

# Robust signal reconstruction for condition monitoring of industrial components via a modified Auto Associative Kernel Regression method

Piero Baraldi<sup>1</sup>, Francesco Di Maio<sup>1</sup>, Pietro Turati<sup>1</sup>, Enrico Zio<sup>1,2,3,\*</sup>

<sup>1</sup> Energy Department, Politecnico di Milano, Via Ponzio 34/3, 20133 Milan, Italy.

<sup>2</sup> Aramis S.r.l., Via Viviani 8, Milano, Italy.

<sup>3</sup> Chair on Systems Science and the Energetic Challenge, European Foundation for New Energy-Electricité de France, Ecole Centrale Paris and Supelec, France.

\* Corresponding author: [enrico.zio@polimi.it](mailto:enrico.zio@polimi.it), [enrico.zio@ecp.fr](mailto:enrico.zio@ecp.fr), [enrico.zio@supelec.fr](mailto:enrico.zio@supelec.fr)

## Abstract

In this work, we propose a modification of the traditional Auto Associative Kernel Regression (AAKR) method which enhances the signal reconstruction robustness, i.e., the capability of reconstructing abnormal signals to the values expected in normal conditions. The modification is based on the definition of a new procedure for the computation of the similarity between the present measurements and the historical patterns used to perform the signal reconstructions. The underlying conjecture for this is that malfunctions causing variations of a small number of signals are more frequent than those causing variations of a large number of signals. The proposed method has been applied to real normal condition data collected in an industrial plant for energy production. Its performance has been verified considering synthetic and real malfunctioning. The obtained results show an improvement in the early detection of abnormal conditions and the correct identification of the signals responsible of triggering the detection.

Keywords: AAKR; signal reconstruction; fault detection; kernel; robustness.

## 1. INTRODUCTION

Condition monitoring is used to assess the health state of industrial components and identify possible incipient faults (Venkatasubramanian et al., 2003a), (Nandi et al., 2005), (Jardine et al., 2006), (Wei, et al., 2013). For this, a model is usually built to reconstruct the values of the monitored signals expected in normal conditions of the components (Hameed, et al., 2009). During operation, observed signal measurements are compared with the reconstructions provided by the model: abnormal components conditions are detected when the reconstructions are remarkably different from the measurements. Data-driven (empirical) models are employed in those cases in which analytical models of the component behavior are not available and cannot be easily developed, whereas historical data collected during operation are available and limited number of hypotheses are required for building a data-driven model (Venkatasubramanian et al., 2003b), (Venkatasubramanian et al., 2003c), (Hines, et al., 2007).

Reconstruction methods are used in very different sectors, ranging from missing data reconstruction with various applications such as seismic data (Duijndam et al., 1999), (Liu et al., 2004), genetics (Kim et al., 2005), (Brock et al., 2008), climatology (Guiot et al., 2005), to financial applications, image analysis (Candès et al., 2006a), (Candès et al., 2006b) and condition monitoring of industrial components (Guglielmi et al., 1995), (Dunia et al., 1996), (Nabeshima et al., 1998), (Jack et al., 2002), (Harkat et al., 2007), (Chevalier et al., 2009), (King et al., 2009), (Baraldi et al., 2011a), (Baraldi et al., 2011b), (Baraldi et al., 2013b).

With respect to condition monitoring of industrial components, several methods have been shown to provide accurate reconstructions of the measured signals under normal operations. However, it has been noticed that these methods tend to suffer of high computational costs (Hunsop, 2011) and to be not robust (Baraldi et al., 2012). By robustness, here we intend the property that in case of abnormal conditions the reconstructions of the signals are properly estimating the values of the

signals expected in normal conditions of the components (Baraldi et al., 2011a). In particular, it has been shown that, especially when the measured signals are highly correlated, the reconstruction provided by AutoAssociative Kernel Regression (AAKR) method (Baraldi et al., 2011a) of an anomalous transient characterized by a drift of one signal can be not satisfactory for two reasons: 1) the reconstruction of the signal affected by the drift tends to assume values in the middle between the drifted and the expected values of the signal in normal conditions; 2) the reconstructions of other signals not affected by the drift tend, erroneously, to be different from the signal measurements, (this latter effect is usually referred to with the term ‘spill-over’). The consequence of 1) is a delay in the detection of abnormal conditions, whereas the consequence of 2) is that the condition monitoring system, although it correctly triggers an abnormal condition alarm, it is not able to correctly identify the signal that triggers the alarm.

These limitations of reconstruction models have been already studied in (Chevalier et al., 2009), (Baraldi et al., 2011a). Solutions to these problems have been proposed, which amount to try to exclude the signals with abnormal behaviors from the set of input signals used to perform the reconstruction. In (Baraldi et al., 2011b), (Baraldi et al., 2012), (Di Maio et al., 2013), the authors have propounded ensembles of reconstruction models handling different sets of input signals. In case of an anomaly impacting the behavior of a generic signal, only the few models fed by that signal provide non robust reconstructions, whereas all the other models provide correct reconstructions. Conversely, in (Baraldi et al., 2013a), an ensemble of different reconstruction models handling the same set of input signals is proposed. Another solution has been embraced in (Feller, 2013), whereby a ponderation matrix iteratively modifies its elements to reduce the contribution of abnormal signals to the reconstruction but the convergence of the method to correct signal reconstructions has not been yet demonstrated and all these solutions come at high computational costs.

The objective of the present work is to propose a robust signal reconstruction method with low computational cost and i) capable of early detection of abnormal conditions, ii) accurate in the reconstructions of the values of the signals impacted by the abnormal conditions and iii) resistant to the spill-over effect.

The proposed method is based on the modification of the measure of similarity used by the AAKR method: instead of measures of similarity based on Euclidean or Mahalanobis distances, the proposed method introduces a penalty vector which reduces the contribution provided by those signals which are expected to be impacted by the abnormal conditions. The rationale behind this proposition of the modification is the attempt to privilege those abnormal conditions caused by the most frequently expected malfunctions and failures. The performance of the proposed method has been verified considering i) synthetic malfunctioning simulated on real healthy data and ii) real abnormal conditions collected from an industrial plant for energy production (Baraldi et al., 2013b). The remainder of the paper is organized as follows. In Section 2, the fault detection problem is introduced. In Section 3, the AAKR method is briefly recalled. Section 4 shows the limitation of the traditional AAKR approach to condition monitoring and states the objectives of the present work. In Section 5, the proposed modification of the traditional AAKR is described and discussed. In Section 6, the application of the proposed method to a real case study concerning the monitoring of 6 signals in an industrial plant for energy production is presented. Finally, in Section 7 some conclusions are drawn.

## **2 FAULT DETECTION**

We consider condition monitoring scheme for fault detection as shown in Fig. 1. The (empirical) model reproducing the plant behavior in normal conditions receives in input the vector,  $\vec{x}^{obs}(t)$ , containing the actual observations of the  $J$  signals monitored at the present time,  $t$ , and produces

in output the reconstructions,  $\hat{\vec{x}}_{nc}(t)$ , i.e. the values that the signals are expected to have in normal conditions (Baraldi et al., 2012). If the actual conditions at the time  $t$  are instead, the residuals  $\Delta\vec{x} = \vec{x}^{obs}(t) - \hat{\vec{x}}_{nc}(t)$ , i.e. the variations between the observations and the reconstructions, are larger and can be detected by exceedance of a prefixed thresholds by at least one signal.

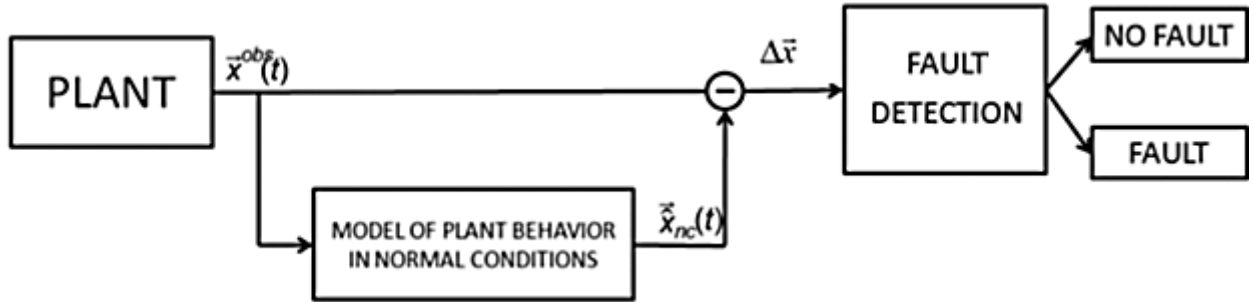


Fig. 1 Scheme of condition monitoring for fault detection

### 3. AUTOASSOCIATIVE KERNEL REGRESSION (AAKR)

Different empirical modeling techniques have been applied to the problem of signal reconstruction, such as AutoAssociative Kernel Regression (AAKR) (Hines et al., 2006), (Baraldi et al., 2011b), Principal Component Analysis (PCA) (Dunia et al., 1996), (Harkat et al., 2007), Robust Principal Component Analysis (Hubert et al., 2005), Fault-relevant PCA (FPCA) (Zhao et al., 2014), Partial Least Squares (PLS) (Muradore et al., 2012), Evolving Clustering Method (ECM) (Zio et al., 2011), Parzen Estimation (King et al., 2009), (Chen et al., 2013), fuzzy-logic-based systems (Marseguerra et al., 2003), AutoAssociative (AA) and Recurrent (R) Neural Networks (NN) (Bishop, 1995), (Guglielmi et al., 1995), (Nabeshima et al., 1998), (Samanta, 2004), (Worden et al., 2011). In this work, we consider AAKR which has been shown to provide more satisfactory performance than ECM and PCA (Chevalier et al., 2009) and is less computationally demanding than AANN, RNN (Baraldi et al., 2013a) and Parzen Estimation

(Hunsop 2011). Furthermore, notice that in many industrial applications of condition monitoring such as those in energy production plant it is common to periodically retrain or update the reconstruction models in order to follow the gradual modifications of the signal behavior due to slow degradation processes, maintenance interventions and minor plant reconfigurations. Thus, it is important to develop reconstruction models, such as AAKR, that are easy to train and characterized by few parameters which can be set by plant personnel.

The basic idea behind AAKR is to reconstruct at time  $t$  the values of the signals expected in normal conditions,  $\hat{x}_{nc}(t)$ , on the basis of a comparison of the currently observed signals measurements (also referred to as test pattern),  $\vec{x}^{obs}(t) = [x^{obs}(t, 1), \dots, x^{obs}(t, J)]$ , and of a set of historical signals measurements collected during normal condition of operation. In practice, AAKR performs a mapping from the space of the measurements of the signals  $\vec{x}^{obs}(t)$  to the space of the values of the signals expected in normal conditions,  $\hat{x}_{nc}(t)$ :

$$\hat{x}_{nc}(t) = \varphi(\vec{x}^{obs}(t) | \bar{X}^{obs-nc}) : \mathbb{R}^J \rightarrow \mathbb{R}^J \quad (1)$$

where  $\bar{X}^{obs-nc}$  indicates a  $N \times J$  matrix containing  $N$  historical observations of the  $J$  signals performed in normal-conditions. Since the mapping is independent from the present time,  $t$ , at which the signals observations are performed, the present time  $t$  will be omitted from the notations. Thus,  $x^{obs}(j)$ ,  $j = 1, \dots, J$ , indicates the value of signal  $j$  at the present time. The reconstruction of the expected values of the signals in normal conditions,  $\hat{x}_{nc} = [\hat{x}_{nc}(1), \dots, \hat{x}_{nc}(J)]$ , is performed as a weighted sum of the available historical observations; for the generic  $j$ -th element of  $\hat{x}_{nc}$ , we write:

$$\hat{x}_{nc}(j) = \frac{\sum_{k=1}^N w(k) \cdot x^{obs-nc}(k, j)}{\sum_{k=1}^N w(k)} \quad (2)$$

The weights,  $w(k)$ , measure the similarity between the test pattern,  $\vec{x}^{obs}$ , and the  $k$ -th historical observation vector,  $\vec{x}^{obs-nc}(k)$ . They are evaluated through a kernel,  $Ker$ , i. e., a scalar function which can be written as a dot product (Burges, 1998), (Müller et al., 2001), (Widodo et al., 2007). From the mathematical point of view, a Kernel is a function:

$$Ker: \mathbb{R}^J \times \mathbb{R}^J \rightarrow \mathbb{R} \text{ s. t. } \exists \boldsymbol{\phi}: \mathbb{R}^J \rightarrow \mathcal{H}, Ker(\vec{x}, \vec{y}) = \langle \boldsymbol{\phi}(\vec{x}), \boldsymbol{\phi}(\vec{y}) \rangle \quad (3)$$

where  $\boldsymbol{\phi}$  is a map from the observation space  $\mathbb{R}^J$  in a (possibly countable infinite dimensional) Euclidean space  $\mathcal{H}$  and  $\langle ., . \rangle$  denotes the dot product. Traditional AAKR adopts as  $Ker$  function the Gaussian Radial Basis Function (RBF) with bandwidth parameter  $h$ , i.e.:

$$w(k) = Ker(\vec{x}^{obs}, \vec{x}^{obs-nc}(k)) = \frac{1}{\sqrt{2\pi h^2}} e^{-\frac{\|\vec{x}^{obs} - \vec{x}^{obs-nc}(k)\|^2}{2h^2}} \quad (4)$$

Notice that, according to Mercer's theorem (Burges, 1998), eq. (4) can be seen as a dot product in a countable infinite dimensional Euclidean space:

$$e^{-\frac{1}{2}\|\vec{x}^{obs} - \vec{x}^{obs-nc}(k)\|_2^2} = \sum_{i=0}^{\infty} \frac{\left(\vec{x}^{obsT} \vec{x}^{obs-nc}(k)\right)^i}{i!} e^{-\frac{1}{2}\|\vec{x}^{obs}\|_2^2} e^{-\frac{1}{2}\|\vec{x}^{obs-nc}(k)\|_2^2} \quad (5)$$

In fault detection applications, Euclidean and Mahalanobis distances are typically used to compute the distance in the Gaussian RBF. In this work, in order to account for differences in the scale and variability of the different signals, a Mahalanobis distance is used, defined by the covariance matrix,  $S$ , such that:

$$\|\vec{x}^{obs} - \vec{x}^{obs-nc}(k)\|_{mahal}^2 = \left(\vec{x}^{obs} - \vec{x}^{obs-nc}(k)\right)^T S^{-1} \left(\vec{x}^{obs} - \vec{x}^{obs-nc}(k)\right) \quad (6)$$

Assuming independence between the signals,  $S$  is given by:

$$S = \begin{bmatrix} \sigma_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_j^2 \end{bmatrix} \quad (7)$$

where  $\sigma_j^2$  indicates the estimated variance of signal  $j$  in the historical observations. Alternatively, instead of using (6) and (7), one can obtain the same results by mapping the data into a normalized space:

$$y(j) = \frac{x(j) - \mu_j}{\sigma_j} \quad (8)$$

where  $\mu_j$  is the mean value of signal  $j$  in the historical dataset, and by applying a Gaussian kernel with Euclidean distance in the normalized space.

#### **4. LIMITATIONS IN THE USE OF AAKR FOR SIGNAL RECONSTRUCTION**

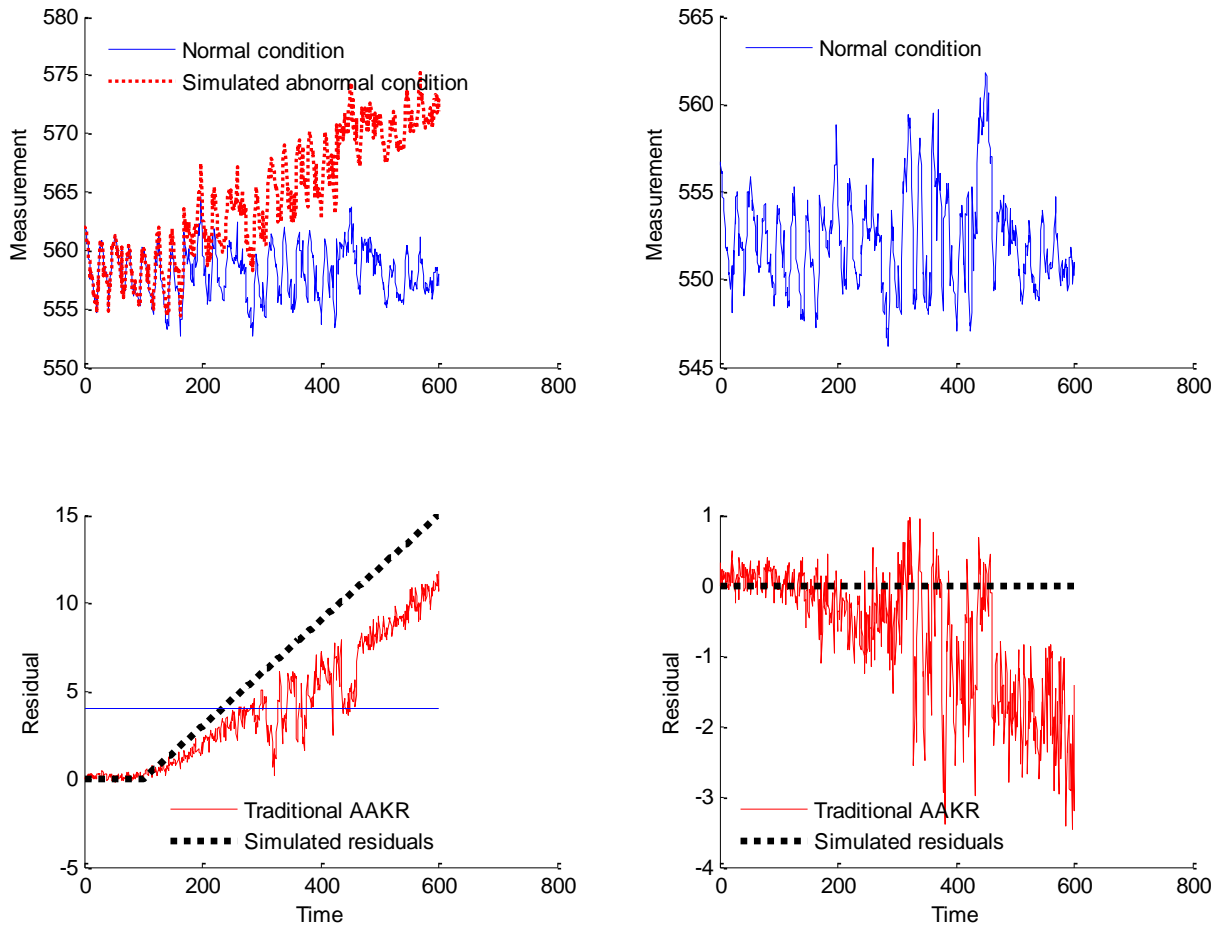
In the context of safety-relevant industries like oil & gas, nuclear and avionics, the need of high reliability targets and the availability of relatively affordable monitoring technology has pushed the installation of a great number of sensors to facilitate the monitoring of physical quantities (i.e. temperature, pressure, flows, vibrations, etc. ) and of the operational condition of their components (Lee, 1997), (Roverso et al., 2007). Among all these sensors, due to the redundancy of sensors mounted on each component, it is not rare that the collected signals are characterized by high Pearson correlation coefficients. In this context of correlated signals, AAKR reconstructions performed in abnormal conditions have been shown to be not satisfactory from the point of view of the robustness: the obtained reconstructions are not accurately estimating the values of the signals expected in normal conditions (Baraldi et al., 2012). This effect is well illustrated by the following case study concerning the monitoring of a component of a plant for the production of energy (Baraldi et al., 2013b). A dataset containing the real evolution of 6 highly correlated signals (Table 1) monitoring the temperatures in Celsius degree in different location of the component for a period of 1 year with sampling frequency of 5200 measurements/year has been used to develop an AAKR reconstruction model. Then, in order to



artificially simulate an abnormal condition, a linearly increasing drift has been added to the real values of one of the six signals for a period of 600 time steps: these drifted data have not been used to develop the AAKR model. Fig. 2 (top) shows the drifted signal, whereas Fig. 2 (bottom) shows the residuals of the reconstruction of the drifted signal (left) and of another signal not impacted by the abnormal conditions, i.e. not drifted (right). Notice that the obtained reconstructions are not robust: 1) the residuals of the drifted signal are not following the applied drift and 2) the residuals of the other signal are erroneously deviating from 0.

	<b>S1</b>	<b>S2</b>	<b>S3</b>	<b>S4</b>	<b>S5</b>	<b>S6</b>
<b>S1</b>	1	0.97	0.98	0.98	0.99	0.98
<b>S2</b>	0.97	1	0.95	0.99	0.98	0.96
<b>S3</b>	0.98	0.95	1	0.96	0.99	0.99
<b>S4</b>	0.98	0.99	0.96	1	0.98	0.97
<b>S5</b>	0.99	0.98	0.99	0.98	1	0.99
<b>S6</b>	0.98	0.96	0.99	0.97	0.99	1

**Table 1 Pearson correlation coefficients between 6 measured signals in an industrial plants for energy production.**



**Fig. 2** Top (left): the continuous line represents signal 1 values in normal conditions, the dotted line the signal values in the simulated abnormal conditions; top (right): evolution of signal 3 (not affected by the abnormal conditions). Bottom: the continuous line represents the residuals obtained by applying the traditional AAKR reconstruction, the dotted line the residual which would be obtained by a model able to perfectly reconstruct the signal behavior in normal conditions.

From the practical point of view of the fault detection, two problems arise in relation to the low robustness of the reconstruction:

- 1) delay in the detection of the abnormal condition (an alarm is usually triggered when the residuals exceed prefixed thresholds).
- 2) detection of abnormal conditions on signals different from those which are actually impacted by the abnormal behavior (spill-over).

With regards to the latter, the identification of the signals impacted by the abnormal conditions is critical since it can allow to identify the cause of abnormality and, thus, to properly plan the maintenance intervention.

## 5. MODIFIED AAKR

In order to enhance the AAKR robustness, we propose to modify the computation of the weights as the traditional AAKR (eq. (2)). The basic ideas underling the proposed modification are (a) to identify the signals affected by abnormal behaviors and (b) to reduce their importance in the computation of the similarity between the test pattern and the historical observations.

With respect to (a), we assume that the probability of occurrence of a fault causing variations on a large number of signals is lower than that of a fault causing variations on a small number of signals:

$$P(S_{fault\ 1}) \leq P(S_{fault\ 2}) \text{ if } |S_{fault\ 2}| \leq |S_{fault\ 1}| \quad (9)$$

where  $S_{fault\ 1}$  and  $S_{fault\ 2}$  indicate the sets of signals affected by variations due to the abnormal (faulty) conditions, and  $|S_{fault\ 1}|$  and  $|S_{fault\ 2}|$  their cardinality. If we consider, for example, the problem of sensor failures, it is reasonable to assume that the probability of having  $N_1$  faulty sensors at the same time is lower than that of having a lower number of faulty sensors,  $N_2 \leq N_1$ , at the same time. Notice that this assumption is realistic for several abnormal conditions such as those affecting sensors or small failures which can occur in single components of a complex industrial systems or in systems characterized by redundancies in the design. On the other hand, notice that the assumption does not possess universality, since it does not apply to other cases such as common cause failures or to some failures in network systems such as electrical grids. The proposed procedure computes the similarity measure between the observation,  $\vec{x}^{obs}$ , and the

generic  $k$ -th historical observation,  $\vec{x}^{obs-nc}(k)$ , according to two steps: (a) a pre-processing step consisting in the projection of  $\vec{x}^{obs}$  and  $\vec{x}^{obs-nc}(k)$  in a new space defined by a penalty vector,  $\vec{p} = [p(1), \dots, p(J)]$ , with increasing entries, i.e.,  $p(1) \leq \dots \leq p(J)$  and (b) the application of the Gaussian RBF kernel in the new space.

Step (a) is based on:

- computing the vector of the absolute values of the normalized differences between  $\vec{x}^{obs}$  and  $\vec{x}^{obs-nc}(k)$ :

$$|\vec{x}^{obs} - \vec{x}^{obs-nc}(k)|_{\sigma} = \left( \left| \frac{\vec{x}^{obs}(1) - \vec{x}^{obs-nc}(k, 1)}{\sigma_1} \right|, \dots, \left| \frac{\vec{x}^{obs}(J) - \vec{x}^{obs-nc}(k, J)}{\sigma_J} \right| \right) \quad (10)$$

- defining a permutation matrix,  $P_{perm}$ , i.e. a matrix which, when multiplied to a vector, only modifies the order of the vector components; in our procedure, we define a matrix,  $P_{perm}$ , such that when it is applied to the vector  $|\vec{x}^{obs} - \vec{x}^{obs-nc}(k)|_{\sigma}$ , the components of the obtained vector are the same of that of  $|\vec{x}^{obs} - \vec{x}^{obs-nc}(k)|_{\sigma}$ , but they appear in a decreasing order, i.e. the first component is the one with the largest difference in  $|\vec{x}^{obs} - \vec{x}^{obs-nc}(k)|_{\sigma}$ ;
- defining a diagonal matrix having increasing entries on its diagonal:

$$D_{\vec{p}} = \begin{bmatrix} \sqrt{p(1)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sqrt{p(J)} \end{bmatrix} \quad (11)$$

where the vector  $\vec{p} = [p(1), \dots, p(J)] = tr(D_{\vec{p}})$  will be referred to as penalty vector;

- projecting  $\vec{x}^{obs}$  and  $\vec{x}^{obs-nc}(k)$  in a new space defined by the transformation:

$$\begin{aligned} \boldsymbol{\psi} : \mathbb{R}^J &\rightarrow \mathbb{R}^J \\ \boldsymbol{\psi}(\vec{x}^{obs}) &= D_{\vec{p}} P_{perm} \vec{x}^{obs} \end{aligned} \quad (12)$$

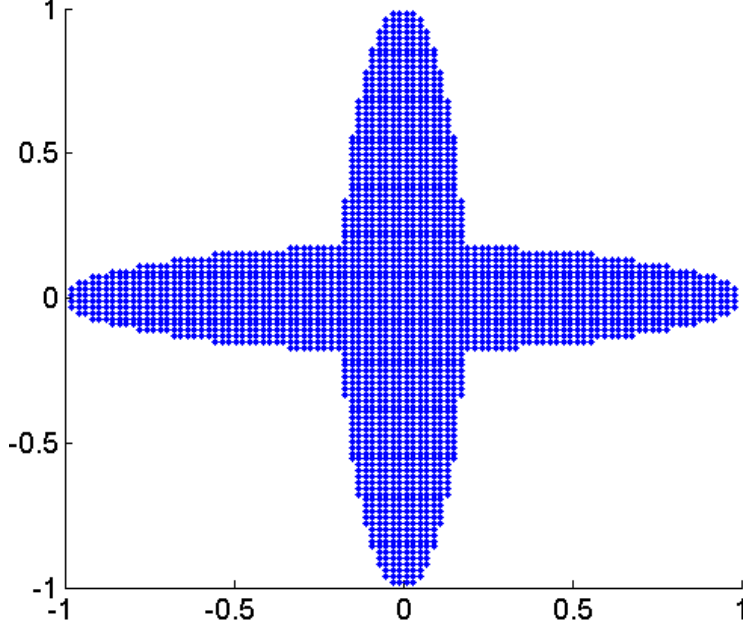
$$\boldsymbol{\psi}(\vec{x}^{obs-nc}) = D_{\vec{p}} P_{perm} \vec{x}^{obs-nc}$$

In step (b), we apply to  $\boldsymbol{\psi}(\vec{x}^{obs})$  and  $\boldsymbol{\psi}(\vec{x}^{obs-nc})$  the Gaussian kernel with Euclidean distance:

$$w(k) = Ker(\boldsymbol{\psi}(\vec{x}^{obs}), \boldsymbol{\psi}(\vec{x}^{obs-nc})) = \frac{1}{\sqrt{2\pi h^2}} e^{-\frac{\|\boldsymbol{\psi}(\vec{x}^{obs}) - \boldsymbol{\psi}(\vec{x}^{obs-nc})\|^2}{2h^2}} \quad (13)$$

Notice that, given a couple of observations  $\vec{x}^{obs}$  and  $\vec{x}^{obs-nc}$ , the effect of step (a) is to equally magnify the same signals in both the observations. In this work, we do not investigate neither if the sequential application of steps (a) and (b) defines a kernel function according to eq. (3) nor the mathematical properties of the non-homogeneous transformation  $\boldsymbol{\psi}$ . Conversely, we limit our scope in showing the effects of the new procedure for similarity weights calculation: we consider two different historical patterns,  $\vec{x}^{obs-nc}(k_1)$  and  $\vec{x}^{obs-nc}(k_2)$ , characterized by similar Euclidean distances from a test pattern, but with  $\vec{x}^{obs-nc}(k_1)$  characterized by a lower number of signals remarkably different from that of the test pattern. In this case, the effect of the penalty vector is to assign to  $\vec{x}^{obs-nc}(k_1)$  an higher similarity measure value than that assigned to  $\vec{x}^{obs-nc}(k_2)$ .

Fig. 3 shows the locus of points characterized by the same similarity to the origin (0,0) in a 2-dimensional space using a penalization vector  $\vec{p} = [1, 30]$ . The obtained surface is very different from the circle which would be obtained by using an Euclidean distance. As expected, such modification introduces a preference during the reconstruction of  $\vec{x}^{obs}$  for vector of  $\vec{X}^{obs-nc}$  that deviate from it in a lower number of components.



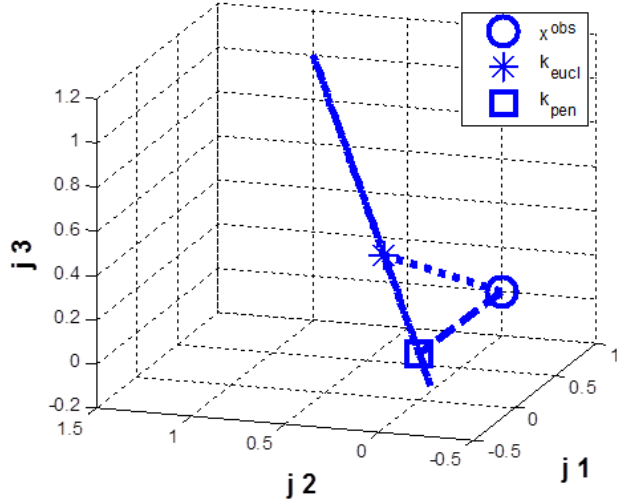
**Fig. 3 2D locus of points having similarity from the origin greater than a set value according to a penalty vector  $\vec{p} = [1 \ 30]$ .**

The application of the proposed method is also exemplified with reference to a numerical conceptual case study. Let us assume to have available an infinite dataset of historical data containing all the possible normal conditions of 3 signals with degree of correlation equal to 1, i.e.  $(\vec{x}^{obs-nc}(k) = [k, k, k] \in \mathbb{R}^3, k \in (-2, 2))$  and to have a test pattern  $\vec{x}^{obs} = [1, 0, 0]$  containing the three signal measurements at the present time. According to the traditional AAKR procedure, the Euclidean distance between the test pattern and the training pattern  $[k, k, k]$  is:

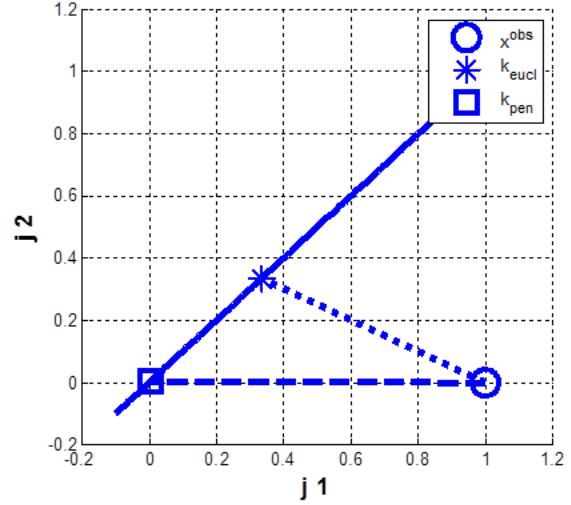
$$d^2(k) = (k - 1)^2 + k^2 + k^2$$

whose minimum value is obtained for  $k_{min} = \arg \min\{d^2(k)\} = 1/3$ . Thus, the pattern  $\vec{x}_{nearest\_eucl} = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right]$  is the nearest pattern to the test pattern (Fig. 4 and Fig. 5). Hence,  $\vec{x}_{nearest\_eucl}$  will be characterized by the highest associated weight, and the signal reconstruction will be close to it. This reconstruction suggests that there is an abnormal condition impacting all three signals at the same time. Let us now consider the reconstruction performed by using the

proposed method with a penalization vector  $\vec{p} = [1, 10, 100]$ . In this case, the most similar pattern obtained by applying the same procedure is  $\vec{x}_{nearest\_pen} = \left[ \frac{1}{111}, \frac{1}{111}, \frac{1}{111} \right]$  and the signal reconstruction will be close to it. Hence, according to this approach, the only significant residual is detected on the first signal (Fig. 4 and Fig. 5).



**Fig. 4** Reconstruction of an abnormal condition (circle) performed with the Euclidean AAKR (star) and the penalty AAKR (squared).



**Fig. 5** Projection of the reconstruction on the plain described by two signals.

Notice that the basic difference between the reconstructions provided by the traditional and the modified AAKR algorithm is the hypothesis of the latter that an abnormal condition involving few signals is more probable than an abnormal condition involving a lot of signals. Coherently with that hypothesis, the modified AAKR identifies an abnormal behavior only on one signal.

## 6. APPLICATION OF THE METHOD TO A SYNTHETIC AND A REAL CASE STUDY

### 6.1 Real normal and synthetic abnormal conditions

The data previously introduced in Section 4 are here used to test the performance of the modified AAKR method. The dataset containing the measurements of 6 highly correlated signals in 5200 consecutive time steps has been divided into 3 subsets:

- Training set  $\bar{X}_{train}^{obs-nc} \in \mathbb{R}^{2500 \times 6}$ , used as historical dataset to perform the reconstruction;
- Validation set  $\bar{X}_{val}^{obs-nc} \in \mathbb{R}^{1700 \times 6}$ , used for the setting of the optimal parameter values;
- Test set,  $\bar{X}_{test}^{obs-nc} \in \mathbb{R}^{1000 \times 6}$ , used for testing the performance of the method.

Since all the available data refer to components operating in normal conditions, abnormal conditions have been simulated in the test set by considering sensors failures. Assuming that sensors failures are independent events with constant probability,  $q$ , the following procedure has been applied:

- for each signal and each test pattern, a random number,  $r$ , has been sampled from an uniform distribution in  $[0,1]$ . If  $r < q$ , the sensor is assumed to be failed and a deviation is simulated from a bimodal uniform distribution  $U([-10, -4] \cup [4,10])$  and added to the sensor measurement.

Thus, the number of signals affected by a deviation in each test pattern is distributed according to a Binomial distribution  $B(q, 6)$ . The deviation intensity has been sampled from the uniform distribution  $U([-10, -4] \cup [4,10])$  in order to avoid to confuse the added deviations with the measurement noise which has been estimated to be a Gaussian noise with standard deviation equal to 1.5.

The obtained test set,  $\bar{X}_{err}$ , containing both normal and abnormal conditions patterns,  $\vec{x}_{err}(k)$ , has been used to verify the performance of the traditional AAKR method and the modified AAKR with different choices of penalization vector:

- Linear  $\vec{p} = [m, 2m, \dots, 6m]$ , with  $m \in \{2, 4, 8, 16, 32, 64\}$



- Exponential  $\vec{p} = [\kappa, \kappa^2, \dots, \kappa^6]$ , with  $\kappa \in \{2, 5, 10, 15, 20\}$
- Cliff Diving Competition ranking  $\vec{p} = [8 \ 20 \ 50 \ 90 \ 160 \ 350]$ .

In all cases, the optimal bandwidth parameter,  $h$ , has been identified by minimizing the Mean Square Error (MSE) of the reconstructions on the validation set,  $\bar{X}_{val}^{obs-nc}$ :

$$MSE(h) = \frac{\sum_{k=1}^{N_{val}} \sum_{s=1}^J \left( \hat{x}_{val}(k, s) - \bar{x}_{val}^{obs-nc}(k, s) \right)^2}{N_{val}} \quad (14)$$

Then, for each test pattern,  $\vec{x}_{err}(k)$ , of  $\bar{X}_{err}$ ,  $k=1, \dots, 1000$ , the reconstruction  $\hat{x}_{nc}(k)$  of the signals values expected in normal conditions has been performed, and the residuals  $\hat{r}(k)$  computed. Typically, an abnormal condition is detected and an alarm is triggered when at least one of the signal residuals exceeds a prefixed threshold. According to the level of conservatism demanded by the problem at hand, and the fractions of false and missing alarms which can be accepted in the considered industrial context, different strategies to fix the thresholds can be applied. In this work, we have set the threshold equal to 4°C for all the signals, i.e. if  $|\hat{r}(k, j)| > 4$ , an abnormal condition involving signal  $j$  is detected. The value of 4 °C has been arbitrarily chosen by looking for a satisfactory trade-off between the fraction of false and missing alarms on the validation set. Further investigations on the strategy for abnormal condition detection and alarm triggering, which are outside the scope of the present work, can be found in (Hines et al., 2006) and (Di Maio et al., 2013).

In order to evaluate the performance in the reconstruction, we have considered 4 different possible cases for each test pattern:

- 1) presence of both false and missed alarms. There is at least one signal for which an abnormal condition is erroneously detected ( $|\hat{r}(k, j)| > 4$ ) when no sensor failure has been simulated (false alarm) and, at the same time, at least one signal for which an

abnormal condition is erroneously not detected ( $|\hat{r}(k, j)| \leq 4$ ), when actually a sensor failure has been simulated (missed alarm);

- 2) presence of only a false alarm. At least one false alarm, no missed alarms;
- 3) presence of only a missing alarm. At least one missed alarm, no false alarms;
- 4) correct identification (OK). Neither false nor missed alarms.

Table 2 reports the performance of the traditional and modified AAKR reconstruction methods in terms of fraction of test patterns in which the application of the detection scheme leads to one of the four categories (1-4), considering different choices of the penalty vector. For the cases of exponential and linear penalty vectors, only the results obtained for  $k=10$  and  $m=8$ , which correspond to the best performance, are reported.

	<b>PENALTY VECTOR</b>	<b>OK</b>	<b>MISSED ALARMS</b>	<b>FALSE ALARMS</b>	<b>MISSED AND FALSE ALARMS</b>
<b>MODIFIED AAKR</b>	<b>EXPONENTIAL <math>\kappa = 10</math></b>	0.885	0.089	0.008	0.018
<b>MODIFIED AAKR</b>	<b>DIVING COMPETITION RANKING</b>	0.669	0.300	0.001	0.030
<b>MODIFIED AAKR</b>	<b>LINEAR <math>m = 8</math></b>	0.585	0.375	0.001	0.039
<b>TRADITIONAL AAKR</b>	<b>EXPONENTIAL <math>\kappa = 1</math></b>	0.500	0.446	0.003	0.051

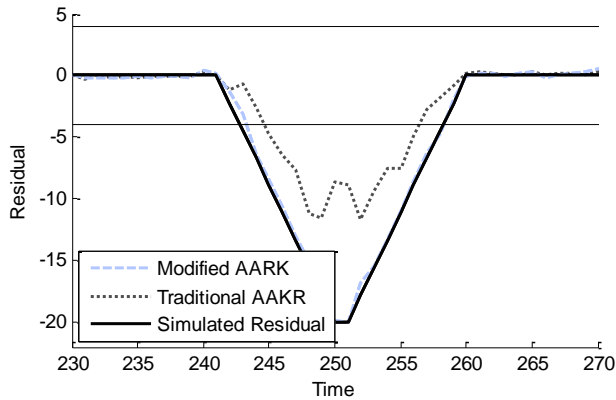
**Table 2 Fraction of test patterns correctly identified (OK), missed, false and both missed and false alarms.**

The most satisfactory method from the point of view of the highest fraction of correct identification and the lowest fraction of missed and false alarms is the modified AAKR with an exponential penalty vector, whereas the less satisfactory is the traditional AAKR. Furthermore, notice that the lowest total fraction of false alarms, which can be obtained from the sum of the false and false and missed alarms (columns 5 and 6 of Table 2), is obtained by the modified AAKR with exponential weight vector.

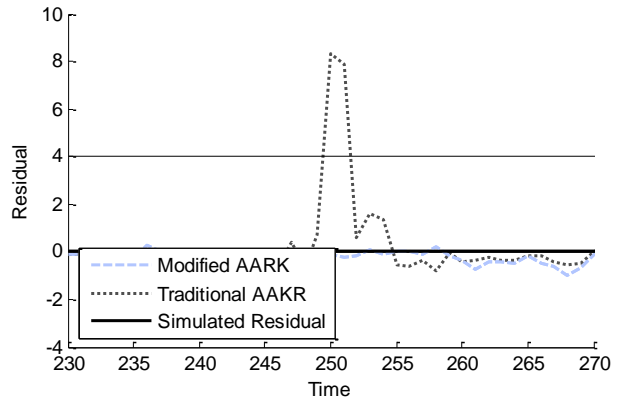
The performance of the modified AAKR method with exponential penalty vector and  $k=10$  has also been verified on different types of synthetic malfunctioning which causes abnormal signal

behaviors such as spikes, increasing noises and gradual drifts. In practice, these abnormal conditions have been simulated following the same procedure described in Section 4 (Fig. 2), i.e. by adding to the normal condition data the simulated failure on one single signal.

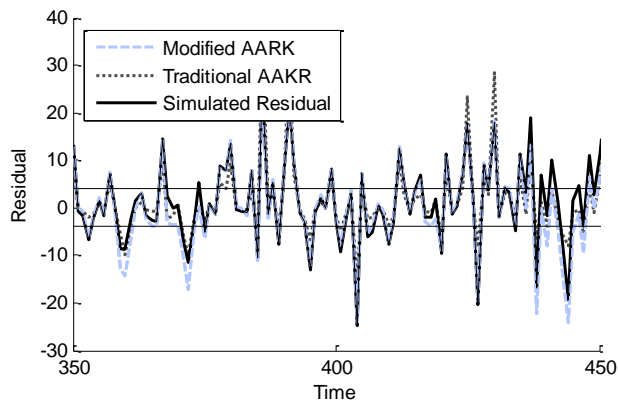
Fig. 6, 7, 8,9,10 and 11 show the obtained residuals for the signal affected by the synthetic malfunctioning (left) and another signal not affected by malfunctioning (right). Notice that the modified AAKR provides an early detection of the abnormal conditions when dealing with synthetic drift (with the obtained residuals almost overlapping the simulated drift) and provides very accurate reconstructions of the signal not affected by any of the synthetic malfunctioning (with residuals close to 0). Hence, the modified AAKR allows to reduce the time necessary for the malfunction detection with respect to the traditional AAKR and is more resistant to the spill-over effect.



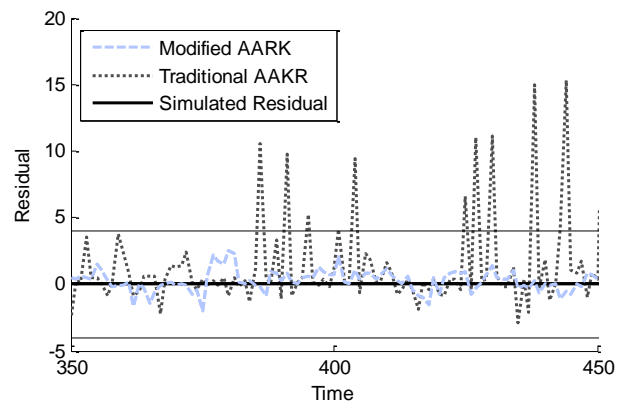
**Fig. 6** Residuals obtained when the measured signal is affected by a spike. Black dotted line: residual provided by the traditional AAKR, dashed grey line by the modified AAKR. The continuous line represents the spike added to the signal value.



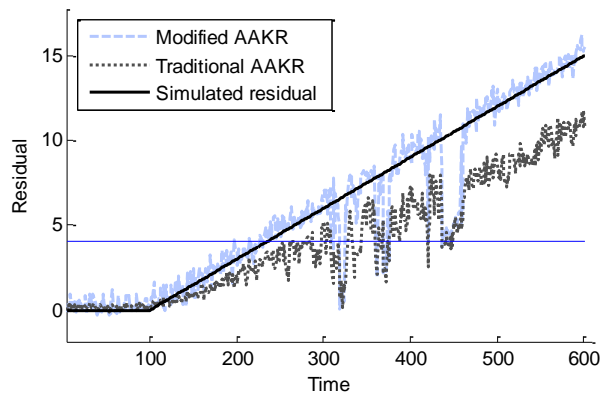
**Fig. 7** Residuals obtained for a signal not affected by any anomalies in the case in which another input signal of the reconstruction model is affected by the spike considered in Figure 6.



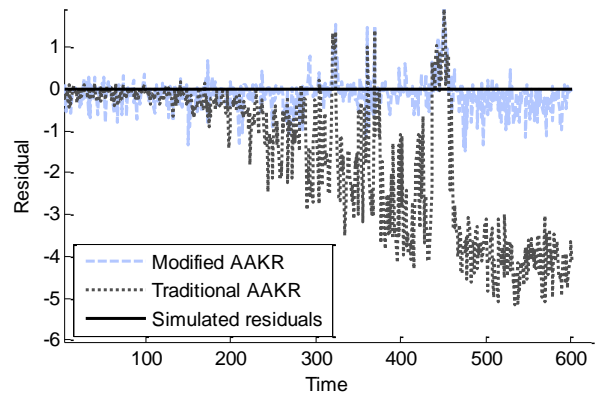
**Fig. 8.** Residuals obtained when the signal is affected by a random Gaussian noise characterized by a variance increasing with time. Black dotted line: residual provided by the traditional AAKR, dashed grey line by the modified AAKR. The continuous line represents the added noise.



**Fig. 9.** Residuals obtained for a signal not affected by any anomalies in the case in which another input signal of the reconstruction model is affected by the noise considered in Figure 8.



**Fig. 9.** Residuals obtained when the signal is affected by a linearly increasing drift. Black dotted line: residual provided by the traditional AAKR, dashed grey line by the modified AAKR. The continuous line

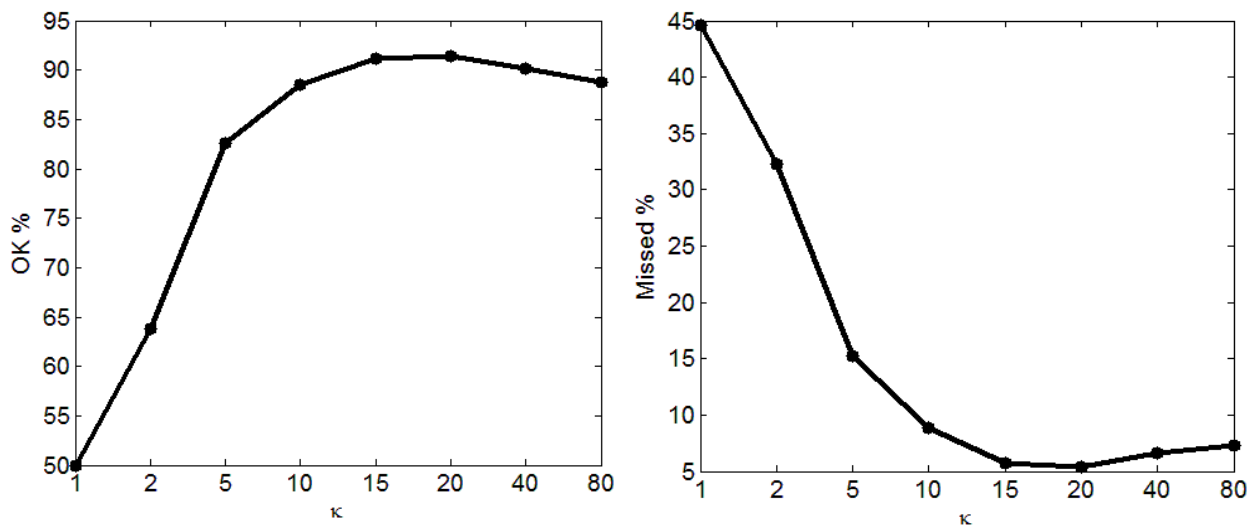


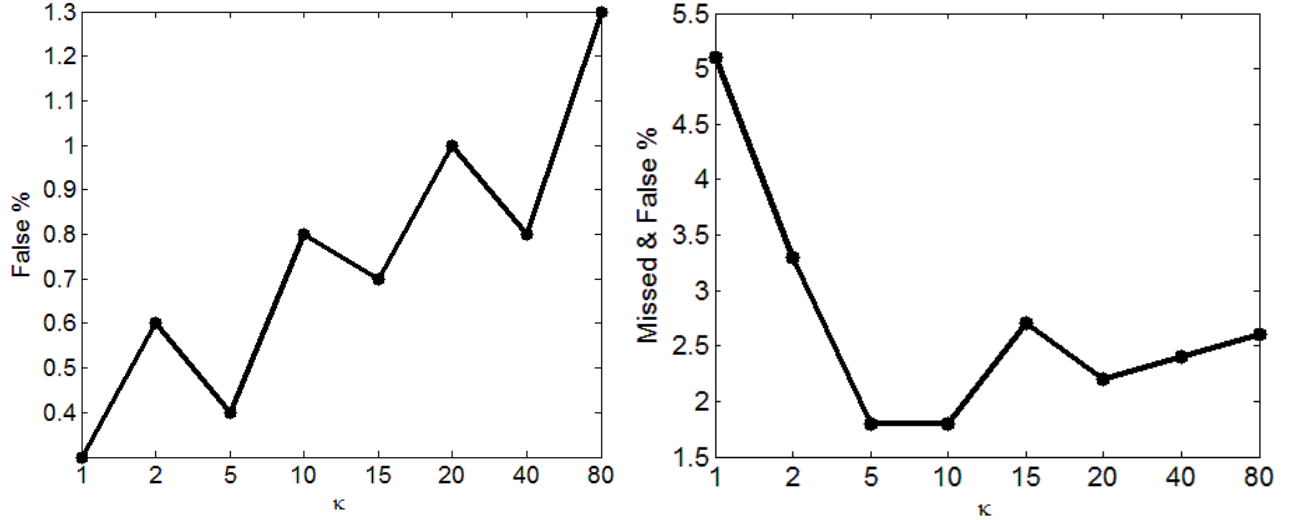
**Fig. 10** Residuals obtained for a signal not affected by any anomalies in the case in which another input signal of the reconstruction model is affected by the drift considered in Figure 10.

represents the added noise.

Further analyses have been performed in order to identify the sensitivity of the modified AAKR method to the setting of the exponential penalty vector parameter, to the number of simultaneous sensor failures and to the intensity of the failure.

With respect to the setting of the parameter  $\kappa$  of the exponential penalty vector, Fig. 11 reports the fault detection performance obtained by varying its value. Notice that for  $\kappa=1$  the modified AAKR method becomes the traditional AAKR method, and, thus, its performance, as reported in Table 2, is less satisfactory. Values of  $\kappa \geq 20$  are associated to an increasing percentage of false and missed alarms since the method tends to identify failures on healthy sensors. It is, however, interesting to note that the performance is not very sensitive to variation of  $\kappa$  in the range of  $\kappa = [10,40]$ , where the percentage of correct detection remains larger than 90% and the percentage of false and false and missed alarms are subject to small variations.





**Fig. 11 Sensitivity to the parameter  $\kappa$  defining the exponential penalty vector: top-left) correct detection; top-right) missed alarms; bottom-left) false alarms; bottom-right) missed and false alarms.**

In order to assess the performance of the method considering different numbers of simultaneous sensor failures, we have simulated 2 new test datasets of 1000 patterns each one obtained by adding deviations on 2 and 4 signals to the normal condition measurements, respectively. The obtained results are reported in Table 3. The modified AAKR method tends to perform remarkably better than the traditional AAKR method in the cases of two simultaneous sensor failures. However, the performance of the two methods tends to decrease if the number of simultaneous sensor failures increases. In particular, in the case of four simultaneous sensor failures, the traditional AAKR method performance is slightly more satisfactory than that of the modified method. This is due to the consequences of the modified AAKR method hypothesis that signal behaviors characterized by several signals affected by abnormal conditions are expected to be rare and so the number of missed alarms increases. This can be explained by considering the same numerical conceptual case study discussed in Section 4. Let us assume that we have to reconstruct the test pattern  $\vec{x}^{obs} = [1,1,0]$  which is characterized by a failure of sensors 1 and 2 (normal condition measurements would be  $[0,0,0]$ ). Considering a training set made of patterns ( $\vec{x}^{obs-nc}(k) = [k, k, k] \in \mathbb{R}^3, k \in (-2,2)$ ), the traditional AAKR reconstructs the test pattern in

the neighborhood of  $\left[\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right]$ , which is the closest training pattern using an Euclidean distance, whereas the modified AAKR with penalty vector  $\vec{p} = [1, 10, 100]$  reconstructs it in the neighborhood of  $\left[\frac{110}{111}, \frac{110}{111}, \frac{110}{111}\right]$ . Hence, the traditional AAKR method detects an anomaly on all the three sensors, providing a false alarm on sensor 3, whereas the modified AAKR detects a failure on sensor 1, providing a false alarm on sensor 1 and missed alarms on sensors 2 and 3. In general, it seems that the performance of the modified AAKR is satisfactory when the total number of sensor failures is lower than half of the number of sensors  $|S_{faults}| \leq \frac{|S_{tot}|}{2}$ . Nevertheless, it is worth noting that when  $|S_{faults}| > \frac{|S_{tot}|}{2}$ ; the performance of the traditional AAKR are themselves not satisfactory since less than 20% of the observations is correctly identified. Thus, if the working hypotheses at the basis of the method, i.e. faults impacting a small number of signals are more probable than those impacting a high number of signals, does not apply, then both the classical and the modified AAKR methods provide unsatisfactory performances.

<b>2 Simultaneous Errors</b>				
	<b>OK</b>	<b>MISSED</b>	<b>FALSE</b>	<b>MISSED &amp; FALSE</b>
<b>Traditional AAKR</b>	335	577	9	79
<b>Exponential <math>\kappa = 10</math></b>	837	158	2	3

Table 3 Quantitative results for 2 simultaneous error.

<b>4 Simultaneous Errors</b>				
	<b>OK</b>	<b>MISSED</b>	<b>FALSE</b>	<b>MISSED &amp; FALSE</b>
<b>Traditional AAKR</b>	170	637	16	177
<b>Exponential <math>\kappa = 10</math></b>	38	302	87	573

Table 4 Quantitative results for 4 simultaneous error.

The sensitivity of the method to the intensity of the deviations has been investigated by adding to the signals of the dataset  $\bar{X}_{err}$  deviations of intensity sampled uniformly in the interval [9,10] instead of [4,10]. The results reported in Table 5 show that:

- the number of correct detections performed by the traditional AAKR method increases due to the decrease of the number of missing alarms
- both the numbers of *false alarms* and *missed and false alarms* increase due to the spill-over effect on the remaining signals.

With respect to the modified AAKR method, in case of higher intensity of the deviations, the number of missed alarms decreases, whereas the increase of the number of false alarms is lower than that of the traditional AAKR. This confirms that the modified AAKR is more resistant to spill-over than the traditional AAKR.

		OK	MISSED	FALSE	MISSED & FALSE
$\kappa = 10$	$E_i \sim U([4, 10])$	887	85	10	18
	$E_i \sim U([9, 10])$	936	2	26	36
$\kappa = 1$	$E_i \sim U([4, 10])$	489	440	7	64
	$E_i \sim U([9, 10])$	757	64	70	109

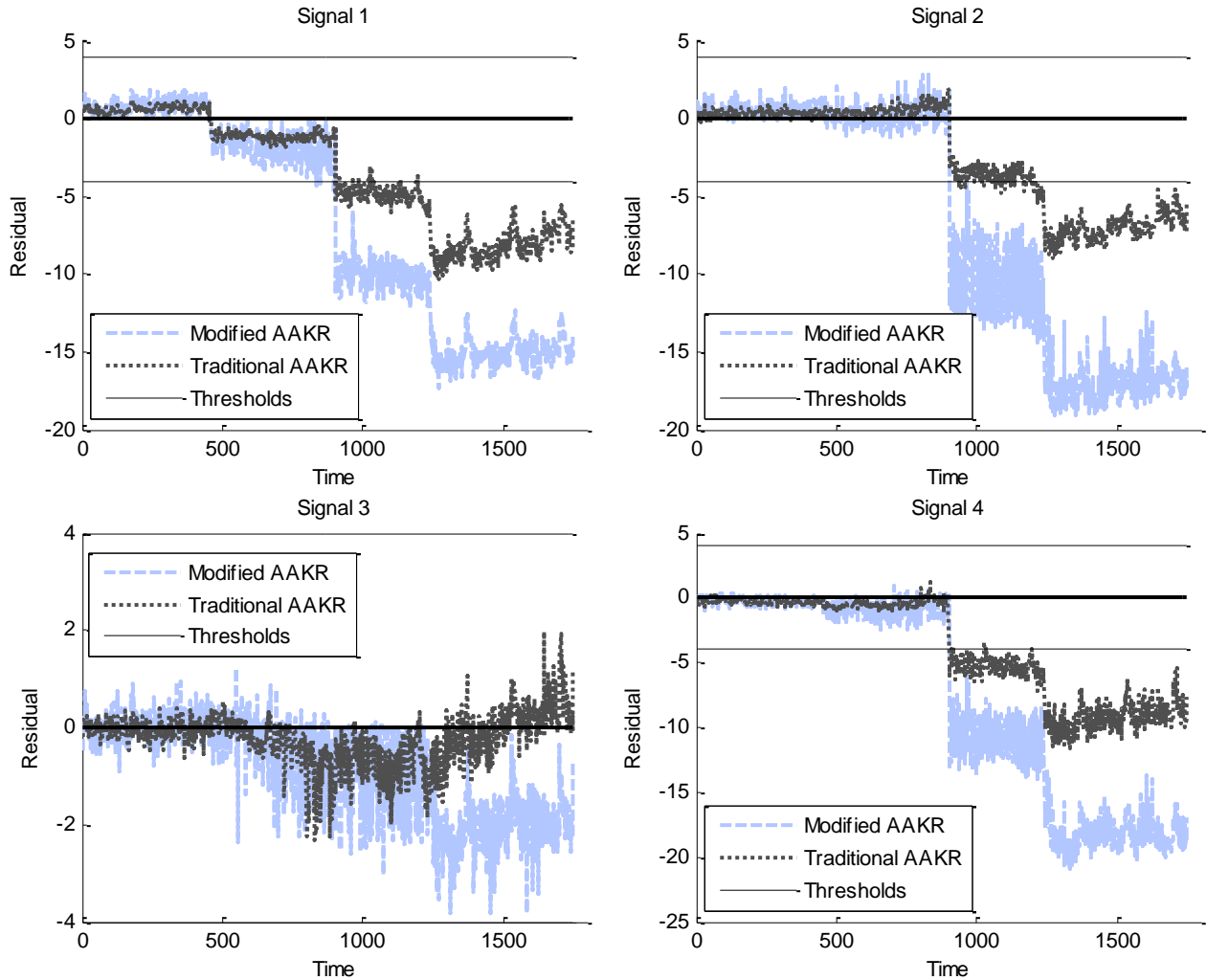
Table 5 Quantitative detection results with increasing deviation intensity for the Euclidean AAKR  $\kappa = 1$  and for the penalized AAKR  $\kappa = 10$ .

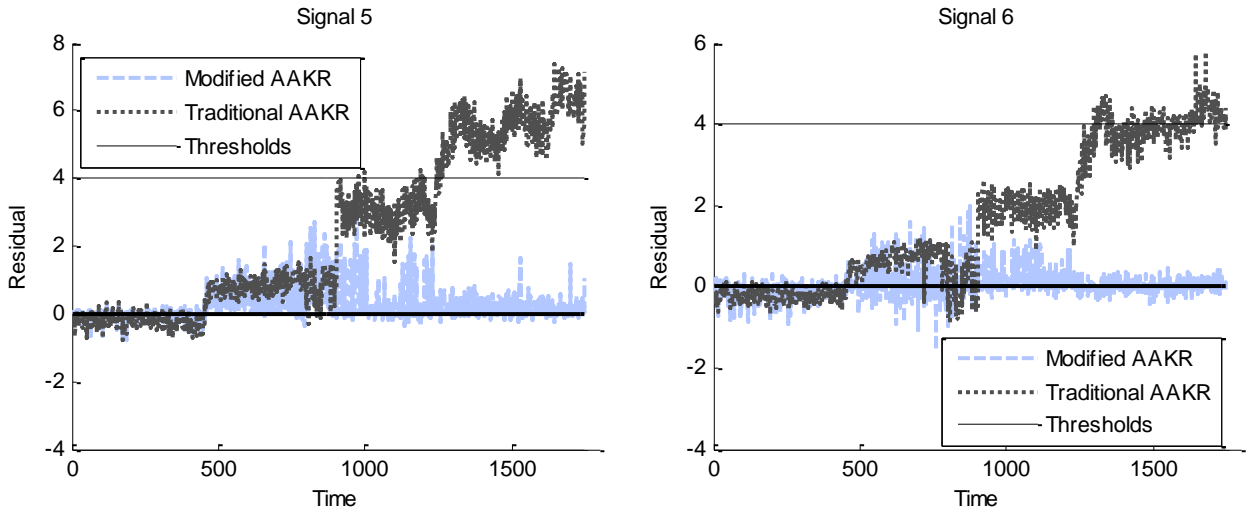
## 6.2 Application to real normal and abnormal conditions

The method is here applied to a real abnormal condition occurring in the same industrial component previously considered. The available 5200 observations collected in normal conditions have been used to train and validate the traditional and modified AAKR reconstruction models, whereas we have considered a test set made by 1752 observations which, according to plant experts, contains signal values collected during a component malfunctioning. Fig. 12 shows the reconstruction of the test set provided by the traditional and the modified AAKR methods. Notice that both methods identify around the time step 900 an abnormal condition with



consequences on signals 1 and 4 (the residuals exceed the threshold of 4). At the same time step, the modified AAKR finds that the abnormal conditions has consequences also on signal 2, whereas the traditional AAKR delays this detection until time step 1250, when it triggers the alarm also for signals 5 and 6.

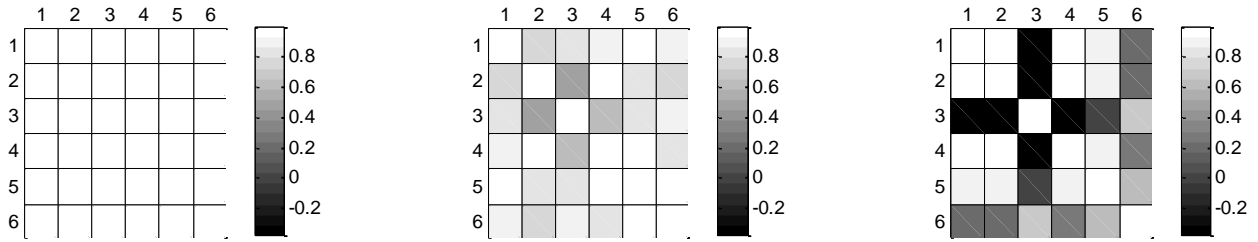




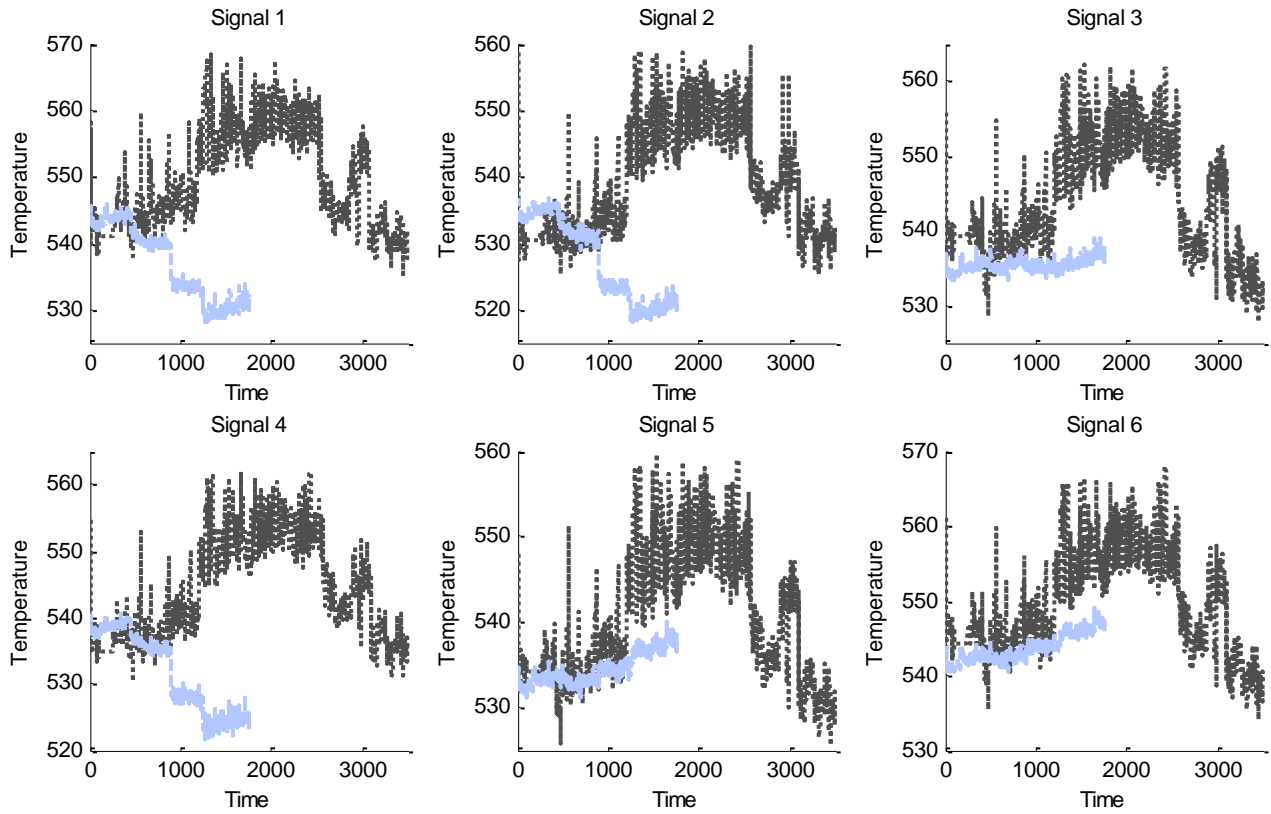
**Fig. 12 Residuals of the temperature signals obtained by the traditional AAKR (dark-dot line) and the Modified AAKR (light-dash line).**

Since information on the type of malfunctioning occurred in the plant are not available, the evaluation of the correctness of the obtained results is based on the following two observations:

- 1) Fig. 13 (left) shows the Pearson correlation coefficients between the signals collected in normal conditions, whereas Fig 14 (middle and right) shows the correlation between the same signals when we consider the reconstruction of the malfunctioning provided by the modified and the traditional AAKR methods, respectively. It can be observed that the reconstructions provided by the modified AAKR are characterized by correlation values similar to those in normal conditions, whereas the correlations between the signal reconstruction provided by the traditional AAKR are very different.
- 2) Fig. 14 shows that the values collected for signals 1, 2 and 4 during abnormal conditions (light dashed line) do not belong to the typical ranges of the respective signals in normal conditions (dark dotted line), whereas the values collected for signals 3, 5 and 6 belong to the normal condition ranges. This suggests that the modified AAKR correctly identified 1, 2 and 4 as signals impacted by a malfunctioning, whereas the reconstructions of signals 5 and 6 provided by the traditional AAKR are affected by the spill-over effects.



**Fig. 13** Visual representation of the Pearson correlation between the signals considering normal condition data (left), and the reconstruction of the abnormal conditions with the modified AAKR method (center) and with the traditional AAKR method (right).



**Fig. 14** Normal condition evolution of the signals in the training set (dark-dotted line) and corresponding evolution in the test set (light-dashed line).

## 7 CONCLUSIONS

In this work, we have considered condition monitoring for fault detection in industrial components. In order to obtain robust reconstructions of the values of the monitored signals expected in normal conditions, we have proposed to modify the traditional AAKR method. The modification is based on a different procedure for the computation of the similarity between the

present and the historical measurements. In particular, before the computation of the Kernel between the two vectors, which is performed as in the traditional AAKR method according to a Gaussian RBF function, the data are projected into a new signal space defined by using a penalty vector which reduces the contribution of signals affected by malfunctioning. The procedure is based on the hypothesis that the probability of occurrence of a fault causing variations on a large number of signals is lower than that of one causing variations on a small number of signals.

The modified AAKR method has been applied to a real case study concerning the monitoring of 6 highly correlated signals in an industrial plant for energy production. The possibility of detecting synthetic and real malfunctioning has been investigated. The obtained results have shown that the reconstructions provided by the modified AAKR are more robust than those obtained by using the traditional AAKR. This causes a reduction in the time necessary to detect abnormal conditions and in a more accurate identification of the signals actually affected by the abnormal conditions.

Future work will be devoted to the demonstration that the whole procedure proposed for the evaluation of the weights, i.e. the pre-processing step and the application of a Gaussian kernel, is itself a kernel function.

## **ACKNOWLEDGEMENTS**

The participation of Enrico Zio to this research is partially supported by the China NSFC under grant number 71231001. The participation of Piero Baraldi and Francesco Di Maio is supported by the European Union Project INNOVATION through Human Factors in risk analysis and management (INNHF, [www.innhf.eu](http://www.innhf.eu)) funded by the 7th framework program FP7-PEOPLE-2011-Initial Training Network: Marie-Curie Action.

## REFERENCES

- (Baraldi et al., 2011a) Baraldi P., Canesi R., Zio E., Seraoui R., Chevalier R., 2011, “*Genetic algorithm-based wrapper approach for grouping conditional monitoring signals of nuclear power plant system*”, *Integrated computer-Aided engineering*, vol. 18, pp. 221-234.
- (Baraldi et al., 2011b) Baraldi P., Gola G., Zio E., Roverso D., Hoffman M., 2011, “*A randomized model ensemble approach for reconstructing signals from faulty sensors*”, *Expert systems with application*, vol. 38, No. 8.
- (Baraldi et al., 2012) Baraldi P., Di Maio F., Pappaglione L., Zio E. and Seraoui R., 2012, “*Condition Monitoring of Power Plant Components During Operational Transients*”, *Proceedings of the Institution of Mechanical Engineers, Part O, Journal of Risk and Reliability*, 226(6) 568–583, 2012.
- (Baraldi et al., 2013a) Baraldi P., Di Maio F., Genini D., Zio E., “*Comparison of data-driven reconstruction methods for fault detection*”, under review, *IEEE Transactions on Reliability*.
- (Baraldi et al., 2013b) Baraldi P., Di Maio F., Genini D., Zio E., “*Missing data reconstruction by a data-driven fuzzy based similarity approach*”, under review, *Applied Soft Computing*.
- (Bishop, 1995) Bishop C., 1995, “*Neural Networks for Pattern Recognition*”, Oxford University Press, NY, USA.
- (Brock et al., 2008) Brock G.N., Shaffer J.R., Blakesley R.E., Lotz M.J., Tseng G.C., 2008, “*Which missing value imputation method to use in expression profiles: A comparative study and two selection schemes*”, *BMC Bioinformatics*, vol. 9.
- (Burges, 1998) Burges C. J.C., 1998, “*A Tutorial on Support Vector Machines for Pattern Recognition*”, *Data Mining and Knowledge Discovery*, vol. 2., pp. 121-167.
- (Candès et al., 2006a) Candès E.J., Romberg J., Tao T., 2006, “*Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*”, *IEEE Transactions on Information Theory*, vol. 52, No. 2, pp. 489-509.
- (Candès et al., 2006b) Candès E.J., Romberg J.K., Tao T., 2006, “*Stable signal recovery from incomplete and inaccurate measurements*”, *Communications on Pure and Applied Mathematics*, vol. 59, No. 8, pp. 1207-1223.
- (Chen et al., 2013) Chen X., Xu G., 2013, “*A self-adaptive alarm method for tool condition monitoring based on Parzen window estimation*”, *Journal of Vibroengineering*, vol. 15, No. 3, pp. 1537-1545.
- (Chevalier et al., 2009) Chevalier R., Provost D., and Seraoui R., 2009, “*Assessment of Statistical and Classification Models For Monitoring EDF’s Assets*”, *Sixth American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation*.
- (Di Maio et al., 2013) Di Maio, F.; Baraldi, P.; Zio, E.; Seraoui, R., 2013, “*Fault Detection in Nuclear Power Plants Components by a Combination of Statistical Methods*”, *IEEE Transaction on Reliability*, 62 (4) , pp. 833-845, 2013.
- (Duijndam et al., 1999) Duijndam A.J.W., Schonewille M.A., Hindriks, C.O.H., 1999, “*Reconstruction of band-limited signals, irregularly sampled along one spatial direction*”, *Geophysics*, vol. 64, No. 2, pp. 524-538.
- (Dunia et al., 1996) Dunia R., Qin S.J., Edgar T.F., McAvoy T.J., 1996, “*Identification of Faulty Sensors Using Principal Component Analysis*”, *AIChE Journal*, vol.42, No.10, pp. 2797-2811.
- (Feller et al., 2013) Feller S., Todorov Y., Jaroszewski D., Chevalier R., 2013, “*Taming Anomaly Detection for Industrial Applications*”, *AIP Conference Proceedings*, vol. 1558, pp. 2257-2260.

- (Guglielmi et al., 1995) Guglielmi G., Parisini T., Rossi G., 1995, "*Fault Diagnosis And Neural Networks: A Power Plant Application*", Control Engineering Practice, vol. 3, No. 5, pp. 601-620.
- (Guiot et al., 2005) Guiot J., Nicault A., Rathgeber C., Edouard J.L., Guibal F., Pichard G., Till C., 2005, "*Last-millennium summer-temperature variations in western Europe based on proxy data*", Holocene, vol. 15, No. 4, pp. 489-500.
- (Hameed et al., 2009) Hameed, Z., Hong, Y. S., Cho, Y. M., Ahn, S. H., & Song, C. K., 2009. "*Condition monitoring and fault detection of wind turbines and related algorithms: A review*", Renewable and Sustainable Energy Reviews, vol. 13, No. 1, pp. 1-39.
- (Harkat et al., 2007) Harkat M.F., Djelal S., Doghmane N., Benouaret M., 2007, "*Sensor Fault Detection, Isolation and Reconstruction Using Nonlinear Principal Component Analysis*", International Journal of Automation and Computing, vol. 4, No. 2, pp. 149-155.
- (Hines et al., 2006) Hines J. W., Garvey D., 2006, "*Development and Application of Fault Detectability Performance Metrics for Instrument Calibration Verification and Anomaly Detection*", Journal of Pattern Recognition Research, vol. 1, No. 1, pp. 2-15.
- (Hines et al., 2007) Hines J. W., Garvey D., 2007, "*Process and Equipment Monitoring Methodologies Applied to Sensor Calibration Monitoring*", Quality and Reliability International, vol. 23, No. 1, pp. 123-135.
- (Hubert et al., 2005) Hubert M., Rousseeuw P.J., Vanden Branden K., 2005, "*ROBPCA: A new approach to robust principal component analysis*", Technometrics, vol. 47, No. 1, pp. 64-79.
- (Hunsop, 2011) Hunsop H., 2011, "*Kernel Based Density Estimation*" in Constrained and hierarchical density estimation for image reconstruction and sensor networks, ProQuest UMI Dissertation Publishing, ISBN 9781243584212.
- (Jack et al., 2002) Jack L.B., Nandi A.K., 2002, "*Fault detection using support vector machines and artificial neural networks, augmented by genetic algorithms*", Mechanical Systems and Signal Processing, vol. 16, No. 2-3, pp. 373-390.
- (Jardine et al., 2006) Jardine A., Lin D., Banjevic D., 2006, "*A review on machinery diagnostics and prognostics implementing condition based maintenance*", Mechanical Systems and Signal Processing, vol. 20, No. 7, pp. 1483–1510.
- (Kim et al., 2005) Kim H., Golub G.H., Park H., 2005, "*Missing value estimation for DNA microarray gene expression data: Local least squares imputation*", Bioinformatics, vol. 21, No. 2, pp. 187-198.
- (King et al., 2009) King S., Bannister P.R., Clifton D.A., Tarassenko, L., 2009, "*Probabilistic approach to the condition monitoring of aerospace engines*", Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, Vol. 223, No. 5, pp. 533-541.
- (Lee, 1997) Lee J., 1997, "*Strategy and challenges on remote diagnostics and maintenance for manufacturing equipment*", Proceedings of the Annual Reliability and Maintainability Symposium, pp. 368-370.
- (Liu et al., 2004) Liu B., Sacchi M.D., 2004, "*Minimum weighted norm interpolation of seismic records*", Geophysics, vol. 69, No. 6, pp. 1560-1568.
- (Marseguerra et al., 2003) Marseguerra M., Zio E., Baraldi P., Oldrini A. 2003, "*Fuzzy logic for signal prediction in nuclear systems*", Progress in Nuclear Energy, vol. 43, No. 1-4 SPEC, pp. 373-380.
- (Muradore et al., 2012) Muradore R., Fiorini P., 2012, "*A PLS-based statistical approach for fault detection and isolation of robotic manipulators*", IEEE Transactions on Industrial Electronics, vol. 59, No. 8, pp. 3167-3175.

- (Müller et al., 2001) Müller, K.-S, Mika, S., Rätsch, G., Tsuda, K., Schölkopf, B., 2001, “*An introduction to kernel-based learning algorithms*”, IEEE Transactions on Neural Networks, vol. 12, No. 2, pp. 181-201.
- (Nabeshima et al., 1998) Nabeshima K., Suzudo T., Suzuki K., Turcan E., 1998, “*Real-time Nuclear Power Plant Monitoring with Neural Network*”, Journal of Nuclear Science and Technology, vol. 35, No. 2, pp. 93-100.
- (Nandi et al., 2005) Nandi S., Toliyat H.A., Li X., 2005, “*Condition monitoring and fault diagnosis of electrical motors - A review*”, IEEE Transactions on Energy Conversion, vol. 20, No. 4, pp. 719-729.
- (Roverso et al., 2007) Roverso D., Hoffmann M., Zio E., Baraldi P., Gola G., 2007, “*Solutions for plant-wide on-line calibration monitoring*”, Proceedings of the European Safety and Reliability Conference ESREL 2007, Risk, Reliability and Societal Safety, vol. 1, pp. 827-832.
- (Samanta, 2004) Samanta, B. 2004, “*Gear fault detection using artificial neural networks and support vector machines with genetic algorithms*”, Mechanical Systems and Signal Processing, vol. 18, No. 3, pp. 625-644.
- (Venkatasubramanian et al., 2003a) Venkatasubramanian V., R. Rengaswamy, K. Yin, S. N. Kavuri, 2003, “*A review of process fault detection and diagnosis Part I: Quantitative model-based methods*”, Computers and Chemical Engineering, vol. 27, pp. 293-311.
- (Venkatasubramanian et al., 2003b) Venkatasubramanian V., R. Rengaswamy, S. N. Kavuri, 2003, “*A review of process fault detection and diagnosis Part II: Qualitative models and search strategies*”, Computers and Chemical Engineering, vol. 27, pp. 313-326.
- (Venkatasubramanian et al., 2003c) Venkatasubramanian V., R. Rengaswamy, S. N. Kavuri, K. Yin, 2003, “*A review of process fault detection and diagnosis Part III: Process history based methods*”, Computers and Chemical Engineering, vol. 27, pp. 327-346.
- (Wei et al., 2013) Wei Z., Tao T., Ding Z. S., Zio E., 2013, “*A dynamic particle filter-support vector regression method for reliability prediction*”, Reliability Engineering and System Safety, vol. 119, pp. 109-113.
- (Widodo et al., 2007) Widodo A., Yang B., 2007, “*Support vector machine in machine condition monitoring and fault diagnosis*”, Mechanical Systems and Signal Processing, vol. 21, No. 6, pp. 2560-2574.
- (Worden et al., 2011) Worden K., Staszewski W., Hensman J., 2011, “*Natural computing for mechanical systems research: a tutorial overview*”, Mechanical Systems and Signal Processing, vol. 25, No. 1, pp. 4-111.
- (Zhao et al., 2014) Zhao C., Gao F., 2014, “*Fault-relevant Principal Component Analysis (FPCA) method for multivariate statistical modeling and process monitoring*”, Chemometrics and Intelligent Laboratory Systems, vol. 133, pp. 1-16.
- (Zio et al., 2011) Zio E., Baraldi P., Zhao W., 2011, “*Confidence in signal reconstruction by the Evolving Clustering Method*”, 2011 Prognostics and System Health Management Conference, PHM-Shenzhen 2011.