Leakage Detection and Localisation in Drinking Water Distribution Networks by MultiRegional PCA

Kazimierz Duzinkiewicz[1], Adam Borowa[1], Krzysztof Mazur[1], Michał Grochowski[1], Mietek A. Brdys[2], Krzysztof Jezior[1]

[1]Faculty of Electrical and Control Engineering, Gdansk University of Technology, ul. Narutowicza 11/12, 80-952 Gdansk, Poland

email: {k.duzinkiewicz; aborowa; k.mazur; m.grochowski; m.brdysl; k.jezior}@ely.pg.gda.pl;

[2]Department of Electronic, Electrical and Computer Engineering, University of Birmingham, Birmingham B15 2TT, U.K.

email: m.brdys@bham.ac.uk

Abstract: Monitoring is one of the most important steps in advanced control of complex dynamic systems. Precise information about systems behaviour, including faults indicating, enables for efficient control. The paper describes an approach to detection and localisation of pipe leakage in Drinking Water Distribution Systems (DWDS) representing complex and distributed dynamic system of large scale. Proposed MultiRegional Principal Component Analysis (MR-PCA) skilfully takes full advantage of well known PCA method and enables not only for detecting the leakages but also supports their localisation. The main idea of MR-PCA is presented on example of small water network. Next the method is applied to DWDS in Chojnice, northern Poland. DWDS Chojnice is decomposed into suitable subnetworks what makes that the monitoring process is easier and require less sensors. The subnetworks and corresponding PCA monitoring models are selected based on the network operational knowledge and information regarding its topology.

Keywords: Monitoring, large-scale systems, network systems, fault detection algorithms, water leakage detection, statistical methods.

Kazimierz Duzinkiewicz received M.Sc. degree in Electrical Engineering and Ph.D. in Control Engineering from Faculty of Electrical and Control Engineering at Gdansk University of Technology, in 1973 and 1982, respectively. He has been employed as a university teacher starting his work in 1973 from the post of Assistant to the current position of Senior Lecturer in Department of Control Engineering. During his research work he has published over 80 scientific papers and over 50 scientific and technical reports, mainly dealing with following problems: a) production scheduling and operational control of technological systems with switchable processes (refinery type systems), b) computer control of electric power station in emergency conditions, c) safety and reliability analysis of hazardous systems, d) mathematical modelling of complex systems, e) multihorizon and multilevel optimisation, control and decision support structures and algorithms with applications to petroleum industry and environmental systems (drinking water and wastewater systems). He was NOC Chair of the 11th IFAC Symposium on Large Scale Complex Systems, Gdansk, July 23-25, 2007.

Adam Borowa received his M.Sc. degree in Control Engineering in 2002 from Electrical and Control Engineering Department at Gdansk University of Technology. Soon after he became a Ph.D. student in this Department, between 2001 and 2002 he served one's apprenticeship on Wastewater Treatment Plant at Swarzewo. From 2002 he has been a member of Intelligent Decision Support and Control System Group at Gdansk University of Technology. He was a member of organizing committee of the 11th IFAC Symposium on Large Scale Complex Systems, Gdansk, July 23-25, 2007. He is co-author of 8 publications and 4 chapters in the books. Mainly he focuses on modelling and monitoring of large scale systems, especially processes with many time scales. In 2008 he received the Ph.D. degree in Automatic Control and Robotics.

Krzysztof Mazur received his M.Sc. degree in Control Engineering from Electrical and Control Engineering Department at Gdansk University of Technology in 2005. Currently a Ph.D. student in this Department. A Member of Intelligent Decision Support and Control System Group at Gdansk University of Technology. He was a member of organizing committee of the 11th IFAC Symposium on Large Scale Complex Systems, Gdansk, July 23-25, 2007. His research interests are in the areas of modelling, control and diagnostics of large scale systems. Co-author of 5 papers and 4 chapters in the books.

Michał Grochowski received his M.Sc. degree in Control Engineering in 2000 from Electrical and Control Engineering Department at Gdansk University of Technology. In 2001 he finished M.Sc. study at Faculty of Management and Economics of Gdansk University of Technology. From 1999 to 2003 he was a Ph.D. student in Electrical and Control Engineering Department at Gdansk University of Technology. In 2004 he received the Ph.D. degree in Automatic Control and Robotics. From 1999 is a member of Intelligent Decision Support and Control System Group at Gdansk University of Technology. From 2004 he held the posts of Assistant and Assistant Professor at Control Engineering Department at Gdansk University of Technology. He was a member of organizing committee of the 11th IFAC Symposium on Large Scale Complex Systems, Gdansk, July 23-25, 2007. He is author and co-author of about 15 refereed papers and four chapters in the books. His researches are focused on intelligent control of complex systems, supervisory control, faults detection, softly switched robustly feasible model predictive control.

Mietek A. Brdys received the M.Sc. degree in Electronic Engineering and the Ph.D. and the D.Sc. degrees in Control Systems from the Institute of Automatic Control at the Warsaw University of Technology in 1970, 1974 and 1980, respectively. In 1992 he became Full Professor of Control Systems in Poland. Between 1978 and 1995, he held various visiting faculty positions at the University of Minnesota, City University, De Montfort University and University Polytechnic of Catalunya. Since January 1989, he has held the post of Senior Lecturer in the School of Electronic, Electrical and Computer Engineering at The University of Birmingham. Since February 2001 he has held the post of Full Professor of Control Systems at Gdansk University of Technology. He has served as Consultant for Honeywell Systems and Research Center in Minneapolis, GEC Marconi and Water Authorities in UK, France, Spain, Germany and Poland. His research is supported by the UK and Polish Research Councils, and industry and European Commission. He is author and co-author of about 220 refereed papers and six books. His current research includes intelligent decision support and control of complex uncertain systems, robust monitoring and control, softly switched robustly feasible model predictive control. The applications include environmental systems, technological processes, autonomous intelligent vehicles and defence systems. He is a Chartered Engineer, a Member of the IET, a Senior Member of IEEE, a Fellow of IMA and a Chair of IFAC Technical Committee on Large Scale Complex Systems. He was IPC Chair of the 11th IFAC Symposium on Large Scale Complex Systems, Gdansk, July 23-25, 2007.

Krzysztof Jezior received his M.Sc. degree in Control Engineering from Electrical and Control Engineering Department at Gdansk University of Technology in 2007. Since then a Ph.D. student in this Department and a Member of Intelligent Decision Support and Control System Group at Gdansk University of Technology. Focused in his research on modelling and diagnostics of large scale systems. Co-author of 1 publication and 2 chapters in the books.

1. Introduction

Nowadays, monitoring systems besides data gathering are able to pre-process the data, to recover and estimate not directly measured variables. However, in large scale systems there is very large quantity of information that are hard to handle and sometimes almost impossible to properly process and hence to efficiently utilised it in the control process. An example of such systems is Drinking Water Distribution Systems (DWDS) the representatives of the class of network systems. The DWDS are usually, very complex (lots of pipes, connecting nodes, pumps, tanks etc.) and distributed (in space). It entails measuring of very large number of variables necessity, in order to possess information about the system state that is necessary for efficient system control. In such situations special methods enabling for analysis of large amount of data (e.g. faults detection and isolation) are required. Advanced monitoring systems should not only visualize desired data but also be able to detect devices faults and/or the unusual system behaviour. The paper proposes an approach to detecting and localisation of water leakage in pipes by using the Principal Components Analysis (PCA) method [1]. The PCA is a method that looks for multidimensional correlation between the variables and uses it to reduce the dimensionality of problems simultaneously remaining most of original information. Mostly, large amount of real data process do not provide large amount of important information. Hence, PCA explores data to find out very meaningful ones and include them into statistical models. Moreover, these models clearly indicate the abnormal state of the system thanks to specially calculated measures (T^2 and *SPE*). In case of DWDS such a situation might be caused by device faults (e.g. sensor or pump break down), water leakage in pipe, significant increasing of the water uptake (e.g. caused by fire brigades) etc. Detecting of fault is important however, in case of DWDS the system operator still does not know its type and localisation. Leakages detection and localisation issue is a very important and complex problem that has been widely investigated [2] – [6]. However, available active leakage control methods are basically unpractical due to costs or long leak detection and location time [4]. In the paper the novel approach the MultiRegional Principal Component Analysis (MR-PCA) method is used to detect and to locate the water leakage based on measurements from limited number of measuring devices [5]. MR-PCA tries to join operational experience of staff working in water companies and advanced mathematical analysis. Moreover, this method compromises between detection efficiency and a number of measuring devices.

The method is explained based on simple water network and followed with its application to real town case study DWDS Chojnice (northern Poland).

2. Monitoring and Diagnostics in Advanced Control Systems

Monitoring and diagnostics, which purpose is the fault detection and identification issue, are essential elements of advanced control of complex systems (Figure 1) [7] – [9].

Monitoring and diagnostics utilize a variety of methods for solving the fault detection and identification issue. Basically these methods can be divided into tree classes (Figure 2), which are quantitative model-based, qualitative model-based and process history based, also known as data driven methods [7], [10], [11].

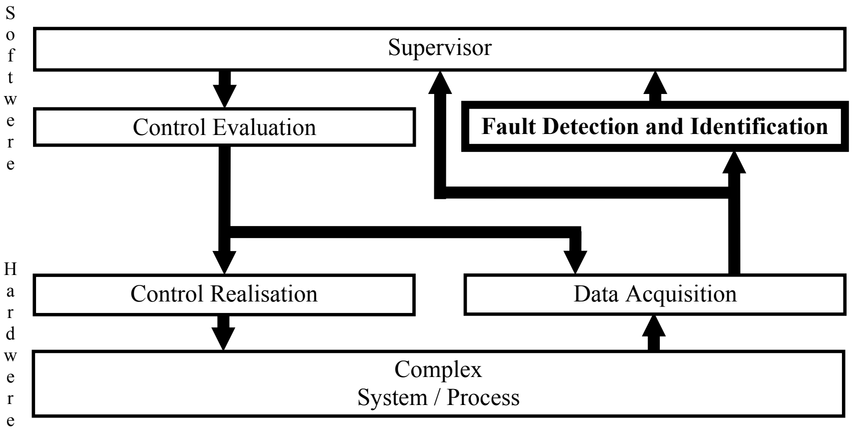

Figure 1. Monitoring and diagnostics (Fault Detection and Identification Unit) in advanced control system structure

Hybrids of monitoring and diagnostic methods can satisfy requirements imposed on a Detection and Identification Unit in a more natural way, since they utilize a set of elements, each fitted to a particular need [7]. Especially if the resulting mixture, consists of different class members, which is the case of MulitiRegional Principal Component Analysis (MR-PCA) [5], [6]. MR-PCA, dedicated to Distributed Systems (DSs) following the network structure, combines Structural Decomposition (SD) and Principal Component Analysis (PCA), where the latter belongs to the Multivariable Statistics methods (Figure 2).

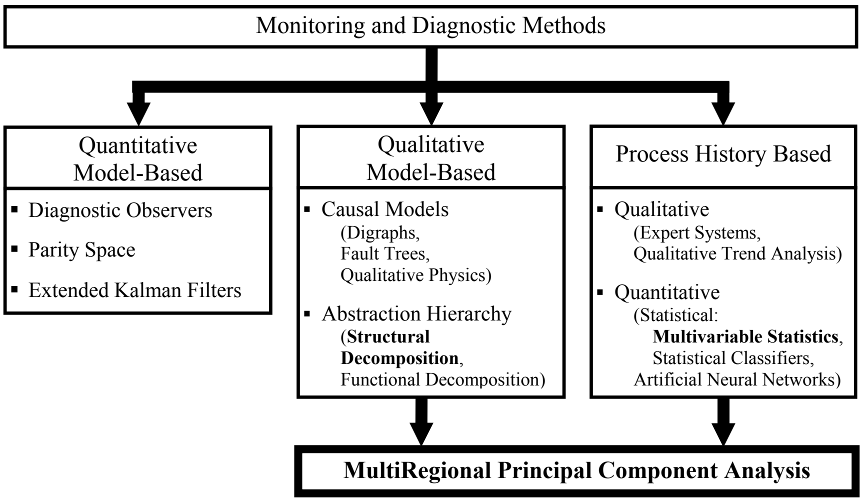

Figure 2. Monitoring and diagnostic methods classification (based on [10])

The main idea behind SD is to conclude about the conditions of system / process in question by means of its subsystems analysis [11]. PCA is described in the next subsection.

2.1 Principal Component Analysis

Principal Component Analysis is a method, which identifies linear dependencies among $n > 1$ variables $x_{i=1,\ldots,n}$, resulting in $s \le n$ decorrelated and linearly related variables $t_{i=1,\ldots,s}$ and a residuals $\tilde{t}_{i=1,\ldots,n-s}$ minimised in the sense of Mean Squared Error (MSE) [1]. Variables $x_{i=1,\ldots,n}$ are assumed to be normally distributed, with independently, identically distributed (IID) Gaussian noise contamination. Due to statistical consistency condition, PCA can model only quasi-static processes, i.e., with unnoticeable transients, because only cross-correlations between variables $x_{i=1,\ldots,n}$ are took into account during the identification.

In more details, given a matrix $\mathbf{X} \in \Re^{N \times n}$ consisting of data collected from the identified process (variables standardized to zero mean and unit variance) and $N >> n$, PCA leads to the following decomposition of $\mathbf{X}$:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \widetilde{\mathbf{T}}\widetilde{\mathbf{P}}^T \tag{1}$$

where $\mathbf{T} \in \Re^{N \times s}$ is the scores matrix containing new data vectors $\mathbf{t}_{j=1,\ldots,N} \in \Re^s$ corresponding to original data samples $\mathbf{x}_{j=1,\ldots,N} \in \Re^n$ and residuals $\widetilde{t}_{j=1,\ldots,N} \in \Re^{n-s}$ collected in residual matrix $\widetilde{\mathbf{T}} \in \Re^{N \times n-s}$. Orthonormal block matrix $\begin{bmatrix}\mathbf{P} & \widetilde{\mathbf{P}}\end{bmatrix} \in \Re^{n \times n}$ plays in the decomposition (1) a key role leading to decorrelation of original cross-correlated data. Its first element $\mathbf{P} \in \Re^{n \times s}$, so-called loadings matrix, which column vectors $\mathbf{p}_{i=1,\ldots,s} \in \Re^n$ contain linear relations indentified in the data $\mathbf{X}$, spans s-dimensional Principal Component Space (PCS), while column vectors $\widetilde{\mathbf{p}}_{i=1,\ldots,n-s} \in \Re^n$ of the second $\widetilde{\mathbf{P}} \in \Re^{n \times n-s}$ span Residual Space (RS) and both spaces are orthogonal (Fig. 3). Thus $\widetilde{\mathbf{t}}_{j=1,\ldots,N}$ and $\mathbf{t}_{j=1,\ldots,N}$ are projections of $\mathbf{x}_{j=1,\ldots,N}$ on the RS and PCS respectively, where the latter (or $\mathbf{P}$ as its basis) is considered as the PCA model.

In order to obtain $\begin{bmatrix}\mathbf{P} & \widetilde{\mathbf{P}}\end{bmatrix}$ one can perform diagonalisation (e.g. using Eigen Decomposition (ED)) of approximated data correlation matrix $\hat{\mathbf{R}}_{\mathbf{X}} \in \Re^{n \times n}$:

$$\hat{\mathbf{R}}_{\mathbf{X}} = \frac{1}{N-1}\mathbf{X}^T\mathbf{X} \tag{2}$$

resulting in:

$$\hat{\mathbf{R}}_{\mathbf{X}} = \begin{bmatrix}\mathbf{P} & \widetilde{\mathbf{P}}\end{bmatrix}\begin{bmatrix}\mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \widetilde{\mathbf{\Lambda}}\end{bmatrix}\begin{bmatrix}\mathbf{P}^T & \widetilde{\mathbf{P}}^T\end{bmatrix}^T \tag{3}$$

where $\mathbf{\Lambda} \in \Re^{s \times s}$ and $\widetilde{\mathbf{\Lambda}} \in \Re^{n-s \times n-s}$ are both diagonal matrices containing eigenvalues of (2): $\lambda_{i=1,\ldots,s}$ and $\widetilde{\lambda}_{i=1,\ldots,n-s}$ corresponding to appropriate block matrix of $\begin{bmatrix}\mathbf{P} & \widetilde{\mathbf{P}}\end{bmatrix}$ respectively and all eigenvalues are proportional to the variance of original data $\mathbf{X}$ in corresponding directions $\begin{bmatrix}\mathbf{P} & \widetilde{\mathbf{P}}\end{bmatrix}$ (Fig. 3). In the notation (3) it is assumed that $\lambda_{i=1,\ldots,s}$ and $\widetilde{\lambda}_{i=1,\ldots,n-s}$ are sorted in descending order and in particular $\lambda_s \geq \widetilde{\lambda}_1$. From (2), (3) it is clear why only cross-correlation process structure can be modelled using PCA approach.

Assumption of IID Gaussian noise contamination leads to equal values of $\widetilde{\lambda}_{i=1,\ldots n-s} = \widetilde{\sigma}^2$, where $\widetilde{\sigma}^2$ is the variance of noise in question in all $n-s$ residual dimensions. This enables for clear and doubtless separation of $\mathbf{\Lambda}$ and $\widetilde{\mathbf{\Lambda}}$, and in consequence $\mathbf{P}$ and $\widetilde{\mathbf{P}}$. However in practice this is the rare case and one has to choose approximation $\hat{s}$ rather than s on the basis of some of available methods [12]. The least sophisticated, though quite effective, is the Captured Percent Variance (CPV):

$$CPV(s) = \frac{\sum_{i=1}^{s}\lambda_i}{\sum_{i=1}^{s}\lambda_i + \sum_{i=1}^{n-s}\widetilde{\lambda}_i}100\% \tag{4}$$

where the choice of $\hat{s}$ depend on the assumed minimal captured by the PCA model percent of data variance $CPV_{\lim}$:

$$\hat{s} = \arg\min_{s}\left(CPV(s) \geq CPV_{\lim}\right) \tag{5}$$

Because of PCA ability of modelling only the linear part of the processes, all nonlinear dependencies contained in data $\mathbf{X}$ would be linearly approximated minimising MSE of residuals. In such case linearization errors are included in the RS, which becomes a PCS non-fitting data container.

Data analysis after decomposition (1) into PCS and RS, can be performed by norms derived for each of subspaces separately, which are Hottelings T^2 and Squared Prediction Error (SPE) respectively [13]. The former, defined as:

$$T^2(j=1,\ldots,N)=\left(\mathbf{t}_j\right)^T \boldsymbol{\Lambda}^{-1}\mathbf{t}_j=\left(\mathbf{x}_j\right)^T \mathbf{P}\boldsymbol{\Lambda}^{-1}\left(\mathbf{P}\right)^T \mathbf{x}_j \tag{6}$$

is the squared Euclidean distance of data sample $\mathbf{x}_{j=1,\ldots,N}$ projected onto PCS to the subspace origin, weighted proportionally to $\mathbf{t}_{j=1,\ldots,N}$ variance (Fig. 3). The letter norm, i.e. *SPE*:

$$SPE(j=1,\ldots,N)=\left(\tilde{\mathbf{t}}_j\right)^T \tilde{\mathbf{t}}_j=\left(\mathbf{x}_j\right)^T\left(\mathbf{I}-\mathbf{P}\left(\mathbf{P}\right)^T\right)\mathbf{x}_j \tag{7}$$

measures the squared Euclidean distance of the residual $\tilde{\mathbf{t}}_{j=1,\ldots,N}$ to the PCS (Fig. 3), where $\mathbf{I}\in\Re^{n\times n}$ is the identity matrix. For both norms (6), (7) thresholds $T^2_{\lim}(\hat{s},\beta)$ and $SPE_{\lim}(\hat{s},\alpha)$ can be computed corresponding to data $\mathbf{X}$ variability in each of subspaces respectively (Fig. 3). In the first case threshold is defined as a chi-squared deviate for the level of significance β and $\hat{s}$ degrees of freedom [1]

$$T^2_{\lim}(\hat{s},\beta)=\chi^2_{\beta}(\hat{s}) \tag{8}$$

which comes from the assumption of data normal distribution. The similar assumption leads to the equation for $SPE_{\lim}(\hat{s},\alpha)$ [13]:

$$SPE_{\lim}(\hat{s},\alpha)=\Theta_1\left(\frac{c_\alpha\sqrt{2\Theta_2 h_0^2}}{\Theta_1}+1+\frac{\Theta_2 h_0\left(h_0-1\right)}{\Theta_1^2}\right)^{\frac{1}{h_0}} \tag{9}$$

where:

$$\Theta_{i=1,2,3}=\sum_{j=1}^{n-\hat{s}}\tilde{\lambda}_j^i,\quad h_0=1-\frac{2\Theta_1\Theta_3}{3\Theta_2^2} \tag{10}$$

and c_α is a standard normal deviate corresponding to the upper percentile $1-\alpha$. At this point should be indicated that α and β are tuning parameters.

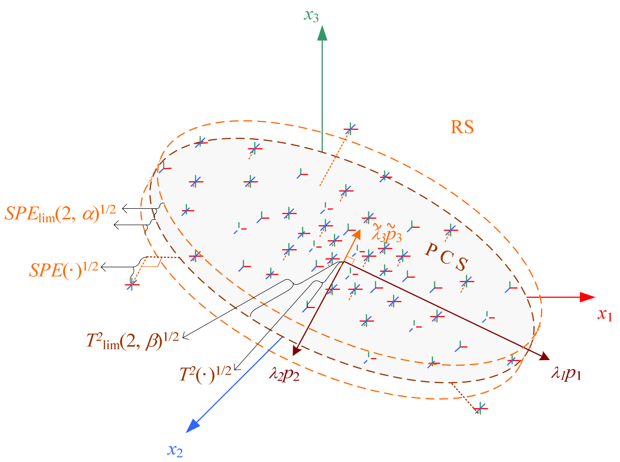

Figure 3. Example of PCA data decomposition for $n=3$, $\hat{s}=2$ and certain α and β.

PCA can be used in process monitoring field for abnormality detection. For this reason PCA model is to be identified from process data $\mathbf{X}$. Due to monitoring purposes it not only means PCS basis derivation, i.e. $\mathbf{P}_{\hat{s}}$ (thus computing $\hat{s}$ simultaneously), but also $\boldsymbol{\Lambda}_{\hat{s}}$, $T^2_{\lim}(\hat{s},\beta)$ and $SPE_{\lim}(\hat{s},\alpha)$ calculation. These last two quantities are required for bounding set of operational states $\mathbf{x}_{j\in\{1,\ldots,N\}}$ considered as desirable/allowable, with respect to α and β chosen values. This implies that process data should contain reasonably the largest representation of process desirable/allowable operational states (by the means of data samples $\mathbf{x}_{j\in\{1,\ldots,N\}}$) and $\mathbf{X}$, because of PCS identification source, is called training data. The quadruple $PCA^{MM}=\left\{\mathbf{P}_{\hat{s}},\ \boldsymbol{\Lambda}_{\hat{s}},\ T^2_{\lim}(\hat{s},\beta)\ \ SPE_{\lim}(\hat{s},\alpha)\right\}$ will be referred to as PCA Monitoring Model.

After PCA^{MM} is obtained off-line from training data, the next step is the actual process monitoring performed on-line for data samples $\mathbf{x}(k)$ constructed analogously to $\mathbf{x}_{j=1,\ldots,N}$. Since $\mathbf{P}_{\hat{s}}$ corresponds to the model of the process under monitoring, current norms values $T^2(k)$:

$$T^2(k) = (\mathbf{x}(k))^T \mathbf{P}_{\hat{s}} \mathbf{\Lambda}_{\hat{s}}^{-1} (\mathbf{P}_{\mathbf{s}})^T \mathbf{x}(k) \tag{11}$$

and $SPE(k)$:

$$SPE(k) = (\mathbf{x}(k))^T \left(\mathbf{I} - \mathbf{P}_{\hat{s}} (\mathbf{P}_{\hat{s}})^T\right) \mathbf{x}(k) \tag{12}$$

measure the (quadratic) distances of current operational state from their, training data based, expected values. Thus ratios $T^2(k)/T^2_{\lim}(\hat{s},\beta)$ and $SPE(k)/SPE_{\lim}(\hat{s},\alpha)$ can be used for abnormal operational state indication/detection in case of unity violation by either of them:

$$\text{abnormal opertional state at time instatnt } k: \quad \frac{T^2(k)}{T^2_{\lim}(\hat{s},\beta)} > 1 \quad \vee \quad \frac{SPE(k)}{SPE_{\lim}(\hat{s},\alpha)} > 1 \tag{13}$$

and the indication/detection magnitude (values of ratios $T^2(k)/T^2_{\lim}(\hat{s},\beta)$ and $SPE(k)/SPE_{\lim}(\hat{s},\alpha)$) depends on the abnormality magnitude measured by the (11), (12) relatively to the closest operational state concerned as a desirable/allowable, represented by the thresholds (7), (8).

As far as the monitored process fulfil PCA assumptions, there is a fundamental difference between abnormality detection by the $T^2(k)/T^2_{\lim}(\hat{s},\beta)$ and $SPE(k)/SPE_{\lim}(\hat{s},\alpha)$ ratios. The former value indicates abnormal operational states, which preserve cross-correlation structure of the process, hence caused mainly by the operation point changes, while the latter ratio is responsible for abnormalities detection of PCA modelled process. However in case of nonlinear process under PCA monitoring, which can be often found in practice [5], [6], [14] – [19], since PCS contains only linearised and originally linear part of dependencies among variables $\mathbf{x}_{j=1,\ldots,N}$, current $T^2(k)/T^2_{\lim}(\hat{s},\beta)$ values indicate an abnormality being a mixture of operation point as well as whole process changes, both not perpendicular to the $\mathbf{P}_{\hat{s}}$ directions. The same applies to the second ratio $SPE(k)/SPE_{\lim}(\hat{s},\alpha)$, with an exception, that this measure detects abnormal operational states (again a mixture of operation point and process changes) not captured by the PCS.

3. MultiRegional Principal Component Analysis

Distributed Systems (DSs) can be decomposed into regions such that operational state of DS follow this decomposition resulting in a set of local (regional) operational states. Because DSs posses network process structure, any local operational state is mutually dependant of its neighbouring regions, where the dependencies are defined by network topology. Hence any source of a abnormal operational state of DS can be analysed through the local operational states, which indicate a local abnormality, if and only if, the abnormality in question, has significant influence on them (local operational states). Significance of analysed abnormality is given by some measure of desirability/allowability of operational state.

Moreover, it is the network topology that defines, which of local models (representing local operational states) are sensitive to abnormal operational state of DS with respect to its particularly placed source. This knowledge can be used to establish a methodology for abnormality source localization.

In case of PCA chosen as a basis monitoring method, as many regions R of DS should be distinguished as it is possible, subject to j-th region compactness and minimal number $n_j > 1$ of monitored variables $x_{i_j=1,\ldots,n_j}$. For each of R regions a local (regional) PCA Monitoring Model $PCA^{MM}_{j=1,\ldots,R}$ and $PCA^{MM}_j = \left\{\mathbf{P}_{\hat{s}_j,j},\ \mathbf{\Lambda}_{\hat{s}_j,j},\ T^2_{\lim,j}(\hat{s}_j,\beta_j),\ SPE_{\lim,j}(\hat{s}_j,\alpha_j)\right\}$ is to be derived following the methodology stated in the previous section (2.1). Defining norms for the j-th region $T^2_j(k)$ and $SPE_j(k)$ analogously to (11), (12) computed at time instant k with respect to local data sample $\mathbf{x}_j(k) \in \Re^{n_j}$, ratios: $T^2_j(k)/T^2_{\lim,j}(\hat{s}_j,\beta_j)$ and $SPE_j(k)/SPE_{\lim,j}(\hat{s}_j,\alpha_j)$ are considered as measures of desirability/allowability of local operational state and again (13) it assumed, that for the j-th region if either of these ratios violates unity, there is an abnormality in DS causing local abnormal operational state. Thus abnormal operational states indication by particular local models still depends on its magnitude (relatively to the closest operational state concerned as a desirable/allowable), however in this case abnormality magnitude is understood locally, i.e. corresponds to particular regional PCA Monitoring Model.

It is important to notice, that it is possible to distinguish between ‘process’ abnormality of DS and a sensor fault. While the latter is detected only by one regional PCA Monitoring Model (assigned to the specific sensor), the abnormality of DS radiates into all regions causing changes in $T_j^2(k)$ and $SPE_j(k)$ in more then one adjacent PCA Monitoring Model.

Network process structure of DSs can be illustrated by the means of nodes and links, which connections structure follows the network topology. From this point of view any distinguished region should consist of variables measured at a node and in all connected links. In special case of isotropic topology (Fig. 4) the highest desirable/allowable regional operational state violation, thus also the highest abnormality indication, is in the neighbourhood of abnormality source, since the grater is the distance of local model from the abnormality source localization, the less its operational state depends on abnormality ‘injected’ into the network.

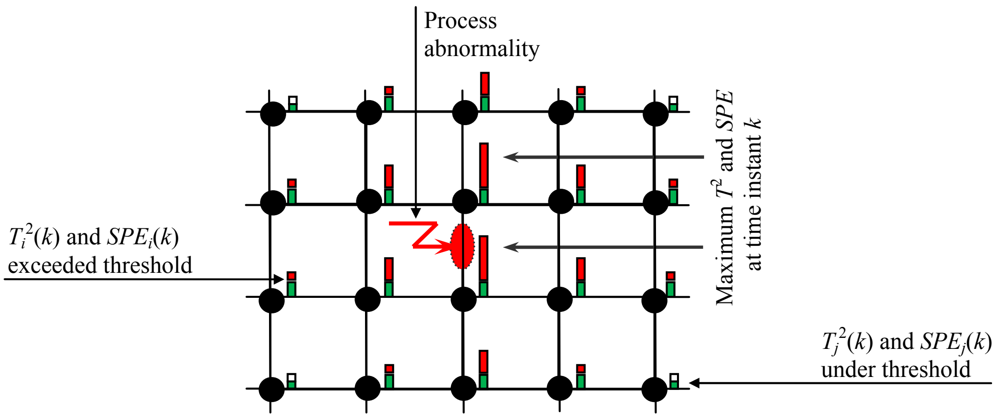

Figure 4. Visualisation of MR-PCA approach

Detection and localisation of abnormality source in DS can be briefly described as follows:

1. During the process operation both measures $T_j^2(k)$ and $SPE_j(k)$ for all R regional PCA Monitoring Models are monitored.
2. If any measure exceeds corresponding threshold $T_{\lim,j}^2(\hat{s}_j,\beta_j)$ or $SPE_{\lim,j}(\hat{s}_j,\alpha_j)$ respectively, then it is said that a certain abnormality (including sensor faults) occurs.
3. If it is only one regional PCA Monitoring Model that indicate abnormality, first check for sensor fault among n_j locally measured variables $x_{i_j=1,\ldots,n_j}$. Else, at once consider detected abnormality as affecting the process.
4. Regional PCA Monitoring Models with the largest $T_j^2(k)/T_{\lim,j}^2(\hat{s}_j,\beta_j)$ or $SPE_j(k)/SPE_{\lim,j}(\hat{s}_j,\alpha_j)$ values determine the localisation of the process abnormality source.

The main idea of proposed method is quite similar to Multi Block Principal Component Analysis (MB-PCA) presented in [20]. However, this name appeared earlier and was dedicated for quite different approach [21]. Therefore, it was proposed [5], [6] to name the method MultiRegional PCA (MR-PCA).

4. Drinking Water Distribution Systems

Drinking Water Distribution System (DWDS) is a good example of a DS. In this section the general description of DWDS is presented.

Nowadays, DWDS is one of the most important systems in community. Its efficient control requires advanced method e.g. predictive control [22], [23] or adaptive control and reliable monitoring system. Proposed approach is applied to detection and localisation of failures in DWDS. Usually in DWDS, drinking water is introduced into the network by using pumps (pumping station) and transported through the network by pipes. Pipes connect in nodes where delivered water is mixed and transported farther. Flows through the pipes are enforced by nodal pressure differences. These are caused by pumps or/and by the water tanks. Tanks are used to store the water in periods when water production is greater then its consumption [23].

The network mathematical model is composed of two parts: static and dynamic. The static part is typically available in an implicit form represented by the element algebraic equalities and the interconnection equalities. This is described for water networks by Brdys and Ulanicki [25]. In general, the element algebraic equalities are described by non-linear functions. The interconnection equalities can be written based on conservation equations. Using the energy loss-gain relationships for the different elements of water distribution system, the conservation equation can be written in three forms: the node, the loop and pipe equations [26].

Unlike the node and loop equations, the pipe equations are solved for the vector of pipe flows Q and hydraulic head h simultaneously. Formulating the static part of water distribution network mathematical model we use the pipe form. The dynamic part of the network mathematical model is represented by differential equalities describing tanks. Because the measurements are available at discrete moments, the water distribution network model is formulated in discrete form.

Paper considers detection and localization of water pipe leakages. The DWDS is modelled in simulation packages Epanet [27]. The leakages are modelled as an emitter. The flow rate through the emitter varies as a function of pressure available at the node [27]:

$$Q = C\mathrm{p}^{\gamma} \tag{14}$$

where: Q is the flow rate, p is the pressure and C is the discharge coefficient (emitter coefficient), finally γ is the pressure exponent.

5. MultiRegional Principal Component Analysis in Application to Drinking Water Distribution Systems

Most often faults in DWDS are pump breakdowns and water leakages from the pipes. The former is easy to identify while the leakage detection and localisation is harder as it is placed underground. The faults might be detected based on the system measurements.

When diagnosis of DWDS is under the consideration monitoring and diagnostics methods are directly divided into two groups [4]: 'measurements and model based', e.g. Inverse Transient Method [2] and 'measurements based' [7], [10], [11]. A group of measurement based methods utilizes statistical data analysis [28], but these methods are still in the stage of research and development [4]. Moreover available active leakage control methods are basically unpractical due to costs and time consuming or having the long leak detection and location times [4]. This is not the case of MR-PCA, which can join not only relatively uncomplicated statistical analysis dedicated to DSs, such as DWDS (due to SD) but also experience of staff working in water companies. Moreover, this method compromises between detection and localisation efficiency and a number of measuring devices [5], [6].

In the DWDS one can measure the water pipe flows, pressure in the nodes, water level in the tanks and water quality (e.g. chlorine concentration). For DWDS approach region (when regional PCA Monitoring Models are taken into consideration) means a measurement nodes together with all adjacent links, while the abnormal DWDS operational state is said to be caused by the single water leakage.

5.1 Fundamentals

Local j -th PCA Monitoring Model PCA_j^{MM} for each of R regions consists of the locally identified PCS basis $\mathbf{P}_{\hat{s}_j,j}$ with corresponding modelled variance information in diagonal matrix $\mathbf{\Lambda}_{\hat{s}_j,j}$, as well as thresholds computed for local abnormality indication. All these elements are derived with respect to chosen PCS dimensionality $\hat{s}_j$ (e.g. using CPV criteria (4), (5), thus also $CPV_{\lim,j}$) and α_j, β_j parameters.

Local PCA models use nodal heads (heads is a sum of nodal pressure and its geodetic heights) and adjacent pipe flows as process variables (of course for PCA models identification as well as on-line

DWDS regions monitoring). Since each local model (PCA model or PCA Monitoring Model) in all cases includes one quantity connected to a particular node (namely nodal head), the 'Identification number' of the particular node will be referred to as the model name and replaces its j-th index. It means, e.g. (Fig. 5) that PCA Monitoring Model $PCA_{88'}^{MM}$ is related to the head in the node '88' and flows in the pipes linking this particular node with its neighbours, which are: '44', '55' and '66'.

The minimal percent of variance captured by the j-th PCA model, $CPV_{\lim,j}$, in all cases is set to 95%. However due to existing nonlinearities this ensures only capture of the linearization results by PCS. Values α_j and β_j are selected separately to ensure abnormality detection (with respect to given local PCA model) and simultaneously false alarms reduction. This implies in the non-normal and nonlinear case of the DWDS, that one must pay special attention while tuning α_j and β_j. Authors in most of further presented case studies assigned values from the range $0{,}6 \div 0{,}7$ and $0{,}7 \div 0{,}9$ for α_j and β_j respectively.

5.2 Simple water network case study

Paper presents MR-PCA method in application to leakage detection in Chojnice drinking water distribution network. Large numbers of potential leakages in the network as well as potential monitoring points might be investigated, there. Hence in order to explain the fundamentals of the method and to illustrate well its efficiency, the simulations were carried out on small testing network (Fig. 5), where the leakages were modelled in different but very meaningful places, as an extra node with pressure dependent demand (Emitter).

Fig. 5 presents three case studies when the leakage was modelled in different places in the testing network. First leakage was modelled in pipe between nodes '44' and '88' (Section 5.2.1), the second one in pipe between nodes '29' and '30' (Section 5.2.2) and the last one in pipe between nodes '25' and '28' (Section 5.2.3). All leakages took place between 46th and 52th hour.

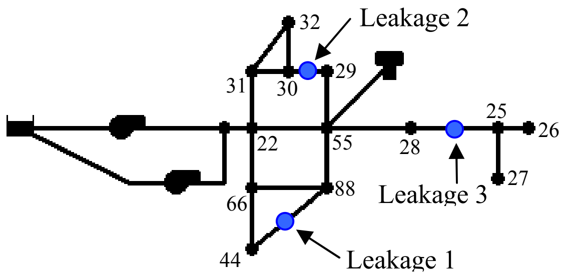

Figure 5. Simple water network. Modelled leakages marked with blue dots

The following figures present the effects of monitoring process by using PCA Monitoring Models designed based on flow rates delivering water to particular node and its nodal head.

5.2.1 Water main screening effect

The dashed blue lines in Fig. 6. represents results of quality measures T^2 and SPE simulated on the training data, without any leakage. The solid red lines represents an effect of monitoring for the same simulation period but with modelled leakage (in pipe between nodes '44' and '88').

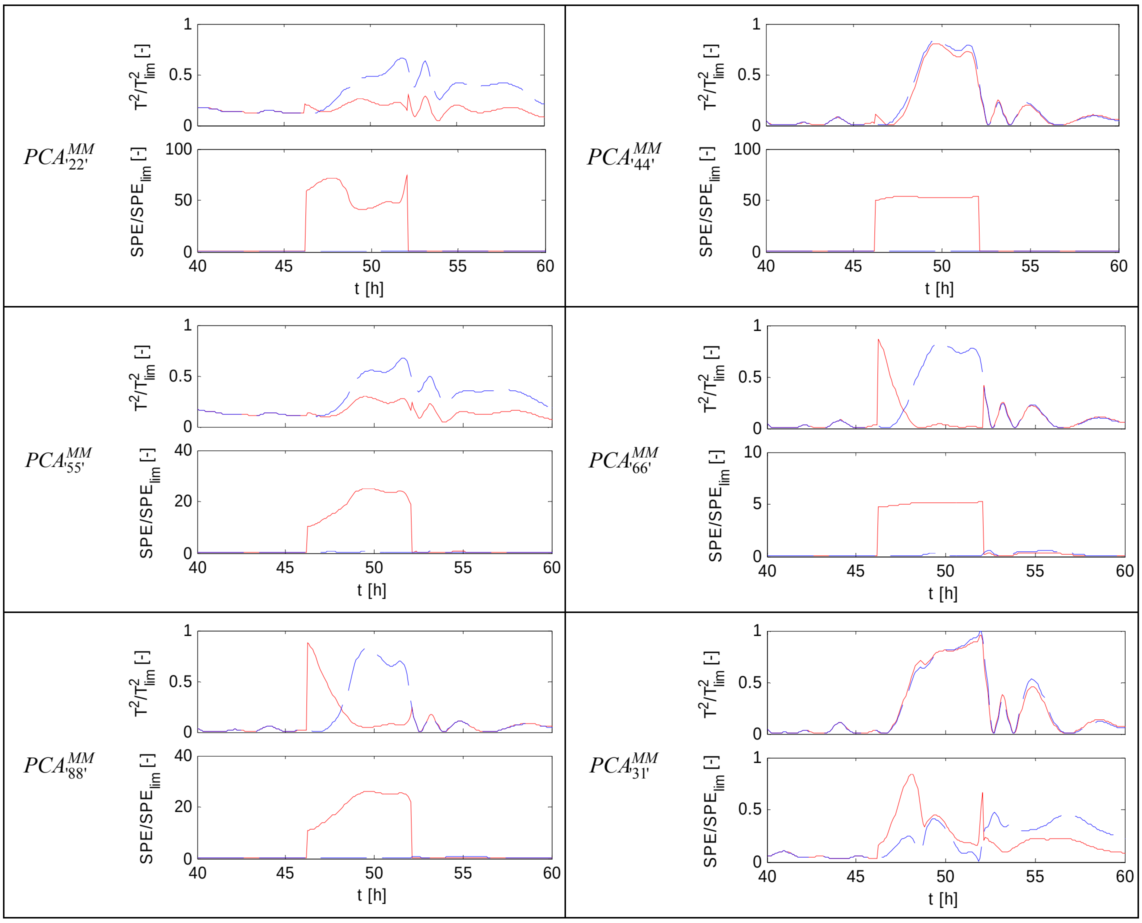

Figure 6. Leakage monitoring results by PCA^{MM} designed at selected nodes – water main screening effect (simulation without leakage (training) - blue; simulation with modelled leakage – red)

Differences between the lines representing PCA models being identified on training data and its current (simulated) responses indicate that something unusual has took place in the network. In this particular example it means pipe water leakage. Notice how considerable increase of the T^2 and SPE measures is generated by PCA Monitoring Models '44', '22', '66' and '88' ($PCA^{MM}_{'44'}$; $PCA^{MM}_{'22'}$; $PCA^{MM}_{'66'}$ and $PCA^{MM}_{'88'}$). It simply means that the leakage has significantly disturbed the water flow rates and pressures being the base for particular PCA Monitoring Models (PCA^{MM}). There are also PCA^{MM} e.g. $PCA^{MM}_{'31'}$ that do not indicate any abnormalities in the network in spite being in similar distance to the leakage. Such a phenomenon is rather strange at the first glance. However, these models are designed based on measurements gathered from points laying 'behind' the water main. Hence, the significant amount of water flowing by the water main causes that modelled leakage did not significantly affect the usual (operating states represented by training data) flow rates and nodal heads on the other side of water main. This implies that the PCA Monitoring Models are unable to detect the leakage. In the paper such an effect is called 'screening effect' of the water main.

5.2.2 Water tank screening effect

This section presents a case when the leakage was modelled in pipe between nodes '29' and '30'. Similarly to the previous case study some of the PCA Monitoring Models produce quality measures significantly exceeding the assumed threshold (unity in the case of ratios $T_j^2(k)/T^2_{\lim,j}(\hat{s}_j,\beta_j)$ and $SPE_j(k)/SPE_{\lim,j}(\hat{s}_j,\alpha_j)$ monitoring) e.g. $PCA^{MM}_{'55'}$; $PCA^{MM}_{'28'}$; $PCA^{MM}_{'29'}$, $PCA^{MM}_{'31'}$ and $PCA^{MM}_{'32'}$, what indicates the leakage (Fig. 7). In opposite, there are models that generate measures similar to their training values what suggests that noting unusual happened.

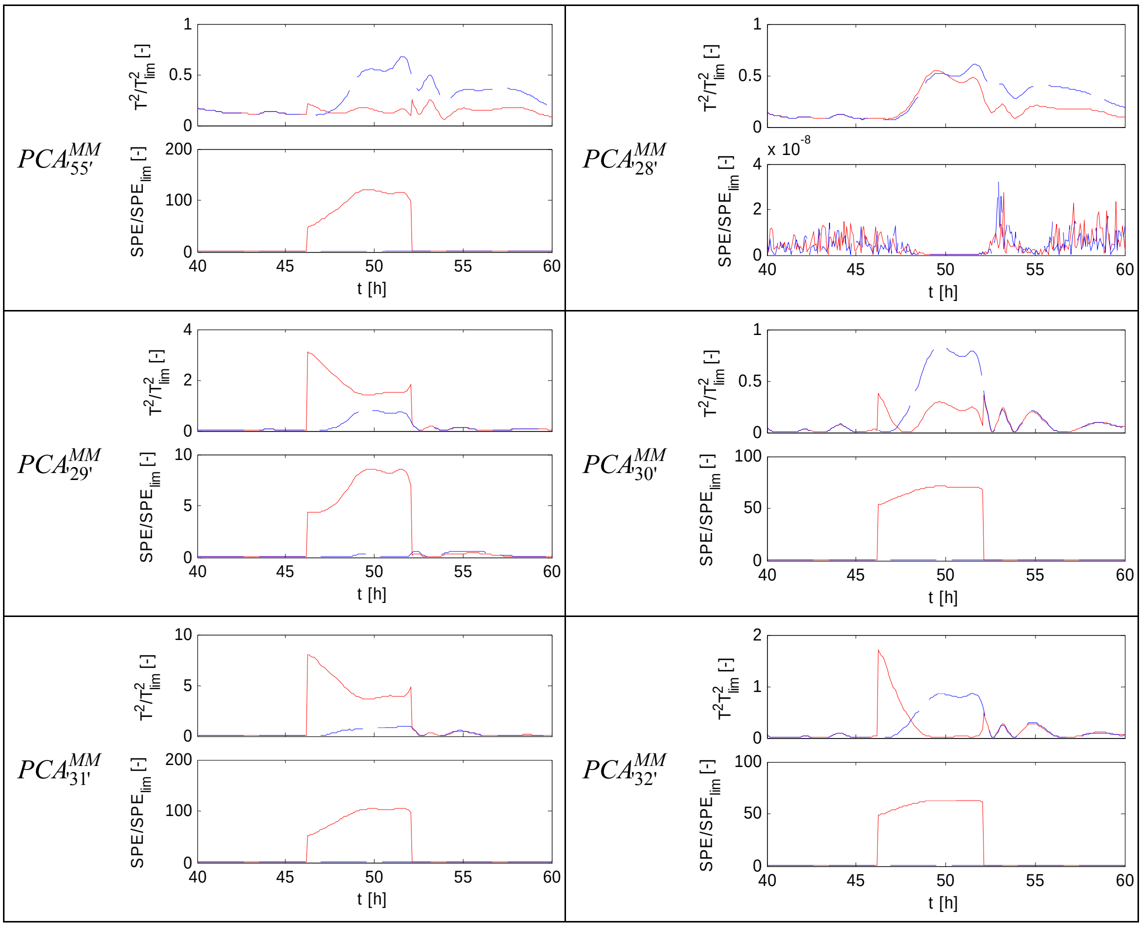

Figure 7. Leakage monitoring results by PCA^{MM} designed at selected nodes - tank screening effect (simulation without leakage (training) - blue; simulation with modelled leakage – red)

Notice that these models are identified based on measurements from two sides of retention tank. This is due to the fact that significant amount of water flowing in and out of the retention tank damps the influence of the modelled leakage on flow rates and nodal heads on the other side of the tank. Hence, e.g. $PCA^{MM}_{'28'}$ is not able to detect the modelled leakage. Since the values of components that the Monitoring Model composes of does not deviate much from the training data.

5.2.3 Isolated subnetwork effect

The third case (Fig. 5) illustrates the situation that the leakage is modelled (pipe between nodes ‘25’ and ‘28’) inside subnetwork that is isolated from the rest of the network (e.g outskirts of the town). The ‘isolated subnetwork’ means that it slightly (if ever) supply any other subnetworks and hence the abnormalities taking place in its interior do not radiate outside it. The isolation effect is often enhanced by screening effects water main and/or retention tank related.

Fig. 8 shows quality measures produced by PCA Monitoring Models located inside the isolated subnetwork. Notice that all of the PCA^{MM} indicate the abnormality. On the other hand, the example $PCA^{MM}_{'22'}$ identified based on measurements almost entirely gathered from outside of the isolated subnetwork does not detect any symptoms of abnormal operational state.

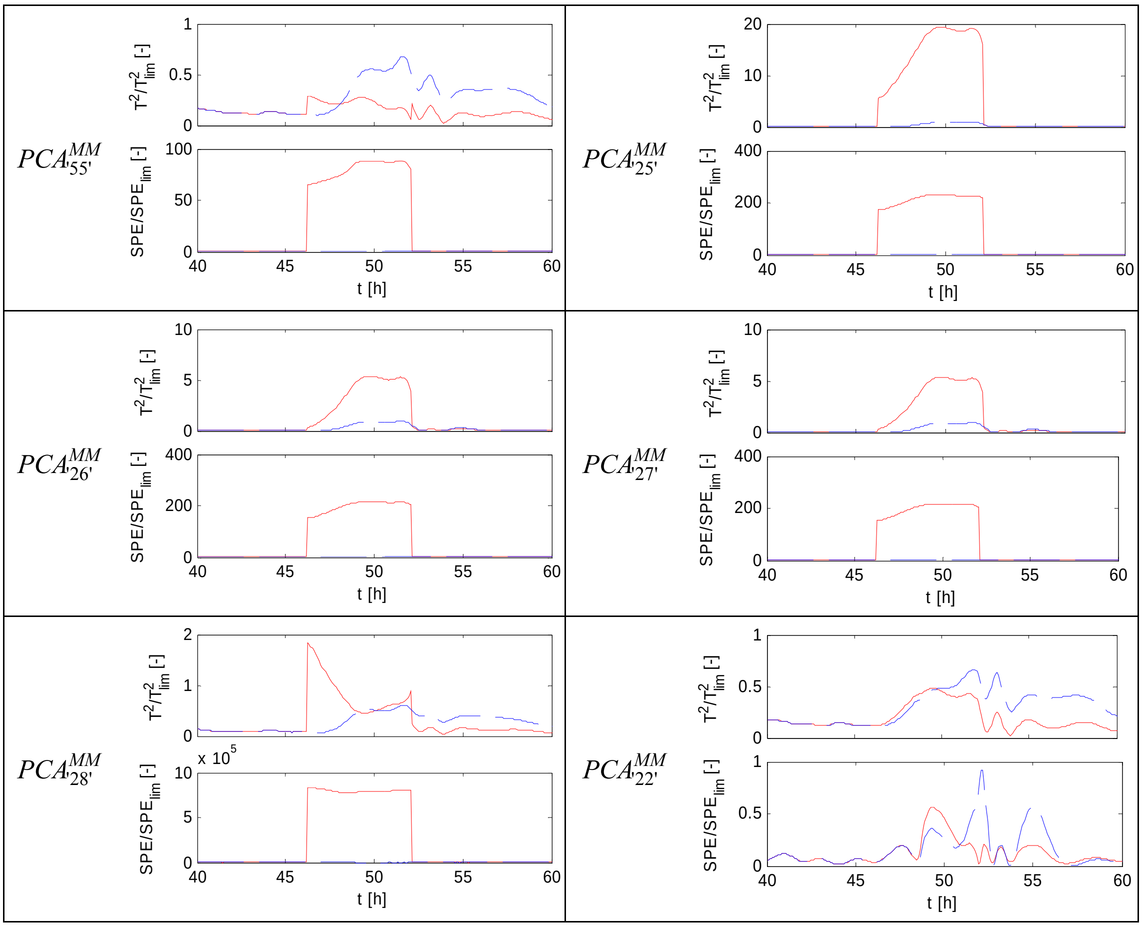

Figure 8. Leakage monitoring results by PCA^{MM} designed at selected nodes – isolated region effect (simulation without leakage (training) - blue; simulation with modelled leakage – red)

5.2.4 Distinguishing the subnetworks

Screening effects of the water mains, water tanks and isolated subnetwork are the serious disadvantage in leakage detection and localisation process at the first glance. However, skilfully utilisation of these features allows for improving the efficiency of faults monitoring process. Possessing the knowledge about DWDS characteristics (pipe flow rates, velocities, nodal heads, diameters of pipes, tank localisations etc.) one can divide the network into several independent subnetworks. The subnetworks should be selected in such a way to enable for independent (with respect to abnormalities significance indication) leakage detection inside them. Such an approach enables for placing much less measuring devices for detecting and then localising the leakages. As the results number of simulations and confronting them with network topology, three regional subnetworks have been distinguished (Fig. 9).

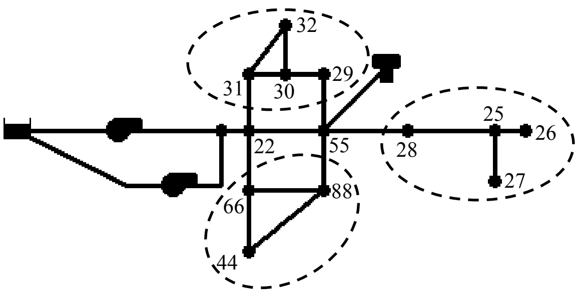

Figure 9. Selecting the regional subnetworks within the simple network

Another step after the detection of the leakage is its localization. It is clearly seen that there is strong correlation between values of both quality measures and distance from leakage and PCA Monitoring Model (Fig. 6, 7 and 8). Namely, in case of leakage the highest values of T^2 and SPE are produced by

PCA^{MM} in close neighbourhood to the source of leakage. Therefore, employment of at least two monitoring nodes within a single subnetwork enables for preliminary leakage localisation. Of course, the more PCA^{MM} are used, the more precise localisation is.

5.3 Chojnice case study

After presenting the fundamentals of described method, the MR-PCA is tested on Chojnice case study network.

5.3.1 Chojnice Drinking Water Distribution Systems

Chojnice is a city of forty thousand of citizens in northern Poland. Model of Chojnice DWDS [29] structure that sufficiently accurate for mentioned purposes is presented in Fig. 12 (Fig. 10 and Fig. 11). This model consists of 188 nodes, 284 pipes, two supply reservoirs in the system and one tank. Water is extracted from main reservoir by five pumps and provided to water treatment station. Model of Chojnice DWDS was built in Epanet simulator, while all the monitoring algorithms were implemented in Matlab. The monitoring points and leakages were selected to present the best advantages and disadvantages of proposed method. During the experiments pipe flows and nodal heads are determined by Model Predictive Control.

Based on the rules and observations described in Section 5.2 one may distinguish the subnetworks inside the network. Fig. 10 illustrates the example of such a subnetwork. Notice that PCA Monitoring Model 'placed' at node '147' is able to detect potential leakages inside selected area, only.

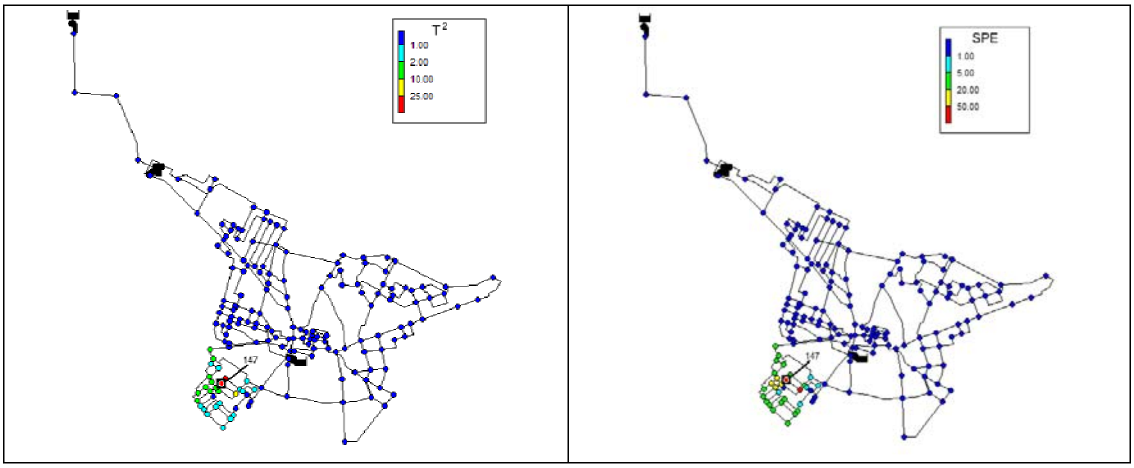

Figure 10. Monitoring at node '147' – range of possible leakages detection

The experiment has been carried out in such a way that water leakages were modelled at each of the network nodes/pipes one by one and hence selected monitoring node tried to detect these. The colours of the nodes in the figure indicate values of the ratios: $T^2_{'147'}(k)/T^2_{\lim,'147'}(\hat{s}_{'147'},\beta_{'147'})$ and $SPE_{'147'}(k)/SPE_{\lim,'147'}(\hat{s}_{'147'},\alpha_{'147'})$ what determine the ability to detect the abnormality by PCA Monitoring Model '147' ($PCA^{MM}_{'147'}$). The red nodes mean the highest detectability while the dark blue ones, the lowest.

5.3.2 Selection of monitoring nodes

Based on number o simulations it has been noticed that there are several PCA Monitoring Models that are able to detect leakages in much wider area then local subnetwork, only. In these models, measurements of water flows are performed in pipes with relatively small water flows comparing to their potential possibilities (regarding its diameters) and that they are located in close neighbourhoods to main streams. The examples of such place are PCA Monitoring Models designed at nodes: '029', '051', '068', '167'.

Fig. 11 illustrates the areas of potential leakages that are detectable by mentioned PCA Monitoring Models. Notice that potential leakages located in almost entire network might be detected by these nodes.

Another important observation is that any leakage at main streams are easily detectable by any PCA Monitoring Models, however it is not recommended to place the monitoring points at the main streams because of their limited ability to detect the other leakages. Nevertheless, the main streams are the crucial places of the network for the operators and hence they are most often under monitoring.

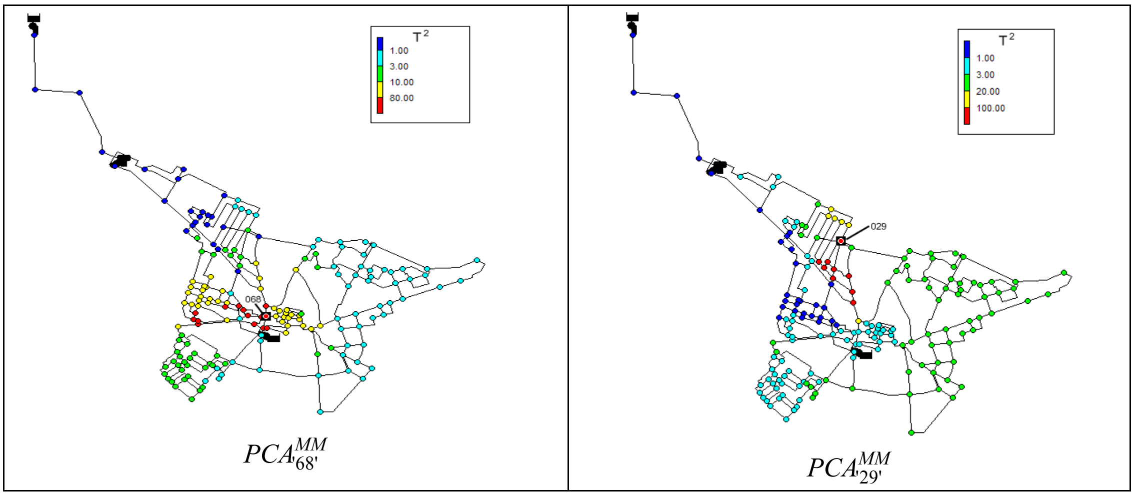

Figure 11. The leakages detectability by PCA Monitoring Models designed at nodes '068' and '029'

In the result of such an analysis five subnetworks has been selected for Chojnice network and two monitoring places at of the regions have been indicated. These are presented at Fig. 12.

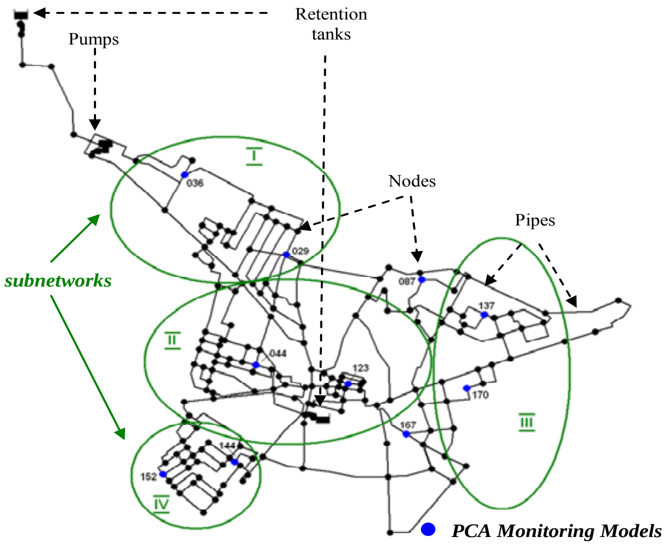

Figure 12. The Chojnice DWDS. Green ellipsoids mark the selected subnetworks while blue dots mark assumed PCA Monitoring Models.

5.3.3 Simulation results - leakage detection and localization

Fig. 13 presents the results of monitoring the Chojnice DWDS by PCA Monitoring Models designed at nodes marked in Fig. 12. The situation without leakage (training data) is marked blue lines in the figures, while simulations with modelled leakage are marked red. Simulations show that only PCA Monitoring Models build at nodes '152' and '144' unambiguously have indicated the failure (both *SPE* and T^2 have significantly exceeded the thresholds). These models consist of nodal heads and pipe flows measurements gathered inside one selected subnetwork, namely 'IV' (Fig. 12). It leads to conclusions that simulated leakage is located within this area. Moreover, *SPE* and T^2 produced by $PCA^{MM}_{'152'}$ are much greater then measures of $PCA^{MM}_{'144'}$, hence it might suggests that the abnormality took place closer to the node '152' (and it indeed is in this case), however one cannot treat this as a straight rule without taking into account the network topology and actual water distributions and its 'trace'. Example of tool enable for analysing the routes of water flowing inside the network is an algorithm making possible for water paths (routes) finding [30].

Besides of $PCA^{MM}_{'152'}$ and $PCA^{MM}_{'144'}$, also PCA Monitoring Model '029' indicates the leakage by its T^2 measure but as mentioned earlier it is the model of special sensitivity.

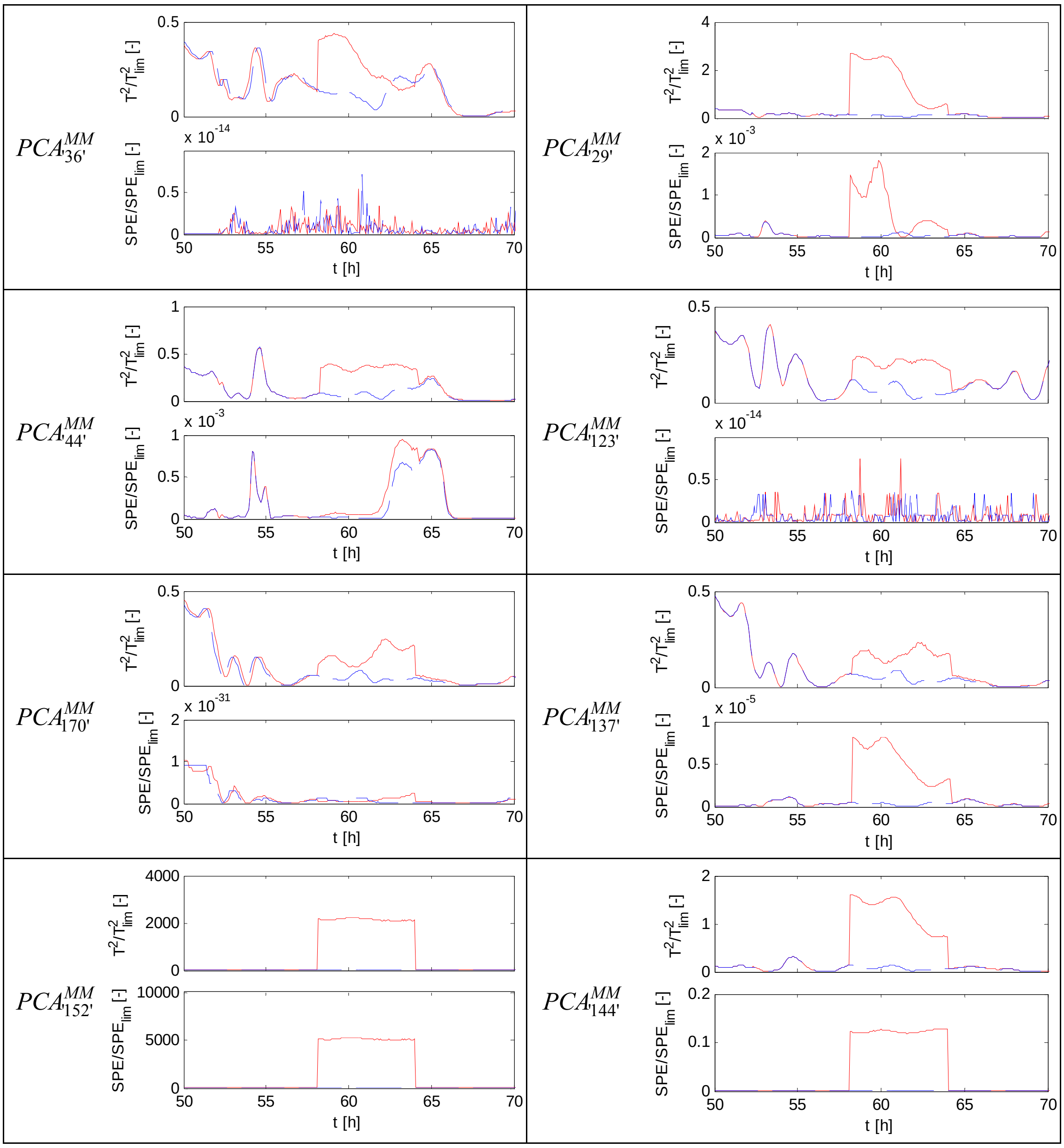

Figure 13. Leakage monitoring results by PCA Monitoring Models designed at selected nodes (simulation without leakage (training) - blue; simulation with modelled leakage – red)

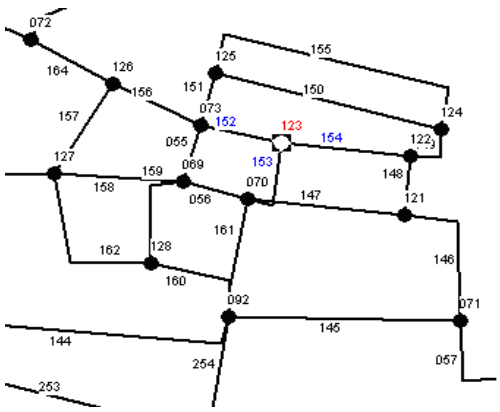

Figure 14. Fragment of the Chojnice network with broken water pipe flow sensor in pipe ‘152’

5.3.4 Simulation results - sensor fault

Very important for the monitoring process is quick and correctly distinguishing the process abnormalities (e.g. pipe water leakage, pump failure) from the sensor faults (e.g. sensor drift, outliers, missing data). The proposed method enables for distinguishing of such cases.

Following simulations present situation when flow rate sensor in pipe ‘152’ (Fig. 14) broke down. Broken sensor delivers measuring data for PCA Monitoring Model ‘123’. The modelled fault took place at about 50th hour of simulation.

Fig. 15 shows the monitoring results from all selected regional PCA Monitoring Models. The situation without failure is marked blue lines in the figures, while simulations with modelled sensor fault are marked red.

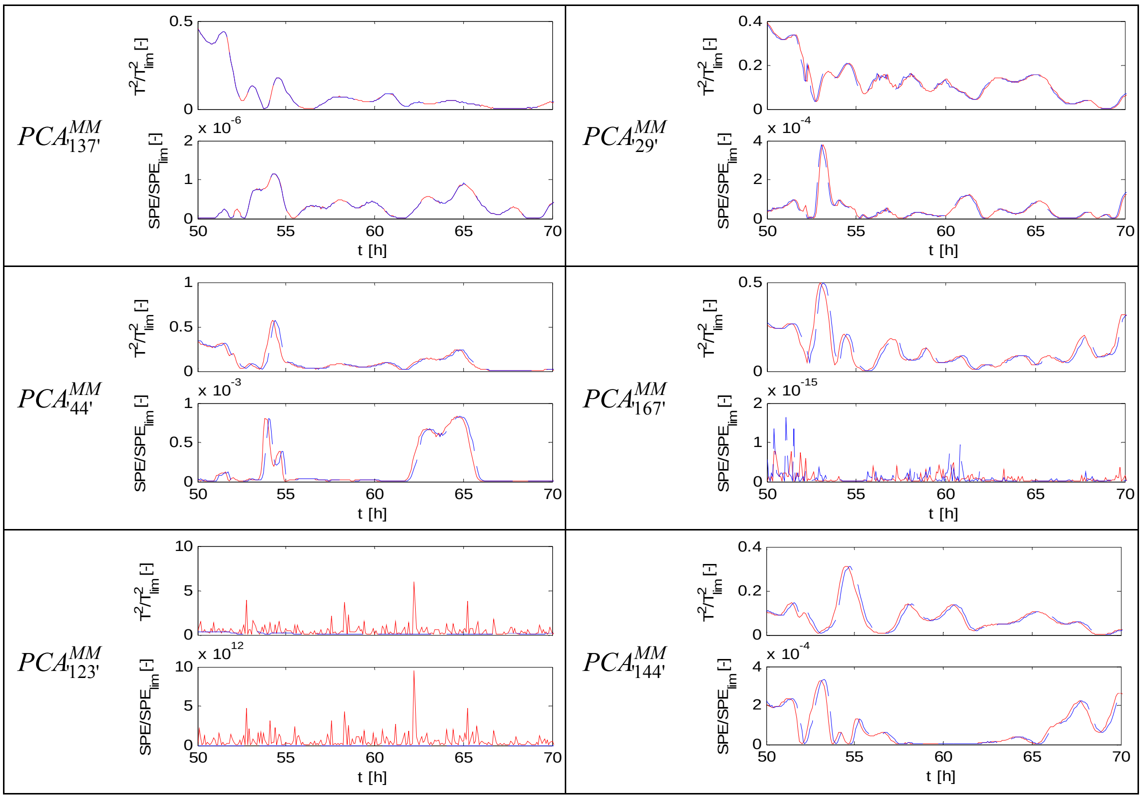

Figure 15. Sensor in pipe ‘152’ fault - monitoring results by PCA Monitoring Models designed at selected nodes (simulation without leakage (training) - blue; simulation with modelled leakage – red)

Based on the rules derived in the paper one might try to detect the sensor fault by analyzing the values of SPE and T^2. If the abnormalities detected by PCA^{MM} are caused by significant changing the operational state of the plant (e.g. by leakage) it should be noticeable at least by all sensors inside one of the selected subnetworks. Moreover, some of the very sensitive PCA^{MM} (e.g. ‘029’ as it was in case study when leakage was simulated inside this particular region) should detect it, as well. However, only PCA Monitoring Model ‘123’ indicates the fault. It suggests that one of the measurements is abrupt. In this case $PCA^{MM}_{'123'}$ consists of measurements from pipes ‘152’, ‘153’ ‘154’ and nodal head ‘123’ and hence we are able to state that one of these sensors probably broke down. Having more measuring devices in this subnetwork and so PCA Monitoring Models would enable for precise indicating the broken sensor by utilizing PCA features [31] or logical elimination.

6. Conclusions and Future Work

The paper has introduced a new approach to PCA based methods utilisation in detection and localisation of pipe leakage in Drinking Water Distribution System, namely MultiRegional Principal Component Analysis. In the first place MR-PCA approach to monitoring and diagnostics of network structured Distributed Systems was stated. The key idea is to use several regional PCA models (PCA Monitoring Models) identified on the basis of spatially local, available measurements to conclude about DS operational state, instead of single ‘global’ model. In particular, this enables for abnormality detection at least. Moreover the network topology of DS may be imbedded into the MR-PCA structure resulting in better diagnostic capabilities.

Since DWDS is a representative of network systems, its abnormal operational states detection and identification can be realised by MR-PCA. In the paper pipe leakages are assumed to be the only one process abnormalities. MR-PCA methodology illustrating simulations ware presented on a simple example of water network first. These, performed for a number of demand scenarios, confirmed the ability of MR-PCA to conduct the system diagnosis with regard to leakages detection and localisation. Furthermore additional phenomena were observed, providing abnormalities localisation complexity to be reduced, due to DWDS subnetworks distinguishing criteria. These are water main and retention tank screening effects, as well as isolated subnetwork effect. In practice only small part of the network variables can be directly measured. However, this does not constrain MR-PCA capabilities of leakages detection and localisation, as long as one is able to ensure at least two local PCA Monitoring Models per distinguished DWDS subnetwork, placed in their certain regions. Choice of these regions, i.e. DWDS sensors allocation is suggested from the MR-PCA abnormalities detectability point of view. Obtained results were successfully applied to case study DWDS Chojnice (Northern Poland).

At this state of the research the accurate localisation of the leakages is supervised by a man. In the future the neural networks and/or fuzzy clustering will be used to complete the process of automatic abnormalities localisation. Moreover, MR-PCA will be a part of supervised Fault Tolerant Model Predictive Control. Another subject of research in the field of fault detection and localisation will focus on identifying the process abnormality type e.g. pipe leakage, pump or valve breakdown. Obtained results are promising and have rather generic nature, hence might be transferred into other DSs e.g. pipeline systems, telecommunication systems, power systems etc, known as a network systems.

REFERENCES

1. JACKSON, J.E., **A User's Guide to Principal Components**, Wiley-Interscience (New York), 1991.
2. VÍTKOVSKÝ, J.P., SIMPSON, A.R. and LAMBERT, M.F., **Leak Detection and Calibration Using Transients and Genetic Algorithms,** J. Water Resour. Plng. and Mgmt 126, 2000.
3. POULAKIS, Z., VALOUGEORGIS, D. and PAPADIMITRIOU, C., **Leakage Detection in Water Pipe Networks Using a Bayesian Probabilistic Framework**, Probabilistic Engineering Mechanics 18, 2003.
4. MISIUNAS, D., **Failure Monitoring and Asset Condition Assessment in Water Supply Systems. Ph.D thesis**, Department of Industrial Electrical Engineering and Automation, Lund University (Lund), 2005.
5. BOROWA, A., MAZUR, K., GROCHOWSKI, M., BRDYŚ, M.A. and JEZIOR, K., **MultiRegional PCA for Leakage Detection and Localisation in DWDS – Approach**, (Eds.) KORBICZ J., PATAN K. and KOWAL M., **Fault Diagnosis and Fault Tolerant Control***,* Academic Publishing House EXIT (Warsaw), 2007.
6. JEZIOR, K., MAZUR, K., BOROWA, A., GROCHOWSKI, M. and BRDYS, M.A., **MultiRegional PCA for Leakage Detection and Localisation in DWDS – Chojnice Case Study,** (Eds.) KORBICZ J., PATAN K. and KOWAL M., **Fault Diagnosis and Fault Tolerant Control**, Academic Publishing House EXIT (Warsaw), 2007.
7. VENKATASUBRAMANIAN, V., RENGASWAMY, R., KAVURI, S.N. and YIN, K., **A Review of Process Fault Detection and Diagnosis Part III: Process History Based Methods**, Computers and Chemical Engineering 27, 2003.
8. ZHANG, X., PARISINI, T. and POLYCARPOU, M., **Adaptive Fault-tolerant Control of Nonlinear Uncertain Systems an Information Based Diagnostic Approach**, IEEE Transactions on Automatic Control 49, August 2004.
9. BRDYS, M.A., GROCHOWSKI, M., GMINSKI, T., KONARCZAK, K. and DREWA, M., **Hierarchical Predictive Control of Integrated Wastewater Treatment Systems**, Control Engineering Practice 16, 2008.
10. VENKATASUBRAMANIAN, V., RENGASWAMY, R., KAVURI, S.N. and YIN, K., **A Review of Process Fault Detection and Diagnosis Part I: Quantitative Model-based Methods**, Computers and Chemical Engineering 27, 2003.
11. VENKATASUBRAMANIAN, V., RENGASWAMY, R. and KAVURI, S.N., **A Review of Process**

Fault Detection and Diagnosis Part II: Qualitative Models and Search Strategies, Computers and Chemical Engineering 27, 2003.

12. LI, W., YUE, H.H., VALLE-CERVANTES, S. and QIN, S.J., **Recursive PCA for Adaptive Process Monitoring**, J. Process Control 10, 2000.
13. JACKSON, J.E. and MUDHOLKAR, G., **Control Procedures for Residuals Associated with Principal Component Analysis,** Technometrics 21, 1979.
14. ROSEN, C. and LENNOX, J. A., **Multivariate and Multiscale Monitoring of Wastewater Treatment Operation**, Wat. Res. 35, 2001.
15. LENNOX, J. A. and ROSEN, C., **Adaptive Multiscale Principal Components Analysis for Online Monitoring of Wastewater Treatment**, Water Science Technology 45, 2002.
16. BOROWA, A., MAZUR, K. and BRDYŚ, M.A., **Identyfikacja i diagnostyka procesu oczyszczania ścieków przy użyciu metody MS-PCA**, PAK 9 (BIS), 2005. (in Polish)
17. MAZUR, K., BOROWA, A. and BRDYŚ, M.A., **Diagnostyka i identyfikacja procesu przy użyciu AdMS-PCA**, PAK 9 (BIS), 2005. (in Polish)
18. MAZUR, K., BOROWA, A. and BRDYS, M.A., **Condition Monitoring Using PCA Based Method and Application to Watewater Treatment Plant Operation**, Proc. 1st IFAC Workshop on Applications of Large Scale Industrial Systems, ALSIS´06, Helsinki-Stockholm cruise liner, 2006.
19. BOROWA, A., BRDYS, M.A. and MAZUR, K., **Detection of Unmeasured Process Abnormalities in Wastewater Treatment Process Using MS-PCA**, Proc. 11th IFAC/IFORS/IMACS/IFIP Symposium on Large Scale Systems Theory and Applications, LSS'07, Gdansk, Poland, 2007.
20. QIU, B., PRINET, V., PERRIER, E. and MONGA, O., **Multi-Block PCA Method for Image Change Detection**, Proc. 12th International Conference on Image Analysis and Processing, ICIAP'03, Mantova, Italy, 2003.
21. WESTERHUIS, J.A., KOURTI, T. and MACGREGOR, J.F., **Analysis of multiblock and hierarchical PCA and PLS models**, Journal of Chemometrics 12, 1998.
22. BRDYS, M.A., DUZINKIEWICZ, K., CHANG, T., POLYCARPOU, M., WANG, Z., UBER, J., PROPATO, M., **Hierarchical control of integrated quality and quantity in drinking water distribution systems**, Proc. 1st IFAC International Conference on Technology, Automation and Control of Wastewater and Drinking Water Systems, TiASWiK'02, Gdansk Sobieszewo, Poland, 2002.
23. DUZINKIEWICZ, K., **Integrated Control of Drinking Water Supply and Distribution Systems. D. Sc. thesis**, AGH University of Science and Technology Publishers, 2005. (in Polish)
24. BRDYS, M.A., DUZINKIEWICZ, K., GROCHOWSKI, M, RUTKOWSKI, T., **Robust Estimation of Integrated Hydraulics and Parameters in Water Distribution Systems**, Proc. 4th ASCE Annual Water Distribution Systems Analysis, World Water and Environmental Resources Congress, Orlando, 2001.
25. BRDYS, M.A. and ULANICKI, B., **Operational Control of Water Systems: Structures, Algorithms and Applications**, PRENTICE HALL (New York, London, Toronto, Sydney, Tokyo), 1994.
26. MAYS, L.W., **Water Distribution Systems Handbook**, McGraw-Hill (New York), 2000.
27. ROSSMAN, L. A., **EPANET 2 users manual**, United States Environment Protection Agency, 2000.
28. POULAKIS, Z., VALOUGEORGIS, D. and PAPADIMITRIOU, C., **Leakage Detection in Water Pipe Networks Using a Bayesian Probabilistic Framework**, Probabilistic Engineering Mechanics 18, 2003.
29. EWALD, G., KUREK, W. and BRDYS, M.A., **Grid Implementation of Parallel Multiobjective Genetic Algorithm for Optimized Allocation of Chlorination Stations in Drinking Water Distribution Systems: Chojnice case study,** IEEE Transactions on System, Man, and Cybernetics-Part C: Applications and Reviews (submitted in 2007).
30. ZIEROLF, M., POLYCARPOU, M. and UBER, J., **Development and Autocalibration of an Input-Output Model of Chlorine Transport in Drinking Water Distribution Systems**, IEEE Transaction on Control Systems Technology 6, 1998.
31. DUNIA, R. and QIN, S.J., **Joint Diagnosis of Process and Sensor Faults Using Principal Component Analysis,** Control Engineering Practice 6, 1998.