Hindawi Publishing Corporation
Science and Technology of Nuclear Installations
Volume 2015, Article ID 693891, 29 pages
http://dx.doi.org/10.1155/2015/693891

Research Article

Risk-Based Clustering for Near Misses Identification in Integrated Deterministic and Probabilistic Safety Analysis

Francesco Di Maio,[1] Matteo Vagnoli,[1] and Enrico Zio[1,2]

[1] *Energy Department, Politecnico di Milano, Via La Masa 34, 20156 Milano, Italy*
[2] *Chair on System Science and Energetic Challenge, Fondation EDF-Electricite de France, Ecole Centrale Paris and Supelec, Chatenay-Malabry Cedex, 92295 Paris, France*

Correspondence should be addressed to Francesco Di Maio; francesco.dimaio@polimi.it

Received 16 December 2014; Accepted 16 February 2015

Academic Editor: Arkady Serikov

In integrated deterministic and probabilistic safety analysis (IDPSA), safe scenarios and prime implicants (PIs) are generated by simulation. In this paper, we propose a novel postprocessing method, which resorts to a risk-based clustering method for identifying Near Misses among the safe scenarios. This is important because the possibility of recovering these combinations of failures within a tolerable grace time allows avoiding deviations to accident and, thus, reducing the downtime (and the risk) of the system. The postprocessing risk-significant features for the clustering are extracted from the following: (i) the probability of a scenario to develop into an accidental scenario, (ii) the severity of the consequences that the developing scenario would cause to the system, and (iii) the combination of (i) and (ii) into the overall risk of the developing scenario. The optimal selection of the extracted features is done by a wrapper approach, whereby a modified binary differential evolution (MBDE) embeds a K-means clustering algorithm. The characteristics of the Near Misses scenarios are identified solving a multiobjective optimization problem, using the Hamming distance as a measure of similarity. The feasibility of the analysis is shown with respect to fault scenarios in a dynamic steam generator (SG) of a nuclear power plant (NPP).

1. Introduction

Integrated deterministic and probabilistic safety analysis (IDPSA) attempts to overcome some limitations of deterministic safety analysis (DSA) and probabilistic safety analysis (PSA). The former is solidly founded on the multibarrier and defense-in-depth concepts and aims at verifying the capability of a nuclear power plant (NPP) to withstand a set of postulated design basis accidents (DBA) [1, 2]. To account for the uncertainties in the model representation of the actual plant behavior, conservatism is introduced in the calculations by thermal-hydraulics (TH) codes under DBA conditions [3]. The latter aims at considering a wider set of possible accidental scenarios and includes the quantification of accident probabilities [4, 5].

Both DSA and PSA are scenario-based analyses, where scenario selection and definition are done by expert judgment. State of the art of DSA and PSA approaches can provide relevant and important insights into what is already known to be an "issue," but they are not capable of revealing what, and to what extent, is not known (i.e., scenarios which are not expert-selected in the DSA and PSA inputs), with the risk of neglecting or underestimating potentially dangerous scenarios [6]. This is due to the difficulties of the static structure of the classic DSA and PSA approaches in treating dynamic variations that usually occur during the operational time of a process [7] due to (i) stochastic disturbances (e.g., equipment failures), (ii) deterministic plant responses (i.e., transients), (iii) controls, and (iv) operator actions [6, 8, 9]. Indeed, the order and timing of the events occurring along a scenario and the values of the process variables at the time of event occurrence are critical in determining the evolution of the scenario itself [10].

The development and application of IDPSA in practice must meet the challenge of computational complexity, in both model construction and implementation and in postprocessing for the retrieval of the relevant information from the scenario outcomes. The number of dynamic scenario branches generated in IDPSA increases in power law with the number of occurring events and, thus, is much larger than in classical PSA based on event trees (ET) and fault trees (FT). The a posteriori information retrieval (postprocessing) then becomes quite burdensome and difficult [11, 12]. Continuous event trees (CETs) [13, 14] and dynamic event trees (DETs) [15, 16] provide realistic frameworks for IDPSA. However, their application is limited by their computationally intensive nature, by the need of tailoring the algorithms to the system under consideration and by the need of processing a massive amount of data for any single initiating event considered [17].

Postprocessing, in general, consists in classifying the generated dynamic scenarios into safe scenarios and prime implicants (PIs), that is, sequences of events that represent minimal combinations of accident failures necessary for system failure and cannot be covered by more general implicants [18]. Among the safe scenarios, Near Misses are important scenarios to be identified, because they are those sequences of events that reach values of the safety parameters close to, but not exceeding, the corresponding acceptable thresholds [19]. They can, thus, be relevant contributors to the "hidden" risk of the system and should not be neglected, as a small deviation may transform them into accidental scenarios.

In the literature, several authors introduce the concept of Near Misses as accident precursors [20, 21]. We here consider Near Misses as sequences of events that incidentally keep the system in a safe state but endangered and insecure. For the purpose of the analysis, they are here defined as sequences of events similar to those leading the system into fault conditions, except for one characteristic which is missing or is slightly different (e.g., sequence time lag, different failure magnitude, and different involved component in an event) [22].

The postprocessing analysis entails a "Forward" classification of the dynamic scenarios into classes, that is, safe, PIs, and Near Misses and a "Backward" identification of the similarities of the features of the scenarios (i.e., stochastic event occurrence and deterministic process variables values), which characterize the groups of Near Misses among the whole set of safe scenarios.

For the "Forward" classification of the Near Misses sequences, we look at two factors of risk: the probability of occurrence of an undesired event and the severity of the consequence caused by the event [23]. Thus, we describe the sequences of events by (i) the probability (p) that the developing scenario is an accidental scenario, (ii) the consequence (c) that the developing scenario can cause to the system, and (iii) the overall risk (r) of the developing scenario that we compute synthetically as $r = p\times c$ (expected consequence).

The optimal features for discerning the Near Misses from the safe scenarios are extracted from the profiles of p, c, and r of the accidental scenarios and selected by a wrapper algorithm, which takes into account six statistical indicators of p, c, and r, and, through a modified binary differential evolution (MBDE) optimization algorithm, selects the best features, which are fed to a K-means clustering algorithm, which is a simple and well-known clustering algorithm (other classical clustering algorithms, such as mean-shift [24, 25] or fuzzy C-means [19, 26]).

The outcomes of this "Forward" classification is, then, interpreted by a "Backward" identification of the similarities of the features of the Near Misses scenarios: the acquired knowledge can be exploited in an online integrated risk monitoring system that can rapidly detect the problem and set up a repair strategy of the occurring failures before the system reaches a fault state.

The proposed approach is illustrated with reference to scenarios occurring in the steam generator (SG) of a NPP [27]. We use multiple-valued logic (MVL) theory for modeling the behavior of the system, where timing and sequences of component failure events are determining the system behavior [4]. By using MVL, we increase the limited description capability of binary variables in modeling the different component operational states (e.g., a valve that can be closed, partially closed, or fully open or can fail at different times) and, therefore, perform an IDPSA postprocessing analysis on the whole set of simulated accidental scenarios [17].

The paper is organized as follows. In Section 2, the SG model used to generate the scenarios for the reliability analysis is presented [27], along with multistate representation of the system dynamics. In Section 3, the PIs are identified and the risk-based "Forward" and "Backward" Near Misses identification method is introduced with reference to the case study considered. In Section 4, conclusions and remarks are drawn.

2. Case Study

2.1. The U-Tube Steam Generator (UTSG) Model. The U-tube steam generator (UTSG) under consideration is sketched in Figure 1. The improper control of the water level, whose difficulties arise from nonminimum phase plant characteristics, that is, plant strong inverse response behavior, particularly at low operating power, due to the so-called "swell and shrink" effects [28], is a major cause of NPP unavailability [28–30].

The reactor coolant enters the UTSG at the bottom and moves upward and then downward in the inverted U-tubes, transferring heat to the secondary fluid before exiting at the bottom. The secondary fluid, the feedwater (Q_e), enters the UTSG at the top of the downcomer, through the space between the tube bundle wrapper and the SG shell. The value of Q_e is regulated by a system of valves: a low flow rate valve, used when the operating power (P_o) is smaller than 15% of nominal power (P_n), and a high flow rate valve when $P_o > 0.15P_n$ [27]. In the secondary side of the tube bundle, water heats up, reaches saturation, starts boiling, and turns into a two-phase mixture. The two-phase fluid moves up through the separator/riser section, where steam is separated from liquid water, and through the dryers, which ensure that the exiting steam (Q_v) is essentially dry. The separated water is recirculated back to the downcomer. The balance between the exiting Q_v and the incoming Q_e governs the change in the water level in the SG. Because of the two-phase nature, two types of water level measurements are considered, as shown in Figure 1, each reflecting a different level concept: the narrow

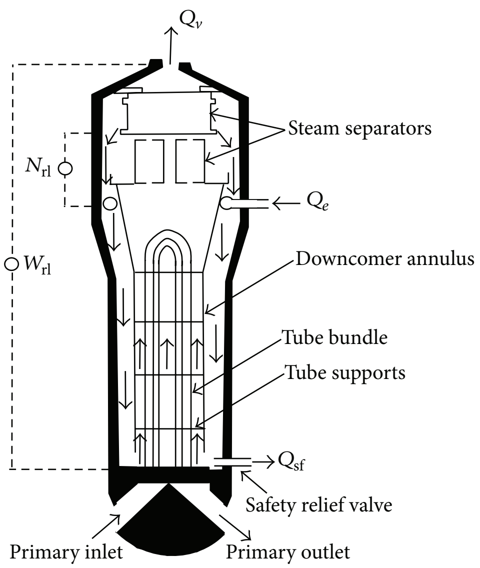


Figure 1: Schematic of the UTSG [31].

range level (N_{rl}) is calculated by pressure difference between two points close to the water level and indicates the mixture level, whereas the wide range level (W_{rl}) is calculated by pressure difference between the two extremities of the SG (steam dome and bottom of the downcomer) and indicates the collapsed liquid level that is related to the mass of water in the SG.

"Swell and shrink" phenomena are also modeled to reproduce the dynamic behavior of the SG: when Q_v increases, the steam pressure in the steam dome decreases and the two-phase fluid in the tube bundle expands causing N_{rl} to initially swell (i.e., rise), instead of decreasing as would have been expected by the mass balance; contrarily, if Q_v decreases or Q_e increases, a shrink effect occurs. A similar model has been presented in [27].

The N_{rl} is governed by Q_e and Q_v across the tube bundle region of the SG as shown by the following transfer function:

$$N_{\text{rl}}(s) = \frac{1}{T_n s}\left(Q_{ef}(s) - Q_{GV}(s)\right), \tag{1}$$

where Q_{ef} is the flow rate of the incoming water in the tube bundle, (2), Q_{GV} is the equivalent steam-water mixture flow rate exiting the tube bundle region, (3), T_n is a time constant that accounts for the N_{rl} dynamics.

The incoming water flow rate Q_{ef} is proportional to Q_e:

$$Q_{ef}(s) = \frac{1}{(1 + T_h s)(1 + \tau s)} Q_e(s), \tag{2}$$

where the lag $1/(1 + \tau s)$ accounts for the feed-water valve dynamics and $1/(1 + T_h s)$ accounts for the water mass transportation dynamics: their values are reported in Table 1.

The exiting steam-water mass Q_{GV} is proportional to Q_v:

$$Q_{GV}(s) = \frac{\left(1 - F_g T_g s\right)}{\left(1 + T_g s\right)} Q_v(s), \tag{3}$$

where the first-order lag $1/(1 + T_g s)$ accounts for the elapsed time from the turbine steam demand and the increase of Q_{GV}, and the nonminimum phase term $(1 - F_g T_g s)$ accounts for the two-phase swell and shrink effects.

Combining (1), (2), and (3), N_{rl} is equal to

$$N_{\text{rl}}(s) = \frac{1}{T_n s}\left(\frac{Q_e(s)}{(1 + T_h s)(1 + \tau s)} - \frac{\left(1 - F_g T_g s\right)}{\left(1 + T_g s\right)} Q_v(s)\right) \tag{4}$$

and W_{rl}, that is, the overall water mass in the steam generator, is

$$W_{\text{rl}}(s) = \frac{1}{T_{\text{int}} s}\left(Q_e(s) - Q_v(s)\right), \tag{5}$$

where T_{int} is a time constant that accounts for the W_{rl} dynamics.

We assume $y_1 = N_{\text{rl}}$ and $y_2 = W_{\text{rl}}$, and $u = Q_e$ and $d = Q_v$; the state space representation of the SG model is, thus,

$$\begin{gathered}
\dot{x}(t) = \begin{pmatrix} 0 & 0 & 0 & \dfrac{1}{T_n} \\ 0 & -\dfrac{1}{T_h} & 0 & -\dfrac{1}{T_n} \\ 0 & 0 & -\dfrac{1}{T_g} & 0 \\ 0 & 0 & 0 & -\dfrac{1}{\tau} \end{pmatrix} \mathbf{x}(t) \\
+ \begin{pmatrix} 0 \\ 0 \\ 0 \\ \dfrac{1}{\tau} \end{pmatrix} \mathbf{u}(t) + \begin{pmatrix} -\dfrac{1}{T_n} \\ 0 \\ \dfrac{1 + F_g}{T_n} \\ 0 \end{pmatrix} \mathbf{d}(t), \\
y(t) = \begin{pmatrix} 1 & 1 & 1 & 0 \\ \dfrac{T_n}{T_{\text{int}}} & 0 & 0 & \dfrac{\tau}{T_{\text{int}}} \end{pmatrix} \mathbf{x}(t).
\end{gathered} \tag{6}$$

The values of the parameters T_h, T_n, F_g,τ, T_g, and T_{int} change depending on the power P_o, as shown in Table 1.

The goal of the system is to maintain the SG water level at a reference position (N_{ref}): the SG fails if the N_{rl} rises (falls) above (below) the threshold N_{high} (N_{low}), in which case automatic reactor or turbine trips are triggered. Indeed, if the N_{rl} exceeds N_{high}, the steam separator and dryer lose their functionality and excessive moisture is carried in Q_v, degrading the turbine blades profile and the turbine efficiency; if N_{rl} decreases below N_{low}, insufficient cooling capability of the primary fluid occurs. Similarly, the W_{rl} is relevant for the cooling capability of the primary circuit [28]. Prealarms are triggered when N_{rl} exceeds N_{hl} (N_{ll}) if a small deviation from N_{ref} occurs or when N_{rl} exceeds N_{vh} (N_{vl}), when the deviation is large. Set points of N_{ref} and of N_{rl} depend on P_o, as shown in Figure 2, and, thus, also the alarms thresholds depend on P_o. The N_{rl} set point is low at low P_o, to

Table 1: Parameters of the UTSG model at different power levels [27].

P_o	$0.03 \times P_n$	$0.04 \times P_n$	$0.09 \times P_n$	$0.24 \times P_n$	$0.30 \times P_n$	$0.50 \times P_n$	P_n
T_n	36	56	63	44	40	40	40
F_g	13	18	10	4	4	4	4
T_h	170	56	30	10	8	5	5
τ	10	10	10	30	30	30	30
T_g	10	10	10	10	10	10	10
T_{int}	140	140	140	140	140	140	140

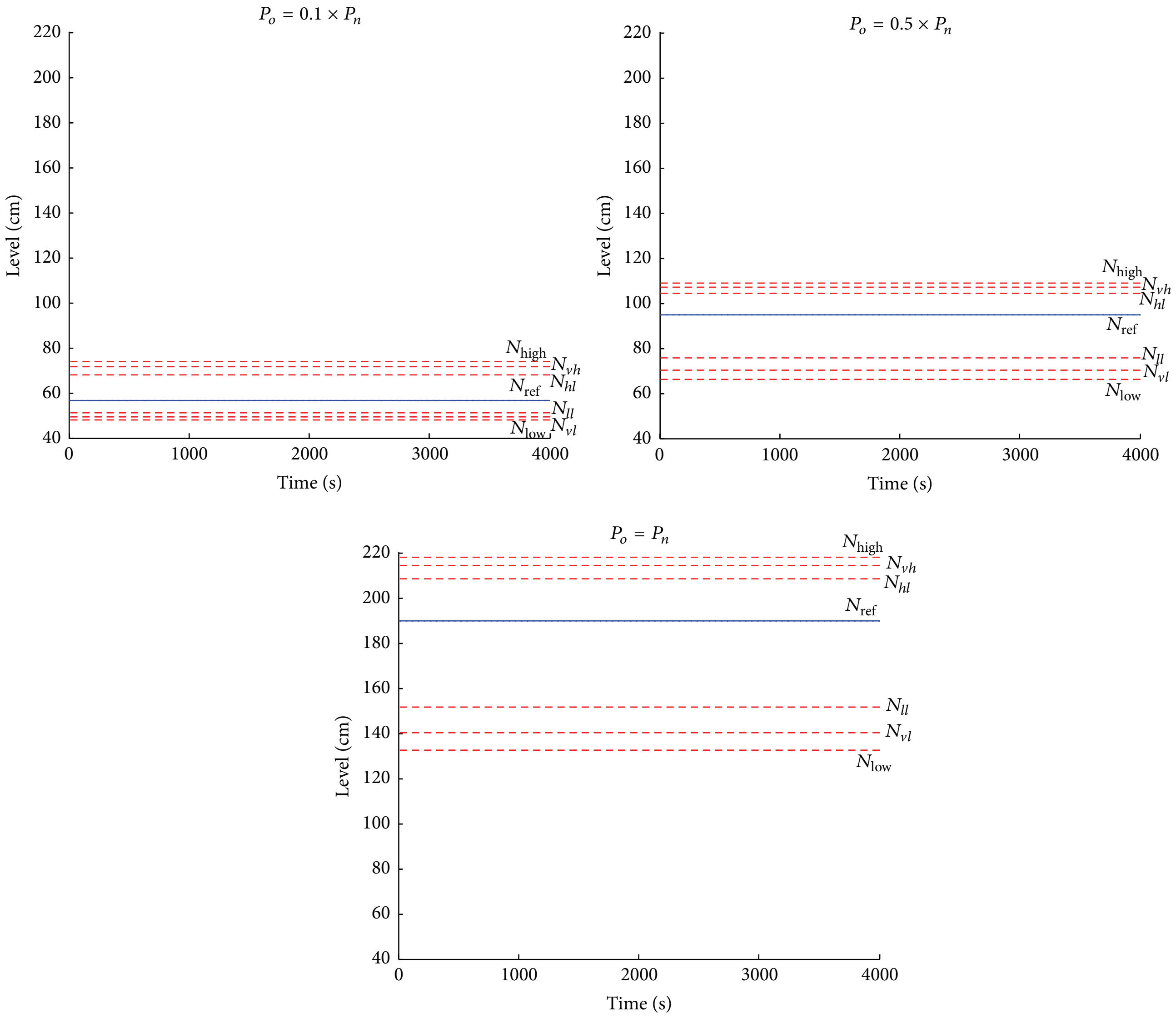


Figure 2: Set point for N_{rl} at different power rate P_o values.

partially account for the strong inverse response of N_{rl} [28]; thus, the low level thresholds are more restrictive than the high level thresholds at low P_o.

A dedicated model has been implemented in SIMULINK to simulate the dynamic response of the UTSG at different P_o values. Both feedforward and feedback digital control schemes have been adopted. The feedback controller is a PID that provides a flow rate Q_{pid} resulting from the residuals between N_{rl} and N_{ref}, whereas the feedforward controller operates a safety relief valve that is opened if and only if N_{rl} exceeds the N_{hl} and removes a constant flow safety flow rate (Q_{sf}). The block diagram representing the SIMULINK model of the SG is shown in Figure 3: the controlled variable is N_{rl}, whereas the control variable is Q_e.

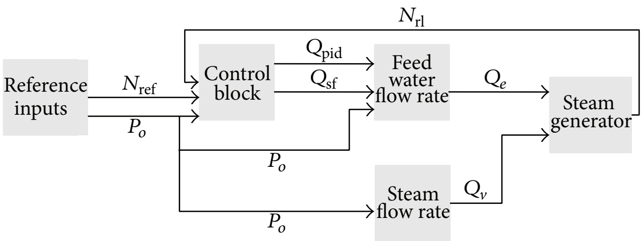


Figure 3: Block diagram representing the SIMULINK model of the SG.

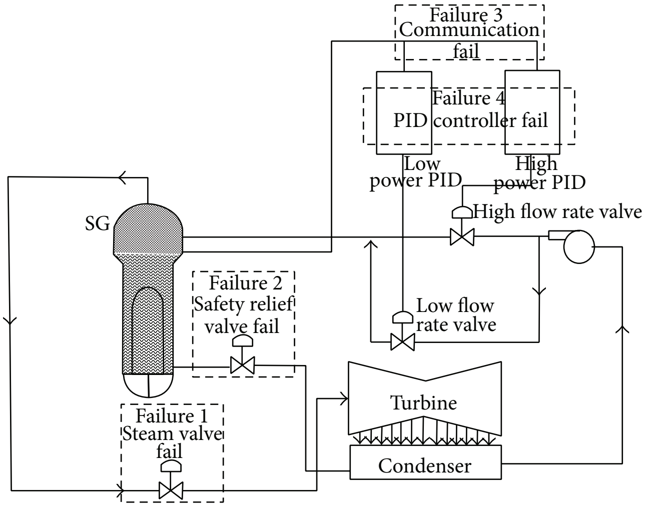


Figure 4: Sketch of the failures that can be injected into the system.

2.2. The Set of Possible Failures. The set of multiple component failures that can occur during the system life are shown in Figure 4.

(1) The outlet steam valve can fail stuck at a random time in $[0, 4000]$ (s) in three different positions: (i) closed; (ii) stuck open at 50% of the nominal Q_v that should be provided at P_o; (iii) stuck open at 150% of the nominal Q_v that should be provided at P_o.

(2) The safety relief valve can fail stuck at a random time in $[0, 4000]$ (s) at a uniform random value Q_{sf} in the range $[0.5, 50.5]$ (kg/s).

(3) The communication between the sensor that monitors N_{rl} and the PID controller can fail at random times in $[0, 4000]$ (s), in which case the PID is provided with the same input value of the previous time step.

(4) The PID controller can fail stuck at random times in $[0, 4000]$ (s), providing a uniform random flow rate Q_{pid} belonging to $[-18, 18]$% of the nominal Q_e that should be provided at P_o.

It is worth noticing that in the UTSG there are two PID controllers and, thus, two communications between the sensors measuring N_{rl} and the PIDs (one for high power feedback control and the other for low power feedback control). The selective action of the PIDs depending on P_o hides some of the failures. For example, if the power profile of the scenario under investigation is a ramp, both PIDs are called in operation: if anyone (or both) failed, their fault state is detectable. On the contrary, if we consider scenarios with constant power profile, for example, low power rate ($P_o < 15\%\ P_n$), the occurrence of a high power feedback control failure cannot be detected, and, thus, the fault remains hidden.

Choices and hypotheses for modeling the failures (i.e., the mission time, the number and type of faults, the distributions of failure times, and magnitudes) have been arbitrarily made with the aim of generating multiple failures in the sequences and capturing the dynamic influence of their order, timing, and magnitude. The choice of a mission time (T_{miss}) equal to 4000 (s) has been made, because it is a long enough interval of time to allow the complete development also of slow dynamic accident scenarios.

2.3. The Multistate Representation of System Dynamics. For realistically treating the dynamic behavior of the UTSG when component failures occur, we go beyond the binary state representation and adopt a multiple value logic (MVL) [17, 32] for an approximated description of the continuous time of occurrence of component failures and their magnitude. The MVL allows describing that the components can fail at any (discrete) time (not only the initial time) along the scenario, with different (discrete) magnitudes (not only the most conservative). The discretization of the time and magnitudes values is as follows:

(i) time discretization: we use the labels $t_{\mathrm{mvl}} = 1$, $t_{\mathrm{mvl}} = 2$, $t_{\mathrm{mvl}} = 3$, and $t_{\mathrm{mvl}} = 4$, for failures occurring in the intervals $[0, 1000]$ (s), $[1001, 2000]$ (s), $[2001, 3000]$ (s), and $[3001, 4000]$ (s), respectively; if the label $t_{\mathrm{mvl}} = 0$, the component does not fail within the time of the whole scenario, T_{miss};

(ii) magnitude discretization:

(a) the steam valve magnitude is indicated as 1, 2, or 3 for failure states corresponding to stuck at 0%, stuck at 50%, and stuck at 150% of the Q_e value that should be provided at P_o, respectively; if the steam valve magnitude is indicated as 0, the component does not fail in T_{miss};

(b) the safety relief valve fails with magnitude indicated as 1, 2, 3, and 4, if it is stuck between $[0.5, 12.6]$ (kg/s), $(12.6, 25.27]$ (kg/s), $(25.27, 37.91]$ (kg/s), and $(37.91, 50.5]$ (kg/s), respectively; if the safety relief valve magnitude is indicated as 0, the component does not fail in T_{miss};

(c) the communication between the sensor measuring N_{rl} and the PID controller is labelled 0 if the communication works, 1 otherwise;

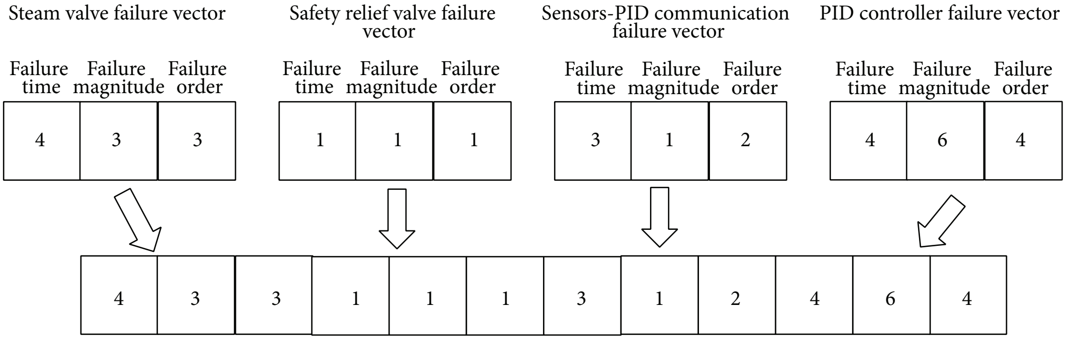

Steam valve failure vector
Safety relief valve failure vector
Sensors-PID communication failure vector
PID controller failure vector
Failure time
Failure magnitude
Failure order
4 3 3
1 1 1
3 1 2
4 6 4
4 3 3 1 1 1 3 1 2 4 6 4

Figure 5: Sequence vector representing a scenario.

Table 2: System configurations.

System configurations	Failure of the outlet steam valve	Failure of the safety relief valve	Level sensor-PID controller communication interruption	Failure of the PID controller
1	—	—	—	—
2	X	—	—	—
3	—	X	—	—
4	—	—	X	—
5	—	—	—	X
6	X	X	—	—
7	X	—	X	—
8	X	—	—	X
9	—	X	X	—
10	—	X	—	X
11	—	—	X	X
12	X	X	X	—
13	X	X	—	X
14	X	—	X	X
15	—	X	X	X
16	X	X	X	X

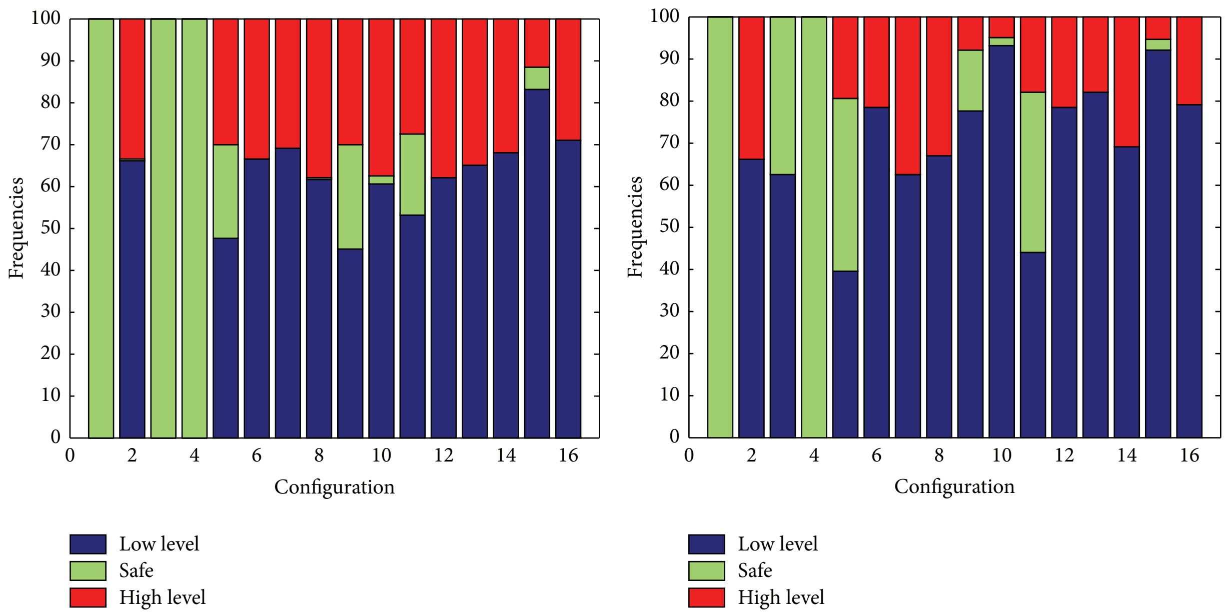
(d) the PID controller failure magnitude range is discretized into 8 equally spaced magnitude intervals, labelled from 1 to 8, representative of failure states corresponding to discrete intervals of output value belonging to $[-18, 18]$% of the Q_e value that should be provided at P_o; if the PID controller magnitude is labelled as 0, the component does not fail in T_{miss}.

The values of time and magnitude and order of failure occurrence for each component are included into a sequence vector that represents a scenario. As an example, the sequence vector of Figure 5 represents a scenario where the steam valve fails stuck at its maximum allowable value at a time in $[3001, 4000]$ (s) and it is the third event occurring along the sequence; the safety relief valve fails first in $[0, 1000]$ (s), with a magnitude belonging to $[0.5, 12.6]$ (kg/s); the communication between the sensor measuring N_{rl} and the PID controller is the second failure event in the sequence and occurs in $[2001, 3000]$ (s); finally, the PID controller fails stuck in $[3001, 4000]$ (s), with a magnitude belonging to $[6, 10]$% of the Q_e value that should be provided at P_o.

The number of possible sequence vectors that arise from the MVL discretization is 100509, each one evolving towards either safe or faulty conditions. To investigate this, a Monte Carlo-driven fault injection engine is used to sample combinations of discrete times and discrete magnitudes of components failures.

The (dynamic) analysis has been performed with respect to the two constant power scenarios, 5% P_n (low power level) and 80% P_n (high power level). The system configurations considered are listed in Table 2.

The dynamic analysis shows that the same combination of components failures does not unequivocally lead to only one system end state but rather it depends on when the failures occur and with what magnitude. This is shown in Figure 6, where the frequencies of occurrence of the three system end states (“High,” “Safe,” and “Low”) are plotted for the 16 dynamic system configurations of Table 2.



(a)

(b)

Figure 6: Histograms for high power level (a) and low power level (b) of the frequencies of occurrence of the end states for each of the 16 system configurations of Table 2, simulated by sampling discrete failure times and magnitudes of components failures.

Figure 7 shows that, at high power operation, the timing of the events is quite important, because with the same system configuration but different times of failure occurrences, the system end state change. Specifically, in Figure 7(a), the safety valve fails stuck at 100% of Q_{sf} after 1020 seconds and the communication between the sensor measuring N_{rl} and the PID controller fails at

(i) 1052 seconds (solid line),

(ii) 1063 seconds (dashed-dotted line).

The two scenarios lead to low and high failure modes, respectively, whereas they would be considered as minimal cuts sets (MCS) in a static reliability analysis presented in Appendix A.

Figure 7(b) shows the effects of different failures magnitudes on the system end state: the safety relief valve fails stuck in its maximum position at 2000 seconds, the communication fails at 2010 seconds, and the PID controller fails at 2020 seconds with two different magnitudes:

(i) magnitude equal to 13% of the nominal Q_e value that should be provided at 80% P_n (dashed-dotted line),

(ii) magnitude equal to 12% of the nominal Q_e value that should be provided at 80% P_n (solid line).

The low power scenarios also present dynamic effects, as shown in Figure 8. In particular, Figure 8(a) shows the effects of the timing on the system end state: the safety relief valve fails stuck at $Q_{sf} = 50.5$ (kg/s) at 1005 (s) and the steam output valve fails stuck at 150% of the nominal Q_v value that should be provided at 5% P_n at

(i) 1046 seconds (dashed-dotted line),

(ii) 1047 seconds (solid line).

Figure 8(b) shows the effects of the order of components failure occurrence on the system end state: the safety relief valve fails stuck at $Q_{sf} = 50.5$ (kg/s) and the PID controller fails stuck at its minimum allowable value.

(i) The PID controller failure is the first failure event along the sequence of events (dashed-dotted line).

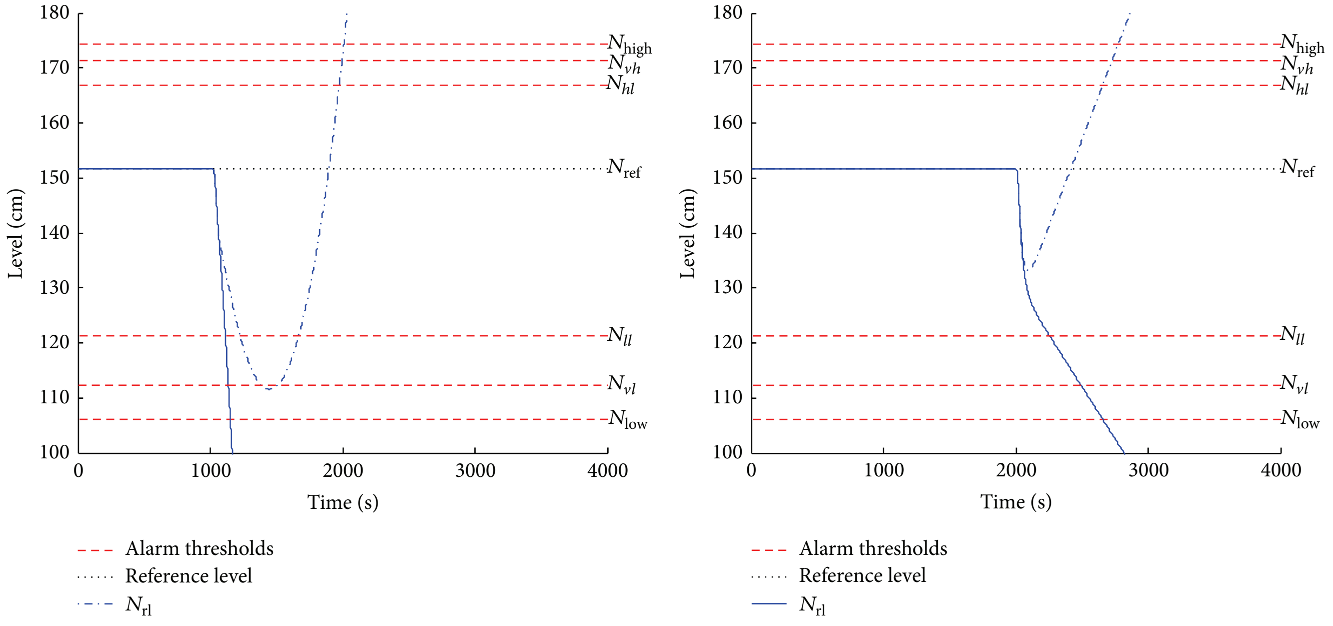
(ii) The safety relief valve failure is the first failure event along the sequence of events (solid line).

Hereafter, without loss of generality, among the system configurations of Table 2, we focus only on the classification of the PIs and Near Misses of the high level failure mode at high power level ($P_o = 80\%\ P_n$).

3. Near Misses Identification

The Near Misses identification is here treated as a classification problem, in which Near Misses are sorted out from the safe scenarios, among the whole set of accidental transients simulated. In practice, the PIs are first identified among the whole set of 100509 possible scenarios and, then, the Near Misses are separated out among the remaining safe scenarios.

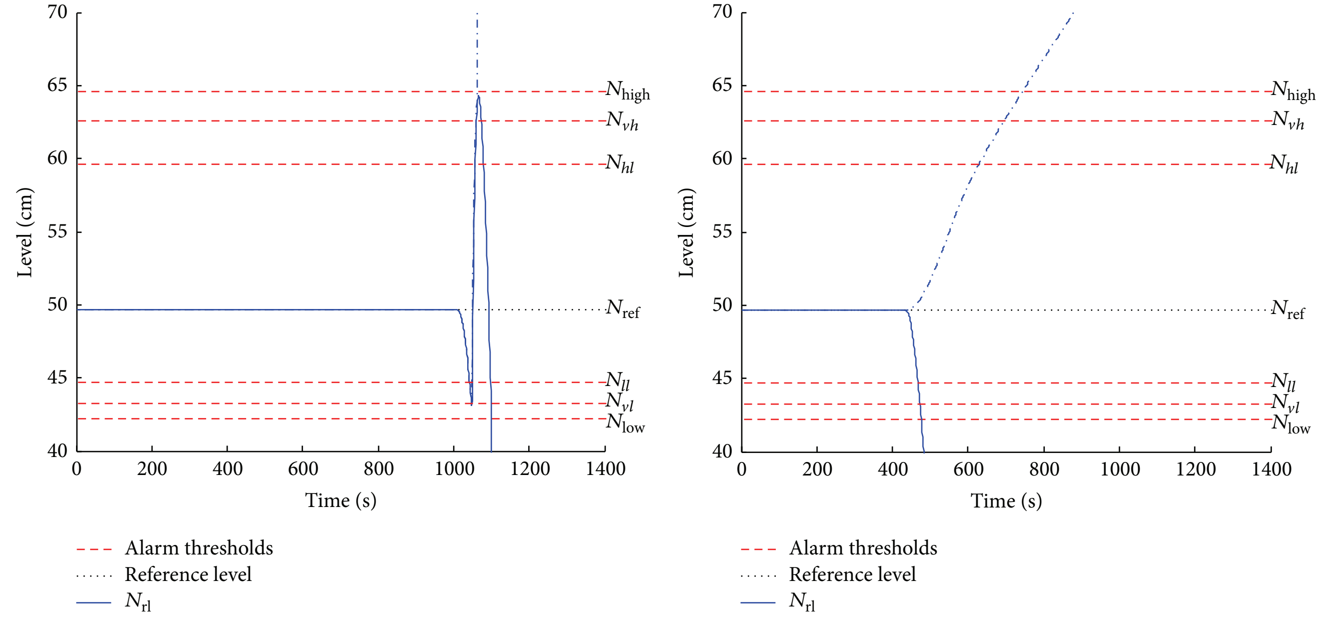
3.1. Prime Implicants Identification. A PI is a set of variables that represents a minimal combination of accident component failures necessary for system failure and cannot be covered by a more reduced implicant [17, 18]. Note that in our case the “PIs” identification task may consider noncoherent structure functions, for which both failed and working states of the same components can lead the system to failure. In such circumstances, traditional methods, for example, based on minimal cut sets analysis, cannot be applied, whereas



(a)

(b)

Figure 7: Example of dynamic system behavior at 80% P_n.



(a)

(b)

Figure 8: Example of dynamic system behavior at 5% P_n.

dynamic reliability methods need to be applied for the identification of the PIs [33, 34].

The PIs identification among the whole set of 100509 possible scenarios is performed by means of the visual interactive method presented in [34]. The basic idea it relies on is that PIs are those scenarios with as few as possible events that are capable of leading the system into a failure state [35]; then, we first select as most important feature for the PIs identification the literal cost of the sequence vector (i.e., the number of components whose behavior is specified in the accident sequence) and then the accident sequences associated with the lowest literal cost are selected and stored as PIs. In fact,

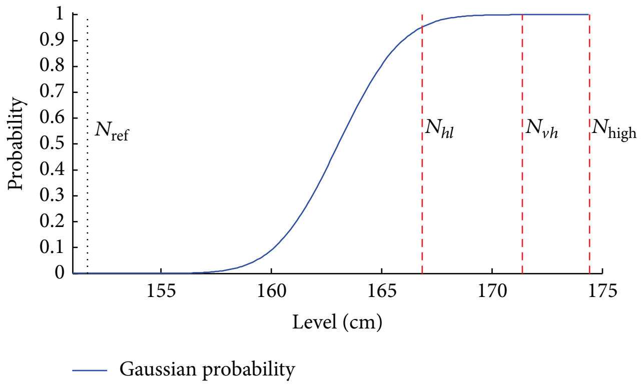


Figure 9: Probability function $p(t)$ for the definition of risk.

these are the most reduced sequences (i.e., with least number of events) that cannot be covered by any other implicant, and, thus, these are PIs by definition. The selected PIs and the implicants covered by them are deleted from the set of implicants and the procedure is repeated for the remaining implicants until all are covered. By so doing, 1255 PIs are identified for the high level failure mode, covering 36128 minterms. The total computational time approximately required for the identification of the PIs is 780 (s) on an Intel Core 2 Duo T9300 CPU @2.50 GHz.

3.2. The "Forward" Classification. Once the (1255) PIs for the SG high level failure mode have been identified, they are removed from the set of all possible scenarios, which is left with 64381 safe scenarios. For the identification of Near Misses among these, we resort to their definition as sequences of failure events that indeed keep the system in a safe condition but endangered (i.e., a quasifault system state). To this aim, we introduce a risk-based characterization of these remaining scenarios, calculating their associated risk, at each time instant t, as [23]

$$\mathrm{Risk}\,(t) = p\,(t) \times c\,(t)\,, \qquad (7)$$

where $p(t)$ is the probability that at time t the scenario can lead the system into an accidental scenario and $c(t)$ is the consequence that the developing scenario would cause to the system.

In this view, we build a functional relationship such that p increases as N_{rl} moves further away from the reference level N_{ref}, in a way that $p = 0$ if N_{rl} is equal to N_{ref} and $p = 1$ if N_{rl} reaches N_{high}. Such relationship is given in (8) below, assuming that scenarios whose $N_{\mathrm{rl}}(t)$ approaches N_{high} are more prone to failure than those with $N_{\mathrm{rl}}(t)$ close to N_{ref}; that is, (8) "filters out" (i.e., neglects) scenarios whose $N_{\mathrm{rl}}(t)$ is close to N_{ref} and "mines" (i.e., weighs more) scenarios whose $N_{\mathrm{rl}}(t)$ is close to N_{high}:

$$\begin{aligned} p\,(t) &= \varphi\left(\frac{N_{\mathrm{rl}}\,(t) - (\mu + 5\sigma)}{\sigma}\right) \\ &= \int_{N_{\mathrm{ref}}}^{N_{\mathrm{rl}}(t)} \frac{1}{\sqrt{2\pi}\sigma} e^{(N_{\mathrm{rl}}(t)-(\mu+5\sigma))^2/2\sigma^2}\, dN_{\mathrm{rl}}, \end{aligned} \qquad (8)$$

where φ is the cumulative probability function of the Gaussian distribution with mean $\mu = N_{\mathrm{ref}}$, and standard deviation $\sigma = (N_{\mathrm{high}} - N_{\mathrm{ref}})/5$. Figure 9 shows the trend of $p(t)$.

The consequence $c(t)$ of a scenario increases as N_{rl} approaches the failure threshold N_{high}, and $c(t)$ can be calculated at time t as [23]

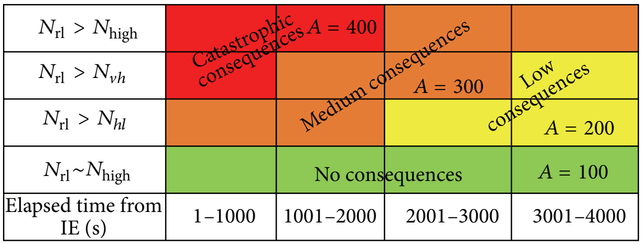
$$c\,(t) = A^{(\mathrm{NRL}(t)-(\mu+3\sigma))/(\mathrm{NRL}(t)-\mu)}, \qquad (9)$$

where A is the intensity coefficient that accounts for the closeness of N_{rl} to the thresholds N_{hl}, N_{vh}, and N_{high}, and for the exceedance time between the first event of the failure sequence (hereafter called initiating event (IE)) and the time of exceeding the threshold: the shorter this time, the more critical the scenario. Thus, the larger A is, the faster and closer N_{rl} approaches a threshold; we assume $A = 100$ (no consequences) if no threshold is exceeded; $A = 200$ (low consequences) if N_{rl} exceeds N_{hl} after at least 2001 (s) from IE or if N_{rl} exceeds N_{vh} after at least 3001 (s) from IE; $A = 300$ (medium consequences) if N_{rl} exceeds N_{hl} within 2000 (s) from IE, if N_{rl} exceeds N_{vh} and the elapsed time is in [1001, 3000] (s), and if N_{rl} exceeds N_{high} after at least 2001 (s) from IE; $A = 400$ (catastrophic consequences) if N_{rl} exceeds N_{vh} within 1000 (s) from IE or if N_{rl} exceeds N_{high} and the elapsed time from IE is in [1, 2000] (s). A matrix representation of the intensity coefficient is shown in Figure 10.

By so doing, the available 64381 remaining safe scenarios are fully described at each time instant $t = 1, 2, \ldots, 4000$ [s] by their values of probability $p(t)$, consequence $c(t)$, and overall risk $r(t)$. An example of the $p(t)$, $c(t)$, and $r(t)$ evolutions for two generic trends of $N_{\mathrm{rl}}(t)$ is shown in Figure 11. More specifically, the $N_{\mathrm{rl}}(t)$ behaviors represented in Figure 11(a) are due to

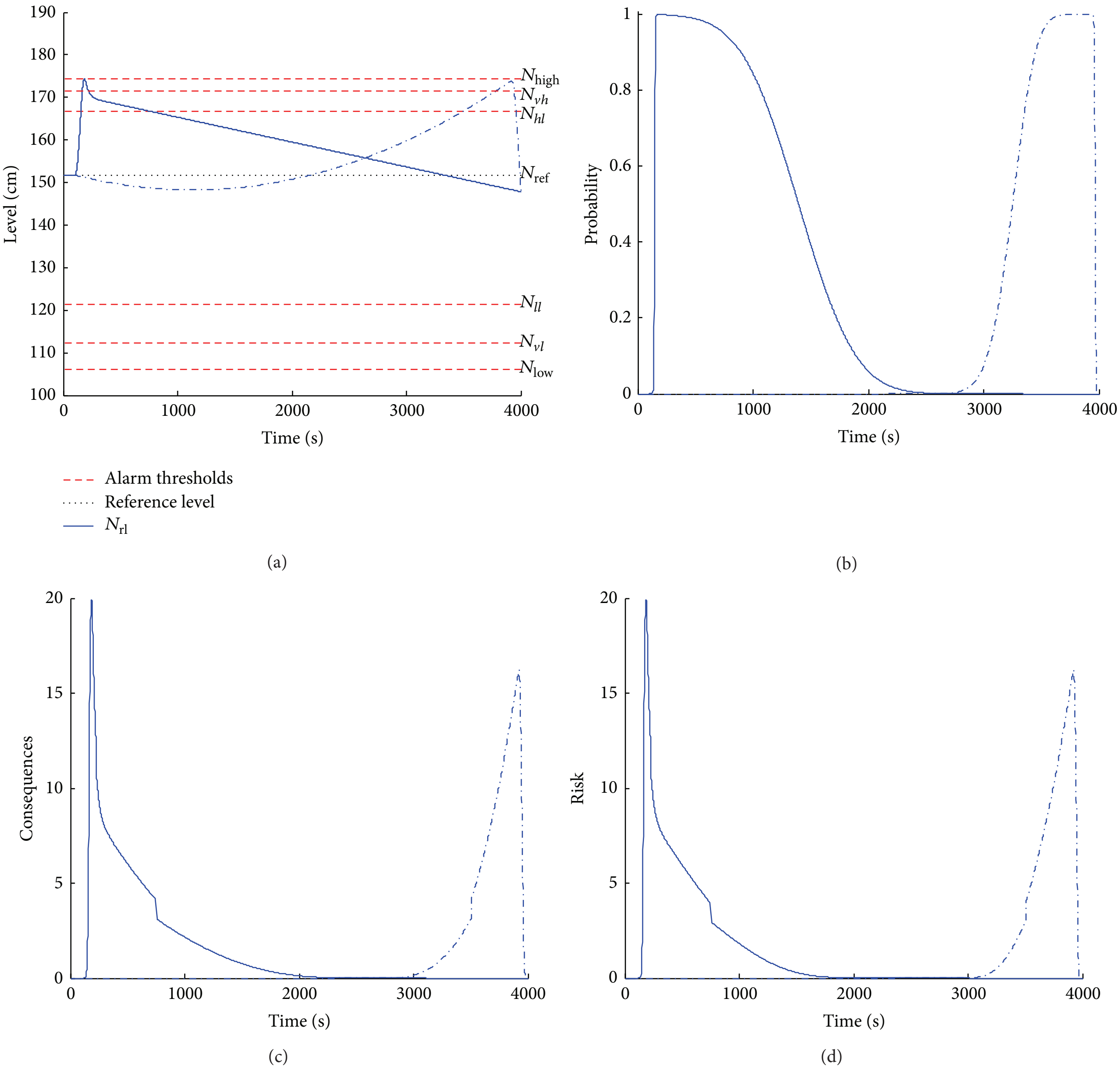
(i) solid line: the PID controller fails at 100 (s) with magnitude 4 and the safety relief valve fails at 190 (s) with magnitude 2;

(ii) dashed-dotted line: the safety relief valve fails at 100 (s) with magnitude 1, the communication between the sensor measuring N_{rl} and the PID controller is interrupted at 136 (s), and the PID controller fails at 3917 (s) with magnitude 5.



$N_{\mathrm{rl}} > N_{\mathrm{high}}$	🟥 Catastrophic consequences	🟥 $A = 400$	🟧 Medium consequences	🟧
$N_{\mathrm{rl}} > N_{vh}$	🟥	🟧	🟧 $A = 300$	🟨 Low consequences
$N_{\mathrm{rl}} > N_{hl}$	🟧	🟧	🟨	🟨 $A = 200$
$N_{\mathrm{rl}} \sim N_{\mathrm{high}}$	🟩	🟩 No consequences	🟩	🟩 $A = 100$
Elapsed time from IE (s)	1–1000	1001–2000	2001–3000	3001–4000

FIGURE 10: Matrix representation of the intensity coefficient A.



(a)

(b)

(c)

(d)

FIGURE 11: Probability $p(t)$, consequences $c(t)$, and risk $r(t)$ for two sequences of events.

It is worth analysing the behavior of $p(t)$, $c(t)$, and, thus, $r(t)$ (Figures 11(b), 11(c), and 11(d), resp.): all three above-mentioned functions increase as $N_{\mathrm{rl}}(t)$ moves further away from N_{ref} and decrease as $N_{\mathrm{rl}}(t)$ approaches N_{ref}. The steps shown in the consequences and risk plots (around 800 [s] for the solid line scenario and around 3500 [s] for the dashed-dotted line scenario) are due to the change of the discrete consequence intensity coefficient A along the scenarios. The solid line scenario is faster than the dashed-dotted line scenario (upper plot) and, thus, the value of the parameters A for the former scenario is 400 (catastrophic consequences, see Figure 10), due to the fact that $N_{\mathrm{rl}}(t)$ exceeds N_{vh} within 1000 (s), whereas, $A = 300$ (medium consequences, see Figure 10) for the dashed-dotted scenarios, because $N_{\mathrm{rl}}(t)$ exceeds N_{vh} within [1001, 3000] (s). Thus, the solid line scenario is more abrupt in its development towards failure and expected to have more catastrophic consequences, and, thus, more overall risk, than the dashed-dotted scenario, because the time between IE and the exceedance of N_{vh} is shorter (i.e, less grace time).

3.2.1. Features Selection. The identification of the Near Misses is treated as an unsupervised classification problem and addressed by clustering, where (i) the number of clusters is unknown and (ii) the features that enable the best clustering according to the risk-based characteristic profiles of $p(t)$, $c(t)$, and $r(t)$ of the accidental scenarios are unknown. Unsupervised clustering, thus, entails identifying the number K of clusters in which similar scenarios can be grouped according to similar values of some scenario features. To do this, from the profiles $p(t)$, $c(t)$, and $r(t)$, we extract some statistical indicators as features [36]:

(1) mean value:

$$\mu_n = \frac{1}{T_{\mathrm{miss}}}\sum_{t=1}^{N} n(t), \tag{10}$$

(2) peak value:

$$\max = \max_{t=1,2,\ldots,T_{\mathrm{miss}}} n(t), \tag{11}$$

(3) standard deviation:

$$\sigma_n = \sqrt{\frac{1}{T_{\mathrm{miss}}-1}\sum_{t=1}^{T_{\mathrm{miss}}} (n(t)-\mu)^2}, \tag{12}$$

(4) root mean square:

$$\mathrm{RMS} = \sqrt{\frac{1}{T_{\mathrm{miss}}}\sum_{t=1}^{T_{\mathrm{miss}}} (n(t))^2}, \tag{13}$$

(5) skewness:

$$\mathrm{SK} = \frac{\sum_{t=1}^{T_{\mathrm{miss}}} (n(t)-\mu)^3}{T_{\mathrm{miss}}-1}, \tag{14}$$

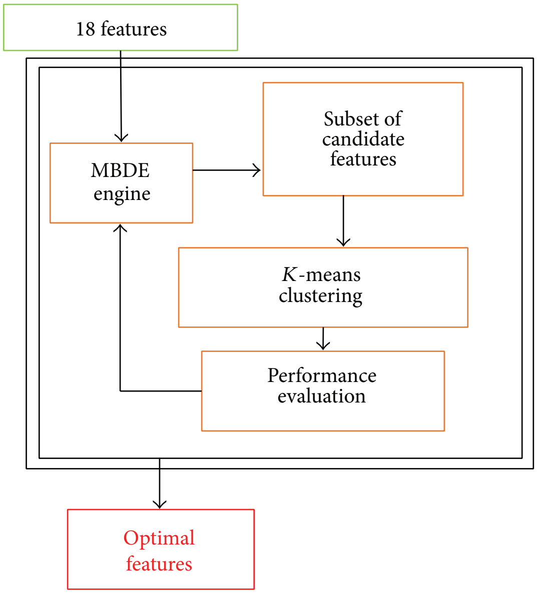


Figure 12: Wrapper approach for optimal feature subset selection based on a MBDE optimization algorithm and a K-means classifier.

(6) kurtosis:

$$\mathrm{KU} = \frac{\sum_{t=1}^{T_{\mathrm{miss}}} (n(t)-\mu)^4}{T_{\mathrm{miss}}-1}, \tag{15}$$

where $n(t)$ is alternatively equal to $p(t)$, $c(t)$, and $r(t)$ and, thus, the total number of features is equal to $6 \times 3 = 18$. Among these 18 available features, we search for those that are optimal for clustering the 64381 scenarios in Near Misses and safe scenarios.

We resort to a wrapper framework [37, 38], whereby a modified binary differential evolution (MBDE) search engine [33, 39] searches candidate groups of features sets that are fed to a K-means clustering algorithm [40]; eventually, the wrapper evolves so that among these candidate groups, the group retained is that which makes the K-means clustering algorithm perform best (most compact and separate clusters). The idea behind the wrapper approach is shown in Figure 12. During the features search by MBDE, the K-means clustering is run on the $N = 0.80\times64381 = 51505$ (training) safe scenarios with sets of features (F) that are randomly selected by the MBDE algorithm. The optimal number (K) of clusters is also unknown and it is determined by looking at the clustering performance obtained by the K-means with reference to the Calinski-Harabasz (CH) index [41], which accounts for the ratio of the overall between-cluster variance (separation) and the overall within-cluster variance (compactness). The search proceeds iteratively until the CH index is maximised and the number of clusters K is fixed. Then, the results of the wrapper algorithm are evaluated on an independent test set (T_{set}), that is, the $0.2 \times 643281 = 12876$ safe scenarios that have been left out during the training phase.

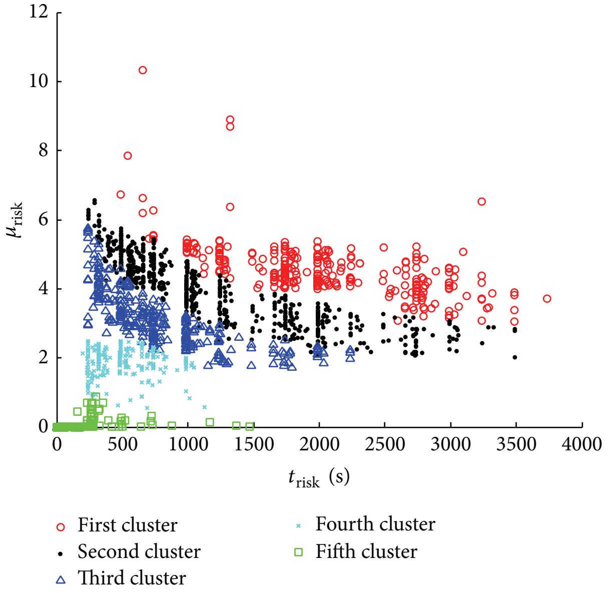


FIGURE 13: Clustering results.

The CH index for a number K of clusters, $k = 1, 2, \ldots, K$ is equal to [41]

$$\text{CH} = \frac{SS_b}{SS_w} \times \frac{(N \times F) - K}{K - 1}, \tag{16}$$

where SS_b is the overall between-cluster variance:

$$SS_b = \sum_{k=1}^{K} n_k \left\| m_k - m \right\|^2, \tag{17}$$

and SS_w is the overall within-cluster variance:

$$SS_w = \sum_{k=1}^{K} \sum_{x \in k} \left\| x_c - m_k \right\|^2, \tag{18}$$

where n_k is the number of scenarios x_c assigned to the kth cluster, m_k is the centroid of the kth cluster, that is, the mean of the selected features belonging to the kth cluster, m is the mean of the selected features, and $\|m_k - m\|^2$ and $\|x_c - m_k\|^2$ are the L^2 norms, that is, Euclidean distances, between the two vectors.

The optimal features selection provides as best features the standard deviation of $c(t)$, the standard deviation of $r(t)$, and the root mean square of $r(t)$; the best performance is obtained with CH = $9.35e + 04$ and $K = 5$.

3.2.2. The Clustering Results. The $K = 5$ obtained clusters of the safe scenarios are shown in Figure 13 with reference to the features of mean risk (μ_{risk}) and time elapsed from the instant t_{risk} at which $r(t)$ starts to deviate from zero, that is, the time interval during which the system is exposed to risk. The rationale behind this choice is that the larger μ_{risk} and the longer t_{risk}, the more dangerous the scenarios. In Figure 13, clusters 3, 4, and 5 (triangles, crosses, and squares, resp.) are well separated, that is, the low level risk scenarios clusters are widened by the adoption of (8) for the quantification of the risk profile $r(t)$. It is possible to distinguish the scenarios having the lowest risk level from the scenarios having low risk level, and, thus, the highest risk scenarios are well separated from the lower risk scenarios. The good performance obtained when (8) is adopted instead of other $p(t)$ profiles, for example, linear probability function ($p(t) \propto N_{\text{rl}}(t)$) that would give the same importance to any level N_{rl}, for the quantification of the risk profile $r(t)$, is due to the fact that (8) "filters out" (i.e., neglects) scenarios whose $N_{\text{rl}}(t)$ is close to N_{ref} and "mines" (i.e., weighs more) scenarios whose $N_{\text{rl}}(t)$ is close to N_{high}: the 332 circles in Figure 13 (listed in Appendix B) can, thus, be considered the Near Misses scenarios, that is, scenarios that incidentally keep the system into safe state, although in endangered and insecure operational conditions.

3.3. The "Backward" Approach. Once the Near Misses for the SG high level failure mode have been identified by clustering, we can search for similarities among them in terms of their multiple value sequences, that is, order and timing of event occurrences and deterministic process variables values. This "Backward" approach can lead us to finding the minimum conditions, that is, minimum μ_{risk} and minimum t_{risk}, that lead the system into a quasifault state. The problem can be framed as a multiobjective optimization problem (MOP) [42] that looks for the set of scenarios $\overline{\mathbf{x}}$ to dominate any other scenarios with respect to the fitness function f:

$$f(x) = \left[f_1(\mu_{\text{risk}}), f_2(t_{\text{risk}}) \right], \tag{19}$$

where f_1 and f_2 are the objectives functions of the defined MOP, that is, minimum μ_{risk} and t_{risk}, respectively. The solution of the MOP of (19) is the Pareto set shown in Figure 14, where 12 solutions are plotted (squares lined by continuous

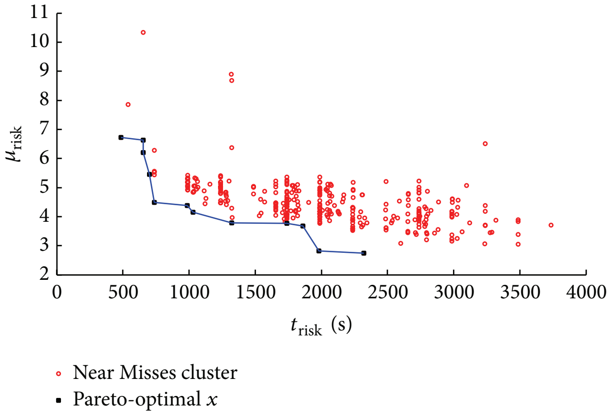
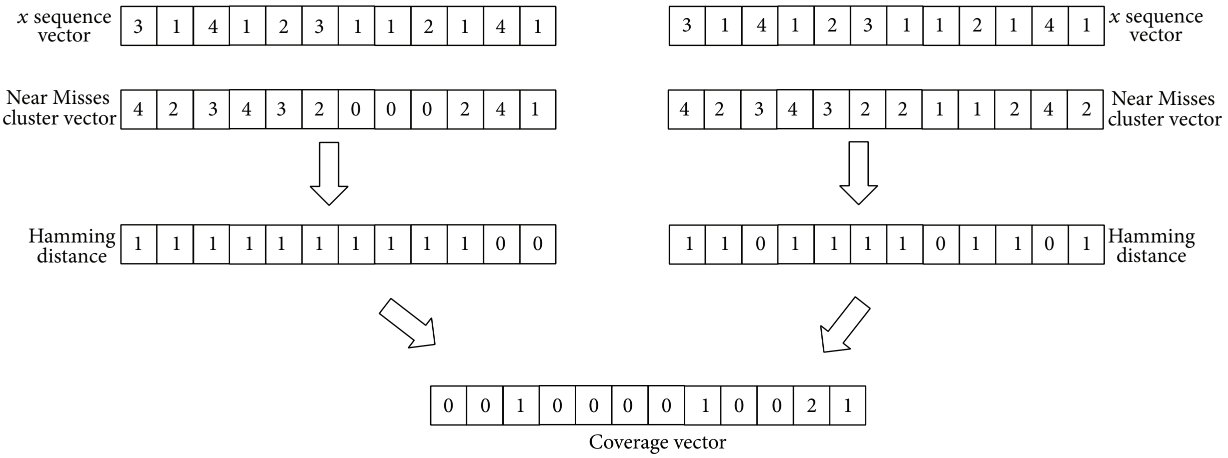


FIGURE 14: Pareto front for the cluster of Near Misses.


x sequence vector: 3 1 4 1 2 3 1 1 2 1 4 1 | 3 1 4 1 2 3 1 1 2 1 4 1

Near Misses cluster vector: 4 2 3 4 3 2 0 0 0 2 4 1 | 4 2 3 4 3 2 2 1 1 2 4 2

Hamming distance: 1 1 1 1 1 1 1 1 1 1 0 0 | 1 1 0 1 1 1 1 0 1 1 0 1

0 0 1 0 0 0 0 1 0 0 2 1

Coverage vector

FIGURE 15: Coverage vector computation by Hamming distance.

line) and listed in Table 3. These scenarios $\bar{\mathbf{x}}$ of minimum $f(x)$ are expected to cover all failure of Near Misses scenarios cluster.

The coverage can be verified by, first, identifying the most similar characteristics of the sequence vectors belonging to the Near Misses cluster with the Pareto set scenarios $\bar{\mathbf{x}}$, and, then, by solving a set covering problem (SCP) [43, 44].

The most similar characteristics can be computed by coverage vectors (one for each scenario belonging to $\bar{\mathbf{x}}$): this entails calculating the Hamming distance [45] between each sequence vectors in $\bar{\mathbf{x}}$ and each one of the other sequence vectors in the Near Misses cluster [46]. The entries of the coverage vector (in our case twelve entries, one for time, magnitude, and order of occurrence of each component failure; see Figure 5) are increased if the Hamming distance between one same entry of the considered scenario belonging to $\bar{\mathbf{x}}$ and of the Near Misses vectors is equal to zero, as shown, without loss of generality, in Figure 15 for 1 sequence vector of $\bar{\mathbf{x}}$ and only 2 Near Misses vectors.

Table 4 lists the 12 coverage vectors, where each entry is the percentage of Near Misses vectors having the same stochastic behavior of the optimal set $\bar{\mathbf{x}}$ shown in Table 3. It can be seen that, for each scenario belonging to $\bar{\mathbf{x}}$, columns 8, 11, and 12 (e.g., sensor-PID communication failure magnitude, PID failure magnitude, and PID order of failure, resp.) have the largest values of the coverage vectors: this means that the majority of the sequence vectors of the Near Misses clusters can be well represented by (only) these failures. Furthermore, the analysis of the MVL values of the scenarios belonging to $\bar{\mathbf{x}}$ (Table 3) where the largest coverage values of these columns are registered (i.e, 87%, 98.5%, and 85.2% for columns 8, 11, and 12, resp.) highlights that these failures are characterized by the same MVL values that can be summarized as follows:

(i) the failure of the communication between the sensor monitoring the N_{rl} and the PID controller;

(ii) the failure of the PID controller with magnitude belonging to $[-5, -1]\%$ of the Q_e value that should be provided at P_o, that is, magnitude equal to 4 in MVL framework, and it is the first accident occurring along the sequence of events in over 85% of the Near Misses scenarios.

A SCP can, thus, be solved for verifying that these latest characteristics are the minimum set of stochastic event

Table 3: List of the Pareto-optimal $\overline{\mathbf{x}}$ sequence vectors.

$\overline{\mathbf{x}}$	Steam valve failure time	Steam valve failure magnitude	Steam valve failure order	Safety valve failure time	Safety valve failure magnitude	Safety valve failure order	Sensor-PID failure time	Sensor-PID failure magnitude	Sensor-PID failure order	PID failure time	PID failure magnitude	PID failure order
1	0	0	0	2	3	1	4	1	2	4	3	3
2	3	1	4	1	2	3	1	1	2	1	4	1
3	3	1	4	2	2	3	2	1	1	2	4	2
4	3	1	3	2	2	2	4	1	4	1	4	1
5	3	1	4	3	3	3	2	1	2	1	4	1
6	3	2	4	1	2	2	1	1	3	1	4	1
7	3	2	4	2	3	1	2	1	2	3	3	3
8	4	1	3	3	2	2	4	1	4	2	4	1
9	4	2	3	1	3	2	0	0	0	1	3	1
10	4	2	4	2	1	3	2	1	1	2	4	2
11	4	2	3	4	2	2	0	0	0	4	4	1
12	4	2	4	4	2	3	2	1	1	4	4	2

Table 4: List of coverage vectors for each scenario belonging to the Pareto set $\overline{\mathbf{x}}$.

$\overline{\mathbf{x}}$	Steam valve failure time	Steam valve failure magnitude	Steam valve failure order	Safety valve failure time	Safety valve failure magnitude	Safety valve failure order	Sensor-PID failure time	Sensor-PID failure magnitude	Sensor-PID failure order	PID failure time	PID failure magnitude	PID failure order
1	11.4	11.4	11.4	8.1	30.4	0.6	25.3	**87**	46.4	1.2	1.5	0.6
2	30.1	44.3	27.1	1.2	19.9	38	19.6	**87**	46.4	68.1	**98.5**	**85.2**
3	30.1	44.3	27.1	8.1	19.9	38	22.3	**87**	14.2	27.4	**98.5**	14.2
4	30.1	44.3	42.2	8.1	19.9	19.9	25.3	**87**	12.7	68.1	**98.5**	**85.2**
5	30.1	44.3	27.1	28	30.4	38	22.3	**87**	46.4	68.1	**98.5**	**85.2**
6	30.1	36.1	27.1	1.2	19.9	19.9	19.6	**87**	13.9	68.1	**98.5**	**85.2**
7	30.1	36.1	27.1	8.1	30.4	0.6	22.3	**87**	46.4	3.3	1.5	0.60
8	52.4	44.3	42.2	28	19.9	19.9	25.3	**87**	12.7	27.4	**98.5**	**85.2**
9	52.4	36.1	42.2	1.2	30.4	19.9	13	13	13	68.1	1.5	**85.2**
10	52.4	36.1	27.1	8.1	9.04	38	22.3	**87**	14.2	27.4	**98.5**	14.2
11	52.4	36.1	42.2	51.2	19.9	19.9	13	13	13	1.2	**98.5**	**85.2**
12	52.4	36.1	27.1	51.2	19.9	38	22.3	**87**	14.2	1.2	**98.5**	14.2

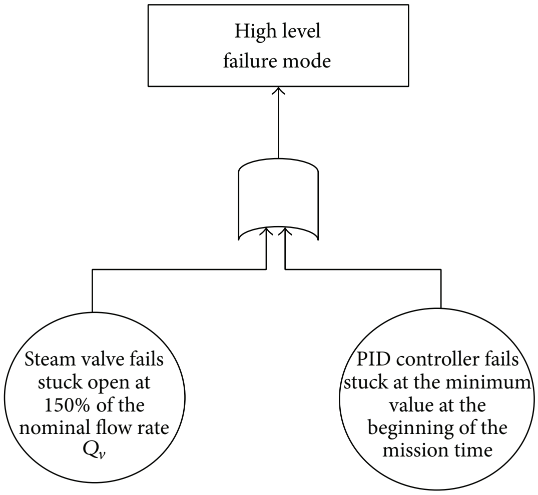

High level failure mode
Steam valve fails stuck open at 150% of the nominal flow rate Q_v
PID controller fails stuck at the minimum value at the beginning of the mission time

FIGURE 16: Fault tree for the high level failure mode.

TABLE 5: Possible system configurations to be considered in the static reliability analysis with constant power profile.

System configurations	Failure of the outlet steam valve	Failure of the safety relief valve	Level sensor-PID controller communication interruption	Failure of the PID controller
1	—	—	—	—
2	X	—	—	—
3	—	X	—	—
4	—	—	X	—
5	—	—	—	X
6	X	X	—	—
7	X	—	X	—
8	X	—	—	X
9	—	X	X	—
10	—	X	—	X
11	X	X	X	—
12	X	X	—	X

occurrences and deterministic process variables values of $\bar{\mathbf{x}}$ that exhaustively describe the scenarios belonging to the Near Misses cluster: if a Near Miss sequence vector is characterized by (at least) one of the common characteristics, this is covered by the optimal set $\bar{\mathbf{x}}$. In the present application we have verified that all the scenarios belonging to the identified Near Misses cluster are covered by the minimal conditions that lead the system into a quasifault state, that is, the optimal set $\bar{\mathbf{x}}$. In conclusion, the occurrence of one of the common characteristics listed above is sufficient to lead the system into endangered and insecure operational conditions.

4. Conclusions

In this paper, a risk-based clustering approach for Near Misses identification has been proposed. The approach includes a risk-based feature selection task, where by each safe scenario it is described in terms of probability, consequence, and overall risk. The optimal features set is identified by a wrapper approach based on the combination of a MBDE algorithm with K-means clustering. The characteristics of the Near Misses scenarios are, then, identified solving a multiobjective optimization problem and Hamming distance as a measure of similarity.

The application of the approach to a case study of IDPSA of a UTSG has shown the possibility of retrieving relevant information for risk monitoring.

Appendices

A. Static Reliability Analysis

For a static reliability analysis of the UTSG, we conservatively assume that component failures occur at the beginning of the scenario, with magnitudes equal to their extreme (either maximum or minimum) plausible values [19]. We analyze the dynamic response of the system at constant P_o values (P_o = 5% P_n and P_o = 80% P_n) and identify the minimal cuts sets

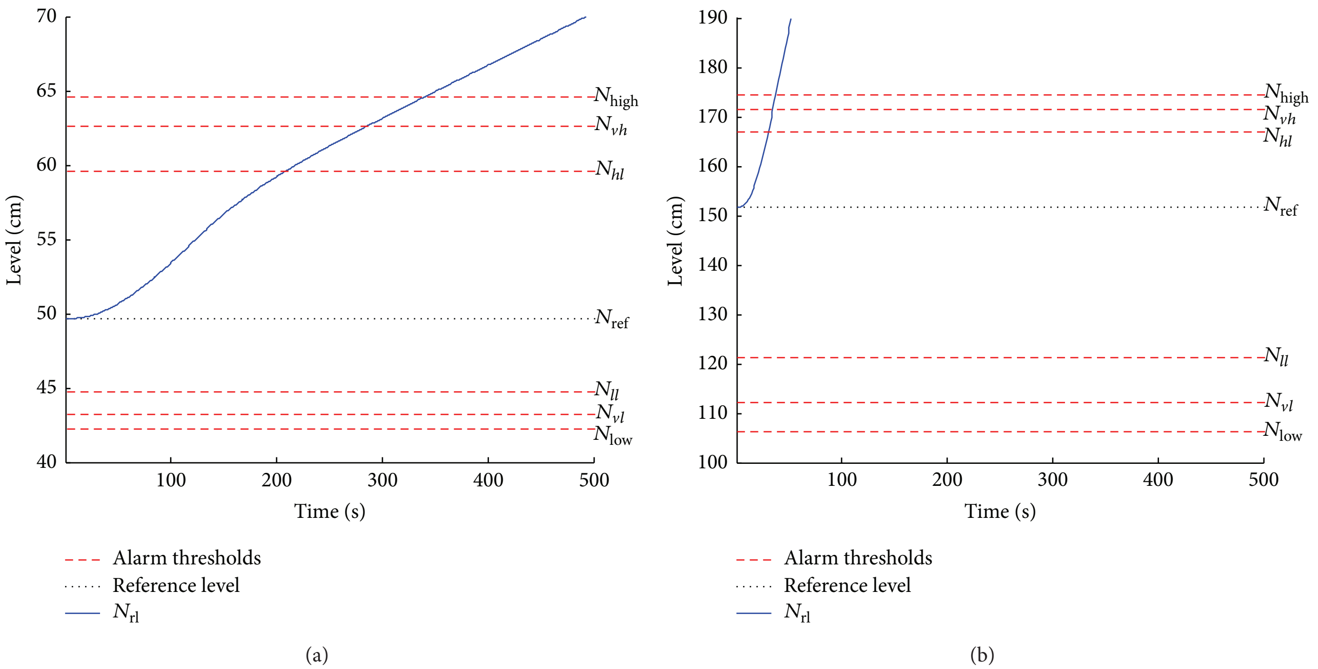


FIGURE 17: N_{rl} evolution when the PID controller output is stuck at time $t = 0$ at the minimum allowable value of −18% of nominal Q_e that should be provided at 5% P_n (a) and at 80% P_n (b).

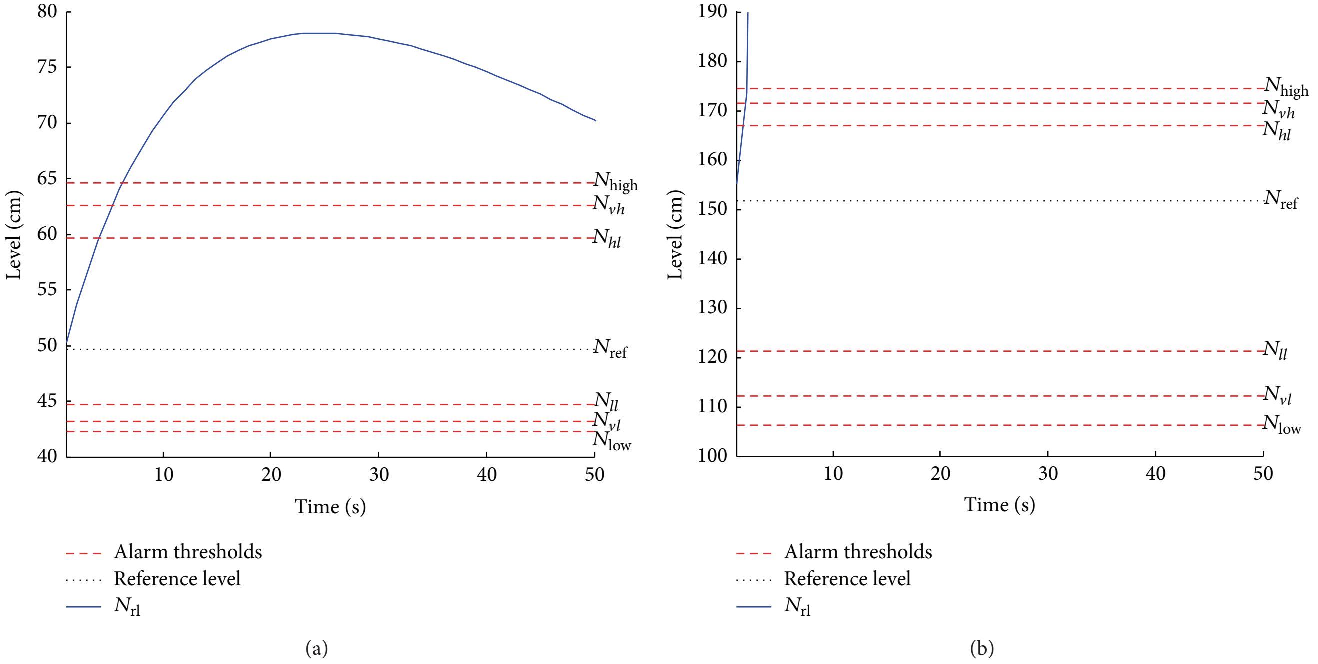
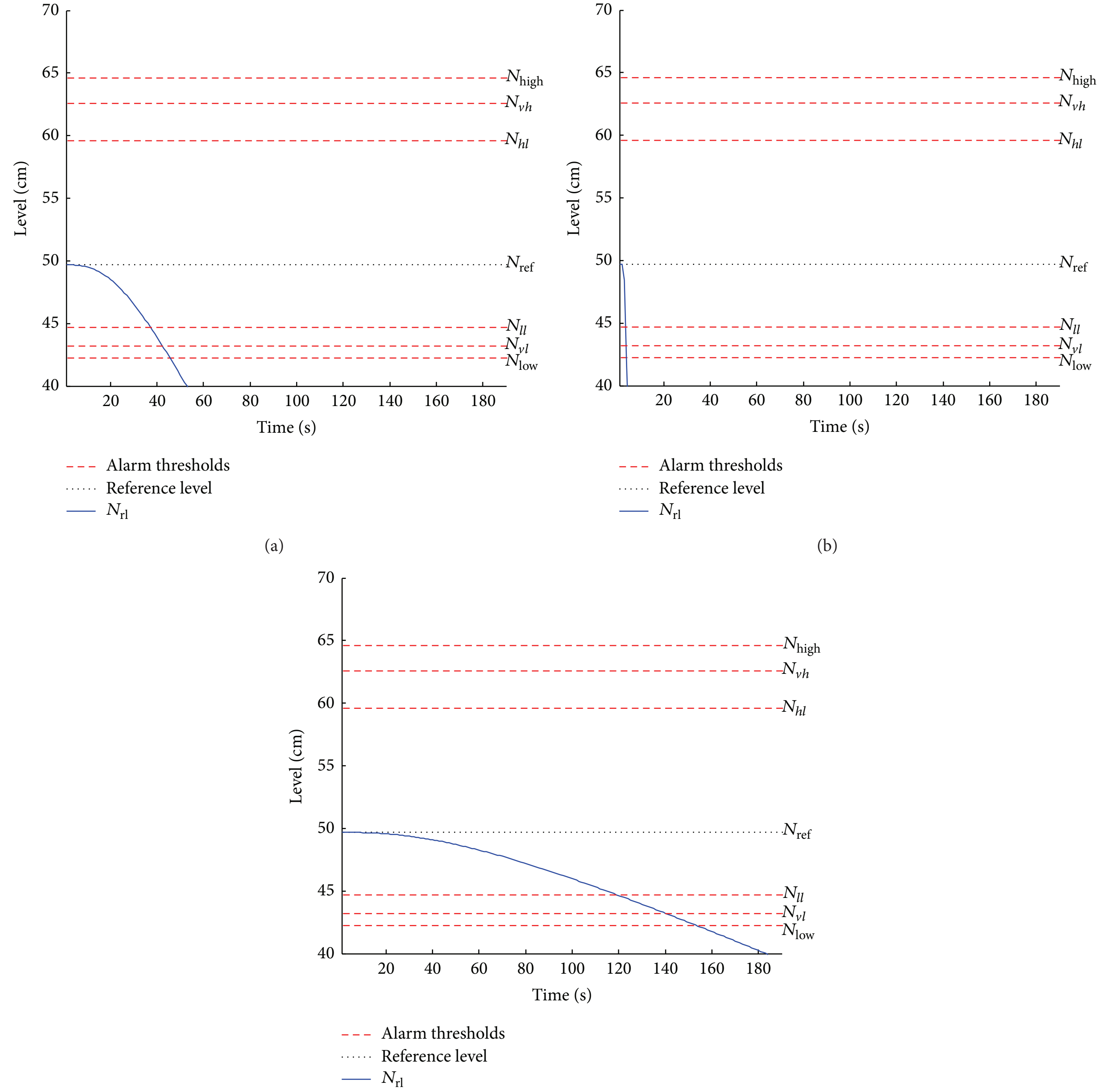


FIGURE 18: N_{rl} evolution when the steam valve fails stuck at time $t = 0$ at the maximum allowable value of 150% of nominal Q_v that should be provided at 5% P_n (a) and at 80% P_n (b).

(MCS) with respect to the low and high level failure modes. Considering the binary, safe or faulty, states of the 6 components (component state is 0 if it works and 1 if it is failed), the number of possible system configurations is equal to 2^6. However, many configurations are not detectable in constant power scenarios, for example, simultaneous occurrence of low and high power communication failures, whereas some others are not important when event occurrence timing is not considered; for example, PID and communication failures occur simultaneously, because, in this case, the feedback control output would always be the same as a stand-alone PID failure. Thus, the possible system configurations to be



(c)

FIGURE 19: (a) N_{rl} evolution when the safety relief valve fails stuck with Q_{sf} = 50.5 (kg/s); (b) the steam valve fails stuck closed; (c) the PID controller fails stuck at 18% of nominal Q_e that should be provided at 5% P_n.

considered in a static analysis with constant power is equal to 12 for each power level (Table 5).

To identify the system MCS, the different system configurations of Table 5 have been simulated by the SIMULINK model, at low and at high (constant) power levels. It turns out that the MCSs for the high level failure mode are the same at both power levels (Figure 16): the failures of the PID controller at its minimum values (i.e., –18% of the nominal Q_e that should be provided at P_o) and of the steam valve at its maximum value (i.e., 150% of the nominal Q_v value that should be provided at P_o) are two first-order MCS. The N_{rl} evolutions when these MCSs occur are shown in Figures 17 and 18.

The analysis of the low level failure mode provides different MCSs at different P_o. At 5% P_n, there are three first-order MCSs represented by the following: (i) safety valve fails stuck at the maximum allowable value: that is, Q_{sf} = 50.5 (kg/s); (ii) steam valve fails stuck closed; (iii) PID

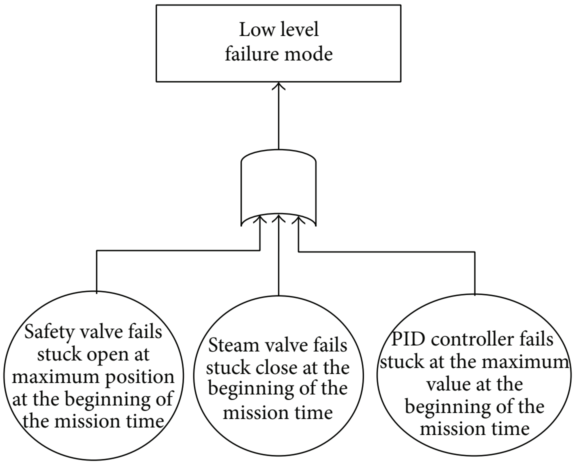

Low level failure mode
Safety valve fails stuck open at maximum position at the beginning of the mission time
Steam valve fails stuck close at the beginning of the mission time
PID controller fails stuck at the maximum value at the beginning of the mission time

Figure 20: Fault tree for the low level failure mode at low power.

controller fails stuck at its maximum values, (i.e., 18% of the nominal Q_e value that should be provided at 5% P_n). The N_{rl} evolution when these MCSs occur and the relative FT are shown in Figures 19 and 20, respectively.

At 80% P_n, three MCSs are found: (i) a second-order MCS that combines the failure of the safety relief valve at its maximum allowable value: that is, $Q_{\mathrm{sf}} = 50.5$ (kg/s), and the failure of the communication, (ii) the steam valve failure in a closed position, and (iii) the PID controller fails at its maximum value (i.e., 18% of the nominal Q_e value that should be provided at 80% P_n). The N_{rl} evolution when these MCSs occur and the relative FT are shown in Figures 21 and 22, respectively.

B. Near Misses Sequence Vector Scenarios

See Table 6.

Abbreviations

CET:	Continuous event tree
CH:	Calinski-Harabasz index
DBA:	Design basis accident
DET:	Dynamic event tree
DSA:	Deterministic safety analysis
ET:	Event tree
FT:	Fault tree
IDPSA:	Integrated deterministic and probabilistic safety analysis
IE:	Initiating event
MBDE:	Modified binary differential evolution
MCS:	Minimal cuts set
MOP:	Multiobjective optimization problem
MVL:	Multiple-valued logic
NPP:	Nuclear power plant
PIs:	Prime implicants
PSA:	Probabilistic safety analysis
SCP:	Set covering problem
SG:	Steam generator
TH:	Thermal-hydraulics
UTSG:	U-tube steam generator.

Symbols

p:	Probability that the developing scenario is an accidental scenario
c:	Consequence that the developing scenario can cause to the system
r:	Overall risk of the developing scenario
t:	Time instant
$p(t)$:	Probability that at time t the scenario can lead the system into an accidental scenario
$c(t)$:	Consequence that at time t the developing scenario is predicted to cause to the system
$r(t)$:	Overall risk of the developing scenario at time t
Q_e:	Flow rate of fresh feed-water entering the steam generator
P_o:	Operating power
P_n:	Nominal power
Q_v:	Flow rate of dry steam exiting the steam generator
N_{rl}:	Narrow range steam generator water level
W_{rl}:	Wide range steam generator water level
T_n:	Time constant for the N_{rl} dynamics
Q_{ef}:	Flow rate of incoming water in steam generator tube bundle region
T_h:	Time constant for the water mass transportation dynamics
τ:	Time constant for the feed-water valve dynamics
Q_{GV}:	Flow rate of steam-water mixture exiting the steam generator tube bundle region
T_g:	Time constant for the dynamics relating Q_V to Q_{GV}
F_g:	Constant in the nonminimum phase term of the dynamics relating Q_V to Q_{GV}
T_{int}:	Time constant for the W_{rl} dynamics

Table 6

Near Miss	T_s	M_s	O_s	T_{sa}	M_{sa}	O_{sa}	T_c	M_c	O_c	T_p	M_p	O_p
1	0	0	0	0	0	0	0	0	0	2	4	1
2	0	0	0	0	0	0	1	1	1	1	4	2
3	0	0	0	0	0	0	1	1	2	1	4	1
4	0	0	0	0	0	0	2	1	1	2	4	2
5	0	0	0	0	0	0	2	1	2	2	4	1
6	0	0	0	0	0	0	2	1	1	3	4	2
7	0	0	0	0	0	0	4	1	2	2	4	1
8	0	0	0	2	2	2	4	1	3	2	4	1
9	0	0	0	2	3	1	4	1	2	4	3	3
10	0	0	0	2	4	2	4	1	3	1	4	1
11	0	0	0	3	2	3	3	1	2	1	4	1
12	0	0	0	3	3	3	1	1	1	1	4	2
13	0	0	0	3	3	3	1	1	2	1	4	1
14	0	0	0	3	3	3	2	1	2	1	4	1
15	0	0	0	3	3	3	2	1	2	2	4	1
16	0	0	0	3	3	2	3	1	3	1	4	1
17	0	0	0	3	3	2	4	1	3	1	4	1
18	0	0	0	3	4	3	1	1	1	2	4	2
19	0	0	0	3	4	3	3	1	2	1	4	1
20	0	0	0	4	2	2	0	0	0	4	4	1
21	0	0	0	4	2	3	2	1	2	1	4	1
22	0	0	0	4	2	3	3	1	2	1	4	1
23	0	0	0	4	2	3	4	1	2	1	4	1
24	0	0	0	4	3	2	0	0	0	1	4	1
25	0	0	0	4	3	2	0	0	0	2	4	1
26	0	0	0	4	3	3	2	1	1	2	4	2
27	0	0	0	4	3	3	3	1	2	2	4	1
28	0	0	0	4	3	2	4	1	3	1	4	1
29	0	0	0	4	4	2	0	0	0	1	4	1
30	0	0	0	4	4	2	0	0	0	2	4	1
31	0	0	0	4	4	3	1	1	1	1	4	2
32	0	0	0	4	4	3	1	1	2	1	4	1
33	0	0	0	4	4	3	1	1	1	2	4	2
34	0	0	0	4	4	3	2	1	2	1	4	1
35	0	0	0	4	4	3	2	1	1	2	4	2
36	0	0	0	4	4	3	3	1	2	3	4	1
37	0	0	0	4	4	3	4	1	2	1	4	1
38	0	0	0	4	4	3	4	1	2	2	4	1
39	1	1	3	1	2	2	1	1	4	1	4	1
40	2	1	2	0	0	0	0	0	0	1	4	1
41	2	1	3	0	0	0	1	1	2	1	4	1
42	2	1	3	2	2	2	0	0	0	2	4	1
43	2	1	2	2	2	3	3	1	4	1	4	1
44	2	1	4	2	4	2	2	1	3	1	3	1
45	2	1	3	2	4	4	2	1	2	1	4	1
46	2	1	3	3	2	4	2	1	2	1	4	1

Table 6: Continued.

Near Miss	T_s	M_s	O_s	T_{sa}	M_{sa}	O_{sa}	T_c	M_c	O_c	T_p	M_p	O_p
47	2	1	2	3	2	3	4	1	4	1	4	1
48	2	1	3	3	3	4	2	1	2	1	4	1
49	2	1	2	3	3	3	3	1	4	1	4	1
50	2	1	3	3	4	4	1	1	2	1	4	1
51	2	1	2	3	4	3	3	1	4	1	4	1
52	2	1	3	4	1	4	2	1	2	1	4	1
53	2	1	3	4	3	4	1	1	2	1	4	1
54	2	1	2	4	3	4	3	1	3	1	4	1
55	2	1	2	4	3	4	4	1	3	1	4	1
56	2	1	2	4	4	3	0	0	0	1	4	1
57	2	2	3	3	4	4	1	1	2	1	4	1
58	2	2	3	4	4	4	2	1	2	1	4	1
59	3	1	2	0	0	0	0	0	0	1	4	1
60	3	1	3	0	0	0	1	1	1	1	4	2
61	3	1	3	0	0	0	1	1	2	1	4	1
62	3	1	3	0	0	0	2	1	2	1	3	1
63	3	1	3	0	0	0	2	1	2	1	4	1
64	3	1	3	0	0	0	2	1	2	2	4	1
65	3	1	4	1	2	3	1	1	2	1	4	1
66	3	1	4	2	2	2	2	1	3	2	4	1
67	3	1	4	2	2	3	2	1	1	2	4	2
68	3	1	3	2	2	2	4	1	4	1	4	1
69	3	1	4	2	3	3	2	1	2	1	4	1
70	3	1	4	2	4	3	1	1	2	1	4	1
71	3	1	4	2	4	2	3	1	3	1	4	1
72	3	1	3	3	1	4	2	1	2	1	4	1
73	3	1	2	3	1	3	3	1	4	1	4	1
74	3	1	3	3	1	4	3	1	2	1	4	1
75	3	1	2	3	1	3	4	1	4	1	4	1
76	3	1	2	3	2	3	0	0	0	1	4	1
77	3	1	3	3	2	4	2	1	2	1	4	1
78	3	1	4	3	2	3	2	1	2	1	4	1
79	3	1	2	3	2	4	3	1	3	1	4	1
80	3	1	3	3	2	4	3	1	2	2	4	1
81	3	1	3	3	2	2	4	1	4	1	4	1
82	3	1	2	3	3	3	0	0	0	1	4	1
83	3	1	4	3	3	3	2	1	2	1	4	1
84	3	1	3	3	3	4	3	1	2	1	4	1
85	3	1	4	3	3	2	3	1	3	1	4	1
86	3	1	3	3	3	2	4	1	4	1	4	1
87	3	1	2	3	4	3	0	0	0	1	4	1
88	3	1	3	3	4	4	1	1	2	1	4	1
89	3	1	4	3	4	3	2	1	2	1	4	1
90	3	1	2	3	4	4	3	1	3	1	4	1
91	3	1	2	3	4	3	4	1	4	1	4	1
92	3	1	3	4	1	4	1	1	1	1	4	2
93	3	1	3	4	1	4	2	1	2	1	4	1
94	3	1	2	4	1	3	4	1	4	1	4	1

Table 6: Continued.

Near Miss	T_s	M_s	O_s	T_{sa}	M_{sa}	O_{sa}	T_c	M_c	O_c	T_p	M_p	O_p
95	3	1	3	4	2	4	2	1	2	1	4	1
96	3	1	3	4	2	4	2	1	2	2	4	1
97	3	1	3	4	2	4	3	1	2	1	4	1
98	3	1	3	4	2	4	3	1	2	2	4	1
99	3	1	3	4	3	4	1	1	2	1	4	1
100	3	1	3	4	3	4	3	1	2	1	4	1
101	3	1	3	4	3	4	3	1	2	2	4	1
102	3	1	2	4	3	4	4	1	3	2	4	1
103	3	1	2	4	4	3	0	0	0	1	4	1
104	3	1	3	4	4	4	1	1	1	1	4	2
105	3	1	3	4	4	4	2	1	2	1	4	1
106	3	1	3	4	4	4	2	1	2	2	4	1
107	3	1	3	4	4	4	3	1	2	1	4	1
108	3	1	2	4	4	4	4	1	3	1	4	1
109	3	2	2	0	0	0	0	0	0	1	4	1
110	3	2	3	0	0	0	1	1	1	1	4	2
111	3	2	3	0	0	0	1	1	2	1	4	1
112	3	2	2	0	0	0	3	1	3	1	4	1
113	3	2	2	0	0	0	4	1	3	1	4	1
114	3	2	4	1	2	2	1	1	3	1	4	1
115	3	2	3	2	3	2	0	0	0	1	4	1
116	3	2	4	2	3	1	2	1	2	3	3	3
117	3	2	3	2	3	2	3	1	4	1	4	1
118	3	2	3	3	1	4	1	1	1	1	4	2
119	3	2	2	3	1	3	3	1	4	1	4	1
120	3	2	2	3	1	3	4	1	4	1	4	1
121	3	2	3	3	2	4	1	1	1	1	4	2
122	3	2	4	3	2	3	2	1	2	1	4	1
123	3	2	3	3	2	4	3	1	2	1	4	1
124	3	2	2	3	3	3	0	0	0	1	4	1
125	3	2	3	3	3	2	0	0	0	1	4	1
126	3	2	4	3	3	3	1	1	1	1	4	2
127	3	2	3	3	3	4	2	1	2	1	4	1
128	3	2	4	3	3	3	2	1	2	2	4	1
129	3	2	2	3	3	4	3	1	3	1	4	1
130	3	2	3	3	3	4	3	1	2	1	4	1
131	3	2	4	3	3	2	3	1	3	1	4	1
132	3	2	4	3	3	3	3	1	2	1	4	1
133	3	2	3	3	3	4	3	1	2	2	4	1
134	3	2	2	3	3	3	4	1	4	1	4	1
135	3	2	3	3	3	2	4	1	4	1	4	1
136	3	2	3	3	4	4	1	1	2	1	4	1
137	3	2	3	3	4	4	2	1	2	1	4	1
138	3	2	4	3	4	3	2	1	2	1	4	1
139	3	2	2	3	4	3	4	1	4	1	4	1
140	3	2	2	3	4	3	4	1	4	2	4	1

Table 6: Continued.

Near Miss	T_s	M_s	O_s	T_{sa}	M_{sa}	O_{sa}	T_c	M_c	O_c	T_p	M_p	O_p
141	3	2	2	4	1	4	3	1	3	1	4	1
142	3	2	2	4	1	4	4	1	3	1	4	1
143	3	2	3	4	2	4	1	1	2	1	4	1
144	3	2	3	4	2	4	3	1	2	1	4	1
145	3	2	3	4	2	4	3	1	2	2	4	1
146	3	2	2	4	2	3	4	1	4	1	4	1
147	3	2	2	4	2	4	4	1	3	1	4	1
148	3	2	3	4	3	4	1	1	1	2	4	2
149	3	2	2	4	3	4	3	1	3	1	4	1
150	3	2	3	4	3	4	3	1	2	1	4	1
151	3	2	3	4	4	4	1	1	1	1	4	2
152	3	2	3	4	4	4	1	1	2	1	4	1
153	3	3	3	2	3	2	4	1	4	1	4	1
154	3	3	4	2	4	3	1	1	2	1	4	1
155	3	3	4	3	3	3	2	1	2	1	4	1
156	3	3	3	3	3	2	3	1	4	1	4	1
157	3	3	4	3	4	3	2	1	2	1	4	1
158	3	3	4	3	4	3	3	1	2	1	4	1
159	4	1	2	0	0	0	0	0	0	1	4	1
160	4	1	2	0	0	0	0	0	0	2	4	1
161	4	1	3	0	0	0	1	1	1	1	4	2
162	4	1	3	0	0	0	1	1	2	1	4	1
163	4	1	3	0	0	0	2	1	1	2	4	2
164	4	1	3	0	0	0	2	1	2	2	4	1
165	4	1	3	0	0	0	3	1	2	3	4	1
166	4	1	2	0	0	0	4	1	3	1	4	1
167	4	1	3	0	0	0	4	1	2	1	4	1
168	4	1	3	0	0	0	4	1	2	2	4	1
169	4	1	4	2	2	3	1	1	2	1	4	1
170	4	1	4	2	2	2	3	1	3	2	4	1
171	4	1	4	2	3	3	1	1	2	1	4	1
172	4	1	4	2	4	2	4	1	3	1	4	1
173	4	1	4	3	2	3	1	1	1	1	4	2
174	4	1	3	3	2	2	4	1	4	2	4	1
175	4	1	3	3	3	2	0	0	0	1	4	1
176	4	1	4	3	3	3	1	1	2	1	4	1
177	4	1	4	3	3	3	2	1	2	1	4	1
178	4	1	3	3	4	2	0	0	0	1	4	1
179	4	1	4	3	4	3	1	1	1	1	4	2
180	4	1	4	3	4	3	2	1	2	2	4	1
181	4	1	4	3	4	3	3	1	2	1	4	1
182	4	1	4	3	4	3	3	1	2	2	4	1
183	4	1	3	3	4	2	4	1	4	2	4	1
184	4	1	2	4	1	3	0	0	0	1	4	1
185	4	1	2	4	1	3	0	0	0	2	4	1
186	4	1	3	4	1	4	1	1	1	2	4	2
187	4	1	3	4	1	4	2	1	2	1	4	1
188	4	1	3	4	1	4	2	1	2	2	4	1

Table 6: Continued.

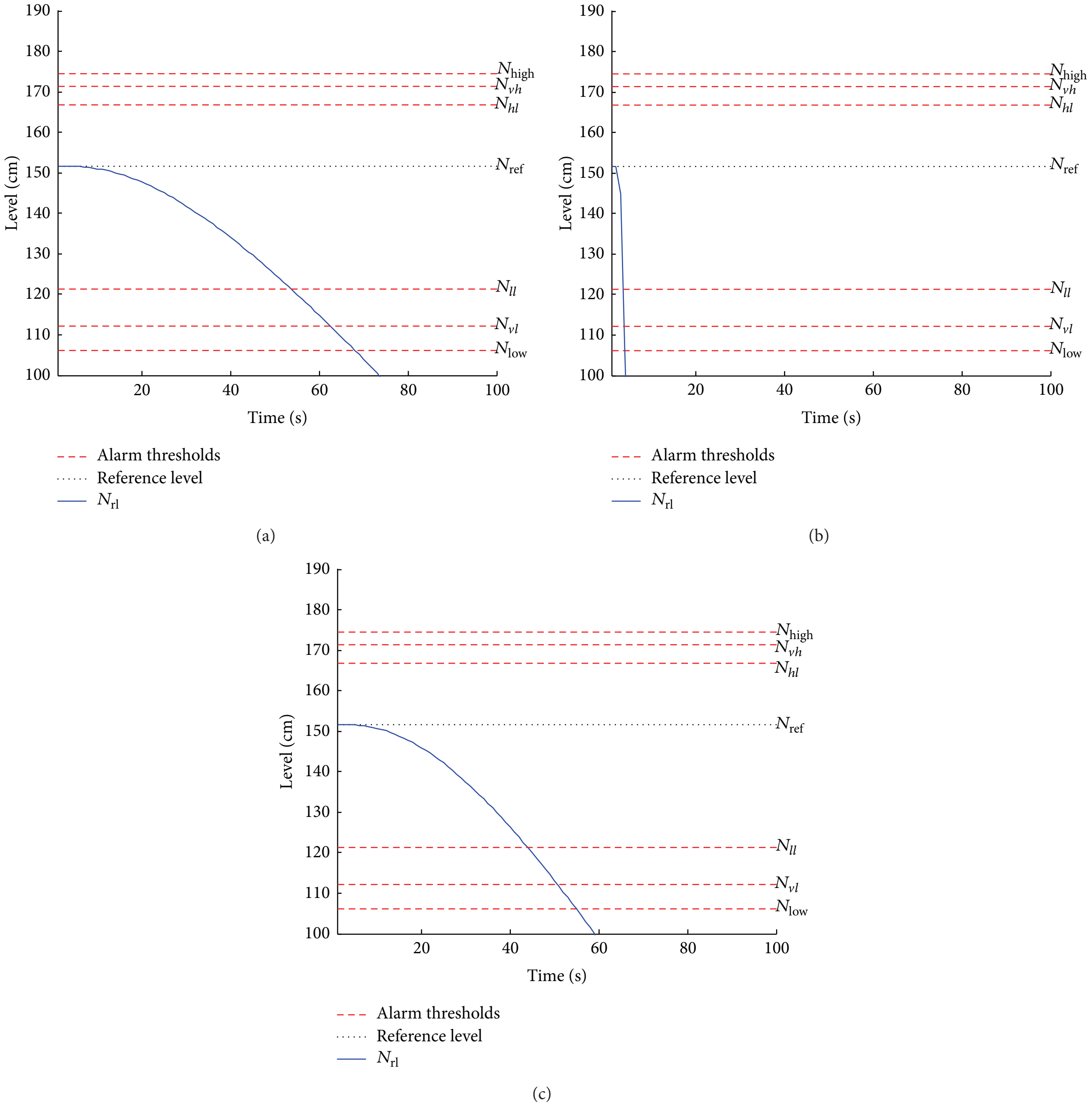
Near Miss	T_s	M_s	O_s	T_{sa}	M_{sa}	O_{sa}	T_c	M_c	O_c	T_p	M_p	O_p
189	4	1	3	4	1	4	3	1	2	1	4	1
190	4	1	2	4	1	3	4	1	4	1	4	1
191	4	1	3	4	1	4	4	1	2	1	4	1
192	4	1	2	4	1	3	4	1	4	2	4	1
193	4	1	3	4	1	4	4	1	2	3	4	1
194	4	1	2	4	2	3	0	0	0	1	4	1
195	4	1	3	4	2	4	1	1	1	1	4	2
196	4	1	4	4	2	3	1	1	2	1	4	1
197	4	1	3	4	2	4	2	1	2	1	4	1
198	4	1	4	4	2	3	2	1	1	2	4	2
199	4	1	3	4	2	4	2	1	2	2	4	1
200	4	1	3	4	2	4	3	1	2	2	4	1
201	4	1	2	4	2	3	4	1	4	1	4	1
202	4	1	3	4	2	2	4	1	4	1	4	1
203	4	1	3	4	2	4	4	1	2	1	4	1
204	4	1	2	4	2	4	4	1	3	2	4	1
205	4	1	3	4	2	4	4	1	2	2	4	1
206	4	1	4	4	2	2	4	1	3	3	4	1
207	4	1	2	4	3	3	0	0	0	1	4	1
208	4	1	3	4	3	2	0	0	0	1	4	1
209	4	1	4	4	3	3	1	1	1	1	4	2
210	4	1	3	4	3	4	1	1	2	1	4	1
211	4	1	4	4	3	3	1	1	2	1	4	1
212	4	1	3	4	3	4	2	1	2	1	4	1
213	4	1	4	4	3	3	2	1	2	1	4	1
214	4	1	4	4	3	3	2	1	1	2	4	2
215	4	1	3	4	3	4	2	1	2	2	4	1
216	4	1	4	4	3	3	2	1	2	2	4	1
217	4	1	3	4	3	4	3	1	2	1	4	1
218	4	1	3	4	3	4	3	1	2	3	4	1
219	4	1	2	4	3	3	4	1	4	1	4	1
220	4	1	2	4	3	4	4	1	3	1	4	1
221	4	1	3	4	3	2	4	1	4	1	4	1
222	4	1	3	4	3	4	4	1	2	1	4	1
223	4	1	2	4	3	3	4	1	4	2	4	1
224	4	1	3	4	3	4	4	1	2	2	4	1
225	4	1	4	4	3	2	4	1	3	2	4	1
226	4	1	2	4	4	3	0	0	0	2	4	1
227	4	1	3	4	4	4	1	1	1	1	4	2
228	4	1	4	4	4	3	1	1	1	1	4	2
229	4	1	4	4	4	3	1	1	1	2	4	2
230	4	1	3	4	4	4	2	1	2	1	4	1
231	4	1	4	4	4	3	2	1	2	1	4	1
232	4	1	4	4	4	3	2	1	2	2	4	1
233	4	1	3	4	4	4	3	1	2	1	4	1
234	4	1	4	4	4	3	3	1	2	1	4	1
235	4	1	3	4	4	4	3	1	2	3	4	1
236	4	1	3	4	4	4	4	1	2	1	4	1
237	4	1	4	4	4	2	4	1	3	2	4	1

Table 6: Continued.

Near Miss	T_s	M_s	O_s	T_{sa}	M_{sa}	O_{sa}	T_c	M_c	O_c	T_p	M_p	O_p
238	4	2	2	0	0	0	0	0	0	2	4	1
239	4	2	3	0	0	0	1	1	2	1	4	1
240	4	2	3	0	0	0	2	1	1	2	4	2
241	4	2	3	0	0	0	2	1	2	2	4	1
242	4	2	3	0	0	0	3	1	2	1	4	1
243	4	2	3	0	0	0	3	1	2	2	4	1
244	4	2	3	0	0	0	4	1	2	1	4	1
245	4	2	2	0	0	0	4	1	3	2	4	1
246	4	2	3	1	3	2	0	0	0	1	3	1
247	4	2	4	2	1	3	2	1	1	2	4	2
248	4	2	4	2	3	2	4	1	3	1	4	1
249	4	2	3	2	4	2	0	0	0	1	4	1
250	4	2	4	3	2	3	1	1	2	1	4	1
251	4	2	4	3	2	2	3	1	3	3	4	1
252	4	2	3	3	2	2	4	1	4	1	4	1
253	4	2	3	3	3	2	0	0	0	1	4	1
254	4	2	4	3	3	3	1	1	1	2	4	2
255	4	2	4	3	3	3	2	1	2	1	4	1
256	4	2	4	3	3	3	3	1	2	1	4	1
257	4	2	3	3	3	2	4	1	4	1	4	1
258	4	2	4	3	3	2	4	1	3	1	4	1
259	4	2	4	3	4	3	1	1	1	1	4	2
260	4	2	4	3	4	3	1	1	1	2	4	2
261	4	2	4	3	4	3	2	1	2	1	4	1
262	4	2	4	3	4	3	3	1	2	1	4	1
263	4	2	2	4	1	3	0	0	0	2	4	1
264	4	2	3	4	1	4	1	1	1	1	4	2
265	4	2	3	4	1	4	3	1	2	1	4	1
266	4	2	2	4	1	4	4	1	3	1	4	1
267	4	2	2	4	1	3	4	1	4	2	4	1
268	4	2	3	4	1	4	4	1	2	2	4	1
269	4	2	3	4	2	2	0	0	0	4	4	1
270	4	2	3	4	2	4	1	1	1	1	4	2
271	4	2	3	4	2	4	1	1	1	2	4	2
272	4	2	4	4	2	3	2	1	1	4	4	2
273	4	2	4	4	2	3	3	1	2	1	4	1
274	4	2	3	4	2	4	3	1	2	2	4	1
275	4	2	3	4	2	4	3	1	2	3	4	1
276	4	2	2	4	2	3	4	1	4	1	4	1
277	4	2	3	4	2	2	4	1	4	1	4	1
278	4	2	3	4	2	4	4	1	2	1	4	1
279	4	2	4	4	2	3	4	1	2	1	4	1
280	4	2	2	4	2	4	4	1	3	2	4	1
281	4	2	3	4	3	2	0	0	0	1	4	1
282	4	2	3	4	3	2	0	0	0	2	4	1
283	4	2	4	4	3	3	1	1	1	1	4	2
284	4	2	3	4	3	4	1	1	2	1	4	1
285	4	2	4	4	3	3	1	1	2	1	4	1

Table 6: Continued.

Near Miss	T_s	M_s	O_s	T_{sa}	M_{sa}	O_{sa}	T_c	M_c	O_c	T_p	M_p	O_p
286	4	2	3	4	3	4	2	1	2	1	4	1
287	4	2	4	4	3	3	2	1	2	1	4	1
288	4	2	3	4	3	4	2	1	1	2	4	2
289	4	2	4	4	3	3	2	1	1	2	4	2
290	4	2	4	4	3	3	2	1	2	2	4	1
291	4	2	4	4	3	3	3	1	2	1	4	1
292	4	2	3	4	3	4	3	1	2	2	4	1
293	4	2	2	4	3	4	4	1	3	1	4	1
294	4	2	3	4	3	4	4	1	2	1	4	1
295	4	2	2	4	3	3	4	1	4	2	4	1
296	4	2	2	4	3	4	4	1	3	2	4	1
297	4	2	4	4	3	3	4	1	2	2	4	1
298	4	2	3	4	4	2	0	0	0	1	4	1
299	4	2	2	4	4	3	0	0	0	2	4	1
300	4	2	3	4	4	2	0	0	0	2	4	1
301	4	2	4	4	4	3	1	1	2	1	4	1
302	4	2	4	4	4	3	1	1	1	2	4	2
303	4	2	3	4	4	4	2	1	2	1	4	1
304	4	2	4	4	4	3	2	1	2	1	4	1
305	4	2	4	4	4	3	2	1	1	2	4	2
306	4	2	3	4	4	4	3	1	2	1	4	1
307	4	2	2	4	4	4	4	1	3	1	4	1
308	4	2	3	4	4	2	4	1	4	1	4	1
309	4	2	4	4	4	2	4	1	3	1	4	1
310	4	2	2	4	4	4	4	1	3	2	4	1
311	4	2	3	4	4	2	4	1	4	2	4	1
312	4	3	3	2	4	2	0	0	0	1	4	1
313	4	3	4	2	4	3	2	1	2	1	4	1
314	4	3	4	3	3	2	3	1	3	1	4	1
315	4	3	4	3	3	3	3	1	2	1	4	1
316	4	3	3	3	4	2	0	0	0	1	4	1
317	4	3	4	3	4	3	1	1	2	1	4	1
318	4	3	3	3	4	2	4	1	4	1	4	1
319	4	3	4	3	4	2	4	1	3	1	4	1
320	4	3	3	4	3	2	0	0	0	2	4	1
321	4	3	4	4	3	3	3	1	2	1	4	1
322	4	3	4	4	3	2	4	1	3	2	4	1
323	4	3	3	4	4	2	0	0	0	2	4	1
324	4	3	4	4	4	3	1	1	1	1	4	2
325	4	3	4	4	4	3	1	1	2	1	4	1
326	4	3	4	4	4	3	2	1	2	1	4	1
327	4	3	4	4	4	3	2	1	1	2	4	2
328	4	3	4	4	4	3	2	1	2	2	4	1
329	4	3	4	4	4	3	3	1	2	1	4	1
330	4	3	3	4	4	2	4	1	4	1	4	1
331	4	3	3	4	4	2	4	1	4	2	4	1
332	4	3	4	4	4	3	4	1	2	3	4	1



(a)

(b)

(c)

FIGURE 21: (a) N_{rl} evolution when simultaneously the safety relief valve fails stuck with $Q_{\text{sf}} = 50.5$ (kg/s) and the communication fails; (b) the steam valve fails stuck closed; (c) the PID controller fails stuck at 18% of the nominal Q_e that should be provided at 80% P_n.

$\mathbf{x}$: System state
$\dot{x}$: Derivative of system state
N_{ref}: Narrow range steam generator water level at a reference position
N_{high}: Automatic reactor trip threshold
N_{low}: Turbine trip threshold
N_{hl}: First prealarm automatic reactor trip threshold
N_{ll}: First prealarm turbine trip threshold
N_{vh}: Second prealarm automatic reactor trip threshold
N_{vl}: First prealarm turbine trip threshold
Q_{pid}: Water flow rate provided by PID controller
Q_{sf}: Water flow rate removed by safety valve
T_{miss}: Mission time
t_{mvl}: Time steps in MVL discretization
φ: Cumulative probability function of the Gaussian distribution

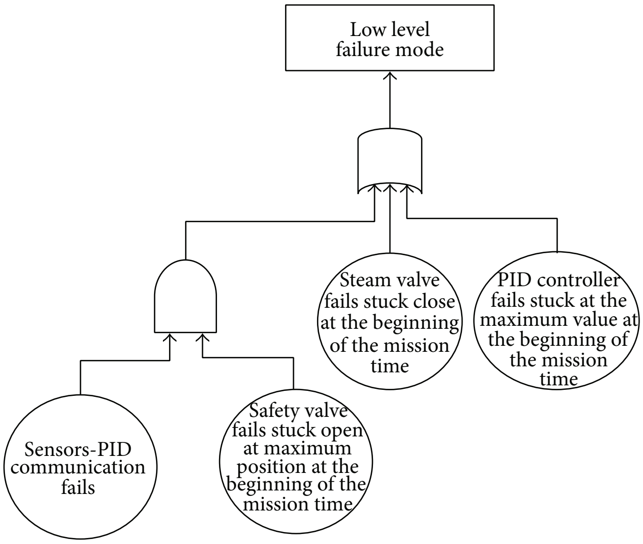

Low level failure mode
Steam valve fails stuck close at the beginning of the mission time
PID controller fails stuck at the maximum value at the beginning of the mission time
Sensors-PID communication fails
Safety valve fails stuck open at maximum position at the beginning of the mission time

Figure 22: Fault tree for the low level failure mode at high power.

μ: Mean value of the Gaussian distribution
σ: Standard deviation of the Gaussian distribution
A: Intensity coefficient
K: Number of clusters
n: Index of the profile of p, c, and r
μ_n: Mean value of the nth profile
max: Peak value of the nth profile
σ_n: Standard deviation of the nth profile
RMS: Root mean square of the nth profile
SK: Skewness of the nth profile
KU: Kurtosis of the nth profile
N: Number of scenarios to the training set
F: Dimension of the set of features
T_{set}: Number of scenarios to the test set
SS_b: Overall between-cluster variance
SS_w: Overall within-cluster variance
n_k: Number of scenarios assigned to the kth cluster
x_c: Generic scenario
m_k: Centroid of the kth cluster
μ_{risk}: Mean risk of the clustered scenarios
t_{risk}: Time elapsed from the instant at which r starts to deviate from zero of the clustered scenarios
f: Fitness function of the MOP
f_1: First objective function of the MOP
f_2: Second objective function of the MOP
$\overline{\mathbf{x}}$: Sequence vector belonging to the Pareto set of the MOP.

Conflict of Interests

The authors declare no conflict of interests.

References

[1] D. G. Kang, S.-H. Ahn, and S. H. Chang, "A combined deterministic and probabilistic procedure for safety assessment of beyond design basis accidents in nuclear power plant: application to ECCS performance assessment for design basis LOCA redefinition," *Nuclear Engineering and Design*, vol. 260, pp. 165–174, 2013.

[2] E. Zio and F. Di Maio, "The needs and dreams for methodologies of IDPSA," in *Proceedings of the Integrated Deterministic and Probabilistic Safety Analyses Workshop*, KTH, Stockholm, Sweden, November 2012.

[3] E. Zio, F. Di Maio, and J. Tong, "Safety margins confidence estimation for a passive residual heat removal system," *Reliability Engineering and System Safety*, vol. 95, no. 8, pp. 828–836, 2010.

[4] T. Aldemir, "A survey of dynamic methodologies for probabilistic safety assessment of nuclear power plants," *Annals of Nuclear Energy*, vol. 52, pp. 113–124, 2013.

[5] W. Keller and M. Modarres, "A historical overview of probabilistic risk assessment development and its use in the nuclear power industry: a tribute to the late Professor Norman Carl Rasmussen," *Reliability Engineering and System Safety*, vol. 89, no. 3, pp. 271–285, 2005.

[6] Y. Vorobyev and P. Kudinov, "Development and application of a genetic algorithm based dynamic pra methodology to plant vulnerability search," in *International Topical Meeting on Probabilistic Safety Assessment and Analysis*, vol. 1, pp. 559–573, 2011.

[7] N. Khakzad, F. Khan, and P. Amyotte, "Dynamic risk analysis using bow-tie approach," *Reliability Engineering and System Safety*, vol. 104, pp. 36–44, 2012.

[8] M. Marseguerra and E. Zio, "Monte Carlo approach to PSA for dynamic process systems," *Reliability Engineering and System Safety*, vol. 52, no. 3, pp. 227–241, 1996.

[9] J. Kirschenbaum, P. Bucci, M. Stovsky et al., "A benchmark system for comparing reliability modeling approaches for digital instrumentation and control systems," *Nuclear Technology*, vol. 165, no. 1, pp. 53–95, 2009.

[10] T. Aldemir, S. Guarro, J. Kirschenbaum et al., "A Benchmark implementation of two dynamic methodologies for the reliability modeling of digital instrumentation and control systems," NUREG-CR Report Draft, 2008.

[11] P. E. Labeau, C. Smidts, and S. Swaminathan, "Dynamic reliability: towards an integrated platform for probabilistic risk assessment," *Reliability Engineering and System Safety*, vol. 68, no. 3, pp. 219–254, 2000.

[12] E. Zio, "Integrated deterministic and probabilistic safety assessment: concepts, challenges, research directions," *Nuclear Engineering and Design*, vol. 280, pp. 413–419, 2014.

[13] J. Devooght and C. Smidts, "Probabilistic reactor dynamics I: the theory of continuous event trees," *Nuclear Science and Engineering*, vol. 111, no. 3, pp. 229–240, 1992.

[14] V. Kopustinskas, J. Augutis, and S. Rimkevičius, "Dynamic reliability and risk assessment of the accident localization system of the Ignalina NPP RBMK-1500 reactor," *Reliability Engineering and System Safety*, vol. 87, no. 1, pp. 77–87, 2005.

[15] E. Hofer, M. Kloos, B. Krzykacz-Hausmann, J. Peschke, and M. Sonnenkalb, *Dynamic Event Trees for Probabilistic Safety Analysis*, Gesellschaft für Anlagen- und Reaktorsicherheit, Garsching, Germany, 2004.

[16] A. Hakobyan, R. Denning, T. Aldemir, S. Dunagan, and D. Kunsman, "A methodology for generating dynamic accident progression event trees for level 2 PRA," Tech. Rep. SAND2008-4746, Sandia National Laboratories, Albuquerque, NM, USA, 2008.

[17] F. Di Maio, S. Baronchelli, and E. Zio, "A computational framework for prime implicants identification in non-coherent dynamic systems," *Risk Analysis*, vol. 35, no. 1, pp. 142–156, 2015.

[18] W. V. Quine, "The problem of simplifying truth functions," *The American Mathematical Monthly*, vol. 59, pp. 521–531, 1952.

[19] E. Zio and F. D. Maio, "Processing dynamic scenarios from a reliability analysis of a nuclear power plant digital instrumentation and control system," *Annals of Nuclear Energy*, vol. 36, no. 9, pp. 1386–1399, 2009.

[20] V. M. Bier and W. Yi, "The performance of precursor-based estimators for rare event frequencies," *Reliability Engineering and System Safety*, vol. 50, no. 3, pp. 241–251, 1995.

[21] J. W. Johnson and D. M. Rasmuson, "The US NRC's accident sequence precursor program: an overview and development of a Bayesian approach to estimate core damage frequency using precursor information," *Reliability Engineering and System Safety*, vol. 53, no. 2, pp. 205–216, 1996.

[22] J. H. Saleh, E. A. Saltmarsh, F. M. Favarò, and L. Brevault, "Accident precursors, near misses, and warning signs: critical review and formal definitions within the framework of Discrete Event Systems," *Reliability Engineering and System Safety*, vol. 114, no. 1, pp. 148–154, 2013.

[23] O. Zadakbar, S. Imtiaz, and F. Khan, "Dynamic risk assessment and fault detection using a multivariate technique," *Process Safety Progress*, vol. 32, no. 4, pp. 365–375, 2013.

[24] K. Fukunaga and L. D. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Transactions on Information Theory*, vol. 21, no. 1, pp. 32–40, 1975.

[25] D. Mandelli, A. Yilmaz, K. Metzroth, T. Aldemir, and R. Denning, "Scenario aggregation and analysis via Mean-Shift Methodology," in *Proceedings of the International Congress on Advances in Nuclear Power Plants (ICAPP '10)*, vol. 2, pp. 987–991, June 2010.

[26] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Kluwer Academic, Norwell, Mass, USA, 1981.

[27] J. F. Aubry, G. Babykina, A. Barros et al., "Project APPRODYN: APPROches de la fiabilité DYNamique pour modéliser des systèmes critiques," Tech. Rep., Collaboration CRAN, EDF R&D, INRIACQFD, UTT-ICD, 2012.

[28] M. V. Kothare, B. Mettler, M. Morari, P. Bendotti, and C.-M. Falinower, "Level control in the steam generator of a nuclear power plant," *IEEE Transactions on Control Systems Technology*, vol. 8, no. 1, pp. 55–69, 2000.

[29] H. Habibiyan, S. Setayeshi, and H. Arab-Alibeik, "A fuzzy-gain-scheduled neural controller for nuclear steam generators," *Annals of Nuclear Energy*, vol. 31, no. 15, pp. 1765–1781, 2004.

[30] M. Marseguerra, E. Zio, and F. Cadini, "Optimized adaptive fuzzy controller of the water level of a pressurized water reactor steam generator," *Nuclear Science and Engineering*, vol. 155, no. 3, pp. 386–394, 2007.

[31] [IAEA-TECDOC-981; 1997], "Assessment and management of ageing of major nuclear power plant component important to safety: Steam," IAEA, IAEA-TECDOC-98, Vienna, 1997.

[32] S. F. Garribba, E. Guagnini, and P. Mussio, "Multiple-valued logic trees: meaning and prime implicants," *IEEE Transactions on Reliability*, vol. 34, no. 5, pp. 463–472, 1985.

[33] F. Di Maio, S. Baronchelli, M. Vagnoli, and E. Zio, "Prime implicants determination by differential evolution for dynamic reliability analysis of non-coherent systems," *Reliability Engineering and System Safety*. Under review.

[34] F. Di Maio, S. Baronchelli, and E. Zio, "A visual interactive method for prime implicants identification," *IEEE Transactions on Reliability*, no. 99, pp. 1–11, 2014.

[35] C. M. Rocco and M. Muselli, "A machine learning algorithm to estimate minimal cut and path sets from a Monte Carlo simulation," in *Probabilistic Safety Assessment and Management*, pp. 3142–3147, Springer, London, UK, 2004.

[36] F. Di Maio, J. Hu, P. Tse, M. Pecht, K. Tsui, and E. Zio, "Ensemble-approaches for clustering health status of oil sand pumps," *Expert Systems with Applications*, vol. 39, no. 5, pp. 4847–4859, 2012.

[37] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial Intelligence*, vol. 97, no. 1-2, pp. 273–324, 1997.

[38] P. Baraldi, F. Cannarile, F. Di Maio, and E. Zio, "Hierarchical *k*-nearest neighbours classification and binary differential evolution for fault diagnostics of automotive bearings operating under variable conditions," *Reliability Engineering and System Safety*. Under review.

[39] L. Wang, X. Fu, M. I. Menhas, and M. Fei, "A modified binary differential evolution algorithm," in *Life System Modeling and Intelligent Computing*, vol. 6329 of *Lecture Notes in Computer Science*, pp. 49–57, Springer, Berlin, Germany, 2010.

[40] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 281–297, Berkeley, Calif, USA, 1967.

[41] T. Calinski and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics*, vol. 3, no. 1, pp. 1–27, 1974.

[42] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[43] J. E. Beasley and P. C. Chu, "A genetic algorithm for the set covering problem," *European Journal of Operational Research*, vol. 94, no. 2, pp. 392–404, 1996.

[44] F. Di Maio, S. Baronchelli, and E. Zio, "Hierarchical differential evolution for minimal cut sets identification: application to nuclear safety systems," *European Journal of Operational Research*, vol. 238, no. 2, pp. 645–652, 2014.

[45] R. W. Hamming, "Error detecting and error correcting codes," *The Bell System Technical Journal*, vol. 29, no. 2, pp. 147–160, 1950.

[46] A. Popa and J. J. McDowell, "The effect of Hamming distances in a computational model of selection by consequences," *Behavioural Processes*, vol. 84, no. 1, pp. 428–434, 2010.