LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-847098

Detection of System Drift for the Health Monitoring of an X-ray CT Scientific Instrument

M. Skeate

April 3, 2023

Disclaimer

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Detection of System Drift for the Health Monitoring of an X-ray CT Scientific Instrument

SE 296 – Capstone Project

Advisors: Prof. Michael Todd (UCSD)

Dr. Harry Martz (LLNL)

Student:

Michael Skeate

Winter Quarter 2023

LLNL-TR-847098

Picture showing the Micro-Computed Tomography X-ray System (MCT) operated by the LLNL Nondestructive Characterization Institute

Table of Contents

Abbreviations

DSF – Damage sensitive feature
LAC – Linear attenuation coefficient
MCT – LLNL's Micro-computed tomography system
NCI – Nondestructive Characterization Institute at LLNL
NDE – Nondestructive Evaluation
QC – Quality control
ROI – Region of interest sensors
ROR – Region of Responsibility, the statistical tolerance interval used for test specimen quality control
RUL – Remaining useful life
SFD – Sensor Fault Detection
SHM – Structural Health Monitoring
SPC – Statistical Process Control
SPR – Statistical Pattern Recognition
SPRT – Wald Sequential Probability Ratio Test

1. Introduction

This SE296 Capstone Project technical report is being submitted as a final requirement of the UCSD Master of Science in Structural Engineering with specialization in Structural Health Monitoring (SHM) and Nondestructive Evaluation (NDE). The Capstone provides students the opportunity to apply knowledge in their technology areas towards the solution of an SHM or NDE problem. As an employee of the Lawrence Livermore National Laboratory and NDE/NCI team member, I chose to apply the SHM design paradigm taught at UCSD to improve the health monitoring of the X-ray Micro-Computed Tomography (MCT) system. I would like to thank LLNL's Dr. Harry Martz for serving as my mentor during this project and the entire LLNL MCT technical team for answering my questions and contributing to my knowledge. I would also like to thank Prof. Michael Todd for recruiting me to the UCSD NDE/SHM program and serving as my graduate advisor.

1.1 Project Background and Motivation

The LLNL Nondestructive Characterization Institute (NCI) under Dr. Harry Martz develops and supports a portfolio of six MCT X-ray systems for quantitative nondestructive evaluation and material characterization[1]. Each MCT system functions as a high precision metrological instrument and it is critical that measured values stay within quality control bounds. MCT systems contain sensors that may lose calibration or become damaged through operational use. This poses an interesting online condition monitoring problem to decide when MCT system measurements have grown unreliable.

Figure 1 shows the MCT instrument components. The X-ray tube (the large upright cylinder on the right) radiates a focused and collimated beam through a test specimen (fixed on the center carousel) which is then measured by an amorphous-silicon flat-panel X-ray detector (supported by the black panel holder on the left). The carousel is rotated and translated by

[1] The LLNL NCI website homepage is available online at https://nci.llnl.gov/home

high-precision motion stages under computer control. This allows X-ray measurements of the test sample to be acquired using a "step-and-shoot" procedure over a 360-degree angular range. A typical computed tomography scan requires 720 different angular measurements to reconstruct high quality volumetric 3D images for the test sample. Figure 2 shows the data processing steps used to transform the raw MCT measurements into digital images. The digital images can be interpreted as 2D maps of the linear attenuation coefficients (LACs) of the materials at the given scan energy.

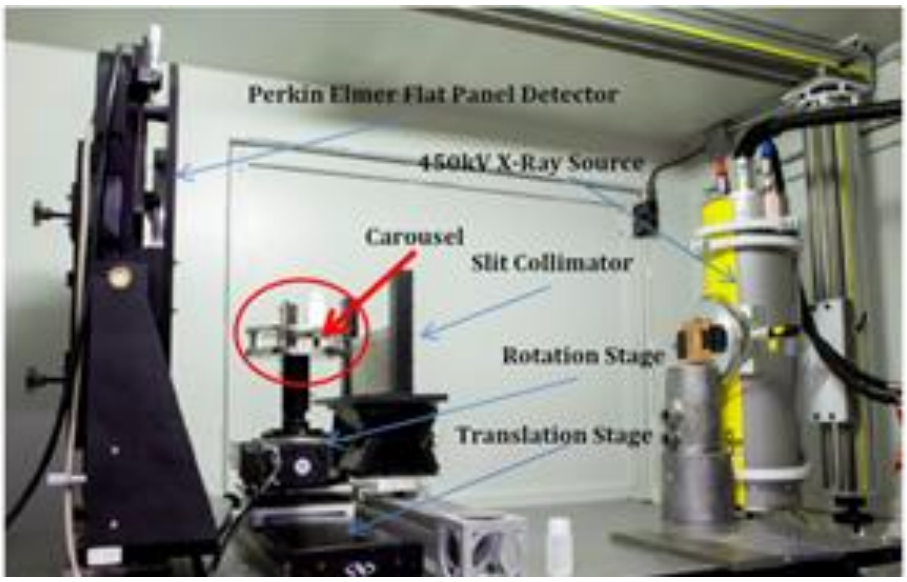


Figure 1 - The Micro-CT X-ray System is a scientific instrument for characterizing material samples. The instrument acquires dual-energy measurements which are then reconstructed using X-ray computed tomography CT algorithms to identify the material composition of a test specimen placed on the center Carousel. Reference [Martz, 2019].

The LAC value of a material sample changes slightly as a function of the X-ray energy spectra. As a result, the material composition of the test sample can be characterized by repeating the measurement and computing an image reconstruction procedure at two different X-ray energies. This measurement technique is called dual-energy computed tomography. The MCT system uses X-ray tube measurements of 100 kV and 160 kV.

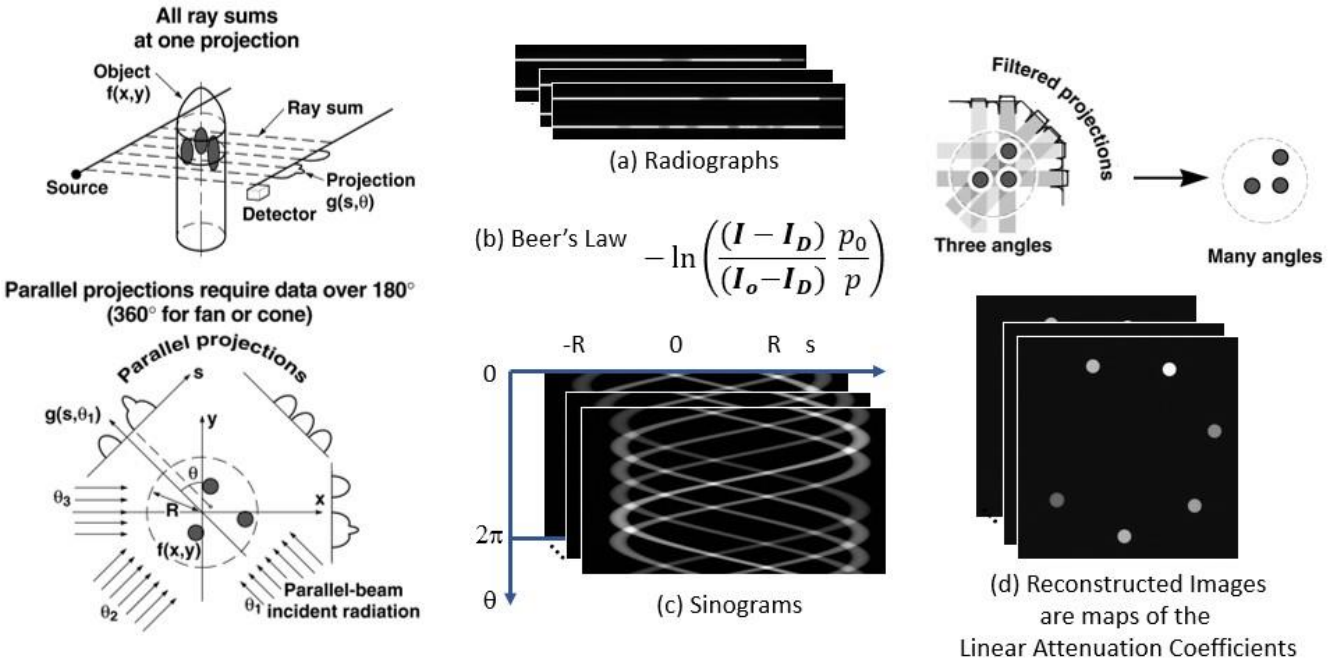


Figure 2 - MCT Data processing steps: (a) the test specimen is interrogated using X-rays at 720 angular views or "projections" resulting in raw digital radiographs; (b) data processing operations are applied to the radiographs to subtract dark current, normalize for variations in the X-ray flux, and transform using Beer's Law; (c) sinograms are created by sorting the processed radiographs; (d) CT reconstruction algorithms based on the Radon transform are applied to create digital images representing slices through the test specimen. Reference [Martz, 2019]

1.1.1 MCT Health Monitoring System

Through repeated operation, the MCT system components become damaged. The damage manifests as an observed drift in the LAC values and other statistics being monitored. Table 1 identifies the underlying causes of MCT system drift based on conversations with the MCT technical team. Damage to the X-ray source and X-ray panel detector are suspected to be the two main contributors of system drift.

Component	Damage mechanism	Other causes of variability
X-ray tube	Damage to the anode, filament or high-voltage cables or generator can alter beam flux, energy spectrum, or direction.	Temperature, Warmup procedure used by operator
X-ray panel	Radiation dose eventually damages the panel sensors causing elevated dark current readings, nonlinear response to incident radiation, or scintillator browning.	Temperature

Table 1 - Underlying causes of system drift in MCT system measurements

The MCT instrument has a built-in health monitoring system to assure that reference measurements remain within acceptable tolerance bands. During normal MCT operation, measurements of six "ground-truth" reference materials are continuously monitored at the two X-ray energy levels of 100 kV and 160 kV. Figure 3 shows a series of control charts produced by the MCT health monitoring system at the lower energy of 100 kV. Excursions beyond the two horizontal tolerance bands indicate that damage to the MCT instrument has grown severe enough to cause reference measurements to drift. Note the reference materials exhibit excursions violating the tolerance bands at different points in time. There are six additional LAC charts at the higher energy of 160 kV. This complicates the assessment of MCT system stability as the twelve chart patterns often provide conflicting indications. Another complication is that an excursion beyond the tolerance band may reverse bringing the LAC values back within the tolerance bands. An ideal SHM damage indicator should incorporate all twelve LAC measurements into a single monolithic damage assessment. Sections 2.2 and 2.3 propose several alternatives to consolidate these measurements in a principled manner.

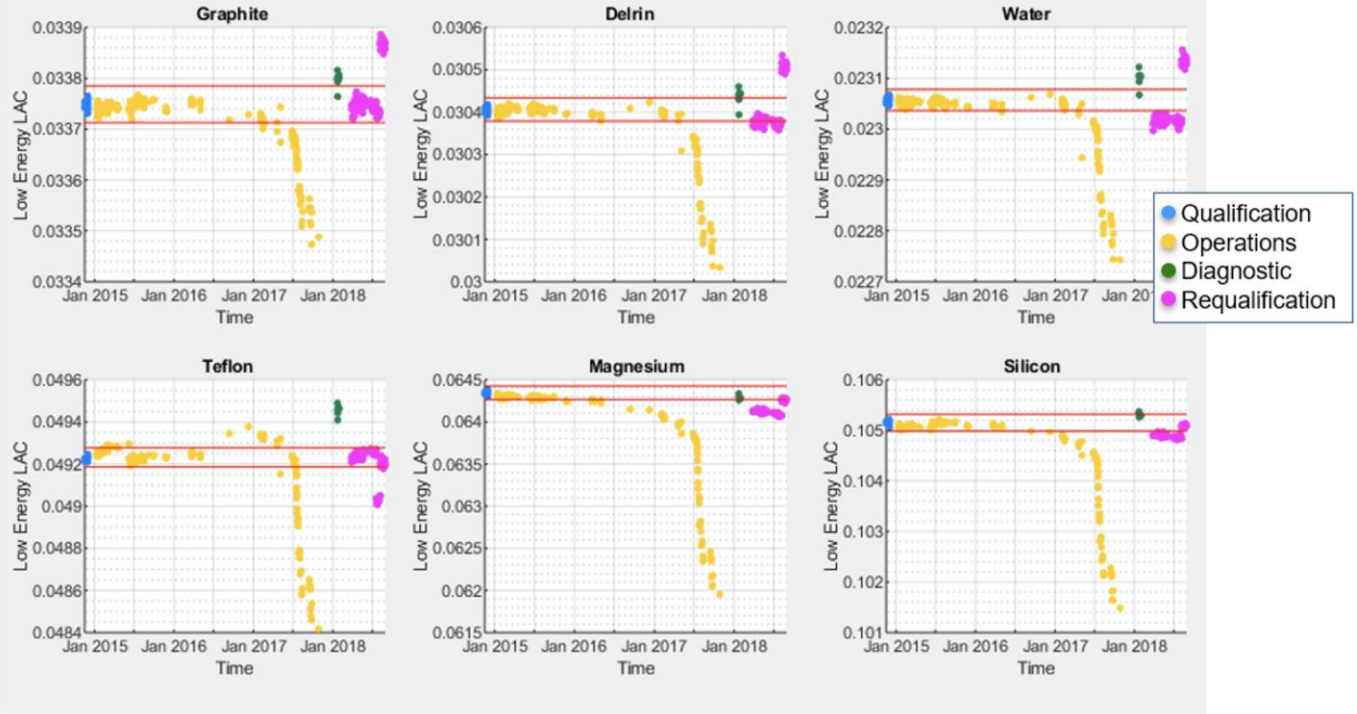


Figure 3 - Example of the control charts used to monitor the linear attenuation coefficient (LAC) values of six reference materials measured at 100 kV: graphite, Delrin®, water, Teflon®, magnesium and silicon. Not shown are six additional control charts constructed for the reference materials measured at 160 kV. System health is confirmed when measurements remain within the quality control bounds (appearing on the chart as horizontal line pairs).

Figure 4 shows the carousel designed to rotate the test sample through the X-ray beam during scanning. During normal operation, a two-slit collimator is positioned in front of the X-ray source. Figure 4 also shows the upper slit that acts as the primary MCT sensor measuring the test sample. The lower slit acts as a secondary sensor and simultaneously measures the six reference materials. The MCT health monitoring system has additional region-of-interest (ROI) sensor areas defined to collect information used for data normalization and monitoring. Statistics calculated from these ROI areas are monitored as a secondary indicator of MCT system health. The secondary sensors are discussed further in Section 2.1.

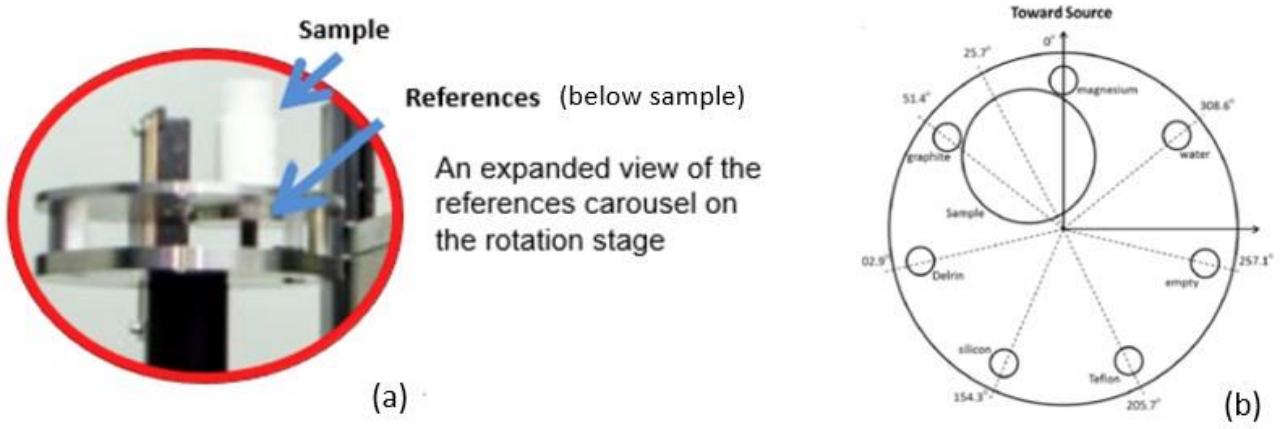


Figure 4 - (a) The carousel holds the test sample during scanning with the reference materials located directly underneath. (b) Top-down view of the carousel showing the large cylindrical test specimen and six smaller reference materials.

Due to collimation, only a fraction of the X-ray panel is exposed to radiation during operation. A simple preventative maintenance procedure called a "panel move" can be used to restore the system back to health. Specifically, a 3/8" spacer is inserted at the panel base causing the illuminated regions of the panel to be shifted. This allows the damaged sensor region to be abandoned and future measurements to be relocated to an undamaged region.

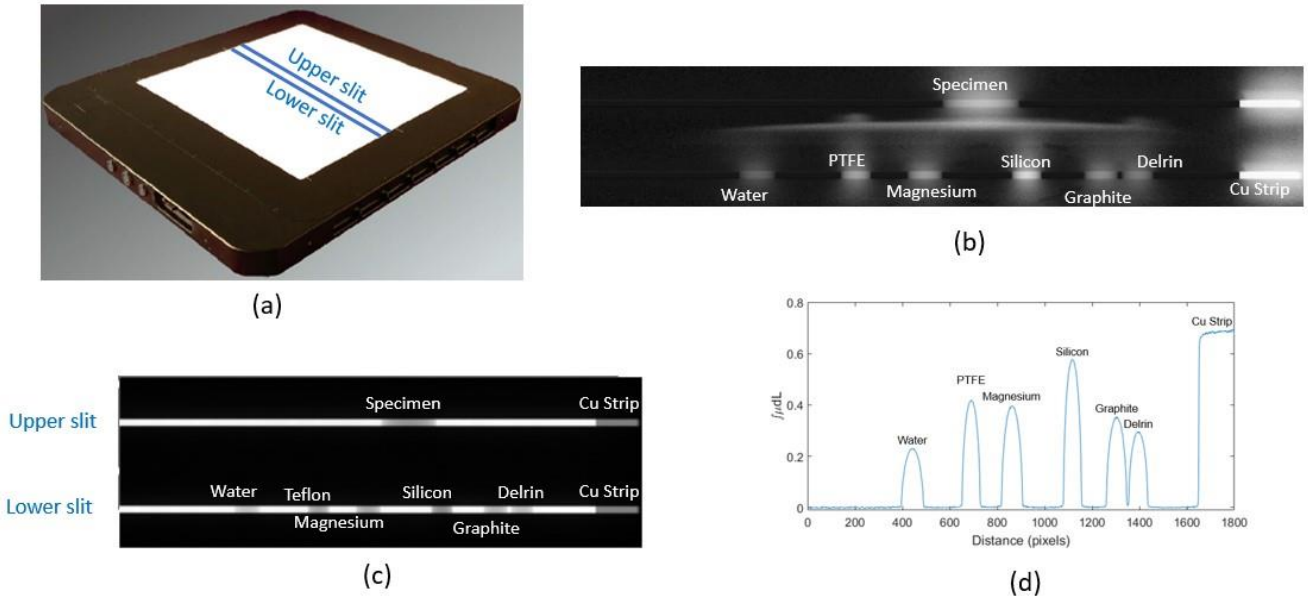


Figure 5 - (a) X-ray panel detector showing approximate location of upper and lower slits; (b) X-ray radiograph of carousel and samples without collimator in place; (c) X-ray radiograph with two-slit collimator in place; (d) Plot of lineout taken along lower slit.

The damage occurring in the X-ray panel detector plays an important role in the behavior of the MCT health monitoring data. This will be discussed further in Section 2.1 Data Acquisition.

1.1.2 Study Goals

This technical report provides an appraisal of the current MCT health monitoring system and explores how SHM can be applied to make improvements. In support of my SE296 project, LLNL NCI provided four years of MCT system health monitoring data. This report documents my findings from working with this dataset as well as my exploratory conversations with MCT technical team members. Table 2 defines the specific study goals adopted at the start of the project.

Study Goal	**Description**
Detect MCT system damage	Determine if SHM methods can be used to detect damage when MCT measurements are observed to drift out of quality control bounds
Locate the damage	Determine how SHM methods might pinpoint the cause of system drift by locating damage to one of the MCT system components (e.g., X-ray detector or source, or other system components)
Predict remaining useful life RUL for condition-based maintenance	Explore how SHM might use MCT health monitoring data to predict RUL. The ultimate objective is to devise a decision support tool to prompt the scheduling of MCT system preventative maintenance.

Table 2 - SE296 Project Study Goals

1.2 The Structural Health Monitoring Design Paradigm

Structural health monitoring (SHM) originated as a term referring to damage detection strategies using vibration measurements from large sensor arrays deployed to planes, bridges, buildings, and wind turbines. During the 1990's, Charles Farrar and Ken Worden began demonstrating a data-driven approach to SHM system design by integrating concepts

from statistics, signal processing, structural dynamics and pattern recognition. Since that time, the Farrar-Worden SHM design paradigm has been adapted for condition-based monitoring of a much wider class of man-made systems, structures and scientific equipment. The paradigm defines structural health monitoring as "the process of developing an automated and online damage assessment capability for all types of engineered systems" [Farrar-Worden, 2021]. This broader definition is adopted for the purposes of this report which uses SHM, health monitoring, and condition-based monitoring as synonymous terms.

While often employing the interrogation methods of nondestructive evaluation (NDE), SHM differs from NDE in that the structure typically remains in service throughout the inspection. NDE methods provide a "snapshot" indication of whether or not damage is present; SHM methods by comparison provide ongoing time-series measurements monitoring the progression of damage in man-made structures. However, the distinction between SHM and NDE is sometimes fuzzy as in the case of in situ inspection methods used for additive manufacturing quality control.

There are other important distinctions between SHM and NDE. SHM typically generates very large datasets due to continuous and longstanding measurement. This has caused SHM to prioritize the development of automated data compression, analysis and interpretation methods. Second, SHM systems are characteristically made "in the field" necessitating solutions that can address a wide variation in environmental and operating conditions (EOC's). For example, daily or seasonal temperature variation may create a serial correlation in measured data that must be corrected to avoid being misinterpreted as evidence for damage.

Figure 6 summarizes the SHM design paradigm employed during the SE296 project and referenced throughout this report. The paradigm was pioneered by Farrar and Worden and further developed over the last twenty years through a partnership between the UCSD Jacobs School of Engineering, the Engineering Institute at Los Alamos National Laboratory, and the University of Sheffield.

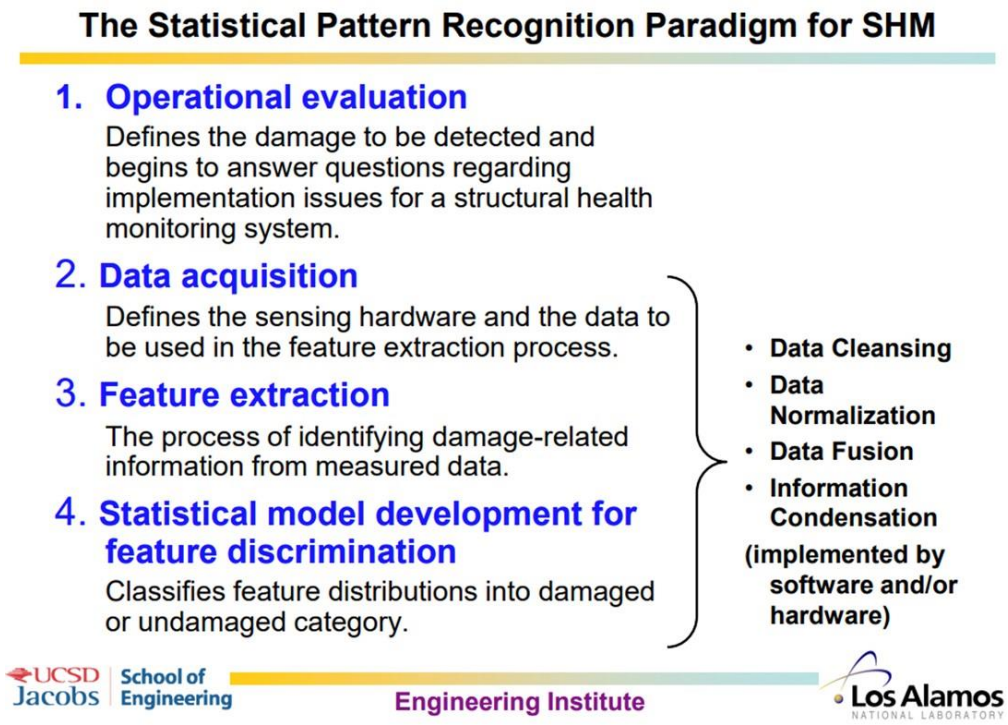


Figure 6 – The four-step paradigm for architecting health monitoring solutions [Farrar, 2019]

Some of the key paradigm teaching points include:

- *SHM is best approached as an engineering design problem in statistical pattern recognition using data-driven methods and machine learning*. The resulting statistical models may be further improved by whatever physics-based knowledge is available.
- Some level of damage is always present in any man-made structure or material. What matters is when the level of damage reaches a critical level where the equipment or structure being monitored no longer performs at an acceptable level. Therefore, *damage must always be defined relative to some normal operation state or condition*.
- *Sensors are not capable of measuring damage directly*. It is always necessary to extract damage sensitive features from the measured data using signal processing and statistical classification. The extracted features are used subsequently by statistical inference models constructed for the purpose of detecting damage, locating damage, classifying damage type, assessing damage severity, and ultimately estimating remaining useful life (RUL).
- A further codification of the general principles accepted by the SHM community is known as the Fundamental Axioms of Structural Health Monitoring (See Appendix A).

The earliest SHM applications originated in the structural engineering community using vibration monitoring to detect incipient damage in structures, vehicles and rotating machinery. However, I will demonstrate that the paradigm adapts well to the health monitoring of the MCT X-ray system. This assertion is a key contribution of this report and applies to the health monitoring of scientific metrology equipment more generally.

I will also compare the classical statistical inference approach currently employed by the MCT technical team, with the more general SHM paradigm, as a secondary research contribution. Throughout the report, I support the thesis that the statistical pattern recognition paradigm offers a much richer framework for attaining the MCT system health monitoring goals.

1.2.1 Damage identification levels

Table 3 defines five damage identification levels according to the SHM design perspective. Each level builds progressively on the previous levels. This means that sensors must be selected, damage sensitive features constructed, and classification models must be built at each previous level before ascending to the next. SHM systems at the higher damage levels are more difficult to construct and require the availability of damage state data. Damage state data is needed to train supervised learning algorithms and must be provided either experimentally or through simulations paired with a known physics-model.

Damage Level	Damage question	Applicable Machine Learning Methods and Data Requirement
1 Detection	Is damage present?	Novelty detection methods can be used when damage state data is unavailable
2 Location	Where is the damage present?	Relies on detection and may be either unsupervised (if novelty detection is used) or Supervised (if classification is used)
3 Classification	What type of damage is present?	Damage state data is required to support supervised learning
4 Assessment	What is the damage severity or extent?	Relies on Classification and definition of a monotonically increasing damage indicator.
5 Prognosis	What is remaining useful life?	Relies on Assessment and requires ability to track speed of increasing damage

Table 3 - Damage Identification Levels. Levels 3-5 require pattern classification which requires damage data to be available. Level 2 may be achieved in an unsupervised manner if the sensor is local to the detected damage. References: [Rytter, 1993], [Farrar, p296]

The Damage Identification Levels provide a fundamental bound on what health monitoring goals are achievable given the data available. My SE296 project necessarily prioritized "Damage Level 1 - Detection" due to the lack of damage state information in the MCT health monitoring dataset. This level is unique from the higher levels of damage identification in that it can be undertaken using unsupervised learning [Farrar, p321].

However, developing an optimal damage detector is only the first step towards building an SHM system capable of achieving the ultimate goal of preventative maintenance. Another contribution of this technical report is to propose Future Work for progressing through the higher levels of damage characterization. Recommendations are provided for the damage state data that must be acquired as a prerequisite.

1.3 Operational Evaluation (SHM Step One)

Operational Evaluation represents the very first step in SHM system development. It corresponds to the requirements gathering phase of model-based systems engineering. The SHM project to be undertaken should be examined in terms of the costs, benefits, and implementation constraints. However, there are two aspects of Operational Evaluation making it unique relative to other engineering lifecycles: (a) the emphasis on defining the damage to be detected and the unique damage sensitive features to be exploited; and (b) the identification of Environmental and Operational Conditions (EOC's) that can confound sensor readings leading to false indications of damage.

Farrar and Worden describe the Operational Evaluation step as an exercise in answering four types of questions:

- Life-safety and/or economic justification for performing SHM;
- Questions about the type of damage to be detected;
- Environmental conditions (temperature, moisture, radiation fields) and operational conditions (changes to equipment settings or configurations, loading);
- Constraints on sensors and other data acquisition equipment.

Answers to these four types of questions are compiled below based on conversations with the MCT technical team. The issues raised during Operational Evaluation inform and guide the SHM design steps discussed in later sections of this report.

- Performance measures: The purpose of an SHM system is to ensure that the structure being monitored maintains an adequate level of operational performance. As a metrological scientific instrument, the two most important measures of the MCT system are measurement stability and accuracy.
- Economic justification: Recalibration requires a minimum of thirty MCT scans which can take as long as two months to acquire. This unproductive use of scanner time is both costly and disruptive to programs depending on timely measurements. It also sometimes happens that the cause of MCT system instability is misdiagnosed causing further delays. For example, it is critical to know whether an observed system drift is being caused by

damage to the X-ray source or panel so that maintenance targets the malfunctioning component.

- Life safety justification: Maintaining MCT data quality is important to national security programs. This justification could be leveraged for R&D funding and retrofit costs leading to a more capable health monitoring solution.
- Damage definition and priority: System drift is the most readily observed evidence of damage. It is also an SHM priority because system drift indicates the loss of measurement stability and accuracy. As discussed in Section 1.1.1, the MCT technical team monitors system drift using continuous measurements of the LAC coefficient values for six material references. Control charts show the tendency for the LAC coefficients to be become unstable after a certain number of scans. Constructing an optimal detector of system drift is therefore the first SHM step, “Damage Level 1 – Detection”, and the focus of this SE296 project.
- Damage location: As mentioned above, locating damage to the X-ray source or panel is critical to preventative maintenance “Damage Level 2 – Location”.
- Estimation of remaining useful life: “Damage Level 5 – Prognosis”, represents the ultimate MCT system goal. Preventative maintenance would maintain measurement fidelity and help mitigate disruptions to program schedules.
- Environmental conditions: Radiation has a dual role for the MCT system. It functions as the interrogating signal for MCT measurements. However, radiation may also be treated an environmental condition because it damages the X-ray panel and eventually necessitates the "panel move" procedure discussed in Section 1.1. The effects of radiation are visible even during normal operating periods. As discussed later, amorphous silicon X-ray panels are subject to a thermal annealing process that reverses the effects of radiation damage over time. The damage and its reversal cause measurements to drift in an oscillatory manner that can be characterized as an autocorrelation. The autocorrelation confounds statistical measures of system drift and therefore must be corrected. Section 3.3 applies an SHM data processing technique called Data Normalization which is often used to correct the confounding effects of environmental conditions.
- Secondary environmental conditions: As with most SHM systems, temperature fluctuation can introduce variability into MCT system measurements. Temperature can

alter the flux and spectral output of the X-ray tube. Temperature also influences the rate of annealing in the X-ray panel which is a suspected cause of measurement drift. High humidity has been raised as a concern at some MCT system locations as a potential contributor to X-ray tube failure.

- Operator changes: Measurement variability can be accidentally introduced by the radiography technicians operating the MCT system. Measurement outliers can occur if the technician changes the standard operating procedures. Examples of procedural changes known to have occurred in the past include: (a) changes to the X-ray tube warmup procedure; (b) changes to the data acquisition settings; (c) alignment changes introduced by a failure to properly position the carousel or X-ray panel.
- Operating loads: The MCT system can be subject to "undocumented" scans that may add equipment operating hours not logged in the scan records. Technicians perform warmup, alignment and test procedures that can add undocumented time. Also, the MCT system is occasionally loaned out for nonprogrammatic use.
- Data acquisition limitations: Radiation requires shielding or rad-tolerance for any added sensors or data acquisition equipment. However, the MCT system operates in a laboratory environment so exposure to weather or other uncontrolled environmental conditions is not a concern. One disadvantage is that some of the MCT systems are located in facilities that do not permit wireless equipment. One advantage is that MCT data acquisition occurs slowly relative to many SHM systems so data may be analyzed and interpreted in an offline manner.

1.3.1 Damage to be detected

Operational Evaluation begins by identifying the damage to be detected and monitored by the SHM system. As discussed in Section 1.1.1, the MCT technical team believes there are two primary damage mechanisms affecting MCT measurements: (a) damage to the X-ray tube through operational use; and (b) damage to the X-ray detector caused by radiation dose. Section 2.1.3 will revisit the concern of X-ray tube damage as a potential topic for future study. As will now be discussed, damage to the X-ray detector is believed to occur on

a much shorter timeframe and also to be more dominantly manifest in the MCT measurement data. It is therefore adopted as a focus of the current SE296 study.

Early developers of amorphous silicon (a-Si:H) X-ray panel detectors performed experimental studies to characterize the effects of radiation damage [Boudry, 1996]. Each pixel in the 2048x2048 panel detector contains a two-device sensor comprised of a photodiode coupled to a field-effect transistor (FET). These semiconductor devices exhibit a leakage current (or "dark current") that varies according to the presence of impurities, lattice defects, or ambient temperature. While the photodiode was empirically determined to be relatively immune from radiation [Antonuk 1990], it was determined that radiation dose increases the leakage current of the FET [Boudry and Antonuk, 1996]. The leakage current increase can be explained by the tendency of radiation to create dislocation defects within the FET semiconducting regions. The leakage current increase was also associated with increased measurement noise. Over time, thermal energy reverses the dislocation defects resulting in an "annealing" effect which accelerates with increasing temperatures. Annealing occurring at room temperature was observed to reduce the leakage current between 10-50% in the first day with a logarithmic decrease in subsequent days [Boudry, 1996]. Although X-ray panel manufacturers have engineered ways to mitigate this effect, there is evidence of the damage-annealing cycle becoming manifest as an autocorrelation in the MCT measurement data. This will be revisited in Section 3.3 on data normalization.

1.4 Literature Review

Computed tomography equipment manufacturers have long offered procedures for system-level performance measurement (e.g., ASTM standard E1695), but the new trend is to monitor and diagnose system health at the component level to predict against imminent failures. Medical CT vendors are starting to incorporate health monitoring as a standard system feature. GE HealthCare advertises[2] Tube Watch™ using AI and Digital Twin technology to monitor imminent failures in X-ray tubes. Industrial CT is considered a critical enabling technology for advanced manufacturing. The European Union's Industry 4.0 initiative is funding the "xCTing" MSCA ITN research program to develop condition

[2] https://www.gehealthcare.com/products/tube-watch

monitoring of in-line CT equipment[3]. The xCTing website promotes CT as "the only known technology that can certify nondestructively the quality of internal complex structures, such as those produced by additive manufacturing or found in assemblies".

There are interesting cases of health monitoring systems designed to detect damage and anomalous operating behavior in scientific instruments and equipment. Farrar discusses an SHM system using vibration monitoring to detect damage in a telescope drive mechanism [Farrar, 2019]. MAINTLET is an NSF-funded project at the University of Illinois seeking to develop health monitoring solutions for scientific instruments located in campus laboratories [Nahrstedt]. A sensor network of vibration monitors, water flow meters, and contact temperature sensors streams data that is assimilated into digital twin simulations to predict potential instrument failures. An engineer designed a health monitoring system for radiological equipment used at the Department of Energy Y-12 uranium enrichment facility in Tennessee [Harrison, 2004]. The system employs a fault detector based on the sequential probability ratio test SPRT. Section 3.1 discusses results obtained by applying the SPRT to MCT health monitoring data.

Scientific instruments are carefully calibrated and controlled to produce accurate and consistent measurements. The loss of calibration or variability in measurements can be treated as a type of damage known as system drift. Methods for detecting system drift originated at Bell Laboratories in the 1920's through the work of quality engineer Walter Shewhart, and have matured into the field of statistical process control SPC. Venkatasubranmanian [2003] provides a comprehensive literature survey of SPC methods including a discussion of the Shewhart control charts employed by the MCT health monitoring system. It warns that the use of univariate control charts can lead to misleading results when the parameters being monitored are correlated. This point will be revisited in Section 2.2.2. There are many more publications in the SPC literature that are relevant to the MCT health monitoring system. *Introduction to statistical quality control* is considered a standard reference by SPC practitioners [Douglas, 2020].

More recent work in the process control industry directly converges with SHM data analysis methods. The TÜV SÜD National Engineering Laboratory developed a data science

[3] https://xcting-itn.eu/

framework for the condition-based monitoring CBM of the flowmeters used inside oil pipelines [Lindsay, 2022]. The TÜV SÜD framework can predict component failure, detect calibration drift, and reduce unscheduled downtime. While the framework closely mirrors the data-driven methods of the Farrar-Worden SHM paradigm, it also emphasizes the importance of exploratory data analysis (data mining) in the search for data sensitive features. Another emphasis of the framework is the three-stage progression of SHM design over time:

- Stage 1 - Detect anomalies;
- Stage 2 - Use detected anomalies as input to classify specific fault conditions;
- Stage 3 - Quantify the effects of each fault condition on the overall system measurement uncertainty.

The close tie between structural health monitoring and uncertainty quantification is of great importance to the MCT health monitoring system.

Sensor fault detection (SFD) has been widely studied for over 50 years being critical to the health monitoring of aeronautics, space vehicle, and automotive instrumentation. SFD provides methods for detecting sensor drift and other anomalous measurements. Many of these methods employ system identification algorithms. De Silva [2020] proposes an SFD architecture that combines the Kalman filter with machine learning methods. This approach will be examined in Section 2.3.

A final body of knowledge that was surveyed is the related field of measurement science (metrology). Measurement science provides statistical methods for quantifying equipment precision, accuracy and uncertainty propagation. As discussed in Section 2.2.1, the MCT technical team employs calibration procedures for statistical tolerance intervals as published in the *NIST Engineering Statistics Handbook.*[4]

[4] Available online at https://www.itl.nist.gov/div898/handbook/

1.5 Contributions

This section summarizes the contributions of this technical project and final report to both the MCT technical team and the SHM engineering literature.

Contributions to the MCT Technical Team

In February of 2023, UCSD student Michael Skeate gave a tutorial presentation to the MCT technical team on the Structural Health Monitoring paradigm and recommendations were provided for using the paradigm to improve the MCT health monitoring system. This technical report provides additional information on the topics presented including:

- An assessment of the statistical procedures currently used for MCT health monitoring including statistical control charts and tolerance intervals;
- An evaluation of the method used to specify the ROR and QC tolerance limits and reasons these limits should be widened: (a) reference given to published k-factor tables that account for the presence of autocorrelation in MCT data [Knoth]; (b) reference given to construct multivariate statistical tolerance intervals [Polhemus];
- Spotlighting the damaging aspects of system drift including the invalidation of the MCT tolerance intervals, and thereby motivating the need for rapid drift detection;
- Integration of an optimal detection algorithm employing sequential hypothesis testing into the statistical control charts currently used for MCT health monitoring;
- Demonstration of the efficacy of autoregressive models to mitigate the serial correlation present in the MCT data measurements being monitored, thereby reducing the false alarm rate during drift detection;
- MATLAB codes providing reference implementations of the sequential probability ratio test and autoregressive model correction;
- Roadmap for future work intended to evolve the MCT health monitoring system from Damage Level One to the higher levels in the Rytter-Farrar-Worden damage hierarchy.

Contributions to the SHM Literature

- Demonstration of applying the SHM paradigm to the health monitoring of scientific instrumentation in general and X-ray computed tomography equipment in particular;
- Spotlighting the important role that SHM will play in developing online condition monitoring for scientific equipment, including X-ray equipment used for in situ inspection and process control for Industry 4.0 and lot-size-one advanced manufacturing.

2 Methods

Section 1 posed the question motivating this SE296 project, "how might SHM methods be applied to improve the current MCT health monitoring system?" Section 2 discusses topics addressing this question including theory, models, datasets, key concepts, design methods, and statistical analysis methods. Section 3 presents the project technical results.

2.1 Data Acquisition (SHM Step Two)

This section discusses step two in the SHM paradigm, Data Acquisition. During this step, the SHM designer gives attention to aspects of measurement engineering: (a) method of damage interrogation and sensor excitation; (b) sensor type, location and number; (c) equipment used for data acquisition; (d) the data collection rate; and (e) required signal processing. The resulting SHM system must provide the data needed to achieve the damage detection goals established during Operational Evaluation.

Section 2.1.1 comments on two fundamentally different data acquisition design strategies. Section 2.1.2 discusses the MCT system health monitoring dataset provided by LLNL for the benefit of this SE296 project. Section 2.1.3 appraises the quality and completeness of the available data and makes recommendations for additional data needed for future studies.

2.1.1 SHM Instrumentation Strategies

Data acquisition lies at the foundation of SHM system design [Farrar, p54]. The SHM system designer must employ sensors that are sensitive to the damage of concern with a high degree of correlation. This is not to say that the sensors will measure damage directly as

that would violate the Fundamental Axioms (See Appendix A). It is rather that the chosen sensors and resulting measurements must enable the extraction of damage sensitive features through the corrective lens of signal processing, mitigation of environmental and operational variability (the EOC's), and interpretation by a sufficiently powerful statistical model. Given all these design factors, it follows that ideally the choice of SHM sensors should be made as a fundamental design concern of the structure to be inspected. This integrated approach to SHM or NDE architecture is often referred to as "design for inspection" [Argyll Ruane, 2023].

As outlined by Farrar p54, the retrofit strategy is by far the most common approach for most SHM projects. This is also the approach taken in this study given the limited time available to complete the project. Table 4 contrasts the "Retrofit" instrumentation strategy with the more comprehensive "Design for Inspection" approach that would make optimal sensor choices based on SHM first principles. It is recommended that future studies move towards the Design for Inspection approach.

Design Aspect	Retrofit Strategy	"Design for Inspection" Strategy
Sensors employed	Use previously available sensors and data acquisition equipment	Chosen sensors on the basis of sufficient sensitivity and correlation to the postulated damage mechanisms
Sensor location	Distributed in an ad hoc manner to provide a general sampling	Placement decided on the basis of analysis, experiments, or past experience
Damage-sensitive data features	Constructed "after the fact" using archived sensor data, trial and error, and ad hoc algorithms	Engineered on the basis of numerical simulations or experimental validation for the postulated damage mechanisms
Treatment of environmental and operation conditions	No measurement of the parameters necessary to correct the effects of EOC's	Addition of secondary sensors to quantify changing EOC's
Excitation	Relies on measurements of the ambient operational environment	Active sensing employed when it can be exploited for optimal damage detection
Data coverage	Lack of damage data relegates retrofitted systems to Level One Damage Identification as the SHM system must rely on unsupervised learning methods (novelty detection)	Damage data needed for supervised learning methods is provided by numerical simulations or experiment. This opens the door to higher levels of damage identification.

Table 4 - Comparison of SHM Instrumentation Strategies. Reference [Farrar and Worden, 2012]

In spite of its many limitations, Farrar notes that the Retrofit approach can still be made effective at damage detection especially when a good historical database of measured system responses and associated damage states is available. Such is the case for this study which benefitted from the four-year compilation of MCT health monitoring data discussed next.

2.1.2 MCT Health Monitoring Dataset

Table 5 describes the MCT dataset provided for this study. More detailed information on the dataset may be found in [Martz, 2019]. The dataset includes almost four years of MCT health monitoring data acquired from a single MCT scanner.

Data record	Description
Manifest	The dataset contains a total of 440 complete CT scan sets acquired over 3 years and 9.5 months (starting 11/17/2014 through 8/30/2018. Each scan set includes radiograph and recon images at both low 100 kV and high 160 kV energies. Including metadata files, the full dataset contains 1824 files stored in HDF5 format totaling 931 GBytes.
Metadata	A "System Operation Data" spreadsheet records over 40 different data acquisition settings for all scans. These settings are held constant for the most part but will have slight variations for some settings (e.g., alignment parameters, technician acting as system operator). A "System Changes" spreadsheet contains a dated record of all changes made to the MCT system including preventative maintenance, equipment repairs, and software updates.
LAC values	This spreadsheet contains a history of the LAC values of the six reference materials for all 440 CT scans at both the low 100 kV and high 160 kV X-ray energies. The spreadsheet also contains the MCT calibration results on three different dates 11/17/2014, 7/13/2015 and 3/28/2018 including the tolerance limits QC bounds).
Recon files	There are 440x2 (low and high energy) recon files in HDF5 format. The content of these files is depicted in Figure 8 and is used to reconstruct the LAC values.
Radiograph files	There are 440x2 (low and high energy) radiograph files in HDF5 format. The content of these files is depicted in Table 6 and are used to calculate the ROI statistics used as a secondary indicator of MCT system health.

Table 5 – Data records in the MCT Health Monitoring Database provided for the study

All MCT data measurements originate as radiographic digital images acquired by the X-ray panel detector. The radiographs are digital 16-bit grayscale images of dimension 2048 pixels x 2048 pixels. Table 5 depicts the data processing steps converting the radiographs into the LAC values which are monitored as the primary indicators of MCT system health.

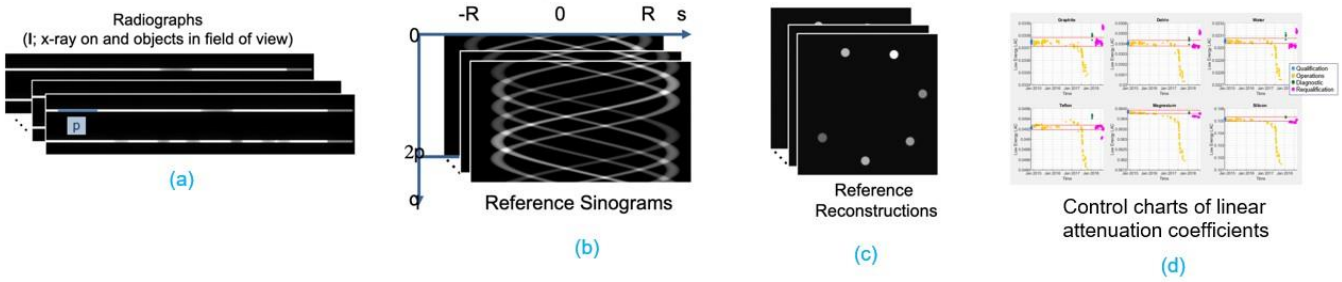


Figure 7 - Data processing steps for MCT Health Monitoring: (a) Measurements originate as radiographic digital images acquired by the X-ray panel detector; (b) the radiographs are spatially-arranged into digital images called sinograms; (c) CT reconstruction algorithms reconstruct the sinograms input images into images of the six reference materials; (d) the reconstructed images are decomposed into the linear coefficient values that are monitored in the form of control charts.

For a typical SHM project, an array of sensors must be added to the structure being monitored. The MCT system is unique in that the X-ray panel detector acts as a high-spatial resolution sensor array. An important benefit afforded by the X-ray panel detector is the configurability of many different sensor regions. Figure 8 shows examples of the configured region-of-interest ROI sub images including the postage stamp and copper strip regions that are located inside the illuminated upper and lower slits discussed in Section 1.1.1.

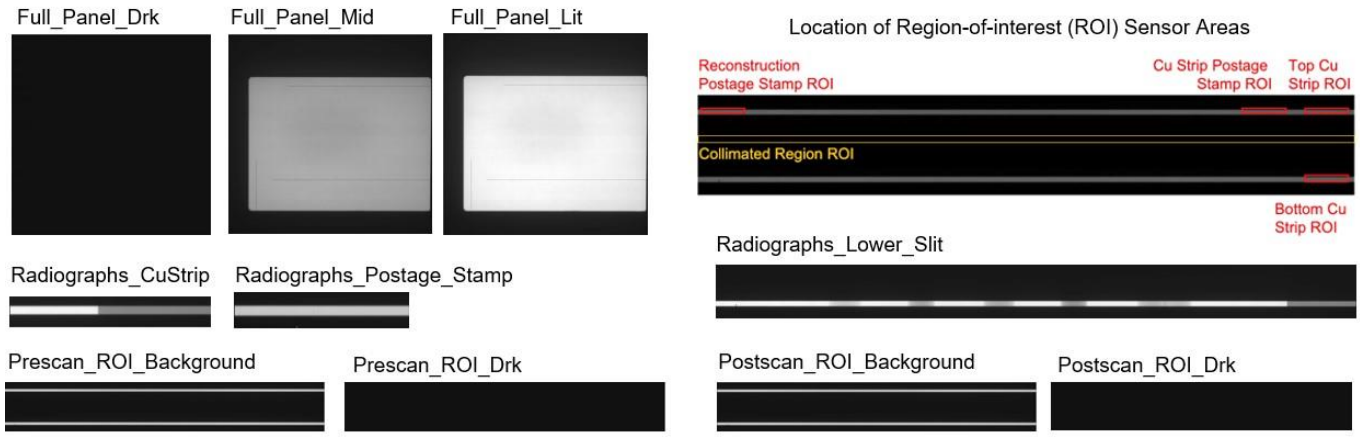


Figure 8 - Region-of-interest (ROI) data areas extracted from the X-ray panel detector. The ROI regions function as secondary sensor and are used for data normalization and health monitoring.

Table 6 discusses the different properties and dimensions of the ROI data measurements. These region-of-interest ROI sensor regions are used for data normalization and provide additional indicators of MCT system health. The ROI sub images therefore function like the secondary sensor commonly used by SHM designers to correct for variable environmental and operation conditions.

Name	Dim	Size	Description
Full_Panel_Drk	2	2048x2048	Full-panel image taken with x-ray source off
Full_Panel_Lit	2	2048x2048	Full-panel image taken at full current
Full_Panel_Mid	2	2048x2048	Full-panel image taken at 2/3 current
Postscan_ROI_Background	2	275x1808	Post-scan ROI image taken at full current
Postscan_ROI_Drk	2	275x1808	Post-scan ROI image with x-ray source off
Prescan_ROI_Background	2	275x1808	Pre-scan ROI image taken at full current
Preschan_ROI_Drk	2	275x1808	Pre-scan ROI image with x-ray source off
Radiographs_CuStrip	3	720x51x344	Right edge of upper slit (copper strip)
Radiographs_Lower_Slit	3	720x157x1808	Lower half of ROI images
Radiographs_Postage_Stamp	3	720x51x250	Left edge of the upper slit (used for per-radiograph flux normalization)

Table 6 - Region-of-interest dimensions and measurement conditions. The full panel, dark and background measurements are acquired once per CT scan as a one-time calibration. The copper strip, lower slit and postage stamp are acquired synchronous with all 720 angular measurements made during a CT scan.

2.1.3 Opportunities for MCT Dataset Data Mining

Data mining is the use of data science methods to find patterns and gains insights into data that has already been collected [Witten, 2017]. Data mining has been used successfully by SHM researchers [Gordan, 2022] and is the approach taken during this project.

This section revisits the levels of damage identification introduced in Section 1.2.1. At each level, an appraisal of the data mining opportunities is made by answering these questions: (a) what specific MCT damage identification goals can be defined? (b) what is the relevant data available in the MCT dataset? (c) what patterns might be discovered using data mining

that can address the MCT damage detection goals? (d) what missing data is needed to achieve higher levels of damage identification? (e) what new statistical methods and models could be developed to interpret the data?

Damage Level 1: Damage Detection
SHM Goals: Improve the speed to detect system drift. This task is undertaken during this SE296 study.
Data Availability: The LAC values provided in the MCT health monitoring dataset is sufficient for this task.
Expected Data Patterns: System drift can be detected using the novelty detection method discussed in Section 2.2.1.
Impact of Environmental and Operation Conditions: Radiation dose creates a cycle of damage and annealing in the X-ray panel detector as discussed in Section 1.3.1. This is believed to explain the high autocorrelation measured in Section 3.2. The autocorrelation requires correction using the data normalization technique discussed in Section 3.3.
Statistical Methods and Models: The current MCT health monitoring system uses statistical control charts with thresholds defined by tolerance intervals. Control charts are an unsupervised learning method. Section 3.1 demonstrates that the Wald Sequential Probability Ratio test can help with earlier detection and confirmation of system drift.

Damage Level 2: Damage Location
SHM Goals: Locate damage to either the X-ray panel detector or X-ray source.
Data Availability: The dark current measurements and postage stamp region of interest data provided in the MCT health monitoring dataset might be sufficient for this task using unsupervised learning methods.
Expected Data Patterns: The dark current measurements are taken in the absence of an X-ray signal so any changes observed using novelty detection methods would be an indicator of X-ray panel detector damage. The postage stamp measurements are known to be good indicators of changes occurring in the X-ray flux and spectrum.
Impact of Environmental and Operation Conditions: The X-ray tube is known to be sensitive to temperature fluctuations. Adding temperature probes as a secondary SHM sensor has been proposed to help monitor these effects. The X-ray tube operator can introduce variability if he or she does not follow a consistent warmup procedure.

Statistical Methods and Models: To be confirmed as damage sensitive features, the dark current and postage stamp measurements must be correlated with damage detected by the control charts or the more sophisticated Kalman filter anomaly detector discussed in Section 2.3.2.

Damage Levels 3,4 and 5: Damage Classification, Severity, and Prognosis

SHM Goals: Alert the MCT system operator when preventative maintenance is needed

Data Availability: Data required for supervised learning methods needs to be collected by the MCT technical team. This should include building a database of water scans to help determine how changes in the material references propagate to the test specimen. Radiation dose measurements could be used to correlate system drift with damage occurring in the X-ray panel and could be compared to the expected component life provided by the panel manufacturer.

Expected Data Patterns: As discussed in Section 2.3.3, damage sensitive features may manifest as clusters corresponding to different damage states. Ideally, these clusters could be monitored as indicators of system drift direction, severity and speed. This information could be used to predict drift dynamics.

Statistical Methods and Models The discovery of effective damage sensitive features should follow the data mining techniques discussed in Section 2.3.3. The ultimate goal is a unified health indicator for making preventative maintenance decisions and estimating remaining useful life. Assessing damage severity requires correlating damage manifesting in the reference material and ROI measurements to a corresponding drift in the test specimen measurements. This should be framed as an uncertainty quantification problem.

2.2 MCT Health Monitoring using Statistical Process Control (Current Approach)

This section provides an appraisal of the current health monitoring approach used by the MCT technical team. Section 1.1.1 introduced the control charts used as the primary means for monitoring MCT system stability. The control charts monitor the MCT measurements LAC values) of six reference materials at two different X-ray energies. Methods for constructing and interpreting control charts are discussed by practitioners of statistical process control, or SPC [Montgomery, 2020]. SPC is the use of statistical techniques for monitoring changes in a measurement process. Control charts can be an effective SHM tool when the measurements being tracked are highly correlated to the onset of damage [Farrar, Chapter 10].

2.2.1 Feature Extraction (SHM Step Three)

Step Three of the SHM design paradigm focuses on the extraction of damage sensitive features. This section identifies the damage sensitive features when control charts are used for SHM. It also discusses the underlying theory of novelty detection to suggest a potential improvement to the control charts used for the MCT health monitoring.

Statistical process control charts originated in the quality control community to help maintain stable processes for the prevention of defects or errors. Control charts monitor a process over time by plotting the sample means and/or standard deviations. Although derived from the principles of statistical inference and hypothesis testing, control charts do not require a formal statistical model. Control charts are easy to use but can be an effective tool for detecting changes to a process.

Control charts employ a general concept known as novelty detection. In novelty detection, a discordancy measure is defined using statistics calculated from measurements taken when the system, structure or equipment is assumed to be operating in an undamaged state. The discordancy measure is then used to compare each new measurement with a threshold used to flag which samples are considered to be outliers (discordant or novel). For univariate data measurements x_t, the appropriate discordancy measure is the z-statistic defined in Equation-1.

$$z = \frac{|x_t - \bar{x}|}{\sigma_x} \quad (Eq1)$$

When control charts are used for SHM, the z-statistic functions as a damage sensitive feature. For the current MCT system, the z-statistic is calculated for each of the derived LAC values (high and low LAC values for six different reference materials) and twelve different SPC charts must be interpreted.

The identification of a discordant measurement (i.e., damage) depends on where the threshold is set in the control plot. The thresholds can be chosen in different ways but the "three sigma" approach is a common SPC practice that defines the outlier thresholds at the mean value plus/minus three standard deviations, or $x \pm 3\sigma_x$. The current MCT health monitoring system refers to these thresholds as the upper and lower Quality Control (QC) bounds. As discussed in Section 2.2.2, the QC bounds are computed using statistical tolerance intervals which yields a similar result to the "three sigma" approach.

In calculating the mean $\bar{x}$ and standard deviation σ_x appearing in Equation-1, it is assumed that the sample size is large enough to ensure that these statistics are computed with a sufficient confidence level. A well-known practical choice based on statistics theory is to include a minimum of thirty samples. The MCT team follows this practice by ensuring that the mean and variance of the LAC values are calculated using a calibration dataset of at least thirty CT scans after a new MCT system is commissioned or following a repair or software update.

The discordancy measure for novelty detection can be extended to multivariate data. In Equation-2, the measured time series data $|x_t|$ is an n-dimensional feature vector instead of a scalar value. Both the $nx1$ mean feature vector $|\bar{x}_t|$ and the nxn feature covariance matrix $|\Sigma|$ should be calculated when the system is assumed to be operating in a normal state.

$$D^2 = (|x_t| - |\bar{x}|)^T |\Sigma|^{-1} (|x_t| - |\bar{x}|) \quad (Eq2)$$

The discordancy measure D appearing on the left-hand-side of Equation-2 is known as the Mahalanobis distance. This could be used as a principled method to consolidate the twelve MCT control charts into a single chart. For each new CT scan, the twelve LAC values would be assembled into a feature vector so the Mahalanobis distance could be calculated. The calculated values can then be plotted in a control chart for monitoring. One complication is deciding the thresholds to use with this approach. Farrar discusses how to use a Monte Carlo "bootstrap" method to calculate the thresholds required to achieve the desired confidence level [Farrar 2020, Chapter 6.10].

2.2.2 Statistical Model (SHM Step Four)

The fourth and final step of the SHM design paradigm focuses on developing the statistical model which takes as input the damage sensitive features defined in step three. The purpose of the statistical model is to discriminate between the damaged and undamaged states of the system, structure or equipment. This section examines the statistical model employed by the current MCT health monitoring system. However, before discussing the statistical model, I first review the basics of statistical tolerance intervals and how they are used to interpret MCT measurement data.

As mentioned in Section 1, the MCT system functions as a metrological instrument used to certify manufactured material test specimens. The specimens can vary on the basis of their designed material composition or due to manufacturing process variability. Tolerance intervals can be used to establish the acceptable statistical range of values for each measured characteristic of a test specimen. Figure 9 shows how the low- and high-energy tolerance intervals (for a given test specimen) can be combined into a plot called the "region of responsibility" or ROR. RORs are constructed to monitor a range of different material properties including density (ρ), high- and low-energy linear attenuation coefficients $(\mu_{high}, \mu_{low)}$, electron density (ρ_e), and effective atomic number (Z_e). The RORs are constructed by scanning a small batch of test specimens which have been chosen as representative of the "ground truth" (i.e., exemplars) by which future test specimens will be certified.

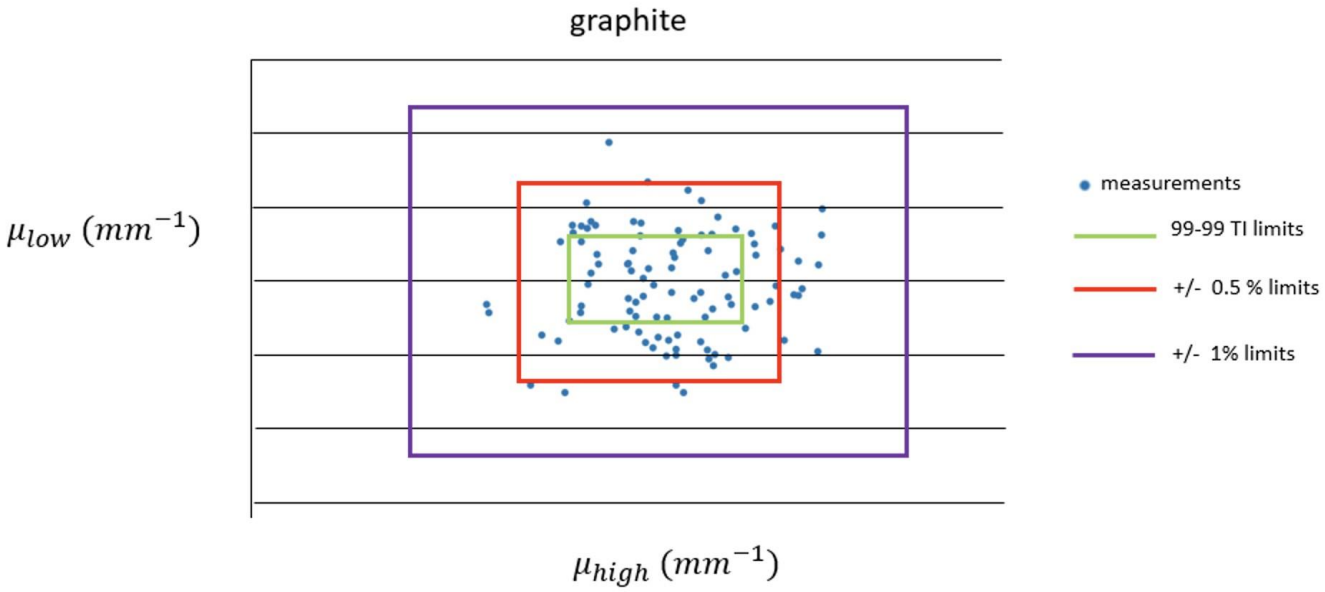


Figure 9 – Example of the "Quality Control" or QC bounds used for MCT data interpretation. The QC bounds are drawn using two univariate tolerance intervals (TI's). The TI for the low-energy LAC value, μ_{low}, is drawn along the y-axis. The TI for the high-energy LAC value, μ_{high}, is drawn along the x-axis. This results in the green box-shaped region. The red and purple box regions are created in an ad hoc manner to contain the outliers which are frequently observed. A separate QC bound plot is created for each of the six reference materials. In a similar fashion, two univariate tolerance intervals for the test specimen are used to draw a plot called the Region of Responsibility (ROR).

An additional set of tolerance intervals called the "Reference Material Quality Control (QC) bounds" are constructed for each of the six reference materials in the MCT health monitoring system. Figure 9 illustrates an example showing the QC bound for graphite. The QC bounds are used to confirm the MCT scanner is operating normally. For example, when a test specimen measurement is found to lie outside of its ROR, the question naturally arises as to whether the specimen is flawed or the MCT system is malfunctioning. The MCT health monitoring system described in Section 1.1.1 was designed to answer this question. The reference materials are always being scanned continuously along with each test specimen. The simple assumption is that the MCT system is operating normally if these reference measurements fall within the QC bounds. Any ROR outlier measurements may then confidently be attributed to the test specimen.

The QC bounds are determined by the two-sided tolerance interval formula shown in Figure 10(b). A tolerance interval is a statistical interval that brackets a fixed proportion of the population (i.e., the coverage value P) at a given confidence level, alpha. The convention is

to specify the confidence level first and the coverage last, so a 90-95 tolerance interval would bracket 95% of the population values with a confidence level of [1-alpha], or an alpha of 10%. The k-factor multiplier appearing in the tolerance interval formula determines the width of the QC bounds and is a function of the proportion P, sample size n, and confidence level alpha. The k-factor can be determined using statistical software or using published reference tables [NIST]. The bounds are drawn in the MCT control charts as the mean LAC value plus and minus the percent allowed variance. A typical bound may be on the order of 0.5% to 0.75%.

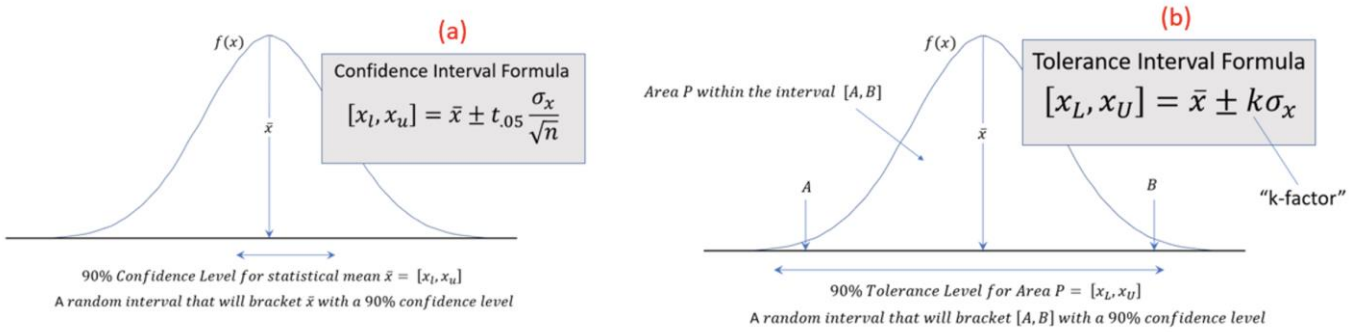


Figure 10 – Comparing definitions for the (a) Statistical Confidence Interval and (b) Statistical Tolerance Interval. The k-factor appearing in the tolerance interval formula is a function of the proportion P, sample size n, and confidence level alpha.

One of the challenges faced by the MCT technical team is setting thresholds for the RORs and QC bounds. As shown in Figure 9, the observed measurement variability often exceeds the range in the published tolerance interval tables. The team has resorted to widening the thresholds using trial and error, selecting empirically a plus-and-minus percent error bound. I now mention several references that offer potential explanations and guidelines for adjusting these thresholds.

Section 3.2 will show that MCT measurements have a high degree of serial correlation. Knoth [2003] demonstrates that autocorrelation leads to higher variability in the observed measurements requiring a correction to the published k-factors. Table 7 shows that the tolerance limits are wider as the amount of serial correlation ρ increases, and diverges more for smaller sample sizes N. It needs to be noted that the values published by [Knoth 2003] apply to the construction of univariate tolerance intervals.

N	ρ = 0	ρ = 0.3	ρ = 0.6
10	3.379	3.754	4.754
30	2.549	2.684	3.093
100	2.233	2.280	2.438

Table 7 – K-Factors for two-sided 95-95 tolerance limits for increasing autocorrelation, ρ. Values taken from [Montgomery] and [Knoth].

The tolerance intervals used to construct the QC and ROR plots are assumed to be independently observed, univariate quantities. However, the MCT measured values are in fact highly correlated, and thus require multivariate treatment. Polhemus [2017] discusses multivariate tolerance intervals and shows that a further widening of the tolerance limits may be expected. Figure 11 shows a hypothetical result. Polhemus describes how to use Monte Carlo simulation to bootstrap the multivariate tolerance region (shown as the purple ellipse). The TI will be ellipsoidal if the underlying distribution is approximately multivariate Gaussian, and extended along a diagonal with a slope proportional to the correlation coefficient. A second approach discussed by Polhemus is to use the Bonferroni method shown in red. This does not require simulation but results in a more conservative estimate of the true bounds.

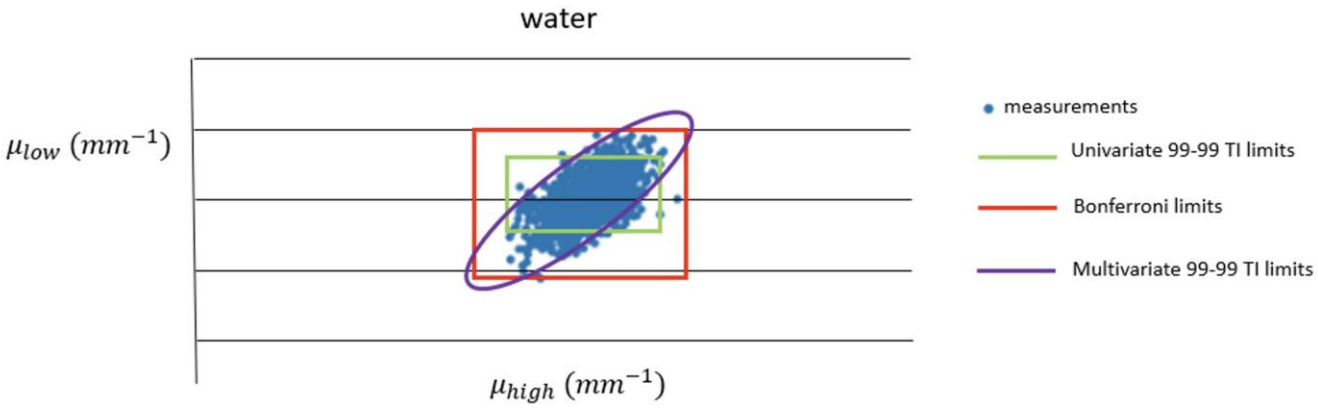


Figure 11 - Construction of Multivariate Tolerance Intervals for MCT data interpretation. The correlation between the μ-values causes the data to be stretched along the diagonal as shown. This will leave outliers unaccounted for by the construction of univariate tolerance intervals (green box) but will be accounted for by the multivariate tolerance intervals (purple ellipsoid).

I now return to the originally stated goal of this section which is to affirm that the statistical model employed by the current MCT health monitoring system is none other than the Gaussian or normal distribution model. The univariate Gaussian model is assumed by the MCT technical team, but the multivariate Gaussian is the more correct model for the reasons discussed. The validity of the inferences made from this model depend on how well the probability density has been estimated from available data. Density estimation is well known to be one of the most difficult problems in statistical/machine learning. As shown in Equation-3, the joint probability distribution needed for MCT health monitoring is realistically a 14-dimensional random vector consisting of the high- and low-μ values for six reference materials and test specimen:

$$X = (\mu_1^H, \mu_1^L, \mu_2^H, \mu_2^L, \mu_3^H, \mu_3^L, \mu_4^H, \mu_4^L, \mu_5^H, \mu_5^L, \mu_6^H, \mu_6^L, \mu_{test}^H, \mu_{test}^L)^T \quad (Eq3)$$

The direct learning of the joint PDF is referred to as the generative approach in statistical learning and would require too many sample measurements to be considered tractable. Discriminative machine learning algorithms represent an alternative approach. These algorithms rely on supervised learning to make accurate predictions and require far less data by comparison. Discriminative learning approaches will be discussed in Section 2.3 as a potential avenue for making improvements to the current MCT health monitoring system.

2.3 MCT Health monitoring using Statistical Pattern Recognition (Proposed Approach)

Classical statistical inference methods including tolerance intervals rely fundamentally on a large amount of measurement data to learn the probability density with a high confidence level. This may be intractable for the MCT health monitoring system given the large number of variables. These methods also depend on assumptions about the underlying data distribution (e.g., multivariate Gaussian, independent and identically distributed). In this section, I explore alternative approaches that can be exploited by the statistical pattern recognition (SPR) health monitoring paradigm. This includes machine learning algorithms that can make accurate predictions with far less training data. It is proposed that the SPR approach offers a richer framework for MCT system diagnosis and prognosis.

2.3.1 Motivation

The current MCT health monitoring system employs control charts to detect system drift caused by damage to MCT components. While simple to use, control charts have a number of limitations:

Control charts:

- Are an unsupervised learning method appropriate for damage detection, but unable to be used for damage classification or severity assessment;
- Are limited to the monitoring of univariate quantities and are prone to misleading results when correlated multivariate quantities are monitored;
- Require manual updates to the control limits (decision thresholds) when the process changes;
- May produce a high false alarm rate when the measurements being monitored are highly corrupted by noise;
- Monitor primary quantities directly and do not offer a model for incorporating relevant physics or secondary measurements.

These limitations motivate the development of new statistical models for MCT health monitoring data interpretation. As discussed in Section 2.1.2, the MCT health monitoring system collects secondary data that might be exploited for SHM diagnosis and prognosis. The MCT technical team refers to this secondary dataset collectively as the "region of interest" statistics. The MCT technical team believes that damage to the X-ray panel may manifest as subtle changes to the dark current measurements. Damage to the X-ray panel may manifest in a region of interest referred to as the "postage stamp".

The development of new statistical models to improve the current MCT health monitoring system is a complex task. It will require more data to be collected, and also a greater time commitment than the current SE296 project allowed. However, the literature survey uncovered several statistical modeling approaches that appear promising as the starting point for future work. These are discussed in the next two sections.

2.3.2 Statistical Modeling based on General Anomaly Detection

An alternate approach posits that the loss of measurement fidelity can be treated as a sensor fault detection SFD problem. Automated SFD methods are widely researched due to the critical role of sensors for monitoring and control in industry. Dynamical systems researchers at the University of Washington and Boeing demonstrated the robust SFD architecture shown in Figure 12 [de Silva, 2020]. The architecture combines methods from system identification (Kalman Filter), data-driven modeling (dynamic mode decomposition) and machine learning (decision trees). Section 4 provides an outline for how this modeling approach could be applied to the MCT system.

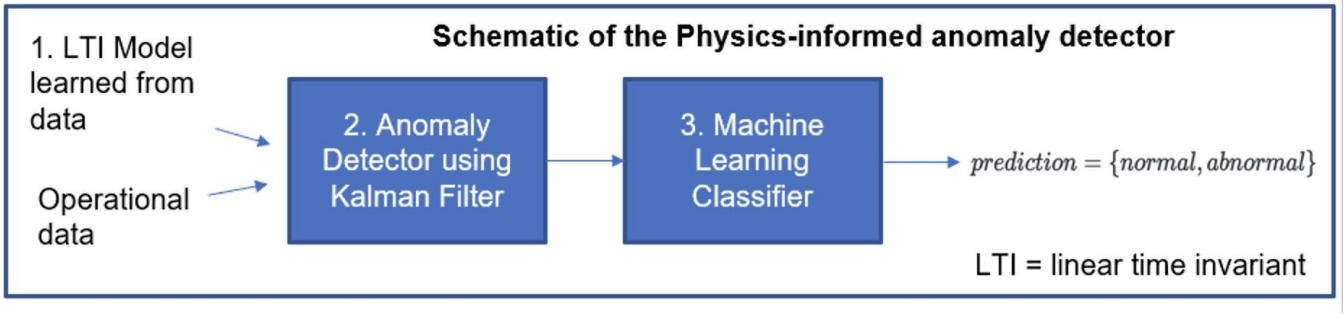


Figure 12 - A general architecture for physics-informed anomaly detection. (a) Dynamic mode decomposition is used as a data-driven approach to learn the LTI model that the Kalman filter requires for state estimation; (b) The Kalman filter is used for monitoring and can be an effective algorithm for discriminating between "normal" and "abnormal" system operating behavior; (c) Decision trees are used to automate the threshold used to decide between the "normal, or undamaged" and "abnormal, or damaged" system state. Reference: [de Silva, 2020]

The Kalman filter can be used to detect system drift and might overcome the limitations of control chart as highlighted in Section 2.3.1. The Kalman filter can analyze multivariate data: the observation data can be of arbitrary dimension and take into account the correlations between different measurement input variables. The Kalman filter also has a high immunity to noise relative to control charts. In order to use the Kalman filter, a model is needed to specify the dynamics and noise present in the system or process being monitored. This makes the Kalman filter more difficult to use than control charts, but also presents an opportunity to inject known physical or empirical constraints as well as measurements taken from secondary sensors.

The general anomaly detection architecture shown in Figure 12 overcomes the limitations of control charts in two additional ways. The data-driven model used by the Kalman filter can

be developed using time series methods or more advanced machine learning approaches. De Silva discusses the use of an algorithm called Dynamic Mode Decomposition that is able to detect subtle coherent spatial or temporal patterns. If these patterns can be correlated with damage, they might be effective as damage sensitive features in an SHM system. The SFD architecture also employs machine learning to adjust the decision thresholds as the process changes. This overcomes another limitation of control charts where the control limits must be updated manually.

As discussed in Section 1.2.1, damage state data is needed to employ the supervised learning methods required to ascend to higher levels of damage identification. Control charts provide a damage detection mechanism for labelling time periods or epochs where the system is observed to be operating "normally" or 'abnormally'. The SFD architecture can serve the same purpose while at the same time incorporating observational data from more sensors allowing more damage patterns to be discovered.

2.3.3 Statistical Modeling based on Cluster Detection and Tracking

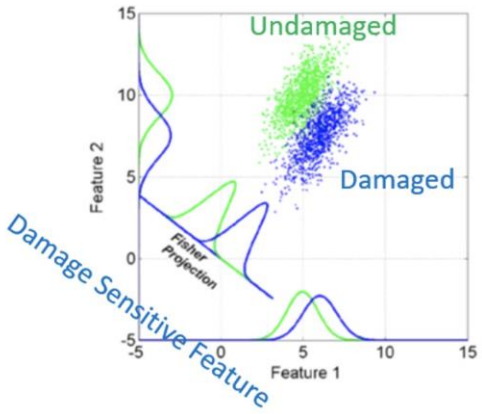
Farrar and Worden pioneered the treatment of SHM design as a problem in statistical pattern recognition and machine learning. A preliminary step in machine learning is to assemble a “data matrix” where each column might contain n measurements (e.g., sensor data, statistics, or transforms) and each row might contain a time series of these measurements at m successive instances.

$$X_{nxm} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \quad (Eq4)$$

Images can be incorporated into this schema by flattening the rxc image data into a tall and skinny vector of dimension r $*$ c. Each column can in general contain data from multiple sensors and image detectors fused together and sampled synchronously in a single time step. Each row represents the measurement of some "feature" that can be correlated to damage by itself or more generally might be fused with other row measurements to achieve

the correlation. In either case, the result becomes a damage sensitive feature that can be incorporated into statistical models for monitoring health state changes.

Assembling data in this manner has the advantage that the dynamical state of the structure or equipment to be monitored may be treated as patterns or clusters to be detected and analyzed in an n-dimensional feature space. This opens the door to the mathematical methods of linear algebra, matrix decomposition, and machine learning. Constructing damage sensitive features often requires employing multiple data science techniques including data fusion, matrix transforms, and dimensionality reduction using projections. This process is referred to as feature engineering.



(a) Fisher projection (pictured in figure to the left): a supervised dimensionality reduction method that maximizes the distance between two labelled classes

(b) Principal component analysis: an unsupervised dimensionality reduction method that projects data along the orthogonal axes that maximize the variance

(c) Dynamic Mode Decomposition: a data-driven algorithm to approximate dynamical systems as a superposition of coherent structures (modal representations) that grow, decay and/or oscillate in time

Figure 13 - Matrix decomposition methods that can be used for feature engineering and the data mining of damage sensitive features. Figure is derived from [Farrar, 2012]

Figure 13 illustrates the general aspects of how feature engineering is applied to finding damage sensitive features. The clusters represent measurements taken from a two-dimensional feature space. As depicted in the figure, it is often necessary to fuse together and transform data before the damaged and undamaged states of a system can be discriminated. In the example shown, the damage sensitive feature (DSF) is the Fisher Projection of the combination of Feature 1 and Feature 2. The resulting DSF could be used as input to a diagnostic or prognostic statistical model.

The association of damage states with clustering introduces a new interpretation of system drift: "drift is a gradual change in the distribution of patterns that is mainly caused by an incipient fault" [Chammas]. If the detection and monitoring of changes to the distribution parameters is possible, the drift speed, direction and severity can all be characterized. This

would enable the development of a complete health monitoring solution, not only one with the capability of classifying damage, but one capable of prognosis and the estimation of remaining useful life.

3 Results

This section presents the results of the exploratory data analysis performed on the MCT system data provided by the LLNL Nondestructive Characterization Institute.

3.1 Optimal detection using the Wald SPRT

This section reviews the theory underlying the Wald sequential hypothesis testing method and demonstrates how it can be used to enhance the diagnostic value of the control charts used by the MCT team.

3.1.1 Outline of algorithm used for sequential hypothesis test

The Wald sequential probability ratio test SPRT is the optimal algorithm for sequential hypothesis testing [Wald, 2004]. As with standard hypothesis testing, two hypotheses are considered with at a certain confidence level (1 − α) and power (1 − β). Figure 10 depicts the sampling distributions of two possible hypotheses to be tested.

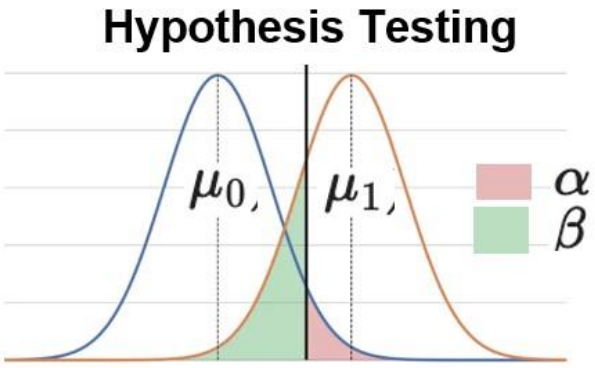


Figure 14 - Standard hypothesis testing as a tradeoff comparison of the null hypothesis H0 with mean value μ0 and the alternate hypothesis H1 with mean value μ1. The probability of Type 1 error (or False positive rate) is determined by parameter α which is also referred to as the significance level. The probability of Type 2 error (or False negative rate) is determined by the parameter β. These parameters require the statistician to make a tradeoff between the two types of error. This is done by choosing the p-value, which is the decision boundary location represented by the vertical line in the figure.

With standard hypothesis testing, the number of measurement samples needed to achieve the desired Type 1 and Type 2 errors is fixed. This is not the case with sequential testing as the number of measurement samples will vary depending on the measurement sequence. Wald demonstrated that sequential testing could be decided on the basis of a likelihood ratio as explained below.

The Sequential Probability Ratio Test is based on the likelihood ratio:

$$\Lambda(X_i) \equiv \frac{\mathcal{L}(\mu_1|\mathcal{X})}{\mathcal{L}(\mu_0|\mathcal{X})} = \frac{\prod_{i=1}^{n} p_1(X_i)}{\prod_{i=1}^{n} p_0(X_i)} \quad (Eq5)$$

Assuming one performs a fully specified hypothesis test and the data follows a normal distribution $X_i \sim N(\mu, \sigma^2)$, the log-likelihood has the form:

$$\log(\Lambda(X_i)) = \left(\sum_{i=1}^{n} \frac{(X_i - \mu_0)^2}{2\sigma^2}\right) - \left(\sum_{i=1}^{n} \frac{(X_i - \mu_1)^2}{2\sigma^2}\right) \quad (Eq6)$$

This is equivalent to updating the ongoing sum for every new sample i, $S_i = S_{i-1} + \log(\Lambda_i)$. Wald 2004 shows that if one defines upper and lower decision thresholds given by:

$$a \approx log\frac{\beta}{1-\alpha} \qquad b \approx log\frac{1-\beta}{\alpha} \quad (Eq7)$$

It follows that the hypothesis test can then be decided using the following decision rules:

$S_i \geq b$: Accept the alternative hypothesis H_1
$S_i \leq a$: Accept the null hypothesis H_0
$a < S_i < b$: the test remains inconclusive so continue monitoring

Wald also proved the optimality of the SPRT algorithm in the sense of requiring the fewest measurements before a hypothesis decision is reached. On average, the Wald SPRT requires half as many measurements when compared to classical hypothesis testing at the same significance levels.

3.1.2 Results

MATLAB codes were developed to explore application of the SPRT algorithm to MCT LAC data (see "*Detect_Drift_Using_SPRT"* in the Appendix). The MCT data is assumed to be Gaussian distributed for all likelihood calculations. Also, the MCT system is assumed to be operating normally during the calibration period. The calibration period corresponds to the first thirty LAC measurements and these measurements are used to calculate the mean and standard deviation statistics. To use the SPRT as a drift detector for the LAC measurement values, one must formulate both the null and alternate hypothesis. The null hypothesis is set with a distribution mean value equal to the calibration mean. To test for an upward system drift, the alternate hypothesis is set with a distribution mean value equal to the calibration mean plus one standard deviation. A second hypothesis test is used to test for downward system drift with the alternate hypothesis located one standard deviation below the mean.

Figure 15 shows the results of the sequential test for all six reference materials. In each plot, the upper and lower Wald decision boundaries are shown as horizontal lines. The red vertical line shows the date when the MCT system is observed to be operating abnormally.

Each point in the plot represents the likelihood sum. In accordance with sequential testing, each test begins with a zero-value. With each new measurement, the log-likelihood is calculated and summed until one of the decision boundaries is reached. Coloration has been added to help interpret the state of the sequential test. Blue points indicate that the sequential test has not yet reached a decision. Green points indicate confirmation of the null hypothesis meaning that no system drift is detected. Red points indicate confirmation of the alternate hypothesis meaning that system drift has been detected. Each time a decision boundary is reached, the likelihood sum is reset to zero and a new sequential hypothesis test is started.

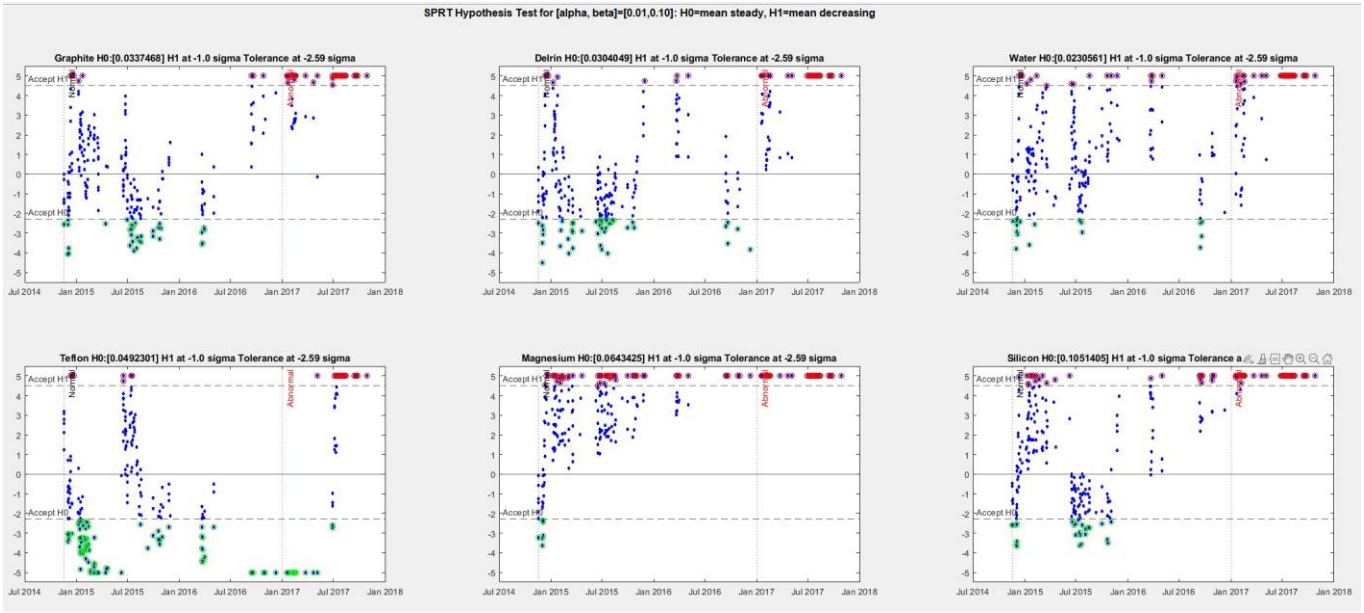


Figure 15 - Sequential Probability Ratio Test for downward drift in the LAC 100 kV measurements from July 2014 through July 2017. The plots for all six reference materials are shown, (top row): graphite, Delrin, water; (bottom row): Teflon, magnesium, silicon.

Figure 16 demonstrates how to integrate the SPRT calculation with the LAC control charts already in use by the MCT technical team. Each LAC measurement sample is colored to reflect the SPRT prediction at each time step. This provides additional diagnostic information to help decide if the MCT system measurements are starting to drift. The SPRT predictions could be incorporated into the QC bounds chart by combining the results from the 100 kV and 160 kV SPRT calculation.

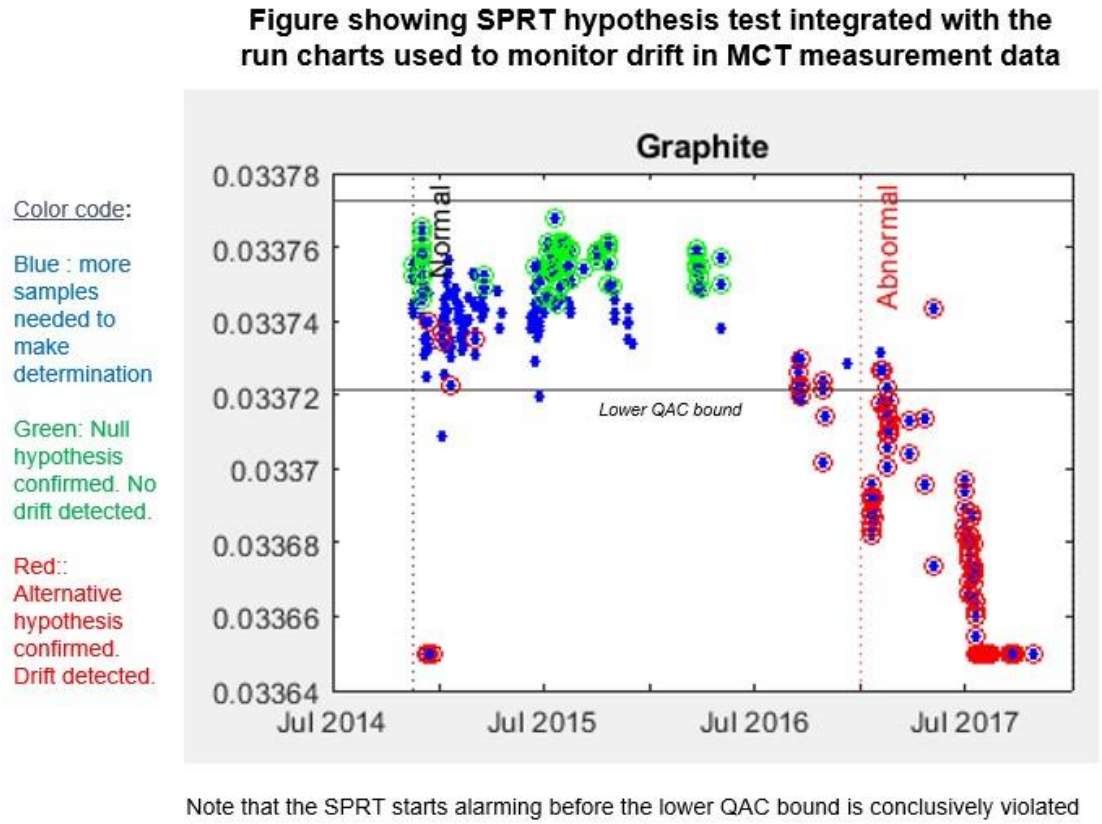


Figure 16 - Plot showing SPRT hypothesis test integrated with the LAC control chart for graphite. The LAC control charts are used to monitor drift in MCT measurement data

3.1.3 Discussion

The α and β parameters can be adjusted to move the Wald decision boundaries. This adjustment alters the tradeoff between Type 1 and Type 2 errors. The SPRT charts were observed to have an elevated false alarm rate due to the autocorrelation present in the MCT system data. This complication is addressed in Section 3.3.

3.2 Evaluation of statistical assumptions

The serial correlation observed in the LAC control charts can be quantified by applying statistical measures of autocorrelation to MCT data. MATLAB code was developed to measure the autocorrelation. The presence of serial correlation greater than 0.2 will cause many false alarms in control charts [Montgomery, 2015]. Table 8 shows that the autocorrelation present in the LAC data far exceeds this threshold over many time lags.

Autocorrelation coefficients (lags 1-10) for the low-energy LAC values

lag	Graphite	Delrin	Water	Teflon	Magnesium	Silicon
1	0.58159	0.63809	0.49121	0.91174	0.91432	0.60521
2	0.55509	0.61373	0.46708	0.86173	0.88471	0.58191
3	0.52921	0.58559	0.44319	0.80074	0.85057	0.55488
4	0.50185	0.55591	0.41317	0.74515	0.81565	0.52783
5	0.46854	0.53078	0.38917	0.69083	0.78297	0.50325
6	0.43972	0.50096	0.36237	0.64067	0.75312	0.47908
7	0.40801	0.47277	0.33721	0.58772	0.72391	0.4586
8	0.38976	0.44833	0.31443	0.53728	0.69366	0.43378
9	0.35443	0.41791	0.28588	0.48464	0.66462	0.41178
10	0.33025	0.39687	0.27577	0.4546	0.63825	0.39394

Table 8 - Calculated autocorrelation between MCT values at different time lags. The table was created using MATLAB code "Calculate_AutoCorrelation_Function.m" provided in Appendix B.

3.3 Data normalization

This section demonstrates that the autocorrelation (trending patterns) present in the MCT LAC value data are the likely cause of the high false alarm rate observed in the Wald SPRT detector. Farrar discusses a data normalization algorithm that uses an autoregressive (AR) model to mitigate autocorrelation. Section 3.3.1 presents measurements of the partial autocorrelation coefficient needed to construct the model. Section 3.3.2 then presents the results following the method discussed by Farrar [2021, Chapter 10.5.1] and [Entezami, 2021].

3.3.1 Procedure for choosing autoregressive model order

Montgomery [2015] advises that the partial autocorrelation function (PACF) should be used when building autoregressive models. This function measures autocorrelation at a given time lag after removing the measured autocorrelation of all previous time lags. As shown in Table 9, the calculated coefficients suggest that using an AR(1) or AR(2) autoregression model would correct most of the autocorrelation present in the MCT data.

Partial Autocorrelation coefficients (lags 1-5) for the low-energy LAC values

Graphite	Delrin	Water	Teflon	Magnesium	Silicon
0.581590	0.638092	0.491210	0.911744	0.914317	0.605205
0.327676	0.348441	0.297605	0.180489	0.297091	0.340275
0.205835	0.207529	0.196451	-0.047091	0.069225	0.208087
0.130270	0.123333	0.123833	-0.017683	-0.001818	0.130525
0.070529	0.080341	0.084887	-0.013611	-0.002661	0.085396

Most of the serial correlation occurs in lags 1 and 2 suggesting the use of an AR(1) or AR(2) autoregressive model.

Table 9 - Calculated partial autocorrelation coefficients between MCT values at different time lags. The coefficients were calculated using the Minitab Statistical Software package.

3.3.2 Results

MATLAB codes were developed to explore how an AR model might be used to mitigate the autocorrelation present in MCT data. Figure 17 shows the form of the model used for each of the twelve LAC time series (six reference materials at both low and high energies).

AR model:

$$x(t) = a + \sum_{j=1}^{p} a_j x(t-j) + \varepsilon(t)$$

residuals

Figure 17 - AR model used for the correction of autocorrelation

Plots of the residual error between the MCT LAC measurement values and the autoregressive model estimates are shown in Figure 18. Both the AR (1) and AR (2) models were applied to the data with similar results. The plot shows that the AR model effectively removes the trending patterns. The result is a damage sensitive feature that appears as white noise during the normal period of MCT operation, but changes markedly to become one-sided when the MCT operation enters the abnormal period of operation. This information may be exploited to lower the false alarm rate in the LAC control charts.

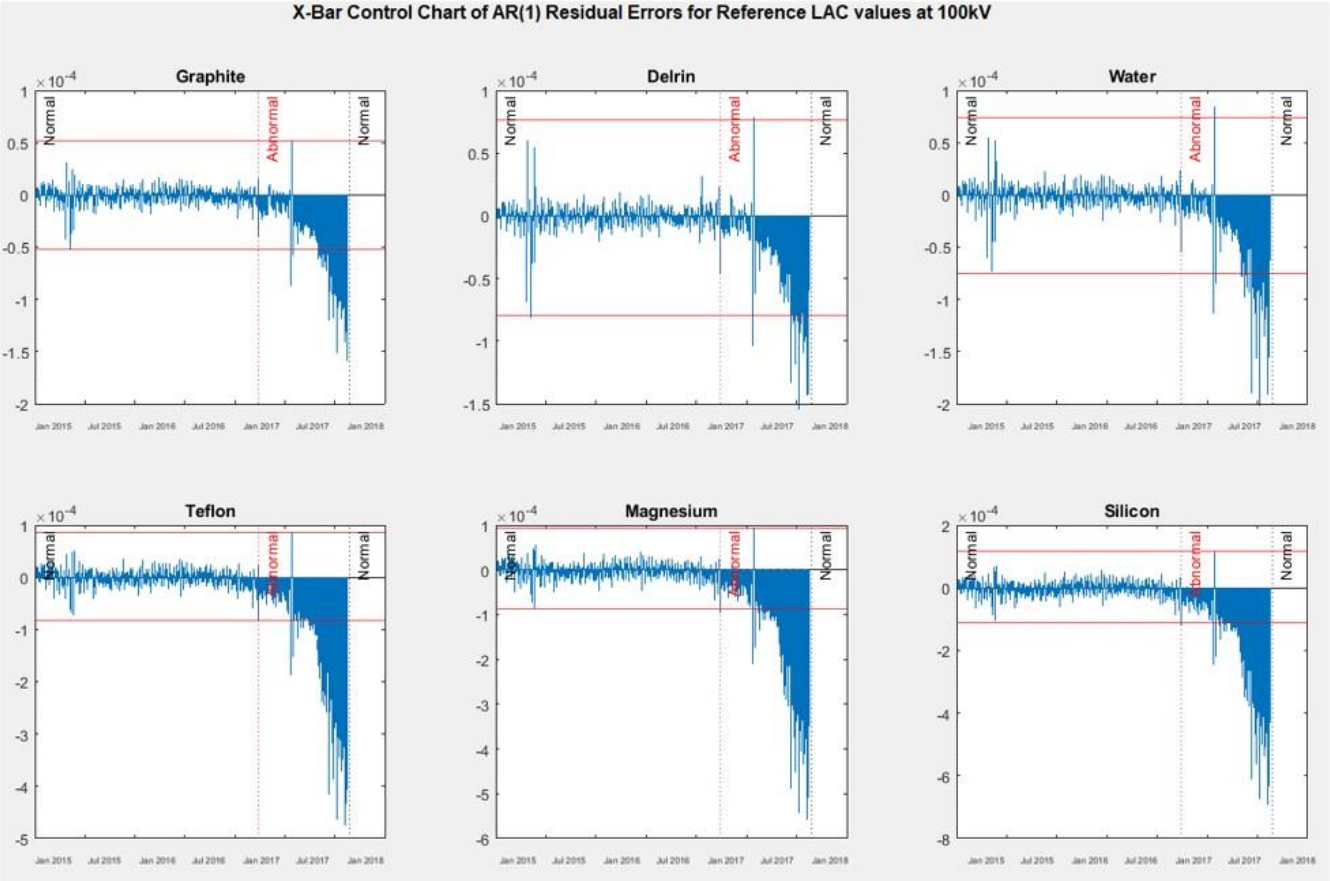


Figure 18 - Plots of the residuals resulting from the AR (1) corrected models are shown for the six reference materials at 100 kV. The residuals show a "white noise" pattern when the LAC values are in control. The residuals show a markedly one-sided pattern once the LAC values begin to drift out of control.

4 Discussion

Section 2.2 provided an appraisal of the current MCT health monitoring systems which uses statistical control charts and tolerance intervals. These tools rely on the classical theory of statistical inference and hypothesis testing. The limitations of these tools were also outlined in light of the correlations and nonstationary effects present in MCT data. While these tools may be capable of detecting system drift (i.e., Level 1 - Damage Detection), they are not capable of achieving higher levels of damage identification.

Improving the current MCT health monitoring system will require two next steps: (a) collecting damage state data that can be used for supervised learning; (b) developing more sophisticated statistical methods and models to interpret all the data collected.

The MCT technical team plans to gather secondary measurements that may prove useful for supervised learning methods. A database of "water scan" data is being collected. This involves performing a CT scan where the test specimen is a "ground truth" water sample. This information will enable the correlation between system drift occurring in the material reference LAC values and ROI measurements to be studied. Radiation dose measurements are also starting to be collected at different locations in the X-ray panel detector as a function of collimation and X-ray source settings. The collection of temperature data is also being considered.

New statistical methods and models will be needed to make sense of all the data being collected. It may be the case that damage cannot be discriminated using the LAC values of a single reference material. The damage sensitive features may require some combination of reference material data, data extracted from the ROI sensor regions, and/or statistical transforms of these feature combinations. This would require using the techniques described in Section 2.3.3 including data fusion, dimensionality reduction, and other matrix transformations.

In Section 3.3, it was shown that the residuals are an effective indicator of sensor drift. The anomaly detection model discussed in Section 2.3.2 might be used to exploit a generalized version of this effect: the variance of the innovations matrix for a linear-time-invariant model of the MCT system measurements. Figure 19 provides a schematic of the proposed approach. The anomaly detector would allow all twelve of the LAC value measurements to be incorporated into a single damage model. The model would also allow experimentation with the ROI sensor data to see if incorporating these into the damage model could lead to a better prediction of system drift. The objective would be to identify which ROI data might be effective candidates for constructing damage sensitive features.

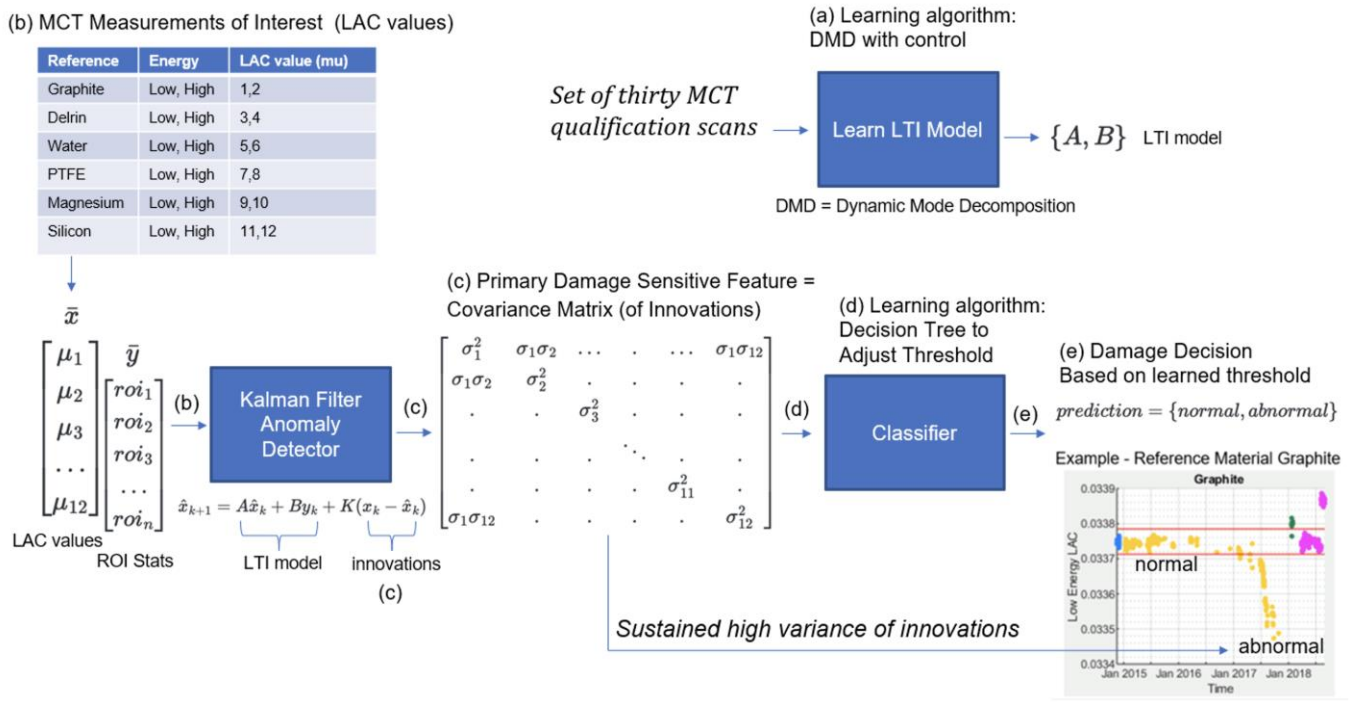


Figure 19 - Proposed use of the General Anomaly Detector as the MCT Statistical Model: (a) A linear-time-invariant (LTI) model of "normal" operating behavior is learned from the set of thirty MCT calibration scans; (b) the twelve LAC values and secondary ROI sensors serve as inputs to the Kalman filter anomaly detector; (c) abnormal behavior manifests as sustained high variance in the innovations matrix; (d) decision trees can be used as a machine learning algorithm to adjust the threshold between normal and abnormal variances; (e) the final damage decision is made based on the variances levels relative to the decision threshold.

5 Conclusion

During this SE296 independent study project I have applied the SHM paradigm to appraise the current MCT health monitoring system and to explore ideas for improvement. I have written MATLAB codes implementing an optimal detector for the onset of drift in MCT system measurements. I have outlined the strengths and weaknesses of using statistical process control charts which are limited to the monitoring of univariate quantities. I have examined the confounding effects of correlation and shown how to mitigate using an SHM data normalization technique. I have demonstrated that treating health monitoring as a problem in statistical pattern recognition introduces new models and methods for MCT damage diagnosis and prognosis. I have completed a literature survey to create a roadmap for future work. These contributions carryover to the health monitoring of scientific and metrological instruments more generally.

6 Annotated References

1. Antonuk, Larry E., et al. "Radiation damage studies of amorphous-silicon photodiode sensors for applications in radiotherapy X-ray imaging." *Nuclear Instruments and Methods in Physics Research* Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 299.1-3 (1990): 143-146.
2. Argyll Ruane Engineering Training Centre, Sheffield. *Nondestructive Testing 2023: Designing for full-lifecycle inspection*. Seminar to be held April 2023, https://events.imeche.org/ViewEvent?e=7567
3. Boudry, John Moore. *"Evaluation of Hydrogenated Amorphous Silicon Photodiodes and Field-Effect Transistors for Use as Elements of Two-X Ray Imaging Arrays."* Ph. D. Thesis 1996. Note: The MCT system uses X-ray panel detectors based on amorphous silicon technology. Boudry's PhD thesis discusses the radiation damage mechanisms for the photodiode and FET circuits used in these detectors.
4. Boudry, J. M., and L. E. Antonuk. "Radiation damage of amorphous silicon, thin-film, field-effect transistors." *Medical physics* 23.5 (1996): 743-754.
5. Chammas, Antoine, et al. "Drift detection and characterization for fault diagnosis and prognosis of dynamical systems." *Scalable Uncertainty Management: 6th International Conference*, SUM 2012, Marburg, Germany, September 17-19, 2012. Proceedings 6. Springer Berlin Heidelberg, 2012. Note: This technical article discusses an architecture that uses clustering for detecting drift and characterizing its direction, speed and severity. It provides an alternative statistical approach to the control charts currently used by the MCT technical team.
6. de Silva, Brian M., et al. *"Physics-informed machine learning for sensor fault detection with flight test data."* arXiv preprint arXiv:2006.13380 (2020). Note: These researchers propose an automated sensor fault detection method using algorithms from system identification (Kalman Filter), data-driven model building (dynamic mode decomposition) and machine learning (decision trees).
7. Entezami, Alireza. *Structural health monitoring by time series analysis and statistical distance measures*. Springer International Publishing, 2021. Note: Discusses techniques to correct for nonstationary trends as a data processing step preceding feature extraction.
8. Farrar, Charles R., and Keith Worden. *Structural health monitoring: a machine learning perspective*. John Wiley & Sons, 2012. Note: The standard reference for architecting structural health monitoring systems. Chapter 10 discusses the use of statistical process control charts as a tool for anomaly detection. Also discusses how to reduce the control chart false alarm rate using an autoregressive mode to correct process data that is nonstationary.
9. Farrar, Charles R., Lecture notes for UCSD engineering course, "SE265 Structural Health Monitoring".

10. Gordan, Meisam, et al. "State-of-the-art review on advancements of data mining in structural health monitoring." *Measurement* 2022 110939.
11. Harrison, Thomas Jay. *"The Sequential Probability Ratio Test (SPRT) in Feature Extraction and Expert Systems in Nuclear Material Management."* (2004). Note: This thesis describes an X-ray equipment health monitoring system based on SPRT. It also has MATLAB code listings for SPRT and other plots.
12. Knoth, Sven, and Raid Amin. "Autocorrelation and tolerance limits." *Journal of Statistical Computation and Simulation* 73.7 2003 467 489.
13. Lindsay, Gordon. "Further elevating condition-based monitoring through data-driven modeling" *Process Instrumentation*, Accessed October 2022, available online at: https://www.piprocessinstrumentation.com/monitoring/article/21238067/further-elevating-conditionbased-monitoring-through-datadriven-modeling.
14. Martz, Harry. Document *IM 961271_LLNL PRES 770160_LEDP*, "Micro Computed Tomography Data Acquisition and Reconstruction Overview", March 22, 2019. Note: This document discusses the design and functioning of the MCT system, measurements, and datasets curated for this SE296 study.
15. Montgomery, Douglas C. *Introduction to statistical quality control*. John Wiley & Sons, 2020. Chapter 10 discusses serial correlation and statistical control charts.
16. Montgomery, Douglas C., Cheryl L. Jennings, and Murat Kulahci. *Introduction to time series analysis and forecasting*. John Wiley & Sons, 2015. Note: Chapter 5 discusses the steps for building Autoregressive Integrated Moving Average ARIMA models.
17. Nahrstedt, Klara, et. Al. "Maintlet: Advanced Sensory Network Cyber-infrastructure For Smart Maintenance In Campus Scientific Laboratories", website available online at: https://monet.cs.illinois.edu/maintlet-advanced-sensory-network-cyber-infrastructure-for-smart-maintenance-in-campus-scientific-laboratories/
18. NIST Engineering Statistics Handbook, available online at: https://www.itl.nist.gov/div898/handbook/ Note: Best practices for the application of statistical methods for engineers and scientists. Discusses the tolerance intervals used by the MCT technical team.
19. Polhemus, Neil W. *Process capability analysis: estimating quality*. Chapman and Hall/CRC, 2017
20. Venkatasubranmanian, V., et al. "A review of process fault detection and diagnosis, part iii: Process history-based methods." *Comput. Chem. Eng* 27.3 2003 327 346
21. Wald, Abraham. *Sequential analysis*. Courier Corporation, 2004. Note: The standard reference on implementing the sequential probability ratio test for hypothesis testing.
22. Witten, Ian H., et al. "Practical machine learning tools and techniques." Data Mining. Vol. 2. No. 4. 2017.

Appendix A: Fundamental Axioms of Structural Health Monitoring

The Fundamental Axioms appear in the standard reference for SHM design engineering, *Structural health monitoring: a machine learning perspective* [Farrar & Worden, 2012]. The axioms encapsulate general design principles that have come to be accepted throughout the SHM research community. The axioms are referenced throughout this report and are reproduced here for the benefit of the reader.

Axiom I. All materials have inherent flaws or defects.

Axiom II. Damage assessment requires a comparison between two system states.

Axiom III. Identifying the existence and location of damage can be done in an unsupervised learning mode, but identifying the type of damage present and the damage severity can generally only be done in a supervised learning mode.

Axiom IVa. Sensors cannot measure damage. Feature extraction through signal processing and statistical classification are necessary to convert sensor data into damage information.

Axiom IVb. Without intelligent feature extraction, the more sensitive a measurement is to damage, the more sensitive it is to changing operational and environmental conditions.

Axiom V. The length and time scales associated with damage initiation and evolution dictate the required properties of the SHM sensing system.

Axiom VI. There is a trade-off between the sensitivity to damage of an algorithm and its noise rejection capability.

Axiom VII. The size of damage that can be detected from changes in system dynamics is inversely proportional to the frequency range of excitation.

Axiom VIII. Damage increases the complexity of a structure.

Appendix B: MATLAB Codes

This Appendix publishes the core MATLAB source codes developed during the SE296 project. The intent is to provide reference implementations for various statistical calculations that will be of general technical interest. All of the project source codes will be transferred to the LLNL MCT technical team including those utility functions omitted from this Appendix.

Main.m

```
% main.m - generate plots for SE296 MCT Study
% Plot 0 - Recreate Code provided with MCT Presentation
LAC_Figure("MCT Data\LAC_141117_180830.xlsx");
%% 1A - Utility function: Import LAC Spreadsheet into MCTData.mat
clc; close all; clear;
% SKEATE; modified spreadsheet sdate so MATLAB could read it
Import_MCT_Dataset("MCT Data\LAC_141117_180830.xlsx", "MCT Data\system_141117_180830_SKEATE.xlsx");

% sampleNumbers 443x1, LAC100 443x6, LAC160 443x6, References 1x6, Tolerances100 2x6,
% Tolerances160 2x6, Timestamps 443x1, TDelta 443x1
load("MCTData.mat");

disp("done");
%% 2. Plot Multiple Control Charts with Sample # along X-Axis and Time Since Last Scan as an Overlay
clc; close all; clear; load("MCTData.mat"); % reset environment

dataRange = 1:6;

% Plot 2a Control Chart with Time Difference between Samples as an Overlay
H = Plot_Control_Chart("Control Chart for References at 100kV", sampleNumbers, LAC100(:,dataRange), ...
    References(dataRange));
Overlay_TDelta (H, sampleNumbers, TDelta);

%% 2.1 Plot Single or Multiple Control Charts with Scan date along X-Axis
clc; close all; clear; load("MCTData.mat"); % reset environment

% Shared settings
dataDomain = 1:size(LAC100,1);  plotIndices = 1:6;  numplots = size(plotIndices,2);
data100_upper = [0.0338 0.0305 0.0231 0.0494 0.0646 0.106];
data100_lower = [0.03365 0.0303 0.0229 0.049 0.0635 0.104];
data160_upper = [0.0271 0.023 0.017 0.0324 0.0309 0.0448];
data160_lower = [0.0269 0.0228 0.01685 0.0321 0.0306 0.0446];


% Plot 2a Control Chart with Time Difference between Samples as an Overlay
H1 = Create_Chart_v2("Control Chart for Reference LAC values at 100kV", numplots);
Plot_Control_Chart_v2(H1, TimeStamps, LAC100, References, Tolerances100, dataDomain, ...
    plotIndices, data100_upper, data100_lower);
Plot_LAC_Bounds(H1, LAC100(1:30, :), plotIndices); % use samples 1:30 to calculate LAC bounds
Add_Operating_State_To_Plot(H1);

H2 = Create_Chart_v2("Control Chart for Reference LAC values at 160kV", numplots);
Plot_Control_Chart_v2(H2, TimeStamps, LAC160, References, Tolerances160,dataDomain, ...
    plotIndices, data160_upper, data160_lower);
Plot_LAC_Bounds(H2, LAC160(1:30, :), plotIndices); % use samples 1:30 to calculate LAC bounds
Add_Operating_State_To_Plot(H2);


```

Main.m (continued)

```
%% Plot 2.3 Plot Individual Charts with an Overlay of Days Since Last Operation
clc; close all; clear; load("MCTData.mat"); % reset environment

for i = 1:6
 H = Plot_Control_Chart("Control Chart for References at 100kV", sampleNumbers, LAC100(:,i), References(i));
 Overlay_TDelta (H, sampleNumbers, TDelta);
end

%% Plot 2 - xMR Chart: use first 200 samples to set limits (based on Control Chart)
clc; close all; clear; load("MCTData.mat"); % reset environment

ylimits_low = [0.0336];
ylimits_high = [0.0338];

dataRange = 1:1;

% Plot 2b xMR Chart: use first 200 samples to set limits (based on Control Chart)
[H,H2] = Plot_xMR_Chart("mR", sampleNumbers, LAC100(:,dataRange), References(dataRange), Tolerances100(:,dataRange), ...
    ylimits_low, ylimits_high);
Overlay_TDelta (H, sampleNumbers, TDelta);
Overlay_TDelta (H2, sampleNumbers, TDelta);

%% Plot 3 - Wald Sequential Probability Ratio Test (SPRT)
clc; close all; clear; load("MCTData.mat"); % reset environment

% Calculate mean and standard deviation over the calibration period (first 30 measurements)
calibration_samples = 1:30; % graphite mu = 0.0337468
mu_H0 = mean(LAC100(calibration_samples,:),1); % first 200 values are in control
sigma = std(LAC100(calibration_samples,:),1);

% Experiment with different decision thresholds
% alpha=0.05; beta=0.2;  % [a:-1.558, b:2.773]  original 17Dec22
% alpha=0.05; beta=0.1;  % [a:-2.251, b:2.890]
 alpha=0.01; beta=0.1;  % [a:-2.293, b:4.500
%  alpha=0.01; beta=0.2;  % [a:-1.599, b:4.382], baseline 18Dec22

deviation = [1 1 1 1 1 1];

% H1: mu > mu_0.  Test for mean is drifting higher
mu_H1 = mu_H0 + (deviation .* sigma);

dataRange = 1:314;  % index 314 corresponds to Jan 2018 (post Abnormal operation)

[result, a, b] = Detect_Drift_Using_SPRT(LAC100(dataRange,:),mu_H0, sigma, mu_H1, alpha, beta);
titleString = sprintf("SPRT Hypothesis Test for [alpha, beta]=[%0.2f,%0.2f]: H0=mean steady, H1=mean increasing", alpha, beta);
plotHandle = Plot_SPRT_Chart(TimeStamps(dataRange), result, a, b, titleString, References, mu_H0, deviation, sigma, Tolerances100(2,:));
Add_Operating_State_To_Plot(plotHandle);

% H1: mu < mu_01 Test for mean is drifting lower
mu_H1 = mu_H0 - (deviation .* sigma);

[result, a, b]  = Detect_Drift_Using_SPRT(LAC100(dataRange,:),mu_H0, sigma, mu_H1, alpha, beta);
titleString = sprintf("SPRT Hypothesis Test for [alpha, beta]=[%0.2f,%0.2f]: H0=mean steady, H1=mean decreasing", alpha, beta);
plotHandle = Plot_SPRT_Chart(TimeStamps(dataRange), result, a, b, titleString, References, mu_H0, -deviation, sigma, Tolerances100(1,:));
Add_Operating_State_To_Plot(plotHandle);

disp("done");
```

Main.m (continued)

```
%% Plot 4 - Overlay SPRT Alarms on Control Chart (combine 2.1 and 3)
clc; close all; clear; load("MCTData.mat"); % reset environment

% Calculate alarms
calibration_samples = 1:30;
[mu_H0,sigma] = Calculate_Data_Statistics(LAC100(calibration_samples, :));
alpha=0.01; beta=0.1;  % [a:-2.293, b:4.500
deviation = [1 1 1 1 1];
% H1: mu < mu_0
mu_H1 = mu_H0 - (deviation .* sigma);
[result, a, b] = Detect_Drift_Using_SPRT(LAC100(:,:),mu_H0, sigma, mu_H1, alpha, beta);

% Plot 2a Control Chart with Time Difference between Samples as an Overlay
dataRange = 1:size(LAC100,1);  plotIndices = 1:6;  numplots = size(plotIndices,2);
data100_upper = [0.0338 0.03046 0.0231 0.0494 0.0646 0.106];
data100_lower = [0.03365 0.0303 0.02298 0.049 0.0639 0.1046];

dataRange = 1:314;  % index 314 corresponds to Jan 2018 (post Abnormal operation)

H1 = Create_Chart_v2("Control Chart for Reference LAC values at 100kV", numplots);
Plot_Control_Chart_v2(H1, TimeStamps(dataRange), LAC100(dataRange,:), References, Tolerances100, dataRange, plotIndices, ...
    data100_upper, data100_lower);
%Plot_LAC_Bounds(H1, LAC100(dataRange, :), plotIndices); % use samples 1:30 to calculate LAC bounds
% ylim(H1,[-1 1]);

Highlight_Alarms(H1, TimeStamps(dataRange), LAC100(dataRange, :), result(dataRange,:,:), data100_upper, data100_lower);
Add_Operating_State_To_Plot(H1);

disp("done");

%% Plot 5 - Scatter plot showing autocorrelation
clc; close all; clear; load("MCTData.mat"); % reset environment

dataDomain = 1:size(LAC100,1);  plotIndices = 1:6;  numplots = size(plotIndices,2);
data100_upper = [0.0338 0.0305 0.0231 0.0494 0.0646 0.106];
data100_lower = [0.03365 0.0303 0.0229 0.049 0.0635 0.104];
data160_upper = [0.0271 0.023 0.017 0.0324 0.0309 0.0448];
data160_lower = [0.0269 0.0228 0.01685 0.0321 0.0306 0.0446];

H1 = Create_Chart_v2("Scatter Plot for Reference LAC values at 100kV", numplots);
lag = 2;
Plot_Scatter_Chart_v1(H1, lag, LAC100, References, plotIndices, TDelta, data100_upper, data100_lower);

%% Compute Autocorrelation coefficients at K=50 lags
clc; close all; clear; load("MCTData.mat"); % reset environment
xdata=LAC100; K=50;
rho = calculate_AutoCorrelation_Function (xdata, K);
% rho = calculate_AutoCorrelation_Function (xdata(80:110,:), K);
T = array2table(horzcat((1:length(rho))' , rho),'VariableNames',horzcat('lag',References));
disp(T)
```

Main.m (continued)

```
%% Compute Autocorrelation coefficient at k=1 over rolling window of 30
clc; close all; clear; load("MCTData.mat"); % reset environment

% remove outliers
LAC100(31,1) = LAC100(29,1);
LAC100(35,1) = LAC100(32,1);
LAC100(36,1) = LAC100(33,1);
LAC100(37,1) = LAC100(34,1);

xdata=LAC100; K=50;

calculate_AutoCorrelation_Function_window = @(wstart) calculate_AutoCorrelation_Function (xdata(wstart:(wstart+50),1),1);

rho = arrayfun(calculate_AutoCorrelation_Function_window, 1:343);
plot(rho);
%% Plot residuals for AR(1) model; run previous section to compute rho
close all;
dataRange = 1:312;  % index 314 corresponds to Jan 2018 (post Abnormal operation)

% remove outliers
LAC100(31,1) = LAC100(29,1);
LAC100(35,1) = LAC100(32,1);
LAC100(36,1) = LAC100(33,1);
LAC100(37,1) = LAC100(34,1);

% AR(1)
a = (1 - rho(1,:)) .* mean(LAC100(1:200,:));
estimates = a + rho(1,:) .* LAC100(1:end-1,:);
residuals_all = LAC100(2:end,:) - estimates;
residuals = residuals_all(dataRange,:);
residuals_smooth = movmean(residuals, 5, 2);

% AR(2)
% a = (1 - rho(1,:) - rho(2,:)) .* mean(LAC100(1:200,:)) ;
% estimates = a + rho(1,:) .* LAC100(2:end-1,:) + rho(2,:) .* LAC100(1:end-2,:);
% residuals_all = LAC100(3:end,:) - estimates;
% residuals = residuals_all(dataRange,:);

TimeStamps = 1:443;  % plots show better detail
H1 = Create_Chart_v2("X-Bar Control Chart of AR(1) Residual Errors for Reference LAC values at 100kV", 6);

%sqr = @(i) stem(H1(i), TimeStamps(2:length(residuals)+1), residuals(:,i),"|");
sqr = @(i) stem(H1(i), TimeStamps(2:length(residuals_smooth)+1), residuals_smooth(:,i),"|");
addTitles =@(i) title(H1(i,1), References(i));

%sqr(2, H1);
arrayfun(sqr, 1:6);
%Plot_LAC_Bounds(H1, residuals(1:30,:), 1:6, 6);
Plot_LAC_Bounds(H1, residuals_smooth(1:30,:), 1:6, 6);
arrayfun(addTitles, 1:6);
Add_Operating_State_To_Plot(H1, false);
```

Detect_Drift_Using_SPRT.m

```
function [result, a, b] = Detect_Drift_Using_SPRT(X,mu_H0,sigma, mu_H1, alpha, beta)
% Detect system drift using Wald Sequential Probability Ratio Test
% Inputs:
%  X: N x 6, time series of LAC measurement values for the six reference materials
%  mu_H0 1x6, mu_H1 1x6: mean of LAC calibration values (for six references) used for null H0 and alternate H1 hypothesis
%  sigma 1x6: standard deviation of LAC calibration values (for six references)
%  alpha 1x1: significance level or probability of Type 1 error
%  beta 1x1: probabililty of Type 2 error
% Ouputs:
%  result, N x 2 x 6
%  result(:,1,:) = sum of likelihoods since last alarm for all time series measurements and all six references
%  result(:,2,:) = 1 if H1 alarm, -1 if H0 alarm, otherwise 0
%  a,b: Wald decision boundaries, both scalar 1x1

% Wald upper and lower decision boundaries
a = log(beta / (1 - alpha)); % Accept null, reject H1
b = log((1 - beta) / alpha); % Accept H1, Reject null

fprintf("Optimal bounds: [a:%.3f, b:%.3f]", a,b)

nsamples = size(X,1);  % N, the number of measurement samples in the time series

nreferences = size(X,2);  % 6, the number of reference materials

result = zeros(size(X,1),2,size(X,2));  % [SPRT_sum , alarm (1,0,-1) , Reference]
likelihood_count_index = ones(nreferences,1);   % start all counts at 1

for i=1:nsamples

    for j=1:nreferences
        % Calculate rslt = sum of likelihoods since last alarm
        rslt = calculate_LogLikelihood(X(likelihood_count_index(j,1):i,j),mu_H0(j), sigma(j), mu_H1(j));

        result(i,1,j) = rslt;

        if rslt >= b
            result(i,2,j) = 1;  % record H1 alarm
            likelihood_count_index(j) = i + 1; % restart likelihood count on next sample

        elseif rslt <=a
            result(i,2,j) = -1;  % record H0 alarm
            likelihood_count_index(j) = i + 1; % restart likelihood count on next sample
        end

    end

end

end
```

calculate_LogLikelihood.m

```
function logLikelihood = calculate_LogLikelihood(X, mu_H0, sigma, mu_H1)
% Calculate log-likelihood used in the computation of the Wald sequential probability ratio test
% Assumes the statistical test is fully specified and X is Gaussian distributed
   sum_H0 = sum((X - mu_H0).^2,1);
   sum_H1 = sum((X - mu_H1).^2,1);

   logLikelihood = (sum_H0 - sum_H1) ./ (2*sigma.^2);
end
```

Plot_SPRT_Chart.m

```
% Inputs: timeStamps, SPRT result, Wald a and b decision boundaries, etc
function [H] = Plot_SPRT_Chart(timeStamps, result, a, b, chartTitle, References, mu_H0, deviation, sigmas, tolerances)

% H = Create_Chart(chartTitle, sampleNumbers, X, References);
numplots = 6;
H = Create_Chart_v2(chartTitle, numplots);

% Calculate tolerance band in sigmas
tolerance_bound_in_sigmas = (tolerances - mu_H0) ./ sigmas;

for i=1:size(H,1)

    clip_result = min(result(:,1,i), 5);  % clip outliers to make better use of plot real estate
    clip_result = max(clip_result, -5);  % clip outliers to make better use of plot real estate

    plot(H(i),timeStamps,clip_result,'o','MarkerSize',3,'LineWidth',.5,'MarkerEdgeColor','b',...
        'MarkerFaceColor','b');


    hold(H(i));

    alarm_H1 = result(:,2,i) > 0;
    scatter(H(i),timeStamps(alarm_H1)',clip_result(alarm_H1,1),'r');


    alarm_H0 = result(:,2,i) < 0;
    scatter(H(i),timeStamps(alarm_H0)',clip_result(alarm_H0,1),'g');


    yline(H(i), 0);
    yline(H(i), [b a], '--', {'Accept H1', 'Accept H0'}, 'LabelHorizontalAlignment', 'left');  % yline(H(i), a, 'Accept H0');

    ylim(H(i),[-5.5 5.5]);


    chartTitle = sprintf("%s H0:[%0.7f] H1 at %0.1f sigma Tolerance at %0.2f sigma", References(i), mu_H0(i),deviation(i), ...
        tolerance_bound_in_sigmas(i));
    title(H(i,1), chartTitle);  % need to add title here as ylim seems to erase if done in Create_Chart
end

end
```

Calculate_Data_Statistics.m

```
% Calc mean and sdev then plot bounds at +- 3 sdev
function [m,sdev] = Calculate_Data_Statistics(data)
m = mean(data, 1);
sdev = std(data,1);
end
```

calculate_AutoCorrelation_Function.m

```
function rho = calculate_AutoCorrelation_Function (xdata, K)

rho = zeros(K,size(xdata,2));

% x = xdata(1+lag: end, :);
% xlag = xdata(1:end-lag, :);
% numerator = (x-m) .* (xlag-m);

m = mean(xdata);   % m = sum(xdata) / length(xdata);
numerator = @(lag) sum((xdata(1+lag: end, :) - m) .* (xdata(1:end-lag, :) - m));

% lags = 1:K;
% numerators = numerator(1);

denominator = sum((xdata-m) .* (xdata-m));

for lag = 1:K
  rho(lag,:) = numerator(lag) ./ denominator;
end

% v = var(xdata);
%
% v2 = sum((xdata - m).^2) / (length(xdata)-1);
%
% v3 = sum((xdata - m).^2) ;

end
```