Accepted for publication
The Transactions of the Institute of Measurement and Control 24(4):333-353 01 Jan 2002

PRINCIPAL COMPONENT ANALYSIS OF SPECTRA

With Application to Acoustic Emissions From Mechanical Equipment

C.C. Tan and N.F. Thornhill

Department of Electronic and Electrical Engineering, University College London, Torrington Place, London WC1E 7JE, UK

R.M. Belchamber

Process Analysis and Automation Ltd, Fernhill Road, Farnborough, Hampshire GU14 9RX, UK

Abstract

This paper discusses principal component analysis (PCA) of integral transforms (spectra and autocovariance functions) of time-domain signals. It is illustrated using acoustic emissions from mechanical equipment. It was found that acoustic signals from different stages of operation appeared as distinct clusters in the PCA analysis. The clusters moved when machinery faults were present and the modelling errors also increased under fault conditions, thus each type of fault had a distinctive signature and could be diagnosed. PCA using autocovariance functions that were derived from the full power spectrum had better performance than spectral PCA using averaged periodograms, and both gave a significant improvement over time domain PCA.

Keywords: Acoustic emission; Autocovariance function; Fault diagnosis; Integral transform; Machinery condition monitoring; Principal component analysis (PCA); Spectral PCA; Spectrum.

1. Introduction

Principal component analysis (PCA) detects co-linearity between signals. It is often the case that there are just a few types of underlying behaviour (Wise *et.al*., 1990; Kresta *et.al.*, 1991) which can be captured during normal operation and exploited in the detection of abnormal situations. Phase shifts caused by time delays can, however, make it difficult to detect co-linearity thus causing an overestimate of the true dimension of the data set. For instance, a phase lag of a quarter cycle destroys the correlation between harmonic signals because $\cos(\omega t)$ and $\sin(\omega t)$ are orthogonal

functions. Spectral PCA addresses this problem. It is insensitive to phase delays because the calculation of the power spectrum removes phase information. The same comment applies to the autocovariance function. The purpose of this paper is to present methods for PCA using power spectra and autocovariance functions and to demonstrate their benefits through an application in on-line diagnosis of faults in mechanical equipment. It gives advances on previously published work because it quantifies the behaviour of clusters in a PCA plot under fault conditions and suggests physical explanations for the changes observed. It also highlights a problem with the use of averaged power spectra and gives a solution using PCA with autocovariance functions.

The methods are illustrated with acoustic emissions from a washing machine in the laundry room of a student residence. A washing machine has characteristics similar to a batch processing unit. For instance, the spin cycle is a centrifugation operation while the wash cycle resembles a mixing operation. The application therefore has relevance to chemical process industries as well as to the consumer goods industry.

The following section places the project in the context of previous work. Section 3 presents the methods for PCA and for the acoustic emission application. It also explores data pre-processing issues. Section 4 discusses the results and demonstrates diagnosis of faulty operation of the washing machine. A comparison with time-domain dynamic PCA is also presented. The paper ends with a conclusions section.

2. Background

Spectral Analysis: Near infrared spectra and infrared spectra are routinely analysed by PCA or used in Partial Least Squares (PLS) models for estimation of analyte concentrations in unknown samples (for example, Karstang and Henrikson, 1991; Riley et.al., 1997; Yeung et.al., 1999). Seasholtz (1999) described how the application of multivariate calibration in analytical methods of NIR and NMR spectroscopy at Dow Chemical has made money for the company. In these cases the instruments themselves created the spectra, but the same benefits can be achieved through numerical calculation of the spectra.

Spectral analysis gives several advantages compared to analysis in the time domain. It gives improved signal to noise ratio if the spectral content of the wanted signal is narrow-band compared to the noise. The power spectrum is insensitive to time delays or phase shifts caused by process dynamics. It is also insensitive to missing values or to outliers, because the transforms of such effects are spread across all frequencies in the spectrum.

Acoustic emissions: Acoustic emissions have been exploited for machinery condition monitoring and non-destructive testing (Swindlehurst, 1973; Dornfield, 1992; Holroyd and Randall, 1993; Sharif and Grosvenor, 1998; Holroyd and Bradshaw, 2000) and used for the integrity monitoring of process equipment such as pressure vessels (Wood and Harris, 2000).

Applications of audio and ultrasonic acoustic monitoring in the process industries have been reviewed by Wade *et.al*. (1991) and Boyd and Varley (2001), and identified as a powerful non-invasive technique for extraction of process information (Belchamber and Collins, 1999). Those authors discussed chemical reactions, crack formation, particle impact, and phase transitions as sources of acoustic emissions in processes. Hou *et. al*. (1999) used non-invasive passive acoustic monitoring for analysis of the flows of dense slurries of silica particles. Cody *et. al*., (2000) analysed the wall vibrations caused by impacts of small particles in a fluid bed catalytic cracking unit. Whitaker *et. al*. (2000) described a new non-invasive monitoring technique which could detect changes in physical properties of powders during granulation and indicate the end-point of the process. These examples show that the use of acoustic monitoring is established in the chemical process industries both for equipment condition monitoring and for process monitoring.

Multivariate statistical analysis of acoustic emissions: Principal component analysis (PCA) was used on specific features of acoustic signals such as intensity and duration by Betteridge *et.al*., (1981), who also mentioned the benefits of an inspection of the cepstrum (derived from a Fourier transform of the logarithm of the spectrum) and of autocorrelation analysis. They did not, however, use the whole spectrum in the PCA analysis. The classification of the spectra of acoustic signals using multivariate statistical analysis was described by Belchamber and Collins (1993). The patent cited the use of an approach to PCA known as SIMCA and gave examples of the classification of the acoustic spectra in 24 frequency channels from a pump and an industrial blender. Tabe *et.al*., (1998) presented a process application of spectral PCA while Wu *et.al*. (1999) used a method known as "eigenfaces" in the recognition of sounds from car engines. The method was used to determine whether an unknown vehicle sound was among the training data set by examination of the distance of its spectrum from the weightings (score) plot of the calibration model. They did not, however, examine clustering of spectra in the PCA score plots.

Aldrich and Theron (2000) related the power spectral densities of acoustic emissions from a laboratory-scale ball-mill to particle size distributions by use of multivariate continuum regression. They reported improved modelling of particle size distributions compared to Kalman filtering and PCR (principal components regression). Principal component analysis has proved useful in the

analysis of the relationship between the crispness of apples, which was modified by storage conditions, and recorded chewing sounds (De Belie, et. al., 2000).

The conclusion from this review is that an application showing fault diagnosis by means of PCA on the spectra and autocovariance functions of acoustic emissions can add to an existing body of work in a field in which there is already some interest.

3. Methods

3.1. Formulation of spectral and autocovariance PCA

Spectral PCA: Descriptions of PCA may be found in Chatfield and Collins, (1980), Wold *et.al*., (1987); Wise *et. al*., (1990), Kresta *et.al*., (1991); and Wise and Gallagher (1996), and others. The novel feature in spectral PCA is that the columns of the data matrix, $\boldsymbol{X}$, are the power spectra $P(f)$ of the signals being analysed.

$$\boldsymbol{X} = \begin{array}{c} \begin{matrix} m\ process\ variables \rightarrow \end{matrix} \\ \begin{pmatrix} P_1(f_1) & \ldots & P_m(f_1) \\ .. & .. & .. \\ P_1(f_N) & \ldots & P_m(f_N) \end{pmatrix} \begin{matrix} N\ frequency \\ channels \\ \downarrow \end{matrix} \end{array}$$

Smoothed power spectra were calculated using the averaged periodogram method (Welch, 1967). An averaged periodogram may use considerably fewer frequency channels than the original number of data points and therefore data compression is achieved. For example, the sound samples used in this work had 44000 samples while their periodograms had 1024 frequency channels.

A full PCA decomposition reconstructs the $\boldsymbol{X}$ matrix as a sum over m orthonormal basis functions $\boldsymbol{t}_1$ to $\boldsymbol{t}_m$:

$$\boldsymbol{X} = \boldsymbol{t}_1 \begin{pmatrix} w_{1,1} & \ldots & w_{1,m} \end{pmatrix} + \boldsymbol{t}_2 \begin{pmatrix} w_{2,1} & \ldots & w_{2,m} \end{pmatrix} + \ldots + \boldsymbol{t}_m \begin{pmatrix} w_{m,1} & \ldots & w_{m,m} \end{pmatrix}$$

In spectral PCA the column vectors $\boldsymbol{t}_i$ are spectrum-like functions having N frequency channels. When post-multiplied by a row vector of coefficients such as $\begin{pmatrix} w_{1,1} & \ldots & w_{1,m} \end{pmatrix}$ the result is a matrix of the same dimensions as $\boldsymbol{X}$. It is usual in PCA to normalise the vectors of weightings (e.g. Wold *et.al*., 1987; Wise and Gallagher, 1996). In spectral PCA, however, it is convenient to normalise the

$\boldsymbol{t}-$vectors in order to achieve an interpretation of the spectra as a reconstruction over an orthonormal basis set. This formulation of PCA is also known as classical scaling (Chatfield and Collins, 1980). Therefore the $\boldsymbol{t_i}-$vectors are formally the *m* independent normalised eigenvectors of the *N* by *N* matrix $\boldsymbol{XX'}$, where $\boldsymbol{X'}$ is the transpose of $\boldsymbol{X}$. They are ordered according to the size of the associated eigenvalue. The ratio between the eigenvalue and the sum of all the eigenvalues also gives a measure of the total spectral variation captured by that eigenvector.

The above expression may be written compactly as $\boldsymbol{X} = \boldsymbol{T}\boldsymbol{W'}$ where the columns of $\boldsymbol{T}$ are $\boldsymbol{t}_1$ to $\boldsymbol{t}_m$ and the rows of $\boldsymbol{W'}$ are $\left(w_{1,1} \quad ... \quad w_{1,m}\right)$ to $\left(w_{m,1} \; ... \; w_{m,m}\right)$. The orthonormality of the columns of $\boldsymbol{T}$ means that $\boldsymbol{W'} = \boldsymbol{T'X}$. Singular value decomposition $\boldsymbol{X} = \boldsymbol{UDV'}$ provides a means for computation of the requisite vectors with $\boldsymbol{T} = \boldsymbol{U}$ and $\boldsymbol{W'} = \boldsymbol{DV'}$. Matrix $\boldsymbol{D}$ is diagonal and its elements are the positive square roots of the *m* non-zero eigenvalues of $\boldsymbol{XX'}$.

The majority of the variation in $\boldsymbol{X}$ can often be captured by truncating the PCA description. If all the variables had similar spectral features then one term would describe most or all of the variability in the spectra. In other cases more terms may be needed. The following is a three-term PCA model in which the variation of $\boldsymbol{X}$ that is not captured by the three terms appears in an error matrix $\boldsymbol{E}$:

$$\boldsymbol{X} = \boldsymbol{t}_1\left(w_{1,1} \quad ... \quad w_{1,m}\right) + \boldsymbol{t}_2\left(w_{2,1} \quad ... \quad w_{2,m}\right) + \boldsymbol{t}_3\left(w_{3,1} \quad ... \quad w_{3,m}\right) + \boldsymbol{E}$$

The issue of the correct number of terms was discussed by Chatfield and Collins (1980), Valle *et.al*., (1999) and elsewhere. In the work reported here the models needed two or three terms, the decision to truncate being made when the eigenvalue associated with the next $\boldsymbol{t}$-vector represented less than 5% of the sum of all the eigenvalues.

Each spectrum may be represented as a point in a reduced space. For instance, when three $\boldsymbol{t}$-vectors are in use the *i*'th spectrum maps to a point having the co-ordinates $w_{1,i}$, $w_{2,i}$ and $w_{3,i}$ in a three-dimensional space. Similar spectra have similar w_1, w_2 and w_3 co-ordinates. Therefore such groups form clusters. When two $\boldsymbol{t}$-vectors are used, the weightings plot is two dimensional.

<u>*Autocovariance PCA*</u>: The autocovariance function of a zero-mean signal *y* having *n* samples is shown below. The index ℓ is known as the lag and the quantity being summed is the signal times a time lagged version of the same signal. The range of lags used was $\ell = 0$ to $\ell = 1023$, a total of 1024 channels.

$$A(\ell) = \frac{1}{n-\ell} \sum_{i=\ell+1}^{n} y[i]\, y[i-\ell]$$

It can be shown that the autocovariance is the inverse Fourier transform (*iFFT*) of the full two-sided power spectrum (i.e. not averaged). The autocovariances in this work were derived from the power spectra rather than by direct calculation because that route enabled low frequency interference to be removed, as discussed below. The autocovariance functions were truncated to the first 1024 channels of the *iFFT* to give the same size ***X*** matrix as for spectral PCA.

Principal component analysis using autocovariance functions was formulated in the same way as was described for spectral PCA except that the columns of the data matrix were the autocovariance functions of the acoustic signals.

<u>*Data pre-processing*</u>: It is usual in time-domain PCA to mean centre the data sets and to scale to unit variance. In spectral PCA the following data pre-processing may be applied:

- Mean centre the time trend and remove linear trends before calculation of the spectra;
- Remove known interferences from the spectra;
- Scale the spectra to the same total power such that:

$$\sum_{i=1}^{N} P(f_i) = constant$$

where $P(f_i)$ is the spectral power in the i'th channel;

or

- Scale the autocovariance functions so that the covariance at zero lags is unity.

The first of these steps removes steady offsets or linear ramps from the time domain data which would otherwise appear in the $f = 0$ channel of the spectrum. It can be applied when the features of interest are the deviations from the mean. The scaling step is equivalent to scaling to constant variance in the time domain because the signal variance is proportional to total power. Alternatively the spectra may be left unscaled if a meaning can be attributed to the intensity of the signal.

Interference such as mains pick-up may be dealt with by setting the values of spectral power to zero in the frequency channels where the interference is located. The ease of such an approach is a benefit of spectral PCA. In time domain PCA, by contrast, interference must be removed by analogue or digital filtering but the design of a narrow band filter is a skilled task. For the

autocovariance functions the interference must be removed from both positive and negative frequencies in the two-sided power spectrum before the inverse FFT.

Fault detection with spectral PCA: This section describes fault detection using spectral PCA. The methods also apply to PCA using the autocovariance functions. When new acoustic sound samples become available, the matrix $\boldsymbol{W}'_{new}$ below gives the weightings of the of the new spectra (or autocovariance functions) in the data matrix $\boldsymbol{X}_{new}$. The expression is the least squares projection of $\boldsymbol{X}_{new}$ onto the ***t***-vectors used in the calibration model, which takes a simple form because of the orthonormality of the ***t***-vectors:

$$\boldsymbol{W}'_{new} = \left(\left(\boldsymbol{T}'_{cal}\, \boldsymbol{T}_{cal} \right)^{-1} \boldsymbol{T}'_{cal}\, \boldsymbol{X}_{new} \right) = \boldsymbol{T}'_{cal}\, \boldsymbol{X}_{new}$$

where $\boldsymbol{T}_{cal}$ is a non-square matrix whose columns when three basis functions were in use would be $\boldsymbol{t}_1$, $\boldsymbol{t}_2$ and $\boldsymbol{t}_3$. Thus the three term model for a matrix $\boldsymbol{X}_{new}$ containing *p* spectra in *p* columns is:

$$\boldsymbol{X}_{new} = \boldsymbol{t}_1 \left(w_{new\,1,1} \;\cdots\; w_{new\,1,p} \right) + \boldsymbol{t}_2 \left(w_{new\,2,1} \;\cdots\; w_{new\,2,p} \right) + \boldsymbol{t}_3 \left(w_{new\,3,1} \;\cdots\; w_{new\,3,p} \right) + \boldsymbol{E}_{new}$$

When a two term model is used then the columns of $\boldsymbol{T}_{cal}$ would be $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ and the model for $\boldsymbol{X}_{new}$ would have no $\boldsymbol{t}_3$ term. Previously unseen spectra may arise if the process develops a fault. It may be the case that the new spectra are different linear combinations of the same ***t***-vectors found during calibration. Then the spectra representing a fault condition would form a distinct cluster in the weightings plot. But the new spectra may also be qualitatively different from the calibration spectra, for instance if the fault generated spectral components at other frequencies. In that case the ***t***-vectors may have little in common with the new spectra and the modelling error $\boldsymbol{E}_{new}$ would be large. $\boldsymbol{E}_{new}$ can be expressed as:

$$\boldsymbol{E}_{new} = \boldsymbol{X}_{new} - \hat{\boldsymbol{X}}_{new}$$

where $\hat{\boldsymbol{X}}_{new} = \boldsymbol{T}_{cal}\, \boldsymbol{W}'_{new}$ is the best fit of the new spectra onto the calibration vectors $\boldsymbol{t}_1$ to $\boldsymbol{t}_3$ (onto $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ only in a two-term model).

The columns of $\boldsymbol{E}_{new}$ are the unmodelled part of the spectra in $\boldsymbol{X}_{new}$. Therefore in a fault detection application the magnitudes of the columns of $\boldsymbol{E}_{new}$ should be inspected. The following vector of squared prediction errors (SPE) in which each term is the sum of squares of the elements of a column of $\boldsymbol{E}_{new}$ (Jackson, and Mudholkar, 1979; Kourti *et.al*., 1996)

$$SPE = \left(\sum_{i=1}^{N} \left(E_{new_{i,1}} \right)^2 \quad \sum_{i=1}^{N} \left(E_{new_{i,2}} \right)^2 \quad \cdots \quad \sum_{i=1}^{N} \left(E_{new_{i,p}} \right)^2 \right)$$

In order to detect a significant SPE, a threshold may be set by examination of the normal variability during normal running. Then any spectra that map into the PCA space with a larger error would be classified as abnormal.

3.2 Fault diagnosis of mechanical equipment

Acoustic emission application: The acoustic emissions used for the demonstration were the sounds made by a washing machine. The sounds were captured by a laptop computer using an audio microphone and an A/D converter sampling at 22k*Hz* and thus each sound sample of duration 2*s* comprised 44000 samples. The files were stored in *wav* format and converted to text using LabVIEW (National Instruments, Austin, TX).

For calibration, five sound samples were taken from each of the fill, wash, drain and spin stages during normal operation of the machine, a total of 20 recordings. Further samples were taken for model validation and also with the machine running under the following fault conditions: with a foreign object (a metal belt buckle) in the laundry, with no laundry in the drum, and with cold water only. Sound samples were also taken of the ambient background sounds with the washing machine switched off.

Calibration ***t***-vectors were derived using a data matrix whose columns were the twenty spectra (or autocovariance functions) of the sound samples recorded during normal operation. The weightings for spectra (or autocovariance functions) from the validation run and the faulty runs were determined by projection onto the calibration ***t***-vectors. Their clusters were inspected for systematic shifts from normal operation and the SPE vector was inspected for modelling errors.

Retrospective fault diagnosis: The diagnosis of faults required distinct signatures for each fault. As discussed above, faults may be characterised by modelling errors (the SPE) or by shifts within the PCA weightings plot, or both.

The SPE plots were inspected for SPE values larger than a threshold value. To quantify the shifts of the clusters in the PCA plots the sound samples from each run of the washing machine were arranged as "sound-sets" comprising one sound sample from each of the fill, wash, drain and spin stages. For instance, the calibration run generated five sound-sets because five sound samples were recorded during filling, five during washing and so on.

Hierarchical classification trees were generated to quantify shifts of these {fill wash drain spin} sound-sets within the PCA plot by means of a scaled Euclidian measure of the distances between them. In autocovariance PCA, where two terms were used in the PCA model, each sound-set was described by a vector with eight elements:

$$\begin{pmatrix} w_{f_1} & w_{f_2} & w_{w_1} & w_{w_2} & w_{d_1} & w_{d_2} & w_{s_1} & w_{s_2} \end{pmatrix}$$

where w_{f_1} and w_{f_2} are the weightings in the two-dimensional PCA model of the fill sound samples, and w_{w_1} and w_{w_2} the weightings for the wash samples. Likewise w_{d_1} and w_{d_2} refer to the drain samples and w_{s_1} and w_{s_2} to the spin samples. The numerical ranges of the weightings are influenced by the diagonal elements in the $\boldsymbol{D}$ matrix. Hence it is necessary to normalise the weightings just as the axes are scaled to occupy a square in the visual weightings plots. The scaling factor is the square root of the eigenvalue and the scaled Euclidian distance between two sound sets of the machine having weightings denoted by symbols w and v is thus given by:

$$d = \sqrt{\left(\begin{array}{l} \dfrac{\left(w_{f_1} - v_{f_1}\right)^2}{\lambda_1} + \dfrac{\left(w_{f_2} - v_{f_2}\right)^2}{\lambda_2} + \dfrac{\left(w_{w_1} - v_{w_1}\right)^2}{\lambda_1} + \dfrac{\left(w_{w_2} - v_{w_2}\right)^2}{\lambda_2} + \dfrac{\left(w_{d_1} - v_{d_1}\right)^2}{\lambda_1} \\ \qquad + \dfrac{\left(w_{d_2} - v_{d_2}\right)^2}{\lambda_2} + \dfrac{\left(w_{s_1} - v_{s_1}\right)^2}{\lambda_1} + \dfrac{\left(w_{s_2} - v_{s_2}\right)^2}{\lambda_2} \end{array} \right)}$$

where λ_1 and λ_2 are the eigenvalues associated with the first and second principal components. Spectral PCA required three principal components and in that case the vectors had twelve elements since weightings w_{f_3}, w_{w_3}, w_{d_3} and w_{s_3} scaled by $\sqrt{\lambda_3}$ were also needed.

The distances between every sound-set and every other sound-set were calculated and arranged in a symmetric matrix with zeros on the diagonals. A hierarchical classification tree was generated as described by Chatfield and Collins (1980). At the first step the smallest non-zero distance $d_{i,j} = d_1$ in the matrix was identified. Its row and column indexes *i* and *j* indicated which two sound-sets were most similar and these were grouped as a cluster characterised by the distance d_1. A smaller matrix was then generated from the original. It did not have rows and columns for the two similar sound-sets identified at step 1. Instead, it had one row and column that gave the distances of all the other sound-sets from the cluster. For the *n*'th set, the distance would be $\min\left\{d_{i,n}, d_{j,n}\right\}$, i.e. the distance between the *n*'th sound-set and whichever member of the cluster was closer. The procedure was then repeated until all the sound-sets had been placed in clusters within the classification tree. At any stage, the outcome of the next step would be either another sound-set added to a cluster already identified or the combining of two sound-sets to start a new cluster.

The classification tree is quantitative and able to characterise shifts in the PCA plot because it shows the closest distance between a given sound-set and other sound-sets.

On-line fault diagnosis: On-line fault diagnosis can use only the sound samples collected up to the present time. For example, at the end of the fill stage an on-line sound-set would contain only the recordings from the filling operation. Hence the distance measures for the classification tree would be:

- After acoustic monitoring of filling:

$$d = \sqrt{\frac{\left(w_{f_1} - v_{f_1}\right)^2}{\lambda_1} + \frac{\left(w_{f_2} - v_{f_2}\right)^2}{\lambda_2}}$$

- After acoustic monitoring of filling and washing:

$$d = \sqrt{\frac{\left(w_{f_1} - v_{f_1}\right)^2}{\lambda_1} + \frac{\left(w_{f_2} - v_{f_2}\right)^2}{\lambda_2} + \frac{\left(w_{w_1} - v_{w_1}\right)^2}{\lambda_1} + \frac{\left(w_{w_2} - v_{w_2}\right)^2}{\lambda_2}}$$

- After acoustic monitoring of filling and washing and draining:

$$d = \sqrt{\frac{\left(w_{f_1} - v_{f_1}\right)^2}{\lambda_1} + \frac{\left(w_{f_2} - v_{f_2}\right)^2}{\lambda_2} + \frac{\left(w_{w_1} - v_{w_1}\right)^2}{\lambda_1} + \frac{\left(w_{w_2} - v_{w_2}\right)^2}{\lambda_2} + \frac{\left(w_{d_1} - v_{d_1}\right)^2}{\lambda_1} + \frac{\left(w_{d_2} - v_{d_2}\right)^2}{\lambda_2}}$$

The data from the washing machine application were analysed as if on-line. The calibration model was the same as before. Then classification trees were created using the three distance measures given above and the appropriate SPEs were also examined. For instance, after filling and washing the fill and wash SPEs were used in the on-line diagnosis.

4. Results and discussion

4.1 Inspections of the data

Acoustic signals: Figure 1 shows normalised time domain plots of the acoustic emissions during normal operation of the washing machine. The actual magnitude of the sounds before normalisation was different for each stage because, for instance, filling was quieter than washing. Within each panel are shown two short sound samples having the same vertical axis scaling but offset so that they can be viewed. Some oscillatory behaviour can be observed but the samples do not overlay cleanly because there are many fluctuations and because the oscillations are phase shifted relative to one another. The signals from different stages of operation do look different, but one would find it hard to make a reliable visual classification of a new signal.

Spectra: The left hand column in Figure 2 shows normalised spectra plotted on a linear *y*-axis for calibration samples from normal operation. The frequency axis extends up to 1*kHz*. There were five spectra for each of the fill, wash, drain and spin stages of the machine cycle. Figure 2 also shows spectra from the *empty* fault condition in which the machine was run with no load of laundry. The presence of peaks in the spectra show that the acoustic emissions contained a number of tones of well defined frequency. For instance, distinct peaks appeared at 100, 200, 300, 400 and 600*Hz* in the drain spectra. The spectral peaks at integer multiples of 100*Hz* are most likely the harmonics of a non-sinusoidal 100*Hz* acoustic waveform. Comments below show that the laptop computer itself

contributed a pure 100*Hz* tone with negligible harmonics. The conclusion is that the laptop and the machinery independently gave acoustic emissions at 100*Hz*. This coincidence may be related to the fact the British mains frequency is 50*Hz* and gives a 100*Hz* signal when rectified by a d.c. power supply unit or a motor drive.

Figure 2 shows that spectra from one stage of operation resembled one another more than they resembled the spectra from another stage, thus some clustering on the PCA plots can be expected. The spectra from the empty machine differed from those of the normal operation. Therefore changes in the PCA clusters can be expected for the runs with faults.

All the spectra in Figures 2 were normalised to the same total power. The reason why the *fill* spectra appear visually to have a smaller magnitude is that the frequency axis extends beyond 1*kHz*. There were frequency channels in the 4-6*kHz* range where the *fill* signals had spectral content and these contributed to the total power.

The bottom left panel in Figure 2 shows the spectra of ambient background sound. The ambient noise was a quieter signal than the others but it appears similar in magnitude when normalised to the same total power. A feature of the background sound was its content at low frequency and a peak at 100*Hz*. It was assumed, therefore, that ambient interference appeared in the frequency channels up to 100*Hz* in the spectra of the sounds from the washing machine. The lowest frequency channels up to and including the 100*Hz* peak were therefore removed from all the spectra, as described in the Methods section. The origin of the 100*Hz* ambient sound was the lap top computer itself. The origin of the lower frequency sounds is unknown, it could be building services and other inner city sounds such as traffic.

Autocovariance functions: Figure 3 shows examples of the autocovariance functions. The horizontal axes show the lag in seconds. The vertical axis has the same scaling (-1 to $+1$) for all the plots and the 100*Hz* interference has been removed. The calculations eliminated phase information so the autocovariance functions, like the spectra, overlay one another and form distinct groups.

The oscillating patterns in Figure 3 are characteristic of signals with harmonic behaviour. As mentioned, the acoustic signals in this application contained several well defined tones. The autocovariance shows a peak whenever such a signal is lagged by a whole number of cycles.

4.2 Discussions of PCA results

Comparison of PCA models: Analysis of twenty calibration sound samples from normal operation is summarised in Table 1 which shows the amount of variation in the calibration set captured by the

first to fourth principal components. Truncation of the model when the next principal component offered less than 5% led to an autocovariance PCA model having two terms and explaining in total over 91% of the variation. The spectral PCA model required three terms and explained 95%, while a two term spectral PCA model gave 89%. PCA of an ***X*** matrix containing the time-domain sound samples as its columns showed little ability to model the data since even four terms only explained 32% of the variation. Time domain PCA is ineffective for dynamic data because phase lags destroy correlations between the data trends. By contrast, PCA using the spectra or autocovariance functions can detect similarities between the signals because these transforms are invariant to signal phase.

PCA using autocovariance functions outperformed spectral PCA in this application is because calculation of the spectra as a periodogram involved an averaging step. The autocovariance functions, by contrast, used the full two sided spectra and were truncated only after all transform steps were completed. Therefore, if a reduced size data matrix is needed then PCA using autocovariance functions would be preferred to spectral PCA. In another project using shorter data sets no averaging was applied. It was found there that the performance of spectral PCA and PCA using autocovariance functions was identical. Thus if no size reduction is needed then either may be used.

	PC1	PC2	PC3	PC4
autocovariance PCA	82%	9.2%	(4.5%)	(2.1%)
spectral PCA	83%	5.9%	5.3%	(2.6%)
time domain PCA	10%	9.0%	7.0%	5.9%

Table 1. Variance in the ***X*** matrix captured by the first four principal components

<u>*Normal operation*</u>: The PCA analyses of the five runs from normal operation are shown as the solid black symbols in Figure 4. The meanings of the symbols are: (ℓ) fill; (⌣) wash; (⌣) drain; (◆) spin. The left hand panels are for spectral PCA, the right hand pair are for PCA using the autocovariance functions (there is no significance in the fact the axis ranges are different, this is the result of the scaling in use). The clusters using autocovariance PCA were more distinct than those using spectral PCA. That is because the third principal component was needed in the spectral case. A three dimensional plot for spectral PCA (Figure 5) shows that three components gave better clustering.

The white symbols in Figure 4 are the spectra from sound samples from other normal runs of the machine which were used for model validation. The additional wash and spin samples were from the same machine on a different occasion while the fill and drain samples were from a different machine.

These additional spectra were projected onto the ***t***-vectors derived from the 20 calibration spectra. Ideally the white clusters would coincide with their black counterparts.

The results from PCA using autocovariance showed little movement in the fill, drain or spin clusters. By contrast, the wash cluster shifted so the conclusion is that there was some variability in the acoustic emissions from the wash stage even during normal operation. The two dimensional spectral PCA plot, however, was difficult to interpret because the clusters were not very distinct. The three-dimensional plot (figure 5) shows better clustering and demonstrates that three principal components were need to model the spectra.

Lower panels in Figure 4 show the squared projection error plots. The black symbols represent the modelling errors in the calibration model and the white symbols are the SPEs for the validation samples. The validation drain samples, which were from a different machine, showed a small modelling error, as did the wash samples from the same machine. The fill samples were also from a different machine but they had no SPE error, suggesting that the normal filling sound was consistent from one machine to another. Nevertheless, the variation in the drain samples suggests that each machine may need its own calibration model.

Fault conditions: Figures 6 to 8 show the black clusters of the calibration spectra from normal operation and white projected clusters for the fill, wash, drain and spin stages for the same machine under various fault conditions. The SPE plots are also shown. In all cases the clusters were more distinct when autocovariance functions were used than for spectral PCA. The clusters on the weightings plots showed shifts and the SPE plots also showed that some conditions gave significant modelling errors. Visual observations from the weightings plots for autocovariance PCA are as follows:

- Foreign object (Figure 6): The SPE plot shows the spectra were well modelled by the calibration ***t***-vectors. The wash, drain and spin clusters moved, however, so the *foreign* fault should be distinguishable from normal operation where only the wash cluster was variable. It is thought that the belt buckle (the foreign object) was hitting the metal drum during draining and the early stages of spinning.
- Empty (Figure 7): The wash and spin clusters moved and there was a large SPE error in the drain stage. The reason for the large SPE for the drain samples can be seen in Figure 3 which shows that the autocovariance functions for the drain sounds in the *empty* fault had a rapid oscillation with a frequency of 600 *Hz* that was not well modelled because it was not present in the calibration samples. The explanation for this observation may be that the water flow was faster during draining when there were no clothes present in the drum.

- Cold water fill (Figure 8): The fill cluster (circles) shifted and so did the spin cluster (squares). The reason for this is that the cold water supply pressure was higher than that of the hot water supply so the filling sound was different when cold water was used. The sound of the spin stage may have been different because the equipment was cool or perhaps the clothes were less soft and flexible after a cold wash.

The signatures of the faults are sufficiently distinct to suggest that PCA using autocovariance functions of acoustic emissions can distinguish fault conditions.

Quantitative comparison of spectral and autocovariance PCA: Figure 9 compares hierarchical classification trees using the two-term autocovariance PCA model (left hand panel) and the three-term spectral model (right hand panel) . Each path through the tree ends at an individual sound-set. Distances on the vertical axis represent the distances between individual sound-sets or between clusters of sound-sets. Classification is better when a cluster branches into individual sound-sets low down in the tree because this means the distances between the clusters are large. Therefore classification using autocovariance PCA (Figure 9, left panel) was better than for spectral PCA because the clusters branched low and were distinct.

Classification should also reflect *a-priori* knowledge about the samples. Classification using the autocovariance model was also better by this criterion because it showed the sound-sets from the validation and calibration runs clustering as nearest neighbours and well separated from any of the faulty clusters.

Classification using spectral PCA suggested that the calibration and validation sound-sets were further apart than were some of the calibration sets and the sound-sets from faulty runs. As mentioned earlier, it is known that if 44000 frequency channels had been used in the spectra the results would be the same as those using autocovariance functions. Therefore the classification errors in the right hand plot of figure 9 can be attributed to the loss of information caused by use of an averaged periodogram. The failure of the classification tree for spectral PCA to match *a-priori* knowledge reinforces the recommendation that PCA should use the autocovariance functions when size reduction of a large data matrix is required.

4.3 Fault diagnosis

On-line analysis of SPE: The threshold for SPE was taken in this application to be the maximum SPE observed in the validation run. The reason for this choice was that the available recordings were

not numerous enough to determine the statistical distribution of the SPE for a hypothesis test. The maximum SPE in the autocovariance PCA modelling of the validation samples was 25. Therefore the SPE threshold was set at 25 and alarm given if a run had an SPE exceeding 25. The threshold is dimensionless because the autocovariance functions were normalised.

- Fill stage: Three recordings of the filling sounds out of five from the run with the *cold* fault gave an SPE alarm (Figure 8, lower right panel). Therefore an SPE alarm during filling diagnoses a *cold* fault but the absence of an SPE alarm during filling does not rule out a *cold* fault.
- Wash stage: None of the faulty runs gave SPE alarms during the wash stage.
- Drain stage: All five recordings of the draining sounds from the run with the *empty* fault gave SPE alarms because the SPE value was much larger than the threshold (figure 7). Therefore SPE alarms during draining are diagnostic of the *empt*y fault.
- Spin stage: No runs gave SPE alarms during the spin stage.

On-line analysis of classification trees: Figure 10 shows a diagnosis run as if on-line. It also shows five additional sound-sets marked "new" which will be discussed later. The top right panel shows a classification tree using only acoustic monitoring of the fill stage, top left used the sounds from the fill and wash stages, lower left used fill and wash and drain, while the lower right used all. These classification trees quantify the visual observations made earlier about movement of clusters with the PCA plots.

- Fill stage: At the fill stage most of the sound-sets were separated by only small amounts. The exception were the sets from the run having the *cold* fault which were clustered together and separated by a distance of almost 0.2 from any other run. Therefore the *cold* fault can be identified during filling. This finding quantifies the shift that was noted in the fill stage in Figure 8 (circle symbols, top left panel).
- Fill and wash: During the wash stage all the clusters became distinct. The validation cluster was closest to the calibration cluster (they branched at a distance of 0.12, see the top right panel of Figure 10). All the sound-sets from the faulty runs were separated from the calibration and validation clusters by a distance of 0.16 or more. Therefore after the fill and wash stages the presence of a fault can be detected.

  The faults cannot, however, be diagnosed reliably because the separation between the *foreign* fault and the *empty* fault was less than that between the calibration and validation runs. It is therefore not feasible to distinguish *empty* faults from *foreign* faults during the wash stage because the separation could be explained by natural variation between runs. The reason for

variability in the wash stage is because each load of laundry had a different composition and made different sounds as it was agitated in the drum.

- Fill, wash and drain: The lower left panel in Figure 10 shows the calibration and validation clusters separated from each other by 0.13 and separated from all other runs by about 0.2. The *foreign* and *empty* faults became distinct during the draining stage, separated from each other by a distance of 0.17. The cluster of sound-sets from the *cold* run, which was diagnosed during filling, was still distinct. The *empty* run can also be diagnosed by the SPE value during draining.

It is concluded that all the faults would be detected once the wash stage had been monitored, and diagnosed once the drain stage had been monitored. Table 2 gives a summary of the signatures for each of the faults.

stages	classification tree	SPE alarm	diagnosis
fill	*cold* cluster isolated	*cold* fault	*cold* fault
fill and wash	presence of fault detected	none	none
fill wash and drain	*foreign* and *empty* clusters isolated	*empty* fault	*foreign* and *empty* faults

Table 2: On-line fault diagnosis using classification tree and SPE signatures

Diagnosis of an unseen run: Figure 10 shows an on-line diagnosis of a new run with a fault where the nature of the fault had not been disclosed to the engineer running the analysis. The five sound-sets from this run were marked with the white dots labelled as "new".

After monitoring of the fill and wash it was clear that the new run had a fault, but it was not possible to tell if it was an *empty* fault or a *foreign* fault because its sound-sets were separated by almost the same distance from both those clusters.

The new run was correctly classified as an *empty* fault once the drain stage was monitored. The lower left panel in Figure 10 shows the new samples were closest to the *empty* runs. Indeed, one of the new sound-sets was closer to the previously-recorded *empty* sets than it was to the other new sound-sets. An inspection of the SPE confirmed the diagnosis. The SPE values during draining were significantly above 25 and the fault was therefore diagnosed as an *empty* fault because the SPE was above the alarm threshold.

The lower right panel in Figure 10 shows the classification when the spin stage was also included. This classification tree was almost the same as that shown in the left hand panel of Figure 9 but some of the distances at which the clusters separated were smaller. The reason for that is because

the new sound samples and the *empty* samples joined in Figure 10 to form a super-cluster. Distances between this super-cluster and other clusters would therefore generally be smaller than in Figure 9 because the super-cluster was bigger than the *empty* cluster alone.

5. Conclusions

Spectral principal component analysis (PCA) is routinely applied in cases where the primary measurement is itself a spectrum, for instance in near infra red spectroscopy. The contribution of this paper has been to demonstrate an application of PCA to the spectra of time-domain signals and also to autocovariance functions derived from the spectra. Benefits included the removal of unwanted phase information in dynamic systems and the ease of removal of interference.

The concepts were illustrated using acoustic emissions from a washing machine running normally or with faults. PCA showed clear clustering of the autocovariance functions of acoustic emissions from the fill, wash, drain and spin stages. Spectral PCA also showed some clustering, but needed one additional term in the principal components model. The reason for the inferior performance of spectral PCA compared to PCA using autocovariance functions was that averaged periodograms with 1024 channels were used for the spectra to reduce the size of the data matrix whereas the full two-sided power spectrum with 44000 frequency channels was used in the calculation of the first 1024 lags in the autocovariance functions. This finding shows that PCA using autocovariance functions better retains the relevant information if the size of the data matrix needs to be reduced. It was noted in another application when no size reduction was required that the performance of spectral PCA was identical to that of autocovariance PCA.

Sound samples from three faulty runs were projected onto a PCA calibration model developed from sound samples recorded during a normal run of the washing machine. Squared prediction errors (SPE) were calculated. SPE indicates the size of the error when a PCA model with a reduced number of principal components attempts to reconstruct the spectra or autocovariance functions.

Shifts of the fill, wash, drain and spin clusters on the PCA weightings plot were also observed. Clustering was quantified using a scaled Euclidian measure of the distance between sets of sound samples from the fill, wash, drain and spin stages. Hierarchical classification trees were presented which together with SPE alarms enabled the detection and diagnosis of all faulty runs of the washing machine. It is therefore concluded that principal component analysis of spectra and autocovariance functions can make a contribution to the acoustic monitoring of machinery condition.

Acknowledgements

The authors would like to thank Dr M.T. Flanagan, Department of Electronic and Electrical Engineering at UCL for suggestions about the programme of acoustic monitoring. Chian Chern Tan was sponsored by the Public Service Commission of Singapore; the contribution of that organisation is gratefully acknowledged.

References

Aldrich, C., and Theron, D.A., 2000, Acoustic estimation of the particle size distributions of sulphide ores in a laboratory ball mill, *J. South African Inst. Mining and Metallurgy*, 100, 243-248.

Belchamber, R.M., and Collins, M.P., 1993, *Method for monitoring acoustic emissions*, European Patent Office, Publication Nº 0 317 322 B1.

Belchamber, R.M., and Collins, M.P., 1999, Sound for process monitoring, *American Lab.(Sept)*, 31, 48-52.

Betteridge, D., Joslin, M.T., and Lilley, T., 1981, Acoustic emissions from chemical reactions, *Anal. Chem*., 53, 1064-1073.

Boyd, J.W.R., and Varley, J., 2001, The uses of passive measurement of acoustic emissions from chemical engineering processes, *Chen. Eng. Sci*, 56, 1749-1767.

Chatfield, C., and Collins, A.J., 1980, *Introduction to multivariate analysis*, Chapman and Hall, London, UK.

Cody, G.D., Bellows, R.J., Goldfarb, D.J., Wolf, H.A., and Storch, G.V., 2000, A novel non-intrusive probe of particle motion and gas generation in the feed injection zone of the feed riser of a fluidized bed catalytic cracking unit, *Powder Technol.*, 110, 12-142.

De Belie, N., De Smeldt, V., and De Baerdemaeker, J., 2000, Principal component analysis of chewing sounds to detect differences in apple crispness, *Postharvest Biol. Technol.*, 18., 109-119.

Dornfield, D., 1992, Application of acoustic emission techniques in manufacturing, *NDT&E*, 25, 259-269.

Holroyd, T.J., and Randall, N., 1993, Use of acoustic emission for machine condition monitoring, *British J. of NDT*., 35, 75-83.

Holroyd, T.J., and Bradshaw, C., 2000, A critical appraisal of monitoring elastic waves as a means of detecting and diagnosing machine faults, *Insight*, 42, 26-28.

Hou, R., Hunt, A., and Williams, R.A., 1999, Acoustic monitoring of pipeline flows: particulate slurries, *Powder Technol.*, 106, 30-36.

Jackson, J.E., and Mudholkar, G.S. (1979). Control procedures for residuals associated with principal component analysis, *Technometrics*, 21, 341-349.

Karstang, T.V., and Henriksen, A., 1992, Infrared spectroscopy and multivariate calibration used in quantitative analysis of additives in high density polyethylene, *Chemometrics and Intell Lab Syst*, 14, 331-339.

Kourti, T., and MacGregor, J.F., 1996, Control of multivariate processes, *J. Quality Control*, 28, 409-428.

Kresta, J.V., MacGregor, J.F., and Marlin, T.E., 1991, Multivariate statistical monitoring of process operating performance, *Can. J. Chem. Engng*., 69, 35-47.

Riley M.R., Rhiel M., Zhou X., Arnold M.A., and Murhammer D.W. 1997. Simultaneous measurement of glucose and glutamine in insect cell culture media by near infrared spectroscopy. *Biotechnology and Bioengineering*. 55: 11-15.

Seasholtz, M.B., 1999, Making money with chemometrics, *Chemometric Intell. Lab. Syst*., 45, 55-64.

Sharif, M.A. and Grosvenor, R.I., 1998, Internal valve leakage detection using an acoustic emission measurement system, *Trans. Institute of Measurement and Control*, 20, 233-242.

Swindelhurst, W., 1973, Acoustic emission – Introduction, *Non-Destructive Testing*, 152-158.

Tabe, H.T., Chow, K.C., Tan, K-J., Zhang, J. and Thornhill, N.F., 1998, Dynamic principal component analysis using integral transforms, *AIChE Annual Meeting*, Miami Beach.

Valle, S., Li, W.H., and Qin, S.J., 1999, Selection of the number of principal components: The variance of the reconstruction error criterion with a comparison to other methods, *Ind. Eng. Chem. Res*, 38, 4389-4401.

Wade, A.P., Sibbald, D.B., Bailey, M.N., Belchamber, R.M., Bittman, S., McLean, J.A., and Wentzell, P.D., 1991, An analytical perspective on acoustic emission, *Anal. Chem*., 63, 497-507.

Welch, P.D., 1967, The use of fast Fourier transforms for the estimation of power spectra. *IEEE Trans Audio & Electroacoustics*, AU-15, 70-73.

Whitaker, M., Baker, G.R., Westrup, J., Goulding, P.A., Rudd, D.R., Belchamber, R.M., and Collins, M.P., 2000, Application of acoustic emission to the monitoring and end point determination of a high shear granulation process, *Int. J. Pharm*., 205, 79-91.

Wise, B.M., and Gallagher, N.B., 1996, The process chemometrics approach to process monitoring and fault detection, *J. Proc. Control*, 6, 329-348.

Wise, B.M., and Ricker, N.L., Veltkamp, D.F., and Kowalski, B.R. (1990). A theoretical basis for the use of principal components models for monitoring multivariate processes, *Proc. Control. Qual*., 1, 41-51.

Wold, S., Esbensen, K., and Geladi, P., 1987, Principal Component Analysis, *Chemometrics and Intell Lab Syst*, 2, 37-52.

Wood, B.R.A., and Harris, R.W., 2000, Structural integrity and remnant life evaluation of pressure equipment from acoustic emission monitoring, *Int. J. Pressure Vessels Pip.*, 77, 125-132.

Wu, H.D., Siegel, M., Khosla, P., 1999, Vehicle sound signature recognition by frequency vector principal component analysis, *IEEE Trans. Instr. Meas*., 48 1005-1009.

Yeung, K. S. Y., Hoare, M., Thornhill, N.F., Williams, T., Vaghjiani, J.D., 1999, Near infra red spectroscopy for bioprocess monitoring and control, *Biotechnology and Bioengineering*., 63, 684-693.

	PC1	PC2	PC3	PC4
autocovariance PCA	82%	9.2%	(4.5%)	(2.1%)
spectral PCA	83%	5.9%	5.3%	(2.6%)
time domain PCA	10%	9.0%	7.0%	5.9%

Table 1. Variation in the ***X*** matrix captured by the first four principal components

stages	classification tree	SPE alarm	diagnosis
fill	*cold* cluster isolated	*cold* fault	*cold* fault
fill and wash	presence of fault detected	none	none
fill wash and drain	*foreign* and *empty* clusters isolated	*empty* fault	*foreign* and *empty* faults

Table 2: On-line fault diagnosis using classification tree and SPE signatures

FIGURE CAPTIONS

Figure 1. Acoustic signals from normal operation. The trends are scaled to unit standard deviation.

Figure 2. Acoustic power spectra from normal operation and the *empty* fault condition. The spectra of ambient sounds are also shown. Each panel contains five spectra normalised to unit power plotted with linear *y*-axis scaling.

Figure 3. Autocovariance functions from normal operation and the *empty* fault condition. Each panel contains five autocovariance plots with a *y*-axis scale of -1 to $+1$.

Figure 4. Weightings and SPE for normal operation: (ℓ) fill; (←) wash; (_) drain; (□) spin. Black symbols are the calibration samples, white symbols are validation samples from normal operation.

Figure 5. Three dimensional weightings plot for normal operation.

Figure 6. Weightings and SPE for the *foreign object* fault.

Figure 7. Weightings and SPE for the *empty* fault.

Figure 8. Weightings and SPE for the *cold* fault

Figure 9. Hierarchical classification trees for autocovariance and spectral PCA.

Figure 10. A demonstration of on-line fault diagnosis.

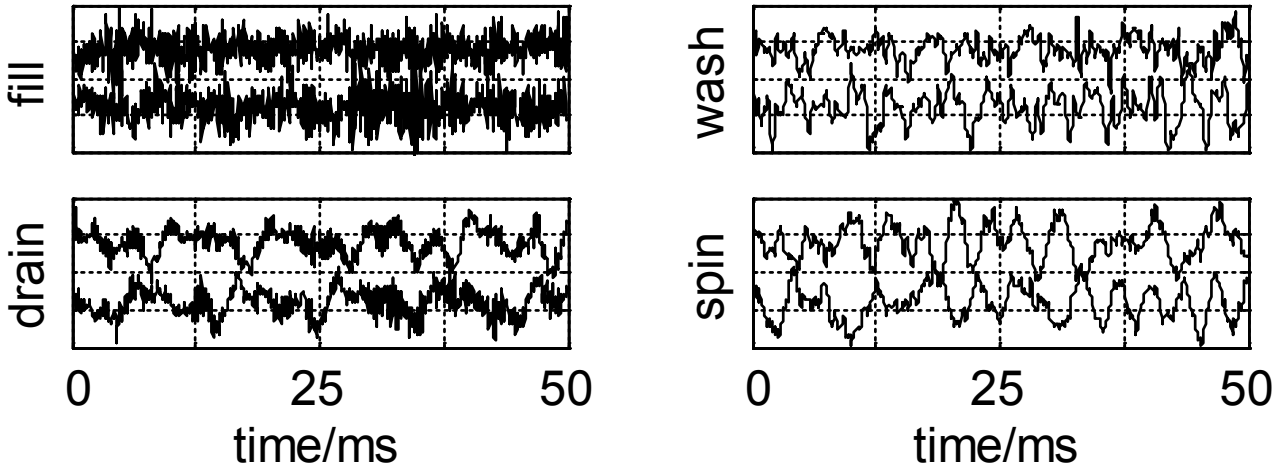

Figure 1. Acoustic signals from normal operation. The trends are scaled to unit standard deviation.

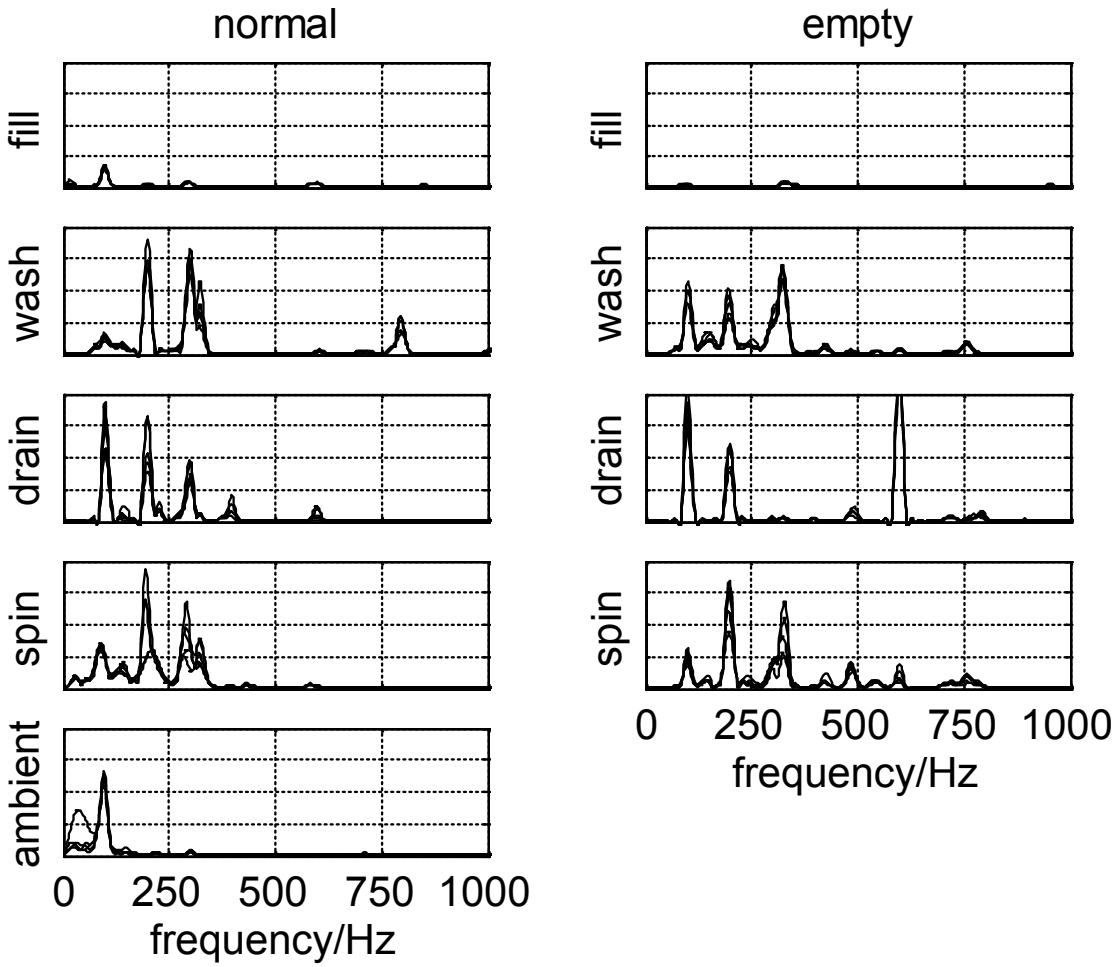

Figure 2. Acoustic power spectra from normal operation and the *empty* fault condition. The spectra of ambient sounds are also shown. Each panel contains five spectra normalised to unit power plotted with linear *y*-axis scaling.

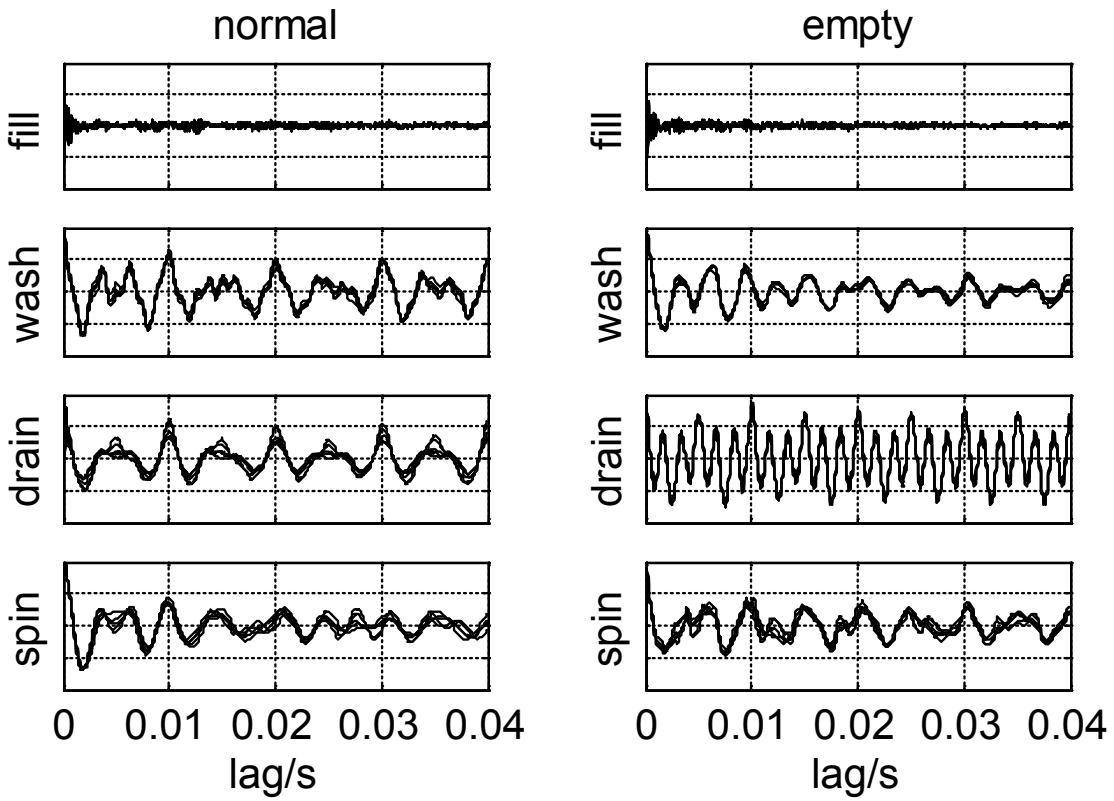

Figure 3. Autocovariance functions from normal operation and the *empty* fault condition. Each panel contains five autocovariance plots with a y-axis scale -1 to $+1$.

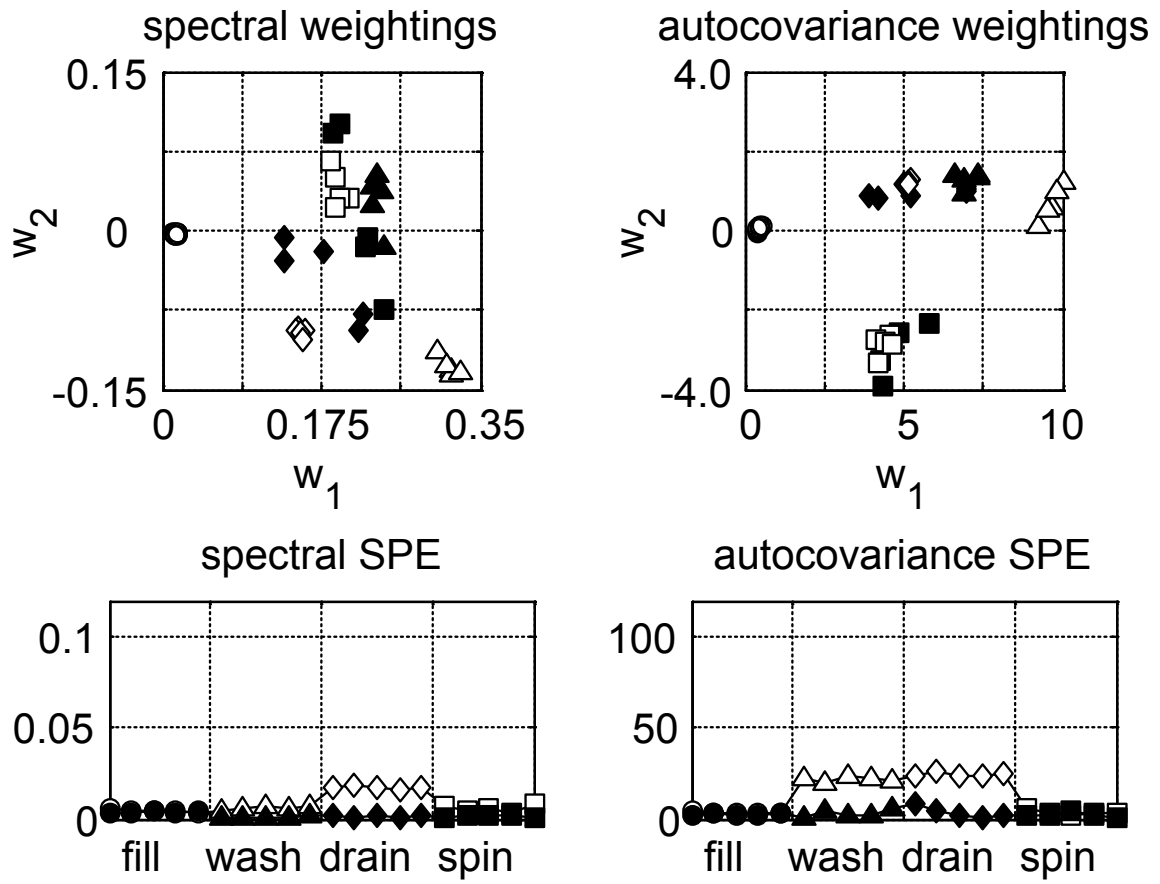

Figure 4. Weightings and SPE for normal operation: (ℓ) fill; (⌙) wash; (_) drain; (◆) spin. Black symbols are the calibration samples, white symbols are validation samples from normal operation.

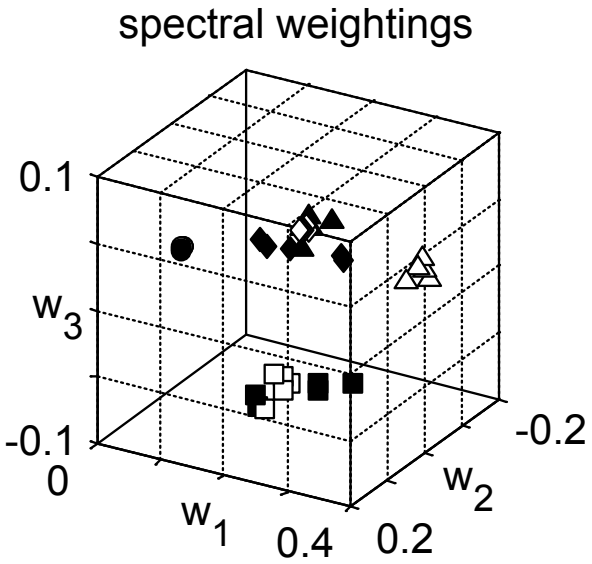

Figure 5. Three dimensional weightings plot for normal operation.

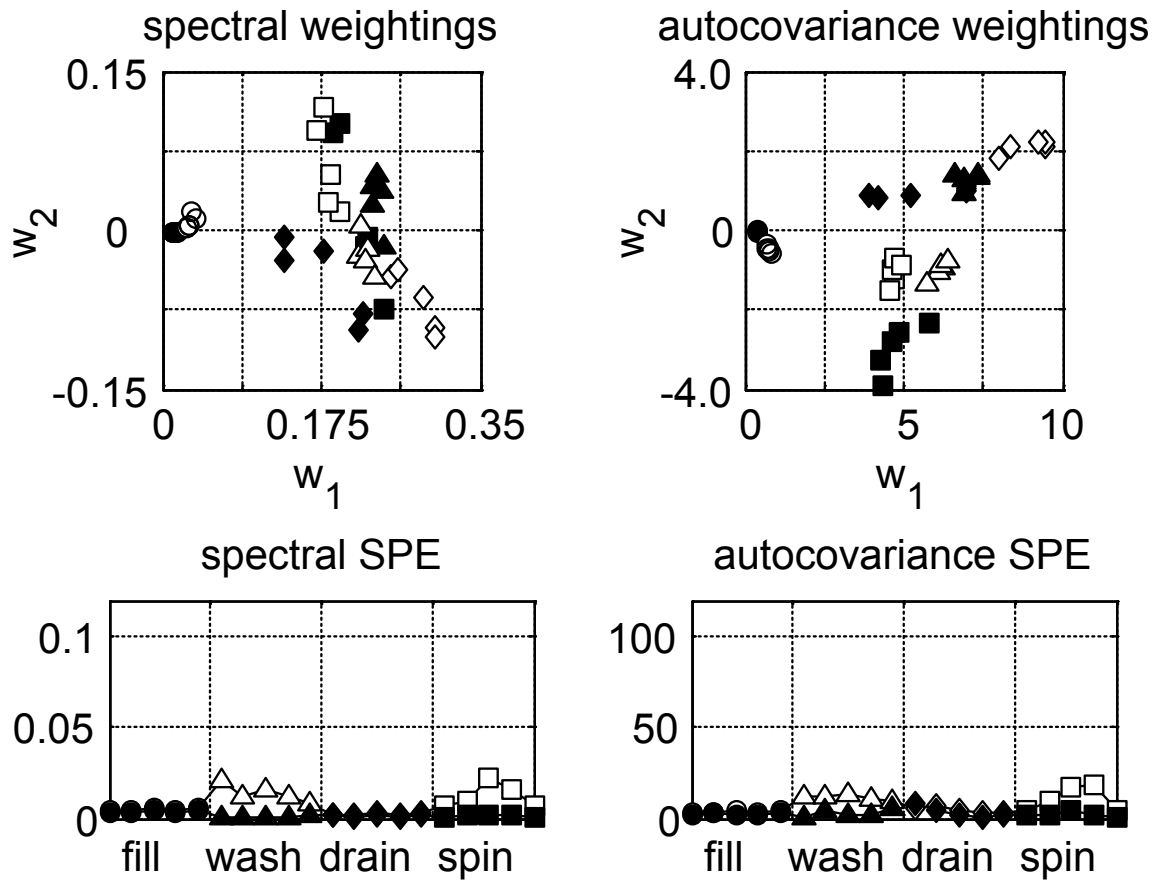

Figure 6. Weightings and SPE for the *foreign object* fault.

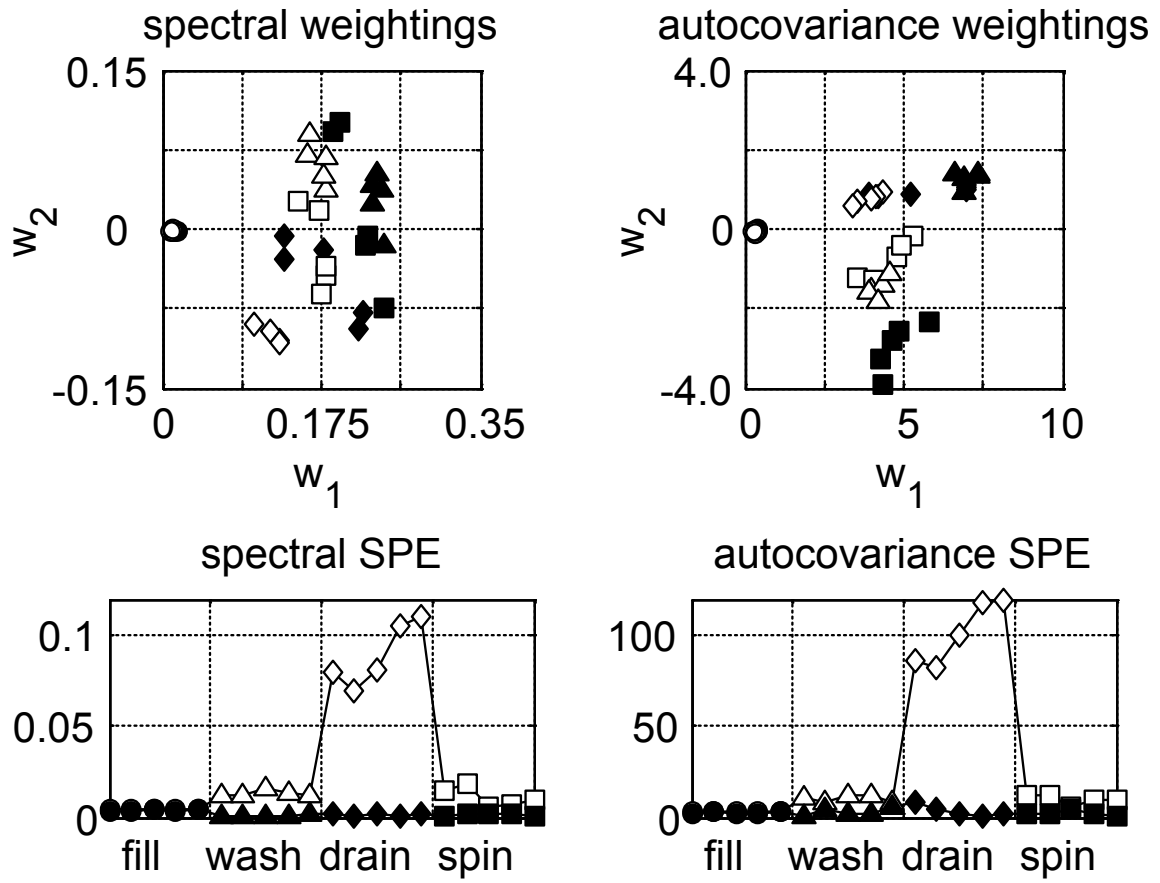

Figure 7. Weightings and SPE for the *empty* fault.

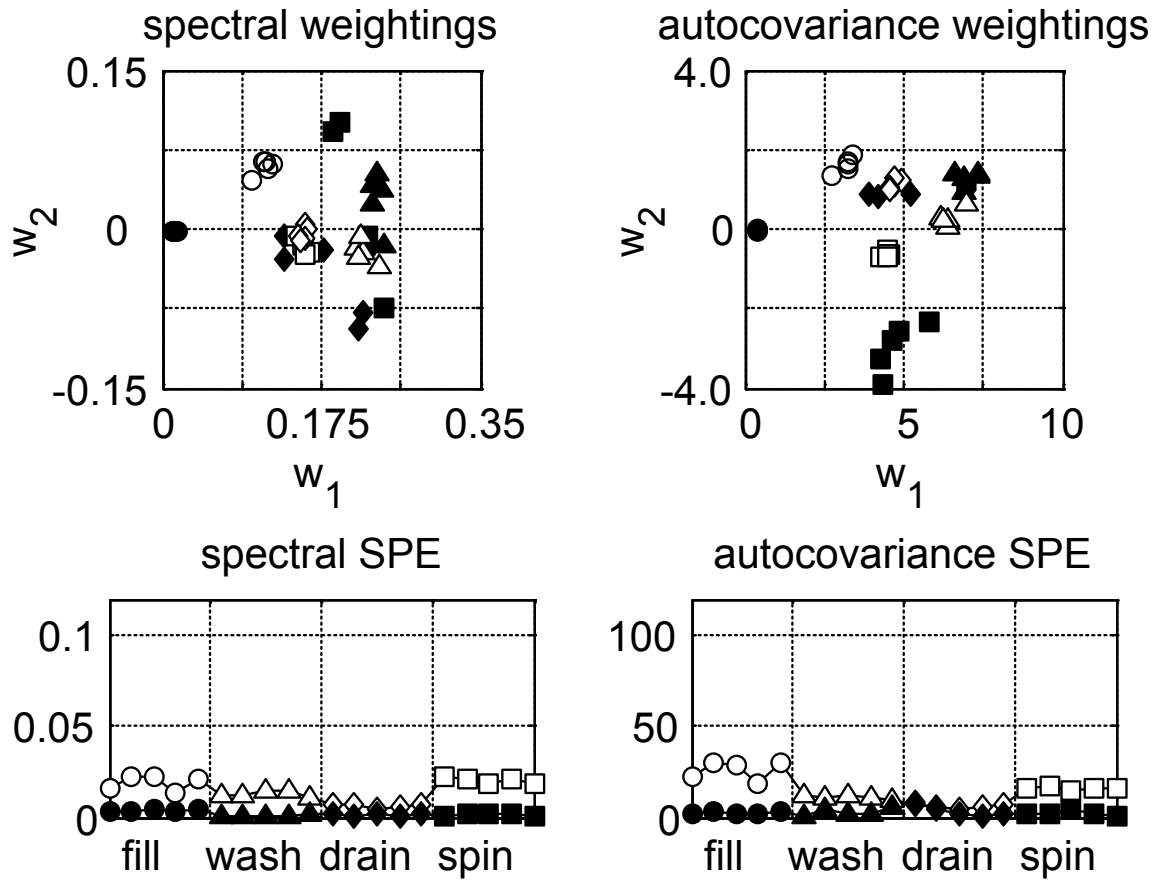

Figure 8. Weightings and SPE for the *cold* fault.

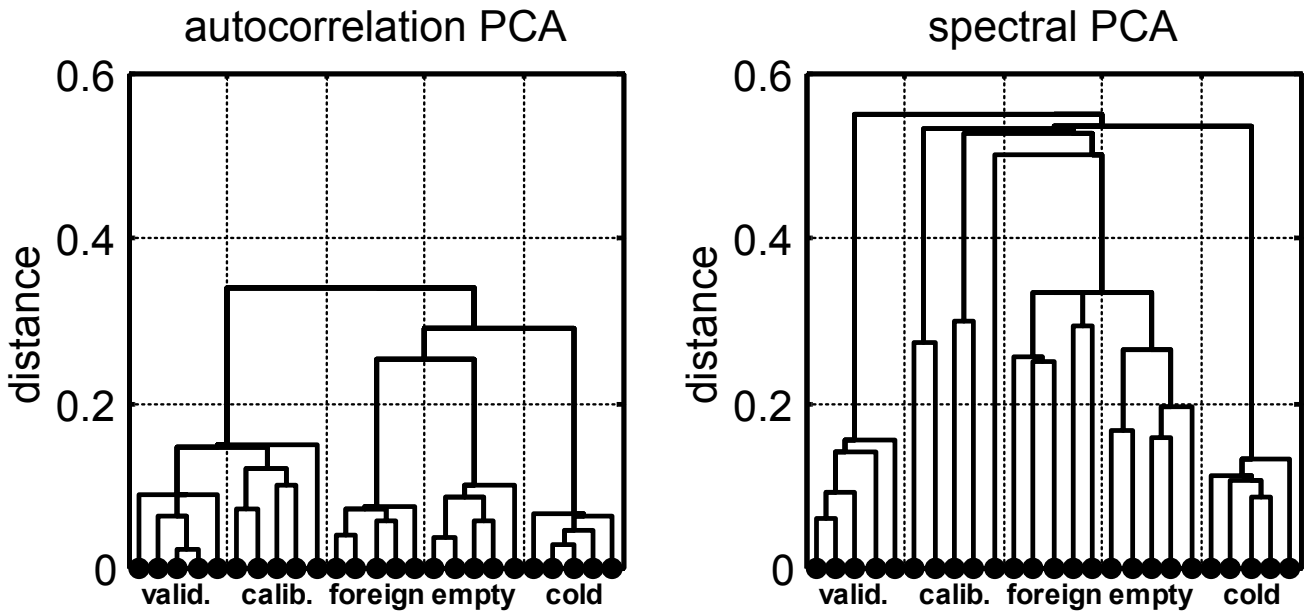

Figure 9. Hierarchical classification trees for autocovariance and spectral PCA.

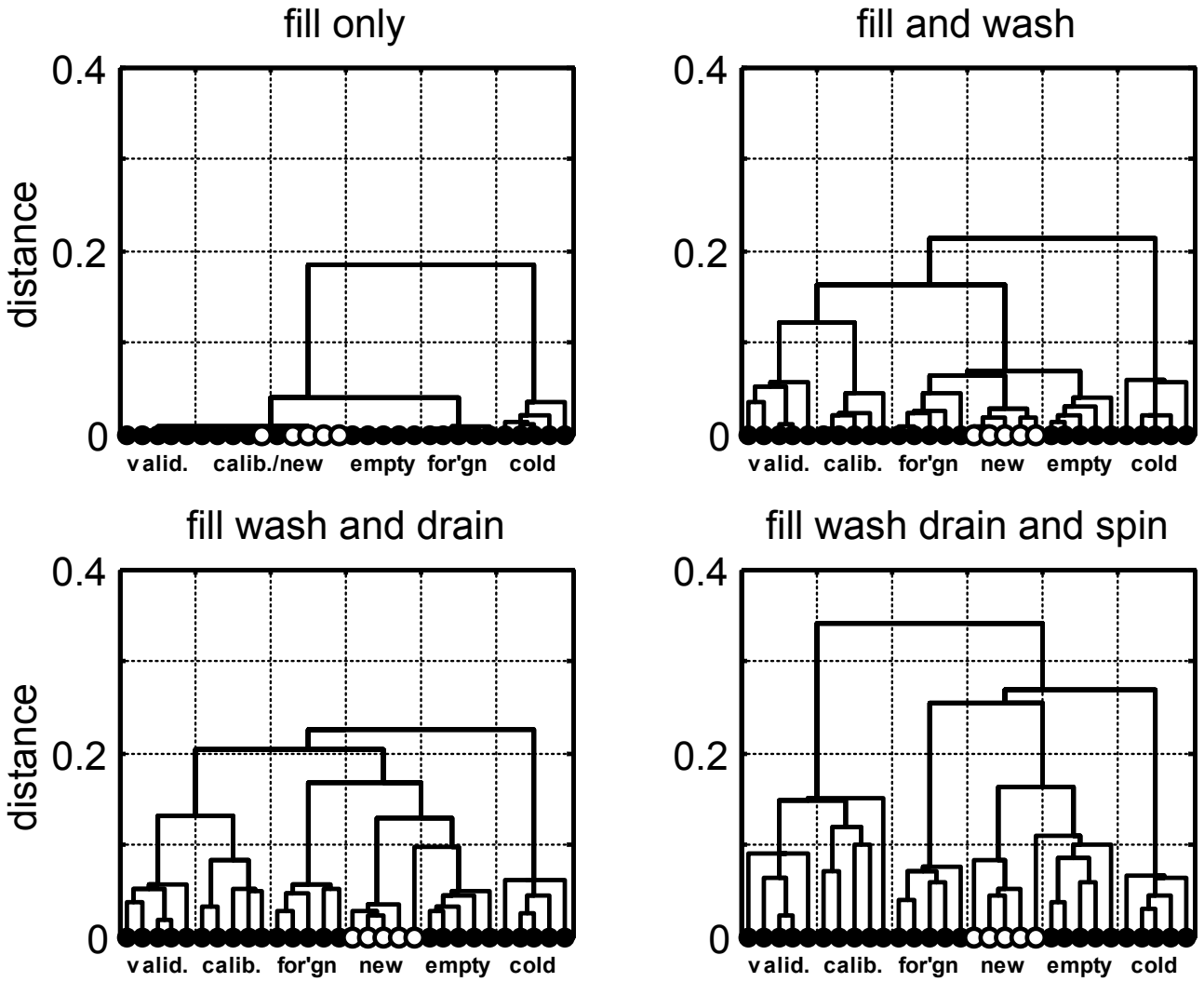

Figure 10. A demonstration of on-line fault diagnosis.