Vermasvuori, Mikko; Vatanski, Nikolai; Jämsä-Jounela, Sirkka-Liisa

DATA BASED FAULT DETECTION OF THE ONLINE ANALYSERS IN A DEAROMATISATION PROCESS

Published in:
1st Workshop on Networked Control System and Fault Tolerant Control October 6-7th, 2005, Ajaccio, FRANCE

Published: 01/01/2005

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Vermasvuori, M., Vatanski, N., & Jämsä-Jounela, S.-L. (2005). DATA BASED FAULT DETECTION OF THE ONLINE ANALYSERS IN A DEAROMATISATION PROCESS. In *1st Workshop on Networked Control System and Fault Tolerant Control October 6-7th, 2005, Ajaccio, FRANCE* European Union.

Author's accepted manuscript, published in Workshop on Networked Control System and Fault Tolerant Control (NeCST), 2005, France

DATA-BASED FAULT DETECTION OF THE ONLINE ANALYSERS IN A DEAROMATISATION PROCESS

M. Vermasvuori, N. Vatanski, and S-L. Jämsä-Jounela

Aalto University
Laboratory of Process Control and Automation
Kemistintie 1. FIN-02150 Espoo, FINLAND

Abstract: Fault diagnosis methods based on process history data have been studied widely in recent years, and several successful industrial applications have been reported. In this paper a comparison of four monitoring methods, PCA, PLS, subspace identification and self-organising maps, for fault detection of the online analysers in a dearomatisation process is presented. The effectiveness of different statistical process monitoring methods in FDI of the online analysers is evaluated on the basis of a large number of simulation studies. Finally the results are presented and discussed.

Keywords: Fault detection, Input variable selection, Feature construction, Subspace identification.

1. INTRODUCTION

During the last decades, the ever-increasing competition caused by globalisation and the continuously tightening environmental restrictions are forcing industry to optimise their production processes. The operating conditions must be controlled precisely in order to ensure the high quality of the product. Process monitoring can help operators to detect and avoid undesired process conditions, thus decreasing the amount of produced off-spec products and improving the quality of the product. To this end, researchers have applied a range of process monitoring and fault detection methods in industrial processes. A large number of industrial applications with successful results have been reported (e.g. Komulainen *et al.* 2004 and Jämsä-Jounela *et al.* 2003) and reviewed, e.g. by Isermann and Ballé (1997) and Patton *et al.* (2000).

The monitoring methods can be classified into two main categories: methods based on first principle models, and methods based on models created with process history data. The data-based methods often utilize the reduction of the dimensionality of the original data set. Data compression is achieved by removing redundant information included in the variables while, at the same time, preserving as much as possible of the information in the original dataset. The most important statistical monitoring methods are based on principal component analysis (PCA) and partial least squares regression (PLS). Dynamic versions of PCA and PLS have been developed for continuous processes by (Ku *et al.* (1995), Chen *et al.* (1998), as well as recursive variants by Li *et al.* (2000) and Qin (1998). The PCA has been combined with wavelet analysis to deal with the autocorrelation of the data, and with neural networks to suit non-linear data sets. Most of the above-mentioned methods assume static process behaviour, but real industrial processes usually exhibit dynamic characteristics. The dynamic properties of processes can be modelled with state-space models, the system matrices of which can be determined from the process history data with subspace model identification methods (SMI). During the last decade a number of different methods have been proposed; canonical variate analysis (Larimore, 1990), N4SID (van Overschee. and de Moor, 1994), MOESP (Verhaegen 1994) and, more recently, a PCA based method by Wang and Qin (2002). Neural network architectures can be divided into three categories: feedforward, recurrent and self-organizing networks. According to Kohonen (1990), neural networks are the most applicable to classification and regression problems which do not need perfect precision. The self-organizing map, introduced by Kohonen, is an unsupervised neural network that has been compared to non-linear PCA, because it adapts to the structure of the data.

Before creating the statistical models for process monitoring, the dimensionality of the input variable set is usually reduced. The reduction can be done by selecting a suitable subset of the measured variables or by forming features from the original data. In

addition to reducing the dimension of the input, features are added that include useful process specific knowledge in the monitoring system. They can e.g. describe certain physico-chemical phenomena of the process, thus containing fundamental information about the processes in a compact form. Features can be constructed using clustering methods e.g. K-means and hierarchical clustering (Duda *et al.* 2001). The idea of distributional clustering (Pereira *et al.* 1993) is combined with an SVM-based classifier and applied to text categorisation problem in Bekkerman *et al.* 2003. Besides clustering, features can be constructed using linear transformations of the input variables (PCA/SVD, LDA), spectral transforms (Fourier, Hadamard), wavelet transforms, convolutions of kernels, or by applying simple functions to subsets of variables, e.g. products. (Guyon and Elisseeff, 2003). Sufficient dimensionality reduction, SDR, introduced in Globerson and Tishby (2003), is a new method for forming the most informative features by reducing the dimensionality of the data while losing as little information of the original dataset as possible.

In this paper a fault detection methods for analysers of a dearomatisation process is presented. Faults are detected using four data-based monitoring methods: PCA, PLS, subspace identification (SMI) and self-organising maps (SOM).

The paper is organised as follows: Chapter 2 is dedicated to introducing the dearomatisation process. In Chapter 3 the steps involved in setting up the FDI systems are presented. Chapter 4 consists of the results and discussion. The paper ends with a concluding section in Chapter 5.

2. DESCRIPTION OF THE DEAROMATISATION PROCESS

Dearomatisation processes are widely used in the oil refining industry. In these processes, aromatic compounds in the feedstock are removed by hydrogenation. The process consists of two trickle-bed reactors with packed beds of catalyst, a distillation column, a filling plate stripper, several heat exchangers and separation drums, and other unit operations. The process diagram of the LARPO process is presented in Fig. 1.

The cold, liquid feedstock fed to the unit is heated with streams from the two reactors in heat exchangers EA1 and EA2, and then fed to reactor DC1 together with hydrogen and recycle liquid. Exothermic saturation reactions in the first reactor remove most of the aromatic compounds when the catalyst is new, while most of the reactions occur in the second reactor when the catalyst is older and has been partly deactivated.

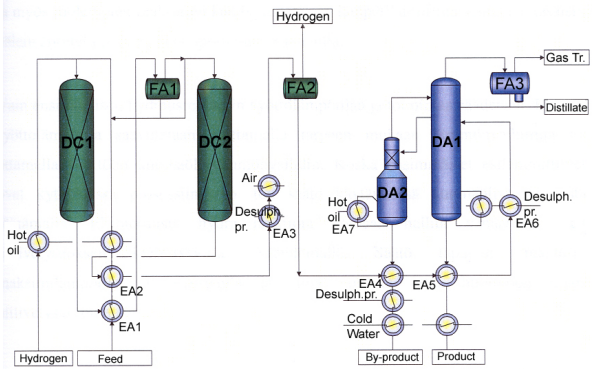

Fig. 1. The Naantali dearomatization process.

After dearomatisation in reactor DC1, the reaction product is cooled in heat exchanger EA1 and then fed to gas separation drum FA1. Gaseous and liquid reaction products are separated in the drum. Part of the liquid is circulated back to reactor DC1. The rest of the liquid, together with separated gas and fresh hydrogen, are fed to the second reactor DC2, where the aromatics level of the product drops to near zero.

After the second reactor, the reaction product is cooled in heat exchangers EA2 and EA3 and fed to the second gas separation drum FA2. Gas separated from the liquid mainly consists of unreacted hydrogen, which is recycled back to the first reactor, and the rest of the gas is removed. The separated liquid is heated with by-product and product streams in heat exchangers EA4 and EA5. Part of the liquid is further heated in heat exchanger EA6 in order to achieve the final temperature before the stream is fed to distillation column DA1.

The overhead of the column is cooled in a cooler and then fed to separation drum FA3, where the gaseous part is removed and the liquid is divided into reflux and distillate. The distillate consists of the lightest compounds of the reaction product. Heat exchanger EA6 produces heat for reboiling the bottom stream. From the upper part of the column DA1, a side stream is conducted to a stripper, which is heated with heat exchanger EA7. A by-product stream is drawn off from the bottom of the stripper. The non-aromatic main product is drawn off as the bottom product of the distillation column DA1 and cooled in heat exchanger EA5. The quality of the cooled product is measured online by flash point and distillation curve analysers. Laboratory measurements of the product are performed twice a day.

The six different types of feedstock used as inputs for the process are petroleum oil cuts with clearly differing properties. A change of feed type is made, on the average, once every 4 days. The dearomatisation process has no noticeable effect on the heaviest part of the distillation curve of the feedstock, but the properties of the lightest cuts are strongly affected by the distillation.

2.1 Analysers and analyser faults

According to a study by Kämpjärvi *et al.* (2006), most significant potential for improving the operation of the dearomatisation process lies in the earlier detection of analyser faults. If the fault could be detected earlier, then the corresponding changes to the operation could be made sooner and the quality of the end product could be kept within the production limits, thus improving the plants economical performance.

The analyzer faults are usually caused by one of the following problems; water contamination of the analysed sample, carbonisation of the flask of the distillation analyser, or fouling of the gas chromatograph. The first fault type causes the analyser result to drop abruptly and the others cause the output to drift slowly away from the correct value. These fault types are simulated in the test data.

3. SET UP OF THE FDI SYSTEMS

Industrial applications of process monitoring have usually been based on well-known and established methods such as PCA, PLS, SOM. The effectiveness of these methods for detecting analyser faults is researched in this study. The fourth method of the study, subspace model identification, is a newer method. In this method the measurement data are used to determine the matrices of state-space models. Here a simplified SMI algorithm proposed by Hyötyniemi (2001) is used. In this method, the state sequences are first formed using windows of past time series. The preliminary system states are then determined and the dimension of the states is reduced using PCA. Finally, the system matrices are solved in the least square sense.

3.1 Selection of the input variables for the FDI

The proper selection of process variables is a critical factor for a successful FDI system application. In this study selecting the outputs is straightforward, as only the variables of interest, in this case the measurements given by the online analysers, are chosen. Input variable selection, which includes the following systematic approach, is a more challenging task. First, the dimension of the complete set of the recorded process measurements is reduced by removing the variables that clearly do not affect the output. Next, the obtained set of variables is further compressed by removing variables, which correlate only slightly with the output. In this case study, almost the same set of 25 most strongly correlating variables were chosen to form the input variable set as in (Kinnunen, 2004). As a final step, process knowledge is added to the input variable set by constructing the calculated variables. The variables in this case study were temperature differences over reactors 1 and 2, temperature difference over the distillation column, and the ratio of bottoms and feed rates of the column. The first two features describe the speed of the reaction, and the last two features give information about the distillation section of the process. The final set of variables used in monitoring is listed in Table 1.

Table 1. The input and output variables.

Input variables – measurements
Flow rate of solvent feed to the reactor 1
Temperature of the feed to the reactor 1
Temperature of the feed from the reactor 1
Flow rate of recycled solvent from separation tank 1 to the reactor 1.
Flow rate to the reactor 2
Temperature of the feed to the reactor 2
Temperature of the feed from the reactor 2
Flow rate of recycled hydrogen into the reactor 1
Temperature of the feed to column 1 before heating
Temperature of the feed to column 1 after heating
Flow rate of the solvent to the column 1
Temperature difference over the heat exchanger that heats the feed to the column 1
Pressure at the top of the column 1
Temperature of the overhead of the column 1
Flow rate of the vapour distillate of the column 1
Flow rate of the liquid distillate of the column 1
Temperature of the liquid distillate of the column 1
Flow rate of the reflux in the column 1
Level of the bottom of the column 1
Three temperatures of the column 1 (on different plates)
Flow rate of reflux from the stripper to the column
Level of the tank 3
Temperature of the hot oil to heat up the stripper before the heat exchanger
Temperature of the hot oil to heat up the stripper after the heat exchanger
Flow rate of the bottoms of the stripper
Input variables – Features
Difference in temperatures across Reactor 1
Difference in temperatures across Reactor 2
Difference in top and bottom temperatures of the Distillation column
Ratio of bottoms and feed of the column
Output variables – analyser outputs
Bottoms product boiling point (0 % evaporated)
Flashpoint temperature of by-product
Flashpoint temperature of bottoms

3.2 Creating the monitoring models

After selecting the input variables, the monitoring models were created. The data used in this study were created with the PROSimulator software developed by Neste Jacobs Oy. A data set consisting of 2395 data points, covering a period of almost 40 hours, was used for modelling. Every hour the values of 1-3 variables were manipulated in order to create variance in the data. These manipulations initiated changes to process measurements similar to changes in real process data caused by normal operation actions. However, the feed type was not changed during the simulation. The variables used to create

Table 2 Variables manipulated during simulation

Composition of the solvent feed
Flow rate of the solvent feed
Composition of the hydrogen feed
Temperature of the hot oil to heat the boiler of the column 1
Flow of reflux from the separation tank 1 to the reactor 1
Temperature of hot oil to heat up the feed to the reactor 1
Fan speed and louvre opening of air cooling of solvent before the separation tank 2
Flow of reflux to column 1
Temperature of hot oil to heat up the stripper
Setpoint of pressure controller of the column 1
Setpoint of temperature controller of the flow from the stripper to the column 1

the changes are presented in table 2. The basic level control loops of the simulated process were closed, but the higher-level quality control was not active.

Principal Component Analysis. Three PCA models, one for each analyser, were constructed with the direct measurements and the corresponding analyser output as input variables. Another three models were created that also included the calculated variables. The models were constructed of four PCs that explained about 97 % of the total variance of the data.

Partial Least Squares. As with the PCA, two models were constructed for each of the three analysers, one using the direct measurements as inputs and the other also using the calculated variables. The output was the corresponding analyser result. The latent variables of the PLS were calculated with the NIPALS algorithm (Wold, 1975). The PLS models were created with the five latent variables that captured about 95 % of the variance of the input and 82 % of the output.

Subspace Modelling Identification. In creating the state-space models, the maximum dimension of the system matrices was set to five, and in the PCA data reduction phase three PCs were used. As before, two models were created for each analyser, one with the calculated variables and one without.

Self-organising Maps. In this study, two separate types of SOM were created. Both sets consisted of six models; one without and the other one with the calculated variables for each analyser. The first set was trained with the original autoscaled data including the analyser output. The second set of SOMs was trained with analyser results and PCA transformed data. After autoscaling, PCA was performed and four PCs explaining about 97 % of the variation were selected and the data were then projected into the PC space and used to train the second set of SOMs. During the testing, the analyser output was not used in determining the best matching unit (BMU). The BMU's value for the analyser output was compared to the current analyser output and, if the difference was larger than the alarm limit, an alarm was given.

4. TESTING, RESULTS AND DISCUSSION

After the training, the fault-detecting abilities of the FDI systems were evaluated using the simulated testing data set consisting of 1440 data points representing a period of 24 hours. Four faults to be detected were created on the analyser outputs: a large abrupt error (-50 % drop in temperature), a smaller abrupt error (-20 % drop in temperature), rapidly progressing drift (-10°C per hour) and slowly progressing drift (-5°C per hour). The abrupt faults correspond to the water contamination of the analysed sample fault in the real process and the drifting faults correspond to the carbonisation of the flask of the distillation analyser, or fouling of the gas chromatograph.

First, the PCA-based FDI system was tested. The Hotelling T^2 index was calculated and used to detect the faults. When the index alarm limit was set to 14, the PCA-based monitoring method was able to detect the abrupt faults correctly without delay. The first drifting type faults in initial boiling point and flashpoint temperature of the bottom product were detected after delays of 74 minutes. The system was, however, unable to detect the fault in the flashpoint temperature of the by-product. Also, the models were not able to detect any of the slowly progressing analyser faults. In addition, the models gave false alarms after the first and before the fourth fault. Figure 3A shows the faults in the testing data and the alarms given by the six PCA models. The x-axis represents time and the y-axis the monitoring results; values close to zero mean a normal process situation, and values near to one represent given alarms.

The PLS models were tested after the PCA models. The residuals between the model outputs and analyser results were calculated and, when they exceeded the limit value of 1.4, an alarm was triggered. The PLS-based fault detection system performed well and it was able to detect the abrupt faults with no delay, and incipient faults with delays of between 9 and 23 minutes. The first drifting fault was detected when the analyser error was 2-3°C and the second one before it was 2°C. The PLS-based monitoring system gave no false alarms. Figure 3B shows the alarms given by the method against the actual fault.

Next the SMI method was tested. The residuals between the models' and the analyser outputs were calculated, and an alarm was given by the system when the index exceeded the alarm limit. The best results of the subspace identification-based fault detection system were achieved with an alarm limit of 2.8. The system was able to detect the abrupt faults immediately they emerged, and incipient faults with delays of between 8 and 35 minutes when the deviation from the nominal was 1.5-3°C. The SMI model for bottom product's initial boiling point gave

few false alarms before fourth fault. Figure 3C shows the alarms given by the method against the actual fault.

Finally, the FDI systems based on the SOM models were tested. Both the SOM trained with normal measurements and the one trained with PCs were able to detect the abrupt faults immediately. In detecting the drifting type faults PCA pre-treated SOM was superior to the normal SOM. PCA-SOMs average detection delay of the slowly developing fault was 22 minutes compared to the normal SOMs average delay of 45 minutes. The difference in performance derives from the fact that residuals produced by the PCA-SOMs were much smoother than those of normal SOMs allowing the use of lower alarm limit of 4.5 instead of 5. Using the lower alarm limit with the normal SOMs would have produced a number of false alarms. The alarms given by the SOMs are shown in Figures 3D and 3E.

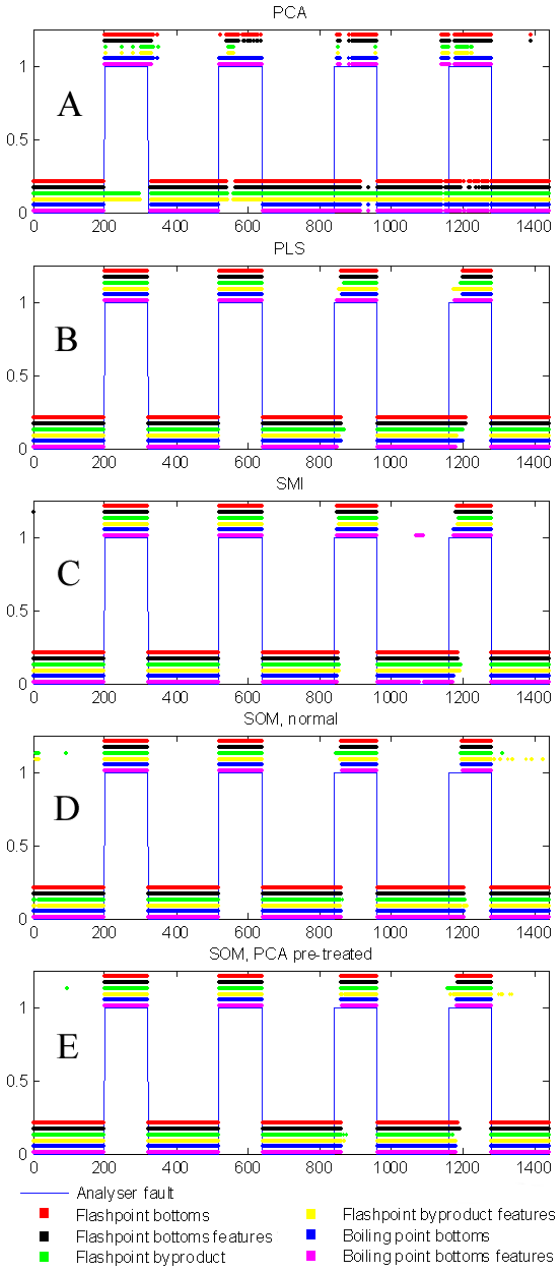

Fig. 3. States of the analysers and monitoring results of the PCA (A), PLS (B), SMI (C), Normal SOM (D) and PCA-SOM (E) models.

4.1 Analysis of the results and discussion

As the results in previous sections showed, the first two abrupt faults proved to be easy to detect. All the tested methods except PCA were able to give an alarm without delays, as shown in Table 2. The alarm signals were fired as soon as the faults occurred and lasted as long as the faults persisted. The subtle differences in the performance of the methods became visible only in detecting the slowly developing faults: On the average the fastest method to detect faults was PLS followed by slightly slower SMI. The PCA pre-treated SOM performed better than the normal SOM and was actually the fastest method to detect the slowly developing fault in by-product flashpoint temperature. The performance of the PCA was worst of the tested methods; its detection delays were longest and it was unable to detect some drifting faults at all.

The PCAs poor performance could be explained by the fact that the detected faults affected only one variable of the input variable set. Had the higher-level quality control been in use, the faults would have propagated into other variables and it would have been easier to detect. PLS on the contrary models both the input and output and performed well as the faults affected the whole set of output i.e. the analyser output. SMI was clearly the fastest method to detect the faults in the by-product flashpoint temperature. The better performance in detecting faults in the stripper with its additional delays and process dynamics is probably explained with SMI models' capability to include dynamic information. The SOMs were used here as nonlinear regression tools and the FDI results were good. The PCA treatment of the data done prior to the training of the maps proved to be beneficial as the PCA-SOMs detected the faults faster than normal SOMs. The training of the maps was also less time consuming as only five variables were used.

Using the process phenomena describing calculated variables together with the direct process measurement data in monitoring improved the results in some cases, but mostly it had only little effect, and on few occasions it actually made the delays longer.

The detection delays of all the methods and faults are presented in Table 3.

5. CONCLUSION

In this paper PCA, PLS, SOM and SMI methods for detecting analyser faults are tested and evaluated. The study showed that the PLS is the best of the tested fault detection methods in this case. Deviations of about 2-3°C between the analyser measurements and the nominal values were detected. This can be considered as a good result, as the nominal analyser measurements were not ideal, but included noise

Table 3. Detection delays

	Fault 1		Fault 2		Fault 3		Fault 4	
	Dir	Feat	Dir	Feat	Dir	Feat	Dir	Feat
Flashpoint temperature, bottoms product								
PCA	0	0	23	23	74	74	-	-
PLS	0	0	0	0	9	9	22	22
SMI	0	0	0	0	11	11	27	27
SOM norm	0	0	0	0	20	20	43	43
SOM PCA	0	0	0	0	20	20	27	31
Flashpoint temperature, by-product								
PCA	100	101	24	24	-	-	-	-
PLS	0	0	0	0	15	14	22	23
SMI	0	0	0	0	14	14	35	31
SOM norm	0	0	0	0	26	18	48	51
SOM PCA	0	0	0	0	38	31	11	20
Initial boiling point temperature, bottoms product								
PCA	0	0	0	0	74	74	-	-
PLS	0	0	0	0	9	9	21	21
SMI	0	0	0	0	9	8	15	13
SOM norm	0	0	0	0	21	22	44	43
SOM PCA	0	0	0	0	20	21	22	22

with amplitude of 1°C. SMI and SOM also proved to be very useful tools in FDI. The performance of the PCA was poor; either the faults were recognised only after their presence was obvious and/or a large number of false alarms were generated. It is worth noting, that the data used in the study was simulated, and even though care was taken to make the data resemble actual process data, the results presented here are probably better than would have been if real process data had been used.

ACKNOWLEDGEMENT

This research has been conducted as a part of the Networked Control Systems Tolerant to Faults (NeCST) project IST-004303 that is partially funded by the EU. The authors gratefully acknowledge the support.

REFERENCES

Chen, G.; McAvoy, T.J.; Piovoso, M.J. (1998), A multivariate statistical controller for online quality improvement, *Journal of Process Control*, **8**, pp. 139-149.

Bekkerman, R., El-Yaniv, R., Tishby, N., and Winter, Y., Distributional Word Clusters vs. Words for text classification, *Journal of Machine Learning Research*, **3**, pp., 1183-1208.

Duda, R., Hart, P., and Stork, D., *Pattern classification*, John Wiley & Sons, USA, 2nd edition, p. 680, 2001.

Guyon, I., and Elisseeff, A., (2003) An Introduction to Variable and Feature Selection, *Journal of Machine Learning Research*, **3**, pp. 1157 - 1182.

Hyötyniemi, H. (2001), Multivariate Regression - Techniques and Tools, *Helsinki University of Technology internal Report 125*, p. 159

Isermann, R.; Ballé P. (1997). Trends in the Application of Model-based Fault Detection and Diagnosis of Technical Processes. *Control Engineering Practise*, **5**, pp. 709-719.

Jämsä-Jounela, S-L., Vermasvuori, M., Endén, P., Haavisto, S., (2003) A process monitoring system based on the Kohonen self-organizing maps,*Control Engineering Practice*,**11**,pp.83-92.

Kinnunen, M., (2004) *Integration of first-principles process models into data-based monitoring*, MSc thesis, TKK, p. 96 (In Finnish)

Kohonen, T. (1990). The Self-Organizing Map, *Proceedings of the IEEE*, **78**, pp. 1464-1480.

Komulainen, T., Sourander, M., Jämsä-Jounela, S-L., (2004) An online application of dynamic PLS to a dearomatization process. *Computers & Chemical Engineering*, **28**, 11, pp. 2611-2619.

Ku, W.; Storer, R.H.; Georgakis, C. (1995), Disturbance detection and isolation by dynamic principal component analysis, *Chemometrics and Intelligent Laboratory Systems*, **30**, pp. 179–196.

Kämpjärvi, P., Sourander, M., Komulainen, T., and Jämsä-Jounela, S-L., (2006) Fault detection and isolation of online analyzer for an ethylene cracking process, *Control Engineering Practice*, submitted.

Larimore, W., (1990) Canonical variate analysis in identification, filtering, and adaptive control. *Proceedings of 29th IEEE conference on Decision and Control*, Honolulu, Hawaii, pp. 596-604.

Li, W., Yue, H.H., Valle-Cervantes, S.; Qin S.J. (2000), Recursive PCA for adaptive process monitoring, *Journal of Process Control*, **10**, pp. 471-486.

Patton, R.J.; Uppal, F.J.; Lopez-Toribio, C.J. (2000). Soft Computing Approaches to Fault Diagnosis for Dynamic Systems: A Survey. *Proceedings of IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes*, Budapest, Hungary pp. 298-311.

Qin, S.J. (1998), Recursive PLS algorithms for adaptive data modeling, *Computers & Chemical Engineering*, **22**, pp. 503-514.

van Overschee,P. and de Moor, B.,(1994) N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems, *Automatica*, **30**, pp. 75-93.

Verhaegen, M., (1994) Identification of the deterministic part of MIMO state space models given in innovations form from input-output data, *Automatica*, **30**, pp. 61-74

Wang, J. and Qin, J., (2002) A new subspace identification approach based on principal component analysis, *Journal of Process Control*, **12**, pp. 841-855.

Wold, H. (1975) Path models with latent variables: The NIPALS approach. *Quantitative Sociology: International perspectives on mathematical and statistical model building,* pp. 307-357.