A novel fault detection system taking into account uncertainties in the reconstructed signals

Sameer Al-Dahidi [a], Piero Baraldi [a], Francesco Di Maio [a,*], Enrico Zio [a,b]

[a] Energy Department, Politecnico di Milano, Milan, Italy
[b] Chair on Systems Science and the Energetic Challenge, European Foundation for New Energy-Électricité de France, École Centrale Paris and Supélec, Paris, France

Received 14 April 2014
Received in revised form 17 June 2014
Accepted 18 June 2014
Available online 10 July 2014

1. Introduction

Over the last few decades, Condition Monitoring (CM) techniques have been strongly developed in terms of measurement devices, and data processing and management capabilities. These developments have encouraged industries like nuclear, oil and gas, automotive and chemical to apply Condition-Based Maintenance (CBM) (Jardine et al., 2006; Campos, 2009) for increasing system availability, reducing maintenance costs, minimizing unscheduled shutdowns and increasing safety (Thurston and Lebold, 2001; Yam et al., 2001; Miao et al., 2010).

A typical CBM scheme is shown in Fig. 1: a fault detection (FD) system continuously collects information from sensors mounted on the component of interest (Jardine et al., 2006; Ahmad and Kamaruddin, 2012) and delivers information on the health state (either normal or abnormal conditions) of the monitored component through an alarm system interface. On the basis of the received information, the decision maker decides whether it is necessary to perform a maintenance action or if it is possible to postpone it (Jardine et al., 2006).

In this work, we only focus on the FD system. This is typically made by an empirical reconstruction model and a decision tool that supports the decision maker. Several methods have been used with success to reconstruct values of the signals expected in normal conditions, for example Artificial Neural Networks (ANNs) (Ebron et al., 1990; Dong and McAvoy, 1994; Fantoni and Mazzola, 1996; Hines et al., 1997; Maki and Loparo, 1997; Xu et al., 1999; Hines and Davis, 2005) and Auto-Associative Kernel Regression (AAKR) (Hines and Garvey, 2006; Yang et al., 2006; Heo, 2008; Baraldi et al., 2012).

The decision tool typically analyzes the differences (residuals) between the measured and reconstructed values of the n measured signals in order to advice on the component health state (normal or abnormal conditions) (Fig. 2). If reconstructions are similar to measurements, then the component is recognized to be in normal conditions (nc) and no alarm is triggered, whereas if reconstructions are different from measurements, then abnormal conditions (ac) are detected and an alarm is triggered (Zhao et al., 2011; Di Maio et al., 2013).

* Corresponding author. Tel.: +39 0223996372.
E-mail address: francesco.dimaio@polimi.it (F. Di Maio).

Notation and list of acronyms

CM	Condition Monitoring
CBM	Condition-Based Maintenance
FD	Fault Detection
ANN	Artificial Neural Networks
AAKR	Auto-Associative Kernel Regression
n	Number of measured signals
j	Index of the generic signal, $j = 1,\ldots,n$
nc	Normal Conditions
ac	Abnormal Conditions
γ	False alarm rate
β	Missing alarm rate
SPRT	Sequential Probability Ratio Test
PI	Prediction Interval
OS	Order Statistics
M	Length of the detection window
NPP	Nuclear Power Plant
N_p	Number of measurement times and/or representative turbine shaft speeds of each signal j, $j = 1,\ldots,n$
t_k	k-th time instant, $k = 1,\ldots,N_p$
$\varepsilon(t_k)$	Prediction error at the k-th time instant
$\vec{x}^{test}(t_k)$	Vector containing the test measurements of n signals at time t_k, $k = 1,\ldots,N_p$
$x^{test}(t_k, j)$	Measured value of signal j, $j = 1,\ldots,n$ at time t_k, $k = 1,\ldots,N_p$
$\vec{\widehat{x}}^{test}(t_k)$	Vector containing the reconstructed values of the test measurements of n signals at time t_k, $k = 1,\ldots,N_p$
$\widehat{x}^{test}(t_k, j)$	Reconstructed value of signal j, $j = 1,\ldots,n$ at time t_k, $k = 1,\ldots,N_p$
$\hat{x}^{lower}(t_k), \hat{x}^{upper}(t_k)$	Lower and upper bounds of PI at time t_k
$1-\sigma$	Confidence level
$\alpha^{95percentile}(t_k)$	Sorted scale factor value at the k-th time instant for 95% confidence level, $k = 1,\ldots,N_p$
NV	Number of measurements/reconstructions in the validation set performed at time t_k after the beginning of the transient used to estimate the PIs
$var_k^{res}(\widehat{x}^{val}(t_k))$	Bias between *NV* measurements and their reconstructions at time t_k of signal j of the validation set
$var_k^{res}(\widehat{x}^{val}(t_k))$	Variance of *NV* reconstructions at time t_k of signal j of the validation set
t_1	1st time instant
N_{train}	Number of time series measurements in normal conditions of a training set
N	Component transients of signals measurements
i	i-th component transient, $i = 1,\ldots,N$
P_1	Probability that 1 signal j, $j = 1,\ldots,n$ is failed and detected for a given detection window of length M
P_n	Probability that at least one out of n signals are failed and detected for a given detection window of length M
P	Probability that a transient is failed and abnormal conditions are detected using $N_p - M + 1$ detection windows
$\gamma_{\max}$	Prefixed maximum limit of false alarm rate
$f_i(x(t,1),\ldots,x(t,4))$	i-th transient of four-dimensional ($n = 4$) signals with $N_p = 101$ time steps
$\vec{x}^{nc}_{i=1:4000}(t_k)$	Time evolutions in nc of the 4 signals in the 4000 transients at t_k, $k = 1,\ldots,N_p$ of the j-th signal (sigmoid behavior)
a, ζ and μ	Random parameters in arbitrary units used to construct transients in nc
$x_i^{nc}(t_k)$	Time evolutions in nc of the i-th transient at t_k, $k = 1,\ldots,N_p$ of the j-th signal (sigmoid behavior)
R	Number of operational zones of a component in normal conditions
h	Gaussian kernel bandwidth
t_f	Random failure time of a signal j, $j = 1,\ldots,n$ in a transient i, $i = 1,\ldots,N$
$x_i^{ac}(t_k)$	Time evolutions in ac of the i-th transient at t_k, $k = 1,\ldots,N_p$ of the j-th signal (different functional behavior)
a^* and μ^*	Random parameters in arbitrary unit used to construct transients in ac based on the signal value
$\overline{\overline{X}}$	Historical measurements performed at past time t_k, $k = 1,\ldots,N_{train}$
$x^{test}(t_k, 1)$	Test value of signal 1 measured at time t_k $k = 1,\ldots,N_p$
$x(t_k, j)$	Historical value of signal j measured at past time t_k, $k = 1,\ldots,N$ of $\overline{\overline{X}}$
$d^2(t_k)$	Euclidean distance between the current test measurements $\vec{x}^{test}(t,j)$ and the k-th observation of $\overline{\overline{X}}$, $x(t_k, j)$
$\mu(j)$ and $\sigma(j)$	Mean and the standard deviation of the j-th signal in $\overline{\overline{X}}$
$w(t_k)$	Similarity measures obtained by computing $d^2(t_k)$, $k = 1,\ldots,N$
$\alpha_i(t_k)$	Scale factor at the k-th time instant of the i-th validation transient, $i = 1,\ldots,NV$ of signal j
$e_i(t_k)$	Residual between the measured value of signal j and its reconstruction in the i-th validation transient, $i = 1,\ldots..,NV$ at time t_k
$\widehat{x}_i^{val}(t_k)$	Reconstructed value of signal j at time t_k, $k = 1,\ldots,N_p$ in the i-th validation transient $i = 1,\ldots,NV$
$x^{test\text{-}normalized}(t,j)$	Normalized values of $x(t,j)$ at time t using $\mu(j)$ and $\sigma(j)$
$x_i^{val}(t_k)$	Value of signal j measured at time t_k, $k = 1,\ldots,N_p$ in the i-th validation transient $i = 1,\ldots,NV$
d	Outcome of the Durbin–Watson test

In threshold-based methods (Puig et al., 2008; Montes de Oca et al., 2012), the presence of abnormal conditions is detected when the residual values exceed a prefixed threshold. A practical difficulty is the setting of the threshold value itself: too high threshold values lead to high missing alarm rates (β), whereas too low values lead to high false alarm rates (γ) (Di Maio et al., 2013). Furthermore, threshold-based methods do not provide any information on the confidence that we should have in the FD system outcomes, such as expected missing and false alarm rates (Zhao et al., 2011). Contrarily, statistical methods which consider the residual as a random variable and analyze its statistical distribution, such as the Sequential Probability Ratio Test (SPRT) (Wald, 1947; Gross and Humenik, 1991; Schoonewelle et al., 1995; Hines and Garvey, 2006; Baraldi et al., 2010; Di Maio et al., 2013), typically allow obtaining the desired level of missing and false alarm rates. However, also they require the setting of some parameters such as those defining the expected statistical distributions of the residuals in abnormal conditions, which can be difficult in practical industrial applications (Emami-Naeini et al., 1988; Di Maio et al., 2013).

Independently from the choice of the reconstruction model and of the method adopted to analyze the residuals, the performance of the overall FD system is influenced by uncertainties which can cause false and missing alarms, and affect the decision on the necessity of performing a maintenance intervention (Helton, 1994; Zheng and Frey, 2005; Weber et al., 2007; Aven and Zio, 2012).

In this context, the objective of the present work is to develop a novel, non-parametric, sequential decision tool that takes into

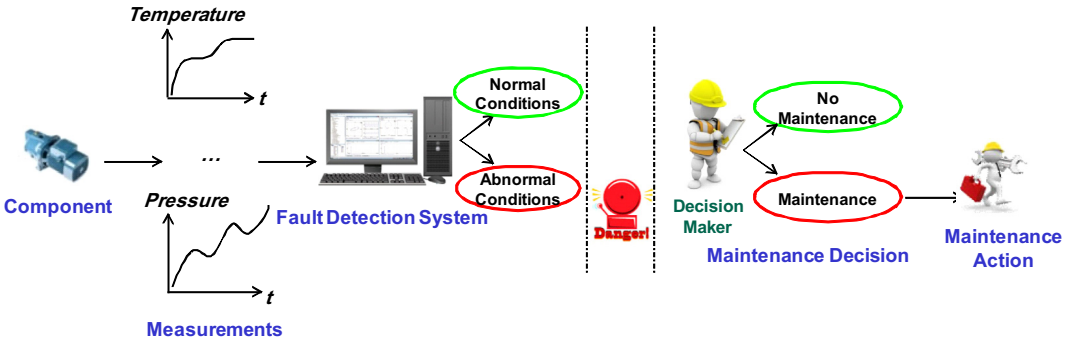

Fig. 1. Scheme of Condition-Based Maintenance (CBM) principles.

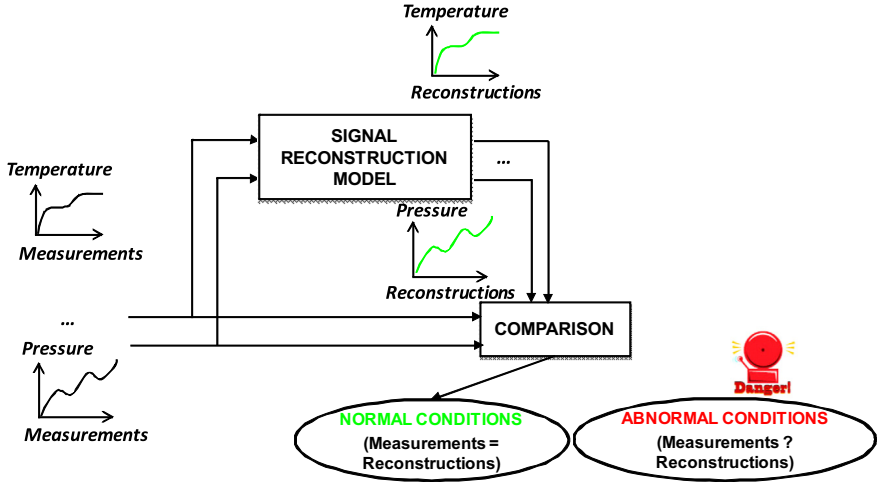

Fig. 2. Traditional FD system.

account the quantified uncertainty on the reconstructions. Methods like Trimmed Sequential Probability Ratio Test (TSPRT) and non-parametric sequential rank-sum probability ratio test (SRPRT) are some examples of similar sequential schemes, in which the residuals are tested sequentially (Yu et al., 2004), but are hardly applicable within a FD system dealing with operational transients.

Contrarily, in this work, the quantification of the uncertainty of the reconstructions is based on the estimation of Prediction Intervals (PIs) by using Order Statistics (OS) theory at a pre-defined confidence level, e.g., 95% (Al-Dahidi et al., 2014) and the novel decision rule consists in properly setting the number M of consecutive measurements (hereafter also called detection window) that have to be covered by the estimated PIs for proper FD during operational transients. In practice: (a) if at least one out of M consecutive measurements falls within the corresponding PIs, the component is assumed to work in normal conditions, (b) if the M consecutive measurements do not fall within the corresponding PIs, abnormal conditions are detected. The proposed decision tool is tested with respect to an artificial case study representing the behavior of a component during operational transients and it is validated using a real industrial case concerning 27 shut-down transients of a nuclear power plant (NPP) turbine. The obtained results show that the approach is able to detect and provide the decision maker with the required information for establishing whether a maintenance intervention is necessary with controlled false and missing alarm rates.

The remaining of this paper is organized as follows. In Section 2, a reconstruction model for signal reconstruction during operational transients is developed, and the method adopted for estimating the PIs on signal reconstructions is presented. In Section 3, the novel non-parametric, sequential decision tool to decide whether the component is in normal or abnormal conditions is proposed. An artificial case study representing the component behavior during typical start-up transients and a real industrial case concerning 27 shut-down transients of a nuclear power plant (NPP) turbine are introduced in Sections 4 and 5, respectively. Furthermore, the results obtained in the application of the novel FD system to the case studies are discussed. Finally, some conclusions are proposed in Section 6.

2. Reconstruction model and uncertainty quantification

A typical reconstruction model receives in input a vector $\vec{x}^{test}(t_k) = [x^{test}(t_k,1),\ x^{test}(t_k,j),\ \ldots,\ x^{test}(t_k,n)]$ containing the test measurements of n signals, $j = 1,\ldots,n$ at time t_k, $k = 1,\ldots,N_p$ where N_p is the number of measurement times of each signal j. On the basis of historical measurements performed in normal conditions, the reconstruction model produces in output a vector $\vec{\hat{x}}^{test}(t_k) = [\hat{x}^{test}(t_k,1), \hat{x}^{test}(t_k,j), \ldots, \hat{x}^{test}(t_k,n)]$ containing the values of the input signals expected in case of normal conditions at the time t_k. For the sake of simplicity, the signal index j will be omitted from the notations $x^{test}(t_k, j)$ and $(\hat{x}^{test}(t_{kj})$, and will be used only when strictly required for understanding.

2.1. Reconstruction of operational transients

Numerous empirical reconstruction models are available in literature, e.g., ANNs and AAKR. Despite that, in (Baraldi et al., 2012),

different approaches to the problem of signal reconstruction during operational transients have been compared using AAKR (see Appendix A) as base model for the reconstruction. The obtained results have shown that, in order to reduce the computational efforts and to increase model reconstruction accuracy, it is sometimes useful to develop an ensemble of reconstruction models, each one dedicated to a different operational zone of the component. Thus, in what follows, we resort to the use of AAKR models tailored on different operational zones for an online signal reconstruction task, as we shall see in Section 4.

2.2. Uncertainty quantification using PIs

The quantification of the uncertainty on the signal reconstructions by AAKR is based on the estimation of Prediction Intervals (PIs) (Bendat and Piersol., 2010) by using Order Statistics (OS) theory (Secchi et al., 2008) and a scale factor (Bouckaert et al., 2011): with respect to a component in normal conditions, the PI with confidence level $1-\sigma$ (e.g., 95%) is defined as the interval $[\widehat{x}^{lower}(t_k), \widehat{x}^{upper}(t_k)]$, such that the probability that the measurement of signal j at time t_k, $x^{test}(t_k)$, falls within the interval is equal to $1-\sigma$ (Al-Dahidi et al., 2014).

The method goes along the following steps. It entails an offline procedure for estimating the PIs and an online procedure for providing the FD system with the proper PIs.

The offline procedure consists in:

- *Step 1: Signal reconstruction*. Using N_{train} training data, the AAKR-built model provides the reconstruction $\hat{x}_i^{val}(t_k)$ of signal j at time t_k, $k = 1,\ldots,N_p$ in the i-th validation transient $i = 1,\ldots,NV$, of length N_p, (*i.e.*, $N_{train} = N_p \times NT$, where NT is the number of training transients each of length N_p). The number NV of validation transients is properly defined by OS theory (Secchi et al., 2008) for guaranteeing a PI with confidence level $1-\sigma$.
- *Step 2: Residual calculations*. At each k-th time, the absolute difference between the measured value and its reconstruction of signal j is calculated as $e_i(t_k) = \left|\widehat{x}_i^{val}(t_k) - x_i^{val}(t_k)\right|$ of the i-th validation transient, $i = 1,\ldots,NV$.
- *Step 3: Prediction error calculations*. At each time k, the prediction error of signal j is calculated as $\varepsilon(t_k) = \sqrt{var_k^{rec}(\widehat{x}^{val}(t_k)) + var_k^{res}(\widehat{x}^{val}(t_k))}$ by calculating the variance of the NV reconstructions $var_k^{rec}(\widehat{x}^{val}(t_k))$ (Eq. (1)) and the variance of the NV residuals (squared bias) $var_k^{res}(\widehat{x}^{val}(t_k))$ (Eq. (2)) of signal j:

$$Var_k^{rec}\left(\widehat{x}^{val}(t_k)\right) = \frac{\sum_{i=1}^{NV}\left(\widehat{x}^{val}(t_k) - \sum_{i=1}^{NV}\left(\widehat{x}_i^{val}(t_k)\right)/NV\right)^2}{NV} \quad (1)$$

$$var_k^{res}\left(\widehat{x}^{val}(t_k)\right) = \frac{\sum_{i=1}^{NV}\left(\widehat{x}_i^{val}(t_k) - x_i^{val}(t_k)\right)^2}{NV} \quad (2)$$

- *Step 4: Scale factor calculations*. At each time k, the scale factor, $\alpha^{95\ \text{percentile}}(t_k)$, is calculated as the 95th percentile of the NV $\alpha_i(t_k)$, $i = 1,\ldots,NV$, where $\alpha_i(t_k) = e_i(t_k)/\sqrt{var_k^{rec}(\widehat{x}^{val}(t_k))}$, to estimate a PI with a given confidence $1-\sigma$. The advantages of using the scale factor are: (1) the trade-off between the coverage and the width is satisfied; (2) the technique is independent from the reconstruction method applied (Bouckaert et al., 2011); and (3) the scale factor deals with the uncertainty caused by the AAKR-built model. In practice, at each time t_k, if the AAKR reconstructions are inaccurate, then, the scale factor values are large (i.e., $e_i(t_k) = \left|\widehat{x}_i^{val}(t_k) - x_i^{val}(t_k)\right|, i = 1,\ldots,NV$ is large) in order to achieve the desired coverage level $(1-\sigma)$, and vice versa.

Then, within the online FD, for any test measurement $x^{test}(t_k)$ of a given signal j at each time t_k, the uncertainty on the reconstruction, $\widehat{x}^{test}(t_k)$, is quantified in the form of PI, as in Eq. (3):

$$\left[\widehat{x}^{test}(t_k) - \alpha^{95percentile}(t_k)\varepsilon(t_k), \widehat{x}^{test}(t_k) + \alpha^{95percentile}(t_k)\varepsilon(t_k)\right] \quad (3)$$

In this way, at each time t_k a PI is yielded with a specified coverage and with an acceptable width (Al-Dahidi et al., 2014).

3. The Novel Decision Tool

In this section, a novel non-parametric, sequential decision tool is proposed to decide whether the component is in normal or abnormal conditions, taking into account the quantified uncertainties.

Once the PI with a given confidence level has been estimated at each time t_k as in Eq. (3), the final decision on the necessity of performing a maintenance intervention can be taken by directly comparing M successive signal measurements with the estimated PIs (Fig. 3). In practice, if the component suffers from some abnormality at time t_k, the measurement at time $t_k + 1$ should start deviating from the normal conditions. If we consider M consecutive measurements, two situations may occur:

- At least one out of M measurements falls within the corresponding PIs, with associated confidence level $1-\sigma$; thus, the component is assumed to work in normal conditions and the alarm is not triggered.
- The M consecutive measurements do not fall within the PIs; thus, abnormal conditions are detected and the alarm is triggered.

The length M of the detection window is the number of measurements that the decision maker should wait before taking a decision regarding the health state of the component (Dixon et al., 2000; Guo and Bai, 2011): this, in principle, should be as low as possible and, at the same time allow for large fault detectability (i.e., small false alarm rates (γ)): that is, the detection window length and the detectability have to be simultaneously optimized.

Assuming that N transients have been measured at different times, each transient being a n-dimensional time series of N_p measurements (yielding to uncorrelated residuals at the Step 2 of the offline PI quantification procedure of Section 2.2), the reconstructions uncertainties can be quantified in the form of PIs with a confidence level $1-\sigma$ such that the probability that the measurement in normal conditions $x^{test}(t_k)$ of signal j, $j = 1,\ldots,n$ at time t_k, $k = 1,\ldots,N_p$, falls within the interval $[\widehat{x}^{lower}(t_k), \widehat{x}^{upper}(t_k)]$ is equal to $1-\sigma$ (Rasmussen et al., 2003; Office of Nuclear Regulatory Research, 2007): then, the optimum length M of the detection window can be calculated using Eq. (4), where γ_{max} is the maximum acceptable false alarm rate (see Appendix B for demonstration):

$$P = 1 - \left[1 - \sigma^M\right]^{(Np-M+1)\times n} \leq \gamma_{max} \quad (4)$$

It is worth mentioning that for more complex cases (e.g., N_p is not fixed), Eq. (4) can be seen as a multi-objective optimization problem whose optimal solution is a compromise between the measurement sampling frequency (that affects N_p) and M that can be found using optimization methods such as evolutionary algorithms (e.g., Genetic Algorithms, Differential Evolution, etc.).

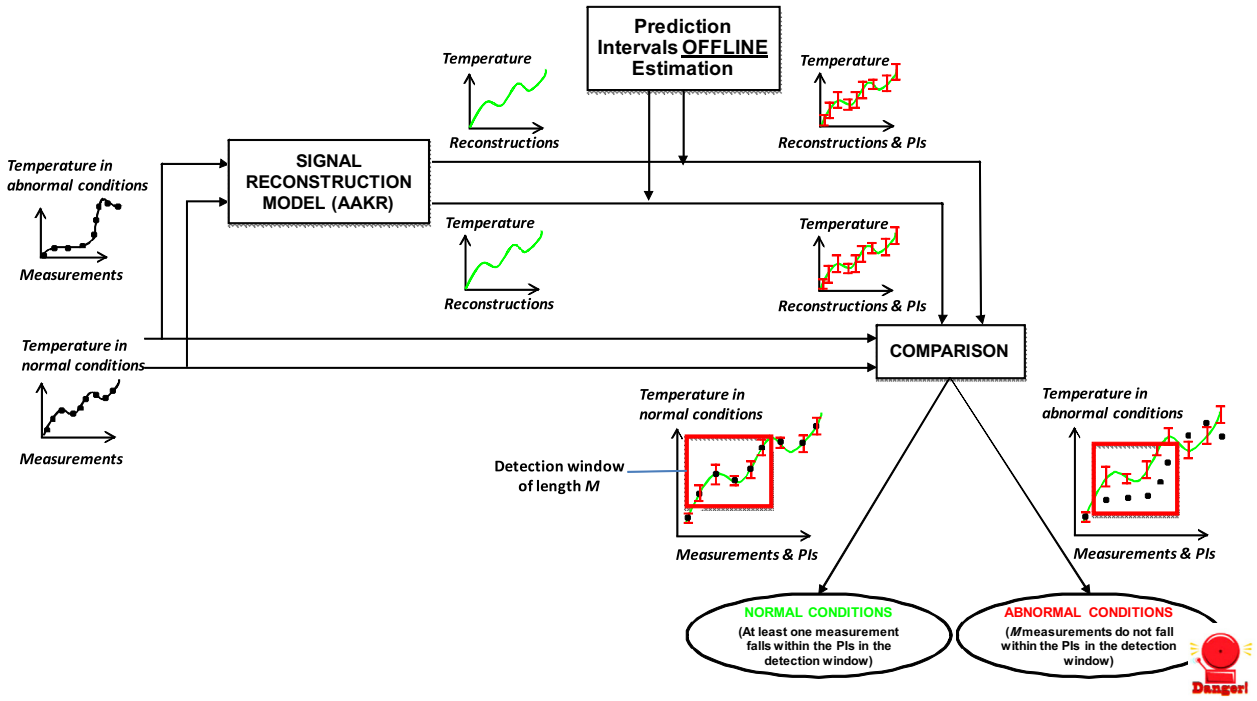

Fig. 3. Novel FD system taking the uncertainty on the reconstructions into account.

4. Artificial Case Study

An artificial case study has been designed to generate transients representative of the start-up behavior of a component (Baraldi et al., 2013a). Each transient, $f_i(x(t,\ 1),\ldots,x(t,\ 4))$, is four-dimensional (i.e., $n = 4$ signals) and has a time horizon of $N_p = 101$ time steps, in arbitrary units of measurements. The choice of testing this novel fault detection system on multivariate signals is due to the fact that the complex industrial components are commonly continuously monitored using redundant sensors measuring different signals taken on different components, i.e., multivariate signals, to collect as diverse and redundant as possible information (Gross et al., 1997).

With respect to normal conditions, 4000 transients representing the start-up of the component have been simulated. The signal evolutions are characterized by a sigmoid behavior $x_i^{nc}(t_k), k = 1,\ldots,101, i = 1,\ldots,4000$ given by Eq. (5):

$$x_i^{nc}(t_k) = 2a\left((1 + erf\left\{\frac{t_k - \mu}{\sqrt{2}}\right\}\right) + 10^{-3\zeta} \tag{5}$$

where α, μ and ζ are random parameters in arbitrary units. In practice, the simulations have been performed by sampling random values of the parameter ζ from a Gaussian distribution $\zeta \sim N(0,1)$, and of the parameters α and μ from uniform distribution functions with lower and upper bounds reported in Table 1.

Fig. 4 shows the obtained evolutions of the four signals in the 4000 transients, $x_i^{nc}(t_k), k = 1,\ldots 101,\ i = 1,\ldots,4000$.

Table 1
Limits of the uniform distributions from which the parameters in Eq. (5) have been sampled.

Parameter	Lower bounds	Upper bounds
a	0.45	0.55
μ	2.2	2.7

Among them, we have used $NT = 300$ transients to train the AAKR-built model, $NV = 59$ transients as validation set to optimize the value of the model parameter, i.e., the kernel bandwidth h (see Appendix C), and to estimate the PIs as described in Section 2.2. Once the detection window M is properly set as in Section 3, the remaining transients are used to verify the performance of the novel FD system by calculating the actual false alarm rate (γ).

In order to reduce the computational efforts and to increase model reconstruction accuracy, an ensemble reconstruction model made by $R = 5$ AAKR-built reconstruction models has been considered in this artificial case, each one dedicated to a different operational zone (Baraldi et al., 2012). The different operational zones are defined according to the time elapsed from the start of the transient and are reported in Table 2. In order to develop the overall reconstruction model, the training patterns are split into different sets, according to the time at which they have been measured. Then, for each operational zone, an AAKR model is built using the corresponding training set. Once the FD system has been built, it can be used on line for the signal reconstruction task by sending the test pattern to the corresponding reconstruction model.

Furthermore, 100 additional abnormal transients have been simulated in order to reproduce the signal behaviours in abnormal conditions by assuming a different time evolution for one signal randomly chosen among the four signals: the signal evolution is characterized by a sigmoid behavior given by Eq. (5) until a failure occurs at a random failure time, t_f; then, the signal evolution is characterized by Eq. (6) with parameters sampled from different uniform distributions due to the fact that the effect of an abnormal condition is usually directly related to the signal value (Baraldi et al., 2012) i.e., to the operational zone at which the failure occurs:

$$x_i^{ac}(t_k) = 2a^*\left((1 + erf\left\{\frac{t_k - \mu^*}{\sqrt{2}}\right\}\right) + 10^{-3\zeta} \tag{6}$$

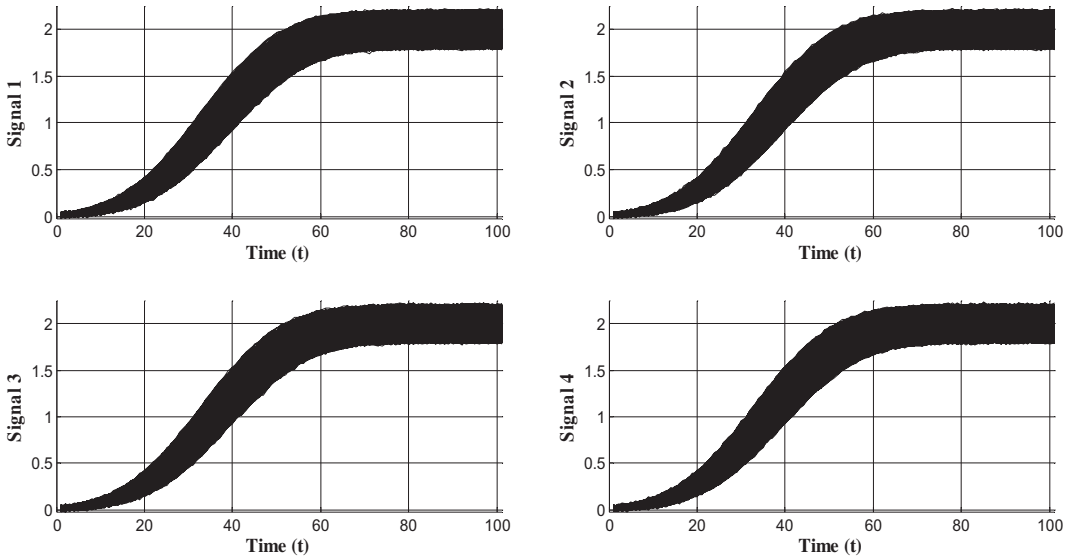

Fig. 4. Simulated time evolution in normal conditions of the 4 signals in 4000 start-up transients.

Table 2
Definition of the five operational zones for the four signals.

Operational zones	Time period	Operative conditions
Zone 1	1–20	Slow start up
Zone 2	21–40	Fast start up
Zone 3	41–60	Start converging to a steady state
Zone 4	61–80	Almost steadiness
Zone 5	81–101	Steady state (nominal value)

Parameters a^* and μ^* are randomly sampled from the uniform distributions of Table 3, that depend on the operational zones of failure occurrence.

As an example, Fig. 5 shows the evolution of the four signals in one of the simulated abnormal transients, $\vec{x}^{ac}_{i=1}(t_k)$, $k = 1, \ldots, 101$, if signal 1 (upper left) is randomly selected to be the failed signal at the randomly chosen time $t_f = 40$ (i.e., operational zone 2).

The AAKR models have been trained and their parameters optimized as described in Appendix C. In particular, the parameter h

Table 3
Limits of the uniform distributions from which the parameters in Eq. (6) have been sampled.

Parameter	Zone 1		Zone 2		Zone 3		Zone 4		Zone 5	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
a^*	0.57	0.59	0.57	0.61	0.57	0.63	0.57	0.65	0.57	0.67
μ^*	2.8	2.9	2.8	3	2.8	3.1	2.8	3.2	2.8	3.3

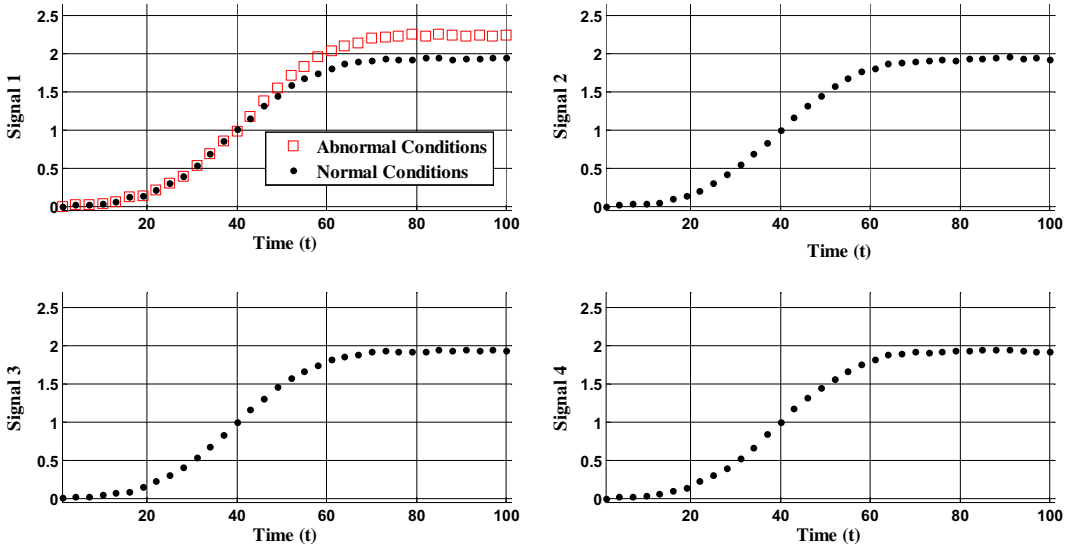

Fig. 5. Example of one simulated abnormal transient.

Table 4
h values for the $R = 5$ operational zones.

Operational zones	Time period	*h* values
Zone 1	1–20	0.05
Zone 2	21–40	0.05
Zone 3	41–60	0.01
Zone 4	61–80	0.009
Zone 5	81–101	0.005

values have been identified by optimizing the accuracy of the signal reconstructions in normal conditions and their robustness in abnormal conditions.

Table 4 reports the obtained optimal values of parameter *h* in the different operational zones.

4.1. PI estimation

The PIs obtained in the reconstructions of signal 1, $x^{test}(t_k, 1)$, of a test transient by considering $R = 5$ AAKR-built reconstruction models and Eq. (3), are shown in Fig. 6. Notice that the PI widths are variable during the time evolution and with an acceptable width increasing from zone 1 (time from 1 to 20) to zone 3 (time from 41 to 60). This is due to the variability of the training patterns used to train the reconstruction model, which is lower at the beginning of the transient.

4.2. Verification of the FD system capability

Results obtained by applying the novel FD system of Section 3 to the artificial case study are presented. Firstly, statistical independence of the residuals is tested with the Durbin–Watson test (Montgomery et al., 2001): small values of the test outcome *d* indicate that residuals are, on average, positively correlated, whereas if $d \gg 2$, they are, on average, negatively correlated. In this case, when the test is applied independently on the residuals of each one of the $n = 4$ signals, *d* turns out to be in the range (1.4–2.6), and, thus, we can conclude that the residuals can be considered as statistically independent; therefore, the hypotheses for applying the proposed sequential decision strategy hold and the optimum length of the detection window, *M*, can be calculated by solving Eq. (4) (see also Appendix B) with $N_p = 101$ time steps, $\gamma_{max} = 0.01$ and the level of confidence on the PI, $1 - \sigma = 0.95$.

$$P = 1 - [1 - 0.05^{M}]^{(101-M+1)\times 4} \leq 0.01 \rightarrow M = 4$$

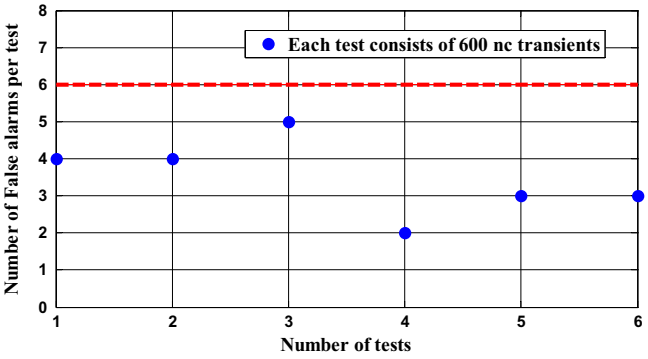

Fig. 7. Number of false alarms of 6 different extensive tests, each with 600 normal transients.

4.2.1. Normal conditions: false alarm rate

In order to verify whether the false alarm rate of the test transients in normal conditions is satisfactory, i.e., less than $\gamma_{max} = 0.01$, we have performed 6 extensive tests each using 600 normal transients. Fig. 7 shows that for all these extensive tests, the number of false alarms obtained is smaller than the maximum allowed limit of false alarms, i.e., $600 \times \gamma_{max} = 600 \times 0.01 = 6$ false alarms.

As an example, Fig. 8 shows the (wrong) detection of one false alarm: $M = 4$ consecutive measurements (time 40–43) of signal 3 fall outside their corresponding PIs within the detection window. Hence, the alarm is triggered and the decision maker should draw the conclusion that an abnormal condition at time $t = 43$ is occurred. The PIs of signals 1, 2, and 4 have been magnified by a factor of 20 and only some points have been shown for ease of visualization, and a zoom-in of signal 3 is shown from time 40 to time 43 at which the abnormality of the transient is declared.

4.2.2. Abnormal conditions: missing alarm rate

In order to verify the ability of the FD system to correctly detect the abnormality of the component, we have performed a test using the 100 available simulated abnormal transients; among these:

1. Abnormality in 87 transients has been correctly detected and the failed signals identified, with a delay in detection at least equal to $M = 4$ measurements. In other words, faults and their causes (i.e., the failed signals) have been correctly isolated/identified. This can help the decision maker to properly plan maintenance interventions. In this regard, this FD system can be, thus, seen as a fault isolation system as well as Non Linear

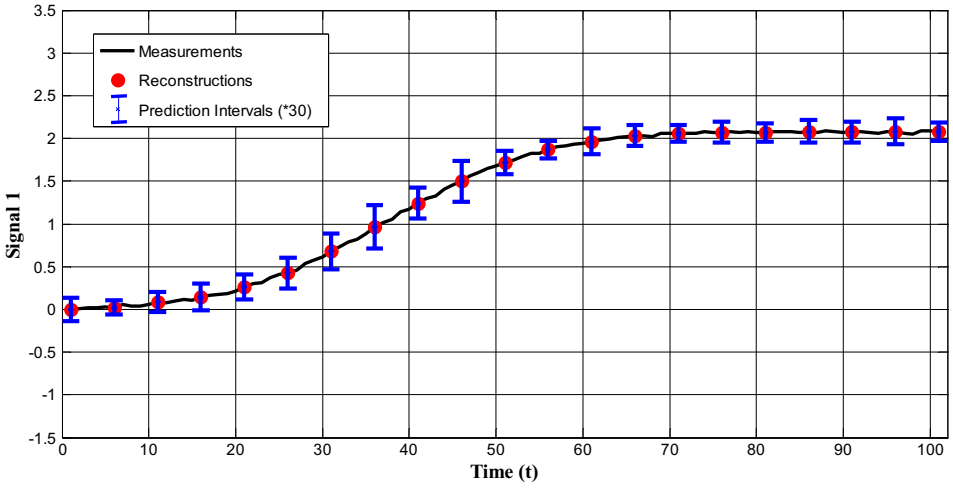

Fig. 6. PIs of the reconstruction of 21 patterns obtained using Eq. (3).

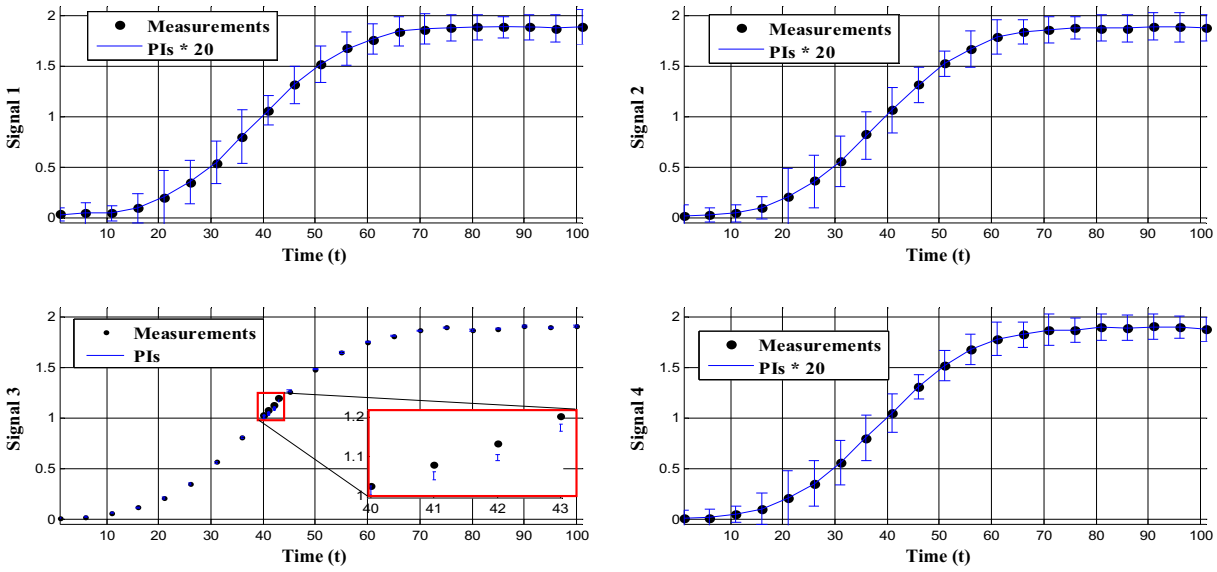

Fig. 8. An example of one false alarm: (wrong) detection for signal 3 at time $t = 43$.

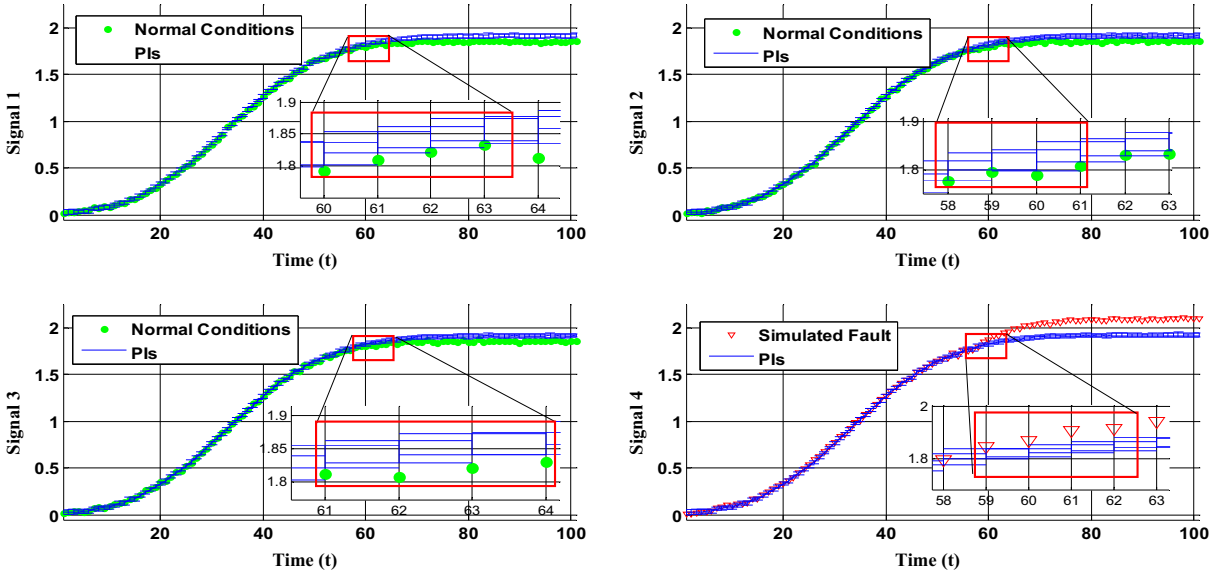

Fig. 9. An example of abnormality detection: the fault is simulated for signal 4 and imputed to signal 2.

Principal Component Analysis (NLPCA) (Harkat et al., 2007) and Principal Component Analysis (PCA) (Dunia et al., 1996; Kerschen et al., 2005).

2. Abnormality has been detected in 3 transients with due delay equal to M, although the failure has been imputed to one of the other (normal) signals: as an example, Fig. 9 shows that although the fault is simulated for signal 4 starting at time $t_f = 58$, it is imputed to signal 2 and detected at time $t_f = 61$, due to the first exceedance of the PIs of $M = 4$ successive measurements among the four signals (for signals 1,3 and 4 t_f should be equal to 63, 64 and 62, respectively).
3. Abnormality in 4 transients has been detected either at the failed signal or the other three normal signals but, the detection is anticipated, and
4. 6 transients have been considered in normal conditions although they are actually in abnormal conditions, i.e., the missing alarm rate (β) is equal to 0.06. This can be explained with the support of Fig. 10: in this case, the simulated failure occurs on signal 2 at random time, $t_f = 78$. However, the sampled anomalous parameters a^* and μ^* do not differ enough from the normal condition parameters a and μ and make this transient behave like the normal historical measurements used as training transients (yellow curves in upper right plot).

In this regard, considering the 90 detected anomalies i.e., case 1 and case 2, we can draw the probability density function (pdf) of the delay time, that is at least equal to $M = 4$ (Fig. 11): in some cases, more time is needed for the detection due to the uncertain behavior of the anomalous transients that do not radically differ from the normal transients. The mean delay in abnormality detection turns out to be equal to 8 measurements (including the $M = 4$ measurements that are necessary for stating the evidence of the anomaly), that is the time that the decision maker should wait before taking a decision regarding the health state of the component.

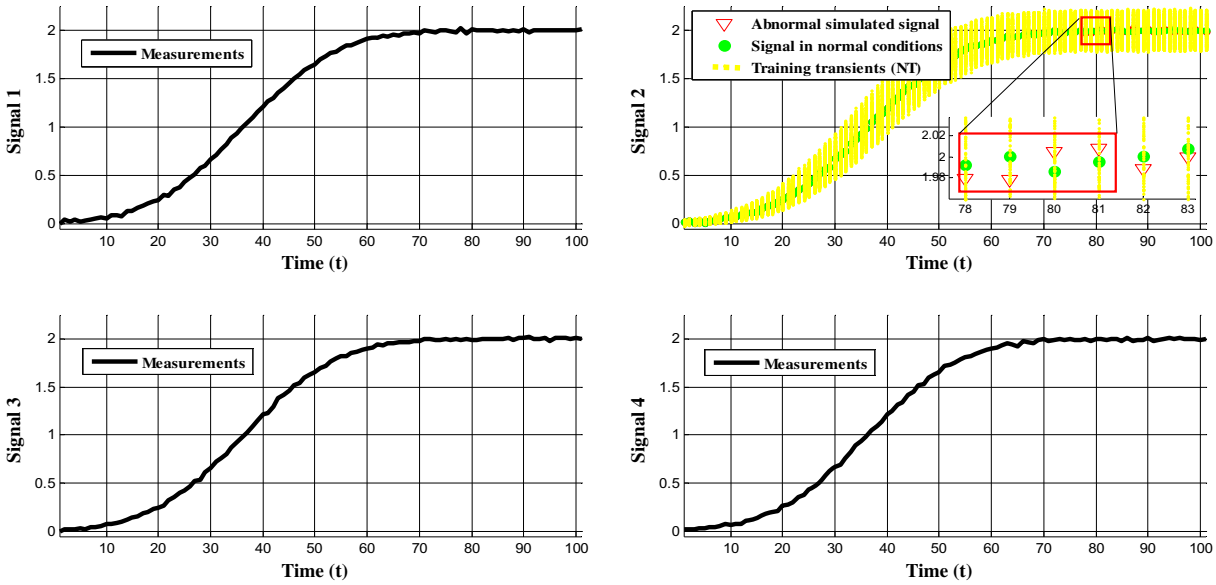

Fig. 10. An example of abnormal transient with the historical normal transients.

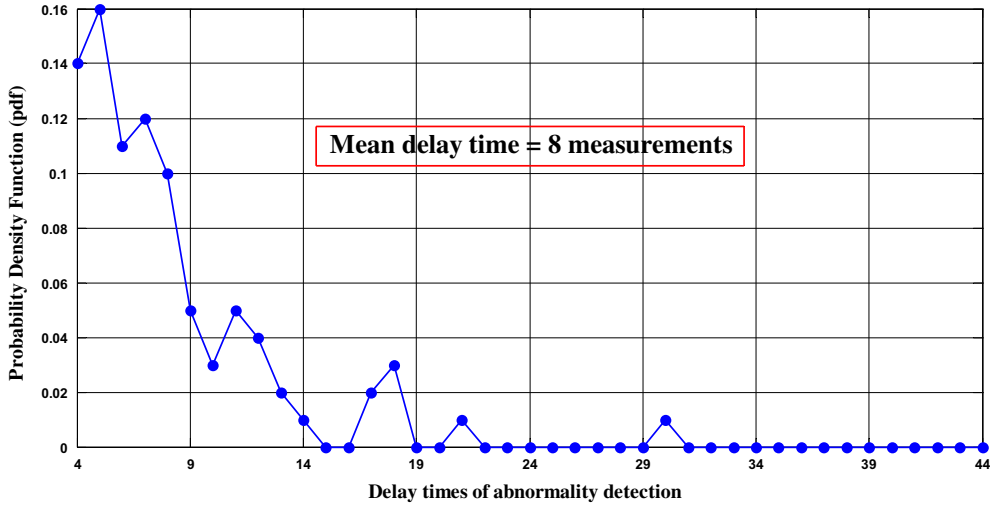

Fig. 11. Probability density function (pdf) of the delay times in abnormality detection.

5. Real Case Study

The novel FD system has been applied to a real industrial case concerning $N = 27$ shut-down multidimensional transients of a nuclear power plant (NPP) turbine. Each transient is three-dimensional (i.e., $n = 3$ vibration signals of the turbine shaft) and has a time horizon of $N_p = 4500$ time steps. Among these, 26 transients with similar functional behavior have been used as representative of normal conditions, whereas 1 transient with a different functional behavior has been used as representative of abnormal conditions. It is, indeed, an actual signature of abnormal conditions of the turbine during operation (whose detailed characteristics cannot be discussed, due to confidentiality reasons). Fig. 12 shows the evolutions of the three signals in the 27 transients and the corresponding turbine speed values.

It is worth mentioning that in order to overcome problems connected with the possibility of misalignment between transients, i.e., amplified and/or delayed transients and, hence, to easily apply the PI estimation method and the novel decision tool to synchronized transients, the signals have been transformed into the "turbine speed-domain" instead of the "time-domain" (Baraldi et al., 2013b). The transformation is obtained by considering only the time interval, where the turbine speed of those transients decreases from 1000 rpm to 600 rpm. In this way, it is possible to focus only on the decreasing part of the shut-down transients, avoiding problems due to the operative oscillations around the steady-state different plant operational conditions and to time misalignment of transients before the shut-down transients start (Baraldi et al., 2013b).

In this work, a single AAKR-built model has been developed, where we have used for training and validation:

- $NT = 4$ transients to train the model to provide the signal reconstructions at 25 representative turbine shaft speeds (equally spaced between 1000 rpm and 600 rpm),
- $NV = 14$ transients to optimize the value of the model parameter, i.e., the kernel bandwidth h, and to off-line estimate the PIs. The number $NV = 14$ of validation transients is properly defined by OS theory (Secchi et al., 2008) for guaranteeing a

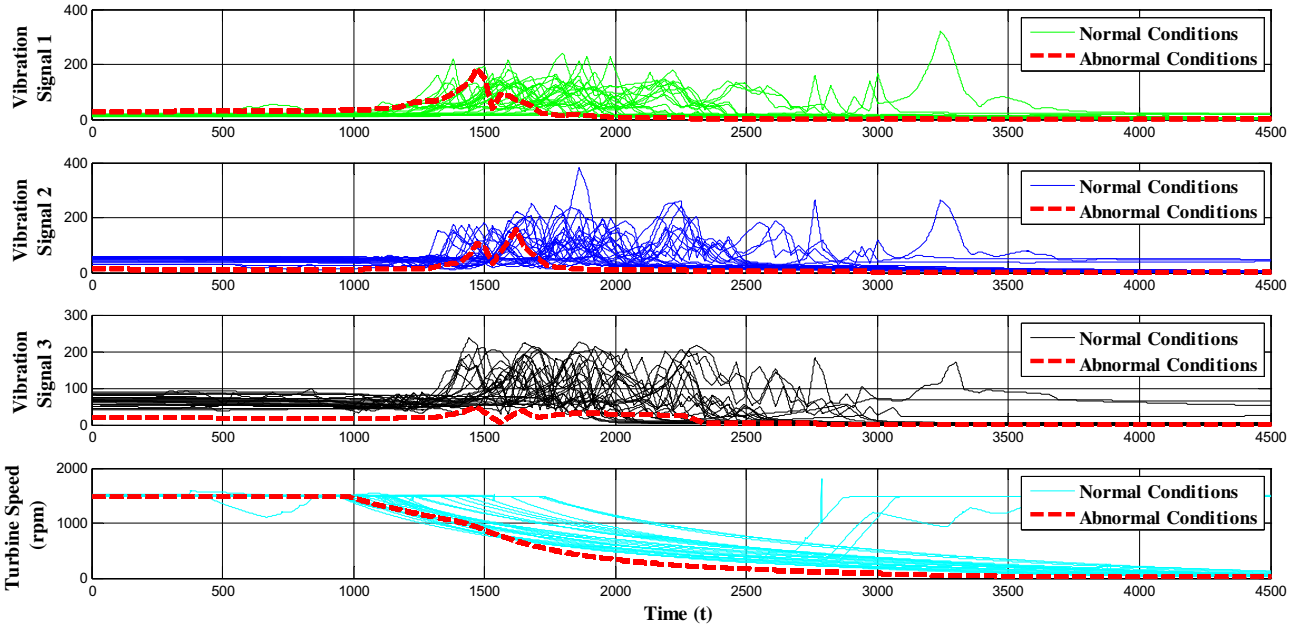

Fig. 12. The evolutions of the three vibration signals in the 27 transients and the corresponding turbine speed values.

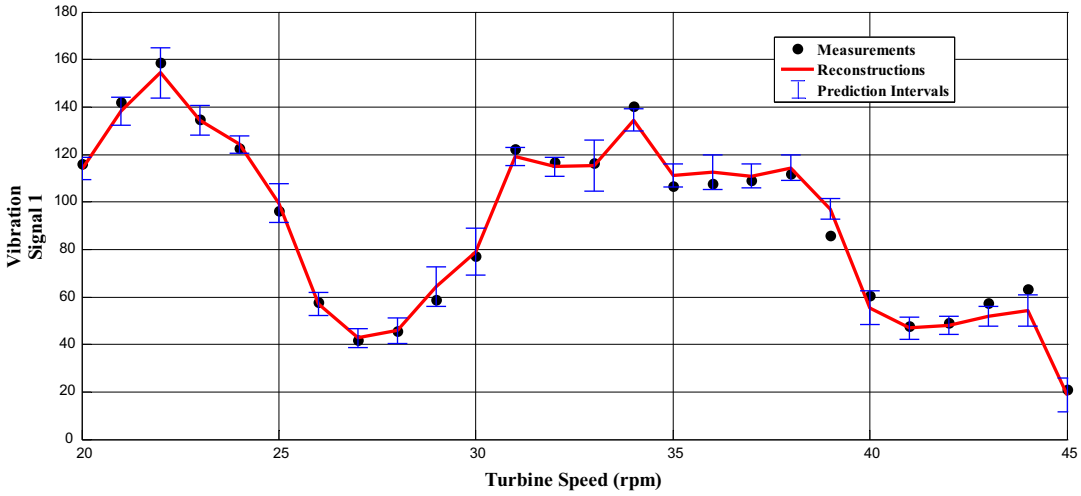

Fig. 13. PIs of the reconstruction obtained using Eq. (3).

PI with a confidence level, $1-\sigma$, that is limited to 80% (Al-Dahidi et al., 2014), due to the scarce number of measurements (16) at each representative turbine shaft speed.

- The remaining transients are used to verify the performance of the novel FD system for a series of 25 representative turbine shaft speeds.
- The parameter h value has been set to 0.1, by optimizing as described in Appendix C the accuracy of the signal reconstructions in normal conditions and their robustness in abnormal conditions.

5.1. PI estimation

The PIs obtained in the reconstructions of vibration signal 1 of a test transient by a single AAKR-built reconstruction model with optimum $h = 0.1$, and Eq. (3), are shown in Fig. 13. Notice that the PI widths are variable during the time evolution and with an acceptable width with respect to the variance of the vibration signal.

5.2. Verification of the FD system capability

Results obtained by applying the novel FD system of Section 3 to the real industrial case are presented. The Durbin–Watson test is applied independently on the residuals of each one of the $n = 3$ signals: d turns out to be in the range of (0.7–1.7), and, thus, we can conclude that the measurements are not statistically independent, and slightly positively correlated. Despite that, the novel FD system shows good performances in terms of false and missing alarm rates as we shall see in Sections 5.2.1 and 5.2.2. In this case, the optimum length of the detection window, M, with $N_p = 25$ representative turbine shaft speeds, $\gamma_{max} = 0.01$ and the level of confidence on the PI, $1-\sigma = 0.8$, is calculated by solving Eq. (4) (see also Appendix B):

$$P = 1 - \left[1 - (0.2)^{M}\right]^{(25-M+1)*3} \leq 0.01 \rightarrow M = 6$$

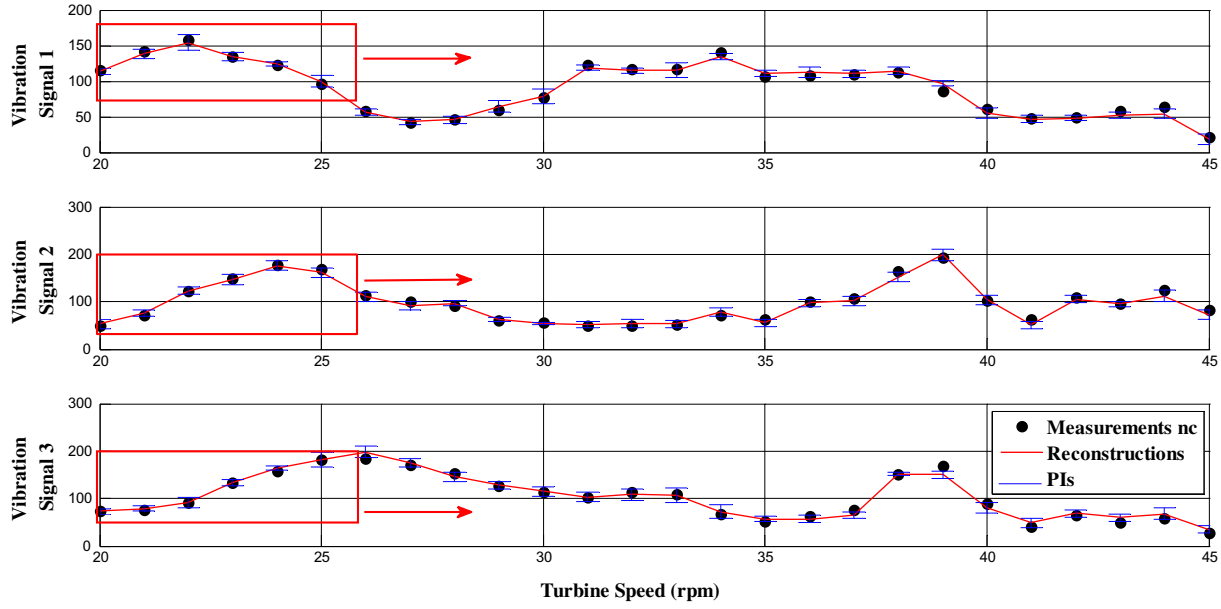

Fig. 14. An example of PIs and the reconstructions in one normal transient.

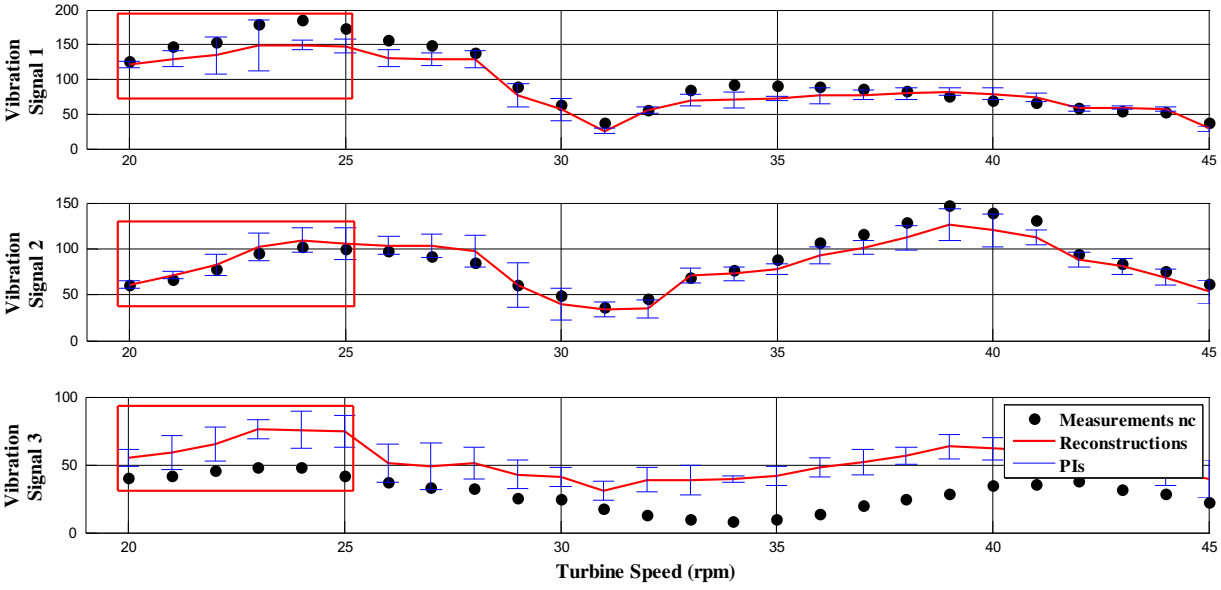

Fig. 15. An example of PIs and the reconstructions in one abnormal transient.

5.2.1. Normal conditions: false alarm rate

In order to verify the ability of the novel FD system to identify the correct health state of the turbine i.e., normal conditions in this case, a cross-validation procedure is used (Bendat and Piersol., 2010). Specifically, a 10-fold cross-validation is implemented to calculate the actual false alarm rate (γ): the test transients are fixed whereas the training and validation are randomly partitioned into 4 blocks of equal size; one of these blocks is used as a training data subset whereas the remaining 3 blocks are combined together to constitute the validation data subset. The cross-validation process is then repeated 10 times (10 folds), each time using different blocks for training and validation sets. Combining the false alarm rates of the 10 folds, the obtained false alarm rate is equal to 0.

Fig. 14 shows an example of the obtained PIs and the reconstructions in one normal test transient for one of the 10 cross-validations. Using a detection window of length $M = 6$, 6 consecutive measurements falling outside their estimated PIs are never found and, hence, the normality of the turbine can be confirmed.

5.2.2. Abnormal conditions: missing alarm rate

For the abnormality detection, the test transient of Fig. 12 has been used. Its reconstructions, provided by the AAKR-built model, and its PIs have been used for comparison and anomaly detection. Fig. 15 shows the ability of the method to promptly detect that this transient differs from those of normal conditions: 6 consecutive measurements at representative turbine shaft speeds fall outside the estimated PIs of the vibration signal 3, as soon as the transient starts.

Again, the isolation of the faulty signal can be useful for the decision maker to identify the cause of abnormality and, thus, to properly plan maintenance interventions.

6. Conclusions

In this work, a novel, non-parametric, sequential decision tool is proposed to help the decision maker in assessing the normal or abnormal conditions of a component, taking into account the quantified uncertainties that affect the component behavior and the FD system itself. The quantification of the uncertainty of the reconstructions is based on the estimation of Prediction Intervals (PIs) by using Order Statistics (OS) theory at a pre-defined confidence level, e.g., 95%. Then, the novel decision rule consists in properly setting the number M of consecutive measurements that have to be covered by the estimated PIs. In practice: (a) if at least one out of M consecutive measurements falls within the corresponding PIs, the component is assumed to work in normal conditions and the alarm is not triggered, (b) if the M consecutive measurements do not fall within the corresponding PIs, abnormal conditions are detected and the alarm is triggered. The optimum length M of the detection window is estimated at a prefixed maximum limit of false alarm rate, γ_{max}. The Auto-Associative Kernel Regression (AAKR) method is adopted to build the empirical model of signal reconstructions.

The novel FD system has been tested using an artificial case study representing the monitoring of a component during typical start-up transients and validated using a real industrial case concerning 27 shut-down transients of a nuclear power plant (NPP) turbine. The obtained results show that the approach is able to guarantee low false alarm rates ($\leqslant\gamma_{max}$, e.g., =0.01) and missing alarm rates and, hence, provide the decision makers with the robust information for establishing whether a maintenance intervention is required or not.

In the overall CBM scheme, other sources of uncertainty such as those related to the decision maker attitude towards the decision of performing a maintenance intervention once the FD system detects abnormal conditions, should be considered. This will be object of future research work.

Acknowledgements

This research has been carried out within the European Union Project INNovation through Human Factors in risk analysis and management (INNHF, www.innhf.eu) funded by the 7th Framework Program FP7-PEOPLE-2011-Initial Training Network: Marie-Curie Action. The participation of Enrico Zio to this research is partially supported by the China NSFC under grant number 71231001.

Appendix A. Auto-associative Kernel Regression (AAKR)

Auto-Associative Kernel Regression (AAKR) is a non-parametric, empirical modelling technique that relies on historical measurements of the signals taken during normal conditions of the component to predict (reconstruct) the current signal measurements vector at a given time t, $\vec{x}^{test}(t) = [x^{test}(t,1), x^{test}(t,j)\ldots, x^{test}(t,n)]$, $j = 1,\ldots,n$; where n is the number of measured signals, e.g., pressure, temperature, vibration, etc. as a weighted sum of those historical observations. The historical measurements performed at past time t_k, $k = 1,\ldots,N_{train}$ are collected into the matrix $\overline{\overline{X}}$ whose generic element $x(t_k, j)$ is the measured value of signal j at time t_k (Baraldi et al., 2011, 2012; Al-Dahidi et al., 2014).

Then the estimated value, $\widehat{x}^{test}(t,j)$, of the measurement, $x^{test}(t,j)$, of the j-th component at time t is given by:

$$\widehat{x}^{test}(t,j) = \frac{\sum_{k=1}^{N_{train}} w(t_k)\cdot x(t_k,j)}{\sum_{k=1}^{N_{train}} w(t_k)} \tag{A1}$$

Weights $w(t_k)$ are similarity measures obtained by computing the Euclidean distance between the current sensor measurement $x^{test}(t,j)$ and the k-th observation of $\overline{\overline{X}}$ (Eq. (A2)):

$$d^2(t_k) = \sum_{j=1}^{n}\left(x^{test}(t,j) - x(t_k,j)\right)^2 \tag{A2}$$

and inserting it in the Gaussian kernel (Eq. (A3)):

$$w(t_k) = \frac{1}{\sqrt{2\pi}h} e^{-\frac{d^2(t_k)}{2h^2}} \tag{A3}$$

where h is the Gaussian kernel bandwidth.

In order to provide in Eq. (A2) a common scale across the different signals measuring different quantities, it is necessary to normalize their values. In the present work, the signal values at time t are normalized according to:

$$x^{test-normalized}(t,j) = \frac{x^{test}(t,j) - \mu(j)}{\sigma(j)} \tag{A4}$$

where, $x^{test}(t,j)$ is a generic measurement of signal j, $\mu(j)$ and $\sigma(j)$ are the mean and the standard deviation of the j-th signal in $\overline{\overline{X}}$ (Eq. (A5)):

$$\mu(j) = \frac{\sum_{k=1}^{N_{train}} x(t_k,j)}{N_{train}}$$
$$\sigma(j) = \sqrt{\frac{\sum_{k=1}^{N_{train}} (x(t_k,j)-\mu(j))^2}{N_{train}}} \tag{A5}$$

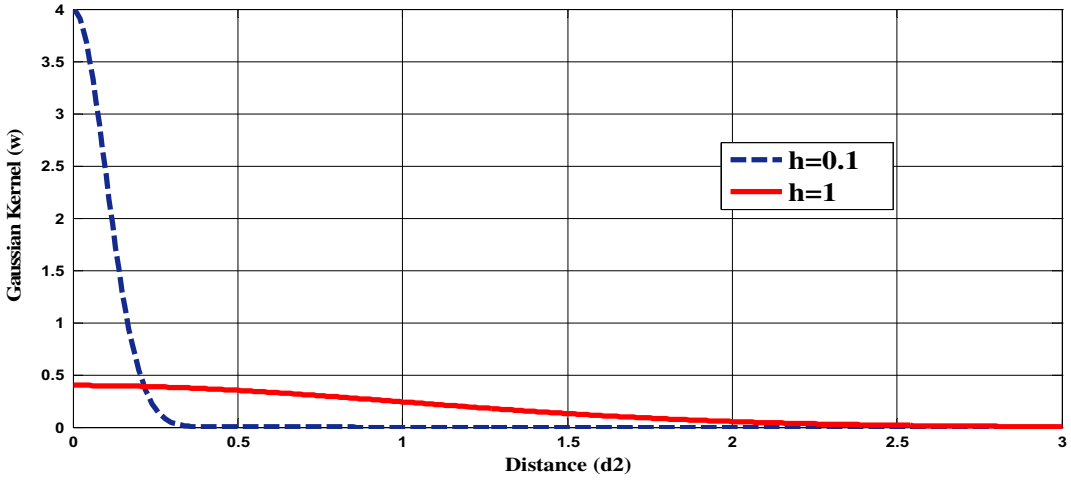

Fig. 16. Gaussian Kernel function with two h values.

Appendix B. Optimization of the decision rule parameter *M*

In case of normal conditions, for each *j*-th signal of the *N* *n*-dimensional transients of time horizon N_p and a given detection window of length *M*, the probability of false alarm P_1 is given by Eq. (A6) (assuming that measurements are statistically independent and that the probability that the measurement in normal conditions $x^{test}(t_k)$ of signal *j*, $j = 1,\ldots,n$ at time t_k, $k = 1,\ldots,N_p$, falls within the interval $[\widehat{x}^{lower}(t_k), \widehat{x}^{upper}(t_k)]$ is equal to $1-\sigma$:

$$P_1 = \text{Probability}\{\text{False alarm for 1 signal taken in a time window of length } M\} = \sigma^M \tag{A6}$$

The probability P_n that a false alarm is given when *n* signals are considered for a given detection window of length *M* is given by Eq. (A7), assuming the signals are statistically independent:

$$\begin{aligned} P_n &= \text{Probability}\{\text{False alarm for } n \text{ signals taken in a time window of length } M\} = = 1 - P\{\text{No false alarms for } n \text{ signals taken in a time window of length } M\} \\ &= 1-(1-P_1)^n = 1-(1-\sigma^M)^n \end{aligned} \tag{A7}$$

Thus, the probability *P* that a false alarm is triggered when a transient of time horizon N_p is observed is given by Eq. (A8):

$$P = 1 - P\{\text{No false alarms for any of the } (N_p - M + 1) \text{ detection windows}\} = 1-(1-P_n)^{(N_p-M+1)} \tag{A8}$$

Substituting Eq. (A7) in Eq. (A8), the probability *P* (Eq. (A8)) can be re-written as:

$$P = 1-\left[1-\left(1-(1-\sigma^M)^n\right)\right]^{(N_p-M+1)} = 1-\left[1-\sigma^M\right]^{(N_p-M+1)\times n} \tag{A9}$$

Thus, the optimum length *M* of the detection window can be calculated using Eq. (A10), where γ_{max} is the maximum acceptable false alarm rate:

$$P\{\text{False Alarm}\} = P \le \gamma_{max} \tag{A10}$$

Appendix C. Kernel's Bandwidth (*h*) Optimization

The value of the kernel bandwidth has to be optimized to have a balance between the AAKR accuracy and robustness following a trial-and-error procedure: (1) the accuracy which is the ability of the model to correctly and accurately reconstruct the signal values of a component in normal conditions: An accurate fault detection (FD) system allows reducing the number of false alarms, and (2) the robustness which is the ability of the model to reconstruct the signal values of a component in abnormal conditions: a robust AAKR model reconstructs the value of a measured signal as if the component is in normal conditions thus, allows reducing the number of missing alarms.

A local optimum value of *h* and a misleading setting of *h* may lead to inaccurate reconstructions that have to be tackled by properly quantifying the reconstructions model uncertainty. As an example, in Fig. 16 it can be seen that with a small bandwidth ($h = 0.1$) large weights (similarities) are assigned to historical data whose distance is very close to zero, whereas with a larger bandwidth ($h = 1$), the weight assignment is less specific (Office of Nuclear Regulatory Research, 2007).

References

Ahmad, R., Kamaruddin, S., 2012. An overview of time-based and condition-based maintenance in industrial application. Computers & Industrial Engineering 63 (1), 135–149.

Al-Dahidi, S., Baraldi, P., Di Maio, F., Zio, E., 2014. Quantification of Signal Reconstruction Uncertainty in Fault Detection Systems, the Second European Conference of the Prognostics and Health Management Society 2014. Nantes, France.

Aven, T., Zio, E., 2012. Foundational issues in risk assessment and risk management. Risk Analysis 32 (10), 1647–1656.

Baraldi, P., Cammi, A., Mangili, F., Zio, E., 2010. An ensemble approach to sensor fault detection and signal reconstruction for nuclear system control. Ann. Nucl. Energy 37 (6), 778–790. http://dx.doi.org/10.1016/j.anucene.2010.03.002, ISSN 0306-4549.

Baraldi, P., Canesi, R., Zio, E., Seraoui, R., Chevalier, R., 2011. Genetic Algorithm-based Wrapper Approach for Grouping Condition Monitoring Signal of Nuclear Power Plant Components. Integrated Computer-Aided Engineering 18 (3), 221–234.

Baraldi, P., Di Maio, F., Pappaglione, L., Zio, E., Seraoui, R., 2012. Condition monitoring of electrical power plant components during operational transients. Proceedings of the Institution of Mechanical Engineers, Part O, Journal of Risk and Reliability 226 (6), 568–583.

Baraldi, P., Di Maio, F., Zio, E., 2013a. Unsupervised Clustering for Fault Diagnosis. International Journal of Computational Intelligence Systems 6 (4), 764–777.

Baraldi, P., Di Maio, F., Rigamonti, M., Zio, E., Seraoui, R., 2013b. Transients analysis of a nuclear power plant component for fault diagnosis. Chem. Eng. Trans. 33, 895–900.

Bendat, J.S., Piersol, A.G., 2010. Random Data: Analysis and Measurement Procedures, Fourth Edition. John Wiley & Sons, Hoboken, New Jersey.

Bouckaert, R.R., Frank, E., Holmes, G., Fletcher, D., 2011. A comparison of methods for estimating prediction intervals in NIR spectroscopy: size matters. Chemom. Intell. Lab. Syst. 109 (2), 139–145.

Campos, J., 2009. Development in the application of ICT in condition monitoring and maintenance. Comput. Ind. 60 (1), 1–20.

Di Maio, F., Baraldi, P., Zio, E., Seraoui, R., 2013. Fault detection in nuclear power plants components by a combination of statistical methods. IEEE Trans. Reliab. 62 (4).

Dixon, W.E., Walker, I.D., Dawson, D.M., Hartranft, J.P., 2000. Fault detection for robot manipulators with parametric uncertainty: a prediction-error-based approach. Rob. Autom. IEEE Trans. 16 (6), 689–699.

Dong, D., McAvoy, T., 1994. Sensor Data Analysis Using Autoassociative Neural Networks. Proc. World Congr. Neural Networks, San Diego, CA 1, 161–166.

Dunia, R., Qin, S.J., Edgar, T.F., McAvoy, T.J., 1996. Identification of faulty sensors using principal component analysis. AIChE J. 42 (10), 2797–2812.

Ebron, S., Lubkerman, D., White, M., 1990. A neural network approach to the detection of incipient faults on power distribution feeders. IEEE Trans. Power Delivery 5 (2), 905–914.

Emami-Naeini, A., Akhter, M.M., Rock, S.M., 1988. Effect of model uncertainty of failure detection: the threshold selector. IEEE Trans. Autom. Control 33 (12), 1106–1115.

Fantoni, P.F., Mazzola, A., 1996. A pattern recognition-artificial neural networks based model for signal validation in nuclear power plants. Ann. Nucl. Energy 23 (13), 1069–1076. http://dx.doi.org/10.1016/0306-4549(96)84661-5, ISSN 0306-4549.

Gross, K.C., Humenik, K.E., 1991. Sequential probability ratio test for nuclear plant component surveillance. Nucl. Technol. 93, 131.

Gross, K.C., Singer, R.M., Wegerich, S.W., Herzog, J.P., VanAlstine, R., Bockhorst, F., 1997. Application of a Model-Based Fault Detection System to Nuclear Plant Signals (No. ANL/RA/CP–92110; CONF-970765–2). Argonne National Lab.

Guo, P., Bai, N., 2011. Wind turbine gearbox condition monitoring with AAKR and moving window statistic methods. Energies 4, 2077–2093.

Harkat, M.F., Djelel, S., Doghmane, N., Benouaret, M., 2007. Sensor fault detection isolation and reconstruction using nonlinear principal component analysis. Int. J. Autom. Comput. 4 (2), 149–155.

Helton, J.C., 1994. Treatment of uncertainty in performance assessments for complex systems. Risk Anal. 14 (4), 483–511.

Heo, G.Y., 2008. Condition monitoring using empirical models: technical review and prospects for nuclear applications. Nucl. Eng. Technol. 40 (1), 49.

Hines, J.W., Davis, E., 2005. Lessons learned from the US nuclear power plant on-line monitoring programs. Prog. Nucl. Energy 46 (3), 176–189.

Hines, J.W., Garvey, D.R., 2006. Development and application of fault detectability performance metrics for instrument calibration verification and anomaly detection. J. Pattern Recognit. Res. 1, 2–15.

Hines, J.W., Wrest, D.J., Uhrig, R.E., 1997. Signal validation using an adaptive neural fuzzy inference system. Nucl. Technol. 119, 181–193.

Jardine, A.K., Lin, D., Banjevic, D., 2006. A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mech. Syst. Signal Process. 20 (7), 1483–1510.

Kerschen, G., De Boe, P., Golinval, J.C., Worden, K., 2005. Sensor validation using principal component analysis. Smart Mater. Struct. 14 (1), 36.

Maki, Y., Loparo, K.A., 1997. A neural-network approach to fault detection and diagnosis in industrial processes. Control Syst. Technol. IEEE Trans. 5 (6), 529–541.

Miao, Q., Wang, D., Pecht, M., 2010. A probabilistic description scheme for rotating machinery health evaluation. J. Mech. Sci. Technol. 24 (12), 2421–2430.

Montes de Oca, S., Puig, V., Blesa, J., 2012. Robust fault detection based on adaptive threshold generation using interval LPV observers. Int. J. Adapt. Control Signal Process. 26 (3), 258–283.

Montgomery, D.C., Peck, E.A., Vining, G.G., 2001. Introduction to Linear Regression Analysis, 3rd ed. John Wiley & Sons, New York.

Office of Nuclear Regulatory Research, 2007. Technical Review of On-Line Monitoring Techniques for Performance Assessment, Vol. 2: Theoretical Issues, ORNL/TM-2007/188, NUREG/CR-6895.

Puig, V., Quevedo, J., Escobet, T., Nejjari, F., de las Heras, S., 2008. Passive robust fault detection of dynamic processes using interval models. IEEE Trans. Control Syst. Technol. 16, 1083–1089.

Rasmussen, B., Wesley Hines, J., Gribok, A.V., 2003. An applied comparison of the prediction intervals of common empirical modeling strategies. In: Proceedings of the 2003 Annual Maintenance and Reliability Conference, Knoxville, TN, May 4–7.

Schoonewelle, H., van der Hagen, T.H.J.J., Hoogenboom, J.E., 1995. Theoretical and numerical investigations into the SPRT method for anomaly detection. Ann. Nucl. Energy 22 (11), 731–742.

Secchi, P., Zio, E., Di Maio, F., 2008. Quantifying uncertainties in the estimation of safety parameters by using bootstrapped artificial neural networks. Ann. Nucl. Energy 35 (12), 2338–2350.

Thurston, M., Lebold, M., 2001. Standards Developments for Condition-Based Maintenance Systems. Ft. Belvoir Defense Technical Information Center, Applied Research Laboratory, Penn State University.

Wald, A., 1947. Sequential Analysis. John Wiley & Sons, New York, NY.

Weber, P., Theilliol, D., Aubrun, C., Evsukoff, A., 2007. Increasing effectiveness of model-based fault diagnosis: a dynamic Bayesian network design for decision making. In: Proceedings of the Conference on Fault Detection, Supervision and Safety of Technical Processes, pp. 90–95.

Xu, X., Hines, J.W., Uhrig, R.E., 1999. Sensor validation and fault detection using neural networks. In: Proc. Maintenance and Reliability Conference (MARCON 99), pp. 10–12.

Yam, R.C.M., Tse, P.W., Li, L., Tu, P., 2001. Intelligent predictive decision support system for condition-based maintenance. Int. J. Adv. Manuf. Technol. 17 (5), 383–391.

Yang, D., Usynin, A., Hines, J.W., 2006. Anomaly-based intrusion detection for SCADA systems. In: 5th Intl. Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technologies (NPIC&HMIT 05), pp. 12–16.

Yu, C., Su, B., 2004. A non-parametric sequential rank-sum probability ratio test method for binary hypothesis testing. Signal Process. 84 (7), 1267–1272. http://dx.doi.org/10.1016/j.sigpro.2004.04.008, ISSN 0165-1684.

Zhao, W., Baraldi, P., Zio, E., 2011. Confidence in signal reconstruction by the evolving clustering method. In: IEEE Prognostics and System Health Management Conference, Shenzhen, China, 23–25 May.

Zheng, J., Frey, H.C., 2005. Quantitative analysis of variability and uncertainty with known measurement error: methodology and case study. Risk Anal. 25 (3), 663–675.