguished by the repetition frequencies of the shock loads sequence.

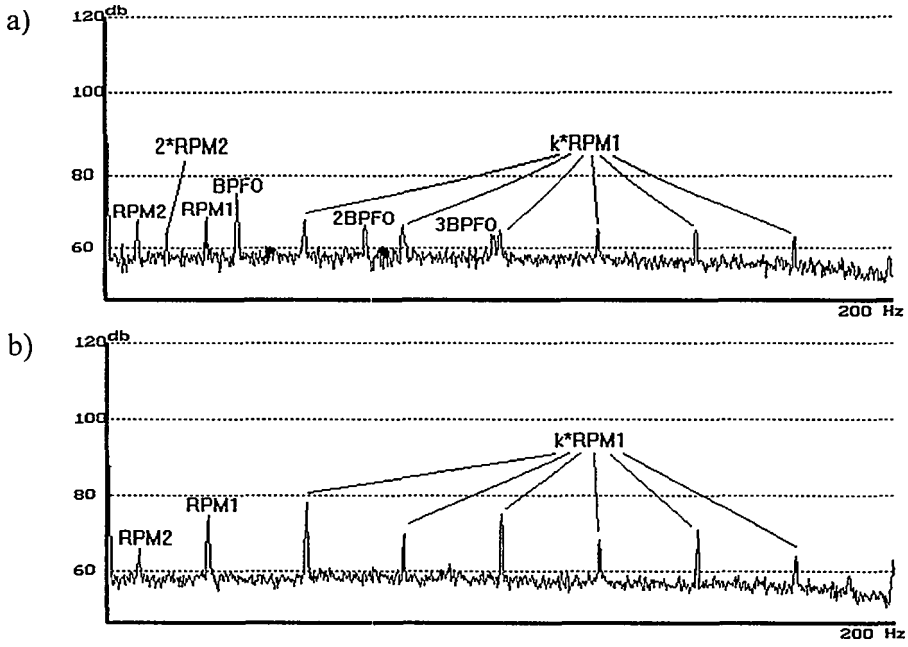


Figure 2. Envelope spectra of a rolling element bearing from a gearbox with one defective gear.

a) Bearing of the shaft with defective gear.

b) Bearing of the adjacent shaft.

The second reason for the complexity of fault detection and identification is connected with the necessity of detecting the envelope spectrum produced by only the random components of bearing vibration and excluding from consideration any of the harmonic components from either the bearing under diagnostics or from other machine units. Special methods for signal processing or careful choice of frequency band for enveloping should be used for this purpose [3].

2. POSSIBILITIES FOR JOURNAL BEARINGS DIAGNOSTICS.

High frequency random vibration is excited by friction forces in both rolling element and fluid film bearings. When defects of friction surfaces develop in these types of bearings, the friction forces and high frequency vibration acquire amplitude modulation and thus bearing defects can be detected by the analysis of envelope spectrum of this vibration. In the same situation, in the case of fluid film bearings, there are much more problems in the detection and identification of defects than in rolling element bearings. The first problem is connected with the limited diagnostic information that can be derived from the modulation frequencies. Rolling element bearings have at least 4 types of friction surfaces with different rotation speeds. These are outer and inner races, rolling elements, and the cage. We can list three fundamental bearing frequencies of rotation and all combinations of them. In case of fluid film bearings, there are only two friction surfaces and one fundamental frequency.

The second problem is due to the characteristics of pressure pulsation formation in the lubrication layer and the vibration excited by this pulsation. The pulsation power is determined by the velocity gradient in the lubrication layer, but it increases not only with rotation speed, but also with the decrease of lubrication layer thickness. The thickness of the lubrication layer, in turn, depends on the bearing design and on the relative position of the shaft axis and bearing shell axis. In a number of bearings, the shaft axis moves and, as a result, the high frequency vibration, may be modulated even in good bearings.

The third problem occurs in the detection of the defect severity. In the case of fluid film bearings, you should take into account not only the modulation index in the envelope spectrum, but the thickness of the lubrication layer as well. Reliable information about the thickness of the lubrication layer, as a rule, is not available for a user, thus the levels for defects have to be adapted for every machine type that has some special design characteristics of the bearings or of the machine.

Following many years of investigations and practical diagnostics of bearings in rotating machines using the envelope spectrum of high frequency vibration, we came to the conclusion that the above problems are not solvable in practice. A number of diagnostic symptoms were found which allow successful condition diagnostics of fluid film bearings [4]. These are based on the characteristics of shaft journal

oscillations in the lubrication layer of the fluid film bearing. Consider three of these characteristics.

The first characteristic is the possible appearance of short pulses with increased oil pulsation during the movement of oil wedge on the bearing shell surface. The rotation of the oil wedge is an indication of, first of all, of shaft wobbling. In a good bearing, such a pulse can appear when the oil wedge passes joints of the shell sectors. In the case of a worn bearing, when the oil wedge passes non-uniform wear zones on the bearing shell, cracks, and so fourth, such short pulses can be considered as shock pulses in fluid film bearings analogous to the shock pulses in rolling element bearings. An envelope spectrum of the journal bearing high frequency vibration in such a case in presented on figure 3 (b).

The second characteristic is possible appearance of shaft vibration at frequencies different from the harmonics of rotation speed. Most often, this is a self-sustained oscillation of the shaft in bearings with loose clearances or defects in the oil supply system. In most practical cases, self-sustained oscillations of the shaft synchronize with one of the sub-harmonics of the rotation speed. Sometimes you can also observe pendulum shaft oscillation in very loose bearings which also, as a rule, synchronize with one of rotation speed subharmonics. An example of the envelope spectrum in this case is presented in figure 3 (c).

And the third characteristic is the appearance of ultra-low frequency random oscillations of the shaft relative to the bearing shell surface. This situation can be found in the bearings with non-uniformly worn shells. These oscillations are defined by the unstable shape of the oil wedge with small and random changes. This can be detected by changes in the shape of the envelope spectrum and background level increases on frequencies below the shaft rotation speed. Such increases of background level in the envelope spectrum should be considered as an effective symptom of the bearing shell wear. An example of such an envelope spectrum is presented on figure 3 (d).

The use of the above diagnostic symptoms allow the fluid film bearings diagnostics without needing a relative displacement transducer installation in the bearing unit. At the same time, these methods do not eliminate the problems stated above, namely, detection of the defect severity and distinguishing between defective bearings and good bearings where the oil wedge moves with the shaft rotation on the bearing shell due to the bearing design peculiarities or shaft faults.

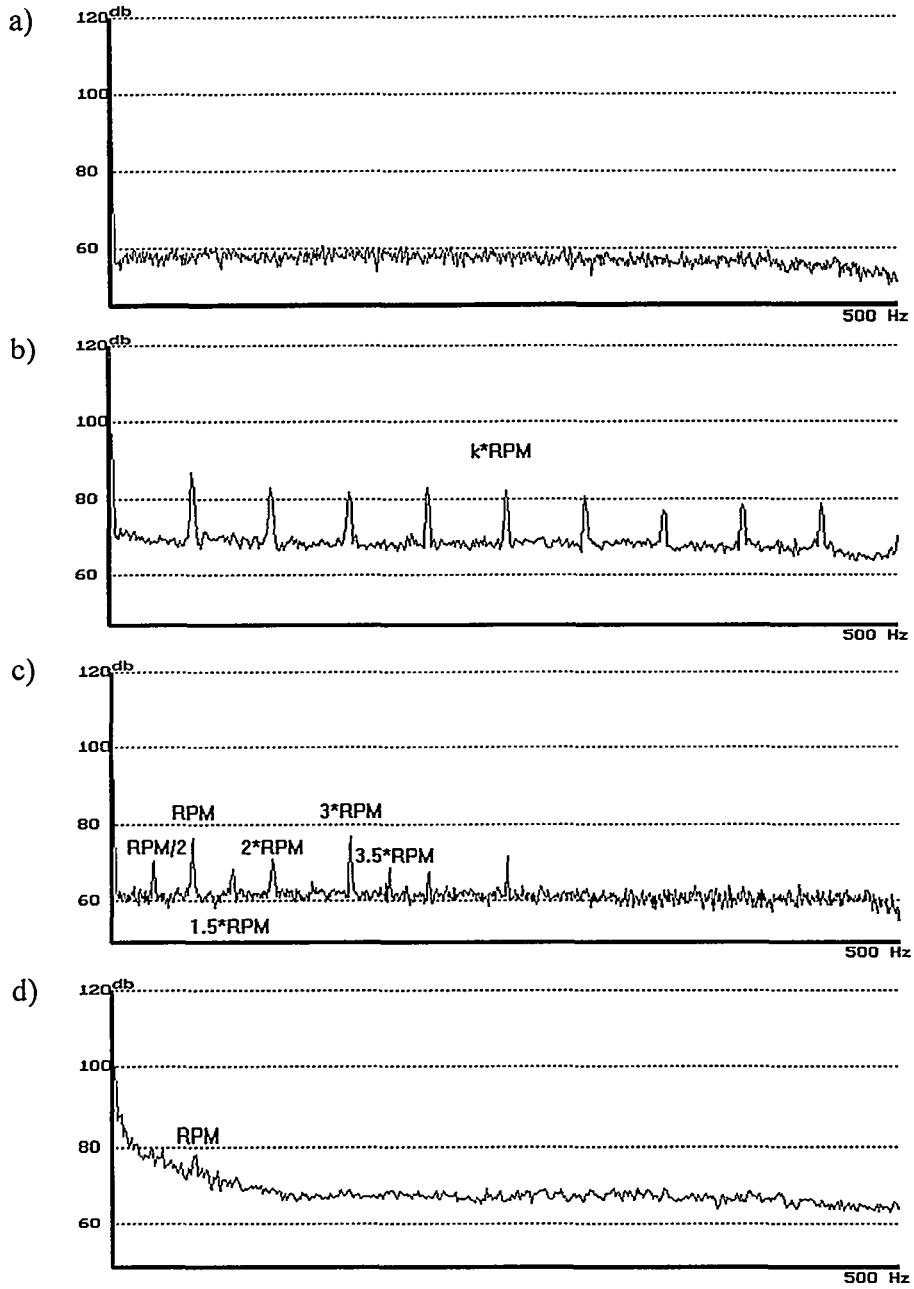


Figure 3. Examples of envelope spectra of journal bearings high frequency vibration.

a) A good bearing.
b) Weared bearing with "shock pulses".
c) Bearing with self-sustained oscillations.
d) Bearing with weared shell.

3. DIAGNOSTIC MEASUREMENTS PLANNING.

The following problems should be considered when the diagnostic measurements are planned:

- Eliminate or decrease the possibility of missing a dangerous defect development.
- Make possible the detection of defects during the incipient stage of their development.
- Allow for division of the defects found into groups according to their development rate.
- Assess the defect severity for the to determine the frequency of additional measurements.

In the case of rolling element bearings, all the above problems can be successfully solved by the analysis of envelope spectrum of high frequency vibration measured on the bearing housing. In this case, no other measurements of vibration or other parameters are needed. The only concern is to comply with the periodicity of measurements. The intervals between measurements can be rather long, about a few months when the service life of the bearing is 2 to 3 years and when the machine is operated in its standard modes. Possible cases of machine operation that should be considered differently occur when there are shock loads on the machine and overheating. If the above conditions are met, it is enough to make 10 to 20 measurements of the high frequency vibration envelope spectra during the whole service life of a bearing to eliminate its un-predicted failure [1].

The levels for defects that are used for the estimation of the detected defect severity in rolling element bearings can be similar for all types of machines and bearings. They monotonously increase with the increase of rotation speed and bearing dimensions. For example, the levels for the bearing dimensions from some millimeters up to a few meters and rotation speeds from a dozen of revolutions per minute up to hundreds of Hertz differ by just 3 to 5 times. According to the importance of a particular machine, the customer can correct the alarm levels. According to our experience with customers in from different industries, such corrections typically do not exceed 2 to 3 times.

A number of different problems occur in the diagnostics of fluid film bearings using an envelope spectrum of high frequency vibration. These problems are solved by the added analysis of diagnostic information from the autospectra of bearing housing vibration.

The first problem is connected to the design characteristics of some machines where the shaft motion oscillates in the stationary bearing shell, even with no defects present. In this case, an envelope spectrum at this bearing will contain harmonic components proportional to the rotation frequency in spite of the absence of defect. The characteristics of these components can undergo spasmodic changes during machine operation and defects in the initial stage of their development may not be detected. In this case, the amplitude changes of these components should be analyzed by using both autospectra and envelope spectra in parallel. This significantly increases the probability of incipient defect detection. The vibration measurements, in turn, should be done more often, i.e. with time intervals of about 1 month or less.

The second problem is closely connected with the variety of fluid film bearings designs, each of which has its own lubrication layer thickness. For this reason, similar defects in the bearings with different lubrication layer thickness lead to different changes in the normalized thickness and thus, different modulation of random vibration. The introduction of corrections for the lubrication layer thickness in the defect levels calculation may not always provide good results. The best solution is to correct levels for defects according to the data derived from the analysis of autospectra, i.e. the increase of vibration components after the detection of severe defects in similar bearings of other machines.

For the reasons discussed above which are different from the rolling element bearing case, fluid film bearings are best diagnosed by the parallel analysis of autospectra and envelope spectra of the bearing shield vibration.

CONCLUSIONS

Comparative analysis of possibilities for rolling element and fluid film bearings diagnostics by the analysis of high frequency vibration envelope spectra allows the following conclusions:

1. Vibration diagnostics is the most efficient method for detection and identification of incipient defects in rolling element and fluid film bearings. A accumulation experience acquired during many years diagnostics of rolling element bearings and during several years of broad investigations on similar type diagnostics of fluid film bearings makes practical similar diagnostic techniques for fluid film bearings.
2. The typical interval between measurements for rolling element bearings in absence of detected defects is several months and for fluid film bearings is about 1 month.
3. The levels for the detection of defects in rolling element bearings and estimation of their severity are nearly independent on the machine and bearing design. Only some small dependence of levels for defects on the rotation speed and bearing dimensions exists. For machines with fluid film bearings, the levels for defects also depend on the bearing design.

REFERENCES

1. A. Barkov, N. Barkova, J. Mitchel, "Condition Assessment and Life Prediction of Rolling Element Bearings", *Sound & Vibration*, 1995, June pp.10-17, September, pp.27-31.
2. A. Barkov, N. Barkova, "Diagnostics of Gearings and Geared Couplings Using Envelope Spectrum Methods", *Proceedings of the 20th Annual Meeting of the Vibration Institute*, Saint Louis, Missouri, USA, 1996, pp. 75-83
3. A. Azovtsev, A. Barkov, D. Carter, "Improving the accuracy of Rolling Element Bearing Condition Assessment", *Proceedings of the 20th Annual Meeting of the Vibration Institute*, Saint Louis, Missouri, USA, 1996, pp. 27-30.
4. A. Azovtsev, A. Barkov, D. Carter, "Journal Bearing Diagnostics by the Envelope Spectra", *Accepted for publication in the 21st Annual Meeting on the Vibration Institute*. June 1997.

SIMULATION OF FAULTS IN ROTATING MACHINES

Erkki Jantunen, Kirsi Vähä-Pietilä, Jari Halme, Kai Katajamäki, Juha Virtanen & Topi Volkov
Technical Research Centre of Finland, Manufacturing Technology, Espoo, Finland
Phone: +358 9 4561, Fax: +358 9 460 627

ABSTRACT

In order to overcome the well-known restriction in the development of expert systems based on neural networks (namely the availability of training data of sufficient quantity and quality), a four-phased data production analysis procedure for training has been developed. The four interlinked approaches for data collection are simple fault simulation, finite element method simulation, test rig simulation and real data gathering. This approach makes it possible to vary such parameters as type and structure of the machinery in question, together with the operational parameters such as running speed, load and environmental factors. The models are optimized to ensure that the data is representative of real machine vibrations, both in healthy and faulty condition. The data is then used to develop diagnostic modules, utilising neural networks, expert systems and other techniques, and to provide diagnostic advice on the health of the machine.

1 INTRODUCTION

The objective of this work is to develop an intelligent, adaptive monitoring and diagnostic system, based on artificial intelligence, which will analyse vibration data in order to sustain a high level of equipment reliability. The core of the development work is simulation, which in the context of this project means the development of mathematical models of specified plant

configurations. The simulation model is verified using data acquired from a test rig and actual machines. The simulation model is optimized to give output as close as possible to the observed data. An intelligent diagnosis system based on neural networks and other techniques is being developed using optimized simulation data together with real data from the test rig and real-world plants at end-user sites. The work is being conducted as part of the VISION project in the Brite EuRam initiative of the European Union (Project No. BE95-1313), and descriptions of this project have been given in the literature [1,2].

2 DEVELOPMENT PROCEDURE

In order to overcome the well-known restriction in the development of expert systems based on neural networks (namely the availability of training data of sufficient quantity and quality), a four-phased data production analysis procedure for training has been developed. The four interlinked approaches for data collection are simple fault simulation, finite element method (FEM) simulation, test rig simulation and real data gathering. With this approach, the amount of data can be extended to fulfil the strictest requirements of a neural network capable of handling almost all-rotating machinery at the user sites. The approach adopted makes it possible to vary such parameters as type and structure of the machinery in question, together with the operational parameters such as running speed, load and environmental factors [3]. The data flow scheme during the development process is shown in Figure 1. In this approach data from the simple fault simulator, FEM model, test rig and real world can be combined and analysed in one module and then treated in the same way when training the neural network. This approach overcomes the problem of gathering a sufficient amount of data from faulty machinery in the field, which for the defined fault and machinery scenarios would take hundreds of years to be adequate for training.

All the data from different sources is stored in a database. The information in the database identifies the machinery, the operation and the measurement connected to calculated parameters (Figure 2). In the development phase of the expert system, the number of calculated parameters is large so as to find out the most appropriate parameters. The neural networks are trained using these parameters.

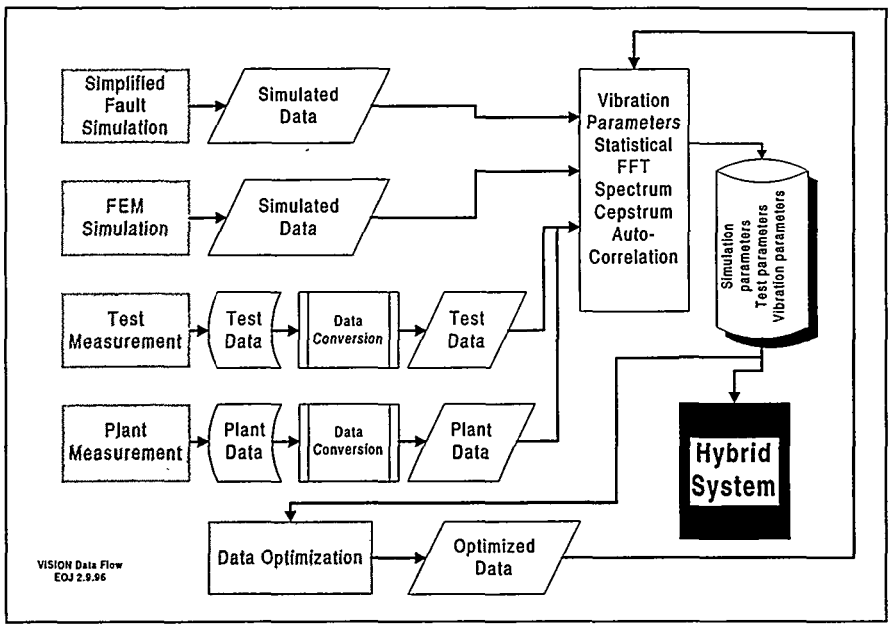


Figure 1. Data flow during the development process.

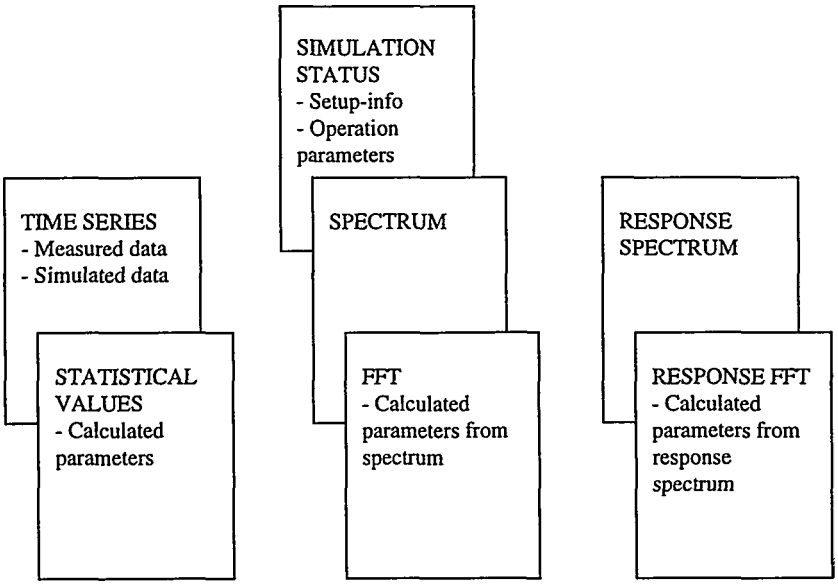


Figure 2. Structure of the database.

3 TEST RIG SIMULATION

In order to acquire data in controlled fault conditions, a small test rig has been built. The test rig consists of a motor, a brake and a shaft with two bearings. The test rig has been used for data collection for fault types such as unbalance, misalignment and different types of bearing faults. All these fault types have been studied at different stages of fault severity. The data from the test rig gives the basic information for the development of the hybrid expert system. Because the data gathered from the test rig is limited for training neural networks, even after running the test rig 24 hour per day, further steps for data development were essential. It should also be noted that one test rig only represents one specific machine, and its structure could not be varied so as to reveal the effect of changes in structural stiffness and consequent natural modes of the vibration response.

4 SIMPLE FAULT SIMULATOR

As a first phase of the simulation process a simple fault simulator (SFS) has been developed. The user interface of the SFS is shown in Figure 3. The SFS can be used to predict the characteristic features of vibration acceleration signals for fault types such as unbalance, misalignment and different types of bearing faults. The SFS can also introduce noise to the signals due to various sources, i.e. the effects of other machinery and amplifier/transducer performance. The amplitude levels of both the predicted fault signals and noise can be adjusted in the SFS to match the real fault severity case. This is done with an automatic fault sequence generator module, where the amplitude of each fault type can be changed one after another. In the fault sequence module, the number of different fault severity cases can be multiplied in a short time. The outputs of SFS can then be used as inputs to the neural network, which needs a great number of different fault severity cases to fulfil its training needs.

Linked to the SFS a module has been developed for the calculation of several parameters, which are used in practice for analysis, and for the diagnosis of the signals. Some of the parameters are statistical, such as rms, max, min, standard deviation, mean deviation, skewness and kurtosis, and they are calculated from the time-domain data. The rest of these parameters are based on more sophisticated analysis methods such as Fast Fourier Transform (FFT) in order to be able to define the

frequency content of the signals. A summary of the different parameter sets is shown in Table 1. The total number of calculated and tested parameters is more than 500 for each set of data. According to the principle shown in Figure 1 these parameters can be calculated for data from the test rig, simple fault simulator, FEM model and real data from the industrial end-users' plant.

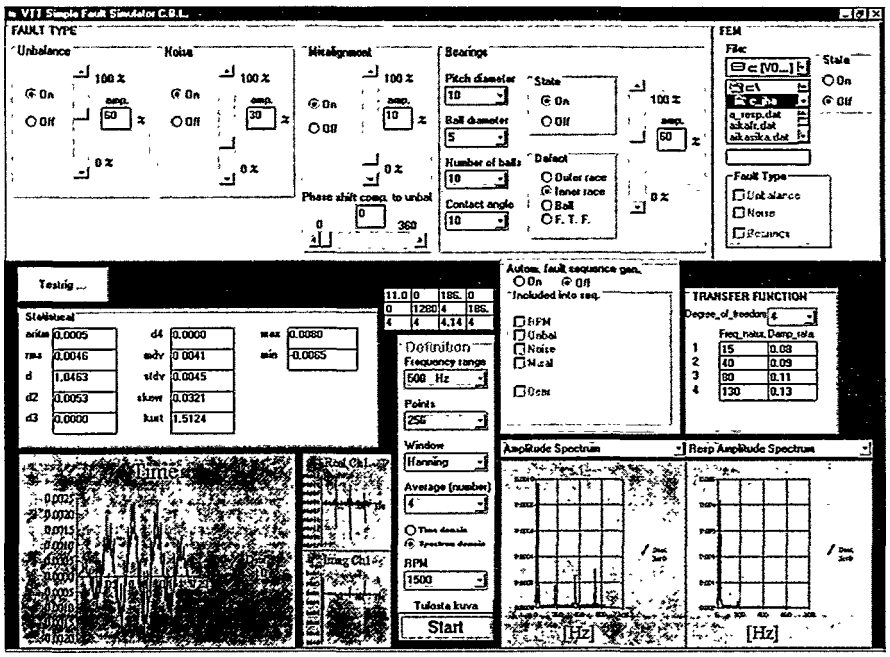


Figure 3. Simple fault simulator

Table 1. Calculated parameter set types

Parameter set	Frequency range [Hz]	Signal type
statistical	0 - 1 000 0 - 10 000	acceleration velocity displacement
envelope	500 - 1 000 2 000 - 4 000 4 000 - 10 000	acceleration
spectrum	0 - 1 000 0 - 10 000	acceleration
cepstrum	0 - 1 000 0 - 10 000	acceleration
autocorrelation	0 - 1 000 0 - 10 000	acceleration

5 FEM SIMULATION

The great advantage of the FEM simulation is that it presents the possibility of making structural changes to the machine with minimum effort, especially if the model has been built for this purpose. In this case, the model has been parameterized using super-element techniques in order to keep the number of degrees of freedom of the model at a reasonable level. The model basically represents the test rig but offers the capability of making dimensional changes, such as shaft diameter variation and bearing distance variation, thus introducing to the response the effect of transfer functions due to the variation of natural frequencies. Figure 4 shows a detail of the model with the positioning of the transducers in the test rig indicated. Compared to the test rig the FEM model has produced two orders of magnitude more vibration data, which, since the model has been verified and optimized against the measured results from the test rig, can be considered very reliable and suitable for training purposes.

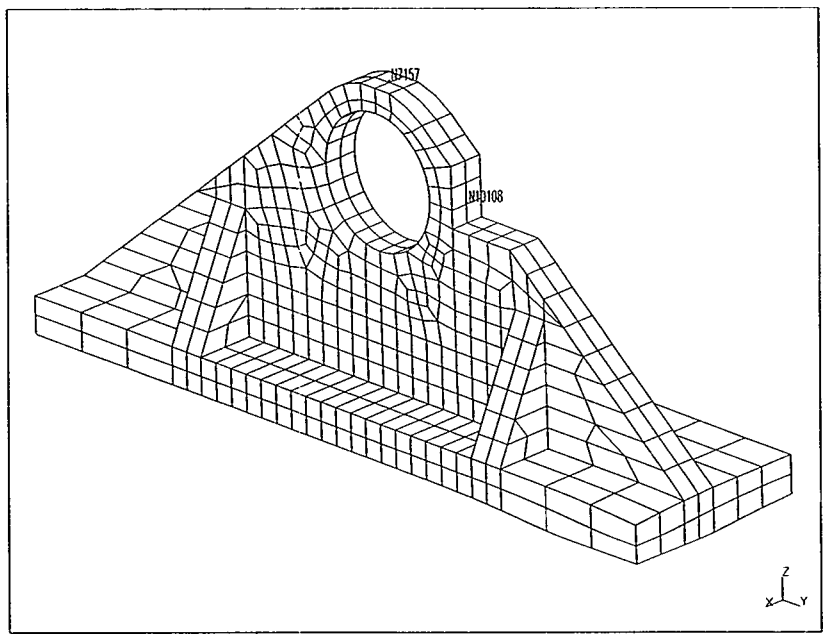


Figure 4. Finite element model of the bearing supports of the test rig

The shaft is modelled with beam, 3d-solid, point mass, spring and dashpot elements. The common frame, which holds the shafting

system, is assumed to be infinitely stiff. The stiffness of the bearings was modelled with spring elements located radially from the shaft to the bearing bracket. Each spring describes the stiffness of one bearing element. The spring elements act between two nodes, with its line of action being the line joining the two nodes. The spring behaviour is non-linear, i.e. the force is assumed to be a non-linear function of relative displacement in the spring. The damping in the bearings was modelled with dashpot elements, which were situated at the same locations as the spring elements. The force versus velocity relation of these elements is non-linear.

The couplings were modelled as frictionless, radially stiff joints having the displacements at the two sides of the joint identical and the angles different, except the axial rotation, which is the same for the two halves.

The effect of lubrication has also been modelled. The differential equation, which is valid for this lubrication situation, has been discretized over a lubrication domain in order to derive proper element equations [4]. In this case, a special user-defined two-dimensional element has been created, which takes into account the effect of lubrication. The element has been coded so that elastic deformations and viscosity changes are included. In addition, time-dependent effects have been taken into account, in order to analyse both stationary and dynamic behaviour of the lubricant.

The lubrication phenomenon is very complex, and accurate modelling requires a large and dense FEM model [5]. In order to keep the analysis times at a reasonable level a simplified procedure must be used. A separate lubrication FEM model is used to analyse the effect of lubrication in one rolling element contact. The results of the lubrication model are used in order to get non-linear stiffness and damping coefficients of the lubricated contact. These coefficients can then be used as parameters for conventional spring and damper elements. In the test rig model, every rolling element contact has been replaced with the combination of a non-linear spring and damper element. The principle of this simplified modelling technique is shown in Figure 5.

To reduce the calculation time super-elements are used. Altogether four super-elements were created, describing the roller and ball bearing brackets and the 3-dimensional idealisation of the shaft in the area of the bearings. Thus the beam elements with concentrated mass

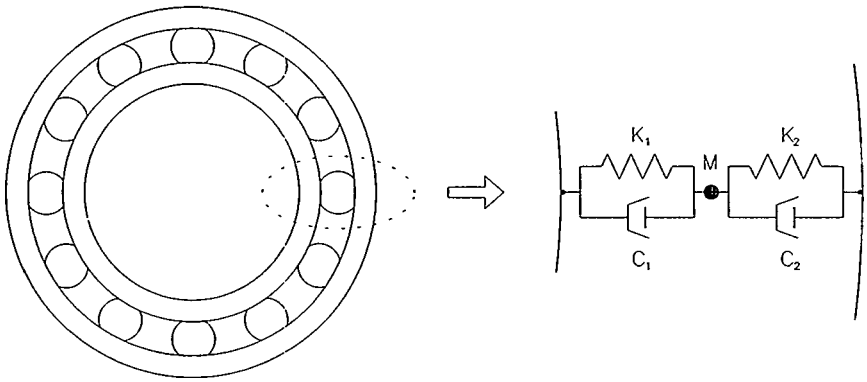


Figure 5. Simplified lubrication contact in rolling element bearings

point elements, together with the spring and dashpot elements describing the bearings, are left active in the model. The super-elements for the bearing bracket consisted each of about 1000 elements and 5500 nodes leading to about 20000 equations for the system matrix. The other two super-elements were much smaller, having about 200 elements and 1000 nodes, and the number of degrees of freedom being about 3000. Based on the above, the final calculation model only needs to have about 240 elements and 200 nodes.

The I-DEAS [6] program has been used in the modelling. The analysis runs have been made with the ABAQUS/Standard general-purpose finite element analysis program [7]. The calculation time for one simulation is about one CPU-hour when using Silicon Graphics Indigo2 workstation. In this case, the dynamic integration scheme consists of 7000 time steps.

6 SIMULATED FAULTS

Mass unbalance is modelled with rotating forces. Mass unbalance occurs when the geometric centre and the mass centre of the shaft do not coincide, and is therefore a once-per-revolution (frequency of rotor speed) fault. Mass unbalance has a fixed phase angle with respect to the shaft and a sinusoidal response is typically found at the running rpm.

Misalignment in a redundantly supported shaft (i.e. a shaft with three or more radially loaded bearings) causes a rotating preload in the bearings, shaft and couplings. The load varies at the frequency of shaft speed; the magnitude of the resulting vibration being dependent on the stiffness of the shafting.

In well-lubricated rolling contacts, such as those in rolling bearings, there is no progressive visible wear due to adhesion or abrasion, but component life is limited by fatigue. Such rolling contact fatigue is characterized by the formation of large wear fragments after a critical number of revolutions. In this approach it is assumed that there are some imperfections in the bearing due to fatigue or some outside source such as hard abrasive particles. These have caused a fault either on the outer or inner race or on the rolling element. These faults, although very small in the beginning, are responsible for impacts, which in turn are responsible for progressive development of fatigue [8].

All the above-described fault types, together with lubrication problems, have been simulated using all of the three simulation approaches, i.e. test rig, FEM and SFS. In total it has been possible with these three tools to simulate about a quarter of a million fault scenarios, i.e. combinations of different faults at different stages with different loads (load, running speed) of the test machinery.

7 CONCLUSION

The work outlined in this paper is developing a new approach to diagnostic systems for vibration data. The key to the success of this approach, and its innovation, is in the use of simulation techniques to validate the diagnostic models developed, and in the use of multiple diagnostic techniques to give a hybrid approach which is more reliable and robust than any single technique. In order to overcome the well-known restriction in the development of expert systems based on neural networks (namely the availability of training data of sufficient quantity and quality), a four-phased data production analysis procedure for training has been developed. The four interlinked approaches for data collection are simple fault simulation, finite element method (FEM) simulation, test rig simulation and real data gathering. With this approach, the amount of data can be extended to fulfil the strictest requirements of a neural network capable of handling almost all rotating machinery at the user sites.

ACKNOWLEDGEMENTS

The efforts, enthusiasm and excellent co-operation of the project teams at Monition Ltd (Ian Jennings, Simon Elliott, Jason Flaxman and Don Heathcote), the University of Sunderland (John MacIntyre, Adam Adgar and Ken McGarry) are gratefully acknowledged. The financial support of the EU, which has made this project a reality, is also gratefully acknowledged.

REFERENCES

1. MacIntyre, J. and Jennings, I., VISION - Using Simulation and Neural Networks for Vibration Interpretation in the Areas of Condition Monitoring and Fault Diagnosis, proceedings of EUREKA MAINE Conference, London, December 1996.
2. MacIntyre, J. and Jennings, I., Condition Monitoring Software that Thinks for Itself, proceedings of MAINTEC 97 Conference, Birmingham, March 1997.
3. Jantunen, E.O. and Poikonen, A.O., Dynamics in Monitoring of gear faults, The Systems Engineering Approach To Mechanical Failure Prevention, Proceedings of the 47TH Meeting of the Mechanical Failures Prevention Group, Virginia, USA, April 13-15, 1993, pp. 279 - 288
4. Huebner K. H. and Thornton E. A., The Finite Element Method for Engineers, 2nd edition, John Wiley & Sons, 1982
5. Hamrock B.J., Fundamentals of Fluid Film Lubrication, McGraw-Hill Inc., 1994
6. I-DEAS, Master Series 4. Structural Dynamics Research Corporation (SDRC).
7. ABAQUS/Standard Version 5.6-1. Hibbit, Karlsson & Sorensen, Inc.
8. Arnell, R. D., Davies, P.B., Halling, J. & Whomes, T. L. Tribology, Principles and Design Applications. London: Macmillan, 1991. 255 p.

CURRENT STATE OF PASSIVE AND ACTIVE VIBRATION CONTROL BY CONFINEMENT

Daryoush Allaei[1], R & D Department, QRDC, Inc., P.O. Box 562, Excelsior, MN 55331-0562, U.S.A.

ABSTRACT

The purpose of this paper is to present an overview of the advances in the fields of localized vibrations and confinement by presenting a brief summary of the past, present, and their future directions. Both basic and applied research papers are discussed and a comprehensive bibliography is presented. Based on a review of the open literature, the first series of articles on the occurrence of the mode localization phenomenon were published in late 1950's during which the phenomenon was explored in the area of solid-state physics. Small irregularities in a periodic lattice were reported to cause localization of energy. In early 1970's, almost twenty years later, the phenomenon was reported to occur in periodic or cyclic engineering structures when they were subjected to small aberrations causing them to become nearly periodic disordered structures. Even though the applicability of the mode localization to vibration control was discussed in mid 1980's, it was not until early 1990's when the phenomenon was identified as an alternative or complementary method to be incorporated with the current passive and active vibration control techniques. It is shown that the concept of vibration control by confinement can overcome some of the shortcomings in the present passive and active noise and vibration control systems.

1 INTRODUCTION

In recent years, it has been shown that the concept of the mode localization phenomenon and vibrational energy confinement can play an important role in various engineering applications such as passive/active noise and vibration control. For example, today's common practices to control structural vibrations are based on the use of a set of actuators with appropriate time-varying magnitudes which are usually controlled based on the displacement, velocity, and/or acceleration vectors measured over the span of the structure. Though sensors, actuators, signal processing and control algorithms, data acquisition systems, and computers have reached a state such that sophisticated active vibration control systems are at least possible, the question of which type of actuator (or even sensor) to use is still vexing. In many cases, structures deal with large forces, accelerations, velocities, and displacements and therefore, there are questions as how to generate counteracting forces efficiently, reliably, and at acceptable financial and energy costs. The actuators must provide a sufficient amount of damping for energy dissipation. In spite of such a large effort in the research community, the full potential of active vibration and noise control systems has not yet been realized because of the limitations of available sensors, actuators, affordable power supplies, and efficient control algorithms that can optimally manipulate the collected data and issue actuation assignments to achieve the required performance. Two of the main reasons why the implementation of active noise/vibration systems has been relatively slow are their energy consumption and high price. In other words, it is desired to solve vibration isolation and suppression problems without requiring unrealistic hardware performance or costs which outweigh the benefits. Recently, it has been shown that an approach based on the mode localization phenomenon and/or vibrational energy confinement [51,57-59,68-74,77,79,81-85,92,94-96] will address these issues and significantly improve the overall benefits of the vibration control systems with the least power, weight, and cost penalty.

[1] The author can also be reached at da-qrdc@msn.com or http://www.mntech.org/mcdc/qrdc.htm.

The purpose of this paper is to provide a review of the recent basic and applied work on the mode localization phenomenon and vibrational control by confinement. Ninety-six papers [1-96] on the subject of localized vibrations are reviewed. Even though, one can group the papers according to their content based on whether the work is basic or applied, passive or active, and linear or non-linear, the author has chosen to sort them chronologically. The reviewed papers are therefore divided into four categories: (1) - 1970 which includes all the reports before 1970; (2) 1970-80; (3) 1980-90; and (4) 1990-present. All the reviewed papers are listed in the list of references.

2 PAPERS BEFORE 1970

Anderson's paper [1] has been recognized as the first published report on mode localization (or Anderson's Localization). In his paper, he studied processes such as spin diffusion or conduction in the impurity band by utilizing a simple model of cellular disorder. His mathematical model was used to prove the theorem that transport does not take place if the density is sufficiently low and wave functions are localized in a small region of space; localization of the mode shapes reduces the metallic conductivity. In other words, solids with regular crystal lattice become semiconductors when imperfection is introduced which results in localization of the eigenvectors. Anderson's analysis was based on the perturbation method. Even though his work was focused on nonmechanical systems, his findings have played a significant role in the understanding of the phenomena in vibrating structures which undergo a similar localization. The confinement of vibration energy in localized structures is similar to reduction of metallic conductivity in solid state physics.

In the 1960's, intensive research [2-4] began on disordered lattices. With the exception of reference [2], most of the reported work was restricted to one-dimensional models and they were quantitative rather than analytical. The interests of the authors were in the determination of the distribution of natural frequencies and the localized mode shapes. The paper published by Ewins [4] in 1969 was concerned with the detuning effects upon the vibrations of bladed disks. He showed that the vibrations in a particular pair of modes depend on the split of the corresponding natural frequencies (curve veering) for given excitation and damping conditions. He also concluded that the presence of small blade imperfections, which are typically found in gas turbines, can increase the resonant stress levels up to 20% above the optimum value. He confirmed his theoretical findings by an experimental study. Since his paper doesn't include any representation of mode shapes, it isn't clear if he observed the mode localization occurring in the bladed disks system.

All of the papers published before 1970 report the occurrence of curve veering and localized modes, and their effects on the vibration response. No consideration was given to the development of a method to predict the occurrence and/or severity of the phenomena, or incorporating it in the design process. However, the authors provided a foundation for further research in the filed.

3 PAPERS PUBLISHED IN 1970-1979

In 1971, in a letter to the editor, Webster and Warburton [5] proved that the previously reported loci crossing of stiffened, curved plates [6] are in fact loci veering. They showed that the first two symmetric-symmetric mode shapes of clamped square panels veer away in a certain range of the dimensionless parameter ab/Rh (a=width, b=length, h=thickness, and R=radius of curvature). The significance of their paper was to show that loci crossing may be the result of the errors in the approximate solution rather than actually occurring in the system. Therefore, one has to be cautious in using a particular numerical method for studying these phenomena.

Nair and Durvasula [7,8] published the results of their extensive study on the free vibrations of skew plates with different edge conditions in 1973. Based on the Ritz variational method, they employed a double series of beam functions to formulate the

problem. They were interested in providing design information with regard to the vibration characteristics of skew plates. By plotting the frequency curves versus the skew angle of the plates, they showed that curve veering is possible when the frequencies belong to the same symmetry group and loci crossing is possible when the modes belong to the opposite symmetry group. The occurrence of loci veering and crossing was determined by the values of the skew angle and the boundary conditions. The representation of their eigenvectors indicate that the mode shapes undergo drastic changes and transition of modes takes place from higher to lower frequency (and vice versa). Also, from their results, one can conclude that the presence of loci veering does not necessary mean that the modes are localized. Such rapid changes in the mode shapes are very important because they are likely to alter the flutter characteristics of a dynamic system. They also pointed out that another implication of the occurrence of these phenomena is that the assumed mode shapes in numerical methods are not valid in the vicinity of loci veering or crossing. The latter could introduce a large error in numerical values generated by computer models.

In 1974, Leissa [9] disputed the reported loci veering in the free vibration characteristics of plates. He concluded that the veering of the frequency curves could be the result of the approximations in the analysis and they were not the actual characteristics of the structure. Furthermore, he pointed out that such errors may be due to the method of analysis, the selected numerical calculations, and/or more importantly the inadequacy of the mathematical model of the real world. Kuttler and Sigillito [14,15] showed that the loci veering really does occur in many cases, it can be an actual phenomena of the mathematical model, and it is not necessarily the result of the errors due to the approximation methods.

Free vibration of disordered periodic beams was reported by Lin and Yang [10] in 1974. They considered two types of imperfections: random deviation of span length and bending stiffness along the beam. Their method of analysis was based on a first-order perturbation procedure. Mode localization was observed when the imperfection was due to variation in the span length. The effect of random bending stiffness was found to be unimportant in the case of multispan beams. More research is needed in order to make a general statement on the role of the bending stiffness variation on the free vibration characteristics of multispan beams.

In the mid 1970's, a set of papers [11-13], on the subject of the localization of the modes of vibration in disordered linear chains, were published by physicists. Their method of analysis was based on the Green's functions and statistical techniques. The main objectives of the authors were to introduce imperfections, such as varying mass ratios and concentrations, in the lattices and to identify and examine two parameters indicating the measure of localization. They formulated the spatial length and the average decay parameters. The former was defined as the length over which the eigenstate had significant magnitude and the latter was defined as the exponential decay length away from the region of appreciable magnitude. Even though these parameters were proven to be effective in certain cases, they failed in accurately indicating the degree of localization in other cases. In particular, they were not applicable to cases when the modes have multiple localization. Loci veering in curved and twisted cylinders transporting fluids was reported, for the first time, by Doll and Mote [13]. They pointed out that such loci veering might have been the result of the discretization errors. The mode shapes were not addressed in their paper. A more detailed study of the free vibration characteristics of the above problem needs to be conducted in order to draw any general conclusion.

4 PAPERS PUBLISHED IN 1980-1989

The number of publications addressing these phenomena in structural dynamics drastically increased in 1980. References [14-41] represent the reports found as a result of our search. Hodges [16] reported an excellent study on the localized vibration in irregular structures. He addressed the importance of the mode localization in vibration

analysis of nonuniform structures and damping measurement. He also showed that confinement of vibration has effects on reflection of waves in strings and beams. In a latter paper, Hodges and Woodhouse [20] reported the results of their study on the effects of irregularities in nearly periodic structures on the isolation of vibrations. Papers published by Huang [17], Kaza and Keilb [18], Stange and MacBain [19], and Valero and Bendiksen [25] were concerned with the effects of irregularities and localization of vibration in turbine blades. Nonuniform ring structures were studied by Schajer [21], and Allaei et al [23,24,27,28,35]. Loci veering and crossing and mode transition were reported in all cases. Allaei's method of analysis was based on exact solution of the ring and the receptance method. Both stiffness and mass irregularities were considered. Similar analysis was also applied to nonuniform tires [33,36,41,51,52]. In the latter case, the finite element model of the tire was coupled with the receptance method. The phenomena was also observed in nonuniform tires. They showed that not only the magnitude of the irregularities, but also their relative locations have significant effects on the occurrence of these phenomena.

The question of whether loci veering is due to the approximation in discretized models or it actually occurs in continuous systems was once again addressed in a paper by Perkins and Mote [22]. They verified the existence of loci veering in continuous models by examining the exact solution of a rotating, guided, circular ring. They stated that loci veering reflects the coupling between the mode shapes. In other words, the frequency curves cross when there is no coupling between the eigenfunctions. Furthermore a criteria, based on two coupling factors, for differentiating veering from crossing was presented. The criteria was successfully tested on a few example problems.

Mode localization was investigated in large space structures by Bendiksen [26] and Cornwell and Bendiksen [39]. Their results indicated that localized vibration is most likely to occur in structures that have a large number of weakly coupled subelements. He recommended that these phenomenon should be considered in designing the control systems. Papers by Pierre, et al [30,32,37,40] confirmed Bendiksen's conclusions. They concluded that structures with nearly identical substructures have strong localization if the coupling frequency between the substructures is of the order or smaller than the spread in natural frequencies of the substructures. Pierre, Tang, and Dowell [30] introduced the ratio of the peak deflections in order to predict the degree of localization due to the mistuning parameter in irregular multispan beams. The indicator varied from -1 to +1 with 0 representing decoupled modes. When the absolute value of the ratio were less than 10%, localization was present. Allaei and Shih [29, 34] reported similar results for the case of an annular circular plate with ring stiffeners and dynamic absorbers. Their paper provided a detailed numerical and graphical representation of loci veering, crossing, mode localization, and transition. Crandall and Yeh [38] presented the vibration analysis of rotor and stator systems. They used an automated component modes procedure for substructures modeled as Timoshenko beam elements connected to other parts by bearings and couplings. Loci veering and crossing was observed in their reported frequency curves versus the rotational speed of the rotor. This was the only paper which has shown that the occurrence of the phenomena depends on the angular speed of the rotor and not the irregularities. As a matter of fact, they did not even consider the presence of irregularities in the system. No mode shapes were presented in their paper. More work needs to be done on this problem in order to better understand the dependency of the rotational speed on mode localization.

Ibrahim [31] reviewed a number of topics related to dynamics of structures with parametric uncertainties. He also addressed the impact of these studies on areas such as sensitivity of structural performance and design optimization. He concluded that even though the mathematical theory of random eigenvalue has reached its maturity, there is a lack of application of the theory to real engineering problems. The implementation of the localization in design problems is certainly one of the shortcomings in the field.

In summary, during the 1980's the presence of these vibration phenomena was investigated in a wide range of engineering applications. It was shown that irregularities,

rotational speed, distribution of supports, and variation of mechanical properties could be responsible for occurrences of one or all of these phenomena. An indicator of strong localization was also introduced.

5 PAPERS PUBLISHED IN 1990-PRESENT

The paper by Shih and Allaei [42] was the only paper found in 1990. By 1991, more researchers were attracted to study the subject. The objectives of most of the authors who published in 1991-92 have been either to explain the phenomena or to formulate a factor which can predict the degree of localization. New terminologies were also established by a few researchers. However, the work reported in references [51,57-59,68-74,77,79,81-85,92,94-96] addressed the applicability of the mode localization and vibration control by confinement to passive and active vibration control problems.

M. Triantafyllou and G. Triantafyllou [45] used a geometrical description to explain the occurrence of loci crossing and veering which were referred to as frequency coalescence and avoided crossing in their paper. They represented the general eigenvalue problem whose solution can yield the natural frequencies in terms of the system parameters. Loci veering or crossing occurs as one of these system parameters is varied over a certain range. Both the natural frequencies and the system parameters were assumed to be complex. The conditions for an m-order frequency crossing point were defined as the simultaneous solution of the eigenvalue equation and its derivatives with respect to system frequencies and parameters. They showed that frequency crossing and veering can be uniquely characterized in terms of the saddle point in the complex frequency plane and a saddle/branch point in the complex parameter plane. Furthermore, it was indicated that the large frequency and mode shape sensitivity, which could characterize mode localization, could be caused by the nth order branch point singularity of the complex loci crossing. This argument also explains the failure of the perturbation theory in the form of a power series expansion which has been applied to localization phenomena. The sensitivity of the system was measured by the ratio of the imaginary part over the value of the parameter at the loci crossing. They successfully applied their formulation to problems such as a disordered double pendulum, a rotating string with attached damper and spring, and an inclined catenary. Their approach shows good potential for adoption in the design of structures and control systems as was demonstrated in references [57-59].

Khader and Masoud [44] reported loci veering, crossing, and split for the case of mistuned bladed disks supported by flexible continuous shafts which resembles a fan stage of a modern gas turbine engine. Mode shapes were not shown and no discussion of their findings in regard to frequency crossing and veering was presented. Allaei and Labadi [46-51] reported their preliminary study on the occurrence of the phenomena in tires and frames. In their study the controlling parameters were the distribution and magnitude of irregularities around tires and the relative orientation of bars in the frame structures. Jei and Lee [52] examined the existence of the curve veering and abrupt changes in the mode shapes in rotor-bearing systems. They showed that in a certain range of the rotational speed of the rotor the phenomena occurs and modes were drastically but continuously changed.

Cornwell and Bendiksen [53] studied the mode localization in disordered cyclic structures with multi-degree-of-freedom substructures. In particular, they proposed a length scale as a measure of the magnitude of the localization. They showed that the strength of the mode localization obeys a power-law-type relationship in the transition region where loci veering occurs. They claimed that their proposed localization length scale provides insight into the transition region where the analytical approach fails. Chen and Ginsberg [54] investigated the relation between loci veering and parameter sensitivity of mode shapes. They derived a perturbation expansion describing the behavior of self-adjoined eigensystems when the difference between two frequencies was small. They also showed that the mode shapes in the veering region were linear combinations of the corresponding modes at the outer limits of the veering zone. However they failed to explain the significance of their claim.

The papers by Ye et al [55] and Luongo [56] are on the subjects of elastic wave localization and mode localization of one-dimensional continuous structures with imperfections, respectively. Free longitudinal oscillations of a rod with small axial stiffness and continuously restrained by springs was used as an example problem. They showed that only the first modes are localized and their number increases as the ration of imperfection to stiffness of the rod increases. They also showed that the phenomenon is governed by a turning point mathematical problem. Occurrence of mode localization in nonlinear periodic lattices and coupled structures with geometric nonlinearities were reported in references [62 and 93]. The effect of fluid loads on the mode localization phenomenon was reported in [54, 61].

In recent years it has been suggested [46, 51, 57-59, 68-85, 94-96] that the theory of mode localization has the potential to improve the active noise and vibration systems by confining the vibrational energy, reducing the control effort, and optimizing the required sensors and actuators. The vibration localization phenomenon can be brought about passively via design modifications, actively by using feedback control, or hybridly by combining the first two approaches. Keltie and Cray [57, 74, 83] investigated the occurrence of mode localization in stiffened plates and its effects on the low frequency sound radiation of such plates. In particular, their study was focused on the feasibility of using submerged stiffened plates as an efficient low frequency (far below the critical frequency) deep water sound projector. They concluded that passive implementation of mode localization is possible by optimally selecting the material and structural parameters of the stiffening ribs. Furthermore, they concluded that such plates can indeed have the potential as enhanced low frequency radiators of sound. Similar study was also reported by Cheng [80]. Passive localization of helical flexural waves of ring-stiffened fluid-loaded cylindrical shells was reported by Photiadis [61]. His results indicated that strong localization effects occur for the helical flexural waves yielding very short localization lengths in the order of a single rib spacing. Furthermore, he concluded that the investigation of the effects of mode localization on cylindrical shells appeared to be very promising. Levine-West conducted a series of experiments on the occurrence of mode localization in space antennas [69]. Based on her findings, she recommended that mode localization has great potential for vibration control in space structures.

Tarnowski et al [79], Yigit and Choura [81], Shelley and Clark [82, 96], and Allaei [84-85, 94-95] have suggested that mode localization can also be implemented via active sensors and actuators. They demonstrated their active localization scheme in rod and beam type structures. They concluded that vibration control by confinement is a viable technique in many real applications.

6 CONCLUSIONS

Table 1 Number of reviewed papers

Period	Number of Publications
- 1959	1
1960 - 1969	3
1970 - 1979	9
1980 - 1989	28
1990 -	55
Total	96

A brief overview of ninety-three published papers and reports was presented in this article. Based on our findings, it appears that the subject has increasingly become more attractive to those conducting basic research and those who are interested in the applicability of the theories. Table 1 shows a systematic increase in the number of published papers. It should be pointed out that the author does not claim that this search has covered all the work on the subject. Those investigations that have not been published or have been kept confidential are not included in his review.

Based on a review of the related literature, the first set of articles on the occurrence of the phenomenon of mode localization were published in late 1950's. The phenomenon was first explored in the area of solid-state physics. Small irregularities in a periodic lattice were reported to cause confinement of energy. In early 1970's, almost twelve years later, the phenomenon was reported to occur in periodic or cyclic structures when they were subjected to small aberrations causing them to become nearly periodic

disordered structures. Even though the applicability of the mode localization to vibration control was discussed in mid 1980's, it was not until early 1990's when the effects of randomly occurring mode localization in periodic structures became a main area of study among researchers. The viability of the mode localization for vibration suppression and vibration control has also been explored by many researchers. It has been shown that occurrence of localized modes and vibrational energy confinement are not limited to disturbed periodic structures, rather they can be passively and/or actively incorporated in any engineering structure. Furthermore, it is evident that the concept of vibration control by confinement can overcome some of the shortcomings in the present active noise and vibration control.

ACKNOWLEDGMENT

The author's research on the subject has been supported by various DOD components during the last five years. Their continuous support of his efforts in this area is greatly appreciated. The suggestions and editorial comments made by Mr. David Tarnowski and Ms. Shohreh Pirzad have been very helpful in the preparation of the article.

REFERENCES

1. Anderson, P. "Absence of Diffusion in Certain Random Lattices," Physical Rev., 109(5), pp 1492-1505, 1958.
2. Rosenstock, H.B. and Mcgill, R.E., "Vibrational Modes of Disordered Linear Chains," J. Math. Physics v3, pp 200-202, 1962.
3. Fox, R.L. and Kapoor, M.P., "Rates of Change of Eigenvalues & Eigenvectors,"AIAA J,V6, pp 2426-2429, 1968.
4. Ewins, D.J., "The Effect of Detuning Upon the Forced Vibrations of Bladed Disks," J. Sound & Vib. 9(1), pp 65-79, 1969.
5. Webster, J.J. and Warburton, G.B.,"Discussion of Approximate Methods for the Determination of the Frequencies of Stiffened and Curved Plates," J.Sound and Vibration 18(1), pp 139-145, 1971.
6. Szechenyi, E., "Approximation Methods for the Determination of the Natural Frequencies of Stiffenes and Curved Plates," J. of Sound and Vibration, 14, pp 401-418, 1971.
7. Nair, P.S. and Durvasula, S., "Vibration of Skew Plates," J. Sound and Vibration 26(1) , pp 1-19, 1973.
8. Nair, P.S. and Durvasula, S., "On Quasi-Degeneracies in Plate Vibration Problems," J. Mech. Sci. V15, pp. 975-986, 1973.
9. Leissa, A.W., "On Curve Veering Aberration," J. Applied Mathematics and Physics. V25 pp 99-111, 1974.
10. Lin,Y.K. and Yang, J.N., "Free Vibration of a Disordered Periodic Beam," J. Applied Mechanics, pp 383-391, 1974.
11. Painter, R.D. and Hartmann,W.M., "Localization of the Vibrational States of Binary Disordered Linear Chains," Physical Review, 12(2) pp 479-491, 1976.
12. Papatriantafillou, C. et al,"Eigenfunctions in One-Dimensional Disordered Systems I & II," Physical Review, 13(2), pp 910-928, 1976.
13. Doll, R.W. and Mote, C.D., "On the Dynamic Analysis of Curved and Twisted Cylinders Transporting Fluids," J. Pressure Vessel Technology v98, pp 143-150, 1976.
14. Kuttler, J.R. and Sigillito,V.G., "Upper and Lower Bounds for the Frequencies of Clamped Orthotropic," J. S. & V. 73, pp 247-259, 1980.
15. Kuttler, J.R. and Sigillito,V.G., "On Curve Veering," J. Sound and Vibration 75(4), pp 585-588, 1981.
16. Hodges, C.H., "Confinement of Vibration by Structural Irregularity," J. S. and V., V82, pp 411-424, 1982.
17. Huang, W.,"Vibration of Some Structures With Periodic Random Parameters," AIAA J. 20(7), pp 1001-1008, 1982.
18. Kaza, K. And Keilb, R., "Flutter & Response of Mistuned Cascade In Incompressible Flow," AIAA20(8), pp1120-1127, 1982.
19. Stange, W. And Macbain, J., "An Investigation of Dual Mode Phenomena in a Mistuned Bladed Disk," J. Vib., Accous., Stress & Reliab. In design 105, pp 402-407, 1983.
20. Hodges, C.H. and Woodhouse, J., "Vibration Isolation From Irregularity in a Nearly Periodic Structure: Theory and Measurement," J. A. Soc. Am. 74(3), pp 894-905, 1983.

21. Schajer, G. , "The Vibration of a Rotating Circular String Subjected to a Fixed Elastic Restraint," J. Sound & Vib. 92(1) pp 11-19, 1984.
22. Perkins, N. And Mote, C., "Comments On Curve Veering in Eigenvalue Problems," J. S. & Vib. 106(3), pp 451-463, 1986.
23. Allaei, D., Soedel, W., and Yang, H.T.Y., "Natural Frequencies and Modes of Rings that Deviate From Perfect Axisymmetry," J. of Sound and Vibration, Vol. 111(1), pp 9-27, 1986.
24. Allaei, D. and Soedel, W., "Documentation of Software Used to Study Vibration of Nonuniform Tires," Report No. 0810-3 HL 86-31p, Ray W. Herrick Laboratories, Purdue University, 1986.
25. Valero, N.A. and Bendiksen,O.O. , "Vibration Characteristics of Mistuned Shrouded Blade Assemblies" J.Eng.For Gas Turbines & Power, 108, pp 293-299, 1986.
26. Bendiksen, O.O., "Mode Localization Phenomena in Large Space Structures",AIAA J, 25(9), 1987.
27. Allaei, D., Soedel, W., and Yang, H.T.Y., "Eigenvalues of Rings With Radial Spring Attachments," Journal of Sound and Vibration, Vol. 121, No. 3, pp 547-561, 1987.
28. Allaei, D., Soedel, W., and Yang, H.T.Y., "Free Vibration of Rolling Structures Carrying Local Non-uniformities by the Receptance Method," Proceeding of the Twentieth Midwestern Mechanics Conference, pp 1347-1353, 1987.
29. Allaei, D., "Vibration Analysis of Circular Plates Using Axisymmetric Finite Element, Proceeding of the Twentieth Midwestern Mechanics Conference, pp 330-336, 1987.
30. Pierre, C., Tang, D.M., and Dowell,E.H., "Localized Vibrations of Disordered Multispan Beams: Theory and Experiment," AIAA J. 25(9), pp 1249-1257, 1987.
31. Ibrahim, R.A.,"Structural Dynamics With Parameter Uncertainties," A. Mech. Rev. 40(3), pp 309-328, 1987.
32. Pierre, C. and Dowell, E.H.,"Localization of Vibrations by Structural Irregularities," J. S. & Vib. 114(3), pp 549-564, 1987.
33. Allaei, D., Soedel, W., and Yang, H.T.Y., "Vibration Analysis of Non-axisymmetric Tires," J. of S. & Vib., Vol. 122, No. 1, pp 11-29, 1988.
34. Allaei, D. And Shih, Y.T.,"Vibration Analysis of Annular Plates Carrying Cocenter Circular Dynamic Elements," Proceeding of the Applied Mechanics and Engineering Science Conference, 1988.
35. Allaei, D., Soedel, W., and Yang, H.T.Y.,"Vibration Characteristics of Rings on Nonuniform Elastic Support," Abstract in 7th ASCE-EMD Specialty Conference, 1988.
36. Allaei, D., Soedel, W., and Yang, H.T.Y., "The Problem of Non-uniformities of Tires," 7th Annual Meeting and Conference on Tire Science and Technology, 1988.
37. Pierre, C., "Mode Localization and Eigenvalue Loci Veering Phenomenon in Disordered Structures,"J. Sound and Vibration, V. 126(3), pp 485-502, 1988.
38. Crandall, S. and Yeh, N., "Automatic Generation of Component Modes for Rotordynamic Substructures,"J. Vib., Accou., Stress & Reliab. In Design 111, pp 6-10, 1989.
39. Cornwell, P. and Bendiksen, O.O., "Localization of Vibrations in Large Space Reflectors," AIAA J. 27(2), pp 219-226, 1989.
40. Pierre, C. and Cha, P., "Strong Mode Localization in Nearly Periodic Disordered Structures," AIAA J. 27(2), 227-241, 1989.
41. Allaei, D., Soedel, W., and Yang, H.T.Y., "Mathematical Models of Stiffness Non-uniformity and Run-Out Effects as Measured on a Low Speed Non-uniformity Machine," Journal of Tire Science and Technology, vol. 17, No. 3, pp 158-174, 1989.
42. Shih, Y.T. and Allaei, D., "On The Free Vibration Characteristics of Annular Plates With Ring-Type Elastic Attachments," J.of S. and Vibration, Vol.140(2), pp 239-257, 1990.
43. Celep, Z. and Guler, K., "Dynamic Response of a Column With Foundation Uplift," J. S. and Vib. 149(2), pp 285-296, 1991.
44. Khader, N. and Masoud,S., "Vibration of Mistuned Bladed Disks Supported by Flexible Continuous Shafts"J. Sound and Vibration 149(3), 471-488, 1991.
45. Triantafyllou,M. & G., "Frequency Coalescence and Mode Localization Phenomena: A Geometric Theory," J. Sound and Vibration 150(3), pp 485-500, 1991.
46. Allaei, D., "Effects of Non-Uniformities in Fasteners on Localized Vibration and Fatigue," 2nd International Conference on Computer Aided Assessment and Control of Localized Damage, Wessex Institute of Tech., Southampton, UK, 1992.
47. Labadi, M. and Allaei, D., "Loci Veering and Mode Localization in Elastic Linkages and Frames," Southeastern Conference on Theoretical and Applied Mechanics (SECTAM XVI), The University of Tennessee Space Institute, Nashville, 1992.
48. Labadi, M. and Allaei, D., "Free and Forced Vibration of the Human Body," SECTAM XVI, Nashville, Tennessee, 1992.

49. Allaei, D. "Mode Localization Due to Tire Non-uniformities," 11th Annual Meeting and Conference on Tire Science and Technology, Akron, Ohio, 1992.
50. Allaei, D., "An Advanced Approach For Vibration Analysis of Nonuniform Tires," 11th Annual Meeting and Conference on Tire Science and Technology, Akron, Ohio, 1992.
51. Allaei, D., "Application of Localized Modes in Vibration Control," 2nd International Congress on Recent Development on Air and Structure Borne Sound and Vibration, Proceedings, 1992.
52. Jei, Y. and Lee, C., "Does Curve Veering Occur in the Eigenvalue Problem of Rotors?" J. Vib. Accou., Stress &Reliability, V. 114, pp 32-36, 1992.
53. Cornwell, P. and Bendiksen, O.O.,"Numerical Study of Vibration Localization in Disordered Cyclic Structures" AIAA J. 30(2), 473-481, 1992.
54. Chen, P.T. and Ginsberg, J.H., "The Effect of Heavy Fluid Loading on Eigenvalue Loci Veering and Mode Localization Phenomena" 2nd Int.Cong. On Recent Development in Air & Structure-Borne Sound and Vibration, 1587-1594, 1992.
55. Ye, L. et al, "Observation of Elastic Wave Localization" 2nd Int.Cong. On Recent Dev. In Air & Structure-Borne Sound and Vib, 1992.
56. Luongo, A., "Mode Localization by Structural Imperfections in One-Dimensional Continuous Systems," J. Sound and Vibration, 155(2), pp 249-271, 1992.
57. Cray, B.A., "Near-Field and Far-Field Sound Radiation From a Line-Driven Fluid-Loaded Infinite Flat Plate having Periodic and Non-Periodic Attached Rib Stiffeners," Submarine Sonar Dep., NUWC-NL Tech. Document 10023, 1992.
58. Allaei, D., "Application of Localized Vibration and Smart Materials in Controlling the Dynamic Response of Structures," Four DARPA Reports, 1992, 93, 94, and 95.
59. Allaei, D., "Application of Localized Vibration and Smart Materials in Controlling the Dynamic Response of Structures, Part II," Two DARPA Reports, 1996 and 97.
60. Happawana, G.S., Bajaj, A.K., and Nwokah, O.D.I., "Eigenvalue Veering, Perturbed Bifurcation Theory and Modal Analysis of Mistuned Linear Systems," ASME Structural Dynamics of Large Scale &Complex Systems, DE-Vol. 59, 1993.
61. Photiadis,D.M., "Localization of Helical Flexural Waves on an Irregular Cylindrical Shell," ASME Structural Dynamics of Large Scale and Complex Systems, DE-Vol. 59, 1993.
62. Vakakis, A.F., King, M.E. and Pearlstein, A.J., "Free and Forced Localization in a Nonlinear Periodic Lattice," ASME Structural Dynamics of Large Scale and Complex Sys, DE-Vol59, 1993.
63. Weaver, R.L., "On Eigenmode Statistics and Power Variances in Randomly Shaped Membranes," ASME Str. Dyn. of Large Scale and Complex Sys, DE-Vol. 59, 1993.
64. Bouzit, D. and Pierre, C., "Localization of Vibration in Disordered Multi-Span Beams With Damping," ASME Str. Dyn. of Large Scale and Complex Systems, DE-Vol. 59, 1993.
65. Lueschen, G.G.G. and Bergman, L.A, "On the Free Vibration of Mono-Coupled Periodic Distributed Parameter Systems," ASME Structural Dynamics of Large Scale and Complex Systems, DE-Vol. 59, 1993.
66. Pabst, U. and Hagedorn, P., "On the Identification of Localized Losses of Stiffness in Structures," ASME Str. Dyn. of Large Scale and Complex Systems, DE-Vol. 59, 1993.
67. Al-Jawi, A.A.N., Pierre, C., and Ulsoy, A.G., "Vibration Localization in Dual-Apan, Axially Moving Beams: Formulation and Results," ASME Structural Dynamics of Large Scale and Complex Systems, DE-Vol. 59, 1993.
68. Allaei, D., "Application of Localized Vibration and Smart Materials in Controlling the Dynamic Response of Structures," International Noise and Vibration Control Conference, St. Petersburg, Russia, 1993.
69. Levine-West, M.B. and Salama, M.A.,"Mode Localization Experiments on a Ribbed Antenna," AIAA J., 31(10), 1993.
70. Allaei, D., "Vibration and Noise Control of Rotating Components Based on Elastic Energy Localization", Third International Congress on Air-and Structure-Borne Sound and Vibration, Montreal, Canada, 1994.
71. Allaei, D., "Vibration and Noise Optimization by Smart Design", Third International Congress on Air-and Structure-Borne Sound and Vibration, Montreal, Canada, 1994.
72. Allaei, D., Jacob, A., Tarnowski, D., "Application of Vibration Localization to Optimize the Service-Free Life of Aircraft Fasteners," Third International Congress on Air- and Structure-Borne Sound and Vibration, Montreal, Canada, 1994.
73. Allaei, D., "Smart by Design ", North American Conference on Smart Structures and Materials, Orlando, Florida, 1994.
74. Cray, B.A., "Acoustic Radiation From Near-Field Periodic and Sectionally Aperiodic Rib-Stiffened Plates," Acoustic Soc. Am., 95(1), 1994.

75. Allaei, D., "A Mathematical Model to Optimize and Predict Service-Free Life of Fasteners," Phase I Final report delivered to the FACTS Office at WPAFB, pp 1-59, 1994.
76. Allaei, D., "The Effects of Non-Uniformities in Fasteners on Localized Vibration and Fatigue", Symposium on Structural Integrity of Fasteners, 1994.
77. Allaei, D., "Performance Comparison Between Vibration Control by Confinement and Conventional Techniques," to be presented Symposium on Active and Hybrid Structural Control, ASME 16th Biennial Conference on Mechanical Vibration and Noise, San Diego, California, 1997.
78. Allaei, D., "A Mathematical Model to Optimize and Predict Service-Free Life of Fasteners," Phase I Option report delivered to the FACTS Office at WPAFB, pp 1-127, 1994.
79. Tarnowski, D. J., Allaei, D., Landin, D., Chen, C., and Liu, S., "An innovative noise and vibration suppression technique based on the combination of vibration control by confinement and layered damping approaches," The 3rd Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu, Hawaii, 1996.
80. Cheng, C.C., "Optimization of Ribbed Plate Vibration," Ph.D. Thesis, North Carolina State Univ., 1994.
81. Yigit, A. and Choura, S., "Vibration Confinement in Flexible Structures via Alternation of Mode Shapes by Using Feedback," J. of Sound and Vibration, 179(4), pp 553-568, 1995.
82. Shelley, F.J. and Clark, W.W., "Active Mode Localization in Distributed Parameter Systems with Consideration of Limited Actuator Placement," 15th ASME Biennial Conf. On Vibration & Noise, 1995.
83. Keltie, R. F. and Cray, B.A., "Low Frequency Sound Radiation from Stiffened Plates," Submitted for publication, J. Acoustic Soc. Am., 1995.
84. Allaei, D., "Passive and Active Vibration Control by Confinement: An Overview," Society of Engineering Science, 3rd Annual Technical Meeting, Tempe, Arizona, 1996.
85. Allaei, D., "Vibration Control by Confinement: An Overview", 4th International Congress on Sound and Vibration, St. Petersburg, Russia, 1996.
86. Pesterev. A.V. and Bergman, L.A, "On Vibrations of a System With an Eigenfrequency Identical to That of One of Its Subsystems," J. Of Vibration and Acoustics,Transactions of the ASME, V. 117, pp 482-487, 1995.
87. Ginsberg, J.H. and Pham, H., "Forced Harmonic Response of a Continuous System Displaying Eigenvalue Veering Phenomena," J. of Vibration and Acoustics, Transactions of the ASME,V. 117, pp 439-444, 1995.
88. Langley, R.S., "Mode Localization Up To High Frequencies In Coupled One-Dimensional," J. Of Sound and Vibration, V. 185(1), pp 79-91, 1995.
89. Al-Jawi, A.A.N., Ulsoy, A.G., and Pierre, C., "Vibration Localization In Band-Wheel Systems: Theory and Experiment," J. of S. and Vibration, V. 179(2), pp 289-312, 1995.
90. Al-Jawi, A.A.N., Pierre, C., and Ulsoy, A.G., "Vibration Localization In Dual-Span Axially Moving Beams, Part I: Formulation and Results," J. of Sound and Vibration, V. 179(2), pp 243-266, 1995.
91. Al-Jawi, A.A.N., Pierre, C., and Ulsoy, A.G., "Vibration Localization In Dual-Span Axially Moving Beams, Part II: Perturbation Analysis," J. of Sound and Vibration, V. 179(2), pp 267-287, 1995.
92. Xie, W.C., "Buckling Mode Localization in Randomly Disordered Mulispan Continuous Beams," AIAA J., V. 33(6), pp 1142-1149, 1995.
93. King, M. and Vakakis, A., "Mode Localization in a System of Coupled Flexible Beams wh Geometric Nonlinearities," ZAMM.Z.angew. Math. Mech. 75(2), pp 127-139, 1995.
94. Allaei, D. and Shih, Y.T., "Influence of Confined Vibration on Sensor and Actuator Optimization," paper 97NV48, SAE Noise & Vibration Conference & Exposition, Traverse City, Michigan, 1997.
95. Allaei, D. and Shih, Y.T., "A comparison between Passive Vibration Control by Confinement and Current Passive Techniques," paper 97NV61, SAE Noise & Vibration Conference & Exposition, Traverse City, Michigan, 1997.
96. Clark, W.W., Sun, F., and Tarnowski, D., "On the Use of Vibration Confinement to Enhance Conventional Vibration Control," to be submitted to *Journal of Vibration and Control,* 1997.

DETECTION AND SEVERITY ASSESSMENT OF BENDING FATIGUE FAILURE IN SPUR GEAR TEETH USING THE CONTINUOUS WAVELET TRANSFORM

I. Yesilyurt and A. D. Ball
Manchester School of Engineering, University of Manchester, Oxford Road, Manchester, M13 9PL, U.K., E-mail: iyesilyurt@fs1.eng.man.ac.uk

ABSTRACT

This paper presents the use of the continuous wavelet transform in the detection and location of weakened gear teeth caused by bending fatigue cracks. Tooth weaknesses were simulated in differing degrees of severity on a real gear tooth by cutting a fine slot into its root. In its most severe form, the slot caused a 45% tooth stiffness reduction. A modal analysis-based technique, developed to assess the stiffness of spur gear teeth, was used to estimate the differing fault severities. With each fault severity, the spur gear was tested on a back-to-back gear test rig and vibration data was recorded. The continuous wavelet transform (utilising the Morlet analysing wavelet) was applied to each set of experimental data, and it was found that the symptoms of the induced faults could be detected even in the instances of low fault severity.

INTRODUCTION

Gear deterioration can, in the main, be attributed to either lubrication or material failures. Effective lubrication of gear systems is of critical importance because it prevents direct tooth contact, reduces friction and vibration levels, removes heat generated in the mesh, and protects the gears from corrosion. If lubrication is insufficient (due for example to high temperature, excessive tooth loading, or lack of lubricant) tooth surface failures such as scuffing, pitting, and wear may result. Lubricant-related gear failure is generally progressive, meaning that the time between the onset of deterioration and the complete loss of serviceability is relatively long.

In contrast to lubricant-related issues, material-related gear failure (primarily bending fatigue leading to tooth fracture) is a more insidious

mode of deterioration. This is because it very often produces no readily detectable symptoms in its early stages, giving little warning of fault progression, and can resulting in unexpected stoppage.

This paper documents the use of the continuous wavelet transform in the detection and location of defected gear teeth caused by bending fatigue. Tooth weaknesses were simulated on a real spur gear by cutting into the root of a tooth with a thin disk cutter. The slot started at one gear face and progressed in two incremental steps across the face width of the gear, with the final most severe cut equating to a 45% tooth stiffness reduction.

A modal analysis-based technique, which has been developed to assess the stiffness of spur gears teeth, was used to estimate the differing fault severities. With each slot length, the spur gear was tested on a back-to-back gear rig and the resulting vibration signals were detected by accelerometers located around the bearing housing closest to the faulty gear. Accelerometer and timing signals were recorded, and the vibration data was averaged over three pinion rotations. The continuous wavelet transform, in which the Morlet wavelet was used as analysing wavelet, was applied to each set of experimental data as the basis for fault detection.

THE TEST GEARS AND DEFINITION OF TOOTH FAULTS

A pair of En32 spur gears, hobbed, non-case hardened, and non-undercut, was used during the tests. These gears had a face width of 30mm, a normal module of 3mm, and the drive pinion had 30 teeth meshing in a 1:1.5 ratio. They were tested on a back-to-back gear rig at their specified full load of approximately 90Nm torque.

The simulated tooth fault was introduced to one of the pinion gear by cutting into the root of the tooth with a thin disk cutter which had a diameter of 47mm and a thickness of 0.5mm. The cut was intended to replicate a crack at the critical tooth section which is growing along the critical stress line. In practice, due to access restrictions caused by adjacent teeth, the crack could be started at the critical section, but its direction deviated slightly from that of the critical stress line.

Figure 1(a) shows a schematic of the slotted 30-tooth spur gear tooth. The cutter access restriction permitted the minimum slot angle to be 42° from the tooth centre line (Z), although the critical stress line is actually at around an angle of 30°.

Figure 1(b) shows an A-A plain view of the faulty section of the test pinion. The introduced slot started at one gear face and progressed in two incremental steps across the face width of the gear. In the first step, the slot had a depth of 1.13mm and a length of 15mm. In the second step, the depth of the first cut was increased to 2.5mm and the slot length was extended across the gear face width.

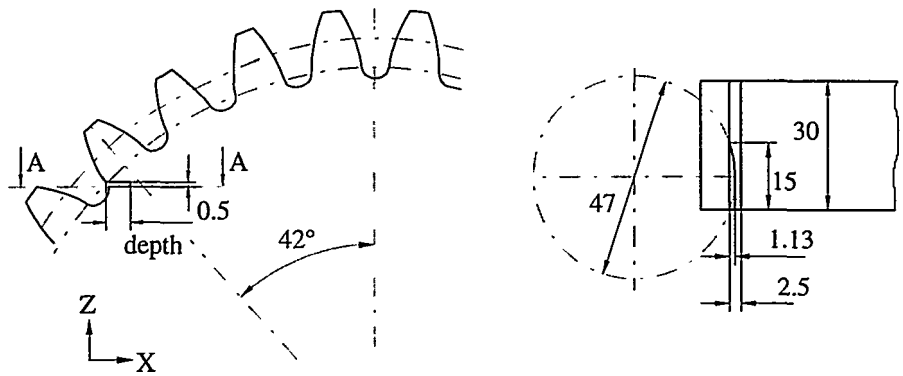


Figure 1: (a) Fault Location, (b) A-A plain view of the fault section

SEVERITY ASSESSMENT OF THE INTRODUCED FAULTS

A modal analysis-based technique, by considering the translational model of the gear tooth [1], has been developed to assess severity of spur gear tooth defects. For the purposes of modal analysis, the test gear was clamped by a pair of diametrically opposite teeth within a specially designed frame. The gear was supported only from the tip corners of the two teeth in contact with the frame, as shown in Figure 2, but across the complete face width. The gear was fixed within the frame by a pre-load screw with a tightening torque of 1.05Nm.

A small metal plate was incorporated between the end of the clamping screw and the gear to prevent damage to the gear teeth and to distribute the pre-load evenly over the contacting teeth. The frame was itself suspended in space by two soft elastic bands so that the whole structure could be assumed to have two degrees of freedom. The test assembly was designed such that, in the absence of excitation, reference axes passing through the suspended teeth of both the frame and the gear, were coincidental.

For simplified vibration analysis, both the gear and the frame can be considered as rigid bodies having lumped masses and inertias located at

their mass centres of gravity, as shown in Figure 3. The bodies can themselves be considered to be connected to each other by four springs, where k_1 and k_2 represent the spring effects in tension of the two points of suspension, k_s represents the equivalent stiffness of the pre-load screw fixing, and k_{tor} represents the torsional spring effect of the pre-load screw during relative rotational motion of the bodies. As a result, the equation of motion of the system can be expressed in terms of the relative displacement X and relative rotation θ between the rigid bodies, i.e.

$$\begin{bmatrix} M_g M_f & 0 \\ 0 & I_g I_f \end{bmatrix} \begin{Bmatrix} \ddot{X} \\ \ddot{\theta} \end{Bmatrix} + \begin{bmatrix} (M_g+M_f)(k_1+k_2+k_s) & (M_g+M_f)(k_1 a - k_2 a) \\ (I_g+I_f)(k_1 a - k_2 a) & (I_g+I_f)(k_1 a^2+k_2 a^2+k_{tor} R_t^2) \end{bmatrix} \begin{Bmatrix} X \\ \theta \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \quad (1)$$

where $M_{g,f}$ and $I_{g,f}$ denote the masses and the mass moments of inertia of the gear and frame respectively, and a and R_t.represent half of dynamic distances between two points of suspension and the tip radius of the test gear. The properties of the gear and frame were as follows: mass of gear 1.33kg, mass moment of inertia of the gear 1.316×10^{-3}kgm^2, mass of frame 6.185kg, mass moment of inertia of frame 0.042kgm^2, axial stiffness of the pre-load screw 449MNm^{-1}.

All tests were performed by exciting the system on the underside centre of the frame, using a steel-tipped impact hammer of mass 0.18kg. Attached to the impact hammer was a PCB type 208B force transducer, and mounted vertically within the bore of the test gear to measure the system response was a B&K type 4397 accelerometer. The frequency response of the system, averaged over 16 measurements with a rectangular window, was obtained using a B&K CF-920 mini-FFT analyser.

To determine the unknown stiffness parameters and to assess the effectiveness of the theoretical model, a second test gear (whose properties were exactly the same as those of the test pinion) was used. This gear was termed the modal reference gear. One of the suspending tooth face widths of the modal gear was removed incrementally in steps of: 7.2%, 13.5%, 19.6%, 29.8%, 39.6%, 49.8%, and 59.7%, and the frequency response function (FRF) of the gear-frame combination was obtained for each removal. Based upon the experimental results, the theoretical model was refined to permit the stiffness reduction (i.e. fault severity) of any kind of tooth damage to be determined [2].

Experimental results with the modal reference gear [2] showed that only the frequency of the translational mode of vibration and the isolated rotational frequency of the system could clearly be seen. The theoretical model, however, only requires that the translational frequency of the system to be known, and therefore only this frequency was considered when assessing gear tooth stiffness reduction.

The test gear with the simulated tooth fault was itself clamped within the modal test frame. It was firstly suspended by two healthy teeth, and Figure 4 shows a resulting frequency response function (FRF). Four FRF's were measured for each circumstance, and modal frequency values were averaged to improve the accuracy of the results. The averaged translational modal frequency for healthy condition was found to be 3609Hz, which is very close to that obtained for the modal reference gear (3618Hz).

The test gear was then re-oriented within the frame, so that it was suspended by one healthy and one slotted tooth, and further frequency response testing was performed. The averaged translational modal frequency values for each tooth condition are presented in Table 1.

Figure 5 shows an FRF of the frame-gear combination for the first introduced fault condition, and comparison with Figure 4 reveals a shift in the translational modal frequency. On average, the frequency had shifted from 3609Hz to 3529Hz, i.e. a reduction to 97.78% of its original value. This frequency shift itself corresponds to approximately a 12.7% reduction in gear tooth stiffness.

The FRF of the gear-frame combination for the second introduced fault is shown in Figure 6. In this case the translational frequency of the system shifted to an average value of 3169Hz, i.e. a reduction to 87.8% of its healthy condition value. For this frequency shift, the theoretical model predicts approximately a 44.9% reduction in gear tooth stiffness.

Condition	Measured Rotational frequencies (Hz)	Average (Hz)	Stiffness Reduction (%)
Healthy	3600, 3609, 3614, 3611	3609	0.00
First Fault	3525, 3524, 3542, 3533	3529	12.7
Second Fault	3205, 3161, 3165, 3164	3169	44.9

Table 1: Measured rotational frequencies for the gear-frame combination and estimated spur gear tooth stiffness reductions for different fault conditions.

APPLICATION OF THE WAVELET TRANSFORM TO THE EXPERIMENTAL DATA

The vibration acceleration of the test gears was detected by an accelerometer located in horizontal orientation on the pinion bearing housing. A position reference signal, producing one pulse per revolution of the test pinion, was incorporated to enable synchronous averaging of the data. The pinion speed was set to 400rpm, producing a fundamental toothmeshing frequency of 200Hz. The vibration signal was analogue band-pass filtered between 160Hz and 1000Hz filter to eliminate irrelevant structural components (these were identified by FRF testing of the gear test rig). After filtering of the accelerometer output, the test and reference signals were sampled at 3.2kHz and stored on computer.

The Morlet wavelet was coded in Matlab v.4.2c.1 and was implemented on a Pentium PC. In the Morlet wavelet code, the sampling frequency of the wavelet function was chosen to be 256Hz, and a wavelet center frequency $f_0 = 1.5\text{Hz}$ was selected. The raw vibration data was collected over 66 pinion rotations, this was then split into 22 three-rotation blocks, and these were averaged together. From the centre of the resulting average, which had a span of 1440 data points, a block of 512 samples was extracted and this was input to the wavelet transform (this is because wavelet analysis requires a time record length equal to a power of 2). Only the wavelet results for the central 480 samples of the 512 sample time record are displayed because these samples represent one exact pinion rotation.

The wavelet calculation was performed in the frequency domain and was based upon octave band analysis because this reveals the fine detail of the signal [3]. However, the number of octaves used in the calculation depends upon the length of the time record, and 18 voices per octave were hence used to obtain sufficient frequency resolution. The result of the Morlet wavelet transform is presented in the form of amplitude contour plot.

Figure 7 illustrates the wavelet transform for vibration from the healthy gear pair. It can clearly be seen that the signal energy is concentrated primarily around the first and second meshing harmonics (200Hz and 400Hz); in addition, some energy is seen to be localized at high frequencies.

Low-frequency amplitude modulation of the first toothmeshing harmonic, which fluctuates twice per pinion rotation is clearly seen. This modulation is most likely caused by non-perpendicular machining of the gear bore. In effect, this causes a sinusoidally varying mesh stiffness which repeats itself

with a frequency of twice the pinion rotation. Similar to the fundamental, the amplitude plot of the second toothmeshing harmonic also exhibits energy level variation along the time axis, in this case however, the fluctuation only repeats itself once per pinion rotation. In addition, the second harmonic exhibits some rapid short-duration frequency variations. These may well be attributable to the freshly-hobbed surface finish of the gears.

Figure 8 shows the wavelet transform of the gear pair when the first fault was introduced. It can be seen that the signal energy is concentrated in four frequency regions around 100Hz, 200Hz, 400Hz, and 800Hz. The 100Hz component is actually due to structural vibration of the test rig.

The presense of the local fault cause a one-sided variation in the band thickness of the first harmonic. This variation reaches to its two maxima around the 30° and 300° positions. Similar variations are also observed at approximately 120° either side of those maxima (i.e. around the 140° and 180° positions, which are when the faulty tooth comes into and leaves the gear mesh). When the gear ratio is considered, the defected tooth on the pinion engages with only three teeth on the wheel gear. As a consequence, these three wheel teeth may suffer surface deterioration.

Figure 9 shows the wavelet transform with the second fault introduced. More pronounced fault symptoms can now be seen, and the interfrequency activities, where the faulty tooth is in mesh, actually link the first and second meshing harmonics. Furthermore, the energy around this angular position is intensified across the whole frequency range.

SUMMARY AND CONCLUSIONS

Severity assessment of tooth defects can be made by observing the shift in translational frequency of a clamped gear wheel. An estimate of the reduction in tooth stiffness can be obtained based upon the shift in translational frequency via the use of a theoretical model.

The Morlet wavelet transform shows signal content as a function of both time and frequency, and with a varying time-frequency resolution. For gear vibration, the principle signal components and their modulations can clearly be observed.

The presence of a gear tooth with a stiffness reduced by 12.7% (i.e. the first introduced fault condition) can be seen in a wavelet amplitude plot as

interfrequency activities attempting to link the toothmeshing harmonics. The the strongest interfrequency activities are localised on either side of the 150° (faulty tooth) position, with a 120° phase lag.

With increased fault severity (a stiffness reduction of 44.9%), clearer indications of fault presence are revealed. These indications manifest themselves as increases in the frequency content of the signal throughout the frequency range at the position when the faulty tooth is in mesh.

REFERENCES

[1] Ozguven, H. N., and Houser, "Mathematical Models Used in Gear Dynamics-A Review", D. R., J. of Sound and Vibration, Vol. 121., 1988, pp. 383-411.

[2] Yesilyurt, I., and Ball, A. D. "The Application of Modal Analysis to the Estimation of Fault Severity in Spur Gears", Third International Conference on Engineering Structural Integrity Assessment, Cambridge, UK, 24-26 September 1996.

[3] Riol, O. and Vetterli, M. "Wavelets and Signal Processing", IEEE SP Magazine, pp: 14-38, October 1991.

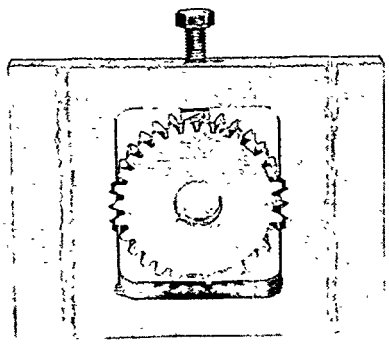


Figure 2: The frame and the test gear used for the modal testing.

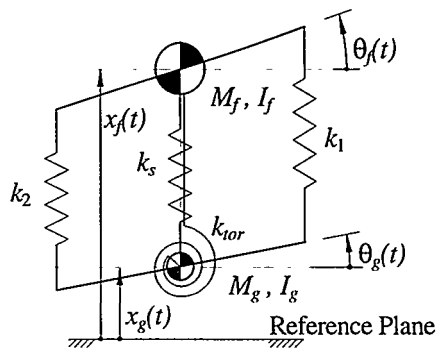


Figure 3: Lumped parameter model

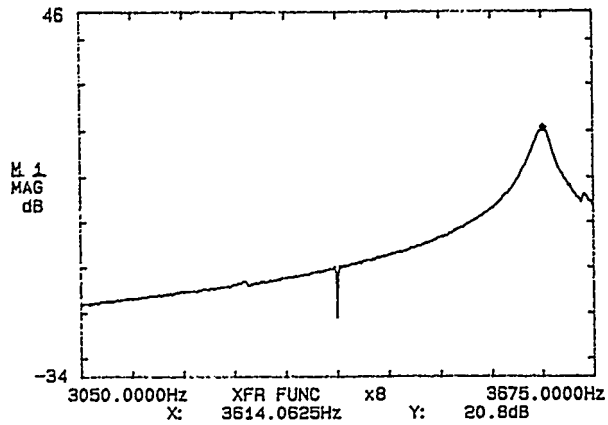


Figure 4: The FRF of the gear-frame combination for healthy tooth.

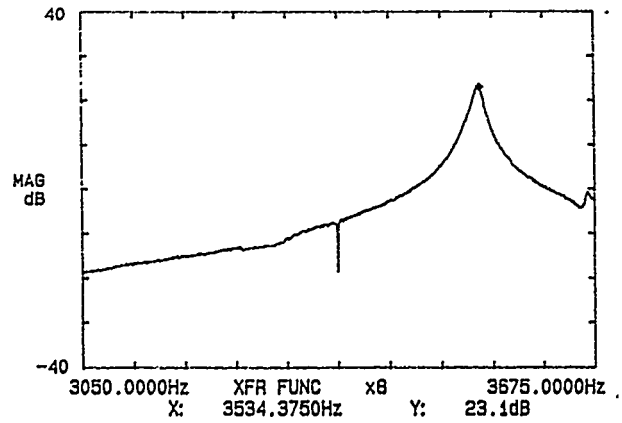


Figure 5: The FRF of the gear-frame combination for the first fault.

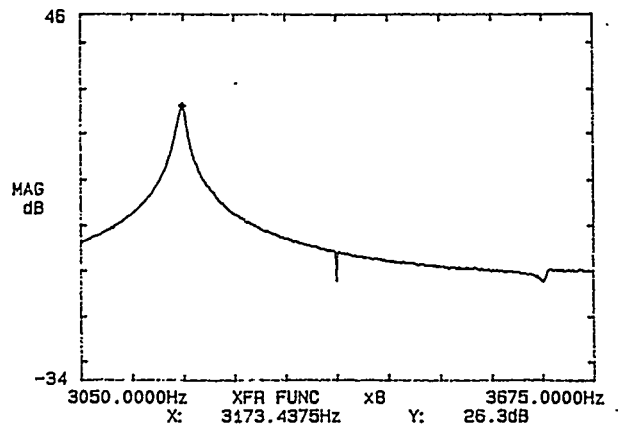


Figure 6: The FRF of the gear-frame combination for the second fault.

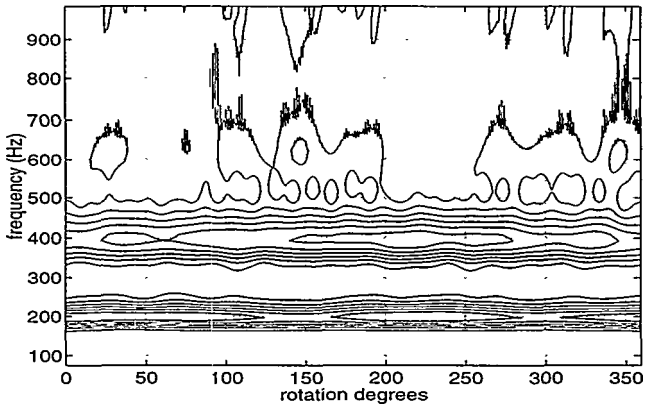


Figure 7: The wavelet transform of the vibration signal of the healthy gear pair

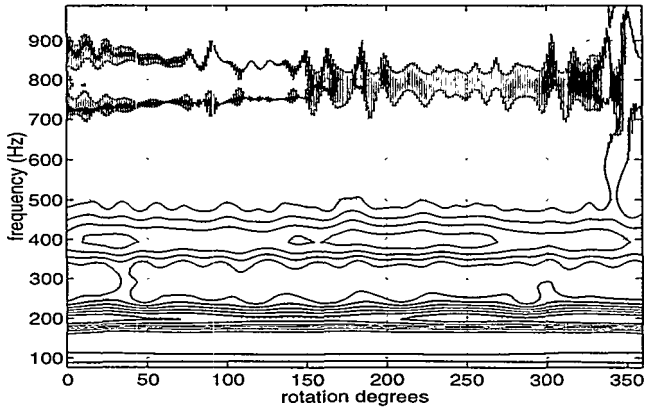


Figure 8: The wavelet transform of the vibration signal of the gear pair for the first fault

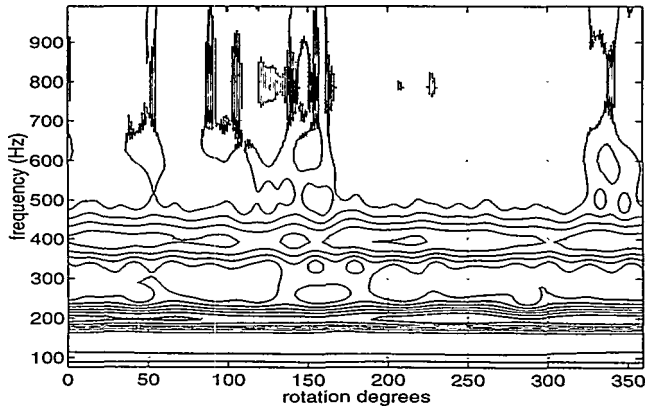


Figure 9: The wavelet transform of the vibration signal of the gear pair for the second fault.

CHANGES IN DEBRIS SIZE DISTRIBUTION DUE TO WEARING CONDITIONS

Yoshitsugu Kimura
Institute of Industrial Science, The University of Tokyo, Tokyo, Japan
Kenichi Hisamura
Lubricant Research & Product Engineering Department, Nippon Oil Company, Yokohama, Japan

ABSTRACT

Possibility of monitoring wearing conditions by quantitative analysis of debris is studied. A series of experiments on lubricated wear of steel is described in which sliding was made on a three-pin-on-disk machine. Sliding conditions were varied by contaminating the lubricant, deteriorating the lubricant, changing normal load and changing temperature. Debris were collected and their size distributions were determined on an image analyzer. It was found that the size of the debris followed similar distribution in any case and that only the normal load governed the size in a steady condition.

1 INTRODUCTION

Tribology has been an important part of failure physics in maintenance engineering. Wear and other tribological failures often limit operating life and sometimes cause unscheduled breakdown of facilities, and tribocomponents like bearings have been major objective items of condition monitoring.

Monitoring of wear has been made in various ways. The spectrometric oil analysis program has been used in practice which gives indication of wear at its very early stage with information about the materials being worn. The analytical ferrography has found its wide application area in plants which visualizes various modes of wear taking place, while the direct-reading ferrography is used to provide a symptom of abnormal wear in a change in the wear-severity index which is a measure of debris concentration and anomalies. Radioactivity has also been used for

monitoring wear as is found in the surface-layer activation technique. On the other hand, technologies for diagnosis of the condition and prediction of future changes of tribocomponents have been developed to make full use of wear data on the quantity, size and morphology of debris [1,2,3].

Wear behavior can be affected by various causes. Most simply, the so-called running-in of sliding surfaces is known to result in favorable changes[4,5]. Light wear and plastic deformation of surface asperities generally smoothen the surfaces resulting in enhanced hydrodynamic lubrication and mitigation of solid contact leading to lowering of wear rates. However, many changes in sliding conditions can be detrimental. Intrusion of dust into interfaces is no matter the worst case, which causes heavy abrasive wear. Contamination of lubricating oil with foreign liquids, fuel or water for example, also markedly impairs lubricating performance of the oil by decreasing viscosity or increasing chemical attack. Deterioration of lubricating oil during use results in depletion of formulated additives and formation of active compounds leading to enhanced corrosive wear. Unexpected increase in actual normal load, which can occur as a result of localization of contact pressure caused by misalignment or uneven wear, provides another class of the changes in the sliding conditions in real tribosystems which is often overlooked in laboratory studies.

The present study describes wear experiments under some of the different conditions as stated above, and tries to correlate those changes in the conditions with the changes in the size distribution of wear debris.

2 WEAR EXPERIMENT

2.1 Apparatus and Specimens

A series of experiments on lubricated wear of steel was conducted on a three pin-on-disk machine. Figure 1 shows specimens. A disk, 40 mm in diameter and 10 mm in thickness, was fixed to a holder and was mounted on the bottom end of a vertical rotating shaft driven by an electric motor. The three pins having flat ends, 4 mm in diameter, were tightly fitted to three holes in a pin holder disk at regular intervals on a circumference of 30 mm in diameter, forming a pin assembly. The assembly was secured to the bottom of an oil bath, at the bottom of which an electric heater was attached to control the oil temperature. The oil bath was supported by a gimbal mechanism mounted at the top of a

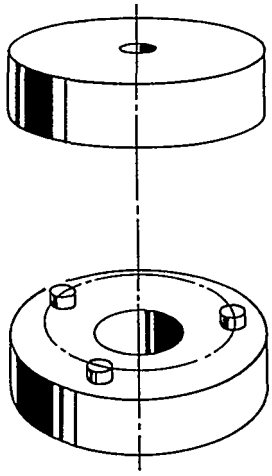


Fig. 1 Specimens

hydraulic loading piston, where the gimbal mechanism ensured an even load distribution over the three pins. The piston was allowed to rotate freely, and a torque arm fixed to the gimbal support pressed against a load cell to provide a measure of friction. Both the disks and the pins were made of 0.35% carbon steel annealed to HV160. The sliding surface of the disk and the pins were ground to 0.1-0.2 μm Ra; the grinding of the pins was made after they were fixed to the pin holder. Five micro-Vickers indentations were made on the sliding surface of a pin for determining wear amount; no measurement was made with wear on the disk surface.

2.2 Lubricating Oil

A lubricating oil used as the standard was paraffinic mineral base oil without additives. Its viscosity was 13.3 mm^2s^{-1} at 40°C and 3.19 mm^2s^{-1} at 100°C; it had a viscosity index of 103. The total acid number was less than 0.01 mgKOH g^{-1} which indicated practically no acidic constituents were included.

2.3 Sliding Conditions

Wear experiments were made under six sliding conditions. A standard condition STD was defined just below, and one of the conditions was changed, while the remaining conditions were kept the same with STD, in the other five sliding conditions, WCL, DET, HL1, HL2 and HT.

-Standard(STD) The load was 392 N which applies a nominal contact pressure of 10.4 MPa. The sliding speed was 47.1 mm s^{-1} at 30 r min^{-1};

this speed was not changed throughout the experiments described in this paper. The oil bath temperature was set at 40°C at the outset of each run.

-Water-Contaminated Lubricant(WCL) The lubricant was changed. Distilled water by 1 vol% was added to the standard lubricating oil. A water-in-oil emulsion was prepared with a homogenizer by adding another 1 vol% of methacrylate copolymer as the emulsifying agent. The bulk viscosity of this emulsion was 14.8 mm^2s^{-1} at 40°C.

-Deteriorated Lubricant(DET) The lubricant was changed. The standard lubricating oil was deteriorated by oxidation at 150°C for 18 days without catalysts. This increased the total acid number to 0.30 mgKOH g^{-1}, and the viscosity up to 14.2 mm^2s^{-1} at 40°C, indicating formation of acidic oxidation products.

-High Load(HL1 and HL2) The normal load was changed. It was increased to 558 N in HL1 and to 784 N in HL2 conditions.

-High Temperature(HT) The temperature was changed. A sliding started after the oil bath temperature was raised to 80°C.

2.4 Procedure

To minimize contamination with dust, the specimens, their holders and the oil bath were thoroughly washed with benzine, which had been previously cleaned with a 0.45 μm pore membrane filter, in an ultrasonic bath before each run. The cup and the blades of the homogenizer used for preparing the emulsion were also washed in the same way. These cleaning procedures were conducted in a clean bench. The lubricating oil was used as received. The deteriorated oil contained insolubles, and was cleaned by the membrane filter.

After the cleaning procedure, the specimens were set on the holders, 30 cm^3 of the lubricant was poured into the oil bath which completely immersed the specimens, and they were mounted on the machine. Then the oil bath temperature was raised to the desired value, the rotation started and the load was applied. For each condition, sliding continued for 15 min for the first two runs, which was followed by a run for 30 min and runs for every 60 min thereafter. Debris were collected as stated below, and the above procedures were repeated to start another run.

3 DEBRIS ANALYSIS

3.1 Collection of The Debris

When a run was completed, the lubricant containing the debris was poured into a cleaned glass bottle. The specimens, the holders and the oil bath were washed with about 20 cm^3 of the benzine cleaned as above to collect the debris adhering to them. They were blended, and a high viscosity oil, VG100, was added by about 80 cm^3, which increased the viscosity of the 'carrier oil' to suspend the debris effectively.

Thus prepared oil sample containing the debris was shaken to ensure uniform dispersion, and the debris were collected on a 0.45 μm pore membrane filter, 35 mm in effective diameter, by employing a slight negative pressure. Only a small part of an oil sample was sufficient for observation of enough number of the debris; excessive amount of debris could make recognition of individuals difficult.

3.2 Image Analysis

The membrane carrying the debris was observed by an image analyzer. The membrane was placed under an optical microscope. The image was provided to a microcomputer system through a CCD camera. Magnification was set so that a frame comprised of 512x511 pixels corresponded to 736x734 μm. It occupied some 1/1780 of the total effective area of a membrane, and 15-25 frames, that included 1500-2000 debris in total, were analyzed for each sample.

Preliminary observation was made with several large debris for STD in an SEM equipped with a profilometry probe. It revealed that most debris were thin flakes, the ratio of the thickness to the dimension perpendicular to it being of the order of 0.1. Then the following analysis was made of the shapes and sizes of the figure projected on the membrane surface by using an image analysis software FRM TOOL-KIT(Ver.2.1). It needed binalization of the concentration levels on an arbitrary basis dependent on the background.

An equivalent diameter of a debris was defined for its projected figure as follows. A major axis was defined by the largest chord of the figure, and a minor axis by the smallest chord which intersected the major axis and passed the figure center. Then the equivalent diameter was defined by the average length of the major and the minor axis. The software provided distribution of the equivalent diameter for each sample of the debris. Here, rather unstable data for the debris having the equivalent diameter less than 2 μm and those larger than 30 μm were omitted, because the

former had little contribution to wear amount, while the latter were erratic although it might contain information on some catastrophic failures other than wear.

4 RESULTS AND DISCUSSION

4.1 Friction and Wear

Examples of the records of the coefficient of friction are reproduced in Fig.2 for the cases of STD and HL2. In every run, the coefficient of friction starts from a low value, reaches a maximum, and then levels off or slightly lowers. In spite of these short-term variation, however, they show a common trend of gradual decrease towards a stationary value around 0.12 to 0.15. This trend is common for all sliding conditions studied here, but WCL alone shows a different short-term variation that the initial coefficient of friction is high and lowers with considerable fluctuations.

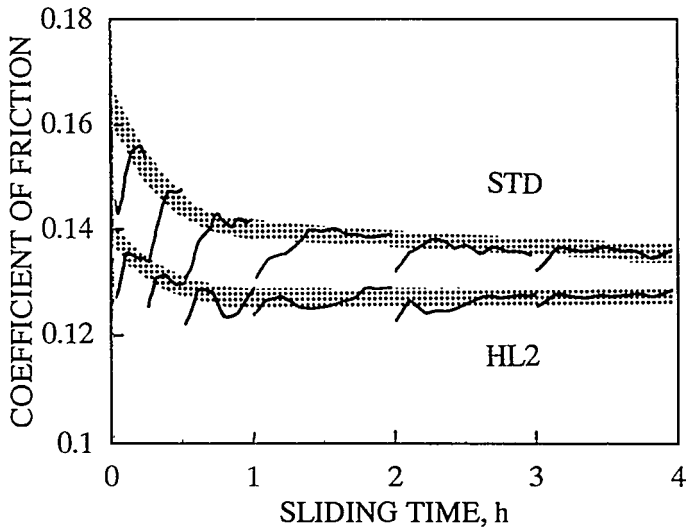


Fig. 2 Expamples of coefficient of friciton

Figure 3 summarizes volumetric wear of the pins as a function of sliding duration, each curve representing the average over one to four separate experiments. Although some are nearly linear while others deviate a little from the linearity, these lack reproducibility. The general amount of wear, however, can be classified into two: that for WCL and HL2 is about twice as large as that for STD, DET and HT. The experiment for WCL was discontinued after 2 h of sliding because the micro-Vickers indentations were worn out.

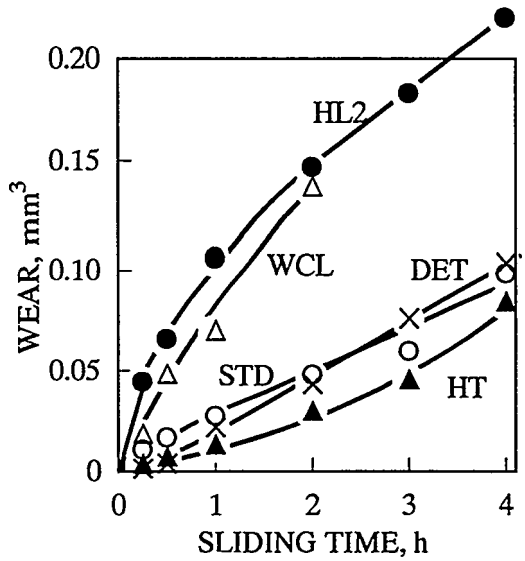


Fig. 3 Wear of three pins

4.2 Distribution of the Debris Size

An example of the distribution of the equivalent diameter of the debris for STD collected after sliding 3 h is shown in Fig.4, in which the fraction of the number of the debris having larger diameter than each value of the abscissa is plotted on the ordinate in a logarithmic scale. The linear relation shows that the equivalent diameters in this sample follow an exponential distribution. It is shown that the samples for all conditions studied here follow exponential distributions. From the slope of these

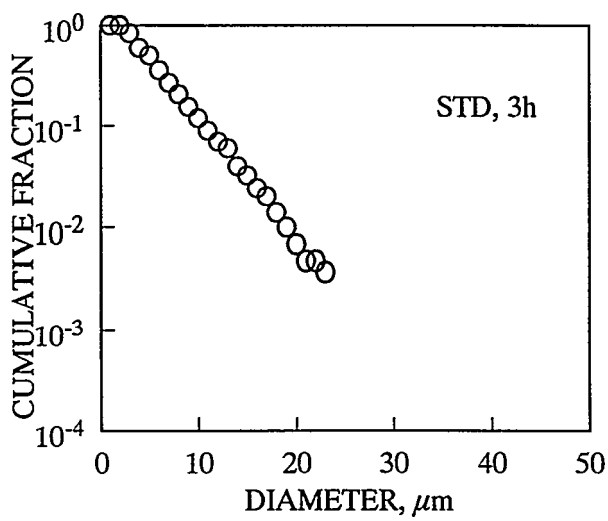


Fig. 4 Expamples of distribution of equivalent diameter of debris

linear relation, the average equivalent diameter is determined for each sample.

4.3 Changes in Average Debris Size

The change in the average equivalent diameter with sliding time is given for STD in Fig.5, which shows a large value at the outset of sliding and a gradual decrease to a steady value. This is a change commonly observed in wear of steel lubricated with lubricating oil without additives due to running-in [4]. It is noted that this change is related with that in the coefficient of friction, but not with that in the progressive wear amount.

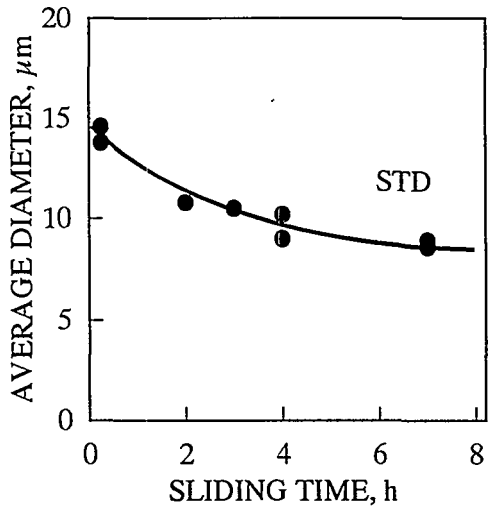


Fig. 5 Change in average equivalent diameter for STD

The curve for STD is compared with those for WCL, DET and HT in Fig.6. The initial values of the average equivalent diameter are different dependent on the conditions. It is the largest for DET, intermediate for STD and HT, and the smallest for for WCL. Further, the average equivalent diameter is increasing with sliding duration for WCL suggesting absence of running-in. Making a contrast to the variation of the initial values, however, the average equivalent diameters of the debris under different sliding conditions tend to converge towards a steady value which is practically the same with that for STD. It is noted the rating of neither the initial value nor the steady value has direct correlation with that of the wear amount.

Similar changes are plotted when the normal load is changed in Fig. 7. The difference in the initial values is somewhat confusing in that it is the largest for HL1, the intermediate load. However, since decrease to the

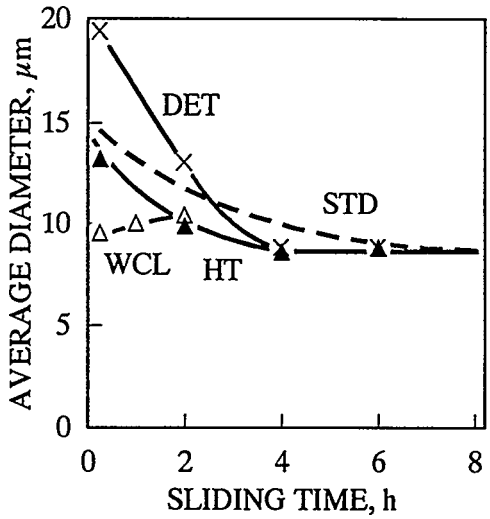


Fig. 6 Changes in average equivalent diameter for different lubricants and temperature

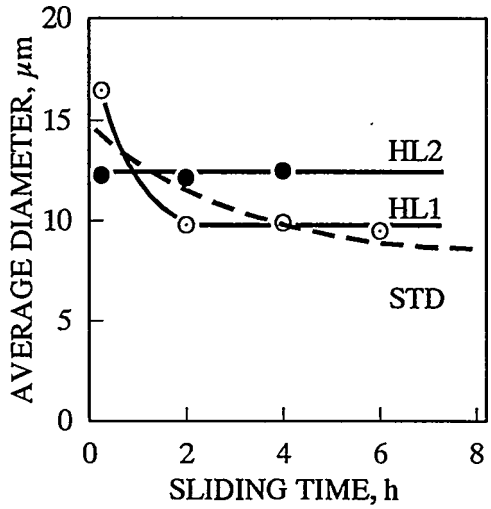


Fig. 7 Changes in average equivalent diameter for different loads

steady value for HL1 takes place quicker than for STD, it can be reasoned that it has taken place for HL2 still quicker from a largest initial value. In this case with different normal loads, it is observed that the steady value of the average equivalent diameter increases with the load. This is a different feature from the cases with different lubricants or temperatures, and supports a theory that the size of wear debris is determined by mechanics while the rate of their removal is determined by the environment [6].

5 CONCLUSIONS

Distribution of debris size was studied for lubricated wear of steel under different wearing conditions. The main conclusions are as follows.

(1) Wear debris have generally flat shapes and, when collected on a filter and observed under a microscope, determination of their size is possible by representing an equivalent diameter for the projected figures.

(2) The average diameter of wear debris generated under a sliding condition follows an exponential distribution.

(3) The average equivalent diameter of wear debris for different sliding conditions has different values at the outset of sliding. Contamination with water and deterioration by oxidation of lubricating oil, as well as change in the lubricant temperature, do not change the steady value, but difference in the normal load affects it. This indicates possibility of detecting causes of anomalous wear by the determination of debris size.

REFERENCES

1. Roylance, B. J. et al. "The Development of a Computer-Aided Systematic Particle Analysis Procedure - CASPA", STLE Preprint, 92-AM-7E-2, 1992, 1-16.
2. Roylance, B. J. et al. "Computer-Aided Vision Engineering (CAVE) - Quantification of Wear Particle Morphology", Lubrication Engineering, **50**, 2, 1994, 111-116.
3. Holmberg, K., Kuoppala, R. and Vuoti, A. "Expert System for Wear Failure Prediction", Proceedings of the 5th International Congress on Tribology, **4**, Espoo, Finland, June 12-15, 1989, 32-37.
4. Kimura, Y. and Sugimura, J. "Microgeometry of Sliding Surfaces and Wear Particles in Lubricated Contact", Wear, **100**, 1984, 33-45.
5. Sugimura, J., Kimura, Y. and Amino, K. "Analysis of the Topographical Changes due to Wear", Journal of JSLE International Edition, **8**, 1987, 69-74.
6. Soda, N., Kimura, Y. and Tanaka, A. "Wear of Some F.C.C. Metals during Unlubricated Sliding, Part II: Effects of Normal Load, Sliding Velocity and Atmospheric Pressure on Wear Fragments", Wear, **35**, 1975, 331-343.

TOWARDS THE DEVELOPMENT OF ADVANCED ANALYTICAL TOOLS IN FIRE SAFETY ENGINEERING

C. Green and Dr. J. MacIntyre
School of Computing & Information Systems, University of Sunderland, UK

Professor T. J. Shields and G. Silcock
Fire Safety Engineering Research & Technology Centre, University of Ulster, UK

ABSTRACT

Fire Safety Engineering represents one of the most interesting and demanding areas for the collection, storage, and analysis of complex sensor data. Patterns of information coming from instruments placed in arrays in fire room simulators generate substantial amounts of data which, if interpreted correctly, could give valuable information to Fire Safety Engineers about how fires develop and spread, and how constructural components such as windows, doors, and walls behave during the course of a fire. However, little in the way of standardised analysis tools currently exist. Work in this area has concentrated on the development of very specialised mathematical tools, which cannot be easily adapted to different test and simulation conditions. The practical use of these tools has therefore been limited.

The Universities of Ulster and Sunderland in the United Kingdom are working together to develop intelligent analytical tools for Fire Safety Engineering, focusing on the use of neural networks and other adaptive techniques. The emphasis of this work is to provide practical, adaptable tools which Fire Safety Engineers can use to learn about the behaviour of elements of structure and materials exposed to various fire conditions. This paper describes in detail the problems faced by Fire Safety Engineers in analysing test data, and how the two Universities are combining their respective skills in the areas of

Artificial Intelligence and Fire Safety Engineering to develop solutions.

1. INTRODUCTION

It is widely recognised that timber plays an important role in both preventing and contributing to the development of fire. Conventionally, forensic analysis of timber in fire investigations is carried out by means of destructive test methods such as wire brushing and full-scale furnace testing. Analysis of the results of such testing relies heavily upon the expertise of the human investigator. Two Universities in the UK, the University of Sunderland and the University of Ulster, are working together in a project which aims to address the problems with current methods, and to introduce improved data collection and analytical techniques, to assist during fire investigations.

2. FIRE DOOR TESTING

2.1 The Role of Fire Doors

Fire door assemblies which have been well-designed and correctly installed should be able to fulfil the primary functions of isolation and security within an enclosure, without restricting the movement of both people and goods [1].

The majority of fires start from small sources and develop quantities of smoke in the early stages, particularly in enclosures having restricted ventilation. One role performed by a fire door is to act as a barrier to the passage of smoke and fire into escape routes thus allowing sufficient time for the occupants to escape to safety. Different parts of a building may need to be separated by fire-resistant construction. Thus the type of door used in these different areas will vary depending on its location in the building, and the fire hazard associated with the building or room, taking into account the contents of the room, occupants etc.

During the early stages, there are generally pressure differences across a fire door, which can be the fire itself. These pressures may force smoke through cracks and gaps which can be found around the

door. As the fire progresses, there may be some damage or deformation to the door, thus allowing greater quantities of smoke to pass through. If the fire gets bigger, then the door will need to be able to withstand exposure to high temperature conditions without losing its integrity and stability [1].

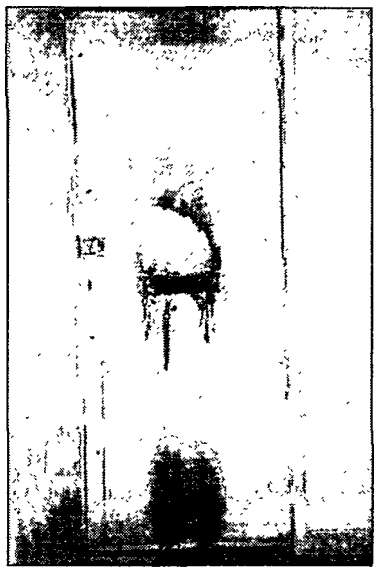


Figure 1. A breach such as this reduces the ability of a fire door to act as a barrier to fire.

We have seen that the precise role of the door will very much depend on whether on not it is intended to protect escape routes from the effects of fire, or to protect the contents and structure of a building by restricting the spread of the fire. It is important then that the correct type of door is installed for the role required in preventing the spread of smoke and fire.

2.2 Fire Door Assembly Performance

Fire door assembly testing in the UK is carried out to the approved British Standard (BS 476 Part 8) which requires a door and frame is exposed to destructive methods using a furnace [2]. This test will certify a particular door for a set period of fire resistance i.e. Half-hour, Hour and Two Hour doors. Providing doors are made to the same standard specification they are accepted as qualifying for that purpose and no further test is required.

Any imperfections in the timber and its inherent reduction in quality can go unnoticed. Subsequently this may lead to a possible reduction in performance during the fire resistance test. Thus, quality control of this product is essential if a high standard of fire resistance is to be provided by this product.

2.3 Difficulties of In-Situ Testing

Installation of the doors also follow strict procedures, since correct installation is as important as the door itself in preventing the spread of fire. Fire safety officers carrying out inspections on fire doors must assess the suitability and fire resisting properties of a door, as well as the installation, before a fire safety certificate can be issued.

A number of difficulties arise for fire officers when trying to carry out safety inspections. Fire doors which have been furnace tested have been tested as a 'door set' complete with hardware and ironmongery, and installed in a particular way. During fire safety inspections, it is very difficult for an officer to know if the installation procedure has been carried out in the same way as the door set which passed the furnace test. It is also difficult for officers carrying out visual inspections to identify a particular door type, its integrity or fire resisting properties. Other countries carry out stringent test procedures on doors once they are installed. In the UK this type of procedure is not followed, and consequently the problems mentioned must be dealt with.

Problems also arise when safety officers inspect 'old and listed' buildings, the inspecting officers have no way of knowing the effects of ageing on the integrity of the doors, or their fire resisting properties.

This project will address these problems through the development of analytical tools to be used by the fire officers at the inspection site. The process of this development work is outlined below.

2.4 Experimental Work to be Carried Out

In order to determine the passive performance parameters relating to the material, geometry and structural form of the door assembly with respect to its performance when subjected to an enclosure fire environment, a series of fire resistance tests will be carried out for a selection of common door sets. Each test will be conducted using the BS476 fire resistance test protocols.

The data which can be referred to as the passive parameters will be determined prior to testing. Such data may include the identification of surface defects such as cracks and door furniture defects. During the fire tests active parameters such as exposed surface temperatures, degree of deformation (bowing and buckling), and crack growth will be recorded using automated data capture equipment and video recording. The performance of a door will be determined when it first fails by the insulation or integrity criterion.

With this data, it should be possible to use neural networks and other advanced computing techniques to relate the performance of the door to the passive parameters and observations. Thus it should ultimately be possible to predict the probably fire resistance performance of doors from a knowledge of the essential passive parameters.

2.5 Advanced Analysis Techniques

The University of Sunderland has considerable experi nce in developing advanced computing techniques for analysis of data in the areas of condition monitoring and non-destructive testing. Various industrial projects have involved the use of neural networks as classifiers to perform pattern recognition and diagnostics; others have used neural networks and expert systems combined in a hybrid intelligent system to perform similar types of task. Other projects have used neural networks and genetic algorithms for various applications in process control and optimisation. This experience leads the University of Sunderland to believe that such techniques can play a valuable role in the analysis of data in the area of Fire Safety Engineering.

Neural networks have become more and more established as a technique for data analysis, especially in the area of non-linear dynamical systems, in recent years. It is not proposed to go into detail

here; much information is available in the literature [3,4]. However, it is the ability of neural networks to perform a multi-variate, non-linear mapping from a set of inputs to a desired output that render them useful in this domain.

As with all industrial data, it is necessary to first understand the nature and characteristics of the data before developing an analytical tool. This is especially true in Fire Safety Engineering, where much of the data is not yet well understood. The Universities of Sunderland and Ulster are working to acquire and analyse data from fire door tests in order to determine which parameters are the most useful in defining the behaviour of doors, and can therefore be measured in-situ to determine whether a door is fit for purpose. Analytical tools are being developed using the techniques mentioned above, with the objective of providing a simple, straightforward, and practical tool which inspects can use to assist them in their job while on site. Such tools would offer benefits of speed, accuracy, reliability and consistency in the examination and assessment of fire doors in situ.

2.6 Objectives and Outcomes

Through the work outlined above the main objectives of this research is to develop analytical tools and techniques which can be used by fire officers to assist with fire door safety inspection. It is expected that these tools will allow fire safety officers identify door types in situ, with a measurable amount of confidence.

Advanced analytical techniques are to be developed, which can be used by fire safety officers to predict the performance of installed doors based on a small set of non-invasive parameters, which the officers themselves can easily measure. Throughout this work, it is hoped that the tools and techniques developed will enable fire safety officers to estimate the feasibility of upgrading from one door type to another, particularly in 'old and listed' buildings, with respect to fire safety performance.

In fire door testing, non-destructive techniques are being considered as a means to identify internal features of a fire door in order that a 'prototype' for a suitable door may be established. Such criteria is envisaged as leading towards improved fire door standards and quality control for door manufacturers.

It is hoped that the tools produced as a result of the research will enable fire safety officers to carry out stringent tests on doors in situ, bringing the UK in line with other countries fire safety inspection and testing procedures.

3. FORENSIC ANALYSIS OF FIRES

3.1 Behaviour of Timber in Fire

Timber and wood-based products such as plywood and chipboard are, and have long been, in widespread use in buildings. They also have a reasonably predictable response when exposed to high temperatures and fire conditions. Indeed, their near-uniform rate of charring when exposed to temperatures typical of a fully-developed fire enables the fire resistance of structural timber elements to be calculated.

Figure 2. Typical scene waiting for fire investigation teams

The combination of these two facts gives the possibility of a potentially valuable investigative technique for determining, after the event, such aspects as the cause and seat of ignition of a fire, its course and speed of development, the sequence of fire spread, the

likely duration of the fire and an indication of the maximum temperatures reached.

3.2 Fire Investigations

A natural fire is a complex phenomenon. Real fires differ widely in extent, duration, spread, intensity, etc., depending on such factors as the types and disposition of combustible materials, the source of ignition, the presence of accelerates, the building configuration and so on; environmental influences such as wind speed and direction also contribute. It follows that the remains of a fire represent a complex system that has been subjected to various heating (and quenching) regimes depending on its location and involvement in the fire. At present, rule of thumb assumptions are made concerning the significance of the appearance of the timber after exposure to fire. Greater depths of charring are assumed to show that the timber was closer to the seat of the fire. For example, a fine pattern of charcoal formation is assumed to indicate a slow, smouldering fire. In reality the charring pattern is as much a feature of the species of timber as of the character of the fire to which it is exposed, and it is known that charring rates are not linear with time at low and high fire temperatures. In order to develop an effective system for forensic investigation of fires based on surviving timber, three aspects need to be addressed:

- the response of different species and sizes/masses of timber to defined types of fire exposure (direction, intensity, oxygen availability etc.). This needs to be investigated experimentally, under controlled conditions and objectively quantified;

- a strategy needs to be devised for selecting initial evidence from the mass of data presented by the remains of the fire. The findings can then be co-ordinated in such a way as to enable the building up of a quantitative fire scenario;

- a guide in the form of a code of practice for use by fire investigators needs to be drawn up and published for use by practitioners.

3.3 Advance to Standards

The two Universities, along with VTT Manufacturing Technology of Finland, are developing a project (acronym FORISST - Intelligent System for Fire Investigation and Safety Standards) to address these issues. The project will explore new methods of acquiring data representative of the effects of fire on timber, through non-destructive methods in the area of fire investigations and fire safety. From this suitable guidelines will be developed to achieve more realistic assessments during fire investigations.

3.4 Objectives and Outcomes

Through the use of non-destructive testing methods and advanced computing technologies (such as expert systems, neural networks, image processing and computer simulation), criteria on which to base tools to assist in fire investigations are proposed. The project will provide benefit to fire authorities, investigation teams and industry in general (throughout the European Community), with the emergence of measurement standards and guidelines for fire safety and forensic work. Other benefits would be support for insurance and forensic services (to achieve more realistic assessments of fire damaged premises), and a positive environmental impact through the reduction of the need for unnecessary furnace testing of timber based components. It is unlikely that the system will be able to provide definitive answers in all cases. Probabilistic conclusions are much more likely in view of the complexity of the situations which it is hoped to reconstruct.

3.5 Partners Needed

The University of Sunderland is currently involved in a number of applied collaborative projects which utilise the advanced technologies previously mentioned. It is intended that the proposed project would use such technologies to provide a solution to what is a very complex problem. We are working closely with fire prevention authorities and timber researchers in this area, but are looking for other partners who would like to work on this exciting project. European partners are sought in the following areas: Fire Door Manufacturers, Fire Door Test Centres, Insurance Companies and Fire Agencies.

4. CONCLUSIONS

We have seen that the collection of accurate and representative data in the areas of fire investigation and fire door testing, and the subsequent analysis of such data by human investigators, are areas which are both problematic and complex. The use of analytical techniques a long with advanced computing technologies to develop tools in fire safety engineering are seen as having potential for assisting in this area. The research described in this paper is believed will be of enormous benefit to fire investigation teams and fire safety officers and industry in general, with the emergence of improved measurement standards and guidelines for fire safety engineering. The authors feel that there is clearly much scope for work in this area.

5. REFERENCES

[1] Building Research Establishment Digest (Digest 320), Department of the Environment. 1988.

[2] British Standards 476, Part 8.

[3] Hinton, G., How Neural Networks Learn from Experience, Scientific American, September 1992.

[4] Green, C., MacIntyre, J., Non-Destructive Evaluation of Wood using Neural Networks, Proc. of 8th International Congress on Condition Monitoring and Diagnostic Engineering Management (COMADEM 95), Kingston, Canada, 1995.

DATA ACQUISITION OF A TEST RIG FOR THE DEVELOPMENT OF A PORTABLE HAEMODIALYSER

R J A Bigsby, Research Fellow, Dept. Mechanical Engineering, University of Leeds, Leeds, LS2 9JT. UK

J G Martin, Experimental Officer, R J Rider, Senior Lecturer, G N Blount, Head of engineering Design, School of Engineering, Coventry University, Coventry, CV1 5FB. UK.

ABSTRACT

One of the main tasks in developing a portable artificial kidney machine is to reduce the volume of dialysate required. The approach adopted in this work was to regenerate spent dialysate by passing it through an activated carbon filter at temperatures down to 0°C. A test rig was developed to explore the feasibility of this idea. In particular, it was ultimately required to determine an optimum machine configuration. The test rig was therefore to be capable of simulating a typical four hour dialysing session.

The test rig was based on two counterflow liquid circuits, one being the blood analogue and the other being the dialysate. The major components of the blood analogue circuit consisted of a reservoir and a heater. The major components of the dialysate circuit consisted of a reservoir, a dialyser, a heat exchanger, the carbon filter and a cooler.

In order to achieve a controlled dialysis simulation and to provide the data necessary to perform an optimisation study, it was necessary to measure and log a number of system parameters. The purpose of this paper is therefore to detail the instrumentation requirements of the test rig and to describe the data acquisition system used. Whilst the majority of the parameters could be adequately treated, the monitoring of solute concentrations seemed much less so and measuring accuracies of these remained unresolved.

1 INTRODUCTION

Haemodialysis treatment involves the transfer of molecules across a semi-permeable membrane. This transfer is carried out by two mechanisms. Water transfer is due to a hydraulic pressure differential across the dialyser membrane known as ultra filtration. Solute transfer is largely by diffusion caused by the presence of a concentration gradient.

Early machines such as the Kolff twin-coil kidney relied on using a large volume of dialysate recirculated from a tank in order to maintain a relatively high concentration gradient and so achieve adequate removal rates of uraemic toxins. Modern systems use a single pass supply of dialysate to maintain even higher concentration gradients. Typically, over 100 litres of purified water is used per session in a modern system.

The first step towards producing a system that dispenses with the restrictions of conventional dialysis equipment would be to maintain a high concentration gradient across the dialyser membrane, but with a greatly reduced volume of dialysate. To achieve this it would be necessary to regenerate a small volume of recirculating dialysate.

As early as 1920, Bock[1] was investigating the adsorptive properties of carbon in urine. His findings showed that it was possible to remove all uric acid and most creatinine. Activated carbon was first used in-vivo by Yatzidis[2] in 1964 to detoxify blood directly.

Unfortunately, urea, the most abundant waste product is particularly unreactive and difficult to adsorb. However, Giordano[3] discovered that urea adsorption was greatly influenced by temperature, and so the test rig developed at Coventry University cools dialysate down to 2°C for adsorption. Using this method, the capacity of activated carbon may be increased as much as 3 times.

2 DESCRIPTION OF TEST RIG

2.1 Requirements

In order to investigate the feasibility of a cold carbon system, it was required to simulate a dialysis session lasting four hours. The dialysate side of the system therefore needed to deliver dialysate to a dialyser at a temperature of 37°C and a variable flow rate from 250ml/min to about 500ml/min. Dialysate pressure at inlet to the dialyser needed to be controlled in order to control ultra filtration. After passing through the dialyser, the dialysate should be cooled prior to entering a carbon cartridge regenerator. Cooling was to be variable from 37°C down to 0°C in order to study the effects of temperature on regenerant effectiveness.

2.2 Proposal

To meet the requirements stated above, a test rig was developed and a system diagram is shown in figure 1.

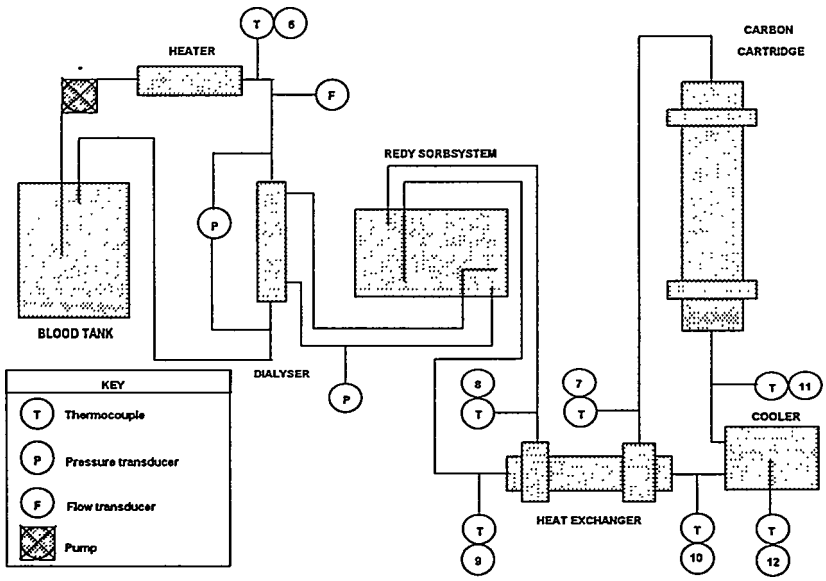


Figure 1 System Diagram

Dialysate was stored in a blood tank reservoir at a constant temperature of 37°C before being pumped at negative pressure through the dialyser. The dialysate would then pass through one side of a heat exchanger to reduce its temperature prior to entering the cooling unit. The cooling unit used peltier effect modules to pump heat from the dialysate. Having been cooled to the required temperature, the dialysate was filtered through the carbon adsorbent. It then returned through the other side of the heat exchanger back to the reservoir.

35 litres of purified water with various solutes added was used to represent the body fluid. During a test, this was pumped through a heater to raise its temperature to 37°C then through the dialyser before returning to the reservoir.

2.3 Instrumentation

There are a host of data acquisition systems on the market and the selection of a system for a particular application can be a difficult decision. The following considerations needed to be made when selecting the system:

1. Number of channels required.
2. Type of signals i.e. single ended or differential.
3. Sampling frequency.
4. Dynamic range of signals.
5. Is signal conditioning included?
6. Alarms for under and over range signals.
7. Local display of real-time data.
8. Data manipulation and storage requirements.
9. Flexibility.
10. Cost.

In the application of data logging to the development of a portable haemodialyser, the parameters required to be measured were as follows:

1. Temperature at nine points on the test rig.
2. Flow-rate at two points.
3. Pressure at two points.

2.4 Temperature

Due to the number of temperature points to be measured, it was decided to use thermocouples. Insulated K type (chromel/alumel) thermocouples were used as they have small linearity errors within the test rig's temperature range and low offset errors. These thermocouples used in conjunction with an Analogue Devices AD595 thermocouple amplifier give a sensitivity of 10mV/°C. Errors in readings due to stagnation conditions were considered negligible due to the low flow rates being investigated. The frequency response was not considered a problem, as the thermocouples' response was limited by the high thermal inertia of the test rig fluid circuits.

2.5 Flow rate

A low cost flow sensor giving a variable pulsed output relative to liquid flow rate was used as the flow transducer. This type of transducer can be used for flow rates from 25ml/min up to 1500ml/min, with minimal pressure drop across the device. The electrical pulses derived from the output of the flow sensor were used to drive a frequency-to-voltage converter. The components for this device could be selected to minimise ripple on the output and to set the gain of the converter. This control of the frequency-to-voltage conversion effectively limited the frequency response of the derived signal to less than 1Hz.

2.6 Pressure

A miniature pressure transducer of the piezoresistive bridge type was chosen due to the relatively high sensitivity of this type of device. As depression was to be measured on one of the devices, a differential pressure transducer was chosen, with the depression being measured with respect to atmosphere on the low side. The sensitivity is in the order of 9mV/0.1bar. The pressure transducers had a response time of 1ms. As the pressures to be measured were steady with manually controlled variations, frequency response was not considered to be a problem.

3 SELECTION OF DATA ACQUISITION SYSTEM

The total number of transducers (channels) required to be monitored was thirteen. Commercial systems normally cater for a binary number of inputs, due to the method of decoding used to control the internal multiplexer, therefore a sixteen channel facility was specified. This would allow for further expansion of input devices.

As all the transducers required some form of signal conditioning, each one was designed to be single ended, which simplified the input requirements to the data logger. The low frequency response of the transducers and the parameters would indicate that a relatively slow acquisition system could be used. This would allow time for the channels to be scanned relatively fast and the data displayed for real-time monitoring. The use of a PC would enable storage of large quantities of data, the display of real-time information and data manipulation.

In the end, a low cost acquisition board (PC27) was selected from the Amplicon range of products. This board plugs directly into the interface bus of an IBM PC/AT or XT, and has an IDC edge connector with individual ground pins for each of the sixteen channels. The board has a sixteen way multiplexer which was software controlled. The 12 bit analogue to digital converter is of the successive approximation type with a sampling time of 10μs. During the sampling period, the data is held constant by an onboard sample and hold amplifier. The period between samples can be software controlled to minimise settling errors.

It was possible to select the input range required, i.e. +/-2V, +/-4V bipolar or 4V unipolar. The latter was chosen, giving a predicted resolution from the 12 bit converter of 1mV. This gave a theoretical resolution of 1°C on the temperature measuring channels, a 0.025% of full scale resolution on the flow measurements and 0.25mbar resolution for the pressure readings.

The data acquisition card came complete with sample software, written in Turbo Pascal; this also provided the necessary routines to control the multiplexer, and reading of data from the PC's data bus. The sample software allowed the system to be tested at various stages of design, therefore control software and hardware manufacture could be developed simultaneously.

4 SOFTWARE

As mentioned in the previous section, sample software and routines came with the acquisition board so the software development requirement was to provide a real-time display of the operating parameters during the test and periodic storage of data onto disk. Some experimentation was also required to set the delay to allow for settling of the selected channel on the multiplexer. A flowchart of the main routines of the software is given in figure 2.

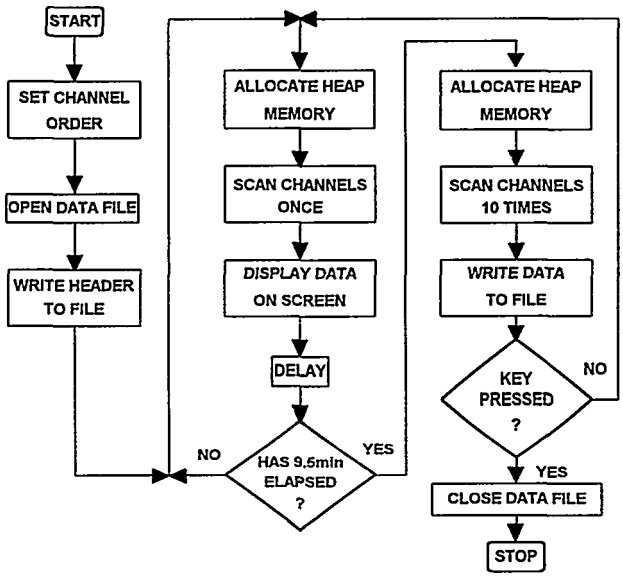


Figure 2 Main Data Acquisition Program

To allow for continuous display of the operating parameters, it was decided to simply scan all the channels being used once, this display was then updated every half a second and allowed the condition of the test rig to be monitored throughout a test. To coincide with fluid samples being drawn off from the dialysate and body fluid, at 10 minute intervals, all the channels were scanned 10 times and this data written to disk for later analysis. Ten samples were taken so that an average reading could be found for each channel. Ten minutes was taken as the time interval as the thermal inertia of the system would mean that no large increases in values would have occurred. Also at the analysis stage there would not be an overwhelming amount of test data.

At the analysis stage, a second Pascal program was developed to calculate heat flow in the heat exchanger and cooling unit. Values of specific heat could be calculated for average fluid stream temperatures by Lagrangian interpolation of a look up table, contained within a second data file of fluid properties. Once these values were known, heat transfer could be calculated from the temperature difference and mass flow rate. Results from the analysis program were then imported into a spreadsheet for graphical presentation (see figure 3).

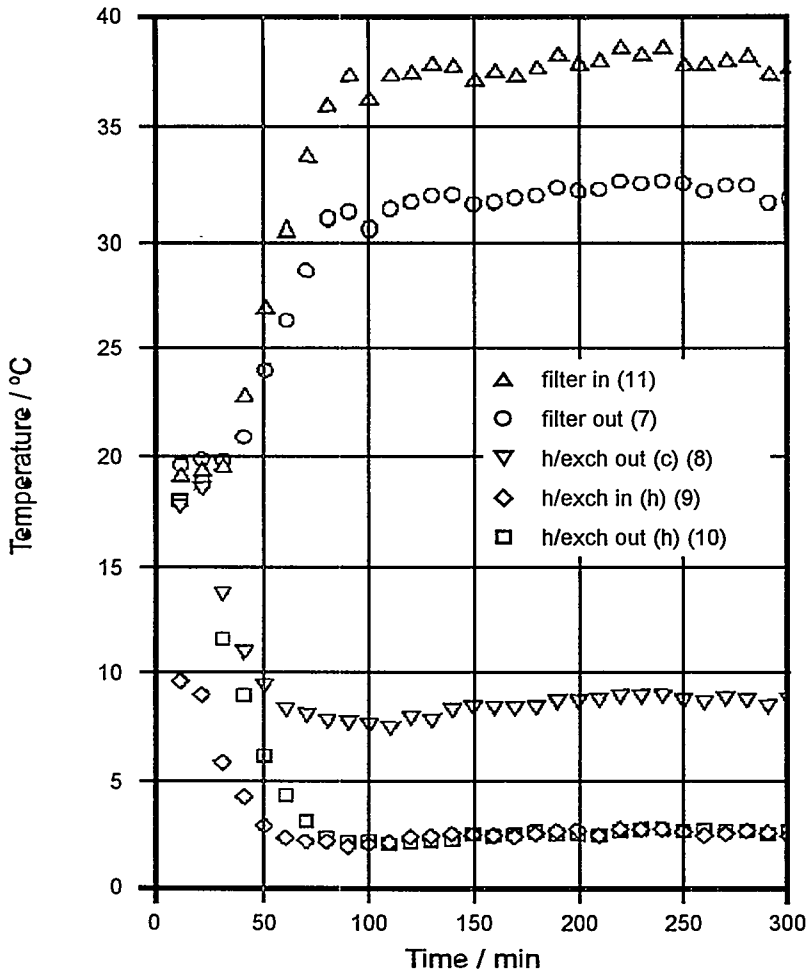


Figure 3 Dialysate Operating Temperatures

5 MONITORING OF SOLUTE CONCENTRATIONS

The main area of interest of this research was the adsorption of urea by activated carbon. Unfortunately, no sensors are available for instantaneous measurement of urea concentration, so on-line monitoring of the performance of a particular test was not possible. Instead, samples were taken periodically from the dialysate reservoir and the body fluid reservoir and the urea concentration found by colorimetric methods at a later date. From these results, graphical output of the variation in urea concentration with time was obtained using a spreadsheet (see figure 4).

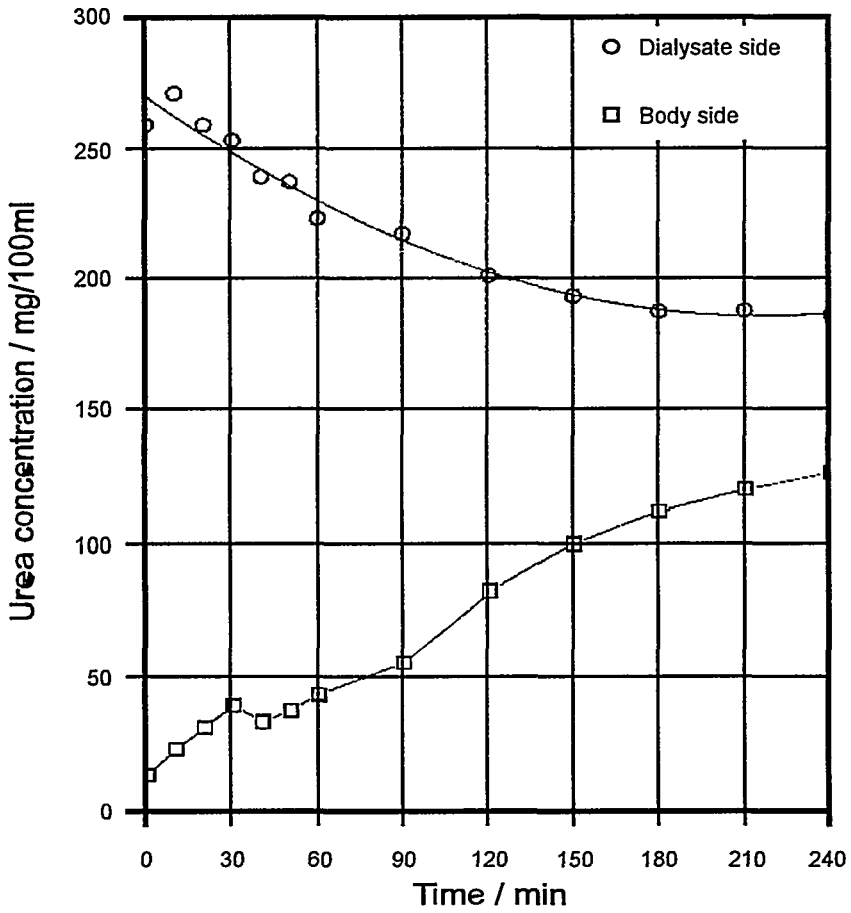


Figure 4 Urea Concentration during 4 Hour Test

6 CONCLUSIONS

From the results obtained using this test rig to simulate haemodialysis, it has been possible to model the effects of temperature, flow rate and dialysate volume on the adsorption of urea. This model used in conjunction with the data obtained by the acquisition system has made it possible to produce an optimised design for a portable haemodialysis machine, in terms of its weight, bulk and cost.

ACKNOWLEDGEMENTS

The authors are very grateful to Dr. David Dukes of the Walsgrave Hospital, Coventry for the many helpful discussions and practical help during the course of the work described. The technical advice of Mr. John Martin of the School of Engineering at Coventry University was of great value.

REFERENCES

1. Bock JC: 'A study of a decolourising carbon' *J. Amer. Chem. Soc,* 42:1564, 1920
2. Yatzidis H: 'A convenient haemoperfusion micro-apparatus over charcoal for the treatment of endogenous and exogenous intoxications. Its use as an artificial kidney' *Proc. Eur. Dial. Transplant Assoc.* 1:83, 1964
3. Giordano C, Esposito R, Bello P: 'A cold carbon depurator for the adsorption of high quantities of urea' *Kidney Intern,* 10:S284, 1976

THE DESIGN ISSUES EFFECTING THE IMPLEMENTATION OF A SMART ELECTRONIC CONDITION MONITORING SYSTEM DESTINED FOR HARSH ENVIRONMENTS

G. Horler, Lecturer.
University of Lincolnshire and Humberside, Cottingham Road Campus, Cottingham Road, Hull. HU6 7RT. United Kingdom.

Abstract

The aim of the paper is to present the considerations made during the design and implementation of a condition monitoring system; the physical constraints placed upon the system design and the choice of implementation technology. It is anticipated that the choices reported are familiar and representative of electronic monitoring system design. Observation of the system design reveals how technologies were adopted or discarded as the design progressed. Consequently the paper provides a summary of the technological attributes of the devices entertained and the supporting tools. This feature summary should be helpful for those wrestling with their own condition monitoring systems.

The background of this project is to realise a system capable of real time acquisition of data from the piston of an internal combustion engine under normal loading conditions. This project has developed from a dedicated hard-wired transducer network installed within an engine [1][2].

1 Introduction

Previous attempts at extracting parameters, such as temperature, from the piston of an engine under load, have enjoyed limited success [3]. Key limitations have been reliability (system failure rate) and flexibility (system adaptability). These limitations have prevented prolonged and repeatable experimentation; hence the effects of in situ environmental and technological variables have not been quantified.

A matrix cataloguing the variables effecting the system is listed in Table 1. Subsequently an operational model for a multiple sample point monitoring system was developed; in turn suggesting a system architecture and design strategy.

System Elements	Electrical	Mechanical	Environmental
Transducer	Analogue or digital. Active or passive. Type (Resistive etc.) Calibration.	Screw or glue mount attachment to piston. Thermal continuity High G force stresses Size, robustness	Thermal linearity, offset and drift. Corrosion, sealed technology.
A/D Conversion	Quantiser linearity. Type (Flash, S.A. etc) Error correction	As transducer	As transducer
Codec	Protocol RTL, BIM, 1X1. Signature and packet generation /recognition. Error correction.	As transducer	As transducer
Transceiver	Modulation AM/FM Digital or Analogue Error correction. Technology (pcb, surface mount etc.)	As transducer plus: Doppler effects. Crankcase topology, line of sight, object piercing, reflections and standing waves.	As transducer plus: Attenuation through absorption, and reflection. Stability
Power	Generation and storage Battery, generator management (Sleep circuit and brownout protection)	As transducer plus: Access	As transducer

Table 1: System Variables

Section 2 outlines the modus operandi of the system. Section 3 describes the design and implementation of selected system elements. Design tools and methods are overviewed in section 4 before a concluding section.

2 System Model and System Development

A signature coded, half duplex protocol was chosen; due to simplicity. An overview of the protocol is presented in section 2.1 and the system blueprint described in section 2.2. Various data coding strategies are considered in section 2.3.

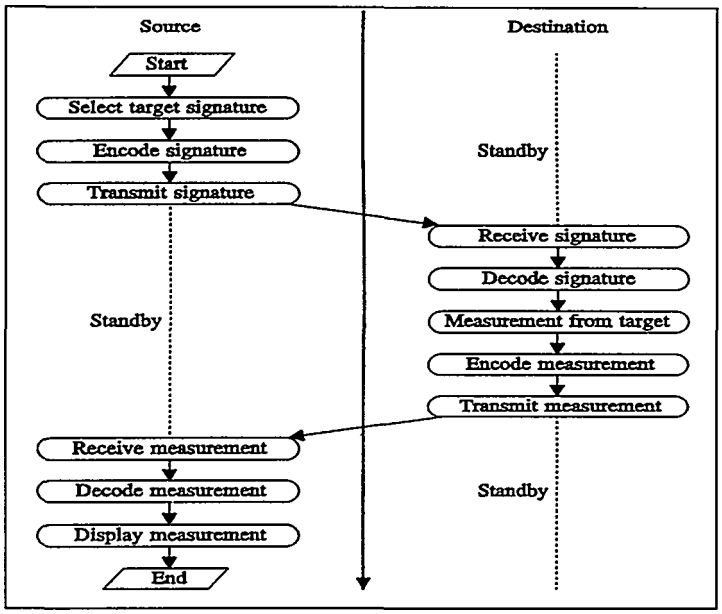


Figure 1: Signature Coded Half Duplex Protocol

2.1 Signature Coded Half Duplex Protocol

Each measurement point in the system is given a unique code (signature). To take a measurement, a signature is transmitted from the source to the destination. The destination electronics provide access to sampling points; however data is only taken from the point which recognises the transmitted signature. The measurement is transmitted back to the source (with or without the signature). This single transmission then reception of data is called half-duplex. Figure 1 graphically depicts the half-duplex strategy.

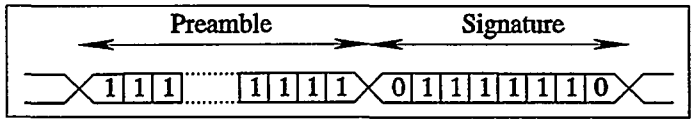


Figure 2: Signature Data Packet

The signature is based on an eight bit code (possible 256 transducers) with a preamble of ones. A preamble and signature once merged is called a packet; packets transmit signature and measurement data around the system. A preamble is required to notify the recognition circuitry that a signature is imminent and also to synchronise bit slicers used in many transceiver designs [4][5].

2.2 System Blue Print

The system is presented in Figure 3. The target signature (source request measurement) is inputted to an encoder. The encoder converts this signature into a format compatible with the transmission system, section 2.3. The encoded signature is transmitted to and decoded by the destination decoder. The encoder adds on the preamble and any error correction codes [6] prior to transmission.

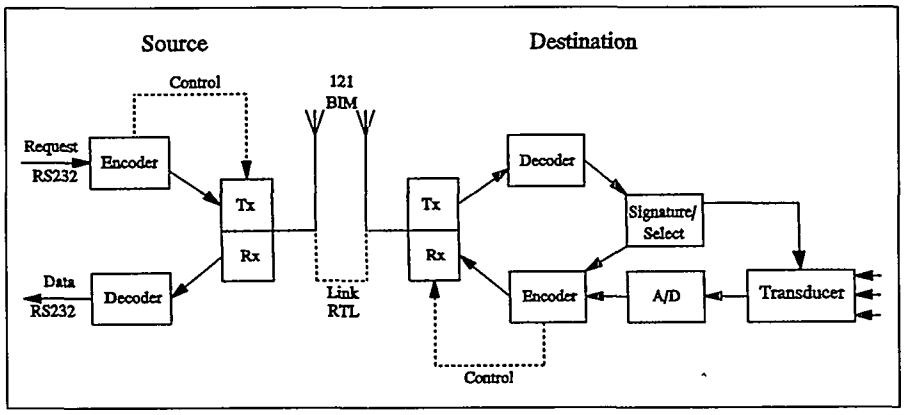


Figure 3: System Blue Print

The destination decoder receives the transmitted data packet and strips off the preamble; subsequent signature and clock recovery, enable the signature recognition circuitry to analyse the signature and invoke a sampling if a signature is recognised. After sampling the data is packetised by the destination encoder prior to transmission back to the source. On reception, the preamble is stripped and the sampled data outputted for analysis or recording.

2.3 Data Coding

The non-return-to-zero (NRZ) data code is ubiquitous. Others exist such as the return-to-zero (RZ), return-to-level (RTL), Biphase Mark (BIM) and the 1x1 codes [7]. These are detailed in Figure 4. NRZ, RZ and RTL are self explanatory. The Biphase Mark code, phase reverses the previous data code bit if a data one is represented. The 1x1 code divides each bit period into fractions; if x equals 2, the period is split into quadrants. The first and fourth quadrants equal zero and one respectively, quadrants two and three take on the data value.

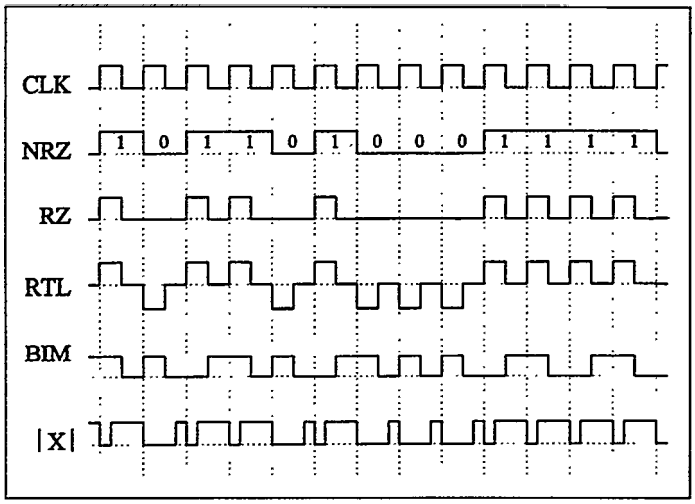


Figure 4: Data Coding Schemes

The RTL, BIM and 1x1 codes incorporate the data clock into the data stream by virtue of thresholding, phase reversal and edge detection respectively. All have advantages and disadvantages affecting encoder, decoder (codec) design, these are investigated in section 3.

3 Codec Design

As seen, encoding of the system clock into the data stream can be accomplished using three techniques. The following sections outline the codec design and implementation for the RTL, BIM and 1x1 formats.

3.1 Return to Level (RTL)

The RTL encoder and decoder circuits are based on analogue electronic elements. Input NRZ data to the encoder Figure 5 is double buffered; the RTL action achieved by switching this input with Vcc/2 under the control of the system clock.

The RTL decoder circuitry is shown in Figure 5. System clock recovery is achieved by using a window comparator; data recovery via a level comparator 'ANDED' with the recovered system clock [8].

The RTL codec has one drawback, the effective d.c. level is lost if a.c. is coupled to a transmitter, hence system clock recovery is impossible. Despite the unsuitability for a.c. coupled systems the RTL codec presented provides a cheap and reliable module for single hard wired

bus systems.

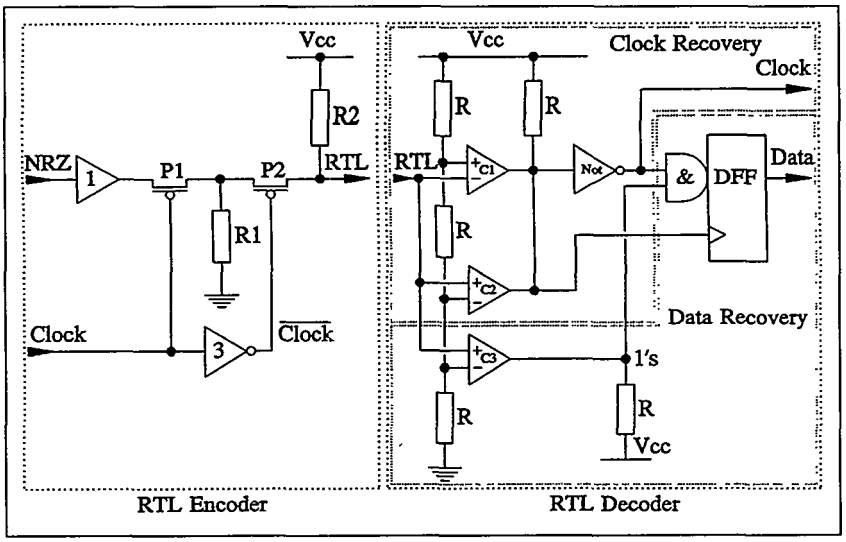


Figure 5: RTL Codec Design

3.2 Biphase Mark (BIM)

The sequence of transmitted levels, (high to low, low to high) is determined by the previous bit phase orientation. If a zero is to be transmitted no change is observed. A data one results in a phase reversal from the previous bit. The beauty of this technique is the elegance of the decoder, a single XOR logic gate, Figure 6.

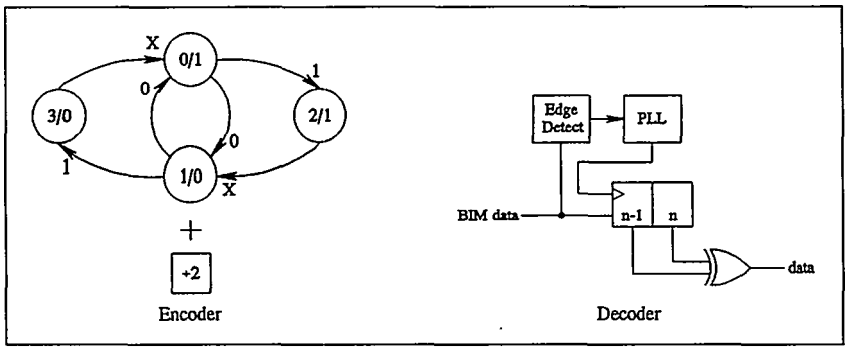


Figure 6: Biphase Mark (BIM) Codec Design

A significant drawback however is clock recovery, which relies upon a Phase Locked Loop (PLL), resulting in a larger more complex circuit, Figure 6. The effectiveness of the BIM scheme is further compromised if long sequences of 'ones' are encountered

(preambles). This is due to transmitted data ones appearing as a clock with twice the period.

3.3 1x1 Data Codec

The 1x1 label defines how each bit period is partitioned, Figure 4. For quadrature partitioning the state machine clock frequency must be four times the data clock frequency. Decoding is achieved by sampling the 1x1 data stream within a window, some period after a high to low edge transition. The encoder state machine and programmable decoder are presented in Figure 7.

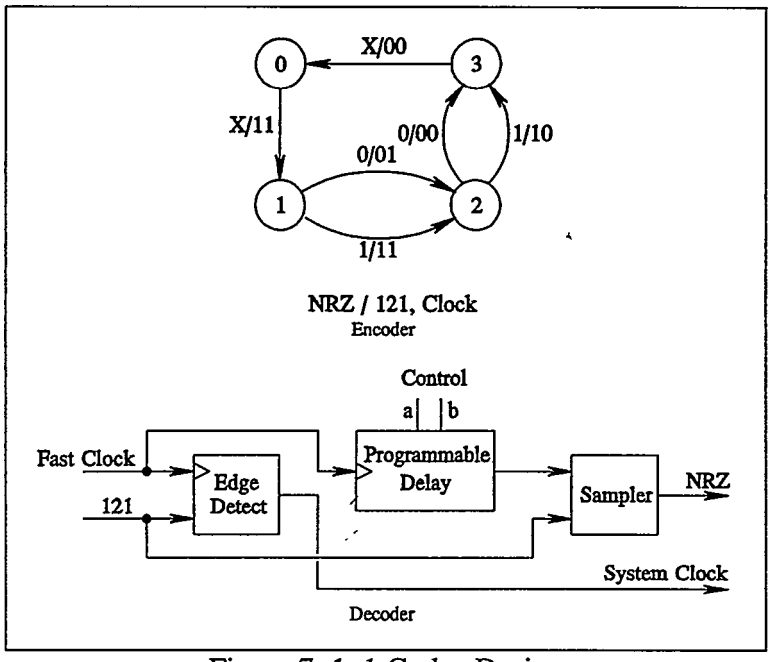


Figure 7: 1x1 Codec Design

The simulation of the decoder in Figure 7 can be found in Figure 8.

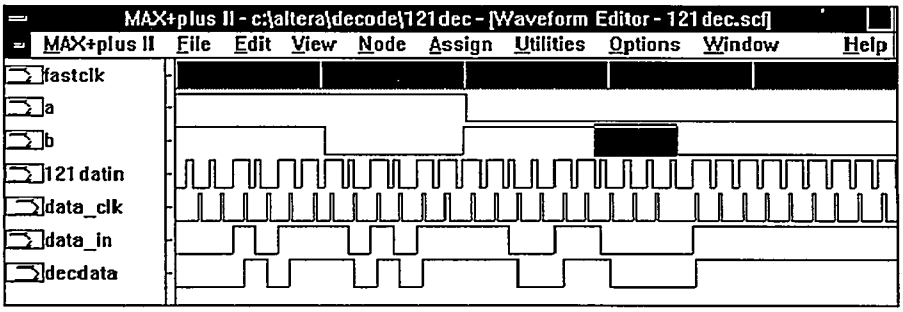


Figure 8: 1X1 Decoder Simulation

4 Technologies and Tools

The sequential nature of the RTL, BIM and 1x1 codes suit formal state machine designs [9], which can be implemented in Altera [10]. Additionally, sequential data lends itself to processing by a microcontroller. The Microchip PIC 16C71 microcontroller [11] was used successfully to implement the 1x1 strategy. The following sections outline the features, virtuous or otherwise of the technologies and tools.

4.1 Altera

Altera provides an integrated environment for the development of digital circuits and systems. Computer aided design of circuits using hardware description languages or place and route are available. What you see is what you get (WYSIWYG) simulation of circuits, Figure 8, is followed by downloading to a family of device types result in efficiency and confidence. Another key feature is the availability of complex devices such as PLL's in the standard libraries. A fast learning curve (hours) and high data processing rates are attractive.

On chip a/d is not possible, furthermore, rudimentary design iterations usually require a full systematic redesign. The Altera design system cost is approximately £3,000.

4.2 PIC 16C71 Microcontroller

Built in a/d conversion permitting a maximum of 4 analogue inputs, a very small device footprint and additional features such as edge and level triggered interrupts, brownout protection and a sleep facility suggests that this device is tailor-made for the task in hand.

These benefits are offset by a much longer learning curve (weeks). However, the PIC is a simple RISC processor and is suitable for both novice and experienced microprocessor users. A low cost (£150) entry level system is favourable. A further £1,000 secures in circuit emulation (ICE); very efficient, though not essential, for timing debugging. A digital analyser or multi channel oscilloscope is a necessity for debugging and fine tuning designs, particularly decoders.

The major drawback of the microprocessor route is the lack of transparency of the design method. Simple tasks such as serial data conversion soon blossom into complicated programs. The clarity lost in the state machine designs are traded for functionality and flexibility.

Microcontroller programming uses a Windows™ based program editor, assembler and simulator provided by ICEPIC™ [12]. The entry level PICSTART™ 16B1 programmer is used to download programs to target devices; a proprietary development board used to develop codec systems is shown in Figure 9.

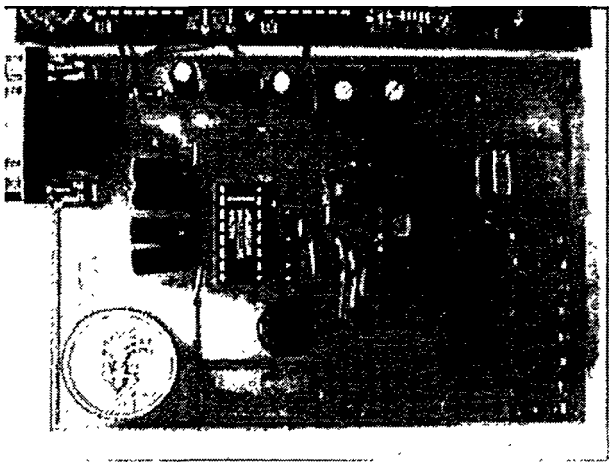


Figure 9: Microcontroller Associated Hardware

Conclusion

The devices and environments considered are excellent. Table 2 provides a comparison of the salient features. Both technologies are suitable for realising the system (in full or in part) under consideration.

Condition monitoring systems using serial communication techniques at low bandwidths, benefit from the PIC simplicity and features. As data rates and processing requirements increase Altera comes into its own.

Thus the PIC 16C71 was chosen as the implementation technology for the bench test system. A system has been built and bench tested and found to work satisfactorily. Future work will see the system transferred to a model and then to an engine. The results will be available for COMADEM 1998.

	Altera	PIC 16C71
System Cost	£3,000	£150 Entry level £1,000 ICE
Tools	Integrated graphical user environment device programmer	Text editor, Assembly language simulator, emulator device programmer
Device cost	£15.10 -> £60.00	£25.20 (flash), £5.88 electronic erase
Design methodology	Art, formalised design techniques (state machine design)	Structured programming
Simulation	WYSIWYG	Pseudo functional
Computing platform	486 or better 16M+ram Windows 3.1 or 95	Entry level 286 DOS ICE as Altera
Virtues	High data bandwidth Complex design capability. Library of standard digital functions. Intuitive, transparent.	Small, cheap, flexible. Well thought out peripheral functions.
Drawbacks	Cost, size, limited peripheral functionality	Maximum 20 Mhz instruction bandwidth. Lack of design transparency. Testing of designs is part of the design process (not wysiwyg)

Table 2 Altera and PIC Features

References

[1] Development of a Piston Temperature Telemetry System. Robert W. Barrah et. al. Electronic Controls and Sensors SAE Special Pub. 1992 Pt 898 pp 1-11

[2] Lubricating oil - extraction and chemical analysis. D. J. Picken et.al.DeMontfort University, U.K.

[3] Accelerated methods for determining the oil deterioration in an internal combustion engine. K. Seare PhD. Thesis, DeMontfort University 1994.

[4] Unpublished feasibility study and proposal for a telemetry system destined for harsh environments. G. Horler University of Lincolnshire and Humberside / DeMontfort University. *

[5] BiM Transceiver Technical Manual. R. F. Solutions U.K. *

[6] Error - control coding and application. Djimitri Wiggert. Artech House. ISBN 0-89006-066-5.

[7] Digital Communication Systems. Peyton Z. Peebles, Prentice Hall International ISBN 0-13-211962-5.

[8] Operational Amplifiers and Linear Integrated Circuits. Thomas L. Floyd pp 98-108, Merril (McMillan Inc.) ISBN 0-02-338641-X.

[9] Fundamentals of Logic Design. Charles H. Roth jnr. Chapter 14 pp 359 West Publishing Co. ISBN 0-314-92218-0.

[10] Altera Databook, Altera Corp. Europe, Solar House, Globe Park, Fieldhouse Lane, Marlow, Bucks. SL7 1TB. U.K. *

[11] Microchip Databook, Microchip Technology Inc., 2355 West Chandler Blvd., Chandler Arizona. 85224-6199 U.S.A. *

[12] ICEPIC, R.F.Solutions and NeoSoft Inc. U. K. *

* For further information contact G. Horler

Telephone: +44 (0) 1482 463376
Fax: +44 (0) 1482 463783
Email: ghorler@humber.ac.uk

OPERATOR AND MAINTENANCE SUPPORT THROUGH PLC LOGIC ANALYSIS AND HYPERTEXT DOCUMENTATION

Johan Plomp, Pertti Huuskonen and Esko Malm
VTT Electronics
P.O. Box 1100, FIN-90571 Oulu, Finland
E-mail: Johan.Plomp@vtt.fi

ABSTRACT

This paper describes a prototype of a support tool for PLC (Programmable Logic Controller) analysis. The tool was motivated by observations regarding the inefficiency of current PLC software debugging tools and the poor availability of cross-referenced documentation and manuals. The prototype analyses the temporal signal dependencies within the PLC logic by means of an automatically extracted high-level PLC logic model and a history of logged values. The user is presented with hypertext explanations enhanced with links for follow-up questions and relevant documentation sections.

1. INTRODUCTION

Industrial machinery is typically controlled by one or a number of communicating PLC's. Operation and maintenance staff usually needs to deal with a great number of devices each controlled by complex PLC programs or program fragments. On the occurrence of a fault, which causes the machinery to halt or to perform an erroneous operation, the operator initially relies on the diagnostics support screen provided with the control system. In certain instances, the support provided by the operator displays is insufficient to solve the problem at hand. The maintenance personnel will subsequently attempt to localise the problem by attaching a debugging device to the PLC and by tracing the signals in the program to find the source of the problem. In this process we have observed the following bottlenecks [1]:

- An operator or maintainer will use the debugger to find input signals which caused the current fault situation. He searches these signals by starting from an output signal which flags the erroneous situation and tracing the signals towards the inputs in order to find the "guilty inputs". Debug tools do not support automatic analysis of the logic circuits, which could help in finding the most likely causes of the error.
- The cause of the error will most likely lie in the past. The signals which caused the error can thus be traced only when a data history is available, which is not the case with conventional debuggers.
- Most PLC code up to date is written in low-level, assembler-like languages, which can at most be graphically presented by function block diagrams or ladder logic. This type of code is difficult to read and the functionality of complex programs is hard to recover. Higher level representations of the program will make the program more clear. When these higher-level design structures have not been provided with the code, they can be reverse engineered by analysing the code. Particularly retrieval of state machines (sequences) is useful in this application domain.
- There is a huge amount of (paper) documentation on the design (mechanical, electrical, operative), most of which is needed to solve the problem. The material is often not available where it is needed and cross-referencing a great number of files simultaneously is a cumbersome task.
- The mental behavioural model formed by the operator based on his observation of the machinery may be wrong or incomplete [2]. In exceptional situations this may cause him to perform incorrect and possibly harmful control operations and may inhibit him from finding the problem cause quickly.

It has been the goal of our LEPO project[1] to create a prototype of a tool dealing with these issues. The prototype includes the following features, which will be elaborated in this paper:

- A logic analyser, which aids the user in his interactive search for "guilty inputs"
- Data history creation by logging the most important variables
- Self-explaining features with context-sensitive on-line help

1 LEPO is short for "Logic Explanation to Process Operators". The research has been carried out in a national research project with industrial partners. Cases from the partner's systems have been analysed with the tool.

- Context sensitive hyperlink creation to relevant sections in the design documents
- (Semi-) automatic extraction of higher-level descriptions of the software, in particular state machines and dependency graphs

2. MODELLING THE PLC LOGIC

We have developed an object-oriented model of the logic code for simulation and analysis. The elementary operations are represented by function blocks which are interconnected according to the data flow in the code [3]. The function blocks are grouped according to the code's block hierarchy and a model of the code preserving the calling hierarchy is constructed (Figure 1). This approach enables us to delegate simulation tasks to the elements of the code and provides us with hooks for data history and help information. The internal model closely resembles the Function Block Diagram, as used in IEC 1131-3 [4, 5].

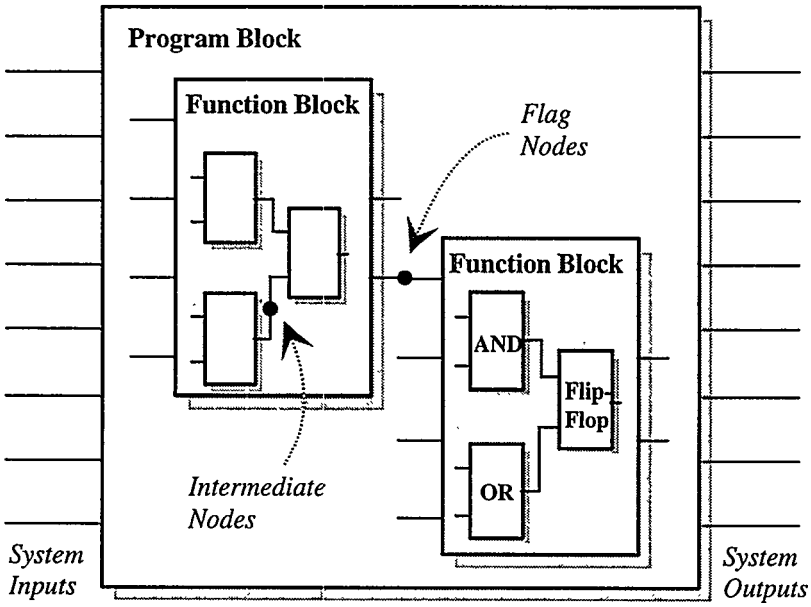


Figure 1. Hierarchical block structure of the internal model

2.1 Enhancing the model with design and maintenance information

We distinguish between three types of information relevant for automation system maintenance:

- Comments in the PLC code, written using the software development tools. These comments are available from the source code or in external files.
- Documentation concerning the PLC code, the controlled device, the electrical wiring, manuals etc., including documentation on the design rationale (design purpose, choices, alternatives, criteria, etc. considered during the design process). This documentation is usually provided in paper form. For efficient use with automated tools, this type of information must be written or converted into an on-line format.
- Maintenance reports; a logbook with the repairs performed by maintenance personnel. This type of reports is rarely available in a structured manner.

This information is attached to the objects (signals and blocks/gates) in our system and is used to enhance our textual code explanations. Not only does this information improve the readability of our explanations, but also relevant links to interesting sections of the hypertext documentation can be provided. The links can be automatically generated for the maintenance reports written with the tool. For other external documents, links can be provided to already existing keyword indices.
The provision of support for context-sensitive fault-report storage was received very well by our industrial partners and seems to eliminate a serious flaw in documentation management.

2.2 Enhancing the model with history information

History data can be collected using a (permanently attached) logging device. This device will have a circular buffer (for testing purposes we currently use a file), so that the history data is available for a limited period of time, depending on the buffer size and the amount of relevant data. The logged data is read into our model and subsequent simulations make use of this data history.
PLC inputs, outputs and flags are most suitable for logging, since their number is limited (in the case of S5) and they determine the PLC

state almost completely. We also chose to log timer signals, since they were implemented with separate bits in S5.
The logged data does not determine the state of the PLC completely, since e.g. memory bits are missing and part of the internal signals are not visible on the outside. Since we know the logic determining the internal signals' state, we have a set of constraints which can be solved in order to determine the state of unlogged signals. This will succeed in most cases, otherwise we have to set the signal to an unknown state. Signal state simulation is done whenever needed, but only for the relevant part of the network at a given time.

3. FINDING "GUILTY INPUTS" BY LOGIC ANALYSIS

A typical operation performed implicitly by operators and explicitly by maintenance personnel is the search for "guilty inputs" starting from "faulty outputs". An error in machine behaviour usually reveals itself to the operator by a machine part being in the wrong position or not moving to the right position, by an error message from an interlock, or by another particular signal not following its normal behaviour. Most of these observed errors can be associated with an output signal of the PLC, which is not in its normal or desired state.
When the fault is not obvious, maintenance personnel will start searching in the logic program for inputs which could have affected the output. In this search they appear to use heuristics, including ruling out signals with the "wrong" state (e.g. a "1" input cannot cause the output of an AND to be "0"), and experience (importance and "normal" state of signals). When "guilty inputs" are found by this iterative method, the devices connected to the inputs (usually sensors) are investigated, and the actual cause of the fault can be deduced.
With the logic model, the tool can automatically analyse the logic and aid the maintenance engineer in performing this demanding and time-consuming task. Upon receiving a query the tool traverses the network backwards and reports on the (input) signal(s) that are causing the faulty behaviour. There are two types of queries that may be of interest for the operator and the maintenance engineer:

- *How did we get here?* - The network is analysed in order to find the signal transitions causing the current situation. History information plays an important role in this task and makes this analysis more powerful than manual searches by conventional tools. Heuristics limiting the search space include disregarding signals not contributing to the current output value and signal transitions not causing the observed output transition.

- *How can we get out of here?* - The network is analysed for what is inhibiting the logic to resolve towards a normal or desired state. This method relies less on history data, but uses graph search techniques while incorporating signal value related heuristics to limit the search space.

3.1 Dependency graphs

The value of the output signal used as a starting point of our query depends on the values of other (input) signals, either at the same time or at some time in the past. This dependency can be modelled as a graph of basic logic functions (AND, OR, NOT) interconnected with temporal signals (a signal with a time span). This dependency graph is a high-level structure which can be used as an intermediate representation for the generation of the "guilty input" analyses. It is still being researched how the graph as such could be used to help the user to understand the PLC code. The graph is first constructed to reflect the logic model's signals in a one-to-one fashion, but may be flattened to a canonical form (sum-of-products or product-of-sums) and subsequently minimised to give an as small dependency graph as possible.

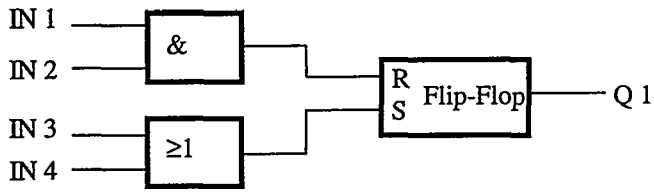


Figure 2. An example logic circuit

In order to illustrate the "guilty input" and dependency graph methods, examine the example circuit in Figure 2, and its timing diagram in Figure 3. The following table shows the answers to the questions *"What caused the current state of Q 1 at the given time"* and *"what is keeping the current state of Q 1 from changing at the given time"*. The relevant time is added to the signal names between parenthesis.

Time	Cause of current state	Keeping current state
5.5	IN 1(3)	IN 1(5.5), IN 2(5.5), IN 3(5.5), IN 4(5.5)
6.5	IN 1(3)	IN 3(6.5), IN 4(6.5)
10	IN 3(7)	IN 1(10), IN 2(10)
13	IN 1(11), IN 2(11)	IN 3(13), IN 4(13)

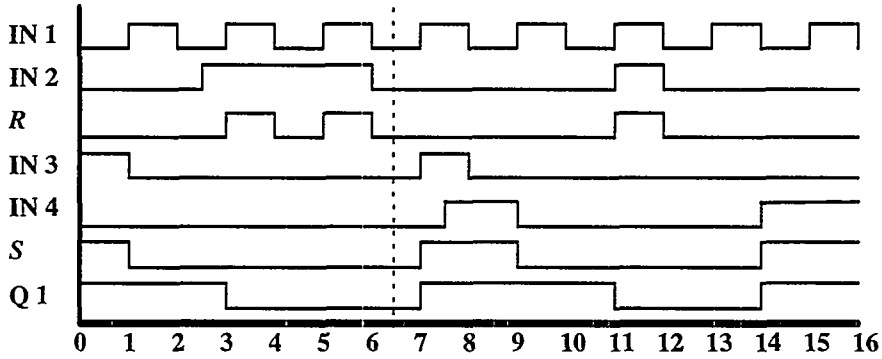
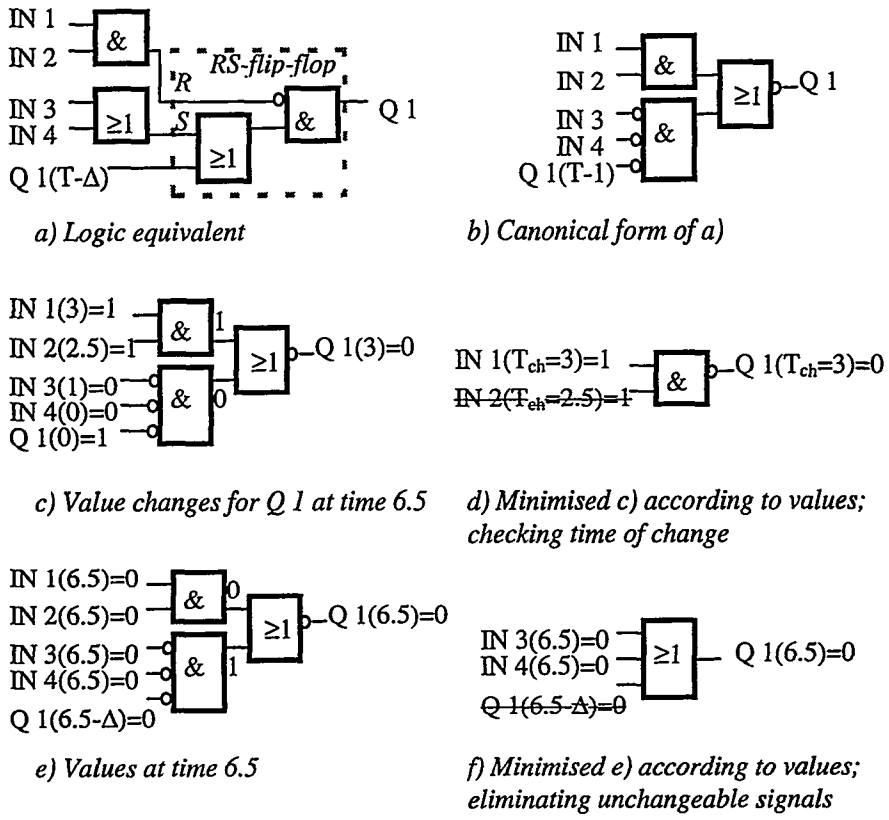


Figure 3: Timing diagram for the example circuit in Figure 2.

The dependency graphs for the questions at time 6.5 would look like depicted in Figure 4. First a general equivalent for the circuit is derived, as shown in a). The equivalent is minimised and transformed to the canonical form, depicted as a sum of products in b), by well-known algorithms and logic transformations.

For the *cause* of the current state, we determine for each signal at what time it changed its value, starting from the output. The relevant signal changes are shown in c), while d) shows the situation after the irrelevant inputs, according to their values, are removed. Finally we can decide the actual cause by only considering the inputs which changed state at the same time as the output of the observed gate. In words, the answer is: *"The state of Q 1 at time 6.5 is '0' (since time 3), because IN 1 changed to '1' at time=3.*

The question what is *keeping* the current state can be answered similarly by observing the equivalent at the current time as shown in e). This equivalent can likewise be minimised according to the values and f) results. Signals which cannot be influenced by us, are irrelevant for the answer and thus we can eliminate the previous value of Q 1. In words, the answer is *"The state of Q 1 at time 6.5 can be changed to '1' by changing either IN 3 or IN 4 to '1'"*.



a) Logic equivalent

b) Canonical form of a)

c) Value changes for Q 1 at time 6.5

d) Minimised c) according to values; checking time of change

e) Values at time 6.5

f) Minimised e) according to values; eliminating unchangeable signals

Figure 4. Dependency graph examples at time 6.5

The presented table shows, that in most cases the actual number of inputs affecting the output in a given time-frame is significantly smaller than the number of inputs connected via the logic network to the output.

4. STATE MACHINE EXTRACTION

State machine extraction is an important feature in our prototype. The state machines provide a higher level description of part of the code, and thus aid the user in understanding the software. Moreover, the state machines allow us to limit our searches for "guilty inputs", since a state can be regarded as a "guilty input". The user can continue the search from the state machine description, if the state was not a sufficient explanation for the query.

There are a great number of different ways to implement state machines using PLC languages. Therefore it is not a trivial task to extract state machines without any knowledge of the implementation

method. We have implemented an algorithm for the retrieval of state machines from the PLC code, which uses some assumptions about the implementation method found to hold in our example cases. The algorithm requires some elementary information from the design process about the mapping of signals to states.
The state machines may also be used to provide support for the simulation process, since they place extra constraints on the values of the signals. This can help to determine the value of signals which would otherwise be undetermined.

5. INFORMATION PRESENTATION

The prototype has been designed to be used with a logic viewer. The viewer is used to present the logic to the user in a suitable and familiar way, to identify objects and to present the signal values at a given time. The explanations generated by our prototype are mostly textual and we use a commercial browser (Netscape / Internet Explorer) to present our information. All presented information include dynamically generated links from which follow-up queries can be started [6]. The explanations also include dynamically generated links (based on keyword search) to design documentation, if available.

6. DISCUSSION

Our prototype tool attacks the important industrial problem of making PLC's more understandable. We use many techniques that originate from the research on diagnosis of digital circuits [7]. These techniques are still relatively unknown in the field of industrial control systems, although a few advances have been made [8, 9, 10, 11].
We envision tools, such as our prototype, to be a welcome extension to the debugging environments for PLC's. Especially in large industrial plants, where a non-functional machine may be very costly, the savings achieved by the fast tracking of faults and the simultaneous improvement of insight in the PLC code functionality, will be substantial. The definition of higher-level languages for PLC software will enhance the quality of the user support that our tool can give, especially when the information needs of the tool are anticipated during the software development process. Providing design documents in electronic form improves their availability, while increasing cross-reference support through hyper-links. The modular

design of our system furthermore allows remote servicing, because the main interface of our system can be used via internet/intranet.

ACKNOWLEDGEMENTS

This project has been funded partly by the Technology Development Centre in Finland (TEKES). We also gratefully acknowledge the support of our industrial partners.

REFERENCES

[1] Huuskonen, P., Kaarela, K., Okkonen, J., Väisänen, A., "Explaining Control Logic to Process Operators", 8th Int. Conf. on Industrial & Engineering Applications of AI & Expert Systems, Melbourne, Australia, 5-9 June 1995, pp 203-211.

[2] Norman, D.A., "The psychology of everyday things", Basic Books, New York, 1988, 257 p.

[3] Plomp, J., Huuskonen, P., Malm, E., Paanasalo, J., "An Object-Oriented Approach to PLC Explanation", STeP '96, Vaasa, Finland, 20-23 Aug. 1996, 10 p (available from the authors or the conference secretary, Univ of Vaasa).

[4] "IEC 1131-3, Programmable Controllers, Part 3: Programming languages", Int. Electrotechnical Commission, 1993, 207 p.

[5] Lewis, R.W., "Programming industrial control systems using IEC 1131-3", IEE., England, 1995, 293 p.

[6] Moore, J.D. & Mittal, V.O. "Dynamically Generated Follow-up Questions" IEEE Computer, Vol.29, No.7, July 1996, pp. 75-86.

[7] Kleer, J.D.: "How Circuits Work", Artificial Intelligence 24, 1984, pp. 205-280.

[8] Falcione A. & Krogh B.H. "Design Recovery for Ladder Logic", IEEE Control Systems, Vol. 13, No. 2, April 1993, pp. 90-98.

[9] Moon, I., Powers, G.J., Burch, J.R., & Clarke, E.M. "Automatic Verification of Sequential Control Systems Using Temporal Logic", AIChE Journal, Vol. 38, No. 1, January 1992, pp. 67-75.

[10] Guasch, A., Milne, R. & Sarrate, R. "Relay Ladder Logic Diagnosis", Artificial Intelligence in Real-time Control 1994. Elsevier Science Publishers, 1995. Pp. 275-280.

[11] Milne, R. & Guasch, A., "Automatic Diagnostic Development Based on a Programmable Logic Controller", Applications and Innovations in Expert Systems II, SGES Pub., pp. 99-110.

SELF-SENSING COMPOSITES INCORPORATING PIEZO-POLYMER BASED ACOUSTIC EMISSION SENSORS

E R Brown, E J Cobb, G D Neill, J A Steel and RL Reuben
Department of Mechanical and Chemical Engineering,
Heriot-Watt University,
Edinburgh.

ABSTRACT

Recent research on thin composite panels has demonstrated the ability to detect and characterise material damage from resultant structure borne stress waves known as Acoustic Emission (AE). The method, involving modal analysis of full AE waveforms, demonstrates a superior means of noise rejection and source location compared to more conventional parameter based AE detection techniques such as amplitude distribution and event counting. This work exploits these advantages for self-sensing composites.

An additional important feature in discriminating the nature of material damage mechanisms from their AE signature is the use of a broadband AE sensor so as to minimise signal frequency distortion. As a low-cost alternative to most commercially available high fidelity AE sensors, this work is aimed at investigating the use of flexible piezo-polymer film as a broadband AE transducing element for condition monitoring of composite plates.

Preliminary test results are presented from a surface-mounted piezo-polymer transducer for simulated AE sources on a thin composite plate. The results show that source discrimination and source localisation are feasible in a number of frequency ranges. Attenuation measurements used to establish sensor element spacing are also provided.

The next stage of development is to integrate an array of polymer sensors into composite panels to provide an accurate means of damage detection, location and diagnosis of structural integrity.

1 INTRODUCTION

Recent research has demonstrated the ability to distinguish different damage mechanisms in composite panels from a detailed analysis of the acoustic emission (AE) signals generated[1,2]. For AE wavelengths much larger than the panel thickness it is observed that failures such as matrix cracking or fibre breakage generate mainly in-plane extensional waves as opposed, for instance, to surface impacts which produce predominantly out-of-plane flexural waves. This method of source characterisation has prompted study examining the feasibility of incorporating piezoelectric polyvinylidene fluoride (PVDF) polymer based sensors directly within composites to detect these signals and to provide a means of damage identification, location and cumulative analysis.

The thin, lightweight and flexible form of PVDF is particularly suited to the unobtrusive integration into the laminated structure of composites. The metallised surfaces of the film are electrically conductive and can be chemically etched to generate multiple arrays of sensors. In addition, polymer based sensors provide the broad bandwidth response and immunity to ringing required to minimise distortion of AE signals[3]. These aspects combined with relatively low-cost, good acoustic matching and robust mechanical properties are all advantageous in this application. The sensitivity of PVDF to AE is however, much lower than more commonly used piezo-ceramic materials, giving a poorer signal-to-noise ratio. To overcome this problem the PVDF sensor used incorporates a local MOSFET amplifier to improve signal strength.

Results are presented from preliminary experiments to establish a workable sensor configuration using simulated AE sources on a glass fabric reinforced plastic (GFRP) plate.

2 PROTOTYPE SENSOR DESIGN

To establish the practicability of the idea, a circular PVDF sensor, approximately 3mm in diameter, was constructed from 0.04 mm thick biaxially polarised film. The metallised surfaces of the film, which form the electrodes, were connected to a BF998 MOSFET device using short conductive copper tabs and the output coupled using longer leads to an

external power supply circuit. The PVDF was bonded to the GFRP surface using a thin layer of cyanoacrylate based adhesive and then encapsulated in insulating and screening layers.

3 EXPERIMENT

The experimental set-up used to test the prototype design is shown in figure 1. The output from the PVDF sensor was fed via the MOSFET power supply unit to a PAC 1220A preamplifier fitted with an integral 0.1 MHz to 1.2 MHz bandpass filter. The signal was further amplified then captured using a PC29 A/D converter controlled by a 486E microprocessor. A sampling rate of 25 MHz was used throughout.

A 3mm thick GFRP plate to BS3953 was used as a test specimen. The plate lay-up is such that the acoustic properties are equivalent in the 0^0 and 90^0 directions, but wave speed and attenuation can vary at intermediate angles. Artificial AE events were generated using standard Hsu-Neilson[4] pencil lead fractures. To simulate out-of-plane impact sources, pencil lead breaks were conducted on the surface at 4 positions increasing in distance from the PVDF sensor along both the 0^0 and 45^0 orientations. To imitate in-plane motion associated with fibre breakage and matrix cracking type events a series of pencil lead breaks were carried out on the edge of the panel. At each position the result was taken as the average over 10 measurements. Although Hsu-Nielson sources have proven to be a fairly repeatable source, the strength of each lead fracture was 'normalised' against a commercial piezo-ceramic sensor adjacent to the break position, and which also acted as the trigger for the data acquisition system. The amplitude of each break signal measured by the adjacent piezo-ceramic sensor was assigned a calibration factor (which varied between 0.98 and 1.01) and the PVDF signals were multiplied by this factor.

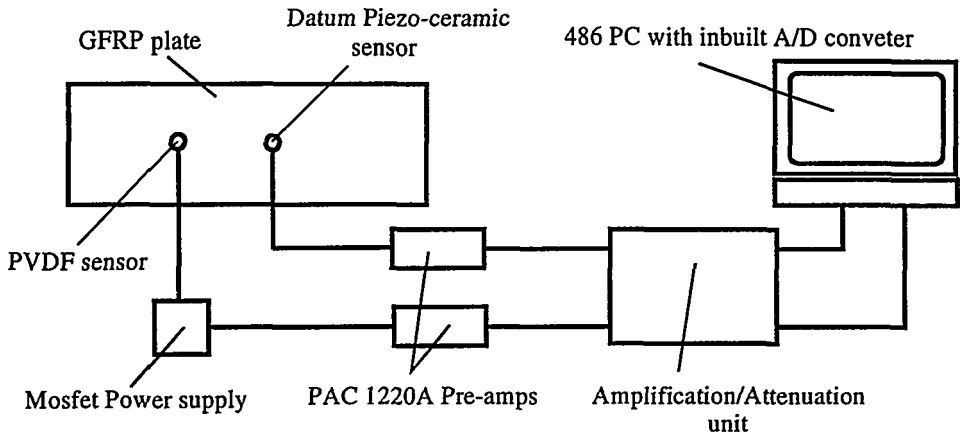


Figure 1 Experimental Equipment

4 DISCUSSION OF RESULTS

A typical AE waveform generated near the sensor on the plate surface is shown in figure 2, with the corresponding result for an edge driven event given in figure 3. In both cases, the dominant mode of plate wave propagation for each type of source is clearly recognisable. In separate experiments using two piezo-ceramic sensors[5] the wave speeds in the plate were found to vary between about 3000 m/s and 1500 m/s for in-plane and out-of-plane waves respectively and wavelengths in the region of 15mm were estimated for out-of-plane waves. Although the sensor is bonded to the surface, the relative magnitude of the two modes suggest it is relatively insensitive to the in-plane motions predominant in extensional waves and responds more to normal displacements generated by the flexural wave.

As a simple measure of attenuation the relative amplitude of the plate waves was obtained for each signal by measuring the maximum peak-to-peak value. These were then averaged for each condition and the results were plotted against distance from the sensor along both the 0^0 and 45^0 directions. These results are shown in figure 4 for flexural waves, and figure 5 for extensional waves. Comparing these attenuation measurements with background noise levels suggests, a maximum sensor spacing of 3 cm would still permit clear interpretation of waveforms. This spacing would achieve a minimum signal-to-noise ratio of approximately 40dB with the limiting mode being higher attenuation of extensional waves along the 0^0 direction. This estimate is considered conservative since it is expected that

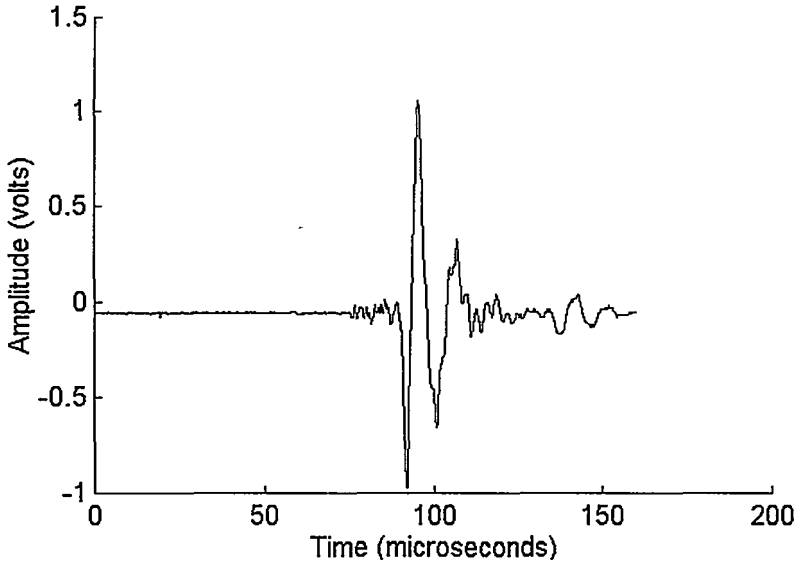


Figure 2 AE waveform generated by pencil lead break on surface of GFRP plate

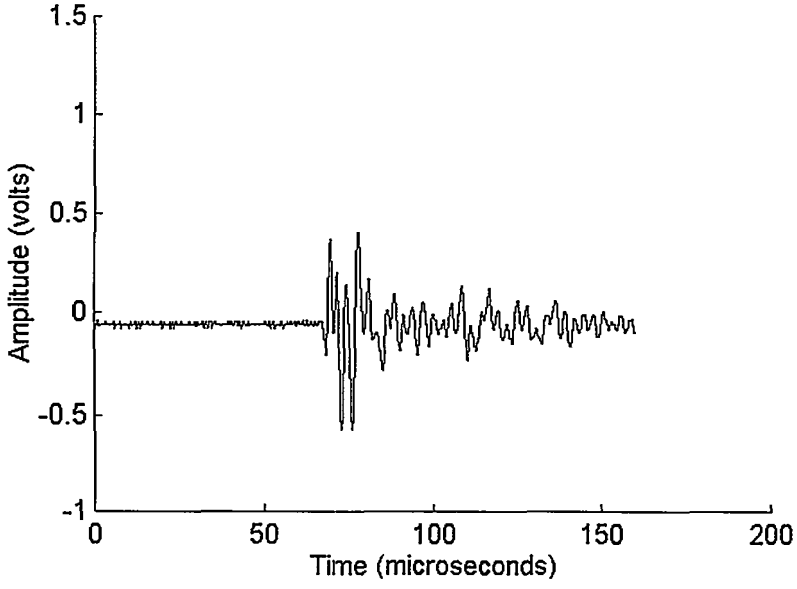
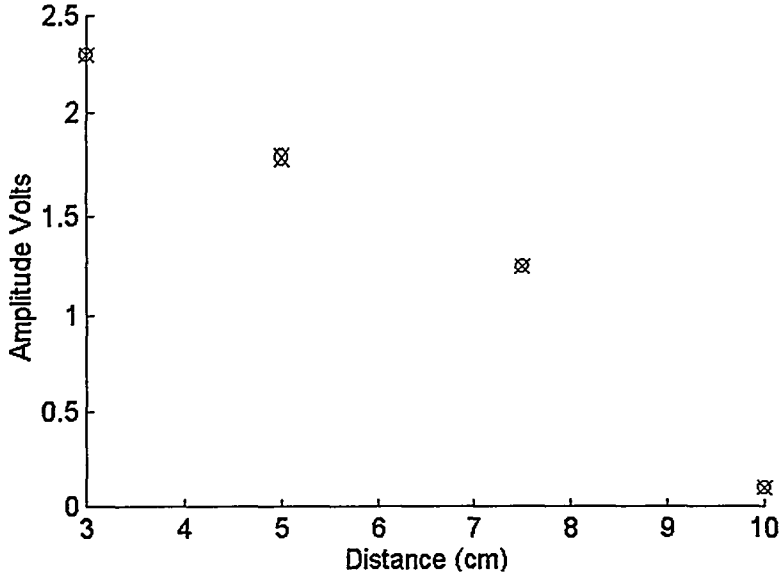
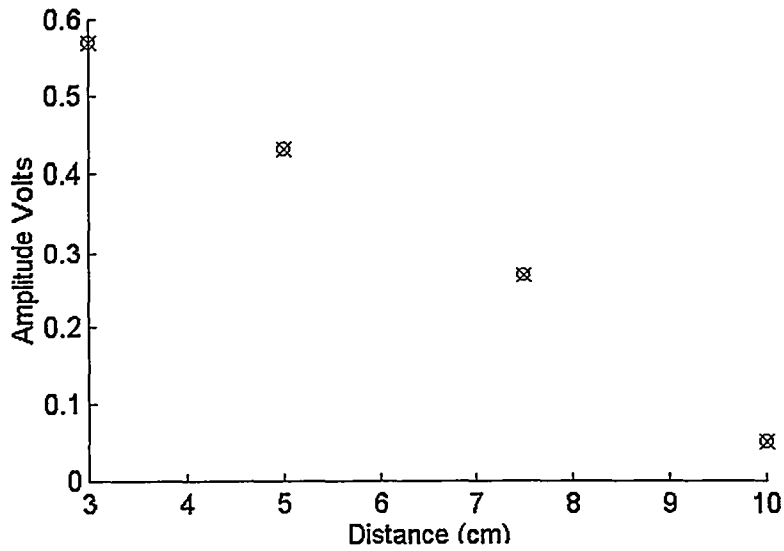


Figure 3 AE waveform generated by pencil lead break on edge of GFRP plate

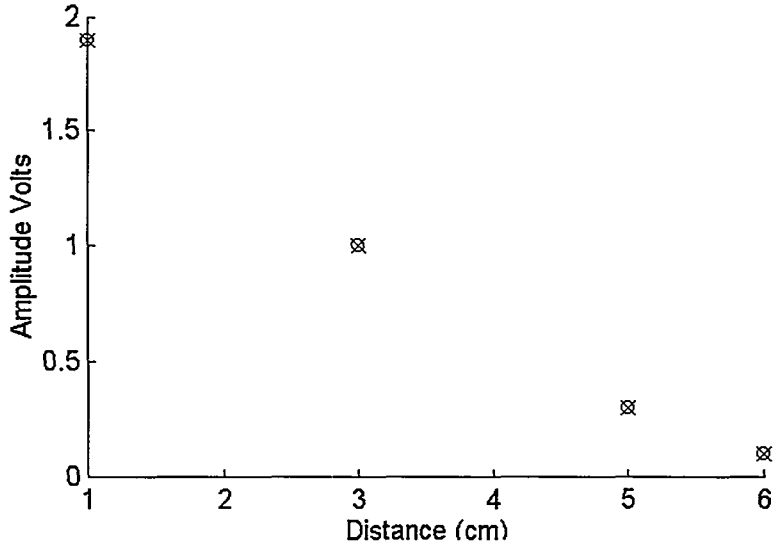


(a)

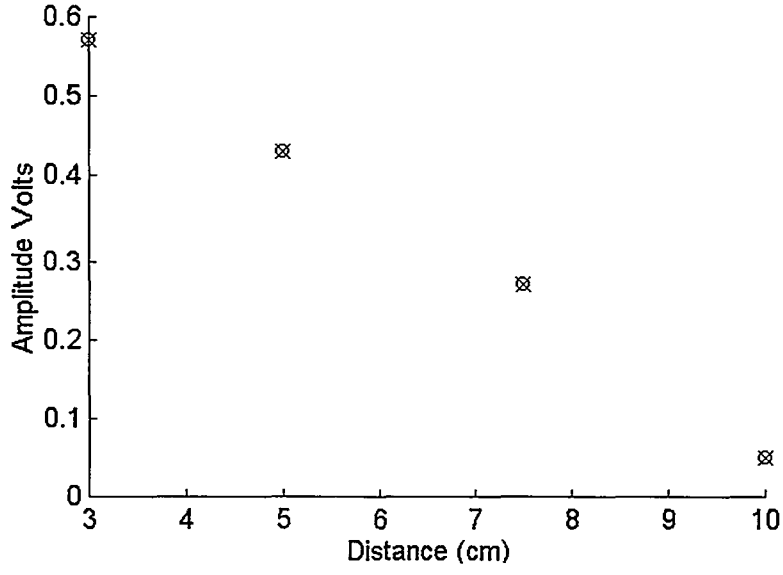


(b)

Figure 4 Flexural wave amplitude attenuation for propagation at (a) 45^0 and (b) 0^0



(a)



(b)

Figure 5 Extensional wave amplitude attenuation for propagation at (a) 45^0 and (b) 0^0

damage related AE events will be more energetic than pencil lead fractures. It does, however, provide the basis for an initial design estimate with the provision that future calibration of a sensor array may be carried out using a using pencil lead source.

5 CONCLUSION

The use of a PVDF film based sensor to detect AE plate waves on GFRP is demonstrated. The broadband characteristic of the sensor allows discrimination between flexural and extensional modes which enable direct identification of different source mechanisms without the need to deconvolute the raw signal. Design trials using simulated AE sources have provided basic information on signal loss through the material and allowed a design configuration with a nominal sensor spacing of 3 cm to be established. The next stage of development will apply this experience in designing integral arrays of PVDF sensors within composite panels along with implementing external hardware support and software algorithms for damage location and identification.

6 REFERENCES

1. W. H. Prosser, K. E. Jackson, S. Kellas, B. T. Smith, J. McKeon an A. Friedman 'Advanced, Waveform Based Acoustic Emission Detection of Matrix Cracking in Composites' Materials Evaluation, **53**(9), Sept. 1995, pp. 1052-1058.

2. Michael R. Gorman 'Plate wave acoustic emission' Journal of the Acoustical Society of America, **90**(1), July 1991, pp. 2551-2556.

3. E. R. Brown, S.C. Kerkyras, G D Neill, RL Reuben and J A Steel 'A Polyvinylidene Fluoride Based Acoustic Emission Sensor For Industrial Applications', Published in Proceedings of Comadem 1996.

4. N. N. Hsu, 'Acoustic Emission Simulator,' US Pat. 4018084 (April 19, 1977)

5. E. J. Cobb 'Acoustic Wave Speed Measurements in GFRP plates' (unpublished).

UNSTABLE SYSTEM CONTROL USING A REAL TIME CENTROID DETECTOR

Greg Horler, Lecturer and Paul Moxon‡.
University of Lincolnshire and Humberside, Cottingham Road Campus, Cottingham Road, Hull. HU6 7RT. United Kingdom.
Email: ghorler@humber.ac.uk

Abstract

The capacity of humans to control unstable systems, such as delivering a toast of champagne upon a flimsy tray, is worthy of reflection. For a machine, such tasks are non trivial. This paper describes how image processing and control techniques may be combined in order to control such a system.

Introduction

The system under consideration is comprised of a plate supported at the centre by a single pivot. The plate may be tilted about the pivot by two servo motors arranged in an orthogonal fashion and suitable linkages. Connected to and in suspension above the plate is a single black and white video camera, Figure 1. The plate, the ball, and the subsequent motion of the ball are the subjects of interest.

The problem is to control the movement. This will include preventing the ball falling off the plate, (avoiding contact with the edges); moving the ball to, then holding it at, a known location; and finally tracking a predetermined path or object.

In order to achieve these aims, a method for determining the displacement, velocity and acceleration of the ball on the plate is required. Detection of the centroid (two dimensional centre of gravity) of an object, via image processing results in a two dimensional co-ordinate for the position of the ball. Calculation of the centroid, for

each video frame enables the movement of the ball to be quantified and used in a suitable controller.

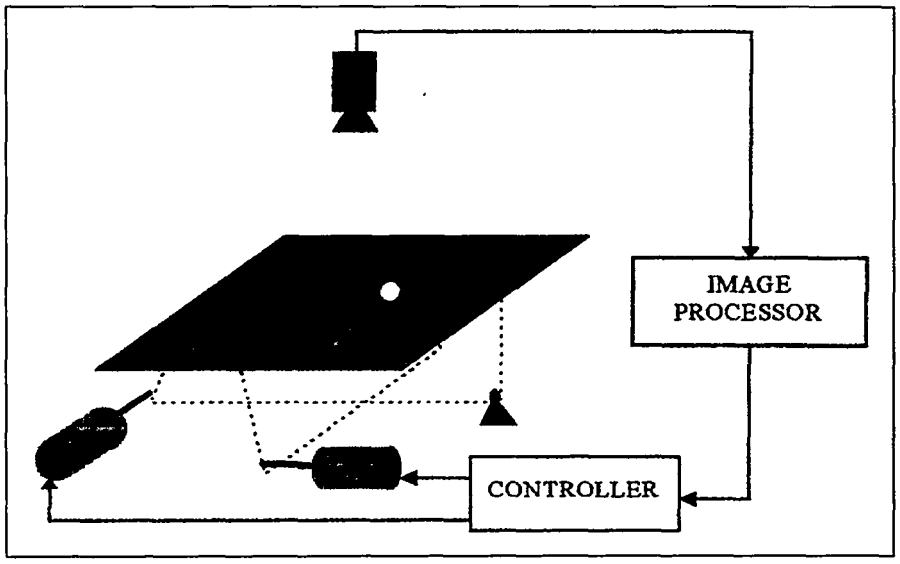


Figure 1: Ball and Plate Apparatus

This paper provides an overview of this project. Section 1 presents a simple controller, the theory and simulation. Section 2 presents the theory and simulation of centroid detection, while section 3 presents an overview of the hardware implementation of the centroid detector.

1 Simple Control Model and Simulation

This section provides an analysis of the ball and plate system which was tested using a computer simulation. The equations of motion can be developed from the acceleration of the ball on the plate.

$$\ddot{x}=K\theta_x \qquad \ddot{y}=K\theta_y$$

Where: $\ddot{x}$ is the acceleration of the ball in the x axis
$\ddot{y}$ is the acceleration of the ball in the y axis
θ_x is the angle of the plate in the x axis
θ_y is the angle of the plate in the y axis
K is adjusted to give the ball a realistic speed of response.

These equations are approximate and are only valid over small plate angles [1] [2].
The velocity and position of the ball can be calculated using the following approximations;

$$\dot{x} = \dot{x} + \ddot{x} \qquad \dot{y} = \dot{y} + \ddot{y}$$

Where: $\dot{x}$ is the velocity of the ball in the x axis
$\dot{y}$ is the velocity of the ball in the y axis

$$x = x + \dot{x} \qquad y = y + \dot{y}$$

Where: x is the position of the ball in the x axis
y is the position of the ball in the y axis

The simulation implemented boundary conditions which limited the calculated x and y positions to the area of the plate. This is to prevent the simulation attempting to control a ball which is not physically on the plate.

The simulated x and y co-ordinates were used to generate incremental rotation angles in the x and y axes respectively. These are used to control the tilt of the plate and hence control the ball. The incremental angles are generated by control equations providing the correct system dynamics. The following incremental angle equations were used in the simulation.

$$\phi_x = \alpha\ddot{x} + \beta\dot{x} + \gamma(x - x_\tau) \qquad \phi_y = \alpha\ddot{y} + \beta\dot{y} + \gamma(y - y_\tau)$$

Where: ϕ_x is the incremental angle of the plate in the x axis
ϕ_y is the incremental angle of the plate in the y axis
x_τ is the target ball position in the x axis
y_τ is the target ball position in the y axis

The constants α, β and γ can be modified to give the controller different characteristics. For example, if γ is set to zero, the effect is to prevent the ball moving; the acceleration and velocity are reduced to zero. A non-zero value of γ forces the equations to drive the ball to the target co-ordinate (x_τ, y_τ).

The magnitude of the angles ϕ_x and ϕ_y are limited to small values in order to ensure that no small scale oscillations occur when the ball approaches the target position. These values are also magnitude limited to simulate the rate limit of the servo motors.

Thus the new plate angles are calculated from the following equations.

$$\theta_x = \theta_x + \phi_x \qquad \theta_y = \theta_y + \phi_y$$

The plate angles would normally be limited in magnitude in order to simulate the physical system. The simulation does not incorporate this facility since the equation dynamics do not cause the plate angles to vary substantially, even at the plate boundary.

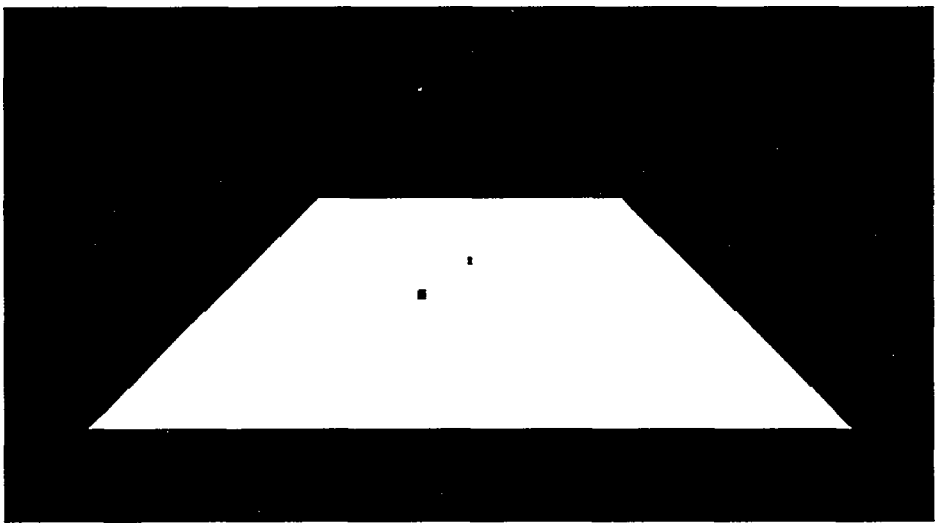


Figure 2: Ball and Plate Simulator

The simulation can be used to investigate the ball on plate system in the following ways.

With γ set to zero the ball should remain stationary. However due to the finite resolution on the incremental angles of the plate, errors are introduced, which prompts the ball to drift on the plate. To remedy this, the full control strategy, using a target marker is used. This provides a referenced focus for the controller and the drift is removed.

In the simulation, the target marker may be placed randomly on the plate and the ball driven to that point, Figure 2. Another function allows the target to follow a predetermined path, the ball then tracks the target.

Other investigations include the removal of the boundary conditions. Here the controller tilts the plate to extremes in order to control a ball which is no longer on the plate. Another example is to reduce the rate limit applied to the incremental angles hence effecting the system damping.

The controller simulation [3] and results confirm theoretically that control of a ball on a plate using only centroid data was possible. The following section examines the theory of centroid generation and investigates various algorithms suitable for hardware implementation.

The Centroid Processor

The centroid may be defined as the centre of gravity of a two dimensional object [4]. In image processing terms it is the pixel co-ordinate which is closest to the objects centre of gravity.

Figure 3 shows an arbitrary shape and the effect of pixelisation, (two dimensional sampling). The co-ordinate system is defined with the origin in the top left hand corner. The x and y co-ordinates defining the centroid are found via the following expressions respectively.

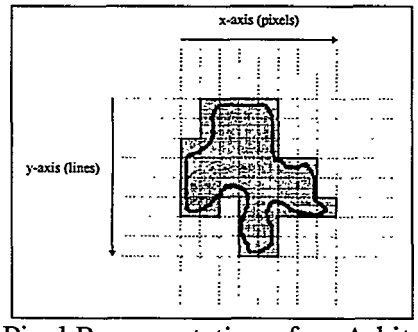


Figure 3: Pixel Representation of an Arbitrary Shape

$$X_{centroid} = \frac{\sum_{n=1}^{mass} x_n}{mass} = \frac{X_{sum}}{mass} \qquad Y_{centroid} = \frac{\sum_{n=1}^{mass} Y_n}{mass} = \frac{Y_{sum}}{mass}$$

where:

X_n is the x co-ordinate of the pixel n within the object
X_{sum} is the sum of the x co-ordinates of the pixel n within the object
Y_n is the y co-ordinate of the pixel n within the object
Y_{sum} is the sum of the y co-ordinates of the pixel n within the object
$mass$ is the total number of pixels comprising the object

Figure 4 shows an object and its centroid.

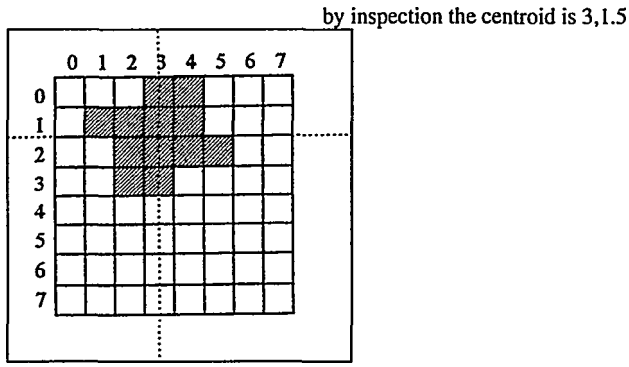


Figure 4: Example Object and Centroid

Direct application of the algorithms yield:

X_{sum} = 3+4+1+2+3+4+2+3+4+5+2+3 = 36
Y_{sum} = 0+0+1+1+1+1+2+2+2+2+3+3 = 18
$mass$ = 18

centroid = (3,1.5)

A fractional pixel is not permitted, hence fractional centroids are rounded down; in this case to (3,1).

In the development of the theory, the effects of noise have been ignored. Noise is commonplace in image processing. In order to reduce noise levels course grain thresholding was used [5]. This was simply achieved by using the most significant bit of each pixel; resulting in a binary (black or white) image. Furthermore the thresholding process also provides segmentation [5] of the image into desirable elements; the ball and the plate.

A simulation of the centroid processor was performed by digitising twenty-four frames of video of the ball on the plate. These may be viewed as an animation, but are presented in stills form in Figure 5. The results of the simulation confirmed the validity of the centroid processor approach. The following section outlines the hardware implementation options.

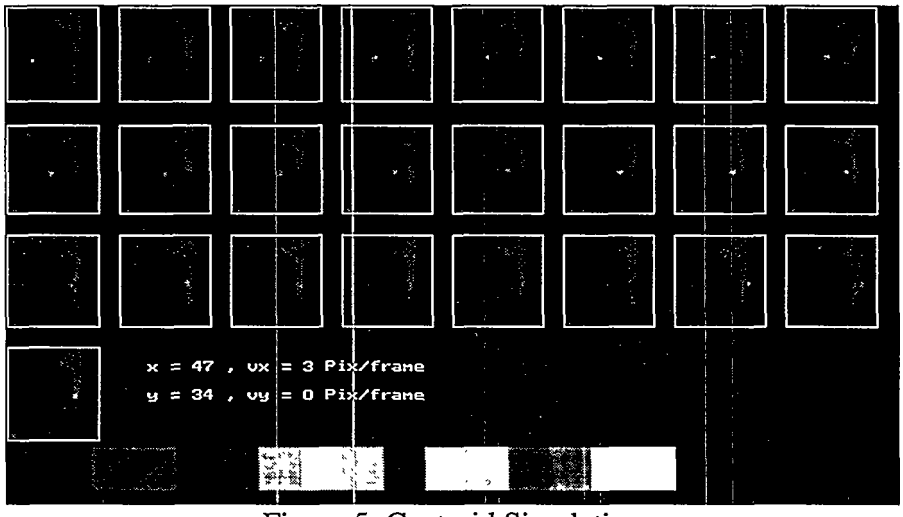


Figure 5: Centroid Simulation

3 Centroid Processor Design

An initial architecture for the centroid processor is shown in Figure 6.

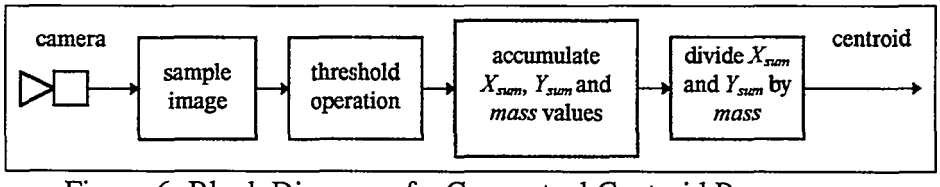


Figure 6: Block Diagram of a Conceptual Centroid Processor

The method of mapping the centroid algorithms to the serial pixel stream determines the processor architecture. The summation operations used for *sum* and *mass* generation can be performed within one pixel period. Division of the *sum* by the *mass* is problematic.

The division period can be reduced by implementing massively parallel dividers. The benefit of this approach is that partial centroids are produced in pixel time. This is necessary if a centroid from more

than one object is required. The drawback with this approach is implementation technology [4]; gate delay and component integration levels.

The tolerances on the divider circuitry may be relaxed if the speed of division is relaxed. Sequential logic based dividers use less circuitry at the cost of speed. Such dividers can be used if the division process is performed 'off line'. This may be achieved by dividing the *sum* by the *mass* of the current video frame during the next frame. This results in a pipelined processor with a centroid latency of one video frame [1].

An input video controller was designed in order to select rectangular portions of the video field of view, Figure 7. This controls the extent of video data inputted into the centroid controller. Thus the plate view could be tailored (shrunk) to the processing capabilities of the centroid hardware.

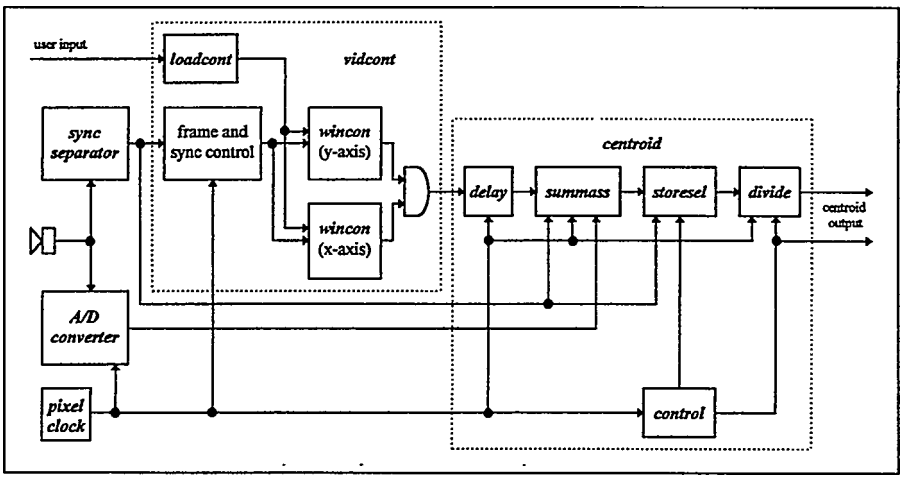


Figure 7: Integrated Video Controller and Centroid Processor

Another method would have been to reduce the sampling rate of the video camera. This method results in large pixel sizes and consequently sacrifices accuracy at the controller.

The precaution of the variable view controller was justified. The division circuitry for the 'off line' implementation was too big for a single Altera device. The controller enabled the design to be tested satisfactorily.

Conclusion

To date centroid control and generation has been simulated. A hardware implementation of a centroid processor has been built and works on a effective screen size; 32 x 32 pixels. The hardware design has been extended to a full screen version; (320 x 256 pixels) and simulated. It is anticipated that this version will be accommodated by the Altera Flex device family. A fully operational centroid processor and controller for the ball and plate problem will be available for COMADEM 98.

References

[1] Ball and Plate Control Using Image Processing Techniques, P. Moxon, 1995. Unpublished BEng. (Hons) Project Thesis. University of Lincolnshire and Humberside.

[2] Ball and Plate Control Analysis, G. Singer, 1995. Unpublished paper. University of Lincolnshire and Humberside.

[3] Ball and Plate Control Using Image Processing Techniques, P. Moxon, 1995. Unpublished BEng. (Hons) Project Thesis. Appendix A. University of Lincolnshire and Humberside.

[4] A Chip Design to Perform Centroid Control, G. Horler, 1988. Unpublished BSc. (Comb.) Hons. Project Thesis. University of Sheffield.

[5] Digital Image Processing and Computer Vision, R. J. Schalkoff, 1989. Chp 6 pp 264. John Wiley & Sons, Inc. ISBN 0-471-50536-6.

‡ British Aerospace, Brough, East Yorkshire, UK

DESIGNING OF MONITORING SETUP FOR VIBRATION SIGNATURE ANALYSIS OF STEAM TURBINE DRIVEN HIGH CAPACITY ROTARY SCREW COMPRESSOR

Pyne T., Dy. Manager (NDT)
Vinod Jose, Engineer (NDT)
Birla VXL Ltd., Porbandar - 360576, Gujarat, India

ABSTRACT

Tracking the behaviour by signature analysis of machines like Screw Compressor having large number of auxiliaries, high power transmissions, variation of process gas properties, changes of load condition, fluctuating revolutions is truly a challenging job. These unavoidable process conditions often disturb the whole setup and there is every possibility to miss important and relevant information. Standards for overall monitoring as well as for peak-amplitudes responsible for root cause identification are not always available because these machines are 'custom designed' and manufacturer's standards are of paramount importance to consider. The health of these machines cannot be assessed by simply comparing with the international standards unlike most common machines such as fans, pumps, motors etc. with minimum number of auxiliaries. There may also be limitations in the features of the instruments used for the purpose. In this paper, an attempt has been made to setup a monitoring approach for screw compressor which will help the industries initially setting base-line data to implement vibration analysis based off-line predictive maintenance programme either with the help of an analyser or with a latest software.

Keywords : Approach, Base-line data, Signature analysis, 'Custom designed' machines, Setup criteria

1. INTRODUCTION

Recent approaches in any predictive maintenance system using vibration analysis technique mostly terminate into predicting the problem of the machine. Regular prediction of a fault is not enough economic for a plant, the root cause is to be identified. Root cause may lie hidden in

design room, in an abnormal process parameter, in a maintenance activity. Vibration signature analysis should aim to relate the symptoms not only to any physical disturbance in the machine but also to its process parameters. Monitoring setup and consequent data collection should be such that the system has got inbuilt sufficient open room flexible enough for storing, analysing and correlating the data. It is to be remembered that on-line monitoring system which just trip every time a machine is under abnormal vibrational force, is not enough for a continuous running process plant where production loss is considerable but an off-line monitoring system is must in order to identify real cause of vibrational forces. Keeping the corrective maintenance management aspect based on signature analysis in mind, it is tried to explore the possible areas and scope of study in monitoring a complex machine like screw compressor.

2. MACHINE DETAILS

2.1 Specifications

Machine	Positive displacement rotary screw compressor
Services required	CO_2 gas compression
BHP	1780 (incl. losses)
Speed	2540 (Manual variation from 1270 to 2667rpm)
Gas handled	26870 cu.m/hr
Inlet pressure	0.91Kg/cm^2
Discharge pressure	4.5 Kg/cm^2
Inlet temp.	50 °C
Discharge temp.	99 °C
Efficiency volumetric	92.1
Casing	Horizontal split, 35 mm thick
Rotor	Helical lobes (male - 4, female - 6), clearance 0.39 mm
Shaft diameter	230 mm
Timing gears	Single helical, Gear teeth (male)-108, Gear teeth (female) - 162
Bearings	Babbit sleeve (Radial bearings), Tilting pad (Thrust bearings)
Vibration instrument	On-line non-contact type proximity probes of Bently Nevada for detecting shaft vibration as well as axial movement.
Coupling	Mounted halves, Spacer required, Keyed
Driver	Back pressure, 4 stages (1 Curtis & 3 Rateau) steam

Turbine speed(rpm)	:9200 rated, Triping - 10626, Critical - 11900 (1st), 29365 (2nd)
Output	:1958 KW
Pressure	:Inlet - 40Kg/cm^2 Exhaust - 2.2Kg/cm^2
Bearings	:Radial - Plain, Thrust - Kingsbury, Forced lubrication.
Mech rating	:2050 KW at 9200/2540
Type	:Helical
# of Teeth	:Pinion - 37, Gear - 134
Oil Qty.	:80 L/min
Critical speed(rpm)	.Pinion - 13842(1st), 30476(2nd) Gear - 15170(1st), 22676(2nd)
Bearings	:Plain journal (radial), Taper land (thrust)

2.2. Sketch : (Refer figure - 1)

3. FEW VITAL INFORMATION

1. Concentrate on acquiring true representative signal from the machine regularly at the same points and by trained engineers so that analysis is fair.

2. Understand vibration analysis as a subject and should not be instrument specific, approach specific. Note the physical phenomenon of each commercial terms of vibration analysis.

3. Start with the same setting. After gaining experience, different setting may be done in all three directions of data collection.

4. Understand the harmonics and their higher than usual peaks.

5. Remember the pattern is more important than the high value of certain defect peaks for root cause analysis.

6. Crosscheck with other symptoms e.g. phase, overall vibration, machine history etc.

7. Specification for spectral analysis is separate for each set or a group of these types of machines. Several sets of data in several sets of measurement points are requied to be acquired.

Sketch with measurement points

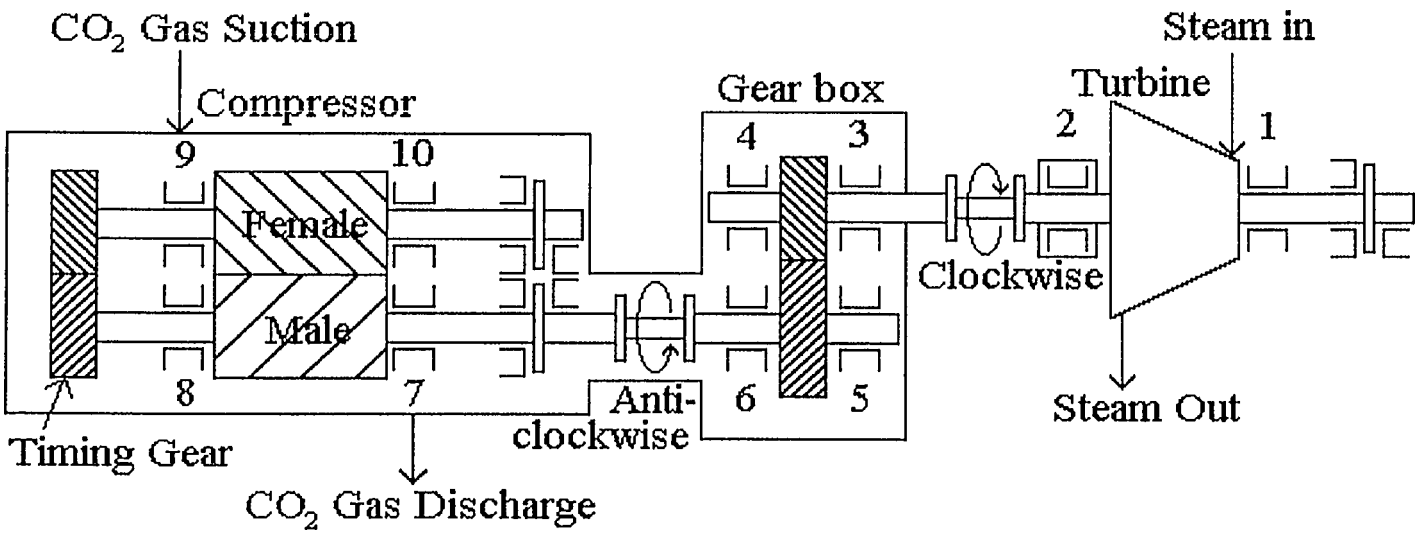


Figure - 1

4. MONITORING SET-UP CRITERIA (Table - 1)

Component / Points	Expected Faults	Spectral Information Required	Tracking ranges to be set	Ultimate setting	Comments
Turbine Points 1&2	Misalignment/ Unbalance / Bent shaft / Resonance	1x > 5x RPM	10x > 12x RPM form 50% RPM	*Overall vel.(rms) Max Freq. 600 > 600K CPM *Zoom spectrum from 1K CPM to 4x compressor	*Rotor RPM related *To search turbine RPM & 3x compressor. (Note gear ratio) (Not required at all points)
	Bearing clearance increase/ wear/ Oil whirl/ Rotor rub	Harmonics/ Subharmonics	Upto 10xRPM from 40% RPM	*Zoom spectrum from 40% RPM to 20xRPM	*Medium frequency, harmonic-pattern is of importance.
	Blade problem of turbine/ resonance with other parts	#blades xRPM (for min. and max. #blades)	Around BPF	*CPB spectra (around 5 > 10%) upto 3x BPF(max) for turbine stage.	*High frequency range depending on #blades (Manufacturer help for min & max #blades)
	Effect of gear on turbine	Gear-Mesh Freq (# teeth xRPM)	Around GMF	*Same CPB upto 3xGMF (Mix setting of BPF and GMF)	*High freq. ranges depending on #teeth. Sidebands at 1x, 2x, 3x should be resoluted well.
	Turbulence and Dynamic forces	Below RPM	Around 1xRPM	*Above zoom spectrum	*Capturing random broad band energy of low frequency zone.

Component / Points	Expected Faults	Spectral Information Required	Tracking ranges to be set	Ultimate setting	Comments
	Shaft movement/ Thrust bearing	# pads xRPM & harmonics of shaft RPM	+/- 12x RPM around Pad Pass Frequency	*Above zoom spectrum	*Depending on number of pads in medium frequency zone
Gear Box Points 3, 4, 5 & 6	Misalignment/ Unbalance / Bent shaft / Resonance	1x > 5x RPM	10x > 12x RPM form 50% RPM	*Overall vel.(rms) *Zoom spectrum from 1K CPM to 4x compressor	*Rotor RPM related *To search turbine RPM & 3x compressor. (Note gear ratio) (Not required at all points)
	Bearing clearance increase/ wear/ Oil whirl	Harmonics/ Subharmonics	Upto 10xRPM from 40% RPM	*Zoom spectrum from 40% RPM to 20xRPM	*Medium frequency, harmonic-pattern is of importance.
	Misaligned/ Mismatch/ tooth wear	GMF and side bands	Around GMF	*CPB spectra (around 5 > 10%) upto 3xGMF.	*High freq ranges. Sidebands at 1x,2x,3x should be resoluted well.
Compressor Points 7, 8, 9 & 10	Misalignment/ Unbalance / Bent shaft / Resonance	1x > 5x RPM	10x > 12x RPM form 50% RPM	*Overall vel.(rms) *Zoom spectrum from 1K CPM to 4x compressor	*Rotor RPM related *To search turbine RPM & 3x compressor. (Note gear ratio) (Not required at all points)

Component / Points	Expected Faults	Spectral Information Required	Tracking ranges to be set	Ultimate setting	Comments
	Bearingclearance increase/ wear/ Oil whirl	Harmonics/ Subharmonics	Upto 10xRPM from 40% RPM	*Zoom spectrum from 40% RPM to 20xRPM	*Medium frequency, harmonic-pattern is of importance.
	Misaligned/tooth wear (timing gear)	GMF	Upto GMF	*CPB spectra (around 5 > 10%) upto 1xGMF	*High frequency ranges. Locating presence of 1x GMF
	Lobe Fault, Lobe wear, rubbing	Harmonics of Lobe Meshing Frequency. (LMF)	Upto 6xLMF	*Above CPB spectra.	*As 6xLMF=24xRPM (Male rotor)
Additional Setting					
	For pinpointing above problems	Order Spectra	15x RPM	*Upto 15xRPM of corresponding Rotor	*Rotor RPM related, to follow speed variation.
	-do-	Magnitude & Phase Angle		*Magnitude and Phase angle (1x, 2x, 3x etc. depending on defects.	Fundamental RPM related (vector and scalar history if tacho facility is available)
	-do-	Crest factor and cepstrum plot (if facility available)		*Upto 60 ms Qufrency (i.e. 800 KCPM Spectrum)	Overall notice of energy level variation

4.1 Alarm Setting (Table - 2)

Defects	Important defect peaks	Alarm Setting
Unbalance, Misalignment, Bent Shaft, Resonance.	Consider atleast upto 5xRPM of corresponding rotor.	Alarm levels are always high similar to common machines. Relative values and pattern are of importance.
Gear Problems	Gear Meshing Frequency.	1x - 70% , 2x - 50%, 3x - 40% of overall value. Sidebands - 25% (Note:- Amplitude will vary depending on ratio, load, teeth)
Journal Bearing Problems	All harmonics upto 10x	2x - 80%, 3x - 40%, rest all - 25% and presence of subharmonics
Compressor lobe	Lobe Meshing Frequency (Pocket Passing Frequency)	Upto 6xLMF (Note:- From few alarm trips, 1x, 2x, 3x are approximated to be 1x - 70%, 2x - 50%, 3x - 20%. Future study on this is required.) Pattern study is similar to common machine. Coincidence of LMF Harmonics with gear / pinion critical speed causing amplification of amplitude is to be noted.

Overall Value in RMS Velocity mm/sec : 7 - normal,
7 > 12* - alert and close monitor,
> 12* danger
(* Old machine - after two major overhauling, value is 10mm/sec)

5. MONITORING AUXILIARIES

Vibration signature analysis of the auxiliary machines e.g. lubrication system, cooling system, steam supply system are equally important like monitoring the screw compressor itself. Separate data acquisition and analysis system on separate data bases depending on these machine's individual specification and service requirement need to be created for the purpose. Other advanced monitoring techniques e.g Wear Debris Analysis, Thermography, Corrosion Monitoring etc. are to be incorporated in order to have an integrated analysis system covering all important parameters. Parameters determining compressor health have got direct bearing on the state of these supporting systems in the following way :-

Lubrication system

1. High temperature of oil damages bearings.
2. Low temperature of oil damages mechanical parts.
3. Low supply pressure damages bearings.
4. Excessive oil flow may be caused by increase in bearing clearances.

Cooling System

1. Low flow rate and high temperature of water increase the temperature of bearings which may damage it.
2. Low flow rate may cause rotor contact due to high temperature-rise of discharge gas.
3. High flow rate may erode casing and rotor.

Steam supply system

1. Low steam inlet and high exhaust pressure cause insufficient power and affect the corresponding vibration spectra in turbine as well as compressor points.
2. High steam inlet pressure has a tendency to increase excessive speed variation.
3. Inlet and exhaust steam temperature determine final hot alignment condition.

5.1 Process gas condition influences

1. High discharge pressure and low suction pressure cause high discharge temperature and rotor contact.
2. High suction temperature may cause high discharge temperature and rotor contact.
3. Ingress of foreign particles / scale deposits.

4. Reverse pressure because of incorrect valve operation may cause rotor contact.

6. SCOPE OF STUDY

1. Co-relation study of shaft radial and axial movement data from on-line Bently Nevada Proximitors with off-line monitor.
2. Radial vibration Vs. transmitted vibration when taken from the casing of the bearings.
3. Bearing temperature Vs. lubrication pressure / temperature
4. Wear debris analysis of lub oil Vs. bearing clearance / wear symptom peaks.
5. Behaviour of process gas condition parameters e.g. temperature, pressure, quantity with Lobe Meshing Frequency / Gear Meshing Frequency.
6. Speed changes Vs. individual defect frequency peak variation.

REFERENCE

1. James E. Berry " Use of Vibration Signature Analysis to Diagnose Machine Problems " Technical Associates of Charlotte Inc., USA. Copyright 1992.
2. Technical Documentation, COMPASS Machine Monitoring System, Bruel & Kjaer A/S, Denmark, Copyright 1992-1994.
3. Instruction Manual for CO2 gas compressor, Kobe Steel Ltd., Tokyo, Japan.

ECONOMIC JUSTIFICATION FOR MAINTENANCE ANALYSIS AND CONDITION MONITORING

Kenneth Culverson
Reliability Engineering Consultant
Alameda, CA, USA

ABSTRACT

With increased global competition and increased emphasis being placed on the "bottom line", it is sometimes difficult to convince the senior management of the necessity of instituting a world class maintenance organization. When maintenance is viewed as a cost center, and not a profit center, gaining the necessary assets to implement an improved maintenance program is extremely difficult. A good deal of the frustration encountered by maintenance and engineering personnel and the outside consultants they hire to assist them, is the difficulty in explaining "simple facts" to those people that control the finances. The lack of a common language, or common background makes the problem just that more difficult. Since the accountants control the money required to do the job properly, it is incumbent on the maintenance and engineering staff to learn and use the language of the accountants to acquire the assets required to accomplish the job. There are other people who do not have the maintenance departments best interest at heart either. To that end, we will present some terms and strategies that have been successfully used to advance the goals of maintenance. It will provide some useful economic equations and methodologies which will put the maintenance organization on the same plane as the money people. The goal is to provide the engineering and maintenance department with the tools required to alleviate the headaches that come at budget time.

1 ECONOMIC TERMS

Knowing the proper terminology is essential to any endeavor. A feel for the terms used by the money people will certainly make life easier in

trying to establish a condition monitoring program. The best way to approach the problem is by using a life cycle cost approach. This approach considers all costs associated with the ownership of a system. So not only is the cost of providing analysis and acquiring the equipment necessary to accomplish condition monitoring looked at, but also the costs of training, spares inventory, personnel, energy usage, lost production, equipment retirement, hazardous waste costs and any other factor that impacts the life of the system. To use a life cycle costing approach requires an estimate of the planned life of the system. To accurately represent the costs and benefits associated with the program some basic accounting methods must be followed.

1.1 Return On Investment

Return On Investment (ROI) gets a lot of use but is usually applied in a simplistic manner of dividing the Investment by the anticipated benefit. While ROI is very easy to use, it has a tendency to overstate the down stream benefits as it does any down stream costs. ROI also neglects to take into effect the time value of money, learning curves, or changes in the technologies used.

1.2 Net Present Value

Net Present Value (NPV) is defined as the present value of all cash inflows and outflows of a project during its life. This form of accounting for costs is generally used as most of the costs tend to occur up front and the benefits are only realized in later years. The equation for NPV is:

$$P = \sum F_n \frac{1}{(1+i)^n}$$

Where
P = Present Value
F = Future value
I = annual interest rate
n = interest period.

This information should be applied to all applicable costs and benefits to determine the payback period.

Table 1 shows a typical spreadsheet set up for a process that has a 6 year life and assumes 10% cost of capital. Figure 1 is a graph of the same spreadsheet showing the payback period of this investment.

	Cash Flow		Present Value	
Year	Benefits	Costs	Benefits	Costs
0		$(15,000)		$(15,000)
1		$(6,000)		$(5,455)
2	$5,000	$(3,000)	$4,132	$(2,479)
3	$12,000		$9,016	
4	$16,500		$11,270	
5	$25,800		$16,020	
6	$23,000		$12,983	
Total	$82,300	$(24,000)	$53,420	$(22,934)

Table 1 Calculation of Net Present Value

1.3 Capitalization

Depending on a firm's accounting standards, and applicable tax laws, the implementation of a condition monitoring program may be capitalized. Capitalization is the recording of expenditures as assets which have future value. Obviously, we can argue that at the engineering costs to set up the condition monitoring program can be prorated into the capital cost of the project. Certainly, any condition monitoring equipment bought for the program can into this category.

1.4 Depreciation

The nice thing about having condition monitoring listed as an asset is that assets can be depreciated over their useful life. Depreciation is the process of allocating the cost of an asset to the periods of benefit. Depreciation has the effect of lowering the life cycle cost of equipment by reducing income by the amount of allowable depreciation. This has the effect of reducing the tax burden on the organization. The formulas for depreciation vary widely depending on the accounting methods used and are well beyond the scope of this paper. In general the salvage value is subtracted from the acquisition cost and the remainder is then apportioned as a deduction from income over a period of years.

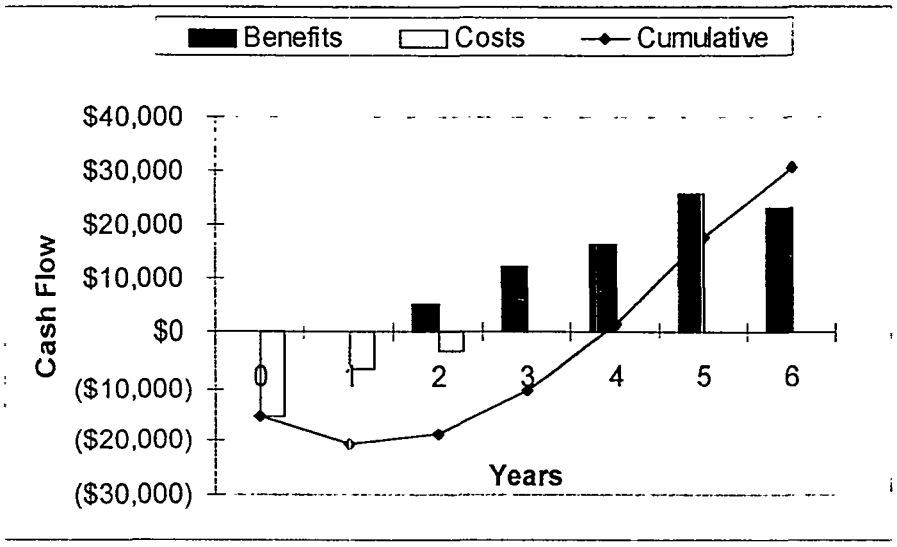


Figure 1. Net Present Value Showing Payback Period

2 OPERATIONS TERMS

In addition to gaining some credibility with the accounting staff, the operations people have to be shown a benefit from the introduction of condition monitoring to the organization.

The two most common complaints that operations has about maintenance are the lack of availability because the machines are not operational, and the poor quality of a product because maintenance did not fix the machine right the first time. While most maintenance professionals would take exception to both statements, this is the perception that must be overcome. We all know the benefits that maintenance gains, but what is in it for operations?

2.1 Increased Availability

As we all know, by monitoring the condition of our machinery on a periodic basis, we can determine the machine's health, and anticipate a potential failure. Just knowing that a machine is going to quit in the near future is usually enough to purchase repair or replacement parts, pre-position them in the proper area, schedule the task and repair the problem

at a convenient time. We will also be able to utilize a greater percentage of the machine's design life in this manner, making it possible for production to have greater availability. In most industries, this translates into the ability to produce more product, making the operations manager look good. The operations manager can tell you with great precision the cost of each hour of lost production, which is a cost avoidance that goes into determining the benefit numbers in the analysis.

2.2 Increased Quality

While the operations manager can tell you the cost of down time, the quality assurance manager can tell you the cost of rework or scrap caused by poor quality in the manufacturing process. Much of the condition monitoring techniques we use everyday can contribute to a lessening of quality defects. Vibration in a drive motor may well translate into an imperfect finish on a product. Oil analysis can detect lubrication problems that cause loss of power in precision drives. Thermographic inspections can detect sources of uneven heating in a product line which cause imperfections in many materials.

3 ESTIMATING COSTS

How do we estimate costs and cost avoidance to implement a world class maintenance program? The costs are relatively straight forward. We will first determine where our major equipment problems are by using a pareto analysis.

3.1 Pareto Analysis

A Pareto analysis is a method used to determine where to apply scarce resources. Essentially, it says that 20% of all problems cause 80% of all headaches. While this is not exactly accurate, it is close enough for our use. If you already have a Computerized Maintenance Management System (CMMS), researching three years of maintenance records will point out those systems which use the most resources to maintain. Three years are recommended to ensure that a particularly bad (or good) year does not skew the results. Although they will probably show up in the list, it is always a good idea to look at those systems that are critical to

operations, but do not require much maintenance, or do not regularly fail. This will provide a short list of systems in the organization which should be analyzed first to determine the appropriate maintenance procedures and steps. Of course, if there is no CMMS then this step can be quite time consuming, but worth the effort. This is the data that can make a program work effectively. This data will also give a consultant a good idea of how big a project might be.

3.2 Use Of Consultants

If there is insufficient knowledge within the organization to complete an analysis and determine appropriate levels of maintenance, you should bring in a consultant to help get the program started. Ideally, this person should be able to present a plan to management and workers, train an internal team in analysis within a short period, and help the team to with the first part of the analysis. The consultant should be available for follow-up visits to assess the progress of the team and provide additional information. The consultant should also be able to bring such information as comparative studies and the cost avoidances they have brought to other firms. Most consultants will be able to provide a reasonable cost and time estimate of the task once they have some knowledge of the size of the problem.

3.3 The Analysis Team

The analysis team will require time to do the studies required, which means time away from maintenance duties. This is a valid cost to add to the job. The makeup of the team will depend in a large part to the systems to be analyzed. It needs to be made up of personnel who are knowledgeable of those systems, and should not be new hires. The team should plan on spending 10-12 hours a week involved in analysis. Whether to use the most senior people on the analysis team is the call of the maintenance manager. If a consultant makes presentations to the workers, many times the right candidates will make themselves known at the presentation. In addition to training on analysis techniques, the team will need some basic training on the technologies available to them to implement the condition monitoring program. Many of these short one and two day seminars are available at no cost from vendors of equipment. The only cost is the time away from the job.

There is a learning curve associated with analysis, as in most endeavors. For example, if the first system analysis requires 4 weeks to complete, the second one will probably require only 3 weeks, the third 2 weeks, the fourth, 1 ½ weeks. This is assuming that the systems are comparable in size. Of course, there is a finite limit to the learning curve and with the proper people the above figures are not unrealistic.

3.4 Technology Survey

Estimating the costs of the equipment required to start the implementation requires determining what technologies are most appropriate to your operation. Up front, you can make some good guesses as to the most appropriate condition monitoring and proactive maintenance techniques will be most appropriate. If the program can be started without acquiring equipment, the analysis team will have a good idea of what is necessary prior to completing the first analysis. With this information, you determine the specifications you require and get price quotes from vendors, including the necessary training. Ensure that training are included as it gives a better overall picture of the true cost of implementation. This step is often overlooked with grave consequences to the program.

3.5 Bench Marking

Bench marking has grown tremendously in recent years as organizations attempt to gain a competitive edge in the marketplace. There is a fair amount of data available in technical journals and from professional maintenance and engineering societies. Bench marking data provides a powerful argument as to where your maintenance program should go, and is extremely useful in presenting a proposal to senior management. It is not a bad idea to contact the author of one of these bench marking studies to participate in a future one. The data is generally shared with all participants prior to publication.

3.6 Cost Avoidance

Cost avoidance is an excellent term to use whenever you want to implement a program that has an up front cost and your department is not

considered a profit center. The current maintenance, operational and quality costs exist. What the maintenance department is attempting to do is to avoid many of those costs by acquiring the proper technology to allow the entire organization to accomplish its goals. Cost saving is a term that should not be used for two reasons. The first is that if you can save costs, then you are not currently doing your job. Secondly, cost savings have a tendency to translate into hard numbers which you have to defend, and will probably be used to cut the budget. Cost avoidance is just nebulous enough to keep either of the above from happening, and help in estimating the effects of condition monitoring.

4 MAKING THE PRESENTATION

Now that you have gathered all this data, how are you going to use the information to get your new program in place? The easy part is now complete, and the real work is just beginning. Within each organization there is a preferred way of doing presentations, such as layout, time allotted, hand-outs, etc. The following are a few general guidelines to go with the internal methods of doing business.

4.1 Using Proper Terminology

It is essential to use proper terminology for the entire presentation. If some of the terms you encounter are a bit foreign to you, check them out to ensure you are using them right, and be prepared to explain them to the audience if necessary.

Never use acronyms or jargon without explaining the terms first. Not only will the presentation be more professional, you won't lose your audience in a sea of terms they do not understand.

4.2 Enlisting Support

Before you put your presentation together for senior management or the budget board, try out parts of it on the operations manager. If you can convince operations that what you are attempting will be good for production, you can probably get that person on your side. Likewise, Quality is another group you would like to get support from before the

presentation. Even the financial people should be given a pitch when you are getting the financial information needed to do the cost analysis. If there are labor concerns, be sure to discuss your ideas with human resources to ensure you do not receive any unpleasant surprises from the workers. And, of course, discuss the ideas with the people who will be most impacted, the maintenance employees.

Anytime there is a change in the way business is conducted within an organization, there is uncertainty and fear. Everyone knows how business is currently done, but what will happen to these employees when this new program comes about. It is absolutely essential that the employees gain knowledge of the scope of the program, the technologies to be used, and the benefits to the firm and the individual. After all, people with the new skills are more marketable.

4.3 Use Graphics

The idea behind making the presentation is to get the funding required to develop a world class maintenance organization. The best way of making your point is to use graphics where ever possible rather than bullet lists. Bullet lists have their place, but clear, concise graphics have greater impact.

Do not overload your presentation with all the detail you have gathered. Have it available as backup, but summarize as much as you can without losing any important information.

4.4 Honesty Works

In attempting to put all this information together to gain funding, it is easy to overlook some item. You will also have to make assumptions about cost avoidance, time to put the program in place. Questions about these things will come up and have to be addressed. It is best to be honest about any gaps in the information. A rational explanation of missing data is better than trying to talk your way out of an embarrassing situation and getting caught on some small detail that ruins any chance of gaining your objective. It is also better to explain the rationale behind estimates rather than trying to defend them as hard numbers. In other words, do not let the "nit pickers" destroy the chance to make a real contribution.

ACKNOWLEDGMENTS

In preparing this paper I drew largely on my own experience in setting up maintenance programs. However, as explained above, I did have to refresh my memory on the precise meaning of some of the terms. To this end I consulted a variety of sources that I keep handy at my desk. The main sources consulted included:

Juran, J. M. & Gryna, F. M. "Quality Planning And Analysis", McGraw Hill Company, 1980

Davidson, S, et. al. "Managerial Accounting", Dryden Press, 1978

Stevenson, W. J., "Production/Operations Management", Richard D, Irwin, Inc., 1986

Blanchard, B. S., "Logistics Engineering And Management", Prentice-Hall, 1986

Maintenance 2000+

Jorma Järviö, Ph.D., Senior Consultant
HM&V Research Oy
Espoo, Finland

ABSTRACT

Industrial maintenance is in turbulence today. By giving up in-house maintenance and starting to buy it, industry has created new business where outside service business professionals are setting new standards for performance. It is not exaggerated to say that the maintenance industry is in transformation.

Two major causes explain the present turbulence. First, maintenance has been traditionally viewed from the engineering perspective (technologies, machines) only. By ignoring or not giving enough weight to the other perspectives (marketing, economical and organizational perspectives) maintenance industry has been more or less defenceless against service providers that sell industrial service.

The second major cause are the rapid changes in environment. Technology assimilation gap is obvious, in-house maintenance crews' capability to maintain competency is difficult and expensive. Industry must either train the people, or buy the needed knowledge (subcontracting). Maintenance organizations need university degree personnel to perform demanding tasks like processor programming and network troubleshooting.

Modern business schools have developed efficient management methods that are widely applied in manufacturing industries, but not in maintenance "industries". By applying these methods, maintenance organizations can increase their efficiencies by 10% - 35%. This development is a must if in-house organizations want to keep their jobs. If this development is not possible or it is too slow, companies very often sell their maintenance department to outside experts, who bring the lacking skills and competence. After this the company starts to buy maintenance from this new company.

1 INTRODUCTION

Three years ago in 1994 there was wide discussions in several issues of *Fortune* magazine about megatrends that affect to employment. The following megatrends were discussed:

1. Average size of enterprises will reduce and they employ fewer people
2. Traditional hierarchical organizations will transform into dynamic, customer serving networks
3. Hierarchy of work classes will change; service oriented professions (like industrial maintenance) will replace manufacturing oriented professions and become *work elite*
4. Conventional top-down organizations will transform into frontline driven ones; customers' role in business will become dominant
5. Services and support role in buyers' decision making becomes more important than those of products
6. Complexity of work will increase and requires continual learning.

When writing this article in early 1997, one can observe that the changes in business environment have followed above trends. There are two major forces behind this development. The first one is the impact of new technologies, and the second one the economic development during recent years. Recession has forced all business towards higher efficiency, and new technologies have provided the means. These developments are especially visible in industrial maintenance. This is the subject matter of this article.

2 MEGATRENDS AFFECTING INDUSTRIAL MAINTENANCE

2.1 First Trend: Average size of enterprises will reduce and they employ fewer people

During the last two decades, the growing pattern of business has changed. Earlier enterprises grew simply by selling more, hiring more people and consequently by selling even more. In other words the growth was based on scale. The result was often fragmentation. Typical example of this kind of development is shown in table 1 (next page):

	lean	*fragmented*
Production (Q'ty/hour)	100	1000
Personnel:		
- production workers	10	100
- supervisors	1	10
- managers	-	1
- assistant managers	-	3
- human resources organization	-	18
- long-range planning	-	19
- audit & control	-	22
- facilitation & expediting	-	23
TOTAL	11	196
Productivity (prod/person)	9.1	5.1

Table 1. Uncontrolled growth of a company [1]

This type of growing pattern has discontinued (exceptions to this development are new businesses. Their growth, however, is result of macrolevel transformation; they replace old technologies or old patterns of work). During 90's the economic growth is more due to increased efficiency. Using above example, it has been now understood that it is far more efficient to operate as ten independent units instead of one large. New information technologies offer equipment, software and connectivity that make traditional headquarter obsolete.

There is a range of managerial tools to use, of which the most common ones are *activity based management* (ABM, also known as activity based costing or value based management), *time based management* (TBM), *supply chain control* (SCC), *business process reengineering* (BPR), and *lean management*. Common denominator to all these approaches is *process management*. Principal idea is to concentrate on those activities that create added value to the end user or customer; all other activities are unimportant or even nonproductive (from end users' viewpoint) and should be given up.

The author has observed several cases in Finland and in the U.S., where the elimination of rules and regulations of the central administration improves efficiency significantly (10% – 30%). These improvements and subsequent cost savings are often anticipated to be result of lost jobs and working outside of own premises (subcontracted by other users or customers). Both of these assumptions are wrong; repair work will reduce, but preventive work will compensate this re-

duction. Contracted work will be performed in those cases where the cost of needed expertise is too high for one party only; by dividing the costs the possession of the expertise becomes feasible. Improved financial result is not due to cost cuttings, but the improved utilization of the production machinery due to smaller (and controlled) downtime.

Another negative attitude deals with demergings, where a company gives up own maintenance and starts to subcontract it. There are two ways to do this. Friendly demerging takes place when the maintenance organization is competent and capable to absorb business expertise. In this case the maintenance organization becomes an independent enterprise but it remains under the "parent company umbrella". In the second demerging case the company sells its maintenance department to an outside service provider and starts to buy it back. Service provider is a company that is specialized in industrial maintenance. In this case, the lacking expertise is acquired by alliance with the maintenance specialist. Old parent may still control the new arrangement through partial ownership. The third possibility, if the two friendly approaches are not possible, is to give up own in-house maintenance organization and outsource instantly. In all three cases, maintenance organization gets the operational independence to develop itself without limitations from other departments.

2.2 The second trend: Traditional hierarchical organizations will transform into dynamic, customer serving networks.

Earlier, technician was waiting for a repair call, which often was a written work order. Tomorrow (even today) the technician must foresee the work. The passive, reactive approach will not work. In order to be competent and competitive, one must be proactive. Further, all actions must be justified and feasible. This requires common sense and businesslike thinking. In-house maintenance is no longer part of some bigger business, it is business itself and must be performed accordingly. Control tools like profitability and customer satisfaction are used daily to guide operations.

Work in manufacturing plants is increasingly performed in independent selfdirecting teams without supervisors. In maintenance this means that there are no longer organizations or supervisors that convoy the maintenance needs from the equipment operator to the maintenance

technician. Instead, the technician must be in direct contact with these teams and put his/her expertise to serve common goals.

Service technician in 2000+ is not a "jack of all trades". Neither is he or she a highly trained specialist (although these kind of specialists are needed, too). Technicians are very skilled to manage information and operate processes, a role that new information technology will introduce. Possibilities of this new technology are introduced in Figure 1:

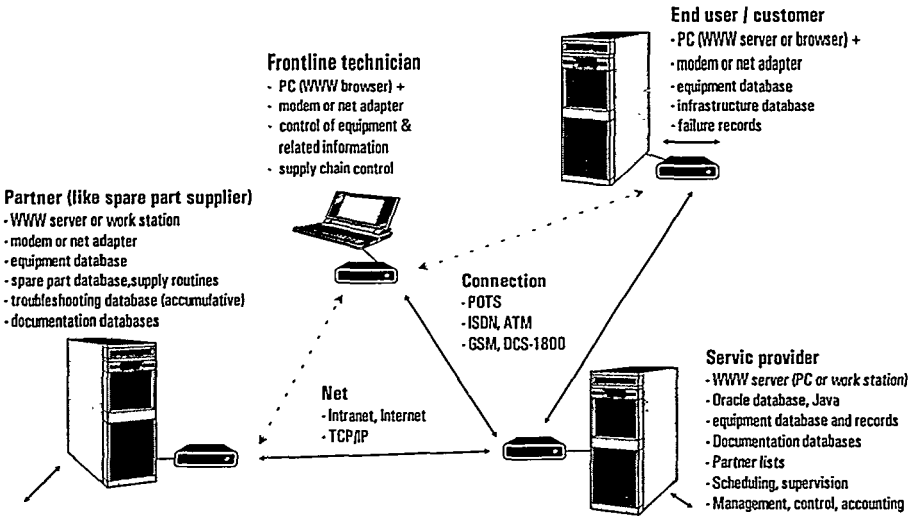


Figure 1 Client/server system in maintenance organization

Documentation in 2000+ will be in electronic form. In this respect the development is rapid; virtually all new design work is performed with CAD systems. High holding costs and laborious management of drawing archives will escalate the conversation of this material into electronic form.

Service technician's tools include portable PC (or any www-browser). With this the technician connects into home office (service provider) to obtain documentation (drawings, service records (history), troubleshooting instructions, spare part catalogues, etc.) that are needed. Technician also can plug the PC into the equipment itself to analyze the build-in sensors and detectors. The media can be POTS (Plain Old Telephony System), wireless GSM or DCS-1800. The net can be Internet or intranet. Service technician does not necessarily know, from which database (or even from which geographical location) the requested information is coming. Key thing is the authorization; who

is allowed to access and on what terms. The question (and the solution) is commercial.

Most communication takes place between the technician and his/her back office. It is technically possible to have direct communication (like access to databases) for instance between the end user and technician, but commercial terms may set restrictions.

The above visions contain three interesting aspects. First, operations are front line driven; the technicians control system, use data and keep it updated. Secondly, the information is stored in servers; the technician obtains only relevant pieces. After using, he or she updates records with information that is considered appropriate, transmits it back to the server and then clears the PC. Thirdly, this type of handling of data affects to the qualifications of the technician. PCs bring the expertise to the frontline; the technician needs not to be expert in troubleshooting, but in troubleshooting information handling (utilization).

2.3 The third trend: Hierarchy of work classes will change; services oriented professions (like industrial maintenance) will replace manufacturing oriented professions and become *work elite*

New technologies are introduced at such pace that the traditional in-house maintenance organizations have problems in keeping at par with this pace. New technologies include electronics, data processing, networking, mekatronics, pneumatics, hydraulics, and so on. It is not oversimplified to state that equipment itself is becoming more and more simple, but the control systems are becoming more and more complex. This development is shown in Figure 2.

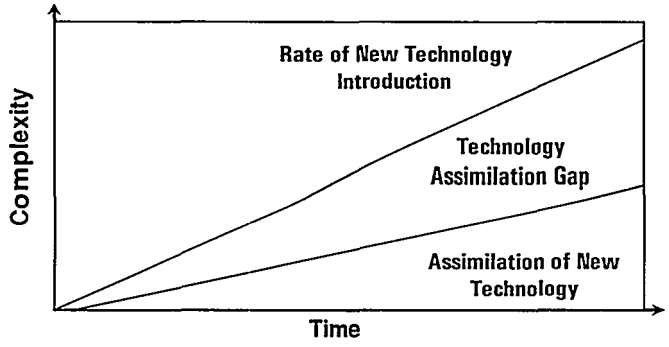


Figure 2 Technology assimilation gap [2]

Technology assimilation gap can be filled with training or by subcontracting. Subcontracting has two advantages. If the quantity of work is small or the life cycle of new technology is short, possession of expertise may not be feasible for one company alone. By sharing the high costs can be reduced to competitive level.

Training needs are not necessarily equipment or even technology related. This problem in shown in Figure 3:

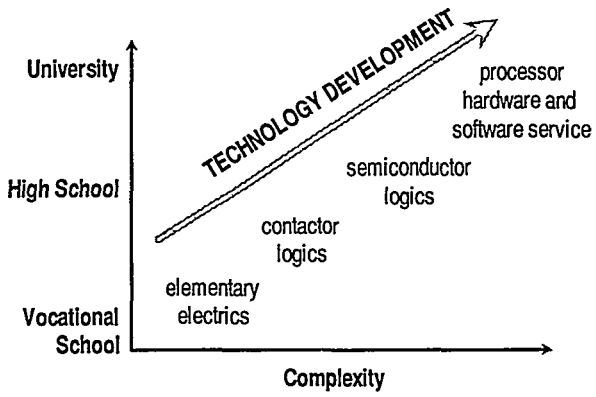


Figure 3 Increased expertise vs. school education

The more complex the task, the more demanding training is needed. Problems appear with the so-called career path; if training does not take place in both dimensions (Fig. 3), the limits are reached very soon resulting frustration.

When a company operates on worldclass level, schools have very little to give in terms of ready programs. In these companies all training must be customized; schools bring the sciences and companies the applications.

More demanding work profile brings new, highly trained people to maintenance organizations. This and the change of work profile as well as realization of the high economic impact of good maintenance will lift the esteem in way that it is not exaggerated to say that services oriented professions (like industrial maintenance) will replace manufacturing oriented professions and become *work elite.*

2.4 The fourth trend: Conventional top-down organizations will transform into frontline driven ones; customers' role in business will become dominant

Today's successful service providers focus their activities on the following:

- *relationship marketing.* Important is to realize that the actual customer is never the machine but the user, the human who selects the service provider. The second important thing is to focus operations on existing customers instead new ones [3]
- own personnel must be motivated to service their customers. Customer satisfaction begins from the satisfaction of own personnel [4]
- local exposure. Global exposure is negative; local customers want local services that are available in a near-by location. None of the worldwide operating service providers advertises size nor worldwide exposure; all are small, local friendly partners and best in town.

Some 10-15 years ago *Turning the Pyramid of Authority Upside Down* (Fig. 5) was a hot management fad.

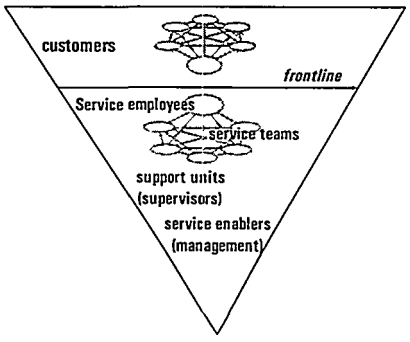


Figure 5 Turning the Pyramid of Authority Upside Down [5]

The driving force of a service company are the customers. Customers are served by *service employees* over the *frontline*. In customers eyes, the service employees shape the image of the service company. The service employees influence directly or indirectly on virtually all business actions that the service company has. From operations viewpoint, service employees are the most important layer of people in service company in all respects.

In real life, three major problem areas exist. The first problem is the lack of empowering in frontline. Managers do not trust service employees. The second problem is opposite, service employees do not want to assume responsibility; commercial matters are simply not their job. The third problem is that the service employees are not equipped with sufficient training or means to carry out their new tasks.

A group of managers of in-house maintenance organizations envisioned the tasks of a service organization 2000+ to be:

- labor (as today)
- information
  - reporting performed work
  - reporting predicted work
  - consulting
    - failure analysis
    - risk analysis
    - recommended actions (best practices)
    - feasibility studies
  - analysis of operations (customer's best interest)
  - maintenance related follow-up
    - environmental control
    - legislation
  - development of cooperation (avoiding overlapping work)
  - joint development of products and processes

2.5 The fifth trend: Services and support role in buyers' decision making becomes more important that those of products.

From equipment users' viewpoint supporting services are more important than the equipment itself. This outcome comes from the survey [2] that Coopers & Lybrand and AFSM International conducted in 1995. Results are shown in Table 2 and 3.

Quality and customer satisfaction related matters are often inadequately perceived. When users consider equipment good and of high quality, this assessment is based on performance rather than the equipment. Gale has studied [6] that the equipment quality or hard quality yields only 30% of total. The rest comes from usability (30%), supporting service (30%) and faultfree paperwork like invoicing (10%).

Rank	Customer view	Senior services executive view
1	Product performance **Services & support** (tied)	Product Performance
2	Product features	Product features
3	Vendor reputation	Vendor reputation
4		**Services & support**

Table 2 Different Views of the Importance of Service Issues in the Product Selection Process (Year 1995)

Rank	Customer view	Senior services executive view
1	**Services & support**	Product performance
2	Product performance	**Services & support**
3	Product features	Product features Vendor Reputation (tied)
4	Vendor reputation	

Table 3 Different Views of the Importance of Service Issues in the Product Selection Process (Year 1997)

2.6 The sixth trend: Complexity of work will increase and requires continual learning

The impact of the sixth megatrend overlaps with the third one continual development of technologies puts pressure on product training as well as on training of people. Continual training is a must, what one has learned in university, will be obsolete in few years. Knowing this, aggressive, successful companies continually train their personnel.

When analyzing successful service providers, clear conclusion is that the most valuable asset they have, are the people, who are motivated and committed to success. This is the only way to stay in business. There are several cases showing that the productivity of a person can be tenfold, when the person wants to train him/herself, wants to make success and is committed to common goals. The training should not only cover maintenance related topics, but also data processing, economics & business, accounting and languages.

3 REFERENCES

1 Michael Hammer & James Champy, *Reengineering the Corporation, A Manifesto for Business Revolution*, HarperCollins Publishers, New York, NY, 1993

2 Joe Trpik, *Future Trends in Service*, Association For Service Management International, Fort Myers, FL, 1995

3 Heskett, Sasser, Hart, *Service Breakthroughs*, The Free Press, New York, NY, 1990

4 Lars Odlén, *Customer Loyalty Management – Making the Difference in Service Business*, IIR Finland (After Sales – Tekninen Palvelu seminaari 29-30.11.1995), 1995

5 modified from Karl Albrecht, *At America's Service*, Dow Jones-Irwin, USA, 1988

6 Gale B T, *Managing Customer Value. Creating Quality & Service That Customers Can See*, The Free Press, New York, NY, 1994

DETERMINATION OF INSPECTION INTERVALS FOR A CONDITION-BASED MAINTAINED SYSTEM

Susumu Okumura, Associate Professor
Department of Mechanical Systems Engineering, University of Shiga Prefecture, Shiga, Japan

ABSTRACT

A method for determining the optimum discrete time points of inspection monitoring for a deteriorating single-unit system under condition-based maintenance (CBM) is developed. The conditions of a system in a normal state, a symptom state or a failed state are considered. The transition between the states is described using a symptom-delay-time (SDT) model. The transition time from a normal state to a symptom state and from a symptom state to a failed state are derived from continuous deterioration processes. Failed-dangerous (FD) and failed-safe (FS) probabilities of imperfect inspection are taken into account. A minimization problem of the long-run average cost per unit time is formulated, an algorithm for obtaining the optimum solution is shown, and the nature and sensitivity of the proposed method are investigated.

1 INTRODUCTION

Failure and deterioration of a maintained system result in increased economic losses as they become more complex. Much importance has therefore been attached to maintenance activities aimed at reducing such losses, and CBM has been widely introduced. When CBM is put into practice, the condition of the maintained system concerned must be identified at discrete time intervals or continuously using condition diagnosis techniques. In this study, we consider CBM involving inspections conducted at discrete time intervals.

When the condition of a system starts to deteriorate rapidly, inspection should be carried out more frequently in order to maximize the probability of detecting a symptom of failure. However, frequent inspections are not always cost effective. Although long intervals be-

tween inspections reduce inspection costs, there is an increased risk of economic losses due to system failure. Therefore, the question of how to determine the optimum discrete time intervals for inspection for cost-effective CBM arises. Inspection intervals are generally determined empirically.

When an inspection policy for CBM is developed, it is not sufficient to consider only the normal state and the failed state of the system, since identifying whether the system is in a failed state or not is not the purpose of an inspection. An intermediate state between the normal and failed states, i.e., a state in which a symptom of impending failure due to deterioration can be observed by inspection (a symptom state), is therefore considered. The transition time from a symptom state to a failed state is called the delay time. Industrial plant maintenance problems were analyzed by Christer [1] and Christer and Wang [2]. They assumed that delay time has an arbitrary density function; however, the transition time from a normal state to a symptom state, which is called the symptom time here, was restricted to have a uniform distribution. Analysis of an SDT model where both symptom and delay times have arbitrary distribution functions was conducted by Okumura *et al.* [4]. However, the developed model was for time-independent probability of an inspection, and the optimum solution was obtained with a decision variable under a constraint on inspection intervals. An SDT model is utilized to describe the transition between the states in this research.

To obtain the distribution function for the time to failure, statistical characteristic values such as failure rate are conventionally used. Lu and Meeker [3] developed a general path model and data analysis methods using degradation measures to estimate a time-to-failure distribution, and reviewed degradation models. However, if the relationship between time and degradation level, which is often shown in a deterioration trend curve, is included in the distribution function we can combine the function with deterioration processes.

The true condition of a maintained system is not always correctly identified by inspection. A system may be identified as being in a normal state by inspection when it is actually in a symptom state (type I error), or a system may be identified as being in a symptom state when it is actually in a normal state (type II error). Therefore, the subsequent time-dependent probabilities are considered for imperfect inspection: FD probability (the probability of a type I error occurring) and FS probability (the probability of a type II error occurring). In the model of Okumura *et al.* [4] the probabilities are time-independent.

A new method for determining the optimum discrete time intervals

for inspection of an SDT modelled system is developed. An inspection policy is formulated as a minimizing problem of the long-run average cost per unit time, and an algorithm for obtaining the optimum solution is given. The behavior of the model is investigated by numerical analysis.

2 FORMULATION OF THE PROBLEM

2.1 Assumptions

(1) A system under condition-based maintenance deteriorates and undergoes states s_0, s_1, ..., s_N over time, where the system in a normal state s_0 is regarded as having no deterioration or being in a situation of neglected deterioration, and the system in a failed state s_N is considered to be in a condition that does not achieve the desired result.

(2) The times when a system transforms from s_{i-1} to s_i ($i = 1, 2, \ldots, N$) are denoted by random variables X_i, each of which is independent, and has a probability density function $f_i(t)$ and a cumulative distribution function $F_i(t)$. As a consequence, the time from s_0 to s_N follows a random variable $X = \sum_{i=1}^{N} X_i$, and has a probability function $g(t)$, a cumulative distribution function $G(t)$ and a failure rate function $\lambda(t)$.

(3) Since CBM is applied to consequential systems, the failure of which may result in huge economic losses, the failure of a system is identified without delay.

(4) The probability that the state of a system in s_i is identified to be in s_j is given by $p_{ij}(t)$ with the relation $p_{ij}(t) = 1 - p_{ij}(t)$. When $i = j$, $i > j$ or $i > j$ the state of a system is diagnosed correctly, failed-safely, and failed-dangerously, respectively.

(5) A system is restored to an as-new condition when it is identified to be in s_θ ($0 < \theta < n$) by inspection or when it fails.

(6) Inspections are executed at discrete points in time t_k ($k = 1, 2, \ldots$) with negligible time required for an inspection. Discrete points t_1, $t_2, \ldots$ are denoted by t, which is called an inspection time vector. New inspection time begins at t_1 after a system is replaced.

2.2 Description of deterioration processes

We call a plot of points $(t, D(t; \beta))$, as shown in Fig. 1, a deterioration trend curve, where $D(t; \beta)$ is the value of the quantitatively measured deterioration of a maintained system and β is a parameter vector. If

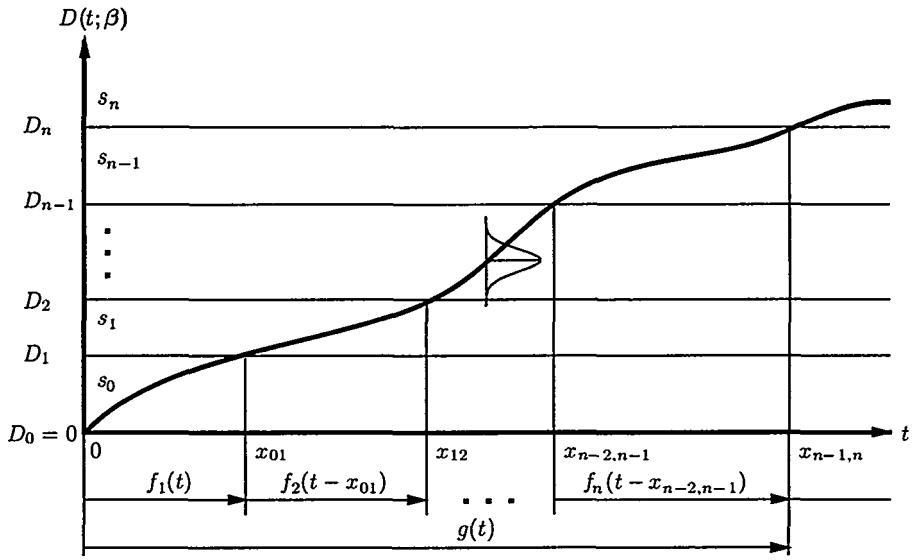


Figure 1. Relationship between deterioration trend curve and probability density functions.

D_i $(i = 1, 2, \ldots, N)$ satisfies $D_0(= 0) < D_1 < \cdots < D_N$, a system is in s_{i-1} when $D_{i-1} \leq D(t; \beta) < D_i$.

Even if $D(t; \beta)$ is known, variability of the measured degradation values and measurement errors can occur. In this research, the degradation quantitatively measured by inspection at time t is assumed to have a normal distribution $N(D(t; \beta), \{\alpha D(t; \beta)\}^2)$, where α is constant over time.

Letting $Y_i = \sum_{j=1}^{i} X_j$ $(i = 1, 2, \ldots, N)$, the cumulative distribution function of Y_i is given by

$$H_i(t) = \begin{cases} F_1(t) & (i = 1) \\ H_{i-1} * F_i(t) & (i = 2, 3, \ldots, N), \end{cases} \tag{1}$$

where $*$ is a symbol of convolution.

Assuming the existence of $D^{-1}(t; \beta)$, we obtain

$$H_i(t) = P\{Y_i \geq D^{-1}(D_i; \beta)\} \tag{2}$$

$$= \Phi(q_i(t; \beta_i))/\Phi(1/\alpha) \quad (i = 2, 3, \ldots, N), \tag{3}$$

where $q_i(t; \beta_i) = (D(t; \beta) - D_i)/\alpha D(t; \beta)$, $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution $N(0, 1)$, $\Phi(1/\alpha)$ is a normalizing constant, and $\beta_i = (\beta, D_i, \alpha)$.

Therefore, the following equations are obtained:

$$F_1(t;\beta_1) = \Phi(q_1(t;\beta_1))/\Phi(1/\alpha) \tag{4}$$

$$F_i(t;\beta_{i-1},\beta_i) = \mathcal{L}^{-1}\left[\frac{\mathcal{L}[\Phi(q_i(t;\beta_i))]}{\mathcal{L}[\Phi(q_{i-1}(t;\beta_{i-1}))]}\right] \quad (i = 2, 3, \ldots, N) \tag{5}$$

$$G(t;\beta_N) = \Phi(q_N(t;\beta_N))/\Phi(1/\alpha) \tag{6}$$

$$\lambda(t;\beta_N) = \frac{q_N'(t;\beta_N)\Phi'(q_N(t;\beta_N))}{\Phi(1/\alpha) - \Phi(q_N(t;\beta_N))}, \tag{7}$$

where $\mathcal{L}[\cdot]$ and $\mathcal{L}^{-1}[\cdot]$ denote the Laplace transform and the inverse Laplace transform, respectively.

2.3 Long-run average cost per unit time

The cases when states are s_0, s_1 and s_2 are investigated as typical examples in this study. We call $s_1 = s_\theta$ a symptom state, and the time between two successive replacements one cycle. Some distinctive operational cases when inspection is carried out k times in a cycle are shown in Fig. 2.

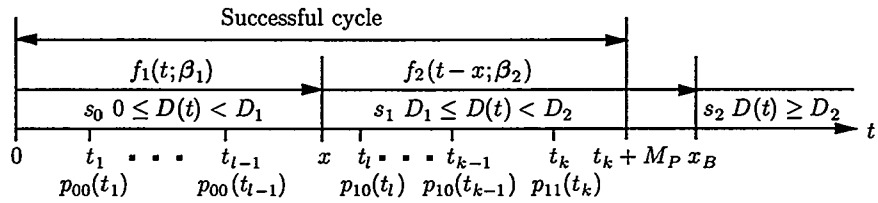
Let random variables C, T and I be the cost incurred in one cycle, the length of one cycle and the number of inspections per cycle, and E_C, E_T and E_I their expected values, respectively. The long-run average cost per unit time is given by E_C/E_T. If P_P is the probability that a cycle ends in preventive replacement, and P_F the probability that a cycle ends in failure replacement, we can write

$$E_C = C_I E_I + P_P C_P + P_F C_F, \tag{8}$$

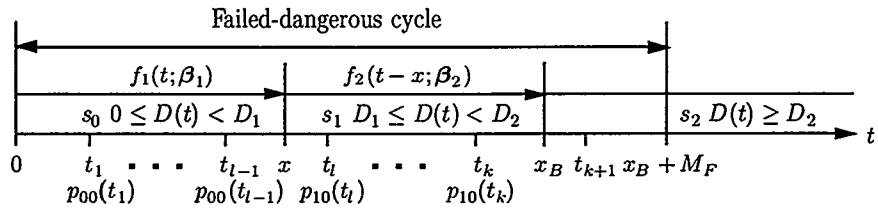
where C_P is the cost of a preventive replacement, C_F is the cost of a failure replacement and C_I is the cost of an inspection.

Since $E_I = \sum_{k=0}^{\infty} k\, p_k$, where p_k is the probability when inspection is conducted k times in one cycle, the following equations are obtained, considering the three types of cycles, (a), (b) and (c) shown in Fig. 2:

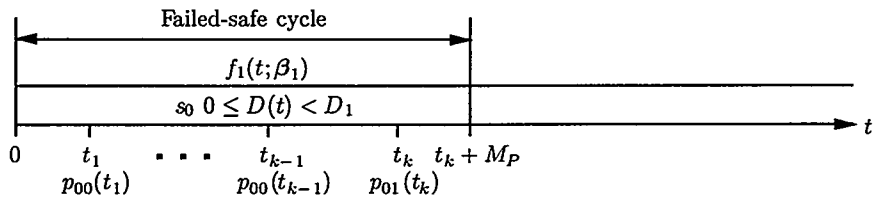
$$\begin{aligned} E_I = &\sum_{k=0}^{\infty} k \sum_{l=1}^{k+1} \prod_{m=1}^{l-1} p_{00}(t_m) \prod_{m=l}^{k-1} p_{10}(t_m) \\ &\cdot \int_{t_{l-1}}^{t_l} f_1(x) \left\{\delta(k-l+1)\overline{F}_2(t_k - x) - p_{10}(t_k)\overline{F}_2(t_{k+1} - x)\right\} dx \\ &+ \sum_{k=0}^{\infty} k \prod_{m=1}^{k-1} p_{00}(t_m) \left\{\overline{F}_1(t_k) - p_{00}(t_k)\overline{F}_1(t_{k+1})\right\}, \end{aligned} \tag{9}$$



(a) An example of successful cycle when a preventive replacement is conducted at t_k



(b) An example of failed-dangerous cycle when a failure replacement is conducted at x_B



(c) Failed-safe cycle when a preventive replacement is conducted at t_k

Figure 2. Some typical examples of operations when states of a system are s_0, s_1 and s_2.

where $t_0 = 0$, $\delta(x) = \{0\ (x = 0), 1\ (x \neq 0)\}$, $\overline{F}_i(\cdot) = 1 - F_i(\cdot)$ and $\prod_{\cdot=x}^{y} = 1\ (y < x)$.

Preventive cycle probability is given by

$$P_P = \sum_{k=1}^{\infty} \sum_{l=1}^{k} p_{11}(t_k) \prod_{m=1}^{l-1} p_{00}(t_m) \prod_{m=l}^{k-1} p_{10}(t_m) \int_{t_{l-1}}^{t_l} f_1(x) \overline{F}_2(t_k - x)\, dx + \sum_{k=1}^{\infty} p_{01}(t_k) \prod_{m=1}^{k-1} p_{00}(t_m)\, \overline{F}_1(t_k), \tag{10}$$

which takes into account both a successful cycle (a) and a failed-safe cycle (c), and $P_F = 1 - P_P$.

The expected cycle length can be written as the probabilistically weighted sum of each type of cycle:

$$E_T = \sum_{i=1}^{3} (\text{Length of type } i \text{ cycle}) \cdot (\text{Probability of type } i \text{ cycle}), \tag{11}$$

where $i \in \{(a), (b), (c)\}$.

We have

$$
\begin{aligned}
E_T = & \sum_{k=0}^{\infty} \sum_{l=1}^{k+1} \prod_{m=1}^{l-1} p_{00}(t_m) \prod_{m=l}^{k-1} p_{10}(t_m) \\
& \cdot \int_{t_{l-1}}^{t_l} f_1(x) \, [\delta(k-l+1) \\
& \qquad \cdot \{t_k + p_{11}(t_k) M_P + p_{10}(t_k) M_F\} \, \overline{F}_2(t_k - x) \\
& \qquad - p_{10}(t_k)(t_{k+1} + M_F) \overline{F}_2(t_{k+1} - x)] \, dx \\
& + \sum_{k=0}^{\infty} \sum_{l=1}^{k+1} \prod_{m=1}^{l-1} p_{00}(t_m) \prod_{m=l}^{k} p_{10}(t_m) \int_{t_{l-1}}^{t_l} f_1(x) \int_{(t_k - x)\delta(k-l+1)}^{t_{k+1}-x} \overline{F}_2(t) \, dt \, dx \\
& + \sum_{k=0}^{\infty} \prod_{m=1}^{k-1} p_{00}(t_m) \, [\{p_{00}(t_k)(t_k + M_F) + p_{01}(t_k)(t_k + M_P)\delta(k)\} \\
& \qquad \cdot \overline{F}_1(t_k) - p_{00}(t_k)(t_{k+1} + M_F) \overline{F}_1(t_{k+1})] \\
& + \sum_{k=0}^{\infty} \prod_{m=1}^{k} p_{00}(t_m) \int_{t_k}^{t_{k+1}} \overline{F}_1(x) \, dx, \qquad (12)
\end{aligned}
$$

where M_P is the expected time required for a preventive replacement, and M_F is the expected time required for a failure replacement. Letting $C_F/C_P \equiv r_{FP}$, $C_I/C_P \equiv r_{IP}$ and $C_P \equiv 1$, the objective is to determine the inspection time vector t which minimizes

$$
\frac{r_{FP} E_I(t;\, \gamma) + (1 - r_{FP}) P_P(t;\, \gamma) + r_{FP}}{E_T(t;\, \gamma)} \equiv \frac{C(t)}{T(t)}, \qquad (13)
$$

where $\gamma = (\beta_1, \beta_2, p_{01}(t_k), p_{10}(t_k))$.

3 OPTIMUM SOLUTION AND CHARACTERISTICS OF THE DETERMINATION METHOD

Let $m_I \in \mathbb{N}$, $t_{m_I} = (t_1, t_2, \ldots, t_{m_I})$ and $D(t; \eta) \equiv C(t_{m_I}) - \eta T(t_{m_I})$. Then the optimum inspection time vector t^* is obtained by the following algorithm.

begin

done := false, $m_I := 1$;

Solve the following nonlinear simultaneous equations for (t_{m_I}, η):

$$
\begin{cases}
\dfrac{\partial}{\partial t_{m_I}} D(t_{m_I}; \eta) = 0 \\
\min D(t_{m_I}; \eta) = 0,
\end{cases}
$$

and obtain the value of $r(t_{m_l})$.

if $t_{m_l} = \emptyset$ **then**

$t^* := \infty$;

else

repeat

Solve the following problem for (t_{m_l+1}, η):

$$\begin{cases} \dfrac{\partial}{\partial t_{m_l+1}} D(t_{m_l+1}; \eta) = 0 \\ \min D(t_{m_l+1}; \eta) = 0 \\ t_i - t_{i-1} \geq \varepsilon, \qquad\qquad i = 1, 2, \ldots, m_l + 1, \end{cases}$$

and obtain the value of $r(t_{m_l+1})$, where ε is a positive constant.

if $t_{m_l+1} = \emptyset$ **or** $r(t_{m_l+1}) \geq r(t_{m_l})$ **then**

$t^* := t_{m_l}$;

done := true;

else $m_l := m_l + 1$;

fi

until *done* = true;

fi

end

We investigate the relationship among the inspection time vector, the cost parameters, parameters of $f_1(t; \beta_1)$ and $g(t; \beta_2)$, and the FD and FS probabilities. An optimum inspection time vector is obtained under various probabilities $p_{01}(t)$ and $p_{10}(t)$, cost ratios r_{FP} and r_{IP}, and parameters of $f_1(t; \beta_1)$ and $g(t; \beta_2)$. The value of E_C/E_T when t^* is denoted by $(E_C/E_T)^*$. The case when $D(t; \beta) = \beta t$ is considered for simplicity, where β is a positive constant. Parameters for $f_1(t; \beta_1)$ and $g(t; \beta_2)$ are $\beta_1 = (1,500, 0.05)$ and $\beta_2 = (1,525, 0.05)$ where $\beta_i = (\beta, D_i, \eta)$. Probabilities $(p_{01}(t), p_{10}(t)) = (0, 0)$, $(0.2, 0)$, $(0, 0.2)$, and times for a preventive replacement M_P and for a failure replacement M_F are set to 1 and 7, respectively.

Relationships among r_{FP}, r_{IP}, R_i and $(E_C/E_T)^*$ when $\varepsilon = 1$ are shown in Fig. 3, where R_i is the region in which i times of inspection minimizes E_C/E_T. Dotted lines correspond to $(E_C/E_T)^*$. Figure 3 (a) is the case when $(p_{01}(t), p_{10}(t)) = (0, 0)$. Figs. 3 (b) and (c) are the cases when $(p_{01}(t), p_{10}(t)) = (0.2, 0)$ and $(p_{01}(t), p_{10}(t)) = (0, 0.2)$, respectively, both of which are omitted to save space. Table 1 shows examples of t^* and $(E_C/E_T)^*$ for 9 combinations of (r_{FP}, r_{IP}).

Let L and S be a large value and a small value, respectively. It is confirmed that when $(r_{FP}, r_{IP}) = (S, L)$ no inspection minimizes

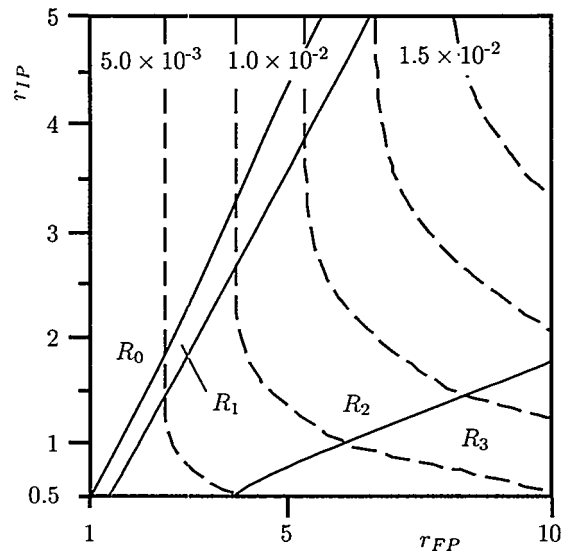


Figure 3 (a). Relationship among r_{FP}, r_{IP}, t^* and $(E_C/E_T)^*$ when $\beta_1 = (1, 500, 0.05)$, $\beta_2 = (1, 525, 0.05)$, $(p_{01}(t), p_{10}(t)) = (0, 0)$.

E_C/E_T. On the other hand, when $(r_{FP}, r_{IP}) = (L, S)$ an appropriate inspection interval minimizes E_C/E_T. The number of inspections becomes large as (r_{FP}, r_{IP}) approaches (L, S).

From Figs. 3 (a) and (c), it follows that $(E_C/E_T)^*$, when FD probability is 0, is smaller than that when FD probability is 0.2. Also, from Figs. 3 (b) and (c), it is clear that $(E_C/E_T)^*$, when the combination of FD and FS probabilities is (0.2, 0), is smaller than that for (0, 0.2). However, from Figs. 3 (a) and (b), we can conclude that only if $r_{FP} = S$ or $r_{IP} = S$, the minimum value of the objective function in

Table 1. Examples of t^* and $(E_C/E_T)^*$; t^*: upper cell, $(E_C/E_T)^*$: lower cell.

r_{IP}	r_{FP}		
	1	5	10
0.5	∞	481, 504, 529	478, 502, 527
	1.8751×10^{-3}	5.3688×10^{-3}	7.2590×10^{-3}
2.5	∞	536, 561	510, 535
	1.8751×10^{-3}	9.0618×10^{-3}	1.3427×10^{-2}
5	∞	∞	529, 554
	1.8751×10^{-3}	9.3752×10^{-3}	1.7429×10^{-2}

Fig. 3 (a) is smaller than that in Fig. 3 (b). Therefore, to reduce the long-run average cost per unit time, there are two solutions. One is to decrease FD probability. In addition to decreasing FD probability, if $r_{IP} = S$ or $r_{FP} = S$ the other solution is to reduce FS probability. Otherwise, a large value of FS probability is acceptable.

4 CONCLUSIONS

In this research, we have developed a method for determining the discrete time points of inspection monitoring for a single-unit system under condition-based maintenance. The transitions between the states are described using an SDT model in which the transition time from a normal state to a symptom state and from a symptom state to a failed state have arbitrary probability density functions. The transition times are derived from continuous deterioration processes based on their trend curve. A method for determining the optimum inspection time vector which minimizes the long-run average cost per unit time is developed considering FD and FS probabilities for imperfect inspection. The relationships among the optimum inspection time vector, the minimum long-run average cost per unit time, inspection cost, preventive maintenance cost and breakdown maintenance cost are investigated.

REFERENCES

1. Christer, A. H. "Delay Time Models of Industrial Inspection Maintenance Problems", Journal of the Operational Research Society, 1984, 401–406.
2. Christer, A. H. & Wang, W. "A Model of Condition Monitoring of a Production Plant", International Journal of Production Research, 1992, 2199–2211.
3. Lu, C. J. & Meeker, W. Q. "Using Degradation Measures to Estimate a Time-to-Failure Distribution", Technometrics, 1993, 161–174.
4. Okumura, S., Jardine, A. K. S & Yamashina, H. "An Inspection Policy for a Deteriorating Single-Unit System Characterized by a Delay-Time Model", International Journal of Production Research, 1996, 2441–2460.

THE NEW CONCEPT FOR INTEGRATED CONDITION MONITORING & PREDICTIVE MAINTENANCE TOOL IS AN ENVIRONMENTAL ADVANTAGE

Leif Berthagen, Managing Director, Eng.
ACCEDE AB, Infrared Tech., Box 7, S-26108 Glumslöv, Sweden
Tel +4641873377, Fax +4641873144, E-mail: berthagen@swipnet.se

Göran Jonasson, Managing Director, Eng.
DIATEK Vibrationsteknik AB, Rattg.1, S-262 62 Landskrona, Sweden
Tel +46 418 20465, Fax +46418 20475, E-mail: cm@diatek.se

ABSTRACT

Machine Condition Monitoring has been practised for more than 30 years using Vibration Measurements. Many different techniques and instruments have been developed, and a lot of individuals has developed experiences as well as theoretical competence in how to make use of the information collected. Many attempts to develop software assistance, help for the evaluation and diagnostic work are known. Today still most systems require an experienced expert and a lot of time to carry out the analysis and conclude if it is a problem or not, what type and how serious. Next steps is to compare with previous measurements, conclude the severity, carry out the trend analysis in order to tell how long time the machinery still can operate safe, and finally advise the action required.

The US Navy wanted to reduce the maintenance costs for the big aircraft carriers but was not very impressed by the systems available on the market. A project was launched with the aim to develop a system with an automated diagnostic software. The requirements was hard ; easy to set up and learn, little or no requirements for education, to be used by both novice and experienced people, to include extensive educational and help functions, easy and safe to change operator but still with a high accuracy of results, fast, reproducible and reliable measurements, automated diagnostic software easy to use, CD-ROM for training and trouble shooting.

The result after 5 years of hard work for the DLI Engineering was the DLI Watchman™ System; 3500 vibration technical rules was originally developed and the system capable giving 650 different answers about machine failures, 10 for each machine with trend analysis for 16 types. More than 1000 new rules have since been developed, this work is continuously going on and the new rules are checked versus a database with 23.000 machines with known history. The Watchman™ System is used as a mobile round datalogger measurement system which has further been developed and as well to an on-line system using the same diagnostic software.

The further development and market requirements has today resulted in the Condition Assessment System MOM™ (Mother of Maintenance) and further developed to be the VOYAGER™ which is a fully integrated, multiple-technology Condition Monitoring system. Future systems will be more integrated condition information systems linked to the machines with the CAS, Condition Assessment System and further linked to the Computerised Maintenance Management Systems. The required performance of this concept must have a very reliable software diagnostic system with the capacity to determine the type and severity of the problems and giving the recommended actions required in plain language.

This paper and presentation describes how this is possible, and by case reports give examples of practical experiences. Hence, how this kind of system can be used in order to improve the operating efficiency, safety and environmental advantages. The importance to consider each machine object individually as well as the different load conditions. The technique is applied to machinery of offshore and onshore drilling platforms, shipping, submarines, in mines, various industries as e.g. power plants, wind power, pharmaceutical, chemical, pulp and paper, etc. Using an efficient tool for the maintenance work it is possible to cover many more and as well smaller objects than ever before using less manpower and hence be able to focus more on solving and preventing, i.e. the Pro-Active Maintenance. The cost - benefit analysis for buying and using CM is an important issue which had to be combined with the company policy and strategy of the total integrated maintenance influencing organisation and personnel. The demands and requirements of the "Tool" used are high in order that the major part of the CM work shall be easy routine and qualified experienced people can work proactive, continuously reducing future maintenance costs.

1 INTRODUCTION

Machinery troubleshooting is a mixture of engineering talent and technological knowledge. It is not a desired subject in engineering schools and the expertise is often only gained through direct experience. This experience have been accumulated in our companies and the personnel involved since around 1970 at an increasingly rapid rate and today we are able to provide quick, accurate analysis of a full range of dynamic systems. We have originally specialised in vibration, which with the collaboration with PREDICT/DLI reached another dimension of the topic, and hence introducing wear particle and oil analysis as an integrated tool to determine rotating machinery problems a feel of being master of recommending repair and maintenance actions was visualised. Utilising finite element analysis, we design modifications to correct structural problems in machinery, structures and mechanical components to eliminate vibration and fatigue problems.

The practical problem-solving expertise has led to the design and implementation of many types of custom solutions. The proven ability by Predict/DLI to manage large projects cost effectively has won major contracts with the US Navy and led to the development of the award-winning automated vibration diagnostic system DLI Watchman™ and the further engineering innovations for machinery condition monitoring have led to a broad line of products in this field used by engineers and consultants around the world.

Development of specific periodic testing and analysis programs for machinery asset management as well has been a major part of the work, and the work with original equipment manufacturers to incorporate CM systems in their production lines to assess product quality combined with the condition diagnostic and predictive maintenance.

Engineers are today involved to an increasing extent in developing, manufacturing and installing custom designed instrumentation and systems for automated on-line CM of critical machinery. The modular designed Fulltime™ system for continuous monitoring use the expert diagnostic software, ExpertALERT™ previously developed, to screen huge amounts of data and produce specific plain language repair recommendations that will minimise unscheduled downtime and maintenance costs.

However, even if the products, hard and software have been developed to an ultimate level both with regard to vibration and used oil & wear particle analysis the need for integration of those and other types / methods of CM systems and techniques has became more interesting. Further the software systems available for Maintenance Management have been more capable together with the CP computer developments. Systems for electrical motor Current analysis has been further developed and especially Thermography is more widely used for CM purposes.

The future requirements when plant optimisation and environmental protection shall be the key issues has encouraged PREDICT/DLI to develop the next generation of Predictive Maintenance tools, i.e. hardware and software concepts. This has been possible because of big US government contracts, a number of big US company customers requirements and the big US market area. Important fact is as well that it was necessary to integrate the vibration engineering from DLI and the UOA-WPA from PREDICT and then it became natural to include the other items in one concept and update the software to Windows 95 and NT.

The result is the *CAS*, Condition Assessment System: **Voyager™**, the fully Integrated multi-technology Condition Monitoring System Linking Predictive Maintenance and Computerised Maintenance Management Systems.

2 VIBRATION SYSTEMS PRACTICAL EXPERIENCE

The DLI Watchman™ system has been in use for almost 10 years and close to 1.000 systems supplied all over the world. The ExpertALERT™ automated diagnostic software has reached version 8.0 and by the ViewALERT version entering the Windows software compatibility. By the input of information from a big number of big customers use of the system the software has been continuously improved and is today known to be very reliable. The follow up of the automated diagnostic results versus the verified cause of failure by inspecting the machinery of many thousands of machines indicate 97 % of accuracy.

The high accuracy is achieved by means of the complete system design, which can be divided into 3 parts; -the hardware, -the software, -the combined use which is much related to the collection of data at the machine location. The system uses the triaxial accelerometer attached to the fixed installed brass or stainless steel coded measurement pads (glued, screwed or welded to the solid machine surface) each identified with a bar code attached close to each pad.

The key to accurate measurements is the measurement itself and it is often people get confused about the relation between accuracy and repeatability in condition monitoring. Vibration analysis is useful in diagnosing machine faults because one can infer changes in condition or poor condition based on the relative amplitude of spectral peaks. This requires very good repeatability between measurements taken on the same machine over time, or for machine-to-machine comparison. It translates into using good quality instrumentation and strict attention to technique. Technique is where most programs run into problems due to incorrect techniques taught by the vendors of condition monitoring equipment. To make routine data collection easy and fast, most have sacrificed both accuracy and repeatability of the measurements.
Important is that an accurate measurement is fine if it is repeatable, but if it is not it is useless. The key for CM is the detection of variations, sometimes very slight, in the amplitude at the forcing frequencies generated by the components of the machine. The mounting position of the sensor should be such that the path of vibration transmission, especially important for high frequencies, is a minimum and passes through a relatively rigid structure.

With the measurement pad located near the bearing the necessary frequencies are detectable and highly repeatable. This may not, however, be true for Spike Energy and similar readings which may rely on higher frequencies which are rapidly attenuated as they travel through the structure. The higher the frequency, the faster this occurs. Since most defects occur in the load zone or are difficult to detect otherwise, it makes sense to make the measurement there. We have found empirically that in the frequency ranges, from 2 Hz to 10 kHz, the presence of a bearing defect can easily be detected without locating the sensor in the load zone. For vertically oriented machines, this point is invalid and we have no difficulty detecting bearing defects.

In practice, the sensor and sensor mounting method used by the PREDICT/DLI systems is superior to either handheld sensors or magnetic base sensors and the complete concept including software set-up for each machine and the individual measurement locations is outstanding and the true state of the art. The attachment pad mounted triaxial sensor used for the DLI Watchman data collection systems meets and exceeds rigid US military requirements for quality (frequency response, dynamic range, repeatability, and signal to noise ratio) set forth in the DOD, Department of Defence contracts (N00406-94-D-0025 and N00033-91-C-3117) and e.g. no handheld probe even comes close to meeting these requirements.

The US Navy performed follow up of the results from use of the DLI Watchman System and active predictive maintenance during a 7 year period showed that after an initial period of 3 years, when old problems was found and corrected, it was possible to reduce the time and running hours planned maintenance revisions by approx. 60%. The need for corrective actions, as balancing and shaft alignments, was drastically reduced due the specified requirements and frequent follow up measurements. The cost benefit was found to be considerable as not only less revisions was required, but the plant or object availability and machine lifetime increased. The efficiency (less energy loss) increased up to approx. 15 % for machinery which suffered from misalignment problems before.

3 CASE REPORT FROM THE CUSTOMER CEMENTA

by Mr Rolf Olsson, Site Maintenance Engineer, Cementa AB.

In cement factories there is a lot of heavy-duty machines. Very important machines in our plant is the two gearboxes for the driving of our main production line - the kiln.

In October -95 we suddenly heard strange noises from one of the two gearboxes for the kiln drive. Before, we hired a company that performed vibration analyses on critical machinery. The company performed their vibration analyses and found something wrong. They could not specify the exact problem cause. We still did not know what the problem was.

At that time we were about to invest in our own vibration analysis equipment, so we asked different companies to present their equipment and as a field test we challenged them to take measurements at the gearbox. This was our first meeting with the Watchman™ system and its capabilities.

At first we were suspicious about a system that says in plain language what is wrong and what to do about it. The first analysis did not detect any errors in the gearbox (one measurement). Then the measurements from the two similar gearboxes were compared by the software and an error detected (the other measurement was used for creating the average to compare with). When we opened the box we found that the Watchman™ System had been correct in the analysis.

Due to the fact that we could detect the damage, it was possible to check the availability of spare parts and to plan the shutdown for the repair action.
It is very difficult to calculate the money saved thanks to the system Watchman™. The greatest advantage was that we could minimise the shutdown time for our 24 hour running main production line. If the shutdown had occurred without any notice it would have taken at least two days longer to repair the gearbox, meaning approximately 10.000 tonnes of clinker in production losses (clinker is the first step in the cement production).

We are now in our first stage of implementing the system Watchman at the Slite plant. It is a professional system that will help any organisation to lower the maintenance costs and to steer the work towards preventive maintenance.

4 CONDITION MONITORING ASSESSMENT SYSTEM

The future requirements for optimisation both from technical and financial point of view require a fully integrated, multi-technology condition monitoring system capable of as close as possible to be 100% accurate in the conclusion and recommended action. This shall require the combined use of different systems and technologies from different vendors, the key technologies are Vibration, Used Oil & Wear Particle Analysis and Thermography. Those techniques can be combined and is often found to be extremely important to confirm, verify and complement each other, even if several vendors claim their system to be the only and ultimate solution, often the advanced customers and users have learned more.

PREDICT/DLI has had an understanding about those things quite some years because of the many field assignments and service missions. This is the reason for designing the Voyager™ software to fulfil those requirements and the important items are to be;
one software package with one database, combine; vibration, used oil and wear particle analysis, process/performance, infrared and the ability to add customer own technologies and tests; one system for up-to-date equipment status; one system giving access to all information (equipment condition, schedule, history, notes, images, measured data, recommendations); report generator with access to all information; structured access to technicians, engineers and managers; consistent user interface throughout the system; open SQL database supports flexible access of all customer data and information; multi-user, network operation (supporting client/server operation).

The issue being to turn measurement data into useful information. The modern maintenance engineer shall not have any time playing with software or endlessly study graphs and wrestle with data collectors, the target for the work is to keep the plant running to the lowest cost and man-hours. The raw data, all spectra, frequencies and levels is really of no interest to the management. The decisions can be improved and performed timely, the information can be shared with other computerised systems in a network or by internet. It is only one system to learn about and it is easy to grow into new technologies.

The DayToDay menu gives fast access to all commonly used functions, the user configurable toolbar in the Windows environment is easy to use for anyone and the context sensitive help, "Tooltips" with tutorials makes it even more user friendly. The interactive report informs exactly what is happening, includes summary status and details if desired as well as additional analysis and addition of recommendations and notes. The history report includes the chronological list of events as tests, results, alarm conditions, notes, images, and recommendations in a fully interactive mode, i.e. allows to go straight to the data for further analysis. A variety of graphical analysis modes using integrated tools for the complete analysis and a button press create a report. The different interactive analysis is performed by one click and the same to look at a drawing, test point or see the data. Iconized data access for easy and fast use of notes, schedules, reports etc.

The Oil and Wear Analysis are integrated and used uniformly and may be data from an external or own laboratory, small or large and the system is designed for the entry of a large number of samples or connected directly to instruments for automated on-line connection.

The Analyst Station is designed for the analysis of one sample at the time but can as well be used for the entry of test results in a smaller laboratory, allows for viewing and capturing of ferrographic images from the microscope to be defined by the image library.

The Vibration Analysis are integrated by means of the mobile datacollectors or the on-line systems and the Expert software system to present the reports in plain language or powerful graphics as waterfall plots, harmonic cursors, etc. The performing of the system simply requires to run the expert software, normalise, edit the faults and add to the averages, "that's it".

The Process and Performance module is intended for additional inputs for complex machinery, temperatures, pressures, etc., any type of measurements can be collected by the vibration datacollector, stored and trended.

The Thermography module can import images from several camera vendors and perform basic and detailed analysis or simply link to the advanced software supplied by the vendor of the camera.

Special Tests and field data can be included with the other data; motor test, crank web deflection, piston wear, megger readings, ultrasonic, etc. and the screen can include images, instructions, and the spreadsheet for data entry.

5 LINKING PREDICTIVE MAINTENANCE WITH COMPUTERIZED MAINTENANCE MANGEMENT SYSTEMS

This topic conclude the integration of the CM system with the CMMS and simplify the PM database creation, automate the scheduling, generate the work orders, and pass measured data as well as feedback from the CMMS.

6 ACKNOWLEDGEMENTS

The authors wish to thank; Mr Rolf Olsson, Site Maintenance Engineer, Cementa AB, Slite works, the island of Gotland in the Baltic Sea, Sweden for his contribution to this paper with a case report, and to PREDICT/DLI company, the management and all employees involved for the extensive & dedicated support and excellent products, specially Mr Ivan Brown, DLI Engineering Division, Bainbridge Island, WA, for many years of collaboration and support in the CM work.

7 SUMMARY

-The future is in integrated CM and suppliers of PdM systems will develop components that can be used within other systems. The key is to use the standard building blocks as; Windows, ODBC, open database structure type SQL, DDE and OLE

-There is a great deal that can be gained by linking the systems. The biggest obstacle in the short perspective is the lack of a standard "file format" for the transfer of information. Developments of open systems will create opportunities and the MIMOSA group and others in collaborations is on the way to resolve this issue.

-The fully integrated CMS as Voyager is required to meet the future demands, when it is a necessity to be able to convert the raw data fast into useful information in a language that can be understood by anyone. It will allow the grow into new technologies, enable a successful work and to record the success. The environmental demands require less to be replaced, all utilities to be more effectively used, and it will cost more to get rid of the used parts, grease and oil than to buy new.

8 REFERENCES

Further information and references are available by visiting the Web site : **www.predict-dli.com**

ASSET MANAGEMENT SOLUTIONS: A COMPREHENSIVE PLATFORM FOR IMPLEMENTING PREDICTIVE AND PREVENTIVE MAINTENANCE PROGRAMS

Israel E. Alguíndigue and Nancy L. Quist
Fisher Rosemount Systems Inc. 12000 Portland Ave. So,
Burnsville, MN 55337. U.S.A

ABSTRACT

This paper describes the capabilities of Asset Management Solutions (AMS) , a software platform for collecting and managing asset management data from equipment and instrumentation in manufacturing and process plants. The system is a Windows 95/NT platform that allows users to perform online calibration and configuration of devices, as well as document all maintenance related activities. AMS also collects online diagnostic information from smart instrumentation and keeps a list of active alerts to facilitate diagnostics and maintenance of assets in a plant.

To facilitate the implementation of predictive and preventive maintenance programs, AMS is being augmented with capabilities for performing management of diagnostic and maintenance data from mechanical and non-mechanical equipment . AMS gives access to a wealth of monitoring and diagnostic information that makes it the focal point of maintenance activities in a facility. Efforts are also underway to develop interfaces between AMS and condition based maintenance software from leading companies in the industry.

1 INTRODUCTION

The need for predictive and preventive maintenance (PPM) programs has been clearly established and recognized in the literature. Mobley in Reference [1] outlines the benefits of a comprehensive PPM program. Among these benefits is a 50 to 80 percent reduction in maintenance cost, and a productivity increase of 20 to 30 percent. Establishing a comprehensive program poses a number of technical challenges; it requires integrating a wealth of technologies, each having different data collection and data processing requirements. In this paper we describe a new software platform that gives access to performance data from field instrumentation and which provides an excellent means for implementing PPM diagnostics.

The Asset Management Solution (AMS) software consolidates a variety of functions related to monitoring and maintenance of field instrumentation. It provides three important supporting functions: first is serves as a data highway for instrumentation performance data, second it enables archiving and management of such data, and third it makes alerts and other status data immediately available to the user via an Alert Monitor. While the current release of the system focuses on maintenance and monitoring activities related to HART-based field devices and valves, the AMS vision is to broaden the view of a device to include other equipment in a plant (such as motors, fans, compressors, etc.) that impacts the process and to centralize all maintenance and diagnostic activities in a single platform, AMS.

2 THE AMS SYSTEM

AMS is a PC-based software providing online access to field instrumentation, and integrating information about all plant devices into a single database. Traditionally each device manufacturer provided its own point solution application for configuration and maintenance of a device; this approach poses a number of disadvantages including:

- The various applications lack a common interface, making transitions among interfaces difficult.
- Device configuration data, change logs, and device diagnostic data storage are decentralized and cannot be cross-referenced.

- Tasks that could be common to each device are duplicated in separate applications.
- All devices in a plant cannot be viewed simultaneously.

AMS provides an efficient solution to asset management by integrating a number of device management packages into a single platform. The various device applications running under AMS have a common look-and-feel based on the Microsoft Windows operating system, and AMS provides a central repository for data related to all the applications. AMS also allows the users to perform online configuration, calibration, monitoring and troubleshooting of smart field devices from a single computer platform. The current version of the AMS system supports HART® communications. A Fieldbus version of the software will be released in early 1998.

2.1 The HART Communications Protocol

AMS uses the HART protocol to communicate with smart field devices. HART, the Highway Addressable Remote Transducer protocol, is supported by a wide variety of vendors; it is an open communications standard that works with any control system without interrupting the flow of process data. The protocol uses the Bell 202 standard frequency shift keying (FSK) signal to communicate at 1200 baud, superimposed at a low level on the 4 to 20 mA analog measurement signal. Having an average value of zero, an FSK signal causes no interference with the analog signal [2].

With HART, it is possible to collect information from a device, as well as send information to the device. This two-way communication capability makes it possible to calibrate, configure and test devices online, but most importantly to obtain information about the device health. A comprehensive set of diagnostic information can be obtained from the field device. A smart pressure transmitter for example provides a wealth of information beyond basic pressure readings. It can relate information about overpressure condition, overtemperature condition, loss of signals, stuck signals, out of limit variables, failed sensor, failed electronic board, and more. In the case of a control valve, the exact

valve position and its travel profile can be obtained by interrogating the device. This information is of great use in predicting a valve's useful life.

2.2 AMS Architecture

The AMS system is based on open and scaleable technology. Each field device is a node in a network of such devices as illustrated in Figure 1. The AMS Field Server acts as a data highway giving access to the device network and the database information. The Field Server supports a client-server architecture, AMS applications can run on machines remote from the server over a TCP\IP network.

In addition to the Field Server, dedicated SNAP-ON™ applications use process and instrumentation data to provide specific functions for specific devices. This data can also be used to implement comprehensive diagnostic and monitoring activities related to the production process, and the instrumentation running the process. AMS thus provides a hub for diagnostic data as well as a platform for implementing comprehensive PPM programs related to field instrumentation and plant equipment.

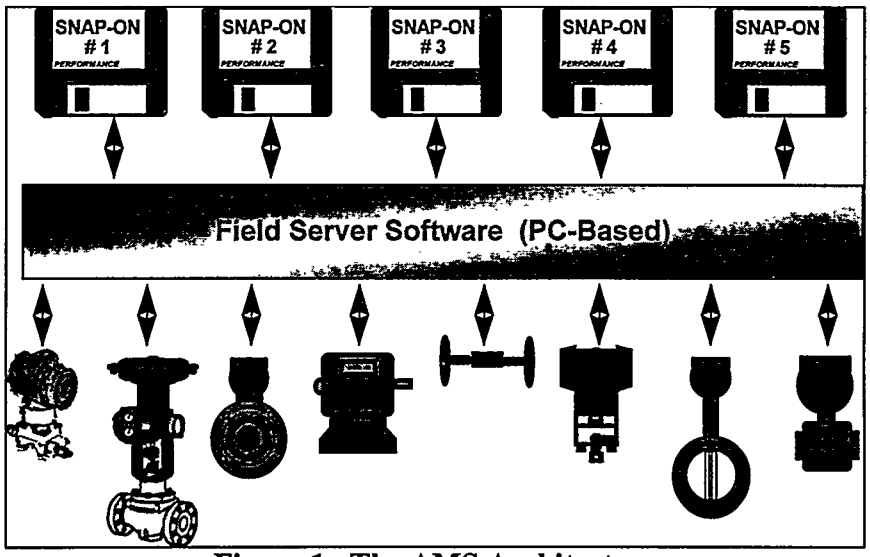


Figure 1. The AMS Architecture

2.3 Current AMS Functionality

AMS allows the user to access device functionality and manage devices without switching applications or learning multiple interfaces and systems. AMS provides some functions that are common to all devices, such as:

- A browser to locate devices on the network
- A global alert list where status (diagnostic) information about all devices is integrated
- Device status monitoring. A standard mechanism for filtering and handling diagnostic conditions identified from devices in the field.
- Streamline maintenance functions (configuration, calibration, self test and loop tests).
- Links to device notes and drawings.
- Audit train for logging maintenance activities and device status alerts.
- DCS interface.

There are also functions in AMS which are device specific. Such information comprises configuration and calibration details that are specific to each family of field devices, as well as status (diagnostic) information and filtering criteria. With these capabilities, it is easy to configure AMS to report and handle status conditions generated by field devices according to customer requirements. Historic information about the device, including its calibration, service, and configuration records can also be obtained using the system.

The following figures illustrate aspects of AMS functionality and user interface.

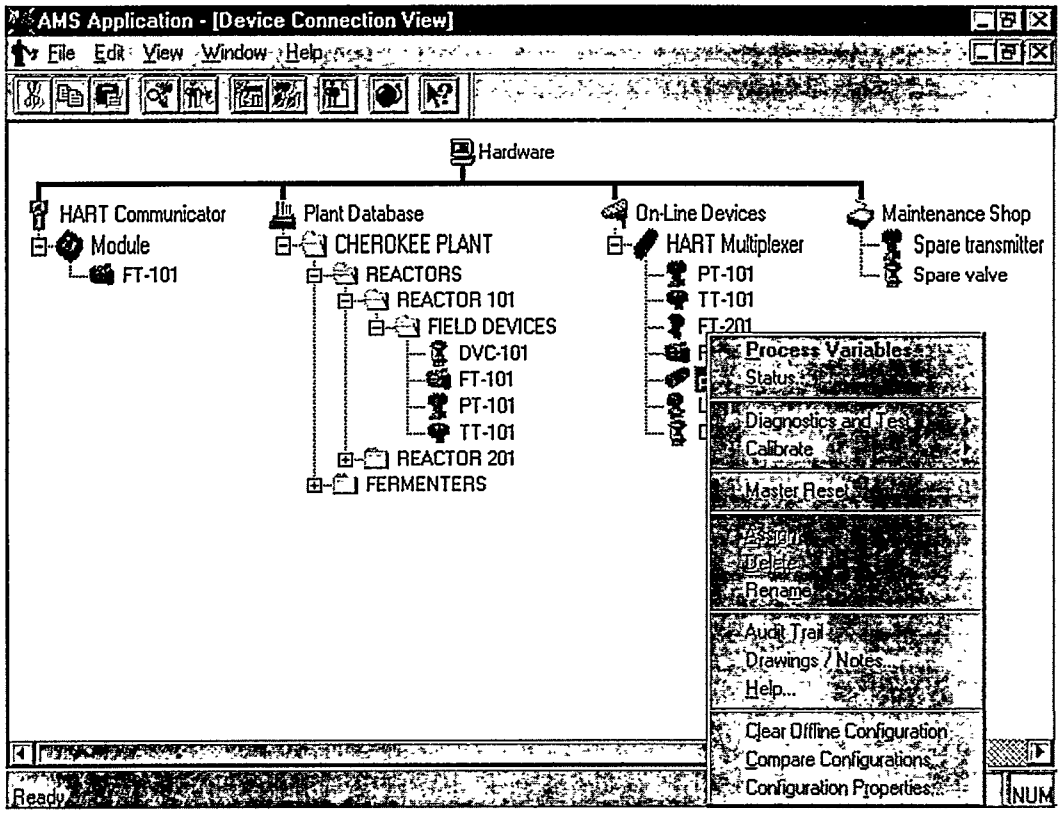


Figure 2. The Device Configuration View of the AMS software provides an easy and intuitive means of organizing and locating field instruments in the plant. The system polls devices to verify active communication links and configuration of the system requires minimal user intervention. The Device Context menu (lower right corner) provides a view to all information related to a device. This includes process variable and status information, as well as configuration and calibration information. The history of device changes is kept in the Audit Trail. Notes/drawings and diagnostics/testing are other available options.

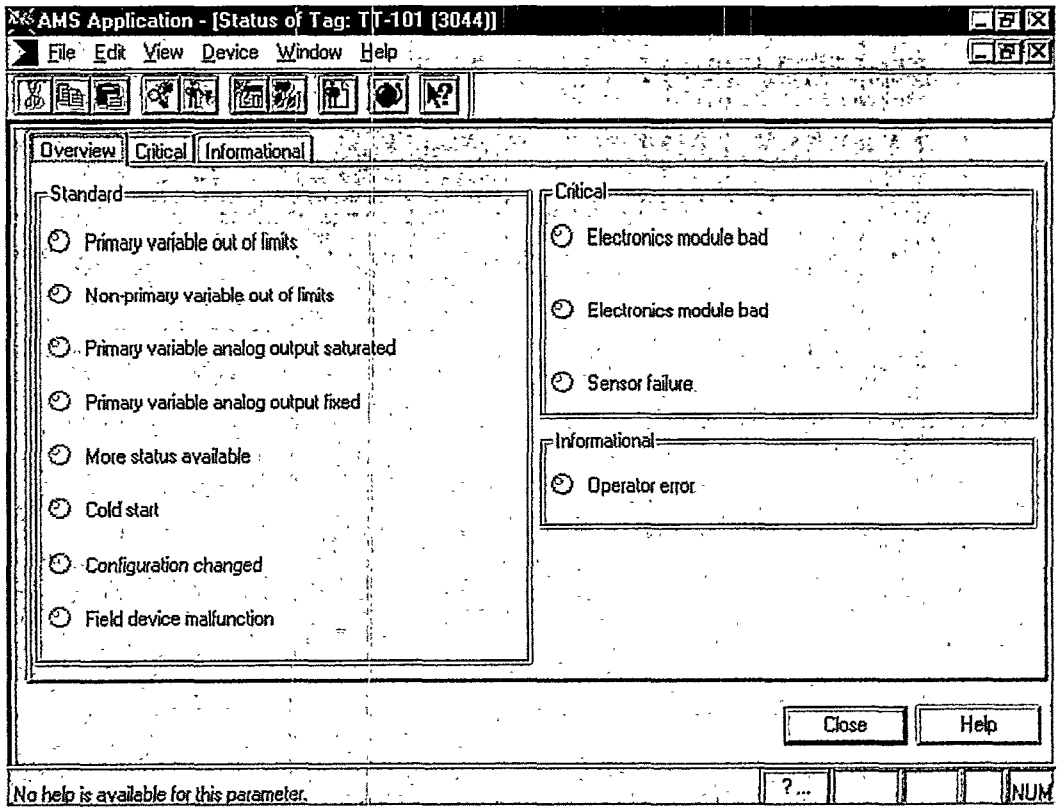


Figure 3. This Status of Tag Screen is used to display information about devices. In the example above, status information for a Rosemount 3044 pressure transmitter is displayed. Status conditions are organized by severity. The status information contains diagnostic information relative to each type of device. This information is generally organized in categories according to its severity. While every device type has its own unique Status Page, the way the information is accessed is standard among all devices in the plant.

Whenever a status condition arises in a device, the information is sent to the Field Server, and displayed in the Alert Monitor (not shown in Figure 3). At any time, the Alert Monitor displays information on current status for devices connected to the AMS Field Server. To prevent nuisance alarms, AMS includes a mechanism for filtering alarms.

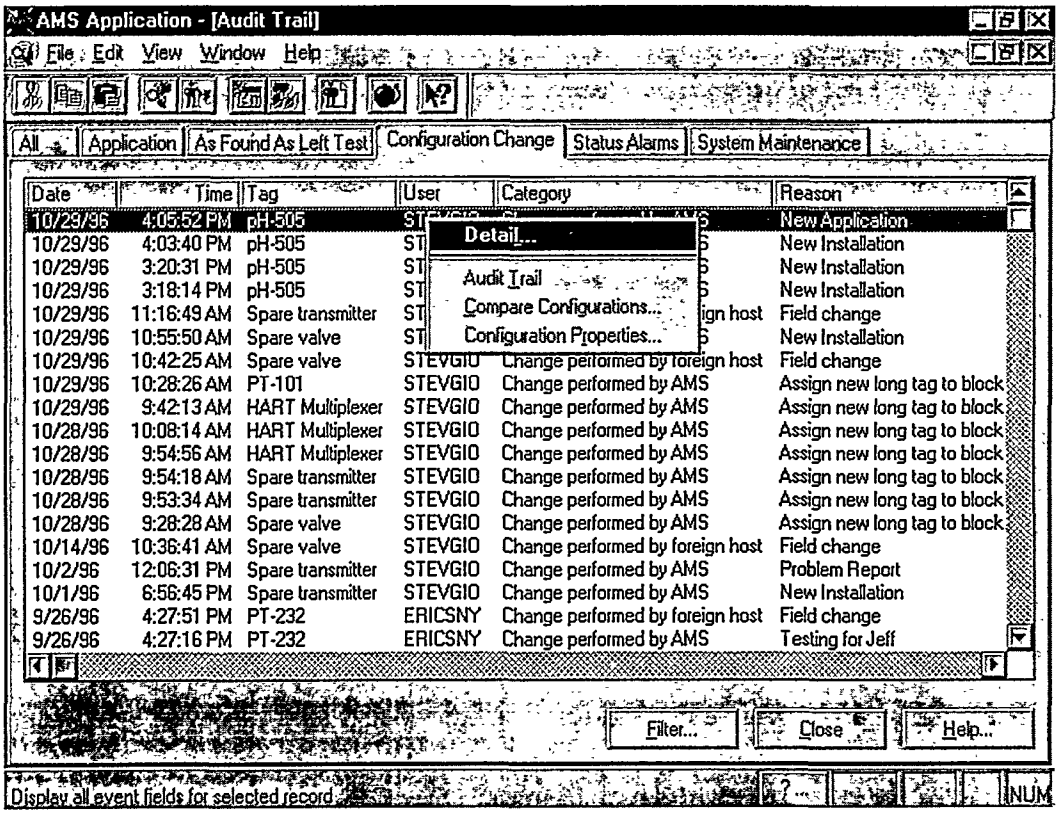


Figure 4. The Audit Trail Screen displays a data log of all configuration and calibration changes to a device. It documents the date and time changes to the device are made, and also indicates event type and reason for the change. Status conditions may also be recorded using the Audit Trail.

2.4 AMS Value Proposition

Below are some examples of substantial savings that can be accrued using the AMS system today

- Because the cost of a valve failure is high, the typical practice is to periodically schedule valves for repair. Thus, they are pulled from service and rebuilt whether it is needed or not. A study of 230 control valves scheduled for turnaround identified that only one-third of them actually required removal from the line. Sixty-four percent could be adjusted or repaired in line, thus reducing significantly the

total cost for repair of the valves [3]. With AMS the status information for valves is easily accessible. The ValveLink SNAP-ON application of AMS will allow the user to perform advanced online valve diagnostics.

- A multinational chemical company reviewed their instrument maintenance work orders to identify that 63% of their effort was expended verifying proper instrument operation. This practice could be eliminated by simple diagnostic tests to identify that the status of the device is OK [3]. AMS provides access to this kind of data.

In general savings of $150/device and $200/valve during start-up, and $ 380 per year for devices and $ 900 per year for valves are not uncommon during normal operations [4]. These savings are due to:

- the implementation of PPM functions based on available diagnostic data,
- the ability to calibrate and configure devices online,
- the availability to automatically document all maintenance and configuration activities for a device. From these historical records compliance and maintenance reports can be generated.

3 THE AMS VISION

The AMS architecture was designed to support the development of stand alone applications which access data from the AMS Field Server and provide diagnostic and monitoring functions. These are SNAP-ON applications to the AMS system. These SNAP-ON applications augment the capabilities of the system by adding a number of device-specific and process related technologies for predictive and preventive maintenance.

A number of SNAP-ON applications are being developed to provide advanced diagnostic capabilities for smart instrumentation. The new applications incorporate state-of-the-art diagnostic technology, and allow for the automation of a number of maintenance functions. Since the status information from a device is readily accessible via the AMS Field Server, monitoring and maintenance application programs are able

to obtain the data needed for diagnostics in a totally autonomous manner. These applications also provide online troubleshooting and repair help.

The vision of AMS goes beyond providing maintenance support for field devices. AMS SNAP-ON applications may also be built to support interfaces to other maintenance and monitoring packages. Particularly, there is interest in supporting other devices in the plant that are not generally considered devices (such as motors, compressors and fans). To this end Fisher- Rosemount is working on establishing partnerships with leaders in the machine conditioning market who are willing to participate in this vision of providing a total asset management solution to their customer.

Figure 5 illustrates the principal building components of this vision. At the bottom of the architecture are the field devices and other plant machinery that impact the process. The AMS Field Server provides access to data from devices in the field, and makes it available to applications. Diagnostic applications use the device diagnostic data to support a number of PPM functions such as motor diagnostics, vibration analysis, oil analysis, thermography, data refining, and others. The AMS system becomes the focal point of all diagnostic-based maintenance in the plant, and allows the maintenance crew to "see" the entire plant operation. This is turn permits diagnostics events to be correlated among measurements and eases the task of identifying root causes. At the very top of the architecture, diagnostic and maintenance data is made available to other applications. CMMS packages, for example, can be interfaced with AMS to provide accounting, scheduling and planning functions.

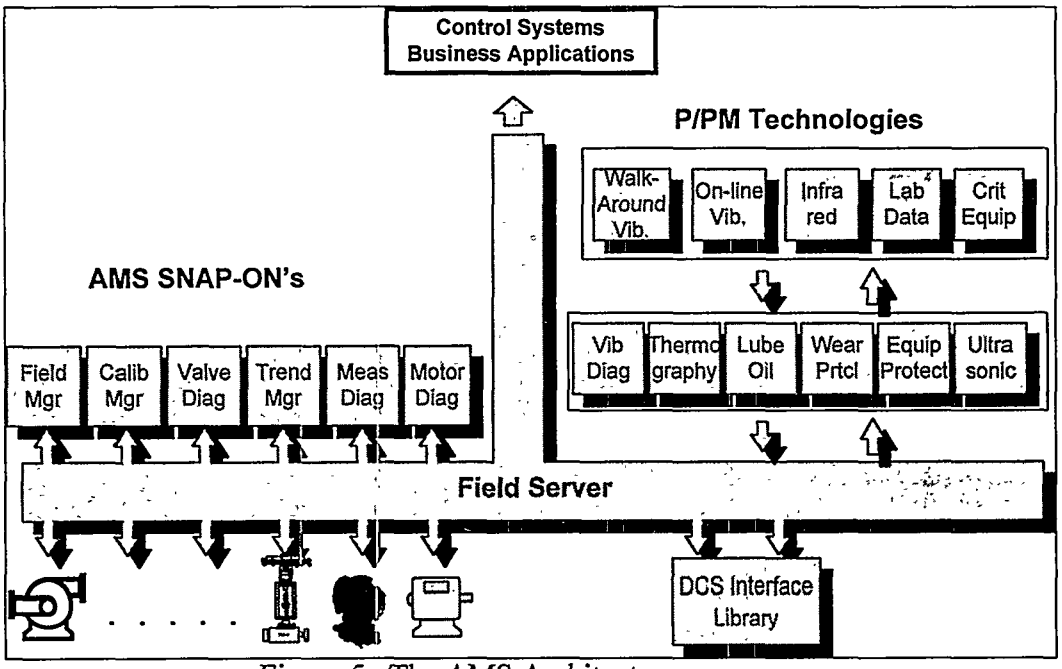


Figure 5. The AMS Architecture

4 CONCLUSIONS

The Asset Management Solution software from Fisher-Rosemount provides a means for managing diagnostic information about field devices, information that can be used to implement PPM strategies. The software also allows maintenance personnel to perform online calibration and configuration, and to keep historic records of all information related to device performance. The Alert Monitor in the AMS package collects status (diagnostic) information from field devices and keeps a list of such active conditions. This feature of the system facilitates the task of monitoring smart instruments. With smart instrumentation rapidly embedding complete diagnostic capabilities in the device, AMS becomes an essential component of any comprehensive PPM program. The current version of the AMS system uses the HART protocol to communicate with devices in the field. A Fieldbus implementation is due early in 1998.

The capabilities of the AMS system are being augmented with SNAP-ON applications to support specific functions for smart field devices, as

well as for mechanical and non-mechanical plant equipment. The open architecture of the AMS system makes it convenient to develop these applications because it makes available vast diagnostic information about the process and the equipment. Alliances with condition maintenance system providers will ensure that condition maintenance functions can also be launched from within the AMS system. Under this scheme, AMS becomes the focal point of all asset management functions in a plant.

5 REFERENCES

[1] Mobley, R., *An Introduction to Predictive Maintenance*, Van Nostrand Reinhold, 1990

[2] HART User Group, HART: A Technical Overview, Rosemount Inc., Publication No. MAN 4692, 1993

[3] Lenz, G., Diagnostics: Current Applications and Future Trends, Proceedings of the Workshop on NDE and Diagnostics Needs For Industrial Impact, Chicago, Illinois, October 7, 1996.

[4] Giovannelli, S., Personal Communications, Fisher Rosemount Systems, March 5, 1997.

TECHNICAL DIAGNOSTIC MANAGEMENT
Terminology of Objectives, Functions, Purposes, Means

Dirk van Schrick, Senior Engineer
University of Wuppertal, Safety Control Engineering,
Wuppertal, Germany

ABSTRACT

This contribution is not concerned with maintenance strategies, monitoring tools, diagnostic procedures or failure prevention techniques as distinct to 99% of papers presented at conferences, congresses or symposia. This contribution is concerned with concepts, terms and definitions within the context of managing failures and faults in technical systems. The paper provides a description of a model of concepts termed *Technical Diagnostic Management Model (TDM Model)* and focusses two partial models based on a *failure and fault conception (FFC)*. The partial models are termed *Technical Management Model (TM Model)* and *Management System Model (MS Model)* resulting from a complex modeling process and intended as a first step for discussing terminological problems within COMADEM and the diagnostic community worldwide.

1 INTRODUCTION

In many technical and non-technical fields concepts just as their terms and definitions constitute the basis for making oneself understood to each other when talking about things of the non-linguistic reality. During congresses or workshops as well as with the study of specialized literature it is noticeable all the time that contributions on terminology in the field of topics of interest are very rarely. In the field of automation for example, this fact is reflected in form of talking and writing about different subjects such as modeling, controller design, identification or diagnosis in technical systems. Contributions in which no methods and no applications but concepts, termed or defined, and their incorporation into a general context are focussed on can be found only here and there.

In the terminology field of system supervision a great number of

concepts can be found which built up on the basic concepts *failure* and *fault*. These concepts often expressed through two-constituent words and their thematically related concepts form the basic set of concepts for developing a complex adaptive model of concepts. The set of concepts has to be realized, classified and structurized to be able to specify an actual, homogeneous, open, comprehensible, generally acknowledged and well-grounded representation of the total of concepts in shape of a model of concepts called *Technical Diagnostic Management Model (TDM Model)*. This model of concepts comprises the concepts of interest and their mutual relations depending on different given aspects, but not their terms and definitions. Thus, a terminology in sense of a fixed system of terms for the communication upon a field of interest can be developed to reduce communicational barriers.

The contribution is structured as follows: Part 2 gives a short listing of the modeling process auxiliaries and a short description of both the *failure and fault conception (FFC)* and the structure of the TDM Model. Part 3 provides a description of the *Technical Management Model (TM Model)* focussing the objectives of an organization in relation to its sections, levels and strategies. Part 4 provides a description of the *Management System Model (MS Model)* focussing the functions, purposes and means of the technical MS part termed TDM system and absolutely necessary for approaching the management objectives.

2 MODELING AND MODEL

Within the TDM context, the modeling process is based on the auxiliaries *aspects, conceptions* and *principles* for both structuring the set of concepts and classifying the concepts itself depending on given aspects. General aspects of interest are *causality, temporality, modality, locality, structure* and *hierarchy*. General conceptions of interest are related to *failure and fault, levels, management, system* and *means*. General guidelines are based on the *principle of comparison, principle of causality, principle of structure, principle of decomposition* and *principle of economy*. As an example and fundamentals for the partial TDM models, the *failure and fault conception (FFC)* will be explained in the following.

The FFC is related to the concept *item*. An item is that which can be individually described and considered, it may be a process, a product, an organization, a system, a person or any combination thereof. Regarding the fulfilment of specified requirements stated

in a functional specification for example, a technical item, such as a machine, can take the two states *conformity* and *nonconformity*. In case of nonconformity, there exists a fault and the item is said to be in a *faulty state (FS)* that can be active or inactive with respect to the possibility to be a cause for another failure. The faulty state FS is characterized by the deviation of a quantity from its desired value. The resulting difference between the desired and the actual value is the content of the concept *error*. The transition from conformity to nonconformity is an event and determines the content of the concept *failure (F)*. The causes for such an undesired state transition, possibly leading to a fault showing up as a faulty state FS that manifests themself as an error, are any undesired phenomena such as disturbances, human errors or inactive faulty states. The totality of these pre-operational or operational phenomena is subject of the concept *fault cause (FC)*. The resulting causal nexus *fault cause - failure - faulty state* can appear repeatedly on one level of consideration or it propagates from lower to higher levels, e. g. a FS on a lower level is a FC on a higher level. Fig. 1 illustrates one realization of the FFC labelled with *Fault Pathology* for the levels *system* and *element*. The faulty state FS

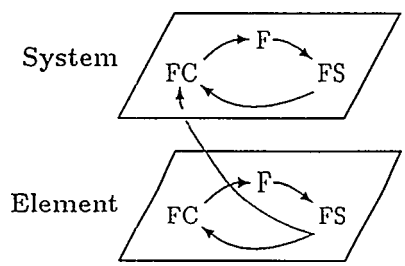


Figure 1: Fault Pathology; Aspect: Faulty State

is subject of consideration in different fields of technology. In both safety and dependability engineering FS is called *failure state* because a failure indicating an undesired state transition is the object of interest. In mechanical engineering FS is called *damage* where an item takes this state after one or more parameters of damage or wear remain under or over a specified limit value of wear capacity. Observing these parameters gives information on the condition of an item, e. g. a machine. The concept *condition* is understood as the totality of characteristics and characteristic values of an item at the time of observation. It is a state in technical sense.

This state-oriented fault pathology is one aspect of the basic partial model of concepts termed *Elements for Consideration Mo-*

del (EC Model) and is not explained further in this contribution, but details can be found in /1/ and different views for example in /2/ or in the literature stated therein. Nevertheless, the basic concepts *failure* and *fault* are of interest according to the TDM objectives and the TDM system. These concepts form a link to the partial models of concepts TM and MS. The link between the TM Model and the MS Model is that a management system is needed to implement the management for approaching its objectives. Consequently, the three partial models are standing in a mutual relation resulting in a structure of the TDM Model of concepts shown in Fig. 2.

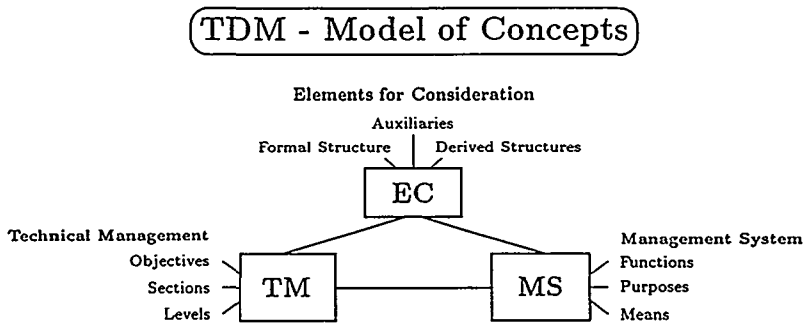


Figure 2: TDM - Model of Concepts; Aspect: Structure

3 TECHNICAL MANAGEMENT

The comprehensive concept *management* can only be reflected correctly with the term "management", if it is evident that *to manage* means the handling comprehensively in all sections and on all levels of an organization and not the activities restricted to top managers. In this case, not only the failure- and fault-related and therefore the diagnosis-related matters are concerned, but also all matters with respect to maintenance, quality, process and organization. The concept *technical* refers to the tasks for the fulfilment of purposes solved with technical means. The tasks are distributed to the organization and concern technical and organizational fields. Moreover, organization-related not desired phenomena affect an item during both the pre-operational and the operational phases and become noticeable as performance insufficiencies of that item, ideally to be detected as early as possible.

All activities of the overall management function of an organization that determine the policy, objectives and resposibilities, and

implement them by appropriate technical means within the management system is understood as *Technical Management (TM)*. To provide a basis for discussion, a partial conceptual model termed *Technical Management Model (TM Model)* is introduced. The structuring of the TM-related set of concepts (TM set) is not only based on an aspect-depended ordering but also on auxiliaries termed *organizational conception, level conception* and *management conception* resulting in a *TM conception (TMC)* including to the concepts *organization, level* and *management*. Fig. 3 illustrates the TMC in form of a TDM building featuring the aspects *objectives*,

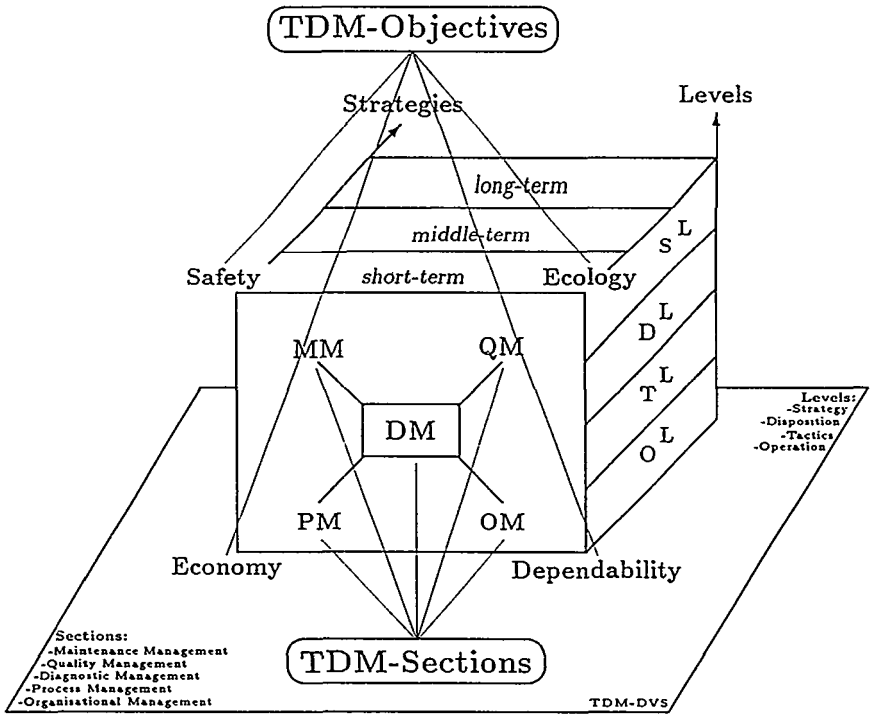


Figure 3: TDM - Building; Aspect: Organization

sections, levels and *strategies*. The objectives of an organisation relating to the sections, levels and strategies will be discussed in the following.

3.1 Objectives

With regard to specified properties of processes and products, the realized ones characterize the degree of approach to the organization-related TDM objectives concerning the aspects *safety*,

ecology, *economy* and *dependability* which represent different constituents of the concept *quality*. The TDM building reflects these objectives with respect to the a structural conception (sections and levels) and the principle of temporality (strategies) according to a general characterization of an organization. In the context of an economical and ecological promising operation, the avoidance of local and global accidents is not only an economic(al) and ecological objective, but it concerns the operation under the aspects *safety* and *dependability* too.

The concept *safety* is associated with the use of an item in health- and life-critical situations. Safety is a principle of the society and is one aspect of its requirements for an item. In contrast, the concept *dependability* is associated with the reliance into the provided service of an item. Dependability is a principle of the customer or user and is one aspect of its requirements for an item. Both concepts *safety* and *dependability* refer to requirements for technical items to be fulfilled with appropriate means to approach to the TDM objectives.

The concepts *Entity Maintenance (EM)*, *Quality Control (QC)*, *Process Control (PC)* and *Replanning (RP)* refer to the sectional means where the concepts *Maintenance Management (MM)*, *Quality Management (QM)*, *Process Management (PM)* and *Organisational Management (OM)* refer to the TDM sections. With respect to the means, strong relations exist between the section DM and the sections MM, QM and OM.

Moreover, all operational techniques, activities and actions have to be applied on all levels of TDM showing the different views onto an organization. A general accepted conception of levels comprises the levels *operation*, *tactics*, *disposition* and *strategy*, cf. /3/. With respects to the aspect *function*, these levels can be assigned to the levels of an organization within manufacturing or production engineering, with levels termed *process*, *process control*, *process supervision*, *production management* and *corporate management* for example. According to the TDM philosophy, all levels are affected to reach the TDM objectives, but both the operational and the tactical level are of main interest because in the TDM context emphasis has been laid on the aspect *technical*. Under the aspect *temporality*, short-, median- and long-term TDM strategies can be stated which contribute to approaching to the objectives, such as planned or unplanned maintenance.

With regard to safety and dependability, both concepts refer to the element for consideration *failure* expressing the link to the EC Model. A failure is understood as an indicator for the undesired

event *transition* of an item from the state of conformity to the state of nonconformity.

The *conception for safety (CS)* is based on the concepts *risk*, *limit risk* and *danger* where the risk quantity R is the parameter defining the state of an item. Depending on R being less or greater than a limit risk R_L, the concept *safety* or *danger* is applicable. Consequently, the concept *safety* refers to state in which the risk of harm to person or damage is limited to an acceptable level.

The *conception for dependability (CD)* consists of both the concept *availability* and the concepts *reliability*, *maintainability* and *maintenance-support*. The latter mark the availability influencing factors. Consequently, the concept *dependability* can be seen as a collective concept used to describe the availability performance and its influencing factors, cf. /4/. In sense of quantitative considerations, appropriate mathematical tools are required to describe the reliability as a time-dependent function, for example.

4 MANAGEMENT SYSTEM

In general, the concept *system* often refers to a totality of technical, organizational or other means for an independent fulfilment of tasks. In the context of TDM, the conception *management system (MS)* comprises the concepts *organizational structure*, *procedures*, *processes* and *means*. The MS, necessary to implement the TM, should be as comprehensive as needed to meet the objectives of the organization. In the following, the technical part of the MS characterized by the concept *TDM system (TDMS)* is considered. Intended as a support for the organization approaching to its TDM objectives, the TDM system has to provide appropriate TDM functions realized with appropriate TDM means to fulfil the TDM purposes. The TDM-related aspects *functions*, *purposes* and *means* are explained in the following.

4.1 Functions

The functions provided by the TDMS have to enable to gain and to process information on faults, being or emerging in an technical item or acting onto it influencing the technical process, to derive appropriate reactions such as maintenance activities or safety-related actions. These functions can be understood as parts of a control loop illustrated in Fig. 4. Within the TDM context, a multi-level

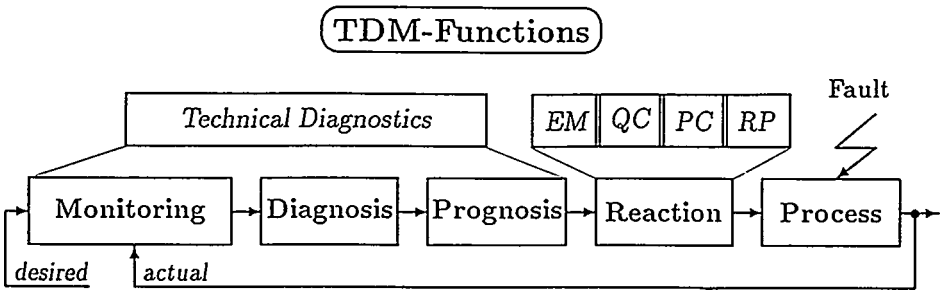


Figure 4: TDM - Control Loop; Aspects: Functions, Section Means

conception of functions exists where the meta-function concepts *monitoring, diagnosis, prognosis* and *reaction* are in superordinate position with respect to different basic-function concepts such as *measuring* or *comparison*. The basic-functions are necessary to form the control error to be processed by the diagnostic and prognostic function, both based also on appropriate basic-functions to maximize fault information. The meta-functions *monitoring, diagnosis* and *prognosis* are realized by the sectional means *Technical Diagnostics.*

The meta-function *reaction*, considered as the controller in the control loop, is realized by the sectional means *EM, QC, PC* and *RP*. With respect to quality control, the feedback is realized not only with process information but also with product information, cf. /5/. In general, the concept *reaction* comprises technical and organizational acticities. The latter have to be provided by the *Organizational Management System (OMS)* based on organizational means for the fulfilment of purposes such as long-term fault prevention, medium-term cause-effect minimization of faults and short-term fault-correction.

4.2 Purposes

The TDM functions fulfil purposes helpful with approaching to the TDM objectives. With regard to the causes of insufficiencies of an item discussed in literature for example, the concept *purpose* very often refers to the element for consideration *fault* expressing the link to the EC Model. The large number of concepts can be structured by the pair of relations *superordination - subordination* resulting in the taxonomy shown in Fig. 5. The main concepts are *fault consideration, fault avoidance, fault sur-*

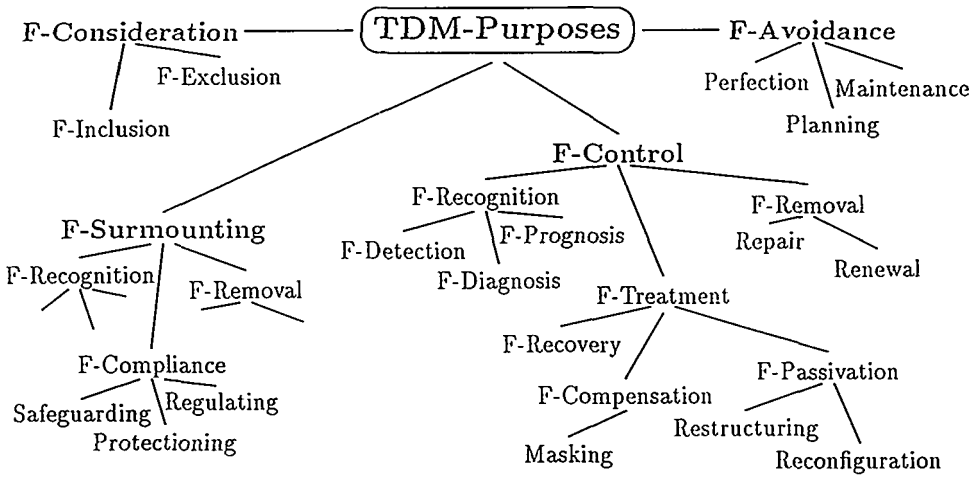


Figure 5: TDM - Taxonomy; Aspect: Purposes

mounting and *fault control.* Both concepts *fault surmounting* and *fault control* subsume the concept *fault recognition* subsuming the fault-related concepts *detection, diagnosis, prognosis* directly associated to the three function-related concepts of Technical Diagnostics. The totality of the remaining concepts is associated to the forth function-related concept *reaction.* With respect to the aspect *structure*, the concept *fault intolerance* subsumes *fault avoidance* and *fault surmounting*, and the concept *fault tolerance* subsumes *fault recognition*, *fault treatment* and *redundancy* where all three concepts are aspects of the joined concept *fault-tolerant control.* The concept *fault avoidance* is associated with short-time pre-operational activities for fault prevention, for example by perfection. The medium-term and long-term operational activities are aspects of the concepts *fault treatment* and *fault removal.* With respect to fault treatment relating to fault causes, this concept includes also the aspect *processing* associated to fault effects, the *errors.*

4.3 Means

For the realization of purposive functions appropriate means have to be provided. The concept *means* refers to the aspects *operation, description, realization, techniques* and *tools.* Additionally, there is a conceptual relation to all phases an item passes through, such as *planning, design, test, operation, maintenance* and *disassembly.*

Actually, more and more attention is paid to the means of description used in all phases and resulting in a large number of promising quantitative and qualitative model-based techniques and relevant information-based tools for neural-net or fuzzy-logic approaches, for example. For further information on technical means for on-line Technical Diagnostics the contributions collected in /6/ are recommended.

5 SUMMARY

In this contribution, a rough outline of a complex conceptional model termed *Technical Diagnostic Management Model* was made. Future work is directed to the concepts *functions* and *means* where both a comprehensive multi-level function-oriented conception and a comprehensive model-based means- and tool-oriented conception are aimed at.

6 REFERENCES

/1/ van Schrick, D. "Ein Begriffsmodell zu den Betrachtungsgrundlagen Fehler, Ausfall, Störung und Versagen", accepted for Presentation and Publishing in VDI-Berichte, GMA-Aussprachetag über Sicherheitstechnik und Automatisierung, 10.-11. April 1997, Langen, Germany, VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik.

/2/ Laprie, J. C. (ed.) "Dependability: Basic Concepts and Terminology", Springer-Verlag, Wien, New York, 1992.

/3/ Erdmann, L., Schnieder, E. and Schielke, A. "Referenzmodell zur Strukturierung von Leitsystemen", Automatisierungstechnik, at 42, No. 5, 1994, pp. 187-197.

/4/ DIN EN ISO 8402. "Qualitätsmanagement; Begriffe", Beuth Verlag GmbH, Berlin, 1995.

/5/ van Schrick, D. "Quality Control Management: Terminology of Functions, Purposes, Means and Methods", submitted for Presentation and Publishing in Proceedings of the 4th Workskop on Intelligent Manaufacturing Systems, July 21-23, 1997, Seoul, Korea.

/6/ Rao, B. K. N. (ed.) "Profitable Condition Monitoring", Kluver Academic Press, 1993.

A FAILURE PREDICTION SYSTEM FOR RECIPROCATING CHILLERS

Dr. Sharayu Tulpule, Manager, and Mr. Paul Duff, Project Engineer,
Service Technology, Carrier Corp., Farmington, CT 06032, USA

Abstract

This paper presents a Failure Prediction and Early Warning Diagnostic System for monitoring the operation of an air cooled Reciprocating chiller, and providing early warning information about the chiller abnormal operation. The system incorporates neural network and thermodynamic model based techniques to detect and predict conditions such as refrigerant leaks, refrigerant overcharge, dirty condenser coils, fan failures, compressor valve failures, sensor noise and bias, and motor overheating. The early warning of impending failures enables initiation of maintenance actions before actual machine failure giving the perception of high reliability.

1. Introduction

Reciprocating chillers provide chilled water for refrigeration purposes, and are required to offer uninterrupted service. Traditionally, Reciprocating chillers have safety features to prevent a unit from operating under improper conditions. The newer models are equipped with electronic controllers with built-in diagnostics. Current diagnostics rely on limit and range checks for a selected set of temperature and pressure values. These limit and range values are based on fixed design conditions. An alarm occurs if a limit value is exceeded. The fixed limit values do not address variations in the normal operation of a chiller due to different ambient conditions or building loads, and sometimes result in nuisance alarms.

In the current controller configuration, before an alarm occurs, there is no indication of any 'trouble brewing' in the chiller. By the time the alarm occurs, it is usually too late and the machine is down. If the alarms are set conservatively to avoid nuisance shutdowns, then, by the time an alarm occurs, the machine may be damaged, requiring major repairs. Minor failures which cause performance degradation often go unnoticed for long periods of time.

This paper presents an Early Warning Diagnostic System (EWDS) to efficiently monitor the performance of a Reciprocating chiller, detect performance degradation, infer impending failures, and provide information to a mechanic to initiate service before a chiller failure. The EWDS is based on an integrated approach combining pattern recognition, and model based approaches to detect, identify, isolate, and predict failures from chiller performance data. The approach utilizes analytical and empirical models to define chiller performance, and provide characterization of chiller operation during nominal and failure modes.

The paper presents this EWDS by first describing the chiller operations and failure modes in Sections 2 and 3. Section 4 describes the EWDS approaches for early detection of chiller abnormal conditions, followed by system description in Section 5. Selected results and conclusions are presented in Section 6 and 7 respectively.

2. Reciprocating Chiller Operation

A Reciprocating chiller typically consists of an evaporator, reciprocating compressor(s), air or water-cooled condenser(s), and expansion valves. In order to produce cold water, the refrigerant follows the vapor compression cycle through the chiller components. The refrigerant gas is compressed to high pressure and high temperature in the compressor. It is allowed to condense to liquid giving off heat to air circulating around the condenser. The design of the condenser allows the liquid refrigerant to cool further to become subcooled liquid. This subcooled liquid passes through the expansion valve - a flow monitoring device, that controls the cooling capacity. The refrigerant evaporates in the evaporator absorbing heat from the water circulating in the evaporator. The cycle is repeated when refrigerant vapor reenters the compressor. The water in the evaporator

gives off heat to the refrigerant, and becomes cold to provide cooling to a building. The controller monitors the expansion device and the amount of air circulating around the condensers to turn the compressors on and off to achieve the required cooling.

Sensors that are typically available on a Reciprocating chiller include:
T1 - Cooler Entering Water Temperature, T2 - Cooler Leaving Water Temperature, T3 - Compressor Suction Temperature, P1 - Compressor Suction Pressure, P2 - Compressor Discharge Pressure, P3 - Compressor Oil Pressure.

3. Reciprocating Chiller Failure Modes

A chiller experiences many failure modes. Only a subset of the failure modes has been selected for early detection (significantly earlier than the current alarms). The selection criteria consist of frequency of occurrence, severity of damage, observability, and the potential impact on maintenance by early detection of a specific failure mode. The selected failure modes for EWDS are:1) Refrigerant Leaks, Undercharge, Overcharge Conditions, 2) Stopped Fans, 3) Plugged Filter Drier, 4) Stuck Expansion Valves, 5) Compressor Valve failure, 6) Compressor Motor Overheating, 7) Sensor Bias and Noise, 8) Water Flow Restriction, 9) Dirty Condenser Coils.

This paper will present descriptions of four failure modes: Refrigerant Leaks, Undercharge, Overcharge conditions, Stuck Expansion Valves, Compressor Motor Overheating, and Dirty Condenser Coils. Other failure modes are presented in [1].

4. Techniques for Early Detection of Failures in EWDS

The primary goal of EWDS is to detect chiller failures as early as possible. Most failures are preceded by growing imbalances in the chiller components which initially manifest themselves through subtle deviations in operating parameters. Incipient failure detection is concerned with observing these deviations from nominal operation in the sensor measurements and relating them to specific failure modes. In order to detect incipient failures, there has to be an understanding of the nominal operational characteristics of a chiller. Chillers operate at a range of

ambient conditions and loads. There is an inherent variability in the nominal profile. For incipient failure detection, the algorithm has to be able to differentiate between deviations in the operating profile as a result of nominal changes and changes introduced by a failure mode.

The EWDS uses neural networks and thermodynamic models to define the nominal operational envelops under various load and ambient conditions. The incipient failure detection algorithms are based on experimental data and the underlying physics.

4.1 Neural Network based Algorithms

A neural network is a massively parallel, dynamic system of interconnected nodes that has the ability to learn in advance the patterns created by the sensor measurements under various normal and abnormal chiller operating conditions. The nodes in the neural network are organized into layers with full or random connections between successive layers. A typical multi-layer network has input, multiple hidden, and output layers. The connections between the nodes in the successive layers have weight values which are adjusted during learning, based on the input sensor patterns and sometimes the desired output values [2].

In theory, a neural network can implement any separating decision surface. The decision surfaces that separate normal and abnormal chiller operational patterns are unique to a given failure mode. In EWDS, a unique decision surface is constructed for each failure mode based on the failure specific pattern of the input sensor measurements.

A unique decision surface for a given failure mode requires a neural network with a unique architecture. The architecture of the neural network is defined by the input parameters, the number of hidden layers, the number of nodes in each layer, and the learning method and rule.

Neural network (NN) modules are utilized in EWDS to detect precursor characteristics of a specific failure mode. Selective input parameters are fed to an individually programmed neural net module to provide an output value indicative of the presence or absence of that failure mode. Neural net architecture for the three failure modes is presented below:

Expansion Valve NN Module utilizes 9 sensor Inputs (T1, T2, T4, {T3, T8 from circuit A, B}, and calculated superheat value for each circuit), 2 Outputs (Estimated expansion valve position in each circuit) and Back Propagation learning scheme with 36 hidden nodes in 1 hidden Layer.

Condenser Coils NN Module utilizes 12 Inputs (T4, T5, {T6, T7, T8 for circuit A, B} , 4 Control Statuses for fans), 1 Output (Cleanliness index) and Back Propagation learning scheme with 10 hidden nodes in 1 Layer.

Refrigerant Charge NN Module utilizes 8 Inputs (T4, T5, {T3, T6, T8 for circuit A, B}, 1 Output (Estimated Refrigerant Charge in pounds given a baseline nominal charge) and Self Organizing Map (SOM) with Back Propagation learning scheme with 16 hidden nodes in 1 hidden Layer, and 64 nodes for SOM arranged in an array of 4 by 16.

4.2 Thermodynamic Model based Failure Detection Algorithms

An accurate proprietary mathematical model of the dynamics of a chiller is used to develop failure detection algorithms based on the Kalman filter approach. The model uses nonlinear algebraic and differential equations to describe the flow, pressure, and temperature dynamics of the chiller, and is utilized to simulate failure modes such as motor overheating, and sensor bias and noise.

The main goal of EWDS is to quickly detect precursors to failures. The high dimensional nonlinear systems of equations utilized in the thermodynamic chiller model make it computationally intensive to be used directly in the failure detection algorithms. Therefore, a reduced order (ninth order), linear, state space model is created from the chiller thermodynamic model. The selection of states and input - output parameters is influenced by the failure modes that are to be detected using the Kalman filter approach. The states, input and output parameters for the state space model are listed below:

State Vector: **X** : X_1 - Refrigerant Enthalpy in Condenser, X_2 - Refrigerant Mass in Condenser, X_3 - Refrigerant Enthalpy in Evaporator, X_4 - Refrigerant Mass in Evaporator, X_5 - Air Enthalpy in Condenser,

X_6 - Refrigerant Saturated Temperature in Condenser, X_7 - Water Enthalpy in Evaporator, X_8 - Refrigerant Saturated Temperature in Evaporator, X_9 -Compressor Efficiency.

Input Vector: **U:** U_1 - Evaporator Entering Water Temperature, U_2 - Condenser Entering Air Temperature, U_3 - Expansion Valve Position.

Output Vector: **Y** : Y_1 - Evaporator Leaving Water Temperature, Y_2 - Condenser Leaving Air Temperature, Y_3 - Compressor Discharge Pressure, Y_4 - Compressor Suction Pressure, Y_5 - Compressor Suction Temperature, Y_6 - Super Heat Temperature, Y_7 - Refrigerant Entering Temperature in Condenser, Y_8 - Refrigerant Leaving Temperature in Condenser, Y_9 - Refrigerant Entering Temperature in Evaporator.

To create the reduced ninth order model, a pseudo random binary sequence (PRBS) [3] is used to drive the three input signals. Linear regression is performed on the chiller data obtained by exciting the exact model to derive matrix coefficients for the ninth order state space model.

The linear reduced order chiller model is represented as:

$$X(k+1) = A\,X(k) + BU(k) + W(k);$$

$$Y(k+1) = C\,X(k) + DU(k) + V(k),$$ where A, B, C, D are the transition matrices.

The Kalman filter equations are given as:

$$\underline{X}(k \mid k) = \underline{X}(k \mid k-1) + K(k)\,\{ y(k) - C\,\underline{X}(k \mid k-1)\},$$

$$\underline{X}(k+1 \mid k) = A\,\underline{X}(k \mid k) + B\,U(k).$$

The Kalman gain matrix K is computed as:

$$K(k) = P(k \mid k-1)\,C^T(k \mid k-1)\,[C(k \mid k-1)\,P(k \mid k-1)\,C^T(k \mid k-1) + R]^{-1}$$

where P(.|.) is the state estimate error covariance matrix.

Known inputs, U, and available sensor signals, Y, from the chiller are processed by the Kalman filter to generate optimal estimates of the chiller states, $\underline{X}$, and measurements, $\underline{Y}$. Subtracting $\underline{Y}$ from the actual measurements, Y forms the error signal vector, **e**, which is used to formulate a failure decision. The normal expected statistics of $e_i(k)$ are:

$$E\{ e_i(k) \} = 0, \quad E\{ e_i^2(k)\} = 1.$$

If a failure occurs, the system dynamics changes. The filter error signal no longer has the zero-mean, white statistics. A failure is detected by testing the error signal element values against a set of predefined thresholds. [4]

5. EWDS System Description

Figure 1 illustrates the functional architecture for the EWDS. The input data are selectively fed to the neural net and Kalman filter modules. Each neural net is programmed to observe precursor symptoms for a given failure mode. Each Kalman filter module is also tuned to a specific failure mode. Each neural net module and Kalman filter module process the input data and create an output value indicating either presence or absence of failure precursor. The outputs from the individual neural nets and Kalman filters are collected and processed in the information fusion module. The information fusion module utilizes the time of detection of precursor information, ambient conditions, and if-then heuristic rules to declare a single initiating failure mode followed by secondary and tertiary failure mode information.

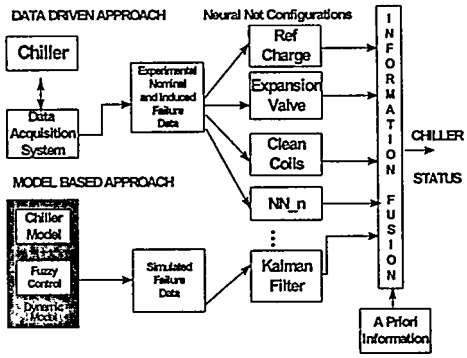


Figure 1: EWDS Functional Architecture

5.1 Failure Mode Experiments and Simulations

Nominal and failure data are acquired to develop the EWDS algorithms via experimentation and simulation. Nominal data are collected over a range of ambient conditions (seasonal variations), operational times (3 to 20 hours), sample rates (1 to 8 minutes), and varying load conditions. Approximately 14000 time instances of nominal data are collected.

In addition to the standard sensor set on the chiller, additional thermistors are installed to better characterize the thermodynamic characteristics of the chiller operation. These added sensors are: T4- Condenser Entering Air/Water Temperature, T5- Condenser Leaving Air/Water Temperature, T6- Compressor Discharge Temperature, T7- Refrigerant Condenser Leaving Temperature, T8 - Subcooled Refrigerant Temperature.

Failure modes that could be introduced in a controlled fashion such as refrigerant charging, fan failures, expansion device failures, etc. are tested. The chiller is subjected to the same load variations as during the normal operations. For each failure test, approximately 2000 to 8000 time instances of data are collected.

The experimental data are used for training the neural networks. The nominal and failure data are divided into a training and a test set with about 60% of the total data being used for training. The nominal data are used in the training of all the neural net modules. Each neural net module utilizes the specific failure data to make it sensitive to that failure mode.

Failures such as compressor motor overheating are difficult to introduce in a controlled fashion due to potential damage to a compressor. The detailed thermodynamic model is used to simulate that failure mode by introducing loss functions to account for motor overheating. The simulated failure data are added to the experimental data, and are used for training and testing EWDS algorithms. EWDS Kalman filter module is tuned and tested with simulated failure data.

5.2 EWDS Implementation

The EWDS exists as a compiled C code running in a 486 PC. The user interface is in visual basic. The system can operate in real time. Currently, it operates in an off-line batch mode where chiller data for every 24 hours is collected and analyzed by the system.

6. Results

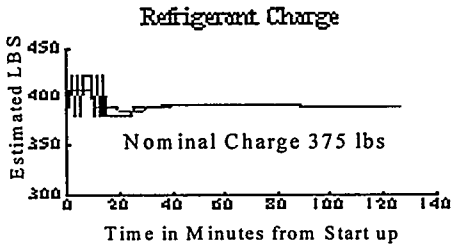
EWDS has been tested on real operational data, and has produced no false alarms. All the failures have been detected successfully and significantly earlier than the current alarms over the test set.

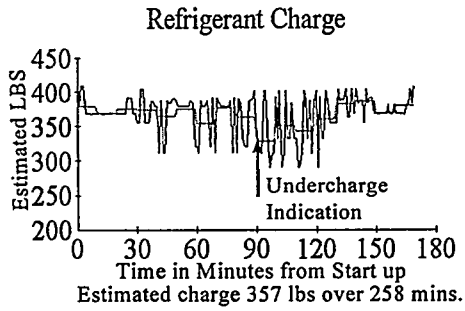
Dirty condenser coils

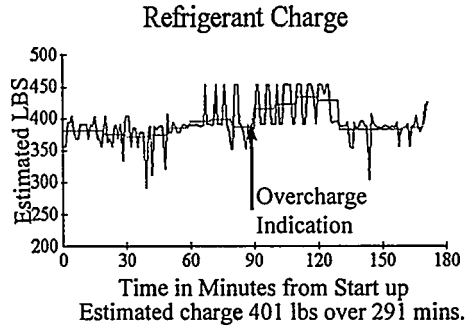
EWDS calculates state of cleanliness for the condenser coils using input sensor patterns and informs a mechanic that 'coils are clean', or 'coils may need cleaning' or 'coils NEED CLEANING NOW'. Currently, there is no alarm associated with this failure mode.

Refrigerant Leaks, undercharge, and overcharge

EWDS estimates refrigerant charge given a nominal charge as a basis. Refrigerant leaks as small as 7% of the total charge are detected. Current alarm does not shut the chiller till it has lost about 33% of its total charge. Thus, for slow leaks, EWDS can declare a failure within a day or so as opposed to few weeks with the current alarm. The top plot shows neural net output for nominal charge. For slight undercharge conditions, EWDS will issue a warning to the mechanic by providing an estimate of undercharge value as shown in the middle plot. The estimate is given at every sample point, and also averaged over a certain time period. The same type of information is provided for overcharge condition as shown in the bottom plot. At present, overcharge condition usually goes unnoticed by the mechanic.

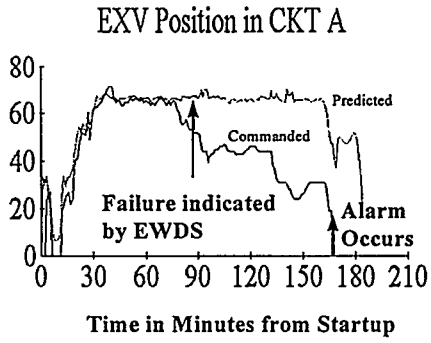






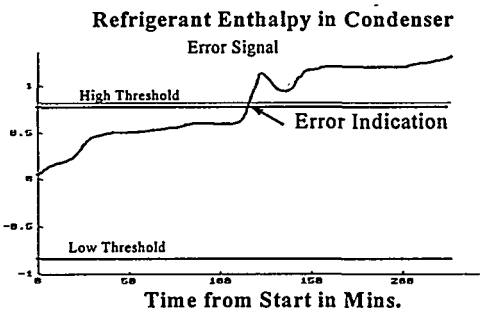
Stuck expansion valve

EWDS gives indication of a stuck expansion valve minutes (during start up and shut down stages) and hours (during constant load operation) before the chiller shuts down due to this failure as shown in the plot. It compares the controller command signal with the neural net predicted signal to issue a warning when the two signals diverge.

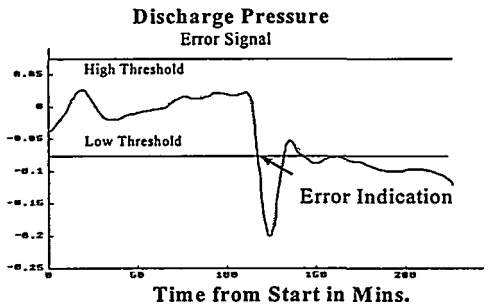


Motor overheating

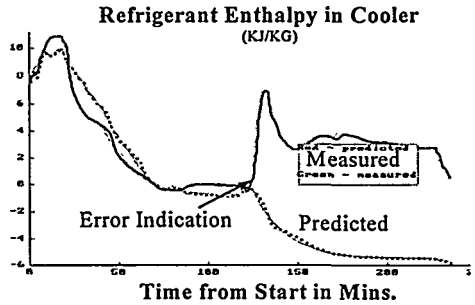
Motor overheating is detected by monitoring the thermodynamic states of the chiller, refrigerant enthalpy in cooler and condenser (bottom plot shows measured and model predicted values) as well as sensors such as discharge pressure.



Kalman filter tracks the predicted and measured signals for a predetermined set of state and output parameters for each failure mode. A failure is indicated if the error signals cross predefined thresholds as shown in the three plots.



7. Conclusions



This paper presents an integrated approach based on neural networks and thermodynamic model based techniques to infer impending failures from subtle changes in the normal operational characteristics of a chiller. The early warning of impending failure enables initiation of service before actual machine failure giving the perception of high reliability.

References

[1] Tulpule, S., et al., "Reciprocating Chiller Prognostics", UTRC Research Report, 1995.

[2] Kosko, B., "Neural Nets and Fuzzy Systems", Prentice-Hall, 1992.

[3] Duyar, A., et al., "Identification of Space Shuttle Main Engine Dynamics", ACC, June 1989.

[4] Gelb, A., "Applied Optimal Estimation", MIT Press, 1974.

PREDICTION, MONITORING AND CONTROL OF NOISE FROM A QUARRYING PROCESS

George J. Trmal, Reader and
Thomas Hill, Environmental Manager*
University of the West of England, Bristol,
*ARC Southern, Quarry Products, Frome,
England

ABSTRACT

Opening of a new quarry has a strong environmental effect and must be approved in advance. One of the main aspects under consideration is the generated noise. The application for planning permission must include a detailed assessment of noise levels. When permission is granted the noise must be monitored and kept within the approved limits. This requires a deep understanding of generation and propagation of noise and experience with methods of noise reduction. The paper describes a improved method of noise prediction, its verification by monitoring and some modifications of machinery designed to keep the noise levels within limits.

1 INTRODUCTION

In almost every human activity the benefits must be considered in conjunction with the environmental damage. Complying with environmental rules might appear to result in an increased cost but the aim is to find a compromise which maximises the benefits to the society. Opening of a new quarry has a strong environmental effect, invariably worries local residents and must be approved in advance by relevant authorities. One of the aspects which receives attention is the likely increase of environmental noise. An application for planning permission must therefore include detailed assessment of likely noise levels at the adjacent residential areas. Modification of quarry machinery, adjustment of working practices and noise screening is often required. This requires a deep understanding of generation and propagation of noise and knowledge and experience with methods of noise reduction.

The noise sources involved in quarrying are related to individual processes. The rock deposits must be first exposed by removing the top and sub-soil. The stripping is achieved using large diesel powered plant. Excavators remove the soil and dumper trucks carry it to site boundaries where bulldozers shape it into screening bunds. Since the operation is carried out without screening, higher noise levels for a limiting period are permitted. Winning the mineral from the deposits requires drilling and blasting. Blasting, which is of a very short duration, is usually not included in the analysis. The main sources of noise include pneumatically powered drills, dumper trucks, especially when climbing inclines, and crushers which reduce the size of the rock pieces. The noise levels in a close vicinity of impact crushers can reach 120 dB. The crushed material can now be handled by conveyors and taken to be graded to different sizes using vibrating screens. These screens act like giant rattles and can produce noise levels in excess of 100 dB.

Accurate prediction of noise levels at the adjacent residential properties is required in the planning stage. At this stage a standard, relatively simple and well tried method is preferred. However, the companies involved in quarrying prefer a more accurate predictions in order to avoid later complaints and associated problems during the quarry operation. More accurate prediction methods which include all most relevant factors can also help to identify the room for improvement in reducing the noise at source and during the propagation. Monitoring noise during the quarry operation helps to verify and improve the prediction method and improve the accuracy of the process.

2 PREDICTION

The theoretical base for predicting the noise levels is simple and logical and it is built into the standard assessment recommendations [2]. However, there are additional factors concerning air and ground absorbtion, screening and reflection which are more difficult to assess. They are determined using empirical formulae, tables and graphs based on experimental work.

2.1 Theoretical base

The noise is assumed to originate from a point source of a given power [W] and to propagate radially from the source. Noise intensity can be then defined as the noise power per unit area through which the noise propagates. As can be seen in fig.1 it is indirectly

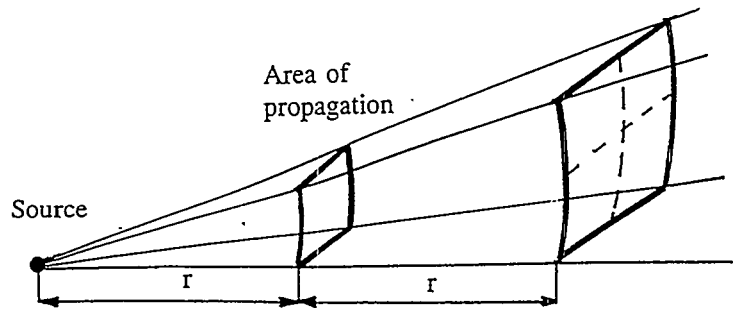


Fig.1. Noise source and noise intensity

proportional to the square distance from the source and is governed by the following formula:

Intensity [W/m^2] = Source_power [W] / area_of_propagation [m]

The intensity is usually expressed in dB according to the following formula:

$L_{Intensity\,[dB]} = 10 * \log (Intensity / Int_{Reference})$

In a 'free noise field and normal conditions the intensity level (in dB) and the sound pressure level are equal. This convenient relationship has been achieved by a suitable selection of the reference values. By combining the above equations it can be written for noise pressure level:

$L_{Pressure\,[dB]} = 10 * \log (A*W*q/(4*T *r^2))$

where: A = attenuation during propagation
W = source power
q = constant (angle of propagation)

or sometimes (by expressing the log function):

$L_{Pressure\,[dB]} = 10*\log(W) - 20 \log(r) + 10*\log(q/4*T) + Atten_{[dB]}$

2.2 Current practice

Department of Environment has released a Mineral Planning Guidance Note 11 (MPG 11) [3] which has four main aims: to recommend a noise prediction model, to recommend a method for setting noise limits, to recommend noise monitoring methods and to discuss possible control measures.

The noise limits are set as $L_{p\,eq}$ = 55 dB_A during the day and 42 dB_A at night. It recommends the prediction model set out in British Standard 5228 [2] which is based on the above theoretical considerations. However, the MPG 11 allows (and recommends) an inclusion of the effect of additional factors such as attenuation by acoustically soft ground, effect of screening barriers and source reflections.

2.3 Additional factors

The noise level is naturally attenuated by distance from the source. Additional attenuation can be caused by air absorbtion, propagation over acoustically soft ground and by screening barriers. The noise level can be also increased by reflection from behind the source.

2.3.1 Air absorbtion

The noise propagating in the air is a absorbed by the friction between the air particles. The losses depend on the temperature and humidity of the air but also on the frequency spectrum of the propagated noise. Higher frequencies are attenuated more then lower frequencies. The effect can be considered negligible for short distances but can be significant over distances of several hundred meters. The formulae and data are available but the process is too laborious for manual calculation.

2.3.2 Ground attenuation

Ground attenuation depends on ground covering (grass etc), form (flat or undulating) and the height of the propagation path. It also depends on the frequency of the propagating noise and the analysis must be made for individual octave bands of the noise spectrum. Empirical equations are available but errors can be made in assessing the acoustic hardness of the ground. Again, the whole process is too complex and laborious for manual calculation.

2.3.3 Screening barrier effect

A sound barrier can be any large object that blocks the line between the source and the receiver. The barrier mass and stiffness must be sufficient to prevent transmission. As shown in figure 2, the sound wave travels in a straight line to the top of the barrier then it 'bends' (diffracts) into the shadow zone. This diffraction depends on

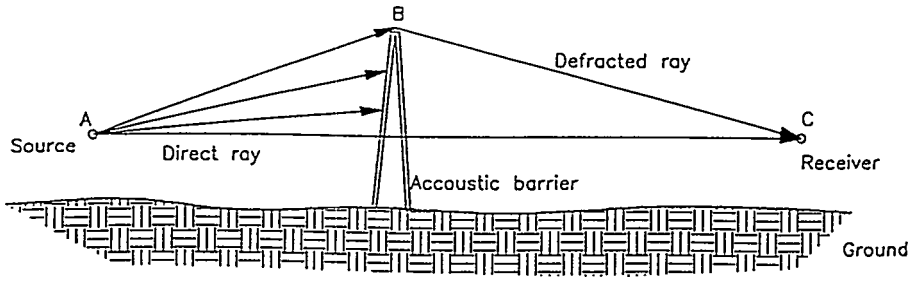


Fig.2. The barrier model

the distances between the barrier, source and receiver and the height of the barrier. It also depends on the noise frequency so the noise spectrum have to be considered. Again, the equations are available but the process is too complicated for manual calculation.

Other factors, such as a reflecting face behind the source, which increases the noise level reaching the receiver, can be considered in a similar way. Inclusion of all these factors enhances the accuracy of the prediction but is unsuitable for manual analysis. Incorporating the calculations into a computer program is an obvious solution.

3 COMPUTER PREDICTION

All described calculation methods has been incorporated in a computer program. The program offers options of predictions according to British Standard 5228, Mineral Planning Guidance Note 11 or a hybrid method which includes most relevant factors. The program does not model the real situation. Modelling must be done by the user and data entered in a simplified form into the program. Help and graphics are available. Fig.3 provides and example for entering the barrier data.

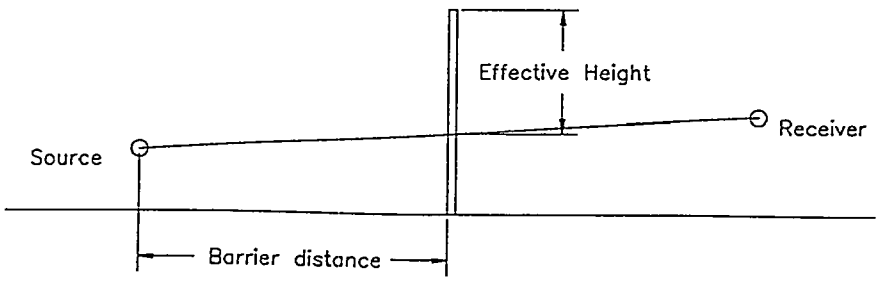


Fig.3. The barrier variables.

The program also incorporates specification of noise sources and evaluates the acoustic power from the results of noise measurement. This can be either in the form of measured value of acoustic pressure from a given distance or can include evaluation of noise spectrum in octave bands.

4 NOISE SOURCES

One of the most common pieces of quarry machinery is a drill rig. It consists of a pneumatic drill mounted on a caterpillar tracks and is driven by a adjacent compressor driven by a diesel engine. There are three main sources of noise on the drill: drill rotational air motor, air driven hydraulic pump and an air driven dust suppression system. Initial measurement identified the air exhausts as major contributors to the overall noise level. The air motor was fitted with an old absorptive silencer with the lining heavily contaminated with debris. The rig manufacturer suggested a new active type of silencer currently supplied with the rig. The new silencer was cheap to buy and easy to fit. Its use resulted in noise pressure level reduction of 5 dB which corresponds to a three fold reduction in noise power. However, the use of silencer resulted in fair amount of back pressure which reduced the efficiency of the drive.

The same silencer fitted to the exhaust of the air driven hydraulic pump reduced the noise pressure level from $93dB_A$ to $83dB_A$. The increase in back pressure reduced the drill rig productivity by estimated 25%. This was considered unacceptable in current situation. The dust extraction system is very sensitive to the back pressure. It was therefore decided to replace the old and contaminated silencer by a new one of the same type. This resulted in the reduction of the noise pressure level from 110 dB_A to 106.7 dB_A. These results clearly show the value of regular maintenance of quarry machinery.

The diesel engine driving the compressor is fully enclosed by thin sheet steel cover. It is suspected that it acts as a resonance board and lining the sheet metal with an absorptive material would bring beneficial results. However, only the engine exhaust was fitted with a new three stage silencer. This brought reduction in noise pressure level of 5 dB_A. The results indicate that the noise from quarry machinery can be significantly reduced by comparatively simple measures.

Frequency analysis of the emitted noise in octave bands was carried out in all cases. Results of this analysis are important when assessing the noise propagation because many noise attenuating mechanisms are frequency dependent.

5 CASE STUDY

An extension of of the existing quarry is required and the impact of the generated nouse is to be assessed in preparation for the application for the planning permission.

A map of the area layout is shown in fig.3. Both the existing quarry and the extension area are clearly shown in the figure. The critical noise path from the quarry extension to the nearest residential property corresponds to section A-A.

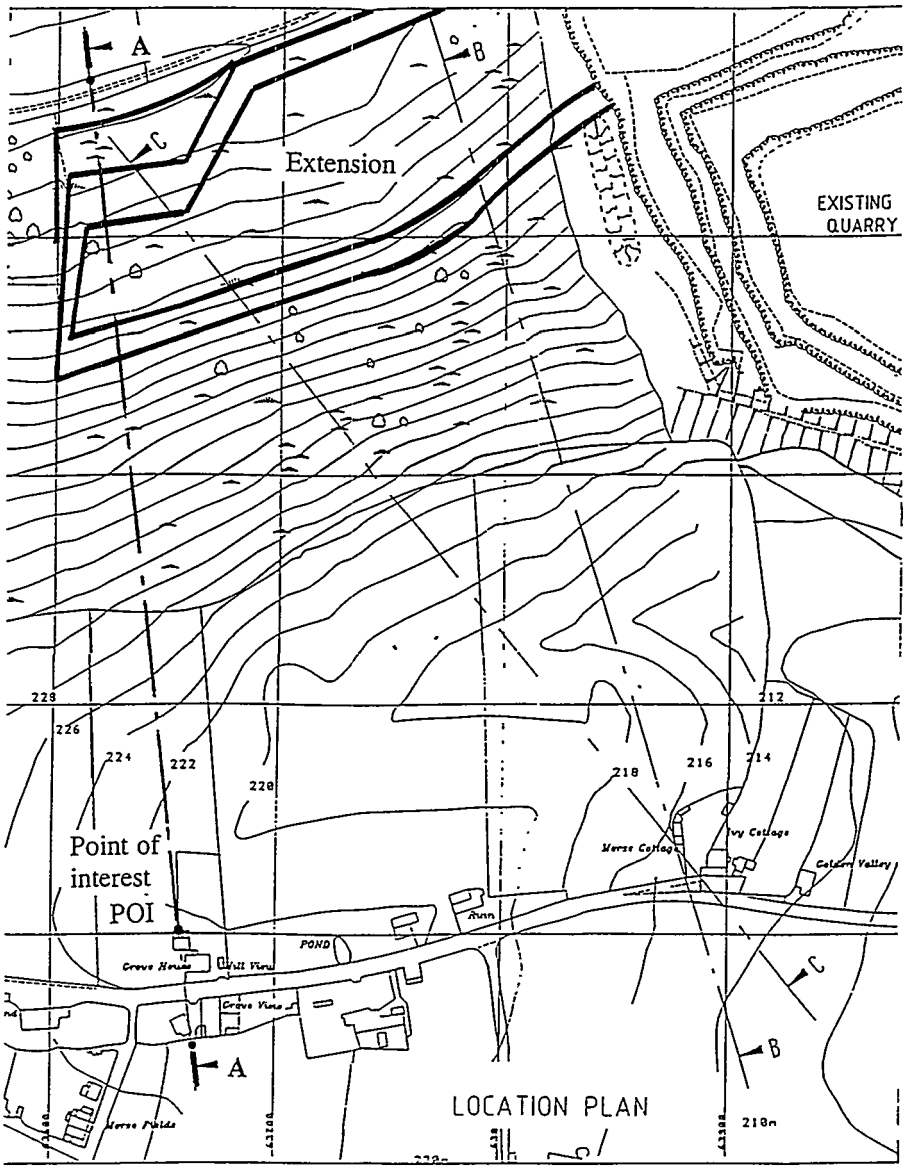


Fig.4. Layout of the proposed quarry extension and the adjacent area.

The profile of this section can be seen in fig.5. The highest level of noise is expected from a pneumatic drill when operating at the top of the extension in position 1. The acoustic power of the source and the

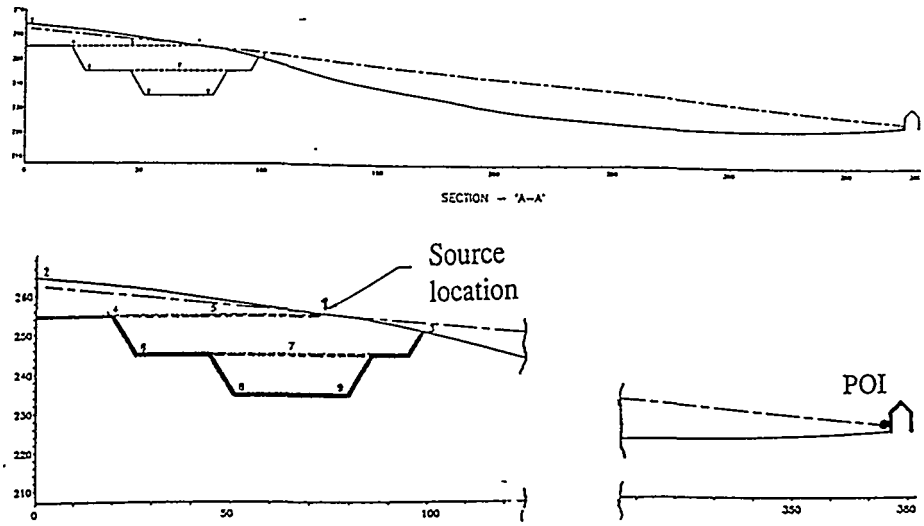


Fig.5. Considered section of the critical noise path

distance from the source to the receiver must be entered into the program so that the distance attenuation can be calculated. The noise path must be modelled for the entry into the prediction program which require the height of the noise source and the receiver, average height of the path above ground, the path length and the acoustic characteristic of the ground in the path, source and receiver area. Frequency characteristic of the source in octave bands is also required. This allows the program to calculate the attenuation of noise due to the ground effect. Characteristic dimensions of a screening barrier and of the reflective wall behind the source must be also entered if applicable.

The computer prediction print-out can be seen in fig.6. The upper part of the print-out is concerned with the entered specification. The noise source is a drill rig which was measured to emit 90.3 dB_A at a distance of 10 m. Noise spectrum in octave bands is also given. The ground specification is provided, there is no screening and no reflection from behind the source. The point of interest, where the noise level is predicted, is at a distance of 275 m. The predicted sound pressure level is 60.5 dB_A, significantly above the limit of 55 dB and reduction of noise at source or screening has to be considered. The table of detailed results shows that the air attenuation

NOISE-PRO
Environmental Noise Prediction Program - By Thomas Hill
Version - 1.0

THEORY PROCEDURE

LOCATION :- FILENAME :- dry-drl.in

Source 1 ..... Drill rig
Face distance 0 Face height 0 Face angle 0.0
Relative humidity 50.0%
Screening .....
Leq ..... 90.3(A)
Measured @ 10.0
Distance to POI.. 275.0
Source height 2.0 m, Source ground 5.0
Receiver height 2.0 m, Receiver ground 5.0
Path height 1.5 m, Path ground 5.0

Source 1 Drill rig
31.5 Hz..... 65.0
63 Hz..... 83.0
125 Hz..... 89.0
250 Hz..... 70.0
500 Hz..... 75.0
1 KHz..... 72.0
2 KHz..... 71.0
4 KHz..... 66.0
8 KHz..... 64.0

Source 1 ..... Drill rig P.O.I. Laeq ..... 60.5 dB(A)

	31.5	63	125	250	500	1 K	2 K	4 K	8 K
Source Laeq	65.0	83.0	89.0	70.0	75.0	72.0	71.0	66.0	54
Air loss (dB)	0.0	0.1	0.1	0.3	0.6	0.9	2.0	6.1	23.7
Dist loss (dB)	28.8	28.8	28.8	28.8	28.8	28.8	28.8	28.8	28.8
Barrier IL(dB)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Soft ground	0.0	0.2	0.9	1.7	2.4	3.2	3.9	4.7	5.4
P.O.I. Laeq	36.2	54.0	59.2	39.2	43.2	39.1	36.4	26.5	6.2

COMBINED LAEQ FOR .= 60.5 dB(A)

Fig.6. Computer print-out of the noise level prediction

is negligible for low frequencies and significant for in higher frequency octave bands. Attenuation due to the ground effect is of similar character.

6 DISCUSSION

Introducing the frequency spectrum of noise sources and frequency dependent noise propagation characteristic is a significant improvement on the existing recommendations. In addition ISO [4] recommends the use of frequency spectrum in assessing the 'noise annoyance' factor. Research has shown that the high frequency noise of the same loudness is more annoying then low frequency noise. This concept is not included in the prediction program but can be easily incorporated.

It would be also possible to create a database of noise data for standard quarry equipment so that competent predictions could be made without detailed measurement. However, it has been shown that poor state of machinery can result in a significant increase in noise emission and regular monitoring is required for maintaining low source levels as well as for verification of the method.

7 CONCLUSIONS

* The proposed method incorporates the current knowledge and can provide an accurate prediction of noise at the adjacent residential locations.
* Monitoring of noise provides an important experience and feedback for improvement of the method.
* The systematic approach of the prediction method contributes to the efficient design and running of the site.
* The information about the generation and propagation of noise allows optimisation of the whole process taking economic as well as environmental aspects into consideration.

8 REFERENCES

1. Hassal J.R, Zaveri K., Acoustic Noise Measurements, Bruel & Kjaer, 1979, ISBN 87-87355-30-2
2. Noise Control on Construction and Open Sites, British Standard 5228, Part 1, 1984
3. Mineral Planning Guidance Note 11, Department of Environment
4. Measurement of Noise with Respect to Community Response. ISO (International Organisation for Standardization) R 1996-1971,

CONDITION MONITORING AS A TOOL FOR CFC REPLACEMENT IN DOMESTIC REFRIGERATOR DESIGN IN DEVELOPING COUNTRIES

Derek M Cranvey, Divisional Manager
HEAT Technology Division
P.A.Hilton Limited
Stockbridge, England

ABSTRACT

The manufactures of domestic refrigerators in Africa, the Middle East and former Eastern Bloc countries have benefited from Internationally funded projects. The projects cover the conversion of the manufactured product from ozone depleting Chlorofluorocarbon (CFC) refrigerants, to approved environmentally friendly Hydrofluorocarbon (HFC) or Hydrocarbon (HC) refrigerants.

Condition Monitoring techniques have been applied to test facilities provided to a number of these manufacturers. These test facilities have been designed to allow cost effective performance verification of product redesign and provide test results in line with relevant ISO standards.

1. INTRODUCTION

Since the signing of the Montreal Protocol [1] in 1987 a series of World Bank and United Nations financed projects have been implemented to eradicate the use of ozone depleting products such as CFC's. These projects have been targeted at manufactures in developing industrial countries. One of the principal uses of CFC's is in the production of domestic refrigerators. CFC R12 is commonly used in the refrigeration circuit and a selection of other CFC products are used to produce the foam insulation for the refrigerator cabinet.

A replacement refrigerant R134a has been selected for the majority of these projects. R134a has a zero Ozone Depleting Potential or ODP meeting the requirements of the Montreal Protocol [1].

The base specification therefore for test facilities was that the converted unit should perform to a standard equal to or better than the CFC based product. The selected standard for the first of the products encountered by the HEAT Technology Division of P.A.Hilton Limited was ISO 5155 [2] covering the essential characteristics and test methods. Using the in-house developed DLS 101 Data Logging System a package of environmental control and test instrumentation was specified. The aim was to provide an 'engineer friendly' computer based condition monitoring system.

2. CONDITION MONITORING

A condition monitoring approach to the system was selected in that a number of parameters existed and each could be recorded for trend or set point data. These parameters could give both individual or combined component performance. Once combined this trend data could be highlighted to identify changes effecting the overall system performance down to component level.

3. REFRIGERANT SELECTION

Despite its selection benefits over CFC R12 the HFC R134a has a relativity high Global Warming Potential or GWP and may contribute to the so called 'Greenhouse effect'. The process of producing R134a is also more complex than R12 requiring more electrical energy. The combination of the GWP and the energy required to produce a refrigerant is classified as the Total Equivalent Warming Impact or TEWI. In some cases 'natural' refrigerants R290 Propane or R600a Isobutane both Hydrocarbons have been selected. These have a zero ODP and almost zero GWP. Hydrocarbon's are however volatile organic compounds or VOC's and have however been linked to environmental pollution problems such as 'low level smog'. They are also flammable under certain conditions posing handling risks. In the majority of target countries both R12 and R134a have to be imported using 'hard currency'. Hydrocarbons however may be available from existing industry or naturally occurring sources in the target country.

4. REFRIGERATOR REDESIGN

To use R134a in the refrigeration system a number of principal components have to be replaced or modified from the R12 system. The majority of the manufacturers concerned have based their designs on European or US designs either by licensed or unlicensed copy. The models range from 1950's to 1970's designs. The conversion from R12 to R134a therefore enabled manufacturers to not only convert but cosmetically update their products. Hydrocarbon systems however require slightly less work on conversion but consideration of the potential for ignition resulting from a flammable mixture of gas and air under pressure.

4.1 Component Level

For HFC R134a systems the following components will almost certainly require replacement when the product is converted from CFC R12.

- Compressor, primarily due to the use of synthetic oils that are miscible in R134a. Hermetically sealed R12 compressors have previously been charged with mineral oil that is unsuitable for R134a use. There are also design and material of construction issues related to the lubricant and refrigerant pressure temperature relationship.
- Refrigerant filter drier, R134a has a different molecular profile and the synthetic oils are hygroscopic.

The following items may also require modification or redesign due to the difference in the pressure temperature relationship between R12 and R134a.

- Expansion device, usually a capillary tube is employed and it dimensions will almost certainly require modification.
- Evaporator, may require modification for capacity changes and possibly cosmetic changes.
- Condenser, may require modification for capacity changes and possibly cosmetic changes.
- The refrigerator cabinet, structural changes to retain and support the replacement and modified components and possibly cosmetic changes.

Hydrocarbon systems can use the same principal components as R12 systems but the electrical components such as relays, thermostats and interior lights may require replacement with intrinsically or extrinsically safe devices. There is however a commercial issue that R12 components may, as CFC 'phase out' nears completion, become withdrawn by the main manufacturers. Labelling of the refrigerator to include the refrigerant used is also critical for maintenance purposes.

5. THE MONITORING SYSTEM

Domestic refrigerators have one prime function and that is to store items, normally foodstuffs, at a temperature that will extend the item's shelf life whilst retaining appearance where appropriate. By reference to the ISO standard ISO 5155 [2] and later standards ISO 7371 [3] and ISO 8187 [4] the principal measurable parameters are power consumption and temperature.

5.1 Sensor and Transducer Selection

The ISO standards however do not directly reflect the performance of the refrigeration system for environmental performance so the coefficient of performance or CoP was selected according to the principles detailed in R.J.Dossart [5]. The more practical CoSP or coefficient of specific performance from A.J.Gigiel [6] was also referenced as this included the total power consumption for the product and not just the refrigeration compressor.

The instrumentation selected was generally according to that used in D.M.Cranvey [7]. The ISO Standard [2][3][4] calls for temperature to be monitored through Tylose M type test packages and solid sensors. Tylose test packages are a gel like mixture wrapped in plastic that represents a food load. Tylose has the characteristics of lean beef. Type T class 1 thermocouples are used throughout both in probe and exposed tip inserted into the Type M (measurement) test packages. These are surrounded by other test packages that are not M type. A kilowatt hour transducer is used to give the power consumption figures for the ISO test. A number of other parameters are under the ISO standards required to be monitored. Measurement of the ambient humidity is taken with a capacitance based probe. A monitor of the period between the thermostat cycle is also required.

This was determined by a relay board included for this purpose and timing was taken through the DLS 101 system software from the host computer internal clock. To determine the power consumption of the unit under test for calculation of CoP figures an instantaneous power transducer is used. Pressure transducers are attached to the compressor suction and discharge lines together with exposed tip thermocouples. Thermocouples were also located on other principal component's inlet and outlets give both an individual assessment and overall performance figure for the unit under test. Current transducers are also employed to test the local power consumption of single electrical components. The transient state of the refrigerant at the entry to the capillary device and the instability of the low flow rates meant that a refrigerant flow meter of the turbine type could not be relied upon to give accurate results.

This package of instrumentation enabled the direct comparison of each system component replaced or modified to be monitored. This forms the instrumentation of the product that is designated HTS 700 Environmental Test Facility.

5.2 Software

Based on an low specification IBM PC AT 386 and above platform the DLS 101 software allows the display of the detailed parameters in numeric form for direct analysis. For analysis against recorded parameters a bar chart or multiple graphs can be displayed with a trend or set point line. The temperature, power, current and humidity readings are processed through the analogue inputs. The relay condition is displayed through the digital inputs in logic format, open or closed.

For final analysis of the collected data files they are converted in the DLS 101 software into a format compatible with the MS Excel spreadsheet software. Here a pre-configured worksheet allows the data relevant to the ISO Standard to be extracted and for CoP calculations to be undertaken. Each component performance can be displayed and recorded as well as a complete system energy balance.

6. TESTING ENVIRONMENT

Testing is undertaken in an insulated test chamber with an artifically controlled environment. The repeatability of the test condition was important so that a range of products could be tested under identical conditions and modifications could be performed during testing.

6.1 Test Chamber

To date all of these test chambers have been fabricated locally partly due to the economic constraints of the projects and partly due to local installation constraints. The final part of the HTS 700 is therefore a close control air conditioner with humidity control. By use of continuously rated components the HTS 700 can also be used for reliability testing and sample production line testing. On lower budget projects the close control air conditioner is substituted with a conventional air conditioner. Instead of providing a wide temperature test range this unit aims to maintain the environment close to the local ambient temperature.

7. PRODUCT ASSESSMENT

In most cases the end users planned to make a fully converted product and then use the test facility to 'fine tune' the redesign. Most of this work centres on the capillary length and evaporator layout. The ISO standards [2][4][5] would appear to be more addressed to the commercial requirement to display the performance of the refrigerator in an 'industry standard' format. They do not directly address the performance of the refrigeration cycle. The standards also call for the refrigerator on its full load test to be filled with Tylose samples. As the majority of domestic refrigerators will never be totally full of lean beef, especially in a developing or Hindu country, the value of this performance is questionable.

7.1 National Standards

Even when compared to other national standards as they were in P.K.Bansal and R.Kruger [8] in their preliminary comparisons of ISO, US, Australisian, Chinese and Japanese standards the ISO standards did not rate highly.

Again each countries test related to Tylose package loaded refrigerators that again would have questionable relevance to a developing country's consumers.

8. DISPLAYING RESULTS

The design of both the hardware and software components leaves the end users free to configure the data presentation to suit their own needs. A typical DLS 101 numeric display is shown in Figure 1. When inserted in to a user configured Excel spreadsheet the data can be analysed further and a typical test sheet is shown in Figure 2.

The end users can either choose to use this analysed data for each refrigerator or take the base line data to configure the set point or trend lines for immediate identification of performance variation.

9. CONCLUSION

The principal problem with comparing converted refrigerators with the R12 version is that these are not like for like comparisons and a number of main components may have been changed. Several models may be produced by a manufacturer and to reduce the conversion time cycle extrapolations of performance may be made. Here the value of individual unit and even component level performance is the key to developing products that have direct advantages over an unconverted version. The HTS 700 system (Figure 3) ensures this analysis is both easily undertaken and has a good standard of repeatability.

It is important that any testing is undertaken both on stand alone and comparative performance trials based on each selected refrigerant and total system suitability. The ISO standard has notable limitations when applied to 'phase out' projects. These are principally the lack of any specific data on the efficient performance of the refrigeration plant. The tests also fail to test for the conventional use of domestic refrigerators. A more appropriate test would include a typical load of foodstuffs and a door opening cycle as part of the full load test.

A standard of environmental performance covering both refrigerant type, cycle efficiency, power consumption and materials of construction would give the potential customer a better indication of product performance.

Figure 1

HEAT96

09:24:58 25-01-1997 Data Point 267

Channel	Label	Value	Unit	Channel	Label	Value	Uni
1	Discharge	73.230	°C	2	Liq Rec In	32.000	°C
3	Exp Val In	25.870	°C	4	Exp Val Out	-3.680	°C
5	Evap Outlet	17.860	°C	6	Suction	29.620	°C
7	Evap Air On	22.370	°C	8	Evap Air Off	5.350	°C
9	Cond Air On	22.870	°C	10	Cond Air Off	27.620	°C
16	1 Mains	214.500	Volts	17	Suction LP	0.916	Bar
18	DischargeHP	7.397	Bars	19	Humidity	22.160	%RH
28	Flow meter	16.724	g/sec	33	1 Current	2.560	Amp

Menu Next screen Previous page

Typical on-line test data numeric display

Figure 2

Microsoft Excel - DMC.CV

File Edit View Insert Format Tools Data Window Help

Times New Roman 12 100%

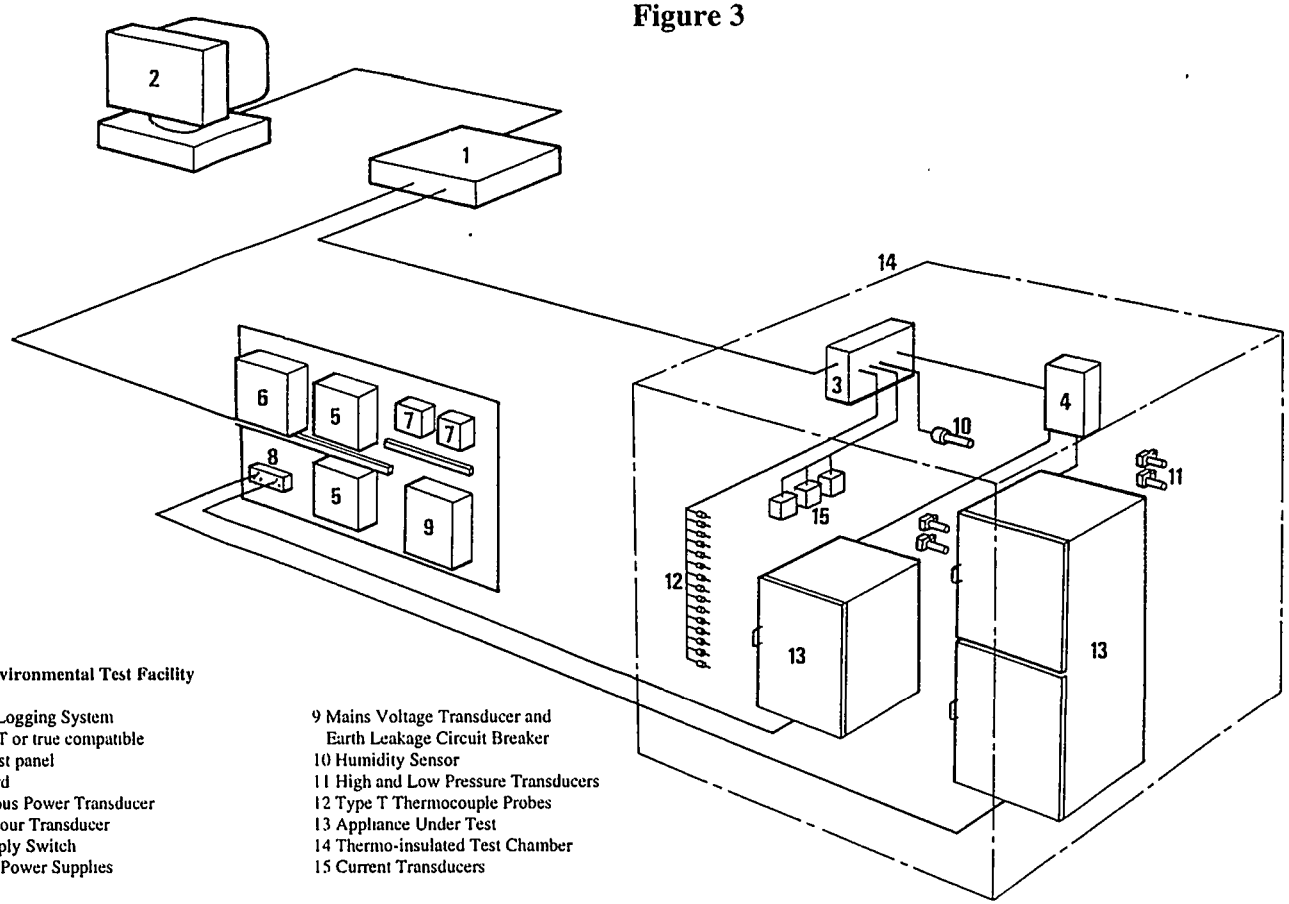
C7 50.61

	A	B	C	D	E	F	G	H
1	Channel No.	***	1	2	3	4	5	
2	Labels	***	Discharge	Liq*Rec*In	Exp*Val*In	Exp*Val*Out	Evap*Outlet	Suction
3	Units	***	°C	°C	°C	°C	°C	°C
4	Upper Lim.	***	20	20	20	-5	325	
5	Lower Lim.	***	-60	-60	-60	-60	-60	
6	Set Point	***	0	0	0	0	0	
7	07:58:36	28/01/97	50.61	21.90	14.89	-12.66	11.89	18
8	07:58:42	28/01/97	50.73	22.03	14.91	-12.65	11.90	18
9	07:58:48	28/01/97	50.86	22.04	14.91	-12.65	11.90	18
10	07:58:54	28/01/97	50.98	22.16	15.03	-12.51	11.90	18
11	07:59:00	28/01/97	50.98	22.16	15.03	-12.51	11.90	18
12	07:59:06	28/01/97	51.11	22.18	15.04	-12.50	11.90	18
13	07:59:12	28/01/97	51.23	22.18	15.04	-12.50	11.92	18
14	07:59:18	28/01/97	51.34	22.30	15.04	-12.50	11.92	18
15	07:59:24	28/01/97	51.47	22.19	15.06	-12.49	11.93	18

DMC

Typical off line test data analysis in MS Excel

Figure 3



HTS 700 Environmental Test Facility

1 DLS Data Logging System
2 IBM PC AT or true compatible
3 Remote Test panel
4 Relay Board
5 Instantaneous Power Transducer
6 Kilowatt Hour Transducer
7 Mains Supply Switch
8 Monitored Power Supplies
9 Mains Voltage Transducer and Earth Leakage Circuit Breaker
10 Humidity Sensor
11 High and Low Pressure Transducers
12 Type T Thermocouple Probes
13 Appliance Under Test
14 Thermo-insulated Test Chamber
15 Current Transducers

This would also make the product more acceptable to export markets and hence aid the manufacturing countries financial development

The use of conditioning monitoring techniques including computer based data collection has enabled a cost effective and local operator acceptable system to be delivered. It has assisted in the process of ensuring that future domestic refrigerator production will be both market acceptable, reliable and environmentally appropriate.

What ever test procedure is used and what ever refrigerant is selected for CFC 'phase out' it should be mentioned that domestic refrigerators properly constructed have a high reliability and very rarely leak refrigerant.

REFERENCES

1. Montreal Protocol on Substances that Deplete the Ozone Layer. *International Meeting*,, Montreal, Canada, September 1987
2. ISO 5155 -1983 (E)
   Household frozen food storage cabinets and food freezers.
   Essential characteristics and test methods.
3. ISO 7371 -1995 (E)
   Performance of household refrigeration appliances.
   Refrigerators with or without low temperature cabinets
4. ISO 8187-1991
   Household refrigerating appliances
   Refrigerator-Freezer Characteristics and test methods.
5. Principles of Refrigeration in SI Units
   R.J.Dossart, John Wiley & Sons, Inc,
   1981, USA, ISBN 0-471-05271-X
6. The Measurement of the performance of Refrigeration Plant
   A.J.Gigiel, Institute of Refrigeration, April 1984, London, UK.
7. Practical Experience of Condition Monitoring in Replacement Refrigeration Systems
   D.M.Cranvey. COMADEM 96 Conference, Sheffield, UK, 1996.
8. Test Standards for Household refrigerators and freezers
   I: preliminary comparisons
   P.K.Bansal and R.Kruger, Department of Mechanical Engineering University of Auckland,
   International Journal of Refrigeration, Vol. 18 No1, 1995.

SUPERVISORY DIAGNOSTIC STRUCTURE USING QUALITATIVE SIGNAL ABSTRACTION AND RULE-BASED INFERENTIAL REASONING: A TWO TANK ILLUSTRATIVE EXAMPLE

Joaquim Melendez-Frigola
Joan Colomer-Llinas
Josep Lluís de la Rosa-Esteva
Grup d' Enginyeria de Control i Sistemes Intel·ligents[+], IIiA-UdG
Girona, Espanya

Julio Waissman-Vilanova
Joseph Aguilar-Martin
Group Automatique Symbolique[+], LAAS-CNRS
Toulouse, France

+ LEA-SICA
Laboratori Europeu Associat-Sistemes Intel·ligents Control Avançat

ABSTRACT

A diagnostic structure based in signal abstraction and expert system is presented by means of an example. Abstraction is introduced as a tool to obtain qualitative significant information from signals and provide expert systems with this qualitative trends. A Matlab/Simulink based framework has been developed to support and assist this kind of realisations. The example here presented has been developed using this framework.

1 INTRODUCTION

Dynamic systems supervision is a field where Artificial Intelligence (AI) techniques give a completely new approach and widens considerably the possibilities to improve the quality of the control as

well as the security of industrial processes. First of all, the possibility of introducing knowledge about experimented types of behaviour of the system, as well from operators as from designers, becomes accessible by using linguistic descriptions that complete, or sometimes replace, theoretical equations. This approach is particularly useful if the mathematical models have not been confirmed by the industrial practice. A former work [Aguilar, 1993] shows that a supervisory structure must include a qualitative interpretation of information conveyed by sensors and actuators, in order to make it compatible with logical reasoning mechanisms.

A framework for Computer Aided Supervisory System Design (CASSD) has been developed with this aim. Supervisory capabilities, as Expert systems (ES) and abstraction tools (*abstractors*), have been incorporated into a popular CACSD framework, Matlab/Simulink, with this purpose. A diagnostic example has been developed using abstraction tools to interface measures and ES. In a first the raw numerical signals from sensors are abstracted, i.e. transformed into symbolic values, representing the tendency and the oscillation degree. In a second stage this information is loaded as facts for reasoning into an ES, the knowledge base (KB) of which contains the production rules proposed by plant or control experts.

In following paragraphs these ideas are expanded by means of a case study. Two coupled tanks are used as a example process. Abstraction of significant information from signals, qualitative *tendencies* and *oscillation degree*, is exemplified in paragraph two. The ES is introduced in paragraph three. Section four introduces the framework resulting of integrating AI tools into Matlab/Simulink and section five shows the application in a laboratory plant.

2 ABSTRACTORS

To take benefit of engineers experience, some qualitative information of process evolution is often needed. The algorithms or tools used with this aim are called abstractors or abstraction tools. They are process dependent and they must be designed as an interface between acquired measures and process engineer knowledge base. As result of these tools, qualitative information is obtained to represent trends of signals (tendencies, oscillation degrees, alarms, degree of transient

states, ...) needed in supervision, specially in fault detection and diagnosis. According to process dynamics several techniques could be used with this aim. They embrace topics as filtering, segmentation, pattern matching, polynomial regression, triangular representation, wavelet transform, histograms and so on. All of them could be used as analysis tools to obtain significant information from process signals at several abstraction degrees. In the same way these tools could be considered as numeric to qualitative interfaces.

Some of these tools have been implemented, to interface a KB described by process engineers and measures, as a set of blocks into Matlab/Simulink. One of these algorithms, used in the case study, is described in detail. The goal is to supply qualitative description of *oscillation degree* and *tendencies* of measures on line to the ES. Qualitative data obtained, as well as quantitative data, will be used later by the ES in a diagnostic task. "Fig. 1" shows an operation scheme of these algorithms [Colomer, 1996]. The final result are two output : the *qualitative tendency* and *qualitative oscillation degree.*

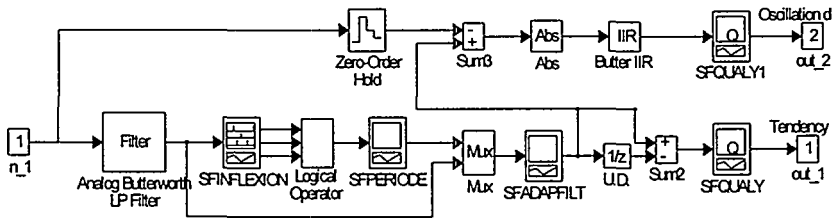


Fig. 1 Simulink implementation of an abstractor to deduce *qualitative tendency* and *qualitative oscillation degree.*

2.1 Qualitative tendency

This useful feature could be extracted from signals in several ways. One of them is explained below. Validity of this method is restricted to signals with damped oscillation. Five steps have been differentiated:

Step 1: Noise filtering. Low-pass filtering of measure is done to remove noise added to the signal and no desired variations. This is needed to guarantee next steps validity.
Step 2: Extrema detection and period estimation. Relative extrema (maxima, minima, and inflection points) of signals are detected to estimate oscillation period of analysed signal. This is obtained every

time a new extrema is detected and it is used to change the cut-off frequency of an adaptive filter (Step 3).

Step 3: Adaptive filtering. The signal is again introduced in a low-pass filter, the cut-off frequency of which is adjusted at a given sample time, depending on natural frequencies determined in step 2. This action smoothes the signal, its oscillations, and overshoots.

Step 4: Gradient. After the filtering process performed in *step 4*, the slope of the signal is obtained subtracting the previous value and the present value.

Step 5: Qualitative tendency. Finally, tendency is obtained splitting the calculated slope in five qualitative levels as follows:

-2	-1	0	1	2
greatly descending	***descending***	***keeps on level***	***rising***	***greatly rising***

This classification is done according to the maximum and minimum values of the gradient obtained in the previous step and takes advantage of the knowledge of the process engineer.

2.2 Qualitative oscillation

Another important feature that could be extracted from the process is the oscillation degree. This qualitative oscillation could be obtained by subtracting the filtered signal (*step 3* of qualitative tendency) from the original. Three steps perform this task :

- *Step 6 : Absolute oscillation signal.* The filtered signal (obtained in *step 3*) is subtracted from the original signal and the absolute value is calculated.
- *Step 7 : Filtering.* The result of previous step is filtered. The resulting signal is smoother than the original one. This signal with less rough changes is easier to qualify.
- *Step 8: Qualitative oscillation degree.* The obtained oscillation degree is qualified into three levels:

0	1	2
No oscillation	***Small oscillation***	***Large oscillation***

Using Simulink blocks, this abstractor has been implemented as is depicted in Fig. 1.

3 THE EXPERT SYSTEM.

CEES (C++ Embedded Expert System, [De la Rosa, 1993]) is a shell developed under object oriented languages that allows definitions of qualitative models and co-operative ESs as independent objects with agent capabilities. This is a rule based ES with fuzzy reasoning and forward chaining inference engine. This utility was firstly designed as stand alone Windows based application. Some examples were developed interfacing this shell by means of dynamic data exchange (DDE) with other Windows based applications , [Sabat, 1996]. The actual version (CEES 2.0) has been adapted as Simulink blocks to design and test knowledge based expert reasoning structures.

Compilation is needed any time a new ES is developed. This allows to use C/C++ syntax in the ES rules. Example of a simple rule is represented in "Fig. 2".

```
Rule 210
Threshold 0.5 TraceHere Yes
Description "valve 1 CLOSED"
  IfAny (PF) PF->result == situation_2
    And PF->cf
    And dif_level->greater(dif_level->low)
    Then deduce(DIAGNOSTIC, v1_close)
    Otherwise deduce(DIAGNOSTIC, v1_open)
  EndIfAny
EndRule
```

Fig. 2 Syntaxis of CEES Rules.

4 INTEGRATION IN A CASSD FRAMEWORK.

AI technologies are added into Matlab/Simulink in order to get a CASSD by using the object oriented paradigm. The aim is to have a framework where control and supervisory systems could be developed without the necessity of using external applications. Matlab/Simulink has been chosen as a platform to develop and to integrate those tools, because it offers openness and many useful services that can help the user to develop (ToolBoxes), and to implant (Real-Time Workshop) supervisory systems. Thus, ES added to this CACSD can take benefit of the existent analysis and representation tools to be supplied with information that matches antecedents used by process engineers when describing knowledge as rules. This is a tool to assist design and developing of supervisory systems. Details of

how these tools have been added into Simulink could be found in [Meléndez, 1996].

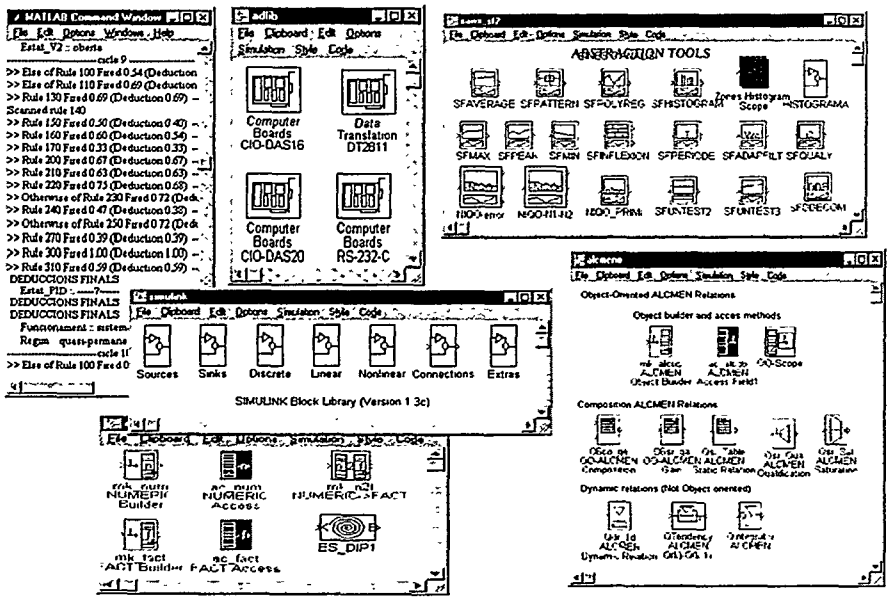


Fig. 3 CASSD framework based on Simulink

Fig. 3 shows the complete set of tools developed to assist supervisory systems design into Simulink. Special interest is focused in the ES because this is a C++ object oriented tool. Then integration of CEES into Simulink implies to adequate data transfer between blocks and memory management to deal with objects. Input and output of ES block are not Simulink standard data since it deal with objects. The data type used by CEES into Simulink are *facts*. This is an object oriented structure that encapsulates fields and methods used by the ES. This capabilites have been implemented as a 'S-function' blocks (standard block provided by MathWorks) containing a C++ based '*.mex' files.

5 APPLICATION TO A REAL PROCESS

The tools described in previous paragraphs have been tested with a Laboratory plant. A model of this plant has been obtained and the diagnostic structure described before has been developed using this. Once the design is finished the same CASSD framework is used to test and validate the design because it is provided by additional tools

(See Simulink blocks in Fig. 3) to access hardware (plug in boards, serial port, ...).

5.1 Laboratory Plant description

The laboratory plant where fault situations are injected is composed by two coupled tanks with two pipes doing the connection between them as depicted in Fig. 4. The goal is to control the level in the second tank by pumping the fluid to the first tank while the liquid flows through valve 2 (V2).

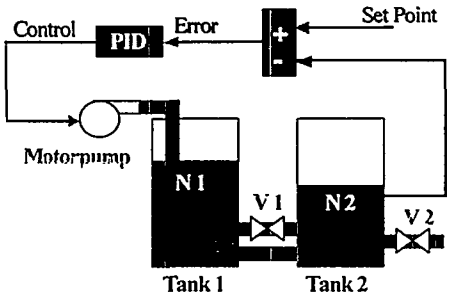


Fig. 4 Laboratory plant.

The control signal of this process is the pump voltage. Normal operation mode is defined as follows: valves (V1 and V2) are open and the process is correctly controlled (good PID tuning and pump working). Then, level 2 tracks the set point. Levels of both tanks and set point are the only available measures. Four possible malfunction situations to detect are typified. Simultaneous situations are not considered.

Situation 1	Situation 2	Situation 3	Situation 4
Valve 1 closed	***Valve 2 closed***	***Pump crashed***	***Bad regulation***

When failures are introduced then the expert diagnostic system has to be able to detect and identify them. Therefore, the goal of the supervisory system is to track the process and detect situations that incite failures or process malfunctions, as well as to know when the process works in the normal operating conditions. In order to detect the failures, the abstractors described in section 2 have been applied to the most representative signals such as the *level difference* (level tank 2 - level tank 1) and the *error* (set point - level tank 2).

5.2 Implementation into the framework

The implementation of a Supervisory structure in Simulink is convenient because its design and test could be done simultaneously in the same CASSD framework. This allows partial validations of the

structure when working with complex systems. Simulink capabilities allow to group set of blocks into a new one simplifying the appearance and allowing knowledge encapsulation. At the same time connectivity between tools is quickly solved by tracing lines between inputs and outputs of blocks.

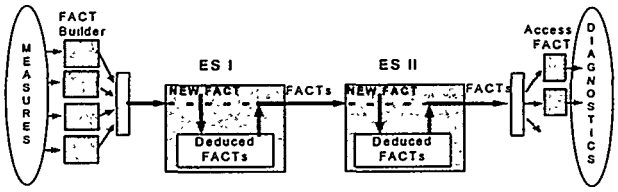


Fig. 5 Modular implementation of ES in two levels. Output *facts* come from deductions or directly from the input.

In the case study presented in this work, the rule base is organised in two levels (Fig. 5). The first one decides about system regime (transient, permanent) and deduces situations of abnormal operation mode assigning a certainty value to these new facts. Deductions are performed by reasoning about qualitative information abstracted from signals as is explained in section two. The second level yields final deductions about behaviour of four physical elements in the real process. This second level uses both deductions of first level and measures (numerical and abstracted information from these measures) of process. The division in two levels is performed after observing the KB given by the process engineer and structuring it as two sets of rules.

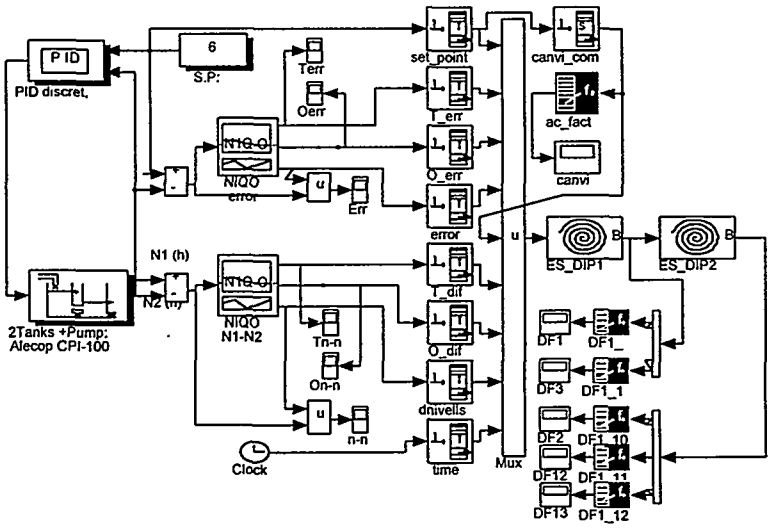
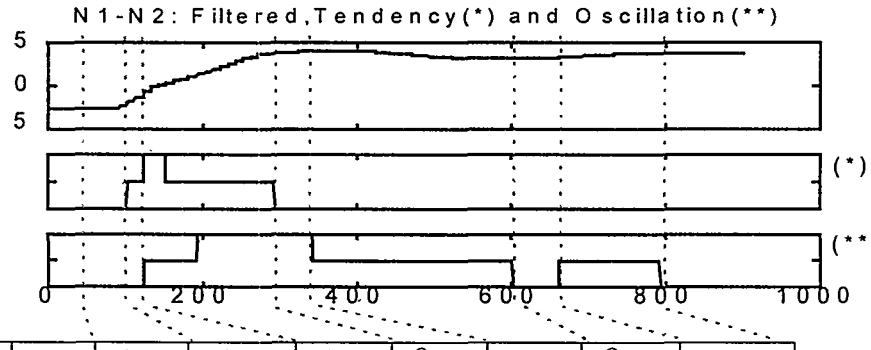


Fig. 6 Simulink implementation of the ES based diagnostic structure.

As Fig. 6 shows, ES is supplied with facts. This facts are built from abstracted data using specialised blocks that force engineers to declare knowledge about measures. In the same way output of blocks are accessible by Simulink blocks (scopes) by means of specialised blocks that allows access to desired fields of facts.

5.3 Results

Evolution of difference of levels could be observed in Fig. 7 when "Valve 1" is closed. Taking benefit of abstracted information diagnostics are given in a permissible delay of time.



Regi.	perma nent	tran sient	tran sient	tran sient	Q-per manent	perma nent	Q-per manent	perma nent
V1	open	--	closed	--	closed	closed	closed	closed
V2	open	open	open	open	open	open	open	open
Pump	good	good	good	good	good	good	good	good
Contr	good	good	good	good	good	good	good	good

Fig. 7 Diagnostic evolution, when valve 1 is closed.

The elapsed time for diagnostics could be reduced but this implies that uncertainty of these deductions increases.

6 CONCLUSIONS

The presented framework is endowed with important capabilities oriented to develop KBs. The example of this work shows how this framework is used for diagnostics but it also could be used in more general supervisory tasks. This implies the dreamt supervision purpose of changing controller parameters and controllers themselves by using algorithms developed under CACSD tools.

The example of this work uses this framework for developing an ES for diagnostic taking advantage of an easy-to-use CASSD environment with a graphical interface (Matlab/Simulink). It is used to obtain significative information to provide an ESs with *facts* forcing engineers to declare knowledge about signals as facts and about process behaviour as rules.

7 ACKNOWLEDGEMENTS

This research is supported by the CICYT project TAP96-1114-C03-03, "*Plataformas Integradas de CAD de Supervisión y Metodologías*", of the Spanish Government.

8 REFERENCES

[Aguilar, 1993] Aguilar-Martin J.,"Knowledge-Based Real_time supervision of dynamic processes: Basic principles and methods. In Spyros G. Tzafestas Ed., Applied Control. Current trends and modern methodologies, Marcel-Decker, Inc., NY, pp. 875-910, 1993.

[De la Rosa, 1993] De la Rosa Esteva, J. Ll.,"Heuristics for Cooperation of Expert Systems", Application to Process Control, PhD thesis, UAB (Spain), 1993.

[Colomer, 1996] Colomer, J, Pous, C., Melendez, J, De la Rosa, J,Ll., Aguilar J., "Abstracting Qualitative information for Process Supervision, in Proceedings of CACSD96, IEEE Int. Symp., pp. 410-415, Dearborn (MI), 1996.

[Melendez, 1995] Melendez J., Colomer J., De la Rosa J.Ll., "Linking G2 and Matlab for developing an expert system diagnostic system. In Supervision de Processus à l'aide du système expert G2TM, pp. 91-110, Ed. Hermès, 1995.

[Melendez, 1996] Melendez J., J. Colomer, J. Ll De la Rosa, J. Aguilar and J. Vehi, "Embedding Objects into Matlab/Simulink for Process Supervision", IEEE International Symposium on CACSD, pp.20-25, Dearborn, MI., 1996.

[Sabat, 1996], Sàbat A., Melendez J., Colomer J., De la Rosa J.L., "Integración de herramientas para el diagnóstico de procesos" in 'Seminario Anual de Automàtica y Electrònica Industrial", pp. 11-13 de Sept., Zaragoza (Spain), 1996.

PC-CONMON: THE DEVELOPMENT OF A PC BASED ARTIFICIAL INTELLIGENCE SYSTEM FOR THE DIAGNOSIS AND PROGNOSIS OF MACHINE CONDITION COMBINING ACOUSTIC EMISSION AND ACCELERATION MONITORING

Ander Hurtado & Aitor Aurrecoechea, Applications Director
Predictive Maintenance department, COINPASA
Guecho, Spain

ABSTRACT

Monitoring the condition of rotating machinery is a matter of capital importance for nowadays factories. One of the most popular methods for monitoring is the use of vibration analysis techniques, which have quite a good performance in most of the machine types. However, this performance may be improved combining simultaneously some other monitoring technique, such as acoustic emission monitoring. The use of these two methods of monitoring has made it possible to develop an expert system for the diagnosis of a wide range of rotating machinery. This system can detect a series of defects in certain types of machines, and has the capability to be improved increasing the number of machines and defects detected.

1. CONDITION MONITORING IN ROTATING MACHINERY

Nowadays, the trend of all industrial sectors is to increase the production rates in all their factories, in order to make them as competitive as possible. The increment in production is achieved through two different ways:

- Increasing machinery speed to manufacture more products in the same amount of time (for example, in paper factories).
- Decreasing the times in which machinery is stopped for planned or unplanned maintenance.

These two methods have a common aspect: the necessity of knowing machinery condition in order to make it work at the top of its capacity. This involves that machinery failures must be prevented, or at least predicted within a reasonable lapse of time, due to the high costs that such a failure would have [1].

There are several methods for determining machine condition depending on the type of machine being monitoring. These methods cover a wide range of techniques, such as vibration analysis, thermography, ultrasound analysis, acoustic emission, lube oil analysis,...

As rotating machinery is a very important group of machinery in almost all industrial facilities, most of the efforts have been employed to develop techniques for monitoring them. The most popular technique used for monitoring the condition of rotating machinery is vibration analysis.

Rotating machinery condition monitoring using vibration is widely used and systems are available commercially at different levels of sophistication. On the other hand, condition monitoring using acoustic emission is at a relatively immature stage, but it's a very promising field.

1.1 Vibration based monitoring techniques

This kind of techniques has been widely investigated and employed for detecting a variety of machine defects. The defects that vibration analysis has proven most successful to find are those related with low frequency vibration , such as:

- Imbalance
- Misalignment
- Machine looseness

Using vibration techniques it is also possible to make a successful control of the condition of gearboxes. The success of this monitoring technique has much to do with a correct monitoring plan in order to establish the trend of vibration in time [2].

There are some international standards which cover the permissible vibration of machinery, such as ISO 10816-1, but they only control the overall vibration level of the machine, being useless to detect individual defects. Besides, there are some kind of machine faults that has a very low level overall vibration, but which can lead to devastating machine failures.

1.2 Acoustic emission related techniques

As stated earlier in this paper, this kind of techniques are at a relatively immature development stage. They have been successfully employed in detecting materials failures, such as cracks in pipes or deposits, or fatigue analysis, but their usage in monitoring rotating machinery condition is quite an unexplored field [3].

One of the objectives of PC-CONMON project was to develop an extensive range of acoustic emission tests in a number of machines in order to check its capability of monitoring rotating machinery.

The failures in which acoustic emission analysis techniques has proven to be more successful have been those related with high frequency signals. These failures include rolling elements bearings defects, and fluid disturbance problems. Acoustic emission has made it possible to detect rolling elements faults at an earlier stage than vibration techniques, which has lead to a more adequate planning of machine overhauls [4].

2. EXPERT SYSTEM FOR MACHINE DIAGNOSIS

Once proven that the simultaneous combination of vibration and acoustic emission techniques could lead to a very powerful machine monitoring tool, the next objective of the PC-CONMON project was to develop an expert system for rotating machinery monitoring. The main objectives of the project were the following:

- The combination of acoustic emission and acceleration signals taking the advantages of each to provide a superior diagnostic capability.
- The development of sophisticated signal processing on acoustic and acceleration signals from machinery in order to optimize the diagnostic capability.
- The development of software using artificial intelligence techniques to diagnose faults and identify deterioration of different machines.

- The use of versatile acquisition and monitoring systems based on PCs and data acquisition boards.

2.1 Hardware requirements

Taking into account these requirements, it was decided to use an industrial laptop as the running platform. It would have two data acquisition boards, with different features. One of them was used to take raw acoustic emission, which implied very high sampling rates. The other one was used to take treated (RMS filtered) acoustic emission and acceleration signals. This hardware structure can be seen in figure 1.

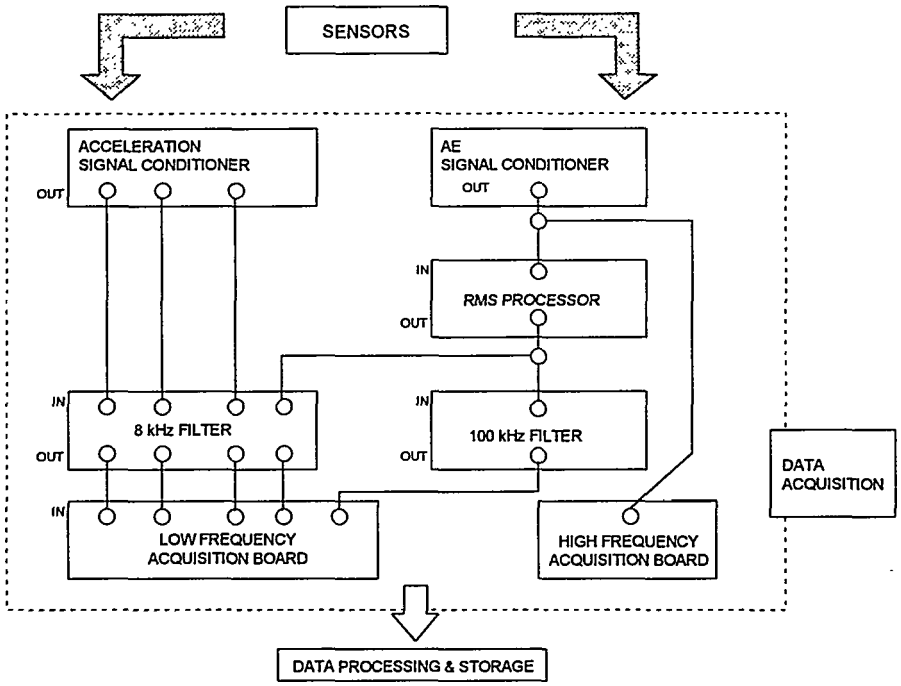


Fig. 1: Hardware structure

The hardware requirements of the system were:

- An industrial laptop computer compatible with a 486 processor or higher, and with a minimum of 8 Mb RAM.
- A fast acquisition board for acoustic emission.
- A slow acquisition board for acceleration.

- Acoustic emission and acceleration sensors.

While developing the software, a card of each type were used, but it was stated that the system would be able to manage any other card, with the only modification of including a new driver. The same requirement was made for the transducers (acoustic emission and accelerometers) employed. The software would be able to support most of the existing transducers of both types.

2.2 Software requirements

When an expert system was decided to be developed, several design requirements were made in order to obtain a simple but powerful software. The main requirements were the following:

- The software should have the capability of grow in knowledge, being an open, flexible system, not a closed one.
- The software should be as simple as possible, in order to make a commercially attractive product, affordable for many factories.
- The software should be user friendly, being accessible for non specialized personnel.
- The hardware needed to run the software should be standard and not expensive.

The software should be able to manage a great number of machine in an arranged database. The structure of this database should be based on a three levels hierarchy:

- Plant (or factory)
- Area
- Machine

This will allow the user to efficiently arrange the machines to monitor, and will permit the existence of a very large number of machines.

2.3 Expert system structure

One of the first things that were taken into account at the start of the project were the huge number of types of machinery that exists in today's factory. For this reason, three main types were chosen to start with:

- Pumps
- Fans

- Gearboxes

The reason for choosing these three types of machines was that, in most of the factories, more than a half of all the machines belongs to one of these types. Besides, these machines are quite simple in their operation, so the rules for treating acoustic emission could be implemented in an easier way.

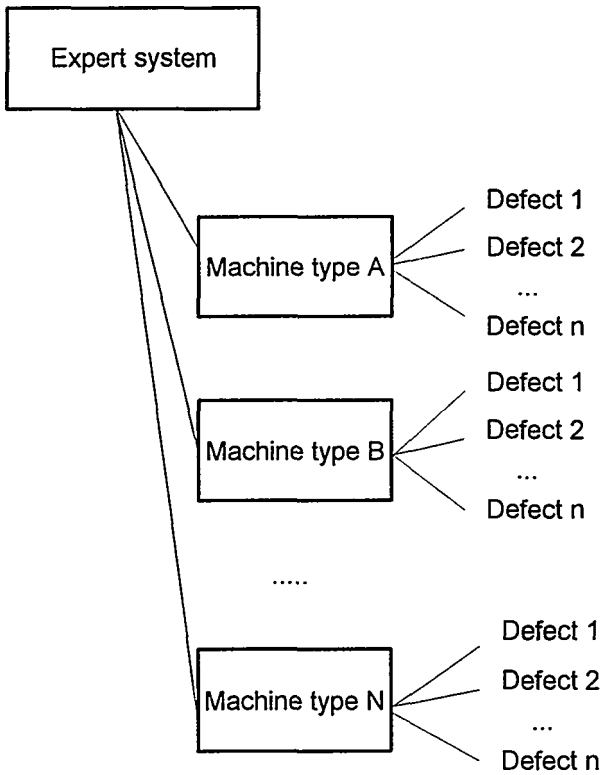


Fig. 2: System structure

To make the system as complete as possible, it was decided to make an expert system with a two level modular structure, as it can be seen in figure 2.. This modular structure would make possible the addition of new machine types in later stages of development, with a minimum of modifications in the already existing software.

The two levels of modular structure are the following:

- Machine level
- Defect level

This kind of structure will make possible to include new types of machines and routines for their monitoring. Besides, once a machine is incorporated to the system, the defects that the system could detect would not be limited. New detected defects can be added when the algorithms are developed, or even the algorithms could be changed in order to detect the defects more effectively.

As stated above, one of the requirements of the system was the ease of use, as it was developed bearing in mind that it could be used by non-specialized operators. This ease of use was achieved through two features:

- The use of a standard Windows interface.
- The 'black box' approach to the results of the expert system.

The first feature implies that a person with a minimum knowledge of Windows operating system, will be able to use the system without having to study a lot of complex information.

The second feature is a specially important one. It implies that the operator doesn't need to know how the system works. The system will automately diagnose the condition of the machine.

There are three levels of diagnosis programmed in the software:

- Automatic report
- Advanced report
- Manual report

So the operator will only have to collect vibration and acoustic emission data, and the expert system will immediately give a figure that represents the state of each of the defects detected on the machine (automatic report). If this figure shows that the machine condition is getting worse, actions could be performed in order to prevent the failure.

This approach has an evident problem: there can be some conditions in which the user would like to know more about how the failure is progressing in order to make a decision. The expert system can give him this information: it can tell the user which symptoms have indicated the state of the machine (advanced report).

The user can even go further and analyze manually the data collected, thus being able to detect some other failure that the expert system cannot detect.

This approach makes possible the achievement of two important features:

- There is a important saving of time due to the automated report of the expert system.
- A specialist that would like to study the condition of the machine, would have all the data available for his task.

As it was stated in the hardware requirements, the selected computer to run the system was an industrial laptop. This implies that the system is a portable one, so it can be moved anywhere in a factory to take measures in machines. The use of the automated report of the expert system gives the chance to study the machine condition on-site, and to take maintenance decisions immediately.

REFERENCES

1. Machine failure analysis and troubleshooting. Heinz P. Bloch, Fred K. Geitner. Gulf Publishing Company, 1986

2. Introduction to machine vibration. Glenn D. White. DLI Engineering Corp., 1995

3. Rotating machinery diagnosis with acoustic emission techniques. I. Sato, T. Yoneyama, S. Sasaki, T. Sazuki. Journal of acoustic emission, Volume 2, number 1/2. 1983

4. Bearing fault detection in the acoustic emission frequency range. M. S. Tavakoli. Condition monitoring, July 1991

DEVELOPMENT OF AN AUTOMATIC FAULT DIAGNOSIS SYSTEM FOR A PLATING PROCESS USING AN EXPERT SYSTEM

S. J. Wilcox[1], P. Bennett[1] , K. J. Baker[1] and J. Rees[2].

[1]Department of Mechanical and Manufacturing Engineering, University of Glamorgan, Pontypridd, UK
[2]The Royal Mint, Llantrisant, UK

ABSTRACT

This paper describes the development an on-line automatic fault diagnosis system for a plating process in an Expert System (ES). The system links with the plant control computer (PLC) and monitors it for plant faults via the use of timers. When an alarm is activated the ES interrogates the PLC to determine the item of plant that has a problem and then uses deductive logic to isolate the most likely cause of the problem. This all happens transparently to the operator of the plant and within a suitably short period of time. The system has consistently diagnosed a number of faults on-line that previously had taken plant personnel variable amounts of time to analyse.

1. INTRODUCTION

The Royal Mint produces all of the coinage for the United Kingdom, which accounts for approximately one third of production. Total production being in the order of 4 Billion coins and blanks per year. A large range of collector coins, medals and precious metal coins are also produced. The processes involved in producing coins utilise large scale plant with distributed control systems. Currently the electroplating of blanks is essential in many cases to reduce material costs. The plating process is potentially very expensive and efficient uninterrupted plant operation is essential. However, the complexity of the plant can make some faults very difficult to trace. Unscheduled plant downtime can cause work in progress to be unusable or contain quality defects and is hence very expensive.

The trend towards automated manufacture has brought benefits in terms of efficiency, especially with regard to machine utilisation, and quality of products. A vital factor in maintaining standards of performance is the minimisation of machine downtime in the event of breakdowns. This imposes great demands on maintenance staff to diagnose and rectify faults rapidly. A paradox that has been created through the general reliability of modern equipment is that maintenance staff are sometimes less familiar with problems than previously. There has thus been a rapid rise in the use of computer assistance to locate faults on complex plant and equipment.

Work in this area has highlighted the application of Expert Systems to similar problems. For example, Ready [1] outlined the use of Artificial Intelligence (AI) and particularly Expert Systems for condition monitoring and fault diagnosis and concluded that some of the available ES development packages are very effective at these tasks. Poon [2] described a low cost automatic condition monitoring system based on the temperature, vibration and power consumption of an electrical machine measures of which were fed into an Expert System for analysis. Conroy, Black and O'Hare [3] described a fault diagnosis and condition monitoring ES that was developed for a combined heat and power unit which determined component faults and poor performance of the system via the monitoring of up to thirty two parameters.

Other relevant work is that of Cooper [4] who described the development of a real time Expert System at a naval dockyard to determine faults, recommend solutions, and provide training (by simulating faults) for 'shore based' electricity supplies to docked ships, Collier *et al.* [5] who detailed the requirements for a real time knowledge based system, and Leith *et al.* [6] who discussed the use of expert systems to monitor the "health" of machines.

Diagnosis is one of the largest application domains for expert systems, but many systems concentrate on the diagnosis and neglect the user interface. Ghallab [7] put forward the view that the usefulness and value of the system as a whole is only good as the ability of the user to understand and implement the recommendations made.

The following problems were identified with the plating plant at the Royal Mint prior to the development of the system. The plant had

been operating reliably with only occasional faults, but due to their infrequency and the complexity of the plant and its associated control system these faults were often very difficult to locate. In addition, it was found that the staff experience and knowledge of the plant was lost as staff relocated resulting in the repeated independent solution of faults. These facts highlighted the need to develop a system that could :

- reliably fault diagnose the plant.
- store the plant and process knowledge of the operations and maintenance staff.
- be effectively used by the operations and maintenance staff.
- be continually updated and refined.

In order to alleviate some of these problems it was decided to develop a real-time fault diagnosis system based around an Expert System (ES) that would utilise as much of the currently available plant information as possible thereby minimising costs as no additional instrumentation would be needed. In order to facilitate this it was necessary to link the ES with the plant control system as this was the only way in which to access information from the plant. This had the disadvantage of not being independent of the PLC but had the advantage of being able to deduce additional knowledge from small pieces of otherwise unrelated information.

2. THE PLANT

The plating plant has thirty stations, around half of which are the plating stations with the remainder being used for pre and post plating rinses such as acid etches and is controlled by an Allen Bradley Programmable Logic Controller (PLC). Coin blanks that are to be plated are fed into perforated plastic barrels with each barrel possessing two dangling electrodes which pass through bearings at either end. Two special cranes called 'transporters' move the flight bars between baths. Figure 1 shows a view along the plating line and Figure 2 is a close up view of a transporter.

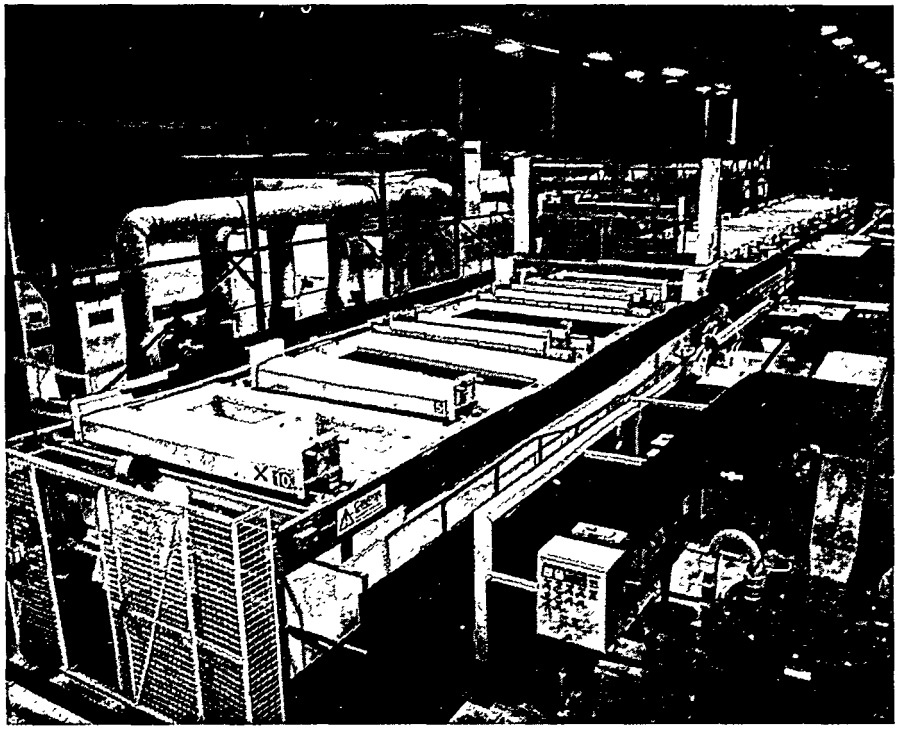


Figure 1: View of the plating line.

Automatic movement of a single transporter is the normal mode of operation for the production plant with two transporters being used when the cycle time is short. The transporters transfer the flight bars between different stations in the process. Eighteen flight bars are used in the process. The process consists of the following steps :

1. filling the barrels with the required load of coin blanks
2. cleaning and then etching of the blanks.
3. plating the blanks.
4. cleaning the blanks.
5. unloading the barrels and drying the blanks.

The plating of the blanks contained in the barrels of a flight bar takes place in one station.

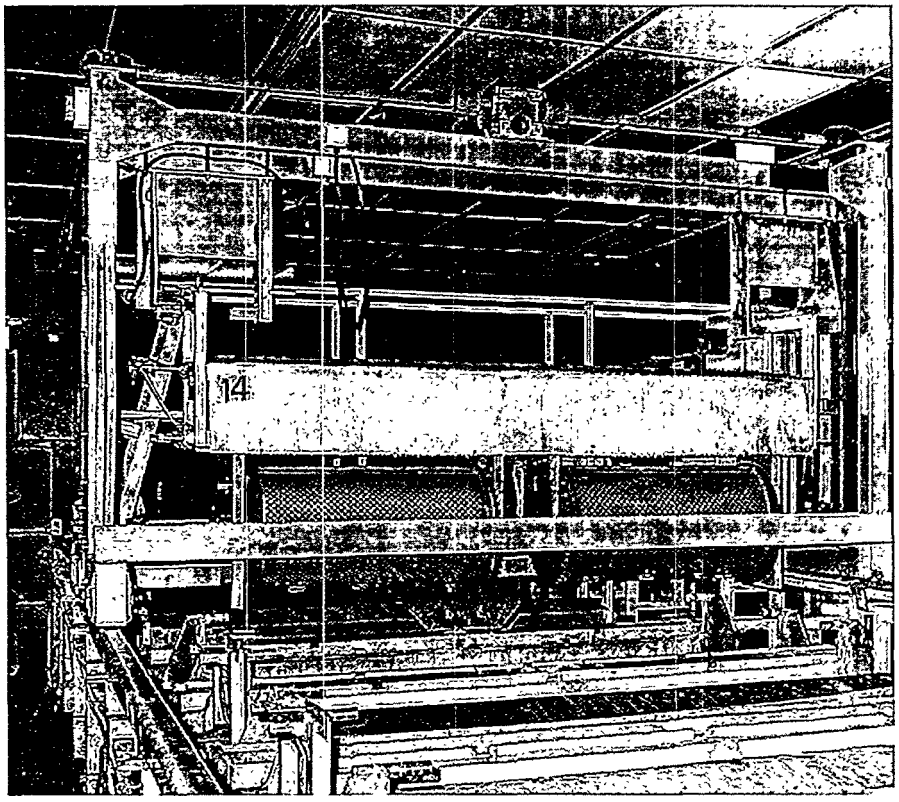


Figure 2: Close up view of a Transporter.

3. DEVELOPMENT AND TESTING OF THE FAULT DIAGNOSIS SYSTEM

It was found that the most effective method for developing the system involved encapsulating in the Expert System the knowledge of maintenance and operations staff as well as logical procedures based on plant drawings and manuals. In order to accomplish this the expert system knowledge base is arranged into 3 areas that perform the diagnosis, hold information relating to the current plant status and to the current fault. In order to structure information in each area the knowledge is arranged in a tree like structure.

The diagnostic tree allows the Expert System logically to reason about the process and has been developed using information obtained from:-

1. the plant operators and maintenance staff?
2. the plant control computer?
3. electrical and mechanical drawings of the plant?
4. knowledge of likely faults?

The diagnostic tree is activated by the Expert System if a fault condition is determined from the plant control computer in the following way:-

1. Has the plant been disabled for some reason.
2. Is automatic operation of the plant inhibited.
3. Have any of the system timer alarms activated.

The ordering is such that a fault in a specific area of the plant will be checked first thus achieving a faster diagnosis.

The alarms are activated by a comparison between the actual time and the expected times to complete an operation, with a disparity causing the alarm. As a fault diagnosis is undertaken the required information is accessed from the plant control system as and when required with the information obtained dictating the subsequent path taken and hence the next node (in the tree structure) to be examined.

On completion of a diagnosis the appropriate node in the fault tree is activated along with the relevant information being placed in the user interface. A typical screen shot is shown in Figure 3.

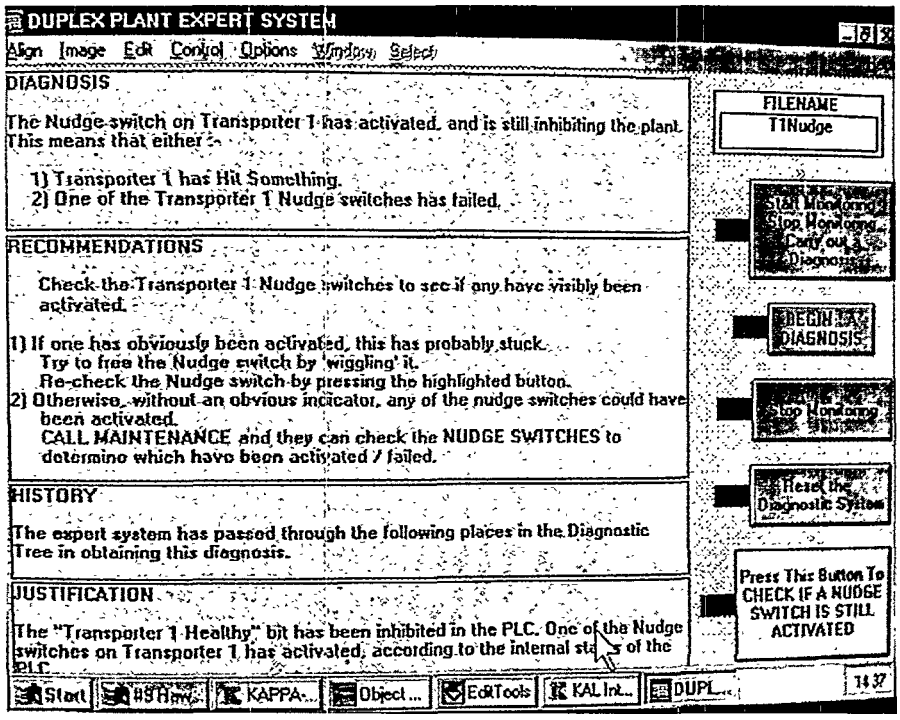


Figure 3 : Screen shot of the ES user interface in use.

The top two panes display the diagnosed fault and make recommendations about how to recover the plant from the current situation and restart it. All of this is performed automatically with no user intervention and within a few seconds of the fault causing a problem. The next pane displays the justification for the current diagnosis in 'plain English' along with plant status information and the reasoning used to make the diagnosis. The final pane displays the list of the nodes in the knowledge base that have been activated in obtaining the current diagnosis which would aid refinement of the system when necessary. Finally, there is a series of buttons that allow the operator to control manually the fault diagnosis system, although it is normally operating continuously.

3.1 An Example Diagnosis

At the extremities of each transporter are safety switches which inhibit the plant if activated. This is necessary to stop the transporters if

they collide with anything. There are a several reasons why the switch could be activated, these being:-

1. activation due to an obstruction.
2. activation due to a sensor failure.
3. spurious activation after overhaul for maintenance.

In order for the Expert System to diagnose this fault condition it first waits for the plant control system to 'time out', then it collects information and determines that safety switch is activated by detecting internal registers that indicate that the plant is inhibited. The Expert System then examines the information in the plc that can cause this condition and then concludes that the safety switches have been activated.

4. CONCLUSIONS AND FUTURE WORK

The Expert System can diagnose a number of faults occurring in the plating plant and has been in operation for several months. Typical faults include sensor failures, overheating motors, inverters tripping out and relays failing. In operation the user is presented automatically with an interface which allows access to the knowledge base, the Expert System interfacing directly with the PLC to automate the fault diagnosis.

Updating the fault diagnosis ES is facilitated by logging the cause of the fault and the corrective action necessary in an existing plant logbook.

5. ACKNOWLEDGEMENTS

The authors are most grateful for assistance from the Royal Mint in conducting the investigations reported in this paper.

REFERENCES

[1] Ready, C.M., 'Expert Systems for Condition Monitoring, Fault Diagnosis and Control', *IEE Colloquium on Condition Monitoring for Fault Diagnosis, Organised by Professional Group*

C9 (Control Techniques and Applications), Friday 18th October 1991, Digest No : 1991/156.

[2] Poon, H.L., 'A Knowledge Based Condition Monitoring System for Electrical Machines', *Computers in Industry*, 1991, **16**, 159-167.

[3] Conroy, G.V., Black, W.J. and O'Hare, G.M.P., 'Fault Diagnosis and On-Condition Maintenance of a Combined Heat and Power Unit', *IEE Colloquium on Expert Systems for Fault Diagnosis in Engineering Applications*, 1989 Digest No 1989/64.

[4] Gerdes, R.K. 'An Expert System for Production Fault Diagnosis', *IEE Colloquium on Expert System for Fault Diagnosis in Engineering Applications,* 4/1-4/2, 1989.

[5] Ghallab, S. and Ef Stathiou, J. 'An Expert System Generating Explanation for Fault Diagnosis in a Water Rig System', *IEE Colloquium on Expert System for Fault Diagnosis in Engineering Applications*, 9/1-9/3, 1989.

[6] Cooper, N., 'A Real-Time Knowledge Based Recovery Advisor for Electrical Power Distribution Networks', *IEE Colloquium on Knowledge Based Advisory Systems for Monitoring and Supervisory Control*, 1987.

[7] Brailsford, J.R., Cross, A.D. and Raven, P.F. 'The Switching Schedule Production Assistant - A Knowledge Based System for the Electrical Power Distribution Engineer', *IEE Colloquium on Knowledge Based Advisory Systems for Monitoring and Supervisory Control*, 1987.

[8] Leith, D., Dears, N.D. and Thomson, W.T., 'Application of Expert Systems in the Condition Monitoring of Electrical Machines', *IEE Colloquium on Knowledge Based Advisory Systems for Monitoring and Supervisory Control*, 1987.

[9] Collier, D. and Neal, P.W. 'The Detection of Abnormal Operating Conditions in a Pulverising Fuel Mill Using an On-line Expert System', *IEE Colloquium on Knowledge Based Advisory Systems for Monitoring and Supervisory Control*, 1987.

ABDUCTIVE DIAGNOSTIC PROCEDURE BASED ON AN AND/OR/NOT GRAPH FOR EXPECTED BEHAVIOUR : APPLICATION TO A GAS TURBINE*

Pilar FUSTER
Dept Mat. & Infor. UIB, Palma de Mallorca, Spain

Antoni LIGEZA
Fac. of El. Engng., Aut. & Elcs. Technical University of Cracow, Poland

Joseph AGUILAR MARTIN
LEA-SICA: LAAS-CNRS Toulouse, France & IIiA-UdG Girona, Spain

* Work partially supported by project TAP96-1114-C03-02

ABSTRACT

In this paper we adress the diagnostic of technical systems by considering first of all their expected normal behaviour. The formalism that we use is called AND/OR/NOT causal graphs. This approach can be regarded as an extension of abductive models. A symptom can be *True*, *False* or *Unknown*, it can be represented by a propositional formula.

The nodes are partitionned into: *Manifestations*, that characterize each type of misbehaviour, *Elementary diagnosis*, that correspond to single faulty states and *Intermediate*, these last are observable or not partial symptoms. An illustration about the on-line diagnosis of a gas turbine shows also some of the steps leading to a causal graph modelisation in which this diagnosis methodology has been applied.

Through an equational physical model, we pretend to fill the gap between the engineer's and designer's knowledge representation as well as the operators "shallow" knowledge. In this application most of the knowledge can be inferred from the thermodynamical equations, even, as it is the most frequent, there are no precise measurements.

1 INTRODUCTION

In this paper we adress the diagnostic of technical systems by considering first of all their expected normal behaviour, completed by the knowledge about other expected behaviours in the case of single faults in their parts.

An illustration about the on-line diagnosis of a gas turbine shows also some of the steps leading to a causal graph modelisation in which this diagnosis methodology has been applied.

The formalism that we use is called AND/OR/NOT causal graphs. The influences between possible states, (situations), of the parts, from elementary diagnosis to manifestation of failures are represented in a connected graph where the links can be positive or negative (NOT) [Fuster & Ligeza 1996]. Those links are combined at each node by means of the logical connectives AND and OR. This approach can be regarded as an extension of abductive models in [Peng & Reggia 1990].

Each node of an AND/OR/NOT causal graph is interpreted as a symptom, it can be *True*, *False* or *Unknown*, it can be represented by a propositional formula. The concept of symptom is not an obvious one and the definition used here has been discussed in [Fuster 1996]. The nodes are partitionned into: *Manifestations*, that characterize each type of misbehaviour, *Elementary diagnosis*, that correspond to single faulty states and *Intermediate*, these last are observable or not partial symptoms.

A Diagnostic problem is stated as solving a given propositional formula in the framework of the corresponding causal graph. The propositional formula has, in the present methodology, a fixed form, it must contain three parts:

- the *manifestation*, i.e. a statement about the truth or falsity of one of the manifestation nodes
- the *observation*, i.e. as many statements as possible about known state of nodes.
- the *diagnostic*.

The thermodynamical modelisation explains the expected normal behaviour, by localizing the functionnalities and expressing the causal

relations so as to be able finally to reproduce qualitativelly the sense of variations and influences.

Fist of all a block/flow description of mechanical turbo-machines, including compressors and motors, is made by using the KBPD (Knowledge-Based Process Description) methodology [Aguilar 1993]. This gives a frame to be filled with relations and equations. In the case of the gas turbine simplified thermodynamical equations give rise to a complete simulation graph obtained by a convenient ordering of the equations.

2 CAUSAL BEHAVIOUR

2.1 Physical relational model

The basics for explaining faulty behaviour in complex systems considers the knowledge about causal relations between symptoms. We show the hierarchy of the categories of symptoms as they have been described in [Fuster & Ligeza 1995] :

F - Failures		symptoms
C - Components U- Control actions Z - Operat. cond.	Initial causes	
V - Interm. symptoms		

F, C, U, Z, and V are the corresponding sets of symptoms. Let D denote the set of elementary diagnoses or items such that their failure produces the ossurrence of a single alarm. We establish a one to one mapping between $C \cup U \cup Z$ and D so that we assign to each component, control action or operational condition some proposition $d \in D$, So we have $|D| = |C| + |U| + |Z|$ and following [Fuster & Ligeza 1995] we admit negative causal influences. So an elementary diagnosis $d \in D$ means a failure of a given component, control action or external signal, and $\neg d$ denotes a correct occurrence. A possible diagnosis is a sequence $\{x_1, x_2, \ldots, x_i, \ldots, x_n\}$ where x_i may be d_i or $\neg d_i$.

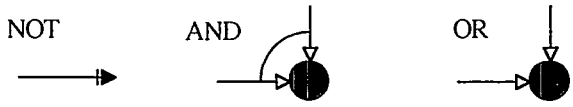
The thermodynamical modelisation explains the expected normal behaviour, by locating the functionalities and expressing the causal relations so as to be able finally to reproduce qualitativelly the sense of variations and influences.

Through an equational physical model, we pretend to fill the gap between the engineer's and designer's knowledge representation as well as the operators "shallow" knowledge. In this application most of the knowledge can be inferred from the thermodynamical equations, even, as it is the most frequent, there are no precise measurements.

Then a block/flow description of mechanical turbo-machines, including compressors and motors, is made by using the KBPD (Knowledge-Based Process Description) methodology [Aguilar 1993].

2.2 AND/OR/NOT causal graph

Once the causal graph has been obtained by the equations and completed by other known links with the elementary components that can produce failures, the straightforfard application of the abductive AND/OR/NOT causal graph methodology yields the diagnosis for the observed misbehaviours of the plant. In the complete graph the nodes correspond to logical clauses involving a single connective as follows:



3 KNOWLEDGE BASED TURBINE MODEL

3.1 Mechanical turbo-machines

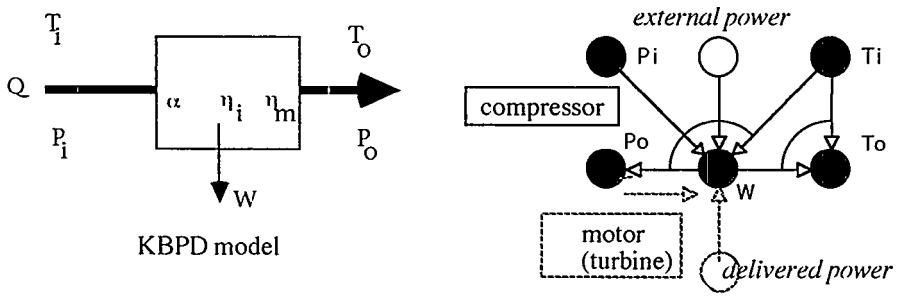


figure 1

A flow of gas Q across a machine produces a mechanical power W, a difference of temperature $\Delta T = T_o - T_i$ and a pressure ratio $\rho = P_o/P_i$.

Temperature transformation : $T_i - T_o = \alpha W$ so $T_o = T_i - \alpha W$, for a receiver (compressor) W<0 and for a generator (turbine) W>0.

Pressure ratio : $\rho = (1 \pm (\alpha W)/(\eta_i . T_i))^{\gamma/\gamma-1}$, $\gamma = C_p/C_v$, ($\gamma \approx 1.4$),

The energy heat conversion coefficient α , is related to the mass flow per unit time, Q and the mechanical efficiency η_m. The energy is then $W = (\eta_i/\alpha) . T_i(1 - \rho^{\gamma-1/\gamma})$ where η_i is the isentropic efficiency.

The energy power is $W = \beta . N^2$ where β is a constant depending of the geometry of the rotation.

In figure 1 the KBPD diagramm is shown together with the logical graph connecting the qualitative events (here OK/notOK) related to the variables.

3.2 Combustion

The combustion process is taken as a power addition, that increases the temperature of the flow of gas. The energy transformation coefficient φ per unit mass of combustible is given by $\varphi = \eta_f . H_f/(Q.C_p)$, and the temperature transformation is $T_o = T_i + \varphi$. Q_f, Q_f is the flow of fuel, and η_f the efficiency of the combustion, H_f is the enthalpy, or transformable energy per unit mass of the fuel, the power actually developped by the combustion is $\eta_f . Q . C_p . (T_o - T_i)$.

3.3 One shaft global turbine KBPD model

Let us consider a gas turbine actionning an inertial mechanical load, as shown in figure 2. The variables will be indexed with respect to the output of the blocks, index $_a$ is the external atmosphere.

Compressor: W_{cM} is the mechanical energy received by the compressor via the transmission (shaft),
energy balance: $W_c = W_{cM} + W_L$, flow balance: $Q_E = Q_a - Q_L$,
temperatures: $T_c = T_a + a_c |W_c|$.

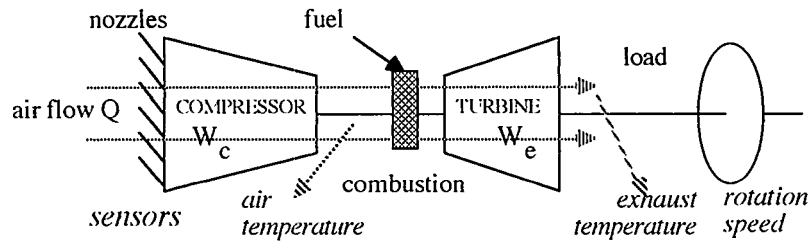


figure 2

Isobar Combustor : $P_B = P_c$, and $T_B = T_c + j.\ Q_f$,

Turbine: Exhaust Pressure: $P_E = P_a$, $\rho_E = P_a / P_B$, $T_E = T_B + \alpha_B |W_E|$,

mechanical rotation speed : $N_E = (W_E / b_E)^{1/2}$,

energy : $W_E = (\eta_{iE} / \alpha_E) .T_B (1 - \rho_E^{\gamma - 1/\gamma})$

Mechanical inertia: W_{cM} transmitted to the compressor via the inertial transmission (shaft), $W_{cM} = W_E - J_M . N_E .dN_E./dt$

4 CAUSAL RELATIONAL GRAPHS

4.1 Causal influence graph

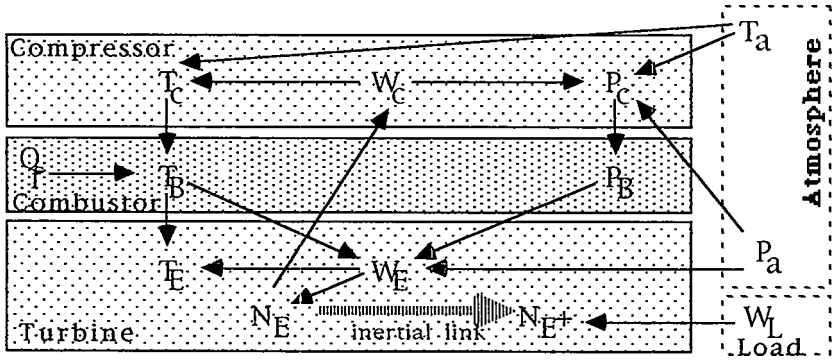


figure 3

A causal graph has been build from the chaining of equations which best suits the common sense physical knowledge. This graph can be considered as a potential simulation model that calculates the future

value of the rotation speed NE+ from the present value NE. It is instanciated by the determination of relations representing the links between nodes. It must be noticed that nodes correspond to variables, and their values can be quantitative or qualitative, but they do not exhibit a logical state. Marginal influences will be deduced from equations taken as oriented assignements, or by instructions in a simulation .

4.2 Causal Diagnostic graph

The above causal graph becomes a Causal DiagnosticGraph by interpreting the nodes, no longer as physical variables, but as logical ones where the values are OK or not-OK. The global graph is obtained by connecting the turbine, compressor and combustion graphs shown above.To the initial graph must be added the primary causes of the possible failures, each primary cause has a fuzzy logic likelihood value.

symbol	meaning	fuzzy likelihood
CoNOZ	incident in the compresor nozzles	0.6
CoLK	leaks in the compressor	0.7
TuLK	leaks in the turbine	0.7
fuel	quantity and quality OK	0.9
load	Correct load	0.4

Manifestations, as logical variables are related to the correct or incorrect values given by the 3 sensors:

symbol	meaning
air	correct temperature of the air delivered by the compressor
exhaust	correct temperature of the exhaust gas
rotation	rotation speed in the normal range

The complete AND/OR/NOT diagnostic graph for the one-sahft turbine is represented in figure 4.

Comparing the diagnosis for rotation failure in figure 5 and for exhaust temperature in figure 5 we notice that, except for some incident in the load, both involve the same possible primary causes, nevertheless the sorting by fuzzy likelihoods gives an additional information that causes a different order in those causes. e.g a bad quality of fuel is more likely to have produced an exhaust anomaly than a wrong rotation spees.

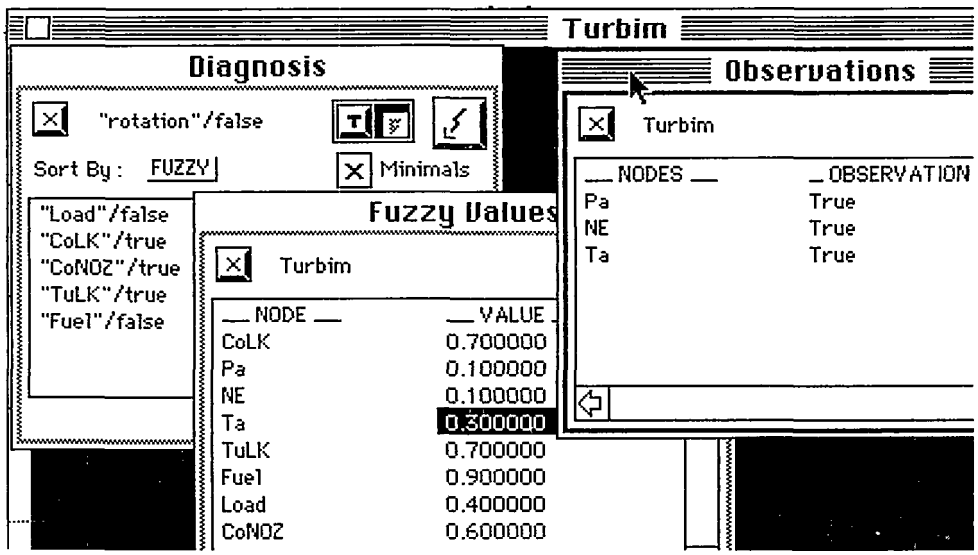


figure 5

The left half of figure 6 shows the diagnosis of an air temperature out of range, and in the right half the diagnosis of an exhaust gas temperature out of range.

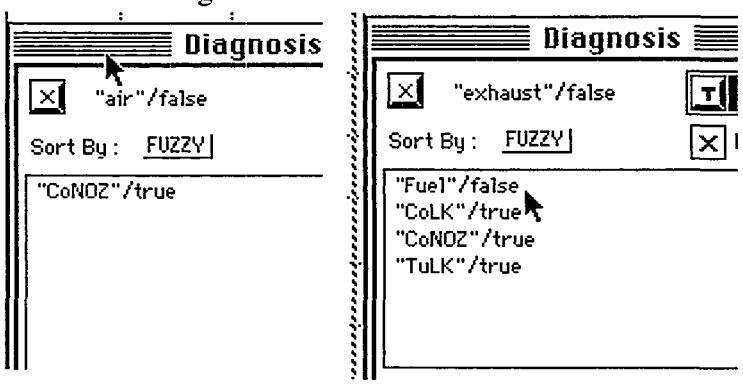


figure 6

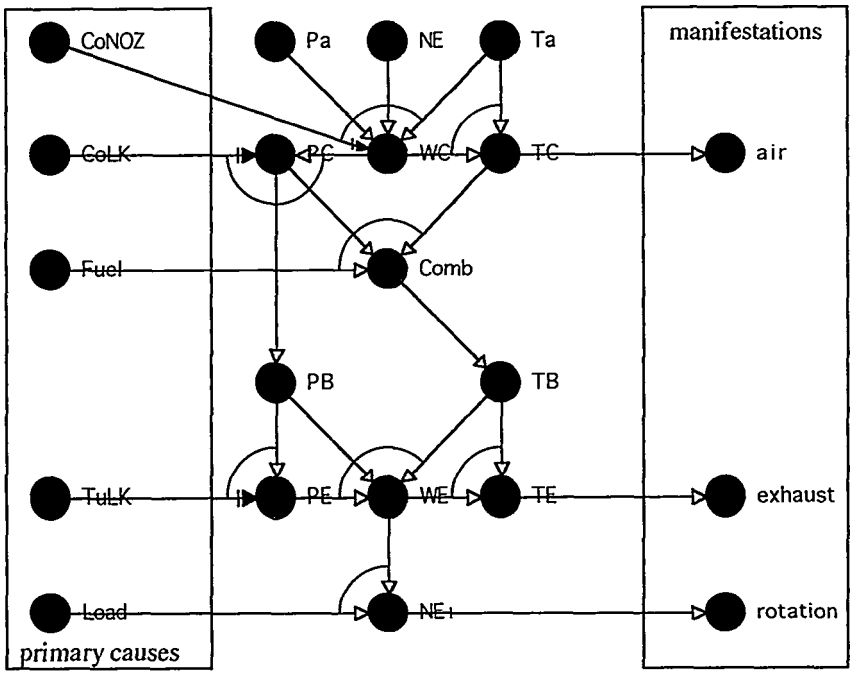


figure 4

5 THE "PILAR" DIAGNOSIS TOOL

Based on the above principles a diagnosis tool is being implemented in Prolog and called "Prolog Inferential Logic Abductive Reasoning" (PILAR). The present version 1.1 gives the possibility of declaring graphically, as shown here, 2 types of nodes: AND and OR, 2 types of links: identity and negation, and to assign fuzzy likelihood values to the primary causes. In PILAR 1.2 arcs will accept also fuzzy transfert coefficients.

Any node can be chosen by the user, but preferably manifestation nodes, and its diagnosis is given in the form of a sorted list with respect to their fuzzy likelyhoods. Figure 5 shows the screen obtained by the diagnosis of a failure in the rotation speed from the graph of figure 5. The fuzzy likelihood values, as well as the observations made about external nodes: meteorogical variables, i.e. athmospheric pressure and temperature are OK, and past rotation speed NE was inside its normal range.

6 CONCLUSIONS

The use of abductive inference in AND/OR/NOT graphs has proved to suit the knowledge based representation of complex systems [Vescovi-Laurent 1992] for diagnostic based in normal behaviour. Several approaches are being experimented nowadays for diagnosis, and it seems that the dificulty of determining causality in dynamic closed loop systems can be partially overcome by applying either fuzzy likelihoods or qualitative probabilities, together with knowledge based models. In other papers as [Fuster & Ligeza 1996] more formal definition of the diagnostic statements in the framework of logic computing can be found, and futher work on that direction is being done.

REFERENCES

[Aguilar 1993]. Aguilar Martin J., Delgado, A. Escobet T. Knowledge-Based Modelization for Diagnosis and Supervision of Industrial Process. *Proceedings of IRTICS* - sec. 14 pp. 1-5, Madrid Oct.1993

[Fuster & Ligeza 1995]. Fuster Parra P. & Ligeza A. : Fuzzy fault evaluation in causal diagnostic reasoning. In *Applications of Artificial Intelligence in Engineering X, G,Rzevski, R.A. Adey, & C. Tasso eds. Computanional mechanics Publications*, Southampton-Boston, pp137-144

[Fuster & Ligeza 1996]. Fuster Parra P. & Ligeza A. : A model for representing causal diagnostic reasoning. In *Proceedings of the 13th European Meeting on Cybernetics and Systems Research*, vol. 2, pp. 1222-1227, Vienna, Austria 1996

[Fuster 1996]. Fuster Parra P. : A model for causal diagnostic reasoning. Extended inference modes and efficiency problems. *PhD Thesis, Universitat de les Illes Balears*,Palma de Mallorca, Spain, 28 Oct.1996.

[Peng & Reggia 1990]. Peng Y. & Reggia J.A. : Abductive Inference models for diagnostic problem-solving. *Springer-Verlag*, 1990

EXPERT SYSTEM KERNEL USED IN PLANT DIAGNOSIS

Adriana Alexandru, Senior Research Scientist and Stefan Ungureanu, Programmer
Research Institute for Informatics, Bucharest, Romania

ABSTRACT

The kernel of an expert system is the most important part of a project concerning the diagnosis of any kind of process. The paper introduces such a kernel that is part of an expert system shell used to develop a diagnosis application for a photovoltaic plant. The following facilities are provided: rule-based expert system, fast inference engine for real-time process diagnosis, flexible interconnection with monitoring software, complete calculator for computing parameters based on values returned by monitoring application, low-effort development and maintenance of rule set.

1 INTRODUCTION

Major work involving the expert system shell design was performed by a team from the Research Institute for Informatics, Bucharest, Romania within the framework of the project "Concerted Actions on PV Systems Technology - Development of expert systems for PV Technology Information and Performance / Diagnosis in PV Plants" from the JOULE II Programme of the European Commission Directorate General (DG) XII. It was coordinated by WIP, Munchen, Germany and assisted in the design specifications and experimental testing by AGSM, Verona, Italy. It is a research regarding the industrial application based on artificial intelligence techniques in similar fields.

The expert system shell can be used as a base for the implementation of a real-time expert system for performance analysis and diagnosis of an

industrial process. It provides a mechanism able to simplify the operator's job of recognition of the emergency state, diagnosis of the fault and initialization of the necessary corrective actions, offering him the alarms list and messages to define the primary cause as well as a list of priorities of potential solutions [1].

The use of such a system has the following results:

- prevent key components damages;
- extend critical components life-time;
- improve overall system performance and reliability;
- minimize off-line analysis efforts and maintenance costs;
- assist the maintenance activities, generating functional reports.
- allow remote plant diagnosis and evaluation.

The kernel of an expert system is the most important part of a project concerning the diagnosis of any kind of process.

2 BASIC CONCEPT OF FAULT DIAGNOSIS

Due to the increasing complexity and risking of modern control systems and the growing demands for quality, cost efficiency, availability, reliability and safety, the call for fault tolerant in automatic control systems is gaining more and more importance [2]. The active approach used to achieve fault tolerance provides fault accommodation, i.e. a reconfiguration of the system when a fault has occurred. To this end, a number of tasks have to be performed, one of the most important and difficult of these being the *diagnosis of the faults*. Faults in an automated system can occur in both the plant and the control units.

The basic tasks of fault diagnosis are to detect and isolate occurring faults and to provide information about their size and source [3]. This has to be done on-line in the face of the existing unknown inputs and without or with only very few false alarms. As a result, the overall concept of fault diagnosis consists of the three subtasks: *fault detection, fault isolation and fault diagnosis*.

In *signal-based* approach of *fault detection*, one extracts proper signals or symptoms from the system, that carry as much information as possible

about the faults of interest [2]. The symptoms are either directly or after proper modifications used for the fault decision and classification. Typical symptoms are the magnitudes of the time functions of the measured signals, arithmetic or quadratic mean values, limit values, trends, statistical moments of amplitude distribution or envelope, spectral power densities or frequency spectral lines, etc. The signal-based methods are used for early fault detection and the detection of faults that occur in the dynamics of the system under consideration.

The *knowledge-based approach for fault diagnosis* is a more suitable strategy in the case of noticeable modeling uncertainty. In this case, knowledge has to be processed which is commonly incomplete and can not be represented by analytical models. The residual evaluation is a complex logical process which demands intelligent decision making techniques like fault trees. Therefore, knowledge-based methods are a quite natural approach in fault diagnosis and the expert systems are applied more successfully here than in the field of control. Knowledge-based diagnosis techniques fall into two categories [2]:

- *symptom-based*, when heuristic symptoms, knowledge about process history or statistical knowledge are used , the evaluation of which being organized in the framework of the diagnosis expert system. The major difficulty is the knowledge acquisition;
- *qualitative model-based*, when the knowledge is derived in terms of facts and rules from the description of system structure and behavior.

3 THE COMPONENTS OF THE EXPERT SYSTEM KERNEL

There are three major components of the expert system kernel: the calculator, the inference engine, and the coordination routines, all together sharing an user-friendly windowed interface.

3.1 THE CALCULATOR

The calculator is used to compute various kind of parameters, named '*Calculated Parameters*', based on the instant values of the '*Acquired Parameters*' provided by the monitoring software. The user has the possibility of defining new intelligent parameters to be used in the diagnosis process. A parser built with dedicated tools (*lex and yacc*) is

used to implement the user interface and the evaluation routines. The computations include useful arithmetic, statistical, both analogic and digital, mathematical functions (figure 1).

A small database is kept when statistical (e.g. average, minimum or maximum value) functions are to be evaluated.

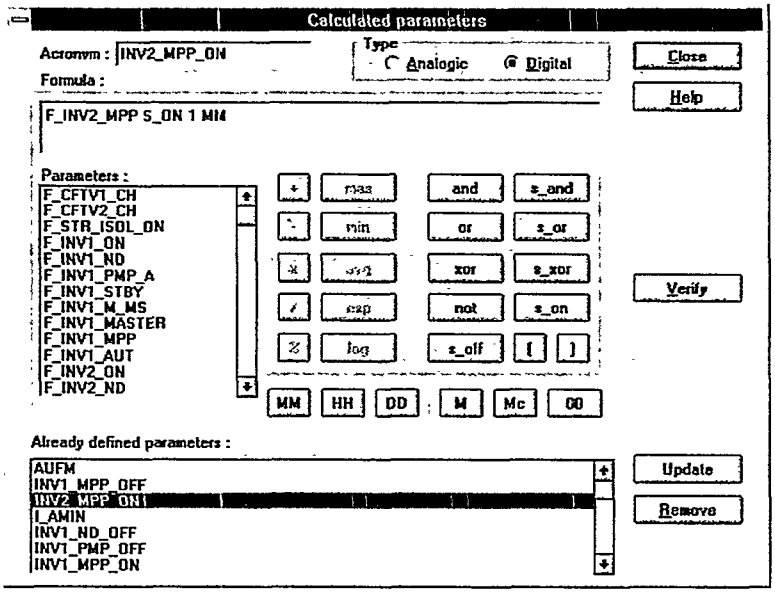


Figure 1. Calculated Parameters Dialog Box

3.2 THE INFERENCE ENGINE

The inputs of the knowledge based fault inferencing mechanism are all available symptoms as facts and the the fault relevant knowledge on the process, mostly in heuristic form [4]. In general, specific rules are applied in order to set up logical interactions between observed symptoms (effects) and unknown faults (causes).

The propagation from appearing faults to observable (detectable) symptoms follows, in general, physical cause-effect relationships, where physical properties and variables are connected to each other quantitatively and also as functions of time. However, the underlying

physical laws are frequently not known in analytical form or too complicated for calculations.

The developments of AI approaches was oriented initially to medical diagnosis and then extended to technical processes. Therefore, also for technical diagnosis, only heuristic symptoms are considered.

The heuristic knowledge in the form of *heuristic (qualitative) process models* can be expressed in the form of *rules* like [4]

IF < condition> THEN <conclusion>

The condition part (premise) contains facts in the form of observed symptoms S_i as inputs and the conclusion part includes events E_k and faults F_j as a logical cause of the facts. Chaining of the rules now establishes the causal dependencies of symptoms and faults in a hirarchical manner. Thus intermediate events E_k are introduced. This procedure results in fault-symptom tree, a directed graph which relate symptoms to events and faults. The backward chaining is then used as a reasoning strategy.

The primary attribute of the rule-based inference engine is the *speed.* Every effort was made to ensure that the engine will run as fast as possible, due to the real-time aspect of the project. This inference engine must complete the 'reasoning' process before the next set of parameter values comes from the monitoring software and is processed by the calculator. The rule architecture was designed with this purpose and resulted also in a low-effort maintenance knowledge-base with compiling capabilities. The inference process uses an internal representation of the rule set and parameter values, resulting in a fast recognition of the photovoltaic plant state.

Some characteristics of the inference process are:

- continuos operation;
- non-interruptible and temporal reasoning;
- capacity of recognizing the deviation of the process states and to investigate the errors' causes in order to offer support for a correct intervention of the user;

- possibility of enlarging the system.

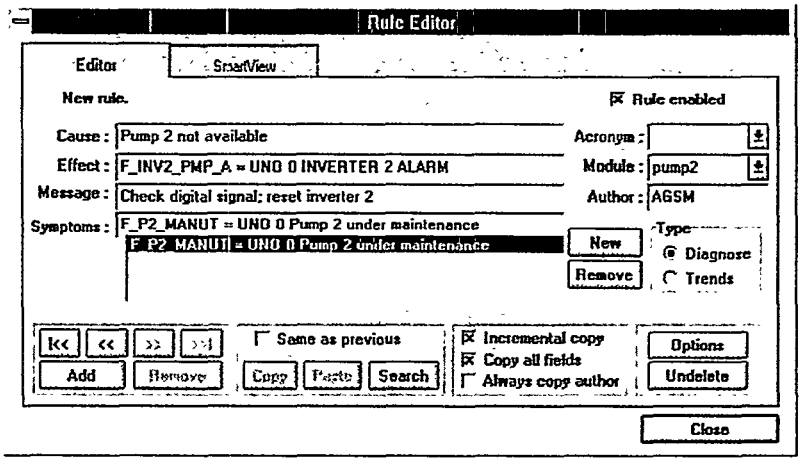


Figure 2 - Rule Editor

The inference engine maintains a parameter history database, and is capable of trace operations with full description of the facts that leaded to a certain decision.

3.3 THE COORDINATION ROUTINES

The last component, but not the least, is the coordination algorithm. It assures reliable communication with monitoring application, internal synchronization and database, both knowledge and parameters, consistency.

The *coordination algorithm* communicates with applications that supply monitoring parameter values in a file, local or networked. The communication is achieved through a reliable channel: the file can be an ASCII text file or a binary file containing parameters' names and values. so that many monitoring applications can be used, leading to architecture flexibility; it also ensures that parameter database is kept consistent in case of communication failure. This file is stored in a shared directory , if the monitoring and diagnosis systems reside on the same computer. Otherwise, the files are stored on the monitoring computer.

In case of a malfunction of the transmission, data are stored and kept for a certain period of time (e.g. one hour). When reestablishing the

connection, the diagnosis system will synchronize itself with the monitoring software and restore its state as if the connection never failed.

Internally, it synchronizes the calculator and the inference engine, passing data from one to another, and to the user interface.

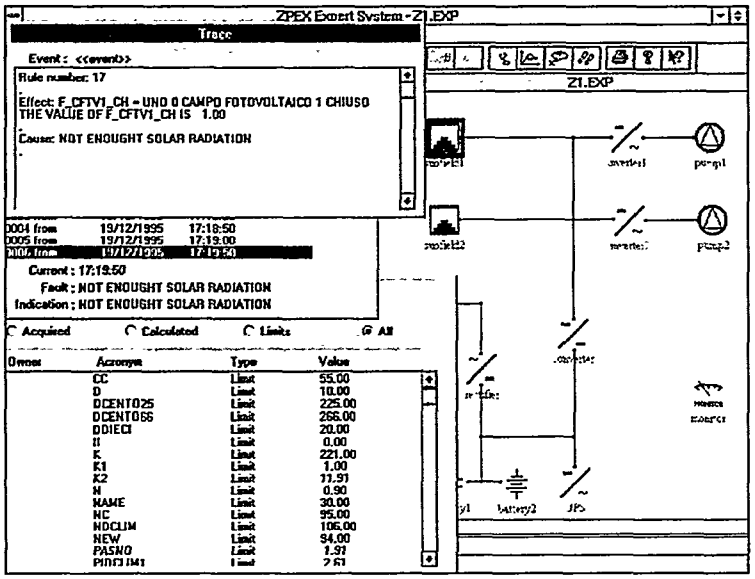


Figure 3. Run-time Screen Shot

The synchronization algorithm will also allow *remote access to the diagnose results*. An expert system running in remote mode is able to display information and alert the user if necessary in the same way the main application is doing on the main machine. This can be very useful when diagnosing processes in hostile environments or when a permanent operator cannot be used at the process site.

For history purposes, daily reports are saved in files and can be printed upon request.

4 HARDWARE AND SOFTWARE REQUIREMENTS

The expert system for performance analysis and fault diagnosis uses the on-line acquired data of the data base built by a conventional real-time monitoring system of the plant. Personal computers became an usual

support for expert systems, but memory and performance limits make them available only for small and medium systems.

The hardware configuration required by the expert system is an IBM PC compatible computer with: 486 DX, 66 MHz, 16 MB RAM, 350 MBB HDD, 3.5" FDD, SVGA video adapter, 14" color monitor, mouse.

The system was developed using Microsoft Visual C++, version 2.0, and targets the 32-bit Windows environments available from Microsoft: Windows NT 3.5 and Windows 95, or any further releases (Windows NT recommended).

5 CONCLUSIONS

High quality software systems is now-a-days an important request that software technology is confronted with. Expert systems are powerful tools that use artificial intelligence techniques for providing information just like a human expert would. The requests related with superior capabilities of 'reasoning' are motivated by the growing number of potential users of such systems.

A concrete implementation of an expert system based on the expert system shell described in the paper was made at the Zambelli PV pilot plant, near Verona, Italy.

6 ACKNOWLEDGEMENT

The authors extend gratitude to the WIP, Munchen for coordinating the project and AGSM, the municipal utility of Verona, for collaboration in the implementation, installation and testing of the diagnosis expert system in Zambelli 70 KW PV Pumping Station.

REFERENCES

1. Alexandru, A., Hogea, S., Jitaru, E. & Trica, A., "Technical aspects Regarding Expert Systems For Technical and Medical Diagnosis", Research Report, ICI, 1993.

2. Frank, P. M., "Analytical and Qualitative Model-based Fault Diagnosis - A Survey and Some New Results", European Journal of Control, 2, 1996, 6-28;

3. Frank P. M. & Köppen-Seliger B., "New Developments using AI in Fults Diagnosis, IFAC/IMACS International Workshop, Bled, Slovenia, 1995;

4. Isserman R., "Integration of Fault Detection and Diagnosis Methods", Prococeedings of IFAC Symposium on Fault Detection, Suppervision and Safety for Technical Processes, Safeprocess '94, Helsinki, June 1994, 597-612.

KNOWLEDGE BASED MAINTENANCE

Prof. Dr.-Ing. A. Sturm, Hamburgische Electricitäts-Werke AG
Hamburg, Germany

ABSTRACT

The establishment of maintenance strategies is of crucial significance for the reliability of a plant and the economic efficiency of maintenance measures. Knowledge about the condition of components and plants from the technical and business management point of view therefore becomes one of the fundamental questions and the key to efficient management and maintenance. A new way to determine the maintenance strategy can be called: Knowledge Based Maintenance. A simple method for determining strategies while taking the technical condition of the components of the production process into account to the greatest possible degree which can be shown. A software with an algorithm for Knowledge Based Maintenance leads the user during complex work to the determination of maintenance strategies for this complex plant components.

1. STRATEGIC CONCEPT

The overall strategy of a plant is made up of a multitude of individual strategic decisions. The practical work of selecting suitable maintenance strategies is influenced by several independent characteristics of the production plant components :

- **recognition of the technical condition of components**
- **importance of the components for the production process**
- **specific costs dependent on life-time**

The recognition of technical condition is determined by the possibilities of diagnosis. The importance of the components for the production process can be labeled with priorities. The decision-making process is ultimately dominated by the desired minimization of costs through the life-time of the plants or their components. By doing so this decision-making process can be dealt with based on knowledge. This new way of linking differing requirements to the determination of the maintenance to be carried out can be called

Knowledge Based Maintenance.

2. CONDITION KNOWLEDGE

Knowledge about the condition of components and plants from the technical and business management point of view therefore becomes one of the fundamental questions and the key to efficient management and maintenance. In the following, knowledge of and about the technical condition and the management characteristics of plants and their components will be called **Condition Knowledge:**

> **Condition Knowledge is the entirety of all the information about the technical and management characteristics of plants and their components in correlation to their use.**

All data on the layout and the design are basic informations. Informations depending on service life are obtained from the management's processing data, from measurements during operation, when servicing is carried out, during inspections and from repair findings.

In the past decades condition knowledge has primarily been investigated over decades by experienced specialists and stored in their heads or in their notebooks. It is hardly conceivable that there will be enough specialists available who are able to register and process this overall knowledge in the future. Industrial management, having less and less active personnel and yet an increasing flood of data and information, requires new methods of presenting condition knowledge for managerial decisions.

When new systems are designed for maintenance, the very first necessity is to analyze how the elements which determine the conditions and the elements relative to managing maintenance have to be prepared and treated. In order to structure information processing programs, one must know what requested decisions can go off **rule-based** or **knowledge-based.** Using examples, figure 1 gives a summary of rule-based and knowledge-based elements of these organizational and decision-making processes. Rule-based elements can be converted to algorithms. Knowledge-based elements can be supported by applying information processing.

Managing an operational process requires various information and data. The information and data is from various sources. Some important **sources of information** in operational processes are:

- **Manufacturer's documentation**
- **Data on initial operation**
- **Measurement data of process-control engineering**
- **Recordings of the operating personell**
- **Diagnostic data**
- **Maintenance findings**
- **Reliability data**
- **Organizational data on the operational and maintenance process**
- **Cost management data**

The source of the information already determines the form and quality of information. Questions to be answered are: When and how information on the production progress and the condition of the plant can be gained; how this information is to be processed; and finally, to what assessment individual pieces of information will be subjected. Figure 2 attempts to show the complexity of this process.

Information from industrial management, from measurements during operation, when servicing is carried out, during inspections and from maintenance findings can be subdivided according to its form, its outer appearance or the time when it occurs. Some useful **forms** of information are:

- **Texts**
- **Images**
- **Signals**

These informations may occur or may be called up **continuously** or **discontinuously**. Beyond this, signals can occur **quasi-constantly** or **fluctuating**. These kinds and forms of information require differing treatment and differentiated assessment. The variety of this supply of information complicates the subsequent information processing procedure and for decades has been the reason for its not being used enough. The developments in technology relating to measuring, diagnosis and information processing presently allow especially intensive further development.

In the simpliest case, information processing uses the procedure of information extraction. Only the portions of information from a signal, image or text (e.g. the mean value of a quantity to be measured) are used that correlate to decisions concerning the process. One can easily go on working with single-parameter quantities.

Often multiple parameters are required for a reliable decision and justifytheuse of information linkages. Models, patterns or coefficients are the result of such manipulations.

For the assessment, one uses determinate or stochastic procedures. The simplest assessment is the direct assignment of a piece of information to an event. This is only possible for a few simple decisions. As a rule the decision-making process is a complex process and requires differentiated algorithms or methods of coming to decisions.

3. CONDITION MONITORING

The first question to answer for establishing a maintenance strategy is when and how information about the condition of a plant can be obtained. There are four possibilities to take into account:

- **Diagnosis during operation**
  The diagnosis of wear is realized through diagnostic systems (stationary - ambulant) continuously or discontinuously (online / offline) and by subjective diagnoses during operation. In cases of overload use, secondary damage is to a great extent prevented.

- **Diagnosis during maintenance**
  This diagnosis is provided through inspection or through measuring during a standstill of the plant, after dismantlement as a rule, in the maintenance process.

- **Diagnosis not necessary**
  With normal stress during the life-time the component will not contribute to any malfunction or to a breakdown.

- **Diagnosis not possible**
  There is wear but a diagnosis cannot be effectively drawn up.

One should not neglect in particular those wear processes that cannot be diagnosed at present but which can lead to considerable aftermath effects of breakdowns. Classifying all subassemblies in these diagnostic groups is a first and inevitable preparatory step for condition-based maintenance.

4. TYPES OF REPAIR WORK

Repair work activities can be realized in widely varying forms of organization and structures. They can be preventative or merely eliminate damage.

But also here, dividing up all the activities into the following three categories can be of help:

- **Unscheduled repair after a breakdown**
  The preparation is carried out after a malfunction occurs.

- **Operative repair before a breakdown**
  The time point of the measure is unknown.
  All activities are prepared in a zero-plan.

- **Scheduled repair before a breakdown**
  The repair is carried out according to schedule and all activities are prepared in advance.

This division of possible repair activities into three categories, which at first appears unusual, is necessary to realize condition-based maintenance.

Not until it is managed to set up organizational forms of maintenance that raise the degree of preparation of maintenance so far that, except for the scheduled time of the activity, everything else is prepared in advance as defined by the zero-plan (work order, material, spare parts, manpower in shift work). In this case assignment will be possible according to diagnostic findings with maximum use of the wear available and minimal standstill costs. This form of repair is called 'operative repair before a breakdown'.

The 'unscheduled repair' and 'scheduled repair' follow the practices that have been carried out up to now.

5. STRATEGY SELECTION

The next practical step in establishing maintenance strategies is to classify all subassemblies, depending on the possibilities of diagnosing them, into the maintenance categories. Figure 3 shows combination possibilities to be decided.

Every classification is dependent on the position of the subassembly in the production process and on the permissible workload for the maintenance measures required. With sufficient knowledge, classification of strategies to different parts can be calculated. For example it is possible to determine reliability factors, assess changes in reliability or calculate the most cost-effective cycles (cf. methods in the sections above).

Presently the data required for calculations are often insufficient. In this case estimates have to be substituted for calculations. Using priority criteria for such estimates is a suitable aid.

The following regularities convey even more global classification.
Subassemblies which are either so insignificant or have an outstanding life-time or degree of reliability that allow to make a judgement of 'no diagnosis required' are classified as 'unscheduled repair'.

A problem arises with the classification of subassemblies for which a diagnosis would actually be required but which is not possible. In such cases the responsible person for the operations will have to classify the subessembl either in the 'unscheduled repair' or the 'operative repair' group. A scheduled replacement is not recommended, or should only be favored for subassemblies which are particularly exposed.

Subassemblies for which a reliable diagnosis can be provided during running operations are classified to the 'operative repair' group if the extent of the activity or the manner of operation allow so. This is also the dominating conversion from a Time Based Maintenance Strategy to a Condition Based Maintenance Strategy.

For subassemblies which can only be diagnosed when maintenance is done, repair in this rigid cycle should also be classified either immediately or for the next (required) 'scheduled repair'. Especially with regard to these subassemblies, the search for and research into diagnostic methods during operation must be intensified.

This covers all the subassemblies of a complex installation in a regime aimed at the most far-reaching application of a Knowledge Based Maintenance.

6. KNOWLEDGE BASED MAINTENANCE

This simple plan for determining strategies while taking the technical condition of the components of the production process into account to the greatest possible degree can be summed up as an algorithm:

1. **Estimation of the life-time of components**
   Rough judgement of possible maintenance

2. **Ascertainment of the diagnostic possibilities for the components**
   Formation of groups and estimation of the permissible diagnostic workload

3. **Definition of possible and realizable types of repair**
   Repair before malfunctions and repair after malfunctions

4. **Assessment of the components through priorities**
   Depth of differentiation according to the installation valency and scope

5. **Classification of components to types of repair depending on the possibility for them to be diagnosed**
   Classification through priority decisions or calculation

A piece of software with an algorithm for knowledge-based maintenance guides the individual responsible for operations when doing the extensive work of establishing the maintenance strategies for the components of a complex plant. Figure 4 shows a flow-chart of the procedure for coming to these decisions.

For maintenance activities for which the point of timerectly from diagnostic information, it is possible to generate automated repair orders. Preconditions are unequivocal addresses of the components to be repaired, schedules of work that are prepared in advance and retrievable and unequivocal algorithms to generate scheduling. The plans for such an order generator are shown in figure 5.

7. BIBLIOGRAPHY

Sturm,A.: Zustandswissen für Betriebsführung und Instandhaltung
VGB Kraftwerkstechnik, Essen 1996.

Sturm,A.: Instandhaltung von Energieerzeugungsanlagen
VGB Kraftwerkstechnik, 76 (1996), No. 6, p. 476-480.

8. Figures

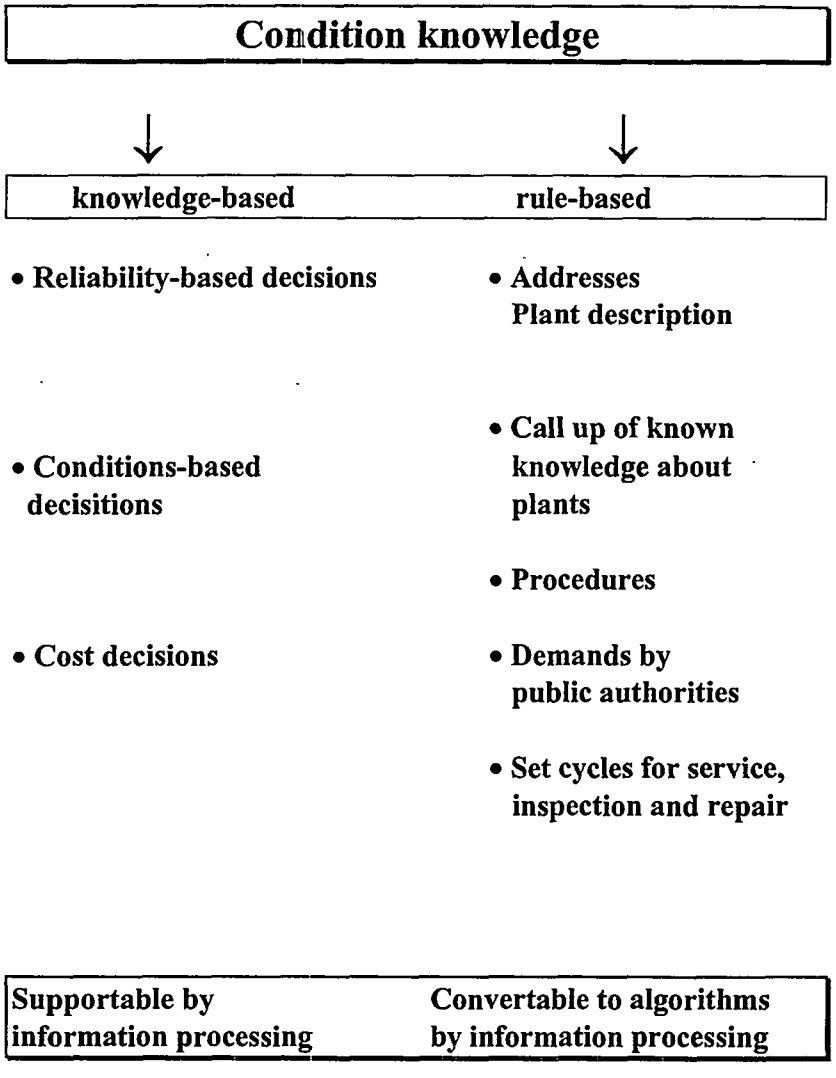


Figure 1 : Structuring of condition knowledge

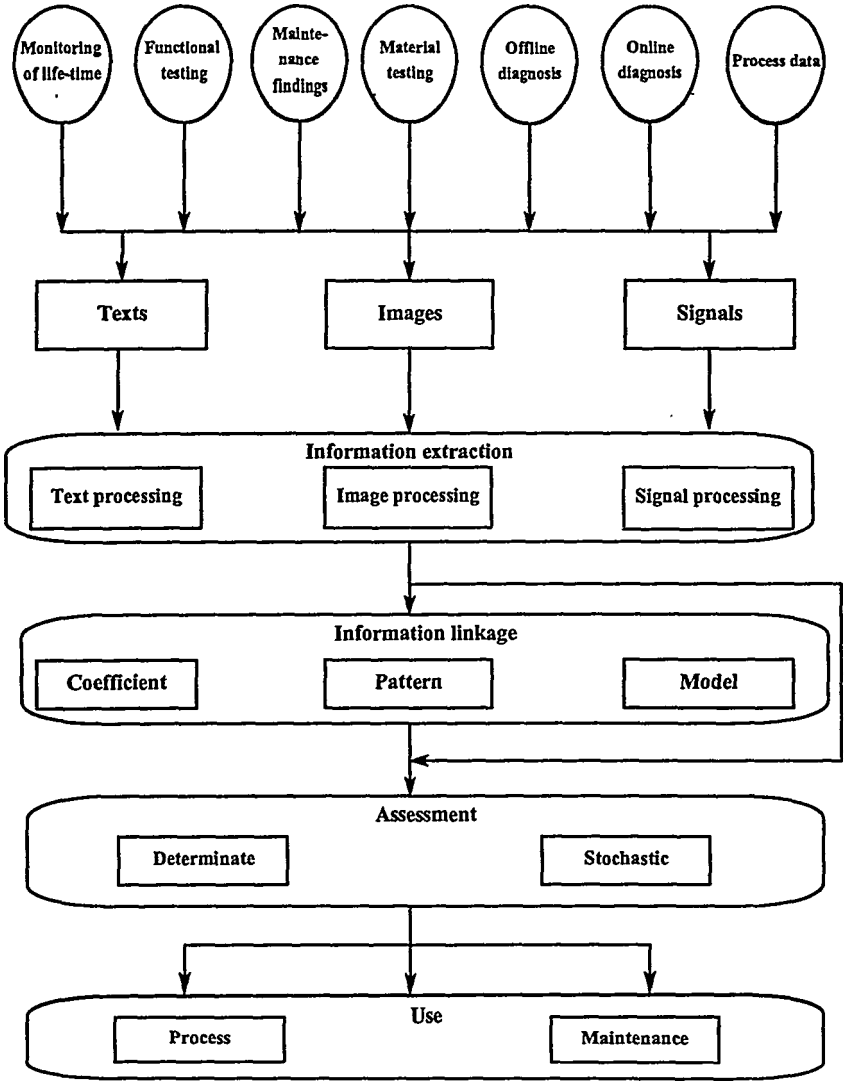


Figure 2 : Information sources of condition knowledge

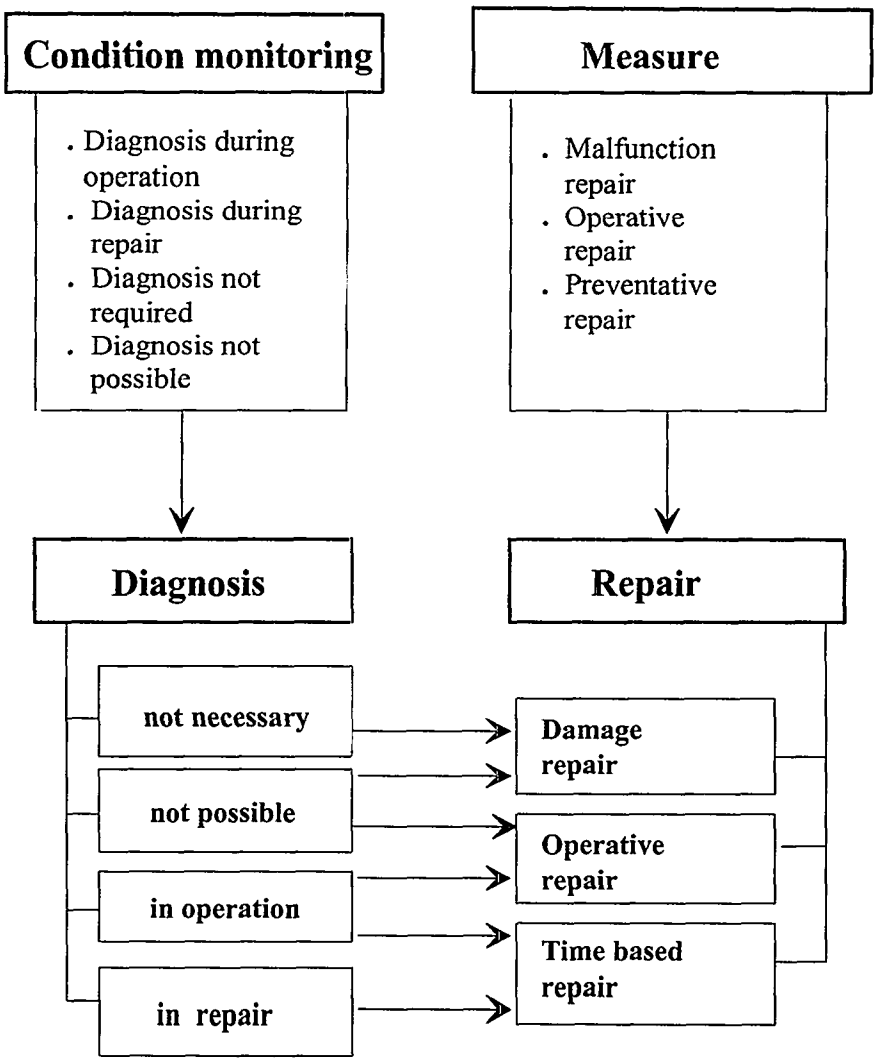


Figure 3 : Classification of diagnostic activities to types of repair

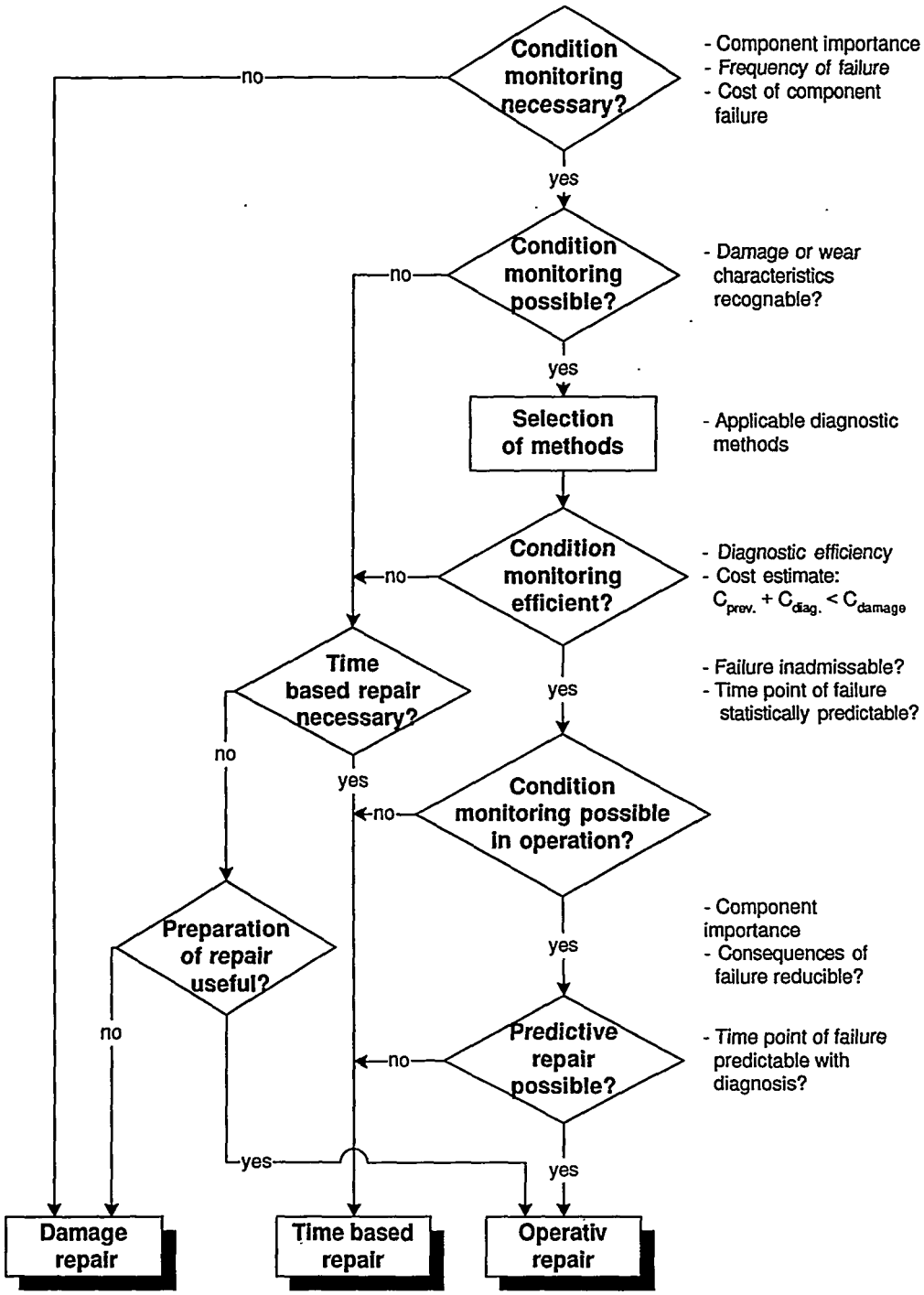


Figure 4 : Coming to decisions for maintenance strategies

Formation of maintenance order

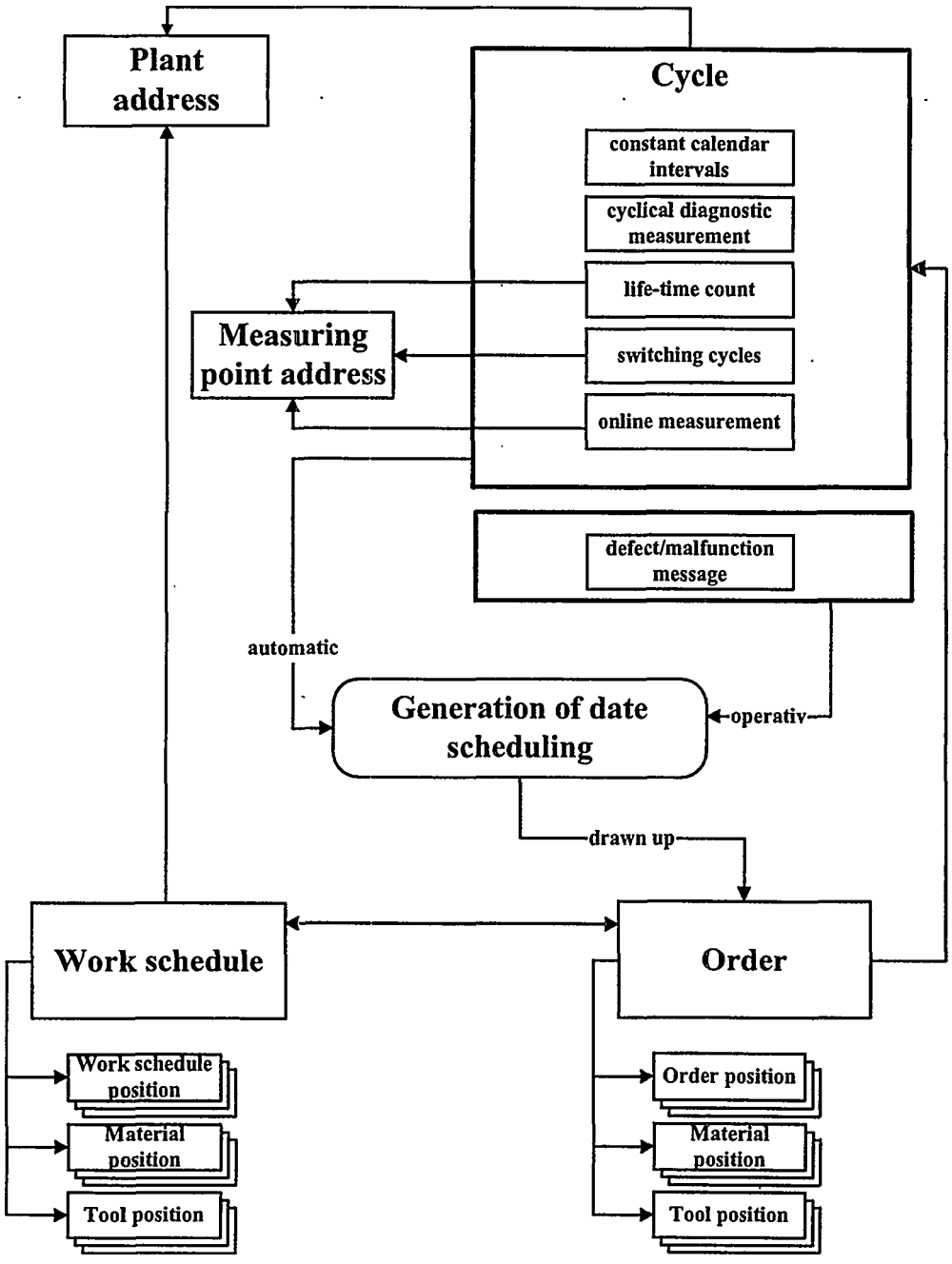


Figure 5 : Generating maintenance orders

DIAGNOSTICS OF ELECTRICAL MOTORS BY MEANS OF LABVIEW PACKAGE

Leon Swędrowski, Adiunkt
Technical University of Gdańsk, Electrical Engineering Faculty,
Gdańsk, Poland

ABSTRACT

Diagnosis of driving systems enables to pass from the scheduled maintenance to the maintenance according to the real machine condition. The best results of diagnoses are obtained using the methods based on spectral analysis of vibration signals or on analysis of supplying currents (in the case of electrical driving systems). The most numerous group of driving system breakdowns is connected with the damages of roll bearings. In this case very effective diagnosing can be obtained using envelope signal processing (ESP) technique. According to the educational program of the Electrical Engineering Faculty (TUG) a special training stand was elaborated. It enables to make research with ESP diagnostic method used for rolling elements bearings as well as supplying current analysis.
The stand can be used also for slip measurement. The measurements of induction motor slip enable to follow-up action of driving systems.
For measuring and signal processing is used computer aided system with the virtual instrument based on LabVIEW software.
The diagnostic system is based on virtual instruments using LabVIEW of National Instruments. This solution is about 50% cheaper than the system realized with specialized instrumentation. The Application Builder delivered with LabVIEW enables to use this expert monitoring system in small enterprises equipped with PC units without LabVIEW software.

1 INTRODUCTION

Diagnostics of driving systems enables to watch their condition in time of operation and to pass from scheduled maintenance to the maintenance according to the real machine condition. For this reason it

is necessary to use the diagnostic methods which allow to predict condition of each machine construction and elements in situ without its disassembly and moreover during its working condition. In this case it is possible to eliminate unexpected breakdowns and usually to elongate the period between two successive maintenance.

The methods based on analysis of generalized spectra of vibration signals or envelope signal processing (ESP) are interesting. In these methods as the source of analyzed signals are the accelerometers fixed on the housing of the investigated unit. Recording and analyzing of these signals demands utilization specialized instruments. The analysis of harmonics and increase of their values in time enables to state the diagnosis of bearings (the most often cause of unexpected breakdowns) and to predict the moment of next examination or time of necessary maintenance.

In the last decade it could be observed the dynamic development of virtual instruments (VIs). Among the main producers of this software there are: National Instruments Corporation (LabVIEW, LabWindows for DOS and LabWindows/CVI for Windows), Hewlett Packard (HP VEE for Windows) and Kiethley Instrument (Test Point). The diagnostic system based on virtual instruments is about 30..70% cheaper than the systems with specialized instruments and is also more convenient in the case of its further transformation.

The stand we use for driving systems monitoring was elaborated by the group from the Electrical Measurements Department of the Electrical Engineering Faculty (TUG). This stand is based on the LabVIEW package which enables to form instrumentation suitable to demands of each customer. Moreover using the Application Builder of the National Instruments it is possible to create VI system for the driving systems diagnosis which can be used with computers do not equipped with LabVIEW package.

This stand consists of numerous motors with determined bearings defects as well as the motors of the same type being in good condition. For this reason it is possible to present the spectra of strictly specified defects and to watch the development of these spectra in time.

2 DIAGNOSTIC SYSTEM

Diagnostic systems is realised using the virtual instrument (VI) basing on LabVIEW package of National Instruments. The system enables to

determine the electrical faults of electrical motors (induction motors) as well as for analyzing the mechanical defects of these motors

2.1 The analysis of mechanical defects

The analysis of mechanical defects is realised in two steps (Fig. 1). In the first step (data collections) there is used the measuring tape recorder TR of Brüel & Kjaer connected to the accelerometer Ac fixed to the bearing housing of investigated unit. In the second step (data evaluation) the signals from the tape recorder are processed and analysed by diagnostic and conditioning VI-system which consists of the PC 486 equipped with LabVIEW package and a/d card of NS, display unit D and printer P.

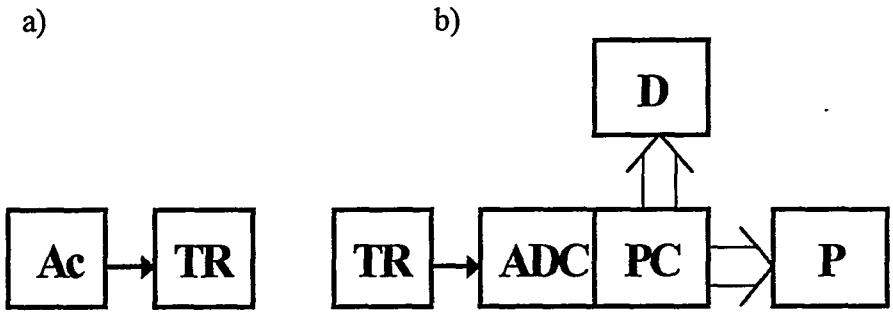


Fig. 1 Bloc diagram of data collections (a) and data evaluation (b)
Ac - accelerometer, TR - magnetic tape recorder,
ADC - a/d card, PC - personal computer, D -display,
P - printers

The diagnostic VI-system processes data registered on magnetic tape for precisely determined bearings.
The main data of numerous bearings are stored in bearing directory. Taking into account the value of rotational frequency f_r at the time when the signals from an accelerometer were registered the main defects frequencies of tested bearing can be determined.
Afterwards it is possible to start the diagnostic process using VI system. The example of virtual instrument, presented as a window of LabVIEW is shown in the figure 2.
The next window presented in the figure 3. shows the values of depth modulation coefficient (DMC) for particular bearing defect frequency specified its fault i.e.:

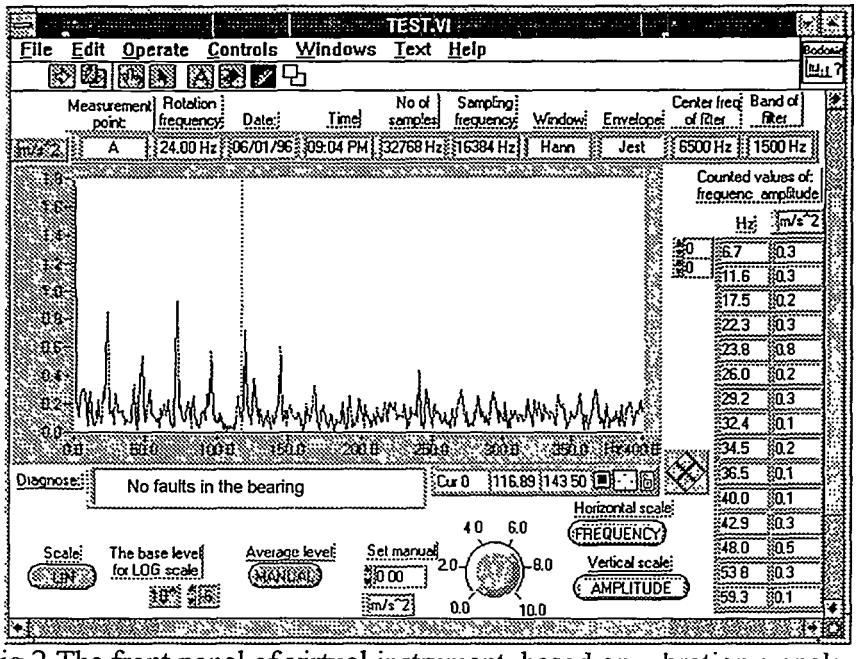


Fig 2 The front panel of virtual instrument, based on vibration signals.

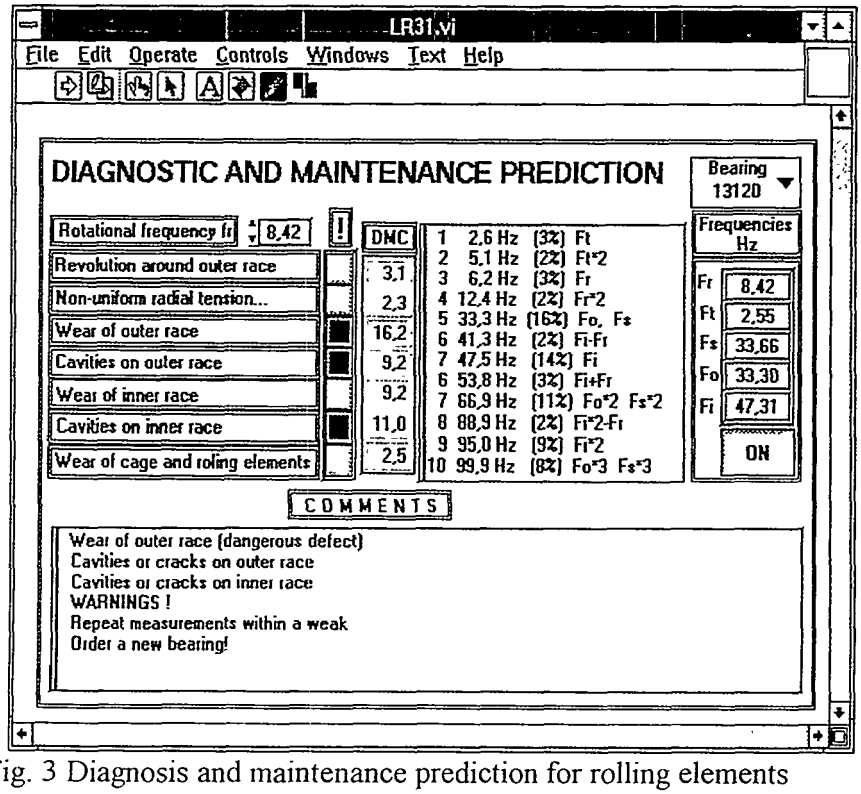


Fig. 3 Diagnosis and maintenance prediction for rolling elements bearings

- revolution around outer race
- non-uniform radial tension of the bearing.
- wear of outer race
- cavities on outer race
- wear of inner race
- cavities on inner race
- wear of cage and rolling elements

If the value of depth modulation coefficient exceeds specified values the colour signal lamp is glimmering in the field of particular defect. The green light if weak defect is detected and the red one for sever defect. The diagnosis is performed for the earlier chosen value of rotational frequency . The diagnostic system specifies also existing in the analysed spectrum frequencies which were bounded with strictly determined faults.

2.2 The analysis of electrical defects

For analysis the electrical defects we use the same stand , equipped with another sensors - current transformers. The signal is next analysed in VI - instrument for such purposes (figure 4).

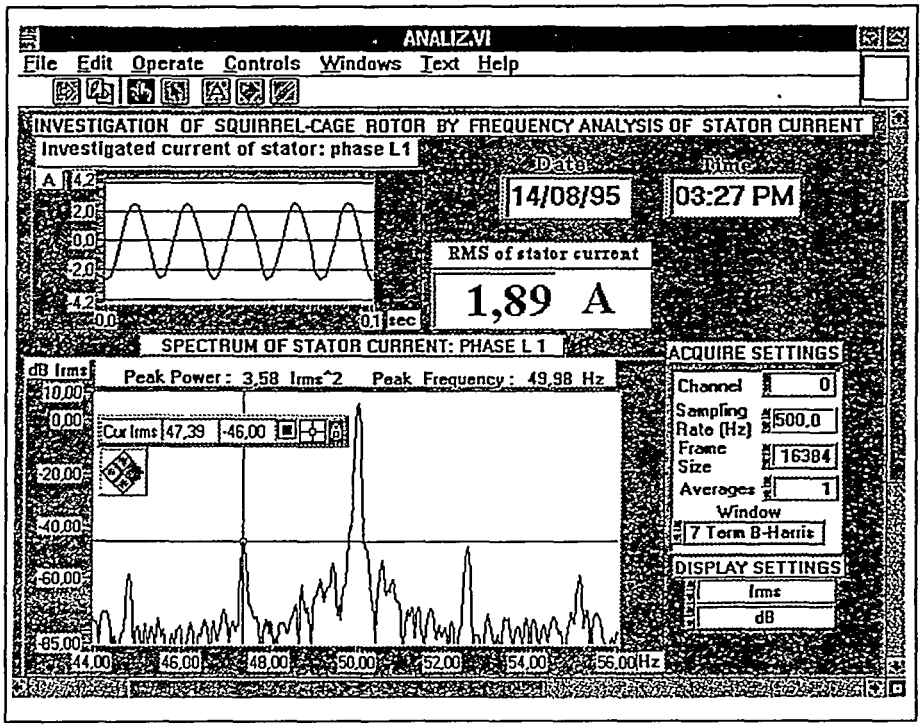


Fig 4 The front panel of virtual instrument, based on supply current measurements.

For analysis of supply current we have not yet so automatized analyse, as for mechanical investigations. The expert system for such analysis is under design.

3 PLANS

We plane to join the methods, which use different input signals (vibration, input voltage and current), in one system, for achieving more precise and reliable diagnoses, as a result of widen range of data. The increase of the range of data sources will extend the diagnostic possibilities of the system. The results of measurements will be shown under the conference.

REFERENCES

1. Barkow, N. et al. : "The certain possibilities of electrical motors monitoring", Proceedings of the International AMSE Conference "Information & System Methods for Engineering Problems" Malta, 1993 Ed. AMSE Press, Tassin la Demi Lune, (Vol. 3, pp. 69-80)
2. National Instruments, : "Programming for Instrumentation", Ed. Nat. Instr. Corp., Austin (USA), 1994
3. National Instruments, : "Instrumentation and Measurement Catalog, Ed. Nat. Instr. Corp., Austin (USA), 1995
4. Johnson, G. J. : "LabVIEW Graphical Programming" Ed. McGraw-Hill Inc., New York, 1994 (522 pp)
5. Referowski, L., Swędrowski, L., Świsulski, D., "The anbalysis of possibilities of rolling element bearings diagnosis using enveloping technique and computer controlled system", Proceedings of the 12th IMEKO World Congress "Measurement and progress", Beijing (China), Ed. International Academic Publishers, 1991
6. Barkow, A., Barkowa, N., "Machine condition monitoring by use of generalized vibration spectra methods", Proceedings of the Seminar "The new solutions of automatic systems for roll bearings and construction units", Scientific Papers of Electrical Engineering Faculty (Technical University of Gdańsk) Nr 5, 1993 (pp. 39-47)
7. Barkowa N., Referowski L., Swędrowski L. „Electrical motors monitoring - different possibilities".Proceedings of the International AMSE Conference SYS'94 „Systems analysis, Control & Design", Lyon (France), July 4-6, 1994. Ed. G. Mesnard. Tassin-laDemi-Lune: AMSE Press 1994 vol. 3, p. 89-97.

DETECTION OF MECHANICAL FAILURES IN INDUCTION MOTORS BY CURRENT SPECTRUM ANALYSIS

Sokanský Karel and Novák Petr and Biloš Jan and Labaj Jiří
Technical University Ostrava
Moraviasilesian Power Stations s.h.c.
Czech Republic

ABSTRACT

From the diagnostic point of view, an electric machine can be understood as an electromechanical system. It means that any manifestations of mechanical failures do not have to show themselves only in mechanical quantities, i.e. vibration in our case. Mechanical failures can also manifest themselves in electrical quantities, namely in electric current in our case. This statement is valid inversely too, which means that faults occurring in electric circuits can be measured through mechanical quantities.
This paper deals with measuring the current spectra of induction motors with short circuited armatures that are drives used in the industries most.

1 INTRODUCTION

Lately, measuring the current spectra of induction motors with short circuited armatures has disseminated into the whole world and hardware and software needed for these purposes have become an integral part of top analyzers. Demonstrably, current spectra may be used for analyzing broken and interrupted rotor bars, eccentricity, faults in magnetic circuits of the stator and/or the rotor, etc.
Literature exists [1] that characterizes manifestations of mechanical failures in current spectra of induction motors. Failures of the teeth of

gearboxes, bearing failures, wrong setting of the machines driven by induction motors, and others can be ranked among them.
The failures presented above can be also analyzed by means of vibration spectra, whose analysis, however, is usually more difficult. If any failure is confirmed by both the current and vibration spectra, the evaluating capability of this method is higher because the failure is corroborated by the measurement of another physical quantity.
Analysis of current, or vibration spectra consists, in the main, in the finding of spectral lines at calculated frequencies. Those frequencies are independent of or, as a rule, dependent on the engine speed. The range of frequencies occurs either at higher frequencies (race harmonics or bearing frequencies), or at lower frequencies, such as the rotation frequency, its multiples, or so-called intermediate harmonic frequencies. The majority of failures show themselves in the appearance of side-bands round the mains or speed frequencies. These lines are caused by magnetic fields of various frequencies in the air gap of the induction motor. These fields deform a time and spatial sinusoids of current of the basic frequency of the mains frequency. In the air gap, interference of magnetic fields of various frequencies occur and amplitude and frequency modulations, or their combinations, take place there as well, which is evidenced with the origin of side-bands. A cause of those phenomena can lie in the failures mentioned above.
The described deformations of the magnetic field manifest themselves in two different effects. The first one shows itself in inducing a voltage in the stator, which results in current deformations (it is modulated and also deformed by higher harmonics). The other is of the force type. The deformed field induces radial forces between the stator and the rotor that are a source of vibration and noise.

2 ANALYSIS OF CURRENT SPECTRA OF INDUCTION MOTORS

There are the following causes that deform the magnetic field in the air gap and thus develop current deformations:

- The supply voltage has not a sinusoidal course.
- The supply voltage is not symmetric in all phases.

- The spatial distribution of magnetomotoric voltage of the winding along the air gap is not sinusoidal.
- The magnetic conductivity of the air gap changes in time and space owing to the races of the stator and the rotor and eccentricity.
- Rotor asymmetry.
- Uneven saturation of parts of the magnetic circuit of the machine, especially of teeth.
- The angular speed of the rotor is not constant.

This paper deals with the influence of inconstant angular speed of the rotor.

3 THE ANGULAR SPEED OF THE ROTOR IS NOT CONSTANT

The fundamental equation of motion of the induction motor is given by the following expression:

$$\frac{J.d^2\vartheta}{p.dt^2} + M_p + M_o + M_i = 0 \qquad /1/$$

where J is the moment of inertia of the induction motor and the mechanism connected with it

p is the number of pole couples

ϑ is an angle between the axes of the stator and the rotor winding

M_p is the moment of load

M_o is the moment of friction

M_i is the moment of the motor.

In the case that all moments in the equation /1/ are constant in time, the angular speed of rotor rotation, which is given by derivation of the angle of rotation , is also constant. From the standpoint of analysis of mechanical failures in current spectra of induction motors, we are interested in changes in the moment of friction M_o in equation /1/. The changes in the moment of friction induce changes in rotor rotation; it can be said that it swings. For revolutions, the next relation can be written:

$$\omega_R = \omega_O + \sum_{\nu=0}^{\nu=\infty} A_\nu.\cos(\nu\Omega t + \varphi_\nu) \qquad /2/$$

A current amplitude can be, thanks to the inconstant angular speed of the rotor, modulated as shown in experiments below.

4 PERFORMED EXPERIMENTS AND THEIR EVALUATION

At the Department of Electric Machines and Devices of the Faculty of Electrical Engineering, many measurements were carried out, whose results are graphically presented in Figs.1,2. The goal of these measurements was to assess the influence of changes in the angular speed of the rotor in the course of time of the feeding current in current and vibration spectra. As already mentioned, a change in the angular speed can be produced by mechanical failures that result in changes in the moment of friction - see relations /1/, /2/. The experiments were conducted on a small induction motor (4AP 112 M4, 2p=4, 4kW, 380V) with a short circuited armature. The motor ran at idle speed and its loading consisted in the mounting of three journals with rubber turned by 120° on the clutch of the motor. These journals struck at each revolution against the rubber band of rectangular shape, whose one side was elastically attached to the frame. The experiment was supposed to imitate changes in the moment of friction during one revolution. When one journal was mounted, the motor braked once per revolution, etc.

From measuring the motor at idle run without strokes, a line at the mains frequency, lines of side-bands distant by the rotation frequency and the third harmonic appear in the current spectra in Fig.1A. From measuring the motor at idle run without strokes, lines of the rotation frequency and its multiples occur in the vibration spectra in Fig.1B.

At one stroke - see Figs.1C, 1D - a rapid increase (30dB) in amplitudes of side-bands distant by the rotation frequency in the current spectrum arises. It is interesting that in the current spectrum, an increase in amplitudes appears in the lines at the 100 Hz and mainly 125 Hz frequencies. These lines do not usually exist at spectral analysis. With the vibration spectrum, an increase in amplitudes in the 125 Hz, 150 Hz and 175 Hz lines can be seen, i.e. at the 5, 6, 7 multiples of the rotation frequency.

In Fig.2 there are records of the current and vibration spectra at two strokes. Amplitudes of current side-bands are rather lower than in the case of one stroke. As for vibration, differences are not apparent.

5 CONCLUSIONS

In conclusion it can be stated that manifestations of mechanical failures inducing variable moments of friction may be analyzed not only in vibration, but also in current spectra of induction motors. Moments of friction increase amplitudes of multiples of the rotation frequency in vibration spectra and amplitudes of side-bands in current spectra.

6 REFERENCES

1. R.C. Kryter, H.D. Haynes: Condition Monitoring of Machinery Using Motor Current Signature Analysis, SOUND AND VIBRATION, September 1989
2. K. Sokanský, D. Kaminský, J. Biloš, J. Schneider: Diagnostika motorů 6 kV pomocí frekvenční analýzy VŠB TU FEI, Ostrava, December 1993
3. J. Schneider, K. Sokanský, K. Chmelík: Current Spectra Analysis-Electrical and Mechanical Faults of a Machines by Inductive Motors, IV. Kolloquium Technische Diagnostik, Dresden 1996

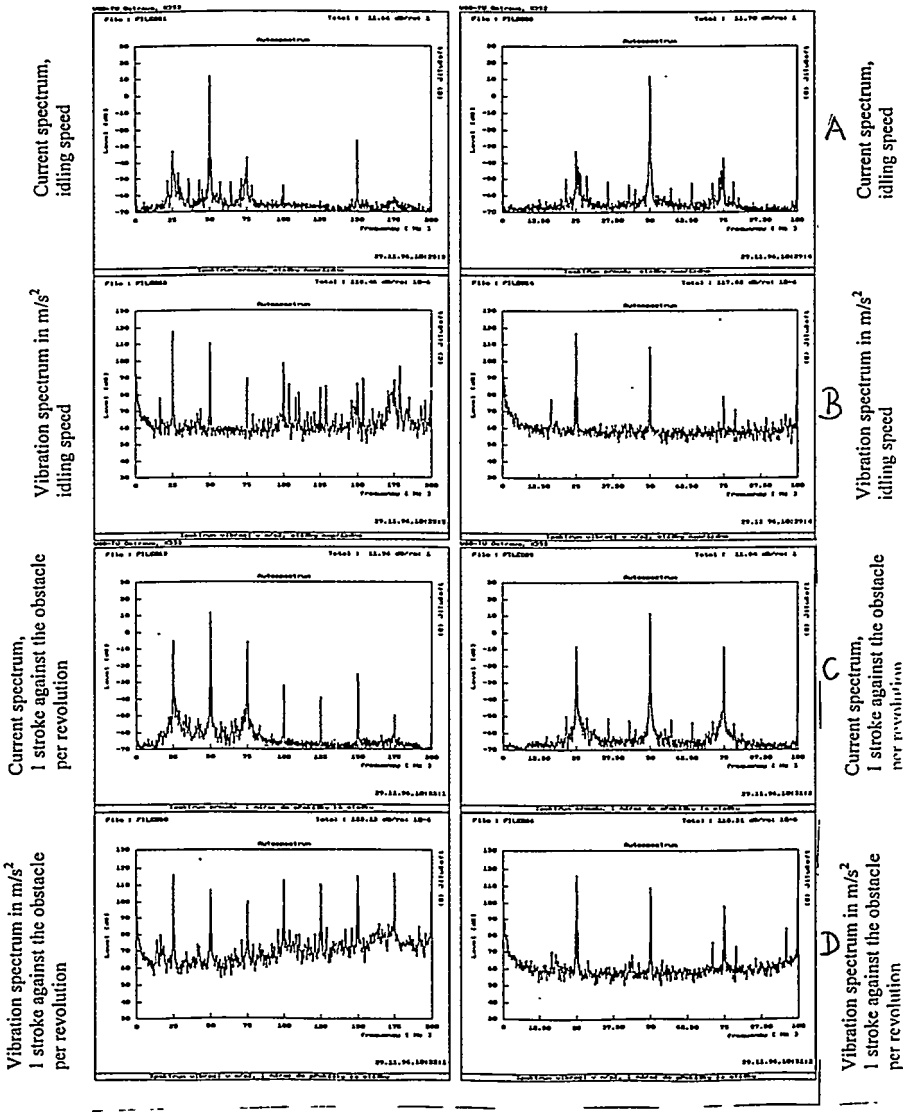


Fig. 1
Current and vibration spectra without strokes and with 1 stroke per revolution

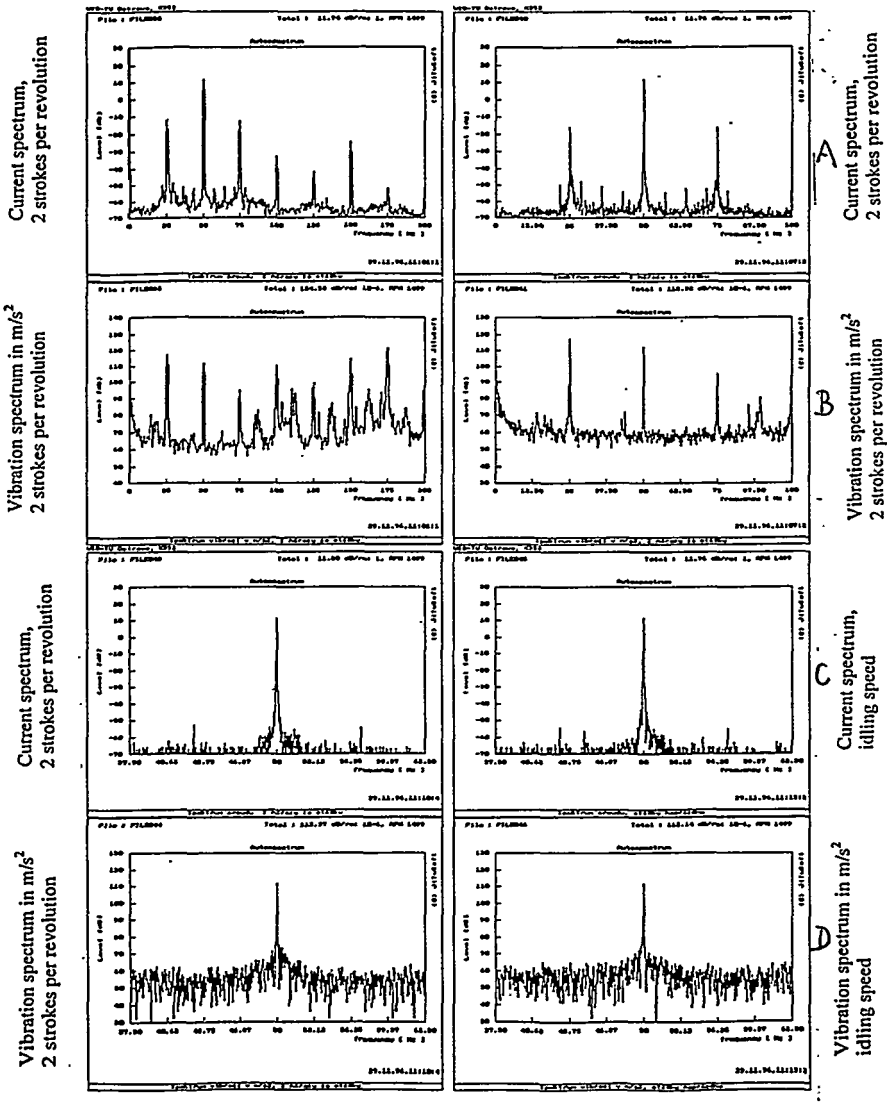


Fig. 2
Current and vibration spectra without strokes and with 2 strokes per revolution

Published by

VTT
Vuorimiehentie 5, P.O.Box 2000, FIN–02044 VTT, Finland
Phone internat. + 358 9 4561
Fax + 358 9 456 4374

Series title, number and report code of publication
VTT Symposium 171 VTT–SYMP–171

Date
May 1997

Project number
V6SU00095

Author(s)

Jantunen, Erkki (ed.)

Name of project
COMADEM

Commissioned by

Title

COMADEM '97
10th International Congress and Exhibition on Condition Monitoring and Diagnostic Engineering Management. Vol. 1

Abstract

COMADEM 97 took place at Helsinki University of Technology from 9th to 11th June 1997. This was the tenth in a series of international Congresses on Condition Monitoring and Diagnostic Engineering Management.

The large number of research contributions, altogether over one hundred papers, at the conference and their scientific and technical level is an indication of how much the importance and economical influence of condition monitoring and diagnostic engineering management is recognized world wide.

The conference proceedings (Vol. 1 and Vol. 2) contains papers on a wide variety of subjects, such as machinery and plant monitoring, instrumentation and control, advances in diagnostic engineering, expert systems, applications of neural networks, modern maintenance management, environmental pollution monitoring and industrial case studies.

Activity unit VTT Manufacturing Technology, Operational Reliability,
Metallimiehenkuja 6, P.O.Box 1702, FIN–02044 VTT, Finland

ISSN and series title
0357–9387 VTT SYMPOSIUM

ISBN
951–38–4562–1

Language
English

Class (UDC)
658.58:061.3

Keywords
condition monitoring, diagnostic engineering, meetings, expositions

Sold by VTT Information Service
P.O.Box 2000, FIN–02044 VTT, Finland
Phone internat. + 358 9 456 4404
Fax + 358 9 456 4374

Pages
553 p.

Price group
L