JERNKONTORETS FORSKNING

Serie	Nr	Datum	Forskningsuppgift nr
D	**807**	**2003-11-27**	**51045**

(TO 51-41)

INTELLIGENT ALARM HANDLING IN THE STEEL MANUFACTURING INDUSTRY

DATA DRIVEN PROCESS CONDITION MONITORING OF FURNACE 301 AT SSAB - REPORT NO 2

Report by Per-Olof Norberg, Advanced Process Control Ltd. and
David Sandoz, Control Technology Centre Ltd.

Key words: monitoring, multivariate statistics, reheating furnaces, process control, nitrogen oxides, iron and steel industry

SUMMARY

It has been proposed that by using predictive condition monitoring on historic data from an industrial steel reheating furnace there will be improved operational understanding, measurement robustness, and incipient fault detection that all together will lead to increased efficiency and productivity. This report discusses the initial analysis and system that has been developed to achieve a versatile predictive condition monitor.

The findings demonstrate the ability to detect abnormal process events and identify subtle deviations from normal process behaviour. These abilities are delivered by a sophisticated multivariable model that gives robust predictions in the presence of highly correlated noisy production data. The analysis also covered the development of on-line predictions of environmental emissions, in particular NOx, and other important economic variables such as scale loss, production rate and specific energy consumption. These estimated parameters can be used to independently backup sensors or included as part of an enhanced control system.

Accurate predictions of NOx, O2, scale loss, production rate and specific energy consumption have been obtained for Furnace 301. Where software based Predictive Emissions Monitoring System (PEMS) is used to monitor NOx (or any other environmental emission), a periodic relative accuracy test is usually required by the relevant government agency. Although PEMS are

usually applied to real time data (as opposed to the averaged data analysed here), this report includes a typical relative accuracy calculation for the NOx prediction to illustrate the precision of the model which may now be used for further evaluation.

The preliminary multivariable model developed for the furnace can be used to detect changes from normal operation. These changes may be considered as abnormal events, although more subtle changes in process behaviour and drift in analyser measurements may simply indicate wear and tear.

The scope of work has been restricted to the RSTF campaigns from shifts 20020218 through 20031704. This campaign has the largest number of samples (compared to VKVH, SPEC and KVLL). A comparison with traditional Univariate Statistical Process Control (SPC) gives a powerful demonstration of the improved sensitivity achieved through the use of the more advanced multivariate condition monitoring approach.

Implementation of this predictive condition monitor would provide operational staff with a versatile tool for post production condition monitoring and NOx estimation. Especially the ability of the predictive condition monitor to determine which variables have contributed the most to a deviation from normality would in itself be extremely useful. Furthermore, process and analyser drift over consecutive campaigns would easily be detected, giving operational staff early warning before this leads to sub optimal production or a possible fault condition.

Serious consideration should be given to the development of a predictive condition monitor and predictive emissions monitoring system based on real time data.

ISSN 0280-249X

Contents

Data Driven Process Condition Monitoring of Furnace 301

1 Executive Summary

It has been proposed that by using predictive condition monitoring on historic data from an industrial steel reheating furnace there will be improved operational understanding, measurement robustness, and incipient fault detection that all together will lead to increased efficiency and productivity. This report discusses the initial analysis and system that has been developed to achieve a versatile predictive condition monitor.

The findings demonstrate the ability to detect abnormal process events and identify subtle deviations from normal process behaviour. These abilities are delivered by a sophisticated multivariable model that gives robust predictions in the presence of highly correlated noisy production data. The analysis also covered the development of on-line predictions of environmental emissions, in particular NOx, and other important economic variables such as scale loss, production rate and specific energy consumption. These estimated parameters can be used to independently backup sensors or included as part of an enhanced control system.

Accurate predictions of NOx, O2, scale loss, production rate and specific energy consumption have been obtained for Furnace 301. Where software based Predictive Emissions Monitoring System (PEMS) is used to monitor NOx (or any other environmental emission), a periodic relative accuracy test is usually required by the relevant government agency. Although PEMS are usually applied to real time data (as opposed to the averaged data analysed here), this report includes a typical relative accuracy calculation for the NOx prediction to illustrate the precision of the model which may now be used for further evaluation.

The preliminary multivariable model developed for the furnace can be used to detect changes from normal operation. These changes may be considered as abnormal events, although more subtle changes in process behaviour and drift in analyser measurements may simply indicate wear and tear.

The scope of work has been restricted to the RSTF campaigns from shifts 20020218 through 20031704. This campaign has the largest number of samples (compared to VKVH, SPEC and KVLL). A comparison with traditional Univariate Statistical Process Control (SPC) gives a powerful demonstration of the improved sensitivity achieved through the use of the more advanced multivariate condition monitoring approach.

Implementation of this predictive condition monitor would provide operational staff with a versatile tool for post production condition monitoring and NOx estimation. Especially the ability of the predictive condition monitor to determine which variables have contributed the most to a deviation from normality would in itself be extremely useful. Furthermore, process and analyser drift over consecutive campaigns would easily be detected, giving operational staff early warning before this leads to sub optimal production or a possible fault condition.

Serious consideration should be given to the development of a predictive condition monitor and predictive emissions monitoring system based on real time data.

2 Methodology

Monitoring the overall condition of a reheating furnace is a significant challenge. Fundamentally we are trying to decide whether the furnace is behaving in a normal or abnormal condition. Normal in this sense simply means what operational personnel consider being normal or best operating practice and abnormal is any deviation from this.

Historical operating data from normal periods are used to 'fingerprint' the process and all analysis stems from this fingerprint. As a first step traditional statistical process control (SPC) should be considered. This is very much a data driven approach and is a univariate approach since strictly speaking the process variables are assumed to be independent of each other and normally distributed. Charts such as Shewart, CUSUM or X-bar can be plotted. Empirical rules are used to identify patterns of behaviour in the charts that indicate abnormal behaviour – more commonly known as 'special cause variation' in SPC analysis – or process drift. The assumptions of normality and independence, coupled with the need for one chart per process variable, often result in failure when SPC is applied to real time continuous processes.

A far more powerful technique is to develop a multivariable model of the process and monitor the errors between the predicted values of process variables and their measured values. This is predictive condition monitoring. Process interactions are taken into account by the multivariable model and robust models can be identified from highly correlated data using a Partial Least Squares (PLS) method. A single, composite, prediction error is calculated and a threshold placed on this value that reflects a high probability of abnormal behaviour. The operator now has just one value to monitor and the assumptions of independence and normality of process variables no longer apply.

Of course, an accurate and robust model is central to the success of a predictive condition monitoring scheme. As this report demonstrates, such a model has been developed for Furnace 301.

2.1 Univariate Approach

The univariate approach to process condition monitoring is to apply traditional SPC. Based on normal operation various charts such as Shewart, CUSUM or X-bar are constructed and appropriate limits calculated that indicate special cause variation. This approach is totally data driven. Our knowledge of the process is the set of limits that indicate special cause variation. When these limits are exceeded or certain patterns are observed in the process data as it is plotted on the SPC charts special cause variation, i.e. abnormal behaviour, is detected.

We will restrict our attention to one of the most commonly used charts - the Shewart chart. The Shewart chart is simply a chart of a variable with limits of +/- 3 standard deviations around the mean value drawn on the chart. Movement of a variable within the standard deviation limits indicates normal, random, variation and is of no concern. Movement of the variable outside the limits is interpreted as a special cause of variation, i.e. abnormal behaviour. The underlying assumption is that the variable has a normal distribution, in which case 99.7% of samples will be within the +/- 3 standard deviations limits.

The mean value and +/- 3 standard deviation limits are calculated from training data and remain fixed until they are re-calculated on new training data. Figure 1 shows Shewart charts for scale loss, O2 and NOx with standard deviation limits recalculated for each period of training data (i.e.

normal operation). During periods of test data the standard deviation limits remain at the values previously calculated as can be seen in Figure 2. In both Figure 1 and Figure 2 excursions outside the standard deviation limits are coloured in red for good data. Bad data are drawn to scale but ignored by the Shewart chart analysis.

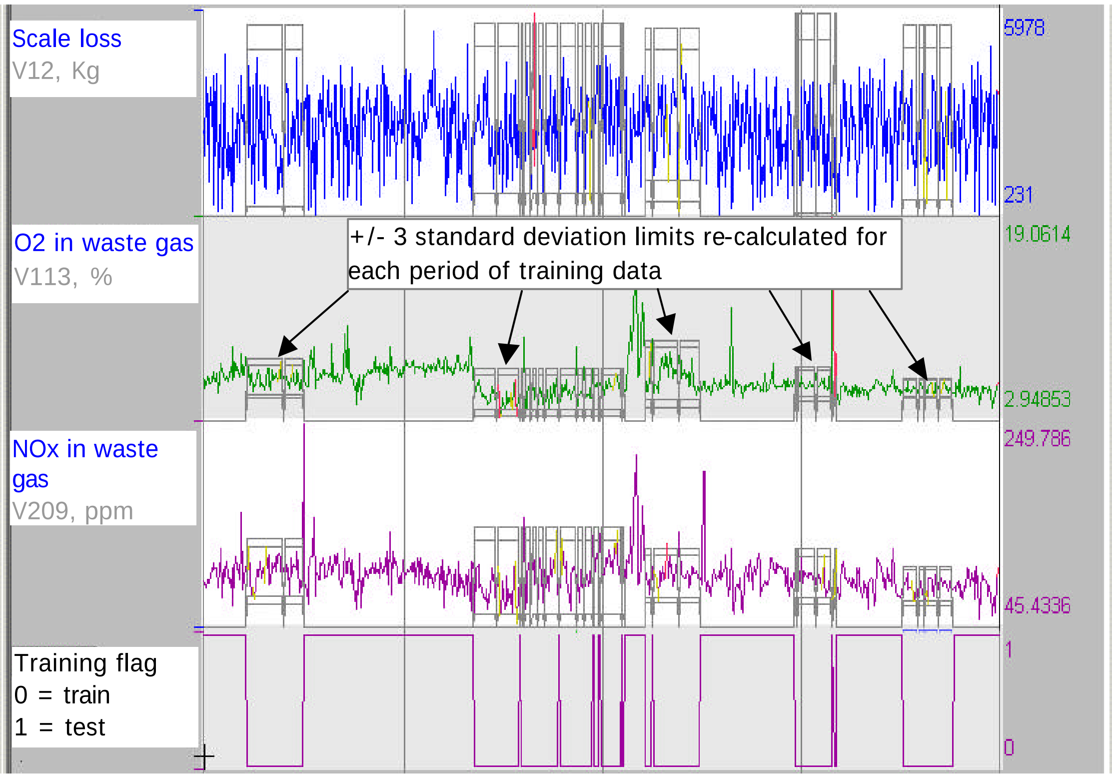


Figure 1 – Shewart chart limits calculated from training data

Figure 1 and Figure 2 show Shewart charts for just 3 variables. There are considerably more variables to consider, particularly variables that are manipulated in order to control the furnace such as air fuel ratios. To ensure a meaningful comparison with the predictive monitoring scheme described shortly, the same variables should be considered. Therefore, another 35 variables need to be considered (the prediction model described later for scale loss, O2 and NOx has a total of 38 variables). These additional 35 charts are not presented here but have been taken into account in the comparison between the univariate and multivariate approaches described in the Analysis section.

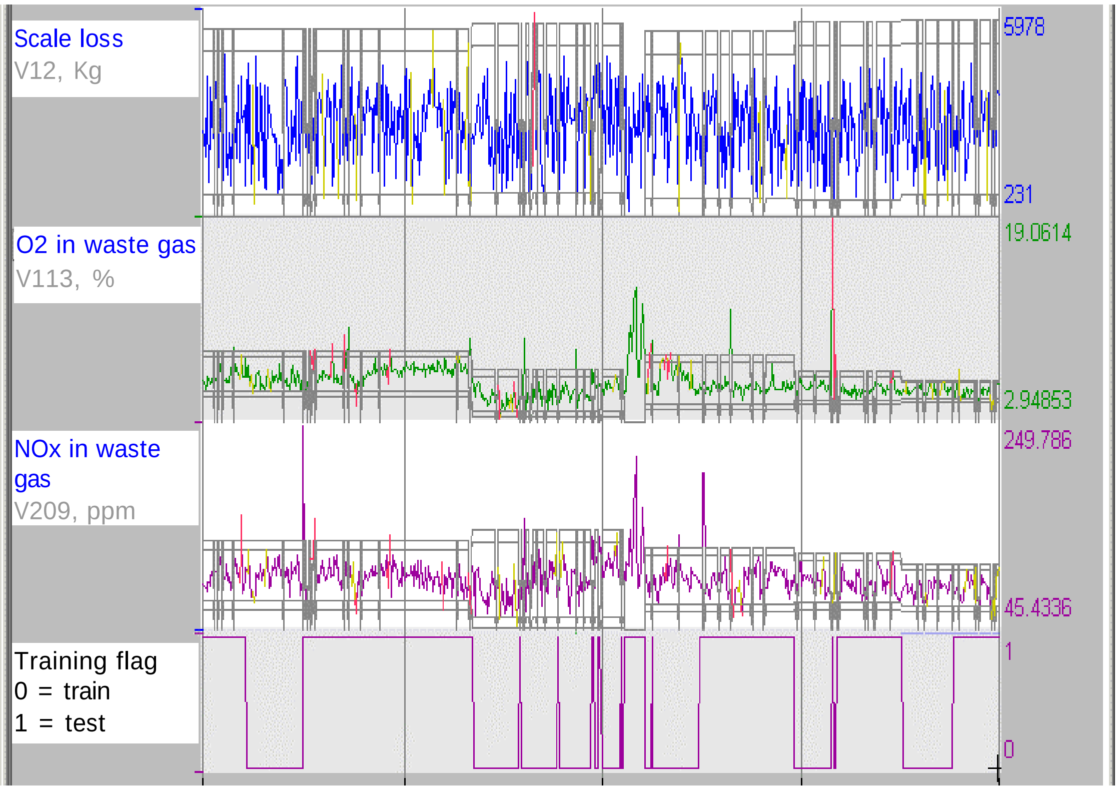


Figure 2 – Shewart chart evaluated for training and testing data

2.2 Multivariate Approach

In the multivariate approach we develop multivariable models that capture cause – effect relationships in normal operation. We compare predicted with actual behaviour and monitor the resulting difference. By placing appropriate limits on the prediction errors and observing patterns in the prediction errors we can infer the condition of the process. This is the essence of predictive condition monitoring.

Identification of an accurate and robust model is of vital importance. The following section on Modelling describes the model developed for Furnace 301. The final part of this description, section 2.3.6, explains how prediction errors are interpreted to monitor the condition of the furnace.

2.3 Modelling

The training data has been revisited to gain a deeper understanding of the furnace. Calculated values in the training data were recomputed to be certain of their validity. Some minor errors in production rate and specific energy consumption were identified and corrected. The air fuel ratio for zone 1 was also fully computed (in the original training data this value was only computed over small sections of the data set). Roll temperatures (N, S and B) were found to have more missing values than actual values.

Note that corrections to the training data have had little if any effect upon correlations that exist in the data. The multivariable models have been recomputed from the corrected training data.

2.3.1 Changes to model variables

This work is concerned with furnace operation during production. A number of process variables included in the original analysis and modelling relate to furnace operation during stoppage. While there may have been correlation between production and stoppage variables, it was felt that the stoppage variables should be excluded from the models. This was done with no discernable impact upon model accuracy. The stoppage variables removed from the model are listed in Appendix 1 – Modifications to Cause Effect Structure of Model. Re-examination of the training data revealed that roll temperatures have a very high proportion of missing values and so are excluded from the model.

The NOx value predicted in the original models is the un-normalised value 115.ME. This has been replaced by a normalised value V209, calculated as below:

V209 Normalised NOx production = V115 * 16 / (21 – V113)

The following effect variables are calculated values: % scale loss, production, total production rate, specific production rate, total energy consumption and specific energy consumption. All are non linear because they involve multiplication and/or division of other process variables. The degree of non linearity was found to be worst in %scale loss (V163) and total specific energy consumption (V162) and so these have been removed from the model structure. Accumulated scale loss (V12) in kg has been added to the model as an effect variable and total fuel consumption (V161) in kWh has been added to the model as a cause variable.

Steam production and total steam production were modelled fairly accurately in the original data set. However, evaluation of the original model over unseen data gives extremely poor prediction of these variables, see Figure 3. As you can see the dominant feature of the steam production variables is their integrating nature. Integrating behaviour cannot be predicted by a steady state model and so these variables have been removed from the model structure.

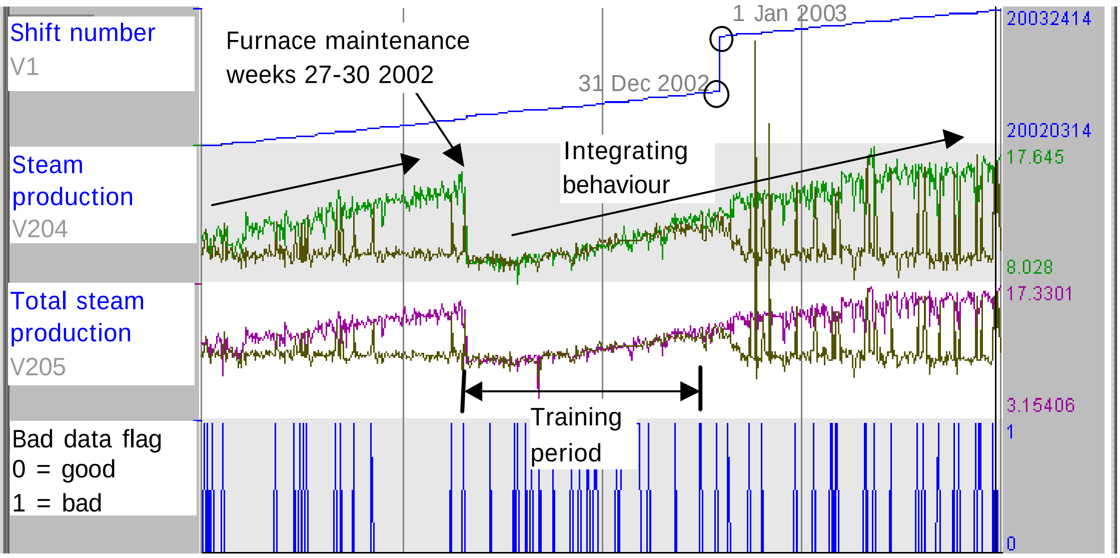


Figure 3 – Integrating behavior of steam production

Gas and air temperatures located after the recuperator were declared as cause variables in the original model structure. These variables should be regarded as effect variables and so were moved to the effect side of the model. However, predictions proved to be inaccurate when the model was evaluated over the complete data set. This observation was repeated when the gas and air temperatures were also on the cause side of the model. Consequently these temperatures have been removed from the model.

2.3.2 Model structure

There is no change to the type of model, which remains as below:

- Zero order dynamics with zero delays on all cause variables. I.e. a steady state model.
- Absolute data format
- Multiple effect format. I.e. any effect variable is a function of all other effect variables in addition to being a function of the cause variables.
- Partial Least Squares (PLS) identification

The changes to the cause and effect variable structure as described previously give the new variable structure summarised in Table 3 in Appendix 1 – Modifications to Cause Effect Structure of Model.

2.3.3 Model Accuracy

Predictions of the effect variables declared in the model are illustrated in Figure 4. To the right of each trace are shown the maximum and minimum values of the measured variable. Overlaid on each trace in brown is the model prediction. The train / test flag shows which data is used to train the model and which data is used to test the model; a value of 0 indicates training data, a value of 1 indicates test data. Approximately half of the data is used to train the model.

Outliers in the data set are excluded in the model training and testing – these are indicated when the bad data flag is 1. Mostly the outliers occur as isolated samples with the exception of a prolonged period of operation where both O2 and NOx are very high. This period could not be modelled successfully at all as the O2 and NOx respond very differently when compared to all other times.

The model predictions are very good, with the exception of O2 where two periods show a sustained drift between the measured value and the prediction. A closer look at the model predictions is shown in Figure 5. Here you can clearly see the drift in the O2 prediction.

Production, total and effective production rates and effective specific energy consumption are calculated from other process variables within the data set. These values are very linear over the ranges calculated. As we might therefore expect the predictions of these calculated variables are more accurate than the predictions of NOx and O2. The standard deviations of the prediction errors for the training and testing data are given in Table 1. Apart from O2, the model accuracy in the test data is similar to that in the training data. The reduced accuracy of O2 prediction in the test data is due to the periods of drift highlighted in Figure 4. The drift in O2 is discussed in section 3.1 - Detecting drift.

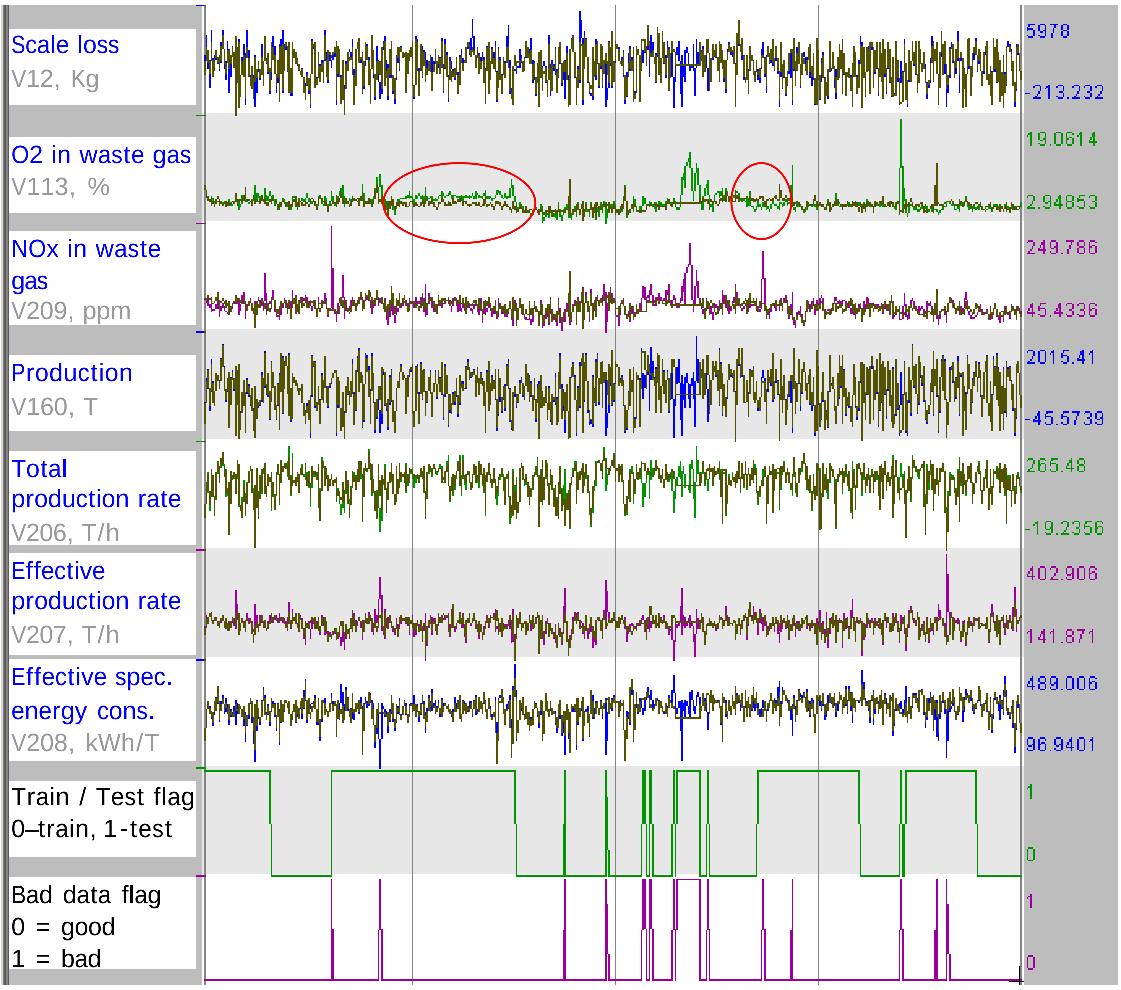


Figure 4 – Prediction of waste gas emissions and economic parameters

	Prediction Error			
	Training data		Test data (excludes training data)	
	Standard deviation	Standard deviation / range	Standard deviation	Standard deviation / range
Scale loss	242 *Kg*	4.2%	337 *Kg*	6.4%
O2	0.6 *%*	9.3%	0.9 *%*	16.2%
NOx	9.2 *ppm*	8.6%	10.8 *ppm*	10.5%
Production	31.5 *Tonne*	1.8%	33.9 *Tonne*	1.9%
Total production rate	6.5 *Tonne/hr*	2.9%	8.3 *Tonne/hr*	3.4%
Effective production rate	8 *Tonne/hr*	6.2%	11.5 *Tonne/hr*	6.8%
Effective specific energy consumption	11.8 *kWh/Tonne*	3.6%	16.2 *kWh/Tonne*	4.5%

Table 1 – Prediction accuracy over training and test data with one model

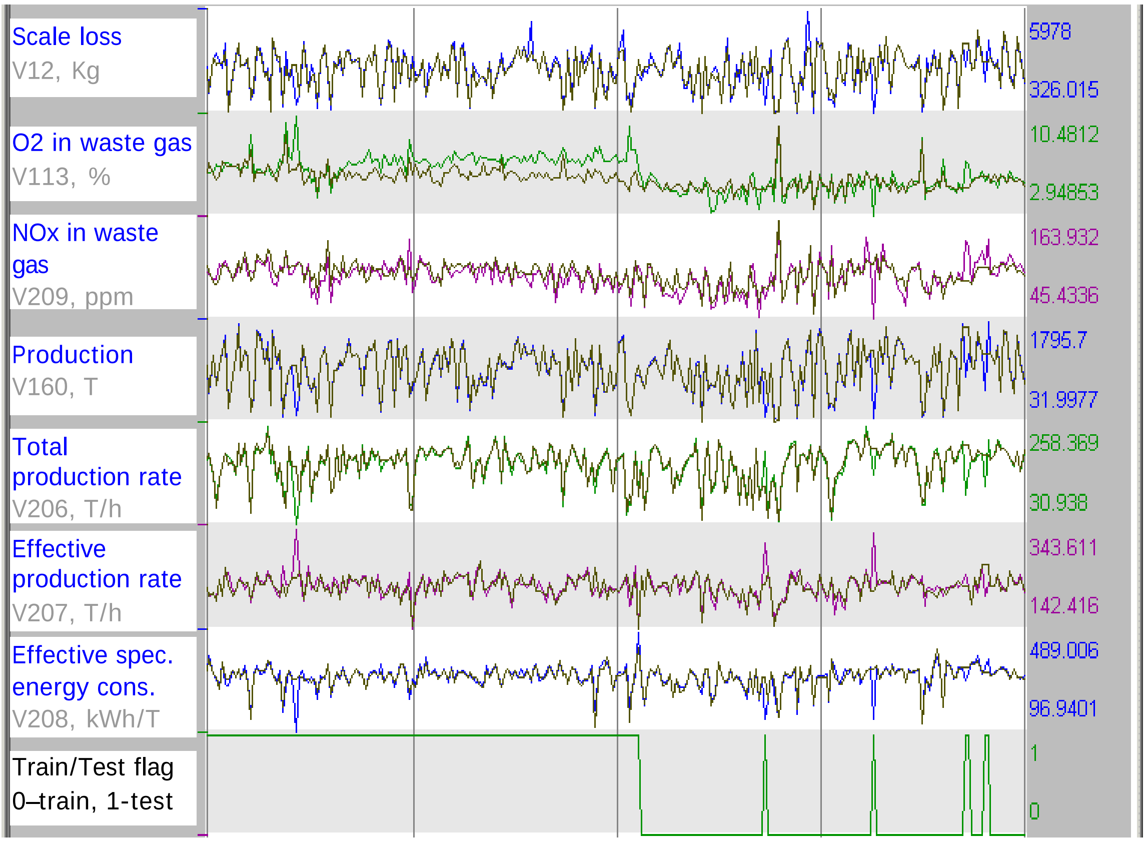


Figure 5 – Close up of model predictions

So far model accuracy has concentrated on prediction of effect variables. In addition to the prediction of effect variables the PLS identification method employed by MonitorMV™ also estimates values of cause variables. For example Figure 6 shows typical estimates for a handful of variables over a region of test data. The cause variable estimates are overlaid in brown on the process values. The estimated values are determined from a reduced score model in a similar fashion to principal component analysis. As you can see the estimates are very accurate, helped by the fact that furnace operation gives rise to a high degree of correlation between cause variables. The ability to estimate cause variables is of great importance in condition monitoring as the breakdown of normal patterns of correlation is indicative of a change from normal behavior.

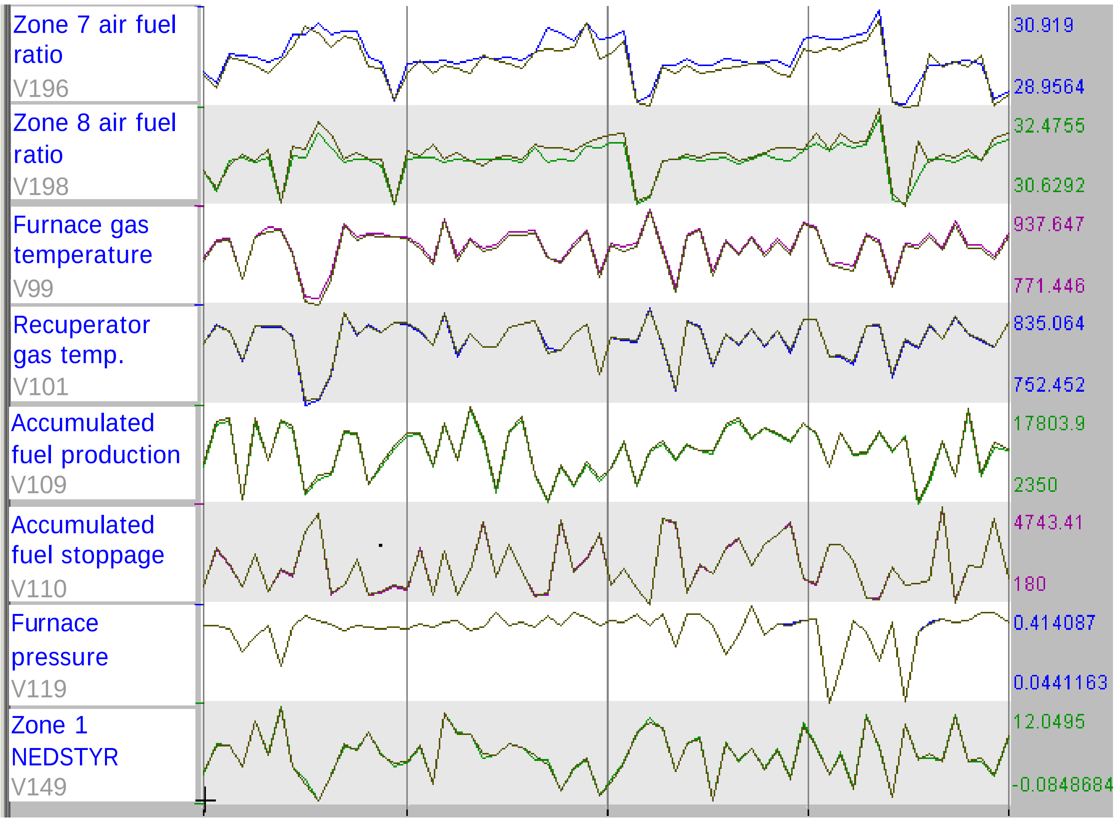


Figure 6 – Typical cause variable estimates

2.3.4 Non Linearity in effective production rate and effective specific Energy Consumption

Effective production rate and effective specific energy consumption are calculated according to the formulae shown below. Although the resulting values are mainly linear over the ranges calculated, these are non linear functions as they involve division of one process variable by another and additionally, in the case of effective specific energy consumption, multiplication of process variables. As the predictive model is linear, we would expect to find reduced accuracy where the non linearity becomes pronounced – when production time and production are small. Fortunately this does not occur very often as we can see by looking at scatter plots of prediction error against production time and production, see Figure 7.

$$Effective production\,rate = \frac{Production}{Production time/3600} = \frac{V160}{V61/3600}$$

$$Effective\,specific\,energy\,consumption = \frac{Effective\,energy\,consumnption}{Production} = \frac{V109 \times V153 \times V154}{V160}$$

The samples circled in Figure 7 and all samples with a production value less than 200 T are now flagged as bad data and will be excluded in the analysis from here onwards. Note that the range of production values is approximately 1750 T, with a maximum value of about 1850 T. The exclusion of production values less than 200T therefore removes only a very small number of samples.

Note that total production rate is also a non-linear calculated value [production/((production time + stoppage time)/3600)] but the sum of production time and stoppage time is never small enough to cause a problem.

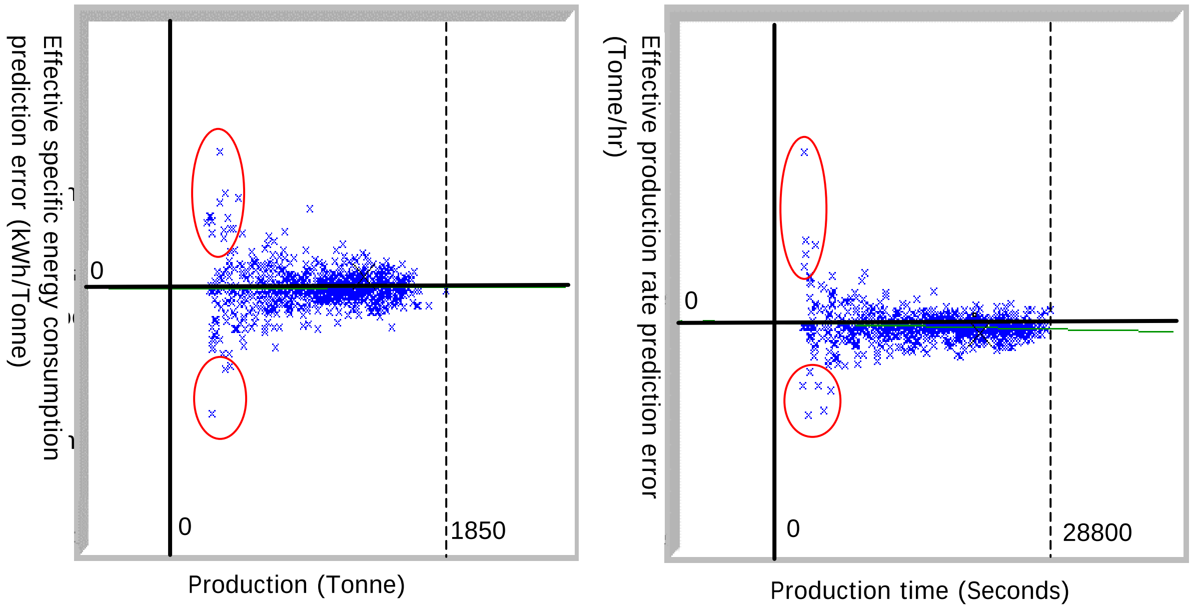


Figure 7 – Scatter plots of prediction error against production and production rate

2.3.5 Scale loss, O2 and NOx model

The effect variables we are modelling fall into two categories: measured values (scale loss*, O2 and NOx) and calculated values (production, total and effective production rates and effective specific energy consumption). *Scale loss is a measured value in the sense that it is not calculated from other process variables in the data set.

The first ten cause variables in Table 3 are associated mainly with the calculated values of production, total and effective production rates and effective specific energy consumption. Through a process of trial and error a small improvement in robustness was achieved for just scale loss, O2 and NOx by removing the calculated effect variables and these ten cause variables, leaving the cause and effect model structure in Table 4. Hearth area ratio was also removed from the variable structure. The accuracy obtained with this reduced model is shown below in Table 2. Comparing against the full model, see Table 1, you will see that prediction accuracy on the training data has actually been degraded slightly. However, the reduced model is a little more robust than the full model. I.e. the prediction errors for the test data and training data are closer to each other.

	Prediction Error Training data		Test data (excludes training data)	
	Standard deviation	Standard deviation / range	Standard deviation	Standard deviation / range
Scale loss	277 *Kg*	5.2%	353 *Kg*	7.3%
O2	0.6 *%*	10.3%	0.8 *%*	14.4%
NOx	9.2 *ppm*	9.9%	10.1 *ppm*	9.9%

Table 2 – Prediction accuracy over training and test data with scale loss, O2 and NOx model

2.3.6 Interpreting Prediction Error

With both cause estimate errors and effect prediction errors we can determine a lot about the condition of the process. We can now begin to infer the nature of a change from normal behaviour:

Cause variable estimates OK, prediction errors in effect variables only.
Normal patterns of correlation exist in the cause variables. This indicates that the process is being driven normally. The presence of prediction error in only the effect variables tells that one of three things is happening:

1) Although the process is being driven normally it has been driven to a region where the model is not longer valid. There is nothing wrong with the process
2) The process is behaving abnormally. Sudden and large errors may indicate an isolated event such as a fault whereas smaller persistent errors across several effect variables may represent process drift.
3) There are faults in effect variable measurements. Persistent prediction error may indicate a problem in a measurement prone to drift/failure such as a thermocouple measurement or an analyser measurement.

Errors in cause variable estimates, effect variable prediction errors OK Normal patterns of correlation have broken down in the cause variables. This indicates that the process is being driven abnormally. Since the effect variable predictions are OK the process is, however, responding normally. This leads to one of the following conclusions:

1) The process is behaving abnormally. Persistent error may indicate a change to new operating mode, sudden and large errors may indicate an isolated event such as a fault or unusual manual intervention.
2) There are faults in cause variable measurements. Persistent prediction error may indicate a problem in a measurement prone to drift/failure such as a thermocouple measurement or an analyser measurement.

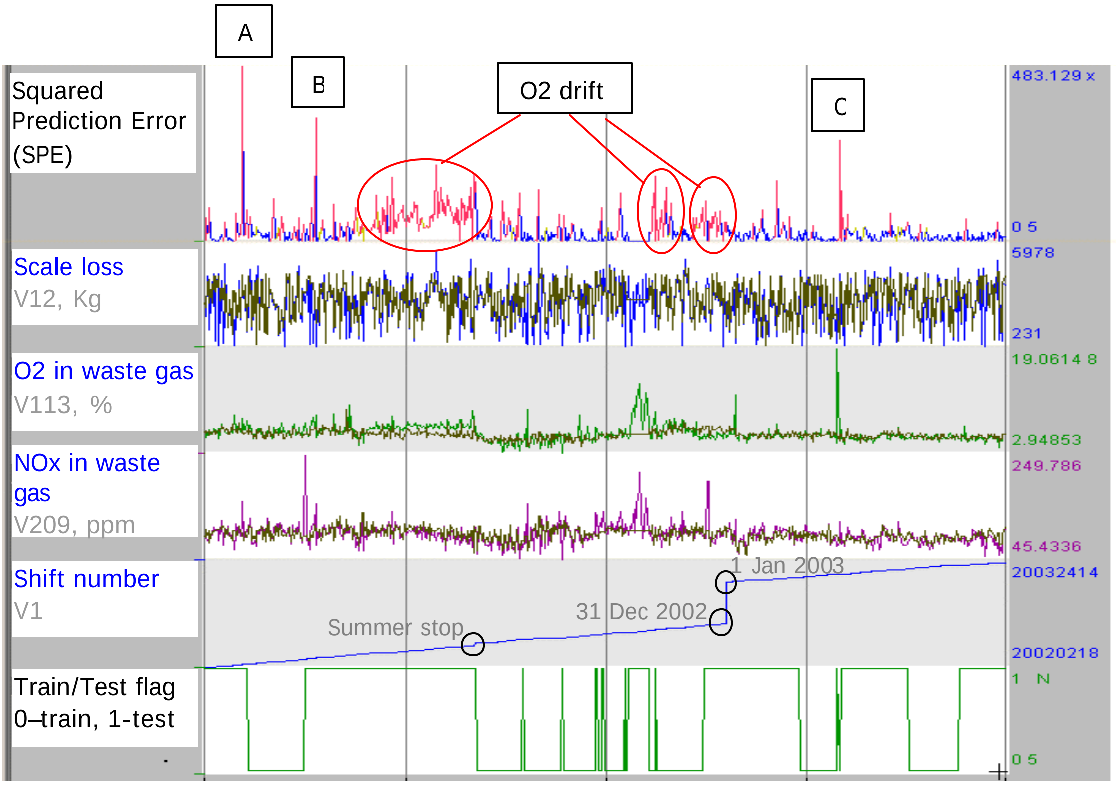
Errors in cause variable estimates and effect variables predictions. Now we have the following possibilities:

1) The model is invalid; there is nothing wrong with the process. Particularly when errors are observed in many variables.

2) The process is behaving abnormally. Where only a small number of cause and effect variables are affected the model may still be valid. Repeated patterns of prediction error could signify a genuine process problem.

The above implies the need to monitor all of the individual estimation and prediction errors. Instead we can monitor a single overall Squared Prediction Error (SPE) that is comprised of the individual errors. From the training data a threshold value for the SPE is determined that indicates a high probability of abnormal process behaviour. Contributions to the SPE from individual variables are readily ranked in order, so that when the SPE threshold value is exceeded the affected variables are easily identified.

Figure 8 shows the SPE value for the scale loss, O2 and NOx model and individual predictions for scale loss, NOx and O2. SPE values plotted in red indicate high probability of abnormal behaviour. As indicated there are three extreme peaks in the SPE value and three periods of persistently high probability of abnormal behaviour. The peak values correspond to abnormally large prediction errors in NOx and O2 – which also occur at high values of NOx and O2. The persistent periods correspond to drift between the O2 measurement and prediction.



A - NOx prediction error -63 ppm, analyser measurement 161 ppm. O2 prediction error -1.9%, analyser measurement 8.4%.

B - NOx prediction error -52 ppm, analyser measurement 156 ppm. O2 prediction error -1.9%, analyser measurement 8.9%.

C - NOx prediction error -37 ppm, analyser measurement 123 ppm. O2 prediction error -2.6%, analyser measurement 8.4%.

Figure 8 – Squared Prediction Error (SPE)

2.3.7 Operating Modes and Product Types

Large changes in the heat content and density of fuel are accompanied by changes in air fuel ratio. These changes are seen as distinct modes of operation by the multivariate model, as shown in Figure 9. This figure plots the first two score variables – the first two 'artificial' variables from the PLS model. In simple terms the score variables capture fundamental process behaviour and so clusters in the scatter plot indicate different operating modes.

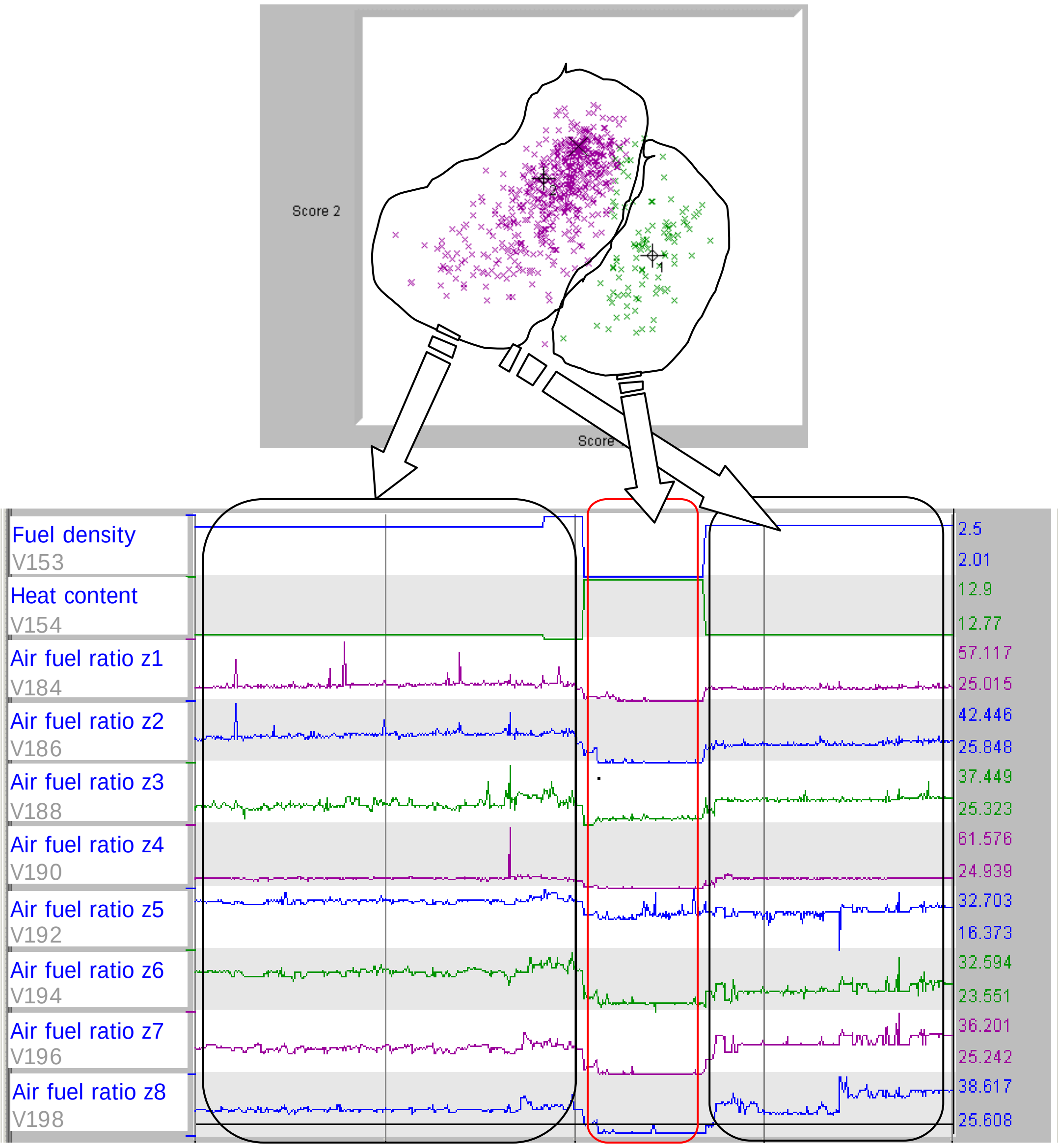


Figure 9 – Operating Modes

Although the work presented in this report is focused on the RSTF product campaigns, initial analysis of all product types was undertaken. A single model for all product types was investigated, but did not show any distinct operational modes that could be identified by product type. However, there are subtle differences in furnace operation between different product types as a model for each product type proved to more accurate than an overall model for all product types.

2.4 Conclusions

Traditional SPC is a data driven approach. This is also a univariate approach since process variables are considered in isolation and strictly speaking they are assumed to be independent of each other and normally distributed. These assumptions are questionable in many industrial processes. Also, an unmanageable number of charts may be required to adequately cover the variables of interest.

In contrast, a multivariate predictive approach uses a multivariable model to predict process behaviour. Process interactions are captured and robust models can be identified from correlated data. Assumptions of independence and normality of process variables are no longer required. Deviations from expected behaviour show up as prediction errors and a single overall composite prediction error is monitored to detect abnormal behaviour. This greatly reduces the monitoring task.

Central to the success of a multivariate predictive approach is a robust and accurate model. Two models have been developed for furnace 301 using a PLS identification method, both of which are accurate and robust. The first model predicts scale loss, O2, NOx, production, total and effective production rates and effective specific energy consumption from 50 cause variables. The second model predicts just scale loss, O2 and NOx from 38 cause variables.

Steam production and total steam production steadily ramp up over time. This integrating behaviour cannot be predicted by a steady state model and so these variables have been removed from the model. Gas and air temperatures located after the recuperator have also been removed from the model as accurate predictions could not be obtained against test data.

3 Analysis

Recall the Shewart charts presented in Figure 2, which are repeated below in Figure 10. For a meaningful comparison with the multivariate predictive approach, there should be one chart for each process variable in the model. That is a total of 54 charts (or 38 charts for the scale loss, O2 and NOx only model). Monitoring such a large number of charts is not a trivial task.

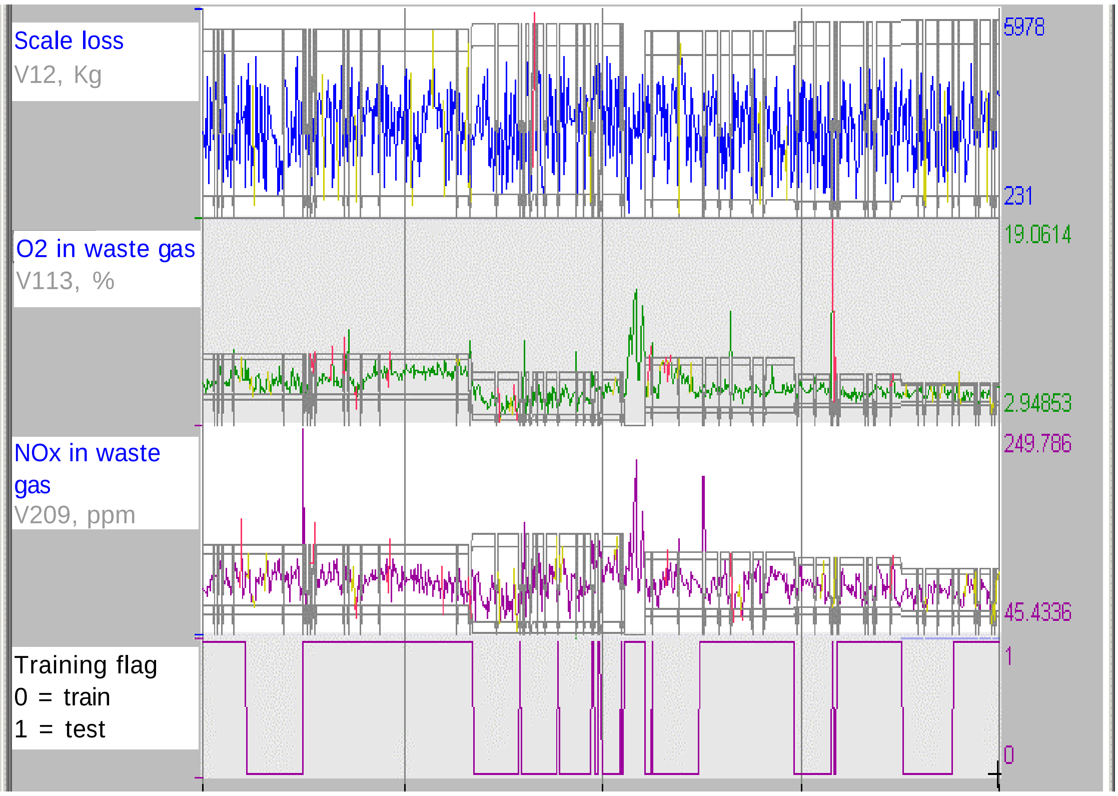


Figure 10 – Shewart charts evaluated for training and testing data

Leaving aside the number of charts required by this univariate approach, do they detect any special cause variation, i.e. any abnormal behaviour? This question is answered in the following sections which compare the univariate and multivariate approaches over periods of long term and short term operation.

3.1 Detecting drift

Recall from Figure 4 that long term drifts were observed between the O2 measurement and the model prediction. The longest period of drift between the O2 analyser measurement and its prediction finishes abruptly where there is a gap in the data set between weeks 27 and 30 in 2002. This also coincides with a step change in steam production as seen in Figure 11. These observations suggest that maintenance was carried out on the furnace between weeks 27 and 30 in 2002. It is likely that the O2 analyser was recalibrated or cleaned and so the measurement falls back in line with the prediction when the furnace is started back up. This apparent drift in the O2 measurement started in week 18 of 2002 and lasted for 9 weeks. By comparison the periods of sustained drift between the O2 analyser measurement and prediction hi-lighted in Figure 4 are not reflected in the Shewart chart for O2, Figure 10. In fairness a CUSUM or X-bar chart would be better suited to detecting a change in average O2 analyser measurement and may well do so in this case. However, even a more appropriate SPC chart cannot tell you whether a shift in the average O2 measurement is an intentional change in operating range or a change in process behaviour. By contrast predictive condition monitoring can make this distinction as O2 predictions will remain accurate at the changed operating range (providing the model is still valid).

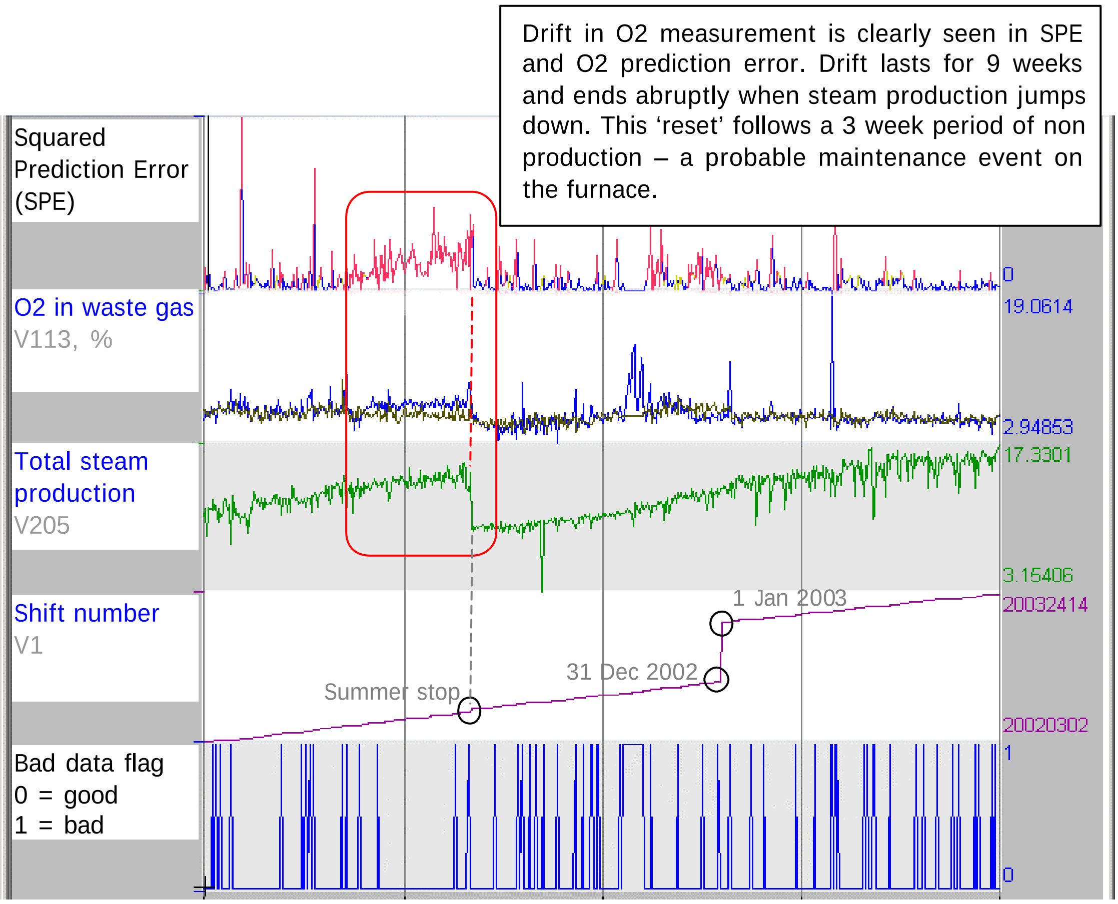


Figure 11 – Drift in O2 measurement

3.2 Detecting Abnormal Events

Changing our focus toward short term special cause variation Figure 12 and Figure 13 illustrate SPC and predictive condition monitoring over a shorter period of 35 samples. To keep the SPC trend, Figure 12, as simple as possible the standard deviation limits are not shown. Instead, excursions outside the limits are simply circled in red. Also, all 38 Shewart charts have been examined and only the variables shown in Figure 12 exceed their limits as indicated over this time span. A solid red vertical line is also drawn to indicate when a limit is exceeded. Each instance is labelled A through J.

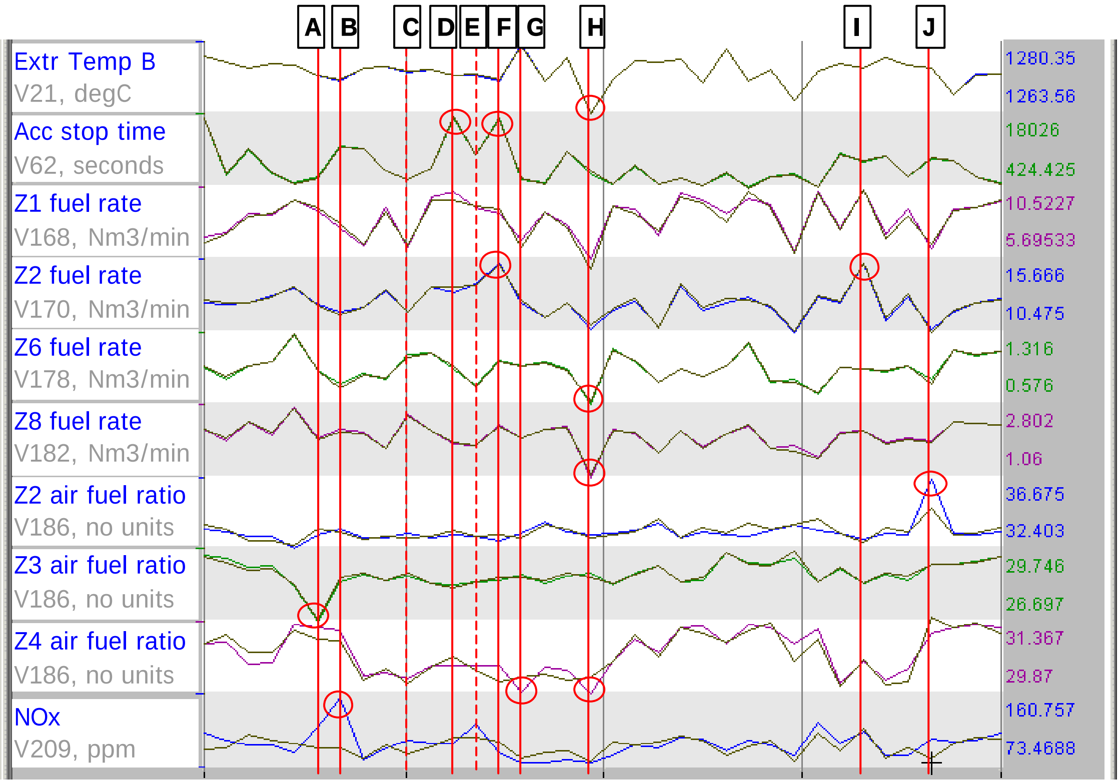


Figure 12 – Special cause variation according to single variable SPC

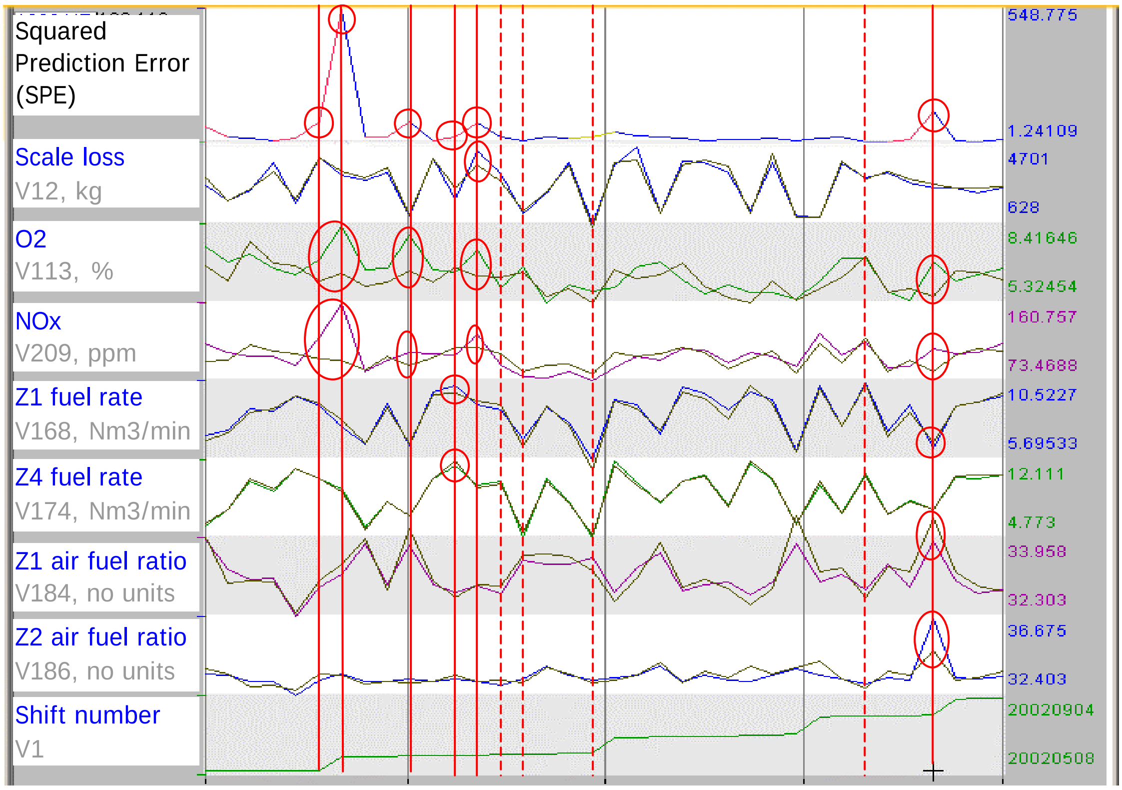


Figure 13 – Abnormal behavior according to Predictive Condition Monitor

The dotted vertical lines labelled 'C' and 'E' signify that all variables are within SPC limits but the predictive condition monitor, Figure 13, has detected abnormal behaviour. Estimates of each variable from the predictive condition monitor are also plotted in brown in Figure 12.

Plotted in Figure 13 are a handful of cause variables in addition to the effect variables (scale loss, O2 and NOx). Estimated values of the cause variables and predicted values of the effect variables from the predictive condition monitor are drawn in brown. The cause variables shown are those that registered the largest estimation errors over the 35 sample window.

Abnormal process behaviour is detected by the predictive condition monitor when the squared prediction error (SPE) exceeds a threshold value. These occurrences are circled in Figure 13, along with the variable(s) that contribute the most to the SPE. A solid vertical line is also drawn in red to indicate the event. The dotted red lines indicate where the process is behaving normally according the predictive condition monitor but special cause variation has been detected by the SPC charts. The vertical lines in Figure 12 and Figure 13 line up with each other. Therefore, events A and B are detected by both the SPC charts and the predictive condition monitor. Event C is detected by the predictive condition monitor but not the SPC charts. Event D is detected by both the SPC charts and the predictive condition monitor Event E is detected by the predictive condition monitor but not the SPC charts. Events F through I are detected by the SPC charts and not the predictive condition monitor. Event J is detected by both the SPC charts and the predictive condition monitor. So, what does this mean? Looking at each event in turn:

- **Events A and B**: Both NOx and O2 increase significantly over two consecutive samples. The condition monitor does not predict these large increases and so flags abnormal behaviour over both samples. At the same time, the condition monitor does not detect any change in correlation between the remaining process variables – they are all estimated accurately. So in all other respects the furnace seems to be behaving normally. In contrast, the SPC charts do not detect any problem with O2 and only detect the high NOx value when it has exceeded its standard deviation limit. A low value of air fuel ration in zone 3 is detected.

- **Event C**: O2 and NOx predictions deviate from analyser measurements, the prediction error in NOX this time is much smaller than previously. The SPC charts do no detect any problem as all variables are within limits.

- **Event D**: The SPC charts detect that accumulated stoppage time is a special cause variation. The predictive condition monitor detects small deviations in average fuel rates to zones 1 and 4.

- **Event E**: Scale loss, O2 and NOx predictions deviate from process values. As in event C the SPC charts do no detect any problem as all variables are within limits.

- **Events F through I**: The SPC charts detect special cause variation in a range of process variables. As Figure 12 and Figure 13 clearly show all process variables are predicted/estimated accurately by the condition monitor and so the process is behaving normally as far as it is concerned.

- **Event J**: A high value of air fuel ratio in zone 2 is obviously detected by the SPC charts. The predictive condition monitor also detects a problem with the same variable as there is a significant estimation error for it. In addition the predictive condition monitor detects abnormal behaviour in the air fuel ratio for zone 1, the average fuel rate in zone 1, O2 and NOx.

3.3 Conclusions

These observations demonstrate a subtle difference between the single variable SPC approach and the predictive condition monitoring approach. Single variable SPC is about detecting changes in single process values. In contrast, predictive condition monitoring is about detecting changes in correlations between multitudes of process variables. Changes in correlation indicate a change in process behaviour and lead to increased prediction error. As we have seen process behaviour can change while individual variables remain within SPC limits. Conversely, individual process values may well exceed SPC limits without affecting process behaviour. These observations lead to the conclusion that predictive condition monitoring is superior to the traditional SPC approach.

4 NOx Estimator

4.1 Objectives

NOx predictions from a computer model can be used to provide estimates of NOx more frequently than analyser measurements, even replacing analyser measurements in some situations. These are real time applications and more commonly known as Predictive Emissions Monitoring Systems or PEMS. PEMS are usually subject to strict accuracy requirements stipulated by government environmental protection agencies. The accuracy requirement is commonly a relative accuracy test.

The data analysed in this report are average/accumulated values per campaign per shift and so the NOx predictions obtained from the models are not suitable for real time use. However, a relative accuracy test is carried out below to provide another measure of the accuracy and robustness of the NOx predictions obtained.

4.2 Relative Accuracy Test for Predictive Emissions Monitoring Systems

The NOx predictions are taken from the scale loss, O2 and NOx model. The relative accuracy is a measure of average prediction accuracy subject to a required confidence interval (e.g. 95% confidence) and is typically calculated as below. See Appendix 2 – Calculation of Relative Accuracy for a full description.

$$Relative\,Accuracy = \frac{|average\ prediction\ error| + confidence\ interval}{average\ reference\ value\ OR\ \ emission\ limit}$$

The $average\ reference\ value$ is the average value of the analyser measurement over the same period that the prediction error is averaged. Detailed performance specifications for PEMS are set out by the appropriate authorities. By way of example, a good description can be found on the American Environmental Protection Agency website http://www.epa.gov/ttn/emc/cem/pems.pdf. This performance specification sets a relative accuracy limit of 20%.

The relative accuracy calculation described in the appendix has been applied using a sample size of 30. Figure 14 shows the relative accuracy calculated for the training data and test data using a moving window of 30 samples. From this plot we can see that the least accurate period of training data gave a relative accuracy of 11.7%. The test data peaks at fractionally over this value (12.0%), the vast majority of test data being comfortably inside 11.7% accuracy. We can safely conclude that the NOx prediction is robust – similar levels of accuracy are obtained with the testing and training data.

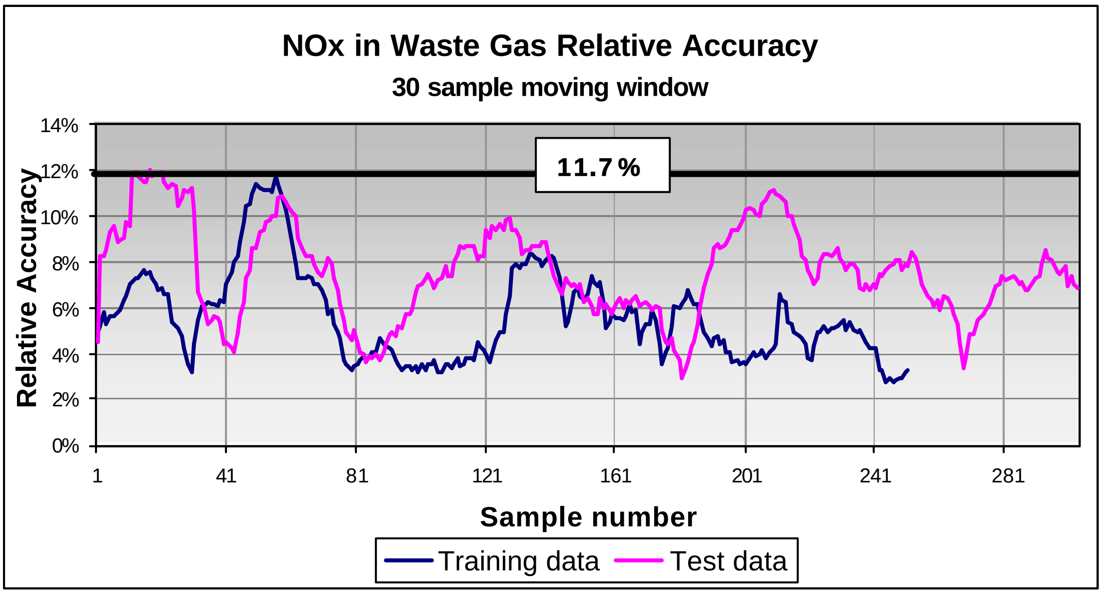


Figure 14 – Relative accuracy of NOx prediction

4.3 Conclusions

Although the NOx predictions do not relate to real time data it is encouraging to see an accuracy of 11.7% in light of a typical PEMS requirement of 20%.

5 Conclusions and Future Work

This initial analysis has shown that a versatile predictive condition monitor can be developed for a steel reheating furnace. The findings have demonstrated the ability to detect abnormal process events and identify subtle deviations from normal process behaviour. In contrast, traditional SPC can fail to detect deviation from normality and conversely indicate special cause variation while the process is behaving normally. A subtle difference between the two approaches is that predictive condition monitoring detects changes in *correlations* between multitudes of process variables whereas SPC detects deviation in absolute *values* in *single* variables. These observations lead to the conclusion that predictive condition monitoring is superior to the traditional SPC approach

The condition monitoring capability is delivered by a sophisticated multivariable model that gives robust predictions in the presence of highly correlated noisy production data. The variables predicted include environmental emissions, in particular NOx, and other important economic variables such as scale loss, production rate and specific energy consumption. These estimated parameters could be used to independently backup sensors or could be included as part of an enhanced control system.

Although this analysis has been concerned with averaged/accumulated data per campaign per shift the NOx prediction has been shown to be robust and a high level of relative accuracy has been demonstrated. Relative accuracy is a commonly used performance measure of real time predictive emissions monitoring systems.

Implementation of this predictive condition monitor would provide operational staff with a versatile tool for post production condition monitoring and NOx estimation. Abnormal events would not be captured in real time owing to the nature of the data upon which this analysis has been carried out. However, the ability of the predictive condition monitor to determine which variables have contributed the most to a deviation from normality would in itself be extremely useful. Furthermore, process and analyser drift over consecutive campaigns would easily be detected, giving operational staff early warning before this leads to sub optimal production or a possible fault condition.

Serious consideration should be given to the development of a predictive condition monitor and predictive emissions monitoring system based on real time data. Some of the variables considered in this analysis may not be suitable for real time consideration such as production, total and effective production rates and effective specific energy consumption as they may be too variable over a short interval. However, this still leaves the analyser measurements and the large number of cause variables to monitor.

Finally, repeatable patterns of prediction error that occur when abnormal behaviour is detected may indicate specific process issues. This idea merits further investigation as a possible diagnostic tool.

Appendix 1 – Modifications to Cause Effect Structure of Model

Removal of stoppage variables

Effect variables removed:
- V114 O_2 in waste gas stop
- V116 NOx in waste gas stop
- V98 Air temp after recuperator stop
- V104 Gas temp after recuperator stop

Cause variables removed:
- V169 Average fuel z1 stop
- V171 Average fuel z2 stop
- V173 Average fuel z3 stop
- V175 Average fuel z4 stop
- V177 Average fuel z5 stop
- V179 Average fuel z6 stop
- V181 Average fuel z7 stop
- V183 Average fuel z8 stop
- V185 Air fuel ratio z1 stop
- V187 Air fuel ratio z2 stop
- V189 Air fuel ratio z3 stop
- V191 Air fuel ratio z4 stop
- V193 Air fuel ratio z5 stop
- V195 Air fuel ratio z6 stop
- V197 Air fuel ratio z7 stop
- V199 Air fuel ratio z8 stop
- V100 Gas temp furnace stop
- V102 Gas temp before recuperator stop
- V120 FCE pressure stop
- V157 sum cooling air recuperator stop

Removal of roll temperatures

Cause variables removed:
- V34 Roll temperature N
- V37 Roll temperature S
- V40 Roll temperature B

Normalisation of NOx

The NOx value predicted in the original models is the un-normalised value 115.ME. This has been replaced by a normalised value V209, calculated as below:

Effect variable added:
- V209 Normalised NOx production = V115 * 16 / (21 – V113)

Removal of total specific energy consumption and % scale loss

Effect variables removed:
- V162 Total specific energy consumption
- V163 % scale loss

Effect variables added:
- V12 Accumulated scale loss

Cause variables added:
- V161 Total fuel consumption

Removal of steam production

Effect variables removed:
- V204 Steam production
- V205 Total steam production

Removal of temperatures after the recuperator

Cause variables removed:
- V204 Steam production
- V205 Total steam production

Revised cause – effect model structure

Cause Variables		Effect Variables	
V4	No. of slabs	V12	Accumulated scale loss (kg)
V5	No. of slabs north	V113	O_2 in waste gas production (%)
V6	No. of slabs south	V209	NO_x in waste gas production (ppm)
V201	% slabs north	V160	Production (ton)
V202	% slabs south	V206	Total production (t/h)
V203	% slabs both	V207	Eff. production (t/h)
V7	Slab weight long	V208	Eff. energy cons (t/h)
V8	Slab weight short		
V9	Slab length long		
V10	Slab length short		
V11	% slab length		
V13	Hearth area ratio		
V14	Slab temp gradient		
V15	Extraction temp north		
V18	Extraction temperature south		
V21	Extraction temperature both		
V25	No of UV data		
V26	No. of UV data north		
V27	No. of UV data south		
V61	Accumulated production time		
V62	Accumulated stop time		
V167	% production time		
V168	Average fuel z1 production		
V170	Average fuel z2 production		
V172	Average fuel z3 production		
V174	Average fuel z4 production		
V176	Average fuel z5 production		
V178	Average fuel z6 production		
V180	Average fuel z7 production		
V182	Average fuel z8 production		
V184	Air fuel ratio z1 production		
V186	Air fuel ratio z2 production		
V188	Air fuel ratio z3 production		
V190	Air fuel ratio z4 production		
V192	Air fuel ratio z5 production		
V194	Air fuel ratio z6 production		
V196	Air fuel ratio z7 production		
V198	Air fuel ratio z8 production		
V99	Gas temp furnace production		
V101	Gas temp before recuperator production		
V109	Accumulated fuel meter production		
V110	Accumulated fuel meter stop		
V119	FCE pressure production		
V149	Sum nedstyr zone 1 recuperator production		
V151	Sum nedstyr zone 2 recuperator production		
V153	Sum fuel density		
V154	Sum fuel heat content		
V155	Sum fuel stochiometric value		
V156	Sum cooling air recuperator production		
V161	Fuel		

Table 3 – Revised cause and effect variable structure

Cause – effect model structure for scale loss, O2 and NOx only

Cause Variables		Effect Variables	
V14	Slab temp gradient	V12	Accumulated scale loss (kg)
V15	Extraction temp north	V113	O_2 in waste gas production (%)
V18	Extraction temperature south	V209	NO_x in waste gas production (ppm)
V21	Extraction temperature both		
V61	Accumulated production time		
V62	Accumulated stop time		
V167	% production time		
V168	Average fuel z1 production		
V170	Average fuel z2 production		
V172	Average fuel z3 production		
V174	Average fuel z4 production		
V176	Average fuel z5 production		
V178	Average fuel z6 production		
V180	Average fuel z7 production		
V182	Average fuel z8 production		
V184	Air fuel ratio z1 production		
V186	Air fuel ratio z2 production		
V188	Air fuel ratio z3 production		
V190	Air fuel ratio z4 production		
V192	Air fuel ratio z5 production		
V194	Air fuel ratio z6 production		
V196	Air fuel ratio z7 production		
V198	Air fuel ratio z8 production		
V99	Gas temp furnace production		
V101	Gas temp before recuperator production		
V109	Accumulated fuel meter production		
V110	Accumulated fuel meter stop		
V119	FCE pressure production		
V149	Sum nedstyr zone 1 recuperator production		
V151	Sum nedstyr zone 2 recuperator production		
V153	Sum fuel density		
V154	Sum fuel heat content		
V155	Sum fuel stochiometric value		
V156	Sum cooling air recuperator production		
V161	Fuel		

Table 4 – Cause and effect variable structure for scale loss, O2 and NOx model.

Appendix 2 – Calculation of Relative Accuracy

Define:

d_i	difference between the analyser measurement and the predicted value for the i^{th} sample.
n	number of samples
$\overline{M}$	mean value of analyser measurement over the n samples
$\overline{d}$	mean difference between the analyser measurement and the predicted value
S_d	sample standard deviation of the difference between the analyser measurement and the predicted value
CI	confidence interval
t	t-statistic
RA	relative accuracy

The mean and sample standard deviation are calculated as below:

$$\overline{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$$

$$S_d = \sqrt{\frac{\sum_{i=1}^{n}\left(d - \overline{d}\right)^2}{n-1}}$$

To calculate the confidence interval the t-statistic t is obtained from a table found in most standard statistical resources. The t value is a function of the number of samples n and the desired confidence. 95% confidence is required here in the NOx relative accuracy calculation. The t value is the two-tailed value.

$$CI = t\frac{S_d}{\sqrt{n}}$$

The relative accuracy is then calculated as follows. In some situations the emissions limit may be used in place of $\overline{M}$:

$$RA = \frac{\left|\overline{d}\right| + CI}{\overline{M}}$$

Appendix 3 – Calculated Variables

The data set supplied for Furnace 301 contains 159 variables. The following additional variables were calculated:

V160	ProduktionTotal	V7*(V4-V5-V6) +V8*(V5+V6)	Tonne
V161	FuelEnergyTotal	(V109+V110)*V153*V154	kWh
V162	EnergyConsumptionProd	V162/V160	kWh/Tonne
V163	Scale loss (%)	100*V12/V160/10	%
V164	AverExtrTempDevN (oC)	V16/V5	oC
V165	AverExtrTempDevS (oC)	V19/V6	oC
V166	AverExtrTempDevB (oC)	V22/(V4-V5-V6)	oC
V167	ProdTime	V61/(V61+V62)*100	%
V168	AverFuelZ1Prod	V63/V61*60	Nm3/min
V169	AverFuelZ1Stop	V64/V62*60	Nm3/min
V170	AverFuelZ2Prod	V65/V61*60	Nm3/min
V171	AverFuelZ2Stop	V66/V62*60	Nm3/min
V172	AverFuelZ3Prod	V67/V61*60	Nm3/min
V173	AverFuelZ3Stop	V68/V62*60	Nm3/min
V174	AverFuel Z4Prod	V69/V61*60	Nm3/min
V175	AverFuelZ4Stop	V70/V62*60	Nm3/min
V176	AverFuelZ5Prod	V71/V61*60	Nm3/min
V177	AverFuelZ5Stop	V72/V62*60	Nm3/min
V178	AverFuelZ6Prod	V73/V61*60	Nm3/min
V179	AverFuelZ6Stop	V74/V62*60	Nm3/min
V180	AverFuelZ7Prod	V75/V61*60	Nm3/min
V181	AverFuelZ7Stop	V76/V62*60	Nm3/min
V182	AverFuelZ8Prod	V77/V61*60	Nm3/min
V183	AverFuelZ8Stop	V78/V62*60	Nm3/min
V184	AirFuelRatioZ1Prod	V79V63	dimensionless
V185	AirFuelRatioZ1Stop	V80/V64	dimensionless
V186	AirFuelRatioZ2 Prod	V81/V652	dimensionless
V187	AirFuelRatioZ2 Stop	V82/V66	dimensionless
V188	AirFuelRatioZ3 Prod	V83/V67	dimensionless
V189	AirFuelRatioZ3 Stop	V84/V68	dimensionless
V190	AirFuelRatioZ4 Prod	V85/V69	dimensionless
V191	AirFuelRatioZ4 Stop	V86/V70	dimensionless
V192	AirFuelRatioZ5 Prod	(V87/V71	dimensionless
V193	AirFuelRatioZ5 Stop	V88/V72	dimensionless
V194	AirFuelRatioZ6 Prod	V89/V73	dimensionless
V195	AirFuelRatioZ6 Stop	V90/V74	dimensionless
V196	AirFuelRatioZ7 Prod	V91/V75	dimensionless
V197	AirFuelRatioZ7 Stop	V92/V76	dimensionless
V198	AirFuelRatioZ8 Prod	V93/V77	dimensionless
V199	AirFuelRatioZ8 Stop	V94/V78	dimensionless
V200	ProdTime	V61/(V61+V62)*100	%
V201	SlabsN	V5/V4*100	%
V202	SlabsS	V6/V4*100	%
V203	SlabsB	(V4-V6-V5)/V4*100	%

V204	SteamProd	(V123/V61)*3600	Tonne/h
V205	SteamTotal	((V123+V124)/(V61+V62))*3600	Tonne/h
V206	Total Production Rate	V160/(V61+V62)/3600	Tonne/h
V207	Effective Production Rate	V160/V61/3600	Tonne/h
V208	Effective Specific Energy Consumption	V109*V153*V154/V160	kWh/Tonne
V209	NOx normalised	V115*16/(21–V113)	ppm