Jämsä-Jounela, Sirkka-Liisa; Komulainen, Tiina

Integrating Process Indicators with Monitoring Method Hybrids

Published in:
Dynamics and Control of Process Systems (DYCOPS 7), Cambridge, Massachusetts USA, July 5-7, 2004

Published: 01/01/2004

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

Please cite the original version:
Jämsä-Jounela, S.-L., & Komulainen, T. (2004). Integrating Process Indicators with Monitoring Method Hybrids. In S. Shah, & J. F. MacGregor (Eds.), *Dynamics and Control of Process Systems (DYCOPS 7), Cambridge, Massachusetts USA, July 5-7, 2004* International Federation of Automatic Control (IFAC).

Author's accepted manuscript, published in 7th International Symposium on Dynamics and Control of Process Systems, 2004, IFAC, USA

INTEGRATING PROCESS INDICATORS WITH MONITORING METHOD HYBRIDS

Sirkka-Liisa Jämsä-Jounela, Tiina Komulainen

Laboratory of Process Control and Automation, Aalto University, P.O.BOX 6100, FIN-02015, Finland

Abstract: In this article the benefits of process monitoring are discussed. The different aspects of data-based monitoring methods and their combinations are reviewed. The role of process indicators is discussed, and the concept of combining a monitoring method library with the process indicators is presented. Finally, the benefits of process indicators and monitoring method hybrids are demonstrated with five industrial process monitoring applications created by the Laboratory of Process Control and Automation, Helsinki University of Technology.

Keywords: monitoring, process indicators, industrial applications.

1. INTRODUCTION

Increasing international competition in the process industries is emphasizing the importance of product quality management; the quality of production is increased by better monitoring and control of the processes. According to several studies, inadequate management of abnormal situations causes annual losses of 20 billion dollars for the petrochemical industry in the USA. This, together with many other similar estimates, has led to extension of the field of diagnostic methods during the last decade. Since then, hundreds of successful applications of different monitoring methods have been reported. Diagnostic methods have proved to be useful and effective in industrial use, for instance in the chemical, mineral and metal, pulp and paper industries. Online process monitoring with fault detection can provide stability and efficiency for a wide range of processes. The current challenge is the automation of abnormal event management using intelligent systems, thereby providing operators with assistance in the most pressing area of need. This has also been viewed as the next major milestone in control systems research by the people working in the process industries. (Lennox and Sandoz 2002)

The diagnostic methods can be divided into model-based and data history-based methods. In the model-based approaches a priori knowledge about the process is needed, whereas in process history-based methods only the availability of a large amount of historical process data is the limiting factor.

The problem in model-based monitoring methods is that it is not always possible to construct dynamical process models that describe the process to an adequate degree of precision. Therefore most of the industrial applications are process history-based methods. The process history-based approach often lacks the ability to capture the process features and, in order to achieve good results, complex methods are required when using direct process measurements.

There are two ways to complement the weaknesses of the data-based monitoring methods. First, different monitoring methods possessing the desired properties can be combined, and secondly, the data-based and model-based approaches can be combined using process indicators as the variables of the monitoring applications.

In the following chapters the different monitoring methods and process indicators are discussed. A concept in which these two approaches are combined in advance in order to develop an operator support system is then presented, and finally some industrial cases using this concept are briefly reviewed.

2. COMBINING MONITORING METHODS

According to Dash and Venkatasubramanian (2000), diagnostic methods can be divided into model-based and process history-based methods, as shown in Figure 1. A fundamental understanding of the functionality of the studied process is necessary for model development in model-based methods. The process history-based approach, which is especially suitable for process monitoring purposes, requires a large amount of data in order to capture and model the features of the process. The history-based models can be subdivided into qualitative and quantitative models. The basis of qualitative models consists of rule-based and trend modeling methodologies, whereas the quantitative methods are divided into statistical and non-statistical, neural networks based on pattern recognition models (Venkatasubramanian et al. 2003).

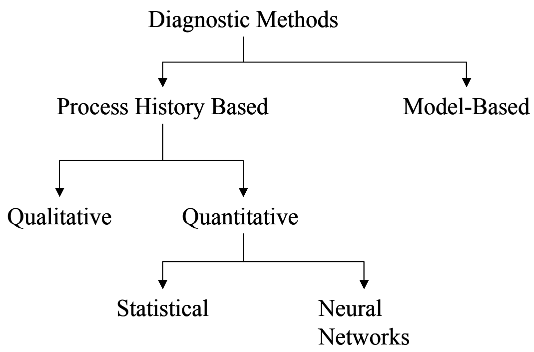

Fig. 1. Categories of the diagnostic methods. (Dash and Vankatasubramanian 2000)

Common features of the statistical methods used are their ability to reduce correlations between variables, compress data, and reduce the dimensionality of the data. These characteristics enable efficient extraction of the relevant information and analysis of the data. The most important statistical monitoring methods are based on principal component analysis (PCA) and partial least squares regression (PLS) (Wold et. al. 2001). Dynamic methods of PCA and PLS consider the dynamic nature of the monitored process and analyze both cross-correlation and auto-correlation. The dynamic methods are especially suitable for continuous processes with long time delays and varying throughputs on process variables (Ku et al. 1995, Chen et. al. 1998). Recursive methods for PCA and PLS have been proposed by Li et al. (2000) and Qin (1998). The recursive methods are especially suitable for time-dependent processes with slow changes. Multi-scale principal component analysis (MSPCA), a combination of PCA and wavelet analysis, removes the autocorrelations of variables by means of wavelet analysis, and eliminates cross-correlations between variables with PCA (Misra et al. 2002). The method is suitable for processes with auto-correlated measurements and time-varying characteristics. Nonlinear principal component analysis (NLPCA) is a combination of neural networks and PCA. Dong and McAvoy (1996) proposed an NLPCA method, which integrates a principal curve algorithm and neural networks. The idea of this method is to fit curves instead of lines to the data with the help of a feedforward network.

Neural network architectures are usually divided into three categories: feedforward, feedback and self-organizing networks. Neural networks are the most applicable to classification and regression problems, which do not need perfect precision. The availability of large amounts of data is especially important.

The self-organizing map, introduced by Kohonen (2001), is an unsupervised neural network that has been compared to NLPCA, because it adapts to the structure of the data, and the weight of the neurons tends to set the densest regions of the data and form an approximation of a curve fitted to the data. A neural net based on adaptive resonance theory differs fundamentally from a self-organizing map in the fact that the size and shape of the map are not determined beforehand. The ART map has many modifications, including combinations of ART maps, like ART3 and ARTnet, and hybrids of ART maps and fuzzy logic in fuzzyARTMAP (Wienke et al. 1996, Rallo et al. 2002). Radial basis function networks, introduced by Leonard and Kramer (1991), fit input data to radial basis functions, whereas traditional feed forward networks usually compare input signals to data vectors. RBFN is suitable for fault diagnosis.

Wavenets are combination of wavelets and neural networks with hierarchical multiresolution learning (Bakshi and Stephanopoulos 1993). This type of neural network is especially suitable for low-dimension dynamic fault diagnostic problems. (Zhao et. a. 1998)

Since no single method has all the desirable features, combinations of the methods have become more common in monitoring applications. The wealth of different monitoring methods is stunning, and selecting suitable methods for industrial problems requires plenty of time and considerable resources. These facts have lead to a situation in which the industrial applications far too often rely on mature methods instead of more sophisticated new approaches. Therefore, a systematic approach for the method choosing process would be necessary. In Table 1 some of the statistical, neural network and hybrid methods are compiled and their theoretical and practical aspects are compared. Compiling an open source extended table of monitoring methods might increase the applications of new monitoring methods in the industry. The theoretical aspects listed are fault detection and diagnostic abilities, robustness to missing measurements etc., immunity to signal noise, accuracy of the results, small amount of modeling data required, capability to handle multivariable data, amount of pre-processing required, and the dynamic, trend separation, nonlinear and adaptive properties, and finally an ability to detect novel states. The practical aspects of usability in industrial applications are: ease of implementation, ease of maintenance, visuality and simplicity of computation.

Table1: The qualities of different monitoring methods.

		Theorethical aspects													**Usability**			
		FAULT DETECTION	FAULT DIAGNOSIS	ROBUSTNESS (MISSING MEASUREMENTS)	NOISE IMMUNITY	ACCURACY	MODELLING DATA LOW	CAPABILITY WITH MULTIVARIABLE DATA	PREPROCESSING NECESSARY	DYNAMIC	TREND SEPARATION	NONLINEAR	ADAPTIVE	NOVELTY DETECTION	EASE OF IMPLEMENTATION	EASE OF MAINTENANCE	VISUALITY	SIMPLICITY OF COMPUTATION
STATISTICAL METHODS																		
PCA	principal component analysis	x	-	x	x	x	x	x	x	-	-	-	-	-	x	x	x	x
PLS	partial least squares	x	-	x	x	x	x	x	x	-	-	-	-	-	-	x	x	x
DPCA/PLS	dynamic	x	-	x	x	x	x	x	x	x	-	-	-	-	-	-	-	x
RPCA/PLS	recursive	x	-	x	x	x	x	x	x	-	x	-	x	x	x	x	x	x
NEURAL NETWORK METHODS																		
SOM	self-organizing map	-	x	x	x	-	-	-	x	-	-	x	-	-	x	-	x	-
RBFN	radial basis function	-	x	x	x	-	-	-		x	-	x	-	-	x	x	x	x
Wavenet	wavelet+neural network	-	x	-	x	x	-	-	x	x	x	x	-		-	-	x	-
ART-NN	adaptive resonance theory	-	x	-	x	-	-	-	x	-	-	x	x	x	x	x	x	x
HYBRID METHODS																		
NLPCA/PLS	nonlinear	x	-	x	x	x	-	-	x	-	-	x	-	-	-	-	x	-
MSPCA/PLS	multi-scale (wavelet)	x	-	-	x	x	-	-	x	x	x	x	-	x	-	-	x	-
DNNPCA/PLS	dynamic, nonlinear	x	-	x	x	x	-	-	x	x	-	x	-	-	-	-	-	-
FuzzyARTMAP	fuzzy logic + ART	-	x	x	x	-	-	-	x	-	-	x	x	x	-	x	x	-

3. COMBINING PROCESS INDICATORS WITH MONITORING METHOD LIBRARY

3.1. Process indicators

Model-based approaches are more suitable for capturing the basic characteristics of industrial processes than pure data-based methods using direct process measurements. Since it is often impossible to construct dynamical or mathematical models to the desired precision, data-based (process history based) methods are usually used. The nonlinearities and dynamical properties of the processes lead to the development of methods of ever-increasing complexity, while a combination of model-based and data-based methods would result in simpler computing and better understanding of the method. Relatively little research is being carried out on how to identify the right variables for process history-based methods and how to capture the essence of process knowledge.

The approach used here is based on process indicators, which are modeling process characteristics. The process characteristics can be captured by means of model-based computed variables that approximate, for example, enthalpy or the mass balance, and variable relations such as differences between temperatures or flow measurement ratios. The use of variables that are suitable combinations of measurements are better able to capture process non-linearities than complex monitoring methods. This ability obviously improves the identification capabilities of most data-based monitoring methods.

These process indicators have to be designed separately for every process, but the idea is universally applicable. Modeling the process nonlinearities with simple relationships and combining these with production goal specific indexes provides better understanding of the overall process.

3.2. The combination concept

The idea is to construct a general scheme of a combination of the monitoring method library and the process indicators, and to use these modules when constructing a monitoring application for a specific process, as shown in Fig. 2.

The monitoring method library module consists of the data-based monitoring methods and the table of theoretical and practical aspects of the methods. The process indicators module consists of higher level modeling rules for developing process characteristic indicators.

For the target process, the suitable monitoring methods and their hybrids are first chosen on the basis of the table of method properties. Then, the process characteristics are specified and modeled. Finally, these are combined together in advance to develop a monitoring application that is a part of the advanced operator support system of the plant.

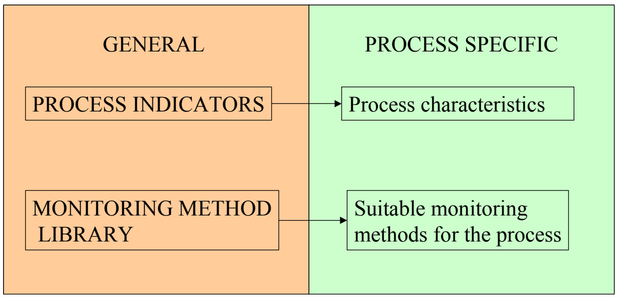

Fig. 2. Combination of process indicators and monitoring method library.

4. INDUSTRIAL EXPERIENCES

In this chapter the industrial applications, which have led to the development of the concept described above, are presented briefly. In the applications, monitoring methods with computed variables describing the process characteristics, have been used. The applications were done by the Laboratory of Process Control and Automation, Helsinki University of Technology in collaboration with industry. The application areas are refining, metals and minerals, and the pulp and paper industries. First, the online monitoring system developed for dearomatization process, is described and discussed in more detail. Then, the other applications are briefly described.

4.1. An online monitoring system for dearomatization process

Komulainen et. al. (2003) have described an online monitoring system for a dearomatization process. The aim of the monitoring system was first, to classify the process state, second, to detect whether the flash point and distillation curve analyzers were working correctly, and third, to provide a reliable prediction of the analyzer measurements. The application used dynamic PLS with the input combination of direct process measurements and process indicators.

The purpose of the dearomatization process is to remove aromatic compounds from the feedstock by hydrogenating them in a continuous process. The process consists of two trickle-bed reactors with packed beds of catalyst, a distillation column, several heat exchangers and separation drums and other unit operations.

Due to the complex, strongly nonlinear, nature of the process, a nine-stage systematic approach was introduced for the development of the online monitoring system. First, the direct process variables that affected the flash point and distillation curve of the product were determined, and the selected variables were then time-lagged. Next, process indicators were identified and the corresponding computed variables were created on the basis of the time-lagged, direct process variables. The combination of direct process variables and computed variables with the strongest influence on the flash point and distillation curve, was then selected. After the combination of variables had been created, the most suitable method for monitoring proved to be dynamic PLS. Different models for monitoring were developed and tested. The offline test was performed with the most suitable models. When the results of the offline tests were satisfactory, an online-monitoring system was developed and tested. Finally, the results of the online test were analyzed.

The process variables affecting the analyzer variables were selected on the basis of process knowledge and correlation analysis. The selected variables included: temperature measurements from reactors and distillation column, direct and recycle flow measurements from several points of the process, pressures and their set points, level measurements, and temperature difference and enthalpy from the reboiler of the column.

The process indicators were identified from the basic equations of the dearomatization process. The characteristics of the distillation especially were investigated, because distillation has a strong effect on the flash point and the initial point of the distillation curve. Based on the selected process indicators, the computed variables were constructed from 36 time-lagged process measurements as follows:

- The heat generated in the reactors, and the heat divided by the flow of the fresh feedstock to the first reactor.
- Several temperature differences in the reactors and the distillation column.
- Flow ratios were determined between feed to the process, feed to the distillation column, distillate, reflux, and product.
- Variables describing enthalpy were represented by simple products of the flow rate and temperature. Instead of direct enthalpies, the enthalpy ratios, i.e. one enthalpy variable divided by another, were used.
- The enthalpy ratios for the distillation column were computed between: feed, distillate, reflux and bottom product
- The process measurement describing the enthalpy of the reboiler, divided by the flow of the feed to the column, was also calculated.

A total of 23 computed variables were created. The combination of direct process measurements and computed variables was formed on the basis of the correlations between the variables and the analyzer variables. The final combination contained 21 direct, time-lagged, process measurements and 23 computed variables, which were constructed from the time-lagged process measurements.

Choosing the most suitable monitoring method the following requirements were specified; The method should be able to handle several dozen variables and the output of the method should give accurate predictions of the analyzer values. The method should be able to distinguish between process transitions at feedstock changes, and the process and analyzer faults. The method should also be able to distinguish between malfunctions of the analyzers and process faults as early as possible, and to give the operator an alarm. A method that could be applied online and based on process history was preferred.

The need for precise prediction ruled out neural nets. The clear relationship between process and computed variables, and the quality variables of the product, led to the use of PLS-based methods. Owing to the dynamic nature of the process, dynamic partial least squares regression (DPLS) was selected. DPLS fulfilled all the requirements, and the relationship between the process variables and the analyzer variables made the structure of the method very simple.

The time-lagged direct process variables and the computed variables formed the input block, and the output block consisted of the analyzer variables. The predictions of the analyzer variables were made using the input block. The residuals between the predicted and the real analyzer values indicated whether the analyzers were functioning correctly. A possible process fault was detected by measuring the Hotelling T2 value of the input block.

The effectiveness of the computed variables was tested by constructing two DPLS models with the original 35 process variables and comparing the results of these models to the results of the DPLS model with computed variables. The teaching data set was the same for all the DPLS models. Testing data set included only data from normal process state. The DPLS model with computed variables had 5 latent variables, and the DPLS models without computed variables had 3 and 5 latent variables. The results with computed variables are excellent for all the analyzer variables where as the models without computed variables fail to classify the normal states correctly for distillation curve variables (Table 2).

During normal process states, the DPLS model with computed variables was used to predict the quality of the final product. The results of the offline test were encouraging; 96 – 99 % of the normal states and 67 – 97 % of the fault states of the analyzers were classified correctly. An online monitoring system was developed and tested in the Naantali Oil Refinery, Finland, for a time period of 144 hours. The monitoring system classified the two feed type changes correctly as normal states, and gave an alarm for an abnormal process state during the disturbance.

Table 2: Comparison of the results with and without computed variables

	With CVs	Without CVs	Without CVs
Latent variables	5	5	3
Captured variance			
Process var.	88.3 %	89.7 %	77.6%
Analyzer var	99.7 %	99.8 %	98.0 %
Correctly detected normal states			
Flash point	100 %	96 %	100 %
Distillation 0%	99 %	0 %	1 %
Distillation 5%	100 %	0 %	59 %
Distillation 10%	100 %	3 %	95 %

4.2. Other industrial applications

Kämpjärvi et al. (2003) developed an online monitoring system, which used a combination of PCA, SOM and RBFN to detect and identify faults. The inputs of the system consisted of direct process measurements and process indicators. The process indicators included relationships between the process variables and model-based computed variables. The main task of the system was to determine when the composition measurements of the online analyzers could be relied on for automatic control purposes. The system was successfully tested online at the Borealis ethylene plant in Porvoo, Finland. The system classified correctly 99,6 % normal states and 77,0 % fault states.

An application of SOM for monitoring the Outokumpu Harjavalta flash smelter was described by Jämsä-Jounela et al. (2003). The system detected equipment malfunctions and monitored process states using SOM in conjunction with heuristic rules. Model-based computed variables had a decisive influence on the success of the application.

Laine et al. (2000) reported an operator support system for ore type identification in the Hitura Mine, Finland. The monitoring system used SOM for on-line identification of the feed ore type and a knowledge database that contained information about how to handle a determined ore type. The key for the successful implementation was the right selection of variables for the ore type determination. The results of the project were remarkable, a commercial product and several industrial implementations.

Rantala (1999) developed a supervisory process monitoring system for the electrolytic refining of copper. This process is characterized by exceptionally large time constants and strong correlations between the variables. The monitoring application was based on a combination of principal component analysis and a self-organizing map. The classification ability of the monitoring system was excellent. The conditions producing high and poor quality copper cathodes were clearly separated in the SOM.

A combination of PLS and computed variables has also been applied successfully to pulp quality control.

5. CONCLUSIONS

Choosing the most suitable monitoring methods and the correct combination of direct process measurements and process indicators play an essential role in the successful development of an industrial process monitoring system. The challenge is to develop a comprehensive monitoring method library and a tool with the ability to choose the most suitable methods given the process description. Having the right method is, however, not enough; the process characteristics also have to be taken into account.

There are already many applications that employ the approach of computed variables modeling the process characteristics, but the research in this field is minimal. The process indicators as a part of the supervisory process monitoring system is now under research in the Laboratory of Process Control and Automation, Helsinki University of Technology.

REFERENCES

Bakshi, B.R., Stephanopoulos, G. (1993), Wavenet: A multiresolution hierarchical neural network with localised learning, *AIChE Journal*, **39**, pp. 57 - 81.

Bergman, S., Sourander, M., Jämsä-Jounela, S.-L., Monitoring of an industrial dearomatisation process, *In IFAC' 02 15th Triennal World Congress*, Barcelona.

Dash, S., Venkatasubramanian, V. (2000), Challenges In the Industrial Applications of Fault Diagnosis Systems, *Computers & Chemical Engineering*, **24**, pp. 785 – 791.

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

Dong, D. and McAvoy, T.J. (1996), Nonlinear principal component analysis--based on principal curves and neural networks, *Computers & Chemical Engineering*, **20**, pp. 65-78.

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.

Kämpjärvi, P., Sourander, M., Jämsä-Jounela, S.-L. (2003), Online analyzer validation and process fault diagnosis for ethylene cracking process under MPC feedback, Industrial Engineering & Chemistry Research, submitted.

Kohonen, T. (2001), *Self-Organizing Maps*, Springer, Berlin, 3rd Ed., 501.

Komulainen, T., Sourander, M., Jämsä-Jounela, S.-L. (2003), An online application of dynamic PLS to dearomatization process, *Computers & Chemical Engineering*, submitted.

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

Laine, S., Pulkkinen, K., S. -L. Jämsä-Jounela (2000), On-line determination of the concentrator feed type at Outokumpu Hitura mine, Minerals Engineering, **13**, pp. 881-895

Lennox, B. and Sandoz, D. (2002), Recent experiences in the industrial exploitation of principle component based fault detection methods, Proceedings of the IEEE International Symposium on Computer Aided Control System Design, Glasgow 2002, pp xxxv – xliv.

Leonard, J.A. and Kramer, M.A. (1991), Radial basis function networks for classifying process faults, *IEEE Control Systems*, **11**, pp. 31 - 38.

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

Misra, M., Yue, H.H., Qin, S.J. and Ling, C. (2002), Multivariate process monitoring and fault diagnosis by multi-scale PCA, *Computers & Chemical Engineering*, **26**, 1281 - 1293.

Rallo, R., Ferre-Giné, J., Arenas, A. and Giralt, F. (2002), Neural virtual sensor for the inferential prediction of product quality from process variables, *Computers & Chemical Engineering*, **26**, 1735 – 1754.

Rantala, A. (1999), Identification of Process Disturbances in Electrolytic Refining of Copper, Master's Thesis, Helsinki University of Technology, Espoo, 103.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N. and Yin, K. (2003), A review of process fault detection and diagnosis Part III: Process history based methods, *Computers & Chemical Engineering*, **27**, pp. 327 – 346.

Wienke, D., van den Broek, W., Buydens, L., Huth-Fehre, T., Feldhoff, R., Kantimm, T. and Cammann, K. (1996), Adaptive resonance theory based neural network for supervised chemical pattern recognition (FuzzyARTMAP) Part 2: Classification of post-consumer plastics by remote NIR spectroscopy using an InGaAs diode array, *Chemometrics and Intelligent Laboratory Systems*, **32** (2) (1996), pp. 165-176.

Wold, S., Sjöström, M. and Eriksson, L. (2001), PLS-regression: a basic tool of chemometrics, *Chemometrics and Intelligent Laboratory Systems*, **58**, pp. 109 – 130.

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

Zhao, J., Chen, B., and Shen, J. (1998) Multidimensional non-orthogonal wavelet-sigmoid basis function neural network for dynamic process fault diagnosis, *Computers & Chemical Engineering*, **23**, pp. 83-92