Fault Detection in Nuclear Power Plants Components by a Combination of Statistical Methods
Francesco Di Maio, Piero Baraldi, Enrico Zio, Rédouane Seraoui

To cite this version:
Francesco Di Maio, Piero Baraldi, Enrico Zio, Rédouane Seraoui. Fault Detection in Nuclear Power Plants Components by a Combination of Statistical Methods. IEEE Transactions on Reliability, Institute of Electrical and Electronics Engineers, 2013, 62 (4), pp.833 - 845. <10.1109/TR.2013.2285033>. <hal-00934649>

HAL Id: hal-00934649
https://hal-supelec.archives-ouvertes.fr/hal-00934649
Submitted on 22 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
FAULT DETECTION IN NUCLEAR POWER PLANTS COMPONENTS BY A COMBINATION OF STATISTICAL METHODS

Francesco Di Maio*a, member IEEE, Piero Baraldi*a, Enrico Zio*a,b, senior member IEEE, Redouane Seraoui*c

*aEnergy Department, Politecnico di Milano, via Ponzio 34/3, Milano, Italy
bChair on System Science and the Energetic Challenge, European Foundation for New Energy – Electricite de France, Ecole Centrale, Paris and Supelec, Paris, France
cEdF R&DSTEP Simulation et Traitement de l’information pour l’exploitation des systèmes de production, Chatou Cedex, France

Abstract – In this paper, we investigate the feasibility of a strategy of fault detection capable of controlling misclassification probabilities, i.e., balancing false and missed alarms. The novelty of the proposed strategy consists of i) a signal grouping technique and signal reconstruction modeling technique (one model for each subgroup), and ii) a statistical method for defining the fault alarm level. We consider a real case study concerning 46 signals of the Reactor Coolant Pump (RCP) of a typical Pressurized Water Reactor (PWR). In the application, the reconstructions are provided by a set of Auto-Associative Kernel Regression (AAKR) models, whose input signals have been selected by a hybrid approach based on Correlation Analysis (CA) and Genetic Algorithm (GA) for the identification of the groups. Sequential Probability Ratio Test (SPRT) is used to define the alarm level for a given expected classification performance. A practical guideline is provided for optimally setting the SPRT parameters’ values.

Index Terms – Condition monitoring, signal grouping, signal reconstruction, auto-associative kernel regression, sequential probability ratio test, nuclear reactor coolant pump.
Acronyms and Abbreviations

ac  Abnormal conditions
AAKR  Auto-Associative Kernel Regression
ANN  Artificial Neural Network
ASN  Average Sample Number
CA  Correlation Analysis
GA  Genetic Algorithm
ICA  Independent Component Analysis
nc  Normal conditions
NPP  Nuclear Power Plant
PCA  Principal Component Analysis
PWR  Pressurized Water Reactor
RCP  Reactor Coolant Pump
SPRT  Sequential Probability Ratio Test
SVM  Support Vector Machine

Notations

\(N\) Number of signals
\(t_i\) \(i\)-th measurement time
\(\hat{x}_{s,t}^{\text{obs}}\) Observed measurements of \(N\) signals at time \(t_i\)
\(\hat{x}_t^{\text{nc}}\) Reconstructions of \(\hat{x}_{s,t}^{\text{obs}}\) in normal conditions (nc)
\(r_i\) Residual between \(\hat{x}_{s,t}^{\text{obs}}\) and \(\hat{x}_t^{\text{nc}}\)
\(\alpha\) False alarm rate
\(\beta\) Missed alarm rate
\(X_M\) Number of historical patterns used to perform the hybrid signal grouping
\(X_V\) Number of patterns belonging to the validation set
\(H_0\) Null hypothesis of the SPRT test
\(H_1\) Alternative hypothesis of the SPRT test
\(\mu_0\) Mean of the Gaussian distribution of \(H_0\)
\(\sigma^2\) Variance of Gaussian distributions \(H_0\) and \(H_1\), i.e., noise of \(\hat{x}_{s,t}^{\text{obs}}\)
\(\mu_1\) Mean of the Gaussian distribution of \(H_1\)
\(\ln(A)\) Lower stopping boundary of the SPRT
\(\ln(B)\) Upper stopping boundary of the SPRT
\(L_n\) Likelihood ratio for the positive mean test
\(\text{SPRT}=\ln(L_n)\) Logarithm of the likelihood ratio \(L_n\)
\(R_n\) Sequence of residuals values \(r_1, r_2, \ldots, r_n\)
\(P\left(r_1, r_2, \ldots, r_n \mid H_0\right)\) Probability that the null hypothesis \(H_0\) given \(R_n\)
\(P\left(r_1, r_2, \ldots, r_n \mid H_1\right)\) Probability of the alternative hypothesis \(H_1\) given \(R_n\)
\(E[n; \mu]\) Number of measurements, i.e., time, necessary for the SPRT to choose among \(H_0\) and \(H_1\)
\(d\) Outcome of Durbin–Watson test
\(x_{s,a}^{\text{obs}}\) Signal measurements of signal 4a
\(x_{s,a}^{\text{nc}}\) Signal reconstructions of signal 4a
\(Q(\mu_i)\) Operating characteristic function, i.e., the probability that \(H_0\) is accepted for different values of \(\mu_i\)
\(P(\mu_i)\) Power function, i.e., the probability that \(H_1\) is accepted, for different values of \(\mu_i\)
1. Introduction

We present a research work aimed at developing a fault detection strategy with the capability of controlling fault misclassification probabilities, i.e., balancing false and missed alarms, for improving Nuclear Power Plants’ (NPPs’) availability and safety. The novelty of the fault detection strategy lies in the combination of a signal grouping technique, a signal reconstruction modeling technique, and a statistical method for determining the signal deviation level for fault alarms.

To set the problem, Fig. 1 shows a typical scheme of condition monitoring of a component for fault detection. Sensor measurements (signals) \( \hat{x}^{\text{obs}} \) are sent in input to a model which reconstructs in output the same signals \( \hat{x}^{\text{nc}} \) as if the component behaviour were in normal conditions (nc). Deviations between the actually measured (observed) signals \( \hat{x}^{\text{obs}} \) and the reconstructed ones \( \hat{x}^{\text{nc}} \) reveal the presence of faults [1]. In simple words, \( \hat{x}^{\text{obs}} = \hat{x}^{\text{nc}} \) under normal conditions, whereas \( \hat{x}^{\text{obs}} \neq \hat{x}^{\text{nc}} \) under abnormal conditions (ac).

Different empirical models have been developed for signal reconstruction. Typical examples include Artificial Neural Networks (ANNs) and Recurrent ANNs [2]-[9], Principal Component Analysis (PCA) and Independent Component Analysis (ICA) [10]-[12], Multivariate State Estimation Technique (MSET) [13]-[14], and Support Vector Machines (SVMs) [15]-[16]. The model considered in this work for reconstructing the component behavior in normal conditions is based on the Auto-Associative Kernel Regression (AAKR) method [17]-[20].

In practice, when a large number of signals is available, a phenomenon called fault propagation effect may occur when reconstructing the component behavior in normal conditions; if one or more signals are degraded by the fault, they may affect the values of other signals, and therefore these latter signals may be incorrectly reconstructed, resulting in a large number of false alarms [21].

To alleviate this problem, one can resort to grouping signals into subgroups, and developing a reconstruction model for each subgroup [21]. Two different types of grouping strategies have been proposed in literature: with overlapping, i.e., the same signal can belong to more than one group [12], [22]-[25], and without overlapping [26]-[28]. In practical applications, the latter strategy tends to be preferred because it allows for a smaller number of models to be developed, at a lower computational effort [28]. Two different approaches to grouping can be implemented: filter, and wrapper. The former bases the grouping on characteristics a
priori judged to be favorable for condition monitoring, e.g., physical and functional homogeneity (i.e.,
groups are made only by temperature signals, or only by pressure signals), irrespective of the signal
reconstruction modeling technique used [24], [27]. The latter uses a search algorithm, e.g., Genetic
Algorithms (GAs) [29], Differential Evolution [30], etc., as a wrapper around the signal reconstruction
model [31]. In this paper, we use a hybrid approach, i.e., filter and wrapper, based on non-overlapping
grouping given by Correlation Analysis (CA) and GA [29] (Fig. 2, bottom left). The motivation of the choice
is that the GA-based wrapper approach allows finding better performing groups for the specific
reconstruction model used, while the contribution of CA leverages the computational burden of the GA (for
further details, the interested reader should refer to [29]).

Fig. 2. The condition monitoring scheme for fault detection adopted in this work.

To balance false and missed alarms in the fault detection, it is possible to perform an analysis of the residuals
\[ r = x_{\text{obs}} - x_{\text{pre}} \] to declare whether a component fault has occurred (Fig. 2, top right). Conventional fault
detection systems are based on deterministic rules, i.e., simple tests that compare observed signal values to given thresholds. This way of proceeding may suffer from either large false alarm rates \( \alpha \) (if thresholds are set too small), or high missed (or delayed) alarm rates \( \beta \) (if the thresholds are set too large), and they can fail dramatically, especially in situations where noisy data are present, or only slight drifts are observed prior to catastrophic faults [32]. On the other side, stochastic approaches regard residuals as random variables described by a probability law. Abnormal conditions are assumed to result in a modification of the residual probability law (for example, transition from a normal distribution with a given mean and variance to another normal distribution with different mean and variance). In the present work, the identification of the modification of the probability law describing the residuals is based on the Sequential Probability Ratio Test (SPRT) [33], [17], a statistical method for triggering an alarm with a controlled balance between false and missed alarms.

This framework of analysis is applied to a real case study concerning \( N=46 \) signals used to monitor the Reactor Coolant Pumps (RCPs) of a typical Pressurized Water Reactor (PWR); the monitoring scheme of Fig. 2 is applied in normal and abnormal conditions, and the occurring faults are detected. A practical guideline is provided for setting the SPRT parameter values for balancing false and missed alarms.

The rest of the paper is organized as follows. Section 2 illustrates the case study and the data available. Section 3 presents the hybrid approach for signal grouping based on CA and GA, embedding the AAKR method for signal reconstruction. In Section 4, the definition of alarm levels for detection of abnormal conditions using the SPRT method is given with respect to the case study of Section 2. Finally, Section 5 concludes the paper with some considerations.

2. The case study

The case study considers \( N=46 \) signals used to monitor the RCPs of a typical PWR. The signal values have been measured every hour for a period of 11 consecutive months on four RCPs, each one on a different line of the primary circuit of the NPP. A dataset has been provided, containing patterns (46-dimensional time series) corresponding to normal conditions, and cleared from outliers. The 46-dimensional patterns (5798) have been divided into a set \( X_M \) of 2798 patterns used to perform the hybrid signal grouping (Section 3), and a validation set \( X_V \) of 3000 patterns used to validate the condition monitoring and fault detection scheme (Section 4).

3. The hybrid approach for signal grouping, with AAKR modelling for signal reconstruction

The AAKR model is used to reconstruct the signal values in normal conditions of operation, \( \hat{x}^{\text{nc}} \), from the observed measurements, \( x^{\text{obs}} = (x^{\text{obs}}(1),..,x^{\text{obs}}(N)) \). The reconstruction is done by a weighted sum of the
observations (see the Appendix for further details on the AAKR). The model is applied to the signals of each of the subgroups in which the signals are partitioned by a hybrid wrapper approach based on CA and GA. In [29], it has been shown that this hybrid method leads to reconstructions that are more tolerant to the fault propagation problem (that has been mentioned in the Introduction) than proceeding with the reconstruction based on the single group of all signals [19], [27], or on groups of signals defined by a filter CA approach only [19], [25], or by a wrapper GA-based approach only [28].

The hybrid approach consists of using CA to identify groups of highly correlated signals, and GA to define the group assignment of the remaining signals [29]. Application to the case study of interest leads to CA identifying 5 groups, with one formed by 24 highly correlated signals, and GA searching chromosomes of 22 elements, each one corresponding to one of the 22 signals not assigned to a group by CA. In this way, the dimension of the GA search space is reduced from $5^{46} \approx 10^{32}$ possible group combinations to $5^{22} \approx 10^{15}$. In the end, the hybrid approach identifies 5 groups so formed [29]:

- Group 1 has 30 signals related to the different temperatures measured in different parts of the RCP (e.g., seals, and hydraulic and engine components),
- Group 2 has 4 signals measuring the rotating speed of the engine components of the RCP,
- Group 3 has 3 signals measuring the water mass flow rate inside the RCP,
- Group 4 has 5 signals measuring the water mass flow rate coming from the first seal of the RCP, and
- Group 5 has 4 signals measuring the water mass flow rate flowing to the first seal of the RCP.

4. Fault detection

After grouping of the signals into 5 subgroups, 5 AAKR reconstruction models are developed (one for each subgroup), and a strategy of fault detection is implemented based on SPRT to keep under control the misclassification probabilities $\alpha$, and $\beta$, i.e., the false, and missed alarm probabilities, respectively.

4.1. The SPRT method

Let $r_i$ represent the value of one signal at the generic $i$-th measurement time $t_i$, and let us assume that a sequence of values $\{R_n\} = r_1, r_2, \ldots, r_n$ is available. The SPRT is a binary stochastic hypothesis test which sequentially analyzes the process observations $\{R_n\}$ and decides whether they are consistent with a null hypothesis $H_0$ represented by a given probability law, or rather with an alternative hypothesis $H_1$ represented by a different probability law [32]. The test is driven by user-specified misclassification probabilities $\alpha$ and $\beta$.

In our context of fault detection, $r_i$ is the residual between the signal measurement and its reconstruction at time $t_i$, $r_i = x_{\text{obs}}(t_i) - \hat{x}_{\text{obs}}(t_i)$. The null hypothesis $H_0$ corresponds to the statistical behavior of the residuals when the component is under normal conditions, and the alternative hypothesis $H_1$ corresponds to that case of abnormal conditions.

Without loss of generality, here we illustrate the test considering as the null hypothesis $H_0$ of a Gaussian distribution with mean $\mu_0 = 0$, and variance $\sigma^2$; and as the alternative hypothesis $H_1$ of a Gaussian distribution
with mean $\mu_1 > 0$, and same variance $\sigma^2$. This test is usually referred to as test for positive mean, and can be used for the identification of positive offsets of amplitude $\mu_1$. However, SPRT can also be applied for the identification of negative offsets, with increased or reduced variance [34]. When the assumption of Gaussian distributions might impose potentially misleading behavior on extreme values of the residuals, i.e., tails of the distributions, extreme value statistics can integrate into SPRT for dealing with the difficulty of fault detection [35]. Furthermore, SPRT can be successfully applied to any other distribution, if needed; as an example, in [36] it has been shown that SPRT overtakes other statistical methods when dealing with a Bernoulli distribution of residuals.

The SPRT operates as follows [37]. Every time a new value $r_i$ becomes available, a test index is calculated and compared to two stopping boundaries $\ln A$ (lower threshold), and $\ln B$ (upper upper threshold), with $\ln A < \ln B$, that are related to the misclassification probabilities $\alpha$ and $\beta$ by the expressions

$$\ln A = \ln \left( \frac{B}{1 - \alpha} \right)$$

(1)

$$\ln B = \ln \left( \frac{1 - \beta}{\alpha} \right)$$

(2)

The test index is equal to the natural log of a likelihood ratio ($L_n$), which is the ratio between the probabilities that the alternative hypothesis $H_1$ and the null hypothesis $H_0$ are true:

$$L_n = \frac{\text{probability of observed sequence } \{R_n\} \text{ given } H_1 \text{ true}}{\text{probability of observed sequence } \{R_n\} \text{ given } H_0 \text{ true}}$$

(3)

Three different cases may arise:

- if the logarithm of the likelihood ratio (3) is greater than or equal to the logarithm of the upper threshold (i.e., $\ln(L_n) \geq \ln(B)$), then it can be concluded that the alternative hypothesis $H_1$ is true;
- if the logarithm of the likelihood ratio is less than or equal to the logarithm of the lower threshold limit (i.e., $\ln(L_n) \leq \ln(A)$), then it can be concluded that the null hypothesis $H_0$ is true;
- if the logarithm of the likelihood ratio falls between the two limits (i.e., $\ln(A) < \ln(L_n) < \ln(B)$), then there is not enough information to discriminate (and, incidentally, no other statistical test could reach a decision with the same given probabilities $\alpha$ and $\beta$ [32]).

Given the sequence of signal values $\{R_n\} = r_1, r_2, \ldots, r_n$, the probability that the null hypothesis $H_0$ (i.e., Gaussian distribution with mean 0 and variance $\sigma^2$) is true is given by [37]

$$P \left( r_1, r_2, \ldots, r_n \mid H_0 \right) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} r_i^2}$$

(4)

Similarly, the probability of the alternative hypothesis $H_1$ (i.e., mean $\mu_1 > 0$ and variance $\sigma^2$) is
The ratio of the probabilities in (4) and (5) gives the likelihood ratio $L_n$ for the positive mean test

$$L_n = e^{-\frac{1}{2\sigma^2} \left[ \sum_{i=1}^{n} \mu_i \left( \mu_i - 2r_i \right) \right]}$$

The SPRT index for the positive mean test is finally obtained by taking the logarithm of the foregoing likelihood ratio:

$$SPRT = \ln (L_n) = -\frac{1}{2\sigma^2} \left( \sum_{i=1}^{n} \mu_i \left( \mu_i - 2r_i \right) \right) = \frac{\mu_i}{\sigma^2} \sum_{i=1}^{n} \left( r_i - \frac{\mu_i}{2} \right)$$

See that (1)-(7) are defined according to the hypothesis of a Gaussian distribution of the residuals \{R_n\} = \{r_i, \dots, r_n\}. Obtaining the stopping boundaries and the SPRT index for any other probability distribution $P$ is straightforward [36].

In any case, the application of the SPRT requests the setting of four parameters:

- the residual variance in normal conditions ($\sigma^2$),
- the expected offset amplitude ($\mu_i$),
- the false alarm probability ($\alpha$), and
- the missed alarm probability ($\beta$).

The values of these parameters influence the time necessary to deliver a decision, i.e., the time necessary for the SPRT to choose among $H_0$ and $H_1$. Assuming again for $H_0$ a Gaussian pdf with mean $\mu_0 = 0$ and variance $\sigma^2$, and for $H_1$ a Gaussian pdf with mean $\mu_1 > 0$ and the same variance $\sigma^2$, the expected sample number, hereafter called Average Sample Number (ASN), for a decision when the residuals \{R_n\} are distributed according to a Gaussian distribution with mean value $\mu$ and variance $\sigma^2$, is given by [38]

$$E[n; \mu] \approx \frac{2\sigma^2 \left\{ \frac{A^k - 1}{A^k - B^k} \ln \frac{\beta}{1 - \alpha} + \frac{1 - B^k}{A^k - B^k} \ln \frac{1 - \beta}{\alpha} \right\}}{\mu_i \left( 2\mu - \mu_i \right)}$$

where

$$h = \frac{\mu_i - 2\mu}{\mu_i}$$

Notice that the approximate theoretical ASN

1. is directly proportional to $\sigma^2$, meaning the larger $\sigma^2$, the noisier are the observed values and the more difficult a decision is; and
2. is inversely proportional to \( \mu_i \) meaning it is easier for the SPRT to select a hypothesis when \( \mu_i \), the mean value of \( H_i \), is far away from \( \mu_0 = 0 \), the mean value of the null hypothesis.

With respect to \( \sigma^2 \), if the signal reconstruction model has been applied to past measurements in normal conditions, it can be easily estimated considering the variance of the past values of the residuals. The other three parameters have to be fixed by the analyst according to his or her experience, and needs. In Section 4.3, a practical guideline is provided for this setting.

As a final remark, it is worth noticing that SPRT has been shown to be optimum, in the sense of minimizing ASN under the hypotheses \( H_0 \) and \( H_1 \) for fixed values of false and missing alarms probabilities, \( \alpha \) and \( \beta \), respectively, if the signal values \( \{R_n\} = r_1, r_2, \ldots, r_n \) are i.i.d [34]. In the case of non-i.i.d measurements, as in our case study, no optimality has been established, although the detection performance is still good in terms of false and missed alarm probabilities, as it is shown in [39], and by the results shown in Section 4.2. The procedural steps for the application of the SPRT are given in the pseudo-code of Fig. 3.

---

1. \( 0 \rightarrow i \) Set to 0 the time index \( i \)
2. \( 0 \rightarrow \text{SPRT} \) Set to 0 the SPRT index
3. \( i+1 \rightarrow i \) Increase the time index
4. At time \( t_i \), collect the new residual \( r_i \)
   a. Update the SPRT index (Eq. 7)
   b. Compare the SPRT index to the upper \( \ln(B) \) and lower \( \ln(A) \) decision boundaries
   c. Detect the component condition, depending on the test outcome:
      - if \( \text{SPRT} < \ln(A) \) (lower limit is passed) then
        the component is declared in normal conditions
        Go to 2 and start again the test for the successive times;
      - else if \( \text{SPRT} > \ln(B) \) (upper limit is passed) then
        the component is declared in fault conditions
        an alarm flag is raised
        Go to 2 and start again the test for the successive times;
      - else if \( \ln(A) < \text{SPRT} < \ln(B) \) (neither limit has been passed) then
        no decision is made
        Go to 3

---

Fig. 3. The pseudo-code of the procedural steps of the SPRT.

4.2. SPRT application to the case of RCP of a PWR

We focus on the signal ‘water flowing to the first seal of the pump in line 1’ (hereafter referred to as signal 4a), which belongs to Group 1. The \( X_M \) historical values of the residuals of the signal under normal conditions have been analyzed to set the hypothesis \( H_0 \), which describes the residual distribution under these conditions. The residual distribution has been confirmed to be Gaussian with mean value \( \mu_0 = 0.0094 \), and variance \( \sigma^2 = 0.014 \), by the Kolmogorov–Smirnov test [40]. The very small mean value of the residuals has confirmed that the reconstruction under normal conditions is not biased, and can be approximated by a distribution with \( \mu_0 = 0 \). Residuals have been further tested for \( s \)-independency with the Durbin–Watson test [41]. Small values of the test outcome \( d \) indicate that residuals are, on average, positively correlated; whereas if \( d > 2 \), successive residuals are, on average, negatively correlated. In our case, \( d = 0.84 \), and thus we can conclude that residuals are not \( s \)-independent, and slightly positively correlated.

To test the performance of the proposed SPRT for fault detection, an offset of amplitude equal to 1% of the standard deviation of the signal values has been injected onto the measurements of signal 4a. The drift starts at \( t = 1 \) hour, and continues until \( t = 1000 \) hours. In Fig. 4, the abnormal conditions’ measurements \( s_{4a}^{obs} \) and
their reconstructions \( \hat{x}_{ir}^{rec} \) by the AAKR model are shown. Fig. 5 reports the obtained residuals \( \{ R_n \} = r_1, r_2, \ldots, r_n \), where \( r_i = x_{ir}^{obs} (t_i) - x_{ir}^{rec} (t_i) \), \( i = 1, 2, \ldots, n \). Table I reports the SPRT parameters used.

Table I

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>0.01</td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.01</td>
</tr>
<tr>
<td>( \mu_0 )</td>
<td>0</td>
</tr>
<tr>
<td>( \mu_1 )</td>
<td>0.46</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Fig. 6 illustrates the outcomes of the SPRT in this case study. The SPRT alarm index quickly moves from zero to the upper bound \( \ln(B) \) allowing acceptance of the Hypothesis \( H_1 \), and thus to promptly detect the presence of abnormal conditions. Once the threshold is passed, the SPRT index is reset to zero, and the test continues confirming the alarm in a number of time steps similar to the theoretical ASN (8). In fact, the empirical ASN, i.e. the average time elapsed between two alarms, turns out to be equal to 1.2 samplings, which corresponds to 1.2 hours, with measurements taken every hour. Notice, however, that increasing the measurement frequency will result in decreasing the mean time necessary to give the alarms.

It is also interesting to notice that in the 833 times that a decision is taken, it has been hypothesis \( H_1 \) (correct hypothesis in the present case study) in 827 cases, and hypothesis \( H_0 \) in 6 cases (wrong hypothesis in the present case study). The fraction of missing alarms 6/827=0.007 is smaller than \( \beta = 0.01 \), in accordance with the analyst-defined parameter value (Table I). Note that, in this case study, a clear tendency of the mean \( \mu \) of the residuals to differ from zero is easily detectable by an analyst observing Fig. 5, and the SPRT has performed as expected.
Fig. 4. Abnormal conditions’ measurements and reconstructions.

Fig. 5. Residuals: difference between measurements and reconstructions of Fig. 4.
Fig. 6. Outcomes of the SPRT for the residuals in Fig. 4.

The SPRT has then been tested in the application for the identification of normal conditions. To this purpose, the reconstruction of the signal has been considered during 1000 hours, under normal conditions (i.e., no offset has been applied to the signal measurements).

Fig. 7 shows the measurements $x_{\text{obs}}$ and the reconstructions $\hat{x}_{\text{nc}}$ obtained by the AAKR reconstruction models. Fig. 8 reports the corresponding residuals $\{R_n\} = r_1, r_2, \ldots, r_n$, which in this case have a mean value close to 0.

Fig. 9 shows the SPRT index obtained by applying the test with the same parameters of Table I. Notice that the SPRT alarm index moves quickly from zero towards the lower boundary ln(A), thus indicating acceptance of the hypothesis $H_0$ (component in a normal conditions). In this case, the empirical ASN has been 1.2 samples, and a decision has been taken 835 times, with 831 times the hypothesis $H_0$ (correct hypothesis), and 4 times the hypothesis $H_1$ (wrong hypothesis). The fraction of false alarms 4/831=0.005 is smaller than $\alpha=0.01$, in accordance with the analyst-defined parameter value (Table I).
Fig. 7. Nominal conditions’ measurements and reconstructions.

Fig. 8. Residuals: difference between measurements and reconstructions of Fig. 7.
The application of the SPRT requires setting the values of the parameters in Table I. As previously noticed in Section 4.1, the correct value of the parameter $\sigma^2$ can be estimated by considering the variance of the residuals under normal conditions. Parameters $\alpha$ and $\beta$ relate to the false and missing alarm probabilities, respectively, and their values are set by the analyst. The most troublesome parameter to set is $\mu_1$, the estimate of the mean value of the residuals in abnormal conditions; this quantity is normally not known in practical applications.

In this section, we analyze the influence of $\mu_1$ on the results of the SPRT with the aim of providing some guidelines for setting its value.

The driving factors of the choice of $\mu_1$ are:

- the Average Sample Number (ASN) necessary to deliver a decision,
- the percentage of false alarms when normal conditions’ residuals are tested,
- the percentage of missing alarms when abnormal conditions’ residuals are tested, and
- the tolerance of the method to identify abnormal conditions when the residuals are characterized by probability distributions different from those expected.

The corresponding desiderata are to have small ASN, small false and missed alarm rates, and a test not very sensible to the distribution of the residuals in abnormal conditions.

Therefore, in what follows, we focus on:
the influence of the choice of \( \mu_1 \) on the ASN, and false and missing alarm rates, for a given sequence of residuals \( \{R_n\} = r_1, r_2, \ldots, r_n \); and

- the influence of different distributions of the residuals, for a fixed \( \mu_1 \).

### 4.3.1. Influence of \( \mu_1 \) on the ASN, and false and missing alarm probabilities

The theoretical behavior of the ASN as a function of the parameters \( \mu_1 \) and \( \sigma \) defining the hypothesis \( H_1 \) can be obtained from (8). Fig. 10 shows the ASN behavior, assuming that the distribution of the residuals has a fixed mean \( \mu = 0.46 \) (the same mean value of the residual in Fig. 5), while varying \( \mu_1 \) and \( \sigma \). In Fig. 11, a magnification of the lowest curve of Fig. 10 shows a qualitative agreement between the theoretical ASN obtained from (8) (circles) and the empirical one (crosses). As expected, the ASN increases with \( \sigma^2 \), and decreases as \( \mu_1 \) increases.

The choice of the value of \( \mu_1 \) should also consider the desired balance between false and missed alarm rates. To this purpose, let us define the operating characteristic function, \( Q \), as the probability that \( H_0 \) is accepted, and the power function, \( P \), as the probability that \( H_1 \) is accepted. Theoretically it is expected that \( P + Q = 1 \) because the SPRT should always deliver a decision if an infinite number of residual samples is available. In the remaining part of this section, we will analyze the behavior of \( Q \) and \( P \) as a function of the mean \( \mu_1 \) of the alternative hypothesis \( H_1 \), in both cases in which abnormal and normal conditions are tested. Note that the missed alarm rate is estimated by \( Q \) in an abnormal conditions test, and the false alarm rate by \( P \) in a normal conditions test.

Fig. 12 reports the experimental evolution of \( P(\mu_1) \) and \( Q(\mu_1) \) for different values of \( \mu_1 \), when the abnormal conditions’ residuals of Fig. 5 are tested. Several SPRT tests with different values of \( \mu_1 \) have been performed on the 1000 available residuals, and we have collected the number of times in which hypothesis \( H_0 \) or \( H_1 \) have been delivered. Notice several points.

- For small values of \( \mu_1 \), the SPRT test does not deliver any conclusion considering the 1000 available residuals. This result is confirmed by (8), which provides SPRT values larger than 1000 (outside the range of Fig. 10).
- When \( \mu_1 \approx \mu/4 = 0.12 \), the SPRT starts to give an alarm (accepting \( H_1 \)) relatively often, with probability \( P(\mu_1) > 1 - \beta \).
- When \( \mu_1 > 2 \mu \approx 1 \), the SPRT begins accepting \( H_0 \) relatively often: the alternative hypothesis \( H_1 \) is not realistic for the case under analysis, and it has to be rejected with probability equal to 1 when \( \mu_1 \approx 4 \mu \), while accepting the null hypothesis \( H_0 \).

Analogously to Fig. 12, Fig. 13 reports the experimental evolution of \( P(\mu_1) \) and \( Q(\mu_1) \) for different values of \( \mu_1 \) when the normal conditions’ residuals of Fig. 8 are tested. Notice that the larger \( \mu_1 \), the more probable is not rejecting \( H_0 \).
Fig. 10. Values of ASN for different $\sigma_i$ and different $\mu_1$, assuming $\mu=0.46$.

Fig. 11. Values of ASN for the case study under analysis at different $\mu_1$. 
Fig. 12. $P(\mu_1)$, and $Q(\mu_1)$ testing the abnormal conditions’ residuals of Fig. 5 with $\mu = 0.46$, and $\sigma = 0.12$.

Fig. 13. $P(\mu_1)$, and $Q(\mu_1)$ testing the normal conditions’ residuals of Fig. 5 with $\mu = 0$, and $\sigma = 0.12$. 
In this case, when choosing the value $\mu_1$, the analyst should simultaneously consider Figs. 11, 12, and 13. The criteria of a small ASN and of a low rate of false alarms would suggest taking a large value of $\mu_1$ because both the ASN (Fig. 11) and the false alarm probability ($P$ curve in Fig. 13, obtained through testing normal conditions’ residuals) are decreasing with $\mu_1$. On the contrary, the objective of minimizing the missing alarm rate would suggest using a low value of $\mu_1$ because the missing alarm probability ($Q$ curve in Fig. 12, obtained testing a normal conditions’ residual) is increasing with $\mu_1$, when a set of residuals are tested, and the system is working in abnormal conditions. A compromise solution for $\mu_1$ needs to be found, able to recognize both normal and abnormal conditions in acceptable time. With reference to the case study analyzed, an indication of taking a value of $\mu_1$ equal to $\mu$ can be suggested. The estimated ASN is not the smallest possible (which could be reached with large values of $\mu_1$), but it can guarantee a good trade-off between small ASN and the reliability of the decision provided by the SPRT.

4.3.2. SPRT results in case of different distributions of the residuals

For the case study at hand, we have suggested setting $\mu_1$ equal to $\mu$ to properly balance the delay in the detection and the probability of false or missed alarms. However, because it is usually not easy for an analyst to a priori quantify the consequences of abnormal conditions on the measured signals, and thus the expected behavior of the residuals from which $\mu$ is estimated, this suggestion is not easily applicable. In practice, however, one can identify, for each signal, the amplitude of the minimum offset that he or she is interested in identifying. This offset can be considered equal to the expected mean value of the residuals between the signal in abnormal conditions ($x^{\text{obs}}$) and the expected value of the signal in normal conditions ($\hat{x}^{\text{nc}}$). For example, in Section 4.2, the minimum amplitude of the offset that has been considered is equal to 1% of the signal standard deviation.

An interesting question arises: which is the behavior of the SPRT in the case of residuals with Gaussian probability distributions different from those supposed in $H_1$? Basically, we want to know what is going to happen if we have fixed a given value of $\mu_1$, according to expert judgment, and then we use the SPRT for the detection of abnormal conditions characterized by a residual distribution with a mean value $\mu$ different from $\mu_1$, or a variance $\sigma^2$ different from that of $H_1$.

To this purpose, Fig. 14 shows the theoretical behavior of the ASN for different values of the residual mean $\mu$ and its standard deviation $\sigma$ (8), when $\mu_1$ is fixed. As expected, if the standard deviation of the residual increases, the ASN tends to become larger. Notice also that, $\sigma$-independently from the standard deviation of the residuals, the maximum of the ASN is reached for $\mu$ equal to $\mu_1/2$.

Finally, observe that for $\mu<\mu_1/2$ the smaller $\mu$ is, the smaller is also the time necessary for the SPRT to deliver a decision. Unfortunately, in these cases in which an abnormal conditions with small $\mu$ is tested, the delivered decision tends to be the wrong hypothesis $H_0$(normal conditions) because the residual distribution results are very similar to the normal conditions’ residual distribution (see Fig. 12 for large $\mu_1$). On the other side, for $\mu>\mu_1/2$, the larger the mean value of the residual is, the smaller the time necessary for the SPRT to deliver a decision is. Thus, in this case, the delivered decision tends to be the correct hypothesis $H_1$. 
Consider an additional example to further clarify the SPRT outcomes when testing residuals with distributions having a mean value $\mu$ different from $\mu_1$. To this purpose, we have simulated a new abnormal conditions on signal 4a which starts at $t=200$ hours, and consists in an offset of increasing amplitude: 0.25% of the standard deviation of the signal values in normal conditions in the time [201-400] hours, 0.5%, in the time interval [401-600] hours, 0.75% in the time interval [601-800] hours, and 1%, in the time interval [801-1000] hours. Fig. 15 shows the measurements $x_{an}^{obs}$, and the reconstructions $\hat{x}_{an}^{nc}$ obtained by the AAKR reconstruction models. Fig. 16 reports the corresponding residuals $\{R_n\} = r_1, r_2, \ldots, r_n$. 

Fig. 14. Values of ASN for different $\sigma$ and $\mu$, assuming $\mu_1$ fixed ($\mu_1=0.46$).
Fig. 15. Signal 4a measurements and reconstructions. Normal conditions from 1 hour to 200 hours, and abnormal from 201 hours to 1000 hours.

Fig. 16. Residuals: difference between the measurements and the reconstructions of Fig. 15.
The SPRT with the same parameters of Table I has been applied to the residuals. Notice that the choice of \( \mu_1 = 0.46 \) indicates that the analyst is interested in identifying offsets of amplitude 0.46 or larger.

Fig. 17 shows the outcomes of the SPRT. Table II reports the number of times in which the SPRT has correctly recognized normal and abnormal conditions, and an estimate of the ASN in the different time intervals.

### Table II

<table>
<thead>
<tr>
<th>Time interval</th>
<th>Offset</th>
<th>Estimated ASN</th>
<th>Number of times in which normal conditions have been detected</th>
<th>Number of times in which abnormal conditions have been detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0-200]</td>
<td>No</td>
<td>1.2</td>
<td>150</td>
<td>2</td>
</tr>
<tr>
<td>[201-400]</td>
<td>Yes (amplitude = 0.11)</td>
<td>1.9</td>
<td>70</td>
<td>5</td>
</tr>
<tr>
<td>[401-600]</td>
<td>Yes (amplitude = 0.23)</td>
<td>2.4</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>[601-800]</td>
<td>Yes (amplitude = 0.34)</td>
<td>2.1</td>
<td>0</td>
<td>94</td>
</tr>
<tr>
<td>[801-1000]</td>
<td>Yes (amplitude = 0.46)</td>
<td>1.2</td>
<td>2</td>
<td>142</td>
</tr>
</tbody>
</table>

Notice that, as expected, when in the time interval \( t = [1-200] \) hours, the component is under normal conditions, the SPRT index moves quickly from zero towards the lower boundary \( \ln(A) \), thus indicating to
accept the hypothesis $H_0$ (component under normal conditions) with high probability. Then, when $t = [201-400]$ hours ($\mu = 0.11, \mu_1 = 0.46, \mu < \mu_1/2$), the decisional process is still very fast (see Fig. 14); but, although the case is one of abnormal conditions (an offset of small amplitude is applied to the signal), the delivered decision tends to be the wrong hypothesis $H_0$ (normal conditions) because the residuals distribution is very similar to the normal conditions’ residuals distribution. Later, for $t = [401-600]$ hours ($\mu = 0.23, \mu_1 = 0.46, \mu < \mu_1/2$), the ASN reaches its maximum (see Fig. 14), but the missed alarm rate decreases in favor of a larger number of true alarm flags. For $t = [601-800]$ hours ($\mu = 0.33, \mu_1 = 0.46, \mu > \mu_1/2$), the test becomes much more reliable because the false alarms rate approaches $\beta$. For $t = [801-1000]$ hours ($\mu = 0.45, \mu_1 = 0.46, \mu < \mu_1$), the ASN is very small, and the delivered decision tends to be the correct hypothesis $H_1$.

As a final investigation, we are interested in analyzing the results of the previous test considering an alternative hypothesis $\mu_1 = 0.11$, rather than $\mu_1 = 0.46$ as it was in the previous analysis. This alternative setting agrees with the suggestion given in the previous section, i.e., $\mu_1$ equal to $\mu$, where $\mu$ is the smallest value of deviation from nominal conditions that the experts foresee to occur in the case of anomalies.

Fig. 18 shows the outcomes of the SPRT. Starting from $t = 201$ hours, that is the onset of the abnormal conditions with $\mu = \mu_1 = 0.11$, the alarm flag is already raised with very high probability. See that the ASN is larger than that obtained in Fig. 11 due to the fact that the theoretical ASN is not at its minimum (which can be reached with very large values of $\mu_1$, see Fig. 19). However, assuming $\mu_1 = 0.11$ has allowed a guarantee of a good trade-off between small ASN and the reliability of the decision provided by the SPRT.

![Fig. 18. Outcomes of the SPRT for the reconstructed signals reported in Fig. 15, $\mu_1 = 0.11$.](image-url)
5. Conclusions

The problem of detecting faults in components of NPPs has been considered in this paper. Based on previously published works [29], a condition monitoring scheme capable of identifying the signal groups by a hybrid approach based on CA and GA has been adopted. This approach has been selected because of its capability to limit fault propagation with respect to a CA-based filter approach, and to demand a lower computational effort with respect to a GA-based wrapper approach. The grouping outcomes are then fed to a traditional empirical AAKR method for signal reconstructions.

On the basis of the reconstructions provided by the AAKR models (one for each signal group identified), the SPRT for the detection of abnormal conditions has been adopted for balancing misclassification probabilities, i.e., false and missed alarm rates. A novel procedure for setting the parameters used of the SPRT has also been proposed, for practical use.

A case study regarding the monitoring of the RCPs of a typical PWR has been considered. In the performed tests, the combination of a hybrid signal grouping technique with AAKR and SPRT has proven capable of detecting the onset of abnormal conditions with a controlled low percentage of false and missed alarms. Future work will entail the design of decision thresholds such that the SPRT strategy can guarantee the minimum time for the fault detection (ASN), even in the case of auto-correlated signals.
Appendix

The Auto-Associative Kernel Regression (AAKR) method

Let \( X_{\text{obs}} \) be a matrix of observed data whose generic element \( x_{\text{obs}}(k, n) \) represents the \( k \)-th time observation, \( k = 1, \ldots, T \), of the \( n \)-th measured signal, \( n = 1, \ldots, N \), taken during normal plant conditions. The basic idea of the AAKR method is to reconstruct the signal values in case of normal conditions, \( \hat{x}_{\text{n}} \), given a current signal measurement vector, \( \hat{x}_{\text{obs}} = (x_{\text{obs}}(1), \ldots, x_{\text{obs}}(N)) \), as a weighted sum of the observations in \( X_{\text{obs}} \). Thus, \( \hat{x}_{\text{n}} \), the reconstruction of \( x_{\text{obs}}(n) \), the \( n \)-th component of \( \hat{x}_{\text{obs}} \), is given by

\[
\hat{x}_{n}(n) = \frac{\sum_{k=1}^{T} w(k) \cdot x_{\text{obs}}(k, n)}{\sum_{k=1}^{T} w(k)}
\]

(10)

The weights \( w(k) \) are similarity measures obtained by computing the Euclidean distance between the current sensor measurements \( \hat{x}_{\text{obs}} \) and the \( k \)-th observation of \( X_{\text{obs}} \):

\[
d^2(k) = \sum_{x=1}^{K} (\hat{x}_{\text{obs}}(n) - x_{\text{obs}}(k, n))^2.
\]

(11)

Inserting it in the Gaussian kernel,

\[
w(k) = \frac{1}{\sqrt{2\pi h}} e^{-\frac{d^2(k)}{2h^2}}
\]

(12)

where the signal \( h \) defines the Gaussian bandwidth.

To provide in (11) a common scale across the different signals measuring different quantities, it is necessary to normalize their values. In the present work, the signal values are normalized according to

\[
x(n) = \frac{x(n) - \mu(n)}{\sigma(n)}
\]

(13)

where \( x(n) \) is a generic measurement of signal \( n \), and \( \mu(n) \) and \( \sigma(n) \) are respectively the mean and standard deviation of the \( n \)-th signal in \( X_{\text{obs}} \):

\[
\mu(n) = \frac{\sum_{k=1}^{T} x_{\text{obs}}(k, n)}{T}
\]

\[
\sigma(n) = \sqrt{\frac{\sum_{k=1}^{T} (x_{\text{obs}}(k, n) - \mu(n))^2}{T}}
\]
Acknowledgements
This research has been carried out under contract C935C50010 funded by EdF R&D STEP Simulation et Traitement de l’information pour l’exploitation des systèmes de production, Chatou Cedex, France) that has provided the data on the RCP of the PWR.

References
European Safety and Reliability Association, ESRA, Technical Committee on Fault Diagnosis. He is functioning as Technical Assistant professor of Nuclear Engineering at the department of Energy at the Politecnico di Milano. He is the current chairman of the Review International Journals, and 25 on Proceedings of International Conferences. He is Chair of the Italian IEEE Reliability Section and the IEEE Nuclear Power Plant Committee. He is a member of the International Journal of Nuclear Safety and Risk Analysis of Nuclear Power Plants, and for System Monitoring, Fault Diagnosis, and Prognosis. In 2009-2010 he has been Research Fellow of the Science and Technology Programme (STFP) in China, 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 been collaborator of the Italian Embassy in China, and acted as technical committee member for a number of international conferences. He is co-author of 22 articles in peer-reviewed international journals, and 25 on proceedings of international conferences. He is Chair of the Italian IEEE Reliability Chapter.

**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 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 the safety and risk analysis of nuclear power plants, and for system monitoring, fault diagnosis, and prognosis. 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 been collaborator of the Italian Embassy in China, and acted as technical committee member for a number of international conferences. He is co-author of 22 articles in peer-reviewed international journals, and 25 on proceedings of international conferences. He is Chair of the Italian IEEE Reliability Chapter.
Enrico Zio (B.Sc. in nuclear engineering, Politecnico di Milano, 1991; M.Sc. in mechanical engineering, UCLA, 1995; Ph.D., in nuclear engineering, Politecnico di Milano, 1995; and a Ph.D., in nuclear engineering, MIT, 1998). He is the Director of the Chair on Complex Systems and the Energy Challenge at Ecole Centrale Paris and Supelec, Fondation Europeenne pour l’Energie Nouvelle – EdF, and the full Professor of Computational Methods for Safety and Risk Analysis in Politecnico di Milano, Italy. He is President of Advanced Reliability, Availability and Maintainability of Industries and Services (ARAMIS) ltd., and Chairman of the European Safety and Reliability Association, ESRA. His research topics are: analysis of the reliability, safety and security 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. He is co-author of five international books and more than 200 papers on international journals, and serves as referee of more than 20 international journals.

Redouane Seraoui works for EDF in the Research & Development division. Following the graduation from the French engineering school "Ecole Supérieure d'Electricité," he works now on a series of large research projects involving collaboration between many leading academic and industry partners. These projects include the development of tools and methods for the monitoring of power plants' critical components, applied to the nuclear industry but also to the fossil and renewable sides.