

PCA of Wavelet Transformed Process Data for Monitoring (Hoo)

K. A. Kosanovich*
Department of Chemical Engineering
University of South Carolina
Columbia, SC 29028
kosanoka@sun.che.sc.edu

M. J. Piovoso
Great Valley Campus
Pennsylvania State University
Malvern, PA 19355
mjc@magpage.com

re-submitted to Intelligent Data Analysis Journal, revision 1

November 1, 1996

Abstract

Producing a uniform product is important for several reasons such as maintenance of a competitive position, reduction in the number of shutdowns and startups, and elimination of the sources of variability. Multivariate statistical methods can assist in the identification of process correlations and the development of process monitoring models. This work extends these concepts by demonstrating that the correlations and resulting monitoring models can be improved greatly with the addition of pre-filtering the time signals using a median filter and time-scale decomposition using a multi-resolution wavelet function. After the data are filtered and decomposed, the multivariate statistical method of principal component analysis (PCA) is used to develop a process monitoring model. Data taken from a difficult to operate industrial process are used to demonstrate these ideas.

Keywords: Haar wavelet, FIR/Median hybrid filter, Principal Component Analysis, process monitoring.

1 Introduction

The terms data analysis and process monitoring, as used in the context of process applications, collectively refer to the interpretation and evaluation of sampled process measurements. Data analysis as used in this work is intended to describe how data are manipulated and used together with fundamental understandings to infer the state of a physical process. Monitoring, on the other hand, refers to the classification of the data based upon a calibration model of expected behavior so that abnormal situations can be detected and fault modes isolated. Figure 1 is a simplified view of the on-line process monitoring activity [1].

Any action taken on a process operation generally relies on a description of the state of the operation or events that are occurring. Although there may be hundreds of measurements in a typical chemical process, there are relatively few events generating this information. Timely and correct interpretation of data will lead to improved quality, safer operations, and waste reduction [2, 3].

Unfortunately, much of the process data are of poor quality. Instrument failure, poorly or uncalibrated instrumentation and high noise levels, all contribute to data problems. Without proper pre-treatment, the necessary interpretation is difficult if not impossible. Gross data must be eliminated or modified and noise levels reduced. In many cases, critical information occurs

*author to whom all correspondence should be addressed

over short time duration and hence is difficult to detect. Rioul and Vetterli have described how wavelets can be used to pre-process data in order to better locate and identify significant events [4]. Combining this type of data pre-processing with multivariate statistics holds great promise for generating useful insights into the problem of process monitoring, data analysis, and data interpretation.

Multivariate techniques can be used to identify process variability and to develop monitoring models and multivariate SPC charts for on-line process monitoring and control. Kresta *et al.* used PCA and Projections to Latent Structures (PLS) to analyze a large number of highly correlated variables that defines a continuous chemical process [5]. By comparing new observations with a model that describes *normal* variability, simple control charts can be generated to detect data inconsistencies and processing problems.

Piovosio *et al.* developed an on-line monitoring model and a control strategy based on a PLS/PCA model and implemented it on an industrial, continuous chemical process [6]. They demonstrated the effectiveness of the on-line monitoring model to detect process upsets.

This paper is organized as follows. We review briefly the methods that are employed to pre-process the data and to develop a process monitoring model. Due to the proprietary nature of the industrial example, only a cursory explanation of the industrial process is provided. The data used, however, are real; the results of applying the methods are presented and discussed for process monitoring. Lastly, we summarize our findings.

2 Methods

2.1 Median/Hybrid Filter

Routine process data or on-line sensor data often exhibit significant variability, systematic biases, and "bad" or missing data points. These anomalies can be severe impediments to the development of process models and process understanding, and to design of high performance control system. To extract useful information from these data, a robust method is needed to pre-process the data to eliminate the "noise" prior to use in model development or control efforts [7].

Linear time-invariant filters have been used successfully to filter data,

$$y_n = \sum_k w_k x_{n-k}$$

where y_n is the output sequence, x_n is the input sequence and w_k are the filter impulse response or weighting coefficients that define the filter characteristics. For example, a low pass filter can be designed to allow low frequency components to pass with little or no attenuation and to attenuate high frequency (noise) components. Filters such as these work well when the noise is Gaussian but have undesirable characteristics when dealing with noise distributions that have long tails.

It has been shown that when the input sequence is a constant with independent, zero-mean, exponentially distributed noise, the optimum estimator of the constant value is the median filter, where the median of a sequence x_n is the middle element when the sequence is arranged in either ascending or descending order [8].

The median filter also has other advantages when dealing with process data which are often characterized by data that have jump discontinuities that result from sudden changes in process conditions. The median filter is able to preserve the time and magnitude of the jump, in contrast to linear filters that will distort discontinuous transitions.

The particular median filter, a finite impulse response median hybrid or FMH filter, that is used in this work is described by Heinonen and Neuvo [9]. The main advantage of this filter over other median filters is the reduced computational complexity, that is, it involves the determination

of the median of three elements rather than a sort of $(2K + 1)$ elements, where K is the window half-width.

The FMH filter functions as follows. Select a window half-width of size $K > 0$ but smaller than half the length of the shortest expected region of constant value in the data. This is referred to as a steady-state region. Large values of K will result in the distortion of a signal whose constant region is of a duration shorter than K , however, a large K does contribute to improving the noise suppression capabilities. Ideally, K should be chosen as large as possible without losing steady-state information.

Given a sequence $\{x_1, \dots, x_N\}$, the $(K + 1)^{st}$ FMH filter output is based on the median of the three numbers: \hat{x}_{K+1} , the average of K past samples

$$\hat{x}_{K+1} = 1/K \sum_{k=1}^K x_k$$

\tilde{x}_{K+1} , the average of K future samples,

$$\tilde{x}_{K+1} = 1/K \sum_{k=1}^K x_{n-k}$$

and the $(K + 1)^{st}$ data point, x_{K+1} . Note that \hat{x}_{K+1} is causal¹ whereas, \tilde{x}_{K+1} is not.

Mathematically, $(K + 1)^{st}$ FMH filter output is given as

$$x_{FMH}(K + 1) = \text{median}\{\hat{x}_{K+1}, \tilde{x}_{K+1}, x_{K+1}\}$$

Finally, the filter can be applied iteratively, thereby enhancing noise reduction [8, 9]. Figure 2 illustrates the performance of the FMH filter on a signal that is corrupted with Gaussian noise. Using a window half-width of 9 and 50 consecutive passes of the filter, the noise is essentially filtered from the signal while preserving the steady-state regions. Figure 2

In the process example, the FMH filter is applied using a window half-width of 10 (approximately a 50 minute time constant) with 20 iterative passes. More repeated passes did not improve the signal, and a smaller window size caused noticeable edge effects.

For comparison a lowpass filter is used to filter the same noisy signal (see top right panel in Figure 2). The lowpass filter is a fourth order elliptical filter with a passband edge of 0.0125 normalized hertz. This is a very narrow filter that eliminates the steps reducing the signal to a ramp-like function. The left panel in Figure 3 shows the result of filtering the noisy steps. Increasing the bandwidth to 0.025 normalized hertz permits more of the high frequency content to be passed, however, the filtered signal (see right panel in Figure 3) does not dampen at fast enough rate thereby producing an overshoot that only qualitatively resembles a step. Fig 2

2.2 Wavelets

Wavelets can be viewed as an extension to Fourier analysis that are well-suited for characterizing non-stationary signals, that is signals whose spectral character change with time. Such signals are not well represented in time and frequency by the Fourier Transform methods. For example, a signal such as a Dirac delta function is perfectly localized in time, but has no resolution in frequency because its Fourier transform is a constant. In contrast, a signal such as an ideal sinusoid has perfect frequency resolution since all its energy is located at only one frequency, but has no localization in time. Wavelets offer a technique to localize events in both time and frequency and they can be applied to continuous and discrete-time problems.

¹output does not depend on future inputs

Wavelet theory provides a unified framework for a number of techniques which have developed independently for various signal processing applications, filtering, data compression and image analysis, to name a few [4, 10]. The wavelet transform involves representing general functions in terms of simple, fixed building blocks at different scales and positions [4, 10, 11]. These blocks are actually a family of wavelet functions generated from a prototype function, called a “mother” wavelet, by translation and scaling operations. That is, the signal is mapped to a time-scale plane that is analogous to the time-frequency plane used in the Short-time Fourier Transform.

Multi-resolution analysis provides a formal approach to constructing the wavelet basis. The idea of multi-resolution analysis is to write a function as a limit of successive approximations, each of which is a smoother version of the function. Figure 4 is a simple illustration of the concept of wavelets as a multi-resolution analysis where the subspaces contained within each other are meant to convey the notion of fine to coarse resolution with the smoothness achieved through removal of some level of detail. Sub-space $V_{-1} \subset V_0 \subset V_1 \subset \dots$ and subspace W_0 is the orthogonal compliment of $V_0 \subset V_1$, that is,

$$W_j \oplus V_j = V_{j+1}$$

The W_j s contain the detail information as the granularity of the resolution goes from a finer (larger j) to a coarser (lower j) one. The nested subspaces, V_j , each contain the best approximation at a particular resolution, that is,

$$\lim_{j \rightarrow \infty} V_j = \bigcup_{j=-\infty}^{\infty} V_j$$

and there will be information loss as the resolution gets coarser ($j = \dots, -3, -2, -1, 0$), that is, in the limit of lowest resolution, the signal is approximated by 0

$$\lim_{j \rightarrow -\infty} V_j = \bigcap_{j=-\infty}^{\infty} V_j = \{0\}$$

Wavelets are the basis functions for the W sub-spaces.

There are several types of wavelet transforms: for continuous signals the time and scale parameters are continuous, leading to a continuous wavelet Transform (CWT). If the time and scale parameters are chose to be discrete, this will give rise to a Wavelet Series expansion; finally, if the signal is discrete, a discrete wavelet transform (DWT) results.

The CWT is given as

$$(Wf)(a, b) = 1/\sqrt{|a|} \int f(t) \psi \left(\frac{t-b}{a} \right) dt$$

where a represents the scale, b the translation, and $\psi \left(\frac{t-b}{a} \right)$ the basis function [11, 12]. The reconstructed signal is obtained using the formula

$$f(t) = C_\psi \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (Wf)(a, b) \psi \left(\frac{t-b}{a} \right) \frac{dad b}{a^2}$$

where C_ψ is the admissibility constant and is given by the Euclidean norm of the basis wavelet [11].

As the scale parameter grows, the signal dilates more, and like a map, the image or wavelet transform gives a more global or low frequency view. The translation parameter serves to shift the function along the time axis. A special case is developed by discretization of the time-scale parameters. That is, if $a = 2^{-j}$ and $b = k2^{-j}$, the corresponding wavelets become a function of two integer parameters, j and k . For this case, the wavelets form a dyadic series.

Almost any function can be a prototype function as long as it satisfies certain admissibility conditions [11]. Daubechies introduced a set of orthonormal wavelets, and more recently, Kosanovich *et al.* defined a new family of non-orthogonal wavelets based on the well-known Poisson distribution function [12, 13].

In this work, the simple Haar function, developed around 1910, is used [11],

$$\psi_H(t) = \begin{cases} 1 & 0 \leq t \leq 1/2 \\ -1 & 1/2 < t \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Although not noted for its sharp frequency localization, as compared to the Daubechies and the Battle-Lemarie wavelets, it is particularly well-suited for the process data encountered here (see §3). The Haar wavelet transform (HWT) can resolve the time and amplitude jump discontinuities features common to process data better than other wavelets. In general, the selection of the wavelet that best decompose the data remains a research topic.

Application of the dyadic form of the HWT to mean-centered data is done after the data are pre-filtered by the FMH filter to remove the high frequency (noise) components and gross outliers. The wavelets coefficients are easily calculated for combinations of the parameters j and k .

2.3 PCA & Process Monitoring

Multivariate statistical analysis methods can assist in the identification of process correlations thereby supporting or improving existing process knowledge. Previous researchers have used the multivariate methods such as PCA and Partial Least Squares (PLS) successfully to data analysis, model development, and control variable selection [2, 5, 6]. A tutorial treatment with application of PCA can be found in [14, 15].

Steady-state conditions and linear or mildly nonlinear correlations of the data are assumed when PCA is used. In some cases, suitable transformations to linearize the data prior to PCA analysis can be performed. If the data are correlated, their information content can be captured by a smaller set of variables [2, 5, 14, 15].

For example, a typical process may be instrumented to collect and store hundreds of process measurements. Physico-chemical relationships tell us that there are not hundreds of independent events occurring, therefore the data are correlated. Any technique that can capture the important events based on the variability in the data will provide both a reduction in the data size and a summary of the information contained in the original data set.

PCA decomposes a single, dependent set of data, \mathbf{X} , of dimension $M \times N$ into a transformed space defined by the eigenvectors of the covariance matrix associated with \mathbf{X} . The M rows are usually the sampled process variables at a fixed sampling time, and a column is a uniformly sampled variable. PCA produces a mapping of the data set onto a reduced subspace defined by the span of a chosen subset of eigenvectors, or *loadings* \mathbf{P} , of the variance-covariance matrix of the data. Each new eigenvector captures the maximum amount of variability in the data in an ordered fashion. That is, the first principal component explains the greatest amount of variation, the second the next largest amount after removal of the first, and so on.

Associated with each eigenvector is the corresponding eigenvalue, λ , which relates the amount of variance explained by that eigenvector. In the sense, PCA is a special case of singular value decomposition. The projections onto \mathbf{P} , generate a set of pseudo-measurements or *scores*, \mathbf{T} , that are linearly independent.

The application of PCA usually involves a prior step that mean centers and normalizes the data. Mean centering implies that the average value for each variable is subtracted from the

corresponding measurement. Scaling or normalization is necessary to avoid problems associated with some measurements having large values and others with small ones.

The entire sets of scores and loadings define the process data, and the loadings are the statistical process model. That is,

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{R}$$

where \mathbf{R} are the residuals given by

$$\mathbf{R} = \mathbf{X} (\mathbf{I} - \mathbf{P}_A \mathbf{P}_A^T)$$

and only the first A of N possible principal components are used in the calibration model. The success or failure of PCA monitoring depends to a large extent on the accuracy of the loadings.

A subset of the first few scores, $A < N$, provides information in a lower dimensional space, the *score space*, of the behavior of the process during the period in which the measurements were made. This set of scores and the PCA loadings can be used to determine if the present process operation has changed its behavior relative to the data that were used to define the scores and loadings [7].

In this work, PCA is applied to the Haar wavelet transform of the data, rather than on the time history of the various process measurements. The advantage of this is that the variability defined by the Haar coefficients are more closely associated with process events than when PCA is applied to the time history of the data. For example, small shifts in several process variables may be significant for establishing a process fault or upset. However, if the time duration is small, this contribution is lost when compared to other more significant variations. The HWT concentrates this variability into a few coefficients, thereby enhancing their effect in the Haar wavelet domain as compared to the time domain.

2.4 Statistics

There are several ways of interpreting the PCA results. The ones used here are the Q-statistic, a measure of the model mismatch; the Hotelling T^2 -statistic, a measure of the fit of new observations to the model space; variance plots, a measure of the samples' variability; and score plots, a qualitative representation of the process performance, relative to the calibration model in the model space defined by the calibration model.

The Q-statistic is the sum of squares of the errors between the data and its estimates. The estimate obtained from a reconstruction of the data from a fixed number of loadings and scores. It provides a means of testing whether the process data have shifted outside the normal operating space, that is, a "goodness of fit" of new samples. The Q-statistic for the k^{th} sample is given by:

$$Q_k = \mathbf{x}_k^T (\mathbf{I} - \mathbf{P}_A \mathbf{P}_A^T) \mathbf{x}_k$$

Confidence limits can be computed for the Q-statistic from the χ^2 distribution,

$$Q_\alpha = \Theta_1 \left(\frac{c_\alpha \sqrt{2\Theta_2 h_0^2}}{\Theta_1} + \frac{\Theta_2 h_0 (h_0 - 1)}{\Theta_1^2} + 1 \right)^{1/h_0}$$

$$\Theta_i = \sum_{a=A+1}^N \lambda_a^i \quad i = 1, 2, \dots$$

$$h_0 = 1 - \frac{2\Theta_1 \Theta_3}{3\Theta_2^2}$$

where c_α is the normal deviate corresponding to the upper $(1-\alpha)$ percentile.

The T^2 -statistic measures unusual variability within the calibration model space. That is, if the calibration model data represent process operation at one operating condition, and the process has shifted to a different one, then the T^2 -statistic will show that data at this operating condition cannot be classified with the calibration data. The T^2 -statistic is proportional to the sum of the squares of the scores on each of the A principal components,

$$T^2 = \left[\frac{M(M-A)}{A(M^2-1)} \right] t_a' S^{-1} t_a$$

where S is the estimated covariance matrix of the scores, M represents the total number of samples, and t_a is the a^{th} score vector.

The confidence limits for the T^2 -statistic can be calculated by means of the \mathcal{F} -distribution ,

$$T_{A,M,\alpha}^2 = \left[\frac{A(M-1)}{M-A} \right] \mathcal{F}_{A,M-A,\alpha}$$

Statistical limits on the Q -statistic and T^2 -statistic are computed based on assuming the data are normally distributed in the multivariate sense [15]. In practice, the underlying distribution of the residuals of individual variables can vary substantially from the Gaussian assumption without affecting the results [16]. Thus, the diagnostic limits are useful to establish when a statistically significant shift has occurred. Charts based on these statistics and used in this manner are analogous to conventional SPC charts.

3 Example Process

A process flowsheet of the industrial process is shown in Figure 5. A chemical reaction is carried out in the two parallel reactors that produces the product and several by-products. The components of the stream leaving the reactor are further separated using a flash drum. The vapors are condensed to recover the product and light gases (boiling points less than that of propylene), by-products of the reactions, are vented.

The flash bottoms stream is further cooled before being recycled to the reactors. Both reactors are intended to operate under identical conditions, however, there are circumstances that will cause one to operate more efficiently than the other. Vice versa, there are conditions that will cause one or both to fail.

Indications of failure are difficult to identify due to the large number of monitored variables. In some cases, what was thought to be a failure mode repeats but there is no evidence of failure. The system is well instrumented and is controlled by a distributed control system. In parallel with this effort to provide a process monitoring system, a first principles model is being developed. Both will be used to increase process understanding.

More than 40 process variables are monitored to evaluate the operating conditions and the overall process performance. Figure 6 shows typical variable traces taken at five minute intervals; observe the high frequency noise in the data as well as the qualitative differences in the signals. A calibration model is developed based upon 5 minute samples taken over a period of one month of 24 monitored variables. These 24 are selected based on process understanding.

To develop the calibration model, the data are first filtered using the FMH filter to remove the gross outliers and to locate periods of steady-state operation. The filtered values are then decomposed using the HWT into low and high frequency components to better isolate (localize) events that occur on different time scales. Using the resulting Haar coefficients, PCA is applied.

The process monitoring activity essentially involves the projection of new data onto the loadings of the calibration model. Statistics on the goodness of fit (Q) or on the individual variables (T^2) are computed. The responsibility of how to compensate for new data that are outside of the confidence limits remains at the discretion of the operators.

4 Results

Figure 7 shows the filtered results of the FMH filter applied to a signal after 20 passes with a half window size of 10. Additional filtering is not necessary as it did not improve the signal. Note that the event around day 17 is preserved by the FMH filter.

Since the data contained samples when the process was not operating “normally”, applying PCA as a simple outlier detector to the results of the FMH filter revealed which data could be classified as such. Score plots of the first few principal components can be used to visually detect these outliers. Figure 10 shows the first two principal components plotted against each other. The numbers indicate the samples that appear to be outliers. 128 samples are selected from among those that deviated significantly from the main cluster of points. A plot of the Q-statistic after eight PC is shown in Figure 11; the 95% confidence limit on the Q-statistic is 5.17 and that of the T^2 statistic is 3.84.

Figure 8 shows the result of applying the HWT to FMH filtered data in contrast to Figure 9 where the HWT is applied to the same signal that was not filtered. Clearly, the filtered signal is less noisy and ought to provide a better signal from which to develop a calibration model. Figure 12 shows a score plot of the first two scores of the calibration model; Figure 13 shows the additional six PCs. Eight principal components gave a 90% total explained variance out of a possible 24 indicating a high degree of correlation. Figure 14 depicts the projection of a portion of the test set, labeled by (*) onto the calibration model.

By projecting the Haar wavelet coefficients that pertain to the identified samples in Figure 14 onto the loading vectors, a contribution plot can be obtained which relates the contribution of each variable to that direction of variability.

Contribution plots decompose the scores into their summation operands and graphs them vs the contributing variable. The summation operands are the products of the loadings of variable j and the corresponding value of variable j . A large product associated with a particular variable implies a correspondingly large contribution.

By comparing the contribution plot of a sample taken from the calibration set with one that is outside the confidence limits, differences in the expected variables' magnitude may provide an indication of which variables have exceeded their expected limits and a possible compensation to correct the problem. Figure 14 depicts the projection of a portion of the test set, labeled by (*) onto the calibration model.

An example is shown in Figures 15 and 16. In both figures the right panels are the expected contribution to the first and second principal component, respectively. In the left panels are the contribution from test set samples #1 and #3. #1 is mostly a first principal component effect and similarly #3 is almost all a second principal component effect. Observe the axes size. It appears that variable numbers 21 and 22 are unusually high. In the case of #1, higher than usual levels of cooling occurred, implying a large rise in the temperature of the reactor. Corrective action is necessary to keep the reactor operations within the prescribed safety limits.

5 Summary

This work seeks to extend the notion of developing a process monitoring model that represents expected operations with normal variations. Recent research in this area has shown that the multivariate statistical method of principal component analysis (PCA) can be used to develop such a calibration model. However, the sensitivity of the model to classifying non-normal operations is dependent on the data used.

This work demonstrates that pre-filtering of the data with a finite impulse response median hybrid filter (FMH) and developing the calibration model from a set of coefficients obtained from a Haar wavelet transform (HWT) of the FMH filtered data produces a calibration model with

better (sharper) classification features. This comes about because the FMH filter removes high frequency components in the data while preserving the time and magnitude of steady-state regions that may be distorted by linear filtering. This is an important consideration because these jump discontinuities are common features of continuous chemical process data.

Wavelets in general, provide good time and frequency localization. Since events in a chemical process occur on many time scales, it is appropriate to analyze sensor signals with a technique that can discriminate at different scales. Here, we employ the Haar wavelet because it is well-suited for signals that contain discontinuities.

We develop the calibration model but not by employing PCA on the filtered data, or on the reconstructed data after applying the HWT but on the resulting HWT coefficients themselves. In this fashion, the PCA calibration model is identified in the Haar wavelet domain where the events in the data are more pronounced.

The process monitoring activity is then carried out in the HWT domain. That is, new data are filtered and their HWT coefficients are projected into the HWT domain by using the first A loadings. By employing a combination of statistics and contribution plots, the reasons for deviations from the calibration model may be determined and an appropriate control action implemented.

6 Acknowledgements

The authors would like to thank 3M for supplying the data, A. R. Moser of the DuPont Company for his assistance with the wavelet code, and one of us would like to acknowledge support from ACS-PRF# 29974-G9 grant.

REFERENCES CITED

- [1] Kramer, M. A. and R. S. Mah, Model-Based Monitoring, FOCAP'94, Crested Butte, CO (1994).
- [2] Kosanovich, K. A., K. S. Dahl, and M. J. Piovoso, Improved Process Understanding Using Multi-way Principal Component Analysis, *I&ECR*, **35**, 138 (1996).
- [3] Davis, J. F., B. R. Bakshi, K. A. Kosanovich, and M. J. Piovoso, Process Monitoring, Data Analysis, and Data Interpretation, ISPE'95, Snowmass, CO (1995).
- [4] Rioul, O. and M. Vetterli, Wavelets and Signal Processing, *IEEE Sig. Proc. Mag.*, **8**, 14 (1991).
- [5] Kresta, J. V., J. F. MacGregor, and T. E. Marlin Multivariate Statistical Monitoring of Process Operating Performance, *Can. J. Chem. Eng.*, **69**, 35 (1990).
- [6] Piovoso, M. J., K. A. Kosanovich, and J. P. Yuk, Process Data Chemometrics, *IEEE Transactions on Instrumentation and Measurement*, **41**, 262 (1992b).
- [7] Piovoso, M. J., K. A. Kosanovich, and R. K. Pearson, Monitoring Process Performance in Real-Time, *Proc. of the Amer. Contr. Conf.*, Chicago, IL, **3**, 2359 (1992).
- [8] Gallagher, N. C. and G. L. Wise, A Theoretical Analysis of the Properties of Median Filters, *IEEE Trans. Acoustics, Speech, & Signal Proc.*, **29**, 1136 (1981).
- [9] Heinonen, P. and Y. Neuvo, FIR-Median Hybrid Filters, *IEEE Trans. Acoustics, Speech, & Signal Proc.*, **35**, 832 (1987).

- [10] Motard, R. L. and B. Joseph, *Wavelet Applications in Chemical Engineering*, Kluwer Academic Publishers, Boston (1994).
- [11] Chui, C. K. *An Introduction to Wavelets*, Academic Press, New York, (1992).
- [12] Daubechies, I., Orthonormal Bases of Compactly Supported Wavelets, *Comm. in Pure & Appl. math.*, 41, 909 (1988).
- [13] Kosanovich, K. A., A. R. Moser, and M. J. Piovoso, *J. Process Control*, 4 (1995).
- [14] Martens, H. and T. Næs, *Multivariate Calibration*, John Wiley & Sons, New York (1989).
- [15] Jackson, J. E., *A User's Guide to Principal Component Analysis*, John Wiley & Sons, New York (1992).
- [16] Wise, B. M., N. L. Ricker, D. F. Veltkamp, and B. R. Kowalski, A Theoretical Basis for the Use of Principal Models for Monitoring Multivariate Processes, *Process Control & Quality*, 1, 41 (1990).

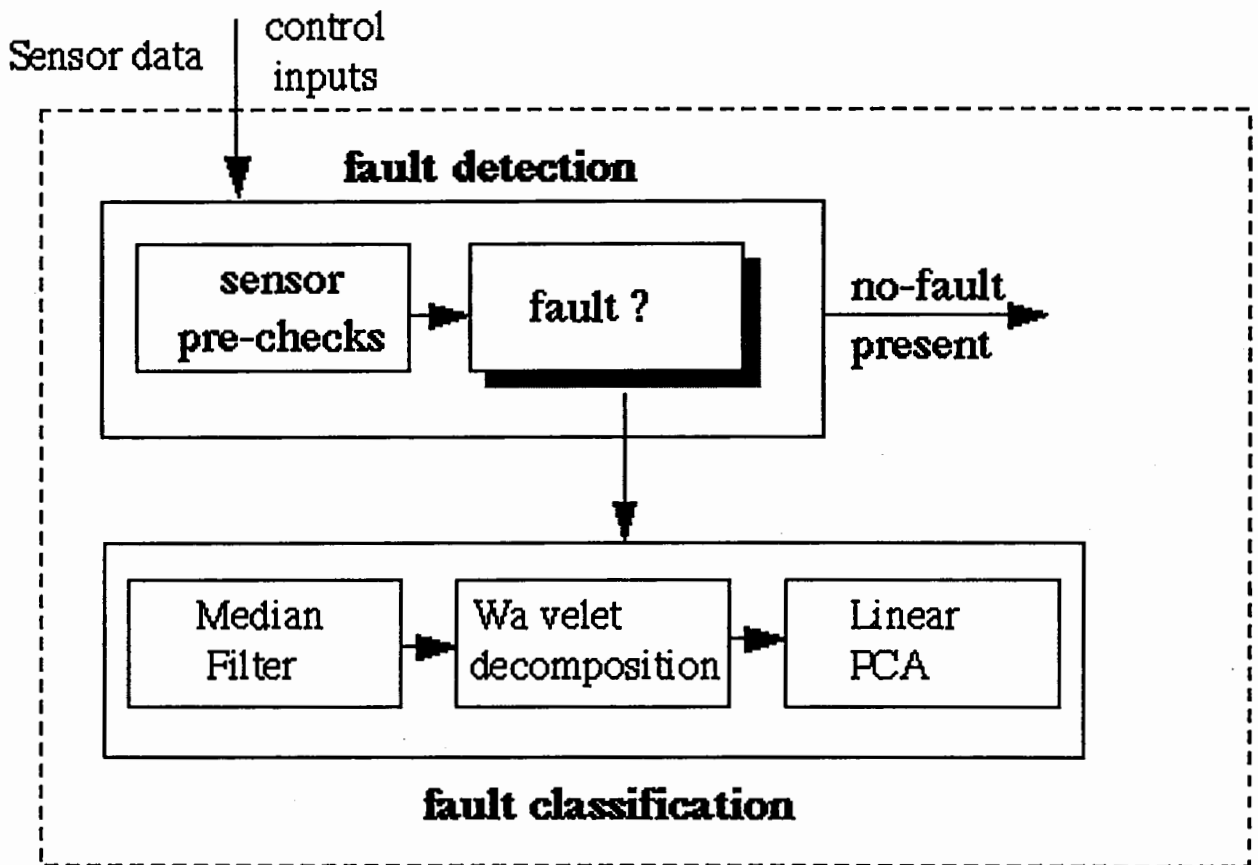


Figure 1: On-line process monitoring activity.

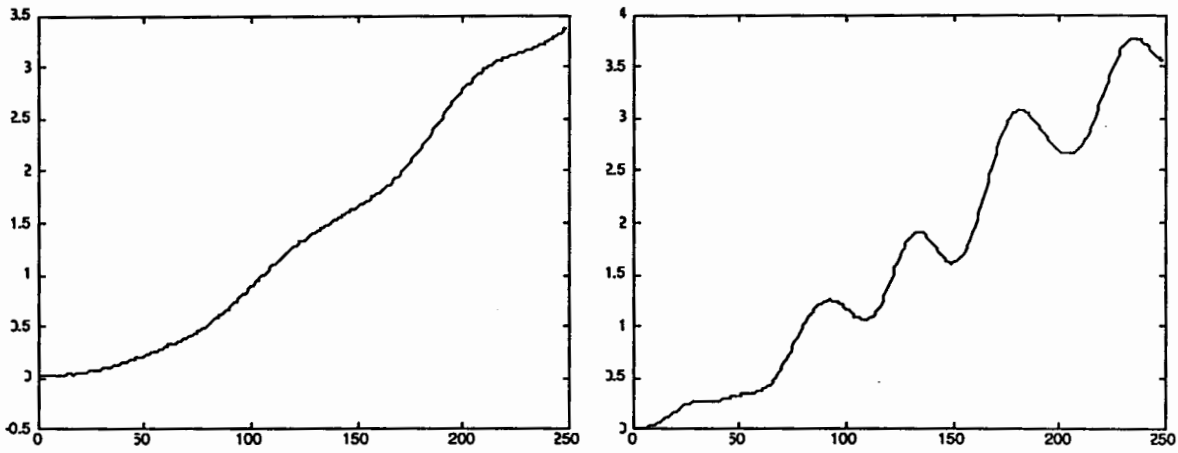


Figure 3: Results of applying a lowpass fourth order elliptical filter on the signal shown in the top right panel of Figure 2. Left panel: passband with 0.0125 normalized Hertz, right panel: passband with 0.025 normalized Hertz

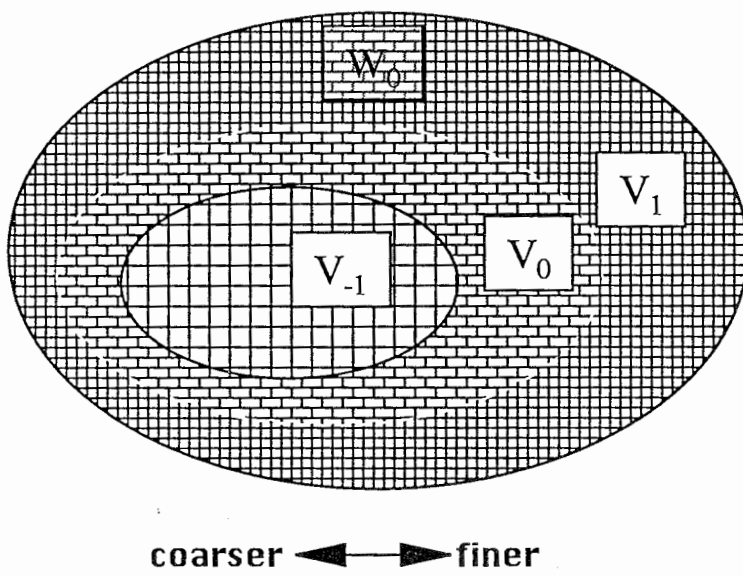
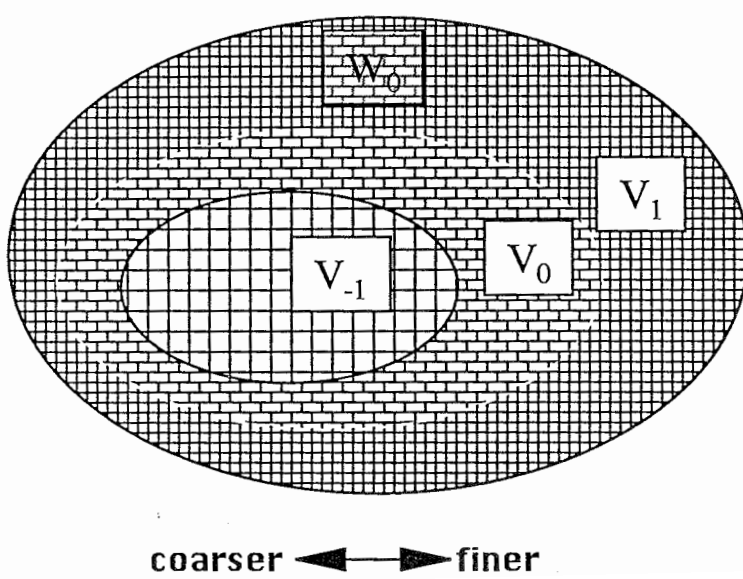


Figure 4: Multi-resolution illustration of wavelet decomposition. The W s contain the detail information and the V s are the filtered versions as the resolution moves from a coarse to a fine granularity



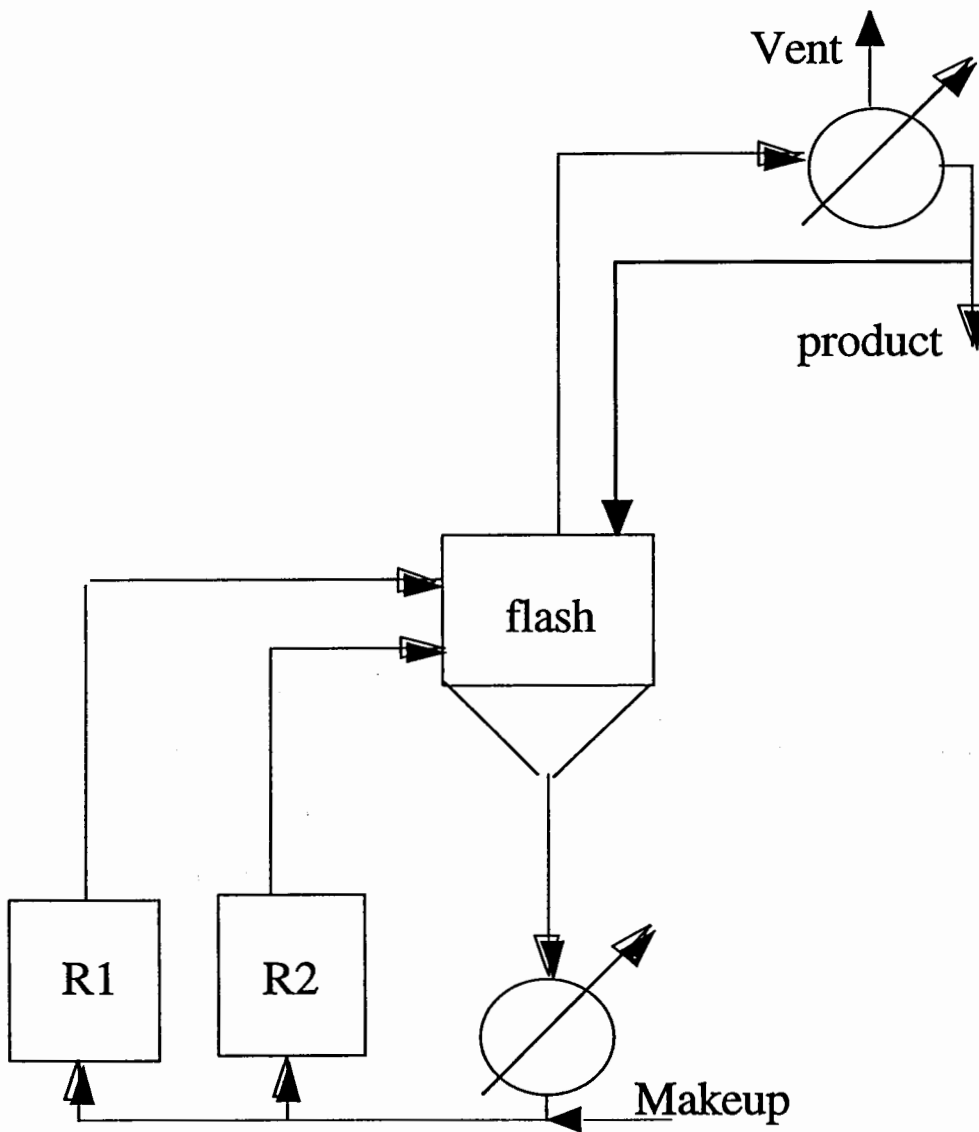


Figure 5: Simplified Process Diagram

pic 5

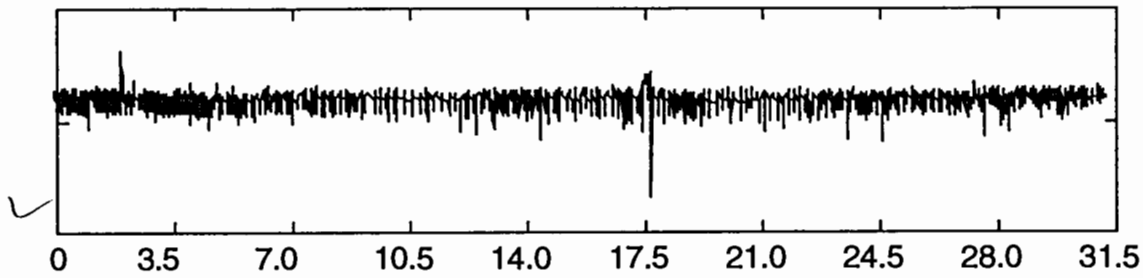


Figure 6: Typical process signal, 5 minute samples.

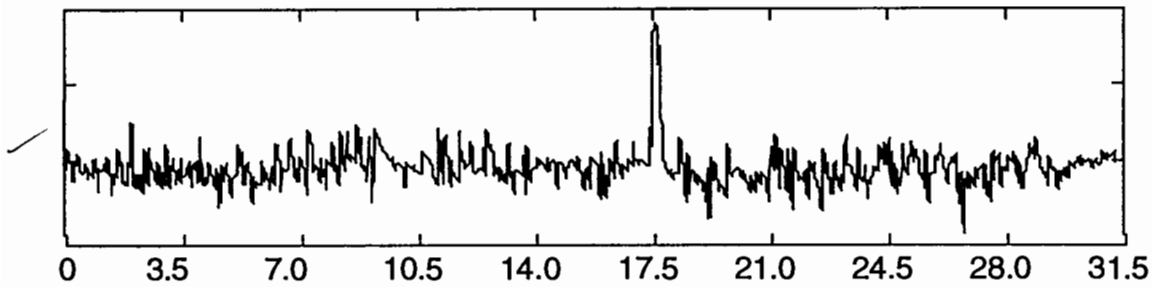


Figure 7: FMH Filtered signal, window half width of 10 and 20 iterative passes.

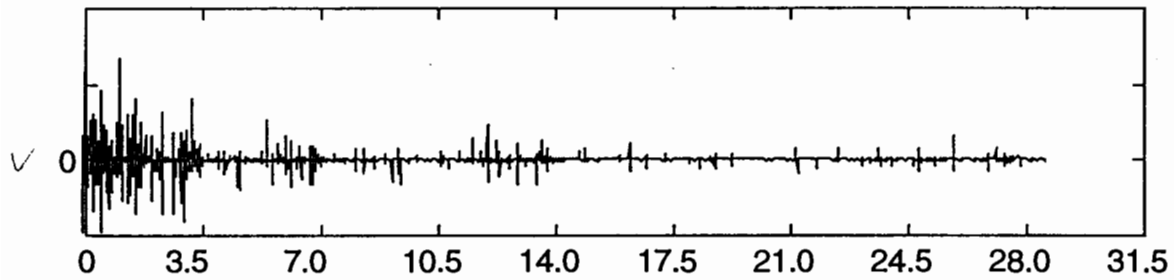


Figure 8: Haar wavelet transform applied to the FMH filtered signal.

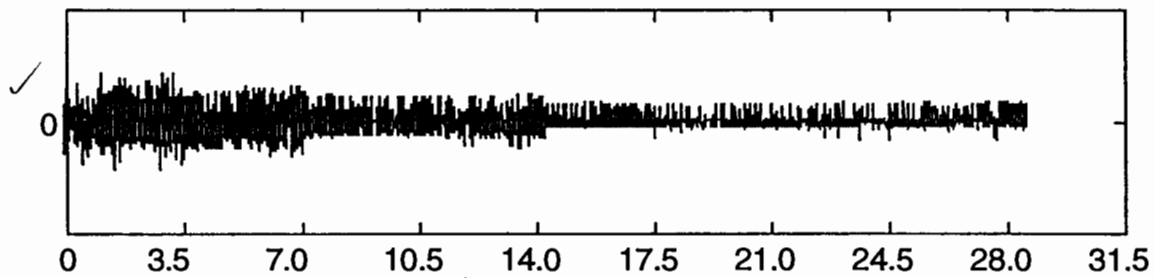


Figure 9: Haar wavelet transform applied to an unfiltered signal.

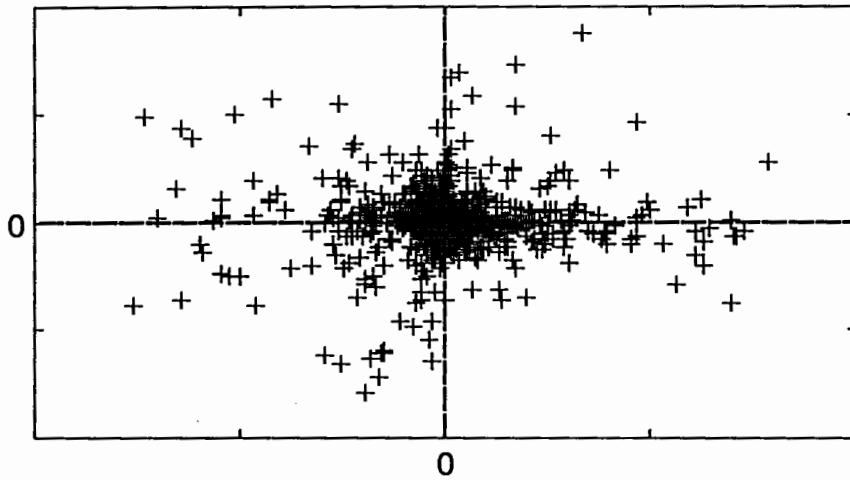


Figure 12: Score plot of first two principal components of Haar wavelet transformed data.

hmod

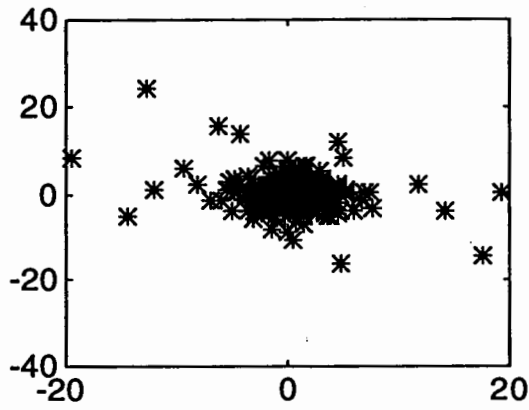
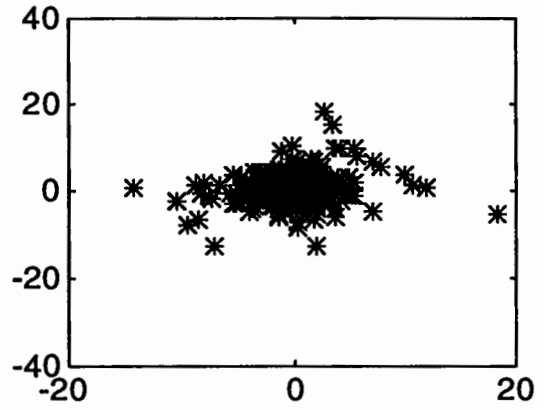
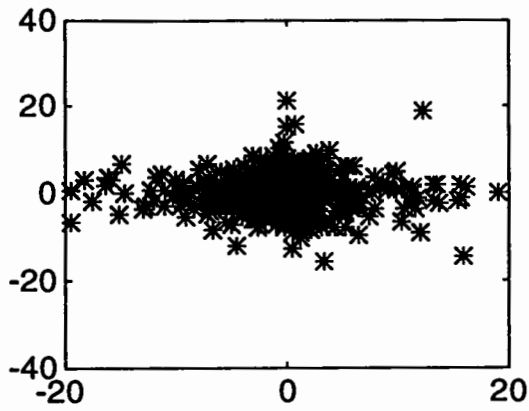


Figure 13: Score plots of the 3rd through 8th principal component of Haar wavelet transformed data.

hmed3

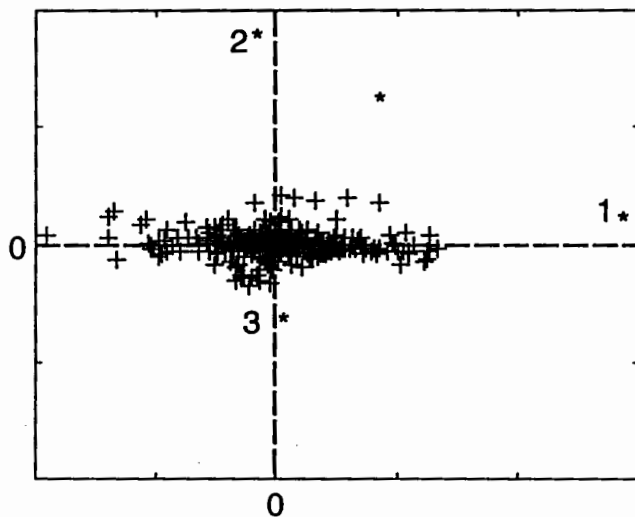


Figure 14: Score plot projecting a portion of the test set onto the calibration model in the score space.

tmod

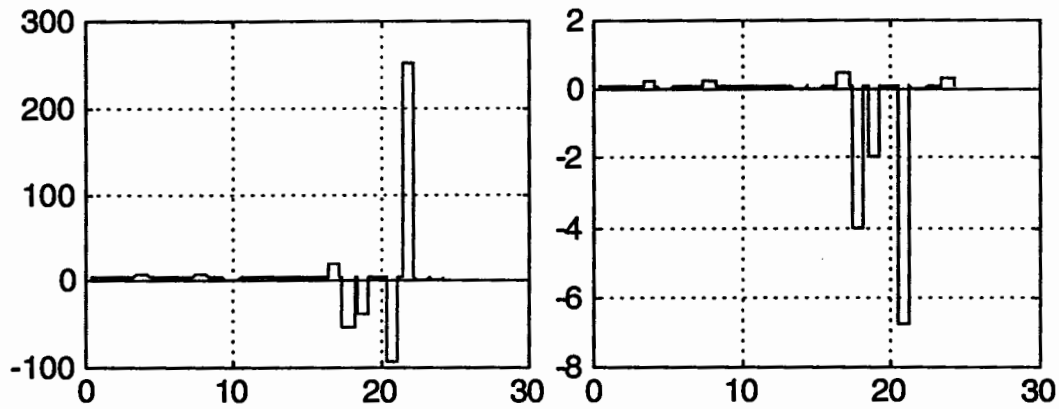


Figure 15: Contribution plots. Left panel: test sample #1, right panel: expected contribution to the first principal component.

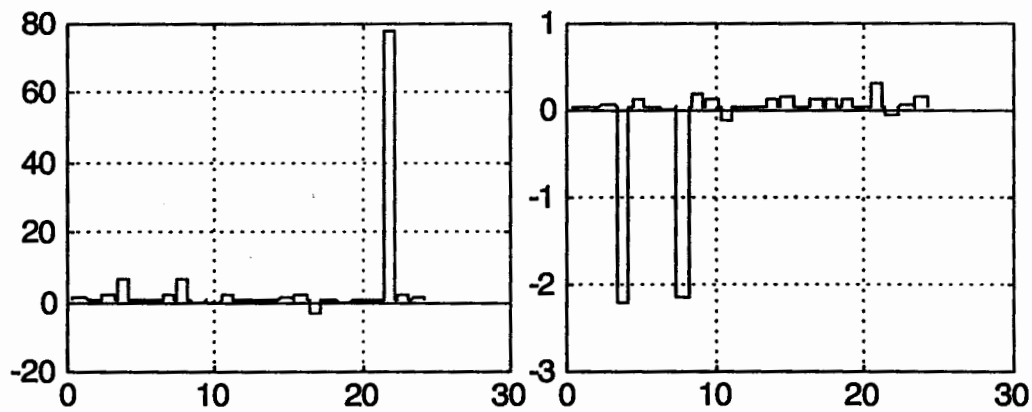


Figure 16: Contribution plots. Left panel: test sample #3, right panel: expected contribution to the second principal component.

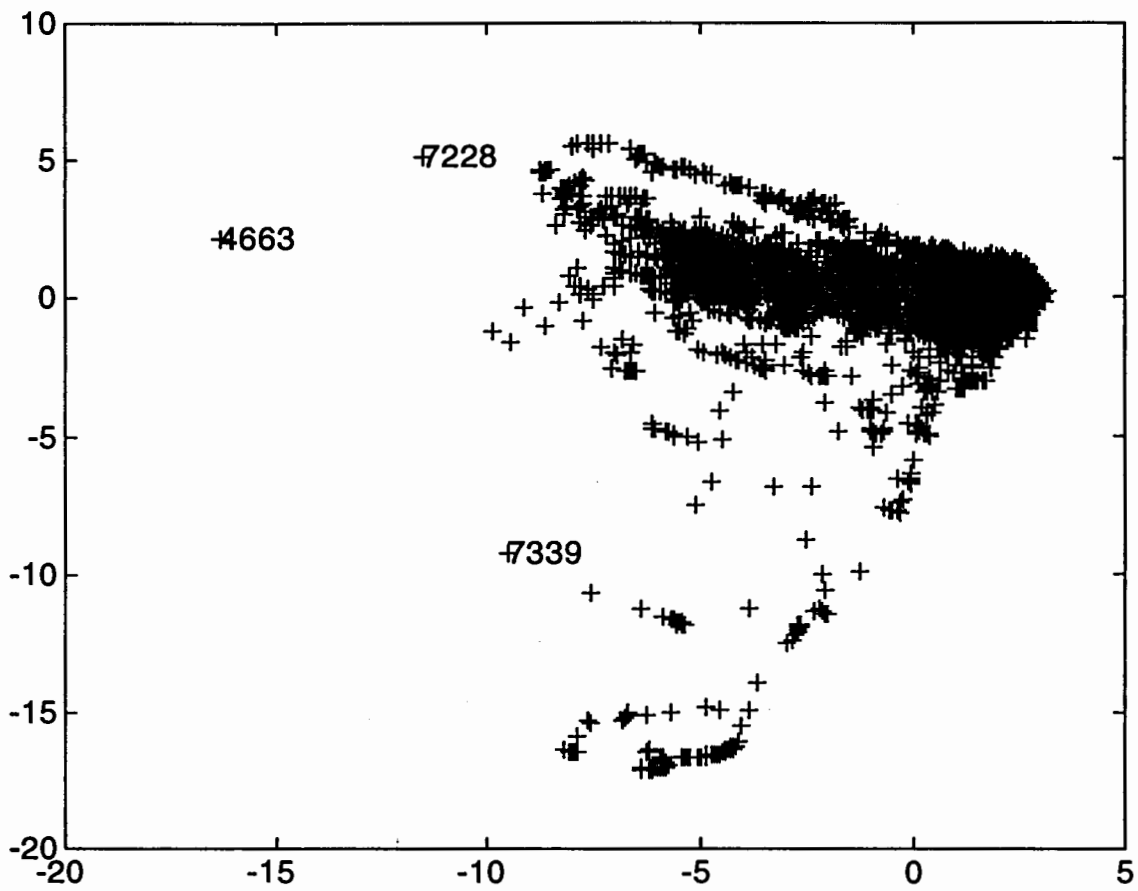


Figure 10: Score plot of first two principal components after applying PCA on FMH filtered data only.

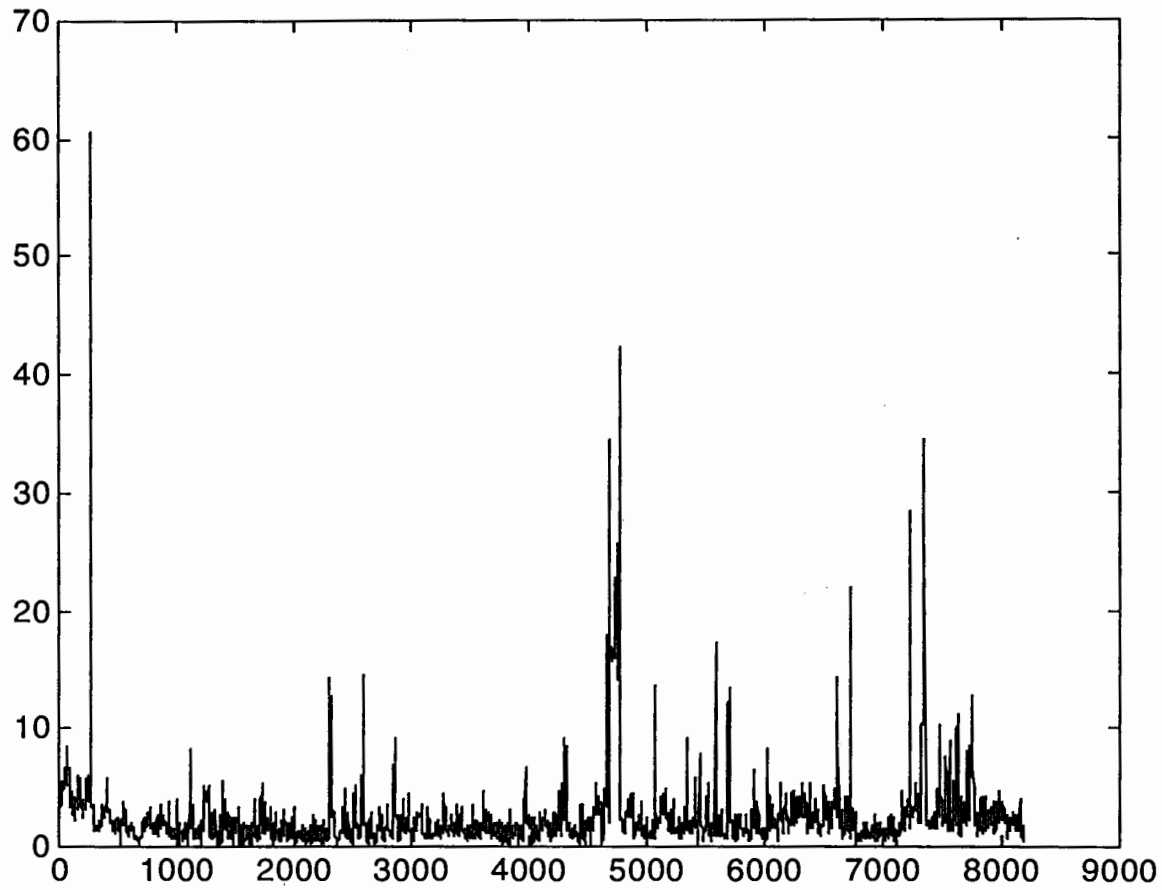


Figure 11: Q-statistic plot after eight principal components calculated after applying PCA on FMH filtered data only.