2003 SME ANNUAL MEETING & EXHIBIT
February 24-26, 2003
Cincinnati, Ohio

SESSIONS NAME : CURRENT PRACTICES IN PROCESS CONTROL

STATISTICAL MULTIVARIATE ANALYSIS FOR IMPROVED PLANT CONTROL AT C. M. DOÑA INES DE COLLAHUASI SCM, CHILE

M. Orchard*, F. Romero† and L. Yacher*

†C.M. DONA INES DE COLLAHUASI SCM, Chile

*CONTAC Engineers LTDA. Avda Américo Vespucio Sur 315. Nunoa. Santiago, Chile.
Phone: (56-2) 270.09.00. e-mail: lyacher@contac.cl

STATISTICAL MULTIVARIATE ANALYSIS FOR IMPROVED PLANT CONTROL AT C. M. DOÑA INES DE COLLAHUASI SCM, CHILE

M. Orchard*, F. Romero† and L. Yacher*

†C.M. DONA INES DE COLLAHUASI SCM, Chile

*CONTAC Engineers LTDA. Avda Américo Vespucio Sur 315. Nunoa. Santiago, Chile.
Phone: (56-2) 270.09.00. e-mail: lyacher@contac.cl

ABSTRACT

Specific software tools for Statistical Multivariate Analysis combined with real time process data were applied to Concentrator Plant at Compañía Minera Doña Inés de Collahuasi SCM in Chile, basically to two Mill Lines, each one with one SAG mill (32'x15' 8MW) and one Ball mill (22'x36' 9.7MW).

The main objective was to empower the operator actions using process patterns references instead of process variable set-points. Pattern clustering and the analysis of pattern deviations regarding normal or desired pattern, allows for improvements in process control and the further implementation of soft sensors for those variables not included in the initial model.

INTRODUCTION

The present state-of-the-art in data acquisition systems, such as PI (Plant Information Systems), enables the use of a wide range of supervision techniques, which implementation had been previously slacken only by technical issues.

Aware of this challenge, Compañía Minera Doña Inés de Collahuasi SCM in Chile -with the support of CONTAC Engineers Ltda.- has implemented interesting sets of statistical mono and multivariate techniques included into the SCAN software, in order to obtain important improvements in its Mill Lines. Those techniques, were carefully selected to enable a complete analysis for the process operation, and therefore, to achieve better performances in the selected plants or specific control objectives.

In order to make use of these techniques, a complete analysis of the SAG mill processes was first needed. As main objective of this first stage, it was considered the detection and characterization of main process operation patterns and disturbances.

The obtained results, which allow to identify variability sources and to characterize the correlation structure between the variables of controlled processes, have proven the efficiency of CONTAC Engineer's SCAN software as an important approach for activity improvement and detection of the most recurrent disturbance sources.

THEORETICAL FOUNDATION

Most of the results that are shown in this paper need a basic knowledge of statistical multivariate analysis to ensure an adequate comprehension. Therefore, a review of the main concepts needed to the interpretation of VFA score and loading plots will be done in this chapter.

<u>Variability factor analysis (VFA)</u>

VFA is an statistical multivariate analysis that considers all predominant sources of variability in a data set and, additionally, helps to reduce the influence of measurement noise. For that reason, VFA is often used for statistical identification of unmeasured disturbances. The procedure basically consists in the reduction of dimensionality of the original data set by selecting the vectors where the variability is maximized as new axis for projection. Thus, it is possible to supervise the entire process by only analyzing these new projection structures called variability factors (VF) as it is shown in figure 1.

Mathematically, the VF vectors determined by the procedure spilt the original data matrix X into several orthogonal structures (see equation 1), according to their capacity to explain processes' data variance.

$$X = t_1 \cdot VF_1 + t_2 \cdot VF_2 + t_3 \cdot VF_3 + \cdots t_q \cdot VF_q \quad (1)$$

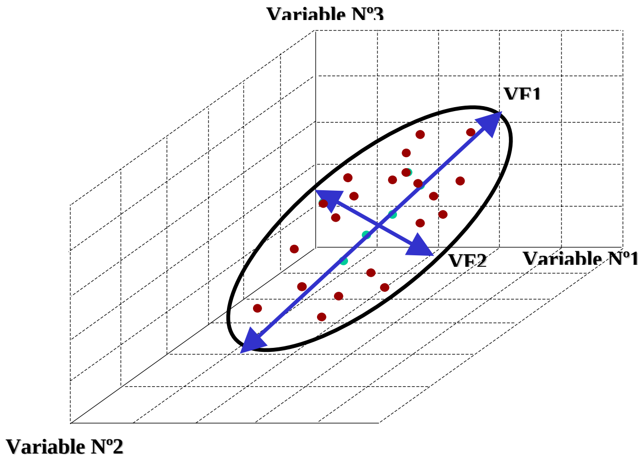

Figure 1. Variability Factor Analysis in a three-dimensional space

The score map (projection of measurement data into the orthogonal VF basis) often displays clouds of data points closer to the origin (which represents the desired operation point for the process) as it is shown in figure 2. If the projected cloud of points is far enough from the score map's origin, it could be considered as a change of the operation point of the entire plant or a fault condition that must be analyzed. Thus, it is interesting to consider a measure for the distance for the projected data from the center of the graph in order to detect disturbances and/or prevent them.

An statistical index, Hotelling's T_i, is in charge of this task by measuring the distance of the score from the pattern mean value according to several centered ellipses. The length of each ellipse's axis depends on the confidence of normal operation if the score is inside it.

When T_i is greater than 1, the corresponding score is outside the 99% confidence ellipse. If several scores appear to be outside this ellipse, the process shall be considered disturbed during those time instants, by an unknown - unmeasured - disturbance. Equation (2) shows the construction of the Ti index for fixed supervision period of time:

$$T_i^2 = t_i \lambda^{-1} t_i^T \quad (2)$$

Where λ^{-1} is the inverse of the covariance matrix for the projected scores t_i..

As the VF vectors are orthogonal, it is possible to obtain important conclusions by observing the contributions of each process' variable to them in the loading map (as figure 4). In fact, if two processes' variables are directly correlated their projections in the loading map must appear very closer.

Moreover, if two processes' variables contribute with complementary and uncorrelated information, their projections in the loading plot will appear very separated and they will only be related to the VF which most explains their corresponding variance. These mentioned characteristics of the loading map turns it into another powerful analysis tool for process analysis that will be considered.

RESULTS FOR PATTERN RECOGNITION IN SAG MILL

The VFA approach looks for the time instants where the SAG's operation adjusts to a recommended performance, generating an operation pattern for the process. Afterwards, the pattern is analysed by using the statistical test included in SCAN software. As a result, the main sources for processes' variability are identified and the principal components for the data set are determined.

Processes' characterization is based in the observation of its historic behaviour (which is registered by PI Systems). As the process consists of two mill lines, it was first desired a comparison between each line performance and a single normal operation pattern, which is shown in figure Nº2.

From figure Nº2 it is clear that each SAG mill (operating at the same conditions) was affected by an unmeasured disturbance.

As an analysis conclusion, and taking into account worker's opinion, it was estimated that the unmeasured variable affecting the SAG's performance was the deterioration of their revetments.

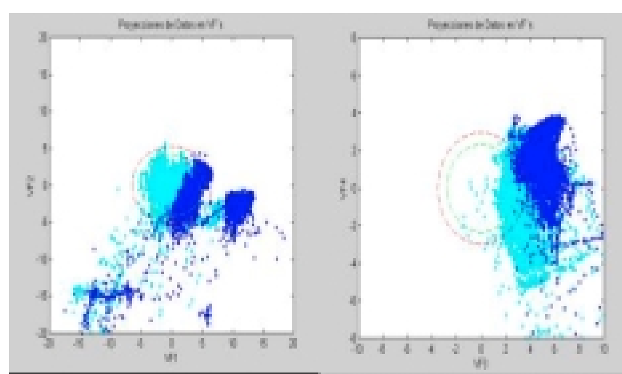

Figure 2. Influence of revetment's deterioration in SAG Nº1 and Nº2 data variance

Therefore, the final analysis considered the existence of two main operation patterns, one of them when the SAG's revetment is new and the other when it is deteriorated. The included process variables are ore and water feed, mill's speed, motor power (two for each SAG mill), pressures (feed and discharge) and pebbles' flow.

SAG Variability patterns, revetment changed

Four (4) variability factors are needed for an adequate supervision of each SAG mill. Those factors considerate approximately the 94% of the total dispersion of the pattern data set (see figure 3).

Figure 3. Explained variability percentage. Pattern for changed revetment

Each factor corresponds to a orthogonal linear combination of processes' variables. The 6% of variance non explained in the four (4) determined factors is related with a typical noise of the process. Thus, in normal operation, this percentage of error should not be an important source of variance. Additionally, from the loading maps (see figure 4), the following conclusions can be stated:

- The *main source* of variability is in the mill's speed, pressures and in the motor's power
- Ore feed, water feed and pebbles' flow give *additional and complementary* information about mill's operation
- There is more residual variability in the discharge pressure than in the feed pressure
- An excellent correlation is appreciated between the ore and water feed loops
- Any unusual increment of variability in the mill speed or in pebbles' flow can be considered as an important disturbance.

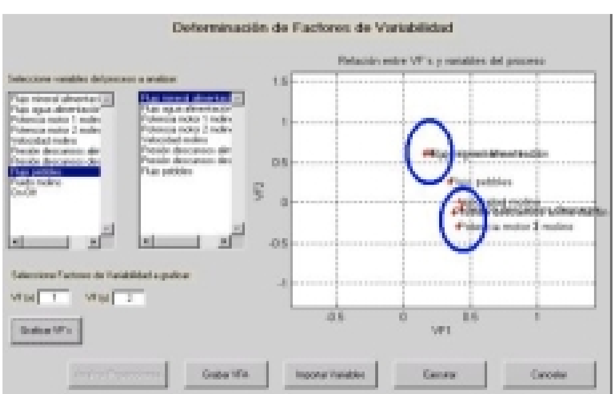

Figure 4. Loading plot. Pattern for changed revetment

The adequate variability range for both SAG mill lines 1 and 2 has been characterized through the same control ellipses. The axis of these ellipses (see figure 5) have been calculated to obtain a 95% and 99% of confidence for normal operation, if the scores are inside them.

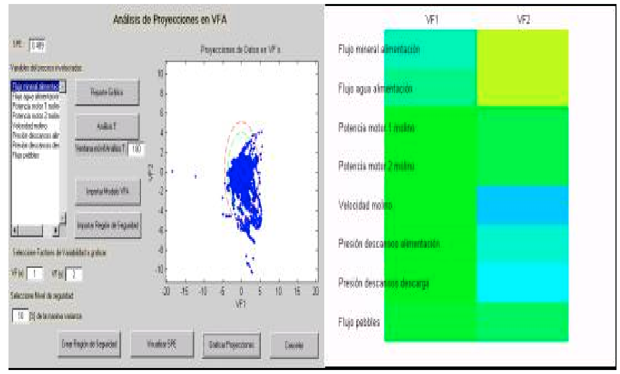

Figure 5. Scores plot (VF 1 and VF 2). Pattern for changed revetment

SAG Variability patterns, deteriorated revetment

Four (4) variability factors are needed for an adequate supervision of each SAG mill. Those factors considerate approximately the 97% of the total dispersion of the pattern data set. In this case, only the 3% of remaining variance is related with a typical noise of the process.

From the loading maps, the following conclusions can be stated (note that some of the typical characteristics of processes' behaviour have changed):

- The *main source* of variability is still in the mill's speed, pressures and in the motor's power
- Ore feed, water feed and pebbles' flow still give *additional and complementary* information about mill's operation
- There is more residual variability in pebbles' flow than in the previous condition
- There is less correlation between the ore and water feed loops than the appreciated in the previous condition
- There exists less residual variance in mill's power

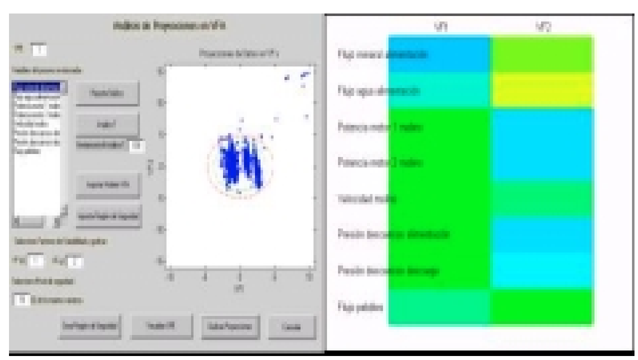

Figure 6. Scores plot. Deteriorated revetment pattern

As in the previous case, the adequate variability range for both SAG mill lines 1 and 2 has been characterized through the same control ellipses. The axis of these ellipses (see figure 6) have been calculated to obtain a 95% and 99% of confidence for normal operation, if the scores are inside them.

Detection of Disturbances

The use of the previously defined patterns for on-line processes' supervision, allows the easy detection of anomalies and, therefore, it helps the corrective decision-making procedure. As validation test, a simulation of an on-line supervision has been performed, for a data set corresponding to a "mine to mill" operation week (October, 15th –19th) where it was known that the mineral hardness index was higher than the usual.

The analysis of the score map for the mentioned data set (which considered in the new revetment pattern), clearly shows a variation for the cloud's position towards the statistical limits of the control ellipse (figure 7). Thus, the proposed projection method can be successfully used to an earlier detection of alterations in the mill's performance, such that changes in the mineralogy characteristics of the feed's ore in comparison to the expected for a normal operation pattern.

Additionally, the utilization of the T_i index (Hotelling's test) allows to analyse the operation's tendency by joining all the available information of the plant into one unique indicator. The increment of that index is directly associated to a deviation of the operation from the desired pattern. In this case, the unitary limit for the T_i index are equivalent to the limits of the ellipse drawn with two variability factors (see figure 7).

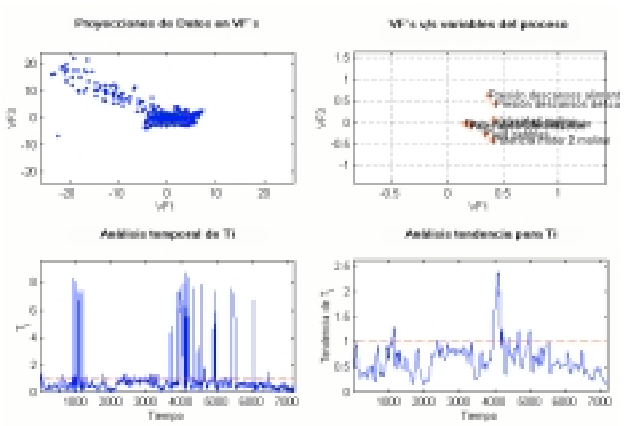

Figure 7. Disturbances in SAG mill Nº1. (Visualization through the T_i index)

According to the obtained results for the validation test (as it appears in figures Nº7 and Nº8), the T_i Hotelling's index has performed as an excellent predictor for future disturbances in the power or pressure signals in the SAG mills. Moreover, the use of this statistical tool has even sometimes achieved a prediction time horizon of several hours, before the expected disturbance in hardness characteristics of ore affected the behaviour of the operation.

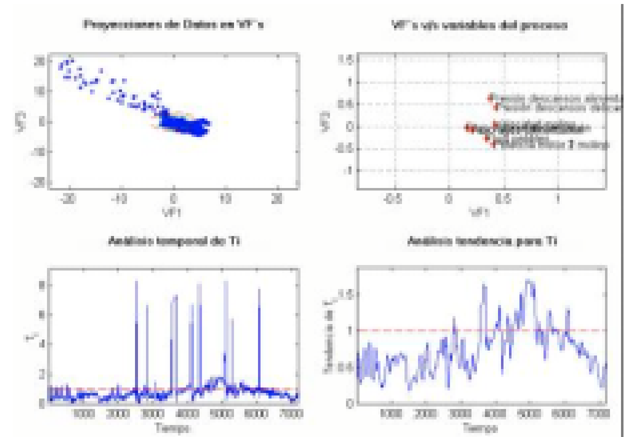

Figure 8. Abnormal condition in SAG Nº2 (Visualization through T_i index)

Additionally, and as a complementary result of the realized study, an intelligent score map has been elaborated foe each one of the identified operation patterns (SAG mills with new or deteriorated revetment), based on the configuration of the importance of each processes' variable to the VF Nº1 and Nº3. Dichos mapas establecen áreas geométricas relacionadas a situaciones concretas de operación en el molino.

In that way, the created geometric areas inside the score map, for the mentioned VF's, ease the visualization of the actual mill's behaviour and the earlier detection of abnormalities according to the position of the score cloud v/s time (see figure 9). This is of an enormous importance for the on-line process adequate supervision and the establishment of future control techniques.

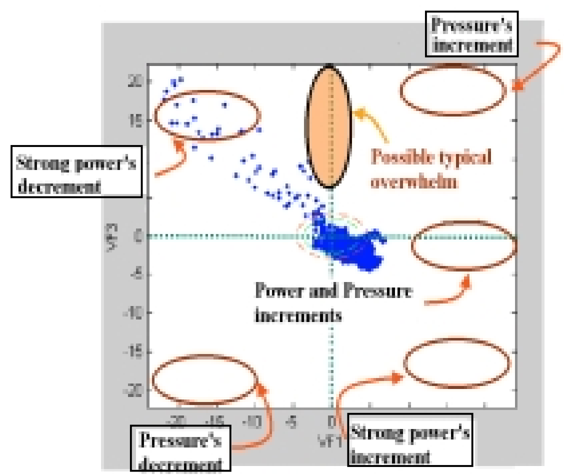

Figure 9. SAG mill's non-secure operation areas

FUTURE APPLICATIONS

The obtained results for the multivariate SAG mill's analysis motivated further investigation for other process' areas where the statistical test and modeling procedures available in CONTAC's SCAN software, can be applied.

This is the case of Collahuasi's flotation plant, where dynamic linear and non linear modeling modules are being applied in order to obtain a prediction for the copper concentrate grade as a function of process measurable variables.

The analysis is based on latent structures in order to reduce model's complexity. Thus, a NIPALS algorithm is used to find the structures of the processes' variable matrix ***X*** which are more correlated with the actual concentrate grade.

However much as linear inner as non-linear relationships are being used in order to explain the variations of the predicted variable (copper concentrate grade) around an operation point. Figures 10 shows the obtained results for variable prediction for 750[min] validation data set over (sample time 5[min]).

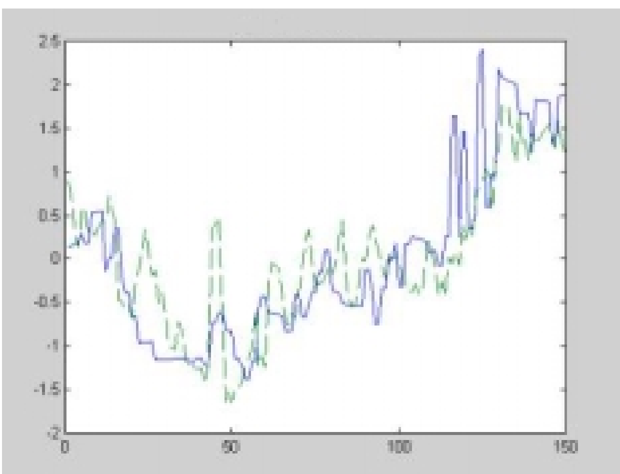

Figure 10. Copper Concentrate grade prediction

Several models were implemented through the analysis of data's latent structures t_i and u, such that:

$$\begin{aligned} X &= t_1 \cdot LS_1 + t_2 \cdot LS_2 + t_3 \cdot LS_3 + \cdots t_q \cdot LS_q \\ Y &= c_1 \cdot u \end{aligned} \qquad (3)$$

Each model considers a relationship between data' latent structures (LS_i). As the *LS* are linear combination of processes' variables, it is always possible to detect the relative importance of every actual variable once the *LS* model is done. For instance, figure 11 shows the inner relationship between the latent structures when linear or non-linear expressions are allowed.

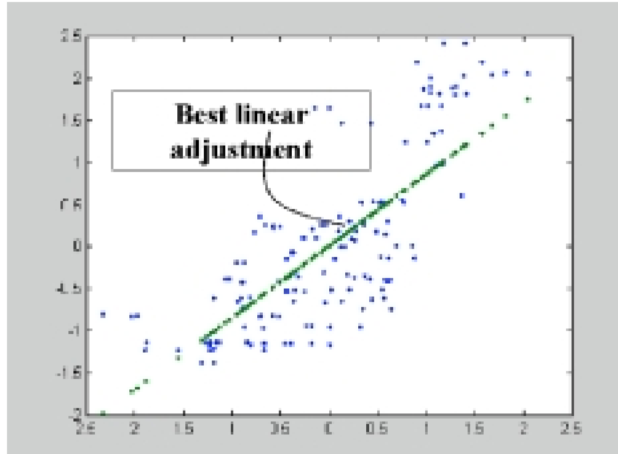

Figure 11. Linear inner relationships

It is important to mention that in the case of linear *LS* models, an explanation of 66.92% for the variations of the concentrate copper grade was obtained by using only the first latent structure. If non-linear components are allowed, then the percentage of explanation increases to 73.83% just using one *LS* (see figure 12).

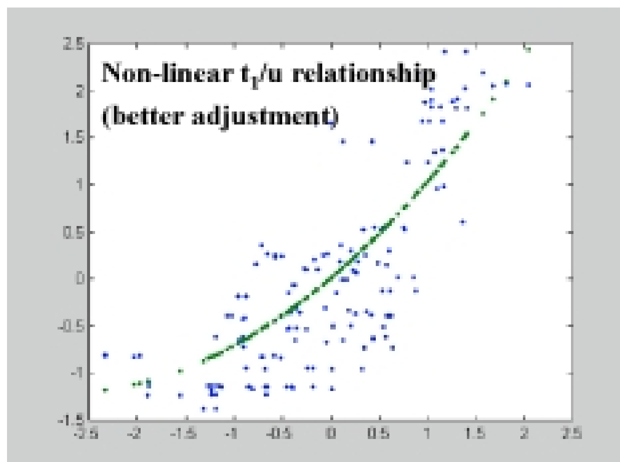

Figure 12. Non-Linear inner relationships

It is important to mention that in the case of linear *LS* models, an explanation of 66.92% for the variations of the concentrate copper grade was obtained by using only the first latent structure. If non-linear components are allowed, then the percentage of explanation increases to 73.83% just using one *LS*.

These excellent results clearly validate the approach and sustain the use of other interesting methodologies focused in the determination of the best *LS* needed to explain data sets (and therefore to explain the behavior of the process), and their use in order to perform even best results in the identified models.

The application of new compression technologies such as wavelet analysis for time-varying process modeling, neural networks for latent structures and fuzzy predictive control are among the following procedures that will be applied with the purpose of obtaining even better prediction results and variable explanation.

CONCLUSIONS

At this first stage, SCAN has demonstrated to be able to identify variability sources which were not initially considered in the characterization of a any operation's behaviour. Specifically, the software allowed to detect the influence of the SAG mill's revetment in its operation and to identify the established relationships between the main variability sources.

Additionally, the knowledge acquired in the pattern identification procedures was successfully incorporated in the on-line processes' supervision in order to detect and predict plant's disturbances. In all presented cases, the implemented methodology was able to help in the definition of the reference operation's patterns, off-line and on-line data analysis, main variability sources identification and in the definition of the existent correlation structures for the controlled process.

The future applications demonstrate a decisive potential in operation improvement, not only for the SAG mill lines but also in other important productive areas of the company, as flotation plants.

Therefore, the use for the information obtained through CONTAC Engineer's SCAN software and from worker's knowledge for the continuum improvement of processes' control and maintenance is drawn as an important line for the future work.

REFERENCES

[1] Hotteling

[2] PLS