Journal of Cybernetics and Informatics

published by

Slovak Society for Cybernetics and Informatics

Volume 10, 2010

http://www.sski.sk/casopis/index.php (home page)

ISSN: 1336-4774

Journal of Cybernetics and Informatics 10 (2010)
http://www.sski.sk/casopis/index.php

MULTIVARIATE STATISTICAL METHODS FOR INDUSTRIAL PROCESS PROGNOSTICS

Bratina Božidar and Boris Tovornik

University of Maribor, Faculty of Electrical Engineering and Computer Science
Smetanova 17, 2000 Maribor, Slovenia
Tel.: +386 2 220 7170 Fax: +386 2 220 7272
e-mail: bozidar.bratina@uni-mb.si

Abstract: The paper deals with multivariate statistical methods used for failure prognostics in industrial processes. Modern on-line process monitoring system should support classic fault detection, isolation and diagnosis (FDI) sub-systems to avoid process down-time, increase production, optimize parameters of the production line, etc. However faults usually demand immediate intervention by operator, therefore by using reliable prognostic system, risks can be avoided, maintenance intervals can be scheduled, operation and production strategy can be updated, etc. Presented methods are intended for operator's visual detection of process deviation (along with automated FDI systems) while process monitoring, diagnosis and data analysis tasks are running. By understanding nominal process operation, a hardly detectable small faults and drifts can be used to predict failure scenarios in process prognostics.

Keywords: Fault detection and isolation, prognostics, principal component analysis, multivariate statistical analysis.

1 INTRODUCTION

Failure prognostics is emerging as the next logical step towards improved system condition based maintenance, beside classic fault detection and diagnostics techniques (FDID). These methods form system health management (HMS) platforms which contribute to longer and reliable operation of systems enable them prediction of process operation, maintenance schedules, remaining useful life of system components, system reconfiguration, optimisation, etc. In the Artificial Intelligence community prognostics is yet becoming popular as a discipline and differentiates from fault detection and isolation objective, as it detects precursor of failures and predicts remaining time to failure to occur. From technical or production point of view such information are important for operator to prevent un-necessary process down-time, therefore reduces considerable money loss (customer penalty, safety violation, reduced production plan).

Technique for prediction of the system can be developed using raw measurement data or suitable models of processes, upon which the prognostics is realized. Each type has its own advantage (transparency, implementation) therefore various methods can be combined. Most of them come from the field of artificial intelligence and soft computing. In survey paper (Schwabacher, Goebel, 2008) many developed algorithms are divided into two groups; model-based and data-based algorithms, similar to FDID concepts. Other authors sometimes have different classification depending on the field and discipline their work relies on. Very popular are multivariate statistical methods, derivations of Monte Carlo method, support vector machine learning algorithms, Kalman filters, neural networks, fuzzy logic, etc. More about development and various prognostic techniques can be found in references at the end.

In the paper multivariate statistical analysis is used (principal components (PCA), nonlinear principal components), to achieve very precise detection and prediction of sensor degradation. Algorithm is developed in Matlab/Simulink and communication to laboratory hydraulic model is realized by OPC interface. Scenario of level sensor degradation (artificial aging) in the tank and pipe clogging (simulation of mineral coating on pipe's inner wall) were tested on the laboratory model, where prognostic system operates under local close-loop.

2 MULTIVARIATE STATISTICAL METHOD FOR PROGNOSTICS

Process industry demands reliable operation therefore any process interruptions and unnecessary changes are usually avoided. However the trend of modern SCADA platforms is integration of various advanced and modern control and FDID algorithms which as a stand-alone system provide critical information to the operator. Today these platforms usually include basic statistical methods with simple pre-processing algorithms to obtain system health information or basic insight into the process behaviour. For example, a faulty measurement of degraded sensor performance (aging) can bring the system into unstable operation, so early detection and prediction will this lead to a failure, when and what effect will the fault have to the process output quality is very important. Imagine a batch in a biochemical industry where the growth process of test cells takes a few months. Cells have to be maintained under certain environment conditions (temperature, pressure...) during the production cycle to comply to Food and Drug Administration regulations. By implementing prognostic methods, prediction, FDI, and online batch monitoring of process deviations can be registered and analysed before the batch is finished (using only a small portion of measured data). In case of unpredicted events a prediction of what-if scenarios can be analysed, or time can be determined to reliably solve the situation before the batch will have to be rejected due to bad quality, etc. Similar scenarios can be projected to other industrial processes.

Multivariate statistical methods proved to be easy to implement and satisfactory for many basic tasks in industry, however simplicity has the consequence to accuracy. So to improve prognostic system instead of linear and simple tools, nonlinear techniques should be taken into consideration. Since modern SCADA systems have implemented data acquisition services, statistical model of the process can be obtained upon these large process history datasets by using various tools e.g. principal component analysis, partial least squares, etc. The classic PCA method does not require much of a processing power and is simple to implement, therefore has been widely used in many fields: image compression, fault detection, dimensionality reduction of data (gene expression, meteorology, medicine), etc. It can handle high dimensional and correlated process variables, provides a natural solution to the errors-in-variables problem and includes disturbance decoupling. However, main drawback lies in linearity therefore a lot of research was invested to nonlinear extension in order to achieve better fittings to process behaviour.

Principal component analysis is in FDI field very popular for extracting information from measured data which can also serve as visual information of process operation changes. By observing changes of principal components system behaviour can be monitored and production quality can be maintained. In mathematical term, PCA is performed from aigenvalue decomposition of the covariance matrix from the original measurements. Data matrix X containing n rows with observation of p correlated variables is transformed into independent variables in score matrix T:

$$\frac{X' \cdot X}{n-1} = P' \cdot D \cdot P \qquad T = XP \tag{1}$$

If sufficient process variation is explained with only first few PCs, some columns of loading matrix P can be eliminated. The PCA estimate of X is then estimated with residual error E:

$$X_{est} = T_k P_k' + E_k \tag{2}$$

Subscript k denotes the number of retained principal components.

For evaluation of unpredicted behavior statistical measures e.g. Hottteling T^2 or Q-norm can be used to define residual bounds for detection any process deviation. Complementary to such distance-based measures, a visual representation can be used in a form of quantitative analysis for reduced performance monitoring. For this task the Euclidean concept of distance is useful (Raich and Cinar, 1995) when observing changes of PCs' angles (comparison in 2D or 3D space). Observing distance between points, the Euclidian angle between points a and b with vertex at the origin, can similarly be defined for higher dimensions using vector products:

$$\cos(\theta_E) = \frac{(a' \cdot b)}{\left(\|a\| \cdot \|b\|\right)} \tag{3}$$

The angle definition is adjusted as weighted distance and the Mahalanobis angle between a and b through the origin can be defined:

$$\cos(\theta_M) = \frac{(a' \cdot D^{-1} \cdot b)}{\left(d(a,0), d(b,0)\right)} \tag{4}$$

by using the Mahalanobis distance for points a and b:

$$d(a,b) = \sqrt{(a-b)^T \cdot D^{-1} \cdot (a-b)} \tag{5}$$

where D presents dispersion. A constant Mahalanobis angle around the line joining point u with the origin is a hyperconical surface, with distortion D. Rescaling the scores in T in a way that each has equal variance is done by the Mahalanobis distance measure, distorting the ellipsoid described by the scatter of data observations into a sphere. Fig.1 shows three principal components presented in 3D where position (centre) and direction of PCs are varied according to different process operation regimes. The picture on the right shows a possible way to visually inspect unfinished batch cycle and predict its output quality, hence the decision about batch rejection can be made.

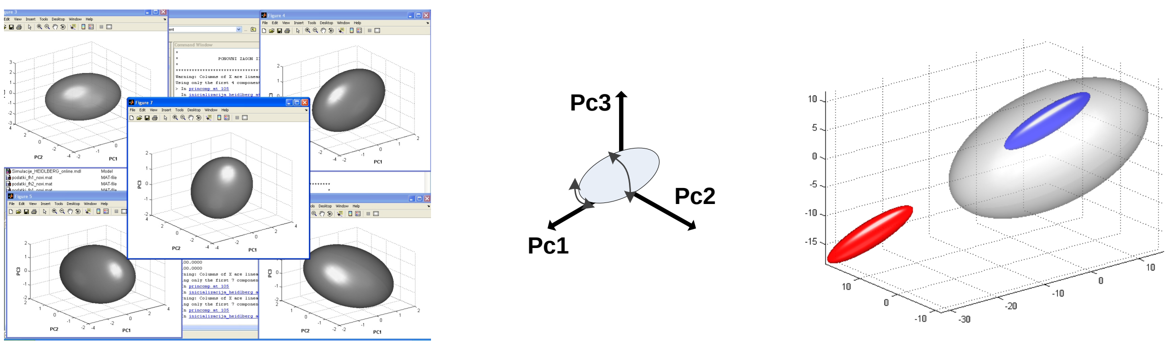

Figure 1: Different PCA models for different process operating regimes(left), and batch prediction (right); rejected batch (red) and accepted batch (blue).

PCA enables quick but rough results, where process deviation needs to be quite large before reliable results can be obtained. To improve statistical model of the process thus enable better prognostics results, nonlinear extension of PCA model is used. NLPCA can be achieved by advanced soft computing algorithms (neural networks, fuzzy logic, genetic algorithm, etc) where auto-associative structure of neural network enables also extraction of nonlinear principal components that can be monitored for process deviations.

In 1991, Kramer (Kramer, 1991) presented a feed-forward neural network to perform identity mappings, where network inputs are reproduced at the output layer. Kramer's NLPCA is a generalization of a classic PCA with enabled nonlinear mappings. To perform NLPCA, the artificial neural network in Fig. 2 consists of three hidden layers of neurons between the input and output layer.

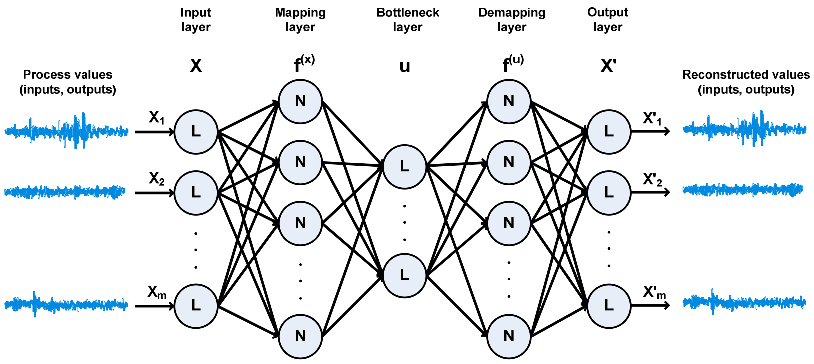

Figure 2: Auto-associative artificial neural network structure.

Next to the input layer there is an encoding layer, followed by a bottleneck layer. Network layers are mirrored to the output so next layer is a decoding layer followed by the network outputs. Nonlinear activation function maps from higher dimension input space to lower dimension - bottleneck space, followed by an inverse transform mapping from bottleneck space back to reconstructed (original) space represented by the outputs. Network outputs produced values are close to the inputs' values by minimizing the objective function.

As described in Kramer's paper, the mappings are achieved by next transformations: a transfer function f_1 maps from x, the input column vector of length l, to the encoding layer, represented by $h^{(x)}$, a column vector of length m,

$$h_k^{(x)} = f_1\left(\left(W^{(x)}x + b^{(x)}\right)_k\right) \quad \textbf{(6)}$$

where, $b^{(x)}$ contains the bias parameters, and $W^{(x)}$ is a weight matrix. A transfer function f_2 maps from the encoding layer to the bottleneck layer containing a reduced number of neurons, which represents the nonlinear principal components u,

$$u = f_2\left(W^{(x)}h^{(x)} + \overline{b}^{(x)}\right) \quad \textbf{(7)}$$

The transfer function f_1 is generally nonlinear, while f_2 can also be identity function. The transfer function f_3 maps from u to the final hidden layer $h^{(u)}$,

$$h_k^{(u)} = f_3\left(\left(W^{(u)}u + b^{(u)}\right)_k\right) \quad \textbf{(8)}$$

followed by f_4 mapping, from $h^{(u)}$ to x', the output column vector of length l, with

$$x_i^{'} = f_4\left(\left(W^{(u)}h^{(u)} + \overline{b}^{(u)}\right)_i\right) \quad \textbf{(9)}$$

The objective function $J = \left\langle \|x - x'\|^2 \right\rangle$ is minimized to find weights and offset parameters of the AANN (optimal values of $W^{(x)}$, $b^{(x)}$, $w^{(x)}$, $\overline{b}^{(x)}$, $w^{(u)}$, $b^{(u)}$, $W^{(u)}$ and $\overline{b}^{(u)}$). Mean squared error between the neural network output and the original data is thus minimized. The choice for the number of hidden neurons in an encoding/decoding layer follows a principle of parsimony, however Kramer recommends using final prediction error (FPE) or Akaike's information criterion (AIC). In case of a small number of mapping nodes accuracy might be low due to limited representational capacity of the network, and in case there are too many nodes, the network will be over-fitted. The algorithms and neural network design were realized in Matlab/Simulink by using neural network toolbox. An extraction of nonlinear components is achieved from bottle-neck layer, where Fig. 3 shows simple example of an extracted nonlinear principal component representation.

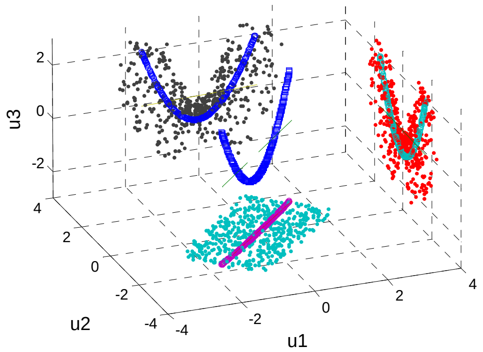

Figure 3: Extracted nonlinear principal component

3 VIRTUAL SENSOR

Instead of just observing changes of process operation by linear or nonlinear visual techniques, another concept presents development of a virtual sensor, which can be used to reconstruct sensor data measurements upon statistical or neural network model (Hines et al., 1998). Output of the AANN and real-time sensor data measurements are compared to generate information about process deviations, trend of deviation, to predict time-to-failure of the process component or its functionality.

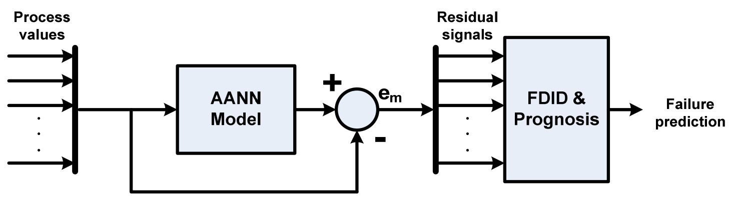

Figure 4: Virtual sensor scheme with FDID and prognosis algorithm as presented in (Hines et al., 1998)

In case of highly dynamic processes a dynamic or recurrent neural network structure is suggested to obtain reliable results.

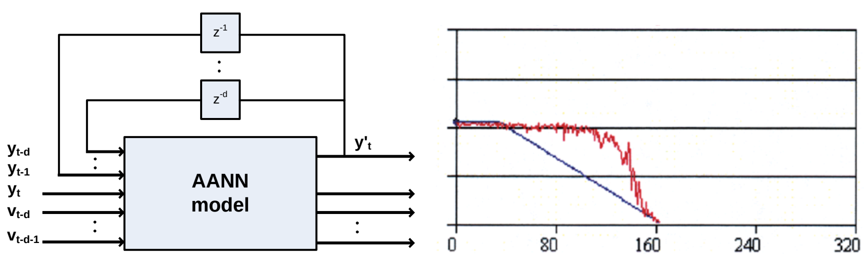

Figure 5: Reconstructed outputs are fed back to the input layer (left) and sensor degradation (right).

4 STUDY CASE: LABORATORY HYDRAULIC THREE-TANK MODEL

Principal component analysis, nonlinear principal component extraction and virtual AANN sensor were realized in Matlab/Simulink and tested on laboratory hydraulic three-tank model.

The process flowsheet of the three-tank laboratory model is depicted in Fig. 6. The upright tanks T_1 and T_2 are mounted above the tank T_3, hence, the inlet to the tanks also depends on the level (hydrostatic pressure) in the tanks T_1 and T_2, respectively (the pumps P_1 and P_2 are not an ideal generators to the system). Also, the outlet pipes are mounted at the bottom of the tank T_3, hence the amount of water in tank T_3 affects the outlet and the inlet flow of the tanks T_1 and T_2. The nonlinear model was derived from the mass balance equations considering the Torricelli's rule and can be conveniently represented as:

$$A_1 \frac{dh_1}{dt} = q_1 - q_{21} - q_{11}; \quad A_2 \frac{dh_2}{dt} = q_2 + q_{21} - q_{22}; \quad A_3 \frac{dh_3}{dt} = q_{22} + q_{11} - q_1 - q_2 \qquad \textbf{(10)}$$

where A_i denotes cross-section of the tank, h_i level in the tank and q_{ij} tank volume inflow or outflow, respectively. The medium in the tanks is fluid, which is taken as an ideal and uncompressible, therefore the specific density of the medium can be neglected (V denotes volume, g denotes gravity constant). For one tank a mass balance equation and the outlet of the tank can be described as:

$$q_{in} - q_{out} = \frac{dV}{dt} = A\frac{dh}{dt}; \qquad q_{ij} = S_{Vi} \cdot sign\left(h_i - h_j\right) \cdot \sqrt{2 \cdot g \cdot \left|h_i - h_j\right|} \qquad \textbf{(11)}$$

where S_{Vi} denotes cross-section of the outlet openness (the valve), h_i and h_j level in the tanks, respectively.

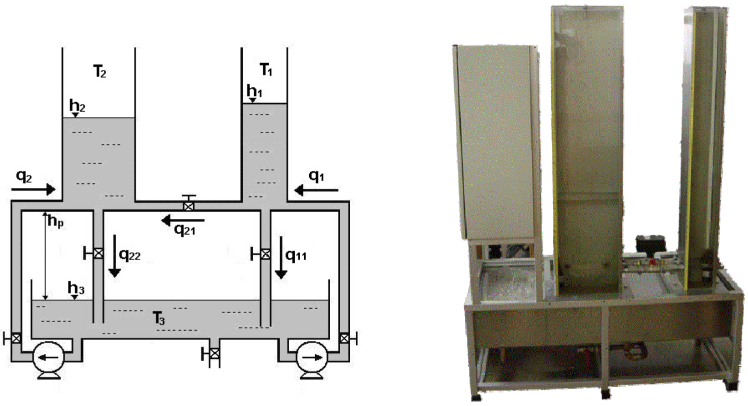

Figure 6: Process flowsheet (left) and image of the laboratory model (right).

To be able to test algorithms developed in Matlab several faults were introduced to the laboratory model while operating under closed loop conditions. Sensor faults were simulated as displacement of the level sensors in the tanks for approximately 2% of measured value, and actuator faults were simulated as partially clogged inlet pipes (closing the inlet valves). The obtained model depended on quality measurements and extraction of important information from noise correlated signals. In order to set up as much as modern real industrial environment an OPC interface (Matlab OPC client) and ethernet communication to PLC was used. The laboratory model was controlled locally by a PLC, while the process variables were fed into and processed in Matlab/Simulink.

In the first case PCA batch (Fig. 7) prognosis was tested. Fig. 8 and Fig. 9 show acceptable (blue) and rejected (red) sets of batch data measurements under different regimes. Level sensor degradation is obvious however small sensor degradation was hard to forecast. Fig. 9 shows forecast upon data measurements of pump, while pipe clogging.

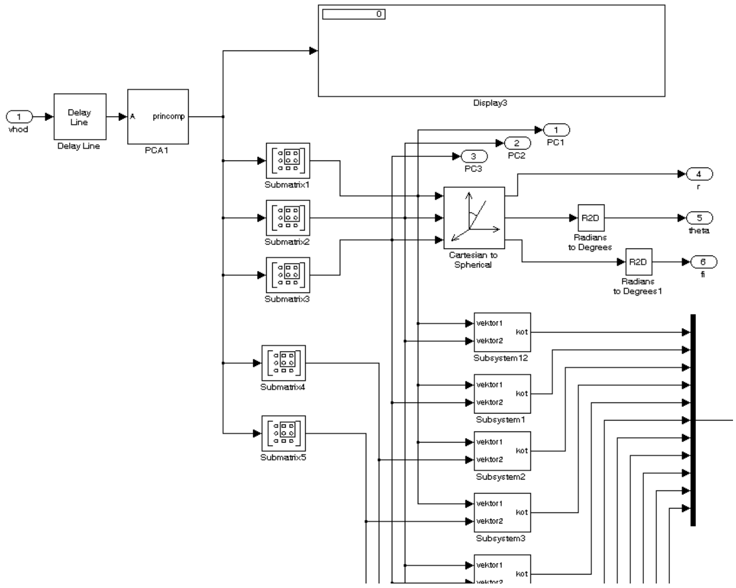

Figure 7: Matlab/Simulink sub-system for PCs batch

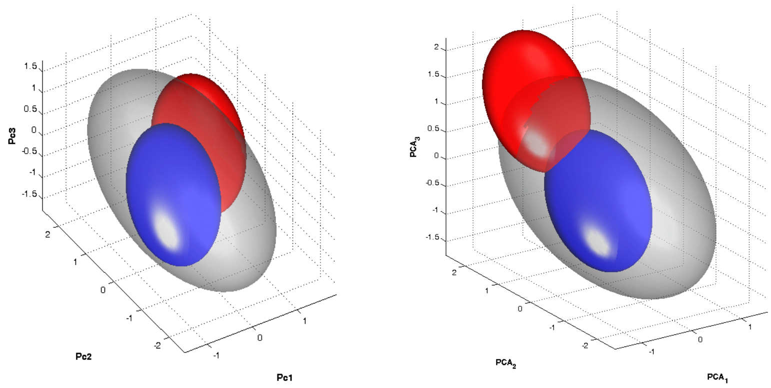

Figure 8: PCA batch prediction for level sensor degradation; small (left), large (right).

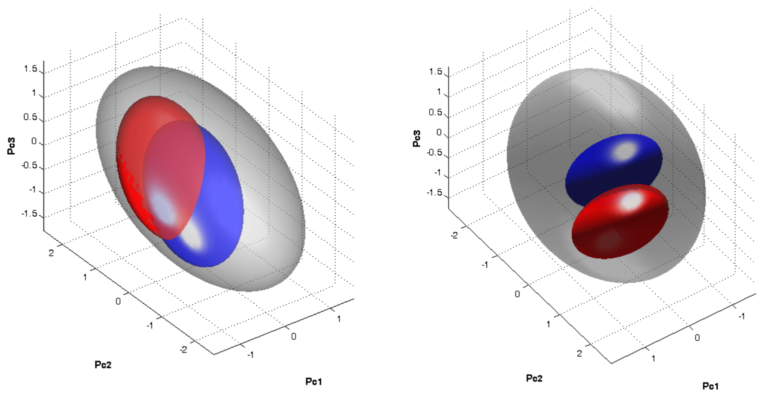

Figure 9: PCA batch prediction for pipe clogging; small (left), large (right).

In the second experiment a concept of nonlinear principal components observation was evaluated. The model was realized in Matlab/Simulink, where a dynamic AANN was trained by Levenberg-Marquardt back-propagation algorithm. The Fig. 10 shows first extracted nonlinear component behaviour when small sensor degradation was introduced to the level measurement. The shape and rotation of the curve is changing upon changes of system characteristics (nonlinearities) or due to changed working point of operation (sensor fault). Nonlinear principal components more accurately describe the behaviour of the process therefore also smaller process deviations can be detected, forecasted and therefore avoided however complexity and design procedure for NLPCA demands more effort.

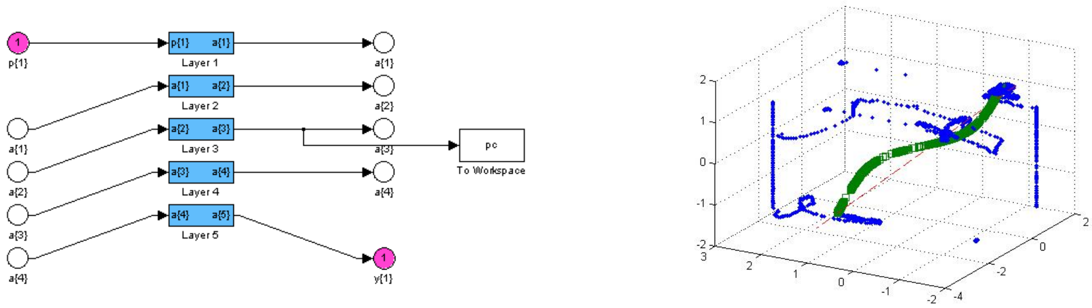

Figure 10: Auto-associative neural network structure in Matlab/Simulink (left) and extracted first NLPC.

In the third experiment a so called data reconciliation scheme was realized. By using AANN whole process operation or any component of the process can be modelled upon measurement data and observed for changes (on-line detection and trend prognosis). In the Fig. 11 a Simulink scheme is shown where artificial degradation of level sensor is realized. Scheme was tested for very small sensor performance degradation (2-4% of measured signal).

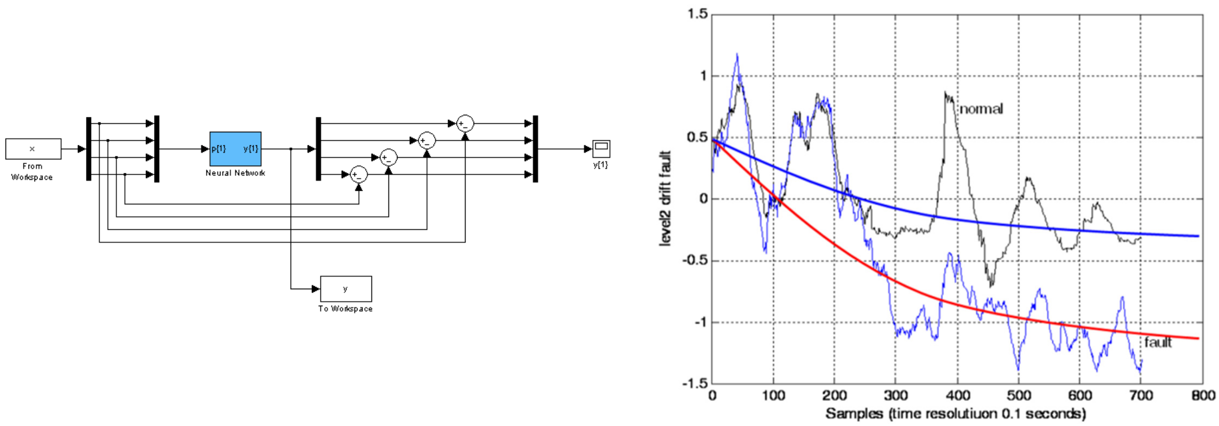

Figure 11: Data reconciliation scheme (left) and detection of slow drift with trend prediction (right)

Recently many SCADA and monitoring system developers started introducing tools and functionalities for basic signal processing and analysis in their commercial products. At first mostly basic statistics functions such as average values, deviation, median, were possible but modern concepts emerged as more standalone add-ons for detailed insight analysis of the process behaviour. A simple application of level sensor reduced performance due degradation was tested with GE Proficy Troubleshooter. The software among others enables development of classic PCA model upon imported process variables from SCADA system historian.

As similar to development in Matlab, the Troubleshooter offers better GUI support and development procedure, which can be appropriate for plant engineers and simple realization of various basic analysis tasks. Upon results, necessary information is passed to the operators to deal with the issue. In Fig. 12 a PCA model and signal pre-processing is shown, realized by Troubleshooter design environment. Fig. 13 shows PCA score graphs of normal and faulty level sensor operation, and also slow drift detection at very small sensor displacement.

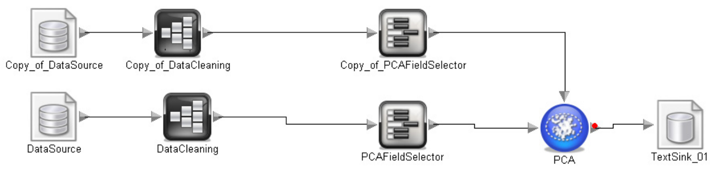

Figure 12: realization of PCA model in GE Proficy Troubleshoter.

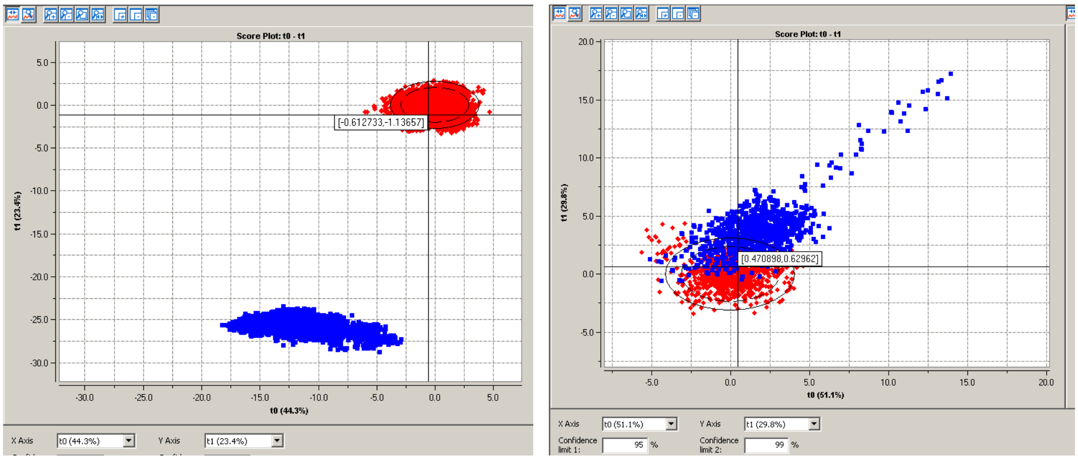

Figure 13: PCA score graph; normal operation (red) and faulty sensor (blue) scores (left), and obvious sensor drift detection at small sensor displacement (right).

5. CONCLUSION

In the paper practical use of some multivariate statistical analysis tools is presented. The emphasis was given to operator's visual support while monitoring processes, with some analysis and prediction functionality. Monitoring of nonlinear principal components and virtual sensors can improve early establishment of root causes in process deviations, therefore operator's shorter reaction time enables quicker treatment of process issues and contribute to less process down-times. However, presented methods are merely at the doorstep of true prognostic algorithms emerging such as remaining useful life prediction (RUL), time to failure, etc.

According to surveys in the field of prognostics, large scale systems remain an area with many unsolved issues in FDI where much more research is needed. Artificial intelligence and soft computing methods can offer great results, especially if combined into hybrid platforms. Design procedure comparison of basic signal processing tasks in mathematical or industrial software has indicated that commercial SCADA developers are catching up functionalities of special mathematical tools. More advanced algorithms (data fusion, neural networks, fuzzy logic, soft-sensors) will probably emerge in future versions of software, where plant operator with appropriate know-how will be able to develop advanced monitoring task of specific variables of the process.

Our research will continue towards advanced algorithms for prognostics that can be easily implemented into commercial process industry equipment or software that enables advanced mathematical computations to achieve better results for product quality monitoring.

REFERENCES

BARALDI, P., ET AL. (2008): Predicting the time to failure of a randomly degrading component by a hybrid Monte Carlo and possibilistic method. *International conference on prognostics and health management*, USA.

COBLE, J., HINES, W. (2008): Prognostic algorithm categorization with PHM challenge application. *International conference on prognostics and health management*, USA.

GOEBEL, K., SAHA, B., SAXENA, A. (2008): A Comparison of Three Data-Driven Algorithms for Prognostics, *The Intelligence Report: MFPT 62/Society for Machinery Failure Prevention Technology, NASA*, USA.

HEIMES, F. (2008): Recurrent neural networks remaining useful life estimation. *International conference on prognostics and health management*, USA.

HINES, J.W., UHRIG, R.E, WREST D.J., (1998). Use of auto-associative neural networks for signal validation. *Journal of Intelligent and Robotic Systems*, **21**, pp. 143-154.

KATIPAMULA, S., BRAMBLEYSAVÁ, M.R. (2005): Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems - A Review, Part I, *HVAC&R RESEARCH*, **11** (No.1), 3-25.

KRAMER, M. A., (1991). Nonlinear principal component analysis using auto-associative neural networks. *AIChE Journal*, (37), pp. 233-243.

RAICH, A., CINAR, A. (1995): Diagnosis of process disturbances by statistical distance and angle measures, *Computer Chemical Engineering*, **21** (No. 6), pp. 661-673.

RAZAVI, F., JALALI-FARAHANI, F. (2008): Ant colony optimization: a leading algorithm in future optimization of petroleum engineering processes. ICAISC, LNAI 5097, pp, 469-478.

RUFUS, F., ET AL. (2008): Health monitoring algorithms for space application batteries. *International conference on prognostics and health management*, USA.

SAMANTA, B., NATARAJ, C. (2008): Prognosis of machine condition using soft computing. Robotics and computer-integrated manufacturing, **24**, pp. 816-823.

SAXENA, A., ET AL. (2008): Metrics foe evaluating performance of prognostic techniques. *International conference on prognostics and health management*, USA.

SCHWABACHER, M., GOEBEL, K. (2007): A Survey of Artificial Intelligence for Prognostics. *The Intelligence Report: AAAI 2007 Fall Symposium, NASA*, California, USA.

JOLLIFFE, I. T., (2004). Principal Component Analysis. *Second Edition. Springer*, New York.

YOUREE, R., ET AL. (2008): A multivariate statistical analysis technique for on-line fault prediction. *International conference on prognostics and health management*, USA.