MULTIVARIATE DATA ANALYSIS: INDUSTRIAL APPLICATIONS

MULTIVARIATE DATA ANALYSIS FOR PROCESS EVALUATION, PREDICTION AND MONITORING AT INCO'S COPPER CLIFF SMELTING AND REFINING OPERATIONS

By

JENNIFER BRADLEY, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

In Partial Fulfillment of the Requirements

For the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (2006) McMaster University

(Chemical Engineering) Hamilton, Ontario

TITLE: Multivariate Data Analysis for Process Evaluation, Prediction and Monitoring at INCO's Copper Cliff Smelting and Refining Operations

AUTHOR: Jennifer Bradley, B.Sc. (Queen's University)

SUPERVISORS: Dr. J.F. MacGregor
Dr. Theodora Kourti

NUMBER OF PAGES: viii, 105

Abstract

Industrial processes generate large quantities of process and product quality data. Most of this data is stored and is analyzed in a univariate fashion. However important information may be lost through the implementation of univariate analysis methods. This information is contained in the correlation structure amongst the process and product quality variables and between these two types of variables. Through multivariate analysis this information is retained. As a result process evaluation, prediction and monitoring are more effectively performed.

Multivariate data analysis techniques were therefore applied to data sets that summarized three of INCO's Copper Cliff smelting and refining processes. In the first instance the analysis of historical data pertaining to a batch leaching process was undertaken and the time required for leaching was predicted. In the second a multivariate soft sensor was developed in order to predict the concentration of nitric oxide contained in the feed gas to the smelter acid plant. The final project involved the analysis and monitoring of a continuous nickel carbonyl process.

The resulting models were evaluated and significant variables with respect to the variation in the process and product quality data and the correlation between them were identified. The product quality data was also well predicted using new process data only that was input to the models. Finally new data was input to the models and the process was monitored using a reduced number of latent variables. Contribution plots were used to identify the original variables that contributed most to the observations that exceeded the established control limits.

Acknowledgements

Without the guidance and support of several people this work would not have been achieved. First I would like to thank INCO Ltd. and McMaster University for allowing me the opportunity to complete this degree. At INCO I would like to thank all of the people I have worked with over the last five years who took the time to pass the knowledge they have accumulated on to me. Particularly I would like to thank Dr. Ken Scholey who fully supported my scholastic efforts with the right mix of encouragement and intelligent question. At McMaster I would like to thank my advisors Dr. John MacGregor and Dr. Dora Kourti who are two of the best examples of teachers that I have ever encountered. The ability to clearly and concisely present material while assisting students in reaching their full potential is truly a gift.

I would like to thank my family who has been a constant and unwavering source of support. I would like to thank my original and best teacher – my mother. Any success I have ever achieved is a direct result of the parenting I have received from her and my late father. Finally I would like to thank my husband for never doubting me – even when I did. The challenges we have faced and will continue to face in the future are made easier through our partnership.

Table of Contents

Table of Figures

1. Chapter 1 – Introduction

During the operation of industrial plants vast quantities of data related to both the processes and the quality of the products produced are generated and stored and INCO's Copper Cliff Smelting and Refining plants are no exception. In general many of the process variables are correlated with one another, as is the case for the product quality variables. In addition the process and product quality variables may be correlated with one another. Important information regarding the state of the operation is contained in the correlation structure between the process and product quality variables. When univariate methods are used for the purpose of process evaluation, prediction and monitoring the information contained in the correlation structure is lost. However when multivariate methods such as principle component analysis (PCA) and partial least squares projection the latent structures (PLS) are implemented this information is retained.

In addition through PCA and PLS, multivariate processes are summarized by a reduced number of new or latent variables. Upon evaluation of the latent variable model the original variables that are significant with respect to the variation in both the process and product quality may be identified. Accurate prediction of the quality variables may also be achieved since the variability in the process and product quality data are modeled along with the correlation structure between the data. Finally multivariate processes may be monitored in real time using PCA and PLS. Since the multivariate process is summarized by a reduced number of variables, these new variables may be monitored as opposed to the monitoring of a large number of the original process variables.

Multivariate methods were applied to data obtained from three INCO plants located in Copper Cliff Ontario. In the first instance PLS was used to predict the time required for oxygen pressure leaching in one of two autoclaves. In addition the batch make-up and autoclave process variables included in the model were evaluated in order to determine those that contributed significantly to the variation in leach time. In general the predicted and actual leach times corresponded well with increased error observed at the most elevated leach times. With respect to variable significance, several groups were identified and included the autoclave in which the leach took place, the density of the autoclave leach slurry, the batch make-up characteristics and finally the autoclave process variables

related to oxygen and cooling water flow. Finally several variables assumed to be significant prior to the analysis, such as batch make-up tank level indicator type and the concentration of certain elements contained in the solids to be leached were found to be insignificant.

In the second instance PLS was used for the development of a soft-sensor. The results of a temporary gas analysis campaign were combined with process data related to three pyrometallurgical sources and PLS was applied. The resulting model predicted a significant amount of the variation in the nitric oxide (NO) concentration in the feed gas treated by the smelter acid plant. In addition several variables related to the pyrometallurgical processes were identified as significant with respect to the variation in NO concentration. Several of the variables initially assumed to contribute to the variation, such as oxy-fuel burner use were shown to be significant. However a number of other variables related to the ratio of oxygen to solids charged to the furnace and to furnace maintenance work were also deemed significant.

The final project involved the PLS analysis of process and product quality variables related to the operation of a nickel powder decomposer reactor. Two nickel powder physical characteristics were predicted based on the decomposer operating conditions and in general the predicted and actual values corresponded well. In addition the process variables that contribute to the variation in the powder physical characteristics were identified. As expected the six variables manipulated by plant operations personnel in order to affect changes in the powder characteristics were shown to be significant. The correlation between the process variables was also displayed.

Finally through PLS modeling of the nickel powder decomposer process two specific modes of operation were identified. Individual PLS models were fit to the observations contained in each of the two groups. Following extreme outlier removal these PLS models were used as the basis for the good operation of the decomposer under each of the two modes. As such new data was input to the PCA models developed and the process was monitored in a multivariate manner. Based on a number of control charts that was small compared to the number of process variables included in the model, process outliers were identified. Contribution plots were then used to determine those variables that contributed to the process variability.

2. Chapter 2 – Overview of PCA and PLS

Enormous amounts of data are typically generated through industrial processes. Data relating to hundreds or thousands of measured process variables may be recorded and stored using data historian software. In many instances the data is analyzed through visual inspection or through the plotting of a few variables on graphs. However multivariate data analysis is increasingly being used in order to extract meaningful information from industrial databases. The data contained in these sets are typically highly correlated and the amount of information contained in any single measurement is low. Through multivariate data analysis methods such as principle component analysis (PCA) or partial least squares projection to latent structures (PLS) the movement of industrial processes over time may be summarized using a small number of reduced or latent variables. The latent variable models may then be investigated in order to determine the relationships between groups of observations and between groups of variables contained in the data set. Finally through PLS the prediction of one group of data such as product quality may be achieved based on another group, typically process data. The resulting multivariate prediction model may then be investigated in order to determine the process variables that contribute most to the variation in product quality.

2.1. Geometric Interpretation

For the case of principle component analysis (PCA) a data matrix X with K variables and N observations may be depicted graphically. A K dimensional space called X-space is defined whereby each variable corresponds to a unique coordinate axis, the length of which is determined by scaling. When the observations are plotted in X-space a swarm of points is created. The first principal component is defined as a line in X-space which accounts for the greatest amount of variation in the data swarm and which passes through the average point. The second principal component is defined as a line orthogonal to the first principal component that explains the second greatest amount of variation in the data and also passes through the average point. Additional components may be added in order to increase the amount of explained variability. The number of principle components may be determined through cross validation whereby a portion of the data is not included in the model development (Wold 1978). The data not included in model development is then

predicted and is compared with the actual values. Components are added until a balance is achieved between goodness of fit and over fitting.

Together the principal components define a plane or hyperplane in X-space and the loading vectors (p_1, p_2, p_3, etc.) define the orientation of the plane with respect to the original K variables. Score values (t_{11}, t_{12}, t_{13}, etc.) are defined as the projection of an observation onto the plane. Based on the variability in the data matrix X, a confidence interval is calculated in order to determine which scores may be considered outliers. Finally the perpendicular distance between an observation and the plane is defined as the distance to the model in X-space (DMODX) and defines the model residual. Therefore through PCA it is possible to summarize a large multivariate data set using a reduced number of new variables, namely the loadings and scores. A graphical depiction of PCA is given below for a case with three original variables.

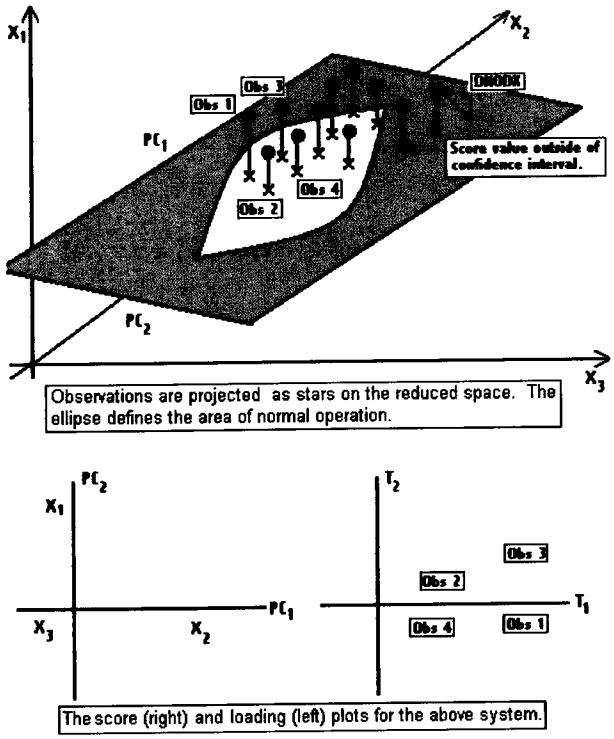

Figure 1: Geometric Overview of PCA with Three Principle Components

For the case of partial least squares projection the latent structures (PLS) a data matrix X with K variables and N observations may be depicted graphically along with a data matrix Y with M variables and N

observations. A K dimensional space called X-space is defined whereby each X-variable corresponds to a unique coordinate axis, the length of which is determined by scaling. When the observations are plotted in X-space a swarm of points is created. Similarly an M dimensional space called Y-space is defined whereby each Y-variable corresponds to a unique coordinate axis of length determined by scaling. When the observations are plotted in Y-space a swarm of points is created. The first principal component is defined as a line in X-space and a line in Y-space which passes through the average points, accounts for the greatest amount of variation in the data swarms and which correlates the projections of the points in the X and Y-spaces well. The second principal component is defined as a line orthogonal to the first principal component which passes through the average points, explains the second greatest amount of variation in the data and which improves the correlation between the projections of X and Y. Principle components may be added in order to increase the extent to which the variation in the X and Y-spaces is explained and to improve the correlation between their projections.

The principle component lines in X-space are orthogonal to one another but the lines in Y-space are not. Together the principal components define two planes or hyperplanes. The loading vectors (p_1, p_2, p_3, etc.) define the orientation of the plane in X-space with respect to the original K variables. The weight vectors (c_1, c_2, c_3, etc.) define the orientation of the plane in Y-space with respect to the original M variables. X-score values (t_{11}, t_{12}, t_{13}, etc.) are defined as the projection of an observation onto the plane in X-space and the Y-score values (u_{11}, u_{12}, u_{13} etc.) are defined as the projection of an observation onto the plane or in Y-space.

2.2. Mathematical Interpretation

A mathematical overview of PCA is given as (Wold 1987):

$$\mathbf{X = TP' + E}$$

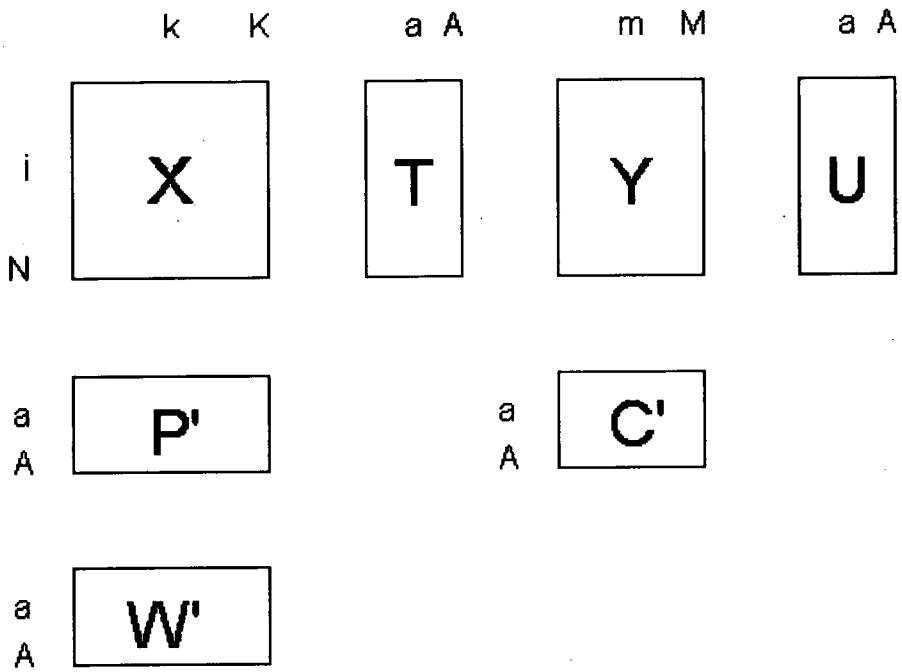

Figure 2: Overview of PCA and PLS Matrices

Where:

X = a data matrix comprised of:
i = 1…N observations
k = 1…K variables
T = a score matrix comprised of:
a = 1…A score vectors t
P' = a loading matrix comprised of:
a = 1…A loading vectors p'
E = a residual matrix
A = the number of model principle components

The score values are linear combinations row-wise of the matrix X and the loading vector p. They are weighted averages of all of the values in the row and result in the greatest explanation of variation. For example for the i^{th} observation and the first principle component:

$$t_{1i} = \mathbf{x_i}'\mathbf{p_1}$$
$$= x_{i1}*p_{11} + x_{i2}*p_{12} + x_{i3}*p_{13} + \ldots + x_{iK}*p_{iK}$$

The score vectors are the eigenvalues of XX^T and model the covariance of X (Hoskuldsson 1988).

The loading values are linear combinations column wise of the matrix X and the score vector t. They are weighted averages of all of the values in the column. For example for the k^{th} variable and the first principle component:

$$p_{1k} \propto \mathbf{t_1}'\mathbf{x_k}$$
$$\propto t_{11}*x_{1k} + t_{12}*x_{2k} + t_{13}*x_{3k} + \ldots + t_{1N}*x_{NK}$$

The loading value p_{1k} gives the importance of variable k to the 1^{st} principle component while the loading vector $\mathbf{p_1}$ gives the direction of greatest variation. The loading vectors are the eigenvectors of X^TX.

Similarly a mathematical overview of PLS is given as (Geladi and Kowalski 1986):

$$\mathbf{Y = TQ' + F}$$

Where:

X = a data matrix comprised of:
- i = 1…N observations
- k = 1…K X-variables

Y = a data matrix comprised of:
- i = 1…N observations
- m= 1…M Y-variables

T = an X-score matrix comprised of:
- a = 1…A X-score vectors t

U = a Y-score matrix comprised of:
- a = 1…A Y-score vectors u

P' = an X-loading matrix comprised of:
- a = 1…A Y-loading vectors p'

W' = an X-weight matrix comprised of:
- a= 1…A X-weight vectors w'

C' = a Y-weight matrix comprised of:

a = 1...A Y-weight vectors c'
F = a residual matrix
A = the number of model principle components

The Y-score values are linear combinations row wise of the matrix X and the X-weight vector w. They are weighted averages of all of the values in the row and result in the greatest explanation of variation in the Y-data. For example for the i^{th} observation and the first principle component:

$$u_{1i} = \mathbf{x_i}'\mathbf{w_1}$$
$$= x_{i1}*w_{11} + x_{i2}*w_{12} + x_{i3}*w_{13} + \ldots + x_{iK}*w_{iK}$$

The Y-score vectors are the eigenvalues of YY^TXX^T and model the covariance of Y.

The X-weight values are linear combinations column wise of the matrix X and the Y-score vector u. They are weighted averages of all of the values in the column. For example for the k^{th} variable and the first principle component:

$$w_{1k} \propto \mathbf{u_1}'\mathbf{x_k}$$
$$\propto u_{11}*x_{1k} + u_{12}*x_{2k} + u_{13}*x_{3k} + \ldots + u_{1N}*x_{NK}$$

The weight value w_{1k} gives the importance of variable k to the 1^{st} principle component while the weight vector $\mathbf{w_1}$ gives the direction of greatest variation. The X-weight vectors are the eigenvalues of X^TYY^TX and maximize the covariance between X and Y.

The Y-weight values are linear combinations column wise of the matrix Y and the X-score vector t. They are weighted averages of all of the values in the column. For example for the k^{th} variable and the first principle component:

$$c_{1m} \propto \mathbf{t_1}'\mathbf{y_m}$$
$$\propto t_{11}*y_{1m} + t_{12}*y_{2m} + t_{13}*y_{3m} + \ldots + t_{1N}*x_{NM}$$

The weight value c_{1m} gives the importance of variable m to the 1^{st} principle component while the weight vector $\mathbf{c_1}$ gives the direction of greatest variation. The Y-weight vectors are the eigenvalues of YY^TXX^T.

Once the a = 1…A component PLS model is formed, the new Y-matrix may be calculated based on a new X-observation collected over k = 1…K variables and the model parameters (Kourti 2005). The calculation steps are as follows:

1. The average ($xbar_k$) and standard deviation ($std(x_k)$) for each of the K variables contained in the X-matrix are calculated.
2. The average ($ybar_m$) for each of the M variables contained in the Y-matrix is calculated.
3. The new X-observation: x'_{new} is scaled and centered as follows:
$$x^{*'}_{new} = (x'_{new} - xbar_k)/ (std(x_k))$$
4. The new X-scores are then calculated using the model X-weights and the new X-observation for all A components:
$$t_{anew} = w'_a x_{anew}$$
$$T_{new} = t_{1new} \ldots t_{Anew}$$
For the first component $x_{1new} = x^*_{new}$ and for a>1:
$$x_{anew} = x_{(a-1)new} - t_{(a-1)new} * p_{(a-1)}$$
5. The new Y-matrix is calculated using the new X-scores and the model Y-weights:
$$Y_{new} = Ybar_M + T_{new} C'$$

2.3. PCA and PLS Modeling Diagnostics

For PCA the extent to which the variation in the X-data is explained by the model is given by the R^2X value. The amount of explained variation increases as this value approaches one and the expression for R^2X is given as:

$$R^2X = 1 - RSS/SSX_{tot.corr.}$$

Where:
RSS = the residual sum of squares
$SSX_{tot.corr.}$ = the total variation in the X-matrix after mean centering

The predictive ability of the model is given by the Q^2 value. This value indicates the extent to which data not used to construct the model may be predicted and the expression for Q^2 is given as (Eriksson et. al. 1999):

$$Q^2 = 1 - PRESS/ SSX_{tot.corr.}$$

Where:
PRESS = the predicted residual sum of squares
$SSX_{tot.corr.}$ = the total variation in the X-matrix after mean centering

R^2X and Q^2 are therefore summary statistics that reveal how well the X-data is fit by the PCA model.

In order to determine the amount of variation explained by each variable the R^2VX statistic is evaluated. Once again, as this value approaches one the explanation of variance for the variable increases. The amount of variation explained by each component may be examined for each variable by plotting the component R^2VX values versus the variables. The expression for the cumulative explained variation after the A^{th} component of a variable is given as (Eriksson et. al. 1999):

$$R^2VX(cum) = [1-SSVX[A]]/[SSVX[0]]$$

Where:
SSVX = the residual variable variance
$= \sum_i e_{ik}^2$

R^2X and Q^2 are used along with the R^2Y to summarize the extent to which the X and Y-data are fit by the PLS model. R^2Y is calculated similar to R^2X and outlines the extent to which the variation in the Y-data is explained by the model. The amount of variation in the Y-data accounted for by each variable, R^2VY is also calculated similar to R^2VX and may be examined for changes with the inclusion of additional model components.

The X-scores may be plotted against one another for the purpose of identifying outliers in the X-data. For example t_1 vs. t_2, t_1 vs. t_3 and t_2 vs. t_3 may be plotted against one another and strong outliers may be identified. These strong outliers are observations that lie well beyond the elliptical confidence region, the boundary of which is calculated based on the Hotelling's T^2 statistic. An example of strong outliers in the t_2 vs. t_3 plot is shown below.

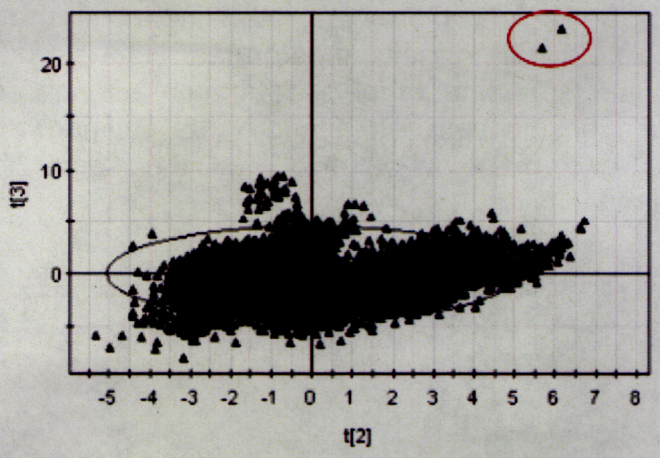

Figure 3: Extreme Outliers Example on t_2 vs. t_3 X-Score Plot

The Y-scores may also be plotted against one another for examination. For example u_1 vs. u_2 and may be plotted in order to identify strong outliers with respect to the Y-data. The X and Y-scores (t_1 vs. u_1) may also be plotted against one another in order to identify outliers in the relationship between X and Y. The slope of this plot may also be used to gauge the extent to which the correlation between X and Y is modeled. As the slope of the t_1 vs. u_1 plot approaches one, the modeling of the correlation improves. Examples of outliers in the relationship between X and Y as outlined on the t_1 vs. u_1 plot and of outliers in the Y-data as outlined on the u_1 vs. u_2 plot are given below.

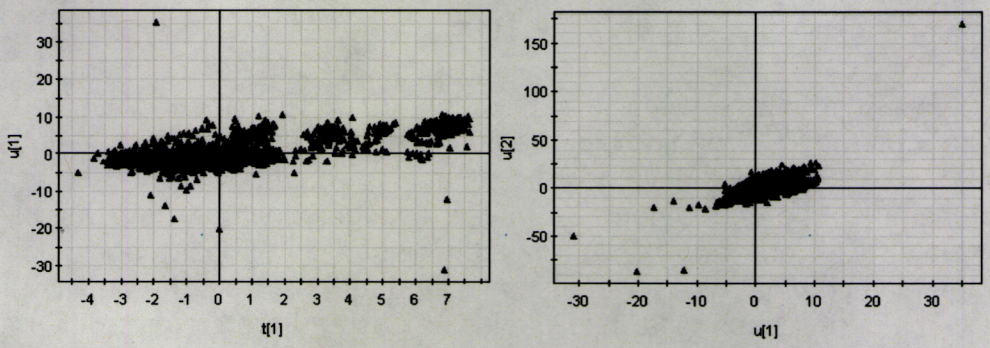

Figure 4: Outlier Examples on t_1 vs. u_1 Plot (left) and u_1 vs. u_2 Y-score Plot (right)

The X and Y-score plot (t_1 vs. u_1) may also be used to identify non-linearity with respect to the relationship between X and Y. A non-linear or curved slope of the u_1 vs. t_1 plot reveals non-linearity in the X and Y-data. In this case the Y-data is normally transformed and the model is refit. For example the natural logarithm of the Y-data may be taken. An example of a non-linear relationship between X and Y, as outlined on the t_1 vs. u_1 score plot is shown below along with the result following log-transformation of the Y-data.

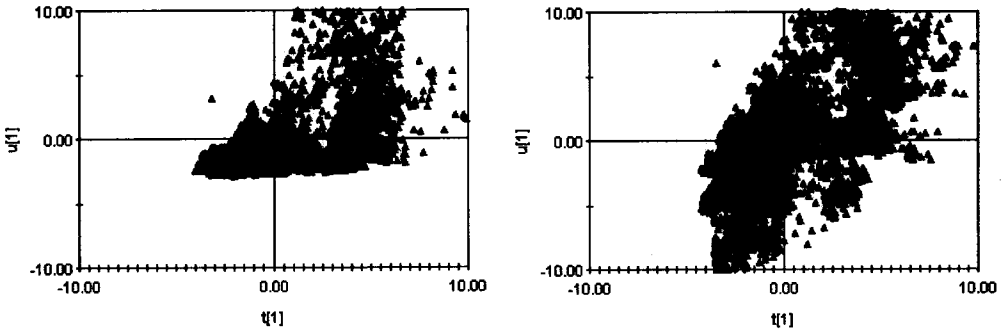

Figure 5: Y-Data Before (Left) and After (Right) Log-Transformation

The relationship between the observations and variables may be evaluated through examination of the score and loading plots. Observations located close to one another on the score plot share similar properties while those that are far away from one another do not. Observations that are located close to the origin of the score plot have properties that are similar to the average of all of the observations contained in the data set.

As is the case with the score plot, variables located close to one another on the PCA loading plot are positively correlated with one another. For example, when the value of one of the variables within the group increases so do the others. Conversely variables that are inversely correlated with one another are located diagonally opposite one another on the loading plot. Finally variables located furthest from the origin of the loading plot possess the largest loading values and are therefore the most important with respect to the model.

For the case of PLS the weights of the X and Y variables, given by w* and c, are plotted together. The Y-weights w combine the residuals of

the X-variables to form the scores t, while the weights w* combine the X-variables to form the scores t. The weight plot is interpreted in a similar manner to the PCA loading plot in order to determine which variables contribute most to the PLS model. X-variables with high weights and are located away from the origin on the w^*c_1 vs. w^*c_2 plot are the most highly correlated with the Y-variables. Conversely those variables located in close proximity to the origin of the weight plot contribute little to the PLS model. Finally X-variables located close to one another on the weight plot contribute similar information to the PLS model. Similarly Y-variables located close to one another on the weight plot are correlated with one another.

The distribution of the observations on the score plot may be understood through examination of the loading or weight plot. For example using the score plot the direction in which the process is moving compared to normal operation may be observed. By comparing the direction of movement on the score plot with the location of the variables on the loading or weight plot, variables that participate in the move may be identified.

For outlying observations that extend beyond the score plot confidence ellipse, contribution plots may be used to identify the variables that contributed to the scores (MacGregor et. al. 1994). In the case of the contribution to the score values between the average observation and observation i for component a, the contribution is given as:

$$\text{Score Contribution}_{(i-avg)} = p_{ak}^*(x_i - x_{avg})$$

Where p_{ak} is the loading of variable k with respect to component a.

For the case of PLS the score contribution is calculated similarly but with w_{ak} replacing p_{ak} in the above equation where w_{ak} is the weight of variable k with respect to component a. Variables with the highest contribution and are of the same sign as the score are of greatest interest as those of the opposite sign contribute to a reduction in the score value.

Although contribution plots reveal which variables make the greatest numerical contribution to a shift in the process they do not indicate the cause of the shift. Only variables that may be related to the cause are identified through contribution plots. However direct cause may be identified using designed experimentation on the process.

Although the PLS weights reveal the importance of the X-variables with respect to the model, evaluation of the PLS regression coefficients may be used for model interpretation, in particular when the number of model components is high. The coefficients reveal the nature of the correlation between the X and Y-variables and the magnitude. In addition variables with little influence of the model are revealed as having small magnitude with large confidence intervals that include zero.

One set of regression coefficients is provided for each Y-variable. Therefore one set of coefficients may be examined compared to many sets of weights, depending on the number of model principle components. However by examining the regression coefficients the correlation between the Y-variables is not observed, as is the case with the weights.

The Y matrix may be expressed as a function of the coefficient matrix B (Manne 1987):

$$Y = XB + F$$

Where:

$$B = W(P'W)^{-1}C'$$
$$= W^*C'$$

PCA and PLS may also be used for the purpose of online monitoring. In this instance the multivariate process may be monitored using a small number of control charts. The Hotelling's T_i^2 provides a summary all of the scores in all of the model principle components (Kourti and MacGregor 1996). It illustrates the distance between the observation i and the origin of the plane or hyperplane defined by the principle component model. For a model containing A components the T^2 value for observation i is given by:

$$T_i^2 = \sum_{a=1}^{A} \frac{t_{ia}^2}{s_{ta}^2}$$

Where s_{ta}^2 is the variance of t_a.

Upon construction of a principle component model a 95% confidence interval for T^2 is calculated. Any observation that exceeds this interval would lie outside the elliptical confidence region found on the score plot. A plot of T^2 therefore reveals strong outliers in the data. Since T^2 is a squared term the magnitude of the outlier deviation is revealed but not the direction. In order to gauge the direction of deviation the predicted scores may be plotted against one another. In addition the time series of individual scores may be plotted.

The distance to the model in X-space, or $DMODX_i$ is the perpendicular distance between observation i and the principle component model plane or hyperplane in X-space and defines the model residual. This residual value outlines the variation in the X-data that is not accounted for by the model. Moderate outliers may be identified as those that exceed the D-crit value. However a continual increase in observations above D-crit may be indicative of a shift in the process. The expression for DMODX is given as (Umetrics 2002):

$$DMODX = \sqrt{\frac{\sum e_{ik}^{\,2}}{K-A}} \times v$$

Where:

e_{ik} = the X-residuals of observation i
v = a correction factor which is a function of the number of observations and components
K = the number of variables
A = the number of components

3. Chapter 3 – Autoclave Leaching PLS Analysis

3.1. Introduction

Using a hydrometallurgical process electrowon copper cathode is produced at INCO's Copper Refinery Electrowinning Department (CRED). The second stage of the operation at CRED involves oxygen pressure leaching in order to solubilize the copper contained in the feed to the plant. The process is comprised of batch make-up and autoclave leaching unit operations and its aim is to reduce the copper content of the feed to less than ten percent. Although the copper leaching efficiency of the process is good, the time required for leaching is somewhat variable. In order to investigate the factors that contribute to the variability with respect to leaching time PLS analysis was performed.

During batch make-up copper sulfate electrolyte and water are added to a mix tank. Solids from the preceding processing step are then added in order to increase the mix tank level to the desired set point. The batch is then added to one of two autoclaves, identified as #2 and #3 autoclave for leaching. Once the batch has been added to the autoclave oxygen is introduced. The pressure within the autoclave, whose set point is 12 kg/cm^2, controls the oxygen flow rate. Since the oxygen used is 95% O_2 and 5% N_2, gas is continually bled from the autoclave in proportion to the oxygen flow rate. This is to avoid the accumulation of nitrogen within the autoclave, which would smother the reaction thus decreasing the extent to which copper is leached. Since the leaching reaction is exothermic cooling water is used to maintain the appropriate autoclave temperature. Once the temperature within the autoclave reaches 115°C the cooling water flow is initiated. At 120°C a high temperature interlock is reached and the flow of oxygen is stopped. Once the temperature within the autoclave decreases sufficiently the oxygen flow is initiated once again. This oxygen flow cycle may be repeated many times. When the pressure within the autoclave has reached the set point the oxygen and cooling water flows decrease until they reach zero and the leach may be considered complete. The trajectories of three typical batches with respect to cooling water and oxygen flows along with the autoclave pressure and temperature are given below. A calculated variable called energy is also included which indicates the amount of heat generated during the leaching reaction.

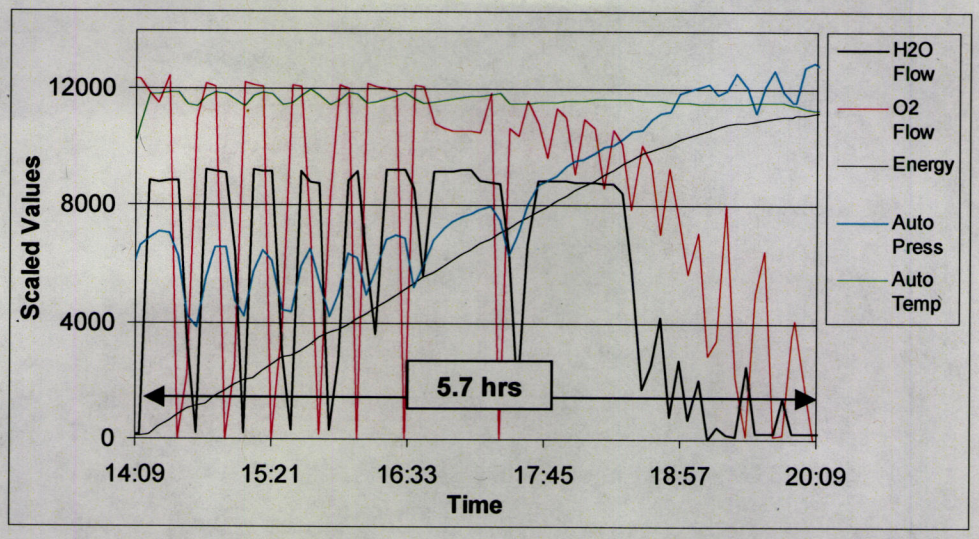

Figure 6: Sample Batch Trajectory - #2 Autoclave

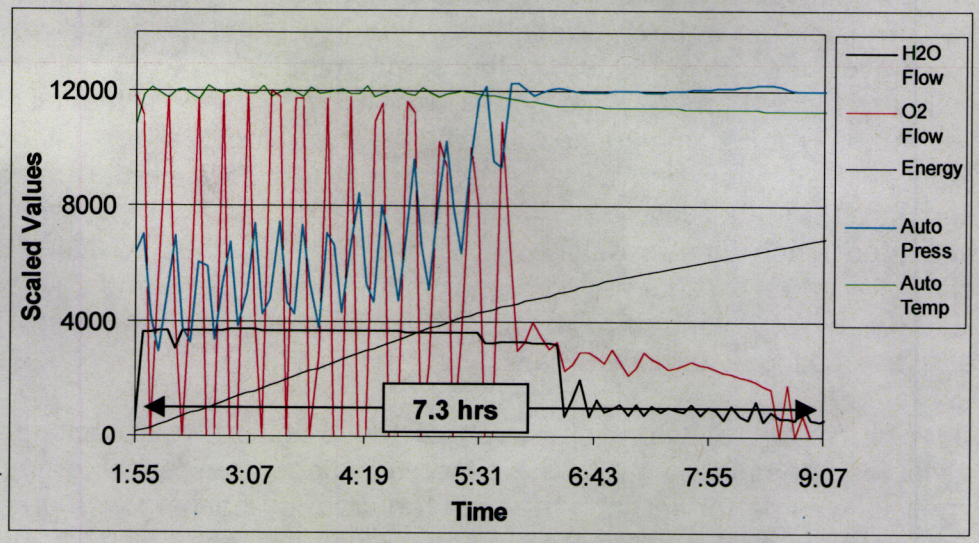

Figure 7: Sample Batch Trajectory - #3 Autoclave

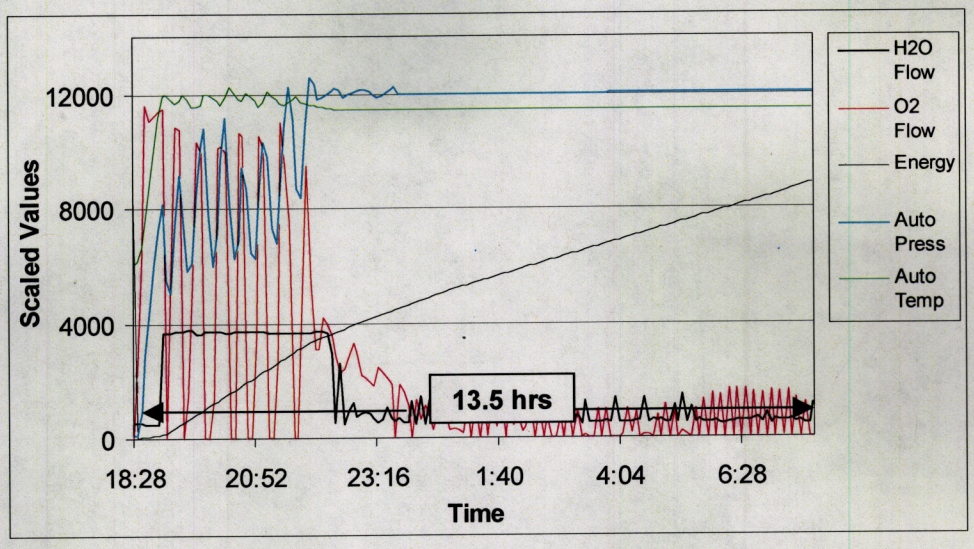

Figure 8: Sample Batch Trajectory - Long Cook

The trajectory depicted in Figure 6 is typical of a batch leached using #2 autoclave. The cooling water flow rate is higher compared to the batch outlined in Figure 7, which was leached using #3 autoclave. The autoclaves are not operated in the same manner as two cooling water coils are utilized for #2 while one is used for #3. Since a higher cooling water flow rate is maintained for #2 the high temperature interlock is reached less often and as a result the oxygen flow is interrupted less compared to the case for #3. Finally a trajectory representing an extended leach time is outlined in Figure 8. Leaches of this type are characterized by a brief period of maximum cooling water and oxygen flow followed by a prolonged period during which time the oxygen flow is low and the cooling water flow cycles.

In order to determine which of the batch make-up and process variables contribute most to the observed variation in leaching time PLS analysis was performed and the time required for leaching was predicted. Seventeen variables related to the batch make-up operation were included in the analysis along with seven related to the autoclave leaching process. A list of the batch make-up and process variables used in the analysis is provided below. Data was obtained for 702 leaches performed using #2 autoclave and for 641 leaches performed using #3. The data collected spanned the time period from January 1st 2004 to December 18th 2004.

The data set was therefore comprised of twenty-four X-variables, one Y-variable and 1343 observations. The predictive model was then investigated in order to determine which of the variables made the greatest contribution to the variation in leaching time.

Batch Make-Up Variables	
Bubbler	Mix tank bubbler type level indication
Sonic	Mix tank sonic type level indication
Density	Mix tank slurry density
Water	Water volume added to mix tank
Spent	Spent electrolyte volume added to mix tank
Cu	Copper concentration - solids added to mix tank
Ni	Nickel concentration - solids added to mix tank
Co	Cobalt concentration - solids added to mix tank
Fe	Iron concentration - solids added to mix tank
S	Sulphur concentration - solids added to mix tank
As	Arsenic concentration - solids added to mix tank
CuS	Copper to sulphur ratio - solids added to mix tank
SpntCu	Spent electrolyte copper concentration
SpntAcd	Spent electrolyte sulphuric acid concentration
CuUnits	Mass of copper added to mix tank (via spent electrolyte addition)
AcdUnits	Mass of sulphuric acid added to mix tank (via spent electrolyte addition)
AsExt	Arsenic extraction in previous processing step
Process Variables	
Auto	Autoclave number (Value = 2 or 3)
O2/Bld	Average oxygen flow to average bleed vent flow ratio
EngSlope	Slope of calculated energy term over first four hours of leach
AvgAgitAmps	Average agitator amperage draw
AvgClH2OTmp	Average cooling water inlet temperature
AvgO2Fl	Average oxygen flow
AvgClH2OFl	Average cooling water flow

3.2. Model Development

The initial PLS model fit to the autoclave data resulted in a four component model with $R^2X = 0.35$, $R^2Y = 0.71$ and $Q^2 = 0.68$. As shown below the t_1 vs. u_1 score plot indicates non-linearity with respect to the relationship between the X and Y-data. The natural log of the Y-data was taken and the model was refit.

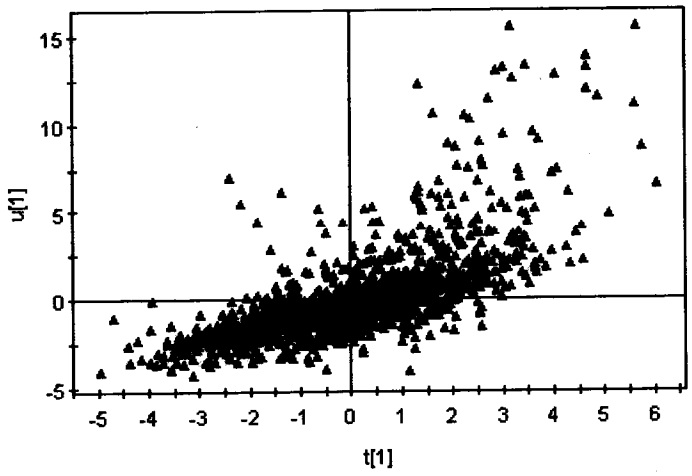

Figure 9: Autoclave Leaching t_1 vs. u_1 Score Plot – Indication of Non-Linearity

The resulting model consisted of four principle components with $R^2X = 0.35$, $R^2Y = 0.76$ and $Q^2 = 0.73$. The majority of the variation was explained and predicted by the first three components. The score plots for the first three components are given below and as illustrated there are no extreme outliers and as such no further observations were removed from the data set.

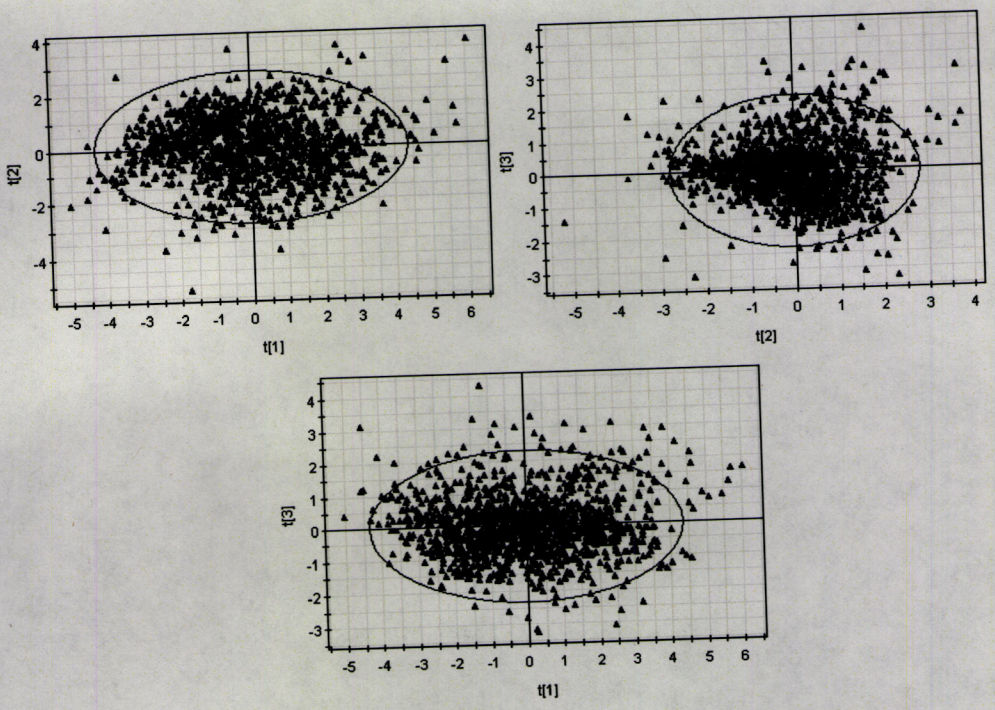

Figure 10: Autoclave Leaching X-Score Plots for First Three Components

Finally the t_1 vs. u_1 and u_1 vs. u_2 score plots do not indicate the presence of outliers in the X or Y-data. In addition the u_1 vs. u_2 plot for the transformed Y-variable appears more linear compared to the non-transformed case.

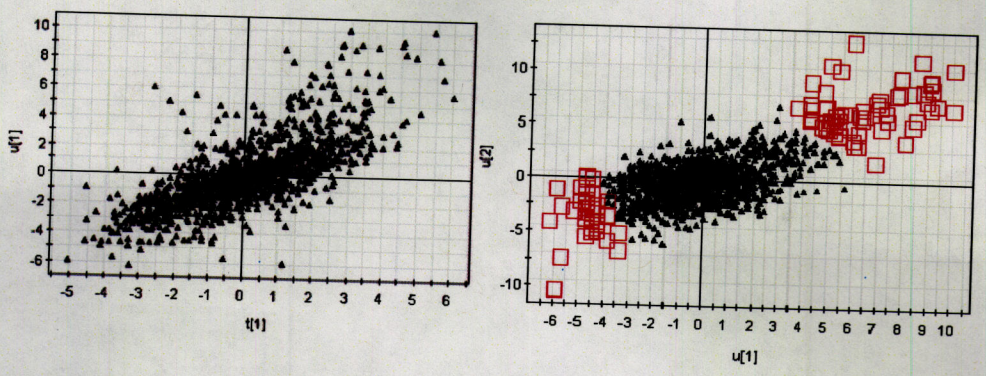

Figure 11: Autoclave Leaching t_1 vs. u_1 (Left) and u_1 vs. u_2 (Right) Score Plots

3.3. Model Interpretation

The weight plot for the first two components is given below. The first variable contained in the data set, Auto displays significance with respect to the second component. Since this variable is not located in close proximity to the origin of the weight plot the use of either #2 or #3 autoclave contributes to the variation observed with respect to leaching time. As shown the variables Bubbler and Density are located close to one another on the score plot as expected since the bubbler type level indicator is affected by the actual slurry density in the mix tank. While these two variables are located close to the Y-variable lnTime, the variable Sonic is located in close proximity to the origin of the weight plot. As such Bubbler and Density are more significant variables in the first and second components compared to Sonic. This result is significant since during the batch make-up operation solids are added to the mix tank until the level set point is reached by the sonic level indicator as opposed to the bubbler indicator.

In addition to Sonic all of the variables related to the composition of the solids added to the mix tank are located close to the origin of the weight plot with the exception of CuS and S. The effect of the copper to sulphur ratio (CuS) of the solids added to the mix tank on leaching time has been studied empirically in the past. In plant practice it is this ratio along with the acid concentration of the spent electrolyte, which is used to determine the volume of spent electrolyte added to the mix tank during batch make up. In general as the copper to sulphur ratio increases the amount of spent electrolyte to be added to the mix tank also increases. This

relationship between CuS and Spent is also apparent upon inspection of the weight plot as the two variables are located somewhat close to one another and are therefore positively correlated with one another.

Finally all of the process variables are located peripherally on the weight plot with the exception of ClH2OTmp. The variable ClH2OTmp was included in the analysis in order to determine whether seasonal changes in cooling water inlet temperature contribute to leaching time variation. The variable AvgAgitAmp was included in order to provide an indication regarding the density or viscosity of the autoclave leach slurry. Although this process variable displays less significance compared to some of the other process variables it displays a positive correlation with Density and Bubbler given the location of these three variables on the weight plot. The variable O2/Bld represents the ratio of oxygen introduced to the autoclave to the gas bled from it during leaching. It displays significance with respect to the second component and is negatively correlated with lnTime. Under normal operation the bleed vent flow is controlled in a set proportion to the oxygen flow. During an extended leach the bleed vent is typically opened beyond its normal level thus decreasing the oxygen to bleed vent flow ratio. This additional venting reduces the pressure in the autoclave, which in turn causes the oxygen flow to increase and may expedite the completion of the leach. The negative correlation between O2/Bld and lnTime is therefore expected. The correlation between AvgO2Fl, AvgClH2OFl and lnTime is also expected. These two variables display the most significance with respect to all of the process variables, since the leaches of shortest time are characterized by brief periods of maximum oxygen and cooling water flow as in Figure 6, whereas long leaches are characterized by extended periods of low flow as in Figure 8. In addition through the autoclave process control system, the amount of oxygen supplied is related to the amount of cooling water supplied. Finally the EngSlope variable is the slope of the calculated energy term, which is indicative of the amount of heat generated during leaching and does not appear to be correlated with any of the other variables included in the analysis. However this variable and AvgClH2OTmp are located in close proximity to lnTime on the $w*c_2$ vs. $c*c_3$ weight plot which is not shown. The closeness of AvgClH2OTmp and EngSlope is expected as the cooling water inlet temperature is used in the calculation of the energy term.

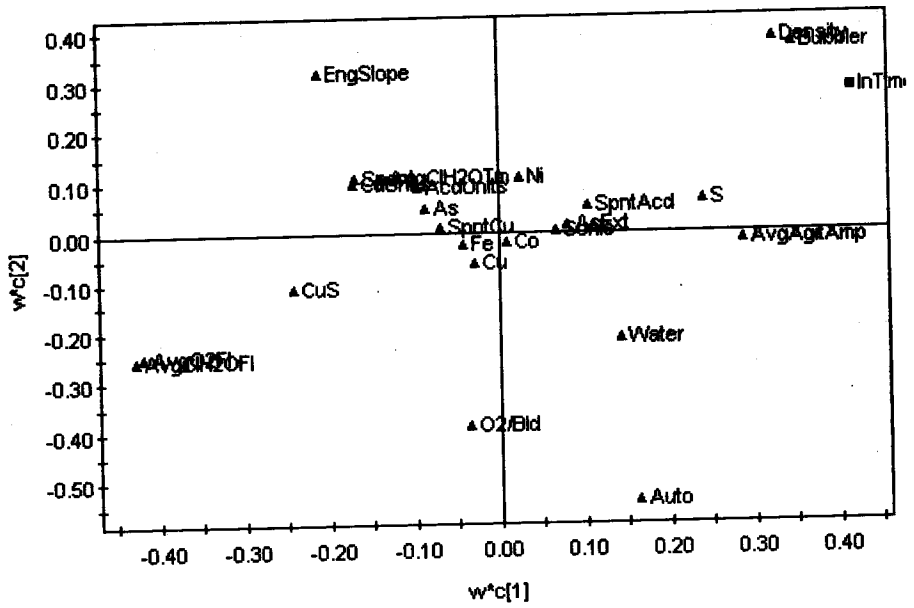

Figure 12: Autoclave Leaching Weight Plot – First Two Components

Although outliers are not present on the u_1 vs. u_2 score plot shown in Figure 11, the general appearance of the plot is unexpected. In general the scattering of observations on the u_1 vs. u_2 score plot is expected to be similar to that observed on the t_1 vs. t_2 score plot illustrated in Figure 10. As such observations on the upper right hand and lower left hand quadrants of the u_1 vs. u_2 score plot were selected and their corresponding batch leach times were examined. Observations located in the lower left hand quadrant of the score plot correspond to the lowest batch leach times while those in the upper right hand quadrant correspond to those with the highest as shown in Figure 11. This is the expected relationship between score plot observation location and leach time given the location of lnTime on the w^*c_1 vs. w^*c_2 weight plot illustrated above. The extreme deviation in leach time for the few highlighted observations compared to the large number of observations whose leach time is close to the average may account for the skewed appearance of the u_1 vs. u_2 score plot.

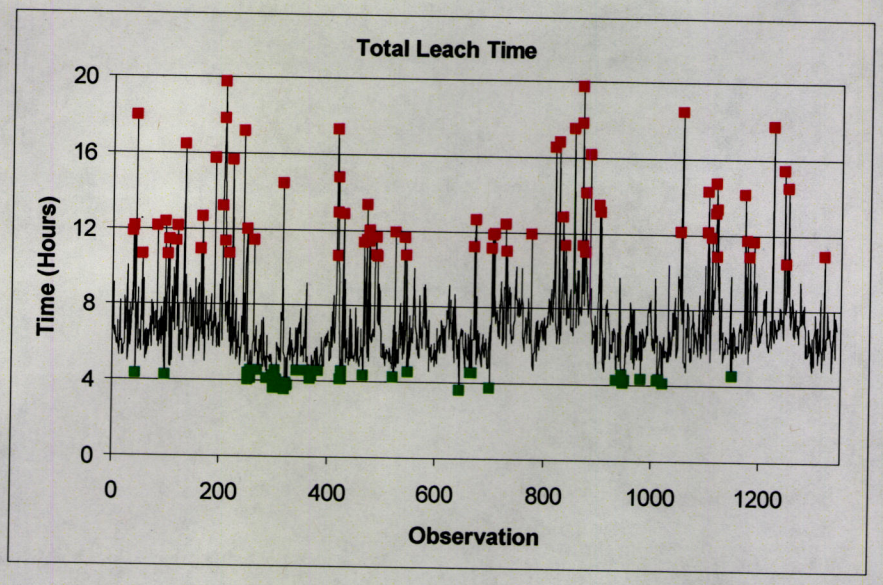

Figure 13: Total Leach Time by Observation

The contribution to the first two scores for the observations with extended leach times, located in the upper right-hand quadrant of the score plot was compared to the contribution for the observations with low leach times, located in the lower left-hand quadrant. The resulting contribution plot is shown below.

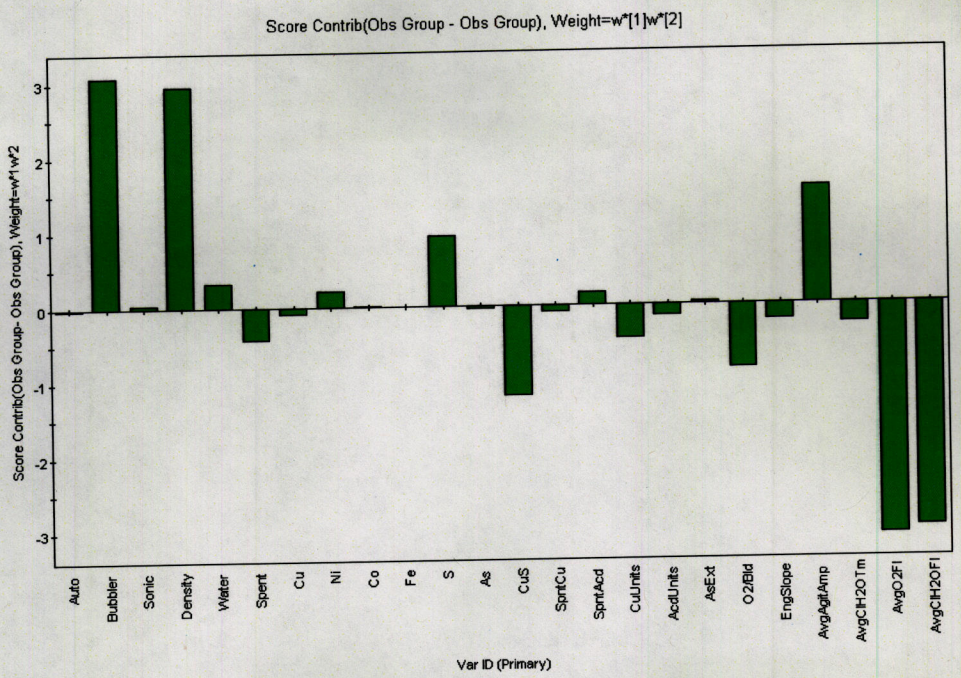

Figure 14: Contribution Plot Between Extended and Low Leach Time Observations

As shown the variables related to the density of the slurry added to the mix tank (Bubbler, Density, AvgAgitAmps) make strong positive contributions. This is expected since as the amount of solids added to the autoclave increases so does the time required for leaching. A normalized variable such as time/solids was not used as the Y-variable since an accurate measure of the amount of solids added to the mix tank is not available. The mix tank level indicators relate the amount of slurry added, however the slurry is comprised of solution and solids and it is the solids portion that affects the leaching time. In addition the density or viscosity of the slurry added to the autoclave may affect the extent to which oxygen is incorporated into the agitated slurry during leaching. For example oxygen may not be incorporated as well for highly dense or viscous slurries. This poor incorporation may reduce the oxygen efficiency of the leach therefore increasing the time required for leaching.

Similar to the case for the $w*c_1$ vs. $w*c_2$ weight plot, S and CuS are the only variables related to the composition of the solids added to the mix

tank which make significant contributions. Water and Spent also contribute to the scores. The significance of these four variables may be indicative of the importance of proper batch make-up practices. For example for a given CuS ratio the appropriate amount of water and spent electrolyte must be added depending on the spent electrolyte acid concentration. The amount of soluble copper added to the mix tank via spent electrolyte addition is shown to contribute to the variation in leach time and as such this parameter may also be considered during the batch make-up operation. Since the contribution shown is negative attention may be made to ensure that sufficient electrolyte is added depending on the copper concentration as well as the acid concentration, which is the current practice.

Finally the oxygen and cooling water flow rate variables make strong negative contributions. This result is expected since short periods of maximum oxygen and cooling water flow are followed by extended periods of low oxygen and cooling water flow for leaches of extended time. The negative contribution of O2/Bld is also typical of the trajectory of this variable during an extended leach.

The model coefficient plot is given below and as shown the autoclave number is negatively correlated with leaching time. Of the batches contained in the data set the average leach time was 6.7 hours for #2 autoclave and 7.4 hours for #3 autoclave. Therefore the negative correlation displayed on the coefficient plot appears to contradict the average leaching times. The distributions of the leaching times for each of the autoclaves were then examined and are given below.

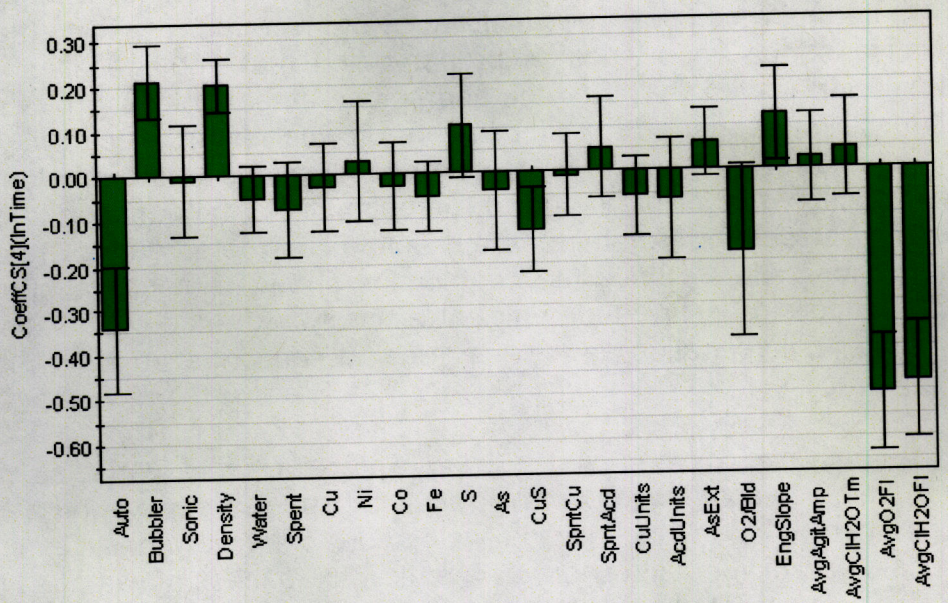

Figure 15: Coefficient Plot

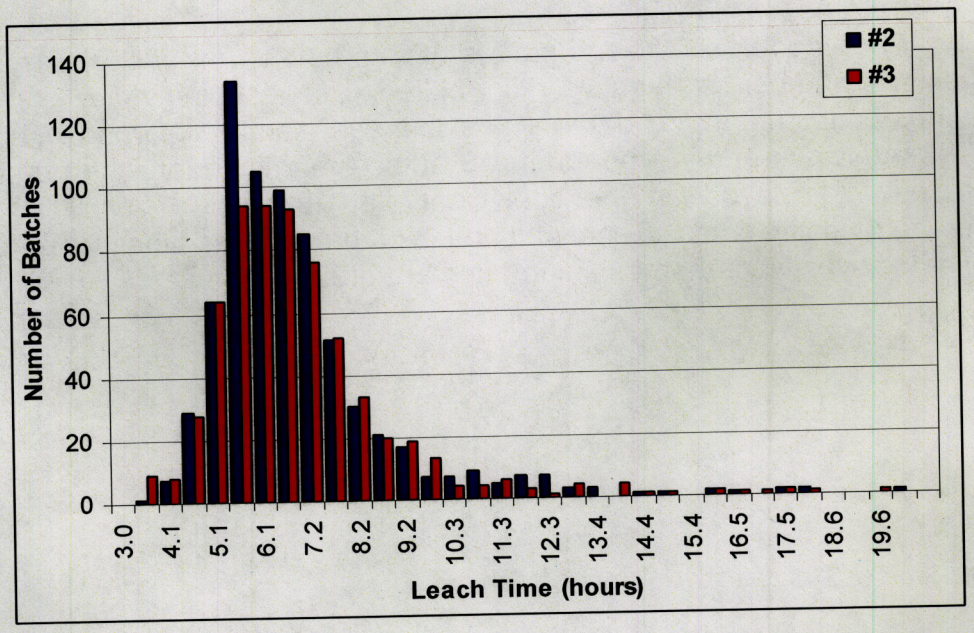

Figure 16: Leach Time Distribution by Autoclave

As shown there are more leaches corresponding to higher leach times for #2 autoclave compared to the #3. For example the area under

the #2 autoclave curve is 363 hours while the area under #3 is 327 hours. In general it is believed that the time required for leaching is less for #2 autoclave compared to #3. Since #2 autoclave is operated with two cooling coils compared to #3 which is operated with one, the autoclave temperature may be maintained below the high temperature interlock more often thus resulting in fewer periods when the oxygen flow is ceased. However for the batches contained in the data set more batches leached using #2 autoclave displayed greater leach times compared to those leached using #3. This is also evident through examination of the observations included in the contribution comparison outlined in Figure 14. For example of the sixty-two observations contained in the extended leach time group forty were leached using #2 autoclave and twenty-two were leached using #3. As a result of the number of long leaches performed using #2 for the data set analyzed, the autoclave number and leaching time are negatively correlated.

As opposed to Auto, the variables Bubbler and Density are positively correlated with leaching time as expected. As the amount of solids added to the mix tank increases, the bubbler level indication increases along with the density and the time required for leaching. The lack of significance displayed by the variable Sonic is not expected since it is this level indicator that is used as a cut-off point during the batch make-up operation. The accuracy of the sonic level indicator may be affected by the presence of foam in the mix tank, which is attributed to downstream reagent addition to the electrolyte.

Of the variables related to the composition of the solids added to the mix tank CuS and S display the most significant coefficients. In general extended leaching times are observed during plant operation when insufficient spent electrolyte is added to the mix tank and the CuS ratio is high. However in this analysis CuS displays a negative correlation with leach time along with Spent. This may be indicative of the effect of insufficient spent electrolyte addition at low Cu:S ratios. The actual relationship between the Cu:S ratio, spent electrolyte addition and leach time may only be established through designed experimentation.

The process variable O2/Bld displays a negative correlation and may be attributed to the plant practice of increasing the bleed vent flow and subsequent oxygen flow during long leaches. Although the oxygen flow is increased during this time the oxygen:bleed vent flow ratio decreases below its normal operating level. Although O2/Bld displays a

negative correlation EngSlope displays a positive one. The average value of the energy slope was twice as small for batches leached using #3 autoclave compared to #2. Since the #3 autoclave leach time data was distributed around a lower leaching time compared to #2, more observations with low energy slope are correlated with those with low leaching time for the data set analyzed.

The coefficients for the variables AvgAgitAmps and ClH2OTmp are small and their confidence intervals contain zero. These variables were included in the analysis in order to gauge whether the variations in agitator amperage draw and cooling water inlet temperature contribute to the variation in leach time. It appears to for the overall data analyzed they do not make a significant contribution. This may be a function of the large number of observations with average leaching time that were included in the analysis. For example during the contribution comparison between the extended leach time group and the low leach time group AvgAgitAmps was found to make a significant positive contribution.

Significantly large coefficients are associated with the variables AvgO2Fl and AvgClH2OFl, both of which display negative correlations with leach time. Therefore when the average flows of oxygen and of cooling water are low the corresponding leach time is expected to be high. This is attributed to the tendency for the trajectories of long leaches to exhibit extended periods of low oxygen and cooling water flow as illustrated in Figure 8.

The model observed and predicted plots are shown below. As illustrated the correlation between the predicted and actual lnTime is good with increased scatter at elevated actual lnTime values. This is also evident on the Time vs. Observation graph. As shown in some instances the predicted time does not correspond fully to the actual time at the most elevated time values. Variables not included in the model such as mixing efficiency, the particle size distribution of the solids charged to the autoclave and the error in measured values such as the mix tank level indicator may account for a portion of the unexplained variation. These variables were not included was they are not currently measured. In addition the amount of soluble copper entrained in the solids charged to the mix tank was not included in the analysis. It is believed that insufficient soluble copper addition may contribute to the occurrence of extended leaching time.

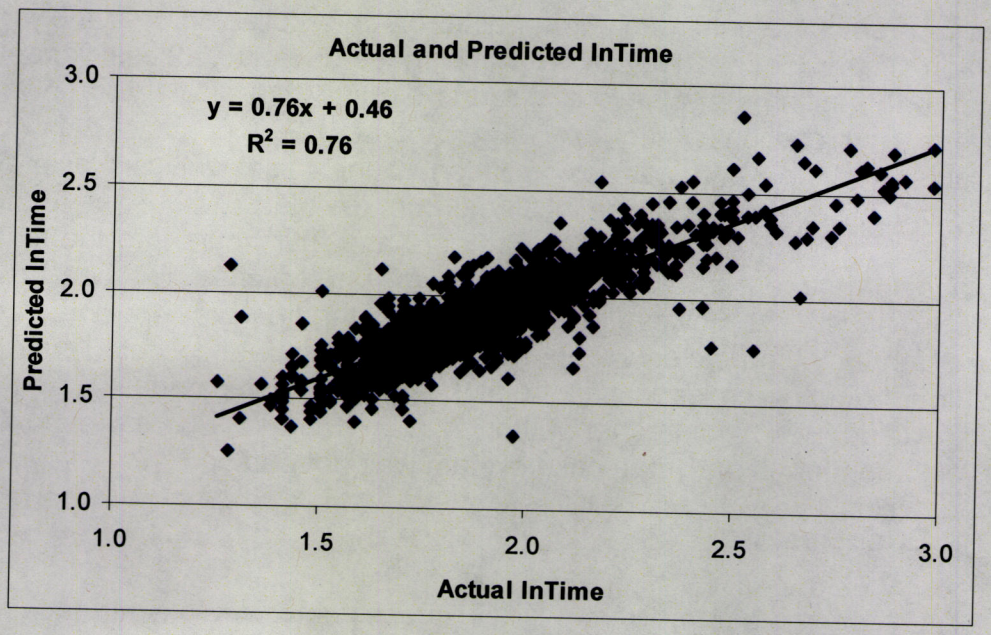

Figure 17: Actual vs. Predicted Leach Time

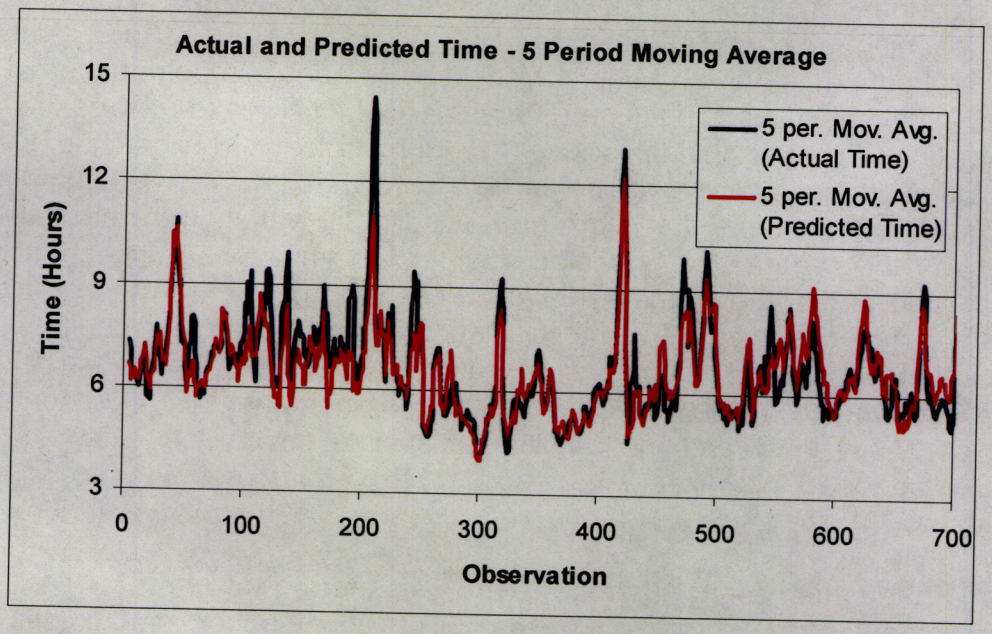

Figure 18: Actual and Predicted Leach Time by Observation

In an effort to develop an early detection system for the occurrence of long leaches a number of models were fit. The purpose of the analysis was to predict the final leach time based on data collected over the initial portion of the leach. The same batch make-up and process variables used in the initial analysis were included along with the totalized flows of oxygen and cooling water. Average and totalized process data was collected for the 1343 batches included in the initial analysis over five different time intervals. These included: from the start of the leach to thirty minutes, one hour, two hours, three hours and four hours into the leach. In addition separate models were fit to the batches leached using autoclave 2 and 3. However the prediction of the final time was not good. For example the predicted variation achieved using data collected over the first four hours of the leach was the highest for all of the models fit and did not exceed 50% for either autoclave 2 or 3. Therefore further development of an early detection system was not pursued.

3.4. Conclusion

The second stage of INCO's CRED process involves oxygen pressure leaching in order to solubilize copper contained in the feed to the plant. This stage of the process involves batch make-up and autoclave leaching operations and although the overall leach extent is good the time required for leaching is somewhat variable. PLS analysis was used to determine which batch make-up and autoclave leaching variables contribute most to the variation in leaching time. Seventeen batch make-up and seven autoclave process variables used to summarize 1343 batches were included in the analysis. The model fit to the data consisted of four principle components that explained 35% of the variation in the X-data, 76% in the Y-data and predicted 73% of the variation. In general the predicted and actual leaching times corresponded well with the greatest discrepancy observed at the most extended leaching times. Variables not included in the analysis such as mixing and oxygen efficiency, the particle size distribution of the solids to be leached, instrumentation sensor and solids assay error and the addition of soluble copper entrained in the solids added to the mix tank may account for a portion of the unexplained variation.

Of the variables included in the analysis four groups were outlined as significant upon examination of the weight, contribution and coefficient plots. The first group is comprised of one variable: autoclave number. Although the average leaching time for the batches contained in the data

set was greater for #3 autoclave compared to #2, the variable Auto was negatively correlated with leach time. This may be attributed to the greater number of observations with extended leach times contained in the data set which were associated with #2 autoclave compared to #3. Therefore although #3 autoclave is operated with one cooling coil compared to #2 which is operated with two, the relationship between extended leach time and the operation of #3 autoclave was not displayed by the model for the data set analyzed.

The second group of significant variables was associated with the density and viscosity of the solids added to the mix tank: Bubbler, Sonic, Density and AvgAgitAmps. Bubbler and Density were shown to be two of the most significant variables while Sonic was not significant. This result suggests that perhaps the bubbler type level indicator could be used to determine the end point for solids addition to the mix tank as opposed to the sonic type level indicator, which is currently used during batch make-up. The positive correlation displayed by the autoclave slurry density and autoclave agitator amperage draw is expected since as the amount of solids added to the mix tank increases so does the time required for leaching. However the pronounced positive contribution by these two variables observed during the comparison between extended leach time batches and those with low leach time may suggest that other factors associated with increased solids addition contribute to extended leach time such as decreased mixing and oxygen efficiency. The implementation of a more accurate density measurement method may therefore be warranted as the current value is calculated based on the sonic and bubbler level measurements.

The third group of variables is related to the proportions of spent electrolyte and water added to the mix tank and the copper and sulphur contents of the solids added to the mix tank. The negative correlation between Spent and lnTime suggests that the maximum amount of spent electrolyte be added to the mix tank for a given Cu:S ratio. In general changes to the batch make-up water and spent electrolyte addition are more frequently made as the Cu:S ratio increases. However given the negative correlation between CuS and lnTime, extended leach times may result even at low Cu:S ratios if the appropriate batch make-up changes are not made.

Finally the last group of significant variables is associated with the oxygen and water flow rates: O2/Bld, AvgO2Fl and AvgClH2OFl. The

negative correlation between the average oxygen and cooling water flow and leach time is attributed to the tendency for batches of extended leach time to display long periods of low oxygen and cooling water flow prior to completion. Although changes to the maximum allowable autoclave operating temperature and pressure and therefore oxygen flow may not be made for maintenance and safety reasons, consideration may be made to altering the oxygen to bleed vent flow ratio. Since O2/Bld is negatively correlated with leach time the use of a higher ratio may contribute to less time required for leaching.

4. Chapter 4 – Acid Plant Feed Gas PLS Analysis

4.1. Introduction

INCO's Copper Cliff Smelter Acid Plant receives high strength sulphur dioxide (SO_2) gas from three pyrometallurgical processes in order to produce concentrated sulphuric acid (H_2SO_4). Two of the sources are INCO flash smelting furnaces while the third is a copper reactor. At the acid plant the high strength gas is diluted to approximately 10-12% SO_2 by volume before drying, conversion to sulphur trioxide (SO_3) and absorption to produce concentrated sulphuric acid. Presently the acid plant feed gas is analyzed for SO_2 only. However, the operation of the plant and the final sulphuric acid product quality are affected by other components contained in the gas. Of these unmeasured components, nitric oxide (NO) is known to negatively affect the final product quality. Since there is presently no measurement of the NO concentration of the feed gas, reagents are added to control the final sulphuric acid product NO concentration in proportion to the plant production rate.

In order to investigate the variability with respect to acid plant feed gas NO concentration a gas analyzer was temporarily installed at the acid pant, which provided a continuous analysis of the gas stream. The resulting NO analysis revealed significant variation. Using the data collected from the gas analyzer along with data pertaining to the upstream pyrometallurgical processes, PLS analysis was used to develop a soft-sensor for the purpose of predicting the feed gas NO concentration. The model was then investigated in order to determine which of the process variables contributed most to the NO concentration variability. Finally new data was input to the model and the NO concentration was predicted using only the process variables. Control charts were also developed for the purpose of identifying periods of potentially decreased model predictive accuracy.

The flash furnace and copper reactor operations were initially examined and process variables which may contribute to the formation of NO and to fluctuations in the acid plant feed gas volumetric flow rate were included in the model. Variables believed to contribute to NO formation include oxygen purity, coke addition rate, oxy-fuel burner operation and nitrogen porous plug flow. Variables that may account for the variation in acid plant feed gas volumetric flow rate include vent pressures, fan speeds, concurrent liquid SO_2 plant operation and dilution air addition.

4.2. Model Development

The data set used to model the acid plant feed gas was comprised of fifty-five process or X-variables and one Y-variable, NO concentration. Process and NO data were obtained every thirty seconds for a total of 15 694 observations. The data collection period extended from November 26th 2004 at 7:40 pm until December 2nd 2004 at 7:00 am.

The initial PLS model fit to the acid plant feed gas data resulted in a six component model with $R^2X = 0.57$, $R^2Y = 0.63$ and $Q^2 = 0.63$. The t_1 vs. u_1 score plot was examined and it was determined that the relationship between the X and Y-data is non-linear. The Y-variable was then transformed and the natural log was modeled. The subsequent model fit to the transformed data resulted in a four component model with $R^2X = 0.43$, $R^2Y = 0.73$ and $Q^2 = 0.73$. The score plots for the first three components were then examined. A number of extreme outliers were removed from the data set and the model was refit. The resulting model consisted of four principle components with $R^2X = 0.43$, $R^2Y = 0.73$ and $Q^2 = 0.73$. The majority of the explained and predicted variation is accounted for by the first two model components.

4.3. Model Interpretation

The weight plot for the first two components is shown below. As shown below the variables related to the furnace throughput: F1SO2, F1OCO2, F1OCDSC, F2SO2, F2OCO2 and F2OCDSC cluster together near the middle of the left hand side of the weight plot. As such these variables are negatively correlated with logNO. This result is expected since as the SO_2 concentration of the acid plant feed gas increases the proportion of other components consequently decreases. The furnace variables related to oxy-fuel burner use: F1OFO2, F1OFNG, F2OFO2 and F2OFNG are located in close proximity to logNO on the weight plot and are therefore strongly positively correlated with logNO. This result is again expected as the incomplete combustion of natural gas is known to contribute to NO formation. Conversely the variables related to copper reactor throughput: MKOCO2 and MKMK are located close to the origin of the loading plot and are relatively not influential. The copper reactor throughput is somewhat less variable compared to the furnace throughput and as such the weight result is expected. However similar to the case for the furnace, the copper reactor variables related to the use of oxy-fuel burner use: MKCJO2, MKCJNG, MKOFO2 and MKOFNG are located

close to logNO on the weight plot and are therefore also strongly correlated with the Y-variable. Finally the variables related to the volume of gas fed to the acid plant: DilAir, SO2BlOutFl and SO2BlOutPr are located diagonally opposite to logNO on the weight plot. This strong negative correlation is expected as increases in the amount of dilution air input to the plant increases the overall volume of gas fed to the plant which consequently decreases the percentage of NO contained.

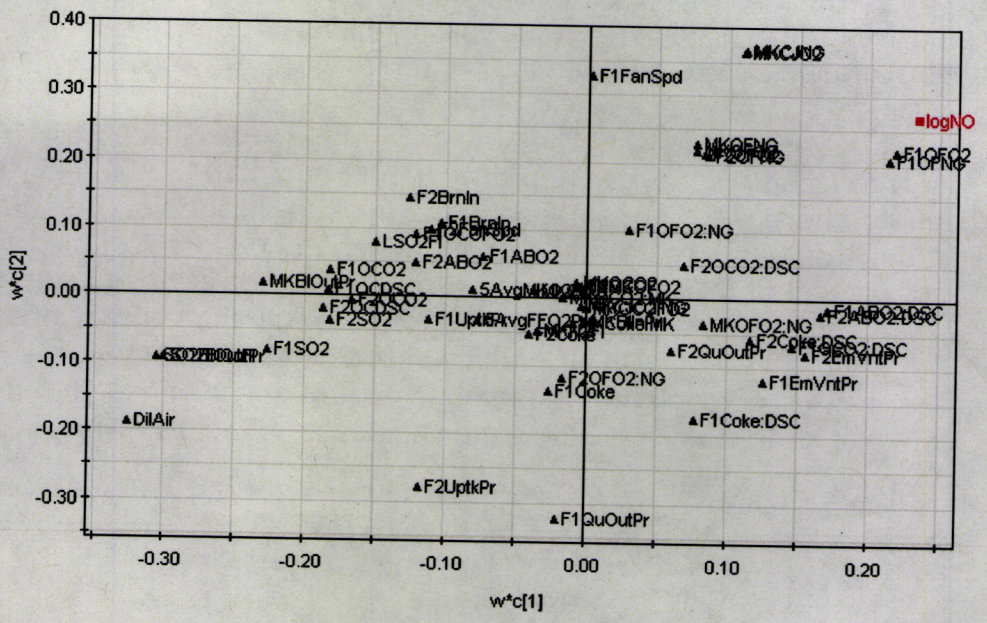

Figure 19: Weight Plot – First Two Components

Variables related to coke addition were included in the model as it is believed that incomplete coke combustion may contribute to NO formation. Of these variables, those related to coke input: F1Coke, F2Coke and MKCoke are located close to the origin while F1Coke:DSC and F2Coke:DSC are located somewhat more peripherally along the w*c1 axis. Measurement error may contribute to the lack of significance displayed by the individual coke variables. Coke addition rate is based on weights determined by load cells and by screw conveyor outputs, both of which may possess some inaccuracy. However the ratio between coke and solids added to the furnace may contribute to the variation in NO as the large fluctuations in the coke to solids ratio may be much greater than the error contained in the coke addition measurement.

Variables related to furnace afterburner operation were also included in the model. Similar to the case for those variables related to coke addition, the ratio between afterburner oxygen addition and solids charged to the furnace: F1ABO2:DSC and F2ABO2:DSC are shown to be influential along the first component while those related strictly to afterburner oxygen addition: F1ABO2 and F2ABO2 are less influential as they are located close to the origin of the score plot. Finally the variable MKN2FI, related to copper reactor nitrogen porous plug flow is also located close to the origin. It was initially believed that the addition of nitrogen through porous plugs to the copper reactor may contribute to NO formation.

The location of logNO in the upper right-hand quadrant may account for the skewed appearance of the u_1 vs. u_2 score plot. As shown below observations located in the upper right-hand quadrant of the score plot correspond to the most elevated levels of NO while those located in the lower left-hand quadrant correspond to those with the least. Since the number of observations with NO values which deviate extremely from the average is small compared to the total number of observations, the u_1 vs. u_2 score plot does not display the expected scatter of points, similar to the t_1 vs. t_2 score plot.

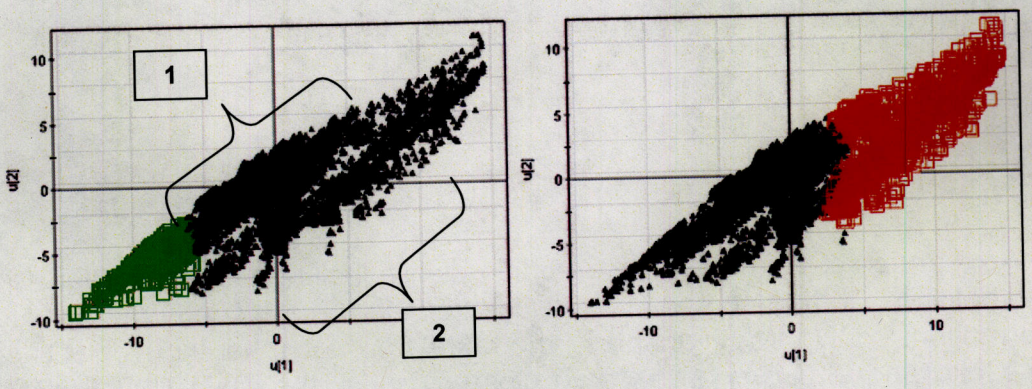

Figure 20: Y-Score Plot - First Two Components

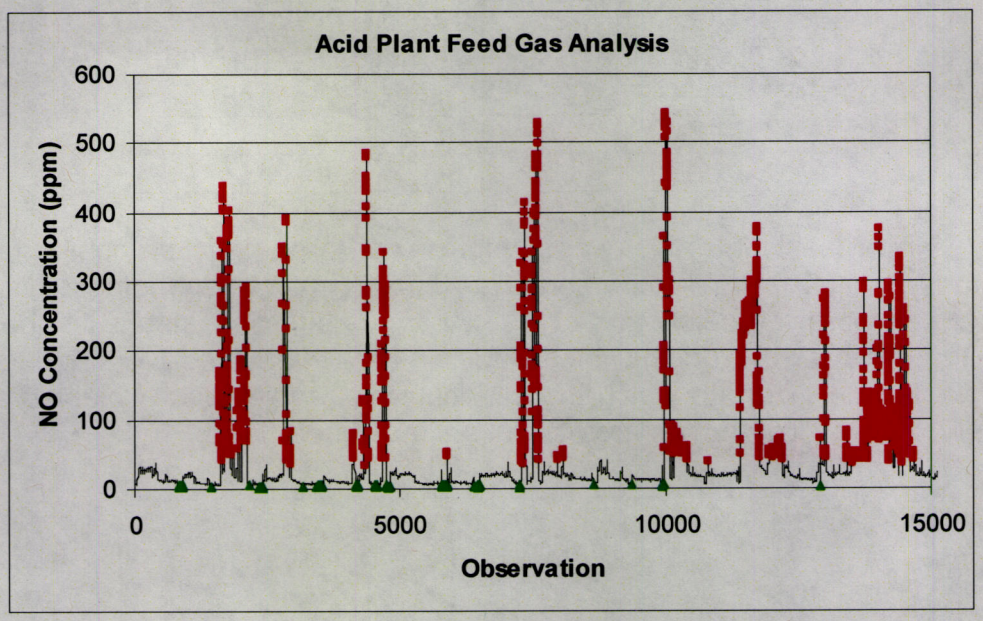

Figure 21: NO Concentration by Observation

Finally the occurrence of clusters on the u_1 vs. u_2 score plot was observed. The clustering appears to occur as parallel groups of observations, labeled 1 and 2, located along a diagonal trend, which moves from the lower left-hand quadrant to the upper right-hand quadrant.

An attempt to de-cluster the score plot was made and as such the model was refit using those variables deemed significant according to the weight plots, a number of summed variables and a number of ratioed variables. A list of the variables included in the revised model is given below.

Variables	
F1SO2	DilAir
F1OCO2	SO2BlOutFl
F1OCDSC	SO2BlOutPr
F1OFO2	TotSO2
F1OFNG	TotSolids
F2SO2	TotNG
F2OCO2	TotSO2/TotSolids

F2OCDSC	TotSO2/TotNG
F2OFO2	TotSolids/TotNG
F2OFNG	TotSolids/DilAir
MKOCO2	TotSolids/SO2BlOutFl
MKMK	DilAir/SO2BlOutFl
MKCJO2	DilAir
MKCJNG	SO2BlOutFl

The resulting model was comprised of three principle components with $R^2X = 0.45$, $R^2Y = 0.65$ and $Q^2 = 0.65$. The corresponding u_1 vs. u_2 score plot is given below and as shown appears less clustered compared to the case for the original model for which all variables were included. However the extent to which the variation in X and Y was explained and predicted was lower compared to the original model.

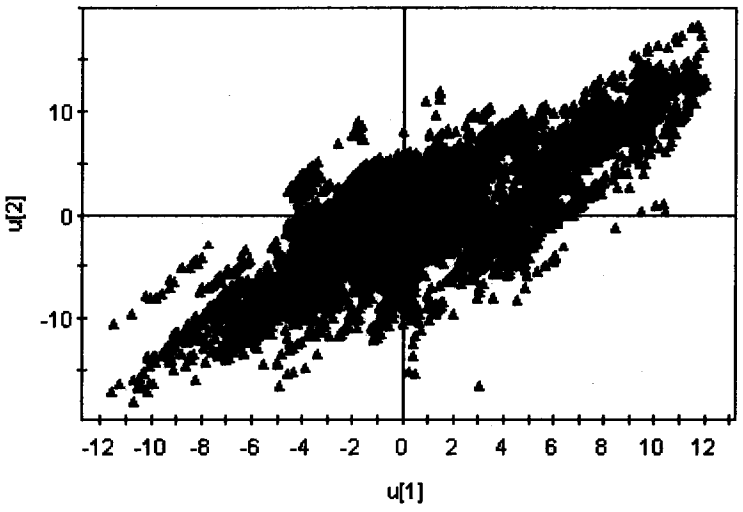

Figure 22: Y-Score Plot Following Model Variable Reduction

The contribution made by the lower cluster on the u_1 vs. u_2 plot, shown in Figure 20 to the first two scores was compared to the contribution made by the average observation. The resulting contribution plot is given below. In general this cluster is characterized by negative contributions from the variables related to furnace off gas SO_2 strength, throughput and number of oxy-concentrate burners in use (F1BrnIn and F2BrnIn) and acid plant throughput. Positive contributions are made by the #1 furnace oxy-fuel burner variables and by the afterburner oxygen to

feed ratios for both furnaces. Finally positive contributions are made by variables related to the uptake for both furnaces. Given the contribution profile, namely the negative furnace and acid plant throughput variable contributions and the negative burner use variables it is estimated that the observations contained in the lower cluster are indicative of a period during which the oxy-concentrate burners were removed for repair or replacement. For example for the observations in the lower cluster, 68% of the eight total burners were in use with a corresponding average actual NO measurement of 68ppm. This may be compared to the observations in the upper cluster with an average burner usage of 98% and average actual NO measurement of 15ppm. Upon inspection of the $w*c_1$ vs. $w*c_2$ weight plot the relationship between observations on the score plot is apparent. Since F1BrnIn and F2BrnIn are located in the upper left-hand quadrant those observations located in the lower right-hand quadrant will have correspondingly low values for these two variables.

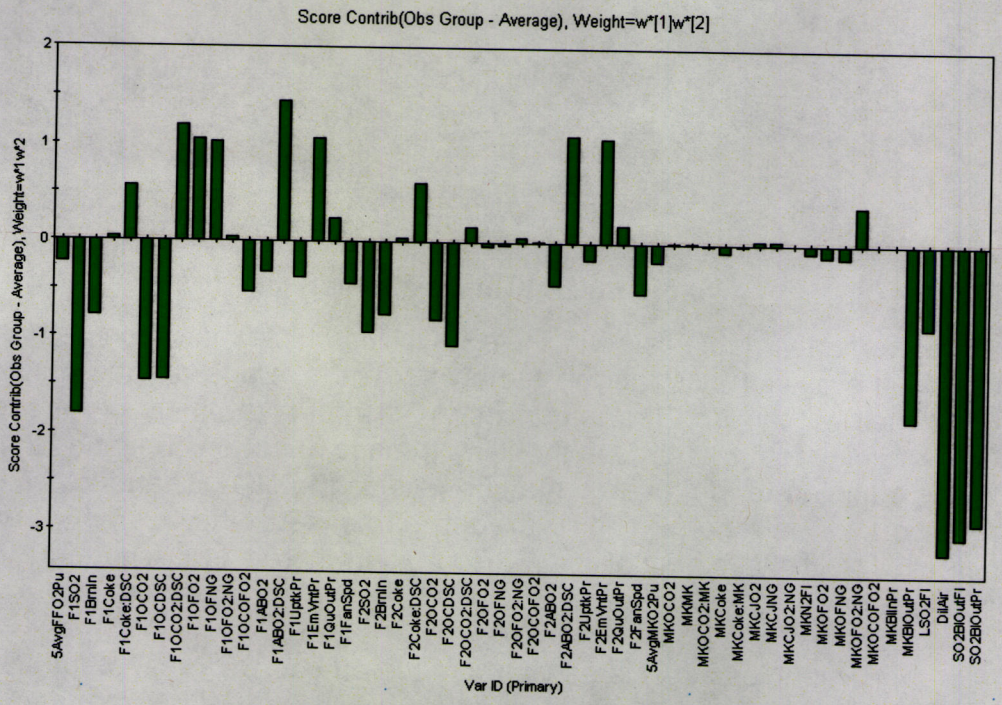

Figure 23: Contribution Plot Between Lower and Upper Y-Score Plot Observations

The coefficient plot is given below and as shown several of the variables display significantly large coefficient values.

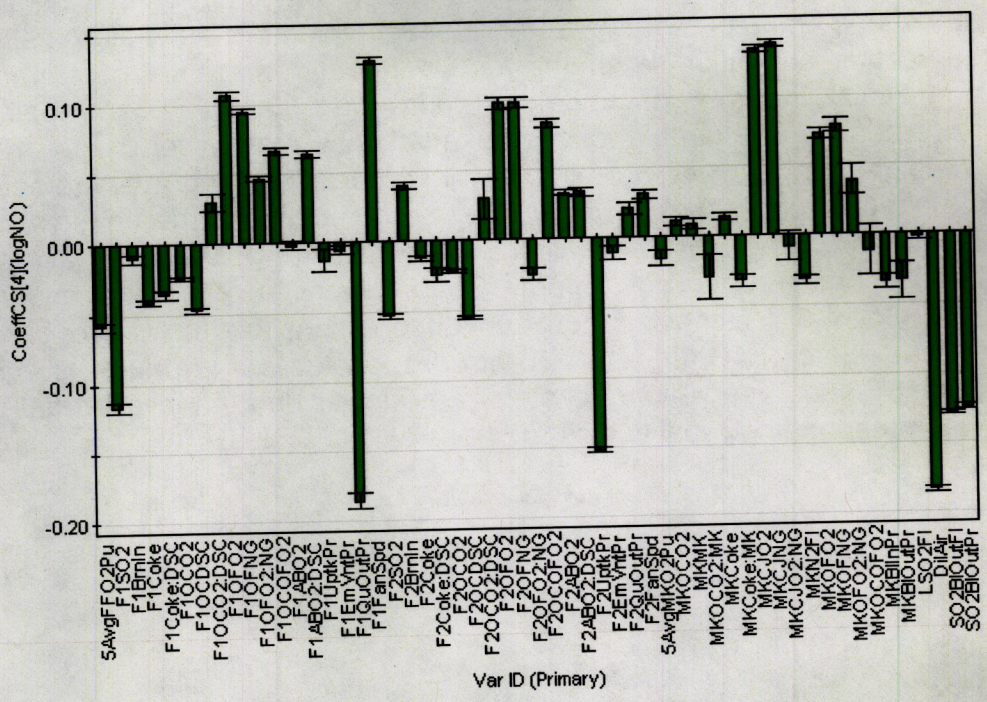

Figure 24: Coefficient Plot

The first variable listed is 5AvgFFO2Pur which is a measure of the purity of the oxygen used in the furnaces. The average oxygen purity is approximately 95% and the impurities are comprised mainly of nitrogen. Given the high temperatures in the flash furnace, nitrogen contained in the oxygen supplied to the furnace may be oxidized to NO. Therefore as the proportion of nitrogen in the oxygen increases, the oxygen purity decreases and the potential for NO formation increases. Therefore the negative correlation between oxygen purity and NO is expected.

The variables associated with furnace off-gas SO_2 strength, F1SO2 and F2SO2 are also negatively correlated with NO. Low off-gas SO_2 strength is attributed to low furnace throughput. During periods of low SO_2 strength the concentrations of other gasses fed to the acid plant, including NO proportionally increase. The oxy-concentrate variables associated with furnace throughput, F1OCO2, F1OCDSC, F2OCO2 and F2OCDSC are therefore also negatively correlated with NO. Periods of low SO_2 strength may also occur when air infiltrates into the furnace. Given the

high furnace temperature NO may be formed through oxidation of nitrogen contained in the air.

During times of low furnace throughput oxy-fuel burners may be used in order to maintain the furnace temperature. NO may form through the incomplete combustion of natural gas and as such all of the variables related to natural gas use: F1OFO2, F1OFNG, F2OFO2, F2OFNG, MKCJO2, MKCJNG, MKOFO2 and MKOFNG, are positively correlated with NO.

The afterburner oxygen to dry solids charged to the furnace ratio variables: F1ABO2:DSC and F2ABO2:DSC are both positively correlated with NO. Afterburners are used to destroy hydrogen sulphide (H_2S) gas, which may form in the furnace uptake. When the furnace throughput or DSC is reduced but the afterburner oxygen rate is unchanged the ABO2:DSC ratio increases. When this ratio is higher than stoichiometrically required for H_2S destruction, the excess afterburner oxygen may oxidize the nitrogen present in the furnace uptake thus forming NO, given the high uptake temperature. Therefore as the ABO2:DSC ratio increases so does the potential for NO formation.

The variables F1UptkPr, F1EmVntPr, F1QuOutPr, F2UptkPr, F2EmVntPr and F2QuOutPr were included in the model as an indication regarding the timing of furnace quench cleanings. The furnace quench is used to cool the off-gas as it exits the furnace uptake and enters the gas cleaning operation. As a result of the drastic temperature reduction, solids build up at the inlet to the quench which must be cleaned. During cleanings the furnace uptake bell damper is opened and gas is directed to atmosphere via a smokestack as opposed to the acid plant. The addition of feed to the furnace is also ceased at this time. Once cleaning is complete the bell damper is closed, gas is once again directed to the acid plant and furnace throughput is resumed. Two of the largest model coefficients are assigned to the variables related to the quench cleaning. As shown below variation in the #1 furnace quench outlet pressure (F1QuOutPr) and #2 furnace uptake pressure (F2UptkPr) coincide with variation in the acid plant feed gas NO concentration. It is believed that NO is formed during quench cleanings as a result of air infiltration and as a result of the addition of oxygen to the furnace in the absence of feed. The second instance is characterized by elevated oxygen to feed ratios (OCO2:DSC) and elevated afterburner oxygen to feed ratios (ABO2:DSC) which also display positive model coefficients The variation in these two

ratios relative to the variation in the uptake pressure for #2 furnace is also given below. This relationship between the oxygen ratio variables and the quench pressure variables is also outlined on the weight plot as the variables are located close to one another.

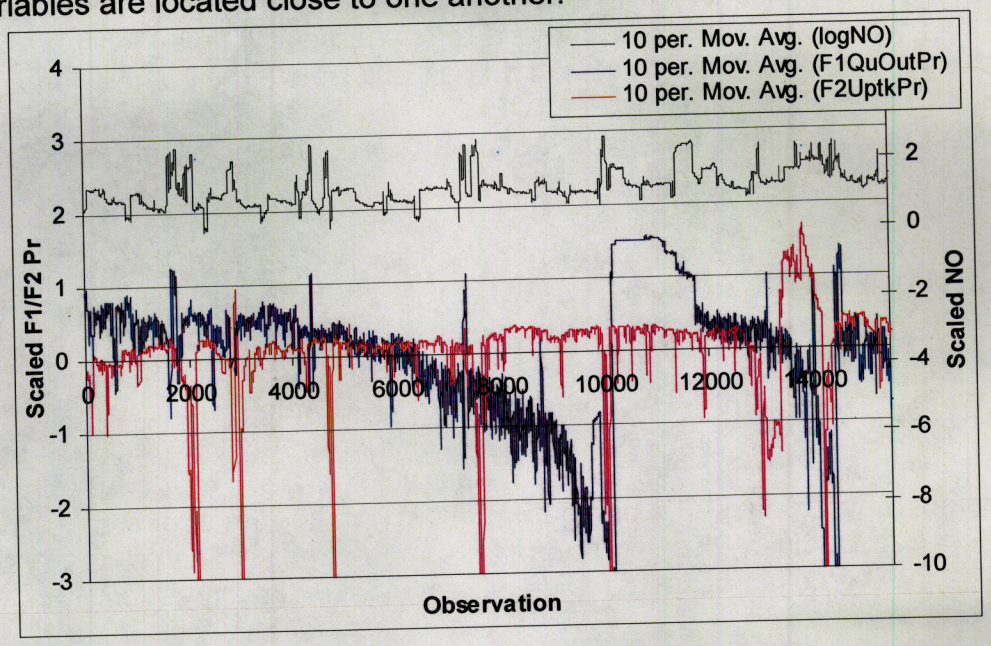

Figure 25: Furnace Uptake Pressures and logNO by Observation

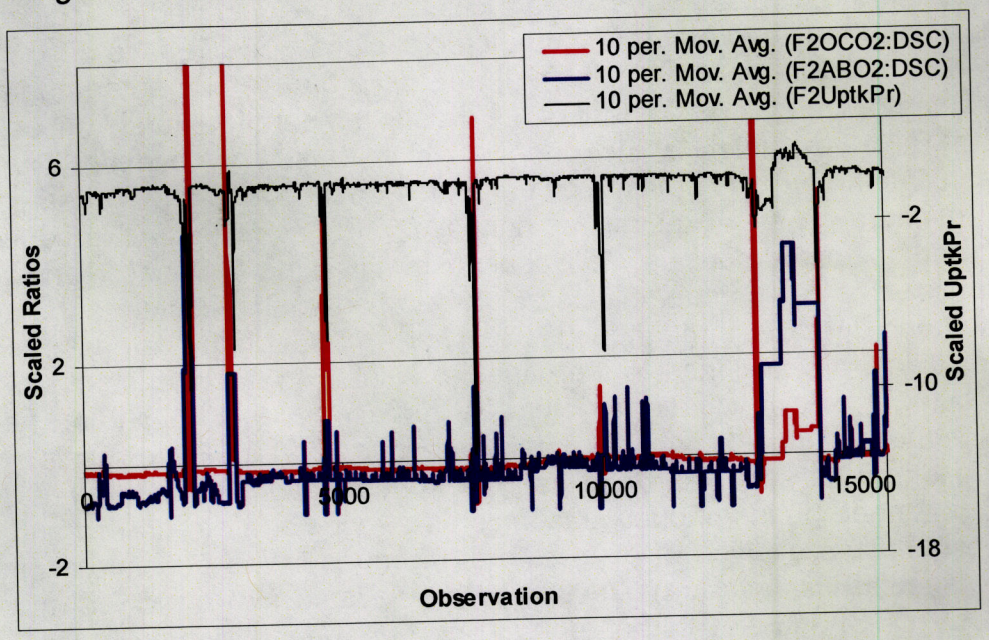

Figure 26: Furnace Uptake and Oxygen Ratios by Observation

Both of the variables related to the furnace fan speeds are positively correlated with NO. This may also be attributed to air infiltration. As the fan speed increases so does the amount of air that is drawn into the furnace and so does the potential for NO formation given the high furnace operating temperatures. High furnace fan speeds are used to ensure the workroom environment is kept free of SO_2.

The variables related to the copper reactor throughput, MKOCO2, MKMK, MKOCO2:MK, MKCoke, MKCoke:MK possess small coefficients as outlined on the plot. The copper reactor may be considered somewhat of a batch process in that feed is introduced continuously over an approximately four-hour period and is then ceased. The contents of the reactor are then transferred to another vessel for further processing. During the four-hour period over which feed is introduced the rate is not highly variable, as is the case with the continuous flash furnace operation. As such the small coefficients displayed by the copper reactor throughput variables are expected. This is also apparent through examination of the weight plot as the variables are located close to the center.

Unlike the throughput variables, the variables related to copper reactor oxy-fuel burner use are strongly positively correlated with NO. This is expected since NO may form upon the incomplete combustion of natural gas.

Finally the variables related to the acid plant throughput: DilAir, SO2BlOutFl and SO2BlOutPr are strongly negatively correlated with NO. If the quantity of NO contained in the feed gas remains constant but the amount of dilution air added to the feed gas decreases the NO concentration will consequently increase.

In order to further examine the process variables that contribute to variation in the NO concentration the contribution plot was examined. Twenty-six groups of observations with elevated NO values were selected and the average concentration for the elevated groups was 144 ppm compared to the overall average of 37 ppm. The contribution to the first two scores for the elevated group of observations compared to the average observation is given below.

The variables that contribute to the variation in NO include those outlined as significant through examination of the weights and coefficients. For both furnaces the off gas SO_2 strength contributes significantly along

with the furnace throughput, burner use, oxy-concentrate and afterburner oxygen to feed ratios, oxy-fuel burner use and the three pressure variables related to the quench cleaning. For the copper reactor the significant variables include oxy-fuel burner use and the reactor off gas cleaning blower outlet pressure. The blower outlet pressure may contribute in a manner similar to the variables related to the acid plant volumetric flow rate, which also make significant negative contributions.

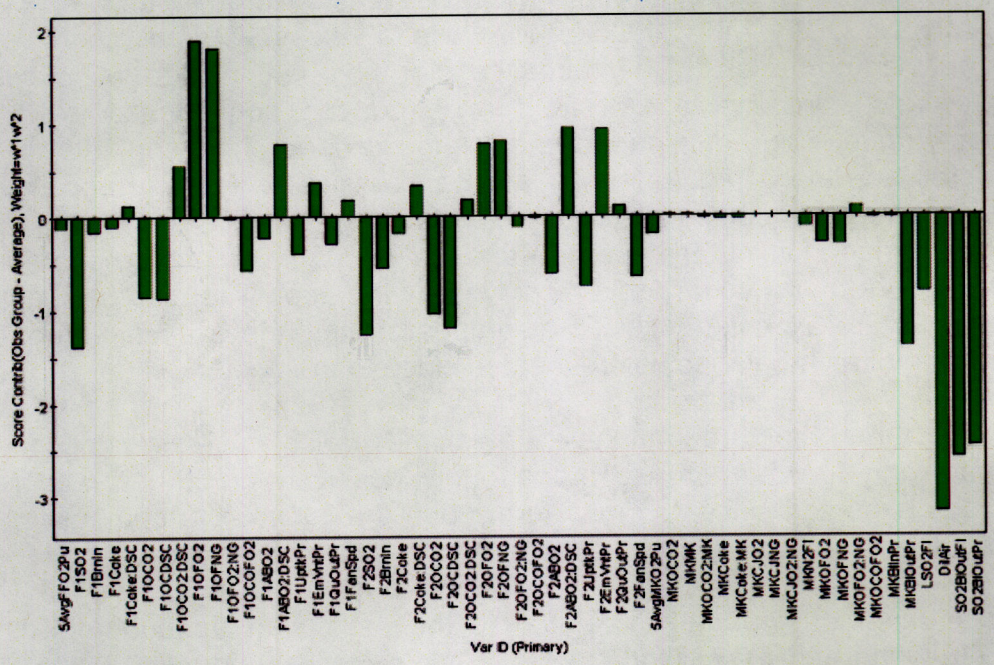

Figure 27: Contribution Plot Between Observations with Elevated NO and Average

In order to improve the predictive ability of the model all of the variables with the exception of those associated with the acid plant were assigned a lag of twenty. This corresponds to an actual lag time of ten minutes. The resulting model was comprised of six principle components with $R^2X = 0.53$, $R^2Y = 0.78$ and $Q^2 = 0.78$. Other lag times were modeled however the best prediction of variation was achieved using twenty lags. The lagged model may be compared to the original model which consisted of four principle components with $R^2X = 0.43$, $R^2Y = 0.73$ and $Q^2 = 0.73$.

The model predicted and actual plots are shown below. The predicted vs. actual logNO displays a good correlation and therefore a large extent of the variation in the Y-data is accounted for by the model.

However as shown on the time series plot, the full magnitude of the extreme positive deviations in the NO value from the average is not captured by the model in a number of instances.

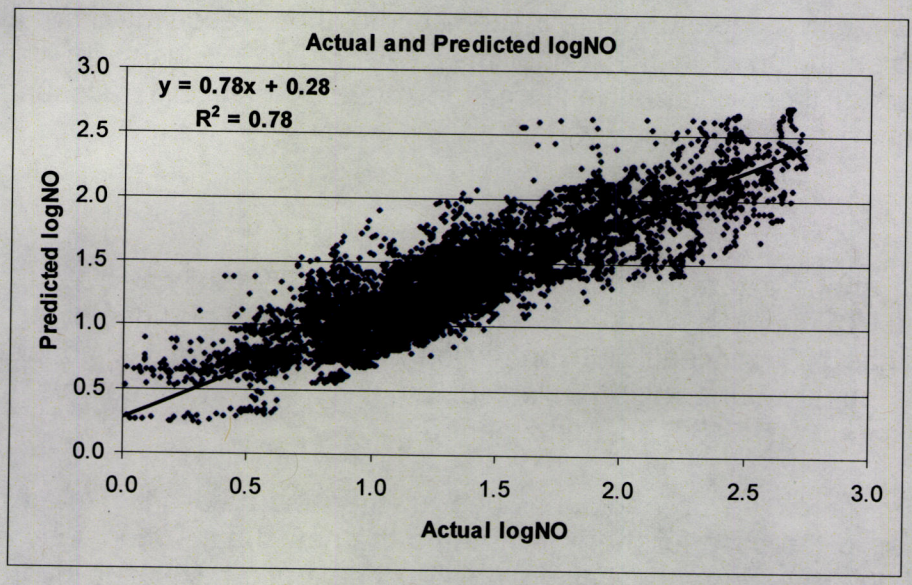

Figure 28: Actual vs. Predicted logNO

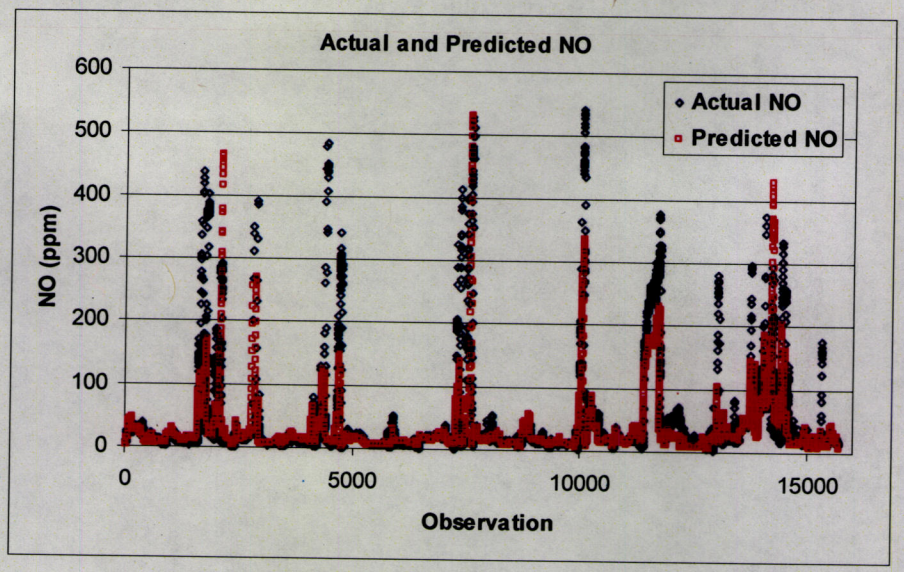

Figure 29: Actual and Predicted NO by Observation

The inaccuracy of some of the measured variables included in the model such as solids and coke addition rates may account for the less than complete explanation of variance in the case of the elevated NO

measurements. Although the temporarily installed gas analyzer was calibrated to within a couple of parts per million of the standards used each day, error inherent to the device may also account for some of the discrepancy. Finally alternative methods for NO formation in the flash furnace and copper reactor in addition to those accounted for by the variables included in the model may exist. The exclusion of variables related to these alternative formation methods may also account for a portion of the unexplained variation.

4.4. New Data Inclusion

Given the extent to which the variation in the Y-data was accounted for by the PLS model a set of new data was input for the purpose of acid plant feed gas NO concentration prediction. The new data set consisted of 5231 observations also collected at thirty second intervals. This corresponds to a 43.6 hour time period.

The predicted and actual values for the new data are plotted below. Similar to the case for the training data the new data is modeled well. In addition the extreme deviations in NO concentration exhibited by the new data are modeled well with the exception of one instance, which is circled in Figure 31.

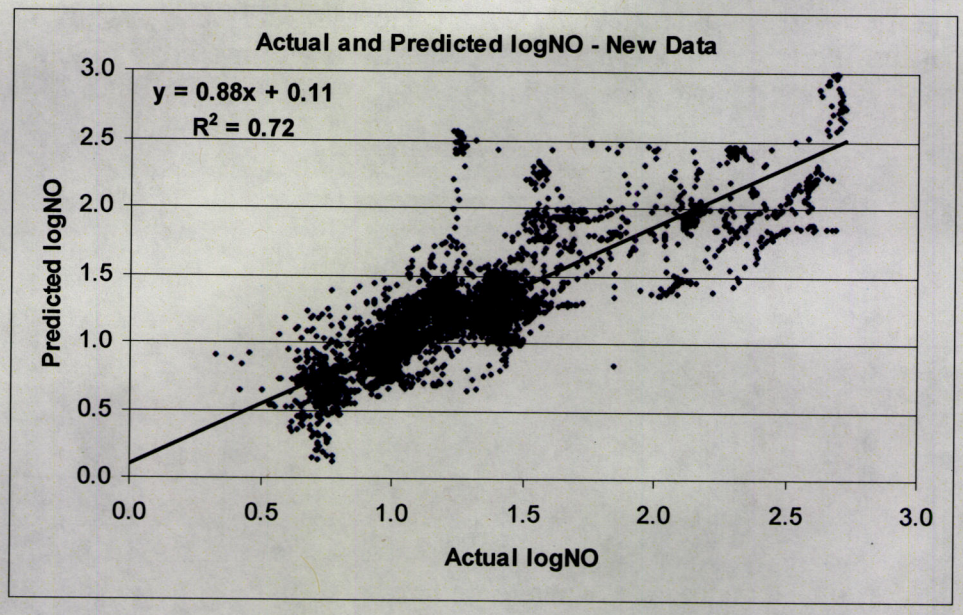

Figure 30: Actual vs. Predicted logNO – New Data

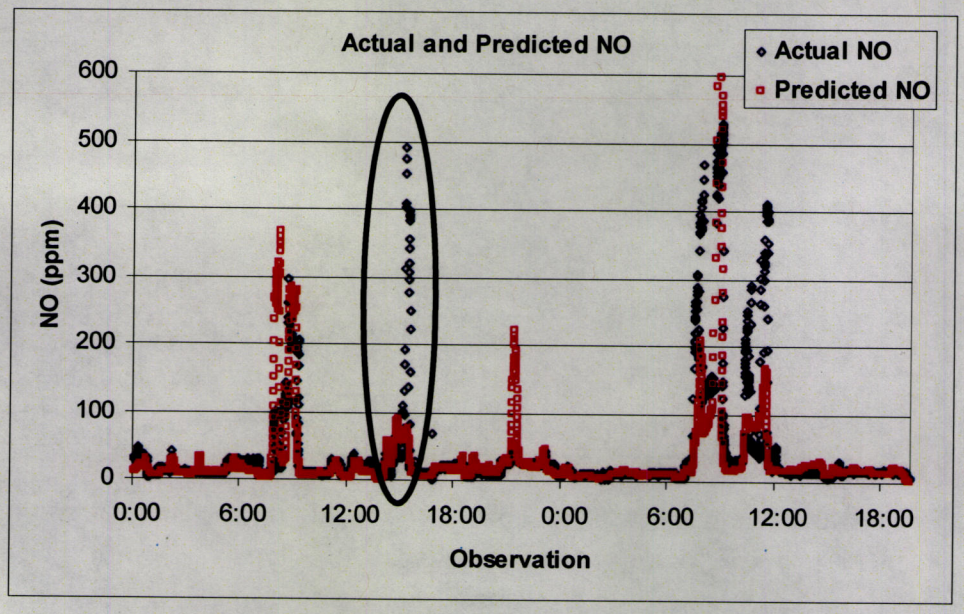

Figure 31: Actual and Predicted NO by Observation – New Data

Control charts were then developed for the purpose of identifying periods of potentially decreased predictive model accuracy. The distance

to the model in X-space ($DMODX_6$) and Hotelling's T^2 (T_6^2) value over all six of the model components are shown below as time series plots along with the predicted NO concentration.

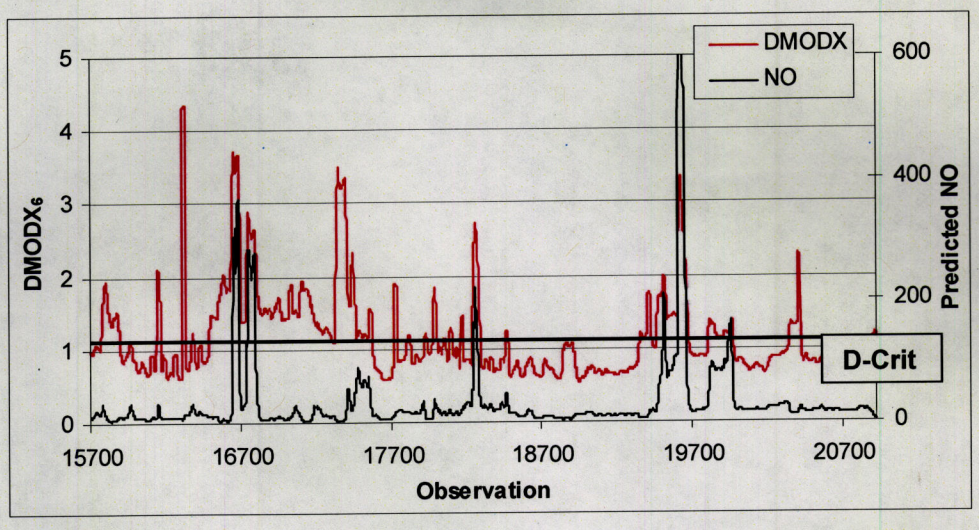

Figure 32: DMODX and NO by Observation

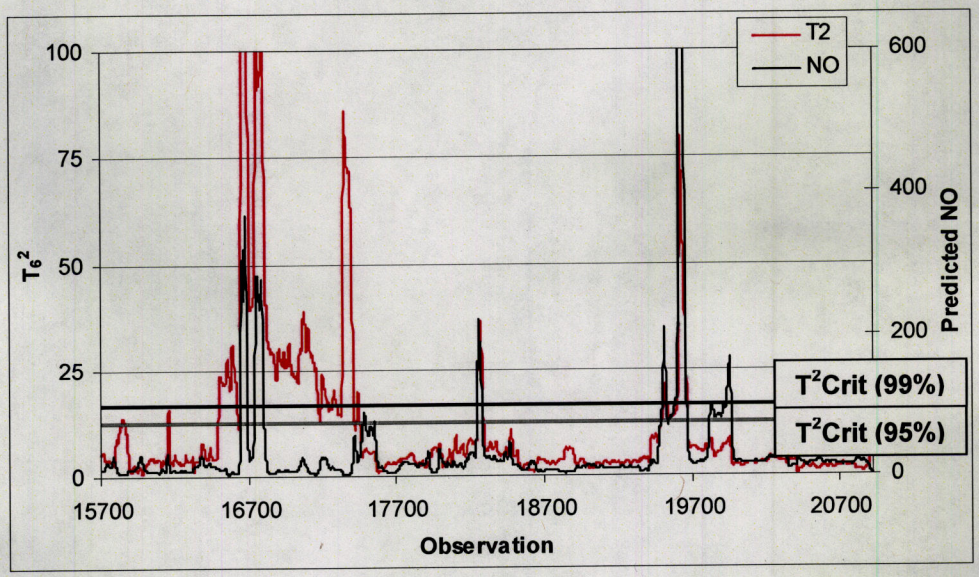

Figure 33: T_6^2 and NO by Observation

As shown above there exists a period of time beginning just before observation 16700, during which the control values are persistently above

D-Crit and T^2Crit (99%). During this period operating conditions within the two flash furnaces, the copper reactor and the acid plant may have differed compared to the conditions which prevailed when the data used for development of the PLS model was collected. As such during this period there exists variation in the data that is not accounted for by the model. Predictions of the NO concentration will be less accurate compared to those made when the control values are below the critical ones. The observations that exceed the T2Crit values shown above may be considered strong outliers and as such would be located outside the score scatter plot confidence ellipse. Similar to the case for DMODX, predictions of the NO concentration are less accurate when the T^2 value is greater than T^2Crit.

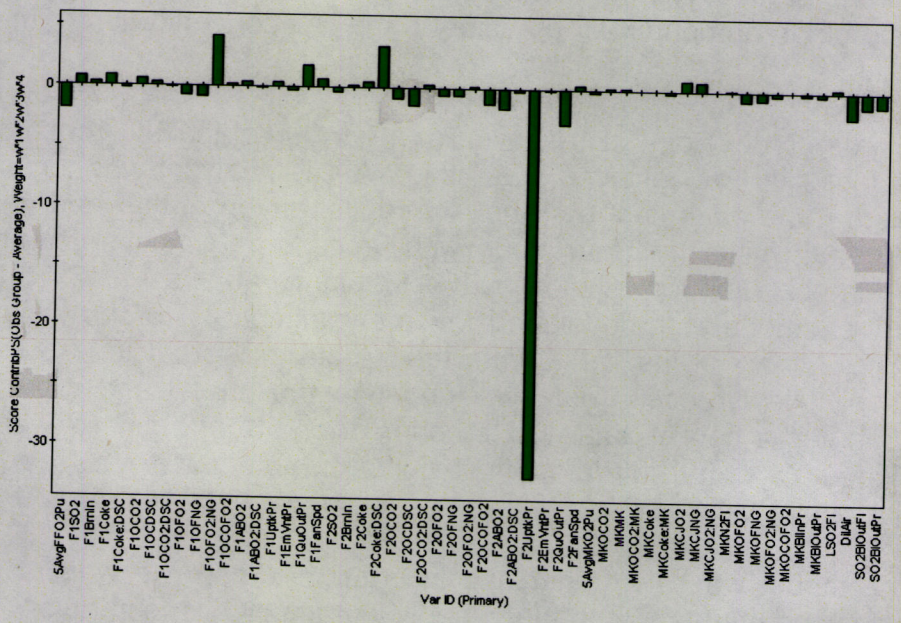

Figure 34: Contribution Plot – First Group of Outliers Compared to Average Observation

The contribution between the first group of outlying observations, occurring prior to observation 16700 was compared to the average contribution to the first four weights for the non-lagged model. As shown on Figure 34 one of the variables related to the quench for #2 furnace, F2UptkPr contributes strongly negatively. As such a maintenance activity related to the furnace uptake may have been performed during this period, which did not occur during the period when the data used to construct the

model was collected. Finally the contribution profile was similar for the two other groups of outlying observations, occurring just after observation 16700 and just before observation 19700.

4.5. Conclusion

At INCO's Copper Cliff Smelter Acid Plant, sulphur dioxide gas from three pyrometallurgical sources is treated and concentrated sulphuric acid is produced. Although the acid plant feed gas is continually analyzed for SO_2, analysis is not provided regarding the concentrations of other components. Of these components nitric oxide is of importance due to the NO specification placed on the final sulphuric acid product. In order to control the product NO concentration reagents are added in proportion to the acid plant throughput as opposed to the NO concentration of the feed gas.

In order to develop a soft sensor for prediction of the NO concentration of the feed gas PLS analysis was used. A gas analyzer was temporarily installed at the acid plant and the NO concentration of the feed gas was measured. Using this data along with process data relating to fifty-five flash furnace, copper reactor and acid plant variables a six principle component model was developed which explained 53% of the variation in the X-data, 78% of the variation in the Y-data and predicted 78% of the variation. New process data was input to the model and 72% of the variation in the new Y-data was accounted for. Control charts were also developed in order to detect instances of potentially reduced model predictive ability.

The model was also examined in order to determine which process variables contribute most to the variability of the NO concentration. Using the weight, coefficient and contribution plots several variables were deemed significant. The first group of significant variables is related to the flash furnace throughput. The furnace off gas SO_2 concentration and oxy-concentrate burner flow rates displayed a negative correlation with NO concentration as did the variables related to the acid plant volumetric flow rate. Instances of low furnace throughput may be correlated with instances of oxy-fuel burner use in order to maintain sufficient temperature. All variables related to natural gas burner use displayed strong positive correlations with NO. The ratio between coke and solids charged to the furnace also displayed a positive correlation with NO as did the furnace fan speeds.

The occurrences of two furnace maintenance activities were also attributed to the variation in NO concentration. The first involves maintenance related to the furnace oxy-concentrate burners, during which time the burners are removed for replacement or cleaning. The number of burners in use was negatively correlated with NO. During a period of burner removal it is estimated that NO is formed via air infiltration to the furnace. The second maintenance activity is related to the cleaning of the furnace uptake or quench since periods of elevated NO concentration correspond to variations in furnace uptake, emergency vent and quench outlet pressures. During these periods of pressure variation, variation was also noted with respect to the afterburner oxygen to solids charged to the furnace ratio and to the oxy-concentrate oxygen to solids charged to the furnace ratio. These two ratios displayed positive correlations with NO.

The acid plant feed gas NO concentration was predicted accurately following the input of new process data only into the PLS model developed. As such this soft sensor may be used by acid plant personnel for indication regarding the NO concentration of the incoming feed gas. In addition to indication, the sensor output may be used in an automated NO reagent control scheme whereby the addition of reagents to control the final product NO concentration is proportional to the feed gas NO concentration as opposed to the production rate as is currently the case. Under this scheme the T^2 and DMODX values may be used to identify periods of reduced model predictive accuracy. During these periods reagents may be added in a proportion greater than those stoichiometrically required for NO control as a safety factor to avoid the production of off-specification acid.

Finally in order to reduce the amount of NO generated in the flash furnace and copper reactor operations the variables deemed significant may be investigated further. Significant efforts have been made in the past and continue to be made with respect to reduction of furnace air infiltration and to efficient burner design. However the evaluation of the furnace operation during burner removal and quench cleanings may also lead to a reduction in the potential for NO formation.

5. Chapter 5 – Nickel Powder Decomposer Models

5.1. Introduction

At INCO's Copper Cliff Nickel Refinery the INCO Pressure Carbonyl (IPC) process is used to produce high purity nickel pellets and powder. The process consists of high-pressure nickel carbonyl ($Ni(CO)_4$) reactor and nickel pellet and powder decomposer unit operations. During the nickel powder decomposer operation nickel carbonyl feed gas produced during the reactor operation, carbon monoxide and oxygen enter the decomposer at the top through a water cooled annulus. The annulus is cooled in order to ensure that nickel carbonyl does not prematurely decompose and build up. The gasses then pass through four independently controlled electrically heated zones. The heat causes the nickel carbonyl to decompose into pure nickel powder and carbon monoxide (CO) gas. The nickel powder then passes through the connecting box and settles in the consolidator. The CO gas exits the connecting box and passes through the water cooled off gas lines to the baghouse. The nickel powder collected on the baghouse filter bags is returned to the connecting box. The CO gas exits the baghouse and is processed further before it is recycled for re-use in the reactor operation. Every four hours the nickel powder is discharged from the consolidator to a discharge container. It is then purged free of nickel carbonyl gas, sampled and transferred to the nickel powder handling system where it is subsequently blended and packaged for shipment to market. The sampled nickel powder is analyzed for bulk density (AD) and particle size (FSSS). These two physical properties are varied through manipulation of six process variables: the four external wall temperatures, the internal zone four temperature and the oxygen addition rate depending on the type of nickel powder targeted for production. A schematic of the nickel powder decomposer is given below.

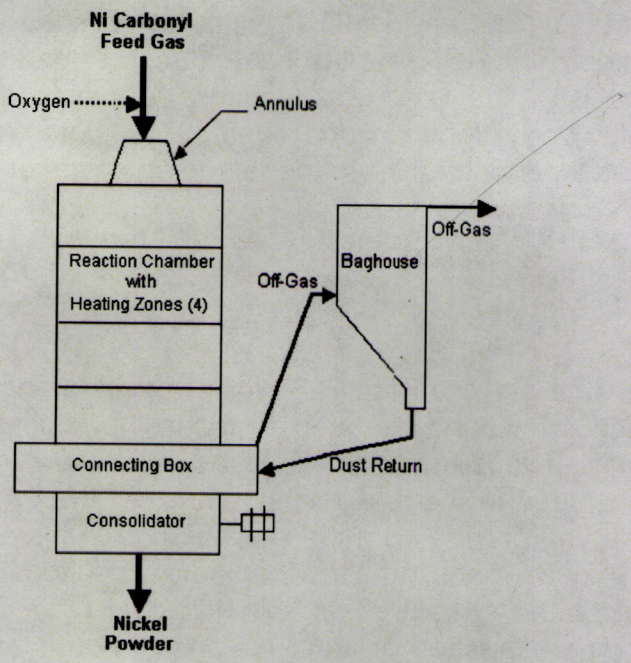

Figure 35: Nickel Powder Decomposer Schematic

Two analyses were carried out using data sets that summarized the operation of the nickel powder decomposer. In the first analysis PLS was carried out in order to determine whether the two measured nickel powder physical characteristics might be predicted based on the process operating conditions. Average process data for thirty variables was obtained during the four-hour period between sampling of the powder and was used along with the measured AD and FSSS values to construct a data set which was analyzed using PLS. In addition the PLS model was used to determine the process variables that are significant with respect to the variation observed for the two powder physical properties. Upon evaluation of the PLS model, two modes of operation were identified: one corresponding to the production of low AD and FSSS nickel powder and the other to the production of high AD and FSSS nickel powder. Individual PLS models were fit to the data pertaining to each mode of production and the extreme outliers were removed. The outlier-free data sets were used as a basis for the good operation of the nickel powder decomposer under the two modes. As such a second analysis was performed whereby PCA was applied to each of the two data sets with the AD and FSSS variables

removed for the purpose of developing an online multivariate modeling scheme. Based on the PCA models formed, control limits for the scores, Hotelling's T^2 statistic and DMODX were established. New instantaneous data collected during each of the two modes of operation were then input to the models. The new observations that exceeded the established control limits were identified and contribution plots were used to identify the variables that contributed to the observed deviation.

5.2. Nickel Powder Decomposer PLS Analysis

5.2.1. Model Development

The data set used for PLS model development consisted of thirty average nickel powder decomposer process variables collected over four-hour intervals, two Y-variables: AD and FSSS and 4 760 observations. The data included in the analysis summarized the decomposer operation over a three year period from the beginning of 2002 until the end of 2005. The initial PLS model fit to the data resulted in seven principle components with $R^2X = 0.86$, $R^2Y = 0.75$ and $Q^2 = 0.74$. The t_1 vs. u_1 score plot was examined and did not reveal non-linearity with respect to the relationship between X and Y. The t_1 vs. t_2, t_2 vs. t_3 and t_1 vs. t_3 score plots were then examined and several extreme outliers were observed. With these outliers removed the model consisted of five principle components with $R^2X = 0.72$, $R^2Y = 0.74$ and $Q^2 = 0.74$. The t_1 vs. u_1 and u_1 vs. u_2 plots were then examined for outliers with respect to the relationship between the X and Y-data. Once again some extreme outliers were identified and were removed from the data set. With these outliers removed the final PLS model was comprised of seven principle components with $R^2X = 0.91$, $R^2Y = 0.79$ and $Q^2 = 0.79$. No outliers were observed on any of the score plots and the explained and predicted variation was nearly equal for both of the Y-variables. Finally the majority of the explained and predicted variation was accounted for by the first three components.

5.2.2. Model Interpretation

The score and weight plots for the first two components are given below. As shown the Y-variables are located close to one another on the weight plot. As such they are correlated with one another and display little independent variation. This is also apparent through examination of the time series plots for the two Y-variables as increases in AD and FSSS

occur simultaneously. Given the location of the Y-variables on the weight plot, observations located in the upper right hand portion of the score plot will have correspondingly elevated AD and FSSS. For example the average AD and FSSS of the observations in the three clusters located in the upper right-hand quadrant of the score plot is 0.77 and 3.05 respectively compared to 0.52 and 2.42 for the overall data set.

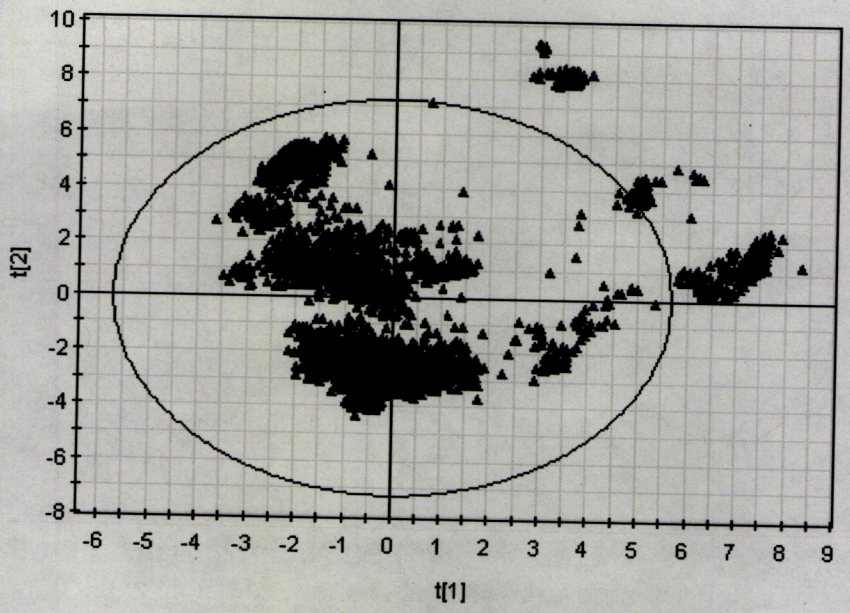

Figure 36: X-Score Plot – First Two Components

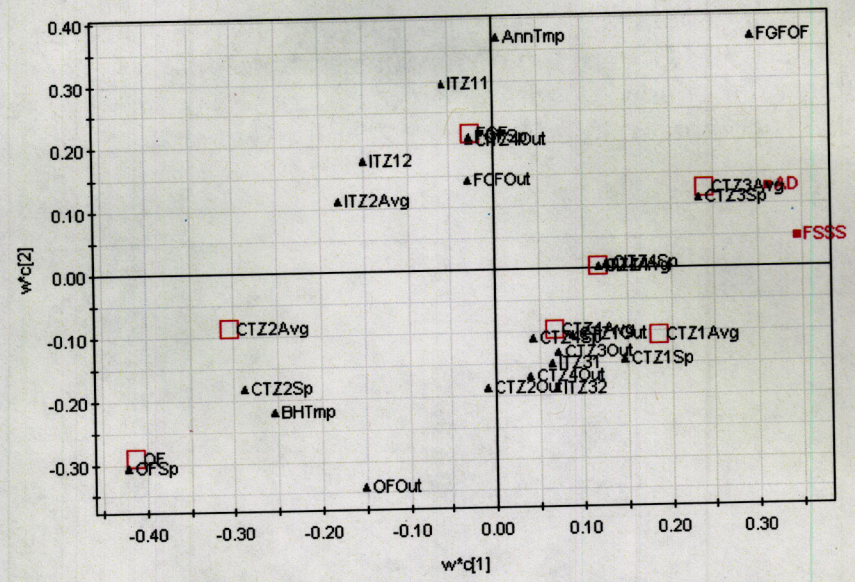

Figure 37: Weight Plot – First Two Components

The six main decomposer process variables manipulated in order to affect the nickel powder AD and FSSS are highlighted on the weight plot. Four of these variables are the external wall control zone temperatures or control zone temperatures. The measured average values of the four control zone temperatures along with the set point and controller outputs were included in the model. The control zone average temperatures are predominantly located on the right hand portion of the weight plot with the exception of zone 2. The zone 3 temperature is located the closest to the Y-variables while zones 1 and 4 are located closest to one another when the first two model components are plotted. The segregation of zone 2 is not unexpected since it is the primary zone in which the decomposition reaction takes place. The variation with respect to its temperature profile therefore differs compared to the other three zones.

The fifth manipulated variable is the oxygen flow rate for which the average measured, set point and controller output values were also included in the model. The location of the oxygen variables diagonally opposite the Y-variables is expected. In plant practice the oxygen flow rate reduced in order increase the AD and FSSS.

The final variable used to alter the nickel powder physical properties is the control zone 4 internal temperature. Once again the average measured, set point and controller output values were included in the model. As shown the average internal control zone 4 temperature is located close to the Y-variables on the weight plot, which corresponds to plant practice. This temperature is increased in order to increase AD and FSSS. In addition through the process control scheme utilized in the plant the feed gas flow is controlled by the control zone 4 internal temperature to ensure full decomposition. The clustering of the variables FGF, FGFSp, FGFOut and CITZ4Out on the weight therefore corresponds with this practice.

The ratio between the feed gas and oxygen flows (FGFOF) was also included in the model and as shown on the weight plot the variable displays significance in both the first and second components. However with this variable removed from the data set there was little change in the explained or predicted variation for both the X and Y-variables in the resulting model.

Although the other internal temperatures are not manipulated they were included in the model in order to improve the overall summary of the decomposer operation. With the variables ITZ11, ITZ12, ITZ2Avg, ITZ31, ITZ32 and ITZ4Avg removed from the data set the resulting PLS model was comprised of five principle components with $R^2X = 0.81$, $R^2Y = 0.77$ and $Q^2 = 0.77$. This may be compared to the seven component full variables model with $R^2X = 0.91$, $R^2Y = 0.79$ and $Q^2 = 0.79$. As shown on the weight plot ITZ11, ITZ12, ITZ2Avg cluster together away from ITZ31, ITZ32 and ITZ4Avg, which are located close to one another. This separation is reflected in the physical location of the internal thermocouples within the decomposer.

Finally the annulus and baghouse temperatures were included in the model. The annulus is the conical water cooled portion of the decomposer located at the top whose purpose is to prevent premature decomposition of the nickel carbonyl feed gas. The baghouse contains filter bags, which are used to remove nickel powder from the decomposer off gas. The significant weights displayed by the variables AnnTmp and BHTmp are somewhat unexpected. In actual plant practice these variables are not monitored with the same frequency compared to the control zone temperatures. However as illustrated on the weight plot, they contribute significantly to the explanation and prediction of variance for AD

and FSSS. For example with these two variables removed the resulting model consists of five principle components with $R^2X = 0.84$, $R^2Y = 0.71$ and $Q^2 = 0.71$. The increase in explanation and prediction may be a result of the correlation structure between these two variables and other process variables contained in the model. Only through designed experimentation may the causal relationship between AnnTmp and BHTmp and the Y-variables be determined. However, it is now postulated by plant personnel that perhaps the heat removed through water cooling is affecting the decomposition.

The coefficients for both of the Y-variables are given below. The close proximity of the Y-variables on the weight plot and subsequent positive correlation is also outlined on the coefficient plot. The sign of the coefficients for each of the variables is the same for AS and FSSS and in addition the magnitude of the coefficients is similar for both AD and FSSS. As such independent variation in these physical characteristics may be difficult. For example it may be difficult to cause AD to increase and FSSS to decrease. The potential for this type of operating scenario may also be investigated through designed experimentation.

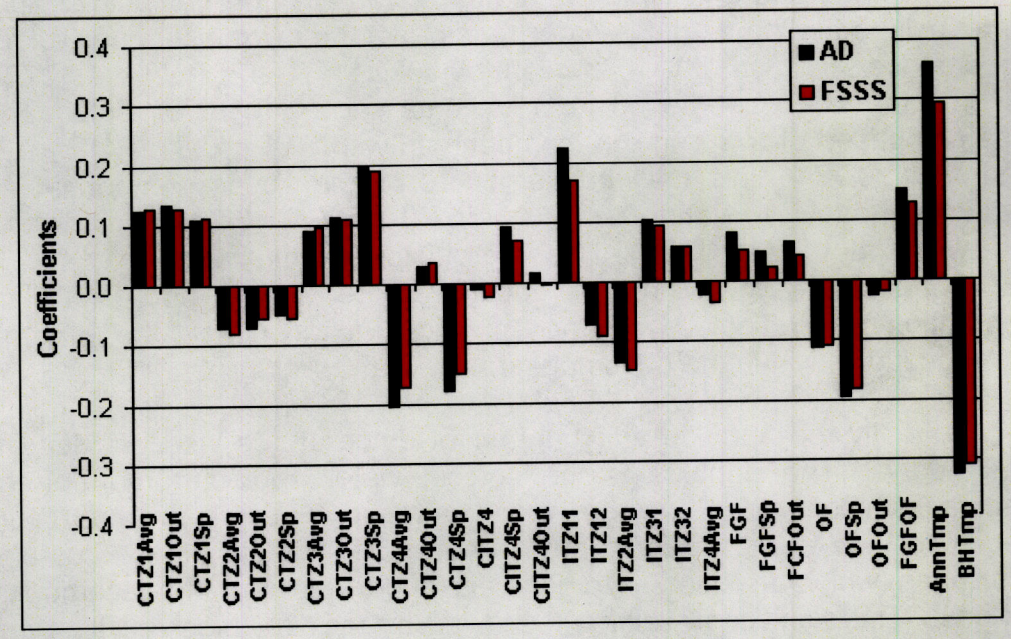

Figure 38: Coefficient Plot – AD and FSSS

According to the coefficient plot control zones 1 and 3 are positively correlated with AD and FSSS while zones 2 and 4 are negatively correlated. The time series plots for these variables were therefore examined. As shown the zone 1 and 3 set points were increased during periods when powder of elevated AD and FSSS was produced with the exception of the last portion of the graph. During this time AD and FSSS increased while the zone 1 and 3 set points did not. The inverse relationship between the zone 2 temperature set point and AD is also illustrated on the time series plot.

The positive correlation between the control zone 4 internal temperature set point and the negative correlation between the oxygen flow and AD as outlined on the coefficient plot is also evident upon inspection of the time series plot for these two variables. The greater weights attributed to the OF variables compared to the CITZ4 variables is also apparent since all of the instances of elevated AD coincide with a reduction in oxygen flow with the exception of one group.

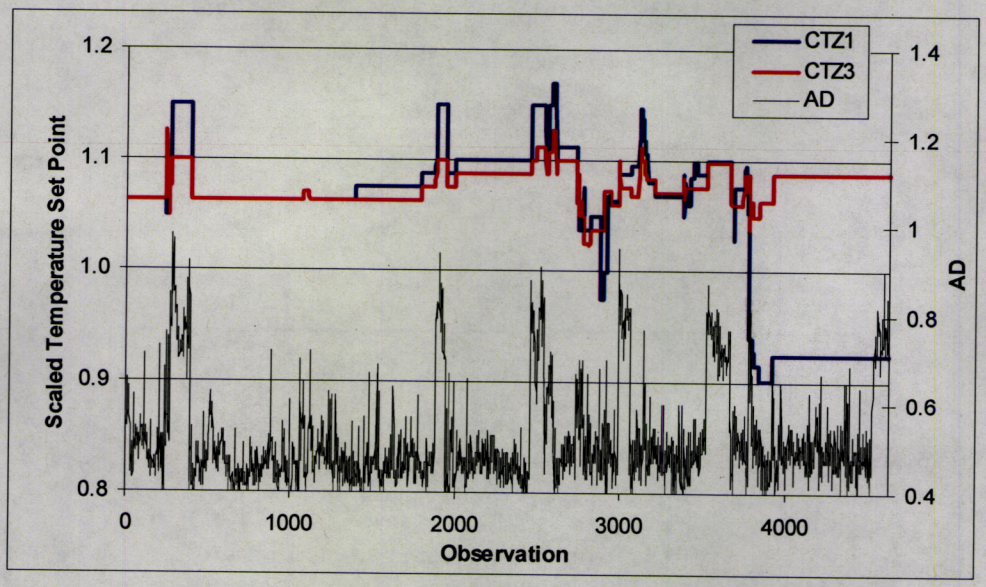

Figure 39: AD, CTZ1 and CTZ3 by Observation

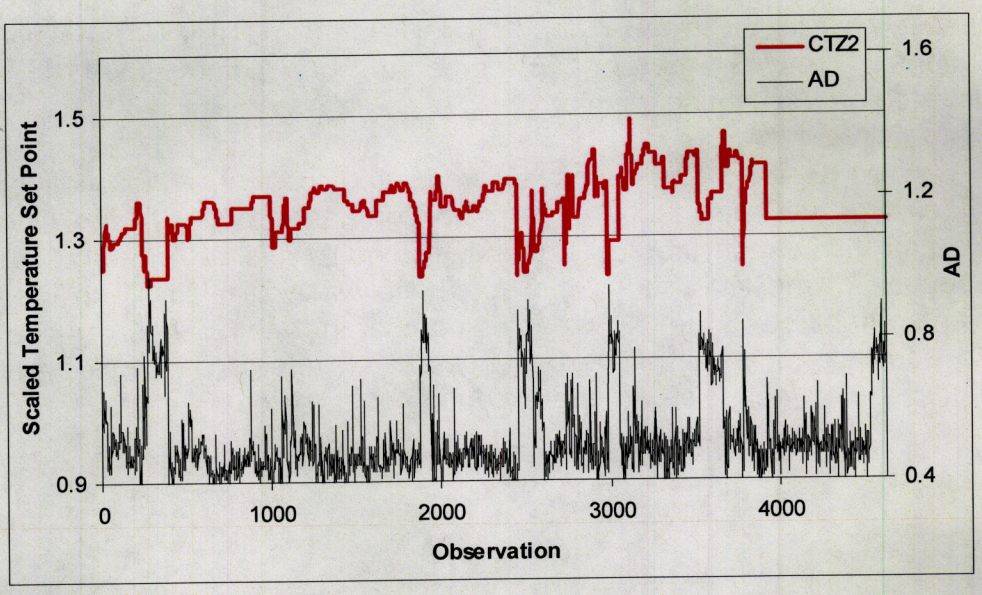

Figure 40: AD and CTZ2 by Observation

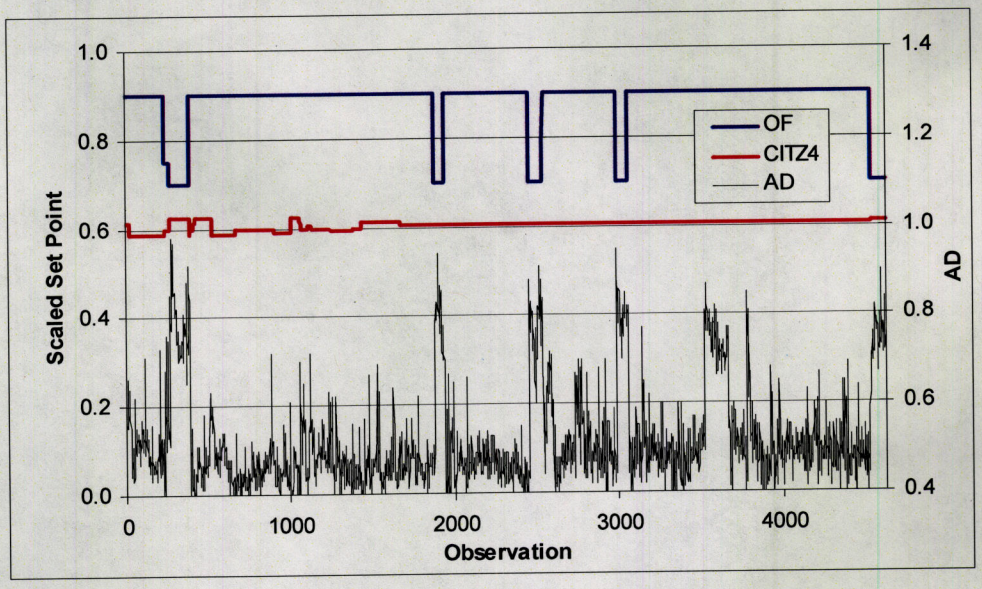

Figure 41: AD, OF and CITZ4 by Observation

The actual and predicted AD and FSSS values are given below along with the corresponding time series plots of the Y-variables. As shown the actual and predicted values correspond well. However the full magnitude of the AD and FSSS values are not realized in some instances. In addition the variation displayed by AD and FSSS at low actual levels is not fully accounted for by the model. Several factors may contribute to the inability of the PLS model to accurately capture the low level variation. These include the physical condition of thermocouples inside the decomposer and the internal surface of the decomposer wall. For example it is known that wall build up and thermocouple deterioration affect the response of the decomposer to changes in the manipulated variables. Since AD and FSSS are physical properties measured by the operators there is also some error inherent in the measurement due to sample collection techniques and measurement instrument accuracy. Finally the nickel carbonyl feed gas strength was not included in the analysis as it is not currently measured. Fluctuations in gas strength may also contribute to variation with respect to the nickel powder physical characteristics.

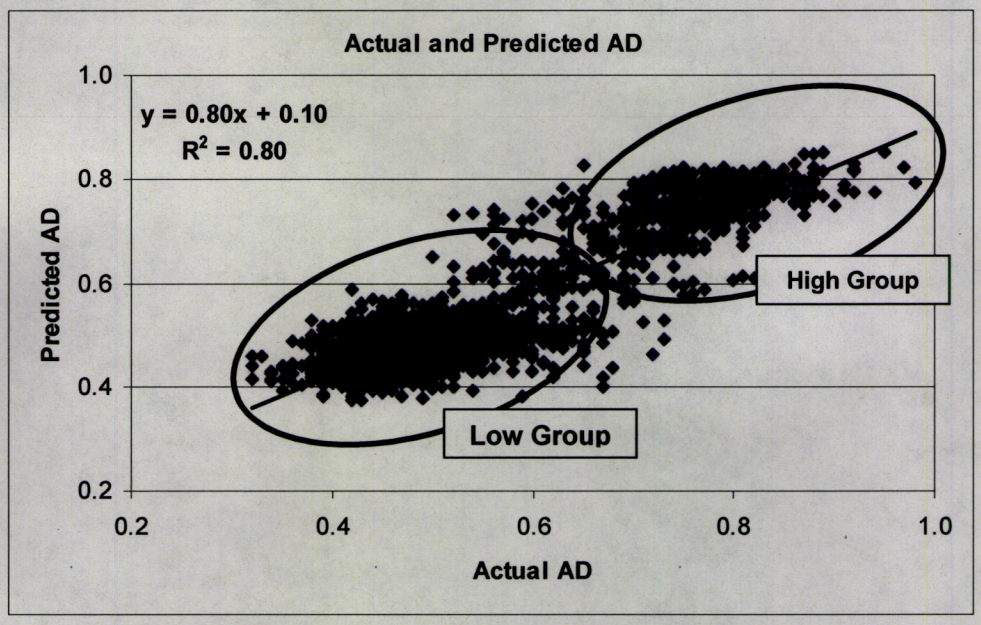

Figure 42: Actual vs. Predicted AD

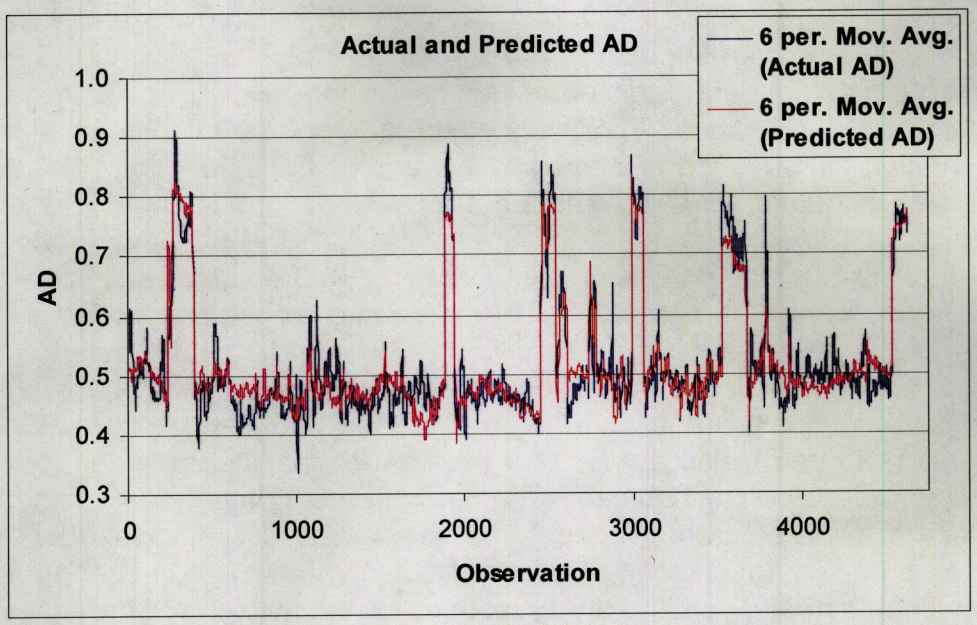

Figure 43: Actual and Predicted AD by Observation

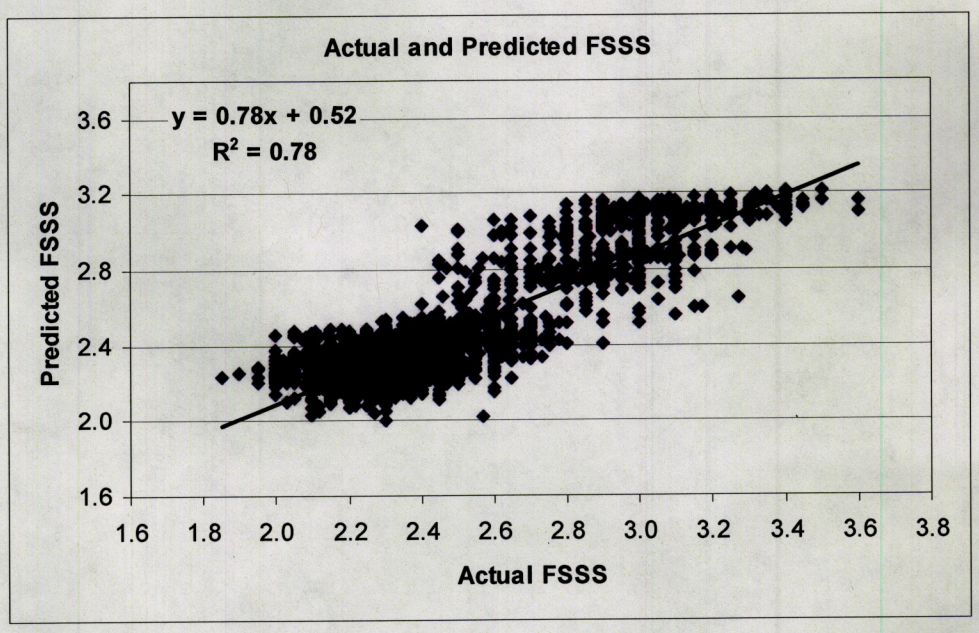

Figure 44: Actual vs. Predicted FSSS

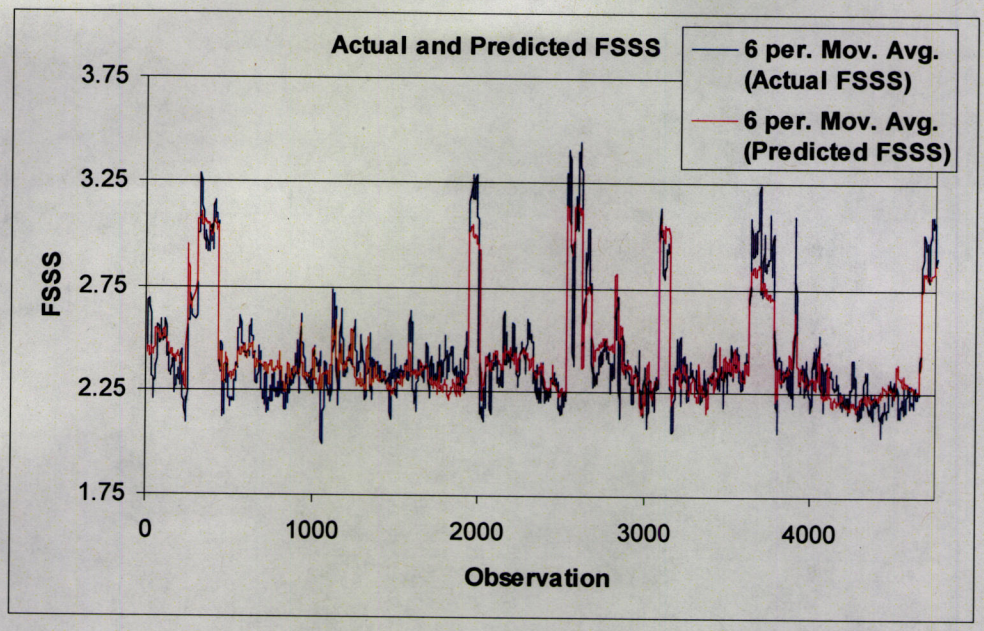

Figure 45: Actual and Predicted FSSS by Observation

As shown on the actual versus predicted AD plot two separate groups of observations are present. The Low Group is characterized by observations with low AD and the High Group by those with high AD. In order to further investigate the differences between these two groups the contributions were examined and the resulting plot is given below. In this case the contribution to the first two scores for the High Group of observations was compared to the contribution for the Low Group.

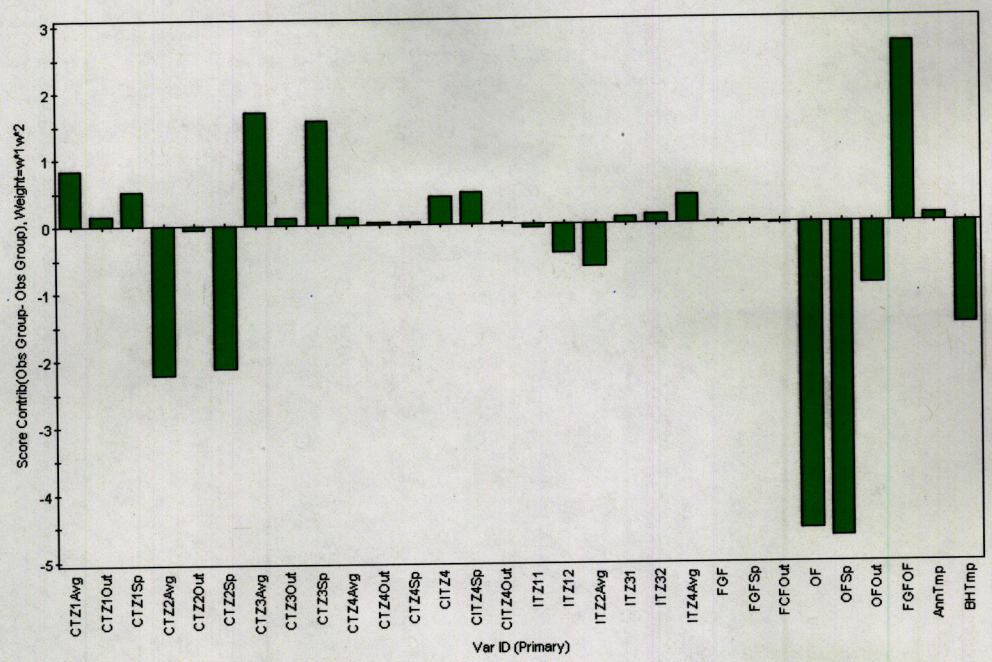

Figure 46: Contribution Plot Between Low and High Groups

As shown the control temperature zones 1 and 3 make positive contributions to the scores for the High Group compared to the Low Group while zone 2 makes a negative contribution and zone 4 makes a negligible contribution. The contribution for zone 4 is evident when the contributions to the third and fourth scores between the Low and High Groups are plotted. In this instance the contribution made by zone 4 is positive which is opposite to the sign of its coefficient. This may reflect the variation with respect to the zone 4 temperature between the two groups as opposed to the variation within the two groups. Similar to the case for zone 4, the positive contributions of the control zone 4 internal temperature and of the feed gas flow are more pronounced when the contribution to the third and fourth scores is examined.

The oxygen flow variables display the most significant contributions for the first and second components. The negative contribution is also evident through examination of the weight and coefficient plots and of the AD time series plot. As evident on the time series plot the occurrence of high AD powder production coincides with periods of low oxygen flow. During periods of significant oxygen flow reduction, the feed gas flow is maintained or is increased moderately and the feed gas to oxygen flow

ratio increases. This increase is reflected in the positive contribution made by FGFOF to the scores between the Low and High Groups. The final variable to make a significant contribution is the baghouse temperature, which contributes negatively. This negative contribution may be indicative of the prevailing conditions within the decomposer during periods of high AD powder production

5.2.3. Conclusion

High purity nickel powder is produced at INCO's Copper Cliff Nickel Refinery through the decomposition of nickel carbonyl gas in a decomposer reactor. At regular four-hour intervals powder is withdrawn from the decomposer, which is subsequently sampled prior to preparation for shipment to market. The nickel powder samples are analyzed with respect to their bulk density (AD) and particle size distribution (FSSS). In order to determine the nickel powder decomposer process variables that contribute to the variation in these two measured physical characteristics PLS analysis was performed. Four-hour average data was collected for thirty process variables related to the decomposer operation. The data set analyzed therefore consisted of thirty X-variables, two Y-variables and 4760 observations. The resulting seven principle component model explained 91% of the variation in the X-data, 79% of the variation in the Y-data and predicted 79% of the variation. Therefore the model accounted for a large extent of the variation present in the X and Y-data. In general the actual and predicted AD and FSSS values corresponded to well to each other; however the full magnitude of the positive deviation was not realized in some instances. Unmeasured parameters such as thermocouple and decomposer wall conditions may account for a portion of the unexplained variation.

Through the evaluation of the weight, coefficient and contribution plots a number of the variables were deemed significant with respect to the variation in the Y-data. In addition the strongly positive correlation between the Y-variables was also observed. As expected the four control zone temperatures were significant with zones 1 and 3 displaying positive correlations with AD and FSSS. Zone 2 displayed a negative correlation with the Y-variables and was also segregated from the other control zone temperatures on the weight plot for the first two scores. This segregation is reflected in plant operation. The 2nd zone differs from the other three in that it is the zone in which the decomposition reaction primarily takes place. The zone 4 temperature also contributed negatively to the scores

when the group of observations with high AD and FSSS was compared to the group with low AD and FSSS.

The two other decomposer parameters that are varied in order to affect changes in AD and FSSS also displayed significance. The oxygen flow was strongly negatively correlated with AD and FSSS as expected. The control zone 4 internal temperature also displayed a negative correlation which is reflected in plant practice. The correlation between this temperature and the feed gas flow was also displayed in the weight plot for the first two components. This correlation corresponds to the decomposer process control scheme as the feed gas flow is controlled by the zone 4 internal temperature. Although not manipulated variables, the annulus and baghouse temperatures were significant with respect to the modeling of the variation in the decomposer process and of the Y-data.

Finally two groups of observations were identified on the predicted AD plot; one corresponding to the production of nickel powder with low AD and the other to production of nickel powder with high AD. Since the correlation structure between the decomposer process variables and the nickel powder AD and FSSS was in general as expected for the six manipulated variables and since the extent to which the variation in both the X and Y-data was explained by the PLS model was good, the two groups of observations identified were used as a basis for the development of an online monitoring scheme.

5.3. Nickel Powder Decomposer PCA Analysis

5.3.1. Introduction

Given the extent to which the variation in the decomposer process parameters and the nickel powder AD and FSSS were accounted for by the PLS model, further multivariate modeling was performed using the initial model as a basis. Separate PLS models were fit to the observations within Low Group and the High Group, previously utilized for contribution plotting. Once the extreme outliers and Y-variables were removed, PCA was preformed on the resulting data set. The PCA models were then used as the basis for the good operation of the powder decomposer during the production of both low and high AD and FSSS nickel powder. These models were investigated in order to determine the relationship between the process variables and the model principle components. New instantaneous data collected at ten-minute intervals was then input into

the models and the operation was monitored using the control charts developed. Finally contribution plots were used to determine the variables contributing to the out-of-control signals.

5.3.2. Model Development – Low Group

The Low Group data set was comprised of 3904 observations with average AD and FSSS values of 0.48 and 2.32 respectively. Since the Low Group is a sub-set of the original data set it also consisted of four-hour average data collected for thirty X-variables. The initial PLS model fit to the Low Group data consisted of five principle components with R^2X = 0.76, R^2Y = 0.23 and Q^2 = 0.23. A number of extreme outliers were removed from the data set following examination of the t_1 vs. t_2, t_2 vs. t_3 and t_1 vs. t_3 score plots. The resulting model was also comprised of five principle components with R^2X = 0.77, R^2Y = 0.24 and Q^2 = 0.23. No outliers were observed on the t_1 vs. u_1 score plot and as such the data set contained no extreme outliers in the relationship between the X and Y-data. The explanation and prediction of variation may be compared to the initial PLS model fit to the complete data set which was comprised of seven principle components with R^2X = 0.91, R^2Y = 0.79 and Q^2 = 0.79. The degree of variation for both the X and Y-variables is much less for the Low Group data set compared to the complete set and accounts for the reduction in explained and predicted variation.

Following the removal of the extreme outliers from the Low Group data set the Y-variables were also removed and PCA was performed on the X-data. The initial PCA model consisted of six principle components with R^2X = 0.89 and Q^2 = 0.74. The t_1 vs. t_2, t_2 vs. t_3 and t_1 vs. t_3 score plots were examined and a number of extreme outliers were removed from the data set. The final PCA model fit to the Low Group data was comprised of six principle components with R^2X = 0.89 and Q^2 = 0.74. The majority of the variation in the data set was accounted for by the first four principle components.

5.3.3. Model Interpretation – Low Group

The explained and predicted variation by component was then examined in order to determine the relationships between the model components and variables. As shown below the first component accounts for the explained and predicted variation for all of the variables with the exception of CTZ3Avg, CTZ3Sp, CTZ4Avg, CTZ4Sp, CITZ4, CITZ4Sp,

ITZ4Avg, OF and OFSp. With the addition of the second component the extent to which the variables CTZ4Avg and CTZ4Sp are accounted for increases. Similarly with the addition of the third component the variables CITZ4, CITZ4Sp and ITZ4Avg increase in the extent to which their variation is explained and predicted. Finally the explained and predicted variation for the variables CTZ3Avg and CTZ3Sp and BHTmp increase with the addition of the fourth component.

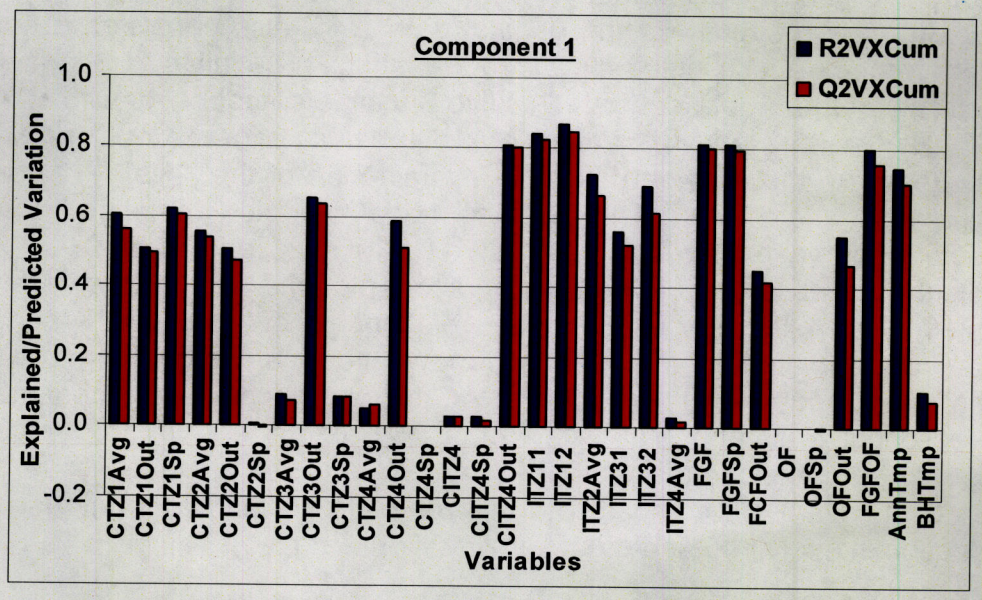

Figure 47: R^2VXCum and Q^2VXCum – Component 1

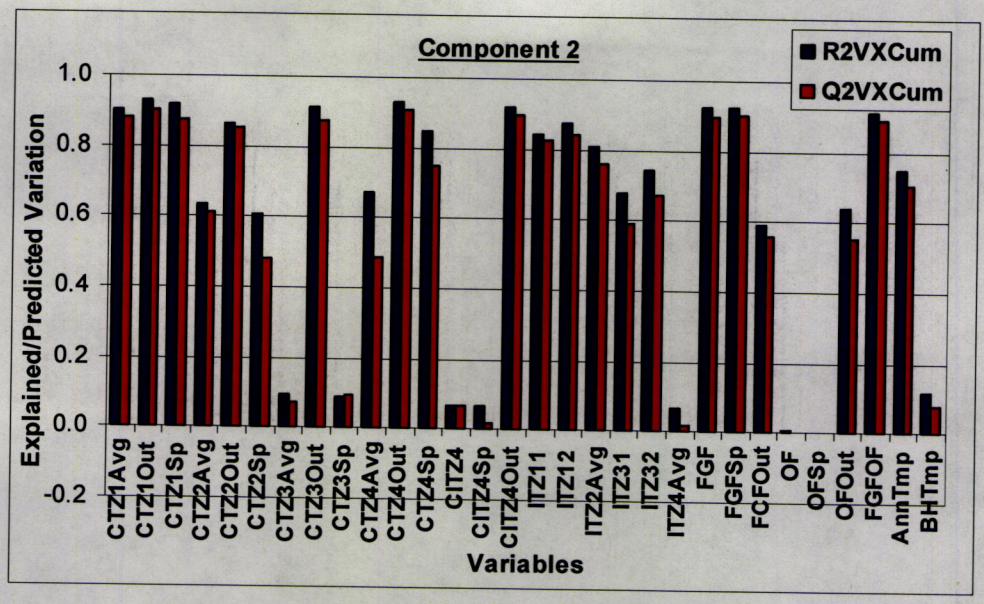

Figure 48: R^2VXCum and Q^2VXCum – Component 2

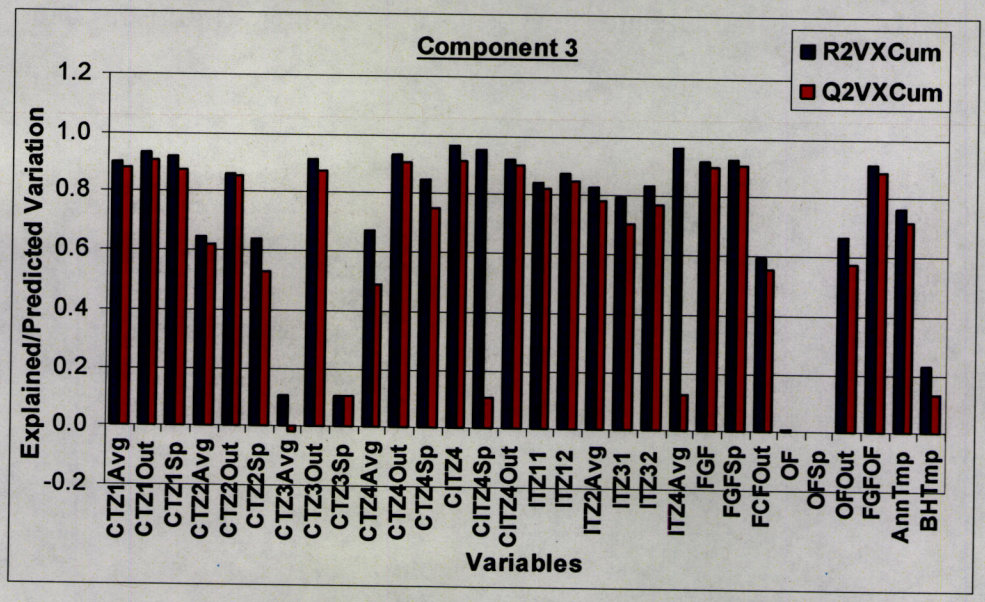

Figure 49: R^2VXCum and Q^2VXCum – Component 3

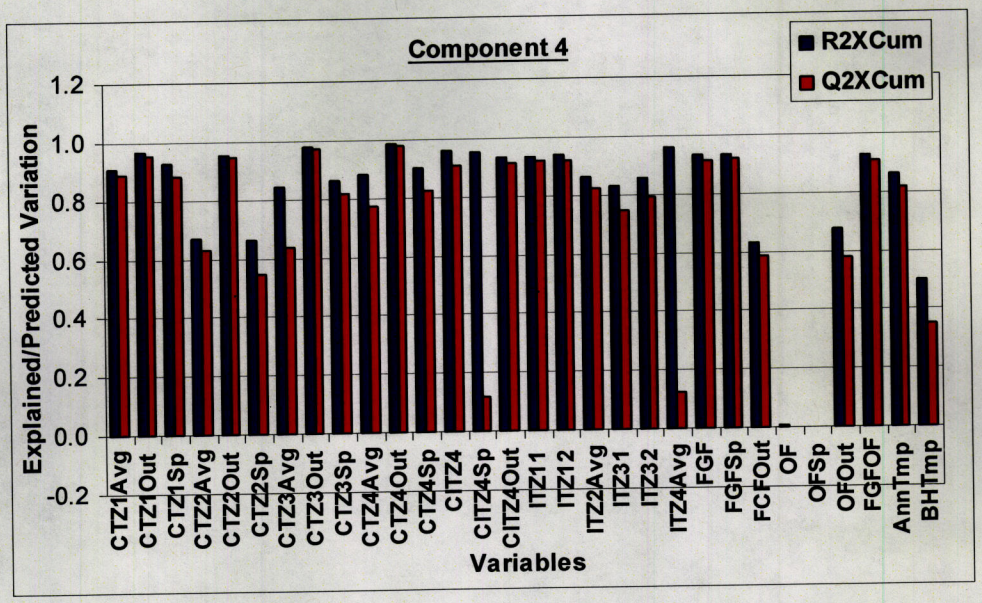

Figure 50: R^2VXCum and Q^2VXCum – Component 4

The score and loading plots were then examined and the plots for the first two components are given below.

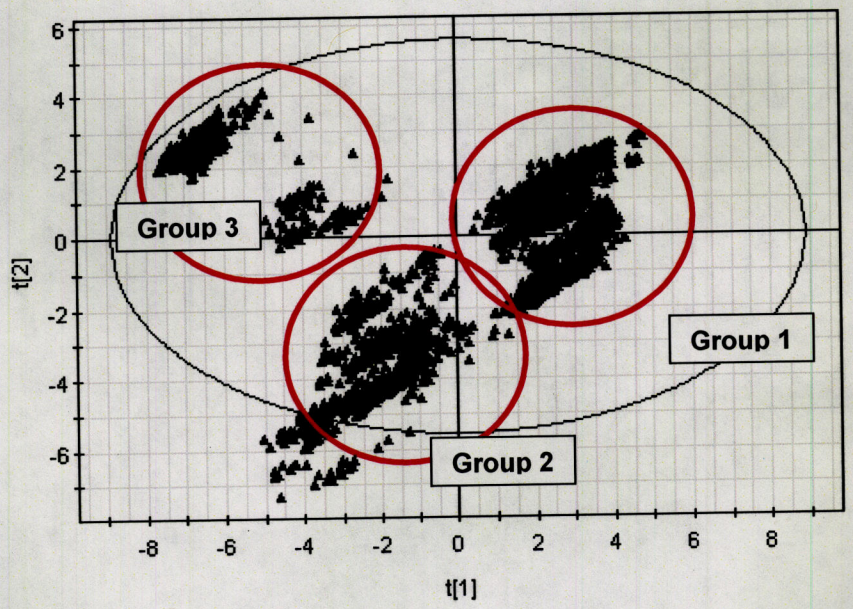

Figure 51: X-Score Plot – First Two Components

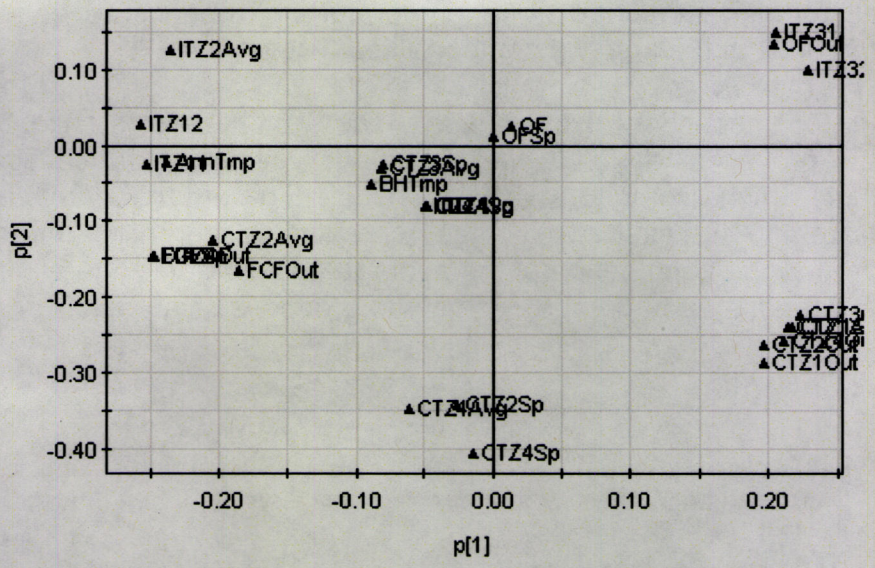

Figure 52: Loading Plot – First Two Components

As evident through examination of the explained and predicted variation by component plots all of the variable outputs display significant loadings in the p_1 direction. However it is only through examination of the loading plot that the correlation between the output variables is observed. For example all of the control zone temperature variable outputs cluster together in the lower right-hand quadrant of the plot while the internal zone four temperature and feed gas flow outputs are located diagonally opposite. The clustering of these two variables is expected as the internal zone four temperature controls the feed flow rate. The oxygen flow output is also not grouped with any of the other outputs on the lower right hand quadrant of the plot. During periods of low AD and FSSS nickel powder production there is expected to be little variation in the oxygen flow rate which accounts for the close proximity of the oxygen flow rate and set point variables to the origin of the loading plot. Although the first and second control zone average temperatures are significant in the first component they also are not clustered together and given their locations at opposite ends of the first component axis are negatively correlated with one another. This is also the case for the ITZ11, ITZ12, ITZ2Avg and AnnTmp variables which are also significant in the first component but which are negatively correlated with ITZ31 and ITZ32. The prevalence of the fourth control zone temperature with respect to the second component is also evident upon inspection of the loading plot. Similarly the close

proximity of the third control zone temperature to the origin reveals its lack of significance in the first two components.

Upon inspection of the t_1 vs. t_2 score plot the existence of three separate groups of observations is evident. In order to investigate the variables which contribute to the clustering of the variables the contribution to the first two scores for each of the three groups was compared to the contribution of the average observation. The contribution plot between the average and each of the groups is given below.

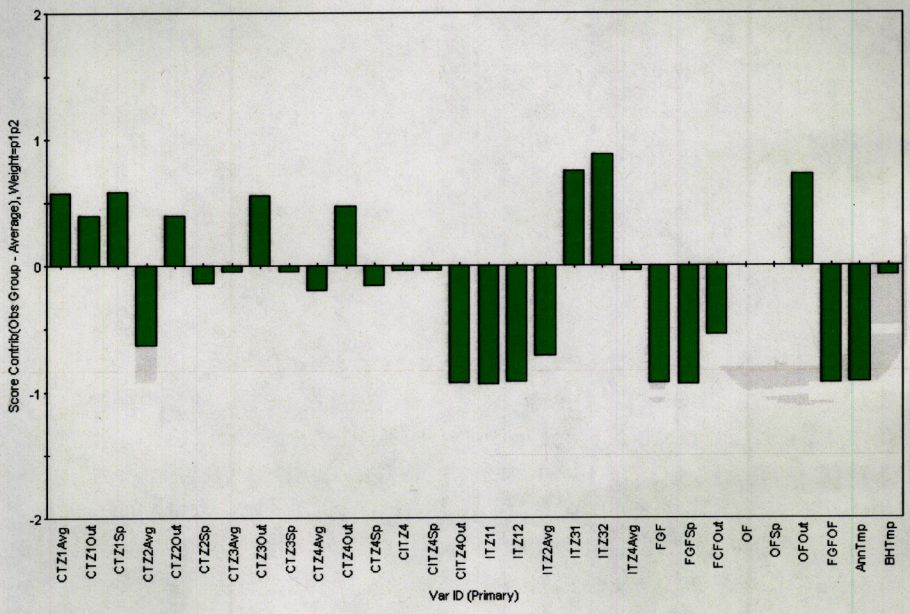

Figure 53: Contribution Plot Between Group 1 and Average Observation

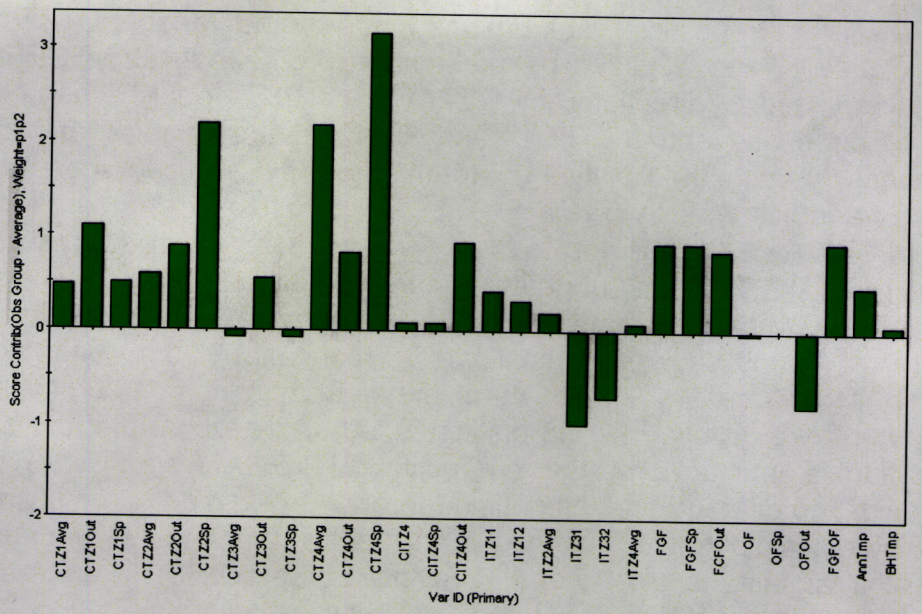

Figure 54: Contribution Plot Between Group 2 and Average Observation

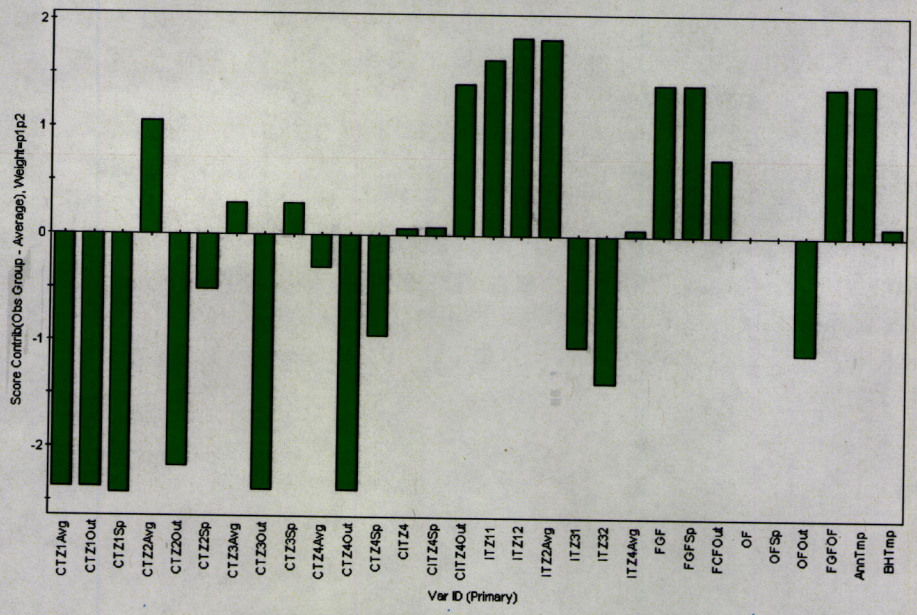

Figure 55: Contribution Plot Between Group 3 and Average Observation

Group 1 is oriented as such on the score plot due to the positive contribution of the variable CTZ1Avg in the positive p_1 direction and negative p_2 direction and to the positive contribution of ITZ31 and ITZ32 in the positive p_1 and p_2 directions. Similarly Group 2 obtains its orientation

as a result of the positive contributions made by the variable CTZ1Avg in the positive p_1 and negative p_2 directions, by CTZ2Avg and FGF in the negative p_1 and p_2 directions, by ITZ2Avg and BHTmp in the negative p_1 and p_2 directions and by ITZ11 and ITZ12 in the negative p_1 and p_2 directions. Finally the positive contributions to the first two scores made by the variables CTZ2Avg and FGF in the negative p_1 and p_2 directions and by ITZ2Avg and BHTmp in the negative p_1 and p_2 directions and by ITZ11 and ITZ12 in the negative p_1 and p_2 directions account for the location of Group 3 on the score plot. Therefore the movement between the three groups shown on the t_1 vs. t_2 score plot is not dominated by contributions from any single variable but from a number of variables. Although three groups may be distinguished on the score plot the overall operation is indicative of the production of low AD and FSSS nickel powder. The differing profiles outlined through evaluation of the score and contribution plots reflect changes made to the manipulated variables in order to maintain the low AD and FSSS powder targets.

The score and loading plots for the second and third components are given below. As shown on the loading plot the third component is dominated by the variables CITZ4, CITZ4Sp and ITZ4Avg as they are all located peripherally along the negative p_3 axis away from the other variables. Similar to the case for the first two components, clustering of the observations is present on the score plot. The contribution made by each of the five groups of observations shown below to the second and third score was compared to the contribution made by the average observation. Group 1 and 2 are dominated by positive contributions from the variables CITZ4, CITZ4Sp and ITZ4Avg while Groups 3, 4 and 5 are dominated by negative contributions from the variables CITZ4, CITZ4Sp and ITZ4Avg.

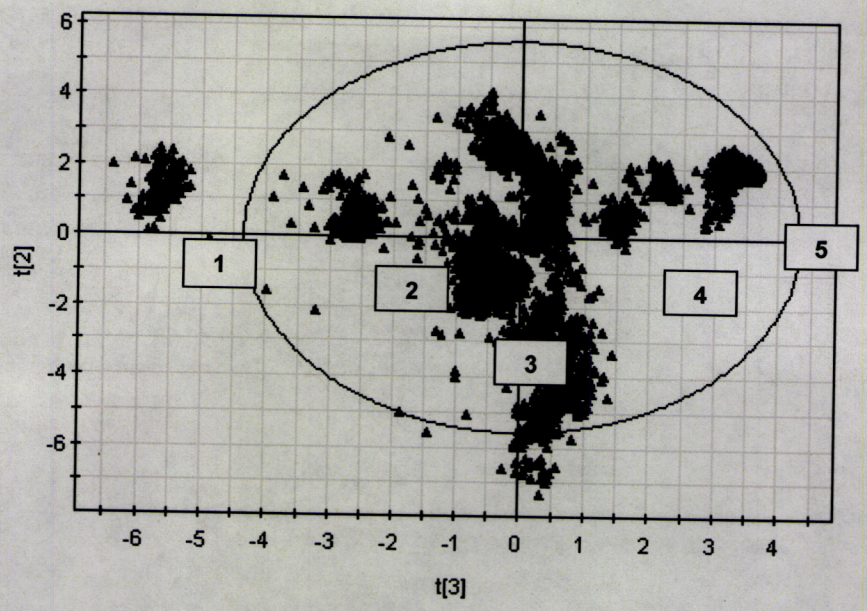

Figure 56: X-Score Plot – Second and Third Component

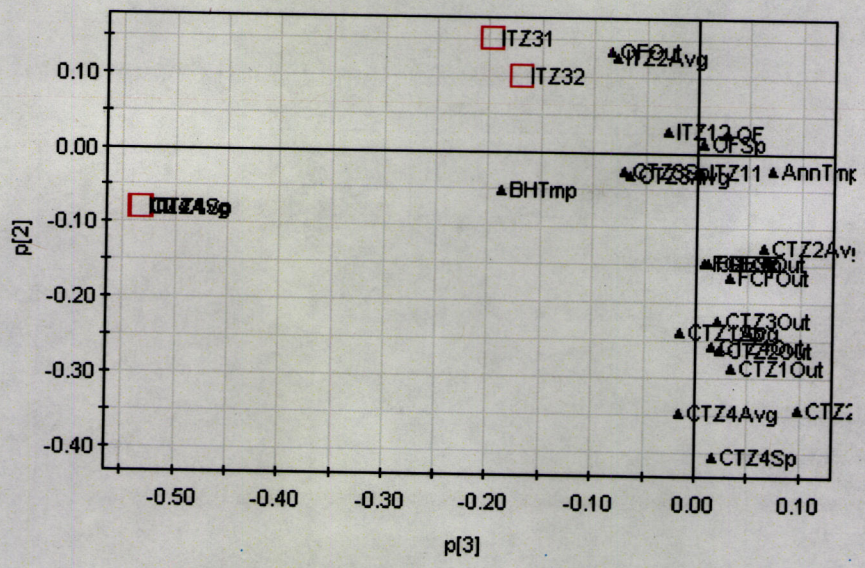

Figure 57: Loading Plot – Second and Third Component

Finally the score and loading plots for the third and fourth components are given below and as shown the fourth component is dominated by the variables CTZ3Avg and CTZ3Sp. Similar to the case for the second and third components the contribution plots were examined for

the three groups of observations shown below. Group 1 is characterized by positive contributions to the third and fourth scores by CTZ3Avg and CTZ3Ap while Groups 2 and 3 by negative ones.

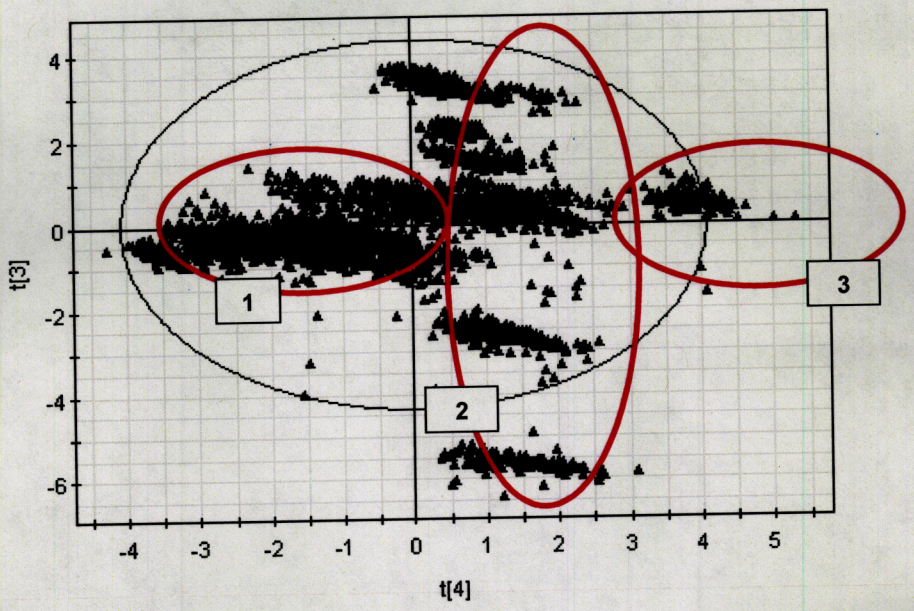

Figure 58: X-Score Plot – Third and Fourth Components

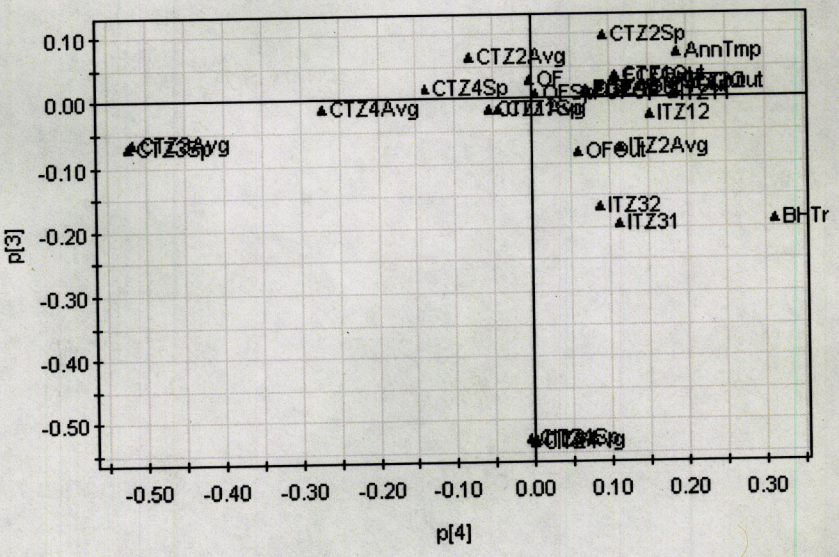

Figure 59: Loading Plot – Third and Fourth Components

5.3.4. New Data Inclusion – Low Group

A set of data was then collected during periods of low AD and FSSS nickel powder production. This data set consisted of instantaneous values obtained for the thirty process variables collected at ten-minute intervals. The data set therefore contained 25 491 observations and thirty X-variables.

The predicted scores for the first two components are given below and as shown the majority of the observations lie within the confidence ellipse.

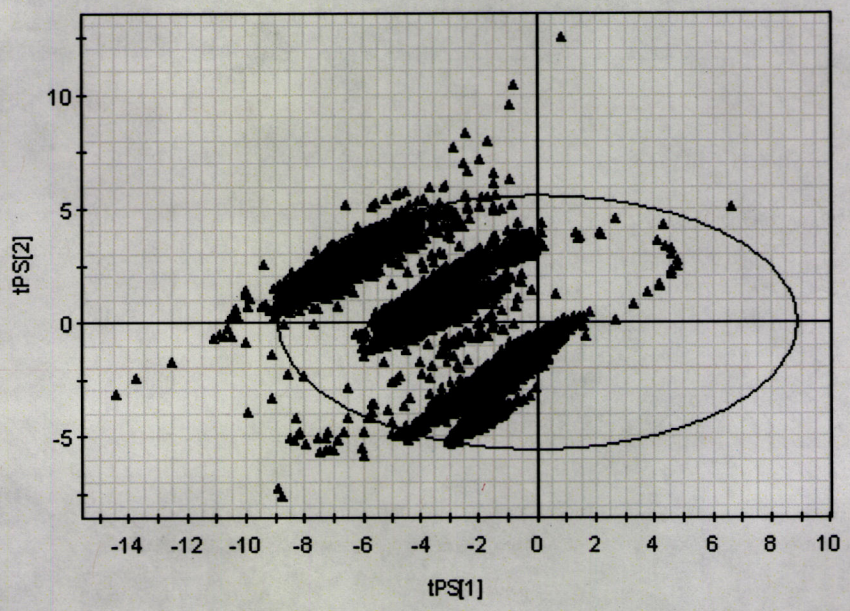

Figure 60: Predicted X-Score Plot - First Two Components

The time series plots for the first score and DMODX are also given below. The outlying observations lying below the general trend are located on the lower left-hand quadrant of the predicted t_1 vs. t_2 score plot. Those lying above the trend are located in the upper right hand quadrant of the score plot. Contribution plots were used to determine which variables contributed to the observed deviation with respect to the first two scores. In general positive contributions were made by CITZ4, FGFOF and AnnTmp for those observations with low first score values, while negative contributions were made by OF and OFOut. In contrast positive contributions were made by CTZ1Avg, CTZ2Avg, CTZ3Avg, CTZ4Avg

and AnnTmp to the observations with elevated first scores. Negative contributions were also made by CITZ4Out, FGFOut, OF and OFOut for these observations.

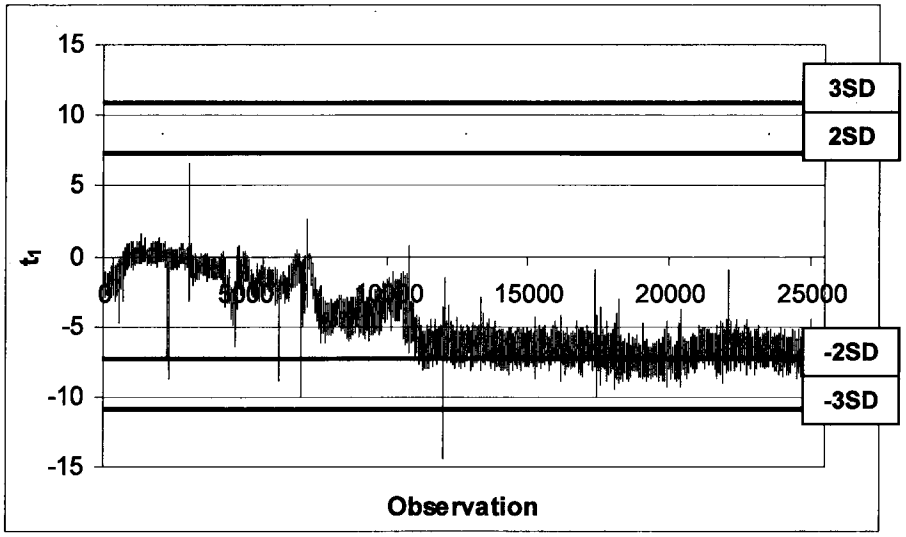

Figure 61: Score Time Series Plot – First Component

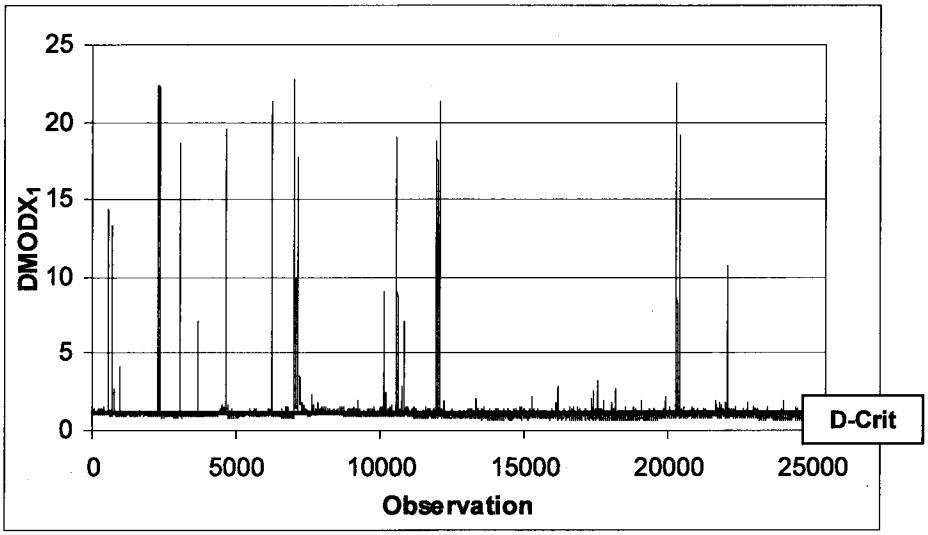

Figure 62: DMODX Time Series Plot – First Component

The time series and DMODX plots for the second component are also given below. Many of the observations identified as outliers on the first score time series plot were also identified as outliers on the second score time series plot. The outlying observations located below the trend of observations were found on the lower left-hand quadrant of the predicted t_1 vs. t_2 score plot while those located above the trend were found on the upper middle section of the score plot.

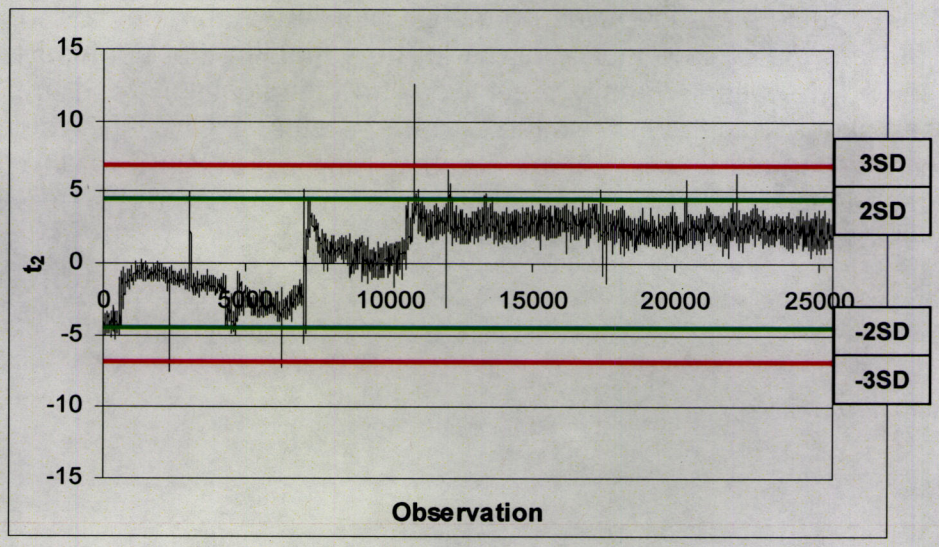

Figure 63: Score Time Series Plot – Second Component

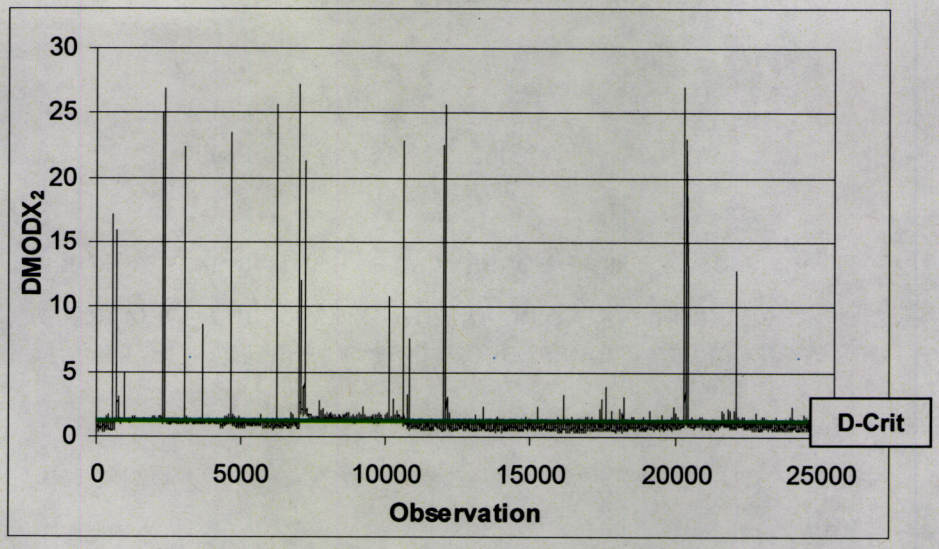

Figure 64: DMODX Time Series Plot – Second Component

A number of variables dominated the explained variation for the first two components as evidenced by the explained variation by variable and loading plots for the first two components. However with the input of new data into the model, several variables with low explained variation in the first two components contributed to the classification of observations as outliers with respect to the first two scores. Therefore the Hotelling's T^2 value for the first two components was plotted. This plot summarizes the observations over the first two components. As shown a number of extreme outliers are identified on the T_2^2 plot. Although they are identified as outliers there is no indication as to the variables that contribute to the observed deviation. Finally there were no consecutively long periods of observations which exceeded the D-crit, +/-3SD or T^2Crit(99%) values. As such the correlation structure for the new data input to the model corresponded well to the structure of the data used for model development.

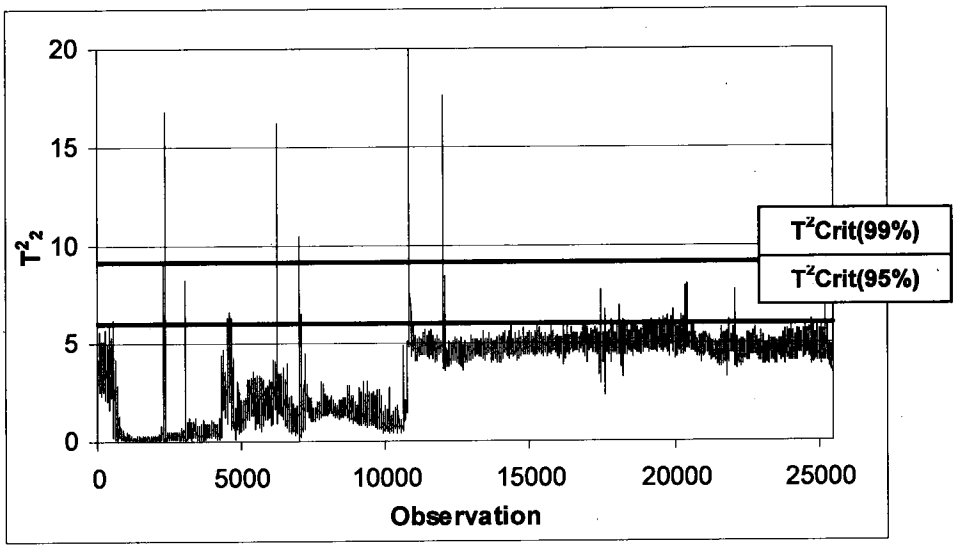

Figure 65: T^2 Time Series Plot – First Two Component

The predicted score plot for the second and third component is given below along with the third score time series plot.

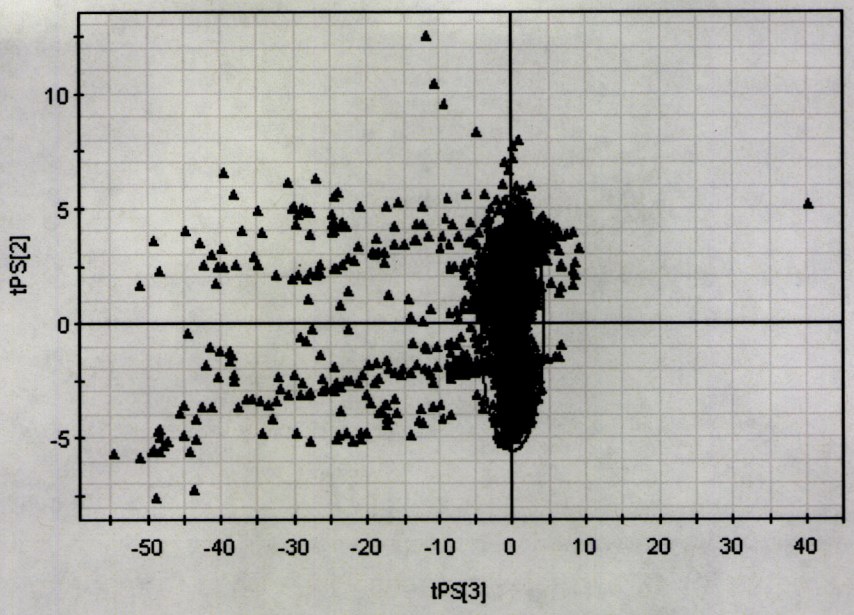

Figure 66: Predicted X-Score Plot – Second and Third Components

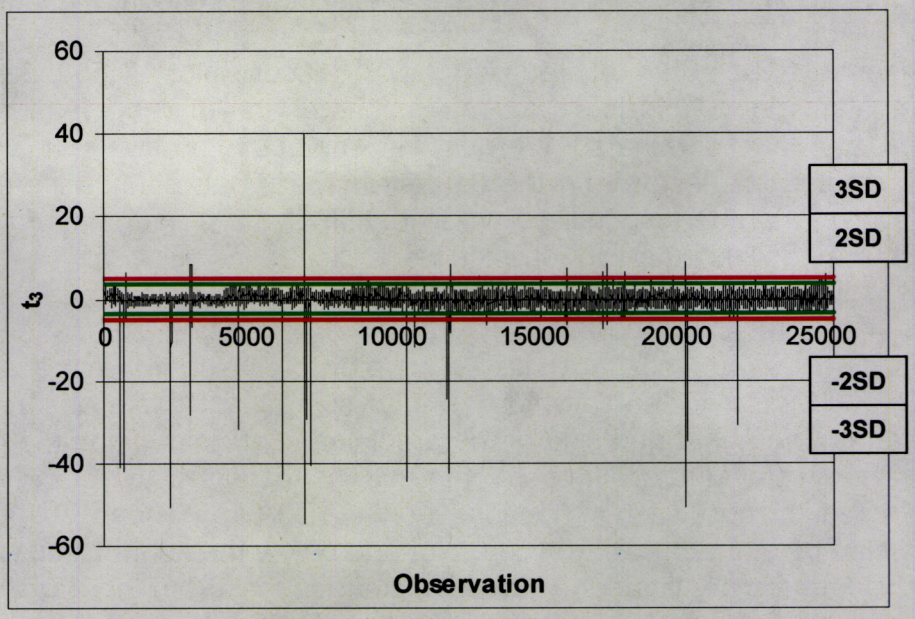

Figure 67: Score Time Series Plot – Third Component

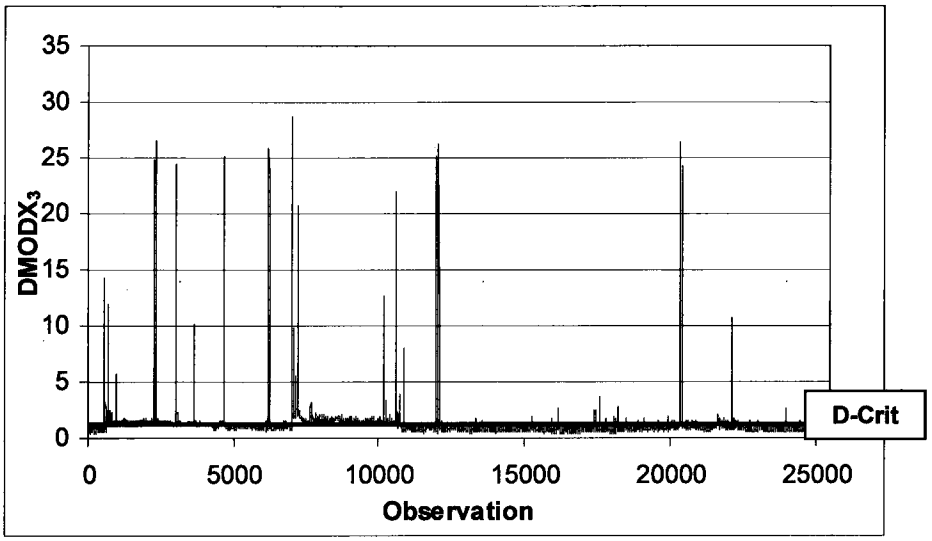

Figure 68: DMODX Time Series Plot – Third Component

According to the contribution plot for the second and third components the observation located above the control limits on the third component time series plot is dominated by negative contributions from CITZ4 and ITZ4Avg and is located on the right-hand side of the confidence ellipse for the predicted t_2 vs. t_3 score plot. Conversely the observations located below the control limits on the third component time series plot are dominated by positive contributions from CITZ4 and ITZ4Avg and are located on the left-hand side of the confidence ellipse. Although a number of outliers were identified there was no prolonged period during which the control limits were exceeded.

The predicted score plot for the third and fourth components are given below along with the time series plot for the fourth score and for DMODX. Similar to the case for the third component those observations located on the right-hand side of the predicted score plot along the p_4 axis and above the control limits on the fourth score time series plot are dominated by positive contributions from CTZ3Avg. Those on the left-hand side by of the predicted score plot and below the control limits for the fourth score are dominated by negative contributions from this variable.

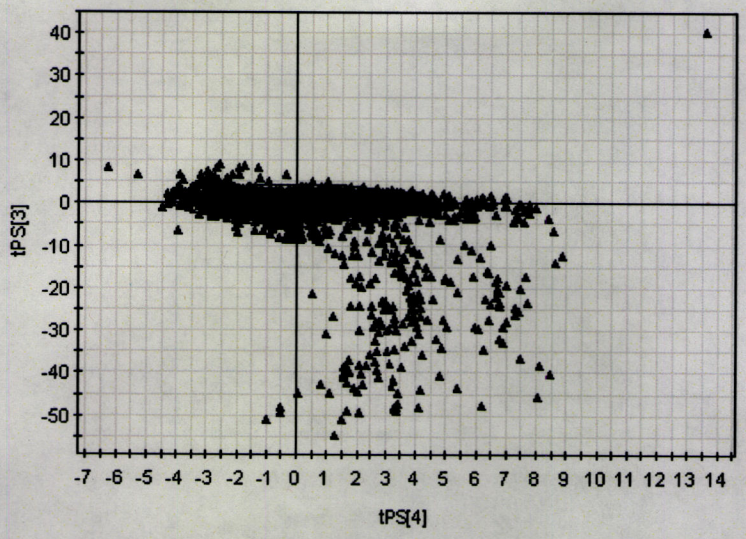

Figure 69: Predicted X-Score Plot – Third and Fourth Components

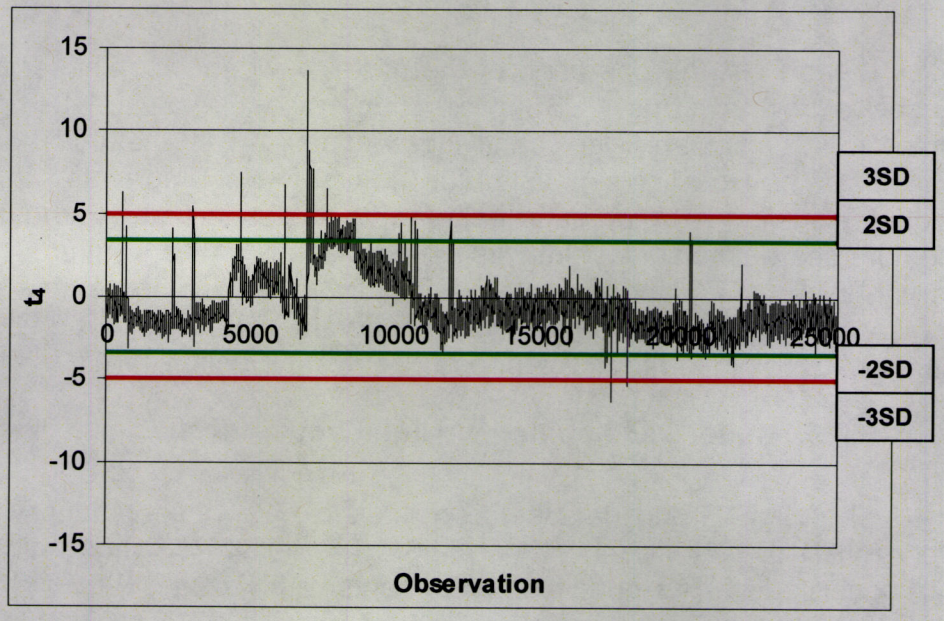

Figure 70: Score Time Series Plot – Fourth Component

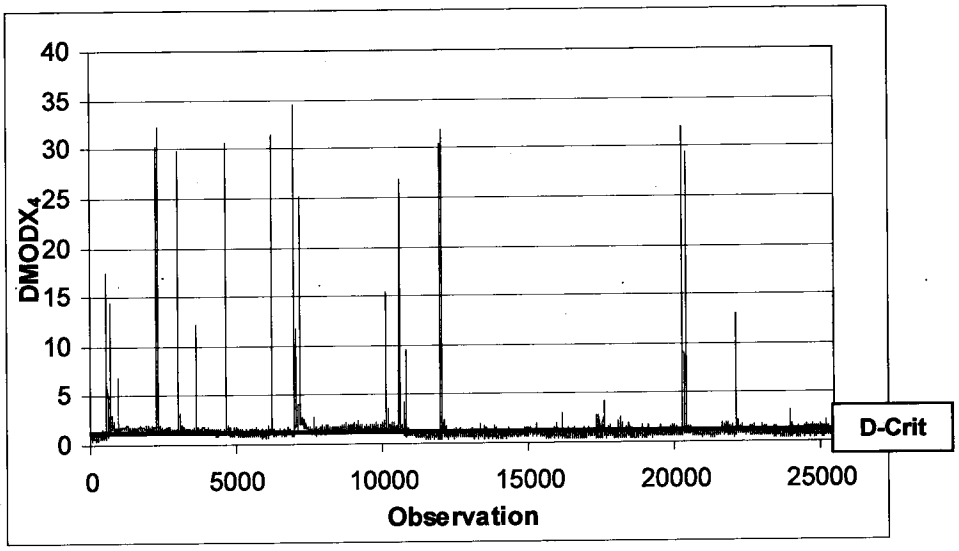

Figure 71: DMODX Time Series Plot – Fourth Component

Using PCA the operation of the nickel powder decomposer during periods of low AD and FSSS production was summarized using four parameters, which are significantly less compared to the thirty original variables. The Hotelling's T^2 value for the first two components was used to identify outliers over the entire range of variables included in the model. The third score time series plot was used to identify outliers with respect to the variables CITZ4 and ITZ4Avg. Similarly the fourth score time series plot was used to identify outliers with respect to CTZ3Avg. The DMODX value was also plotted for the purpose of outlier detection.

5.3.5. Model Development – High Group

The High Group data set consisted of 726 observations with an average AD of 0.74 and FSSS of 2.96. The initial PLS model fit to the data consisted of nine principle components with $R^2X = 0.94$, $R^2Y = 0.57$ and $Q^2 = 0.51$. The score plots for the first three components were examined and a number of outliers were identified. With these outliers removed from the dataset the PLS model contained nine principle components with $R^2X = 0.95$, $R^2Y = 0.58$ and $Q^2 = 0.55$. The two Y-variables were then removed from the data set and PCA was applied. The resulting PCA model consisted of twelve principle components with $R^2X = 0.99$ and $Q^2 = 0.91$. Upon the removal of some extreme outliers the resulting PCA model was comprised of twelve principle components with

R^2X = 0.99 and Q^2 = 0.93. The majority of the explained and predicted variation was accounted for by the first four model components. This model was used as the basis for outlining the good operation of the nickel powder decomposer during periods of high AD and FSSS powder production.

5.3.6. Model Interpretation – High Group

The explained and predicted variation by component was then examined in order to determine the relationships between the model components and variables. As shown below the first component accounts for the explained and predicted variation for all of the variables with the exception of CTZ3Avg, CTZ3Sp, CITZ4, CITZ4Sp, ITZ4Avg, OF, OFSp, FGFOF and BHTmp. With the addition of the second component the extent to which the variables CITZ4, CITZ4Sp, ITZ4Avg, OF, OFSp, FGFOF and BHTmp are accounted for increases. The third component increases the extent to which the overall variation is explained and predicted. With the addition of the fourth component the variables CTZ3Avg and CTZ3Sp explained and predicted variation increases.

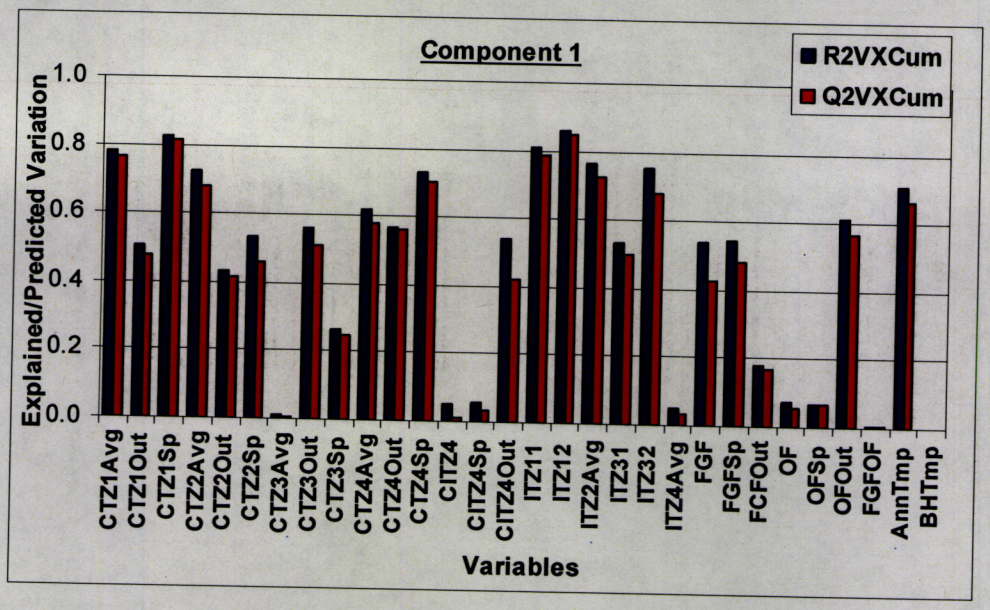

Figure 72: R^2VXCum and Q^2VXCum – Component 1

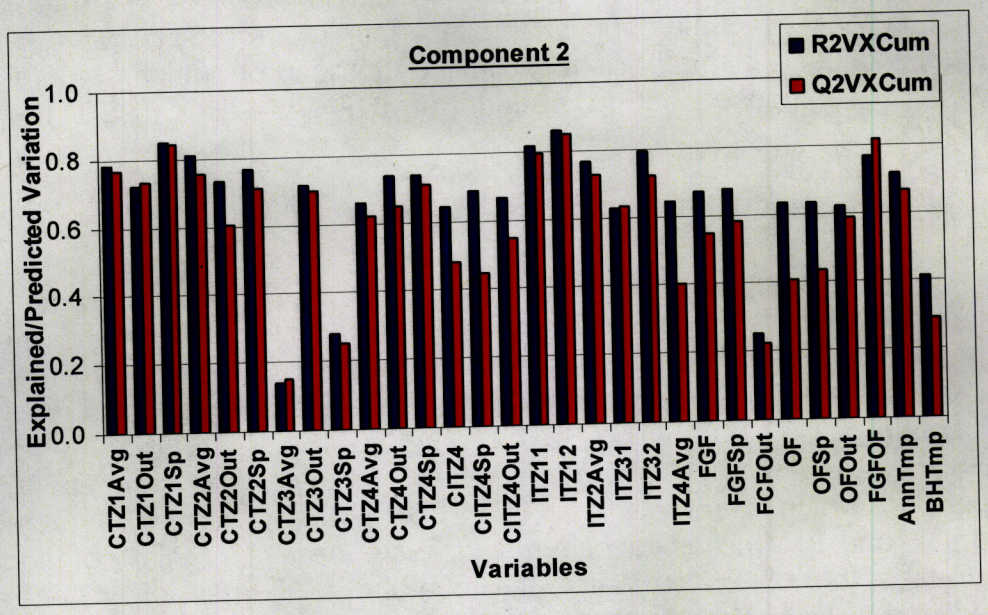

Figure 73: R^2VXCum and Q^2VXCum – Component 2

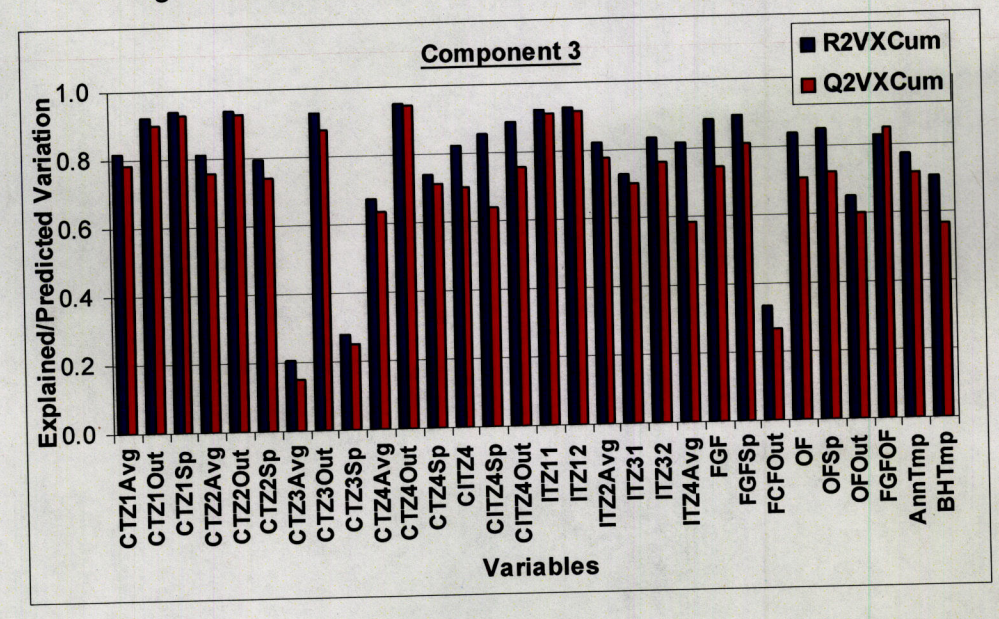

Figure 74: R^2VXCum and Q^2VXCum – Component 3

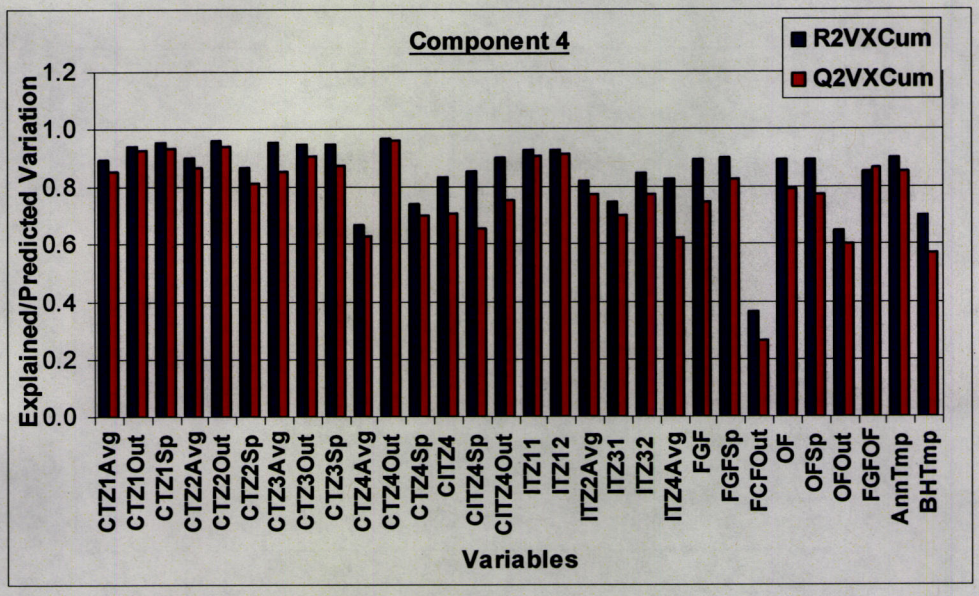

Figure 75: R^2VXCum and Q^2VXCum – Component 4

The score and loading plots for the first two components are given below. Similar to the case for the Low Group all of the output variables cluster together on the loading plot with the exception of FGFOut and CITZOut, which are located close to one another. The control zone 1 and 4 average and temperature set points also cluster together but zone 2 is located on the opposite side of the loading plot. Therefore zone 1 and 4 are positively correlated with one another while zone 2 is negatively correlated with 1 and 4. Given the close proximity to the origin, the zone 3 average temperature is not as significant to the first two components compared to the other control zone average temperatures. The internal temperatures also display this correlation with ITZ31 and ITZ32 located on the right hand side of the loading plot and ITZ11, ITZ21 and ITZ2Avg located on the left-hand side. Variation along the second component is attributed to variation with respect to the oxygen flow and set point, the feed gas to oxygen flow ratio, the internal zone 4 temperature and set point and ITZ4Avg.

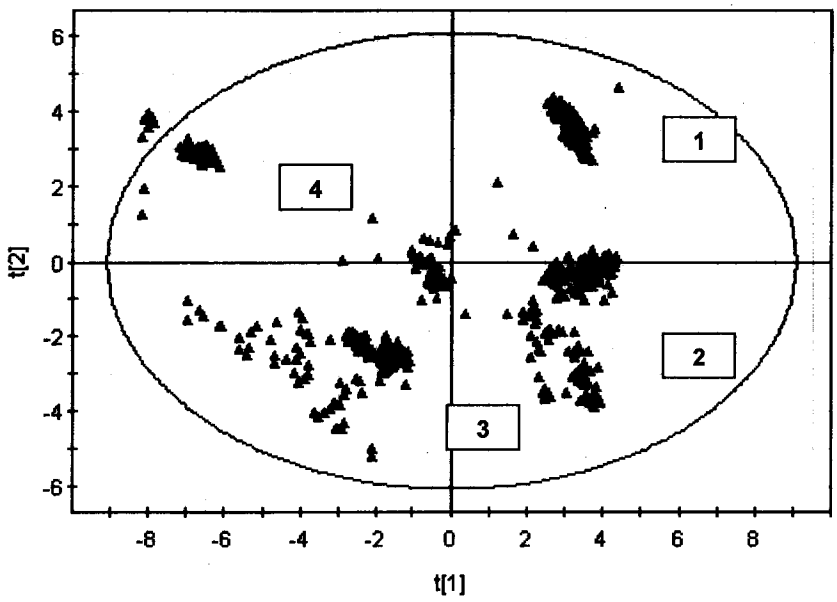

Figure 76: X-Score Plot – First Two Components

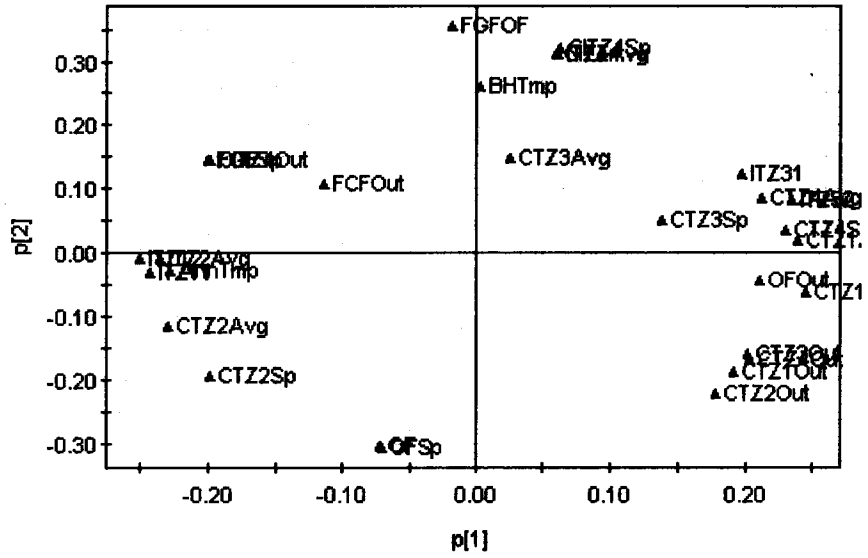

Figure 77: Loading Plot – First Two Components

The contribution plots for each of the four groups of observations outlined on the score plot were then examined and were compared to the contribution to the first two scores made by the average observation.

Group 1 is located in the upper right-hand quadrant as a result of positive contributions from CTZ1Avg, CTZ4Avg, CITZ4, CITZ4Sp, ITZ31 and ITZ32 in the positive p_1 and p_2 directions and from FGF in the FGF in the negative p_1 and positive p_2 directions. Group 2 is given its location as a result of positive contributions from CTZ1Avg, CTZ4Avg, ITZ31 and ITZ32 in the positive p_1 and p_2 directions. Similarly Group 3 obtains its location due to positive contributions from CTZ2, ITZ11, ITZ12, ITZ2Avg and AnnTmp in the negative p_1 and p_2 directions, from FGF in the negative p_1 direction and positive p_2 directions and OF and OFSp in the negative p_1 and p_2 directions. Finally Group 4 is located as such due to positive contributions from CTZ2Avg in the negative p_1 and p_2 directions, CITZ4 and CITZ4Sp in the positive p_1 and p_2 directions, ITZ11, ITZ12, ITZ2Avg and AnnTmp in the negative p_1 and p_2 directions and by FGF in the negative p_1 and positive p_2 directions. Therefore movement with respect to the process in the first and second component is dominated by many of the model variables. As was the case for the Low Group, the clustering observed on the score plot for the first two components along with the resulting variable contributions is indicative of changes made to the manipulated variables in order to produce high AD and FSSS nickel powder which meets the final product physical property targets.

Finally the score and loading plots for the first and fourth components are given below and as shown the fourth component is dominated by the variables CTZ3Avg and CTZ3Sp. The contributions made to the first and fourth scores by each of the three groups of observations outlined below were compared to the contribution made by the average contribution. Group 1 is characterized by large positive contributions from CTZ3Avg and CTZ3Sp while Groups 2 and 3 are characterized by large negative contributions from these variables.

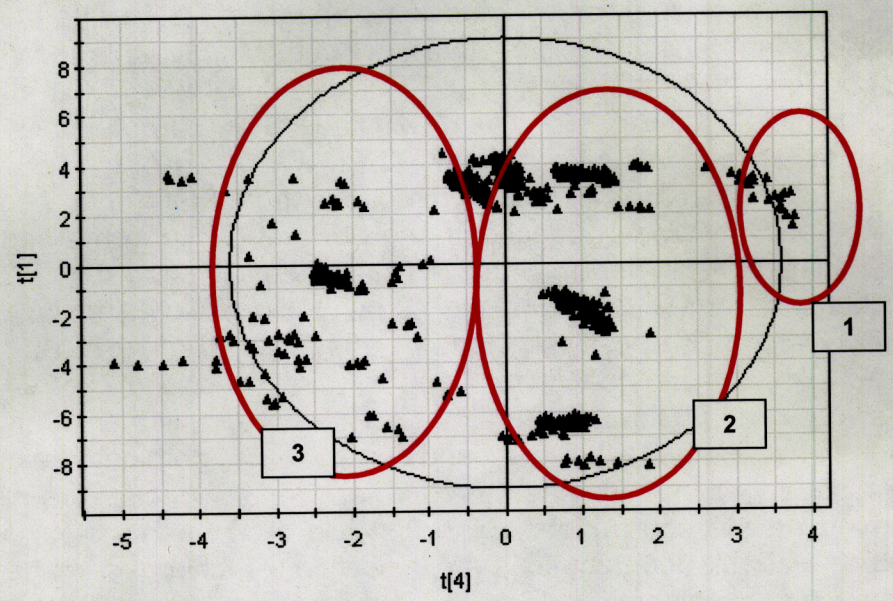

Figure 78: X-Score Plot – First and Fourth Components

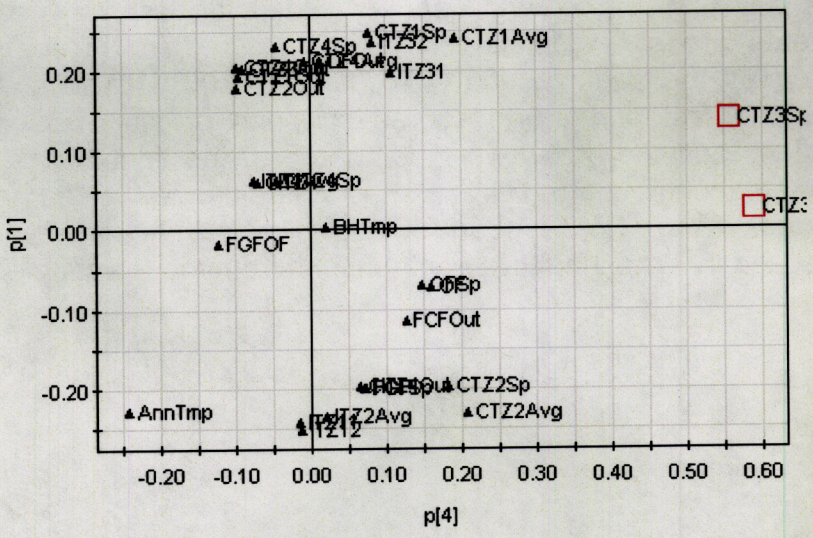

Figure 79: Loading Plot – First and Fourth Components

5.3.7. New Data Inclusion – High Group

A second set of instantaneous data was collected during periods of high AD and FSSS production. This data set contained 2314 observations and thirty X-variables collected at ten minute intervals. The predicted score plot for the first two components is given below along with the time series and DMODX plots for the first score.

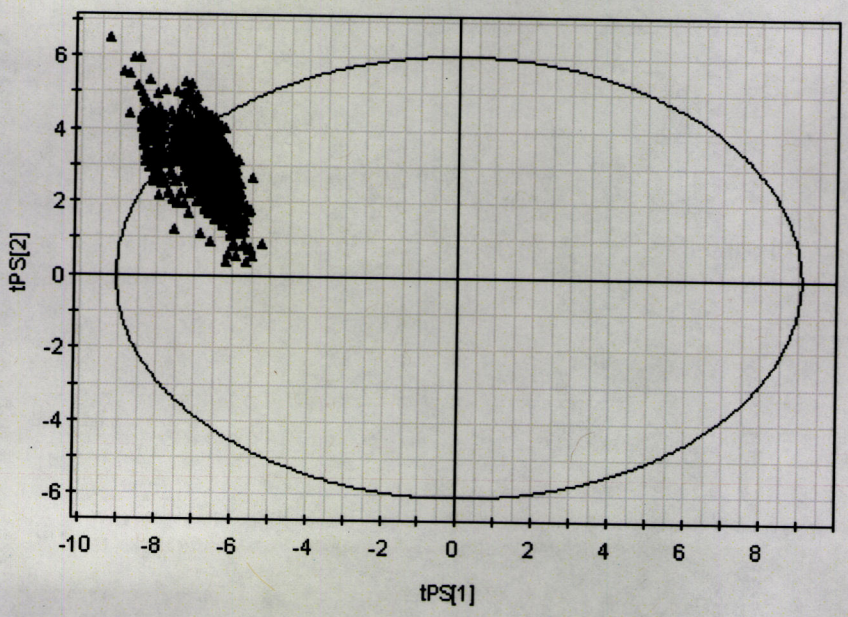

Figure 80: Predicted X-Score Plot – First Two Components

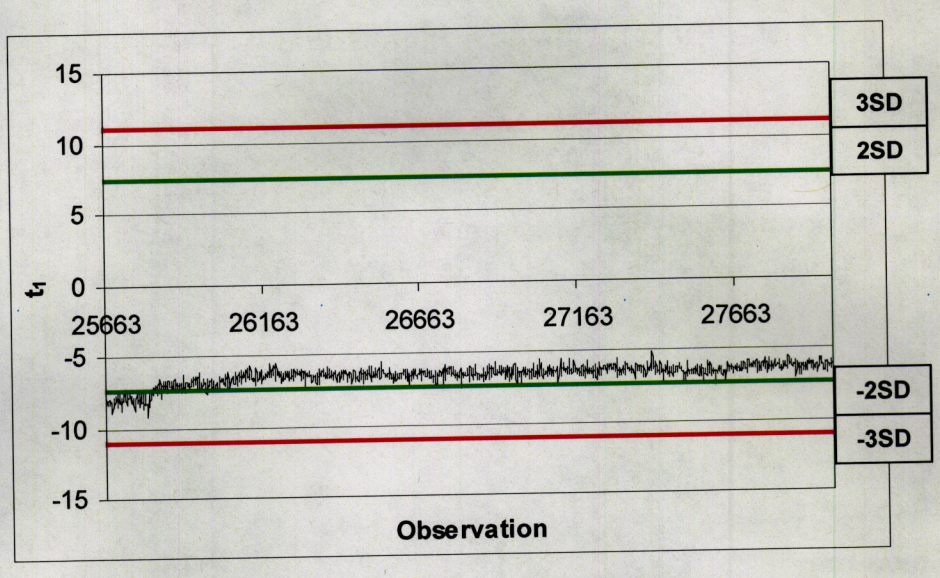

Figure 81: Score Time Series Plot – First Component

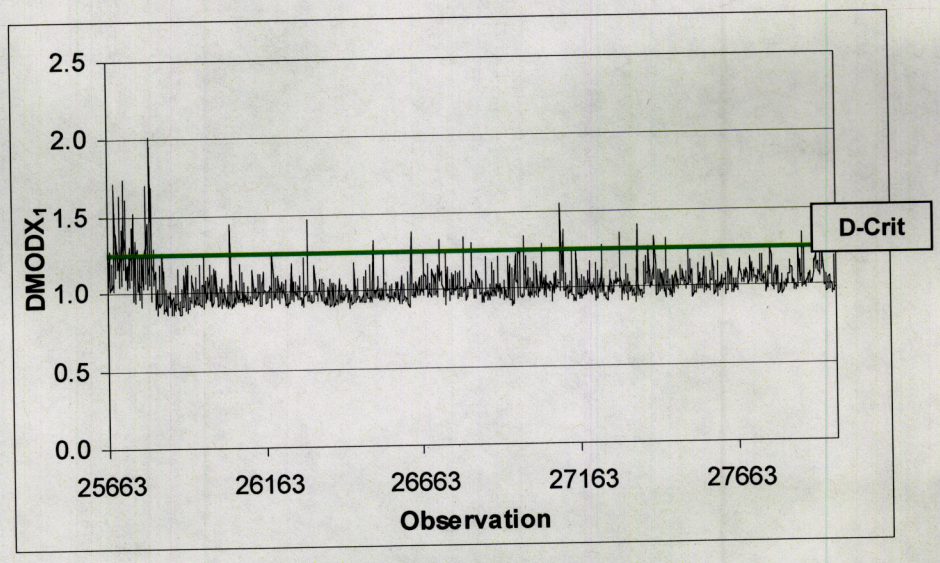

Figure 82: DMODX Time Series Plot – First Component

The score contribution plots for the first two components were examined and the observations which are outliers with respect to the first score time series plot are dominated by positive contributions from CTZ2Avg, CITZ4Out, ITZ11, ITZ12, ITZ2Avg, FGF, FGFSp, FGFOut and

AnnTmp and negative contributions from CTZ1Avg, CTZ4Avg, ITZ31, ITZ32 and OFOut.

The time series and DMODX plots for the second score is also given below. According to the contribution plots, the observations that are outliers in the t_2 direction are dominated by positive contributions from CITZ4, CITZ4Out, ITZ4Avg, FGF, FGFSp, FGFOut, FGFOF and BHTmp. Negative contributions are made by CTZ1Out, CTZ2Out, CTZ3Out, CTZ4Out, ITZ31, ITZ32, OF, and OFSp.

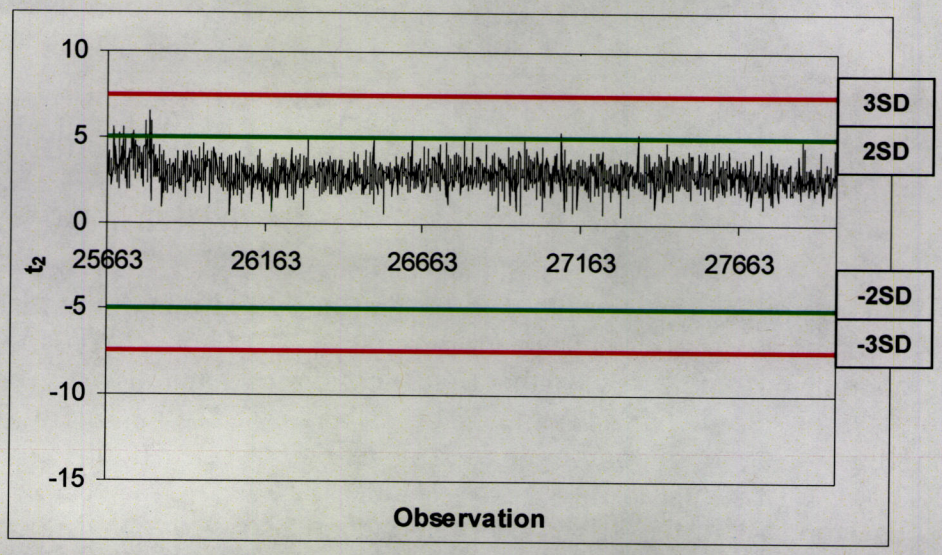

Figure 83: Score Time Series Plot – Second Component

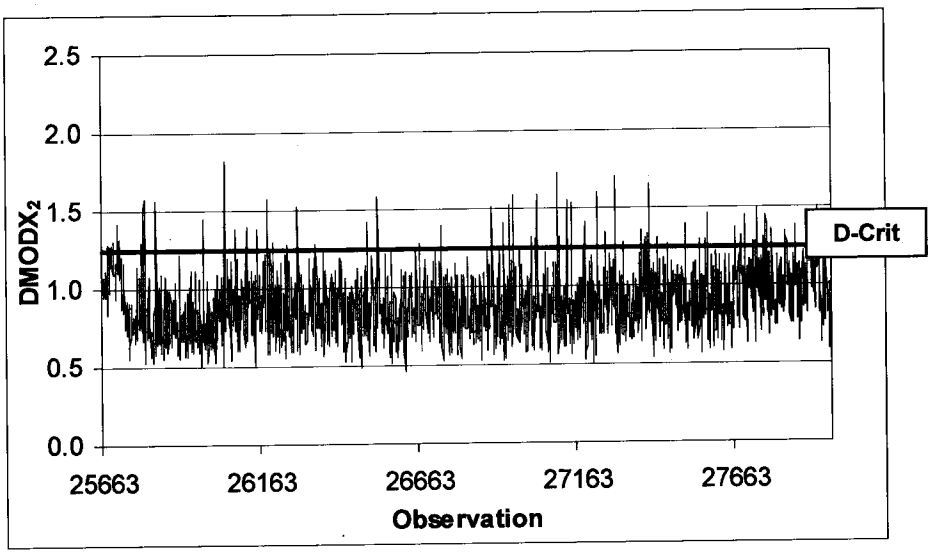

Figure 84: DMODX Time Series Plot – Second Component

The predicted score plot for the first and third component is given below along with the time series and DMODX plots for the third component. The outliers in the third score are dominated by positive contributions from CITZ4Out, ITZ11, ITZ12, ITZ2Avg, FGFOF and AnnTmp. Negative contributions are made by CTZ1Out, CTZ2Out, CTZ3Out, CTZ4Out, CITZ4, ITZ31, ITZ32, ITZ4Avg, OF, OFSp, OFOut and BHTmp. All of the observations identified as outliers were located in the upper left hand quadrant of the score plot beyond the confidence ellipse. Finally no outliers were observed in the fourth score time series plot.

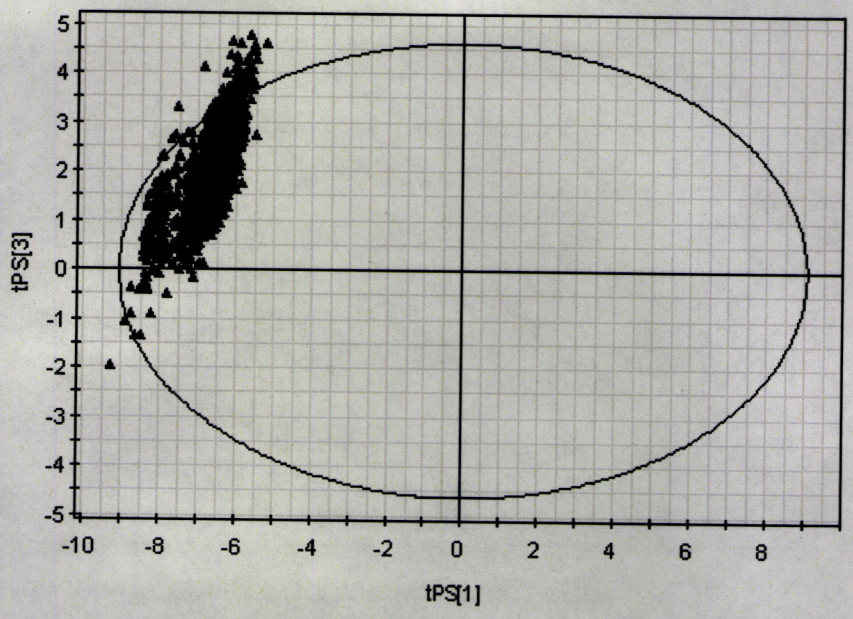

Figure 85: Predicted Score Plot – First and Third Components

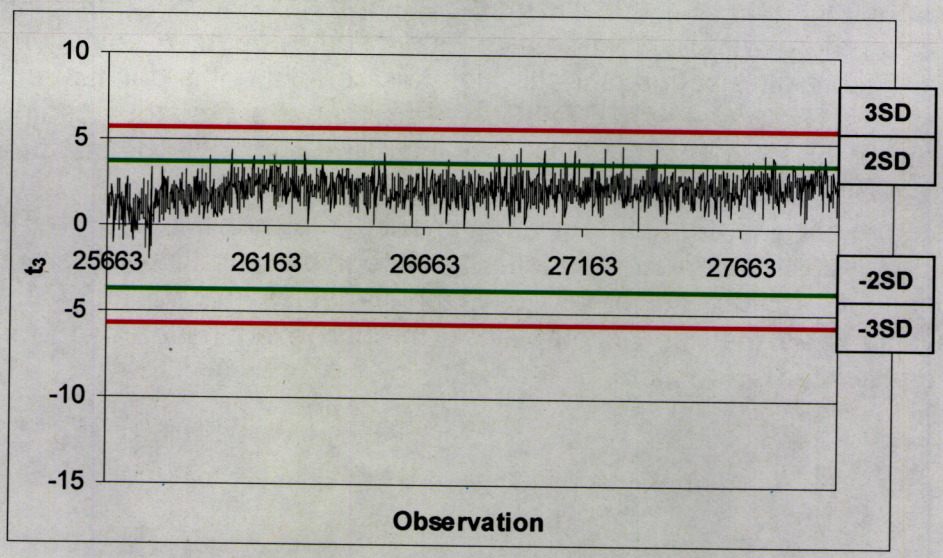

Figure 86: Score Time Series Plot – Third Component

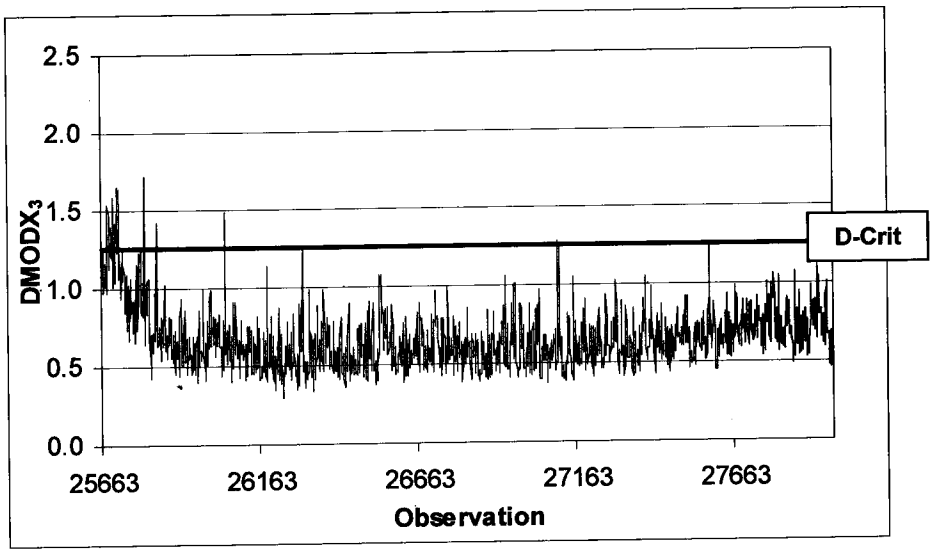

Figure 87: DMODX Time Series Plot – Third Component

Although the variables that contribute most to the explained and predicted variation for the PCA model were identified by component, the outlying new observations input to the model were not dominated by contributions from any specific variables for the first three components. In addition there were not any outliers identified through examination of the fourth score time series plot although this component is dominated by the variables CTZ3Avg and CTZ3Sp. Together the first four components account for 85% of the explained variation and 78% of the predicted variation. Therefore an alternative to monitoring the score time series plots for each of the first four components is to monitor the Hotelling's T^2 value for the first four components. Although the chart identifies outliers it does not provide insight as to which variables may contribute to the deviations as may be accomplished through examination of the loading and predicted score plots.

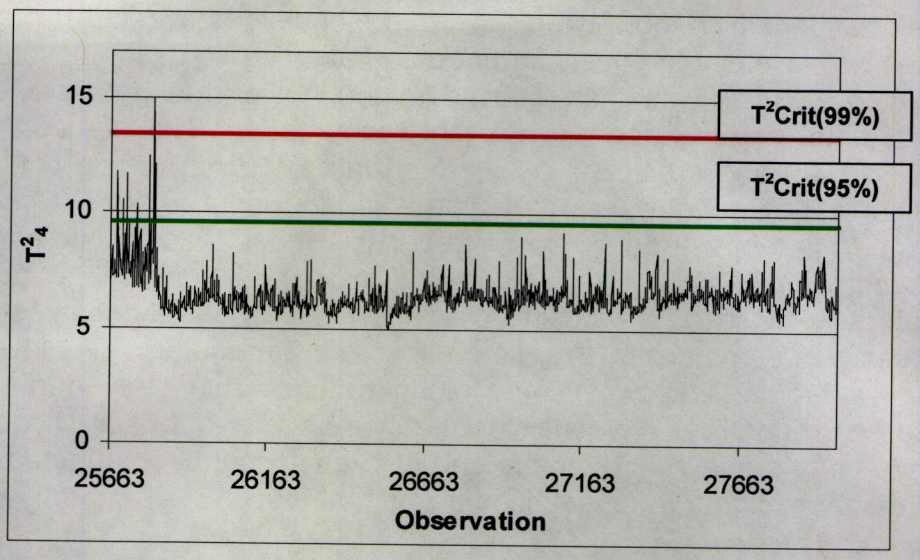

Figure 88: T_4^2 Time Series Plot

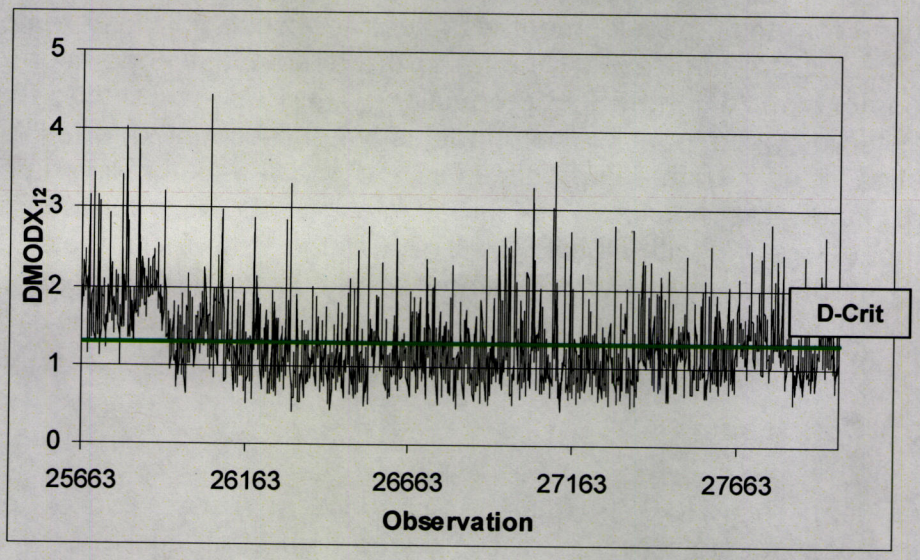

Figure 89: DMODX Time Series Plot – Twelfth Component

5.3.8. Conclusion

Online monitoring schemes were developed for two modes of nickel powder decomposer operation. The first mode of operation corresponded to the production of low AD and FSSS powder. A PLS model previously

developed for which there were no extreme outliers between the X and Y-scores was used as the basis for good operation of the decomposer under this mode. PCA was then applied to the outlier-free data set, which resulted in a six component model that accounted for 89% of the variation in the X-data and predicted 74%. New data collected during periods of low AD and FSSS powder production was then input to the PCA model. The operation of the decomposer under this mode was monitored using the control limits established by the PCA model. Although thirty process variables were used to outline the operation of the decomposer the process was summarized over time using four parameters: the Hotelling's T^2 value for the first two scores along with the third and fourth score values and DMODX. No consistent periods of outlying observations were found in any of the control charts and the variables that contributed to the out-of-control signal were identified using contribution plots.

Similarly a second PCA model was developed in order to monitor the decomposer during periods of high AD and FSSS production. The data set used for model development was again obtained following removal of extreme outliers from a PLS model of the decomposer operation during periods of high AD and FSSS production. The PCA model fit to the data consisted of twelve principle components, which accounted for 99% of the variation in the X-data and predicted 93% of the variation. New data input to the model was monitored using the Hotelling's T^2 value for the first four components along with DMODX. Similar to the case for the low AD and FSSS model no periods of sustained operation above the control limits was observed and outlying observations were examined using the variable contribution plot.

5.4. Conclusion

As part of INCO's Pressure Carbonyl Process nickel powder decomposer reactors are used to produce high purity nickel powder. In order to ensure that the powder meets the required bulk density (AD) and particle size (FSSS) product specifications the appropriate conditions within the decomposer must be set and maintained. As a means of investigating the decomposer process variables which contribute to the variation with respect to AD and FSSS, PLS analysis was carried out. Through this analysis operation of the decomposer under two specific modes was identified, corresponding to production of powder with low AD and FSSS and with high AD and FSSS. In addition the variables that contribute significantly to the variation between these two modes of

operation were identified. Strong correlations were observed between the six manipulated decomposer process variables and the Y-variables. In addition the correlation between the manipulated variables was also observed. For example the control zone 2 temperature was segregated from the other three control zones on the weight plot which reflects the decomposer operation as this is the zone in which decomposition primarily takes place. In addition the correlation between the internal zone 4 temperature and the feed gas flow rate was observed on the weight plot which is reflected in plant operation as the flow rate is controlled by the temperature. Finally the correlation between the Y-variables was observed on the weight plot which indicates that independent variation of AD and FSSS may be difficult.

Following the identification of the two modes of operation, PLS models were fit to the groups of observations corresponding to the production of low AD and FSSS nickel powder and to the production of high AD and FSSS nickel powder. Once the extreme outliers were removed from the PLS models, PCA was applied in order to model the good operation of the decomposer under the two modes of operation. The variables for which a high degree of variation was accounted for by the models were then determined by component. New data was then collected during periods of low and high AD and FSSS production and was input to the models. The statistics which best summarized the operation under the two modes and which provided some insight as to the variables contributing to the observations which exceeded the control limits were selected for monitoring. Contribution plots were also used for variable identification with respect to outlying observations. No periods of sustained outlying observations were noted in the control charts indicating that the decomposer was operated in the generally the same manner during the period for which the data used to construct the PCA model was collected and the period over which the new data input to the model was collected.

Finally nickel powder withdrawn and sampled from each of two decomposers at four-hour intervals is stored in large bins prior to packaging and shipment to market. Powder from the storage bins is blended in order to meet the bulk density and particle size distribution specifications of the final product. The physical characteristics of the final product are closely monitored using statistical quality control charts. Upper and lower control limits are defined against which the final product bulk density and particle size distribution are compared. In order to more clearly define the good operation of the nickel powder decomposer similar

control limits may be placed on the powder that is extracted from the decomposer at four-hour intervals and is sampled. Average process data collected during in-control powder production may then be used as a basis for PCA model development upon which an online monitoring scheme may be based.

6. Chapter 6 – Summary and Conclusion

Multivariate data analysis techniques were applied to the large quantities of data generated in relation to the smelting and refining processes utilized at INCO's Copper Cliff Operations and to the quality of the products produced. These techniques were used for three purposes: process evaluation, prediction and monitoring. In general the multivariate nature of the relationships between the process variables and between the process and product quality variables were modeled well.

The first analysis involved the use of PLS for oxidative pressure leach time prediction. Variables related to both the batch make-up and autoclave leaching operations were included. The model explained seventy-six percent of the variation in the leach time data. In general the observations corresponding to the most elevated leach times were modeled least well. This may be attributed to the exclusion of variables related to the solids charged to the autoclave for leaching from the model along with instrumentation sensor and solids assay error. However several of the variables included in the model were found to be significant with respect to the variation in leaching time. As such further investigation into the effect of variables related to the amount of solids, water and spent electrolyte added to the autoclave on leaching time may be warranted along with the relationship between oxygen flow into the autoclave and gasses bled from it. Finally the practice of using a sonic type mix tank level indicator for batch make-up end point determination should be reconsidered.

During the second analysis a PLS model was again used to develop a soft sensor for the prediction of the NO concentration of the gas treated by the smelter acid plant. Fifty-five variables related to the gas generating pyrometallurgical and acid plant processes were included in the analysis. Following refinement of the predictive model to include the lag time inherent in the operating conditions, new data was input to the model. The NO concentration was then predicted using only the process variables. Seventy-two percent of the variation in the actual NO concentration was explained using the process data only. In addition the control charts developed were used to outline periods of potentially reduced model predictive ability. Finally the PLS model developed was examined and several pyrometallurgical process variables were identified as significant contributors to the variation in the NO concentration of the gas treated by the acid plant. Two of these were related to furnace

maintenance work and as such may be further investigated as a means of reducing the potential for NO formation within the furnace.

The final analysis involved the use of both PLS and PCA. Initially a data set was constructed which contained variables related to the decomposer reactor process and to the quality of the nickel powder produced. PLS was used to predict two quality variables based on the process conditions. The significance of the six manipulated decomposer process variables was displayed along with the correlation between the process variables and between the quality variables. In addition two distinct modes of decomposer operation were identified. Following the construction of data sets that summarized the good operation of both modes, PCA was used to develop an online monitoring scheme. Although thirty process variables were included in the model, the ability to monitor the decomposer based on a significantly reduced number of latent variables was demonstrated. Finally contribution plots were used to identify those variables that contributed to the classification of observations as outliers according to the multivariate control limits developed.

Finally it is believed that two factors contributed to the successful formation and evaluation of the multivariate models that in general accounted for a large degree of the variation in present in both the process and product quality data. First the data sets constructed upon which the models were based summarized both the X and Y-data well. Knowledge of the processes studied was used to select the most relevant process variables for inclusion in the models. In addition data was collected over significantly long periods in order to capture the full range of process operating conditions. Second in all three analyses the Y-data used was characterized by a large number of observations that closely approximated the average and significantly fewer observations that displayed extreme deviations from the average. In all instances evaluation of the extremely deviant observations was of most interest. As such contribution plots were effectively used to compare the small number of extreme observations to the average observation and to identify the variables that contributed to the scores in these instances.

References

Eriksson L., Johansson E., Kettaneh-Wold N. and Wold S. Multi-and Megavariate Data Analysis Principles and Applications. Umeå, Sweden: Umetrics AB, 1999.

Geladi P, and Kowalski B. "Partial Least Squares: A Tutorial." Analytica Chimica Acta 185 (1986): 1-17.

Hoskuldsson A. "PLS Regression Methods." Journal of Chemometrics 2 (1988): 211-228.

Kourti T. and MacGregor J.F. "Multivariate SPC Methods for Process and Product Monitoring." Journal of Quality Technology 28.4 (1996): 409-427.

Kourti T. "Application of Latent Variable Methods to Process Control and Multivariate Statistical Process Control in Industry." International Journal of Adaptive Control and Signal Processing 19 (2005): 213 – 246.

MacGregor J.F., Jaeckle C., Kiparissides C. and Koutoudi M. "Process Monitoring and Diagnostics by Multiblock PLS Methods." AIChE J 40.5 (1994): 826-838.

Manne R. "Analysis of Two Partial Least Squares Algorithms for Multivariate Calibration." Chemometrics and Intelligent Laboratory Systems 2.1-3 (1987): 187-197.

Simca-P+ User Guide. Umeå, Sweden : Umetrics AB; 2002

Wold S. "Cross-Validatory Estimation of the Number of Components in Factor and Principal Component Models." Technometrics 20 (1978): 397-405.

Wold S. "Principle Component Analysis." Chemometrics and Intelligent Laboratory Systems 2 (1987): 37-52.