JERNKONTORETS FORSKNING

Serie	Nr	Datum	Forskningsuppgift nr
D	**806**	**2003-11-27**	**51045**

(TO 51-40)

INTELLIGENT ALARM HANDLING IN THE STEEL MANUFACTURING INDUSTRY

THE APPLICATION OF MULTIVARIATE STATISTICS TO REHEATING FURNACES AND TO AN ACID REGENERATION PLANT - REPORT NO 1

Report by O Marjanovic and D J Sandoz, Control Technology Centre Ltd.
Per-Olof Norberg, Advanced Process Control Ltd.
Björn Jönsson, AvestaPolarit AB
Lennart Klarnäs and Magnus Norberg, SSAB Tunnplåt AB

Key words: multivariate statistics, reheating furnaces, acid regeneration, process control, iron and steel industry

SUMMARY

In this project three specific programmes have been developed, each on a different process and with a different objective. These Programmes are

- Validation of thermocouple measurements in a reheating furnace (AvestaPolarit AB) It is shown clearly that an installed condition monitor would have detected an anomaly between measured and calculated temperatures and would have provided a valid estimate of temperature that could be used to temporarily replace the measurement during the time of anomaly. Such validation must provide the means for more effective energy management of the furnace by avoiding the positioning of control system set points at inappropriate temperatures.

- Development of a prediction model for NO_x emissions in a reheating furnace (SSAB Tunnplåt) A prediction model has been developed which gives a satisfactory level of accuracy. Future developments should focus on incorporating the developed prediction model into the advisory system that would indicate which cause variables should be changed and by what amount in order to minimise NO_x emissions while maintaining high productivity.

- Investigation of the use of Multivariate Statistics for the modelling of an acid regeneration process (SSAB Tunnplåt)
  The investigations with the acid regeneration plant have proven to be less productive than those reported above for the reheating furnaces. However interesting aspects have been shown concerning the capability of the Multivariate Statistics to classify regions of process operation and to relate these regions to variations in key process variables.

ISSN 0280-249X

Table of Contents

1. Introduction and Background

This project was carried out under the auspices of “The Jernkontoret”, the Swedish Steel Producers’ Association, in order to investigate the applicability of various intelligent alarm handling methods in addressing energy related condition monitoring issues within the steel manufacturing industry. The funds for the project have been provided by the Swedish Energy Agency (STEM). Industrial applications have been investigated in two companies in Sweden, SSAB Tunnplåt and AvestaPolarit AB. The work has involved engineers from these two companies and also from Advanced Process Control Ltd(APC) and Control Technology Centre Ltd(CTC). APC have provided expertise and consultancy in the field of reheating furnace operations and related control system design. CTC, a spin out company from the University of Manchester, has provided engineering services and software to exploit advanced alarm handling technology based upon techniques in Multivariate Statistical Process Control (MSPC).

Control Technology Centre Ltd. has been involved for a number of years in developing the software product MonitorMV, a toolbox of technologies for process condition monitoring, fault detection and diagnosis. In this project, MonitorMV has been employed in the development of several condition monitoring solutions for the steel manufacturing industry. The MonitorMV Product has been under continuous development throughout the course of this project. CTC has attracted funds from a variety of companies, from the process control industries and from the mining industries in order to fund this development. The capability of the product has been progressed in part on the basis of the various experiences gained in industrial applications in the chemical, steel and minerals processing industries.

The project commenced in January 2001 and finished in September 2003. During this period work has progressed to address three specific programmes of work each on a different process and with a different objective. These Programmes are

- Validation of thermocouple measurements in a reheating furnace (AvestaPolarit AB)

- Development of a prediction model for NOx emissions in a reheating furnace (SSAB Tunnplåt)

- Investigation of the use of Multivariate Statistics for the modelling of an acid regeneration process (SSAB Tunnplåt)

CTC Ltd made 5 visits to both SSAB Tunnplåt and AvestaPolarit AB during the course of this project in order to gain experience of the processes in question and in order to implement and assess the online capability of the various developments that have been established with MonitorMV.

2. About MonitorMV

MonitorMV is a toolbox of technologies for process condition monitoring, fault detection and diagnosis.

The MonitorMV software includes a range of both standard and state-of-the-art methods, which fall under the heading of multivariate statistical process control (MSPC). Technologies included in MonitorMV are Principal Component Analysis (PCA) and Partial Least Squares (PLS) modelling, clustering, statistical modelling using either Gaussian or Kernel-based methods and multiple model set handling for the real-time monitoring of complex processes. In support of these technologies, MonitorMV offers a range of visualisation options, including 2D/3D contour plots and quality control charts as well as the traditional MSPC plots. A recent development is MonitorMV Batch, an additional range of tools for tackling condition-monitoring issues for specifically batch processes.

MonitorMV is composed of two separate software systems, MonitorMV Design and MonitorMV Online. The primary function of the MonitorMV Design system is that of an analysis and design tool for the creation of Statistical Models to describe process data and for the development of Statistical Condition Monitors on the basis of these models. The primary function of the MonitorMV Online system is that of a real- time data collection and condition-monitoring tool. The MonitorMV Online system can interface with a range of proprietary DCS systems as a basis for the real-time collection of process data. Such data may then be evaluated with respect to one or more condition monitoring models previously created using MonitorMV Design. The MonitorMV Online system also has the capability for executing real-time signal processing calculations using the MonitorBasic programming language.

Communication between MonitorMV Design and MonitorMV Online is achieved through Model Files and Specification Files. Subject to the availability of sufficient memory, MonitorMV Design and MonitorMV Online may be executed concurrently on the same computer, or alternatively they may be on separate computers.

In January 2003, a new company, Perceptive Engineering Ltd., was formed in order to bring MonitorMV to be a commercially available software package, supported by a group of advanced control and condition monitoring consultants and engineers. For further information on MonitorMV, please contact Perceptive Engineering Ltd. (www.perceptive-engineering.co.uk).

3. Technology Primer

The techniques that have been used in this project fall under a general heading of Multivariate Statistical Process Control (MSPC). These techniques include Principal Component Analysis (PCA) and Partial Least Squares, or Projection to Latent Structures, (PLS) modelling.

The primary objective of an MSPC suite of software is to model and monitor a process over time in order to detect if statistically significant events, or abnormalities, occur. This technology relies heavily on the concept of cross- correlation in order to capture the underlying relationships between various process variables that exist during the normal process operation. Both PCA and PLS are introduced, in some detail, in the following two sections of this chapter.

3.1 Principal Component Analysis

Principal Component Analysis (PCA) is a method of extracting the majority of information from a set of measured signals, i.e. process variables, and expressing it using a greatly reduced number of variables, known as principal components. This technique is widely used in areas where large quantities of highly correlated data needs to be consolidated and, as such, has found significant use in the process industries. In addition to reducing the dimensionality of problems prior to, for example, statistical analysis, PCA also tends to eliminate uncorrelated noise from multiple measurements.

PCA is based upon the matrix equation:

$$X = PDQ^T \tag{1}$$

The matrix X is referred to as the data matrix and is of dimensions $N \times m$, where N is the total number of data points in the data set and m is the number of process variables. Note that, generally, there are more data points than process variables, i.e. $N >> m$.

Therefore, the data matrix, X, can be defined as follows:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & & x_{2m} \\ \vdots & & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{Nm} \end{bmatrix}$$

where x_{ij} is the measurement of variable j at time i. PCA breaks this matrix into three other matrices, P, D and Q. D is a diagonal matrix of dimensions $m \times m$, whilst P and Q are orthonormal matrices such that:

$$PP^T = QQ^T = I_{m \times m} \tag{2}$$

Traditionally, PCA is performed using Singular Value Decomposition (SVD). This method allows one to obtain P, D and Q. In particular, the diagonal matrix D contains m nonnegative principal component amplitudes, in descending order, thus:

$$D = \begin{bmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_m \end{bmatrix}$$

where $d_1 \geq d_2 \geq ... \geq d_m$ and $d_i \geq 0$ for $1 \leq i \leq m$.

For the cases with highly correlated variables only the first few diagonal terms in matrix D will have significant values while the remaining ones will be close to zero. This result corresponds to the fact that the first few principal components are capable of explaining the majority of the variation (or information) within the measured signals. In fact, the original PCA relationship, given in equation (1), can be replaced by a greatly reduced set of equations:

$$\hat{X} = P_n D_n Q_n^T \tag{3}$$

where P_n is a matrix of dimension $N \times n$, consisting of the first n columns of the matrix P, Q_n is a matrix of dimension $m \times n$, consisting of the first n columns of the matrix Q and D_n is a diagonal matrix of dimension $n \times n$, containing the first n diagonal elements of the matrix D. Finally, $\hat{X}$ is the PCA prediction of the original data matrix.

Thus, the initial data matrix may be approximated to an arbitrary degree by just n retained principal components.

The score vector t at a time instant i is computed from:

$$t_i = \begin{bmatrix} x_{i1} & x_{i2} & \cdots & x_{im} \end{bmatrix} \cdot Q_n \tag{4}$$

The scores can be plotted in 2D or 3D displays providing graphical representation of the main features in the data set. This feature has been employed in the case of statistical analysis of the acid regeneration process, described in section 7.3.

PCA predictions of measured variables at time instant i are then obtained by:

$$\hat{X}_i = t_i Q_n^T \tag{5}$$

In the context of a general condition monitoring application, PCA predictions of measurements, given in equations (4) and (5), play the key role. In particular, by observing the prediction error for each variable, i.e. $x - \hat{x}$, it is possible to pinpoint the set of variables that deviate from their expected behaviour. Furthermore, in a case of the instrument validation, variable that corresponds to a malfunctioning sensor can then be removed from the calculation of scores, given in equation (4), while the inference of its true value is still achievable through the use of equation 5.

It is important to note that the number of retained principal components needs to be carefully chosen. There are a number of techniques that exist for its appropriate selection. These include, amongst others, auto- correlation, cross- validation, cumulative percent variance, scree test and so on. MonitorMV allows the methods of cross- validation and cumulative percent variance to be employed by the user in selecting a number of principal components. Also, displays of X and $\hat{X}$ are available for visual inspection in order to decide on a number of selected principal components. The basic approach, when performing PCA using MonitorMV, is to observe the display given in Figure 1.

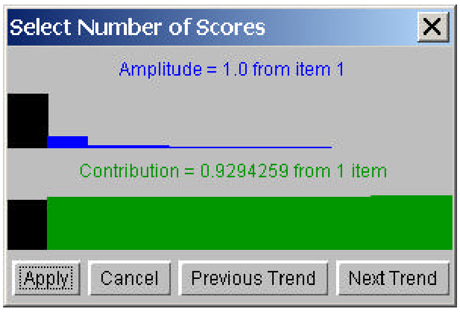


Figure 1

The top graph, coloured in blue, presents the relative value of the diagonal element in the matrix D, defined earlier in this section, that is associated with the selected principal component. The bottom graph, coloured in green, is based upon the extent to which the model explains the data (1 is perfect) with a selected number of retained principal components, sometimes called cumulative variance. In this particular case, it is shown that the first principal component contributes 92.94% to the total training data variation.

In addition, MonitorMV allows the user to view plots of measurement and PCA-based predictions of each signal considered by PCA analysis, as shown in Figure 2. In this display PCA- based prediction trends are coloured brown while the measured signal trends are coloured in blue, green or magenta.

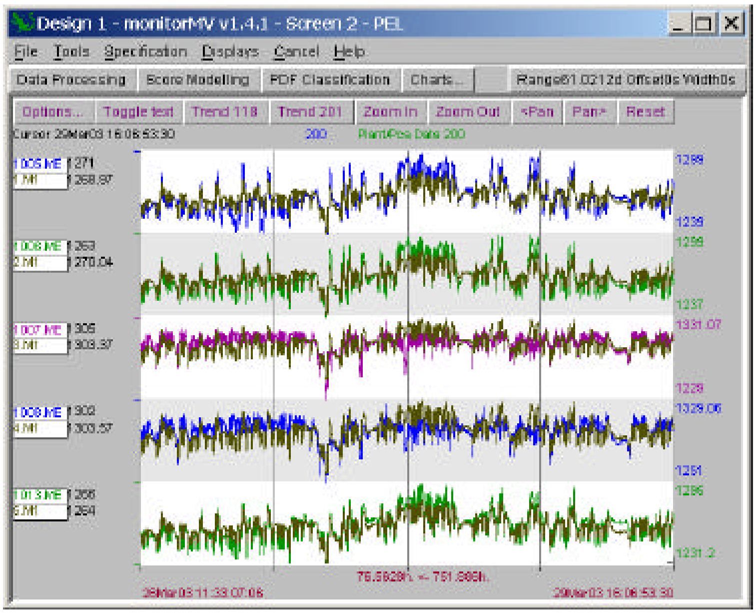


Figure 2

Finally, a computationally intensive cross-validation method can be employed in the MonitorMV Design system in order to assist the user in choosing a number of retained principal components. In this approach, data is subdivided into training and validating sets. The PCA computation is performed upon the training set and the prediction errors are evaluated over the validating set. The data is then 'rotated' to give a different split of training and validating sets and the computation is repeated. As a result of cross- validation calculations a complex statistic of the PCA prediction error, abbreviated as PRESS, is computed and presented alongside the amplitudes and cumulative variance contribution of principal components, as seen in Figure 3.

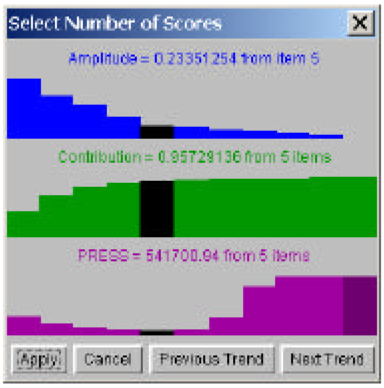


Figure 3

The optimal number of retained principal components is given at a point where the PRESS is at its minimum. In the case of Figure 3, the optimal number is set to 5.

Finally, it should be noted that the number of retained principal components generally reflects the number of independent features in the considered data set. Hence, by using the physical knowledge of the process it may be possible in some cases to estimate the true number of principal components that should be retained.

3.2 Partial Least Squares (Projection to Latent Structures)

Partial Least Squares or Projection to Latent Structures (PLS) is a method of identification that offers certain attractive features, both in providing a more robust identification approach than the Ordinary Least Squares (OLS) approach and as a basis for multivariate condition monitoring. The basic approach of the algorithm is, as with PCA, to identify the principal features in the data. However, unlike PCA, PLS divides the variables into cause and effect. It then identifies the primary features in the cause variables that are able to describe the variation in the effect variables.

The basic cause- effect structure of PLS, given in matrix form, may be written as:

$$Y = AX + e \tag{6}$$

where X and Y represent the input (cause) and the output (effect) matrices respectively and can be defined as follows:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & & x_{2m} \\ \vdots & & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{Nm} \end{bmatrix}, \; Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & & y_{2n} \\ \vdots & & \ddots & \vdots \\ y_{N1} & y_{N2} & \cdots & y_{Nn} \end{bmatrix}$$

The Error term is described by column vector e, where $e = [e_1 \quad e_2 \quad \ldots \quad e_N]^T$.

Note that in this definition of cause and effect matrices it is assumed that there are m cause variables, n effect variables and N measurement samples.

The basic PLS algorithm is recursive in nature, with the aim of breaking the matrices X and Y down into products of the score matrices, U and T, and the loadings vectors, Q and P, thus:

$$Y = U_k Q_k^T + E_k \tag{7}$$

$$X = T_k P_k^T + F_k \tag{8}$$

for the case of k scores or 'latent variables' (LVs).

Note that the issue of choosing the number of latent variables (LVs) is analogous to that of choosing principal components in PCA analysis, and is, therefore, not discussed in any detail in this section.

In the first iteration, values are computed for U_1, Q_1, T_1 and P_1 that maximise the covariance between X and Y. This contribution is then subtracted from the data matrices and the procedure repeated for subsequent values of U_j, Q_j, T_j and P_j, where $1 < j \leq k$.

In the traditional PLS approach, scores for the k retained latent variables are defined from the loading vectors of the cause and effect matrices as follows:

$$T_k = XP_k \tag{9}$$

$$U_k = YQ_k \tag{10}$$

whilst the predictions may be generated from the cause and effect variables by using:

$$\hat{X}_k = T_k P_k^T \tag{11}$$

$$\hat{Y}_k = U_k Q_k^T \tag{12}$$

The matrices X and Y are now indirectly related through their scores by the so-called 'inner model', which is simply a linear regression of t_i on u_i for $1 \leq i \leq k$, yielding:

$$\hat{U}_k = T_k B_k \tag{13}$$

where B_k is a matrix of regression coefficients.

Hence, by substituting (9) into (13) and then substituting (13) into (12) the following input- output relationship is obtained, which relates input matrix to an output matrix:

$$\hat{Y}_k = X\left[P_k B_k Q_k^T\right] \tag{14}$$

In the context of the NOx estimation problem, reported in Chapter 6, PLS has been employed to extract the main features in the cause variables' data and relate these to a single effect, namely NOx emissions. By employing PLS, as opposed to ordinary least squares (OLS) for example, the issue of cross- correlation between different cause variables is appropriately addressed ensuring a robust prediction model. Hence, the aim in developing a prediction model can be stated as an appropriate identification of the expression given in equation (14).

4. Overview of Reheating Furnace Operation

This Overview is presented in order that the reader may properly appreciate the investigations described later in this report.

Reheating furnaces are the first component of the hot-strip rolling mill at the SSAB factory in Borlange. Their purpose is to reheat the steel slabs from ambient temperature to around 1200 degrees C. The source of energy is burning (oxidation) of volatile liquids or gases (heavy fuel oil, LPG, natural gas). As a by- product of oxidation, nitrogen oxides are produced and discharged into the atmosphere through a stack.

Reheating furnaces generally consist of three chambers: preheating, heating and soaking zones. The slabs are fed into the preheating zone, through the charging door, and then slowly moved through heating and soaking zones, sequentially. The slabs are heated roughly to the required temperature in preheating and heating zones. The purpose of soaking zones is to achieve uniform temperature distribution of the slabs.

The key business drivers for the reheating furnace are given as follows:

- Maximise productivity (throughput of slabs)
- Minimise running cost (energy consumption)
- Minimise negative temperature deviation from the ideal heating curve (avoid under- heating of the steel slabs)
- Maintain gaseous emissions (NOx) within legislation limits.

In order to achieve satisfactory reheating of the steel slabs while avoiding excessive energy consumption, the furnace is equipped with a Fuel Optimisation Control System (FOCS). This control system regulates the slabs' temperatures by manipulating set- points of different zone temperature PID- based local control loops. PID controllers, in turn, control the zone temperature by manipulating the air and fuel flow rates into the zone burners.

A Diagram and Specification of the reheating furnace U302 at the SSAB site in Borlange, Sweden, is given in Figure 4.

Reheating Furnace U302
Building year: 1982
Manufacturer: Italimpianti
Capacity: 300 ton/h
Burner zones: 9
Burners: 119
Fuel: Heavy fuel oil
Length: 36.3 m
Slab extraction temp: 1150-1250 C

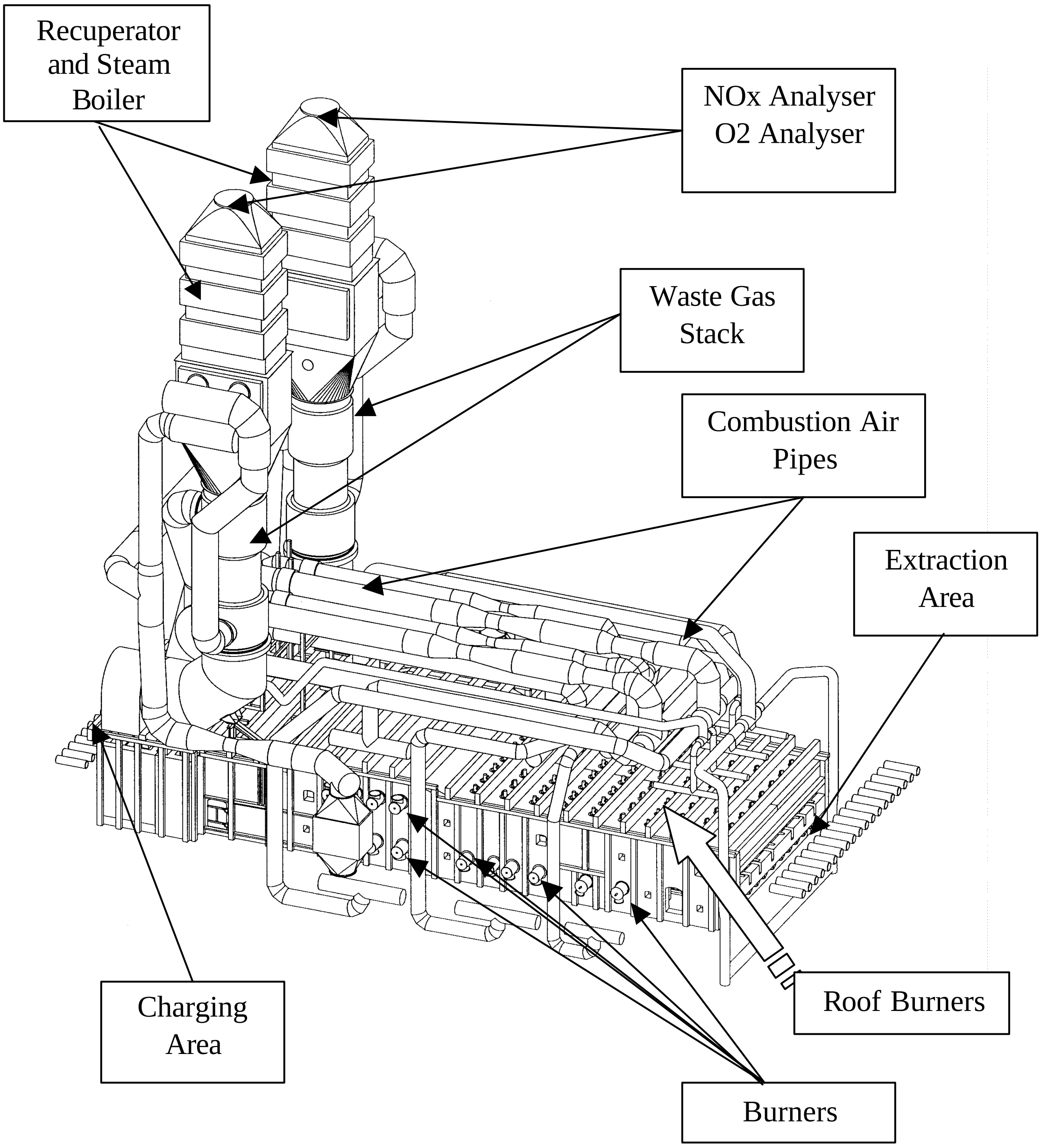


Figure 4

5. Sub-Project 1; The Validation of the Thermocouple Measurements using Principal Component Analysis

5.1 Introduction

A fundamental concept in any automatic control system is the utilisation of feedback, which is achieved by means of real-time measurement of the variables that are to be controlled and the variables that give rise to disturbances but which cannot be controlled. It is essential, for satisfactory control system performance, that these measurements are reliable and consistent. Any improvement in the robustness of control system measurements makes a direct contribution to the overall reliability of the control system.

Furthermore, general sensor equipment is susceptible to long- term drifts and sudden failures. These malfunctions result in off-quality product, less than optimum yields, under utilised capacity and unnecessary emergency shutdowns amongst other things. Therefore, improvement in measurement reliability also has a direct impact on key business drivers in any manufacturing industry.

The Fuel Optimisation Control System (FOCS) scheme, employed in the reheating furnaces of the hot-strip rolling mills, relies heavily on accurate temperature measurement inside the different furnace zones. These thermocouple-measured temperatures are used as measurements in local PID- based control schemes (one controller for each furnace zone) as well as for the initial conditions in slab temperature calculations. Hence, the reliability of these sensors has an important impact on the performance of the overall reheating furnace control system.

The impact of faulty or erroneous temperature measurement in a reheating furnace is twofold. In the case where the measured value is below the actual temperature, excessive fuel is used in the furnace burners. This in turn increases the energy consumption, which is probably the main business driver for this process. On the other hand, if the measured value is above the actual temperature then the product quality may be degraded. Hence, in either case an important business driver is adversely affected by the failure of instrumentation equipment to provide accurate and reliable feedback measurement.

This sub-project is concerned with the validation of such measurement and is based on the principle of redundancy through the utilisation of the Principal Component Analysis (PCA) method. In particular, the presence of cross- correlation between different thermocouple measurements is exploited for the detection of faulty thermocouples and for subsequent estimation of the true values of the associated temperatures.

This sub- project has been carried out at the AvestaPolarit site in Avesta, Sweden. The system, described in this report is currently under trial in AvestaPolarit and is expected to find its way to the control room of the reheating furnace as a valuable tool in addressing the reliability of instrumentation equipment.

5.2 Training Data

The data that is used for the development of PCA models has a direct impact on the performance of the resulting condition monitor. In particular, since the data used for training generally represents the normal state of the affairs, great care must be taken not to include those periods during which problems were encountered with the process itself or with individual instrumentation units.

In this particular case, training data were chosen from a 2 month period covering March and April 2003, taken for 24 thermocouples which are situated in all of the zones of the reheating furnace. Data points that correspond to the periods during which maintenance of the rolling mill was performed have been excluded from the training data since such data are not consistent with those of normal process operation.

Note that, ideally, training data sets should be chosen to correspond to those periods of time that follow immediately after thermocouple re-calibration takes place. Should it be seen that there is significant difference in the statistical interpretation when compared with earlier training, this will indicate there might have been some weakness in the calibration procedure that should be investigated (e.g. such as a thermocouple being displaced within its mounting pocket). Proper calibration is vital if energy management issues of a furnace are to be properly addressed.

5.3 Development of Condition Monitors

5.3.1 Introduction

The first step in the development of statistical models is to observe the cross-correlation between various signals. Such information may help in grouping highly-correlated signals into one set to be considered by a single condition monitor. Also, in the case of mutually uncorrelated sets of highly correlated variables, correlation analysis may provide a clue as to how many principal components are required to adequately represent the training data set. However, caution is in order at this point due to the fact that the training data represents a normal operating regime. There may not be sufficient excitation of the key cause variables of a process in order to bring out inter-relationships between process variables and, therefore, correlation analysis may produce misleading results. Hence, the results from the correlation analysis are to be taken with some caution. So unsurprisingly, process-oriented knowledge may be much more valuable asset in addressing issues concerning statistical modelling rather than 'blind' correlation analysis.

Due to the character of operation and the geometric shape of the reheating furnace, thermocouple measurements from different zones are not as highly correlated as one may expect. In fact, there is very little correlation between thermocouple measurements from different zones. For example, temperature measurements from the preheating zones are almost completely uncorrelated with temperature measurements from the soaking zones.

In order to improve accuracy of the fault detection/diagnosis scheme it has been decided to design three condition monitors that would focus on different sections of the reheating furnace. The criterion for grouping of thermocouples for each condition monitor has been taken to be the level of cross-correlation between these measurements as well as their mutual closeness in physical sense. Also the attempt has been made to minimise the number of PCA-based condition monitors in order to keep the real-time application reasonably simple. Thus three such monitors have been considered – Condition Monitors 1000, 2000 and 3000.

5.3.2 Development of Condition Monitor 1000

This condition monitor considers 4 thermocouples. These are subdivided into 2 mutually uncorrelated sets of almost identical signals. Each set belongs to a particular zone in the reheating furnace. In particular, T1_Oster and T15_Vaster are situated in the bottom of the dark zone while T10_Oster and T24_Vaster are located in the preheating zone 7. Since these signals constitute two sets of perfectly correlated variables, which are mutually uncorrelated, one would expect two principal components, which is indeed the case as shown in Figure 5.

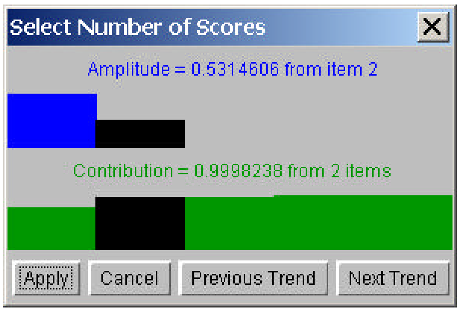


Figure 5

In this particular case it is shown that first two principal components contribute 99.98% to the total training data variation.

Results of the statistical analysis performed on the prediction errors for these four signals over the month of June (validating data set) are given below in Table 1.

Signal ID	Mean	Deviation	Maximum	Minimum
T1_Oster	0.11	0.99	5.01	-6.28
T10_Oster	0.08	0.39	3.36	-2.65
T15_Vaster	-0.11	1	6.33	-5.06
T24_Vaster	-0.08	0.39	2.65	-3.35

Table 1

While Table 1 shows that the developed model is highly accurate in predicting thermocouple measurements, the total number of highly cross- correlated signals is small. As a result, the true value of the temperature related to a faulty thermocouple cannot be estimated by this condition monitor since the level of redundancy is small. Nevertheless, such an accurate condition monitor would be able to accurately identify the particular zone of the reheating furnace within which one of the two thermocouples is malfunctioning. Identification of a particular faulty thermocouple would, however, have to be left to a process engineer at a site.

5.3.3 Development of Condition Monitor 2000

Variables that are considered by this condition monitor represent pairs of highly correlated thermocouples, belonging to the same reheating furnace zone. In particular, thermocouples from the upper dark zone (T2_Oster and T16_Vaster), preheating zone 8 (T11_Oster and T25_Vaster), heating zone 1 (T3_Oster and T17_Vaster) and heating zone 2 (1004.ME and 1017.ME) as well as those situated in the bottleneck between heating and soaking zones (T13_Oster and T27_Vaster) are considered by this monitor.

Amplitudes of individual principal components and their cumulative contribution to the total variation of the training data set are displayed below in Figure 6.

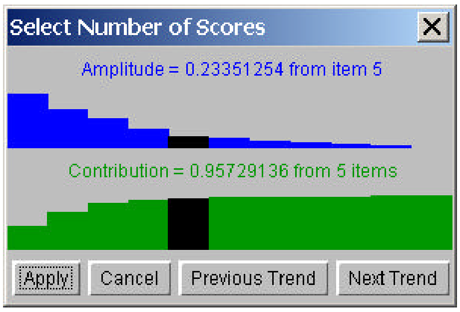


Figure 6

In this case it has been decided to choose 5 principal components, which contribute 95.73% to the total variation in the training data set. Note that the contribution to the total variation does not change significantly as the number of principal components changes from 4 to 5, as seen in Figure 6. There are two reasons why 5 components have been chosen rather than 4. Firstly, these 10 thermocouple measurements are grouped into 5 pairs that are situated in the same zone. Hence, it is expected that there would be at most 5 principal components. Note that any further reduction in a number of principal components would arise from the cross correlation between different pairs of thermocouple measurements. Therefore, it is reasonable to set upper bound on the number of principal components to 5. Secondly, observations of prediction errors, when evaluated over the validating data set (June 2003) reveal the benefit in choosing 5 rather than 4 principal components. Furthermore, by employing cross- validation scheme, through the computation of PRESS statistic, it is found once again that 5 PCs is an optimal choice for this model.

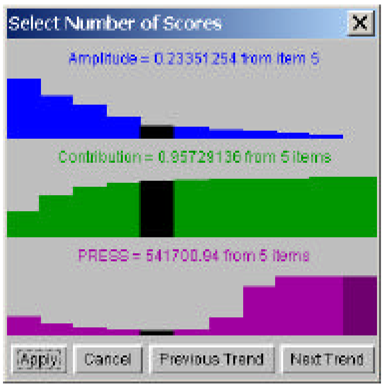


Figure 7

The tables given below contain results of the statistical analysis, performed on the validating data set (June 2003), of prediction errors for both PCA models.

PCA with 4 Principal Components

Signal ID	Mean	Deviation	Maximum	Minimum
T2_Oster	1.04	16.15	98.44	-66.61
T3_Oster	-1.14	5.2	17.04	-27.88
T4_Oster	-2.86	10.12	33.61	-47.21
T11_Oster	-0.03	12.55	44	-40.09
T13_Oster	0.04	5.64	25.8	-41.91
T16_Vaster	-2.06	12.52	65.58	-87.82
T17_Vaster	2.18	4.07	22.76	-14.07
T18_Vaster	-1.11	15.91	58.98	-60.45
T25_Vaster	5.07	13.66	65.52	-42.13
T27_Vaster	0.34	6.28	52.14	-31.99

Table 2

PCA with 5 Principal Components

Signal ID	Mean	Deviation	Maximum	Minimum
T2_Oster	2	14.24	89.7	-75.66
T3_Oster	-1.72	3.5	13.41	-21.12
T4_Oster	-1.84	5.79	19.91	-28.03
T11_Oster	-1.51	5.64	33.47	-30.79
T13_Oster	-0.46	5.07	24.64	-40.88
T16_Vaster	-2.06	12.78	64.76	-88.65
T17_Vaster	2.09	3.78	23.9	-13.99
T18_Vaster	0.48	6.61	43.07	-35.88
T25_Vaster	4.1	8.73	44.33	-22.25
T27_Vaster	0.74	6.34	51.36	-31.11

Table 3

In these tables it is observed that the standard deviation of the PCA model with 5 PCs is reduced, in some cases significantly, when compared with PCA model having 4 PCs. This is especially so in the cases of T4_Oster, T11_Oster, T18_Vaster and T25_Vaster.

Also, it was observed that PCA with 5 PCs is as robust as the PCA model with 4 PCs when a number of thermocouple signals are masked out. Hence, in order to improve predictability and maintain the level of robustness when several thermocouple measurements are inferred (masked) rather than measured it has been decided to choose the PCA model with 5 principal components.

5.3.4 Development of Condition Monitor 3000

Variables that are considered by this condition monitor are 10 thermocouple measurements that are located in the soaking zones of the reheating furnace. This set of signals is highly correlated as reflected in Figure 8. In this display it is observed that taking only the first two principal components contributes 98.62% to the total variation of the training data.

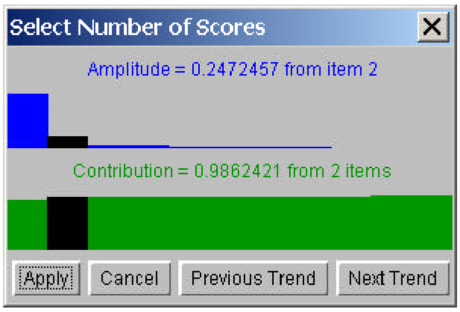


Figure 8

Hence, this PCA model contains 2 principal components. Validation, using the data from June 2003 has generated the following results for the prediction errors.

Signal ID	Mean	Deviation	Maximum	Minimum
T5_Oster	0.79	2.26	12.73	-12.65
T6_Oster	-0.82	3.31	17.61	-12.4
T7_Oster	-16.54	6.22	17.97	-36.14
T8_Oster	4.77	5.43	36.53	-19.55
T14_Oster	-4.85	3.65	8.51	-29.71
T19_Vaster	5.81	2.41	15.33	-6.93
T20_Vaster	-0.03	2.48	19.34	-10.3
T21_Vaster	9.76	4.07	29.7	-24.8
T22_Vaster	3.34	6.87	28.2	-32.16
T28_Vaster	-4.09	3.82	25.86	-45.73

Table 4

It is interesting to note that there is a relatively large mean value in the prediction error of T7_Oster (situated at the bottom of the soaking zones). This factor is explained by the trends of Figure 9 (for T7_Oster), where the coloured line corresponds to the measured signal and the brown line corresponds to the respective PCA prediction signal.

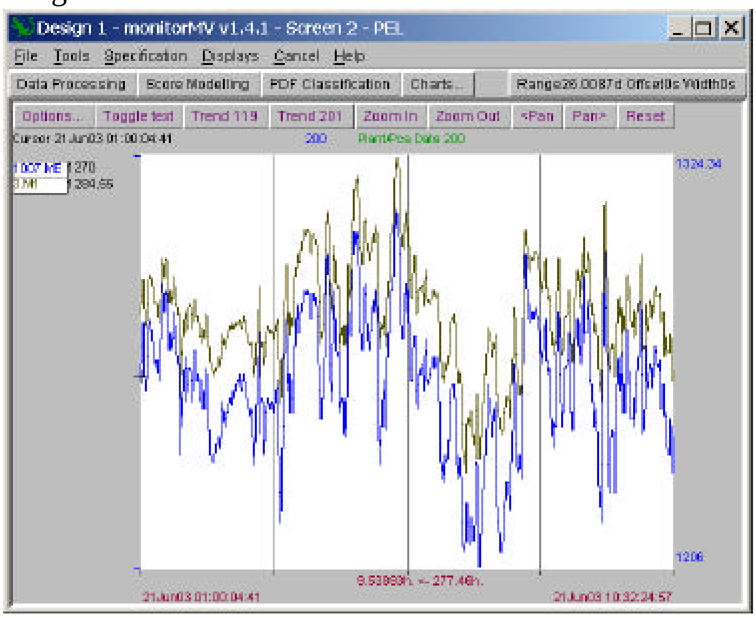


Figure 9

Note that the blue line is consistently below brown line indicating that this thermocouple measurement may suffer from systematic error.

5.4 Improving the Robustness of Condition Monitors

The key challenge in this particular application is the lack of the strong cross-correlation between various thermocouple measurement signals. As a result, prediction errors of the PCA- based statistical model that are routinely encountered are of comparable size to the measurement errors which have consequential impact on the key business drivers, such as the energy consumption (in the case of the negative prediction error) as well as the product quality (in the case of the positive prediction error).

There are many different methods of increasing or decreasing the sensitivity of condition monitors. One method is to reduce the number of false alarms by limiting attention to those events that contain frequency components in a specific range. For example, if the event that is to be detected is the slow drift, representing a general low- frequency signal, then by low- pass filtering information such as prediction error it is possible to solely focus on all those events that belong to this very specific band of frequencies. In that case sudden and short-lived disturbances, observed in prediction error trends, are ignored during the filtering process while the slow disturbances are emphasised.

In this particular case there is no specific and unique spectrum of events that can be detected. In many cases a slowly drifting thermocouple is not easily detectable and yet may have long- term impact on production. On the other hand, sudden and rapid failure of a thermocouple causes complete loss of information that may play a crucial role in the overall automation scheme. Hence, unsurprisingly, there has to be a compromise between emphasising slow drifts (low- frequency) and sudden and rapid changes (high-frequency). In this application such compromise is quantified by means of a first- order filter time constant. General formulation of a simple delay-free and unity gain first- order filter is given in the Laplace Domain as follows:

$$G(s) = \frac{1}{Ts+1} \tag{15}$$

where T represents time constant, expressed in seconds, while s is a complex Laplace variable.

The impact that a filter time constant has on prediction errors is discussed next.
The larger the time constant the greater is the emphasis on the low- frequency components of a filtered signal, as are seen to predominate in Figure 11. On the other hand, reducing the time constant does not shift emphasis from the low to high frequency band. Instead, it increases the bandwidth of a filter and allows more and more of high- frequency components to be present, alongside low- frequency components, in the output (filtered) signal.

If the sole objective of the filter is to reduce variation of the prediction error then the time constant should be very large, eliminating majority of the medium and high-frequency content in the filtered signal, as observed in Figure 11. However, in such case response to a sudden change in prediction error is likely to take a very long time, as it is observed in Figure 11, where the step –response of a filter is plotted as a function of time.

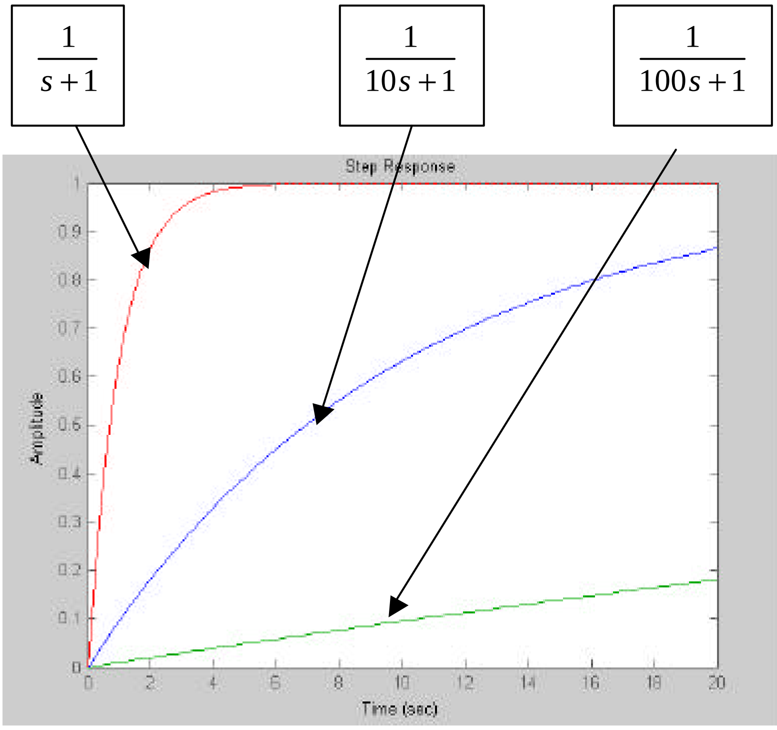


Figure 11

Next, the statistical information concerning the filtered prediction errors, evaluated over the validating data set (June 2003) for all three monitors is displayed. In order to demonstrate the effect that the choice of time constant has on a size of filtered prediction error three different cases were considered.

First of all Monitor 1000 is considered.

Time Constant = 1 minute

Signal ID	Mean	Deviation	Maximum	Minimum
T1_Oster	0.12	0.91	3.69	-4.55
T10_Oster	0.08	0.28	2.83	-1.74
T15_Vaster	-0.12	0.92	4.6	-3.73
T24_Vaster	-0.08	0.28	1.73	-2.82

Table 5

Time Constant = 10 minutes

Signal ID	Mean	Deviation	Maximum	Minimum
T1_Oster	0.12	0.74	2.96	-3.31
T10_Oster	0.08	0.16	0.89	-0.7
T15_Vaster	-0.12	0.75	3.35	-2.99
T24_Vaster	-0.08	0.16	0.69	-0.89

Table 6

Time Constant = 1 hour

Signal ID	Mean	Deviation	Maximum	Minimum
T1_Oster	0.12	0.5	1.96	-1.31
T10_Oster	0.08	0.11	0.45	-0.4
T15_Vaster	-0.12	0.51	1.33	-1.98
T24_Vaster	-0.08	0.11	0.4	-0.45

Table 7

Next, the results for the Condition Monitor 2000 are presented.

Time Constant = 1 minute

Signal ID	Mean	Deviation	Maximum	Minimum
T2_Oster	2.01	14.15	89.1	-74.53
T3_Oster	-1.72	3.46	12.77	-21.02
T4_Oster	-1.84	5.73	19.19	-27.83
T11_Oster	-1.51	5.57	33.04	-29.89
T13_Oster	-0.46	5.01	24.32	-40.14
T16_Vaster	-2.06	12.7	64.17	-87.89
T17_Vaster	2.09	3.72	23.26	-13.64
T18_Vaster	0.48	6.5	40.87	-34.8
T25_Vaster	4.1	8.63	42.62	-21.63
T27_Vaster	0.74	6.27	50.8	-30.64

Table 8

Time Constant = 10 minutes

Signal ID	Mean	Deviation	Maximum	Minimum
T2_Oster	2.01	12.51	80.32	-62.96
T3_Oster	-1.72	3.03	9.91	-18.25
T4_Oster	-1.84	5.35	17.37	-25.88
T11_Oster	-1.51	4.94	24.95	-21.48
T13_Oster	-0.46	4.36	19.87	-35.11
T16_Vaster	-2.06	11.24	54.02	-80.46
T17_Vaster	2.09	3.26	19.55	-11.7
T18_Vaster	0.48	5.72	25.62	-20.05
T25_Vaster	4.1	7.85	38.77	-15.14
T27_Vaster	0.73	5.51	44.93	-25.04

Table 9

Time Constant = 1 hour

Signal ID	Mean	Deviation	Maximum	Minimum
T2_Oster	1.99	8.49	38.67	-33.74
T3_Oster	-1.72	2.13	4.15	-10.09
T4_Oster	-1.83	4.35	9.97	-18.7
T11_Oster	-1.51	3.6	13.32	-14.13
T13_Oster	-0.46	3	8.18	-16.9
T16_Vaster	-2.05	7.51	26.12	-39.44
T17_Vaster	2.09	2.28	11.19	-4.21
T18_Vaster	0.47	4.31	21.65	-9.31
T25_Vaster	4.09	6.16	26.78	-11.03
T27_Vaster	0.73	3.94	22.87	-10.19

Table 10

Finally, the results for the Condition Monitor 3000 are displayed below.

Time Constant = 1 minute

Signal ID	Mean	Deviation	Maximum	Minimum
T5_Oster	0.79	2.16	12.14	-12.13
T6_Oster	-0.82	3.25	17.06	-11.6
T7_Oster	-16.54	5.96	16.46	-34.65
T8_Oster	4.77	5.12	35.2	-18
T14_Oster	-4.85	3.51	8.18	-28.4
T19_Vaster	5.8	2.35	14.74	-6.66
T20_Vaster	-0.03	2.38	16.73	-10.15
T21_Vaster	9.76	3.92	27.86	-18.19
T22_Vaster	3.34	6.67	26.78	-31.08
T28_Vaster	-4.09	3.59	23.7	-41.29

Table 11

Time Constant = 10 minutes

Signal ID	Mean	Deviation	Maximum	Minimum
T5_Oster	0.79	1.69	5.87	-9.12
T6_Oster	-0.82	2.75	12.19	-9.13
T7_Oster	-16.54	5.06	11.2	-29.06
T8_Oster	4.77	3.79	19.07	-13.74
T14_Oster	-4.85	2.71	5.31	-17.66
T19_Vaster	5.81	2.02	11.85	-5.25
T20_Vaster	-0.03	1.9	7.19	-7.83
T21_Vaster	9.75	3.19	21.14	-8.2
T22_Vaster	3.34	5.41	19.09	-22.56
T28_Vaster	-4.09	2.76	16.28	-19

Table 12

Time Constant = 1 hour

Signal ID	Mean	Deviation	Maximum	Minimum
T5_Oster	0.79	1.21	4.84	-5.05
T6_Oster	-0.82	2.12	7.69	-7.29
T7_Oster	-16.5	4.31	6.23	-24.6
T8_Oster	4.75	2.65	12.98	-6.92
T14_Oster	-4.85	1.83	0.34	-11.23
T19_Vaster	5.8	1.63	11.09	-2.69
T20_Vaster	-0.03	1.35	4.46	-4.11
T21_Vaster	9.73	2.4	17	-2.08
T22_Vaster	3.35	3.97	13.22	-15.88
T28_Vaster	-4.09	1.9	8.73	-8.87

Table 13

Note that as the value of the time constant increases, the standard deviation as well as maximum and minimum values of the prediction errors decrease. However, the mean value of the prediction error remains almost unchanged. This is due to the fact that low- frequency components of the prediction error, which are main contributors to the mean value, are unaffected by low- pass filtering.

5.5 Online Implementation of the Thermocouple Validation Scheme

5.5.1 Introduction

In order to fully exploit the benefits of thermocouple validation scheme that has been developed for this project, the MonitorMV Online system has been employed to implement the condition monitoring application in real- time on the reheating furnace at the AvestaPolarit AB site in Avesta, Sweden.

Presently, the MonitorMV Online system application is installed on a computer AvestaPolarit AB site that is remote from the furnace control room. However, it is expected, in the future, to be installed in the control room of the reheating furnace as a valuable tool in addressing the reliability of the instrumentation equipment.

5.5.2 Layout of the MonitorMV Picture

A purpose MonitorMV Picture has been designed in order to properly present the results of the condition monitoring scheme.

The Primary screen that should be observed by operator personnel is available as Picture 1 in the MonitorMV Online system application. The background image of this picture is the physical diagram of the furnace. Additionally, filtered prediction errors of the thermocouples are placed at the relevant locations, as seen in Figure 12.

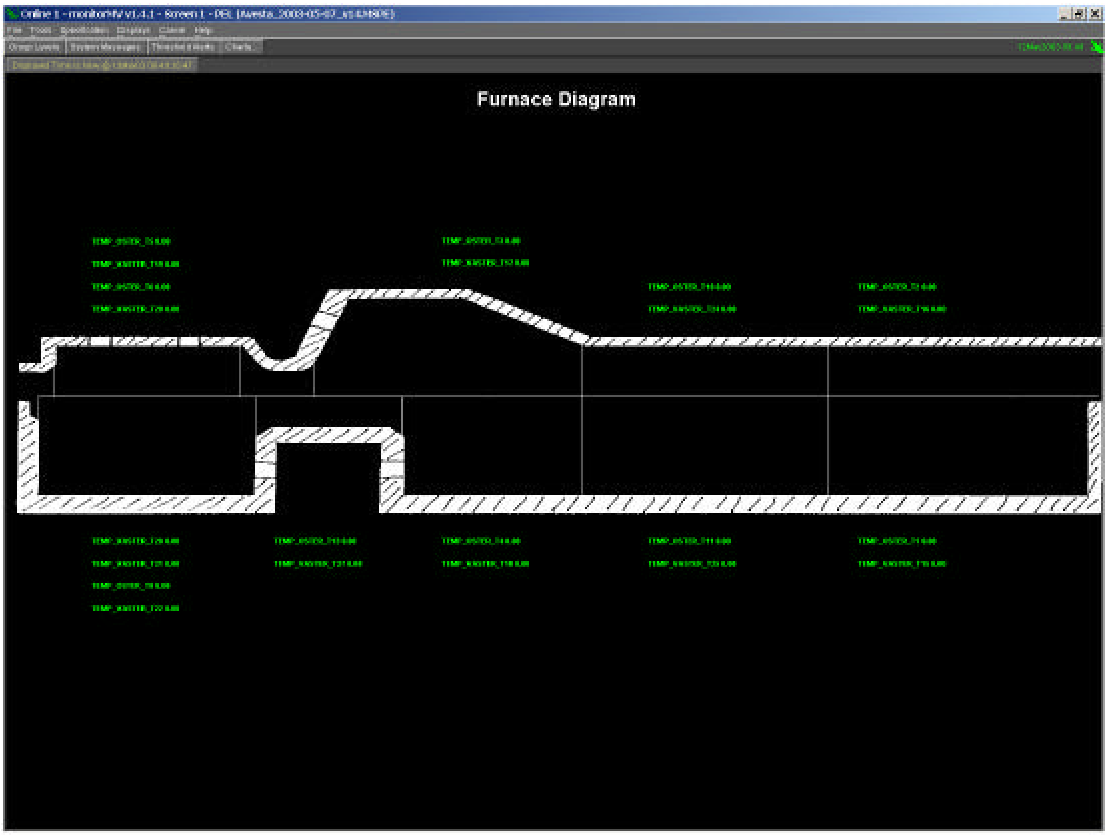


Figure 12

Colouring of these signals is used to indicate the status of a particular thermocouple. In the case where the relevant condition monitor is not active the corresponding filtered prediction errors are coloured blue. If the monitor is active and the alert level is not violated then a filtered prediction error is coloured green. Finally, if the alert level is violated then a corresponding filtered prediction error is coloured red. Displayed values represent the size of a filtered prediction error.

Hence, this picture represents the 'home page' of the application, to provide a clear display of the size and location of the malfunction.

5.5.3 Assignment of Alarm Levels and Filter Time Constants

The Alarm system, available within the MonitorMV Online system, has been applied to the filtered prediction errors of the PCA- based condition monitors. At the present, alarm levels have been set according to the maximum/minimum values of the filtered prediction errors, evaluated over the validating data set. In this way, it is believed that the number of false alarms would be small, to provide satisfactory confidence of the operation personnel in the robustness of the condition monitoring scheme. Depending on the future performance of the overall scheme these limits may be reduced from these somewhat conservative levels in order to increase the sensitivity of the condition monitors.

It was decided that the filter time constant be set, in the first instance, to 10 minutes for all of the signals. Such choice is seen as the compromise between the speed of the response to sudden and rapid changes in prediction errors and the reduction of sensitivity to a short- lived rapid disturbances that would otherwise unnecessarily trigger alarm.

Limits imposed on filtered prediction errors are given in the Table 15.

Signal ID	Positive Alert Level	Negative Alert Level
T1_Oster	5	-5
T2_Oster	85	-85
T3_Oster	20	-20
T4_Oster	30	-30
T5_Oster	10	-10
T6_Oster	15	-15
T7_Oster	30	-30
T8_Oster	20	-20
T10_Oster	5	-5
T11_Oster	25	-25
T13_Oster	35	-35
T14_Oster	20	-20
T15_Vaster	5	-5
T16_Vaster	80	-80
T17_Vaster	20	-20
T18_Vaster	30	-30
T19_Vaster	15	-15
T20_Vaster	10	-10
T21_Vaster	25	-25
T22_Vaster	25	-25
T24_Vaster	5	-5
T25_Vaster	40	-40
T27_Vaster	45	-45
T28_Vaster	20	-20

Table 14

5.6 Case Study

5.6.1 Introduction

This section demonstrates the capability of the MonitorMV system in detecting failure of the thermocouples. In particular, one thermocouple was reported to have failed during 14th of February 2003. However, this malfunction was not observed by operators and was spotted by a process engineer on the 16th February. Note that the system discussed in this report is actually implemented on a different reheating furnace at AvestaPolarit. Nevertheless, this case study has been included in the report as a demonstration of the system capability to detect subtle and non- trivial abnormalities of the instrumentation equipment.

5.6.2 Process Data

Process data that has been utilised for the development of the condition monitoring scheme was collected during January 2003. Portions of the data for which the measurement values were invalid and periods during which the furnace maintenance was performed were not included. Overall, 22,961 data points have been used for the training purposes.

Variables that were considered in this development are all of the thermocouple temperature measurements inside reheating furnace B (27 in total).

5.6.3 Development of a Condition Monitor

12 highly cross-correlated Measured signals have been included in the condition monitor. Following Principal Component Analysis 4 principal components have been chosen for the PCA model, contributing 97.6% to the total variance of the training data set, as seen Figure 13.

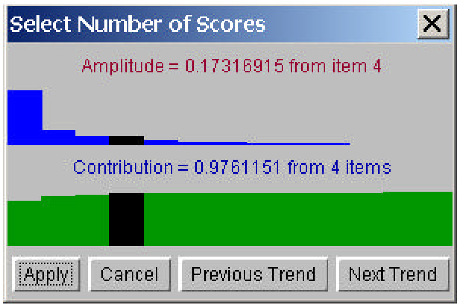


Figure 13

5.6.4 Validation of the Statistical Model

Data from the period between the 1st February and 12th February 2003 was used to assess the validity of the developed PCA model.

Inspection of the squared prediction error (SPE), displayed as the top trend in Figure 14, shows there to be visible periods during which SPE does have excessive values, i.e. it is coloured red, indicating that the model is not fully able to generalise to the situations, which were not present in its training data set. Note that SPE represents the sum of the squares of all the prediction errors associated with each variable considered by a corresponding PCA monitor.

However, it can be argued that the periods of excursion are not typically long enough, especially when compared with length of the periods during which SPE remains at the low level.

Nevertheless, attention should be paid towards developing more accurate models so as to reduce the occurrence of false alarms. Also, development of accurate models would increase sensitivity to the abnormalities and allow validity to be extended over longer periods of time, i.e. reduce the number of re- modelling exercises.

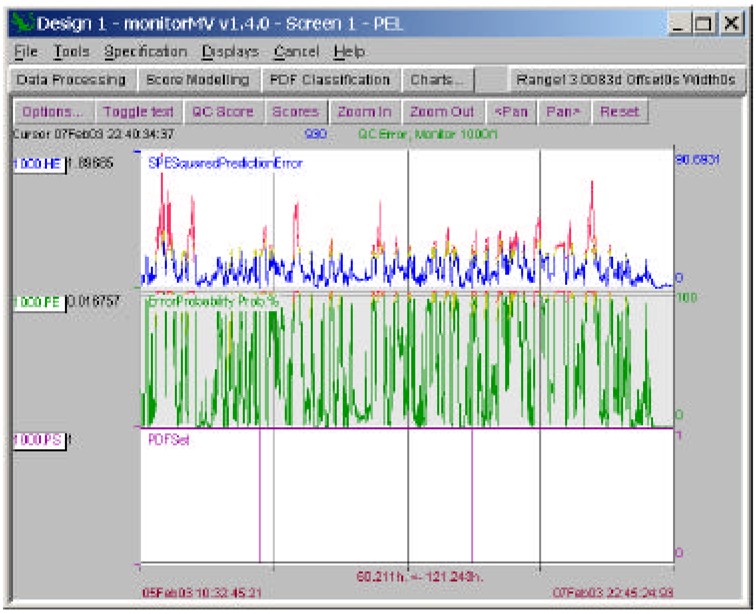


Figure 14

5.6.5 Data Analysis During the 13th of February 2003

In this section the analysis of data during 13th of February is reported. Results are presented in terms of the standard charts/ trends used in MonitorMV for fault detection purposes. These are the SPE plot, explained in the previous section, individual prediction trends and the Contribution chart. The Contribution chart represents scaled prediction errors, i.e. scaled with respect to the training data, of the individual measured signals displayed graphically as a histogram.

Figure 15 shows the associated SPE trend and it is apparent that while the trend of SPE is not consistently high there are a number of relatively long excursions during the 13th of February. Also, the duration of these individual excursions increases towards the end of the day.

Hence there is an indication through the SPE chart that the inter- variable relationship between thermocouples has changed to some extent. This could generally be attributed to a number of possible causes. One possible cause is that the underlying thermal dynamics of the reheating furnace have changed (due to a premeditated effort to improve performance or to a fault that has occurred inside the process). Another possible cause is that one of the measurements is not valid, i.e. that the sensor responsible for measuring a particular signal is malfunctioning. Both of these events are consequential in terms of overall control system performance and should be detected as early as possible using a condition monitoring scheme such as the one implemented in MonitorMV.

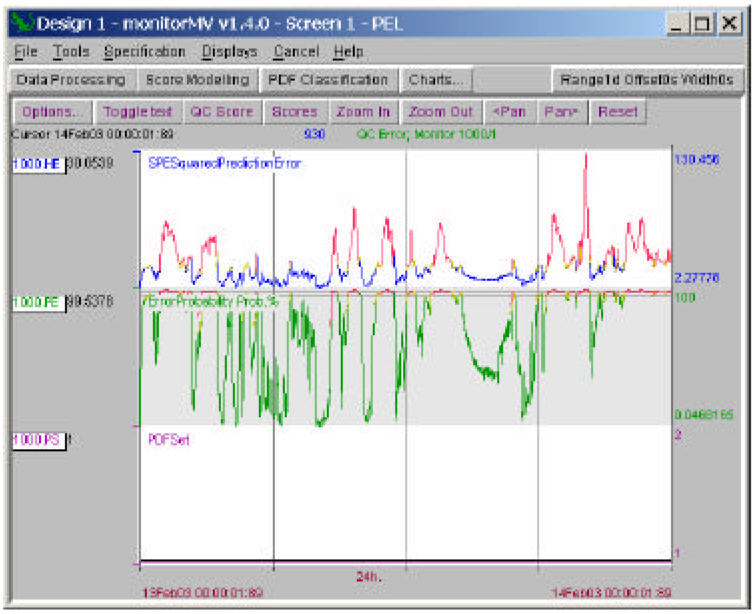


Figure 15

Inspection of the Error Contribution bar chart shows that all of the variables remain for most of the time within 5 standard deviations of the training data error. However, it is observed that from 14:00 (approximately) the signal associated with the failed thermocouple temperature measurement (with a MonitorMV's signal ID of 2034.ME) has a largest although not distinguishable error contribution as seen in Figure 16. Note that MonitorMV signal ID is presented on most of the trends instead of the signal tag name Temp_FOCS_9. However, in all explanations in this report, the signal tag name is used instead of the MonitorMV signal ID number.

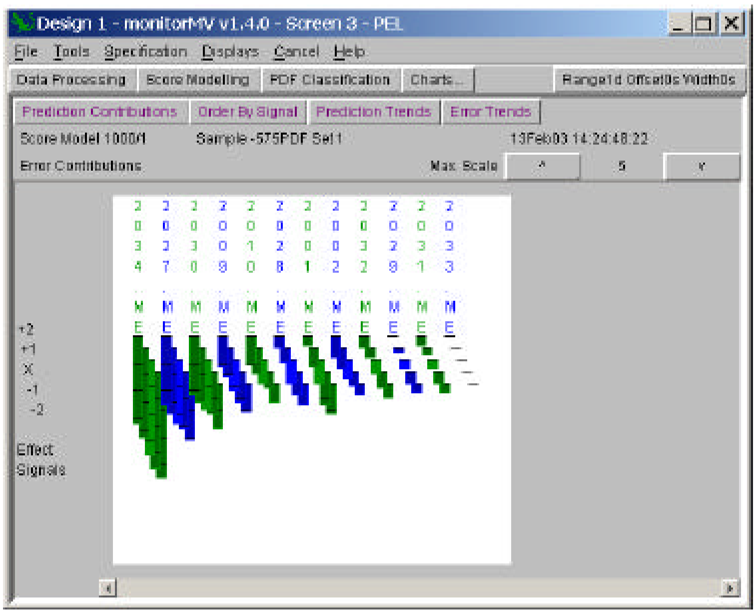


Figure 16

The failed thermocouple temperature measurement crosses the 5 standard deviations threshold in the error contribution bar chart at 22:30, as seen in Figure 17. Since the other error contributions are not as significant, the error contribution bar chart does provide a clear indication that this particular thermocouple may be faulty.

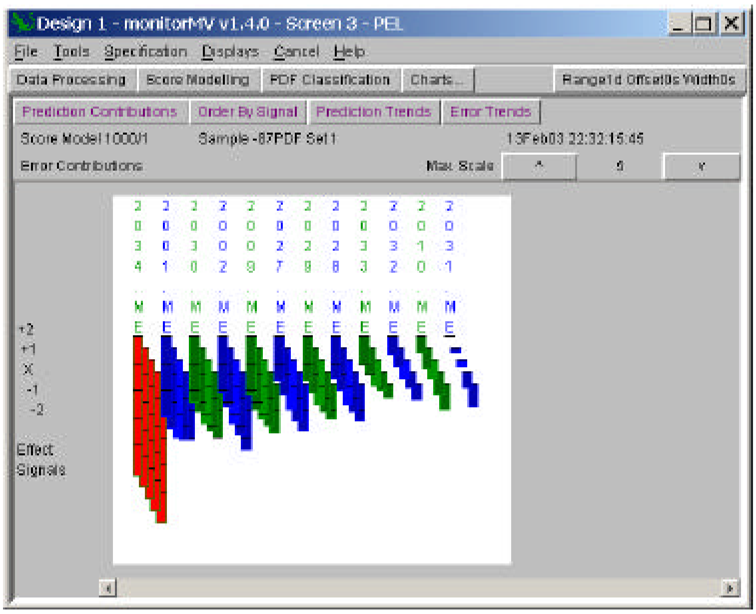


Figure 17

The actual PCA prediction error for Temp_FOCS_9 at this point is equal to –60 degrees C. In other words, the prediction trend indicates that the measurement of Temp_FOCS_9 is 60 degrees lower than expected, as seen in Figure 18, where the blue line represents the measurement while the brown line represents the PCA prediction of Temp_FOCS_9.

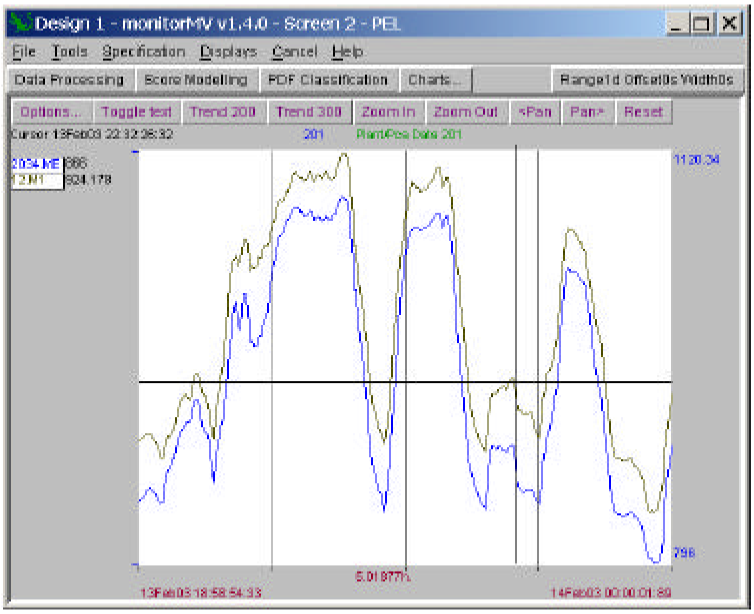


Figure 18

Following 13th February, the SPE is consistently high, indicating an underlying problem. Out of all the variables, Temp_FOCS_9 is the most likely source of the problem as its prediction error is consistently below –25 degrees C. In fact the mean of the prediction error for Temp_FOCS_9 over the last 2 hours of 13th of February is equal to -43.7 degrees C. Note that in normal circumstances the mean of any prediction error should be equal to 0.

Hence, there is a very strong indication that thermocouple associated with Temp_FOCS_9 signal has failed during the later part of the 13th of February.

5.6.6 Data Analysis During the 14th of February 2003

The SPE plot of Figure 19 shows that the deterioration of the thermocouple measurement is exposed by a continuous increase in the overall size of prediction error during 14th February.

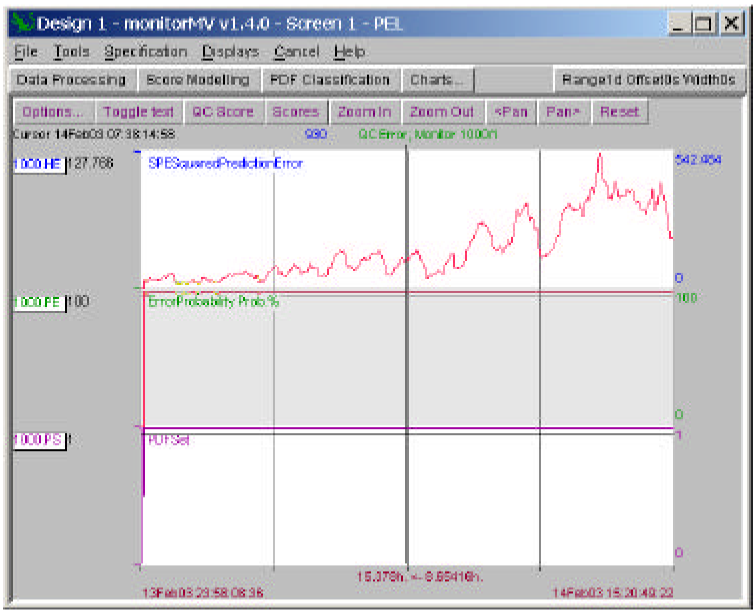


Figure 19

Note that, during the first 4 hours of the day, the mean of the prediction error for Temp_FOCS_9 is equal to –47.12 degrees C. However, during the next 6 hours (from 4:00 until 10:00) the mean of the prediction error is equal to –78.7 degrees C, indicating a persistent and increasing problem with the associated thermocouple.

Deterioration of the faulty thermocouple can be observed in Figure 20 where the PCA prediction of Temp_FOCS_9 (brown line) is plotted alongside the Temp_FOCS_9 signal itself (blue line).

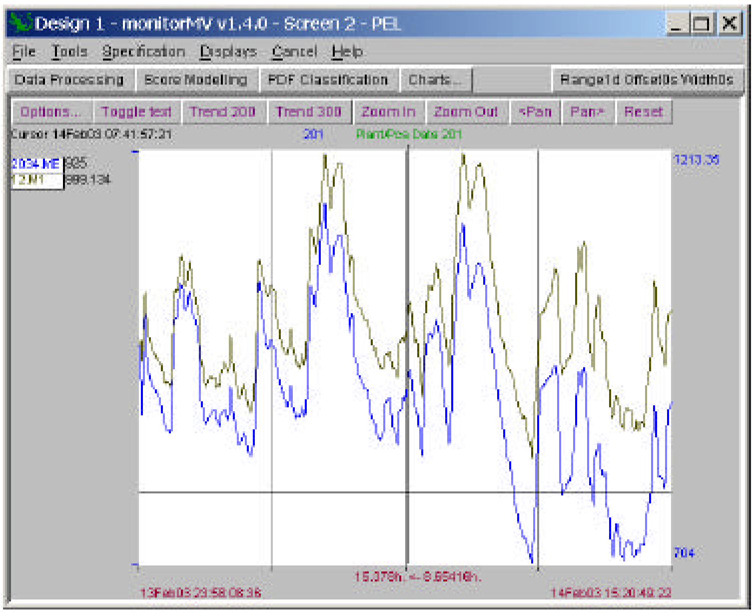


Figure 20

Notice that the vertical distance between brown and blue line, indicating the prediction error, is increasing with time. This is more clearly observed in Figure 21 where the actual prediction error is plotted. Notice the consistent drift of prediction error in the negative direction, which confirms the findings from other monitoring charts that the fault has indeed occurred.

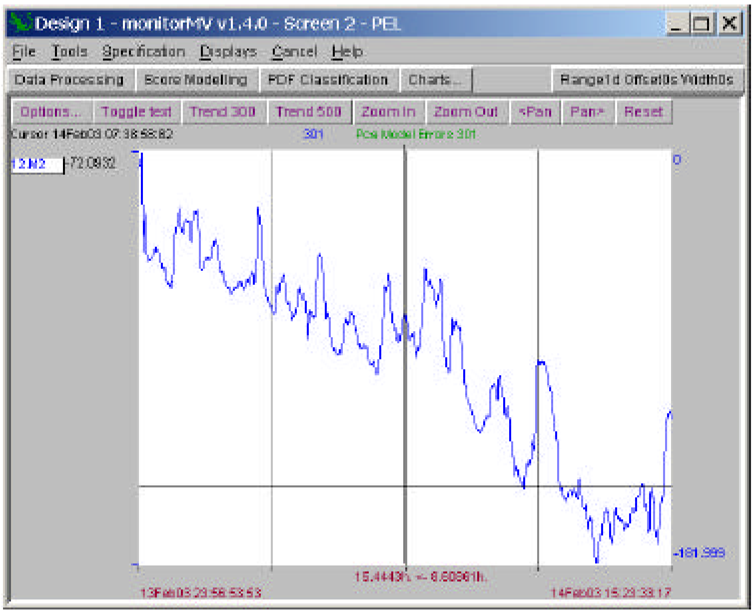


Figure 21

From 2:30 (approximately) onwards Temp_FOCS_9 is clearly and consistently the highest prediction error contributor as seen in Figure 22.

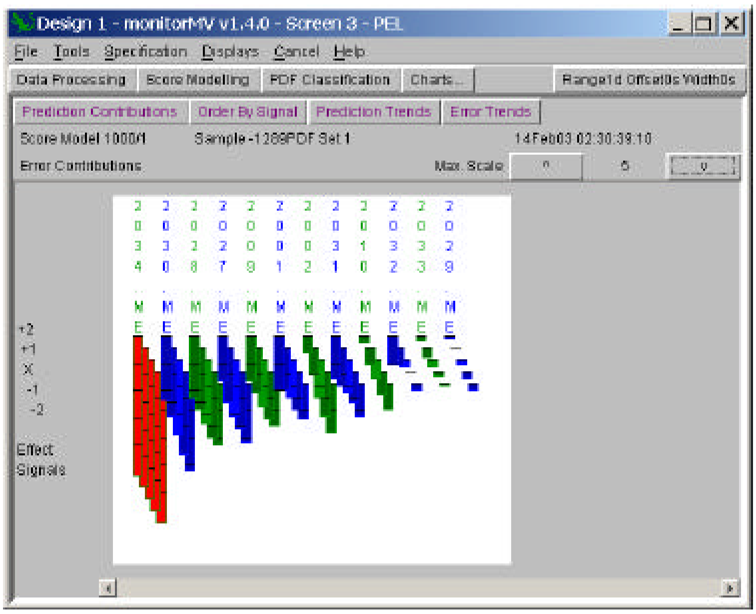


Figure 22

5.6.7 Conclusion

In conclusion, this case study demonstrates that MonitorMV has indicated the existence of a problem as early as the later part of the 13th of February and also clearly identified the failing thermocouple in the early hours of the 14th of February.

Note that these observations have been made using standard techniques available within MonitorMV (SPE chart, prediction trends and contribution bar chart) demonstrating its fault detection capabilities.

However, further improvements in terms of the fault sensitivity as well as the robustness of the condition monitor could be achieved by filtering prediction errors and by means of a careful design of a Threshold Detector system available within the MonitorMV Online system, as described in earlier sections of this chapter.

5.7 Impact of the Faulty Thermocouples on the Energy Consumption

As is already stated in section 5.1, in the case where reheating furnace thermocouples are having negative bias, i.e. the measured value is lower than the actual temperature, the result is over- heating of the slabs. In such case energy consumption of the overall reheating furnace will increase.

In this section an estimate of the excess energy consumption is made for the reheating furnace U302 at the SSAB site in Borlange, Sweden, for which the basic diagram and specifications are given in Figure 4.

The calculations presented in this section develop relationships between increased discharge temperature of the slabs and the energy consumption.

Firstly, consider the following equation for the calculation of the heat.

$$Q = m \cdot c_p \cdot \Delta T \qquad (16)$$

where Q is the amount of heat required to raise the temperature by ΔT degrees Kelvin of the body with mass equal to m kilograms, made of material with a specific heat capacity given by c_p. In the case of steel slabs, c_p for over $1100^\circ C$ is equal to 628 J/kgK.

So, in order to raise temperature of the 1 tonne of steel slab by $1^\circ C$, over its nominal temperature, it is required to use $Q = 10^3 kg \cdot 628 J/kgK \cdot 1^\circ C = 628 kJ = 0.174 kWh$.

Since the heat efficiency of U302 is approximately equal to 58%, it means that the heat required to be delivered by the furnace burners' system is equal to $0.3kWh$ for each tonne of steel slab to be overheated by 1 degree centigrade.

Hence, if the slabs are, for example, overheated by 50 degrees centigrade and such behaviour persists for a day in a furnace with an annual throughput of 1,300,000 tonnes, then the amount of excessive energy consumption in that single day is equal to 53.42MWh. If such problem persisted for a year then a total amount of excessive energy consumption would reach 19.5GWh.

5.8 Summary

The prerequisite for satisfactory control system performance is reliable availability of feedback measurements. In fact, a control system can only be as reliable as its feedback measurement equipment. Also, general sensor equipment is susceptible to long-term drifts and sudden failures. As a result, key business drivers, such as running cost, productivity and product quality, can be adversely affected, compromising the economic viability of the entire processing plant. Therefore, the improvement in feedback measurement reliability has a direct and positive impact on key business drivers in any manufacturing industry.

The Fuel Optimisation Control System (FOCS) scheme, employed in the reheating furnaces of the hot-strip rolling mills, relies heavily on accurate temperature measurements inside different furnace zones. These thermocouple-measured temperatures are used as measurements in local PID- based control schemes (one controller for each furnace zone) as well as for the initial conditions in slab temperature calculations.

The impact of the faulty or erroneous temperature measurements in reheating furnace is twofold. In the case where the measured value is below the actual temperature, excessive fuel is used in the furnace burners. This in turn increases the energy consumption, which is probably the main business driver for this process. Also, if the measured value is above the actual temperature then the product quality may be degraded. Hence, in either case an important business driver is adversely affected by the failure of instrumentation equipment to provide accurate and reliable measurement.

The sub- project that is described in this chapter is concerned with a development of a validating mechanism for the thermocouples used in reheating furnaces. The approach, taken in this project, is based on the principle of redundancy through the utilisation of the Principal Component Analysis method, available within the MonitorMV software system. The presence of strong cross-correlation between different thermocouple measurements is exploited for the detection of faulty sensors and for the subsequent estimation of the true value of the associated process variable. The developed validation scheme is presently under trial on one of the reheating furnaces at the AvestaPolarit site in Avesta, Sweden.

Multivariate statistical process control is a data driven technology. Therefore, the data that is used for the development of statistical models has a direct impact on the performance of the resulting condition monitor. Ideally, training data sets should be chosen to correspond to those periods of time that follow immediately after re-calibration takes place. In this way, it is ensured that a training data set does not contain variables with systematic error.

Due to the character of operation and the geometric shape of the reheating furnace, thermocouple measurements from different zones have been found to be not highly correlated. For example, temperature measurements from the preheating zones are almost completely uncorrelated with temperature measurements from the soaking zones.

In order to improve accuracy of the fault detection/diagnosis scheme it was decided to design three condition monitors that would focus on different sections of the reheating furnace. The criterion for grouping of thermocouples for each condition monitor has been taken to be the level of cross-correlation between these measurements as well as their mutual closeness in physical sense.

The key challenge in this particular application has been the lack of the strong cross-correlation between various thermocouple measurement signals. As a result, prediction errors of the PCA- based statistical model that are routinely encountered are of comparable size to the measurement errors which have consequential impact on the key business drivers, such as the energy consumption (in the case of the negative prediction error) as well as the product quality (in the case of the positive prediction error). Amongst many different methods of reducing sensitivity and, thereby increasing the robustness of the monitoring scheme it has been decided in this particular project to reduce the number of false alarms by limiting attention to those events that contain frequency components in a specific range. For example, if the event that is to be detected is the slow drift, representing a general low- frequency signal, then by low- pass filtering the information such as prediction error it is possible to solely focus on all those events that belong to this very specific band of frequencies. In that case sudden and short-lived disturbances, observed in prediction error trends, are ignored during the filtering process while the slow disturbances are emphasised. Such focus on particular features in the data has been achieved by low-pass filtering the prediction errors of the PCA models. The tuning parameter has been chosen to be the time constant of these filters.

In order to fully exploit benefits of thermocouple validation scheme that has been developed, the MonitorMV Online system has been employed to implement this condition monitoring application in real- time on the reheating furnace A at the AvestaPolarit AB site.

The presented case study demonstrates the capability of the MonitorMV system to detect failure of thermocouples used in reheating furnaces of the hot- strip rolling mill. The situation that is considered took place at the AvestaPolarit plant. One thermocouple was reported to have failed during 14th February 2003. It has been demonstrated that MonitorMV indicated the existence of a problem as early as the latter part of 13th February and also clearly identified the failing thermocouple in the early hours of 14th February. Note that these observations have been made using standard techniques available within MonitorMV (QC charts, prediction trends and contribution bar chart) demonstrating its fault detection capabilities.

Future developments should focus on placing the Online application into the control room of the reheating furnace A, thereby providing accurate and reliable validation of thermocouple measurements that are continuously available to the process operators and the process control systems. Also, a similar scheme should be employed for the reheating furnace B, for which the capability of MonitorMV system to detect faulty thermocouple has been demonstrated in the case study, results of which are given in section 5.6. Furthermore, the concepts that are employed in validation of thermocouples could be employed in the future for other instrumentation equipment that exhibits high levels of cross- correlation. . This would ensure that critical sensors are backed up and monitored automatically, thereby shielding the control system from misinformation and potentially costly maloperation.

6. Sub-project 2: Development of a NOx Estimation Scheme using Partial Least Squares

6.1 Introduction

Nitrogen oxides (NOx) are generated from the combustion process directly by the thermal oxidation of gaseous nitrogen by oxygen (thermal NOx) or by combination of nitrogen compounds in the fuel with oxygen. In the case of reheating furnaces, NOx is produced as a result of burning fuel while reheating steel slabs.

The need to protect the environment from combustion generated emissions, such as carbon dioxide (CO2) and nitrogen oxides (NOx) has led in recent years to considerable demand for improved combustion system design and operation. In terms of the improved combustion system design, there is a number of new burner technologies that aim to minimise NOx emissions, such as air staging or two- stage combustion and pressure atomised oil burners technology. In terms of the improved combustion system operation, advanced process control technology is expected to shed light into methods of reducing NOx while avoiding costly modifications to the actual process equipment and/ or compromise in process performance.

The most important business drivers in economic considerations of the reheating furnace are minimisation of energy consumption and maintenance of high throughput, i.e. extraction rate of the steel slabs. However, it is evident that, with increasingly stringent environmental regulations and heavy penalties for non-conformance, especially since the ratification of Kyoto agreement, furnace emissions are becoming a significant cost driver and may become the most important constraint in coming years. Such environmental considerations are forcing process plants to measure emissions and investigate methods for their cost- effective reduction.

In order to provide continuous measurements of NOx emissions, expensive analysers have to be installed and maintained. On the other hand, software based inference engines, known as ‘soft sensors, may well provide a viable and economic alternative to costly hardware- based analysers. Furthermore, if a developed inference engine is given in appropriate form, i.e. cause- effect structure with relatively simple parameterisation, then a by- product of the soft sensor development is a delivery of prediction model that can be readily utilised in improved closed- loop control of NOx emissions.

Hence, the development of an accurate prediction model for NOx emission is seen as a crucial step in pursuing development of a soft sensor application and/ or implementation of adequate control scheme. Note that such a control scheme can be delivered as either automatic regulation of NOx emissions, i.e. implemented in closed- loop form, or as a non- invasive advisory application, for which the loop is broken at the controller output.

Additionally, sudden and rapid change in terms of NOx emissions that is not accounted for by the developed prediction model may be a symptom of an operational problem of the reheating furnace. Such issue was not covered in this project.

However, development of prediction model clearly benefits attempt to develop the condition-monitoring scheme of the reheating furnace.

During this particular project, the NOx estimation scheme has been developed, by using the MonitorMV package, for a reheating furnace U302 in the hot- strip rolling mill at the Swedish Steel AB (SSAB) site in Borlange, Sweden. Specifications and the diagram of this reheating furnace are presented in Figure 4. The Project has been carried out in collaboration with process control engineers of SSAB and APC Ltd.

6.2 Control of NOx Emissions

Reduction of NOx may be achieved by either primary or secondary means. Primary reduction of the generated NOx takes place in the furnace itself, usually by improved control and/ or modification of the combustion process. Secondary reduction is performed by removing the NOx from the exhaust after leaving the furnace. Secondary reduction can be done by means of ammonia injection, either non-catalytic or catalytic, or by means of flue gas recirculation. Generally, the cost of secondary reduction will depend on the amount of NOx to be removed, therefore it is desirable to have the lowest initial concentration possible to minimise operating costs. Hence 'Low NOx' burners and optimal NOx control are seen as primary means and are desirable even if secondary reduction must be used.

Low NOx burners are generally designed to control fuel and air mixing at each burner in order to create larger and more branched flames. Peak temperature is thereby reduced, and results in less NOx formation. Additional benefit of this type of burners is that the improved flame structure also reduces the amount of oxygen available in the hottest part of the flame, thus improving burner efficiency. Note that 'low NOx' burners can be combined with other primary measures, such as optimal control, in order to minimise excessive NOx emissions and the cost of the secondary NOx reduction methods.

Control of NOx emissions can be achieved by modifying the operating conditions of the burners and the entire reheating furnace. Implementation would, in principle, take the form of an automatic feedback- based control scheme. In this case the model would need to be developed that relates NOx emissions to its main causes, such as burners' fuel and air flowrates, as well as the main business drivers of the reheating furnace, notably productivity and the energy consumption. Then, the control objective could be stated in mathematical optimisation framework as an attempt to minimise NOx emissions while maintaining high productivity and minimising energy consumption.

Hence, the first step in the development of an optimisation scheme, implemented as either a closed- loop or advisory/open- loop solution, is the development of an accurate cause- effect model that relates important issues within a process. In this case, it is the model between the key NOx producing cause variables and the NOx emissions.

6.3 NOx Estimation as a Soft Sensor Application

Soft sensor is a common name for a software-based inference of difficult-to-measure process or product quality variables. For example, some attributes of manufactured products such as polymer melt index, moisture content of food, and resistance to thermal flow in insulation can only be measured by laboratory analysis. With soft sensors these measurements can be made continuous and available on the manufacturing line. Hence, the control of product quality can be increased significantly without large capital cost of installing expensive analysers. Inferential sensors can also be used in conjunction with analysers for redundancy purposes. More specifically, if the prediction model employed by a soft sensor is adequately accurate it can be used to detect instrument failure or systematic error and, therefore, highlight the need for instrument repair/ recalibration. Furthermore, during the maintenance of the analyser, value of the important quality variable is continuously made available through the inferential capability of the prediction model. However, while soft sensors have been applied in the form of so- called predictive emission monitors (PEMs) in the utility boilers and crude oil furnaces, to name but a few, there is very little evidence of their application in the reheating furnaces of a hot strip rolling mill.

6.4 Choosing Cause Variables

In order to develop accurate prediction model important decision in the early stages of model design is the selection of a set of cause (input) variables. Such decision is made by employing process knowledge as well as some statistical analysis methods, notably correlation analysis. In this particular case, process engineers from SSAB and APC Ltd. have provided extensive process knowledge that greatly helped decide on which variables are to be considered as cause signals in the corresponding prediction model.

In the initial development, correlation analysis was hampered by the lack of process excitation and irregular measurements of NOx emissions. Hence, the process knowledge provided a crucial insight into the underlying cause- effect structure of the model to be developed. In particular, impact on the NOx emissions by the variables that are related to the first two zones (preheating zones) of the furnace was highlighted by APC Ltd. However, some additional process variables were also found to have significant impact on NOx emissions, namely flow rate of combustion air in recuperator, total flow rate of oil into the burners as well as the total flow rate of the atomising steam. This last variable is found to be negatively correlated with NOx emissions. In other words, increase in the total flow rate of atomising steam into the burners is found to reduce NOx emissions. Hence, the total flow rate of atomising steam could be seen as a crucial cause variable in any attempt to minimise NOx emissions.

The complete list of cause variables that were included in the model, together with their brief description, are given in Table 15. Note that the most important cause variables from these two zones are the flow rates of air and fuel into the burners as well as the zone temperatures.

Signal ID	Description
9.ME	Temperature in the south of zone 1 used by PID regulator
10.ME	Temperature in the north of zone 1 used by PID regulator
11.ME	Temperature in the south of zone 1 used by FOCS control system
12.ME	Temperature in the north of zone 1 used by FOCS control system
13.ME	Temperature in the south of zone 2 used by PID regulator
14.ME	Temperature in the north of zone 2 used by PID regulator
15.ME	Temperature in the south of zone 2 used by FOCS control system
16.ME	Temperature in the north of zone 2 used by FOCS control system
63.ME	Combustion air temperatuire for zone 1
64.ME	Combustion air temperatuire for zone 2
73.ME	Flowrate of combustion air to the burners in zone 1
74.ME	Flowrate of oil to the burners in zone 1
75.ME	Flowrate of combustion air to the burners in zone 2
76.ME	Flowrate of oil to the burners in zone 2
96.ME	Furnace pressure
97.ME	Flowrate of combustion air in the south sector of the furnace
98.ME	Flowrate of combustion air in the north sector of the furnace
99.ME	Total flowrate of oil
100.ME	Total flowrate of atomising steam
101.ME	Pressure of the atomising steam in zone 1
102.ME	Pressure of the atomising steam in zone 2
121.ME	Walking beam cover ratio for zones 1 and 2
129.ME	Average distance of the slabs to the furnace wall in the south sectors of zones 1 and 2
137.ME	Average distance of the slabs to the furnace wall in the north sectors of zones 1 and 2
146.ME	Air/Fuel ratio for zone 1
147.ME	Air/Fuel ratio for zone 2
148.ME	Status of extraction door (open=1, closed=0)

Table15

6.5 Training Data

Data collection is an essential part of the overall prediction model development since quality data are the only base for building a quality prediction model. In particular, if a prediction model is to be of dynamic form, as it is the case with NOx predictor, then a training data set has to be 'sufficiently excited' in order to reveal information concerning dynamic relationships between cause and effect variables.

In order to sufficiently excite the process it has been decided to perform numerous step tests on some key variables that are believed to be major contributors to the NOx emissions. These have been decided to be the flow rates of air and fuel into the burners of the first two zones of the furnace.

Fuel flow rates have been varied between 90% and 70% of the burner capacity while the air flow rates have been manipulated by changing the air to fuel ratios. During the step tests, automatic temperature control systems (PID controllers) for the first two zones of the reheating furnace were set to manual status and their outputs, i.e. fuel and air flow rates, were directly manipulated. Sample displays of these variables during step- tests are given in Figures 23 and 24.

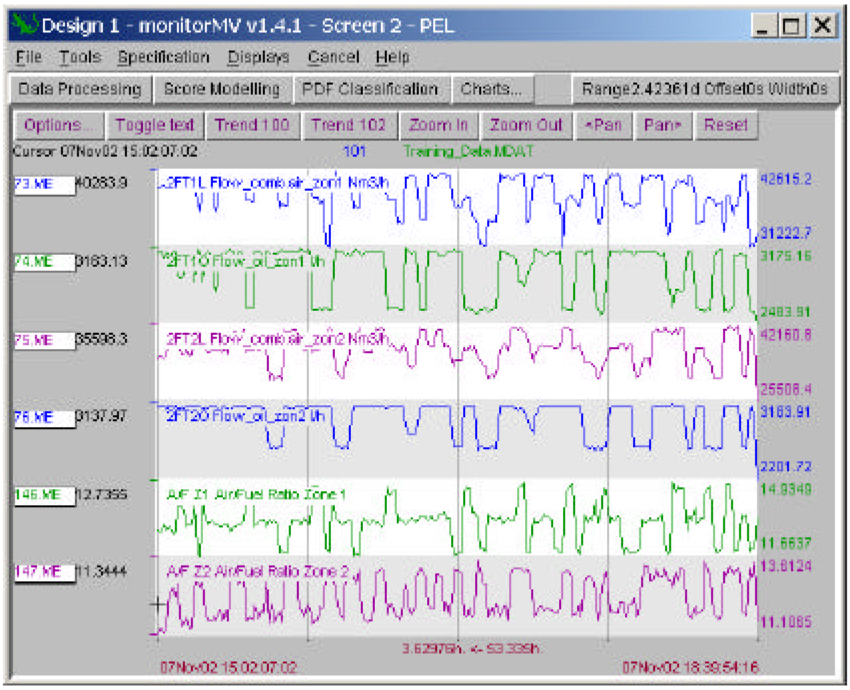


Figure 23

Figure 24

6.6 Dynamic Model Structure

The fundamental principle behind a general discipline of system identification is to develop accurate models of a plant for condition monitoring and control engineering purposes. Development of a process model generally consists of two stages. The first stage is concerned with model structure, whereby the parameterisation of the relationships that exist between various process variables is performed. During the second stage values of parameters, given within a model structure, are estimated using identification technique, such as Partial Least Squares (PLS). In this section two most widely known model structures are introduced.

Models are generally considered, identified and implemented in sampled- data framework, which for a single input, single output linear system have the following general form:

$$\hat{y}(t) = a_1 y(t-1) + a_2 y(t-2) + ... + a_n y(t-n) + b_1 u(t-d) + ... + b_m u(t-d-m+1) + e(t) \tag{17}$$

This equation is referred to as an 'Auto-Regressive with eXogeneous variable' (ARX) model and subsumes the well known model termed a 'Finite Impulse Response' (FIR) model, which has the following structure:

$$\hat{y}(t) = b_1 u(t-d) + b_2 u(t-d-1) + ... + b_m u(t-d-m+1) + e(t) \tag{18}$$

In these two equations, $y(t)$ is the output measurement at time t and a_1, a_2, b_1, b_2 etc, are parameters that are related to the dynamics of the system with n corresponding to the order of the system. Note that $\hat{y}(t)$ is the value of the effect variable that is predicted by the model at time t. This predicted value will differ from the actual measured value of the effect variable, $y(t)$, by an amount $e(t)$, which is termed prediction error. Finally, d is the time delay of the system (in samples).

In terms of the MonitorMV terminology, minimum delay is given by d while the maximum delay is given by $d+m-1$ in both equations (17) and (18). Order of dynamics is specified by n in equation (17). Hence, for example, if maximum delay is equal to minimum delay and the order of dynamics is set to 0, then a resulting model describes static relationship between cause and effect variable.

Many industrial control engineering technologies restrict themselves to the use of the FIR model format. The reasons for this are twofold:

- An engineer is able inspect the pattern of the FIR coefficients to gain a feel for the time constants and gains of the process and to also inspect the accuracy of the model.
- There is no need to be concerned with the selection of the 'order' of a transition matrix and the basis for selection of the number of terms in the driving as well as measurement matrix is clear, as indicated above.

The large model structures imposed by the FIR model format, given in equation (18) do introduce a significant computational burden in solving for control moves, but this is of little consequence except for very large systems, given the state of today's low cost and high performance computer power. However, such structures do create problems for statistical identification methods because of the large number of parameters that have to be determined.

Although the ARX model form does appear to offer some important advantages over the FIR structure it does also have some limitations. The accurate prediction for such a model is dependent upon good reflection of dynamics within the sampled history of the effect variables. Should these signals have significant levels of noise superimposed upon them or be irregularly measured, as in the case of NOx, then the ability of the model to accurately predict can be compromised.

6.7 Development of the NOx Prediction Model

The NOx prediction model was developed by using Partial Least Squares (PLS) approach, available within MonitorMV package and described in section 3.2. Due to the irregular measurements of the NOx emissions, model was decided to be of the FIR (finite impulse response) structure, see equation (18), as opposed to ARX, see equation (17).

The minimum delay was set to 0 minutes, while the maximum delay was set to 3 minutes for each cause signal and the sampling interval was set to 1 minute.

The identified model has been identified with 30 (out of possible 109) scores, contributing 75.8% to the total variation of the training data, as seen in Figure 25. Such choice of a number of scores is made in terms of a compromise between the accuracy of the model and its robustness in dealing with highly cross-correlated process variables. In this particular case, choosing 30 scores allows reasonable predictability, as shown later on in this section, while dealing with co-linearity present among cause variables, particularly thermocouple measurements of the zone temperatures.

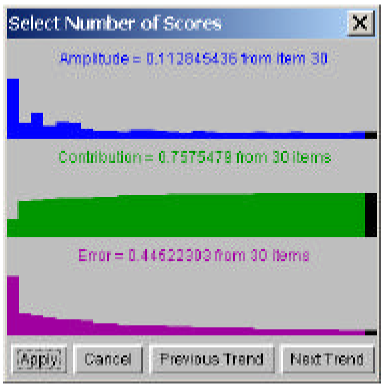


Figure 25

Prediction of the model over the training data set is displayed in Figures 26, 27 and 28.

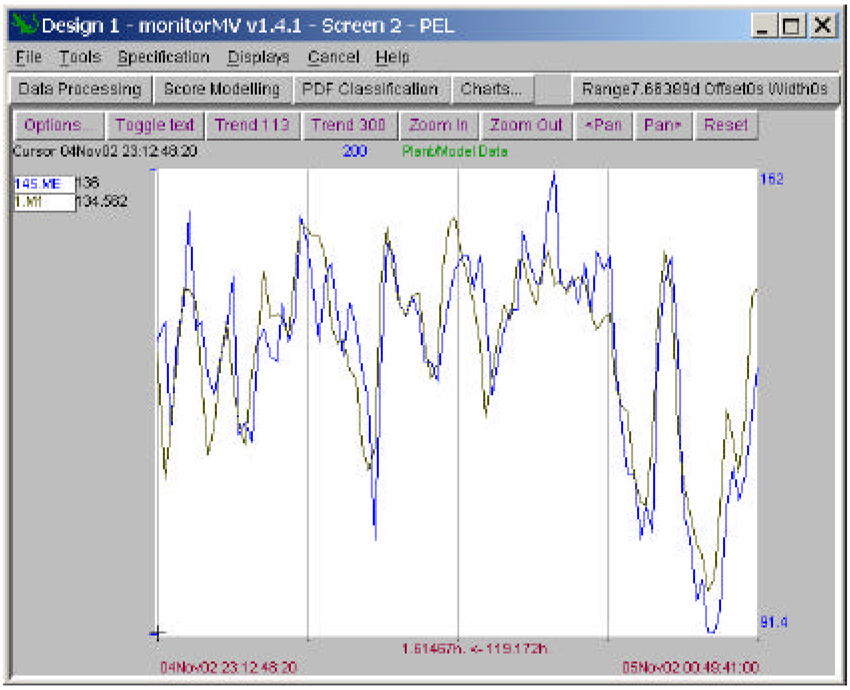


Figure 26

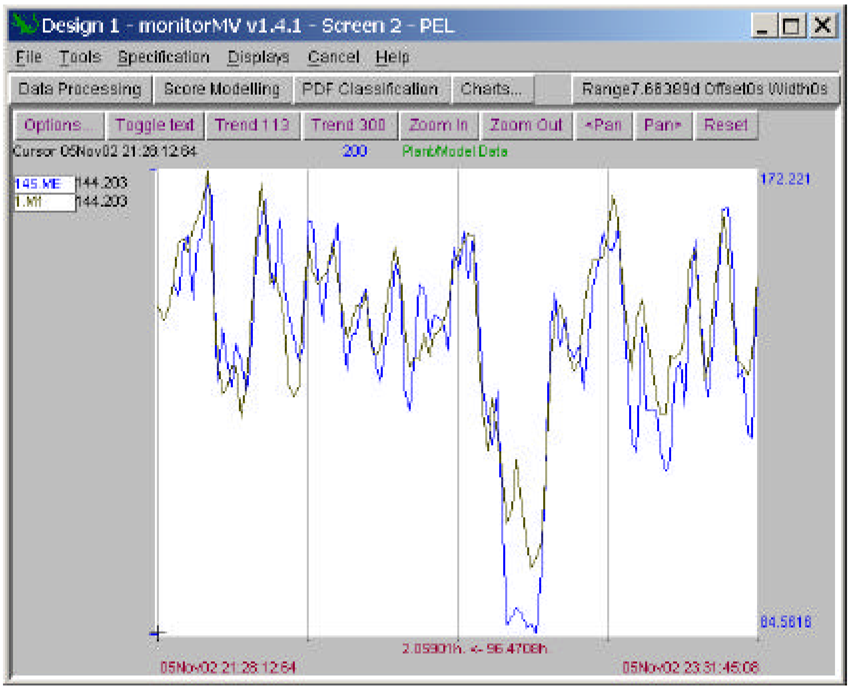


Figure 27

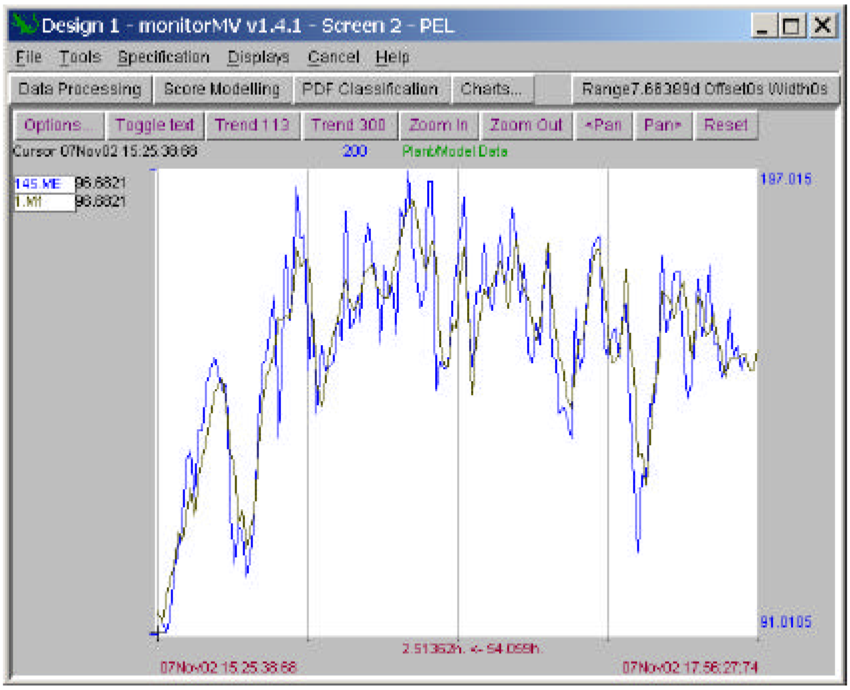


Figure 28

The developed prediction model has been validated on the data set, which was not used in identification but did belong to the period when step- tests were performed. Predictions of the model over this validating data set are displayed in Figures 29 and 30.

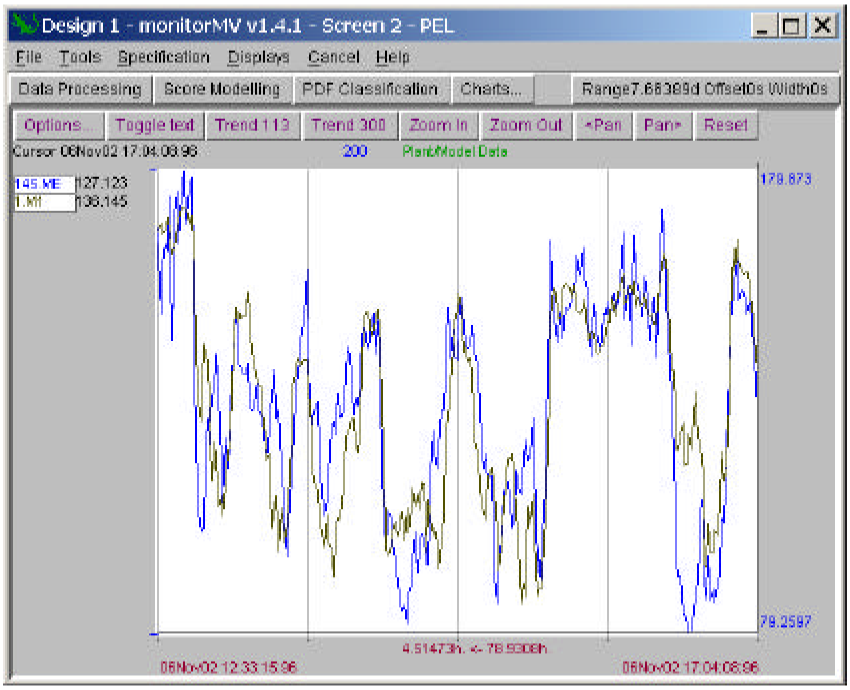


Figure 29

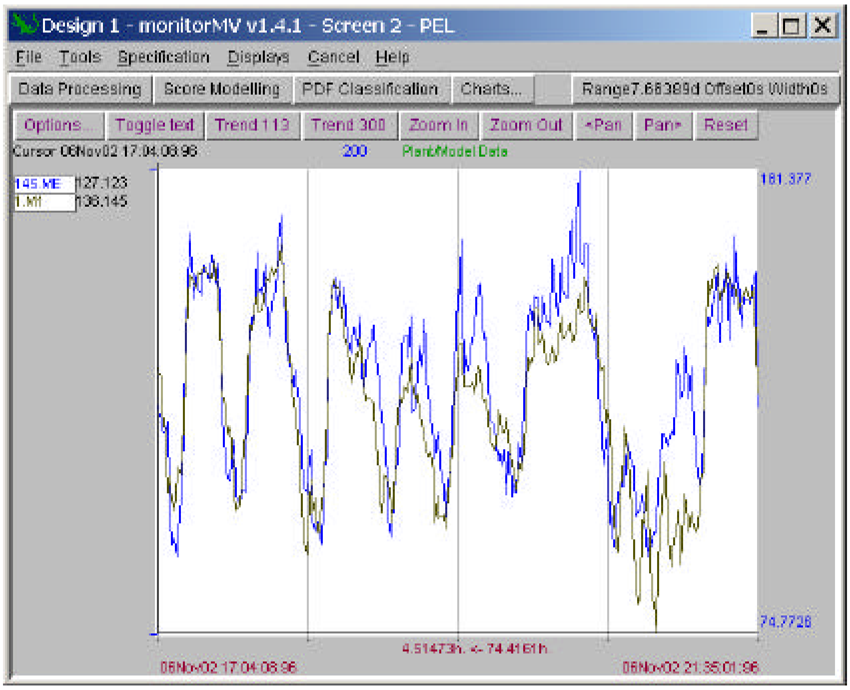


Figure 30

It is observed that the model was able to generalise to the data set that was not used in training. Hence, the conclusion has been made that the prediction model achieved a satisfactory level of accuracy.

Validation has also been performed using the data from January, February, March and April 2003. Statistical information about the prediction error during these periods is presented in Table 16.

Validation Period	Mean	Deviation	Maximum	Minimum
January 2003	24.04	18.2	91.83	-29.45
February 2003	26.35	20.63	89.02	-38.15
March 2003	23.24	21.33	97.09	-81.16
April 2003	5.67	19.96	74.2	-50.55

Table 16

Note that it is the mean value of the prediction error that has dramatically changed during these four months of the validation period. However, standard deviation is also significant indicating that not all of the dynamics present in the process have been depicted by the prediction model. Hence, some form of model adaptation is needed, particularly in order to account for the time- varying nature of the mean change in the prediction error. Development that addressed this issue is discussed in the following section.

6.8 Bias Adaptor

In order to improve the robustness of the developed prediction model it has been decided to employ adaptation of the prediction model in its most simple form - namely, the exponentially weighted moving average of prediction error or the low-pass filtered prediction error. This is evaluated and continuously added to a prediction of a model. In this way, non-zero mean of the prediction error is removed and its standard deviation is decreased. On the other hand, validity of cause-effect information contained in the prediction model itself is reduced, as it will be discussed more thoroughly towards the end of this section.

For a sake of clarity, in the reminder of this report the NOx predictor or prediction model relates only to the actual PLS- based dynamic model. The NOx estimator includes the bias adaptor.

As it is stated earlier, the bias adaptor takes the form of a low- pass filter that is applied to a prediction error whenever the analyser-measured NOx emission value is available. Then, the filtered output is added to the NOx prediction from the model.

The bw-pass filter can be expressed, in discrete- time framework, by the following difference equation:

$$y(k) = (1-a)\cdot y(k-1) + a \cdot u(k) \qquad (19)$$

where $0 \le a \le 1$ represents the 'learning factor', y represents the output variable, which is the filtered prediction error in this particular case, u represents the input variable, which is the raw prediction error in this particular case, and k represents the sampling instant in time.

Note that the expression given in (19) is the discrete- time representation of the Laplace domain transfer function equation given in (15). The reason for expressing it in discrete-time framework is that, unlike temperature measurements, the NOx measurement is not available at all times. Hence, the concept of time constant is less relevant in this case. Also, the term 'learning factor' is much more widely accepted term in adaptive signal processing applications.

By increasing the 'learning factor', i.e. $a \to 1$, more emphasis is placed on adapting the bias value, i.e. filter output, to a latest prediction error value, i.e. filter input. Since the filter output is then added to a prediction itself, it is expected that as $a \to 1$, the output of the NOx estimator approaches the measured NOx value. This is demonstrated in the following tables where the statistical analysis results for three different 'learning factor' coefficients are displayed, using the same set of validating data taken from January 2003.

Learning Factor = 0.01

Validation Period	Mean	Deviation	Maximum	Minimum
January 2003	0.81	16.12	77.45	-47.17
February 2003	0.12	16.43	77.74	-60.8
March 2003	0.24	18.16	81.45	-102.5
April 2003	0.08	16.39	70.93	-68.26

Table 17

Learning Factor = 0.1

Validation Period	Mean	Deviation	Maximum	Minimum
January 2003	0.09	13.34	70.65	-47.06
February 2003	0.03	12.26	65.43	-55.59
March 2003	0.04	14.19	91.59	-88.25
April 2003	0.03	12.71	64.57	-52.91

Table 18

Learning Factor = 0.25

Validation Period	Mean	Deviation	Maximum	Minimum
January 2003	0.04	10.94	56.92	-39.97
February 2003	0.01	9.93	51.84	-39.67
March 2003	0.02	11.64	84.79	-77.88
April 2003	0.01	10.33	58.35	-41.67

Table 19

Learning Factor = 0.5

Validation Period	Mean	Deviation	Maximum	Minimum
January 2003	0.01	7.44	36.81	-32.66
February 2003	0	6.75	38.45	-30
March 2003	0.01	7.98	66.13	-63.59
April 2003	0.01	7.04	52.34	-30.76

Table 20

Note that as the learning factor approaches its maximum value of one, the 'size' of estimation error, measured in terms of any of the four statistical measures, given in Tables 17 through to 20, tends to zero.

Another interpretation of the bias adapter can be made by considering the frequency-domain analysis of signals and systems. In this framework, the bias adapter can be seen as a model uncertainty block. In other words, the bias adapter compensates for those dynamics that are unaccounted for by the PLS- based prediction model. In this context, the learning-factor determines the bandwidth of unmodelled dynamics. More specifically, the smaller the 'learning factor' is the lower the bandwidth of the unmodelled dynamics. Hence, if the learning factor' is close to zero then only the very low- frequency components of unmodelled dynamics are compensated for by the bias adaptor. On the other hand, if the 'learning factor' is close to one then almost all of the frequencies are compensated for by the bias adaptor. In such a case, however, the prediction model is rendered obsolete since the prediction error is forced to zero by the action of bias adaptor alone. Hence, a compromise needs to be struck between reliance on structured information contained in the prediction model and the compensation for the unmodelled dynamics.

In order to further demonstrate effect that a learning factor has on the error of the NOx estimator, distribution functions of the NOx estimator error for three different values of a learning factor are plotted in Figure 31 alongside their corresponding Normal (Gaussian) distribution functions. Data over which prediction error is evaluated has been collected during January 2003.

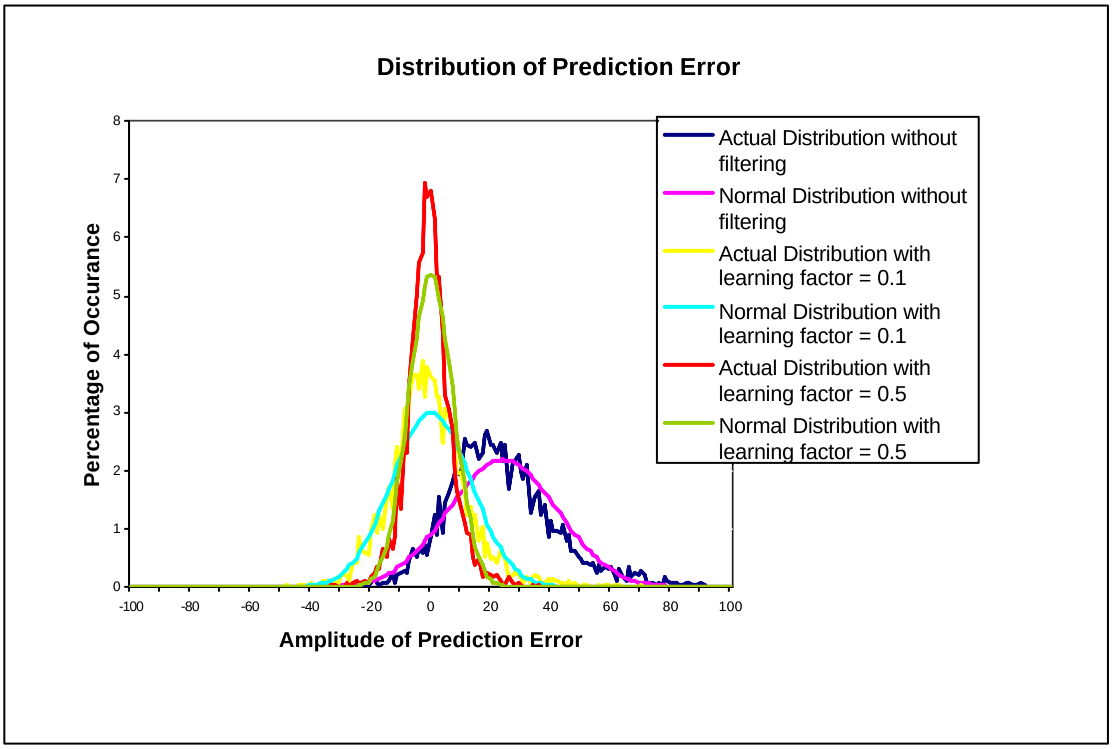


Figure 31

It can be seen in Figure 31 that as the value of the learning factor increases the corresponding distribution functions become steeper, i.e. variance or standard deviation is reduced, and the centre of the distribution function, i.e. the mean value of prediction error, is shifting towards the origin. This figure, therefore, demonstrates further the effect that learning factor has on the character of a prediction error.

Other important information that follows from Figure 31 is, however, the fact that the prediction error distribution function is relatively similar to the Gaussian (normal) distribution for all 3 choices of a learning factor values. This result indicates that the NOx estimator accounts for most of the structured information concerning NOx emissions.

6.9 Validation Monitors

6.9.1 Introduction

In order to improve the reliability/ robustness of the overall NOx estimator, additional validating condition monitors have been implemented. The purpose of these condition monitors is to validate and, in the case of instrumentation failure, infer the values of a subset of cause variables, namely temperatures in zones 1 and 2 and the combustion air temperatures. In this way, the overall reliability of a developed solution is greatly improved in the case of possible instrumentation failure. Note that these validation monitors use the same principle as those developed in AvestaPolarit application.

In this particular application, cause variables that do exhibit high level of cross correlation are the temperatures in the zones 1 and 2 and the combustion air temperatures from the first 7 zones of the furnace. Hence, two PCA- based validation monitors have been designed using the MonitorMV package with a training data set representing three months of data (January through to March 2003). Their respective details are discussed in the following two sub- sections.

6.9.2 Validation of Zone Temperature Measurements

This condition monitor considers temperature measurements from zones 1 and 2, which are highly cross-correlated process variables. These signals are denoted as 9.ME- 16.ME in Table 15.

A PCA model with 3 principal components (PCs) contributes 97.03% to the total variation of the training data set, as seen in Figure 32.

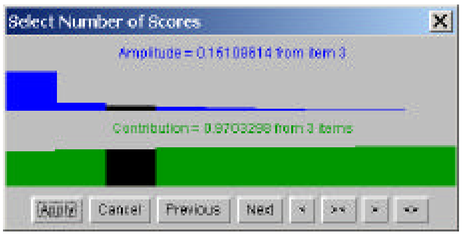


Figure 32

The reason for taking more than 1 principal component in this case is the improvement of the predictability of a model. While 1 principal component seems to be sufficient, by observing Figure 32, it is noticed in the prediction trends that there is a significant improvement in predictability over both the training and the validating data set if more principal components are included. However, in order to maintain good detection and isolation of the faulty thermocouple not too many principal components should be selected.

Results of the statistical analysis, performed on the prediction errors of these temperature measurements over the April 2003 (validating data set), are given in Table 21.

Signal Tag	Mean	Deviation	Maximum	Minimum
2TT1_1	1.08	5.21	27.07	-21.5
2TT1_2	1.52	10.25	42.11	-41.93
2TT1_9	-7.81	11.77	63.35	-54.26
2TT1_10	3.53	6.76	38.17	-36.58
2TT2_1	-1.77	7.72	30	-29.39
2TT2_2	1.81	9.74	37.12	-49.63
2TT2_9	1.23	11.52	59.32	-45.24
2TT2_10	0.29	10.04	56.88	-42.4

Table 21

It is important to note that while the standard deviation of these prediction errors is not very large their maximum and minimum values are. An attempt to improve robustness of this condition monitor is presented in section 6.10.

6.9.3 Validation of Combustion Air Temperatures' Measurements

Variables that are used by this condition monitor are the combustion air temperatures from the first 7 zones of the reheating furnace.

In this case, a PCA model with 3 PCs contributes 96.76% to the total variation of the training data set, as seen in Figure 33.

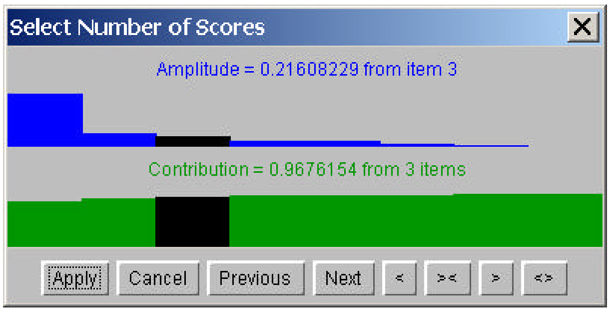


Figure 33

Once again more than 1 principal component is taken in order to improve the predictability of a developed PCA- based model.

The table given below contains results of the statistical analysis, performed on the validating data set (April 2003), of prediction errors for this PCA model.

Signal Tag	Mean	Deviation	Maximum	Minimum
2TT1L	1.02	4.22	26.27	-24.35
2TT2L	-0.57	1.45	4.45	-7.08
2TT3L	1.61	5.3	30.39	-18.92
2TT4L	-2.43	4.26	21.79	-21.86
2TT5L	1.17	4.53	16.18	-24.79
2TT6L	0.11	5.03	21.11	-15.79
2TT7L	-1.48	4.43	13.76	-17.23

Table 22

Once again, maximum and minimum values of the prediction errors encountered for the validating data set are relatively large. Hence, an attempt has been made, as described in section 6.10, to reduce the number of false alarm occurrences associated with this validation monitor.

6.10 Improving Robustness of Validation Monitors

Statistical analysis of the prediction errors for 2 PCA- based validation monitors has revealed that these monitors are not highly accurate, as shown in sections 6.9.2 and 6.9.3. In other words, the 'sizes' of prediction errors are not necessarily as small as one may require. As a result, prediction errors of the PCA- based statistical models that are routinely encountered are of comparable size to the measurement errors which are result of consequential systematic error present in the instrumentation equipment and, therefore, have consequential impact on the accuracy of the PLS- based prediction model.

The method that was employed in this sub- project for improving the robustness of condition monitors is identical to that described in section 5.4.

The statistical information concerning the filtered prediction errors, evaluated over the validating data set (April 2003) for both validation monitors, is displayed in Tables 23 through to 14. In order to demonstrate the effect that the choice of time constant has on a size of filtered prediction error three different cases were considered. Results for the zone temperatures' validation monitor are given in Tables 23, 24 and 25, while the results for the combustion air temperatures' validation monitor are given in Tables 26, 27 and 28.

Time Constant = 1 minute

Signal Tag	Mean	Deviation	Maximum	Minimum
2TT1_1	1.08	5.16	23.5	-21.43
2TT1_2	1.52	10.08	41.06	-41.18
2TT1_9	-7.81	11.61	60.21	-52.93
2TT1_10	3.53	6.63	37.08	-34.55
2TT2_1	-1.77	7.65	28.61	-28.66
2TT2_2	1.81	9.62	36.26	-45.16
2TT2_9	1.23	11.3	57.95	-42.99
2TT2_10	0.29	9.79	52.49	-41.46

Table 23

Time Constant = 10 minutes

Signal Tag	Mean	Deviation	Maximum	Minimum
2TT1_1	1.08	4.55	16.04	-17.93
2TT1_2	1.52	8.17	37.8	-33.24
2TT1_9	-7.81	9.94	46.39	-36.83
2TT1_10	3.53	5.26	22.67	-23.4
2TT2_1	-1.77	6.87	24.04	-22.73
2TT2_2	1.81	8.44	29.1	-35.46
2TT2_9	1.23	9.06	42.47	-32.62
2TT2_10	0.29	7.18	33	-30.92

Table 24

Time Constant = 1 hour

Signal Tag	Mean	Deviation	Maximum	Minimum
2TT1_1	1.08	3.46	9.96	-9.88
2TT1_2	1.51	5.53	23.03	-19.1
2TT1_9	-7.81	6.74	20.79	-25.6
2TT1_10	3.53	3.45	13.71	-11.43
2TT2_1	-1.77	5.74	18.38	-18.5
2TT2_2	1.81	6.41	18.61	-19.77
2TT2_9	1.23	6.48	27.13	-16.49
2TT2_10	0.3	4.48	15.75	-16.45

Table 25

Time Constant = 1 minute

Signal Tag	Mean	Deviation	Maximum	Minimum
2TT1L	1.02	4.18	26.08	-20.08
2TT2L	-0.57	1.44	4.39	-7.01
2TT3L	1.61	5.27	30.15	-18.64
2TT4L	-2.42	4.23	18.49	-21.65
2TT5L	1.17	4.51	16.11	-24.7
2TT6L	0.11	5	20.98	-15.7
2TT7L	-1.48	4.41	13.7	-17.12

Table 26

Time Constant = 10 minutes

Signal Tag	Mean	Deviation	Maximum	Minimum
2TT1L	1.02	3.36	20.45	-13.37
2TT2L	-0.57	1.29	3.67	-5.74
2TT3L	1.61	4.68	22.01	-15.39
2TT4L	-2.42	3.64	9.86	-16.43
2TT5L	1.17	3.88	13.67	-21.97
2TT6L	0.11	4.46	16.96	-13.53
2TT7L	-1.48	3.96	11.37	-14.45

Table 27

Time Constant = 1 hour

Signal Tag	Mean	Deviation	Maximum	Minimum
2TT1L	1.02	2.02	10.12	-5.25
2TT2L	-0.57	0.97	2.79	-3.33
2TT3L	1.59	3.25	11.06	-9.16
2TT4L	-2.42	2.41	5.32	-10.2
2TT5L	1.18	2.48	8.11	-9.47
2TT6L	0.12	3.54	9.64	-8.67
2TT7L	-1.49	2.96	7.75	-8.28

Table 28

As expected, and already discussed in section 5.4, the standard deviation of prediction errors decreases as the time constant of the corresponding filter increases. On the other hand, mean value remains almost unchanged. This is due to the fact that low-frequency components of the prediction error, which are main contributors to the mean value, are unaffected by the low- pass filtering.

6.11 Online Implementation of the NOx Estimation Scheme

6.11.1 Introduction

The developed NOx estimation scheme and the associated bias adaptor as well as the validation monitors have been implemented in real-time at the SSAB site in Borlange, by means of the MonitorMV Online system. The existing online application is expected to be used in any future developments of the NOx control scheme, (most probably implemented in an advisory form) and of a condition monitoring system of the overall reheating furnace.

In the case where the outputs of the PID controllers, i.e. the flow rates of air and fuel, for the first two zones of the furnace drop below 10% of their capacity, all the application components that include the PLS predictor, bias adaptor and validation monitors are switched to 'Manual' state. Otherwise, the application is in 'Auto', i.e. normal operating state. This is due to the inaccurate measurements of air and fuel flow rates, which play dominant role in NOx predictions, when their levels are below 10% of their capacity.

6.10.2 Layout of the MonitorMV Picture

The primary screen that should be observed by operator personnel is Picture 1, displayed in Figure 34. In this picture schematic of the overall NOx estimation scheme is presented.

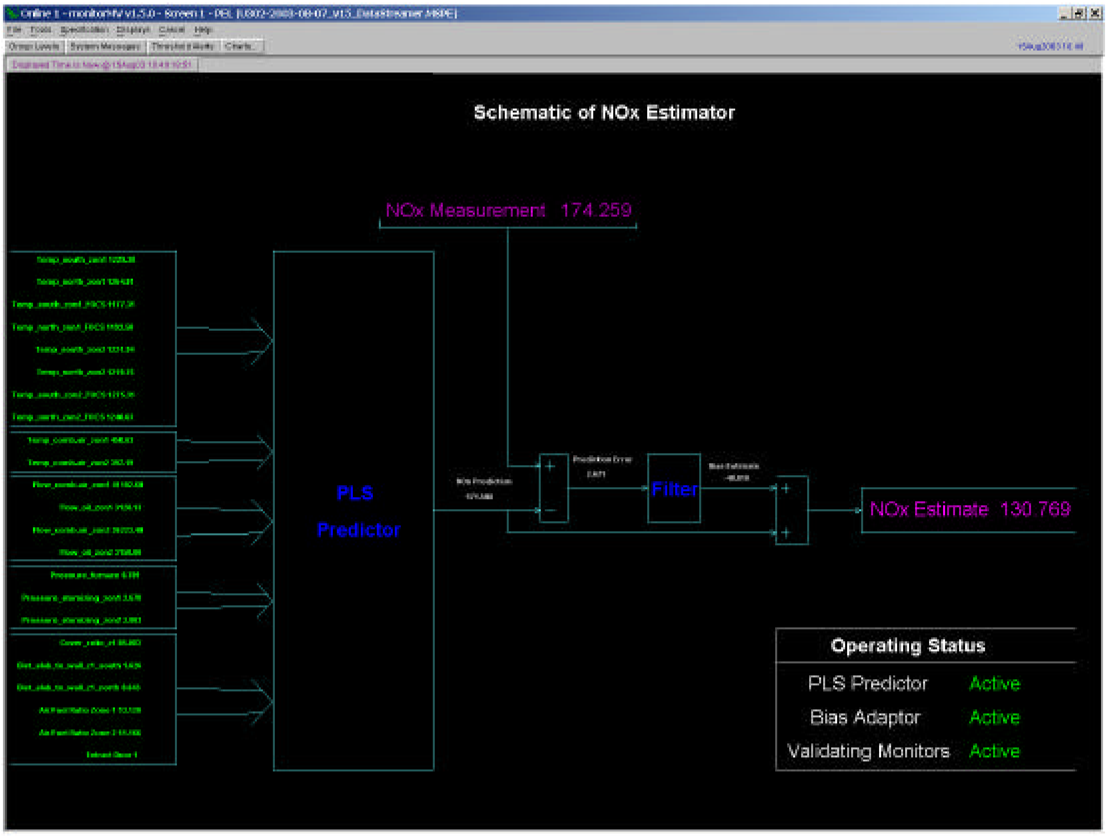


Figure 34

The operating status of PLS predictor, bias adaptor and validation monitors are displayed in the bottom right corner of the picture. In all cases 'Manual' status is coloured red and 'Active', representing 'auto' state, is coloured green, as shown in Figure 34.

Cause signals are listed in the left side of the picture while the NOx measurements are placed at the top of the picture. Instantaneous values of several intermediate variables within the NOx estimation scheme, are displayed in the middle of the picture. These include PLS- based NOx prediction, prediction error and bias estimate. Finally, the value of the NOx estimate, as a result of PLS prediction and bias adaptation, is given in the right side of the picture.

Cause signals that are subjected to validation, namely temperature measurements from the first two reheating furnace zones as well as the combustion air temperatures, are coloured green (normal status) or red (abnormal status) depending on whether or not their respective filtered prediction errors have violated corresponding threshold alert.

6.10.3 Assignment of Alarm Levels, Prediction Error Filter Time Constants and Bias Adaptor's Learning Factor

The alarm system, available within the MonitorMV Online system, has been applied to filtered prediction errors of PCA- based validation monitors. At the present, alarm levels have been set according to the maximum/minimum values of the filtered prediction errors, evaluated over the validating data set. In this way, it is believed that the number of false alarms would be significantly reduced, improving confidence of the operation personnel in the robustness of the condition monitoring scheme. Depending on the future performance of the overall scheme these limits may be reduced from these somewhat conservative levels in order to increase sensitivity of the condition monitors.

As far as the choice of the filter time constant is concerned, it is decided to be initially set to 10 minutes for all of the signals. Such choice is seen as the compromise between the speed of the response to sudden and rapid changes in prediction errors and the reduction of sensitivity to a short- lived rapid disturbances that would otherwise unnecessarily trigger alarm.

Limits imposed on filtered prediction errors of the validation monitors are given in the Table 29.

Signal Tag	Positive Alert Level	Negative Alert Level
2TT1_1	20	-20
2TT1_2	40	-40
2TT1_9	50	-50
2TT1_10	25	-25
2TT2_1	25	-25
2TT2_2	35	-35
2TT2_9	45	-45
2TT2_10	35	-35
2TT1L	25	-25
2TT2L	10	-10
2TT3L	25	-25
2TT4L	20	-20
2TT5L	25	-25
2TT6L	20	-20
2TT7L	15	-15

Table 29

The learning factor of the bias adaptor is set to 0.01. As a result, only the very low frequency components of the unmodelled dynamics are compensated for. The reminder of the process dynamics are to be emulated by means of the PLS- based prediction model. Hence, it is expected for the estimator's error to have zero mean and be largely composed of high- frequency components that have not been modelled by PLS method and have not been compensated for by means of a bias adaptor.

6.12 Summary

The need to protect the environment from combustion generated emissions, such as carbon monoxide (CO) and nitrogen oxides (NOx) has led in recent years to considerable demand for improved combustion system design and operation. And while the most important business drivers in economic considerations of the reheating furnace are minimisation of energy consumption and maintenance of high throughput, it is evident that, with increasingly stringent environmental regulations and heavy penalties for non-conformance, furnace emissions are likely to become a significant if not crucial cost driver. Such environmental considerations are forcing process plants to measure emissions and investigate methods for their cost- effective reduction.

The crucial step in attempting to address the issue of NOx emissions in cost- effective manner is the development of accurate cause- effect prediction model. Such model would not only offer viable and economic alternative to costly hardware- based analysers, in a form of a 'soft sensor', but also provide the basis for a development of a NOx control scheme.

Additionally, sudden and rapid change in terms of NOx emissions that are not accounted for by the developed prediction model may be a symptom of an operational problem of the reheating furnace. Such issue was not covered in this project. However, development of prediction model clearly benefits attempt to develop the condition-monitoring scheme of the reheating furnace.

This chapter details development of a NOx estimation scheme, using MonitorMV Design and Online systems, for a reheating furnace U302 at the SSAB site in Borlange, Sweden. Specification and diagram of this reheating furnace are presented in Figure 4. This sub-project has been carried out in collaboration with process control engineers of the SSAB, in particular Mr Jonas Engdahl, Mr Lennart Klarnäs and Mr Magnus Norberg, as well as Mr Per-Olof Norberg, advanced process control consultant.

In order to develop accurate prediction model important decision in the early stages of model design is the selection of a set of cause (input) variables. Such decision is made by employing process knowledge as well as some statistical analysis methods, notably correlation analysis. In the initial development of this project, correlation analysis was hampered by the lack of process excitation and irregular measurements of NOx emissions. Hence, the process knowledge provided a crucial insight into the underlying cause- effect structure of the model to be developed. In particular, impact on the NOx emissions by the variables that are related to the first two zones (preheating zones) of the furnace was highlighted by Mr Per-Olof Norberg. The most important cause variables from these two zones have been identified as the flow rates of air and fuel into the burners as well as the zone temperatures.

Another essential part of the overall prediction model development is the training data collection. This is due to the fact that the most system identification tools belong to a so- called 'data- driven' technology. Hence, quality training data is truly the only base for building a quality prediction model using these 'data- driven' technologies. In particular, if a prediction model is to be of dynamic form, as it is the case with NOx predictor, then a training data set has to be 'sufficiently excited' in order to reveal information concerning dynamic relationships between cause and effect variables. In order to sufficiently excite the process numerous step tests have been performed on the flow rates of air and fuel into the burners of the first two zones of the furnace.

The NOx prediction model has been developed by using Partial Least Squares (PLS) approach, available in MonitorMV. Due to the irregular measurements of the NOx emissions, model was decided to be of the FIR (finite impulse response) structure. The developed model has shown satisfactory level of accuracy and the statistical analysis of its prediction error has shown that prediction error distribution function is similar in shape to an equivalent Normal (Gaussian) distribution. This finding indicates that the NOx prediction model accounts for most of the structured information concerning NOx emissions.

In order to improve the robustness of the developed prediction model and ensure its validity in a face of non- stationarity of a process, it has been decided to employ adaptation of the prediction model in its most simple form. Namely, the exponentially weighted moving average of prediction error, i.e. the low- pass filtered prediction error, is evaluated and continuously added to a prediction of a model. In this way, non- zero mean of the prediction error is removed and its standard deviation is decreased.

Also, the additional validating condition monitors, based on Principal Component Analysis have been implemented. These monitors are used to ensure availability of measurements for a subset of cause variables. In this way, the overall reliability of a developed solution is greatly improved in the case of possible instrumentation failure. This validation scheme has been employed for those cause variables that exhibit strong cross- correlations. These were found to be the temperatures in the zones 1 and 2 and the combustion air temperatures from the first 7 zones of the furnace.

However, it has been found that prediction errors of the PCA- based validation models that are routinely encountered are of comparable size to the measurement errors which are result of consequential systematic error present in the instrumentation equipment and, therefore, have consequential impact on the accuracy of the PLS- based prediction model. As a result, sensitivity of the condition monitors had to be reduced by performing low- pass filtering of their prediction errors. In this way, focus is placed on slow drifts rather than short- lived rapid and sudden disturbances.

The developed NOx estimator, consisting of the PLS- based prediction model and the bias adaptor, as well as the associated validation monitors have been implemented online, using MonitorMV Online system, at the SSAB site in Borlange, Sweden, providing the continuous estimation of the NOx emissions. This NOx estimator is expected to facilitate further developments of NOx control scheme and aid in a development of a condition-monitoring scheme for a reheating furnace.

7. Sub-project 3: Investigation of the use of Multivariate Statistics for the Modelling of an Acid Regeneration Process

7.1 Introduction

This sub-project is concerned with investigations that relate to an Acid Regeneration process at the SSAB factory at Borlange.

The Acid regeneration process is used primarily for the regeneration of pickling liquor, namely hydrochloric acid, that is used to remove iron oxide on the steel during the continuous annealing of the steel slabs. As a by- product of acid regeneration, iron oxide is created. This iron oxide has a market value.

In 2001, the Acid regeneration process had only recently been commissioned. The operational characteristics of the process were only just beginning to be appreciated. It was not properly understood how to avoid situations that gave rise to large deposits of iron oxide on the walls of the regeneration plant – such deposits being difficult to remove and giving rise to costly maintenance exercises.

For this reason, it was decided to attempt to apply multivariate statistical process analysis in order to gain knowledge into the process and solve operational problems that were causing sub- standard performance of the process. Although some progress in gaining understanding of process operation has resulted from applying MSPC to the acid regeneration plant, the overall success of this sub-project has been more limited than those previously described and there is presently no lasting and adequate condition monitoring solution for this process.

Attempts to exploit the condition monitoring technology to the acid regeneration plant are described in this chapter in chronological order. The main reason for this ordering is that it allows the reader to properly understand the sequence of events and decisions that were taken during the programme of work.

Although the programme of work has been limited in its success, there are a number of positive interpretations that can be made and these have influenced the overall programme, including the manner in which the other sub-projects have been approached.

7.2 Basic Description of the Acid Regeneration Process

The Acid regeneration plant is a process that regenerates used pickle liquor, which results from the pickling of hot- rolled steels using hydrochloric acid. As a by-product, ferric oxide (Fe_2O_3, hematite) is produced, which can be used for subsequent industrial processing.

The operation of the process consists of the following general steps:

1. The spent pickle liquor (waste acid), which is taken from the pickling baths is fed over an installed waste acid filter in order to separate solid particles.

2. The waste acid is pumped to the spray booms. From there it is sprayed into the reactor by nozzles, which are attached to the ends of the spray booms.

3. The reactor is fired by four burners placed tangentially on one level that generate a circulating stream of hot gases in the reactor. The supply of energy is necessary for evaporation of water, for reaching the reactor temperature and for compensating the loss of heat in the system.

4. In normal operation mode, the hot roast gas consists mainly of steam, HCl combustion gas and minor quantities of ferric oxide dust, leaving the reactor with a temperature of about 390 degrees centigrade.

5. The ferric oxide, resulting from the reaction, falls down into the reactor-cone and is carried out by a rotary valve.

6. The roast gas is fed into a venturie and mixed with waste acid. Roast gas is cooled and the concentration of the waste acid is increased.

In the project, further parts of the process have not been considered and are therefore not described in this section.

7.3 Phase I: Initial Developments

This section describes developments that took place during the first stage of the project. In particular, an initial development of PCA-based models for the overall process is described. Problems, which were encountered and decisions which were made, are discussed in this section.

The first attempts to develop a statistical model to describe the acid regeneration process were based around the principle that if all available data is collected over a significant period, covering many days of process operation, then such data should provide a basis for representing the normal profile of the process. Subsequently, if other data is referenced against this profile then there should be a basis for determining if this other data is normal or not.

Following this theme, resulting PCA- based models revealed several clusters in the score space, i.e. the space that is spanned by the retained principal components. These clusters correspond to

- normal acid operation,
- starting up/shutting down ‘water mode’ operation, and
- periods during which the acid plant was not operational.

and are illustrated in figure 35.

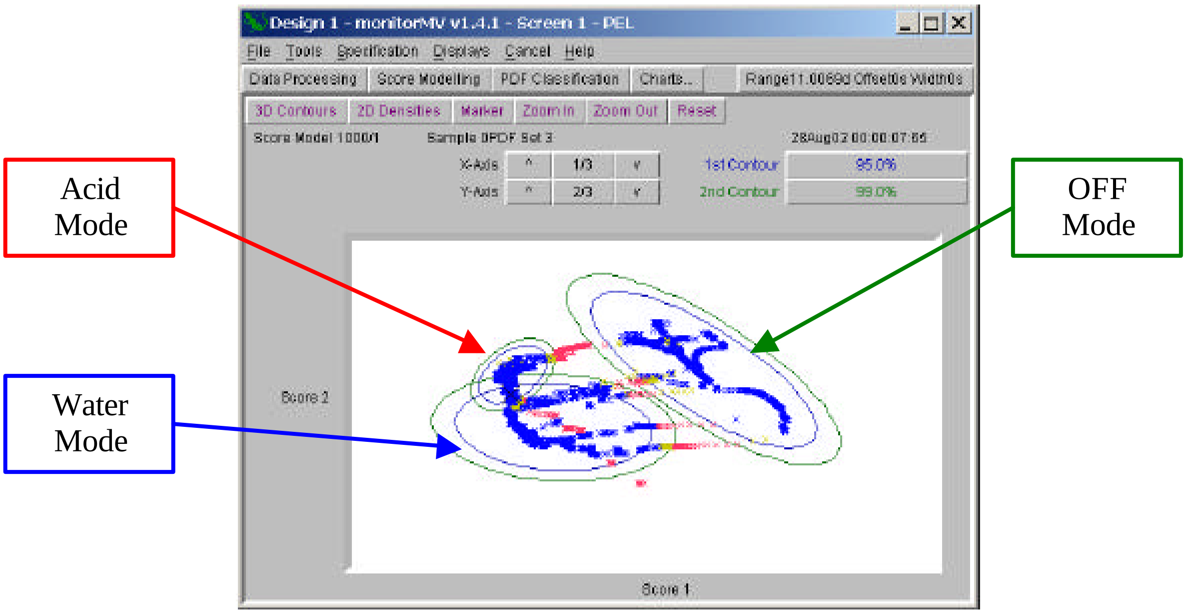


Figure 35

Ellipses in Figure 35 represent multivariate Gaussian- based probability density function (PDFs) boundaries that classify regions of process operation. Each Class (or ellipse) is associated with a particular cluster, which, in turn, corresponds to a particular mode of operation of the acid regeneration plant, as shown in Figure 35.

By means of principal component loadings, described by Q_n in equation (4), it is possible to relate process variables to these three clusters and gain some understanding of how different modes of operation relate to individual process variables. For example, score 1 is dominated by the flow rates of air and fuel into the burners of the acid plant's burner system. Hence, the main distinction between acid mode of operation and 'OFF' mode of operation lies in significant change of absolute value of these process variables, as is to be expected.

Also, temperatures inside the burners' chambers are almost completely uncorrelated with any other process variable and are almost the sole contributors to the second principal component. These variables are responsible for the 'stretched' shape of the 'OFF' mode cluster since during this mode of operation burner temperatures are slowly decreasing, causing the score space trajectory to move from the top to the bottom corner of the 'OFF' mode cluster, as shown in Figure 36.

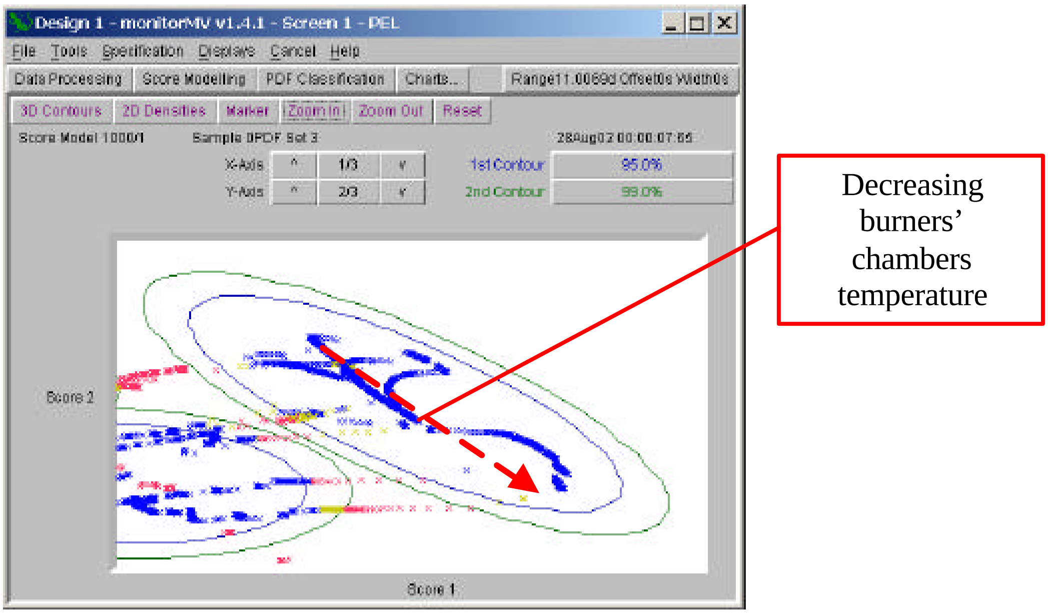


Figure 36

The results described here are of interest and highlight the ability of statistical methods, present within MonitorMV system, to detect different modes of process operation and relate them to the behaviour of individual process variables.

However, the momentum of this initial progress in the project was not maintained – this because it did not prove possible to relate the signals being monitored by MonitorMV to the causes of main concern to the process operating staff. The plant management were engaged in their own campaign to get to an understanding of the process and were making frequent changes to aspects of the process and to the process operating conditions. Information concerning such changes was not being referenced by MonitorMV or was not in a form that could be utilised. The outcome was that, although MonitorMV could detect that the process was operating in a different regime, there was no basis for deciding the basis for the difference or if the difference corresponded to normality or otherwise.

The above considerations became clear after attempting to relate to all three modes of operation simultaneously. It was therefore decided to narrow the scope of examination to only the normal acid mode of operation. In this way it was thought that the sensitivity of the model would be increased and small-scale variations that may differ from the normal would be more clearly highlighted. However, as a result of increased sensitivity of the principal component models, the non-stationary nature of the process became even more apparent. In particular, it was found out that the general statistical model had extremely limited period of validity before it being rendered obsolete by some change in the operating condition of the acid regeneration plant.

The real lesson here is that progress in statistical modelling for condition monitoring is only feasible if a process is settled in its operating conditions. Any changes must be of a consistent and observable nature and must be able to be referenced by the monitor if any progress is to be made. This, unfortunately, was and is not the case with the acid regeneration plant.

7.4 Phase II: Attempt to Develop Cause- Effect Model of the Acid Regeneration Plant

The programme of work progressed in order to try to make some headway in producing models to describe the behaviour of the acid plant. It was decided to investigate the possibility of determining a model that would relate cause signals with effect signals by employing a PLS based model. In this way, non-stationarity that was a direct result of the changes in cause signals would be accounted for by such model and the validity of a model would be extended in time.

Several sets of cause signals were chosen. Notably, the flow rates of air and fuel into the burners, flow rates of used hydrochloric acid into the reactor, as well as the set-points of several PID loops were used to develop prediction models. All the other process variables were treated as effects.

Unfortunately, this approach was hampered by the lack of excitation in the measured cause variables and the unavailability of the dominant cause variable measurements, namely the quality of the incoming acid. As a result, no accurate cause- effect model was developed and this direction was abandoned.

7.5 Phase III: Development of Iron Oxide Condition Monitor

The final focus in this project has been placed on the reactor unit of the process and development of a statistical model to relate to the quality of iron oxide. Laboratory analysis results of the iron oxide quality were obtained and used to select data sets that corresponded to satisfactory operation.

However, the chemical composition of the iron oxide was analysed irregularly and infrequently. Also, there was no guarantee that the information about the timing of the sample collection was correct. Soon it became apparent that in order to develop an adequate statistical description it would be necessary either to analyse iron oxide quality more regularly or to get regular feedback from process engineers concerning the overall quality of process performance.

In collaboration with process operations staff, sets of data were obtained which corresponded to satisfactory and to unsatisfactory behaviour respectively. Also, crucial process variables, considered to reflect problems in the process operation, were identified (namely 6 particular reactor temperature measurements). A PCA model has been developed using the portion of data described as representing satisfactory production. Overall, 749 data points have been used for the training of the statistical model.

Due to the lack of strong cross- correlation between reactor temperatures, 3 out of possible 6 principal components were retained contributing 83.45% to the variation of the training data set, as shown in Figure 37.

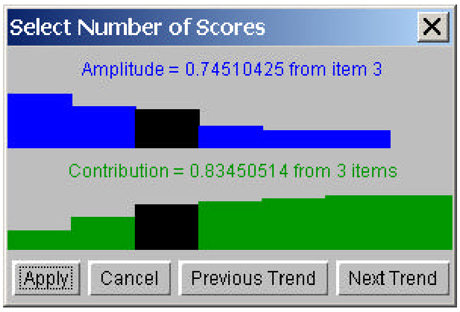


Figure 37

The ability of the developed statistical model to predict reactor temperature measurements is observed in Figure 38. Note that, on this display, trends that are coloured in blue, green and magenta represent actual temperature measurements while the brown lines represent the corresponding predictions of these signals, based on the statistical model. In this way, any change in correlation patterns can be observed on a variable- by- variable basis in order to establish what type of correlation pattern breakdown has occurred. In this particular case, the PCA model has managed to capture the majority of variation in the training data set, as seen on the prediction trend display.

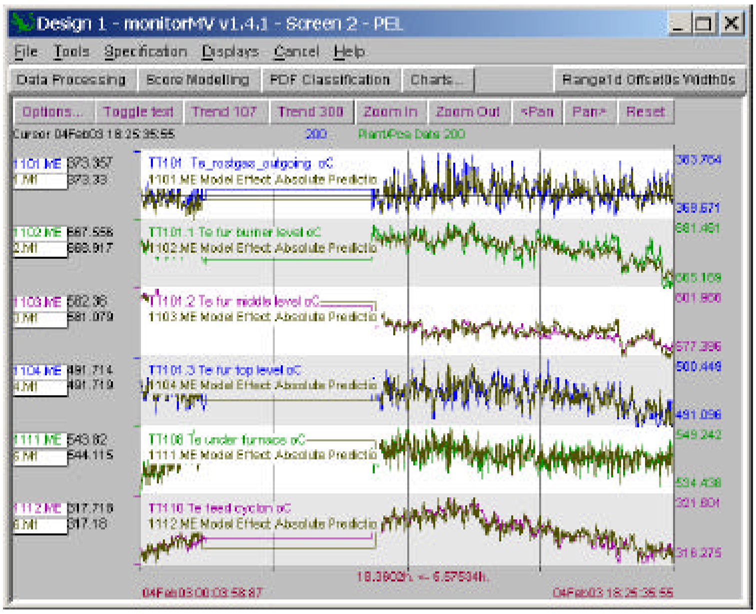
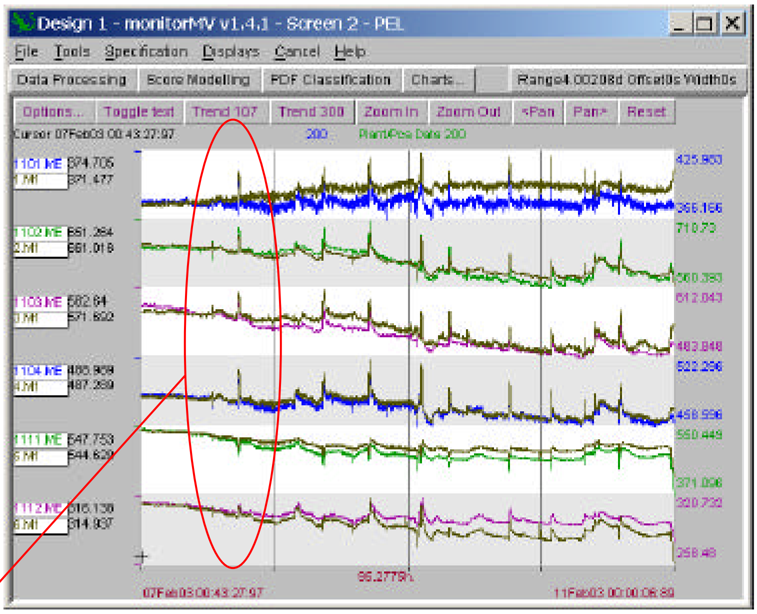


Figure 38

Using the information, provided by the production manager of the acid regeneration plant, about the quality of produced iron oxide the ability of the developed statistical model to detect substandard production was assessed.

In particular, production during 9th and 10th of February 2003 was reported to have been sub- standard in terms of the iron oxide quality. On 8th February the prediction errors from the statistical model start to increase significantly and remain significant throughout 9th and 10th February as seen in Figure 39. This clearly indicates the model's capability to detect deviation in process performance.



Change in correlation pattern

Figure 39

In particular, it is observed that the difference between the prediction and the actual value of the outgoing roast gas temperature (signal ID 1101.ME, tag name TT101) is increasing with time. This indicates that the outgoing roast gas temperature is smaller than expected during this particular period. A similar deviation from expected behaviour is observed in the case of temperature under furnace (signal ID 1111.ME, tag name TT108), while the feed cyclone temperature trend (signal ID 1112.ME, tag name TT110) is seen to be higher than its predicted trajectory.

Hence, in this particular case the prediction trends provide a clear indication that the process performance is continuously deviating from the operating regime that was present in the training data set. Furthermore, the prediction trends indicate which variables have been affected the most and in what way. Such information could be employed to inform operators of deviating performance and provide guidelines, in terms of reactor temperatures, on how to improve process performance.

Also, during the 1st and 2nd of March 2003 iron oxide quality was sub- standard and, therefore, data representing this period was also analysed using the developed condition monitor. The most clear indication of degrading performance, as seen in Figure 40, is presented through a steady increase in predictions of outgoing roast gas temperature (signal ID 1101.ME, tag name TT101), the reactor temperature at the top level (signal ID 1104.ME, tag name TT101.3) and the temperature below the reactor (signal ID 1111.ME, tag name TT108) when compared to the actual values of these signals.

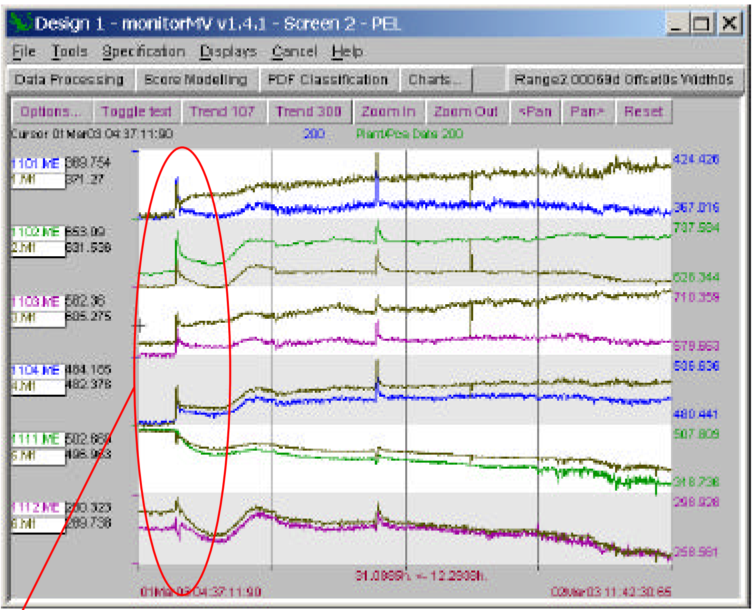


Figure 40

Hence, once again developed PCA model was able to detect change in the correlation pattern in several reactor temperature measurements.

However, a serious limitation of the developed statistical model is its restricted validity. This is due to the limited range of process operation that was captured in the training data. It became necessary, therefore, to extend the training data set to include other operating regimes that deliver satisfactory process performance.

As a result of this case study it was decided to perform three special production runs that would last for 3 weeks. During this period, samples relating to the iron oxide quality were collected every two hours, both new and old acid nozzles were used and pressure as well as the flow rate of the incoming acid would be varied. In this way, it would be possible to establish whether or not different product grades can be characterised as well as be distinguished from each other by means of statistical modelling. If successful, such modelling would pave the way for the optimisation of process performance, ensuring that the iron oxide quality conforms to desired specifications.

This exercise in the project has taken place in the last month of active work. Only around 50% of the iron oxide quality samples have been analysed and this is not enough to provide an effective statistical analysis.

7.6 Phase IV: Final Statistical Analysis of the Acid Regeneration Process

7.6.1 Introduction

In this section, results of the final statistical analysis performed on 10 days of operation, during special process runs that were undertaken during the month of May 2003, are reported.

Variables that were the focus of this analysis were chosen to be reactor temperatures and burner chambers' temperatures. Due to the fact that these two sets of temperature variables are mutually uncorrelated, two PCA models were developed. MonitorMV specific signal IDs, tag names and descriptions of these signals are given in Table 30.

MonitorMV signal ID	tag name	Description
1101.ME	TT101	temperature of the gas leaving the reactor
1102.ME	TT101.1	temperature in the reactor at the burner level
1103.ME	TT101.2	temperature in the reactor at the middle level
1104.ME	TT101.3	temperature in the reactor at the top level
1111.ME	TT108	temperature at the bottom of the reactor
1112.ME	TT110	temperature of the roast gas inside the cyclon
1105.ME	TISA107_3	Temperature inside the burner chamber no. 1
1106.ME	TISA107_4	Temperature inside the burner chamber no. 2
1107.ME	TISA107_5	Temperature inside the burner chamber no. 3
1108.ME	TISA107_6	Temperature inside the burner chamber no. 4

Table 30

Training data were composed of those periods of data that corresponded to a satisfactory iron oxide production. All together 35,681 samples, with sampling interval of 6 seconds, were used for training of the models.

Also, the validating data that corresponded to the satisfactory production was used to observe whether developed models were able to generalise to those periods, which represented satisfactory performance of the process. Overall, 12,103 samples, with sampling interval of 6 seconds, were used for the validation of the models.

Finally, the data that corresponded to sub- standard production of iron oxide was used to test developed models. All together 56,366 samples, with a sampling interval of 6 seconds, were used for the testing of the developed PCA models.

Note that only data that was collected during the acid mode of operation was used.

7.6.2 Reactor Temperatures' Condition Monitor

In the case of reactor temperatures there is no strong cross- correlation, as it was already shown in section 7.5 and observed in Figure 37 where the relative amplitudes of the first few principal components are not significantly larger than the last few principal components. Using the cross- validation technique and PRESS statistic, described in more detail in section 3.1, it was decided to choose two principal components. As shown in Figure 41, first 2 (out of possible 6) principal components contribute 86.59% to the total variation of the training data set.

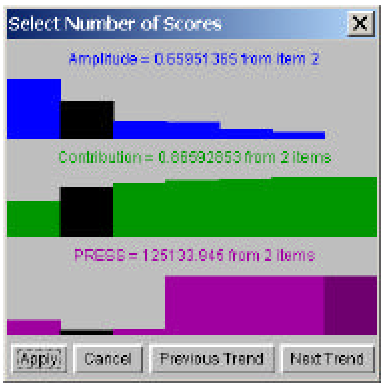


Figure 41

The predictions of the individual reactor temperatures are shown in Figures 42 and 43 for two segments of the training data set.

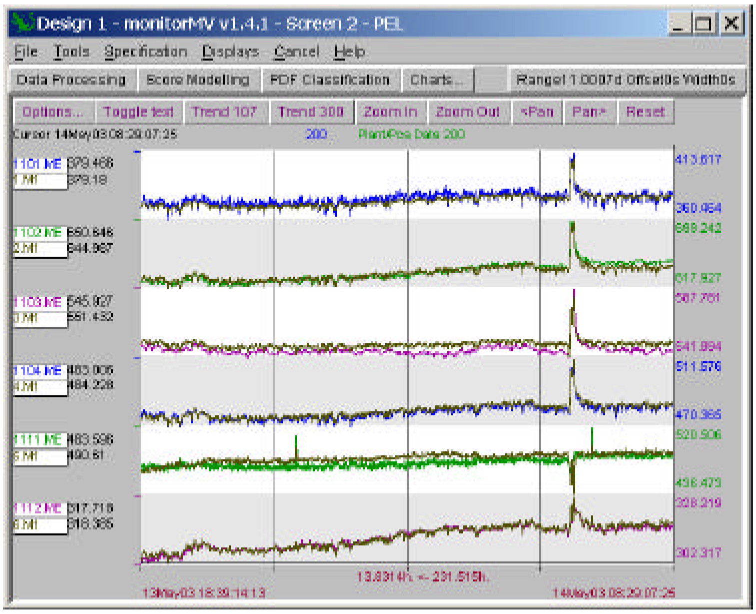


Figure 42

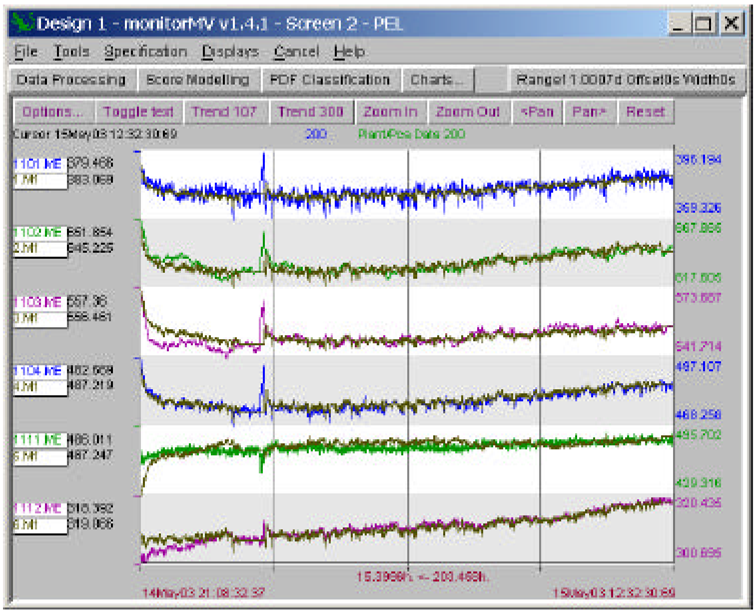


Figure 43

As observed in both of these figures, the PCA model did accurately depict most of the features present in the training data set.

Results of the statistical analysis performed on the prediction errors for the training data set are presented in Table 31.

tag name	Mean	Deviation	Maximum	Minimum
TT101	0	2.47	12.86	-12.96
TT101.1	-0.01	6.03	24.19	-33.44
TT101.2	0	4.78	31.7	-24.14
TT101.3	0	1.62	5.67	-7.35
TT108	-0.01	9.09	44.9	-62.51
TT110	0	1.64	8.67	-15.59

Table 31

PCA predictions for the validating data set are displayed in Figure 44 and the results of the statistical analysis performed on the prediction errors for the validating data set are given in Table 32. The results show that the developed PCA model was not able to accurately predict reactor temperatures at all points for the validating data set but there is good correspondence for certain of the signals across certain portions of the data ranges. There appears to be a singular event, just past the half way stage that gives rise to an offset on all but the last of the temperatures on display.

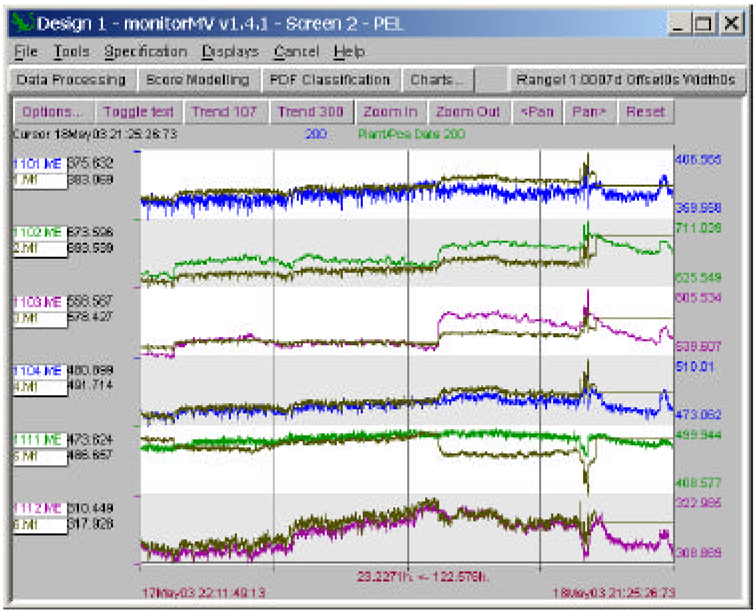


Figure 44

tag name	Mean	Deviation	Maximum	Minimum
TT101	-5.32	3.31	3.55	-19.86
TT101.1	15.69	5.73	38.96	-1.89
TT101.2	3.85	8.28	23.19	-10.36
TT101.3	-2.26	1.63	2.48	-7.32
TT108	11.11	12.72	65.49	-17.62
TT110	-1.13	0.9	2.35	-6.14

Table 32

In the case of the data that corresponded to a sub- standard production of the iron oxide, PCA predictions for three different segments are displayed in Figures 45, 46 and 47.

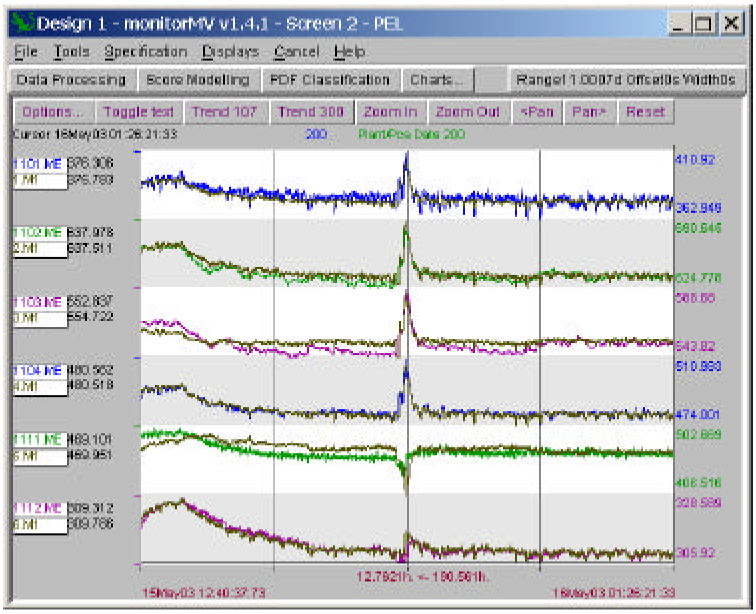


Figure 45

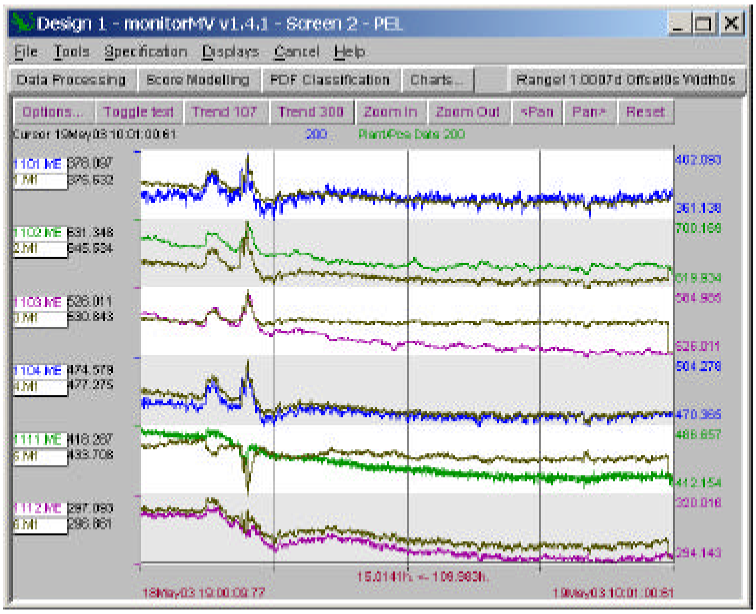


Figure 46

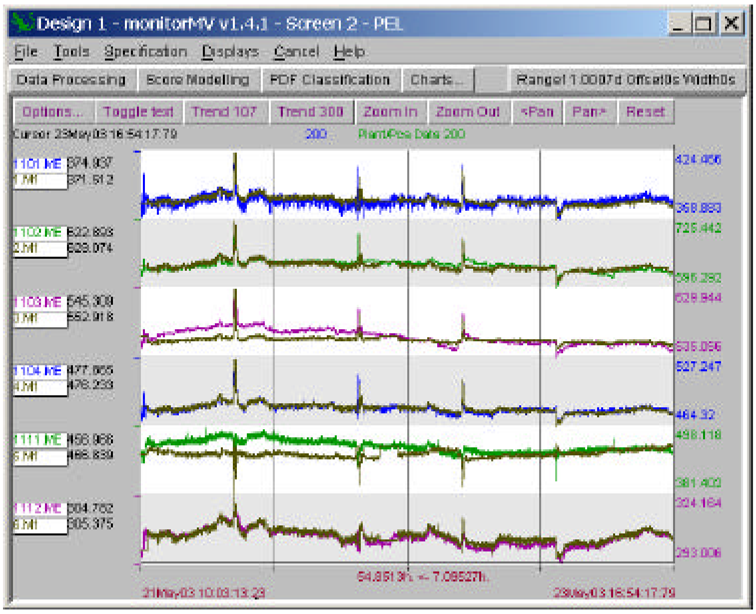


Figure 47

Also, the results of the statistical analysis performed on the prediction errors are presented in Table 33.

tag name	Mean	Deviation	Maximum	Minimum
TT101	-0.04	3.12	14.73	-13.59
TT101.1	2.88	9.46	34.3	-30.56
TT101.2	-0.46	10.33	31.47	-30.13
TT101.3	-0.33	1.76	7.5	-15.7
TT108	5.05	15.77	59.36	-69.61
TT110	-1.11	2	11.74	-10.93

Table 33

It is observed, by looking at Figure 45, that prediction errors for this period of unsatisfactory production are kept low. Hence, in this particular case the analysis is not able to detect a sub- standard quality of iron oxide.

However, there is a more significant deviation in prediction errors for the periods corresponding to Figures 46 and 47. Unfortunately, by comparing results given in Table 32 with those in Table 33 it is shown that such deviation is not significantly larger than the deviation that was observed in the validating data set.

It should be noted, however, that a thorough analysis of the data can only be conducted only after all the iron oxide quality samples, corresponding to three special runs, have been made available. The above investigation is of a preliminary nature only – although there is an indication that the impression of the operations staff that the six chosen temperatures have influence on quality might be misguided.

A proper analysis needs to be progressed once all of the data from the three experiments is made available. The indication is that more than the six temperature signals may need to be included and there is the possibility that more experiments may be needed in order to reach a proper conclusion as to the parameters, which are most influential upon the quality of the iron oxide. Proper interpretations can only be made once there is enough data to provide an effective basis for statistical interpretation.

8. Conclusions and Future Directions

8.1 Thermocouple Validation Scheme

The Validation scheme applied to the Reheating Furnace at the AvestaPolarit factory has been structure into three sections. Each section is comprised of temperature measurements that are correlated with each other. There is no correlation between the temperatures of different sections. This scheme is shown to provide a viable basis for determining the integrity of temperature measurement in the furnace. A particular anomaly has been analysed (section 5.6) and it is shown clearly that an installed condition monitor would have detected this anomaly and would have provided a valid estimate of temperature that could be used to temporarily replace the measurement during the time of anomaly. Such validation must provide the means for more effective energy management of the furnace by avoiding the positioning of control system set points at inappropriate temperatures.

Future developments should focus on placing an the application into the control room of the reheating furnace A, providing accurate and reliable validation of thermocouple measurements. Also, a similar scheme should be employed for the reheating furnace B. Furthermore, the concepts that are employed in validation of thermocouples should be employed in the future for other instrumentation equipment that exhibits high levels of cross- correlation.

8.2 NOx Estimation Scheme

The main cause signals for the NOx emissions were found to be fuel and air flow rates in preheating zones 1 and 2. Additionally, it has been found that the flow rate of the atomising steam into the burners has a significant impact on the NOx emissions. This particular variable is found to be negatively correlated with NOx emissions. In other words, increase in the total flow rate of atomising steam into the burners is found to reduce NOx emissions. Hence, the total flow rate of atomising steam could be seen as a crucial cause variable in any attempt to minimise NOx emissions. Additional causes were chosen to be all of the remaining available process variables that relate to zones 1 and 2.

Due to the irregular measurements of the NOx emissions, the prediction model was selected to be of the FIR (finite impulse response) structure, as opposed to ARX. Identification was performed using the PLS method. The developed prediction model has been validated on the data set, which was not used in identification and it was observed that the model was able to generalise to the data set that was not used in training. Hence, the conclusion has been made that a prediction model achieved a satisfactory level of accuracy. Validation has also been performed using the data from January, February, March and April 2003. However, it was then found that the mean of the prediction error, in particular, was significantly larger than expected. Hence, the decision has been made that some form of model adaptation is needed, particularly in order to account for the time-varying nature of the mean change in the prediction error.

Model adaptation took the form of the exponentially weighted moving average of prediction error, i.e. the low-pass filtered prediction error, which is evaluated and continuously added to a prediction of a model. In this way, non-zero mean of the prediction error is removed and its standard deviation is decreased.

In order to improve the reliability/ robustness of the overall NOx estimator, additional validating condition monitors have been implemented. The purpose of these condition monitors is to validate and, in the case of instrumentation failure, infer the values of a subset of cause variables, namely temperatures in zones 1 and 2 and the combustion air temperatures. In this way, the overall reliability of a developed solution is greatly improved in the case of possible instrumentation failure. Note that these validation monitors use the same principle as those developed in AvestaPolarit application.

Future developments should focus on incorporating the developed prediction model into the advisory system that would indicate which cause variables should be changed and by what amount in order to minimise NOx emissions while maintaining high productivity. Also, the model should be employed in the development of the condition-monitoring scheme for the entire reheating furnace. This is especially so since sudden and rapid change in terms of NOx emissions that is not accounted for by the developed prediction model may be a symptom of an operational problem of the reheating furnace. In order for these schemes to be successful, a diagnostic rule base needs to be established by using the process knowledge, which would relate results produced by such advisory/ condition monitoring schemes and the actual process.

8.3 Condition Monitoring of the Acid Regeneration Plant

The investigations with the acid regeneration plant have proven to be less productive than those reported above for the reheating furnaces. However interesting aspects have been shown concerning the capability of the Multivariate Statistics to classify regions of process operation and to relate these regions to variations in key process variables.

Although it has been shown to be straightforward to describe with accuracy short periods of process operation on the basis of derived Principal Component Models, these models could not sustain accuracy in the longer term because of the high degree of variability in the process. Such variability arose because of the frequent changes to process conditions that were made by process operations staff in order for them to better understand process behaviour and improve quality. The details of such changes were not available to MonitorMV and therefore could not be factored in to the MonitorMV models.

Thus a specific set of experiments was carried out in May in order to determine if particular temperatures are influential on iron oxide quality. These experiments, three in all, have involved the collection of frequent iron oxide samples to be subsequently analysed in the laboratory. Unfortunately the complete set of analysis results has yet to be made available and proper conclusions concerning the experiments cannot be drawn. Early indications are that the signals that were considered to be potentially the most influential upon product quality might not turn out to be so and that the search for meaningful and measurable process signals that can be used to infer quality may have to be widened.