Quantification of Signal Reconstruction Uncertainty in Fault Detection Systems

Sameer Al-Dahidi[1], Piero Baraldi[1], Francesco Di Maio[1], and Enrico Zio[1,2]

[1] *Energy Department, Politecnico di Milano, Milan, 20133, Italy*

sameer.aldahidi@polimi.it
piero.baraldi@polimi.it
francesco.dimaio@polimi.it
enrico.zio@polimi.it

[2]*Chair on Systems Science and the Energetic challenge, European Foundation for New Energy- Électricité de France, École Centrale Paris and Supélec, Paris, 92295, France*

enrico.zio@ecp.fr; enrico.zio@supelec.fr

ABSTRACT

In Condition-Based Maintenance (CBM), Fault Detection (FD) systems monitor the health state of the components and aid the operator to decide whether a maintenance intervention is necessary. A FD system is a decision-aid tool typically based on i) a reconstruction model that estimates (reconstructs) the values of measurable signals in normal conditions, and ii) an analyzer of the differences (residuals) between the measured and reconstructed values: abnormal conditions are detected when residuals are statistically significant. The performance of the reconstruction model is influenced by several sources of uncertainty which can influence the operator decision: 1) measurement errors, 2) intrinsic stochasticity of the physical process, 3) uncertainty on the settings of the model parameters, and 4) uncertainty on the model output due to incompleteness of the training data. The objective of the present work is the quantification of the overall uncertainty affecting the model reconstructions. The proposed novel approach for uncertainty quantification relies on the estimation of Prediction Intervals (PIs) by using Order Statistics (OS) for a pre-defined confidence level. The proposed approach is verified with respect to an artificial case study; the obtained results show that the approach is able to guarantee the desired level of confidence on the correctness of the detection and provide the decision maker with the required information for establishing whether a maintenance intervention is necessary.

Keywords: Signal Reconstruction, Fault Detection, Uncertainty, Prediction Intervals, Auto-Associative Kernel Regression, Order Statistics, Scale Factor.

1. INTRODUCTION

Recent developments in data processing and computational capabilities are encouraging industries such as nuclear, oil and gas, chemical, automotive and aerospace to apply Condition-Based Maintenance (CBM) (Campos, 2009) for increasing system availability, reducing maintenance costs, minimizing unscheduled shutdowns and increasing safety (Thurston & Lebold, 2001).

A typical scheme of CBM can be described as follows: a Fault Detection (FD) system continuously collects information from sensors mounted on the component of interest (Ahmad & Kamaruddin, 2012; Montes de Oca, Puig & Blesa, 2012) and delivers a decision regarding its health state (either normal or abnormal conditions). In case of abnormal conditions, an alarm is triggered and the decision maker decides whether it is necessary to perform a maintenance action or it is possible to postpone it. In this work, we consider a FD system architecture based on an empirical reconstruction model and a decision tool.

Different empirical models have been used with success to estimate (reconstruct) the expected values of the signals in normal conditions. Typical examples include Artificial Neural Networks (ANNs) (Hines, Wrest & Uhrig, 1997; Safty, Ashour, Dessouki & Sawaf, 2004; Rahman, 2010), Auto-Associative Kernel Regression (AAKR) (Chevalier, Provost & Seraoui, 2009; Baraldi, Canesi, Zio, Seraoui & Chevalier, 2010; Baraldi, Di Maio, Pappaglione, Zio & Seraoui, 2012), Evolving Clustering Method (ECM) (Zhao, Baraldi & Zio, 2011), Principal Component Analysis (PCA) (Garcıa-Alvarez, 2009; Baraldi, Zio, Gola, Roverso & Hoffmann, 2011), Independent Principal Component

Analysis (Ding, Hines & Rasmussen, 2003), Support Vector Machines (SVMs) (Zavaljevski & Gross, 2000; Batur, Zhou & Chan, 2002; Laouti, Sheibat-Othman & Othman, 2011) and Fuzzy Similarity (Baraldi, Di Maio, Genini & Zio, 2013).

The decision tool is typically constructed on the analysis of the differences (residuals) between the measured and the reconstructed values of the n signals at time t, $\vec{x}^{test}(t)$ and $\hat{\vec{x}}^{test}(t)$, respectively, in order to decide whether the component is in normal or abnormal conditions (Figure 1). In practice, two possible cases may arise at time t: a) reconstructions are similar to measurements, $\hat{\vec{x}}^{test}(t) \approx \vec{x}^{test}(t)$ b) reconstructions are different from measurements, $\hat{\vec{x}}^{test}(t) \neq \vec{x}^{test}(t)$. In the former case, the component is recognized to be in normal conditions (nc) and the alarm is not triggered, whereas in the latter case abnormal conditions (ac) are detected and the alarm is triggered.

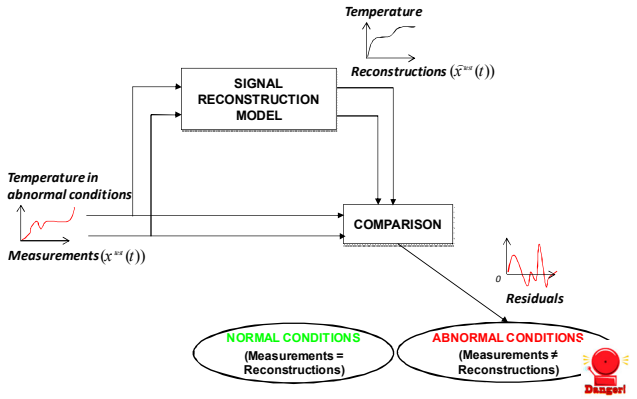


Figure 1. Traditional FD system.

Independently from the choice of the reconstruction model and of the method adopted to analyze the residuals, different sources of uncertainty may influence the performance of the FD system and can cause false or missing alarms (Helton, 1994; Zheng & Frey, 2005; Aven & Zio, 2012).

In this context, the present work focuses on the analysis of the uncertainty in the signal reconstruction phase of the FD process. In particular, we consider the following sources of uncertainty: 1) the measurement errors, 2) the inherent variability (stochasticity) of the physical process, 3) the uncertainty on the settings of the reconstruction model parameters, and 4) the uncertainty on the reconstruction model output due to incompleteness of the training data. The objective is the quantification of the overall uncertainty which the reconstructions provided by the empirical model are subject to. To this aim, we propose a novel method based on the estimate of Prediction Intervals (PIs) by using Order Statistics (OS) theory. For illustration purposes, we adopt the AAKR technique to build the reconstruction model, but the approach proposed is general and can be applied to any other techniques for developing the reconstruction model.

The method for the quantification of the uncertainty on the signal reconstructions is verified with respect to an artificial case study representing the behavior of a component during operational transients. This situation, characterized by a non-stationary behavior of the signals, has been chosen due to the criticality of the FD task during operational transients (Baraldi et al. 2012). In particular, the time evolution of 4 signals during various start-up transients have been simulated and used to assess the performance of the method in the quantification of the uncertainty on the reconstructions. Artificial data have been used in order to allow testing the approach on a large number of different simulated transients and, thus, to evaluate its capability of correctly quantify the uncertainty on the reconstruction.

The remaining of this paper is organized as follows; in Section 2, a description of the four sources of uncertainty to which a FD system is subject is provided. In Section 3, a reconstruction model for signal reconstruction during operational transients is developed, and a method for estimating the PIs of the reconstruction is proposed. In Section 4, an artificial case study representing the component behavior during typical start-up transients is introduced and, in Section 5, the results of the application of the proposed method are discussed. Finally, some conclusions are proposed in Section 6.

2. SOURCES OF UNCERTAINTY IN FD SYSTEMS

The reconstructions provided by an empirical model, e.g., AAKR (Chevalier et al., 2009; Baraldi et al., 2010; Baraldi et al., 2012), are subject to the following 4 sources of uncertainty (Lin & Stadtherr, 2008; Baraldi et al., 2011; Ramuhalli, Lin, Crawford, Konomi, Braatz, Coble, Shumaker & Hashemian, 2013):

1. the measurement errors, which can be due to systematic or random errors of the sensors;
2. the inherent variability (stochasticity) of the physical process, which causes different evolutions of the signal during identical operational transients: e.g., during two different start-up transients of the same component in the same environmental and operational conditions, different signal evolutions are observed.
3. the uncertainty on the correct setting of the AAKR-built model parameters. In practice, according to the AAKR method, signal reconstructions are built on the basis of a measure of similarity between the test pattern and "neighbouring" training patterns (Appendix A.1). The computation of the similarity measure is based on a kernel function characterized by a parameter, called bandwidth, whose value is typically set by following a trial and error procedure on some validation data.
4. the uncertainty caused by the incompleteness of the training data. The performance of an empirical signal

reconstruction model built by AAKR is remarkably influenced by the quality and quantity of the training patterns (Appendix A.1).

3. RECONSTRUCTION MODEL AND UNCERTAINTY QUANTIFICATION

A typical reconstruction model receives in input at time t a vector $\vec{x}^{test}(t) = \left[x^{test}(t,1), x^{test}(t,j), ..., x^{test}(t,n) \right]$ containing the test measurements of n signals, j=1,…,n. On the basis of historical measurements performed in normal conditions, the reconstruction model produces in output a vector $\vec{\hat{x}}^{test}(t) = \left[\hat{x}^{test}(t,1), \hat{x}^{test}(t,j), ..., \hat{x}^{test}(t,n) \right]$ containing the values of the input signals expected in case of normal conditions at the present time t. For the sake of simplicity, the signal index j will be omitted from the notations $x^{test}(t,j)$ and $\hat{x}^{test}(t,j)$, and will be used only when strictly required.

3.1. Reconstruction of operational transients

In Baraldi et al. (2012), different approaches to the problem of signal reconstruction during operational transients have been compared. The obtained results have shown that in order to reduce the computational efforts and to increase model reconstruction accuracy, it is useful to develop a final reconstruction model made by several reconstruction models, each one dedicated to a different operational zone of the component. To this aim, the training patterns are split into different sets, according to the different operational zones. Then, for each operational zone, a dedicated AAKR model is built using the corresponding training set. Once the reconstruction model has been built, it can be used on line for the signal reconstruction task by sending the test pattern, $\vec{x}^{test}(t)$, to the corresponding reconstruction model (Figure 2). In this case, looking at the signal value it is possible to select the corresponding AAKR model. However, for more complex case studies, where discontinuity of the reconstructed variable should be avoided when the model change, one can rely on other algorithms like Takagi-Sugeno concept and Bayes approaches for AAKR model averaging.

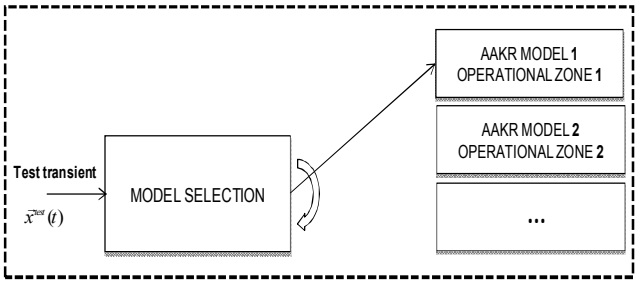


Figure 2. Scheme of AAKR model selection.

It is worth mentioning that abrupt signal changes that might be induced by AAKR model switching have been accommodated in our approach because different models have different thresholds on detection and triggering the alarm.

3.2. Uncertainty quantification using PIs

The uncertainty on the signal reconstruction provided by an empirical model can be quantified by using PIs. With respect to a component in normal conditions, a PI with confidence level 1-σ is defined as an interval, $\left[\hat{x}^{lower}(t), \hat{x}^{upper}(t) \right]$, such that the probability that the measurement of signal j at time t, $x^{test}(t)$, falls within the interval is equal to 1-σ (Eq. (1)) (Office of Nuclear Regulatory Research, 2007; Rasmussen, Wesley Hines & Gribok, 2003). In other words, assuming that the component is in normal conditions:

$$p\left(x^{test}(t) \in \left[\hat{x}^{lower}(t), \hat{x}^{upper}(t) \right] \right) = 1 - \sigma \qquad (1)$$

In order to assess the correctness and effectiveness of the estimated prediction intervals, two indicators are usually considered: the coverage, i.e., the fraction of patterns in a validation set which actually fall within the prediction interval and the prediction interval width. Desiderata are that a PI with confidence 1-σ has coverage of at least 1-σ and width is as small as possible.

Satisfactory PI estimates of time series data have been obtained by using nonlinear regression techniques such as Artificial Neural Networks (ANN), Neural Network Partial Least Squares (NNPLS), Kernel Regression (KR) and Evolving Clustering Method (ECM)) (Rasmussen et al., 2003; Zhao et al., 2011; Ak, Li, Vitelli & Zio, 2013; Zhao, Tao, Ding & Zio, 2013). In applications developed for the nuclear industry, PIs associated to normal component operations have been calculated, using Eq. (2) (Rasmussen et al., 2003; Office of Nuclear Regulatory Research, 2007):

$$\hat{x}^{upper,lower}(t) = \hat{x}^{test}(t) \pm t_N^{\sigma/2} \sqrt{A + B} \qquad (2)$$

$$A = var\left(\hat{x}^{val}(t_{m=1,...,N_{val}}) \right)$$

$$B = \sum_{m=1}^{N_{val}} \left(\hat{x}^{val}(t_m) - x^{val}(t_m) \right)^2 / N_{val}$$

where, $x^{val}(t_m)$ is the value of signal j measured at time t_m after the beginning of the transient of a validation set, $\hat{x}^{val}(t_m)$ is the signal reconstruction value of signal j provided by the empirical model at time t_m of a validation set, N_{val} is the number of patterns in a validation set

containing time series measurements performed in normal conditions, N is the number of training patterns used to train the empirical model, 1-σ is the confidence level ($0 \leq \sigma \leq 1$), $t_N^{\sigma/2}$ is the t-distribution value for a given σ and number of training patterns.

It is important to mention that the patterns in the validation set are different from those in the training set, the former being used to optimize the kernel bandwidth parameter (see Appendix A.3 for more details) and to calculate the PIs, and the latter to train the reconstruction model, and that the quantity $\sqrt{A+B}$ is typically referred to as prediction error. In this work, it is denoted as ε.

In this work, a confidence level, 1-σ equals to 95% is considered. It is worth mentioning that this latter value has been chosen as per the Nuclear Regulatory Commission guidelines that require using the 95th percentile largest uncertainty estimate (Office of Nuclear Regulatory Research, 2007; Denning, Aldemir & Nakayama, 2012). However, setting up the confidence level depends upon the industrial application. In that case, the value of $t_N^{\sigma/2}$ for N>30 is close to 2. Notice that from the point of view of the FD, the higher is the confidence level, 1-σ, the larger is the obtained prediction interval and the lower is the expected false alarm rate (γ). On the other side, the larger is the prediction interval, the higher is the expected missing alarm rate (β) and the longer is the detection delay time.

A drawback of performing PIs quantification using Eq. (2) is that the prediction interval width is independent from the test patterns, $\vec{x}^{test}(t)$. This is not satisfactory since the empirical model performance may vary in different zones of the training space, according to the density and information content of the training patterns available to build the model. Thus, prediction interval widths are expected to be different for different patterns $\vec{x}^{test}(t)$, with smaller PI width when the test pattern is in a zone characterized by a high density of training patterns.

Furthermore, when the AAKR is applied to the reconstruction of operational transients, Eq. (2) typically leads to very large PIs for all measurements. This is due to the term $\mathrm{var}\left(\hat{x}^{val}(t_{m=1,\ldots,N_{val}})\right)$ which, even in the case of reconstructions very close to the signal measurements, can be large due to the variability of the patterns in the validation set.

To overcome these limitations, in the present work we propose to:

1. reduce the variability of the patterns in the validation set by considering, for the computation of the PI at time t_k, k=1,...,N_p, only the reconstructions in the validation set performed at time t_k after the beginning of the transient, with N_p equals to the number of patterns in each test, validation and training transients. Thus, instead of considering, as in Eq. (2), the variance of all the N_{val} reconstructions of the validation set, the variance is computed by considering the $NV<N_{val}$ reconstructions referring to patterns measured only at time t_k.
2. replace $t_N^{\sigma/2}$ with a scaling parameter called scale factor (α) which is used to rescale the prediction error ε, so that, at each time t_k it yields a PI with a specified coverage and with an acceptable width (Bouckaert, Frank, Holmes & Fletcher, 2011). The proper number NV of measurements to estimate the PIs with a given coverage 1-σ is selected relying on Order Statistics (OS), according to Secchi, Zio and Di Maio (2008). In this regard, using the 95% confidence level; the number NV of measurements used to estimate the PIs at each time t_k is estimated and is equal to 59.

In practice, at time t_k after the beginning of the transient, for a reconstructed signal j, $\hat{x}^{test}(t_k)$, Eq. (2) becomes (for large values of NV):

$$\hat{x}^{upper,lower}(t) = \hat{x}^{test}(t) \pm \alpha\sqrt{C+D} \qquad (3)$$

$$C = var\left(\hat{x}^{val}_{i=1,\ldots,NV}(t_k)\right)$$

$$D = \sum_{i=1}^{NV}\left(\hat{x}_i^{val}(t_k) - x_i^{val}(t_k)\right)^2 / NV$$

The method goes along the following steps. It entails an offline procedure for quantifying the scale factor α, and an online procedure for FD.

Step 1: *Offline signal reconstruction.* Using N training data, the AAKR-built model provides the reconstruction $\hat{x}_i^{val}(t_k)$ of signal j in the i-th validation transient of length N_p, i=1,…,NV, *(i.e., $N=N_p*NT$,* where NT is the number of training transients each of length N_p). These historical measurements are collected into the matrix $\overline{\overline{\mathrm{X}}}$ whose generic element $x(t_k,j)$ is the measured value of signal j at time t_k, k=1,…,N_p.

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 $e_i(t_k) = \left|\hat{x}_i^{val}(t_k) - x_i^{val}(t_k)\right|$ of the i-th validation transient, i=1,…, NV.

Step 3: *Prediction error calculations*. At each time k, the prediction error of signal j is calculated as

$\varepsilon(t_k) = \sqrt{\text{var}_k\left(\hat{x}^{val}(t_k)\right) + bias_k\left(\hat{x}^{val}(t_k)\right)}$ by calculating the variance $\text{var}_k\left(\hat{x}^{val}(t_k)\right)$ (Eq. (4)) and the bias $bias_k\left(\hat{x}^{val}(t_k)\right)$ (Eq. (5)) of the *NV* reconstructions of signal *j* (for large values of *NV)*:

$$\text{var}_k\left(\hat{x}^{val}(t_k)\right) = \frac{\sum_{i=1}^{NV}\left(\hat{x}_i^{val}(t_k) - \sum_{i=1}^{NV}\left(\hat{x}_i^{val}(t_k)\right)/NV\right)^2}{NV} \quad (4)$$

$$bias_k\left(\hat{x}^{val}(t_k)\right) = \frac{\sum_{i=1}^{NV}\left(\hat{x}_i^{val}(t_k) - x_i^{val}(t_k)\right)^2}{NV} \quad (5)$$

Step 4: *Scale factor calculations*. At each time *k*, α is calculated as the 95th percentile of the *NV* $\alpha_i(t_k)$, i=1,…, *NV* where $\alpha_i(t_k)=e_i(t_k)/\varepsilon(t_k)$. The coverage capability depends on the number of the *NV* validation transients used. 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) α deals with the uncertainty caused by the AAKR-built model. In practice, at each time *k*, if the AAKR reconstructions are inaccurate, then, the α values are large (i.e., $e_i(t_k) = \left|\hat{x}_i^{val}(t_k) - x_i^{val}(t_k)\right|, i = 1,...,NV$ is large) in order to achieve the desired coverage level (1-σ), and vice versa.

In order to guarantee a certain coverage 1-σ (i.e., (1-σ) of the measurements $x^{test}(t_k)$ of signal *j* in normal conditions are within the PI at each time *k*), we need to find a scale factor such that (1-σ) of the $\alpha_i(t_k)$ are lower and the remainder higher than α. This value is denoted as $\alpha^S(t_k)$ where S stands for "Sorted" and is found by sorting the *NV* available $\alpha_i(t_k)$ (Bouckaert et al., 2011), where *NV* is properly defined by OS (Wald, 1947; Secchi et al., 2008). For σ = 0.05; the correct scale factor may be denoted as $\alpha^{95\,percentile}(t_k)$.

Finally, within the online FD, for any test measurement $x^{test}(t_k)$ of a given signal *j* at each time *k*, Eq. (3) can be re-written as:

$$\hat{x}^{upper,lower}(t) = \hat{x}^{test}(t) \pm \alpha^{95\,percentile}(t)\varepsilon(t) \quad (6)$$

4. CASE STUDY

In this work, an artificial case study has been designed to generate transients representative of the start-up behavior of a component (Baraldi, Di Maio & Zio, 2013). Each transient, $f_i(x(t, 1),...,x(t, 4))$, is four-dimensional (i.e., n= 4 signals) and has a time horizon of N_p=101 time steps, in arbitrary units of measurements.

With respect to normal conditions, 5500 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,…,101, i=1,…,5500 given by Eq. (7):

$$x_i^{nc}(t_k) = 2a\left(1 + erf\left\{\frac{t_k - \mu}{\sqrt{2}}\right\}\right) + 10^{-3\zeta} \quad (7)$$

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, μ from uniform distribution functions with lower and upper bounds reported in Table 1.

Figure 3 shows the obtained evolutions of the four signals in the 5500 transients, $\vec{x}_{i=1:5500}^{nc}(t_k)$.

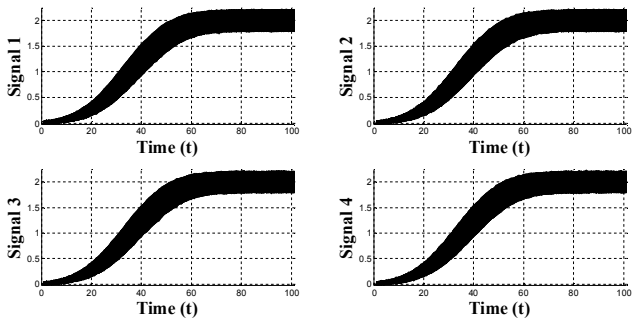


Figure 3. Simulated time evolution in normal conditions of the 4 signals in 5500 start-up transients.

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*, and for calculating the scale factors $\alpha_i(t_k)$. The remaining transients are used to verify the performance of the proposed method.

Furthermore, 50 additional abnormal conditions transients (Eq. (8)) have been simulated in order to reproduce the signal behaviours in abnormal conditions (Figure 4) by assuming a different time evolution for one signal randomly chosen among the four available. It is worth mentioning that this situation, characterized by assuming only one signal in abnormal conditions to create the abnormal transients has been chosen due to the criticality of the FD task under this assumption, i.e., this situation is considered the most challenging case.

$$x_i^{ac}(t_k) = a_\ell t_k + 10^{-3\zeta} \quad (8)$$

where a_ℓ is a random parameter whose values are sampled from a uniform distribution with lower and upper limits reported in Table 1.

Table 1. Limits of the uniform distributions from which the parameters in Eq. (7) and Eq. (8) have been sampled.

Parameter	Lower bounds	Upper bounds
a	0.45	0.55
μ	2.2	2.7
a_ℓ	0.3	0.4

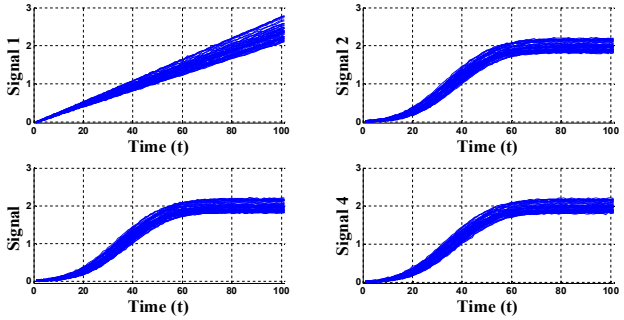


Figure 4. Simulated time evolution in abnormal/normal conditions of signal 1 and the other three signals, respectively, in 50 start-up transients.

4.1. Reconstruction model

The final model for the reconstruction of signals during start-up transients is made by $R = 5$ AAKR-built reconstruction models, each one dedicated to a different operational zone. 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.

Table 2. Definition of the five operational zones and their optimal h values for the four signals.

Zone #	Time period	Operative conditions	h values
1	1-20	Slow start up	0.05
2	21-40	Fast start up	0.05
3	41-60	Start converging to a steady state	0.01
4	61-80	Almost steadiness	0.009
5	81-101	Steady state (nominal value)	0.005

The AAKR models have been trained and their parameters optimized as described in Appendix A.3. In particular, the parameter h values have been identified by optimizing the accuracy of the signal reconstructions in normal conditions and their robustness in abnormal conditions. The obtained optimal values of parameter h in the different operational zones are reported in Table 2.

5. VERIFICATION OF THE PROPOSED METHOD FOR UNCERTAINTY QUANTIFICATION

In this Section, the results obtained by applying the method for PI estimation to the case study of Section 4 are presented. In Subsection 5.1 the PIs obtained by applying a traditional approach for PI estimation, based on a single AAKR-built reconstruction model and Eq. (2), are compared to those obtained by using the proposed method. Subsection 5.2 presents the results of an extensive test performed in order to understand whether the obtained PIs with confidence level 95% provide satisfactory coverage levels, i.e., the fraction of patterns in a validation set that actually falls within the quantified prediction interval is at least equal to 95%, whereas in Subsection 5.3 the ability of the method to properly represent the four sources of uncertainty affecting the signal reconstructions (namely, measurement errors, intrinsic stochasticity of the physical process, uncertainty on the correct setting of the AAKR parameter, and uncertainty caused by the incompleteness of the training data) is discussed.

5.1. PI estimation

The PIs obtained in the reconstructions of signal 1, $x^{test}(t_k,1)$, of a test transient by considering a single AAKR-built reconstruction model and Eq. (2), are shown in Figure 5. Notice that, as expected, the obtained PI widths are constant and very large. This is due to the fact that, according to Eq. (2), the PI widths are independent from the test patterns, $x^{test}(t_k,\ 1)$, and are computed by considering the variance, $\mathrm{var}\left(\hat{x}^{val}\left(t_{m=1,\ldots,N_{val}},1\right)\right)$, of the reconstructions of patterns taken in different zones of the operational transients, and thus characterized by an high variability of signal values.

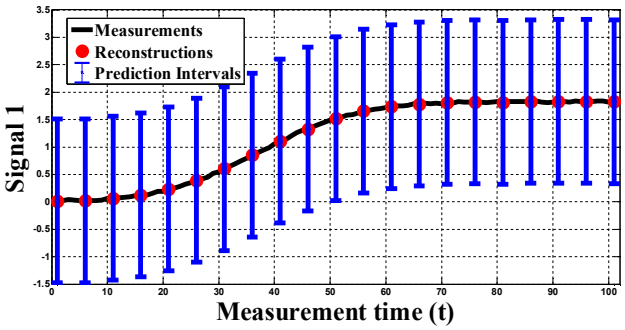


Figure 5. PIs of the reconstruction of 21 patterns obtained using Eq. (2).

Figure 6 shows the results obtained by applying the procedure of Section 3 to a similar transient. Notice that the

PI widths are variable during the time evolution and with a reduced width with respect to those obtained in Figure 5.

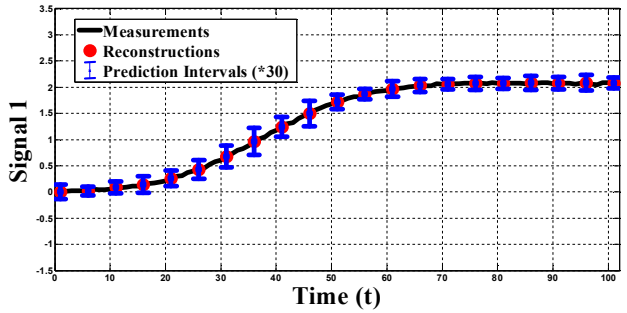


Figure 6. PIs of the reconstruction of 21 patterns obtained using the proposed method.

It is worth noticing that the PI widths of the reconstructions in zone 1 (time from 1 to 20) are smaller than those obtained in zone 3 (time from 41 to 60). This is due to the variability of the training patterns used to train the AAKR-built reconstruction model, which is lower at the beginning of the transient.

5.2. Verification of the prediction interval coverage

In order to verify whether the coverage of the obtained prediction intervals with confidence level 95% is satisfactory, i.e., of at least 95%, we have performed an extensive test using 5000 normal conditions test transients. Figure 7 shows the coverage of the obtained prediction intervals for the first signal, $x^{test}(t_k, 1)$, at different times after the beginning of the transient. The test has been performed using NV value equal to 59. In practice, we have counted how many times the signal measurement falls within the prediction interval at the different times.

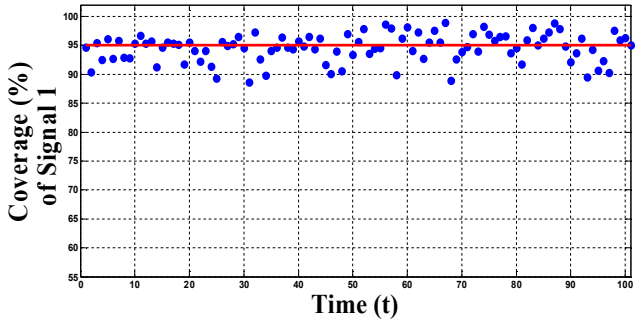


Figure 7. Coverage of the PI with a level of confidence 95% at different times considering 59 validation transients.

Notice that the obtained coverage values are, as expected, close to the confidence level 95%, as it is confirmed by the overall coverage throughout all the transient length which is equal to 94.6%.

To investigate the impact of the number of validation transients to the overall coverage, the same test has been performed with a random number of validation transients, NV=20, lower than 59. As expected, the overall coverage drops down to 88% (Figure 8). This is indeed due to the inadequate use of OS. If the number NV had been taken larger than 59, the overall coverage would be exceed the 95%.

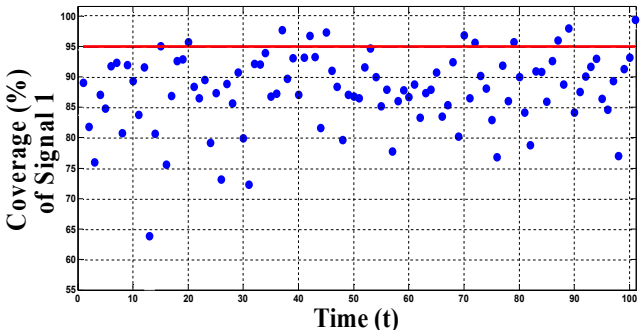


Figure 8. Coverage of the PI with a level of confidence 95% at different times considering 20 validation transients.

5.3. PI capability of quantifying the different uncertainty sources

In this Subsection, without any loss of generality, we focus on the signal reconstruction problem during the first operational zone of the component transient. The evolutions of the NT=300 training transients used to train the AAKR model in zone 1 are shown in Figure 9.

In order to verify the capability of the PI estimates of properly quantifying the effect of different sources of uncertainty, we have performed the following experiments:

1) variation of the measurement error
2) variation of the intrinsic stochasticity of the physical process
3) variation of the AAKR bandwidth parameter value
4) variation of the number of transients used to train the AAKR model.

Experiments 1), 2), and 4) require generating new sets of transients, whereas in experiment 3) different AAKR-built models are generated and trained using the same set of transients illustrated in Section 4.

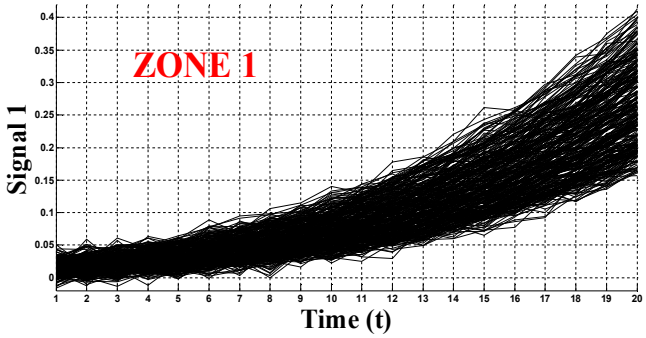


Figure 9. Training transients of signal 1 (zone 1).

5.3.1. Variation of the measurement error

Five different sets of transients characterized by different values of the measurement error have been simulated. In practice, a noise characterized by different standard deviations has been added to the signals generated according to Eq. (7) and Eq. (8). Table 3 reports the five levels of standard deviation considered. The simulated transients have been used to train the AAKR model, to find the prediction intervals according to the proposed method and to compute the overall coverage of the prediction intervals. For each level of noise, we have repeated the AAKR development of the model and the PI estimation five times using different random partitions of the available transients in training, test and validation sets. The same cross-validation procedure is applied also in Subsections 5.2.2, 5.2.3 and 5.2.4. In what follows, we present the average of the five obtained coverage values and their standard deviations.

Table 3. Five levels of standard deviation characterizing the noise in the signals generated by Eq. (7) and Eq. (8).

Noise Levels	Standard Deviations values
1	0.5
2	1
3	1.5
4	2
5	2.5

Figure 10 (top) shows the overall coverage obtained considering the different measurement noise levels. Notice that the obtained coverage values are close to 95% and that the coverage is not influenced by the measurement error. Figure 10 (bottom) shows the average width of the prediction interval. As expected, the higher is the measurement noise, the larger is the prediction interval width. This experiment confirms the ability of the proposed method to properly quantifying the effect of the measurement error on the PI estimate: the method is able to achieve the desired coverage level regardless of the level of the noise, by adjusting the PI width.

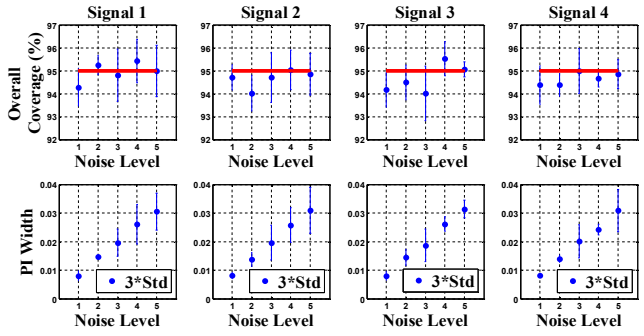


Figure 10. Overall mean coverage (top) and PI width (bottom), considering different measurement noise levels.

5.3.2. Variation of the intrinsic stochasticity of the physical process

In the considered artificial case study, the stochasticity of the physical process is represented by the variation of the parameters α, μ, and a_ℓ in Eq. (7) and Eq. (8), which determines the transients behaviour. In order to simulate different levels of stochasticity in the process, we have sampled the values of these parameters from different probability distributions. Table 4 reports the considered distributions in the four cases: the larger is the range of the uniform distributions, the higher is the stochasticity of the process.

Table 4. Distributions from which the parameters of Eq. (7) and Eq. (8) are sampled, in the considered four cases characterized by different levels of process stochasticity.

Case #	a	μ	a_ℓ
1	U(0.48,0.53)	U(2.33, 2.58)	U(0.33,0.375)
2	U(0.45,0.55)	U(2.2, 2.7)	U(0.3,0.4)
3	U(0.435,0.58)	U(2.08, 2.835)	U(0.28,0.425)
4	U(0.4,0.6)	U(1.95, 2.95)	U(0.25,0.45)

The overall coverage obtained in the four cases is shown in Figure 11 (top): the model achieves satisfactory coverage values regardless the level of stochasticity of the process. As in the previous case, this is obtained by adjusting the PI width (Figure 11 (bottom)): the wider the range of the uniform distributions of the parameters of the equations governing the transients behaviour, i.e., the higher the level of stochasticity in the process, the wider the width of the PIs.

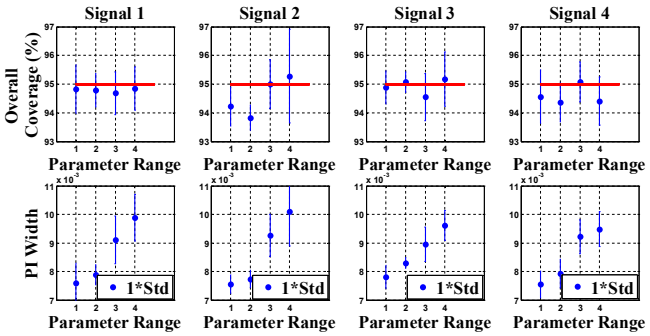


Figure 11. Overall mean coverage (top) and PI width (bottom), considering different cases of process stochasticity.

5.3.3. Variation of the AAKR bandwidth parameter value

In this experiment, the same set of transients illustrated in Section 4 have been used to train eight different AAKR models characterized by different values of the bandwidth parameter, h, (h = 0.005, 0.009, 0.02, 0.05, 0.3, 0.5, 0.9, 1.5).

The overall coverage of the prediction intervals with confidence 95% obtained by the eight different AAKR models is shown in Figure 12 (top). Notice that the obtained coverage values are close to the target of 95%. Figure 12 (bottom) shows that very small and very large values of *h* are characterized by large PI widths. This is due to the fact that the corresponding reconstruction models are characterized by bad performances and, thus, in order to obtain the desired coverage, the prediction interval is enlarged. Furthermore, it is interesting to observe that the PI width is minimum for the value of *h=0.05,* which minimizes the reconstruction error (see Appendix A.3).

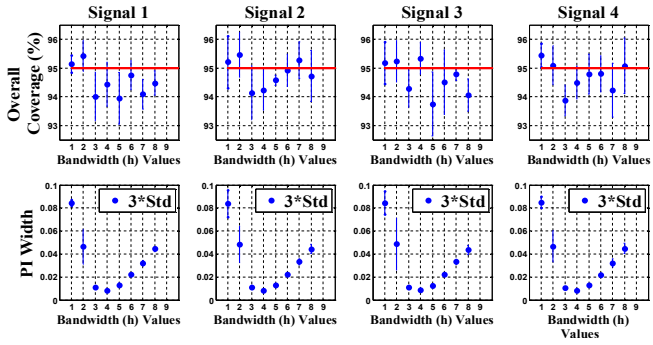


Figure 12. Overall mean coverage (top) and PI width (bottom), considering different AAKR-built models characterized by different values of the bandwidth parameter.

5.3.4. Variation of the number of transients used to train the AAKR model

In order to investigate the effect of the uncertainty caused by the incompleteness of the training data, different AAKR models have been developed using different numbers of training transients. In particular, we have trained three AAKR models based on 100, 300 and 500 training transients, *NT*. In each case, the optimal *h* value has been identified by considering the Mean Squared Error, *MSE* (see Appendix A.3).

The overall coverage obtained in the three cases is shown in Figure 13 (top). As expected, the coverage is close to the target value of 95% and the PI width tends to decrease as the number of training transients increases (Figure 13 (bottom)). This latter effect is due to the fact that model accuracy tends to increase with the number of patterns used to train the empirical model (see Appendix A.1).

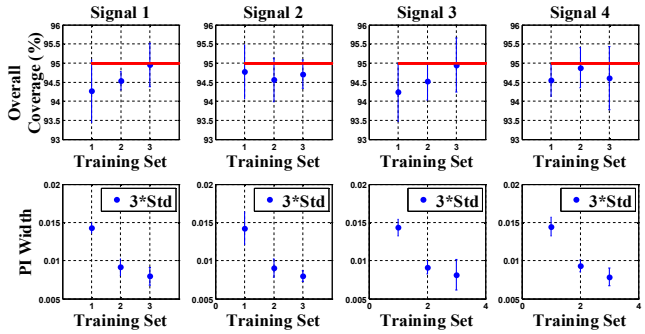


Figure 13. Overall mean coverage (top) and PI width (bottom), considering different number of training transients.

6. CONCLUSIONS

In this work, a novel method to quantify the uncertainty to which signal reconstructions are subject has been developed. Uncertainties are quantified in the form of prediction intervals which have been estimated using Order Statistics (OS) theory. The capability of the methods to deal with measurement errors, intrinsic stochasticity of the physical process, uncertainty on the settings of the model parameters and uncertainty on the signal reconstructions due to incompleteness of the training data has been shown with respect to an artificial case study regarding the monitoring of a component during start-up transients.

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.

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. doi: 10.1016/j.cie.2012.02.002

Ak, R., Li, Y. F., Vitelli, V., & Zio, E. (2013). A Genetic Algorithm and Neural Network Technique for Predicting Wind Power under Uncertainty. *Prognostics and System Health Management Conference PHM-2013* (1-6), September 9-11, Milan, Italy. doi: 10.3303/CET1333155

Aven, T., & Zio, E. (2012). Foundational issues in risk assessment and risk management. *Risk Analysis*, 32(10), 1647-1656. doi: 10.1111/j.1539-6924.2012.01798.x

Baraldi, P., Canesi, R., Zio, E., Seraoui, R., & Chevalier, R. (2010). Signal Grouping for Condition Monitoring of

Nuclear Power Plants Components. *Control and Human-Machine Interface Technologies*.

Baraldi, P., Zio, E., Gola, G., Roverso, D., & Hoffmann, M. (2011). Signal reconstruction by a GA-optimized ensemble of PCA models, *Nuclear Engineering and Design*, 241, 301-309. doi: 10.1016/j.nucengdes.2010.10.012

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. doi: 10.3233/ICA-2011-0375

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. doi: 10.1177/1748006X12463502

Baraldi, P., Di Maio, F., Genini, D., & Zio, E. (2013). A Fuzzy Similarity Based Method for Signal Reconstruction during Plant Transients. *Prognostics and System Health Management Conference PHM-2013 (*889-894*),* September 9-11, Milan, Italy. doi: 10.3303/CET1333149

Baraldi, P., Di Maio, F., & Zio, E. (2013). Unsupervised Clustering for Fault Diagnosis in Nuclear Power Plants Components. *International Journal of Computational Intelligence Systems*, 6(4), 764-777. doi: 10.1080/18756891.2013.804145

Batur, C., Zhou, L., & Chan, C. C. (2002). Support vector machines for fault detection. *In Decision and Control, Proceedings of the 41st IEEE Conference on* (1355-1356), December 10-13, Las Vega, Nevada USA. doi: 10.1109/CDC.2002.1184704

Bouckaert, R. R., Frank, E., Holmes, G., & Fletcher, D. (2011). A Comparison of Methods for Estimating Prediction Intervals in NIR Spectroscopy: Size Matters. *Chemometrics and Intelligent Laboratory Systems*, 109(2), 139-145. doi: 10.1016/j.chemolab.2011.08.008

Campos, J. (2009). Development in the application of ICT in condition monitoring and maintenance. *Computers in Industry*, 60(1), 1-20. doi: 10.1016/j.compind.2008.09.007

Chevalier, R., Provost, D., & Seraoui, R. (2009). Assessment of statistical and Classification model for monitoring EDF's assets. *Control and Human-Machine Interface Technologies*.

Denning, R. S., Aldemir, T., & Nakayama, M. (2012). The use of latin hypercube sampling for the efficient estimation of confidence intervals, in Proceedings of the International Congress on Advances in Nuclear Power Plants (ICAPP '12), Chicago, Ill, USA.

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*, Vol. 62(4), 833 - 845. doi: 10.1109/TR.2013.2285033

Ding, J., Hines, J.W., & Rasmussen, R. (2003). Independent component analysis for redundant sensor validation. *Proceedings of the 2003 Maintenance and Reliability Conference (MARCON 2003), Knoxville, TN*.

Garcıa-Alvarez, D. (2009). Fault detection using principal component analysis (PCA) in a wastewater treatment plant (WWTP). *In Proceedings of the International Student's Scientific Conference.*

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

Hines, J.W., Wrest, D.J., & Uhrig, R.E. (1997). Signal validation using an adaptive neural fuzzy inference system, *Nuclear Technology*, 119, 181–193.

Laouti, N., Sheibat-Othman, N., & Othman, S. (2011). Support vector machines for fault detection in wind turbines. *In Proceedings of IFAC World Congress* (7067-7072), August 28th – September 2nd, Milan, Italy.

Lin, Y., & Stadtherr, M. A. (2008). Fault detection in nonlinear continuous-time systems with uncertain parameters. AIChE journal, 54(9), 2335-2345.

Montes de Oca, S., Puig, V., & Blesa, J., (2012). Robust fault detection based on adaptive threshold generation using interval LPV observers, *International Journal of Adaptive Control and Signal Processing*, Vol. 26(3), 258-283. doi: 10.1002/acs.1263

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

Rahman, S. A. S. A. (2010). Application of Artificial Neural Network in Fault Detection Study of Batch Esterification Process. *International Journal of Engineering and Technology IJET-IJENS*, 10(3).

Ramuhalli, P., Lin, G., Crawford, SL., Konomi, A., Braatz, BG., Coble, JB., Shumaker, BD., Hashemian, HM., (2013). Uncertainty Quantification Techniques for Sensor Calibration Monitoring in Nuclear Power Plants, Pacific Northwest National Laboratory.

Rasmussen, B., Wesley Hines, J., & Gribok, A. V. (2003). An Applied Comparison of the Prediction Intervals of Common Empirical Modeling Strategies. *Proceedings of the 2003 Annual Maintenance and Reliability Conference,* Knoxville, TN.

Safty, S. M. E., Ashour, H. A., Dessouki, H. E., & Sawaf, M. E. (2004). Online fault detection of transmission line using artificial neural network. *IEEE International Conference on Power System Technology* (1629-1632), November 21-24, Singapore. doi: 10.1109/ICPST.2004.1460264

Secchi, P., Zio, E., & Di Maio, F. (2008). Quantifying uncertainties in the estimation of safety parameters by

using bootstrapped artificial neural networks. *Annals of Nuclear Energy*, 35(12), 2338-2350. doi: 10.1016/j.anucene.2008.07.010

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.

Zavaljevski, N., & Gross, K.C. (2000). Support vector machines for nuclear reactor state estimation. *Proceedings of the ANS Topical Meeting on Advances in Reactor Physics and Mathematics and Computation into the Next Millennium*, Pittsburgh, Pennsylvania, USA.

Zhao, W., Baraldi, P., & Zio, E. (2011). Confidence in Signal Reconstruction by the Evolving Clustering Method. *IEEE Prognostics and System Health Management Conference* (1-7), May 24-25, Shenzhen, China. doi: 10.1109/PHM.2011.5939535

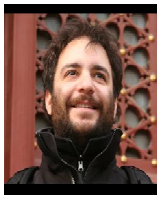
Zhao, W., Tao, T., Ding, Z. S., & Zio, E. (2013). A dynamic particle filter-support vector regression method for reliability prediction. *Reliability Engineering and System Safety,* 119, 109–116.

Zheng, J., & Frey, H. C. (2005). Quantitative analysis of variability and uncertainty with known measurement error: methodology and case study. *Risk Analysis*, 25(3), 663-675. doi: 10.1111/j.1539-6924.2005.00620.x

BIOGRAPHIES

Sameer Al-Dahidi (B.Sc. in Electrical Engineering, The Hashemite University, 2008; M.Sc. in Nuclear Energy - Operations and Maintenance Specialty - Ecole Centrale Paris and Université Paris - Sud 11, 2012) is pursuing his Ph.D. in Energetic and Nuclear Science and Technology at Politecnico di Milano (Milano, Italy). He is an Early Stage Researcher (ESR) in 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. His research aims at analyzing the uncertainty in Condition-Based Maintenance (CBM); developing maintenance decision making taking into account uncertainty and Human Factors; and optimizing the maintenance activities in industrial plants and components which can increase the safety and productivity of industrial plants, while reducing the overall operational and maintenance costs. In 2008-2010, he worked as an Electrical & Instruments Engineer at CCIC in Oil & Gas and petrochemical mega projects in Kuwait and UAE. In 2010 and 2011, he did his internships at AREVA NP in France.

Piero Baraldi (BS in nuclear engineering. Politecnico di Milano, 2002; PhD in nuclear engineering, Politecnico di Milano, 2006) is assistant professor of Nuclear Engineering at the department of Energy at the Politecnico di Milano. He is functioning as Technical Committee Co-chair of the European Safety and Reliability Conference, ESREL2014, and he has been the Technical Programme Chair of the 2013 Prognostics and System Health Management Conference (PHM-2013). He is serving as editorial board member of the international scientific journals: "Journal of Risk and Reliability" and "International Journal on Performability Engineering". His main research efforts are currently devoted to the development of methods and techniques (neural networks, fuzzy and neuro-fuzzy logic systems, ensemble system, kernel regression methods, clustering techniques, genetic algorithms) for system health monitoring, fault diagnosis, prognosis and maintenance optimization. He is also interested in methodologies for rationally handling the uncertainty and ambiguity in the information. He is co-author of 56 papers on international journals, 55 on proceedings of international conferences and 2 books. He serves as referee of 4 international journals.



Francesco Di Maio (B.Sc. in Energetic Engineering, 2004; M.Sc. in Nuclear Engineering, 2006; Double EU-China PhD in Nuclear Engineering, 2010) is Assistant Professor in Nuclear Power Plants at Politecnico di Milano (Milano, Italy). His research aims at developing efficient computational methods and soft computing techniques (Neural Networks, Fuzzy Logic, Genetic Algorithms) for improving a number of open issues relevant for dynamic reliability analysis, system monitoring, fault diagnosis and prognosis, and safety and risk analysis of nuclear power plants. In 2009-2010 he has been Research Fellow of the Science and Technology Programme (STFP) in China, financed by the European Commission, and spent 24 months of practical research at Tsinghua University (Beijing, China). In 2010, he has been appointed as Senior Researcher in City University of Hong Kong. He has published more than 30 articles in peer-reviewed international journals. He is Chair of the Italian IEEE Reliability Chapter.

Enrico Zio (High School Graduation Diploma in Italy (1985) and USA (1984); Nuclear Engineer Politecnico di Milano (1991); MSc in mechanical engineering, University of California, Los Angeles, UCLA (1995); PhD in nuclear engineering,

Politecnico di Milano (1995); PhD in Probabilistic Risk Assessment, Massachusetts Institute of Technology, MIT (1998); Full professor, Politecnico di Milano (2005-); Director of the GraduateSchool, Politecnico di Milano (2007-2011); Director of the Chair on Complex Systems and the Energy Challenge at EcoleCentrale Paris and Supelec, Fondation Europeenne pour l'Energie Nouvelle – EdF (2010-present); Chairman of the European Safety and Reliability Association-ESRA (2010- present); Rector's Delegate for the Alumni Association, Politecnico di Milano (2011- present); President of the Alumni Association, Politecnico di Milano (2012- present); President of Advanced Reliability, Availability and Maintainability of Industries and Services (ARAMIS) srl (2012- present); Member of the Scientific Committee on Accidental Risks of INERIS, Institut national de l'environnement industriel et des risques, France; Member of the Academic Committee of the European Reference Network for Critical Infrastructure Protection, ERNCIP, 2013). His research topics are: analysis of the reliability, safety and security, vulnerability and resilience of complex systems under stationary and dynamic conditions, particularly by Monte Carlo simulation methods; development of soft computing techniques (neural networks, support vector machines, fuzzy and neuro-fuzzy logic systems, genetic algorithms, differential evolution) for safety, reliability and maintenance applications, system monitoring, fault diagnosis and prognosis, and optimal design and maintenance. He is co-author of seven books and more than 250 papers on international journals, Chairman and Co-Chairman of several international Conferences and referee of more than 20 international journals.

APPENDIX

Appendix A.1 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, $k=1,...,N$ 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. 2012; Baraldi, Canesi, Zio, Seraoui & Chevalier, 2011; Di Maio, Baraldi, Zio & Seraoui, 2013).

AAKR technique requires three different sets of data:

1. **Historical data (often called training data)** which are historical measurements of the signals taken during normal conditions of the component used to train/develop the model for accurate reconstructions.
2. **Validation data** which are historical measurements of the signals taken during normal/abnormal conditions of the component used to optimize the model parameters, such as the kernel bandwidth h, as we shall show in the following.
3. **Test data** which are the measurements taken at current time t to perform a real-time health assessment of the component.

In Figure 14, a sketch of the procedure for predicting one test measurement at time t: $\vec{x}^{test}(t) = [x^{test}(t, 1), x^{test}(t, 2)]$ is provided. Historical data which fall within the bandwidth h have a large impact on the reconstructed values $\hat{x}(t)$.

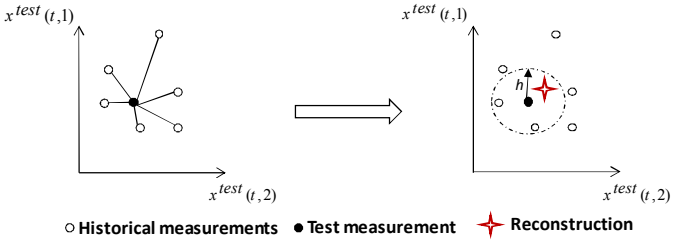


Figure 14. AAKR basic principle.

In more details (Baraldi et al. 2011), the j-th component at time t of $\hat{x}^{test}(t, j)$ is given by Eq. (9):

$$\hat{x}^{test}(t, j) = \frac{\sum_{k=1}^{N} w(t_k) \cdot x(t_k, j)}{\sum_{k=1}^{N} w(t_k)} \quad (9)$$

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. (10):

$$d^2(t_k) = \sum_{k=1}^{N} \left(x^{test}(t, j) - x(t_k, j) \right)^2 \quad (10)$$

and inserting it in the Gaussian kernel Eq. (11):

$$w(t_k) = \frac{1}{\sqrt{2\pi}h} e^{\frac{-d^2(t_k)}{2h^2}} \quad (11)$$

where h is the Gaussian kernel bandwidth.

In order to provide in Eq. (10) 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 Eq. (12):

$$x^{test-normalized}(t,j)=\frac{x^{test}(t,j)-\mu(j)}{\sigma(j)} \quad (12)$$

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 $\overline{\overline{X}}$:

$$\mu(j)=\frac{\sum_{k=1}^{N}x(t_k,j)}{N},\qquad \sigma(j)=\sqrt{\frac{\sum_{k=1}^{N}\left(x(t_k,j)-\mu(j)\right)^2}{N}} \quad (13)$$

A.2 Performance Metrics

In order to evaluate the performance of AAKR model, the following criteria should be considered (Baraldi et al. 2011):

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 (γ). The accuracy metric is typically defined as the Mean Squared Error (*MSE*) between the model reconstructions and the signal measured values.

Let $\overline{\overline{X}}_{nc}^{test}$ be a matrix of measured data whose generic element $x_{nc}^{test}(t_k,j)$ represents the k-th time measurement, k=1,...,N_p, of the j-th measured signal, j=1,...,n, taken during normal conditions, and $\hat{x}^{test}(t_k,j)$ its reconstruction in nc; then, the *MSE* with respect to signal j is given by Eq. (14):

$$MSE_j=\frac{\sum_{k=1}^{N_p}\left(x_{nc}^{test}(t_k,j)-\hat{x}^{test}(t_k,j)\right)^2}{N_p} \quad (14)$$

A global accuracy measure that takes into account all the monitored signals and test patterns is defined by Eq. (15):

$$MSE=\frac{\sum_{j=1}^{n}\frac{\sum_{k=1}^{N_p}\left(x_{nc}^{test}(t_k,j)-\hat{x}^{test}(t_k,j)\right)^2}{N_p}}{n}=\frac{\sum_{j=1}^{n}MSE_j}{n} \quad (15)$$

Notice that, although the metric is named accuracy, it is actually a measure of error and, thus, a low value is desired.

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 (β). The robustness metric is here defined as the *MSE* between the model reconstructions and the mean of the historical data $\overline{\overline{X}}$.

Let $\overline{\overline{X}}_{ac}^{test}$ be a matrix of measured data whose generic element $x_{ac}^{test}(t_k,j)$ represents the k-th time measurement, t_k, k=1,...,N_p, of the j-th measured signal, j=1,...,n, taken during abnormal conditions, and $\hat{x}^{test}(t_k,j)$ its reconstruction in nc and let $\overline{\overline{X}}^{mean}$ be a mean matrix of the NT training transients, with length N_p, computed at each time t_k, k=1,...,N_p whose generic element $x^{mean}(t_k,j)$ represents the mean of the k-th time observations performed at t_k, k=1,...,N_p, of the j-th measured signal, j=1,...,n, taken during normal conditions; then, the robustness *MSE* with respect to signal j is given by Eq. (16):

$$MSE_{ac,j}=\frac{\sum_{k=1}^{N_p}\left(\hat{x}^{test}(t_k,j)-x^{mean}(t_k,j)\right)^2}{N_p} \quad (16)$$

A.3 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. That is, the optimum bandwidth h value that minimizes the product (Eq. (17)) between the global model accuracy, *MSE*, and the global model robustness, MSE_{ac}:

$$Objective\ Function = MSE \times MSE_{ac} \quad (17)$$

Without loss of generality, the optimization of the AAKR model parameter, i.e., the kernel bandwidth h, is hereafter presented with respect to only the operational zone "1". A cross-validation approach can serve the scope of optimizing the objective function; for the sake of saving computational time, in this work a large set of data of the training and validation transients have been used, i.e., we have used NT=300 transients to train the AAKR-built model and NV=59 transients as validation set to optimize the value of the model parameter, h. Figure 15 shows the objective function (Eq. (17)) obtained when 11 potential settings of h (0.005, 0.007, 0.009, 0.01, 0.05, 0.09, 0.10, 0.15, 0.20, 0.25, 0.30) are used. It is worth noticing that the optimal bandwidth value for the first operational zone is close to 0.05. The optimum h values of the remaining four operational zones are estimated using the same procedure. The obtained optimal values of parameter h of the five operational zones are reported in Table 2.

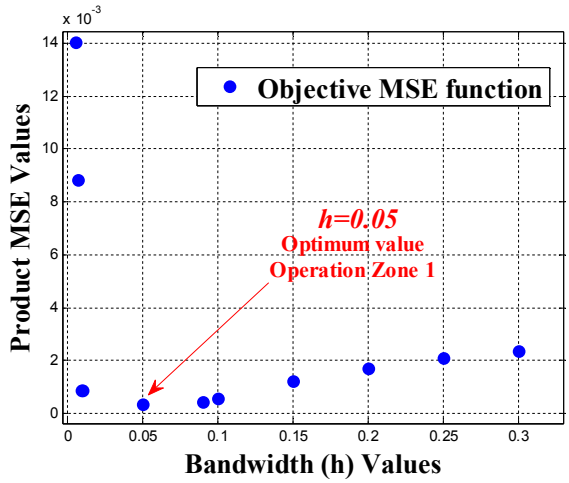


Figure 15. Reconstruction error (objective *MSE* function) versus kernel's bandwidth (*h)* values.

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 Figure 16 it can be seen that with a small bandwidth (*h* = 0.2) large weights (similarities) are assigned to historical data whose distance is very close to zero, whereas with a larger bandwidth (*h* = 1.5), the weight assignment is less specific (Office of Nuclear Regulatory Research, 2007).

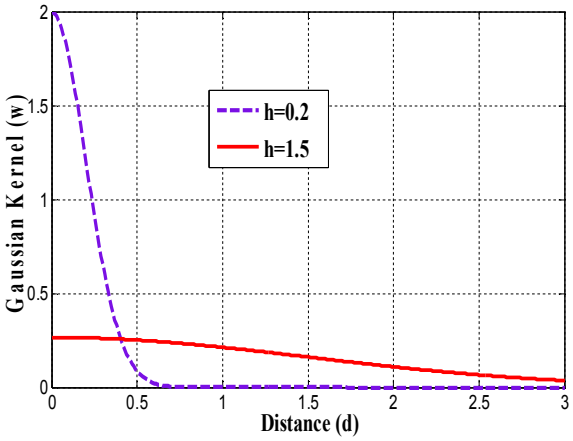


Figure 16. Gaussian Kernel Function with two *h* values.