



Measurement Errors in Multivariate Chemical Data

*Peter D. Wentzell**

Department of Chemistry, Dalhousie University, P.O. Box 15000, Halifax, Nova Scotia, B3H 4R2, Canada

Medidas analíticas modernas são comumente apresentadas na forma de vetores (e.g., espectros) ou estruturas de dados de ordem maior como matrizes, que são normalmente sujeitas a estratégias de análise multivariada de dados para a extração de informações. Um aspecto dessas medidas que normalmente é pouco compreendida é a natureza subentendida dos erros, e como eles afetam a obtenção da informação química. Este *Account* descreve alguns dos métodos que podem ser utilizados para caracterizar erros de medidas multivariadas e como esta informação pode ser usada para melhorar os resultados das análises de dados. A caracterização inclui classificações gerais de erros, representações de domínios Fourier e matriz da covariância dos erros. O cálculo e a interpretação da covariância dos erros e matrizes de correlação são ilustrados usando medidas experimentais e os métodos de análise de dados que fazem uso dessa informação do erro são brevemente resumidos. Um exemplo simples é apresentado para mostrar como a informação sobre medidas de erros permite uma extração bem-sucedida da variância química significativa dos dados.

Modern analytical measurements are commonly presented in the form of vectors (e.g., spectra) or higher order data structures such as matrices, and these are often subjected to multivariate data analysis strategies to extract information. One aspect of these measurements that is often poorly understood is the underlying nature of the measurement errors and how these affect the ability to obtain chemical information. This *Account* outlines some of the methods that can be used to characterize multivariate measurement errors and how this information can be used to improve the results of data analysis. Characterization includes general classifications of error, Fourier domain representations, and the error covariance matrix. The calculation and interpretation of error covariance and correlation matrices are illustrated using experimental measurements, and data analysis methods that make use of this error information are briefly reviewed. A simple example is presented to show how information about measurement errors allows for more effective extraction of meaningful chemical variance in the data.

Keywords: measurement errors, error covariance, Fourier transforms, multivariate data analysis, measurement noise

1. Introduction

Analytical chemistry is the science of chemical measurements, and inherent in any chemical measurement is the principle of measurement error. Since chemistry is, at its core, an experimental science, the idea of measurement error becomes almost intuitive to most students of chemistry at an early stage and they recognize that all measurements inevitably contain some error, whether this is expressed explicitly through the use of statistical confidence intervals or implied through the use of the proper number of significant digits. Despite this imbedded knowledge, which is enhanced in most cases by some rudimentary training in

statistics, many of us are guilty at times of neglecting this component of the measurements by omitting error bars in graphs, choosing to exclude replicate measurements, or reporting too many significant figures, among other transgressions. It is important to remember, however, that distinguishing meaningful changes in measurements from random variations (noise) is essential to the extraction of chemical information.

Historically, early chemists were able to conduct experiments in which variables for model systems could be closely controlled, allowing precise measurements to be made and obviating the need for extensive statistical analysis. In recent decades, however, more challenging problems, such as those involving biological, clinical and environmental systems, have demanded greater attention

*e-mail: peter.wentzell@dal.ca

to experimental design and statistical treatment by methods such as analysis of variance (ANOVA). Coupled with this has been the evolution of analytical instrumentation, which made accessible multivariate measurements in the form of vectors, matrices and higher order structures, and necessitated the need for multivariate statistical approaches. These two factors precipitated the development of chemometrics as a sub-discipline of analytical chemistry in the late 1970s. Many definitions of chemometrics have been proposed,¹⁻⁴ but here it will be defined simply as the field of study that seeks to maximize the amount of information extracted from chemical measurements, whether that is through prescribing optimal measurement conditions or data analysis strategies. The emergence of chemometrics occurred later than parallel disciplines in other fields (e.g., psychometrics, econometrics) where the challenges in data interpretation arose earlier. While chemometrics has borrowed some approaches from those disciplines, it has also developed methods that reflect the unique characteristics of chemical data sets.

While many chemometric tools have been devised to extract information from multivariate chemical measurements, one area that has been somewhat neglected is the role of multivariate measurement errors in this process. Often techniques that are developed to analyze multivariate data make simplistic assumptions about the measurement errors that are carried over from practices or assumptions made about univariate data, but in reality the structure of multivariate measurement errors can be much more complex. In many cases, especially when measurement errors are small or the assumptions are approximately valid, traditional chemometric tools can be applied with excellent results, but in other cases, consideration of the measurement error structure can mean the difference between the success and failure of the data analysis. Studying the nature of multivariate measurement errors has several important benefits. First, it can lead to a better understanding of the origins of errors for a particular instrument or measurement system, knowledge which can then be used to improve the quality of measurements by addressing the limiting sources of error. Second, by knowing the characteristics of the errors, data analysis tools can be designed to treat the errors in an optimal way, leading to more efficient extraction of chemical information. Finally, the error structure inherent in multivariate data can be propagated through different pre-processing and data analysis steps to determine its impact on the final result.

The goals of this article are: (a) to describe some of the terminology associated with the description of errors in general and multivariate errors in particular, (b) to introduce some of the methods used to characterize

multivariate measurement errors, and (c) to present some of the approaches that make use of measurement error information to enhance data analysis. Much of the material presented here is derived from a short course offered by the author at the First Chemometric Winter School held at the Universidade Federal de São Carlos (São Carlos, São Paulo State, Brazil) in August of 2013.

2. Terminology

The concept of measurement error is at the same time both simple and complex. We can define the error, e , as simply the difference between a measured value, x , and its true value, x^0 .

$$e = x - x^0 \quad (1)$$

Although this definition appears simple, the first complication arises from the definition of the “true” value. Setting aside the fact that we don’t generally know the true value (except in simulations), it can be defined in a variety of ways, as illustrated in Figure 1. If the true value sought is the height of an analyte peak, errors can be introduced by chemical interferences, baseline variations, and random errors in the detector response. However, in another context, such as multivariate calibration where interferences are accommodated in the model, the true value could be defined as the sum of the analyte and the chemical interference, or could also include the baseline, in which case errors only arise from the detector noise.

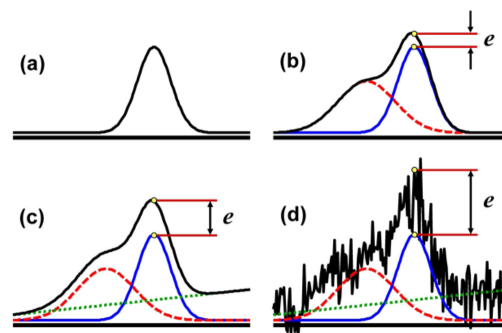


Figure 1. Illustration of different sources of errors in a chemical measurement: (a) pure analytical signal; (b) analytical signal and measured signal (solid lines) in the presence of a chemical interference (dashed line); (c) analytical and measured signals in the presence of a chemical interference and baseline drift (dotted line); and (d) analytical and measured signals in the presence of chemical interference, baseline drift, and measurement noise.

In practice, errors are typically defined by using the mean of several measurements as a surrogate for the “true” value. In this case, the definition of the replicate becomes critical and there are many such definitions possible. As an

example, consider the case in which the measurement is a peak in the nuclear magnetic resonance (NMR) spectrum of a sample of blood plasma in a clinical trial or metabolomics study. The lowest level of replication, the *instrumental replicate*, would be obtained by simply scanning the spectrum of the same sample several times without replacement. Assuming no degradation of the sample, this would reflect the variability of the instrument, including baseline variations and detector noise. A *replacement replicate* is similar, except that it involves replacement of the sample between scans and therefore would include variations associated, for example, with changes in the position of the sample within the field. *Technical replication*, which can be performed at several levels, involves replication of the steps required for preparation of the sub-sample prior to obtaining the spectrum. In this case these could include, for example, separation of the plasma from the blood cells and dilution, thereby introducing any sources of variation associated with those steps. *Preparatory or sub-sampling replicates* involve replication of all of the steps required for the preparation of different sub-samples, which in this case could be considered to be repeated draws of blood from the same subject at the same time. *Sampling replicates* refer to replication of analysis for samples of the same origin (e.g., same batch, organism, or culture) and in the present example may refer to samples extracted from the same subject at different times, thereby incorporating diurnal variations in analyte concentration. *Sub-population replicates*, generally the highest level of replication, would involve replicate scans from different samples in the same class, for example healthy adult males, and incorporates the variations among individuals in addition to variations at all lower levels of replication.

It should be noted that not all of these levels apply to all analytical methods and additional levels of replication may be defined. Moreover, other sources of variability, such as different labs and operators, may be included, and this has given rise to the terms *repeatability* (same lab/analyst/day) and *reproducibility* (different labs/analysts/days) to describe the precision of analytical methods.⁵ Generally the variability in a measurement is expected to increase at each level of replication as factors that were fixed at lower levels begin to introduce a random contribution. This means that the definition of the “true” value, and therefore the errors observed, change depending on the level of replication, so it is important to consider how replicates are to be defined when attempting to characterize measurement errors.

A term that is sometimes used as a synonym for measurement errors is “noise”. While this can have the same meaning, it generally implies a random error source and may also be used to imply a random series of error

contributions in time, space, or some other ordinal variable such as wavelength. Because of this latter usage, the term “measurement noise” is often used to refer to the vector of measurement errors associated with a spectrum or other measurement vector.

An important distinction that is often overlooked in the literature is the difference between “error” and “uncertainty”. The term “error” refers to a specific value associated with one measurement of a particular quantity, and may be positive or negative (or even zero), while “uncertainty” refers to a statistical characterization of the errors in replicated measurements, such as a variance, confidence interval, or standard deviation. The population variance of the measurement error is defined as the expectation value, designated as $E()$, of the squared error, as given in equation 2, where N represents the number of samples.

$$\sigma^2 = E(e^2) = E[(x - x^0)^2] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (x_i - x^0)^2 \quad (2)$$

In practice, the sample variance is based on a finite number of measurements and is calculated using the sample mean as an estimate of the true value, x^0 .

$$\sigma^2 \approx s^2 = \frac{1}{(N-1)} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (3)$$

While these definitions are well-known to most readers and can be applied equally well to individual variables for multivariate data, extension to multivariate systems also requires another consideration, and that is the relationship between errors in two or more variables. If we consider two uniquely defined variables, x and y , for a single replicate, for example intensity measurements at two wavelength channels for a single sample, then we can give the following definitions.

$$\sigma_x^2 = E(e_x^2), \quad \sigma_y^2 = E(e_y^2), \quad \sigma_{xy} = E(e_x \times e_y) \quad (4)$$

Here, σ_{xy} represents the *error covariance* between variables x and y and describes the statistical relationship between errors in the two variables. When $\sigma_{xy} = 0$, the errors in the two variables are said to be *uncorrelated* or *independent*. Alternatively, the errors may be related by a *correlation coefficient*, ρ , such that $\sigma_{xy} = \rho \sigma_x \sigma_y$, where $\rho = 1$ means that the errors are perfectly *correlated* and $\rho = -1$ means they are perfectly *anti-correlated*. Analogous to the sample error variance, the sample error covariance is defined as given in equation 5.

$$\sigma_{xy} \approx s_{xy} = \frac{1}{(N-1)} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \quad (5)$$

Note that the error covariance can be either positive or negative so there is no analog to the standard deviation, which is the square root of the variance.

In a measurement vector consisting of n measured variables, both the error variance and covariance are important in characterizing the errors. When the error variance is the same among a group of variables, the errors are said to be *homoscedastic*, while a non-uniform variance means that the errors are *heteroscedastic*. Independent errors are observed when the error covariances within a group of variables are all zero, while non-zero values means that errors are correlated. When heteroscedastic and/or correlated errors are observed for a series of variables in a measurement vector, they are often characterized according to their pattern of behaviour. Table 1 lists a variety of ways that measurement errors, or noise, can be characterized. This list includes classifications based on heteroscedasticity

(e.g., shot noise), correlations (e.g., $1/f$ noise) or both (e.g., multiplicative noise), as well as classifications based on other factors, such as the origin (e.g., Johnson noise). Note that the classifications listed are not mutually exclusive and that multiple types of noise are generally observed in any system, although one may dominate.

In the analysis of multivariate analytical data, a condition that is often assumed for the methods employed is that the measurement errors are independent and identically distributed with a normal distribution, or *iid normal*. In other words, measurement errors are assumed to be homoscedastic, uncorrelated and normally distributed. For most experimental data, these assumptions are not valid and some degree of heteroscedasticity and/or correlation is present in most error structures encountered in the author's experience. While this does not necessarily negate the validity or effectiveness of

Table 1. Some classifications of experimental measurement errors

Type of noise	Description
Independent errors/ uncorrelated errors	Errors for which the error covariance is zero.
Correlated errors	Errors for which the error covariance is not zero.
Homoscedastic errors	Errors that have the same (uniform) variance.
Heteroscedastic errors	Errors that have a non-uniform variance.
White noise	A vector of uncorrelated measurement errors. May also imply homoscedastic errors and/or errors which follow a normal distribution.
Pink noise or $1/f$ noise	A type of "low frequency" noise in which the measurement errors in adjacent measurements are more correlated than for measurements that are farther apart. Characterized by random variation that occurs slowly. Variations include "brown noise" ($1/f^2$ noise).
Drift noise	Another term for low frequency or correlated noise which implies a slow change in measurement conditions such as temperature.
Source flicker noise	Yet another term for low frequency noise (usually $1/f$) that is specifically associated with variations in source signal intensity (e.g., a lamp in spectroscopy or an ion source in mass spectrometry).
Proportional noise	Heteroscedastic noise in which the standard deviation of the error is proportional to the magnitude of the signal. Often associated with source flicker noise.
Additive noise or offset noise	Correlated noise ($\rho = 1$) that randomly shifts the entire signal up or down by a fixed amount (e.g., shifts in baseline position).
Multiplicative noise	Correlated noise ($\rho = 1$) that randomly shifts the entire signal up or down by an amount proportional to the magnitude of the signal (e.g., variations in path length in absorption spectroscopy). A type of proportional noise.
Interference noise	Can refer to periodic (correlated) noise that occurs at a particular frequency or set of frequencies (e.g., pump noise in chromatography). Can also refer to variable signal interference by other chemical or physical factors.
Baseline noise	Can refer either to the variance introduced by a variable baseline (offset noise, drift noise, interference noise, etc.) or the variance of the noise in the baseline regions where no signal is present.
Shot noise	Heteroscedastic noise where the noise standard deviation is proportional to the square root of the signal. Arises from the Poisson distribution associated with counting statistics (e.g., photomultipliers).
Johnson noise or thermal noise	Variation in a signal arising from the random thermal motion of charge carriers across resistance in an electrical circuit. Often has white noise characteristics.
Detector noise	General term for noise arising at the detector.
Digitization noise or quantization noise	Arises from finite precision of analog-to-digital converters (rarely a concern in modern instruments) or discrete category variables.
Stationary noise	Not a type of noise, but a condition where the joint probability distribution of the errors does not change from one set of measurements to the next.

the methods employed, it may make them sub-optimal for achieving the desired goals. Indeed, many of the pre-processing techniques employed for multivariate data (e.g., scaling, derivative filtering) have the implicit objective of making the errors closer to *iid* normal. The goal of many multivariate analysis techniques is to separate meaningful chemical variance from noise variance, so the assumptions about the latter can be important. For optimal data analysis, it is necessary to quantitatively describe the nature of the measurement error structure. Two common ways to do this are through the use of Fourier analysis and through error covariance matrices, which are described in the sections that follow.

3. Fourier Transforms

Several of the terms in Table 1 refer to the “frequency” of the noise, terminology which is a carry-over from the Fourier analysis of time series signals in engineering disciplines. This terminology may be less familiar to chemists, however, so it warrants some discussion as a means to describe multivariate measurement errors. Most chemists are already somewhat familiar with the Fourier transform from its various applications in the field, such as in Fourier-transform infrared spectroscopy (FTIR). Because signals obtained by modern instruments involve discretely sampled digital values, the discussion here is limited to the discrete Fourier transform (DFT). A comprehensive discussion of the DFT is beyond the scope of this article and has been presented elsewhere,⁶⁻⁸ so only a brief review will be included here.

A time series consisting of N measurements sampled at discrete intervals of Δt can be represented as the measurement vector \mathbf{y} , where the corresponding times are $t_i = (i - 1) \Delta t$. Fourier theory states that the measurements in \mathbf{y} can be reproduced using a sum of $(N/2 + 1)$ sinusoidal signals of predefined frequencies as given by equation 6 (note: N is assumed to be even here).

$$y(t) = \sum_{n=0}^{N/2} C_n \cos(2\pi f_n t + \phi_n) \quad (6)$$

In this equation, C_n and f_n are the amplitude and phase angle of sinusoid at frequency f_n , where f_n is given by,

$$f_n = \frac{n \times f_s}{N} = \frac{n}{N \times \Delta t} \quad (7)$$

and f_s is the sampling frequency, defined as $1/\Delta t$. The Fourier transform can be represented in several ways, but one common approach is to plot the amplitude and phase angle as a function of frequency, referred to as the amplitude and phase spectra. This is illustrated in Figure 2 where the DFT of

a ten point square wave has been generated. Figure 2a shows the original square wave along with the discretely sampled points and the signal reconstructed by the Fourier transform (two cycles of the square wave are shown for clarity). Note that while the DFT exactly reconstructs the sampled points as required, it does not exactly reproduce the underlying signal, which is impossible due to a finite sampling rate. Figure 2b shows the six sinusoids used to reconstruct the signal, including the DC signal and two frequencies with zero amplitude. The amplitude and phase spectra are shown in Figures 2c and d, respectively.

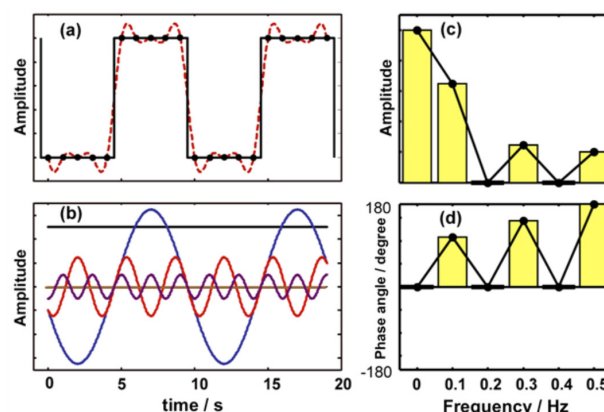


Figure 2. Discrete Fourier transform of a ten point square wave (two cycles shown for clarity). (a) Original signal (solid black line), sampled points (black circles), and Fourier reconstructed signal (dashed red line); (b) sinusoidal components of Fourier transform; (c) amplitude spectrum of Fourier transform; (d) phase spectrum of Fourier transform.

Amplitude spectra from the DFT are useful in studying the characteristics of noise from ordered chemical measurement vectors. The term “ordered” is used here to refer to data sets in which the measurements are associated with some ordinal variable such as time or wavelength that describes a natural sequence in the measurements. This is in contrast to data sets where the order of the measured variables has no particular relevance, such as elemental concentrations in an environmental study. The reason for this distinction is that the DFT is particularly effective in detecting patterns of correlation in measurement errors over a short range in a series of measurements. While conventional descriptions of the Fourier transform use time as the ordinal variable, with frequency as the complementary domain, there is no mathematical reason why this should be the case and any ordinal variable (e.g., wavelength, chemical shift) can be employed. In such cases, it is more appropriate to refer to the complementary domain as the Fourier domain, but the frequency domain terminology persists even when the ordinal variable is not time. Hence the term “low frequency noise” does not necessarily refer to the frequency domain.

The widely used term “white noise” is derived from Fourier analysis of signals where the amplitude spectrum in the Fourier domain is ideally flat, reflecting equal contributions at all frequencies (analogous to white light). In practice, noise is by definition stochastic, so the DFT of a single sample of white noise does not appear flat, as illustrated in Figure 3a where a 100-point sample of white noise and its Fourier transform (amplitude spectrum) are shown. To remove the “noise from the noise”, the amplitude spectra for multiple white noise sequences need to be averaged. This is shown in Figure 3b, where the DFT represents the average of amplitude spectra from 100 white noise sequences. Other methods to improve the appearance of noise spectra are boxcar averaging of amplitude spectra from longer noise sequences and calculating the DFT of the noise autocorrelation function. The autocorrelation function for white noise is a delta function, which has a flat Fourier transform.

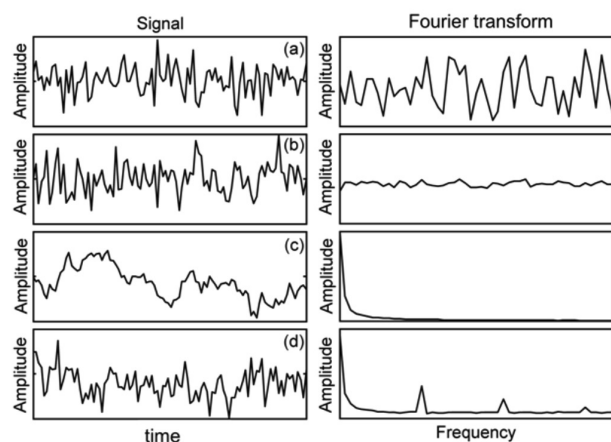


Figure 3. Some examples of noise sequences (left) and their FT amplitude spectra: (a) white noise; (b) white noise with the amplitude spectrum averaged over 100 samples; (c) pink noise ($1/f$ noise); and (d) a combination of white noise, pink noise and interference noise. The amplitude spectra in (c) and (d) are also averaged over 100 samples.

Also shown in Figure 3 are some other examples of noise sequences in the time and frequency domains. Figure 3c shows a typical sequence of pink noise, or $1/f$ noise and its amplitude spectrum (averaged over 100 noise sequences). In the time domain, this type of noise appears to be more slowly varying than white noise and is often associated with drift noise or source flicker noise. The DFT is characterized by a large amplitude at low frequencies that falls off as $1/f$. Figure 3d contains a mixture of noise contributions (white noise, pink noise and interference noise) that might be more typical of a real signal. While the various noise components are not readily apparent in the time domain, the DFT shows them more clearly. The white

noise contribution is represented by the offset from zero in the amplitude spectrum and the pink noise is indicated by the higher contributions at low frequencies. Interference noise appears as a spike at a particular frequency and, as indicated here, can typically exhibit higher order harmonics at fixed intervals.

The frequency domain characteristics of noise can be an important consideration in signal processing of measurement vectors, especially in the context of pre-processing analytical signals. For example, smoothing using polynomial least squares (Savitzky-Golay) filters is widely used in analytical chemistry to reduce noise. These filters act to remove the high frequency content of the signal vector, which is dominated by white noise, while retaining the low frequency components which typically contain the important information. However, this type of signal processing is ineffective when the dominant source of noise is $1/f$ noise, which is mainly present at low frequencies. Likewise, derivative filters are often used to suppress low frequency offset and drift noise that dominate at low frequencies, but can have the effect of amplifying high frequency noise.

While Fourier analysis of noise in analytical signals is useful, it is severely limited in the information it can provide. Since these methods were originally designed for time series analysis of stationary signals that follow a well-defined correlation structure, the more complex correlation patterns that can be present in measurement noise from chemical measurements may not be easily revealed by such techniques. Moreover, noise heteroscedasticity (non-uniform variance) among measured channels is not readily apparent in the Fourier transform. Therefore, a more comprehensive description of the measurement error characteristics is required.

4. The Error Covariance Matrix

If we consider an $n \times 1$ vector of measurements \mathbf{x} , we can define the vector of measurement errors, \mathbf{e} ($n \times 1$), in a manner analogous to equation 1,

$$\mathbf{e} = \mathbf{x} - \mathbf{x}^0 \quad (8)$$

where the vector \mathbf{x}^0 represents the true measurement vector. In the same way that the variance is used to characterize the statistical behaviour of univariate measurement errors, we can use the *error covariance matrix*, Σ ($n \times n$), to describe the statistical characteristics of the measurement error vector. Σ is defined as the expectation of the outer product of the measurement error vector:

$$\Sigma = E(\mathbf{e}\mathbf{e}^T) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_n^2 \end{bmatrix} \quad (9)$$

The error covariance matrix (more precisely called the error variance-covariance matrix) describes the statistical relationship among all of the errors in the measurement vector. The diagonal elements of this matrix give the error variances associated with each measurement channel and will therefore indicate any heteroscedasticity that is present. The off-diagonal elements, σ_{jk} , give the covariance of the measurement errors at channels j and k , as previously defined in equation 4. Note that Σ is a symmetric matrix such that $\sigma_{jk} = \sigma_{kj}$. In the case of independent (uncorrelated) errors, Σ is a diagonal matrix, and when the measurement errors are *iid*, the values along the diagonal of this matrix will all be the same.

When two measurement errors are correlated, the error covariance can also be represented as $\sigma_{jk} = \rho_{jk} \sigma_j \sigma_k$, where ρ_{jk} is the correlation coefficient for measurement errors at channels j and k , and varies between -1 and $+1$. This leads to an alternative way of representing the error structure in terms of the *error correlation matrix*, which contains the correlation coefficients of the elements of Σ , given by:

$$\rho_{jk} = \frac{\sigma_{jk}}{\sigma_j \sigma_k} \quad (10)$$

Note that the diagonal elements of the error correlation matrix are always unity. The error covariance and error correlation matrices provide complementary information, where the former gives the magnitude of the relationships among the errors and the latter indicates the underlying structure of this relationship. Knowing the statistical properties of the errors in measurement vectors allows the application of data analysis tools that can exploit this knowledge to more effectively separate meaningful chemical variance from noise variance, as discussed in Section 5. First, however, estimates of the error covariance matrix must be obtained. In general, three approaches can be used for this purpose: experimental replication, theoretical prediction, and empirical modeling. Each of these approaches is described briefly below.

4.1 Experimental replication

In the same way that replicate univariate measurements can be used to estimate measurement error variance, replicate measurement vectors (e.g., spectra) can be used to estimate the measurement error covariance, using the

mean vector to represent the “true” value. If N replicate measurement vectors, \mathbf{x}_i ($n \times 1$), are obtained, the sample error covariance matrix, \mathbf{S} ($n \times n$), is defined by equation 11.

$$\mathbf{S} = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \quad (11)$$

The definition of a replicate (e.g., replicate scans, replicate subsamples) is very important in the application of this equation as it defines what is considered noise variance and what is considered relevant chemical variance and can alter the structure of the error covariance matrix. Another important consideration is the number of replicates used. Estimates of variance (and covariance) for both univariate and multivariate measurements are characterized by a very high uncertainty, so it is important that a sufficient number of replicates be used to minimize the “noise in the noise”. Otherwise, any potential benefits of including measurement error information in the data analysis may be offset by its poor reliability. Assuming that the measurement errors follow a normal distribution, the relative standard deviation in the variance estimates can be approximated from the χ^2 distribution as 71% for 5 replicates and only drops to 14% with 100 replicates. Given that a typical number of replicates is fewer than ten, this can present a problem in the reliable estimation of Σ for a measurement vector. One solution is *pooling* of error covariance estimates by averaging the calculated \mathbf{S} from different subsets of samples, each of which has a relatively small number of replicates. This is generally viable as long as the measurement vectors do not change substantially between samples as is typically the case for many analytical measurements, such as near-infrared (NIR) spectra measured for samples of similar origin. In these cases, the improved covariance estimation generally outweighs the small differences due to changes in signal magnitude.

An example of an experimentally calculated error covariance and correlation matrices using pooled data is presented in Figure 4. Figure 4a shows 27 NIR reflectance spectra from samples of a polyester resin (9 samples, 3 replicates each), while Figure 4b shows the error covariance matrix and Figure 4c shows the error correlation matrix. The three-dimensional representations of these matrices give a clearer picture of the error structure, but can often be difficult to interpret when there are multiple complex error sources. Note in Figure 4b that the magnitudes of the variance and covariance are largest where the signal is the largest, suggesting a proportional type of error. There are also larger variance/covariance contributions when there are sharp transitions in the signal, which is indicative of sensitivity to variations in the wavelength channel where the transition occurs. The error correlation matrix in Figure 4c shows that

the measurement errors are highly correlated, with ρ close to unity over most of the spectrum, although there is a break in the correlations in the spectral measurements above and below 6000 cm^{-1} . In contrast to the error covariance matrix, there is (generally) a decrease in correlation at points where there is a sharp spectral transition. This is likely due to the fact that variations arising from small shifts in wavelength are more random and less correlated. These observations are consistent with most NIR spectra which tend to be dominated by multiplicative measurement errors that are proportional and highly correlated. It is clear that such representations can offer insights into error structures present in analytical measurements and can potentially provide a wealth of information about the measurement system itself.

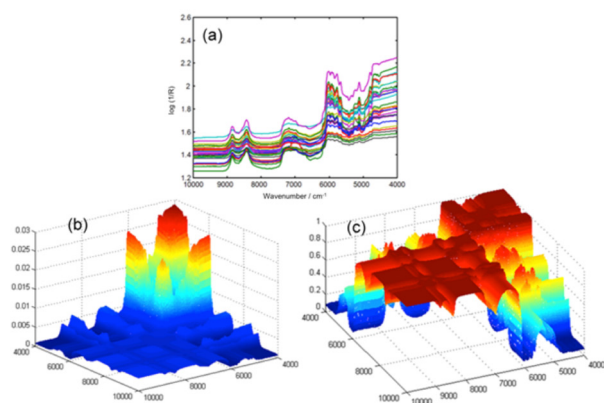


Figure 4. Near-infrared reflectance spectra of polymer resin samples (a) and the resulting pooled error covariance matrix (b) and error correlation matrix (c).

4.2 Theoretical prediction

In principle, if sufficient prior knowledge of the error sources for an experimental system is available, the error covariance matrix can be predicted on that basis alone. For example, measurement errors may be dominated by Poisson statistics or proportional contributions. In practice, however, most measurements involve a complex mixture of error sources that are difficult to predict *a priori*, so accurate theoretical prediction of error covariance is difficult to achieve. Nevertheless, the concept of error propagation can be effectively used to assess the effects of various mathematical operations on the errors in analytical measurements, and this deserves some attention.

If we consider some derived quantity, y , which is a function of several variables such that $y = f(x_1, x_2, \dots)$, then the general formula for propagation of error⁹ is represented by equation (12).

$$\sigma_y^2 = \left(\frac{\partial y}{\partial x_1}\right)^2 \sigma_1^2 + \left(\frac{\partial y}{\partial x_2}\right)^2 \sigma_2^2 + \left(\frac{\partial y}{\partial x_3}\right)^2 \sigma_3^2 + \dots + 2\left(\frac{\partial y}{\partial x_1}\right)\left(\frac{\partial y}{\partial x_2}\right) \sigma_{12} + 2\left(\frac{\partial y}{\partial x_1}\right)\left(\frac{\partial y}{\partial x_3}\right) \sigma_{13} + 2\left(\frac{\partial y}{\partial x_2}\right)\left(\frac{\partial y}{\partial x_3}\right) \sigma_{23} + \dots \quad (12)$$

Here, σ_i represents the standard deviation of the errors in x_i and σ_{ij} represents the covariance of the errors in x_i and x_j . Many representations of this equation do not include the covariance terms, but these are necessary when the errors are not independent. The inclusion of such terms becomes cumbersome as more variables are included, since every pair of variables needs to be included, and a matrix formulation is more convenient. If we define a Jacobian row vector, \mathbf{j} ($1 \times n$), to contain the partial derivatives of y with respect to each x ,

$$\mathbf{j} = \left[\frac{\partial y}{\partial x_1} \quad \frac{\partial y}{\partial x_2} \quad \dots \quad \frac{\partial y}{\partial x_n} \right] \quad (13)$$

then the variance in y is given by:

$$\sigma_y^2 = \mathbf{j} \Sigma_x \mathbf{j}^T \quad (14)$$

where Σ_x ($n \times n$) is the error covariance matrix for vector \mathbf{x} . This generalization can be extended further. If there is a series of values y_i ($i = 1$ to m), all of which are functions of the measurement vector \mathbf{x} , and define a Jacobian matrix, \mathbf{J} ($m \times n$), such that $J_{ik} = \partial y_i / \partial x_k$, then the error covariance matrix for the vector \mathbf{y} ($m \times 1$), is given by:

$$\Sigma_y = \mathbf{J} \Sigma_x \mathbf{J}^T \quad (15)$$

where Σ_y is $m \times m$. This equation describes the changes in the error covariance matrix that take place when a transformation applied to the measurement vector \mathbf{x} produces a new measurement vector, \mathbf{y} . It is particularly useful when the transformation is represented as a linear operation that involves the multiplication of the original measurement vector by some $m \times n$ transformation matrix, \mathbf{T} , such that $\mathbf{y} = \mathbf{T}\mathbf{x}$. In this case, the new error covariance matrix is,

$$\Sigma_y = \mathbf{T} \Sigma_x \mathbf{T}^T \quad (16)$$

This equation is particularly useful because it can be applied a wide variety of situations that include smoothing,¹⁰ differentiation,¹¹ subspace projection¹² and wavelet transforms,¹³ and can be used to track how the errors in the original measurements are carried through different data analysis steps. For example, equation 16 can be used to determine how uncertainties in the original measurements are transformed into PCA scores space.

To illustrate these ideas, Figure 5 shows the effect of row scaling on the error correlation matrix of a hypothetical data set. The simulated data set consists of ten variables with values given in Figure 5a. This may represent, for example, a fatty acid profile in a classification analysis or elemental concentrations in an environmental study. In such cases, where the relative magnitudes are typically more relevant than the absolute values, row scaling (normalization of the measurement vector to a constant sum) is often used. In this case, it was assumed that each measurement was characterized by independent proportional errors of 10%. Prior to scaling, it can be seen in Figure 5b that the measurements are uncorrelated. In contrast, propagation of error shows that, after row scaling, the measurement errors have become correlated as shown in Figure 5c. In particular, there is a strong negative correlation of the errors for most variables with the error in the first variable. This is anticipated, since the first variable has the largest errors, and a positive error will tend to make the sum high, increasing the probability that the contributions of the other variables will be low. The opposite is true for negative errors in the first variable. This example shows that even simple operations like row scaling can have significant effects on the error structure of the measurements which will have implications for subsequent data analysis.

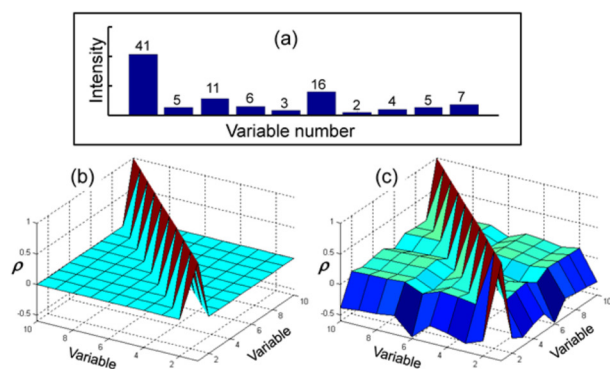


Figure 5. Hypothetical example illustrating the effect of row normalization on error covariance: (a) hypothetical data showing error-free profile of ten variables; (b) error correlation matrix before row normalization assuming independent proportional errors; and (c) error correlation matrix after row normalization.

4.3 Empirical modeling

Empirical modeling of error covariance matrices represents a middle ground between the direct use of replicate measurements, which is tedious and prone to noisy estimates, and theoretical prediction, which is difficult or impossible. In this approach, a limited number of replicates are used to develop an empirical model capable of providing a reliable estimate of the error covariance.

There are several advantages to this approach. First, the model can lead to a better fundamental understanding of the limiting measurement errors in the system under study, potentially revealing improved methods for managing the errors at the data acquisition or analysis stages. Second, a general model reduces the need for replication since the model parameters can be estimated from a smaller number of replicates and extended to new signals measured on the same experimental system. Finally, the use of a parsimonious model will smooth the stochastic variations inherent in the estimation of experimental error covariance matrices that arise from a limited number of replicates, leading to more reliable results.

At first glance, modeling the complexities of error covariance might appear to be a daunting task, but in fact the errors are often dominated by a bilinear structure that can be deduced by principal components analysis (PCA) and target testing. Residual independent error components can then be modeled as a separate term. This approach has been used to develop error covariance models for a number of different experimental systems ranging from NIR reflectance to NMR.^{14,15} While this strategy can undoubtedly be improved upon, the models generated in this manner serve as good approximations to the experimentally derived results.

This approach to modeling error covariance matrices is illustrated in Figures 6 and 7. Figure 6a shows fluorescence emission spectra for mixtures of polycyclic aromatic hydrocarbons (27 mixtures with five replicates each).¹⁶ The corresponding error covariance and error correlation matrices are shown in Figures 6b and 6c, respectively. The error covariance matrix indicates that the error variance and covariance increase with the magnitude of the signal, while the error correlation matrix suggests the presence of both independent and strongly correlated components in the errors due to the presence of a flat but non-zero contribution in the off-diagonal elements. The error covariance matrix is reproduced in Figure 7a for comparison with the modeled error covariance matrix shown in Figure 7b. In this case, the modeled matrix was generated with good reliability with contributions from only three sources. The contribution of each of these sources to the standard deviation of the signal (square root of the diagonal of the error covariance matrix) is shown in Figure 7c, along with the reconstructed and observed standard deviations. The first contribution is variance in the offset of the spectrum (additive noise), which leads to correlated errors and is anticipated due to variations in blank signal during the course of the experiment. The second contribution, which also leads to correlated noise, is a multiplicative effect that is proportional to the square root of the spectrum. This is likely due to variations in the source intensity over the course of the experiment which will lead

to offset errors that depend on the magnitude of the signal. The final contribution is from independent errors, which are also proportional to the square root of the mean spectrum but do not affect the off-diagonal elements of the error covariance matrix. This contribution arises from the shot noise typically associated with fluorescence measurements and characterized by a square root dependence on the intensity. The final model is represented by equation 17, which shows bilinear offset and multiplicative effects and the independent noise contributions.

$$\begin{bmatrix} \Sigma \\ \mathbf{1} \end{bmatrix} = \begin{bmatrix} \vdots \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{bmatrix} \begin{bmatrix} \cdots & 1 & \cdots \\ \cdots & \sqrt{\bar{x}} & \cdots \end{bmatrix} + \sigma_c^2 \begin{bmatrix} \ddots & & \\ & \ddots & \\ & & \ddots \end{bmatrix} \bar{\mathbf{x}} \quad (17)$$

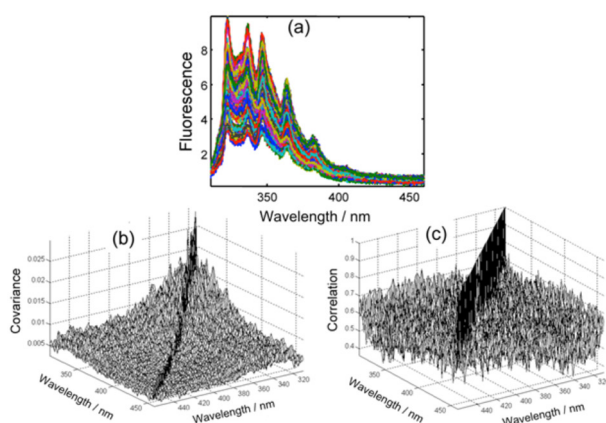


Figure 6. Fluorescence emission spectra (a) for 27 mixtures of three polycyclic aromatic hydrocarbons, each measured in triplicate (see ref. 16 for details) with the corresponding pooled measurement error covariance matrix (b) and error correlation matrix (c). Adapted from reference 14.

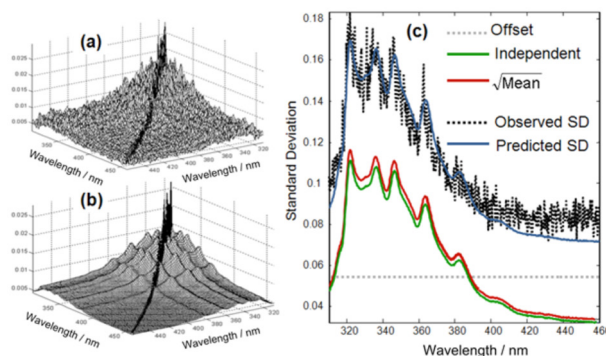


Figure 7. (a) Measured and (b) modeled error covariance matrix for the fluorescence spectra in Figure 6; (c) measured and modeled standard deviation of measurements showing the contributions of from mean and offset terms of covariance along with the independent errors. Adapted from reference 14.

The quantities σ_a , σ_b and σ_c are parameters estimated by the model for each of these contributions. Even a relatively simple model such as this can model error covariance surprisingly well and offers insights into the source of errors in the measurement system.

5. Data Analysis

The characterization of measurement errors for multivariate data sets is of limited utility unless methods are available to exploit such information. The lack of such methods has been one factor that has contributed to practices that do not encourage the acquisition of replicated data and error analysis for multivariate measurements. At the same time, the development of new methods has been impeded by a scarcity of data sets for which measurement error information is available. As the role of measurement errors becomes more widely recognized, however, the situation is changing and more tools and data are becoming available.

A variety of data analysis techniques have been developed in recent years that make use of measurement error information to improve results. Because the problems described here are not unique to chemical measurements, these methods span a variety of disciplines and a comprehensive treatment is beyond the scope of this article, but a few methods will be mentioned briefly, with an emphasis on those most relevant to chemistry. One of the most widely used techniques is called total least squares (TLS), which uses measurement error information in developing optimal solutions for regression problems that follow a classical least squares (CLS) model.^{17,18} Although not widely used in chemistry, TLS provides a general framework for solving regression problems according to a maximum likelihood criterion given measurement error information in the form of error variance-covariance structures. A more specialized technique, also developed with a maximum likelihood formulation, is positive matrix factorization (PMF).¹⁹ PMF was originally created to address the problem of heteroscedastic measurement errors in multivariate curve resolution (MCR) of environmental source-receptor data, and this has been its principal area of application. MCR is an example of subspace modeling in chemistry, where a low dimensional subspace (line, plane, hyperplane) is used to model higher dimensional data. One of the most widely used tools for this type of application is principal components analysis (PCA), which is at the root of many chemometric approaches. This motivated the development of maximum likelihood principal components analysis (MLPCA), which incorporates measurement error information into the subspace modeling procedure to develop more optimal solutions. Although a comprehensive treatment of this method is not possible here and the reader is referred to the literature,²⁰⁻²² it deserves some additional attention because of its relevance to chemical applications.

5.1 Maximum likelihood principal components analysis

At its core, MLPCA is a subspace modeling method that uses principles of maximum likelihood estimation to obtain the optimal estimate of a low-rank model for data in a high dimensional space, using measurement error information to define the optimal solution. It is closely related to TLS in terms of its fundamental assumptions about the problem and its solution,²³ but differs in the formulation of the problem and (typically) in the algorithms employed. Practically, MLPCA can be viewed as a kind of weighted PCA that de-emphasizes noisy measurements, but this perspective is overly simplistic, since the complexity of the error structures accommodated by the method goes beyond simple weighting functions. The basic philosophy of MLPCA is to efficiently extricate the variance/covariance structure of the underlying variables from the noise by using knowledge about the error structure in an optimal way. In other words, the more that is known about the noise, the more effectively it can be excised from the meaningful information.

It is important to recognize that, in spite of its name, MLPCA is neither a true maximum likelihood method nor PCA (except in ideal circumstances), although it aspires to be both. As a statistician would point out, it cannot be a maximum likelihood method unless the measurement error structure is exactly known. This is rarely (if ever) the case, and instead MLPCA relies on estimates of the measurement errors obtained from experimental measurements. It is also not PCA, which by definition is a method that models the total variance (information plus noise). In contrast, MLPCA is a subspace modeling technique that endeavors to separate the two sources of variance. Although PCA is often used in chemometrics as a subspace modeling technique, its use in this regard will only be optimal under *iid* measurement error conditions. MLPCA can be regarded as a superset of subspace modeling techniques of which PCA is a subset. The use of MLPCA for subspace modeling also implies that the linear form of the model is valid and the dimensionality of the subspace is correct. If these conditions are not met, then the dimensionality reduction achieved by MLPCA is as arbitrary as any other projection method that might be used.

A significant advantage of MLPCA over other methods that make use of measurement error information is that it casts the results in a PCA framework, which is the starting point for many chemometric methods. This has allowed it to be adapted to a variety of applications that include multivariate calibration,^{16,24} curve resolution,²⁵⁻³⁰ exploratory data analysis,³¹ calibration transfer,³² and multi-way analysis.³³⁻³⁶ The application of MLPCA has also led to the elimination of complex and often poorly understood

data preprocessing methods in many cases by providing a more rational approach to treating measurement errors.^{10,11} In this regard, it is important to recognize that, historically, one of the underlying motivations for preprocessing multivariate data has been to render the measurement errors to be more uniform and uncorrelated so that the assumption of *iid* errors made for the optimal application of PCA is more valid. For example, this is the purpose of scaling, which is intended to make measurement error variances more uniform, and derivative filtering, which is intended to remove correlations in the noise. With MLPCA, the measurement errors are treated in an optimal manner regardless of their form, so such transformations do not generally affect the results.

One of the fundamental differences between PCA and MLPCA is the objective function to be minimized. If \mathbf{X} ($m \times n$) is a matrix of data where the rows represent samples and the columns represent the measured variables, and $\hat{\mathbf{X}}$ is the data estimated by a p -dimensional model, then PCA minimizes the following objective function.

$$S_{PCA}^2 = \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \hat{x}_{ij})^2 \quad (18)$$

In contrast, the objective function for MLPCA uses the measurement error information to weight the residuals. The actual form of this function depends on the structure of the error, but one common form is given in equation 19, where \mathbf{x}_i represents a row vector of \mathbf{X} and Σ_i is its corresponding error covariance matrix.

$$S_{MLPCA}^2 = \sum_{i=1}^m (\mathbf{x}_i - \hat{\mathbf{x}}_i) \Sigma_i^{-1} (\mathbf{x}_i - \hat{\mathbf{x}}_i)^T \quad (19)$$

It should also be noted that the projected data in MLPCA are different from those in PCA. While the latter uses an orthogonal projection, the former uses a maximum likelihood projection, exploiting the measurement error information to obtain the best estimate of the true measurement.

Although MLPCA is cast in a PCA framework, there are other important differences between the methods as well. One of these is that MLPCA actually consists of six different algorithmic implementations (including PCA) that depend on the error structure of the data matrix. These structures are shown in Figure 8 and range from the simplest case of *iid* errors (where PCA is applicable) to the most complex case where the measurement errors are correlated across all of the rows and columns of the data matrix. Cases A, B and D can be solved in a straightforward manner, while cases C, E, and F involve an optimization that is normally implemented through a more computationally intensive alternating least squares (ALS) algorithm. For

most first-order (vectorial) measurements, such as first-order calibration, where each sample can be considered as a row in Figure 8, the measurement errors between samples can be considered to be independent of one another and cases A-E apply. It should also be noted that these cases also cover those instances where the error correlation exists only within the columns (but not the rows), since the matrix is easily transposed. Case F, where the errors are correlated among both the rows and columns, is most commonly observed for second-order measurements (e.g., fluorescence emission-excitation matrices, liquid-chromatography with diode array detection). In such cases, the error covariance matrix must be estimated on the basis of the unfolded (vectorized) data matrix and consequently tends to be very large ($mn \times mn$). In practice, case F is usually too computationally demanding to carry out without some form of data compression. More details on these different error structures and the algorithms employed to treat them are given in the literature.^{20-22,33-36}

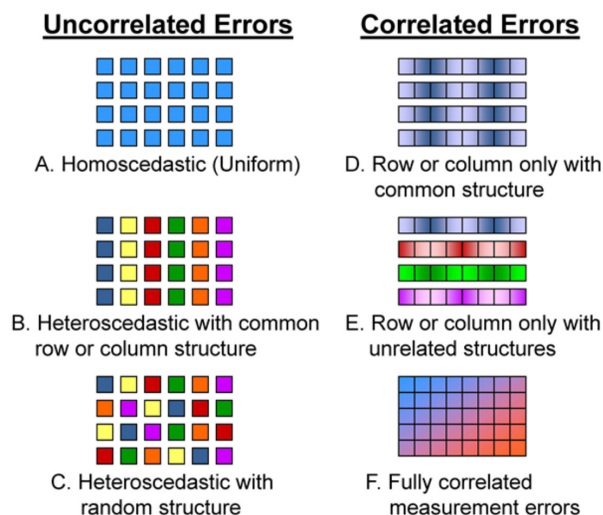


Figure 8. Pictorial representation of different error structures treated by MLPCA algorithms for two-way data. Connectivity of blocks indicates independence or correlation of errors. Colors and shading indicates equality of variances in the measurements.

5.2 A simple example

To illustrate some of the ideas presented here, a simple example involving experimental data will be used. The original data set was obtained by Silva *et al.*³⁷ as part of a study related to forensic document analysis in collaboration with the Brazilian Federal Police and consisted of attenuated total reflectance (ATR) FTIR spectra from 1000 ink samples. The original design included scan, spot and batch replicates from ten different pens, but for simplicity of illustration, only data from one type of pen is included here. The 100 spectra employed for this pen type

consisted of 50 replicates from one batch, and 10 replicates from each of 5 additional batches. In the results presented here, the first 40 replicates from the first batch were used to estimate the measurement error covariance matrix, which was then used in the data analysis of the remaining 60 samples (six batches with ten samples each). In the original paper, the spectra were also preprocessed using the standard normal variate and first-derivative filtering (see the original work for additional details). In the current example, both raw and preprocessed spectra were employed.

Figures 9a and b show the raw and derivative spectra, respectively, of the 100 ink samples, with the first 40 replicates offset from the other samples for clarity. Figures 9c and e show the error covariance and correlation matrices calculated for the replicates of the raw spectra. Figure 9c clearly indicates that the measurement errors are heteroscedastic, while Figure 9e shows a high degree of correlation in the measurement errors. This combination is indicative of multiplicative noise that is commonly observed in IR reflectance spectra. However, there is a dramatic change in the error structure for the preprocessed data. The error covariance matrix of the derivative spectra shown in Figure 9d indicates that heteroscedasticity is still an issue, and in fact the large flat region suggests that the error variance covers a wider range than for the original data. This is not surprising, since derivative filtering will tend to amplify high frequency noise. Negative error correlations are also present at a magnitude similar to the positive ones, although this is difficult to see from the

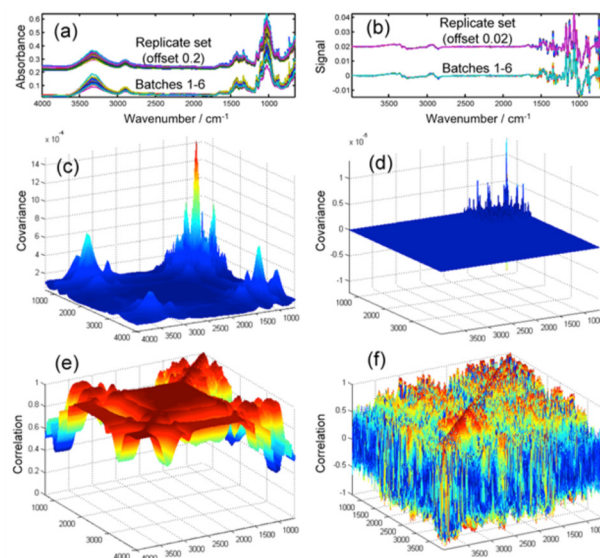


Figure 9. IR reflectance spectra (a) and derivative IR spectra (b) of 60 ink samples (6 batches \times 10 replicates) and 40 replicate ink samples of batch 1 (offset). The error covariance matrices for the 40 batch 1 replicate spectra are shown in (c) and (d) for the raw and derivative spectra, respectively. The error correlation matrices are shown in (e) and (f) for the raw and derivative spectra, respectively.

angle of the figure. This is also a consequence of derivative filtering. Figure 9f shows the error correlation matrix for the derivative spectra and indicates that correlations are still present, although they have a much more complex structure. One of the aims of derivative filtering is remove the correlations in the noise and make it closer to independent (as assumed by PCA and other data analysis methods). While the errors here are not uncorrelated, the correlation structure is certainly more random than that observed for the raw data.

The sixty raw test spectra were analyzed by both PCA and MLPCA, with the estimated error covariance matrix being used in the latter case, and the results are shown as scores plots in Figures 10a and b, respectively. The different symbols and colors in the plot show the projections of ink samples from different batches for this type of pen. While PCA shows some similarities for samples within the same batch, the separation of the batches is not very good. In contrast, MLPCA is able to separate the six batches from one another quite cleanly. This suggests that, by using the measurement error information, MLPCA is better able to distinguish the batch-to-batch variations from the spot-to-spot and scan-to-scan variations, which are considered as measurement noise. Similar conclusions can be drawn for the results of the derivative spectra, shown in Figures 10c and d. Note that PCA on the derivative spectra produces results in Figure 10c that are somewhat better than for the raw spectra in Figure 10a in terms of the batch groupings, but there is still considerable mixing of the batches. These improvements are likely due to changes in the error structure that make PCA more useful. Figure 10d shows the MLPCA results on the derivative

spectra using the corresponding error covariance matrix in Figure 9d. While not quite as good as MLPCA with the raw data, these results retain most of the same clusters as in that case and are clearly superior to the PCA results. The similarity between the MLPCA results in Figures 10b and 10d can be realized by reflecting the points across the diagonal of the latter. The spatial similarity is then apparent in spite of the complex changes that have been introduced in the data. This is not surprising, however, since MLPCA inherently adapts to the changes introduced through preprocessing by incorporating the resulting changes in the error covariance matrix. In this way, it will tend to be more robust to changes introduced by simple preprocessing steps, and preprocessing for the purpose of changing the error structure is rendered somewhat superfluous.

6. Conclusions

One of the main goals of chemometrics is to reveal relevant information hidden within seemingly complex data structures. Many tools have been developed to achieve this end, each with its own implicit or explicit assumptions about the characteristics of the data that allow it to effectively extract the latent information. When such methods fail, it may be because the information sought is not present in the data, or it may be that the assumptions made are not valid. A principal difference between a chemometrician and a statistician is that the former is able to use chemical knowledge to guide the design of experiments and the analysis of data to improve outcomes. Analytical chemists have the ability to assess and interpret the measurement error characteristics of the instruments that they employ, and this knowledge can be a valuable asset in the analysis of multivariate data.

The purpose of this article has been to shed some light on the often neglected role of measurement errors in multivariate measurements, and to describe how measurement errors can be characterized and exploited in data analysis. In most cases, when measurement errors are small or follow an assumed structure, traditional methods of data analysis are likely to yield useful results. However, since many methods are based on modeling the total variance, it is intuitive that knowledge about the error variance should be able to improve the effectiveness of these tools. Moreover, methods that use the error structure are likely to become more important as measurement systems become more complex, expanding beyond traditional spectroscopic and chromatographic applications and into less familiar domains such as high-throughput biological methods. Although some tools have been introduced here, the area of multivariate measurement errors is ripe for

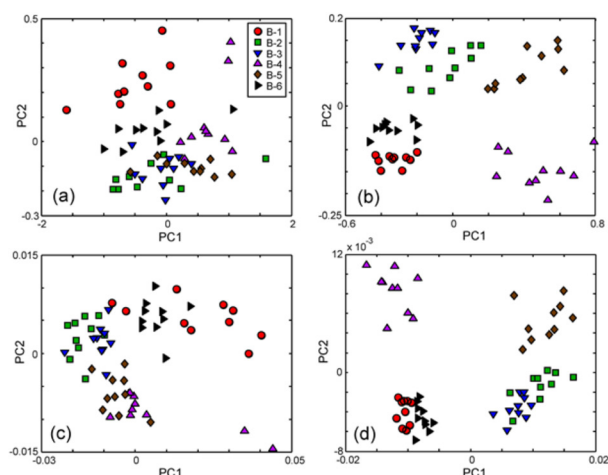


Figure 10. Scores plots (principal component 2 vs. principal component 1) following PCA and MLPCA on the raw and derivative spectra of the 60 samples in Figure 9. (a) PCA applied to raw spectra; (b) MLPCA applied to raw spectra; (c) PCA applied to derivative spectra; and (d) MLPCA applied to derivative spectra.

more research to expand the understanding of measurement errors and their value in multivariate data analysis.

Acknowledgments

The author gratefully acknowledges the financial support of the Natural Sciences and Engineering Research Council (NSERC) of Canada and Dalhousie University, and thanks Maria Fernanda Pimentel and Carolina Santos Silva for allowing use of the ink spectral data for this article.



Peter D. Wentzell obtained his BSc from Dalhousie University (Halifax, Canada) in 1982 and his PhD from Michigan State University (East Lansing, USA) in 1988. Following a Natural Science and Engineering Research

Council Postdoctoral Fellowship at the University of British Columbia (Vancouver, Canada) he joined the faculty in the Department of Chemistry at Dalhousie University where he is currently a Killam Faculty of Science Professor. His research interests include all areas of chemometrics, with a particular emphasis on understanding and exploiting multivariate measurement errors and applications to high throughput biological measurements.

References

1. Wold, S.; Sjöström, M.; *Chemom. Intell. Lab. Syst.* **1998**, *44*, 3.
2. Malinowski, E. R.; *Factor Analysis in Chemistry*, 2nd ed.; Wiley: New York, USA, 1991.
3. Massart, D. L.; Vandeginste, B. G. M.; Buydens, L. M. C.; De Jong, S.; Lewi, P. J.; Smeyers-Verbeke, J.; *Handbook of Chemometrics and Qualimetrics: Part A*; Elsevier: Amsterdam, Netherlands, 1998.
4. Beebe, K. R.; Pell, R. J.; Seasholtz, M. B.; *Chemometrics: A Practical Guide*; Wiley: New York, 1998.
5. Currie, L. A.; *Pure & Appl. Chem.* **1995**, *67*, 1699.
6. Chu, E.; *Discrete and Continuous Fourier Transforms: Analysis, Applications and Fast Algorithms*; Taylor and Francis: Boca Raton, USA, 2008.
7. Sundararajan, D.; *Discrete Fourier Transform: Theory, Algorithms and Applications*; World Scientific: Singapore, 2001.
8. Wentzell, P. D.; Brown, C. D. In *Encyclopedia of Analytical Chemistry*; Meyers, R. A., ed.; Wiley: Chichester, UK, 2000, vol. 11, p. 9764.
9. Kragten, J.; *Analyst* **1994**, *119*, 2161.
10. Brown, C. D.; Wentzell, P. D.; *J. Chemom.* **1999**, *13*, 133.
11. Brown, C. D.; Vega-Montoto, L.; Wentzell, P. D.; *Appl. Spectrosc.* **2000**, *54*, 1055.
12. Wentzell, P. D.; Hou, S.; *J. Chemom.* **2012**, *26*, 264.
13. Leger, M. N.; Wentzell, P. D.; *Appl. Spectrosc.* **2004**, *54*, 855.
14. Leger, M. N.; Vega-Montoto, L.; Wentzell, P. D.; *Chemom. Intell. Lab. Syst.* **2005**, *77*, 181.
15. Karakach, T. K.; Wentzell, P. D.; Walter, J. A.; *Anal. Chim. Acta* **2009**, *636*, 163.
16. Schreyer, S. K.; Bidinosti, M.; Wentzell, P. D.; *Appl. Spectrosc.* **2002**, *56*, 789.
17. Markovsky, I.; Van Huffel, S.; *Signal Processing* **2007**, *87*, 2283.
18. Markovsky, I.; *Stat. Interface* **2010**, *3*, 329.
19. Paatero, P.; Tapper, U.; *Environmetrics* **1994**, *5*, 111.
20. Wentzell, P. D.; Andrews, D. T.; Hamilton, D. C.; Faber, K.; Kowalski, B. R.; *J. Chemom.* **1997**, *11*, 339.
21. Wentzell, P. D.; Lohnes, M. T.; *Chemom. Intell. Lab. Syst.* **1999**, *45*, 65.
22. Wentzell, P. D. In *Comprehensive Chemometrics*; Brown, S. D.; Tauler, R.; Walczak, B., eds.; Elsevier: Amsterdam, Netherlands, 2009; vol. 2, ch. 25.
23. Schuermans, M.; Markovsky, I.; Wentzell, P. D.; Van Huffel, S.; *Anal. Chim. Acta* **2005**, *544*, 254.
24. Wentzell, P. D.; Andrews, D. T.; Kowalski, B. R.; *Anal. Chem.* **1997**, *69*, 2299.
25. Wentzell, P. D.; Karakach, T. K.; Roy, S.; Martinez, M. J.; Allen, C. P.; Werner-Washburne, M.; *BMC Bioinformatics* **2006**, *7*, 343.
26. Tauler, R.; Viana, M.; Querol, X.; Alastuey, A.; Flight, R. M.; Wentzell, P. D.; Hopke, P. K.; *Atmos. Environ.* **2009**, *43*, 3989.
27. Jaumot, J.; Pina, B.; Tauler, R.; *Chemom. Intell. Lab. Syst.* **2010**, *104*, 53.
28. Stanimirova, I.; Tauler, R.; Walczak, B.; *Environ. Sci. Technol.* **2011**, *45*, 10102.
29. Dadashi, M.; Abdollahi, H.; Tauler, R.; *Chemom. Intell. Lab. Syst.* **2012**, *118*, 33.
30. Dadashi, M.; Abdollahi, H.; Tauler, R.; *J. Chemom.* **2013**, *27*, 34.
31. Wentzell, P. D.; Hou, S.; *J. Chemom.* **2012**, *26*, 264.
32. Andrews, D. T.; Wentzell, P. D.; *Anal. Chim. Acta* **1997**, *350*, 341.
33. Vega-Montoto, L.; Wentzell, P. D.; *J. Chemom.* **2003**, *17*, 237.
34. Vega-Montoto, L.; Gu, H.; Wentzell, P. D.; *J. Chemom.* **2005**, *19*, 216.
35. Vega-Montoto, L.; Wentzell, P. D.; *J. Chemom.* **2005**, *19*, 236.
36. Vega-Montoto, L.; Wentzell, P. D.; *Anal. Chim. Acta* **2006**, *556*, 383.
37. Silva, C. S.; Borba, F. S. L.; Pimentel, M. F.; Pontes, M. J. C.; Honorato, R. S.; Pasquini, C.; *Microchem. J.* **2013**, *109*, 122.

Submitted: October 21, 2013

Published online: December 10, 2013