



HAL
open science

A Fuzzy Similarity-Based Method for Failure Detection and Recovery Time Estimation

Enrico Zio, Francesco Di Maio

► **To cite this version:**

Enrico Zio, Francesco Di Maio. A Fuzzy Similarity-Based Method for Failure Detection and Recovery Time Estimation. *International Journal of Performability Engineering*, 2010, 6 (5), pp.407-424. <hal-00609189>

HAL Id: hal-00609189

<https://centralesupelec.hal.science/hal-00609189v1>

Submitted on 26 Jul 2012

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.



HAL Authorization

A FUZZY SIMILARITY-BASED METHOD FOR FAILURE DETECTION AND RECOVERY TIME ESTIMATION

Enrico Zio, Francesco Di Maio

Energy Department, Polytechnic of Milan

Via Ponzio 34/3, 20133 Milano, Italy

enrico.zio@polimi.it

ABSTRACT

In this work, an extension of a data-driven approach for estimation of the available Recovery Time (RT) is presented. The improvement is in avoiding the need of resorting to a fault detection module for the identification of the anomalous system behavior: the algorithm proposed jointly detects the onset of the accidental transient and performs the estimation of the available RT. This is achieved by fuzzy similarity analysis of the currently developing scenario and reference multidimensional trajectory patterns of failure scenarios; the RT remaining before the developing trajectory pattern hits a failure threshold is predicted by combining the times of failure of the reference patterns, weighed by their similarity with the developing pattern.

For illustration purposes, failure scenarios of the Lead Bismuth Eutectic eXperimental Accelerator Driven System (LBE-XADS) are considered.

Key Words: Failure Detection, Recovery Time Estimation, Emergency Accident Management, Lead-Bismuth Eutectic eXperimental Accelerator Driven System (LBE-XADS), Fuzzy Similarity Analysis, Pattern Recognition.

1. Introduction

Nuclear Power Plant (NPP) accident management involves the anticipation of paths of potentially dangerous behaviors, the prediction of the related effects and the setting of actions to avoid undesired impacts on the safety of the NPP [IAEA, 2003].

In case of an accident, or an initiating event that may develop into an accident, the plant personnel must perform various tasks before taking counteracting actions:

- Identification of the plant state: this diagnostic task aims at identifying the cause of the problem and the states of a number of parameters critical for the plant operation and safety;
- Prediction of the future development of the accident: this involves prediction of the future evolution of the states of the critical parameters and of the plant residual Recovery Time (RT), which is the time remaining for taking corrective actions to avoid system failure;
- Planning of accident mitigation strategies, to be activated if the actions to safely control the accident evolution were not successful.

Computerized procedures can be used to aid the operators in the tasks of obtaining reliable information, following procedures, identifying plant states, predicting the future accident progression and planning accident control and mitigation actions. Indeed, it is a shared belief that the complicated phenomena taking place in a NPP during an accident situation need to be supported by computerized procedures rather than left only to human expert evaluation [Øwre, 2001]. On the other hand, the problem of what kind of support to provide to nuclear power plant operators, in particular during transients leading up to accidents, is far from trivial [NEA, 1992; EC, 1999; USNRC, 1999; IAEA, 2003].

Fortunately, NPP personnel have the capability to effectively manage a broad range of accidents; their successful handling of complex accident behaviors requires that they detect the occurrence of the accident, determine the extent of challenge to plant safety, monitor the performance of active, passive, automatic and digital systems, select strategies to prevent or

mitigate the safety challenge, implement the action strategies, and monitor their effectiveness. The capability to effectively carry out these tasks during an accident is influenced by the availability of timely and accurate plant status information and the awareness of the RT available for action. Poor decisions may be taken under a perceived short time available for sorting out the information relevant to the plant status [Glasstone et al., 1998]; on the contrary, timely and correct decisions can prevent an event from developing into a severe accident or mitigate its undesired consequences.

The existing computer-based tools which can aid accident management can be categorized according to their complexity and specific application purpose [IAEA, 2003]:

- Compact simulators;
- Plant analyzers;
- Full-scope training simulators;
- Multifunctional simulators;
- Severe accident simulators;
- Accident management support tools.

In particular, the accident management support tools combine tracking and predictive simulators of the developing accidental transients. The tracking simulators monitor the plant status and provide calculated values also of those parameters that are not directly observable by the monitoring systems.

One of the main quantities of interest to be delivered by predictive simulators for accident management use is the available RT. Approaches to RT prediction can be categorized broadly into model-based and data-driven [Chiang et al., 2001]. Model-based approaches attempt to incorporate physical models of the system into the estimation of the RT. However, uncertainty due to the assumptions and simplifications of the adopted models may pose significant limitations. Moreover, the requirement of high computational speed for on line response necessarily leads to limited details in the phenomena modeled, with consequent limited accuracy in the representation of the actual

plant behavior [Berglund et al., 1995; Serrano et al., 1999]. On the contrary, data-driven techniques utilize monitored operational data related to system health. They can be beneficial when understanding of first principles of system operation is not straightforward or when the system is so complex that developing an accurate model is prohibitively expensive. Data-driven approaches can often be deployed quickly and cheaply, and still provide wide coverage of system behavior.

Data-driven approaches can be divided into two categories: statistical techniques (regression methods, ARMA models, etc.) and Artificial Intelligence (AI) techniques (neural networks, fuzzy logic systems, etc.).

With respect to AI techniques, Neural Networks (NNs) and Fuzzy Logic (FL) techniques have gained considerable attention in the past few years, due to their ability to deal with the uncertainties and non-linearities of the real processes, especially in abnormal conditions [Øwre, 2001]. Successful prediction models have been constructed based on Neural Networks [Barlett et al., 1992; Campolucci et al., 1999; Peel et al., 2008; Zio et al., 2008; Santosh et al., 2009] and Neuro-Fuzzy (NF) systems [Wang et al., 2004]. In spite of the recognized power of neural network modeling techniques, skepticism against their use in safety-critical applications relates to their black-box character which limits intuition with respect to the understanding of their performance [Wang et al., 2008].

An opportunity for increased transparency and openness of data-driven models is offered by fuzzy logic methods, which are increasingly proposed in modern control and diagnostic technologies. Based on the principles of Zadeh's fuzzy set theory, fuzzy logic provides a formal mathematical framework for dealing with the vagueness of everyday reasoning [Zadeh, 1965]. As opposed to binary reasoning based on ordinary set theory, measurement uncertainty and estimation imprecision can be properly accommodated within the fuzzy logic framework [Yuan et al., 1997; Zio et al., 2005].

The work presented in this paper extends the multidimensional approach previously presented by the authors for the prediction of the available RT during an accident [Zio et al., 2009b]. In particular, the need is avoided of a fault detection module that from the beginning matches the evolving trajectory with the trajectories stored in the database; in view of a simplification of the procedure for its practical application and an increase of the reliability of the overall accident management support system, the RT estimation module is devised so as to automatically identify the anomalous behavior of the plant and proceed to the RT evaluation.

The computational framework considers a set of multidimensional trajectory patterns arising from different system failures (hereafter called reference trajectory patterns) and uses a fuzzy-based, data-driven similarity analysis for predicting the RT of a newly developing failure trajectory (hereafter called test trajectory pattern). The pattern matching process is based on a fuzzy evaluation of the distance between the signals of the multidimensional test pattern and the patterns of reference [Angstenberger, 2001]; the fuzzy distances from all reference patterns are then combined to transform the multidimensional data into a one-dimensional similarity indicator, which is used in the estimation of the available RT.

An application of the method is illustrated with regards to the dynamic failure scenarios of the Lead Bismuth Eutectic eXperimental Accelerator Driven System (LBE-XADS) with digital Instrumentation and Control (I&C) [Cammi et al., 2006]. The focus of the application is on the RT estimation and, therefore, the modeling is tailored to the purpose of showing the feasibility of effectively identifying the onset of a failure scenario and estimating the RT during an accident: thus, the analysis does not cover the software failure behavior, the benefits of fault tolerant features, the interactions of the software with the hardware and human components; it is understood that the actual implementation of the method in a qualified tool of accident management will require full dynamic accident calculations, inclusive of comprehensive models of hardware, software and human failure modes and their interactions.

The paper is structured as follows. Section 2 provides a detailed description of the computational algorithm for the multidimensional fuzzy data similarity evaluation and the associated RT prediction. Section 3 presents the mechanistic model of the LBE-XADS, with the description of the monitored signals. In Section 4, the results of the application of the approach to LBE-XADS failure scenarios are presented. Finally, some conclusions are drawn in Section 5.

2. Methodology of the RT estimation by fuzzy similarity analysis

It is assumed that N trajectories (reference trajectory patterns) are available, representative of the evolution of relevant signals during reference failure scenarios. These trajectories last all the way to system failure, i.e., to the time when anyone of the signals reaches the threshold value beyond which the system loses its functionality.

At any time, a multidimensional signal trajectory (test trajectory pattern) is compared for similarity with the N multidimensional reference trajectory patterns stored in the database and the residual lifes of these are used to estimate the RT available in the developing transient.

Figure 1 shows a schematics of the computational framework in the general case of multidimensional trajectories of Z monitored signals $f(x_1, x_2, \dots, x_Z, t)$.

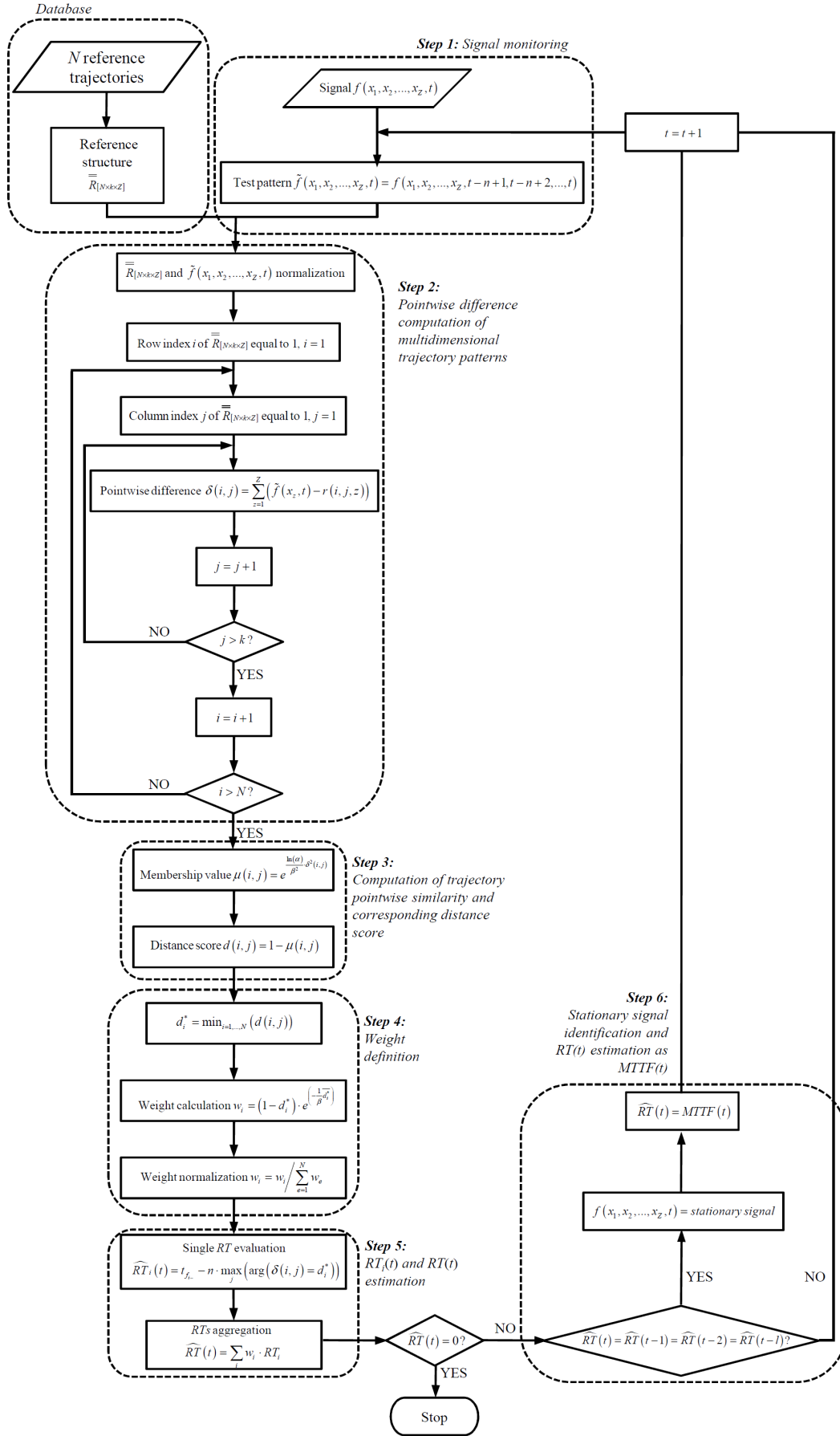


Figure 1 The flowchart of the procedure for RT estimation by fuzzy similarity analysis

For completeness of the paper, procedural Steps 1-5 of the algorithm, already presented in [Zio et al., 2009a], are reported in Appendix. Here, only the key modifications with respect to the previously presented procedure (i.e., the addition of Step 6) are highlighted, together with the motivation behind it.

- ***Step 1: signal monitoring.***

See Appendix.

- ***Step 2: pointwise difference computation of multidimensional trajectory patterns.***

See Appendix.

- ***Step 3: computation of trajectory pointwise similarity and corresponding distance score.***

See Appendix.

- ***Step 4: weight definition.***

See Appendix.

- ***Step 5: $RT_i(t)$ and $RT(t)$ estimation.***

See Appendix.

- ***Step 6: Stationary signal identification and $RT(t)$ estimation as $MTTF(t)$.***

Starting from time $t=1$, at each discrete time t , the devised algorithm is qualified to continuously provide the prediction $\widehat{RT}(t)$. However, in case that the developing signal is actually stationary, $\widehat{RT}(t)$ would be biased towards too conservative estimations because of the “max” assumption adopted in Eq. (10) in the Appendix for the computation of t_{j_M} ; under this assumption, an $(l+1)$ -long stationary segment entails that $\widehat{RT}(t) = \widehat{RT}(t-1) = \widehat{RT}(t-2) = \widehat{RT}(t-l)$, $l=1,2,\dots,T$, as shown in Figure 2. The detection of such string of RT estimates allows for *i*) the classification of the signal as of no abnormal deviation *ii*) the corresponding qualification of the system state as working in nominal conditions (in other words, no fault has been detected and the system reaches safely

the time horizon of observation T) and, therefore, *iii*) the correction of the bias of extremely low (i.e., pessimistic) RT predictions. With respect to the latter point, when at time t the signal is classified as stationary, the correction consists in the substitution of the estimate $\widehat{RT}(t)$ with the system Mean Time to Failure $MTTF(t)$, obtained from the available recovery times $RT_i(t)$ of the entire population of failure trajectories in the reference library:

$$\widehat{RT}(t) = MTTF(t) = \frac{1}{(|i| |t_{f_i} > t|)} \sum_{i|t_{f_i} > t} (t_{f_i} - t) = \frac{1}{(|i| |t_{f_i} > t|)} \sum_{i|t_{f_i} > t} RT_i(t) \quad (1)$$

where t_{f_i} is the system failure time along the i -th trajectory, $(|i| |t_{f_i} > t|)$ is the cardinality of the set of reference trajectories whose failure time is larger than t and $RT_i(t)$ is their residual life starting from t .

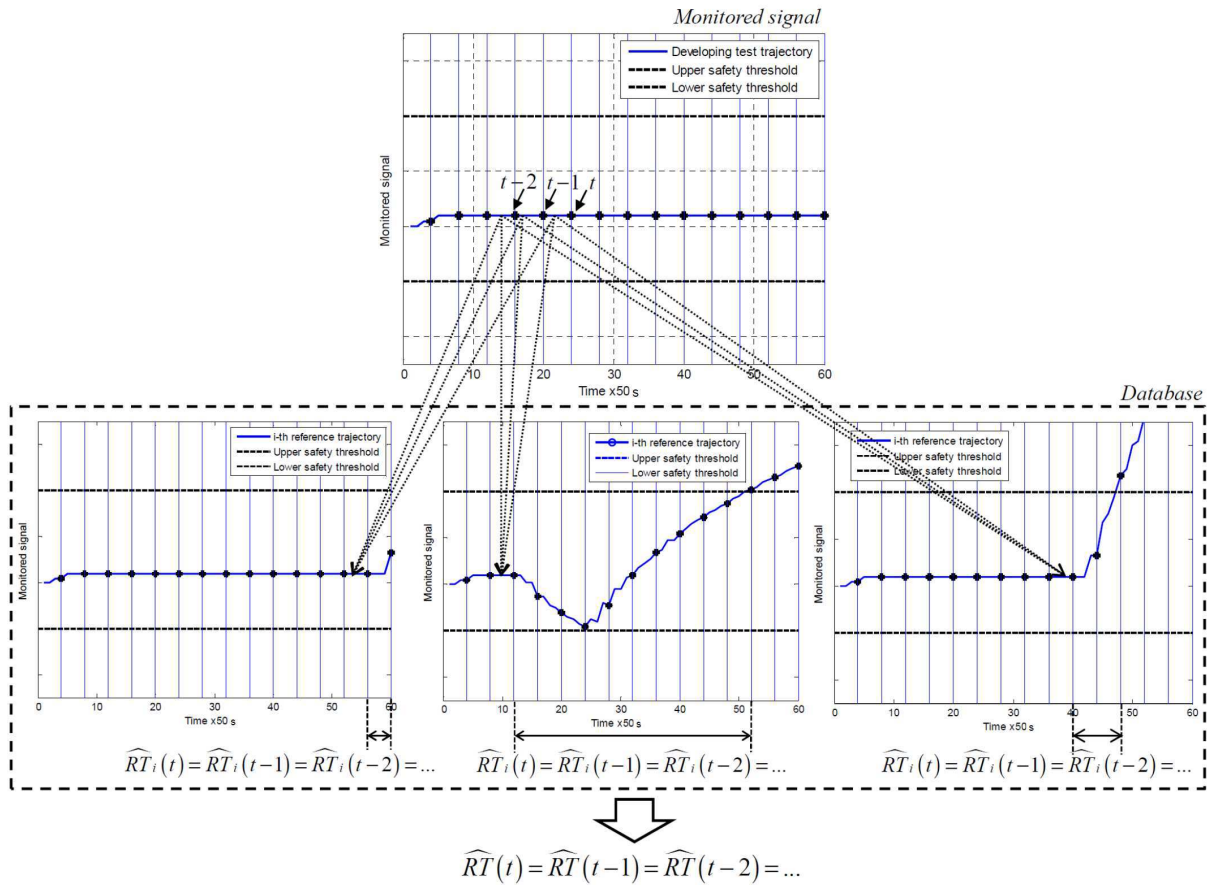


Figure 2 The flowchart of the procedure for RT estimation by fuzzy similarity analysis applied to a stationary signal

In the subsequent time steps, as the scenario evolves, the procedure restarts from Step 1 and continues to update the estimate of $\widehat{RT}(t)$: in case that non-stationarity in the signal is detected, the $\widehat{RT}(t)$ is computed through steps 1-5 of the procedure; otherwise, $\widehat{RT}(t)$ is taken equal to the $MTTF(t)$.

3. Case study

The case study considered is the same as in [Zio et al., 2009a], concerning the RT in failure trajectories of the Lead-Bismuth Eutectic eXperimental Accelerator Driven System (LBE-XADS), a sub-critical, fast reactor in which the fission process for providing thermal power $P(t)$ is sustained by an external neutron source through spallation reaction by a proton beam $Q(t)$ accelerated by a synchrotron on a lead-bismuth eutectic target [Bowman et al., 1992; Carminati et al., 1993; Rubbia et al., 1995; Van Tuyle et al., 1993; Venneri et al., 1993]. The physical description of the LBE-XADS and of the failures considered are here repeated, for completeness of the information in the paper. The primary cooling system is of pool-type with Lead-Bismuth Eutectic (LBE) liquid metal coolant leaving the top of the core, at full power nominal conditions, at temperature equal to 400 °C and then re-entering the core from the bottom through the down-comer at a temperature of 300 °C. The average in-core temperature of the LBE $T_{LB}^{av,C}$ is taken as the mean of the entrance and exit temperatures. The secondary cooling system is a flow of an organic diathermic oil at 290-320 °C, at full power conditions. Cooling of the diathermic oil in each loop is obtained through an air flow $\Gamma_a(t)$ provided by three air coolers connected in series.

A dedicated, dynamic simulation model has been implemented in SIMULINK for providing a simplified, lumped and zero-dimensional description of the coupled neutronic and thermo-hydraulic evolution of the system [Cammi et al., 2006]. The model allows the simulation of the

system controlled dynamics as well as of the free dynamics when the control module is deactivated and the air cooler flow is kept constant.

Both feedforward and feedback digital control schemes have been adopted for the operation of the system. The control is set to keep a steady state value of approximately 300 °C of the average temperature of the diathermic oil $T_o^{av,S}$: this value represents the optimal working point of the diathermic oil at the steady state, full nominal power of 80 MWth. On the contrary, an oil temperature beyond the upper threshold $T_o^{th,u}=340$ °C would lead to degradation of its physical and chemical properties, whereas a temperature below the lower threshold $T_o^{th,l}=280$ °C could result in thermal shocks for the primary fluid and, eventually, for the structural components [Cammi et al., 2006]. Conservatively, no dependence on the duration of exposition to temperatures beyond the threshold values has been assumed: in other words, the system is considered to fail at such temperatures regardless of the time during which it exceeds the thresholds.

Multiple component failures can occur during the system life. To simulate this, the model has been embedded within a Monte Carlo (MC) sampling procedure for injecting faults at random times and of random magnitudes. Samples of component failures are drawn within a time horizon of 3000 s. The set of faults considered are:

- The PID controller fails stuck, with a random flow rate output value m_1 sampled from a uniform distribution in [0,797] Kg/s.
- The air coolers fail stuck in a random position that provides a corresponding air flow mass m_2 uniformly distributed in [0,1000] Kg/s.
- The feedforward controller fails stuck with a corresponding flow rate value m_3 uniformly distributed in [0,797] Kg/s.
- The communication between air coolers actuators and PID controller fails so that the PID is provided with the same input value of the previous time step.

The first three faults are applicable to both analog and digital systems, whereas the last one is typical of digital systems. Furthermore, the fault magnitude probability distributions are assumed to be uniform, even if the components may more likely fail in a certain mode than in others. This includes also rare multiple events in the set of failure scenarios and further tests the robustness of the RT prediction procedure.

The sequence of multiple failures is generated by sampling the first failure time from the uniform distribution [0,3000] s and the successive failure times from the conditional distributions, uniform from the last sampled time to 3000 s. This assumption is conservative, favoring larger number of failures in the sequence [Di Maio et al., 2009].

The evolution of the failure scenarios may lead to three different end states, within the time horizon of 3000 s:

1. Low-temperature failure mode ($T_o^{av,S} < T_o^{th,l}$)
2. Safe mode ($T_o^{th,l} < T_o^{av,S} < T_o^{th,u}$)
3. High-temperature failure mode ($T_o^{av,S} > T_o^{th,u}$)

The following three signals are taken for the estimation of the available RT:

- Mean in-core LBE temperature, $T_{LB}^{av,C}$
- Mean oil temperature of the secondary heat exchanger hot side, $T_o^{av,S}$
- Mean air flow rate at the secondary heat exchanger cold side, $\Gamma_a(t)$

It is important to stress that the procedure for sampling the fault events here implemented is not intended to reproduce the actual stochastic failure behavior of the system components; rather, the choices and hypotheses for modeling the faults (i.e., the time horizon of the analysis, the number and typology of faults, the distributions of failure times and magnitudes) have been arbitrarily made with the aim of favoring multiple failures. In any case, the components considered subjected to fault and their types are not intended to provide a comprehensive description of the system fault behavior

but are only taken as exemplary for generating the dynamic failure scenarios to be used as reference and test patterns.

4. Results

4.1 Application of the procedure for $RT(t)$ prediction

For the present case study, $Z=2$ -dimensional reference trajectories of evolution of $T_{LB}^{av,C}$ and $T_o^{av,S}$, relative to $N = 6400$ failure scenarios (differing in faulty components, times of faults occurrence and faults magnitudes), have been used as in [Zio et al., 2009b]. The database of reference trajectories is organized in the structure $\overline{R}_{[N \times k \times Z]}$, where $k = \frac{T}{n} = 30$. The generic element $r(i, j, z)$ of the reference structure will be compared for similarity with the z -th signal of the test trajectory pattern containing the values of the latest $n=100$ time steps of the trajectory. For each of the test trajectories the six steps of the procedure (see Appendix) are performed as follows:

Step 1: signal monitoring.

The 2-dimensional trajectory $f(x_1, x_2, t)$ is followed starting from time $t = 1$ s to the end of the time horizon $T = 3000$ s, with time steps of 1 s. At each time step t , its value is appended and stored in the matrix containing the $n-1=99$ values of the two trajectory signals collected at the previous times. The fuzzy similarity algorithm for RT estimation is activated.

Step 2: pointwise difference computation of multidimensional trajectory patterns.

The total pointwise difference $\delta(i, j)$ is evaluated by resorting to Eq. (3) in the Appendix, with $z=1, 2$.

Step 3: computation of trajectory pointwise similarity and corresponding distance score.

The pointwise differences $\delta(i, j)$ are mapped into values of membership $\mu(i, j)$ of the “approximately zero” FS. The bell-shaped function of Eq. (5) in the Appendix is taken with parameters values $\alpha = 0.2$ and $\beta = 0.02$, implying strong sharpness in the FS and thus in the similarity requirement [Zio et al., 2009a]. The distance scores $d(i, j)$ are then computed by Eq. (4) in the Appendix, $i = 1, 2, \dots, 6400$, $j = 1, 2, \dots, 30$.

Step 4: weight definition.

The minimum distances d_i^* are evaluated (Eq. (7) in the Appendix), and the relative normalized weights w_i calculated through Eqs. (8) and (9) in the Appendix, $i = 1, 2, \dots, 6400$.

Step 5: $RT_i(t)$ and $RT(t)$ estimation.

For each reference trajectory in the library, an estimate $\widehat{RT}_i(t)$ for the test trajectory is computed (Eq. (10) in the Appendix, $i = 1, 2, \dots, 6400$); then, the values $\widehat{RT}_i(t)$ are aggregated in the weighted sum (Eq. (6) in the Appendix) with the weights w_i previously calculated.

Step 6: stationary signal identification and $RT(t)$ estimation as $MTTF(t)$.

In case a stationary signal is detected (i.e., $\widehat{RT}(t) = \widehat{RT}(t-1) = \widehat{RT}(t-2) = \widehat{RT}(t-l)$, $l = 1, 2, \dots, T$), the Mean Time to Failure $MTTF(t)$ is calculated resorting to Eq. (1) and $\widehat{RT}(t)$ is set equal to it, for each time step.

The estimates of the $MTTF(t)$ for five 2-D test pattern trajectories taken from [Zio et al., 2009a] are plotted in Figures 3-7, in thin continuous lines with the bars of one standard deviation of the samples $(t_{f_i} - t | t_{f_i} > t)$, where t_{f_i} is the time at which the diathermic oil temperature profile of the i -th reference trajectory exceeds either thresholds $T_o^{th,u}$ or $T_o^{th,l}$, with corresponding system loss

of functionality; the $\widehat{RT}(t)$ estimates obtained based on trajectory segments of length $n = 100$ s are also plotted in bold circles. Initially, the estimate of the RT matches the $MTTF(t)$, because the algorithm has not yet acquired enough information about the developing trajectory; then, the algorithm starts to provide the $\widehat{RT}(t)$, based on the results of the fuzzy similarity analysis: the $\widehat{RT}(t)$ values matches the $MTTF(t)$ in case of a stationary signal, or drifts away from the $MTTF(t)$ values and towards the real $RT(t)$ (dashed thick line), if an anomalous signal deviation is detected. In the Figures, the bold vertical line indicates the time of diathermic oil threshold exceedance; notice that none of the estimates exceeds the actual failure time.

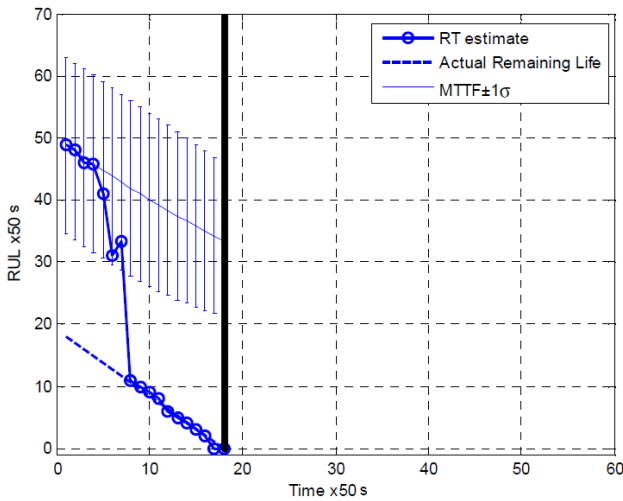


Figure 3 RT estimation for a trajectory belonging to the low-temperature failure mode

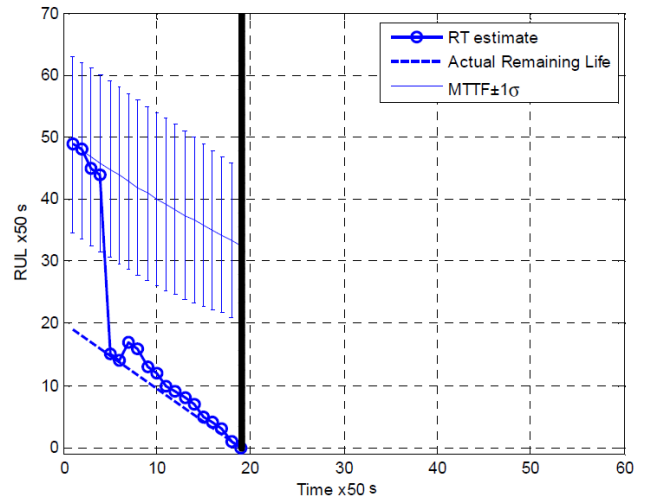


Figure 5 RT estimation for a trajectory belonging to the high-temperature failure mode

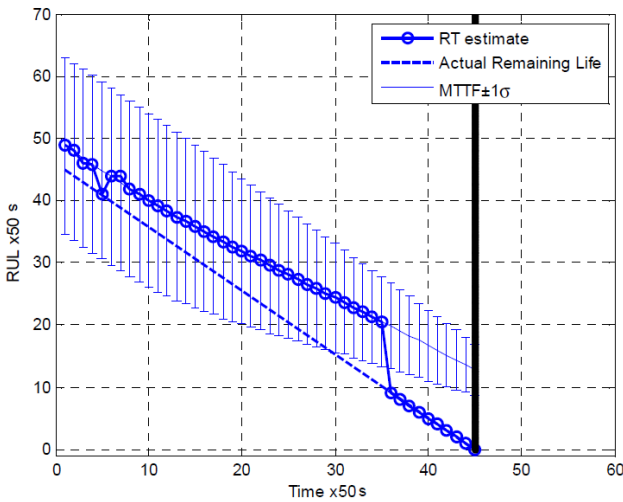


Figure 4 RT estimation for a trajectory belonging to the low-temperature failure mode

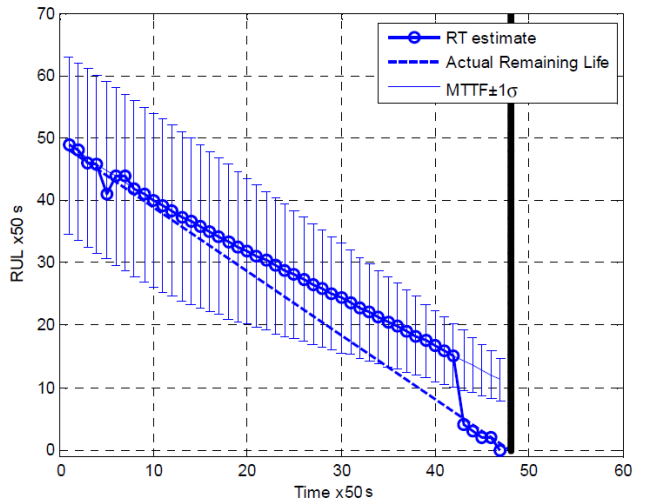


Figure 6 RT estimation for a trajectory belonging to the high-temperature failure mode

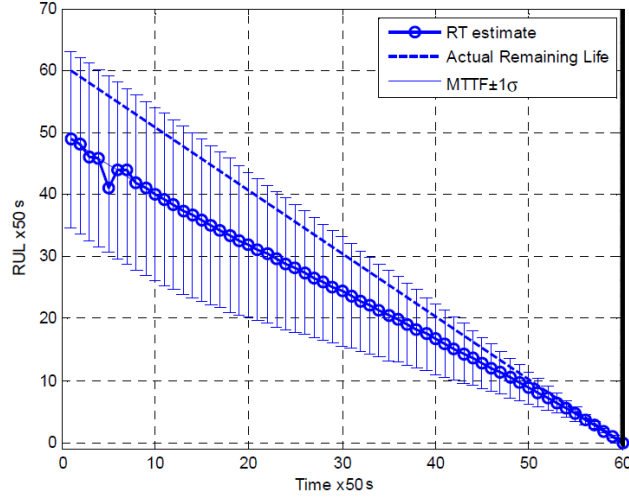


Figure 7 RT estimation for a trajectory which does not exceed any safety threshold value, although a failure sequence has occurred

4.2 Performance evaluation of the $RT(t)$ estimation procedure

The performance of the available RT estimation procedure has been tested extensively on a batch of $P=1280$ multidimensional test trajectories, different from the reference trajectories used in the fuzzy similarity analysis upon which the RT estimation is based. Figure 8 shows the $\widehat{RT}(t)$ predictions (continuous line with dark bullets) compared to the actual remaining lives for 25 failure trajectories (graphically appended in sequence), starting from the time in which the trajectories deviates from the stationary nominal value. The assumption made in step 5 of the procedure for the evaluation of Eq. (7) in the Appendix results in a conservative trend of initial anticipation of the available RT associated to trajectories whose failure actually occurs late in life. Also, as expected, the largest available RT estimation errors occur for those trajectories in which the component failure is of low-magnitude, whose effect only slowly drives the system to failure and $\widehat{RT}(t)$ estimation away from $MTTF(t)$, towards the true $RT(t)$.

To globally quantify the performance of the procedure, the mean relative error (RE) at time t , between the estimate $\widehat{RT}(t)$ and its true value $RT(t)$, is introduced:

$$RE(t) = \frac{1}{P} \sum_{p=1}^P \frac{|\widehat{RT}_p(t) - RT_p(t)|}{RT_p(t)} \quad (2)$$

where $RT_p(t)$ is the actual available recovery time at time t of test pattern p , and $\widehat{RT}_p(t)$ its estimate, $p = 1, 2, \dots, P$.

Figure 9 shows the empirical probability density function of the mean relative error evaluated on 3000 s of evolution of the 1280 test trajectories; the distribution is skewed towards small error values, with mean and median equal to 0.08 and 0.03, respectively: this proves that the procedure most frequently makes small relative estimation errors.

The computational time required for the estimation along one complete test trajectory of 3000 s is of 20 s on an Intel[®] Celeron M of 900 MHz.

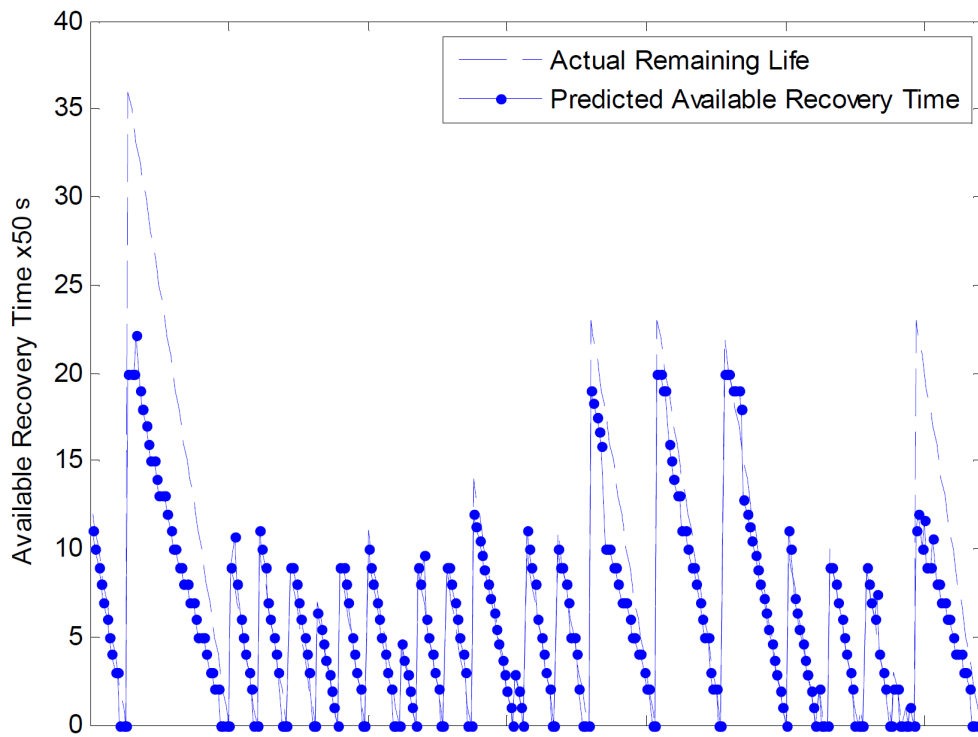


Figure 8 Predicted and actual remaining life for 25 test failure scenarios

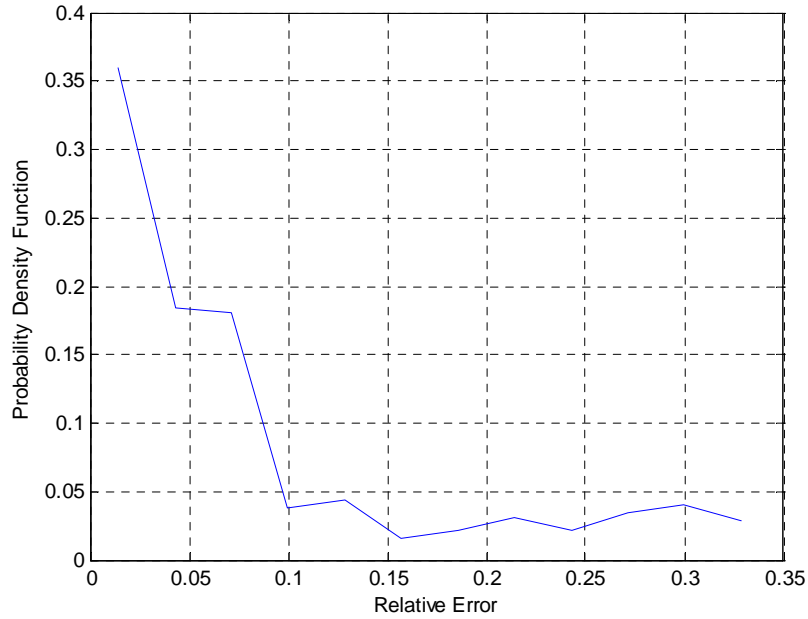


Figure 9 Empirical probability density function of the relative errors on the 3000 s of the 1280 test trajectories

Figure 10 compares the RE evaluated at time steps of 50 s in the last 600 s when the LBE temperature and the oil temperature are simultaneously monitored. In one case, the fault detection module is activated as in [Zio et al., 2009b] while in the other case the results have been obtained with the algorithm as implemented in this work, i.e., without resorting to any “a priori” fault detection module. In both cases, β in Eqs. (5) and (8) of the Appendix is set equal to 0.02. It is seen that:

- the accuracy in the estimation of the available RT improves over time: as the available RT decreases, the relative errors approximate their mean values approximately equal to 0.05.
- In the latest steps, the accuracy in the estimation of the available RT never exceeds a RE value of 0.1.
- The accuracy of the unattended RT estimation algorithm here implemented is comparable with the one that resorts to the fault detection module for the identification of an anomalous transient.

- Also, it is expected that, by avoiding the fault detection task, the RT estimation procedure is made free of possible errors in the fault detection module, so that the reliability of the whole operator support system for emergency accident management is increased.

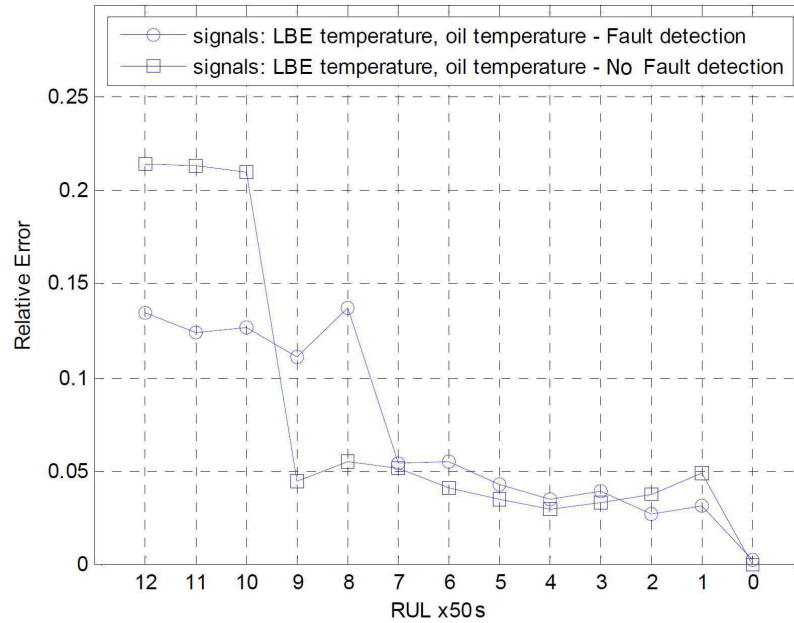


Figure 10 Relative Error evaluated each 50 s starting from 600 s before failure, for 1280 test trajectories, using $Z=2$ signals: in one case the fault detection module is activated [Zio et al., 2009b] (circles), in the other case the fault detection module is not implemented (squares)

5. Conclusions

This paper extends a fuzzy similarity-analysis method for estimating the available Recovery Time (RT) during the evolution of a failure trajectory of a system.

The extension to the method is aimed at freeing it from the need of resorting to a fault detection module for the identification of the anomalous system behavior. This is expected to increase the reliability of the method by avoiding potential errors in the detection module. The method of RT estimation relies on data from different transient failure scenarios to create a library of reference patterns of evolution; for estimating the available RT of a test pattern, its evolution data are, then, matched to the patterns in the library and their known residual life times are used for the estimation, based on a multidimensional fuzzy pointwise similarity analysis procedure.

The RT estimation procedure involves six main steps: (1) monitoring the signal evolution; (2) computing the pointwise difference between the test pattern evolution monitored and the available reference patterns; (3) evaluating their multidimensional fuzzy pointwise similarity and distance score; (4) defining the weights of the individual RT estimates provided by the fuzzy similarity of the reference patterns; (5) aggregating these to evaluate the system residual RT; and (6) setting to the MTTF value the RT estimation, when the system is identified as working in stationary conditions.

Application to a literature case study concerning the RT estimation for fault scenarios in the Lead Bismuth Eutectic eXperimental Accelerator Driven System (LBE-XADS) has given satisfactory results from the point of view of both accuracy of the estimation and computing time. The increased computing time required by matching the trajectories at each time step, and not just after the fault detection module has indicated a fault, is expected to be repaid by the increased reliability of the RT estimation system. Reliability and computational time are two objectives which can be compromised case-by-case through a proper selection of the number of monitored signals upon which to base the pointwise similarity evaluation and with the use or not of a fault detection module.

Acknowledgements

This work was supported in part by Ordine degli Ingegneri di Milano and by the Science & Technology Fellowship Programme in China financed by the European Commission (Reference: EuropeAid/127024/L/ACT/CN).

References

[Angstenberger, 2001] Angstenberger, L., Dynamic Fuzzy Pattern Recognition, International Series in Intelligent Technologies, 17, Kluwer Academic Publishers, 2001.

- [Barlett et al., 1992] Barlett, E.B., Uhrig, R.E., Nuclear Power Plant Status Diagnostics Using an Artificial Neural Network, *Nuclear Technology*, 97, 1992.
- [Binaghi et al., 1993] Binaghi, E., Della Ventura, A., Rampini, A., Schettini, R. (1993) Fuzzy Reasoning Approach to Similarity Evaluation in Image Analysis. *International Journal of Intelligent Systems*, Vol. 8, 1993, p. 749-769, 1993.
- [Berglund et al., 1995] Berglund et.al, The use of CAMS during a safety exercise at the Swedish Nuclear Inspectorate, HWR-423. OECD Halden Reactor project –July 1995
- [Bowman et al., 1992] Bowman, C.D., Arthur, E.D., Lisowski, P.W., Lawrence, G.P., Jensen, R.J., Anderson, J.L., Blind, B., Cappiello, M., Davidson, J.W., England, T.R., Engel, L.N., Haight, R.C., Hughes, H.G., Ireland, J.R., Krakowski, R.A., LaBauve, R.J., Letellier, B.C., Perry, R.T., Russell, G.J., Staudhammer, K.P., Versamis, G., Wilson, W.B., Nuclear energy generation and waste transmutation using an accelerator-driven intense thermal neutron source. *Nucl. Instr. Meth. Phys. Res. A* 320, 336, 1992.
- [Cammi et al., 2006] Cammi, A., Luzzi, L., Porta, A. A. and Ricotti, M. E., Modelling and control strategy of the Italian LBE-XADS, *Progress in Nuclear Energy*, Volume 48, Issue 6, Pages 578-589, 2006.
- [Campolucci et al., 1999] Campolucci, P., Uncini, A., Piazza, F. and Rao, B.D., On-Line Learning Algorithms of Locally Recurrent Neural Networks, *IEEE Trans. Neural Networks*, 10, pp. 253-271, 1999.
- [Carminati et al., 1993] Carminati, F., Klapisch, R., Revol, J.P., Roche, Ch., Rubio, J.A., Rubbia, C., An Energy Amplifier for Cleaner and Inexhaustible Nuclear Energy Production Driven by a Particle Beam Accelerator. CERN Report CERN-AT-93-47(ET), 1993.
- [Chiang et al., 2001] Chiang L. H., Russel, E., Braatz R., Fault detection and diagnosis in industrial systems, Springer-Verlag, London, 2001.
- [Di Maio et al., 2009] F. Di Maio, M. Stasi, P. Secchi, S. Vantini, E. Zio, “Dynamic Fault Scenarios For Training A Fuzzy C-Means Clustering Fault Classifier”, *Proceedings of S.Co.2009 Conference- Complex Data Modeling and Computationally Intensive Statistical Methods for Estimation and Prediction*, Milano, Italy, September 2009.
- [Dubois et al., 1988] Dubois, D., Prade, H., Testemale, C., Weighted Fuzzy Pattern Matching. *Fuzzy Sets and Systems*, 28, 1988, p. 313-331, 1988.
- [EC, 1999] EC, EU research in Reactor safety: Achievements of the 4th and prospects for the 5Th Euratom framework programme. Eurocourse 1999. GRS, May 1999.
- [Glasstone et al., 1998] Glasstone, S, Sesonske, A., Nuclear reactor engineering. New Delhi: CBS Publishers and Distributors; 1998.
- [IAEA, 2003] IAEA, Application of simulation techniques for accident management training in nuclear power plants, ISBN: 9201039034, 2003
- [Joentgen et al., 1999] Joentgen, A., Mikenina, L., Weber, R., Zimmermann, H.-J., Dynamic Fuzzy Data Analysis based on Similarity between Functions. *Fuzzy Sets and Systems*, 105 (1), 1999, p. 81-90, 1999.
- [NEA, 1992] NEA, Severe Accident Management – Prevention and Mitigation., Paris, 1992
- [Øwre, 2001] Øwre, F., Role of the man–machine interface in accident management strategies, *Nuclear Engineering and Design* 209, 201–210, 2001
- [Peel et al., 2008] Peel, L., Data Driven Prognostics using a Kalman Filter Ensemble of Neural Network Models, *International Conference on Prognostics and Health Management*, 2008.
- [Rubbia et al., 1995] Rubbia, C., Rubio, J.A., Buono, S., Carminati, F., Fitier, N., Galvez, J., Gels, C., Kadi, Y., Klapisch, R., Mandrillon, P., Revol, J.P., Roche, Ch., Conceptual Design of a Fast Neutron Operated High Power Energy Amplifier. CERN Report CERN-AT-95-44(ET), 1995.

- [Santosh et al., 2009] Santosh, T.V., Srivastava, A., Sanyasi Rao, V.V.S., Gosh, A. K., Kushwaha, H.S., Diagnostic System for Identification of Accident Scenarios in Nuclear Power Plants using Artificial Neural Networks, *Reliability Engineering and System Safety*, 94, 759-762, 2009.
- [Serrano et al., 1999] Serrano et al., Development of an extension of the CAMS system to severe accident management, HWR-580. OECD Halden Reactor project –May 1999
- [USNRC, 1999] USNRC 1999, NUREG-0700, Rev.1 - Human-System Interface Design Review Guideline, USNRC, Washington, June 1996
- [Van Tuyle et al., 1993] Van Tuyle, G.J., Todosow, M., Geiger, M.J., Aronson, A.L., Takahashi, H., Accelerator-driven subcritical target concept for transmutation of nuclear wastes. *Nucl. Technol.* 101, 1, 1993.
- [Venneri et al., 1993] Venneri, F., Bowman, C.D., Jameson R., Accelerator-driven Transmutation of Waste (ATW) - A New Method for Reducing the Long-term Radioactivity of Commercial Nuclear Waste. Los Alamos Report LA-UR-93-752, 1993
- [Wang et al., 2008] Wang T., Yu, J., Siegel, D., Lee, J., A similarity based prognostic approach for Remaining Useful life estimation of Engineered Systems, *International Conference on Prognostics and Health Management*, 2008.
- [Wang et al., 2004] Wang, W.Q., Goldnaraghi, M.F., Ismail, F., Prognosis of Machine Health Condition using Neuro-Fuzzy Systems, *Mechanical Systems and Signal Processing*, 18, 813-831, 2004.
- [Yuan et al., 1997] Yuan, B., Klir, G., Data driven identification of key variables, In: Ruan, D. (Ed.), *Intelligent Hybrid Systems Fuzzy Logic, Neural Network, and Genetic Algorithms*. Kluwer Academic Publishers, pp.161–187, 1997.
- [Zadeh, 1965] Zadeh, L., Fuzzy sets, *Inform. Control* 8, 338–353, 1965.
- [Zio et al., 2005] Zio, E., Baraldi, P., Identification of nuclear transients via optimized fuzzy clustering, *Annals of Nuclear Energy*, 32, 1068–1080, 2005.
- [Zio et al., 2008] Zio, E., Pedroni, N., Broggi, M., Golea, L., Locally recurrent neural networks for nuclear dynamics modeling, *FLINS - The 8th International FLINS Conference on Computational Intelligence in Decision and Control*, 21-24 September 2008, Madrid, Spain, 2008.
- [Zio et al., 2009a] Zio, E., Di Maio, F., A Data Driven Fuzzy Approach for predicting the Remaining recovery Time in a Dynamic Fault Scenarios of a Nuclear System, *Reliability Engineering and System Safety*, RESS, 10.1016/j.ress.2009.08.001, 2009.
- [Zio et al., 2009b] Zio, E., Di Maio, F., Data-Driven on-line prediction of the Available Recovery Time in Nuclear Power Plant Failure Scenarios, submitted to *Journal of Risk Analysis*, under review process.

APPENDIX

The procedure of fuzzy similarity analysis and RT estimation follows according to six steps:

- *Step 1: signal monitoring.*

The Z -dimensional trajectory of signals $f(x_1, x_2, \dots, x_Z, t)$ is continuously monitored throughout the time horizon of observation T , starting from (discrete) time $t = 1$; at each discrete time t , its values are recorded and appended to the matrix of the values collected at the previous time steps. For reasons which will become clear in the following, the database containing the reference trajectory patterns is organized in a 3-D structure $\overline{R}_{[N \times k \times Z]}$, where $k = \frac{T}{n}$ and its generic element $r(i, j, z)$ is the projection on the z -th signal axis of $r(i, j)$ which is the j -th segment, of length n , of values of the i -th reference trajectory, $i = 1, 2, \dots, N$, $j = 1, 2, \dots, k$, $z = 1, 2, \dots, Z$, normalized in the range $[0.2, 0.8]$. For clarity's sake, in Figures 11 and 12 a 2-D reference trajectory and its partition into 15 elements are shown, respectively (i.e., $Z=2$ and $k=15$).

At each time, the algorithm for the estimation of the available RT matches the similarity of the developing signal trajectory to those in the reference library and combines their end times to provide an estimate of the available RT of the current trajectory.

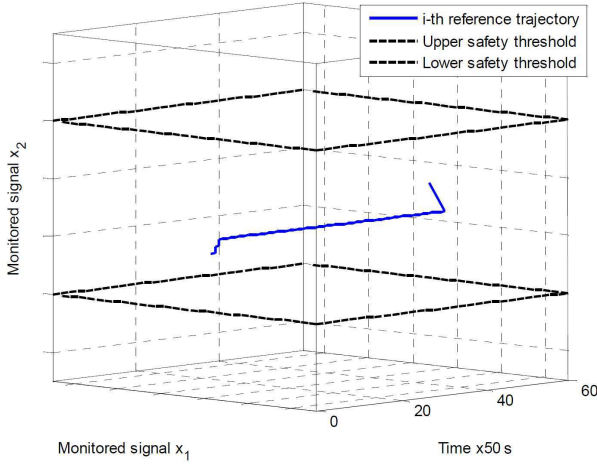


Figure 11 A bidimensional reference trajectory on a time horizon $T=3000$ s

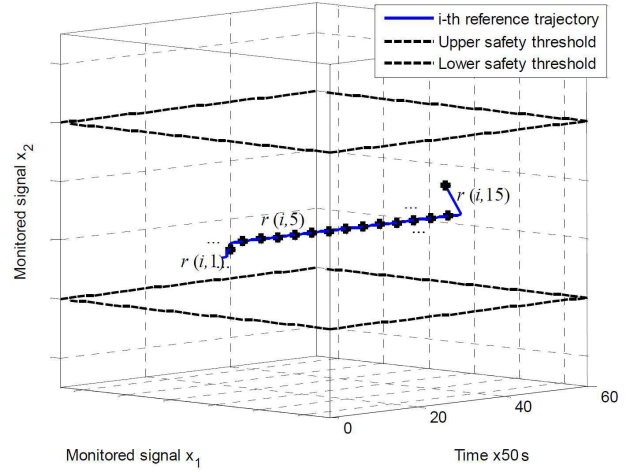


Figure 12 Bidimensional reference trajectory of Figure 11 partitioned into $k=15$ segments of length $n=3000/15=200$ s, $j=1,2,\dots,15$

- **Step 2: pointwise difference computation of multidimensional trajectory patterns.**

At the current time t , the latest n -long segment of values of the test trajectory pattern

$\tilde{f}(x_1, x_2, \dots, x_Z, t) = f(x_1, x_2, \dots, x_Z, t-n+1, t-n+2, \dots, t)$ is normalized in $[0.2, 0.8]$. The

pointwise difference $\delta(\cdot)$ between the $n \cdot Z$ values of pattern $\tilde{f}(x_1, x_2, \dots, x_Z, t)$ and those of the generic reference trajectory segment $r(i, j, z)$, is computed:

$$\delta(i, j) = \sum_{z=1}^Z (\tilde{f}(x_z, t) - r(i, j, z)), \quad i=1, 2, \dots, N, \quad j=1, 2, \dots, k, \quad z=1, 2, \dots, Z \quad (3)$$

The matrix $\overline{\delta}_{[N \times k]}$ contains the difference measures $\delta(i, j)$ between all n -long segments of the Z -dimensional reference trajectories and the test trajectory pattern of the monitored signals.

- **Step 3: computation of trajectory pointwise similarity and corresponding distance score.**

In general, there are numerous practical cases in which the classification of objects as ‘similar’ or ‘non similar’ is gradual; in these cases, then, the similarity measure should allow for a gradual transition [Binaghi et al., 1993; Joentgen et al., 1999]. In the case of interest here, this can be achieved by evaluating the pointwise difference of two trajectories with

reference to an “approximately zero” fuzzy set (FS) specified by a function which maps the elements $\delta(i, j)$ of the difference matrix $\overline{\delta}_{[N \times k]}$ into their values $\mu(i, j)$ of membership to the condition of “approximately zero”. The distance score $d(i, j)$ between two trajectory segments is then computed as:

$$d(i, j) = 1 - \mu(i, j), \quad i = 1, 2, \dots, N, \quad j = 1, 2, \dots, k \quad (4)$$

Common membership functions can be used for the definition of the FS, e.g. triangular, trapezoidal, and bell-shaped [Dubois et al., 1988]. In the application illustrated in this work, the following bell-shaped function is used:

$$\mu(i, j) = e^{-\left(\frac{-\ln(\alpha)}{\beta^2} \delta^2(i, j)\right)} \quad (5)$$

The arbitrary parameters α and β can be set by the analyst to shape the desired interpretation of similarity into the fuzzy set: the larger the value of the ratio $\frac{-\ln(\alpha)}{\beta^2}$, the narrower the fuzzy set and the stronger the definition of similarity [Zio et al., 2009a].

- **Step 4: weight definition.**

The $\widehat{RT}(t)$ is estimated as a similarity-weighted sum of the $RT_i(t)$:

$$\widehat{RT}(t) = \sum_{i|t_{f_i} > t} w_i \cdot RT_i(t), \quad i = 1, 2, \dots, N \quad (6)$$

The ideas behind the weighting of the individual $RT_i(t)$ is that: *i*) all failure trajectories in the reference library bring useful information for determining the available RT of the trajectory currently developing; *ii*) those segments of the reference trajectories which are most similar to the most recent segment of length n of the currently developing failure

trajectory should be more informative in the extrapolation of the occurring trajectory to failure.

To assign the weight w_i , the minimum distance d_i^* along the i^{th} row of the matrix of Eq. (4) is first identified:

$$d_i^* = \min_{j=1,\dots,k} d(i, j), \quad i = 1, 2, \dots, N \quad (7)$$

The weight w_i is then computed as:

$$w_i = (1 - d_i^*) \cdot e^{\left(\frac{1}{\beta d_i^*}\right)}, \quad i = 1, 2, \dots, N \quad (8)$$

and normalized:

$$w_i = w_i / \sum_{e=1}^N w_e \quad (9)$$

Note that the smaller the minimum distance, the larger the weight given to the i -th trajectory [Zio et al., 2009a].

- **Step 5: $RT_i(t)$ and $RT(t)$ estimation.**

Referring to the generic i -th trajectory in the library for which the system failure time $t_{f_i} > t$ (i.e., the time when the signal value exceeds the threshold beyond which the system loses its functionality), the value $RT_i(t)$ is determined as:

$$RT_i(t) = t_{f_i} - t_{j_M}, \quad i = 1, 2, \dots, N \quad (10)$$

where $t_{j_M} = n \cdot \max_j \left(\arg(\delta(i, j) = d_i^*) \right)$ is the final time index of the latest-in-life segment of the i -th trajectory among those with minimum distance d_i^* from the developing test

trajectory (n is the test trajectory pattern length and $\max_j \left(\arg \left(\delta(i, j) = d_i^* \right) \right)$ gives the largest column index j of $r(i, \bullet)$ whose element is equal to d_i^*). Thus, $RT_i(t)$ is the time available before reaching the failure threshold on the reference trajectory starting from the end time of the latest-in-life segment of minimum distance from the developing trajectory (Figure 13).

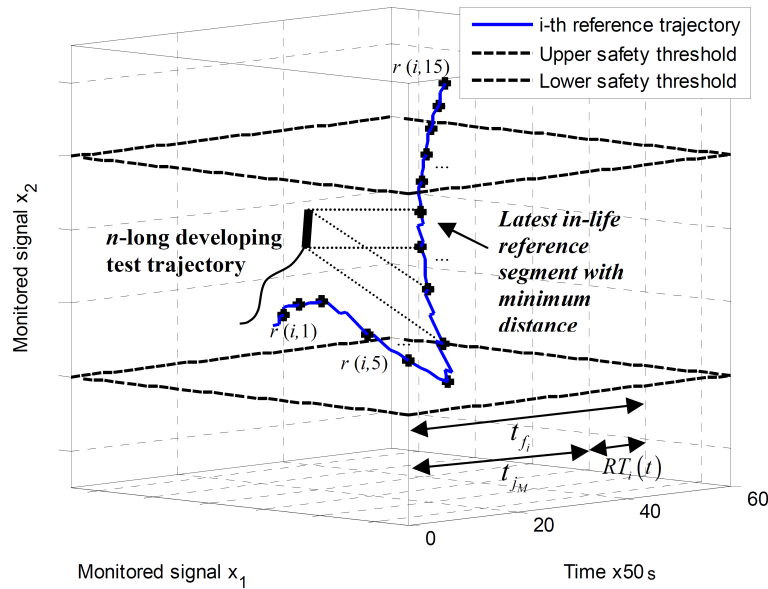


Figure 13 The $RT_i(t)$ for the i -th bidimensional reference trajectory starts from the end time of the latest-in-life segment of minimum distance from the developing trajectory

Then, the estimate $\widehat{RT}(t)$ of the remaining useful life along the developing trajectory is simply computed as in Eq. (6), with weights w_i evaluated by Eq. (9).

It has to be noted that the assumption $t_{j_M} = n \cdot \max_j \left(\arg \left(\delta(i, j) = d_i^* \right) \right)$ allows for a conservative RT estimation, biased towards “pessimistic” predictions of the available RT, because in the case that more than one segment along the i -th reference trajectory is closest to the developing test trajectory, the latest one is taken, i.e., the one closest to failure. This is even more conservative in case of stationary signals: Figure 2 in the main body of the paper

provides a heuristic explanation of this and the implemented modification for its improvement, which leads to the addition of step 6 to the procedure.

- ***Step 6: stationary signal identification and $RT(t)$ estimation as $MTTF(t)$.***

This is the key modification with respect to the previously presented procedure [Zio et al., 2009a]. The details are presented in Section 2 of the main body of the paper.