ON-LINE INDUSTRIAL IMPLEMENTATION OF PROCESS MONITORING / CONTROL APPLICATIONS USING MULTIVARIATE STATISTICAL TECHNOLOGIES: CHALLENGES AND OPPORTUNITIES

Michael Dudzic*, Yale Zhang

Process Automation, Dofasco Inc.
Box 2460, Hamilton, Ontario, Canada, L8N 3J5

Abstract: The global steel industry is striving to improve product quality through excellence in operation. To support this, significant investments have been made in upgrading instrumentation, data acquisition and computing infrastructures. The expectation is that with more process and product data readily available, useful information and better process knowledge can be gained in a timely fashion. The problem that has developed is that with the large volumes of data available, the associated data analysis and modeling have become increasingly complex. As a result, much of the data is either not used or summarized / heavily compressed. This means that a significant amount of the information and knowledge resident in the data is lost, diminishing the returns from the investment made in the information technology infrastructure. A class of technologies that Dofasco has used to meet this data challenge is multivariate statistics (MVS), with a primary focus on Principal Components Analysis (PCA) and Projection to Latent Structures (PLS). These methods have been successfully applied to analyze data for a variety of purposes, which includes the development of on-line predictive models and process monitoring systems. Since 1993, Dofasco has been involved with over 70 off-line / on-line applications of this technology at our steel facility in Hamilton, Ontario, Canada. Through these applications, significant financial returns to the company have been generated.

Keywords: multivariate statistics; model-based control; process monitoring; continuous caster; desulphurization

1. INTRODUCTION

As the ability to collect industrial data increases and improves, the key challenge is to exploit the useful information so as to maximize process understanding and knowledge. The concern then becomes the appropriate selection of data analysis techniques that can be used with large data sets in order to provide production engineers and operators with useful and easy-to-understand results.

From an automation engineer's perspective, improved data analysis and statistical modeling can be used to: (1) maximize information and knowledge generation to aid in both business and process related decisions, (2) develop predictive models for both inference and control, (3) facilitate useful statistical process control (SPC) schemes, (4) provide the ability to supplement first principle modeling activities, and (5) provide reliable results under difficult industrial conditions. The obstacles are that typical industrial data include a large number of highly correlated, noisy and possibly missing measurements. For such data sets, the use of traditional data analysis and modeling technologies can be both time consuming and produce results that do not capture the essential and useful information resident in the data. For any on-line application, robustness and long-term sustainability are key requirements for success.

In this paper, we will present some our industrial experiences in the use of multivariate statistics and the benefits from it. A brief introduction to the MVS technologies will be presented from an industrial perspective. This will be followed by a description of two on-line applications that continue to provide significant value to the company. The first example is an on-line process monitoring system that is used to observe the complete operation of the casting process in the mould area of Dofasco's #2 Continuous Slab Caster. The second example is a predictive model used at Dofasco's Desulphurization facility for determining the optimal amount of reagent needed to accurately remove sulphur from pig iron. The paper

* Author to whom correspondence should be addressed. Tel.: +1(905)548-7200 ext. 6986. E-mail: mike_dudzic@dofasco.ca.

concludes with the following two sections; our thoughts for future applied research in this area that would be of value to industry in general and some reflections on the industrial application of MVS.

2. OVERVIEW OF MULTIVARIATE STATISTICAL TECHNOLOGIES

The practical value of PCA and PLS modeling methods is that these techniques allow for the systematic examination and interpretation of large amounts of highly correlated data. Examination of MVS model outputs can provide insight into the operation of an industrial process during monitoring and quality assurance activities. With PCA and PLS, the systematic interpretation of dominant patterns in the data, and the isolation of the most important contributors to these patterns are also possible. This allows the classification of data relationships according to normal and abnormal operation. Some of the numerous advantages of PCA and PLS latent variable modeling methods have over traditional monitoring and prediction technologies are: (1) provision for data dimension reduction, (2) robustness to highly correlated, noisy and missing data, (3) the meaningful graphical display of model outputs, and (4) applicability to both continuous and batch processes.

The following section describes some concepts of multivariate statistical modeling from an industrial viewpoint. Further excellent theoretical reviews of applying MVS in industrial applications can be found in Kourti (2002), and Kourti and MacGregor (1996).

2.1. An Industrial Perspective on Multivariate Statistical Modeling

In both SPC and empirical modeling applications, the sample covariance matrix, $\mathbf{S_{xx}}=(n-1)^{-1}\mathbf{X}^T\mathbf{X}$, and its inverse, often form the basis of the statistics being calculated, where $\mathbf{X}$ is an n × p matrix of n observations of p variables. When developing linear predictive models, and in the calculation of the Hotelling T^2 statistic in monitoring applications, the inversion of an ill-conditioned $\mathbf{S_{xx}}$ matrix can be problematic. It is precisely when the $\mathbf{S_{xx}}$ matrix is ill-conditioned that multivariate statistics become particularly useful. Our focus on the use of MVS as the basis for SPC and modeling applications in industrial settings, is because in such applications, it is very common to have large data sets that include many variables that are highly correlated, and hence result in $\mathbf{S_{xx}}$ matrices that are ill-conditioned.

The Hotelling T^2 statistic is defined in Eq. (1) as the multidimensional extension of the t^2 statistic.

$$T^2=(\mathbf{x}-\boldsymbol{\mu}_\mathbf{x})^T\mathbf{S_{xx}}^{-1}(\mathbf{x}-\boldsymbol{\mu}_\mathbf{x}) \qquad (1)$$

where $\mathbf{x}$ is a p × 1 vector of variables whose covariance structure is described by $\mathbf{S_{xx}}$, and $\boldsymbol{\mu}_\mathbf{x}$ is a p × 1 vector of the associated sample means of the variables in $\mathbf{x}$. The Hotelling T^2 statistic is often used as the basis for a control chart in applications where there are many variables to be monitored, although there are other options such as the multivariate exponentially weighted moving average (MEWMA) or the multivariate cumulative sum (MCUSUM) statistics. However, due to the correlated nature of many process data, the matrix inversion in Eq. (1) can often produce poor numerical results due to ill-conditioning. To remedy this, the $\mathbf{S_{xx}}$ matrix may be decomposed using linear projection based methods such as principal component analysis (PCA). PCA projects the original process data ($\mathbf{x}$) into a lower dimensional space spanned by the first A eigenvectors of $\mathbf{S_{xx}}$. Once projected into this space, the resulting transformed vectors ($\mathbf{t}$) are orthogonal and have better numerical properties that allow for fault monitoring using the T^2 statistic given in Eq. (1) with $\mathbf{X}$ and $\mathbf{x}$ replaced by $\mathbf{T}$ and $\mathbf{t}$ respectively, where $\mathbf{T}$ is an n × A matrix of projected data ('scores'), and $\mathbf{t}$ is an A × 1 vector of scores for a single observation. The control chart is used by evaluating the value of T^2 at prescribed intervals and comparing this value to a control limit. Control charts can also be constructed to monitor the vector of $\mathbf{t}$ values individually.

The PCA projection of data is accompanied by some loss of information since the dimension of the space spanned by the first A eigenvectors of $\mathbf{S_{xx}}$ is smaller than the dimension of $\mathbf{S_{xx}}$ (i.e. A < p). To measure this, a companion statistic called the 'squared prediction error' (SPE) is also monitored for fault detection when projections are used. This statistic measures the squared orthogonal distance from the un-projected data point ($\mathbf{x}$) to the space spanned by the eigenvectors of $\mathbf{S_{xx}}$ and is also compared to a threshold limit in fault detection. This distance captures the information in the space spanned by the last p-A eigenvectors of $\mathbf{S_{xx}}$.

The definition of these statistics, their distributional properties, and their interpretation can be found in several references including Morrison (1990), Mason and Young (2002), Jackson (1991). Kourti and MacGregor (1996) have advocated a means of rewriting the T^2 and SPE statistics in terms of the elements that are used in the summation equations that define these terms. These elements, when separated and scaled in a particular way, can be related to the un-projected $\mathbf{X}$-data and are said to be 'contributors' to the T^2 and SPE statistics. These contributors can be used to investigate faults in a Pareto fashion by focusing effort on the large contributors.

In many industrial applications, it is necessary to use the information in measured process variables to predict values of output variables that are often related to product quality and performance, or to define values for input variables such as controller setpoints. Approaches to modeling for prediction include the development of first principles modeling, and empirical modeling. While first-principals models offer more promise in developing theoretical understanding of process behaviour, they are often difficult to formulate and costly to develop. In some applications, these models can be too computationally intense to solve for direct on-line use. In comparison, many empirical models, including those based on MVS, offer less theoretical insight but can be developed and solved more rapidly than first-principles models. Furthermore, in contrast to first-principles models, empirical models can include variables that theoretical models do not account for, are generally only valid in the region spanned by the data used to build the model, and should not be expected to produce reliable results when extrapolating. Nevertheless, empirical models are often useful in industrial settings where speed of implementation and robustness to practical issues such as missing data are important.

One useful class of empirical models are based on linear predictions. These predictive models have the following form:

$$\mathbf{y}_p = \boldsymbol{\beta}\,\mathbf{x} \qquad (2)$$

where the output variable $\mathbf{y}_p$ is a k × 1 vector of predicted quantities that result from the combination of the input vector $\mathbf{x}$. Predictions can be made when the elements of the k × p matrix $\boldsymbol{\beta}$ have been computed. There are various approaches to determining values for $\boldsymbol{\beta}$, many of which rely on the inversion of the $\mathbf{X}^T\mathbf{X}$ (i.e. (n-1) $\mathbf{S}_{\mathbf{xx}}$) matrix. If the variables forming the columns of the data matrix **X** are nearly collinear, as is often the case, the inversion of $\mathbf{X}^T\mathbf{X}$ can become numerically unstable and can produce poor predictions that are of little practical use. This phenomenon is well known and remedies such as PLS are available.

PLS is a projection-based method that explicitly employs the covariance structures between the elements of **x** to those of **y**, and those of **x** to **x,** and **y** to **y**, in model development. The differences between this and the objectives in modeling with other methods were analyzed by Burnham et. al. (1999), and the numerical methods associated with PLS are summarized in Nelson et al. (1996). In general, PLS and other projection-based methods can produce estimates that are more stable because the correlation among the variables being modeled has been taken into account.

From a robustness point-of-view, the eigenvalue-eigenvector model structure used in PCA can be used to treat missing data in many ways during both the model estimation phase and during on-line use of the model. One approach to missing data relies on the model estimation algorithm. During the PCA model estimation phase, the application of the Non-linear Iterative Partial Least Squares (NIPALS) algorithm extracts eigenvectors from the **X** data sequentially via the application of an iterative series of linear regressions. When data in any of the columns of the **X** matrix are missing, the regressions are preformed using the data that are present and the missing points are ignored. As long as the number of variables in any row or column is greater than the number of principle components being estimated, the NIPALS algorithm can find a solution. A detailed description of various missing data handling methods for both PCA and PLS is given by Nelson *et. al*. (1996).

3. ON-LINE MONITORING APPLICATION – #2 CONTINUOUS CASTER STABLE OPERATION MONITORING SYSTEM

In this section is a description of selected features of the integrated on-line monitoring system at Dofasco's #2 Continuous Caster. This will highlight the use of PCA in a multivariable SPC monitoring application.

3.1. Continuous Steel Caster Process Description

In the manufacture of steel, the conversion of liquid steel into solid steel slabs is commonly achieved through a process known as continuous casting, which is illustrated in Figure 1.

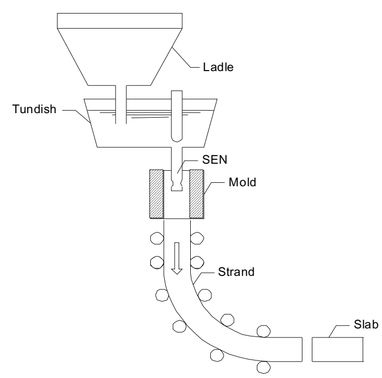

Fig. 1. Schematic of a continuous caster.

In this process, the liquid steel is shipped by a ladle to the continuous caster, and then continuously poured into an accumulator vessel called a tundish. The tundish acts as a buffer, allowing casting to continue while a new ladle is delivered to the caster. The liquid steel is then fed into an open-ended copper mould through a submerged entry nozzle, SEN. Cooling water is circulated within the mould walls so that liquid steel in contact with the copper mould solidifies, forming a steel strand with a solid shell and liquid center. The strand is continuously withdrawn from the mould into additional cooling chambers where the internal liquid steel solidifies under controlled cooling conditions. Once the strand has fully solidified, it is cut into slabs according to customer requirements.

In a continuous caster, the typical operating sequence consists of a brief start-up transition (called a start-cast), followed by a prolonged continuous, run-time operation, and then a shutdown operation. Within the run-time operation, there are occasional process transitions that occur to accommodate equipment or product specification changes.

3.2. Motivation for Improved Monitoring

The focus for improvement was to detect the onset and to prevent the occurrence of caster strand breakouts, with a secondary focus on providing early detection of other abnormal operating situations. A breakout is a catastrophic failure in the steel casting process. It occurs when the solidifying shell of the cast steel strand does not form uniformly on leaving the bottomless mould and ruptures, allowing molten steel to escape. This spillage of molten steel can be potentially hazardous and can cause equipment damage and costly process delays. The onset of a caster strand breakout is known to be related to changes in mould thermocouple signals, cooling water flows, mould level and other process variables that are all interrelated.

Experience shows that breakouts occur during start-cast, run-time or under process transition operation. Breakouts can be avoided by reducing the casting speed thereby giving more residence time in the mould for the steel to solidify. To avoid the occurrence of a breakout, it is critical to detect improper solidification of the steel shell in advance with enough lead-time to appropriately slow down the casting.

Monitoring continuous casting processes and predicting the onset of breakouts in advance has been attracting considerable industrial interest for several decades. To date, commercially available monitoring systems only address those types of breakouts in run-time cast operations that have definite signatures in the process data. Our challenge was to develop an improved monitoring scheme that included provisions for: (1) the ability to detect many different classes of breakouts, (2) advanced warning, as early as possible, to the high probability onset of a breakout under all applicable operating conditions, and (3) easy-to-access diagnostics for real-time troubleshooting of processing problems that could potentially result in a breakout.

3.3. PCA-based Multivariate SPC Monitoring Solution

In the following subsections, are descriptions of selected features of the PCA-based monitoring system under run-time, start-cast and process transition operating conditions.

3.3.1. Run-time Operations Monitoring

PCA has been applied to monitor the continuous run-time cast operations, where a principal component model was developed based on selected historical data from over half a year of normal run-time operations. The model building task involved the selection of significant input measurements, the incorporation of variable lagging to capture process dynamics, and the appropriate pre-treatment of the data to ensure a representative data set. The resulting PCA model is a projection of 323 inputs into a 10 dimensional score space. Two control charts, the Hotelling T^2 (HT) statistic and the squared prediction error (SPE), then monitor the real-time output of this model.

For this application, 6 PCA models were required to cover all the different operating scenarios at the caster. Model selection is automatic so that the operator only views the correct model output. During run-time operation, the system uses the appropriate model to compare current inputs to their historical benchmark in real-time at a sub-second frequency. The resulting HT and SPE control charts provide a continuous measurement of the overall stability of the casting process. In addition, variable contributions are calculated and ranked in order of magnitude to provide direction to operators in troubleshooting their process.

HMI (Human Machine Interface) screens have been custom designed to supply sufficient information for operators to monitor caster operations in real-time and perform further diagnosis if abnormal situations are detected. As shown in Figure 2, the information contained in the main screen of the integrated caster monitoring system includes: (1) general casting information such as heat number, steel grade, casting speed, etc..., (2) SPE and HT control charts indicate the stability of the casting operation, and (3) a graphical representation of the mould and its surrounding thermocouples. The SPE and HT are scaled to [0,1] with respect to the control limits, respectively. There is a warning limit set at 0.8 and an audible alarm is generated if the scaled SPE or HT statistics exceed the value of one. Other screens include contribution plots (see Figure 3), customized trending, etc…, which operators can access via navigation buttons.

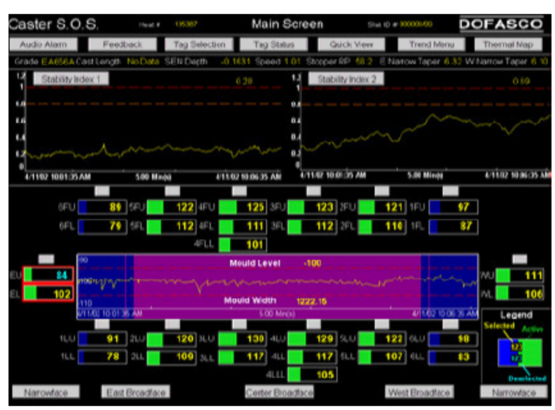

Fig. 2. Main screen of integrated monitoring system for #2 Continuous Caster

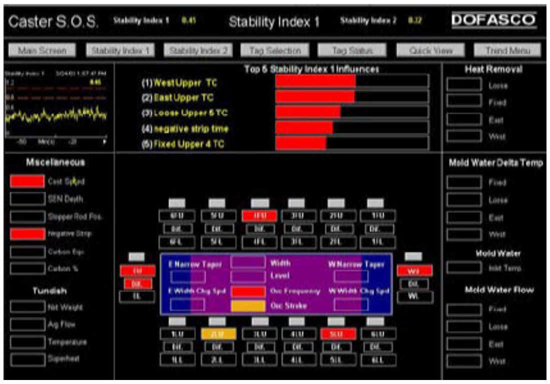

Fig. 3. Contribution plot screen of integrated monitoring system for #2 Continuous Caster

Further details on run-time MVS caster monitoring can be found in Vaculik*, et. al.* (2001).

3.3.2. Start-cast Operations Monitoring

The start-cast is essentially a dynamic process in which most process variables (such as thermocouple temperatures, heat transfer flux through each side of the mould, etc…) show non-linear trajectories versus time and they are highly auto-correlated. In continuous processes, it is expected that the relationship among the variables remains stationary through time. In order to capture the non-stationary, non-linearity and the dynamic nature of the start-cast process, it is necessary to include data through time for all variables.

In approaching this problem, a start-cast is considered a batch process over a pre-defined, finite duration. For our case, the start cast duration begins with the time the casting speed starts to increase, and ends when the strand length exceeds a specific value (as determined by process knowledge). The multivariate statistics technology applicable for this application is multi-way PCA (MPCA). MPCA is an extension of the PCA algorithm, as was demonstrated by Nomikos and MacGregor (1994 and 1995) for batch process monitoring. The MPCA algorithm organizes the data such that each sample of process measurements at different sampling intervals is considered a new variable in the MPCA model. This allows the characterization of the changing relationships among the variables.

Some key issues in applying MPCA to monitoring start-casts are briefly described here. Further details on start-cast MVS caster monitoring can be found in Zhang, *et. al*. (2003). For theoretical details on the use of MVS in dynamic data modeling for batch processes, see Kourti (2003).

To construct the start-cast data, large quantities of historical data for normal start-cast operations are required. The data can be shaped into a 3-dimensional data block in which the 3 dimensions are: a number of normal start-cast *operations* (must cover the full region of operation); the set of process *variables*; and the samples through time (*observations*) over the start-cast duration. In the model development, 147 start-cast *operations,* with 60 process *variables,* over 800 *observations,* were used.

In the actual operation, the start-cast duration is variable. This does not comply with one of underlying requirements in the MPCA algorithm that, all trajectories must be synchronized with the evolution of the batches. To solve this problem, the indicator variable approach (Nomikos and MacGregor, 1994) was used to synchronize the process trajectories, where the strand length acts as the indicator variable since it progresses monotonically in time and has the same starting and ending value for each start-cast. All process trajectories of start-casts are synchronized using interpolation based on a set of predefined scales in the strand length. The synchronization scales are determined by a quadratic function of time such that it possesses small intervals at the beginning of the start-cast duration and large intervals at the end. Such non-uniform scales provide a better opportunity for early detection of abnormal situations.

The core concept of the MPCA algorithm is to handle the resulting 3-dimensional, synchronized dynamic data. The algorithm first unfolds the 3-dimensional data block to a 2-dimensional data matrix to preserve the direction of operations (see Figure 4).

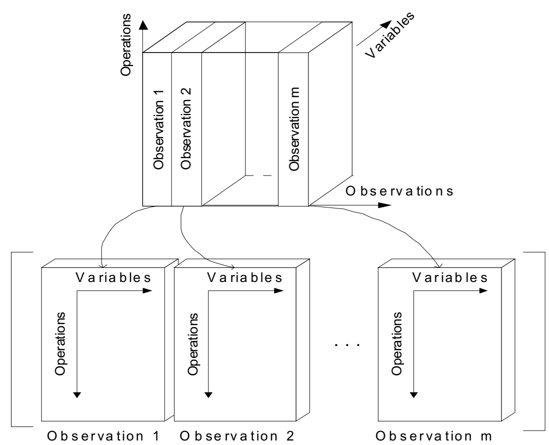

Fig. 4. 3-Dimensional data unfolding

This is accomplished by the following two steps. First, the data block is sliced vertically along the observation direction; next the slices are juxtaposed to build the data matrix with a large column dimension such that each row corresponds to a start-cast. A standard PCA algorithm is then applied to this unfolded data matrix to build a statistical model for the deviation of each process variable from its average trajectory over the start-cast duration, benchmarking the operation-to-operation variance existing in the historical data of normal start-casts. Similar to the run-time operation monitoring, two statistics, SPE and HT, are defined at each observation over the start-cast duration in such a way that they are able to describe how each start-cast is compared with normal operation as described by the model. As a new start-cast evolves, its deviation from the average trajectories is examined by comparing it to the model. If the new start-cast is statistically different from normal operation, then an alarm is generated to indicate the abnormal situation. Once an alarm is generated, the variables of the new start-cast that contribute most to the process deviation are identified in the contribution plots.

3.3.3. Process Transition Operation Monitoring

The remaining operating condition to monitor for involves process transitions. There are some well defined process transitions involved with the continuous casting process related to equipment or product specification changes where abnormal operation resulting in a breakout can occur. Our solution to monitoring one important process transition, the changing of the submerged entry nozzle, is discussed in detail in the companion paper at this conference titled: "Industrial Experience on Process Transition Monitoring for Continuous Steel Casting Operation".

3.4. Integrated Monitoring System Architecture

The schematic of the fully integrated on-line monitoring system for start-cast, process transitions and the continuous, run-time cast operations is depicted in Figure 5. Real-time measurements of various sensors are collected on-line by a data acquisition module, and then sent to an on-line process monitoring module as well as a historical database for data archiving purposes. Once the monitoring module receives the real-time process measurements, a series of computations are performed based on a selected multivariable statistical model from the model set.

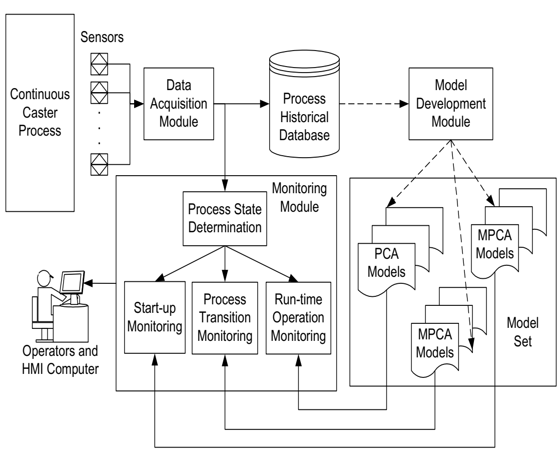

Fig. 5. Schematic of the integrated on-line caster monitoring system

The monitoring module consists of 4 computation functions, i.e., start-up monitoring, process transition monitoring, run-time operation monitoring and process state determination. The latter is able to determine the process state based on the current process conditions (described by casting speed, strand length, etc…), and automatically selects the appropriate model and monitoring function. Five general process states are defined: shut-down, start-up, specific process transition, run-time and idle (which is designed to handle some special operating conditions). Based on current process conditions, a state is determined and the corresponding model results are presented to the operator. For presentation simplicity and ease-of-use for the operators, the HMI screens are identical for all of the operating states and transition seamlessly between them.

3.5. Operational Results and Benefits

Since the deployment of the caster monitoring system at #1 Continuous Caster in April 1997, Dofasco has enjoyed its highest productivity levels with an approximate 50% reduction in breakouts over the last 7 years from its previous best year. The system has also aided in early detection of other critical operational situations such as mould level sensor problems. In general, this system has helped operators have greater confidence in running the caster at higher production rates.

The run-time monitoring application at #2 Continuous Caster was implemented on-line in February 2002. That year, there were 4 high probability cases of breakouts that were prevented through the aid of this new monitoring system. A significant reduction in breakouts was also observed in 2003, as compared to previous years. A portion of the start-cast monitoring feature was implemented in May 2003 and fully commissioned January 2004. In 2003, there was one high probability start-cast breakout case that was prevented through the aid of this new monitoring feature. The process transition monitoring feature is still under development.

The run-time monitoring application is patented with the start-cast and process transition monitoring features patent pending.

4. ON-LINE PREDICTIVE CONTROL APPLICATION – DESULPHURIZATION REAGENT CONTROL (DRC) SYSTEM

4.1. Torpedo Car Desulphurization Process

Dofasco uses a Torpedo Car Desulphurization facility to remove sulphur from most batches of hot metal in order for the steel to meet the specified quality standard. A schematic of the facility is illustrated in Figure 6. Desulphurization refers to the process of injecting the hot metal with chemical agents in order to bind, and thereby remove, a portion of the sulphur in the metal. At this facility, 2 different reagents are utilized.

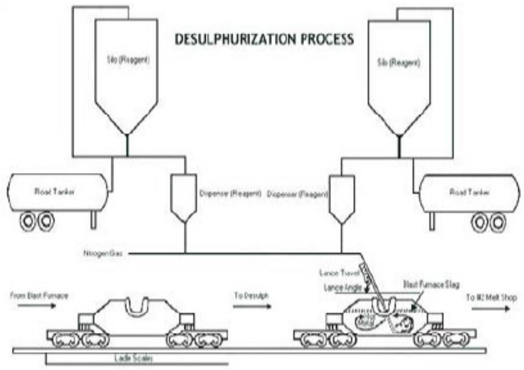

Fig. 6. Torpedo car desulphurization process

4.2. Motivation for Controller Change

The harsh environment at this facility makes it challenging to obtain reliable, high-quality process data. Because of this, only very simple regression models, utilizing a fraction of the available data, were robust enough for the original on-line model-based control system. An additional concern was that the maintenance of these models was a time-consuming, cumbersome task. The operations staff felt there was room to improve the performance of this control system in order to better control the reagent usage and accurately meet downstream sulphur targets.

4.3. Adaptive PLS Control Strategy

An emprirical modeling methodology was required in this application because the mechanisms of the chemical and mechanical effects are not fully understood. As an approach, PLS was favoured because it enabled us to most effectively use the multitude of highly correlated data available for this operation. Also, PLS is able to robustly handle situations where real-time data are missing, a reality in this application. In the desulphurization application, it is known that the process does shift and drift due to changes in the characteristics of the purchased chemical reactants. Therefore, it is important that the modeling technology have an adaptation feature.

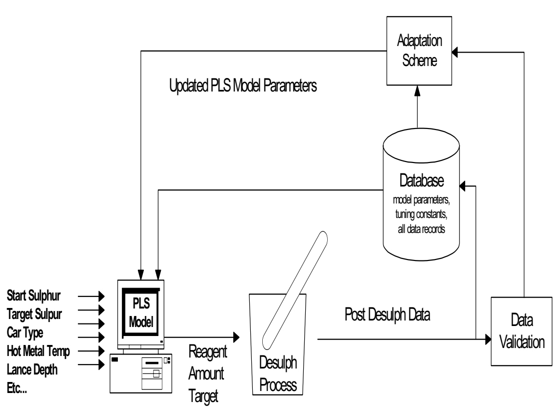

Fig. 7. Adaptive PLS desuphurization reagent control logic

The control structure developed is illustrated in Figure 7. The algorithm used is a modified version of the Adaptive Modified Kernel Algorithm described in Dayal and MacGregor (1997a,b). The PLS models at the Desulph Station are defined in terms of:

1. 28 β values (from Eq. 2) – 14 β values for each of the two reagents, where each value of β is the coefficient associated with one of the 14 input variables (which include initial and target sulphur values, mass of hot metal, all available chemistry data, and process variables such as the type of torpedo car in use).
2. 14 xm values – these are the weighted means for each of the transformed input variables (x).
3. 2 ym values – these are analogous to the xm values but represent the weighted means of the transformed output variables (y)
4. 14 xw values – these are the weighted standard deviations for each of the transformed x variables
5. 2 yw values – these are analogous to the xw values but represent the weighted standard deviations of the transformed y variables

Together, these values fully specify the model and allow the calculation of the required amounts of the two reagents. The adaptation algorithm updates the values of all of these parameters when an appropriate new set of data is available. The model updating is done as follows.

The xm and ym values are updated to reflect the new data according to:

$$\begin{aligned} xm_{j,i} &= (1-\alpha_1)\, xm_{j,i-1} + \alpha_1\, \hat{x}_{j,i} \\ ym_{j,i} &= (1-\alpha_1)\, ym_{j,i-1} + \alpha_1\, \hat{y}_{j,i} \end{aligned} \tag{3}$$

where $xm_{j,i}$ is the new weighted average value for the jth x variable, $xm_{j,i-1}$ is the weighted average from the previous iteration of the adaptation scheme, $\hat{x}_{j,\,i}$ is the mean of the transformed values for the jth x-variable in the new dataset, and α_1 is a tuning parameter that can take on values between zero and one and determines how quickly old data are discounted in the adaptation routine. The y variables are defined in a similar way.

The xw and yw values are updated to reflect changes in the standard deviations over time according to:

$$xw_{j,\,i} = \sqrt{(1-\alpha_2)\,\sigma^2{}_{j,i-1} + \alpha_2\, \breve{x}^2{}_j} \tag{4}$$

where $\sigma^2{}_{j,\,i-1}$ is the weighted variance for the jth x-type variable from the previous time the adaptation algorithm was executed, $\breve{x}_{j,i}$ is variance for the transformed j^th x-variables in the new dataset, $xw_{j,i}$ is the new value of the weighting factor for the j^th x-variable, and α_2 is the discounting factor that determines how much weight is given to previous values of the variances of the variables. The yw values are updated similarly.

The values of the β's are determined based on the kernel PLS algorithm using $\mathbf{X}^{\mathbf{T}}\mathbf{X}$ and $\mathbf{X}^{\mathbf{T}}\mathbf{Y}$. It is the $\mathbf{X}^{\mathbf{T}}\mathbf{X}$ and $\mathbf{X}^{\mathbf{T}}\mathbf{Y}$ matrices that are updated using the new data. These matrices are used to update the "old" covariance structures. This updating is done using a standard moving average scheme as follows.

$$\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)_{updated} = \alpha\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)_{current} + (1-\alpha)\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)_{new} \quad (6)$$

$$\left(\mathbf{X}^{\mathrm{T}}\mathbf{Y}\right)_{updated} = \alpha\left(\mathbf{X}^{\mathrm{T}}\mathbf{Y}\right)_{current} + (1-\alpha)\left(\mathbf{X}^{\mathrm{T}}\mathbf{Y}\right)_{new} \quad (7)$$

The updated correlation matrices are then used to fit a new PLS model. Finally, the kernel PLS algorithm is updated in the following manner:

$$\beta_{updated} = f\left(\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)_{updated}, \left(\mathbf{X}^{\mathrm{T}}\mathbf{Y}\right)_{updated}\right) \quad (8)$$

where $\beta_{updated}$ represents the new set of all β values. Note that together with the parameters of the model, the most recent values for $\mathbf{X}^{\mathbf{T}}\mathbf{X}$ and $\mathbf{X}^{\mathbf{T}}\mathbf{Y}$ must be stored for use in the next call to the adaptation algorithm. One advantage of the kernel algorithm is that only these matrices need to be stored and not the original data matrices $\mathbf{X}$ and $\mathbf{Y}$.

4.4. On-line Implementation

The initial model was fitted using approximately three months worth of historical data, which amounts to a few thousand observations. The initial version of the on-line PLS application was implemented in July 1995. It was integrated into the operation as a PC based control system with an easy-to-use Visual Basic HMI. In 2000, changes to operating practices required a new model algorithm to handle multiple (co-injected) reagents. The updated PLS control strategy, which included the adaptation feature, was implemented in September 2000.

In demonstrating the new control system performance, Figure 8 shows the "before" and "after" histograms for the distribution of the final sulphur values for one aim. Approximately three months worth of data for both periods of operation were used to generate the frequency values for each distribution. The figure shows that the new model provided more precise estimates of the amounts of reagent required (narrowed the distribution) and this allowed Dofasco to move its target sulphur values closer to the constraint.

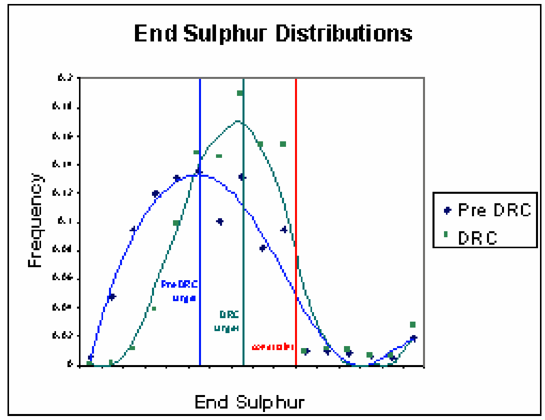

Fig. 8. Distributions of End Sulphur values before and after the implementation of the adaptive PLS model

In demonstrating the adaptation capability, Figure 9 shows a plot of the predicted quantities of reagent required to achieve a desired desulphurization (based on the PLS model) versus the amounts of reagent actually used to achieve the result. If the models were perfect and there was no noise in the data, this plot would be a straight line with a slope of one. There are two groups of data shown on this plot. The red diamonds represent data based on a PLS model with no adaptation. The black stars are points that were generated using an adaptive PLS model. Both groups of data show remarkably good clustering close to the diagonal line indicating that reliable predictions are possible despite noise in the raw data. However, the static results show a slight upwards offset from the adaptive results. This represents a small bias in the static model such that it is consistently predicting slightly too much reagent required. Even a small bias like this can result in substantial over-consumption of reagent over the long term. Model adaptation eliminates the offset and provides assurance that the model is reliable and up-to-date.

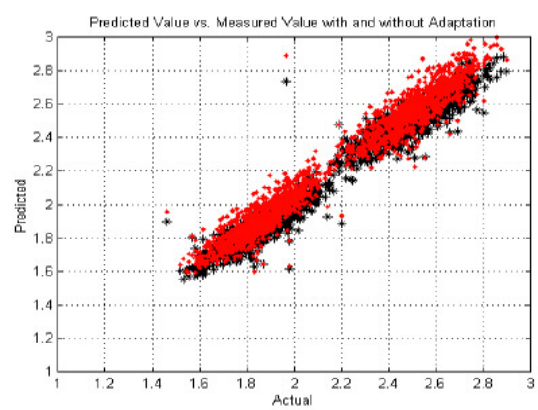

Fig. 9. Predicted versus actual reagent quantities. (KEY: red diamonds - without adaptation; black stars - with adaptation)

4.5. Operational Results

There was immediate positive feedback from operations on the performance of the new PLS reagent control system as compared to the previous model-based control system. At the end of 1995, a formal audit of the new control system was established. The results were: (1) the root mean square (RMS) error between the aim and final sulphur was reduced by 50%, (2) the re-desulphurization rate (i.e. the times where a second reagent addition is necessary to meet target) was reduced by 70%, (3) the amount of reagent purchased was reduced by 8.5%, and (4) the process iron yield was improved by 0.25% due to less iron tied up in slag. The performance of the enhanced adaptive PLS algorithm was also excellent, with minimal model maintenance effort being required. Since the introduction of the co-injection operating strategy, improvement initiatives (i.e. the adaptive PLS model, automation upgrades and new operating practices) have achieved another 17% reduction in reagent consumption (Quinn et al., 2002), while accurately meeting sulphur targets. Improved process understanding and process control have helped in unforeseen ways. For example, our relationships with our supplies have improved because our combined expertise has been maximized - Dofasco is the process expert, while the suppliers are the product experts.

The adaptive PLS desulphurization reagent control system is patented.

5. CONSIDERATIONS FOR FURTHER MVS BASED INVESTIGATIONS

Through the experiences of implementing on-line MVS applications at Dofasco, much has been learned about what it takes to successfully commission such applications. In our opinion, there are several challenges that industrial practitioners will have to overcome to be able to effectively utilize the full potential of this technology in on-line applications. In general, what is needed is a flexible, industrially-focussed, methodology for applying MVS technology. The success of this technology in industry will be directly related to the availability of MVS systems that are useable and understandable by people of varying technical backgrounds. In our experience, MVS solutions in industry need to pay special attention to issues that can be categorized under the following headings: (1) identifying and defining appropriate application scope, (2) data gathering / preprocessing, (3) off-line modeling to on-line solution, and (4) long-term maintenance strategies. These are discussed further below. In identifying these issues, we have also tried to provide suggestions on how to address them; however, we believe that many of these issues remain open areas for research and we hope this paper may open new academic and industrial work in these areas.

5.1. Identifying and Defining an Appropriate Scope

In an industrial setting, it is necessary to be able state the economic justification. In this regard, the practitioner must be able to determine if the capability and applicability of the MVS technology makes sense over alternate technologies. Similarly, one must define the appropriate modeling strategy necessary to handle many complex operating situations such as: many different products produced by a process, large feed rate changes, recycle processes, and more complicated batch processing. We believe that academic studies on applying MVS in these cases would be beneficial, and that results could be compelling in defining project scope for future industrial MVS work. We stress this because we have found that appropriately defining scope for the application greatly improves its chances of long-term success. For example, it is important to determine how small or how big a modeling application should be, and whether it is better to have many small applications versus one large application. Defining these matters at the outset of a project will impact many issues from long-term maintenance to data acquisition and processing.

5.2. Data Gathering and Data Preprocessing

We have found that a key to success in any multivariate statistical application is the early involvement of process experts. Although multivariate statistical models are empirical models, having access to expert advice on which variables are "known" to be important and what transformations or combinations of variables may be important, can affect not only the speed of development of a system but also its ultimate success. Furthermore, process experts are usually helpful in constructing appropriate training and validation data sets if historical data is to be used. Should there prove to be a need for designed experimentation, process experts can help in that too. Ultimately, process experts are required in determining the depth and breadth of data to cover the operating window of the process. In most cases we have encountered, it was important that data pre-processing take into account operation-specific needs which may include: non steady-state process dynamics, non-stationary process conditions, non-linear process conditions, and large process dead-times. We suggest more in-depth collaboration between industry and academia in developing an understanding of industrial data sets and, data presentation and visualization needs.

5.3. Application of Models On-line

Typically, the first step in putting an MVS model on-line is justifying the proposal to do so. This involves quantifying the benefits in terms of cost, revenue, safety and/or quality. Therefore, meaningful measures of the applicable benefits should be developed and computed. What these criteria are will depend on the application, but typically they are related to the performance of the process or expected changes in the performance as opposed to being related to model performance such as the percentage of variability explained by the model. For example, in monitoring and fault detection applications, tuning of the alarm levels is very important. Although some academics advocate setting limits based on a choice of a probability (tail) value, tuning limits based on achieving a desired average run length (ARL) may be more useful in some instances, especially relative to economic costs incurred while using control charts where process adjustments are made based on both T^2 and Q^2. We suggest more academic work be preformed on statistical topics such as this, perhaps following Box and Jenkins (1963), Box and Kramer (1992), and Lorenzen and Vance (1986).

5.4. Long-term Maintenance Strategies

Having long-term maintenance strategies in place is important in ensuring success. Typically, the sophistication of these strategies is determined by the rate at which the process is expected to change (i.e. how often it is expected that the models or alarm limits will need tuning). In cases where the process is known to shift or drift over relatively short time periods, it may be necessary to consider on-line adaptation, among other things. Especially in the area of maintenance strategies, we believe that there is still research to be done. The following are areas we feel require further work: on-line adaptation of alarm limits in monitoring applications, monitoring the performance of MVS systems to determine when tuning / remodeling is required, methods for expediting the remodeling process when a major change is made to the process, and simple and efficient ways to minimize the number of models and limits that need to be maintained. Dofasco has developed internal methodologies to address some of these needs, however, there is a significant opportunity to further address these for the benefit of all process industries.

6. REFLECTIONS, SUMMARY, CONCLUSIONS AND FUTURE DIRECTIONS

This paper has described two important on-line MVS-based applications at Dofasco that continue to provide significant value to the company. Since our first off-line analysis in 1993, we have utilized MVS-based technologies on over 70 applications. These areas can be categorized in the following groups: (1) off-line process analysis / troubleshooting, (2) on-line process monitoring, and (3) on-line model-based control. To date, Dofasco has implemented 9 on-line MVS-based applications, 6 in the area of process monitoring and 3 in model-based control strategies. The remainder involve off-line process analysis / troubleshooting. Off-line process analysis is a prerequisite proof-of-concept activity that provides justification for moving forward with on-line applications. Off-line troubleshooting applications typically involve discovering specific insight hidden within a large dataset to support a team effort to fix a problem. The applicability of this technology over a wide range of problems / opportunities that we deal with make it a very valuable tool for the company.

Multivariate Statistical technologies have been successful at Dofasco because of the support provided at all levels of management. The MVS applications developed at Dofasco stem from a structured approach in developing applications, and have delivered valuable business results that can be related to company profits. The MVS applications have proven to be robust to many practical matters such as missing data, and have proven to be amenable to various maintenance and updating strategies. MVS provides models and control chart schemes that are readily understood by production personnel once training is provided. Although MVS methods use many process variables simultaneously for modeling and fault detection, the methods provide a built-in means of relating SPC chart, and model calculations to the individual variables used for modeling. This has facilitated troubleshooting and process diagnosis on the part of shop floor personnel, and has helped in building acceptance of the MVS methods.

Perhaps more than anything else, operator (final user) acceptance is critical to the long-term success of an on-line MVS system. Early involvement of the operators in the design of the user interfaces has proven beneficial in our experience. Also, seeking advice from the operators early, and throughout the process has helped in acceptance of the system, and has helped alleviate feelings of being intimidated by the technology. Operators often have very good suggestions on how to integrate the new technology into existing control room infrastructure and / or how to incorporate signals and information from other systems into the new MVS system.

This paper has addressed many technical and operational issues in the on-line application of MVS technology. We note here that there are also other corporate issues that need to be considered. The most critical and perhaps the most difficult to quantify or influence is that of having a corporate culture to support this type of development. For Dofasco, this has meant that a willingness to perform MVS work and to manage it in the form of an engineering project was necessary. When coupled with project management methods, periodic project evaluation, and the on-going development of computer infrastructure, the framework we presented has helped to create success and has led to changes in the way analysis work is done in general. The success of the MVS, and other, often much larger, data-driven projects has led to more use of data by shop floor personnel during daily activities like production meetings, more discussion on how to improve quality, and most importantly, more ideas on how to use available data to improve production methods and products.

Dofasco's vision to the future for MVS technologies is to "institutionalize" it throughout the corporation by: (1) maximizing the understanding and benefits of this technology through technical awareness and training, (2) integrating, where appropriate, the technology into our corporate quality initiatives, and (3) continuing to build technical breadth and depth while leveraging leading-edge aspects of the technology with the academic experts (especially with linkages at McMaster University). To facilitate the propagation of this technology throughout the corporation, Dofasco has developed an in-house generic software platform for on-line MVS-based solutions. This platform is architected with all the latest open system features for application on various computing platforms that are resident in the company. This also facilitates and simplifies the long-term maintenance of a growing application base.

In general, Dofasco will continue to leverage multivariate statistical methods, with other advanced automation technologies, to develop new, high-valued on-line control and monitoring applications.

7. REFERENCES

Box, G. E. P., G. M. Jenkins (1963), Further Contributions to Adaptive Quality Control: Simultaneous Estimation of Dynamics: Non-Zero Costs, Bulletin of the International Statistical Institute, Proceedings of the 34-th Session, Ottawa, Canada.

Box, G. E. P., T. Kramer (1992), Statistical Process Monitoring and Feedback Adjustment. A Discussion, Technometrics, **34** (3), pp.251-285.

Burnham, A., J.F. MacGregor, R. Viveros (1999), Latent Variable Multivariate Regression Modeling, Chemometircs and Intelligent Laboratory Systems, **48**, pp.167-180.

Dayal, B., J. F. MacGregor (1997a), Improved PLS Algorithms, Journal of Chemometrics, **11**, pp.73-85.

Dayal, B., J. F. MacGregor (1997b), Recursive Exponentially Weighted PLS and its Application to Adaptive control and Prediction, Journal of Process Control, **7** (3), pp.169-179.

Dudzic M., V. Vaculik and I. Miletic (2000), "On-line applications of multivariate statistics at Dofasco", in *Proc*. IFAC Workshop on Future Trends in Automation in Mineral and Metal Processing, Heisinki, Finland, pp.438-443.

Jackson, J. E. (1991), A User's Guide to Principal Components. John Wiley and Sons, New York.

Kourti, T., J.F. MacGregor (1996), Recent Developments in Multivariate SPC Methods for Monitoring and Diagnosing Process and Product Performance, Journal of Quality Technology, **28** (4), pp.409-428.

Kourti, T., J. Lee, J.F. MacGregor (1996), Experiences with Industrial Applications of Projection Methods for Multivariate Statistical Process Control, Computers in Chemical Engineering, **20**, S745-S750.

Kourti, T. (2002), Process Analysis and Abnormal Situation Detection: From Theory to Practice, IEEE Control Systems, **22** (5), pp.10-25.

Kourti T. (2003), Multivariate dynamic data modeling for analysis and statistical process control of batch processes, start-ups and grade transitions, *J. Chemometrics*, **17**, pp.93-109

Lorenzen, T. J., L. C., Vance (1986), The Economic Design of Control Charts: A Unified Approach, Technometrics, **28** (1), pp.3-10.

Morrison, D. F. (1990), "Multivariate Statistical Methods, Third Edition", McGraw-Hill, New York.

Nelson P.R.C., P.A. Taylor and J.F. MacGregor (1996), Missing data methods in PCA and PLS: score calculations with incomplete observations, *Chemometric Intell. Lab. Syst.*, **35**, pp.45-65

Nomikos P and J.F. MacGregor (1994), Monitoring of batch processes using multi-way principal component analysis, *AICHE J*., **40**, pp.1361-1375

Nomikos P and J.F. MacGregor (1995), Multivariate SPC charts for monitoring batch processes, *Technometrics*, **37**, pp.41-59

Quinn, S.L., V. Vaculik (2002), Improving the Desulfurization Process Using Adaptive Multivariate Statistical Modeling, AISE Steel Technology, **79** (10) pp.37-41.

Vaculik, V., B. MacCuish, M. Dudzic, I. Miletic, M. Hough, K. Smith (2001), "Innovative monitoring technology improves Caster operation at Dofasco", Iron and Steelmaker, **28**, No. 9, pp. 23-28.

Zhang Y. M. Dudzic and V. Vaculik (2003), Integrated monitoring solution to start-up and run-time operations for continuous casting, *Annual Reviews in Control,* **27**, pp.141-149