

Draft 6/20/2000 To be submitted to Spectroscopy

Chemometric Analysis of Nuclear Magnetic Resonance Spectroscopy Data

Todd M. Alam¹ and M. Kathleen Alam²

¹ Department of Organic Materials,

² Information Extraction, Detection and Analysis Department,
Sandia National Laboratories, Albuquerque, NM 87185-1407

RECEIVED
AUG 17 2000
OSTI

Chemometric analysis of nuclear magnetic resonance (NMR) spectroscopy has increased dramatically in recent years. A variety of different chemometric techniques have been applied to a wide range of problems in food, agricultural, medical, process and industrial systems. This article gives a brief review of chemometric analysis of NMR spectral data, including a summary of the types of mixtures and experiments analyzed with chemometric techniques. Common experimental problems encountered during the chemometric analysis of NMR data are also discussed.

M. Kathleen Alam and Todd M. Alam are principal members of the technical staff at Sandia National Laboratories and can be contacted at mkalam@sandia.gov or tmalam@sandia.gov.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

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

Introduction

The analysis of complex mixtures using NMR spectroscopy continues to impact a variety of different fields including agriculture, industry, medicine, materials, chemistry and biochemistry. While the use of numerical techniques is well founded in the analysis of NMR data (For example see Ref. (1), and references cited therein), the use of chemometric multivariate techniques for NMR has seen rather limited use in comparison to other spectroscopies. In this context we will refer to chemometrics as any multivariate technique that extracts chemically relevant information from data produced in chemical experiments. With the utilization of NMR spectroscopy to investigate increasingly more complex problems, advanced chemometric techniques for NMR data analysis will become an important part of the analysis arsenal. The production of very large and/or complex data sets are becoming more common with the introduction of conjugated analytical techniques (such as LC-NMR and LC-MS-NMR), the increased use of automated and/or flow NMR for process control, and the NMR analysis of combinatorial synthesis experiments. In addition, the recent advances in computational power, plus the increased ease of data transfer between computer systems now makes the suite of chemometric techniques more available to the routine user. In this review the use of chemometrics as applied to NMR spectroscopy will be presented, along with a few specific examples to illustrate the general concepts.

The chemometric analysis of NMR spectra begins with the assumption that the observed data set of i spectra, $A_i(\omega)$, covering a frequency range ω , is a linear combination of k pure component spectra, $P_j(\omega)$, with concentrations (intensities) C_{ij} given by

$$A_i(\omega) = \sum_{j=1}^k C_{ij} P_j(\omega) \quad (1)$$

This can be expressed in matrix form as

$$\mathbf{A} = \mathbf{C}\mathbf{P} \quad (2)$$

For example, a data set containing spectra of 32 mixtures ($i = 32$), where each NMR spectra has 8096 (8K) real points (frequencies), resulting from the combination of four different pure component spectra ($k = 4$), the dimensions of the various matrices in Eqns. 1 and 2 would be $32 \times 8K$ for \mathbf{A} , 32×4 for \mathbf{C} and $4 \times 8K$ for \mathbf{P} . A visual representation of these matrices is shown in Figure 1. While this example utilizes the actual NMR spectra for the matrix \mathbf{A} , it should be noted that there are numerous ways to create a multivariate data set from NMR data for possible chemometrics analysis. Examples include; 1) the use of simple peak intensities, 2) number of peak numbers within a given spectral region, or 3) the total spectral integral for a defined spectral region. In the sections below the form of data used in creation of the multivariate data matrix for non-standard situations will be noted.

Chemometric techniques that allow for calculation of \mathbf{C} and \mathbf{P} from a set of NMR experimental spectra \mathbf{A} are of great interest and will be the focus of this review. The numerical solutions to the above matrix relationship (Eqs. 1 and 2) are not unique, and can give rise to abstract factors (spectra-like eigenvectors) or scores (concentration-like eigenvalues), that are linear combinations of the pure spectra and pure concentration

profiles. In many instances, these abstract solutions are sufficient for the identification or classification of the chemical species present within the sample, and can be used without further analysis. This is the situation commonly encountered during principal component analysis (PCA), and is detailed in the next session. Techniques allowing the identification of real or pure component spectra and concentration profiles from these abstract solutions will be discussed in the subsequent self-modeling and three-way methods section.

Principal Component Analysis (PCA)

Principal component analysis (PCA) refers to the transformation of the data into an orthogonal basis set. The variance described by the basis vectors is largest in the first vector and decreases with additional vectors. The data matrix **A** (Eqns. 1 and 2), with i samples, each sample containing n frequency points ($i < n$), can be interpreted as an ensemble of i points in a n dimensional space. PCA is the process of obtaining a series of "lines and planes of closest fit to a system of points in space".(2) The closest fit in this case is the least squares fit. The vectors within this new basis set are referred to as PCA loadings and scores. If the noise within the data is randomly varying, it will be contained in the later loadings and scores of the PCA decomposition. Estimating the data set by using only significant PCA loadings can effectively filter random noise. Without additional constraints, there are an infinite number of mathematical solutions to the eigen problem defined in Eqs. 1 and 2.

These "abstract" solutions are linear combinations of the chemically correct pure component spectra, and for many applications are sufficient for chemical classifications and identification. Principal component (PC) scores plots will occasionally reveal clusters

of related samples allowing sample identification and discriminations. A plot of a three-dimensional PC space is shown in Figure 2, where distinct groupings are clearly evident and could provide a means of sample segregation. Different variations of PCA are available depending on the information desired in the resulting loadings. If the experimental data matrix A is mean-centered, this form of PCA is known as the covariance method, and resulting loadings retain the scale of the original spectral data. If the data in A is mean-centered and normalized, this form of PCA is referred to as the correlation method. In this method the spectral scaling is not retained, with all spectral features influencing the PCA equally, such that minor spectral features can have significant impact.(3) Several chemometric methods are available that will mathematically discriminate the scores plots, placing boundaries and statistical limits on the visual grouping. Often, the group is composed of more than three eigenvectors and cannot be discerned visually. Methods such as the Mahalanobis distance,(4) Linear Discriminant Analysis (LDA),(5) and Soft Independent Method of Class Analogy (SIMCA), (6) can provide delineation of each group. PCA has been used widely through the chemical sciences for pattern recognition including gas chromatography,(7) mass spectroscopy,(8) and throughout the infrared and near infrared community.(9)

PCA has been applied to the analysis of ^1H NMR spectra of apple juices,(3) olive oils,(10) wheat,(11) gasoline,(12) and complex biofluids such as urine.(13) PCA has also been applied to the analysis of ^{13}C NMR of cellulose,(14) pulp fibers,(15) pulp kinetics,(16) olive oils,(17) the attempted classification of soil types,(18) along with the ^{13}C and ^{31}P NMR spectra of kraft black liquor and dissolved lignin.(19) Most of these analyses utilized either integration for well defined spectral regions or the actual NMR

spectra in the formation of the multivariate matrix **A**. By using the intensity of cross-peak correlation in the construction of the multivariate data set, PCA analysis of ^1H - ^{13}C heteronuclear two-dimensional (2D) NMR data sets were used to differentiate grapevine cultivars.(20) PCA analysis of peak intensities from 2D NMR correlation spectroscopy (COSY) experiments has been reported for the pattern recognition of complex spin systems.(21) Coupling of a generalized Bayesian approach with LDA has been used for automated signal class recognition.(22) A multivariate data set built from different relaxation rates, (^1H rotating frame spin-lattice relaxation and the dipolar dephasing relaxation rate), was used for the PCA analysis of solid-state coal samples.(23)

Self-Modeling and Three-Way Techniques

The abstract eigenvectors and eigenvalues from PCA require an additional transformation step to produce real or pure component spectra and concentration profiles for the individual species within the mixture. Self-modeling and three-way modeling analysis techniques allow this transformation to be evaluated without resorting to known standards, and as such, are powerful tools in the analysis of mixtures. Because determination of the pure component spectra allows a more rigorous quantitative analysis of the analyzed mixtures, these self-modeling and three-way method analysis techniques will be discussed in more detail. Utilization of chemical and physical constraints allows these transformation matrices to be determined. For example, pure component solutions can be obtained by requiring such constraints as positive spectra amplitudes, positive concentrations, non-negative diffusion decay profiles, single peak pure components, or unimodal character of concentration profiles. Examples include the Alternating Least-

Squares (ALS) optimization and Multi-Curve Resolution (MCR) analysis of DOSY and pulse-field gradient (PFG) spectra of polymer mixtures,(24) and the global optimization procedure named CORE-NMR (COmponent-REsolved NMR) of pulsed-gradient spin-echo (PGSE) data sets.(25, 26) CORE is a global fitting procedure that utilizes a two level optimization approach for every significant frequency channel in the NMR spectrum. It requires that the relaxation modulation function be defined (multi-exponential or other functionals can be used), that individual components do not overlap or that the components overlap preferentially. In the analysis of PGSE data sets CORE NMR was found to be very robust and stable.(26)

The solutions obtained are only as valid as the imposed constraints, and in many instances can provide ambiguous results or non-unique solutions. Interestingly it has been shown that if two data sets can be obtained that are proportional to each other, only the correct "pure" solution is obtained from the analysis of this eigenvalue problem,(27) without resorting to the use of imposed constraints. Since a single spectra is a vector or one-way array, and a matrix is a two-way array, then the analysis of these proportional data sets (and in general an infinite set of combined matrices) involves three-way arrays, these techniques are commonly referred to as *three-way methods*.

A second data set proportional to Eqn. 2 can be described by

$$\mathbf{B} = \alpha \mathbf{C} \mathbf{P} \quad (3)$$

where α is a diagonal matrix containing the proportionality factor α_{ik} . Use of three-way methods to analyze these two proportional NMR data sets (**A** and **B**) results in only a

single solution for the pure components and concentration profiles. The creation of two proportional data sets can be obtained by either analyzing two different sets of samples fulfilling Eqns. 2 and 3, or by analyzing a single sample under varying experimental conditions. This type of proportionality can be applied to a range of different analysis techniques, including fluorescence excitation and emission, modulated infrared spectroscopy, plus linear and circular dichroism.(27) The elegant work of Windig, Antalek and co-workers,(28-32) as well as Schulze and Stilbs,(33) have demonstrated that for NMR, this type of analysis can be applied to PGSE diffusion matrices, magnetic resonance imaging (MRI) data, and spin-lattice relaxation matrices.

For exponential relaxation processes in NMR, spectra obtained for equally spaced time intervals satisfy the α proportionality in Eqns. 2 and 3. As demonstrated in Figure 3 for an ^{17}O NMR spin-lattice relaxation matrix of equally spaced time delays, the analysis matrices **A** and **B** can also be out of a single data set. For matrix **A** the spectra 1 through $n-1$ are used, while for matrix **B** spectra 2 through n are utilized. A major advantage of being able to create both **A** and **B** from a single data set is the elimination of variations in the instrumental response and stability during the collection of the proportional data sets. By solving Eqns. 2 and 3 for $C\alpha$

$$\begin{aligned} C\alpha &= AP^+\alpha \\ C\alpha &= BP^+ \end{aligned} \tag{4}$$

an equation resembling the generalized eigenvector equation is obtained.

$$\mathbf{A}\mathbf{P}^+\alpha = \mathbf{B}\mathbf{P}^+ \quad (5)$$

This equation can be solved directly as the solution to the generalized eigenvector problem, if the matrices **A** and **B** are square; usually they are not, and steps must be taken to transform **A** and **B** such that Eqn. 5 can be solved. The solution of Eqn. 5 allows the computation of **C** and **P**, the concentration and pure component spectral profiles, respectively.

A method utilizing the direct exponential curve resolution algorithm (DECRA) has been used to analyze NMR data type of this kind,(28-31, 34) where DECRA is based on the generalized rank annihilation algorithm.(35) As an example, the DECRA analysis of the spin-lattice relaxation matrix for the solution ^{17}O NMR spectra of a two-component mixture of 3-methyl propanol and ethanol is shown in Figure 4 and 5. The rapid quadrupolar relaxation of the ^{17}O nucleus results in large line widths producing significant spectral overlap. In Figure 4 the amplitude of the individual *i* components are modulated by the spin lattice relaxation time (T_{1i}), and provide the data set from which the matrices **A** and **B** are formed. The resulting pure component spectra and residuals are shown in Figure 5. More details of this analysis are given in Ref. (34). While this example may appear to be rather simplistic since the components are visually resolved and should be amenable to conventional deconvolution techniques, we have recently demonstrated that significant errors result from the introduction of assumptions about the form of the line shape.(34) DECRA makes no assumptions or constraints about the actual spectral line shape required, and that a significant reduction in residuals is realized. The DECRA algorithm does require that the number of factors (pseudorank) be known.

Attempts to extract more factors than are actually in the data set results in a new component spectra that contains only noise, and can be used as a marker for the determination of the pseudorank.(28)

Partial Least Squares (PLS) and other techniques

Whereas the PCA algorithm uses only the spectral information to derive the model, the resulting loading vectors may not be optimal for concentration prediction. Partial Least Squares (PLS) use the concentration information during calibration to place more useful information for prediction in the first several loading vectors.(36) The PLS algorithm has been extensively used in the chemical sciences, including mass spectroscopy,(37) gas chromatography,(38) and throughout the infrared and near infrared community.(39-41)

PLS algorithms have also been utilized in the analysis of NMR data sets spanning a wide range of applications. PLS analysis of solid-state ^{13}C magic angle spinning (MAS) NMR has been used for the characterization of wood pulp, including the prediction of carbohydrate constituents,(42) lignin content,(16) and alkali resistance.(43) High resolution ^{13}C and ^{31}P NMR has also been analyzed using PLS to predict the combustion properties of softwood and hardwood kraft black liquors.(19) PLS calibration of ^1H NMR has been reported for the analysis of octane number in gasoline,(12) and the 5-day biological oxygen demand (BOD_5) in industrial wastewater.(44) PLS data analysis has also been used as a tool in NMR shift assignments.(45) Principle Component Regression (PCR) and PLS methods have also been applied to the discrimination of olive oil variety using ^{13}C NMR.(17) A combination of PLS and Net Analyte Signal (NAS) analysis has

also been recently reported for the investigation of alcohol mixtures using solution ^{17}O NMR.(46) By identifying those constituents that interfere spectrally using NAS analysis, improved PLS correlations could be obtained. PLS analysis of the NMR free induction decay (FID) have also been reported, including the ^1H analysis of moisture content and basic density in softwoods,(47) the ^1H study of moisture content in meat products,(48) and the analysis of model process NMR data.(49)

Common Experimental Problems

One of the early difficulties encountered with the chemometric analysis of NMR data was the size of the spectral data sets (ω in Fig. 1) which commonly ranged from 1K (1024) real spectral points to 64K (65536) spectral points. The computational difficulty associated with the large data size was further compounded by the slow CPU speed of early computer systems (in comparison to today's systems) used on NMR instruments, along with the difficulty in exporting NMR data to external computational facilities. Advances in computer technology have all but eliminated these problems, with the dramatic increase in today's computation speed making the analysis of NMR data sets easily manageable. The development of improved interface hardware, along with standardized data structure protocols have made data transfer to other computer systems routine. In addition, most of today's NMR instrument manufacturers now employ third party Unix and NT based computer systems to control the instrument, such that chemometrics analysis software can be directly employed on the data sets.

Other difficulties encountered in the analysis of NMR data include the baseline distortions and variations in the spectral phasing between spectra. Baseline distortions,

including the annoying baseline roll, have a large detrimental impact on chemometric analysis, but are commonly eliminated using either manual or automatic baseline fitting routines, or the implementation of linear prediction techniques. Small changes in the spectra phase within the multivariate data matrix can also influence the chemometric analysis. Small phase differences can be eliminated by very careful manual phasing,(13, 14) use of improved automatic phase routines, or by converting the observed spectra to magnitude spectra [$S(\omega) = (\text{Real}^2 + \text{Imaginary}^2)^{1/2}$].(43) Phasing and baseline distortions also affect the quality of resonance integration.(17) For many of the chemometric studies mentioned in this article, reliable integration values were obtained by using only a small portion of the entire NMR spectra during analysis,(42) along with using well-defined frequency ranges.(13, 17, 19) Variations in the integration can also be reduced by use of more complicated deconvolution techniques. Variations in the NMR instrumental response, excitation efficiency, magnetic field or sample homogeneity or differences in spin-lattice relaxation rates also have a large impact effect on the chemometric analysis. Scaling of the total spectral intensity to a constant value,(14, 15) scaling to a distinct spectral resonance,(11) scaling to an internal reference, or using relative areas(11, 43) helps eliminate the errors caused by the instrumental or excitation variations. Careful attention to relaxation effects, including Nuclear Overhauser Effects (NOE) effects,(17) will result in an improved chemometric analysis.

The most prevalent problem encountered in the NMR-chemometric analysis of complex mixtures is the solvent, concentration, pH and temperature induced shifts of the observed resonances, especially apparent at high magnetic field strengths. This difficulty has been dealt with in a variety of different ways, including the use of standardized peak

referencing,(3, 43) or the reduction of spectral resolution through smoothing or data compression.(3) Using integration areas for a defined spectral region in lieu of actual spectral intensity data is also a common processing technique for overcoming solvent shift effects.(13)

Summary

The use of chemometrics for the analysis of NMR data sets allows more complex systems and mixtures to be addressed. While there are still many difficulties encountered in the analysis of NMR data, chemometric analysis has clearly been shown to be a viable technique. With the continued improvements in analysis software along with the increasing automation of NMR instruments, the implementation of chemometric techniques to NMR spectral data will become more common.

Acknowledgements

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94A185000.

References

- (1) D.N. Rutledge, ed. *Signal treatment and signal analysis in NMR*. Data Handling in Science and Technology, ed. B.G.M. Vandeginste and S.C. Rutan. Vol. 18. 1996, Elsevier: Amsterdam. 548.
- (2) S. Wold, K. Esbensen and P. Geladi, *Chemometrics and Intell. Lab. Systems* **2**, 37-52 (1987).
- (3) P.S. Belton, I.J. Colquhoun, E.K. Kemsley, I. Delgadillo, P. Roma, M.J. Dennis, M. Sharman, E. Holmes, J.K. Nicholson and M. Spraul, *Food Chemistry* **61**(1/2), 207-213 (1998).
- (4) H. Mark, *Anal. chem.* **59**, 790-795 (1987).
- (5) B.R. Stallard, *Applied Spectroscopy* **51**, 625- (1997).
- (6) O. Svensson, M. Joefson and F.W. Langkilde, *Applied Spect.* **51**, 1826- (1997).
- (7) H. Wilken and H.R. Schulten, *Fresenius J. Anal. Chem.* **355**, 157-163 (1996).
- (8) K.D. Zissis, S. Dunkereley and R.G. Brereton, *Analyst* **124**, 971-979 (1999).
- (9) C. Robert and F. Cadet, *Applied Spectroscopy Reviews* **33**, 253-266 (1998).
- (10) G. Vlahov, A.D. Shaw and D.B. Kell, *J. Am. Oil Chem. Soc.* **76**, 1223-1231 (1999).
- (11) A. Sacco, I.N. Bolsi, R. Massini, S. Manfred, E. Humpfer and S. Ghelli, *J. Agric. Food Chem.* **46**, 4242-4249 (1998).
- (12) R. Meusinger, *Fuel* **75**(10), 1235-1243 (1996).

- (13) M. Spraul, P. Neidig, U. Klauck, P. Kessler, E. Holmes, J.K. Nicholson, B.C. Sweatman, S.R. Salman, R.D. Farrant, E. Rahr, C.R. Beddell and J.C. Lindon, *J. Pharmaceutical Biomedical Analysis* **12**(10), 1215-1225 (1994).
- (14) K. Wickholm, P.T. Larsson and T. Iversen, *Carbohydrate Research* **312**, 123-129 (1998).
- (15) H. Lennholm and T. Iversen, *Holzforschung* **49**, 462-464 (1995).
- (16) L. Wallbäcks, U. Edlund and Nordèn, *J. Wood Chem. Tech.* **9**(2), 235-249 (1989).
- (17) A.D. Shaw, A. di Camillo, G. Vlahov, A. Jones, G. Bianchi, J. Rowland and D.B. Kell, *Analytica Chimica Acta* **348**, 357-374 (1997).
- (18) H. Wilcken, C. Sorge and H.-R. Schulten, *Geoderma* **76**, 193-219 (1997).
- (19) P. Malkavaara, R. Alèn and E. Kolehmainen, *Magn. Reson. Chem.* **37**, 407-412 (1999).
- (20) L. Forveille, J. Vercauteren and D.N. Rutledge, *Food Chemistry* **57**(3), 441-450 (1996).
- (21) H. Grahn, F. Delaglio, M.A. Delsuc and G.C. Levy, *J. Magn. Reson.* **77**, 294-307 (1988).
- (22) C. Antz, K.-P. Neidig and H.R. Kalbitzer, *J. Biomolecular NMR* **5**, 287-296 (1995).
- (23) D.E. Axelson, *Multivariate Statistical Analysis and Pattern Recognition*, in *Magnetic Resonance of Carbonaceous Solids*, R.E. Botto and Y. Sanada, Editors. 1993, American Chemical Society: Washington, DC.

- (24) L.C.M. Van Gorkom and T.M. Hancewicz, *J. Magn. Reson.* **130**, 125-130 (1998).
- (25) P. Stilbs and K. Paulsen, *Rev. Sci. Instrum.* **67**(12), 4380-4386 (1996).
- (26) P. Stilbs, K. Paulsen and P.C. Griffiths, *J. Phys. Chem.* **100**, 8180-8189 (1996).
- (27) M. Kubitsa, *Chemometrics and Int. Lab. Syst.* **7**, 273-279 (1990).
- (28) W. Windig and B. Antalek, *Chemometrics and Int. Lab. Syst.* **37**, 241-254 (1997).
- (29) W. Windig, J.P. Hornak and B. Antalek, *J. Magn. Reson.* **132**, 298-306 (1998).
- (30) W. Windig, B. Antalek, L.J. Sorriero, S. Bijlsma, D.J. Louwerse and A.K. Smilde, *J. Chemometrics* **13**, 95-110 (1999).
- (31) W. Windig and B. Antalek, *Chemometrics Intelligent Laboratory Systems* **46**, 207-219 (1999).
- (32) N. Zumbulyadis, B. Antalek, W. Windig, R.P. Scaringe, A.M. Lanzafame, T. Blanton and M. Helber, *J. Am. Chem. Soc.* **121**, 11554-11557 (1999).
- (33) D. Schulze and P. Stilbs, *J. Magn. Resonance Series A* **105**, 54-58 (1993).
- (34) T.M. Alam and M.K. Alam, (*In preparation*), (2001).
- (35) E. Sanchez and B.R. Kowalski, *Anal. Chem.* **58**, 496-499 (1986).
- (36) D.M. Haaland and E.V. Thomas, *Analytical Chemistry* **60**, 1193-1202 (1988).
- (37) J.F. Manville and J.R. Nault, *Appl. Spect.* **51**, 721-724 (1997).

- (38) S. Selsbo, P. Almen and I. Ericsson, *Energy and Fuels* **10**, 751-756 (1996).
- (39) D.M. Haaland, H.D.T. Jones and E.V. Thomas, *Appl. Spect.* **51**, 340-345 (1997).
- (40) M.K. Alam, J.E. Franke, T.M. Niemczyk, J.D. Maynard, M.R. Rohrscheib, M.R. Robinson and R.P. Eaton, *Applied Spectroscopy* **52**(3), 393-399 (1998).
- (41) M. Alam, M. Rohrsceib, J. Franke, T.M. Niemczyk, J.D. Maynard and M. Robinson, *Appl. Spect.* **53**, 316-324 (1999).
- (42) L. Wallbäcks, U. Edlund, B. Nordén and I. Berglund, *Tappi J.*, 201-206 (1991).
- (43) K. Elg-Christofferson, J. Hauksson, U. Edlund, M. Sjöström and M. Dolk, *Cellulose* **6**, 233-249 (1999).
- (44) M.B. Seasholtz, *Chemometrics Intelligent Laboratory Systems* **45**, 55-63 (1999).
- (45) U. Edlund and H. Grahn, *J. Pharmaceutical Biomedical Analysis* **9**(8), 655-658 (1991).
- (46) M.K. Alam and T.M. Alam, *Spectrochimica Acta Part A*: **56**, 729-738 (2000).
- (47) L.G. Thygesen, *Holzforschung* **50**, 434-436 (1996).
- (48) A. Gerbanowski, D.N. Rutledge, M.H. Feinberg and C.J. Ducauze, *Sci. Aliments* **17**(3), 309-323 (1997).
- (49) X. Lai, M. Sardashti, B.J. Lane, J.J. Gislason and D.J. O'Donnell, *Appl. Spect.* **54**(1), 54-61 (2000).

Figure 1. Schematic representation of chemometric analysis for a NMR data set (**A**) which is a linear summation of individual pure component spectra (**P**) given a concentration profile (**C**).

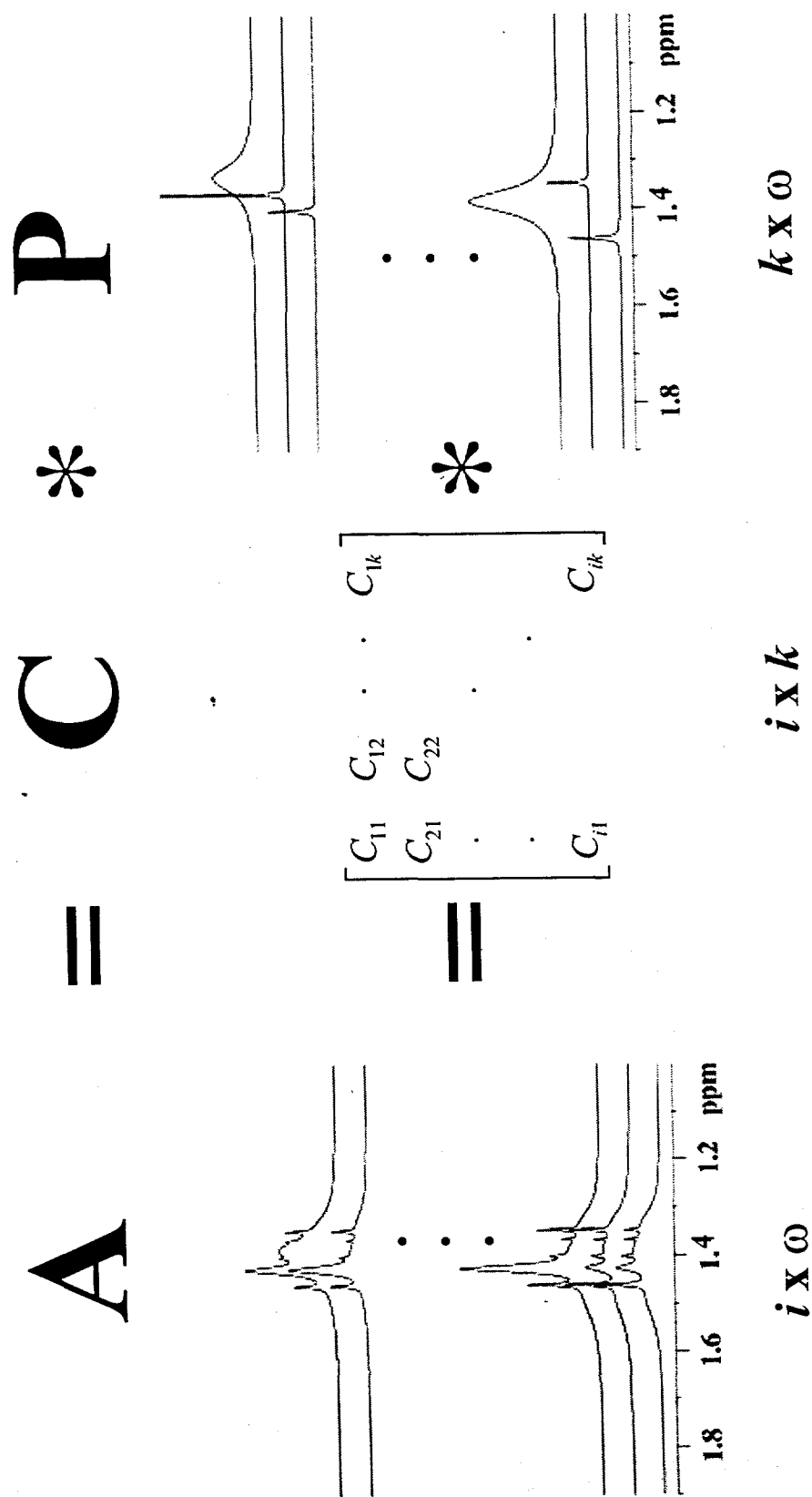


Figure 2. Scores plot for the PCA classification utilizing three principal components (PC).

PCA Discrimination

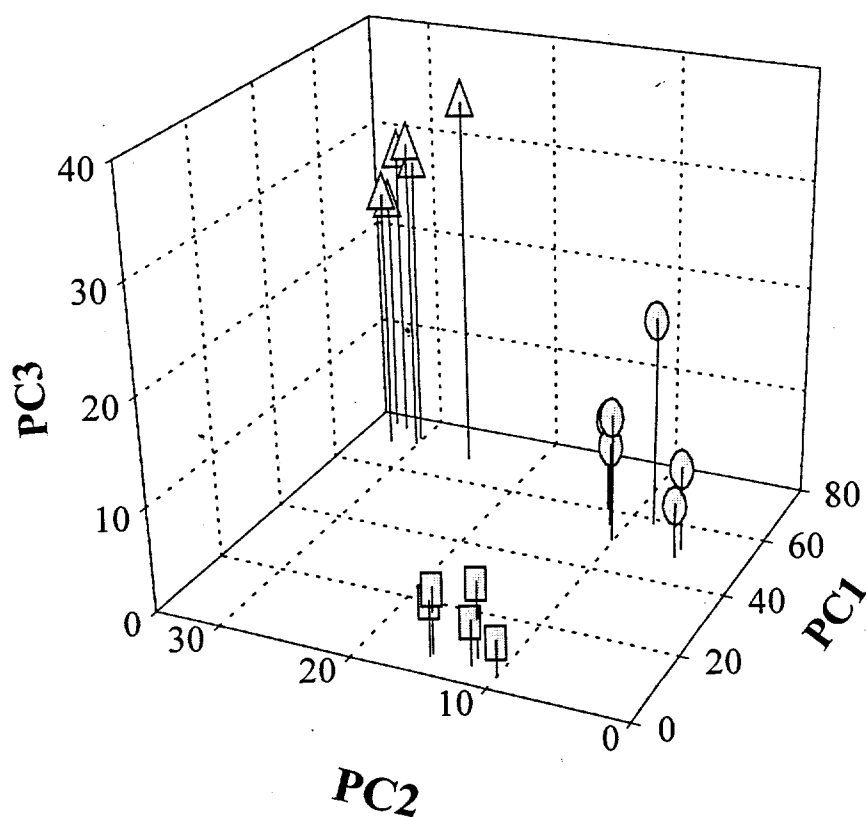


Figure 3. An illustration of how a single relaxation matrix can be split into two proportional data sets for DECRA analysis. More details given in the text.

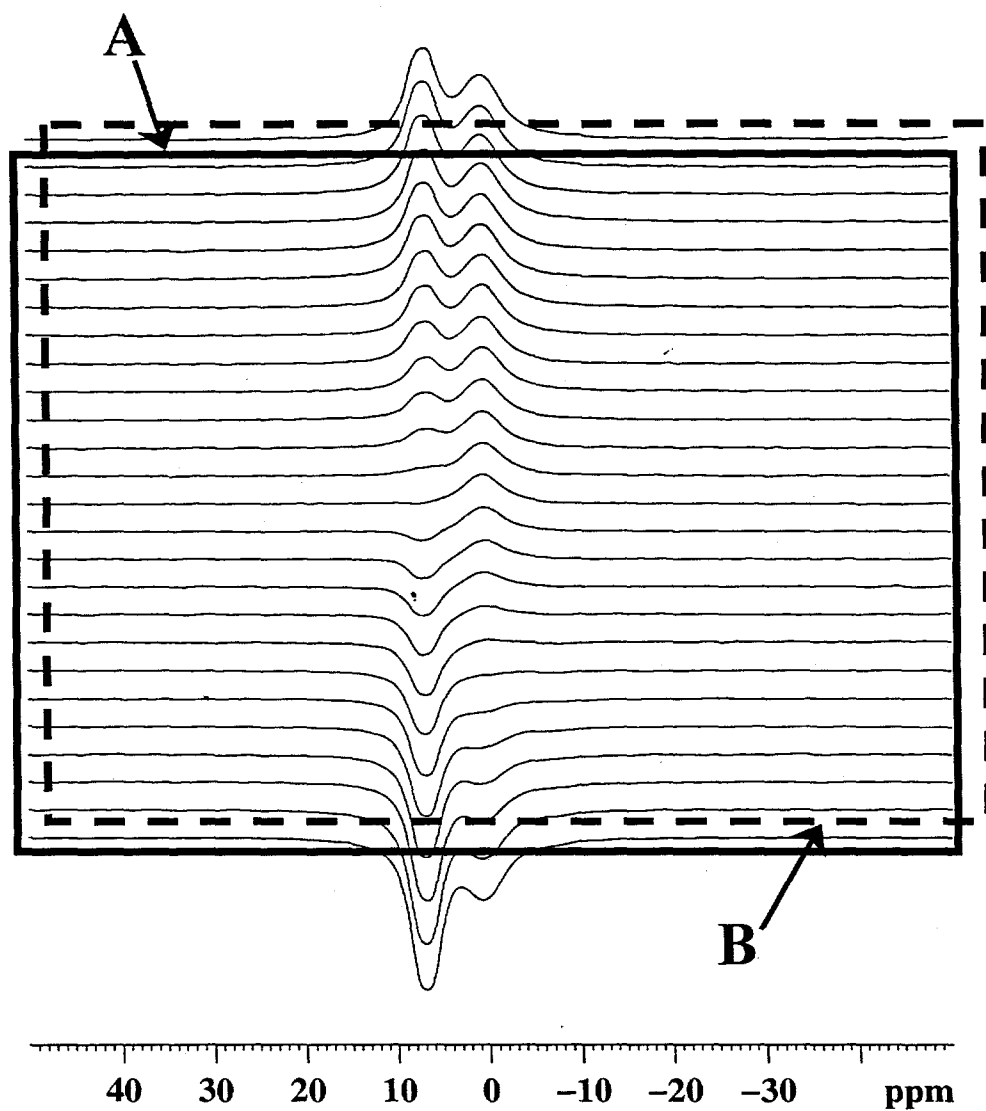


Figure 4. ^{17}O NMR spin-lattice relaxation matrix for an ethanol and 3-methyl-1-propanol mixture, where the relaxation produces an exponentially based signal modulation used to identify the pure component spectra.

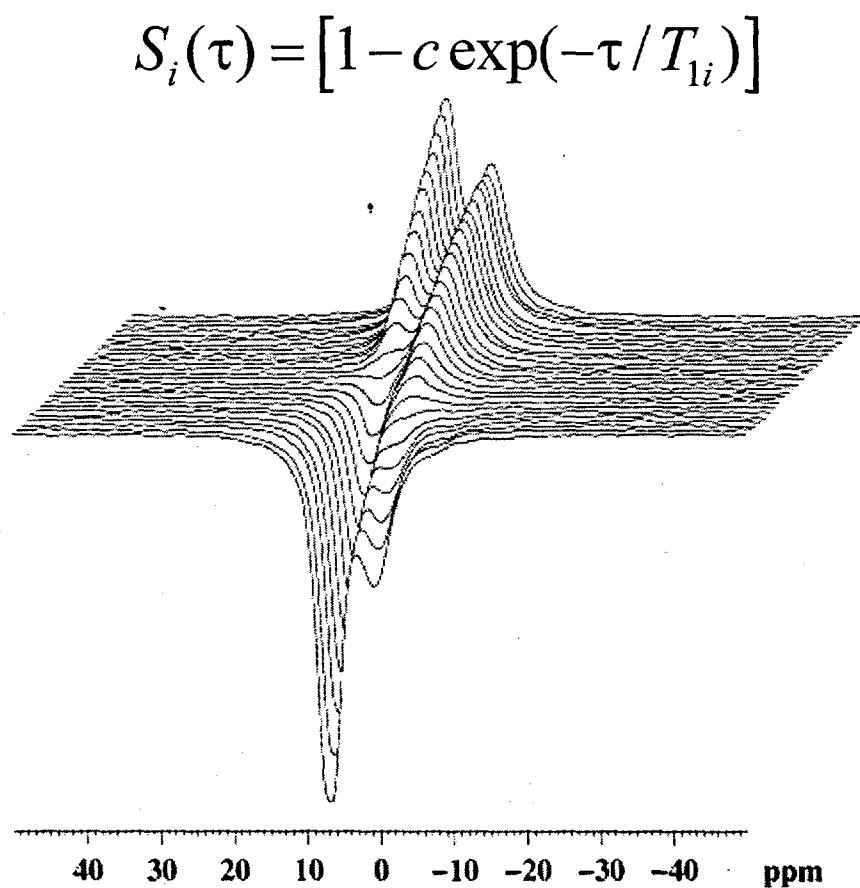


Figure 5. Results of the DECRA analysis of the ^{17}O NMR spin-lattice relaxation matrix for an ethanol and 3-methanol-1-propanol mixture: a) derived pure component spectra and b) spectra residuals.

