INCORPORATION OF EXTERNAL INFORMATION INTO MULTIVARIATE PCA/PLS MODELS

Seongkyu Yoon[1] and John F. MacGregor[2]

Dept. Chemical Engineering, McMaster University, Hamilton Ontario Canada L8S 4L7
[1] yoons@mcmaster.ca
[2] macgreg@mcmaster.ca

Most processes are subject to changes of the operating conditions such as feed rate and composition, product specification, controller status, and so on. These operating condition changes and the resulting process variations largely contribute to the common-cause variation. Sometimes the common-cause variation disguises or distorts the relevant information to faults. If the fault effects are correlated, or small compared to the operating condition changes, then the correlation model would be likely insensitive to the faults. This difficulty can be avoided by using a new correlation model that includes only the process variations relevant for the fault detection and isolation. The new model called a hybrid correlation model is estimated by incorporating prior knowledge within an empirical correlation model. Various approaches to incorporating different types of prior knowledge into the correlation models, and using them for process monitoring and fault diagnosis are presented in this study. The proposed method is used for analyzing a real industrial dataset. It is shown that the hybrid correlation model outperforms the regular correlation model in analyzing the process abnormality and increases the sensitivity for process monitoring.

Keywords: Hybrid models, Prior knowledge, Fault detection and isolation, Monitoring, Principal Component Analysis, Partial Least Squares

1. INTRODUCTION

Multivariate statistical process control (MSPC) methods provide an alternative to the causal model-based fault detection and isolation (FDI). Using the correlation models, one can summarize process variations with a few of principal components. Since the correlation models can be more easily estimated than the causal relationships, the multivariate statistical models have been popularly used for FDI of large or complex processes.

Process operation data for building correlation models is collected when only common-cause variations are present. The common-cause variations include changes of the operating conditions such as feed conditions, product specification, controller status, and so on. When these operating conditions change, they affect many other variables and induce a lot of variation into the process data. These variations are usually described with a few relevant principal components. On the other hand, the correlation model that mainly explains the operating condition changes and the related variations is likely to be insensitive to faults. Especially when the faults are correlated, or small compared to these operating variations, the correlation model may lead to poor FDI. These difficulties can be avoided by using a preprocessed data in which the variations caused by the operating condition changes are removed. This data preprocessing can be done with prior knowledge of a process.

In many cases, one may have partial/incomplete

fundamental knowledge of the process even though complete fundamental models of a process are unavailable. These partially available process knowledge combined with the correlation models may be useful in interpreting questionable events occurring in processes (Yoon and MacGregor, 2000). If a given process has an undesired and disturbing influence and this process is characterized by one specific variable, the influence can be removed by data laundering through target rotation (Christie, 1996). The data laundering removes unnecessary variations caused by the specific variable from the data matrix. On the other hand, one can incorporate the external information into the model rather than removing (Takane and Shibayama, 1991; Gurden et al, 2000). The key concept is the separation of sources of variation which brings greater clarification as to the role of the external information within the measured data. Nomikos and MacGregor (1994) simply augmented the observation matrix with the estimates of the total conversion and the instantaneous rate of energy release and obtained the MPCA model. They claimed that one can increase the information contents of the PCA model such that a special variation related with the key process parameter becomes a dominant contribution in the principal component. This approach enriches the data with prior knowledge. The key parameters used by Nomikos and MacGregor (1994), and Hodouin et al. (1993) are usually calculated from material or energy balance equations. Those equations also can be used for the generation of the residuals from which a PCA model is estimated. Tong and Crowe (1995) applied this idea to data reconciliation and proposed the principal component tests on the material balance residuals for identifying the variable in gross error. Wachs and Lewin (1998) similarly presented a model based PCA. They showed that the model based PCA could deal with parametric and structural uncertainty of the models. Therefore, a hybrid modeling approach which combines available process information with empirical correlation models would offer significant benefits, and make it possible to avoid, or minimize the limitations of the correlation models.

This paper examines what kinds of external information can be used for the hybrid correlation modeling, how one can incorporate partially available prior knowledge into the correlation models, and how one can use the hybrid correlation models for process monitoring and fault diagnosis. As a unifying scheme of incorporating external information into correlation models an approach to hybrid correlation modeling is proposed. The focus is on integrating the available FDI resources under MSPC framework and maximizing the usage of the hybrid models for FDI.

The outline of the paper is as follows. In section 2, various types of prior knowledge are discussed. In the following sections, three types of hybrid correlation methods are presented: the decomposition based method; the augmentation based method; the blocking and transformation based method. In addition, the hybrid model which combines theoretical and empirical relationships is discussed. In section 4, a real industrial data is analyzed using the proposed scheme.

2. INCORPORATION OF PRIOR KNOWLEDGE

A hybrid correlation modeling here is defined as a way of incorporating external information into the multivariate statistical correlation models. Using the hybrid correlation model for FDI is defined as hybrid FDI. Incorporation can be done in a way of either adding process knowledge into the correlation model, or removing unnecessary process variation from the correlation model. The external information available at different levels includes:

- Prior knowledge on variables: disturbances, independent manipulated variables, and so on.
- Additional information such as controller setpoint, mode, controller output and so on
- Mass and energy balances that may be partially or fully known
- Location, or frequency characteristics of process variations

3. INCORPORATION OF PRIOR KNOWLEDGE INTO MULTIVARIATE CORRELATION MODELS

*3.1 Decomposition of **X** with row and column information*

Various types of external information can be represented in the form of either a column or a row for incorporation. Given the data matrix $\mathbf{X}$ ($n \times m$), assume an n by p ($\leq n$) observation information matrix, $\mathbf{G}$, and an m by q ($\leq n$) variable information matrix, $\mathbf{H}$. For example, the q rows of $\mathbf{H}'$ may represent q sets of material balance weights on the mass variables. A column vector in $\mathbf{G}$ may represent the observed values of a nuisance variable such as production rate which introduces a lot of extraneous variation into the variables in $\mathbf{X}$ (Fig. 1). Then the data matrix can be represented as follows (Takane and Shibayama, 1991):

$$\mathbf{X} = \mathbf{GMH}' + \mathbf{BH}' + \mathbf{GC} + \mathbf{E} \qquad (1)$$

where $\mathbf{M}(p \times q)$, $\mathbf{B}(n \times q)$, and $\mathbf{C}(p \times m)$ are matrices of coefficients to be estimated, and $\mathbf{E}(n \times m)$ a matrix of error components. The four terms in (1) explain the decomposed components of the original data matrix, $\mathbf{X}$. The first term pertains to what can be explained by both $\mathbf{G}$ and $\mathbf{H}$, the second term by $\mathbf{H}$, the third term by $\mathbf{G}$, and the fourth term by neither $\mathbf{G}$ nor $\mathbf{H}$. The first three terms are explainable terms by

external information and the last term is unexplainable. The estimates of **M**, **B**, and **C** are:

$$\hat{\mathbf{M}} = (\mathbf{G}'\mathbf{G})^{-}\mathbf{G}'\mathbf{X}\mathbf{H}(\mathbf{H}'\mathbf{H})^{-} \quad (2)$$

$$\hat{\mathbf{B}} = (\mathbf{I} - \mathbf{P}_G)\mathbf{X}\mathbf{H}(\mathbf{H}'\mathbf{H})^{-} \quad (3)$$

$$\hat{\mathbf{C}} = (\mathbf{G}'\mathbf{G})^{-}\mathbf{G}'\mathbf{X}(\mathbf{I} - \mathbf{P}_H) \quad (4)$$

where $(\mathbf{G}'\mathbf{G})^{-}$ and $(\mathbf{H}'\mathbf{H})^{-}$ are *g*-inverses of $(\mathbf{G}'\mathbf{G})$ and $(\mathbf{H}'\mathbf{H})$, respectively. $\mathbf{P}_G$ and $\mathbf{P}_H$ are orthogonal projection operators onto spaces spanned by the column vectors of **G** and **H**. They are calculated as follows:

$$\mathbf{P}_G = \mathbf{G}(\mathbf{G}'\mathbf{G})^{-}\mathbf{G}' \quad (5)$$

$$\mathbf{P}_H = \mathbf{H}(\mathbf{H}'\mathbf{H})^{-}\mathbf{H}' \quad (6)$$

Using orthogonal complement projectors of $\mathbf{P}_G$ and $\mathbf{P}_H$, $\mathbf{Q}_G (= \mathbf{I} - \mathbf{P}_G)$ and $\mathbf{Q}_H (= \mathbf{I} - \mathbf{P}_H)$, the fourth term in (1) can be calculated as follows:

$$\begin{aligned}\hat{\mathbf{E}} &= \mathbf{X} - \mathbf{G}\hat{\mathbf{M}}\mathbf{H}' - \hat{\mathbf{B}}\mathbf{H}' - \mathbf{G}\hat{\mathbf{C}} \\ &= \mathbf{X} - \mathbf{P}_G\mathbf{X}\mathbf{P}_H - \mathbf{Q}_G\mathbf{X}\mathbf{P}_H - \mathbf{P}_G\mathbf{X}\mathbf{Q}_H \\ &= \mathbf{Q}_G\mathbf{X}\mathbf{Q}_H\end{aligned} \quad (7)$$

By substituting the least squares estimates for the corresponding parameters in (1), the data matrix, **X**, is decomposed as follows:

$$\begin{aligned}\mathbf{X} &= (\mathbf{P}_G + \mathbf{Q}_G)\mathbf{X}(\mathbf{P}_H + \mathbf{Q}_H) \\ &= \mathbf{P}_G\mathbf{X}\mathbf{P}_H + \mathbf{Q}_G\mathbf{X}\mathbf{P}_H + \mathbf{P}_G\mathbf{X}\mathbf{Q}_H + \mathbf{Q}_G\mathbf{X}\mathbf{Q}_H\end{aligned} \quad (8)$$

The four terms in (8) are the estimates of the corresponding four terms in (1). Note that some of the terms in (8) may be zero when **G** or **H** is a square matrix of full rank (e.g., **G**= **I** or **H = I**).

Once the data matrix is decomposed according to the external information, principal component analysis may be applied to each component separately. In certain cases, some of the decomposed submatrices may be recombined for PCA. Since the components analyzed are associated with specific meanings of their own, PCA of the components may be more readily interpretable than direct PCA of the original data matrix.

Prior knowledge in form of columns, $\boldsymbol{X} = \boldsymbol{P}_G\boldsymbol{X} + \boldsymbol{Q}_G\boldsymbol{X}$; When prior knowledge is available in form of column(s) **G**, **H = I**, $\mathbf{P}_H = \mathbf{H}(\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}' = \mathbf{I}$, $\mathbf{Q}_H = \mathbf{0}$ (zero matrix). Thus (8) becomes

$$\mathbf{X} = \mathbf{P}_G\mathbf{X} + \mathbf{Q}_G\mathbf{X} \quad (9)$$

Where the first term represents the variation that can be explained by **G**, and the second term that cannot be explained by **G**. A PCA model can then be estimated with $\mathbf{Q}_G\mathbf{X}$ rather than with the original data matrix. This PCA model excludes the variations causing from the nuisance variables expressed by **G**. This scheme generalizes the method of data laundering via target rotation (Cristie, 1995). The data laundering can be considered the case that the target variable column, **y** is to be **G** in (8) and then $\mathbf{P}_y = \mathbf{y}(\mathbf{y}'\mathbf{y})^{-1}\mathbf{y}'$. The data matrix is decomposed into the two terms as $\mathbf{X} = \mathbf{P}_y\mathbf{X} + \mathbf{Q}_y\mathbf{X}$. where $\mathbf{P}_y\mathbf{X}$ is just the least squares projection of the columns of **X** onto **y,** i.e., $\mathbf{y}\hat{\boldsymbol{\beta}}$. The laundered data, $\mathbf{X}_l$ is nothing but $\mathbf{Q}_y\mathbf{X}$ that is a projection of the data matrix onto the orthogonal complement of the space spanned by the target variable. Orthogonal signal correction (Wold et al., 1998) can be understood as the case that one uses the response variable(s), **Y** as a target matrix. Then $\mathbf{X} = \mathbf{P}_Y\mathbf{X} + \mathbf{Q}_Y\mathbf{X}$. One can remove the information in **X** that is not predictive of, i.e., orthogonal to **Y**. The remaining step is to build the PLS model with **Y** and $\mathbf{P}_Y\mathbf{X}$.

Prior knowledge in form of rows, $\boldsymbol{X} = \boldsymbol{X}\boldsymbol{P}_H + \boldsymbol{X}\boldsymbol{Q}_H$; Here we assume certain row information on **X**, ie **H**, is available. Mass and energy balances belong to this type of information. Thus, **G = I**, $\mathbf{P}_G = \mathbf{G}(\mathbf{G}'\ \mathbf{G})^{-1}\mathbf{G}' = \mathbf{I}$, $\mathbf{Q}_G = \mathbf{0}$ (zero matrix). (8) becomes

$$\mathbf{X} = \mathbf{X}\mathbf{P}_H + \mathbf{X}\mathbf{Q}_H \quad (10)$$

where the first term represents the variation that can be explained by **H**, and the second term that cannot be explained by **H**.

The data reconciliation proposed by Tong and Crowe (1995) uses the mass balance equations, $\mathbf{X}\mathbf{B}' = \mathbf{0}$, where **B** is $r \times m$ with $r < m$. The residuals of reduced constraints for a linear steady-state process are defined in the matrix form, $\mathbf{R} = \mathbf{X}\mathbf{B}'$. PCA is estimated with the residual rather than the original observation matrix. Gross errors are identified with a principal component test. The principal component test for the data reconciliation can be interpreted differently with (8). Since the variable relationship, $\mathbf{H} = \mathbf{B}$' is available, the observation matrix is decomposed: $\mathbf{X}\mathbf{P}_B + \mathbf{X}\mathbf{Q}_B$ where, $\mathbf{P}_B = \mathbf{B}'(\mathbf{B}\mathbf{B}')^{-}\mathbf{B}$. Then the process variations are decomposed into two sources of what can be explained and what cannot be explained by the balance equations. The first term, $\mathbf{X}\mathbf{P}_B$ will explain the process variations based on the steady-state and deterministic variable relationships. The PCA model of $\mathbf{X}\mathbf{P}_B$ may be used for the data reconciliation as used. The second term, $\mathbf{X}\mathbf{Q}_H$ explains the process variations in which the deterministic variations are removed.

Row and column information, When both **G** and **H** information are available, all the terms in (8) become relevant. It is noted that row information can be converted into a column, and *vice versa*. For example, the material balance over the variables can be

represented as a residual and the calculated residual can be augmented into the observation matrix as a column, or used as **G** for the data decomposition.

3.2. Incorporation of knowledge by augmentation of the X matrix

When a material or energy balance is estimated based on a mechanistic model, the data matrix used for PCA or PLS models can be augmented with this estimated quantity as follows:

$$\mathbf{X}_{Aug} = [\ \mathbf{X} \mid \mathbf{X}_C\ ] \tag{11}$$

Where, $\mathbf{X}$ and $\mathbf{X}_C$ are the columns of the original data matrix and the estimated quantity, respectively. This augmented quantity enhances the diagnostic features of the multivariate statistical methods when faults occur at related sensors, or units. The same method can be used when a stochastic variable such as the heat exchanger fouling factor, or reaction impurity, is estimated from the available mechanistic models. The isolation of multiplicative faults can be simplified. One can not only overcome the difficulty in isolating multiplicative faults, but also minimize the necessity of process expertise in interpreting the contribution plots of the multiplicative faults. This is because the corresponding contribution of a more meaningful multiplicative parameter is highlighted instead of a group of measurements.

3.3. Using knowledge on process structure and signal frequencies

In very large processes involving many processing units with many variables in each unit, the number of potential errors or faults can be very large, making the diagnosis more difficult. Schemes for process monitoring using multivariate statistical projection methods such as PCA and PLS can be extended to situations where the processes can be naturally blocked into several subsections (MacGregor et al., 1994). The main advantage of such blocking is to allow for easier interpretation of the data by looking at smaller meaningful blocks and the relationship between blocks. The choice of blocking depends upon engineering judgments. Similarly, one can decompose the process variations over several frequency (scale) ranges and estimate the PCA models of each block, or of all the blocks together (Bakshi, 1998). Knowledge of the frequency characteristics of the faults are very crucial for the implementation of the multiscale decomposition approaches.

4. INDUSTRIAL APPLICATION

In this section, a real industrial data is analyzed. Discussed are the issues on how to improve process monitoring and fault diagnosis while minimizing the effects of the disturbances by using some of the proposed methods.

4.1 Process and fault description

Dataset was collected from a styrene monomer production plant. The full scale SM unit consists of a dehydrogenation section with styrene, product recovery system and a distillation section in which benzene-toluene, ethylbenzene and tars are separated from the styrene. In this study, only the dehydrogenation section is considered; its brief process flow diagram is shown in figure 2. Ethylbenzene is dehydrogenated into styrene monomer in the presence of the iron oxide catalyst. This reaction is endothermic and the necessary heat is supplied by the addition of high temperature steam. The reactor produces about 70% of the crude SM and 28 % of the unreacted EB with the condensing waters. The condensing water from the water/oil settling drum (WOSD) is stripped, and becomes fresh water for recycling use in steam boilers.

Hourly average values of 105 process variables were collected for 4 months from 25 Nov. 1999 to 20 Mar. 2000. The data includes normal and abnormal operations. The plant operation was shut down due to the malfunctions of the interface level indicator of the WOSD and the SM product pump at 9AM on Mar. 20, 2000. The interface level of the WOSD had indicated approximately 1% lower value than the setpoint since 5AM on the same day. The feedback control of the level controller was kept on for approximately 4 hours without operator intervention and resulted in a complete closure of the water disposal valve. In the mean time, the viscous material that was considered polymer and existed between the water and SM phases was taken to the SM side of WOSD. This caused a trip of the SM product pump since the viscous material blocked the pump suction line.

The malfunction of the water level sensor of WOSD was actually triggered by the accumulation of the viscous material within the level instrument. This viscous material is usually generated during the condensation of the reactor products at high temperature. It was understood that the viscous material could be monitored with the pressure difference between the reactor outlet and the recycle stream from the condenser stripper. However, the indicator of the pressure difference was not shown sensitive enough to detect the polymer accumulation due to the noisy characteristic of the pressure sensor and the disturbances. Accordingly, the operator intervention was not properly made at an early stage of the abnormal process operation.

4.2 FDI using regular PCA model

Among 105 variables, 52 key variables were selected for a postmortem analysis. With the original data, 18 principal components were identified with cross-validation as a stopping criterion and they explained 97.6 % of the total process variations. However, 8 principal components corresponding to those with eigenvalues are greater than 1 were used for the simplicity of the PCA model. They explained 88.1 % of the total variations. The first two principal components explain more than 55.0 % of the total variations. Figure 3(a) and (b) show the score plots of t_1 vs. t_2, and t_3 vs. t_4 for the training (1~2086) and testing (2087~2790) data. The training and testing data sets on the t_1 vs. t_2 scores space formulate several clusters. Interestingly, the dataset corresponding to the abnormal operation is found within the 99 % control region. On the other hand, the score plot of t_3 vs. t_4 in figure 3(b) properly indicates the abnormality of the plant operation.

Figure 4 shows the T^2 and DModX plots. DModX plot indicates the deviation from the common cause variations around 2240 T_S. It was found the levels of WOSD were mainly responsible for the abnormality at this time. The PCA monitoring plot exactly indicated the initial symptom of the catastrophic plant shutdown about one month in advance even though it was not reported by the plant personnel. Unfortunately, this plant was not equipped with plant monitoring system.

However, the deviation of the level sensors from the common cause variations does not directly explain how it resulted in the malfunction of the level sensor of the WOSD water side (WOSDWLvl) approximately one month later. It is noted in figure 4 that the T^2 monitoring plot violated the 99 % control limit around 2410 T_S. This corresponds to the data cluster outside of the control limit of the t_3 vs. t_4 score plot in figure 3(b). The variable contributions to T^2 at 2412 T_S confirmed a group of variables being responsible for the T^2 deviation at this time: the inlet temperature of the air fan cooler (AFCinTmp); the return line temperature from the stripper (STRrTmp); the bottom valve opening of the 1st separation drum (SD1WSVlv); the water side level of the W/O settling drum (WOSDWLvl). It was noted that the contribution of the ambient temperature (AMBTmp) was small. Thus, the variable contributions to T^2 provide an ambiguous clue to diagnosing the reason of the excessive variation on the hyper model plane.

4.3 FDI using hybrid PCA model

The effects of all the disturbances and their combinations were investigated before building the hybrid correlation model. The effects of the EB flow and the steam flow rates were considered the blocking factors that degrade the sensitivity of the correlation model. These two variables were set as the target variables (**G**) and their effects were regressed from the original data to leave $\mathbf{Q}_G\mathbf{X}$.

With the laundered data, 17 principal components were identified with the cross-validation as a stopping criterion and they explained 97.6 % of the total process variations. However, 10 principal components corresponding to those with eigenvalues are greater than 1 were used for the simplicity of the PCA model. They explained 82.2 % of the total variation. Figure 3(c) and (d) show the score plots of t_1 vs. t_2, and t_3 vs. t_4 for training (1~2086) and testing (2087-2790) data. The testing data sets on the t_1 vs. t_2 scores space clearly shows the abnormal operation. The dataset corresponding to the abnormal operation deviates from the 99 % control region. T^2 and DModX monitoring charts based on the laundered data ($\mathbf{Q}_G\mathbf{X}$) provide essentially the same fault detection as shown by the regular PCA model in figure 4. However, the interpretation of the fault contribution was greatly enhanced.

Figure 5 shows the variable contributions to T^2 at 2412 T_S with the preprocessed data. It indicates a group of variables which were responsible for the T^2 deviation at this time. Especially, the contribution of the ambient temperature (AMBTmp) is shown to be significant. This effect was blocked by the first two principal components of the regular PCA model but is now shown clearly because the blocking factors were removed. A close analysis indicates that the sudden ambient temperature increase is highly correlated with the situation that forced the styrene monomer to polymerize. It is also noticed that the compressed air flow rates to the second and the third stage reactors had been large contributions. Even though the hydrogenation reaction is endothermic, the temperature increase caused by the increased air rates resulted in the increased reactor outlet temperature since the conversion rate is not 100 %. Consequently, the reactor outlet temperature at the last stage experienced an abrupt temperature increase. This temperature increase magnified the effect of the ambient temperature increase. By the combined effect, the polymer accumulation was accelerated. Information given in figure 5 is consistent with this analysis result.

It is noted that there is potential improvement if one uses Wavelet filtering to remove unnecessary fluctuation caused by the ambient temperature while keeping the non-stationary part of the variation.

5. CONCLUSION

This research is to cope with practical situations in industries where the resources are partially available. Issues of dealing with external information for the correlation modeling have been addressed. Several existing methods were explained under the framework of the hybrid correlation modeling. Types

of external information used for the correlation modeling were clarified.

Using the orthogonal projections (or generalized LS) and external information, the observation matrix is decomposed into several terms. In general, it is decomposed into two terms: one including the effect of prior knowledge and the other term including the remaining effects. Depending upon the application, one can use any component of the decomposed terms. It was shown that various types of prior knowledge could be used to better interpret correlation models and achieve better performance. However, one needs to be cautious in using the prior knowledge for correlation modeling. It has to be confirmed that the target variables would be clearly unnecessary process variations for monitoring.

The hybrid correlation modeling and FDI have been assessed with the industrial dataset. It was shown that the proposed method provided better results when the available prior knowledge was properly used.

REFERENCES

Bakshi B., (1998), *AIChE J.*, **44**(7), 1596-1610

Christie, O. H. J. (1996), *Journal of Chemometrics*, **10**, pp.453-461.

Gurden, S. P., J. A. Westerhuis, S. Bijlsma, and A. K. Smilde, (2000), *J. Chemometrics*, **15**, pp.101-121.

Hodouin, D., J. F. MacGregor, M. Hou and M. Franklin. (1993), *CIM Bulletin*, **86** (975)

MacGregor, J. F., C. Jaeckle, C. Kiparissides and M. Koutoudi. (1994), *AIChE Journal*, **40**(5), 826

Nomikos, P and J. F. MacGregor. (1994). , *AIChE J*, **40**(8), p.1361-1375.

Takane, Y., and T. Shibayama (1991), *Psychometrika*, **56**(1), pp.97-120.

Tong, H and C. M. Crowe. (1995), *AIChE J*, **41**(7), p.1712-1722.

Wachs, A. and D. R. Lewin. (1998), *Proc. IFAC Symp. On Dynamics and Control of Process Systems*, Corfu, pp.86.

Yoon, S. and J. F. MacGregor. (2000), *AIChE J*., **46**(9), p.1813-1824.

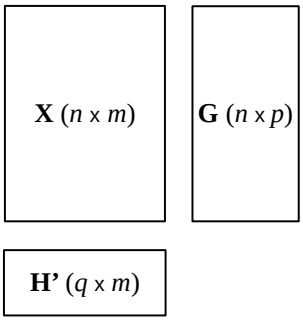

Fig. 1. Row and column information

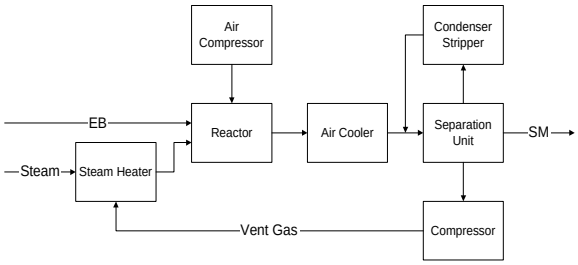

Fig. 2. Dehydrogenation of SM plant,

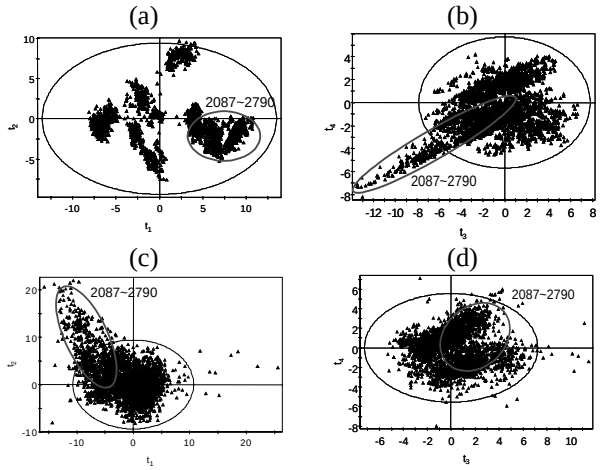

Fig. 3. Score plots; (a) t_1 vs. t_2 of regular PCA model; (b) t_3 vs. t_4 of regular PCA model; (c) t_1 vs. t_2 of hybrid PCA model; (d) t_3 vs. t_4 of hybrid PCA model

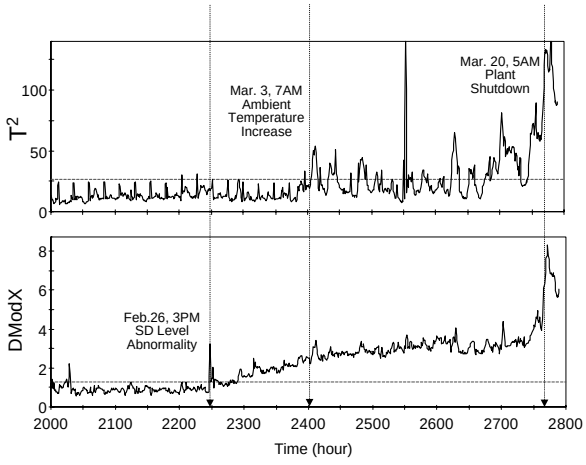

Fig. 4. Monitoring plots; (a) Regular PCA; (b) PCA model with prior knowledge

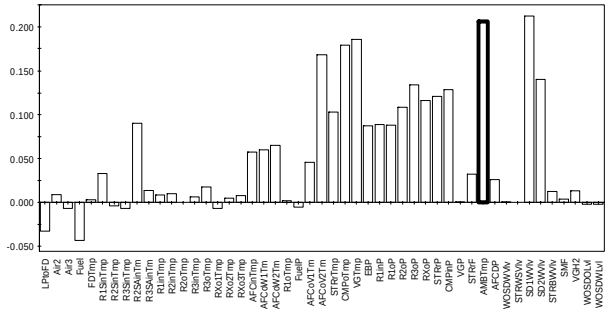

Fig. 5. Variable contribution plots