Available online at www.sciencedirect.com
ScienceDirect
Computers and Chemical Engineering 31 (2007) 419–425
Computers & Chemical Engineering
www.elsevier.com/locate/compchemeng

A systematic approach for soft sensor development

Bao Lin [a], Bodil Recke [b], Jørgen K.H. Knudsen [b], Sten Bay Jørgensen [a,*]

[a] *Department of Chemical Engineering, Technical University of Denmark, Lyngby 2800, Denmark*
[b] *FLSmidth Automation, Valby 2500, Denmark*

Received 15 November 2005; received in revised form 25 May 2006; accepted 25 May 2006
Available online 25 July 2006

Abstract

This paper presents a systematic approach based on robust statistical techniques for development of a data-driven soft sensor, which is an important component of the process analytical technology (PAT) and is essential for effective quality control. The data quality is obviously of essential significance for a data-driven soft sensor. Therefore, preprocessing procedures for process measurements are described in detail. First, a template is defined based on one or more key process variables to handle missing data related to severe operation interruptions. Second, a univariate, followed by a multivariate principal component analysis (PCA) approach, is used to detect outlying observations. Then, robust regression techniques are employed to derive an inferential model. A dynamic partial least squares (DPLS) model is implemented to address the issue of auto-correlation in process data and thus to provide smoother estimation than using a static regression model. The proposed methodology is illustrated through applications to a cement kiln system for estimation of variables related to product quality, i.e., free lime, and to emission quality, i.e., nitrogen oxides (NOx) emission. The case studies reveal the effectiveness of the systematic framework in deriving data-driven soft sensors that provide reasonably reliable one-step-ahead predictions.

Keywords: Multivariate regression analysis; Soft sensing; Robust statistics

1. Introduction

Soft sensors have been reported to supplement online instrument measurements for process monitoring and control. Both model-based and data-driven soft sensors have been developed. If a first principle model (FPM) describes the process sufficiently accurately, a model-based soft sensor can be derived (Prasad, Schley, Russo, & Wayne Bequette, 2002). However, a soft sensor based on detailed FPM is computationally intensive for real-time applications. Modern measurement techniques enable a large amount of operating data to be collected, stored and analyzed, thereby rendering data-driven soft sensor development a viable alternative. Application of standard multivariate statistical approaches to operating data may lead to model degradation due to contaminating outlying observations. Therefore, the objective of this paper is to present a systematic framework for the development of data-driven soft sensors based on robust statistical techniques.

A data-driven soft sensor is an inferential model developed from process observations. Early work on soft sensor development assumed that a process model was available. Joseph and Brosilow (1978) report an inferential model developed using a Kalman filter. In case the process mechanisms are not well understood, empirical models, such as neural network (Qin & McAvoy, 1992; Radhakrishnan & Mohamed, 2000) and multivariate statistical methods, are used to derive a regression model (Kresta, Marlin, & MacGregor, 1994; Park & Han, 2000). Multiple linear regression (MLR) suffers from numerical problems as well as degraded models when a data set is strongly collinear. Principal component regression (PCR), partial least squares (PLS) and canonical variate analysis (CVA) solve this issue by projecting the original process variables onto a low number of orthogonal latent variables (LVs).

An inferential sensor provides valuable real-time information that is necessary for effective quality control. Therefore, soft sensors have been widely applied for the estimation of quality measurements that are normally determined through infrequent sampling, and often with off-line analysis, such as the product composition of a distillation column (Zamprogna, Barolo, & Seborg, 2005) and particle size distributions in a

* Corresponding author. Fax: +45 45932906.
E-mail address: sbj@kt.dtu.dk (S.B. Jørgensen).

0098-1354/$ – see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2006.05.030

grinding circuit (Casali et al., 1998). Soft sensor development in batch/fed-batch processes has been review extensively in Dochain and Perrier (1997) and James, Legge, and Budman (2000). The 1990 Clean Air Act requires continuous emission monitoring devices equipped for NOx, SO_2 and CO_2 for certain large sources, such as industrial boilers and furnaces (Dong, McAvoy, & Chang, 1995; Qin, Yue, & Dunia, 1997). Although costly online analyzers have been installed at many plants, the emission measurement from a sensor may become unavailable due to instrument failure, maintenance or repair. Consequently, applications of multivariate soft sensors to emission monitoring have been increasingly reported (Dong et al., 1995; Qin et al., 1997). Recently, the Food and Drug Administration (FDA) introduces process analytical technology (PAT) into the pharmaceutical industry to ensure high and consistent product quality. An essential component of PAT is the real-time information of product properties. Soft sensors derived with multivariate statistical approaches can be powerful tools for pharmaceutical industry to facilitate process understanding, to monitor process operation and quality, to detect abnormal situations and to improve process reliability (Hinz, 2006; Kourti, 2006).

Online process measurements are often contaminated with data points that deviate significantly from the true values due to instrument failure or changes of operating conditions. Since outlying observations may deteriorate the regression model, robust statistical approaches have been developed to provide reliable results in the presence of abnormal observations. This paper presents a systematic approach using robust multivariate techniques to build a soft sensor from available process measurements. The application examples are the estimation of free lime and NOx emission for cement kilns.

The paper is organized as follows. First, a generic procedure is presented. Data preprocessing in Section 2 includes both univariate and multivariate approaches for detecting outlying observations. Robustified PCR and PLS approaches are described in Section 3. Section 4 contains illustrative applications on development of free lime and NOx soft sensors for cement kilns, followed by conclusions in Section 5.

2. Data preprocessing

Outliers are commonly defined as observations that are not consistent with the majority of the data (Chiang, Pell, & Seasholtz, 2003; Pearson, 2002a), including missing data points or blocks, and observations that deviate significantly from normal values. A data-driven soft sensor derived with PCR or PLS deteriorates in the presence of abnormal observations, resulting in model misspecification. Therefore, outlier detection constitutes an essential prerequisite step for design of a data-driven soft sensor.

Although missing data with regular patterns are not common in data from well-designed experiments, they often exist in operating data. For example, in the cement kiln system near zero drive current data simply correspond to a stop of cement kiln operation. During such a period, other kiln measurements obviously will not be reliable or meaningful. Therefore, a heuristic approach has been implemented in the proposed procedure to detect and handle missing data related to severe operating interruptions. Specifically, a template is defined by using the kiln drive measurement to identify missing observations. In case a small block (i.e., less than 2 h) of data is missing, interpolated values based on neighbouring observations will be inserted. If larger segments of missing data are detected, these blocks will be marked and not used to build a soft sensor.

Missing data do not always show a systematic pattern. A missing segment might exist in only one of the process measurements. In this case, such blocks can be replaced by using model-based interpolation methods that fill the missing gap with a model derived from the data set (Gupta & Lam, 1996; Nelson, Taylor, & MacGregor, 1996). Missing data are one type of outliers. The second type denotes abnormal operating conditions. For example, the malfunction of process equipment might cause a change in process measurements that may affect several successive samples. For detection of these outlying process observations, both univariate and multivariate approaches have been developed.

The *3σ edit rule* is a popular univariate approach to detect outliers (Ratcliff, 1993),

$$|x(i) - \bar{x}| > t \cdot \sigma \tag{1}$$

where $\bar{x}$ is the mean of the data sequence and $t = 3$ is the threshold. This method labels outliers when data points are three or more standard deviations from the mean.

Unfortunately, this procedure often fails in practice because the presence of outliers tends to inflate the variance estimation, causing too few outliers to be detected. The *Hampel identifier* (Davies & Gather, 1981) replaces the outlier-sensitive mean and standard deviation estimates with the outlier-resistant median and median absolute deviation from the median (MAD). The MAD scale estimate is defined as:

$$\mathrm{MAD} = 1.4826\,\mathrm{median}\{|x_i - x^*|\} \tag{2}$$

where x^* is the median of the data sequence. The factor 1.4826 is chosen such that the expected MAD is equal to the standard deviation σ for normally distributed data.

Fig. 1 shows 300 samples of SO_2 measurement from a cement plant during otherwise steady operating conditions. Due to harsh operating conditions, especially the flying dust within the kiln system that may block the measurement probe, the data segment of the gas analyzer measurement contains many outlying observations. It should first be noted that the mean value of the sequence is biased significantly from the nominal value, while the median value is close. In addition, outliers inflate the standard deviation such that most of the outlying observations are treated as normal data. With the threshold of $x_{\mathrm{Med}} \pm 3 \cdot x_{\mathrm{MAD}}$, the *Hampel identifier* identifies most outliers successfully.

A moving window *Hample filter* can be implemented with two tuning parameters: the threshold, t, and the width of the time window, K. The following choices are recommended (Pearson, 2002b): $2 \le t \le 5$, $3 \le K \le 5$, implying that 7–11 points are used for calculating the median and MAD of moving data window.

Since process measurements from chemical processes are not independent, detecting outliers using univariate diagnos-

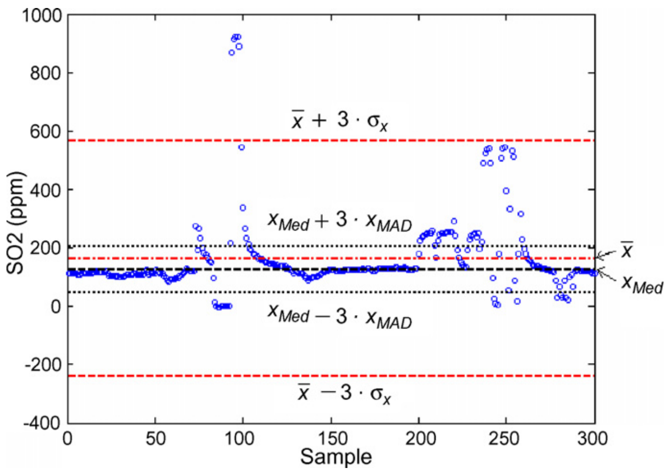

Fig. 1. Comparison between the standard *3σ edit rule* and the *Hampel identifier* on SO_2 data from a cement kiln. is shown with (–·–, red) and median with (– – –, black). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

tics is not sufficient, but may result in *masking* and *swamping*. *Masking* refers to the case that outliers are incorrectly identified as normal samples, while *swamping* is the case when normal samples are classified to be outliers. Three hundred samples of a kiln fuel flow rate measurement are shown in Fig. 2, which contain a short period of reduced fuel flow rate due to mechanic problems. A *Hampel identifier* is able to detect univariate outliers. Although the fuel measurements (marked by a circle) are within the bounds of *Hampel identifier*, the dynamics of transition effect of the kiln operation lasts longer than the period of low fuel supply. Therefore, the observations up to the sample 150 belong to an abnormal operation period. Such outliers can be effectively detected using a multivariate regression model representing the nominal operation condition, which is derived with the proposed approach in the following paragraphs.

Principal component analysis (PCA) is a multivariate analysis method that projects the data matrix to a lower dimensional space spanned by the loading vectors. The loading vectors corresponding to the k largest eigenvalues are retained to optimally capture the variations of the data and to minimize the effect of random noise. The fitness between data and the model can be

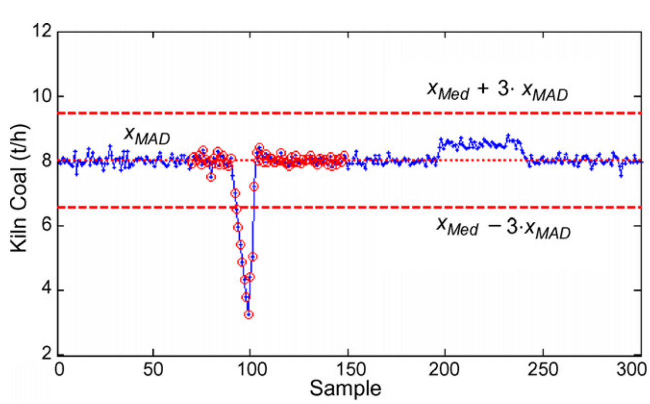

Fig. 2. Kiln fuel flow rate measurement with multivariate outliers that cannot be detected with a univariate approach.

calculated using the residual matrix and Q statistics that measures the distance of a sample from the space of the PCA model (Jackson & Mudholkar, 1979). Hotelling's T^2 statistics indicates how far the estimated sample by the PCA model is from the multivariate mean of the data; thus, this statistics provides an indication of variability within the normal subspace (Wise, 1991).

The combined Q and T^2 tests are used to detect remaining abnormal observations. Given the significance level for the Q and T^2 statistics, measurements with Q or T^2 values over the threshold are classified as outliers. In the proposed procedure, the significance level, α, has the same value in the two tests; however, finding a trade-off between accepting large modelled disturbances and rejecting large unmodelled behaviours for outlier detection clearly needs further investigation.

3. Robust statistics

Scaling is an important step in PCA. Since numerically large values are associated with numerically large variance, appropriate scaling methods are introduced such that all variables will have approximately equal weights in the PCA model. In the absence of a prior knowledge about relative importance of process variables, *autoscaling* (mean-centering following by a division with the standard deviation) is commonly used. Since both mean and standard deviation are inflated by outlying observations, *autoscaling* is not suitable for handling data which are especially noisy. The proposed procedure applies robust scaling approach before performing PCA (Chiang et al., 2003). This procedure replaces mean by median and the standard deviation by MAD.

There are two types of approaches for rendering PCA robust. The first detects and removes outliers using a univariate approach, then carries out a classic PCA on the new data set; the second is multivariate and is based on robust estimation of the covariance matrix. The proposed procedure uses the ellipsoidal multivariate trimming (MVT) approach (Devlin, Gnanadesikan, & Kettenring, 1981). This trimming method iteratively detects bad data based on the squared *Mahalanobis* distance:

$$d_i^2 = (x_i - x_i^*)^{\mathrm{T}} S^{*-1} (x_i - x_i^*) \qquad (3)$$

where x^* is the current robust estimation of the location and S^* is the robust estimation of the covariance matrix. Since the data set has been preprocessed with a *Hampel identifier*, 95% of data with smallest *Mahalanobis* distance are retained in the next iteration. Devlin et al. (1981) suggest that the iteration proceeds until the average absolute change in Fisher z transforms of the elements of the correlation matrix between two successive iterations is less than a predefined threshold, or the maximum number of iteration is reached. In this study, the iterative trimming procedure stops as late as the 10th iteration such that at least 60% of the data is retained for the estimation of the covariance matrix. Chiang et al. (2003) suggest the closest distance to center (CDC) approach where 50% observations with the smallest deviation from the center of the data are used to calculate the mean value. The CDC method is integrated in the initialization step such that the initial covariance matrix is not disrupted by outlying observations.

Principal component regression derives an inferential model with score vectors and the dependent variable. During the regression step, zero weights are assigned to outlying observations identified by the PCA model; a weight value of 1 is assigned to normal data. PLS is another multivariate statistical approach for relating input and dependent data matrices. The input data are projected onto a k-dimensional hyper-plane such that the coordinates are good predictors of dependent variables. The outlying measurements identified with the PCA model are also down-weighted before PLS analysis.

In summary, the systematic procedure of applying robust statistical techniques for soft sensor development consists of the following steps:

1. Handle missing data using a template defined with key process measurements.
2. Detect outliers with a univariate approach (*Hampel identifier*) followed by a multivariate approach (robust PCA) using Q and T^2 tests.
3. Derive regression model with weighted PLS.
4. Validate the soft sensor on independent process data.

The proposed procedure has been applied to many data sets collected from several cement kilns. Results from a few of the cases are given next.

4. Case studies

The rotary kiln is the most operationally complex and energy consuming equipment in the cement industry. The product quality of a cement kiln is indicated by the amount of free lime (CaO) in clinker. The direct off-line measurement is at most available with a time delay of about an hour. The measurement is also very sensitive to operating perturbations within the kiln system, which result in uncertain indication of the average quality. One indicator of the load on environmental quality is measured by the nitrogen oxides (NOx) emission. These oxides are formed in the cement kiln systems due to the high temperature in the burning zone. Traditional continuous emission monitoring is carried through analytical sensors, which are expensive and difficult to maintain (McAvoy, 2002). It is therefore desirable to develop soft sensors that are able to accurately predict NOx and free lime in real time. A soft sensor based on FPM is difficult to derive due to exothermic and endothermic reactions taking place in both solid and gas phases, as well as the large number of components involved. The proposed systematic procedure is employed to derive data-driven soft sensors in the sequel.

4.1. Case 1: free lime soft sensor

The operating data from a cement kiln log system are used to derive a soft sensor of free lime in the clinker. There are totally 19 process measurements available, including kiln drive current, kiln feed, fuel flow rates to calciner and kiln, plus several temperature measurements within the kiln system. Thirteen process

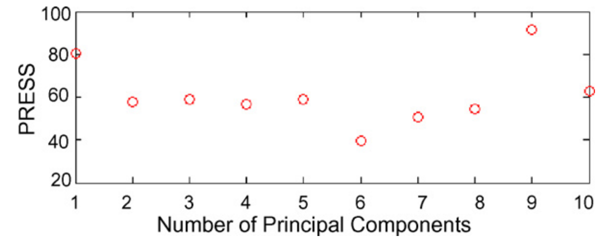

Fig. 3. PRESS of PCR model for CaO during validation period.

variables are selected as inputs based on process knowledge, as well as considering the reliability of process measurements. The standard measurements are logged every 10 min, whereas the laboratory analysis of free lime content of the clinker is logged approximately every 2 h. A data block of 12,500 samples for each of the standard measurements is selected in this study: 6500 samples for modelling and 6000 samples for validation.

One-step-ahead prediction residual sum of squared errors (PRESS) between the model and process measurement evaluated on validation data is used to select the number of principal components (PCs):

$$\text{PRESS} = \sum_{i=1}^{N_\text{V}} (\hat{y}(i) - y_\text{m}(i))^2 \quad (4)$$

where N_V is the total number of samples during the validation period. It is calculated only when a new laboratory measurement is available.

The PRESS of regression models derived with PCR and PLS are shown in Figs. 3 and 4, respectively. The PCR model with six PCs has the minimum PRESS (39.5). A second model is developed with a standard PCR approach that uses the autoscaled data and does not downweigh outlying observations. The PCR model with seven PCs achieves the minimum PRESS of 43.6, which is about 10% larger than that of a robustified PCR model. The PLS analysis shows a minimum of PRESS (42.7) for two latent variables, because PLS finds LVs that describe a large amount of variation in X and are correlated with dependent variables, Y, while the PCs in PCR approach are selected only on basis of the amount of variation that they explain in X.

Given the PCA decomposition, weights of 0 are assigned to abnormal points to downweigh these observations before a regression model is derived. Ninety-five percent significance level is commonly used for Q and T^2 tests. The lower the significance level, the higher the chance to reject outlying points. For the robust PCR model with six PCs, the significance level varies from 100 to 90%. As shown in Fig. 5, downweighing outlying

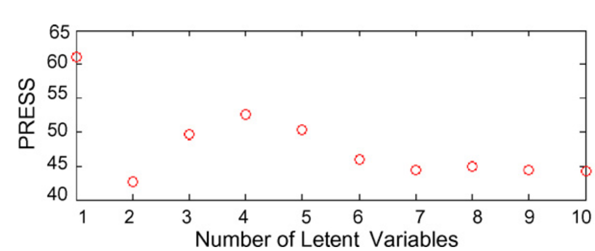

Fig. 4. PRESS of PLS model for CaO during validation period.

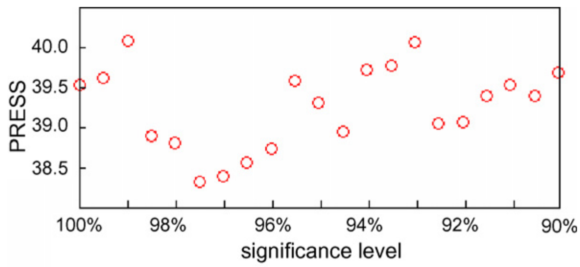

Fig. 5. PRESS of PCR model with six PCs for CaO with significance level varying from 100 to 90%.

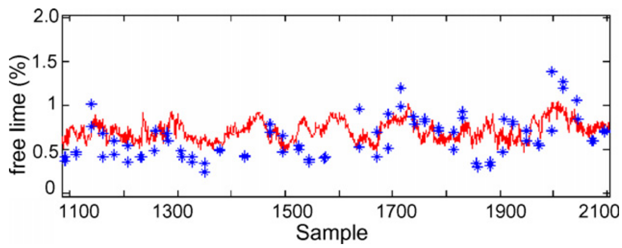

Fig. 6. Validation of robust PCR model (PRESS = 38.3) for CaO with six PCs ((*) laboratory measurements; (solid line) PCR).

observations is able to improve the predictability of the soft sensor. With the choice of an optimal significance level 97.5% for a PCR model, a minimum PRESS of 38.3 is obtained, which is about 12% less than that of a standard PCR model (43.6). The PRESS of a PLS model is reduced to 41.7 by downweighing the outlying observations detected with the PCR model, around 10% less than a standard PLS model (PRESS = 45.0).

Out of 6500 data points in the modelling block, 1289 samples are detected as outliers and downweighted in the PLS regression analysis. It should be noted that the number of outliers detected with the proposed approach depends on several factors. The quality of process measurements determines partially the number of outlying observations. Parameters of the robust PCA algorithm are the second factor. Since the covariance is estimated through an MVT procedure, the PCA model and outliers detected with it, are affected by the number of iterations and the ratio of measurements kept during the trimming procedure. Thirdly, the choice of the significance level is also influential, since it determines the threshold to detect multivariate outliers.

Comparisons of the PCR and PLS models with laboratory measurements during the validation period are shown in Figs. 6 and 7, respectively, where only 1000 samples during the validation period are shown. The PLS model is able to capture more relevant information than the PCR model with a smaller number of LVs. Although the robust PCR approach has a smaller PRESS than that of the PLS model, it is obtained at the cost of using four more principal components and thereby introducing a higher noise level in the regression model.

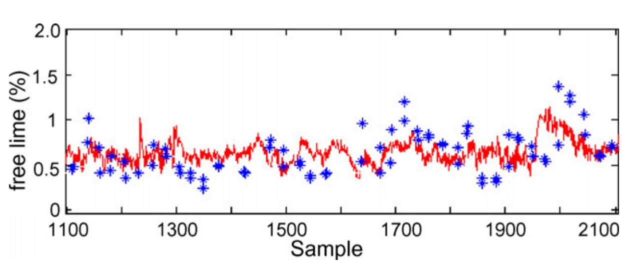

Fig. 7. Validation of robust PLS model (PRESS = 41.7) for CaO with two LVs ((*) laboratory measurements; (solid line) PLS).

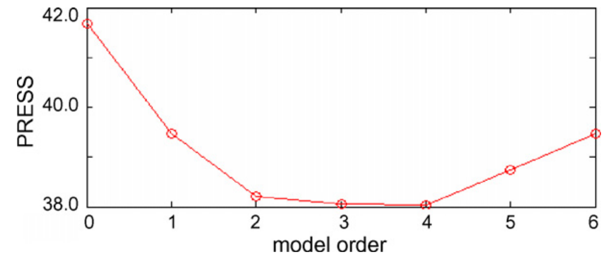

Fig. 8. PRESS of CaO soft sensor with a DPLS model (order from 0 to 6) evaluated on validation data with 6000 samples.

The fundamental assumption of the PLS approach is that the data matrix is not correlated in time. However, operating data commonly exhibit auto-correlation due to process dynamics. The PLS approach only constructs a linear static model from the data matrix, thereby it cannot reveal the dynamic relations between process variables. A dynamic PLS (DPLS) model is obtained by augmenting the original data block with time-lagged variables. The PRESS of the CaO soft sensor using a DPLS model with orders varying from 0 to 6 is shown in Fig. 8. DPLS with a model of order 4 achieves a minimum PRESS of 38.0. As shown in Fig. 9, the CaO soft sensor with a fourth order DPLS model is visibly smoother than a static PLS model (see Fig. 7). Including time-lagged terms initially recovers additional information and leads to a smoother prediction. However, including further time-lagged terms introduces additional noise into the model.

Although deviations are observed when fast dynamics occur in the process, the CaO soft sensor developed with a systematic robust statistical approach captures the slow changes and the trend of laboratory measurements reasonably well, which are important for process operation and control. This type of behaviour has been demonstrated on several cement kilns.

4.2. Case 2: NOx soft sensor

The operating data from a cement kiln log system are used to derive a NOx soft sensor. There are 43 process measurements that are sampled once per minute. A data block of 20,000 sam-

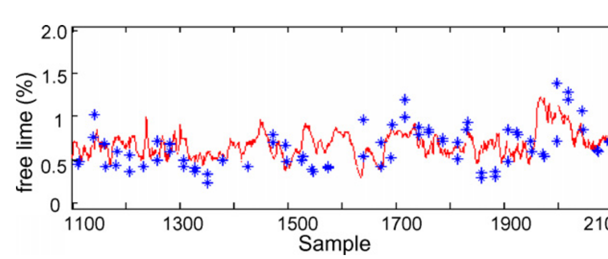

Fig. 9. Validation of CaO soft sensor (PRESS = 38.0) with a DPLS with of order 4 ((*) laboratory measurements; (solid line) DPLS).

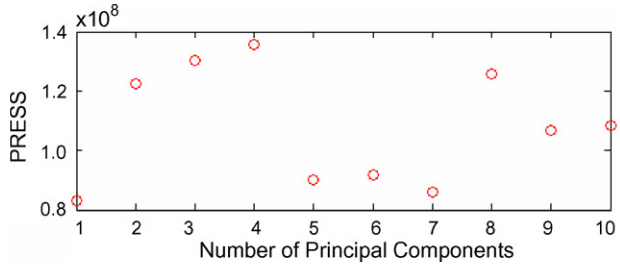

Fig. 10. PRESS of PCR model of NOx soft sensor during validation period.

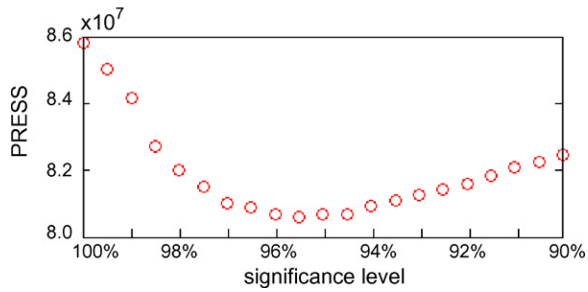

Fig. 11. PRESS of PCR model with seven PCs for NOx with significance level varying from 100 to 90%.

ples is selected: the last 10,000 samples for modelling and the first 10,000 samples for validation.

Fig. 10 shows the relation of PRESS versus the number of PC for a PCR soft sensor. Although PCR model with one PC has the numerically smallest PRESS, the model hardly captures process dynamics. The PCR model of seven PCs with the PRESS of 8.58×10^7 is employed to determine optimal Q and T^2 tests significance levels. As shown in Fig. 11, the minimum PRESS is obtained with an optimal significance level 95.5%, which detect 4155 outliers out of the 10,000 samples for modelling. It is observed that the optimal Q and T^2 tests significance levels achieve the trade-off between rejecting outlying points and essential process dynamics. In addition, optimal Q and T^2 tests significance levels also depend on the quality of the modelling data block.

The PRESS value of the NOx soft sensor developed by the standard PLS procedure (see Fig. 12) is 9.89×10^7. The performance is slightly improved by incorporating the univariate outlier detection procedure (PRESS = 9.48×10^7). Fig. 13 shows the relationship between the number of LVs and the PRESS of the NOx soft sensor from a PLS model following the RPCA outlier detection. The minimum PRESS of 8.40×10^7 is obtained with three LVs, which is around 15% less than that of

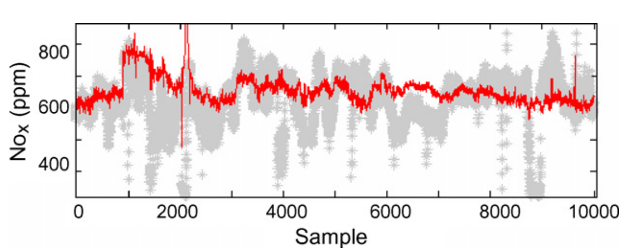

Fig. 12. Validation of a standard PLS model (PRESS = 9.89×10^7) for NOx with one LV ((*) online measurements; (solid line) PLS).

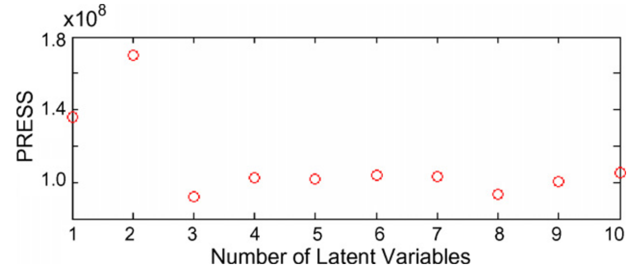

Fig. 13. PRESS of robust PLS model for NOx during validation period.

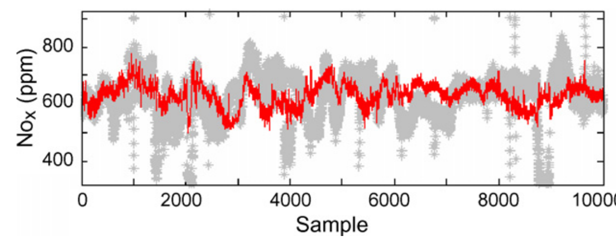

Fig. 14. Validation of robust PLS model (PRESS = 8.40×10^7) with two LVs for NOx ((*) online measurements; (solid line) PLS).

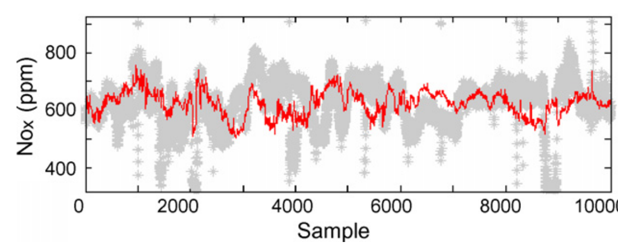

Fig. 15. Validation of NOx soft sensor (PRESS = 8.39×10^7) with a 10th order DPLS of two LVs ((*) online measurements; (solid line) DPLS).

the NOx soft sensor with a standard PLS model. As shown in Fig. 14, the performance of the NOx soft sensor is significantly improved by the univariate and multivariate outlier detection steps.

A dynamic PLS NOx soft sensor is also developed. The study reveals that the PRESS curve levels off after introducing two time-lagged input blocks. As shown in Fig. 15, the 10th order DPLS soft sensor of two LVs (PRESS = 8.39×10^7) provides much smoother prediction than a static NOx soft sensor. Compared to the PCR model with seven PCs, the PLS model demonstrates the advantage of obtaining a similar PRESS with a much lower number of LVs.

5. Conclusions

This paper presents a systematic approach to build data-driven soft sensors. Due to the low signal-to-noise ratio in operating data, data preprocessing is demonstrated to be an essential step in the framework. Robust statistical techniques are integrated to effectively extract process information in the presence of outlying observations. The proposed methodology is applied to predict free lime and NOx emission of cement kiln processes.

Both soft sensors are able to provide reasonably accurate prediction, providing complementary information to online gas analyzers and laboratory measurements. Smooth estimation is obtained with a dynamic model by introducing an appropriate number of time-lagged terms. More importantly, the real-time estimation of free lime shows potential to be used for quality control of the product.

The case studies demonstrate the effectiveness of outlier detection with a robust PCR approach and downweighing outlying observations to enhance the predictability of a regression model. The case studies also indicate the existence of an optimal significance level. However, the issues in finding the optimal Q and T^2 significance levels for regression model development and integrating the information from irregularly sampled off-line quality measurements into the weighting vector need further investigation.

References

Casali, A., Gonzalez, G., Torres, F., Vallebuona, G., Castelli, L., & Gimenez, P. (1998). Particle size distribution soft-sensor for a grinding circuit. *Powder Technology*, *99*(1), 15–21.

Chiang, L. H., Pell, R. J., & Seasholtz, M. B. (2003). Exploring process data with the use of robust outlier detection algorithms. *Journal of Process Control*, *13*(5), 437–449.

Davies, L., & Gather, U. (1981). The identification of multiple outliers. *Journal of the American Statistical Association*, *88*, 782–801.

Devlin, S. J., Gnanadesikan, R., & Kettenring, J. R. (1981). Robust estimation of dispersion matrices and principal components. *Journal of the American Statistical Association*, *76*, 354–362.

Dochain, D., & Perrier, M. (1997). Dynamical modelling, analysis, monitoring and control design for nonlinear bioprocesses. *Advances in Biochemical Engineering Biotechnology*, *56*, 147–198.

Dong, D., McAvoy, T. J., & Chang, L. J. (1995). Emission monitoring using multivariate soft sensors. In *Proceedings of the American control conference*.

Gupta, A., & Lam, M. S. (1996). Estimating missing values using neural networks. *Journal of the Operational Research Society*, *47*(2), 229–238.

Hinz, D. C. (2006). Process analytical technologies in the pharmaceutical industry: The FDA's PAT initiative. *Analytical and Bioanalytical Chemistry*, *384*(5), 1036–1042.

Jackson, J. E., & Mudholkar, G. S. (1979). Control procedures for residuals associated with principal component analysis. *Technometrics*, *21*(3), 341–349.

James, S. C., Legge, R. L., & Budman, H. (2000). On-line estimation in bioreactors: A review. *Reviews in Chemical Engineering*, *16*, 311–340.

Joseph, B., & Brosilow, C. (1978). Inferential control of processes. *AIChE Journal*, *24*, 485–509.

Kourti, T. (2006). The process analytical technology initiative and multivariate process analysis, monitoring and control. *Analytical and Bioanalytical Chemistry*, *384*(5), 1043–1048.

Kresta, J. V., Marlin, T. E., & MacGregor, J. F. (1994). Development of inferential process models using PLS. *Computers and Chemical Engineering*, *18*(7), 597–611.

McAvoy, T. (2002). Intelligent "control" applications in the process industries. *Annual Reviews in Control*, *26*(1), 75–86.

Nelson, P. R. C., Taylor, P. A., & MacGregor, J. F. (1996). Missing data methods in PCA and PLS: Score calculations with incomplete observations. *Chemometrics and Intelligent Laboratory Systems*, *35*, 45–65.

Park, S., & Han, C. (2000). A nonlinear soft sensor based on multivariate smoothing procedure for quality estimation in distillation columns. *Computers and Chemical Engineering*, *24*, 871–877.

Pearson, R. K. (2002a). Outliers in process modeling and identification. *IEEE Transactions on Control Systems Technology*, *10*(1), 55–63.

Pearson, R. K. (2002b). Scrub data with scale-invariant nonlinear digital filters. *EDN*, *47*(2), 71–78.

Prasad, V. V., Schley, M. M., Russo, L. P., & Wayne Bequette, B. (2002). Product property and production rate control of styrene polymerization. *Journal of Process Control*, *12*(3), 353–372.

Qin, S. J., & McAvoy, T. J. (1992). Nonlinear PLS modeling using neural networks. *Computers and Chemical Engineering*, *16*(4), 379–391.

Qin, S. J., Yue, H., & Dunia, R. (1997). Self-validating inferential sensors with application to air emission monitoring. *Industrial & Engineering Chemistry Research*, *36*, 1675–1685.

Radhakrishnan, V. R., & Mohamed, A. R. (2000). Neural networks for the identification and control of blast furnace hot metal quality. *Journal of Process Control*, *10*(6), 509.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*, 510–532.

Wise, B. M. (1991). *Adapting multivariate analysis for monitoring and modeling dynamic systems*. Ph.D. thesis. University of Washington.

Zamprogna, E., Barolo, M., & Seborg, D. E. (2005). Optimal selection of soft sensor inputs for batch distillation columns using principal component analysis. *Journal of Process Control*, *15*(1), 39–52.