Nuclear Engineering and Design 413 (2023) 112502

Contents lists available at ScienceDirect

Nuclear Engineering and Design

journal homepage: www.elsevier.com/locate/nucengdes

Unsupervised anomaly detection in pressurized water reactor digital twins using autoencoder neural networks

S.A. Cancemi [a,*], R. Lo Frano [a], C. Santus [a], T. Inoue [b]

[a] Department of Civil and Industrial Engineering, University of Pisa, Pisa, Italy
[b] JAMSTEC - Institute for Marine-Earth Exploration and Engineering, Yokosuka, Japan

ARTICLE INFO

Keywords:
Neural Network
Autoencoder
Predictive Maintenance
Safety
NPP
Unsupervised Anomaly Detection
LOCA

ABSTRACT

Deep learning (DL), that is becoming quite popular for prediction and analysis of complex patterns in large amounts of data is used to investigate the safety behaviour of the nuclear plant items. This is achieved by using multiple layers of artificial neural networks to process and transform input data, allowing for the creation of highly accurate predictive models. Particularly to the aim the unsupervised machine learning approach and the digital twin concept in form of pressurized water reactor 2-loop simulator are used. This innovative methodology is based on neural network algorithm that makes capable to predict failures of plant structure, system, and components earlier than the activation of safety and emergency systems. Moreover, to match the objective of the study several scenarios of loss of cooling accident (LOCA) of different break size were simulated. To make the acquisition platform realistic, Gaussian noise was added to the input signals. The neural network has been fed by synthetic dataset provide by PCTRAN simulator and the efficiency in event identification was studied. Further, due to the very limited studies on the unsupervised anomaly detection by means of autoencoder neural networks applied for plant monitoring and surveillance, the methodology has been validated with experimental data from resonant test rig designed for fatigue testing of tubular components. The obtained results demonstrate the reliability and the efficiency of the methodology in detecting anomalous events prior the activation of safety system. Particularly, if the difference between the expected readings and the collected data goes beyond the predetermined threshold, then the anomalous event is identified, e.g., the model detected anomalies up to 38 min before the reactor scram intervention.

1. Introduction

Predictive maintenance (PM) is an innovative approach to maintenance that aims to predict potential equipment failures and carry out maintenance work before upsets or accidents occur. This approach can significantly reduce downtime and costs, increase productivity, and, above all, improve the safety of nuclear power plants (NPPs). PM is often contrasted with reactive maintenance (RM), which responds to equipment failures after they occur, and preventive maintenance, which performs maintenance activities at scheduled intervals of time regardless of equipment condition. PM relies on advanced technologies, such as sensors, machine learning, and Internet of Things (IoT), to monitor equipment performance and predict potential failures. The process begins by collecting real-time data on equipment performance, including temperature, vibration, pressure, and other parameters that represent the equipment condition (International Atomic Energy Agency, 2018). Data are then analyzed using statistical models, artificial intelligence (AI) algorithms, and other techniques in order to identify patterns and anomalies that may predict potential equipment failures. The insights from this analysis are used to schedule maintenance activities, ensuring that equipment remains in ideal operating condition. In addition to reduce downtime and costs, predictive maintenance can also be employed to improve the safety in industrial and commercial operations thereby minimizing the risk of accidents and incidents that can result from equipment malfunctions and failures.

Several studies highlighted that predictive maintenance can lead to a 11% of cost reduction, 8% of reduction of safety, health, environment and quality risk, and 7% of lifetime extension of aging asset (PwC., 2017; Deloitte., 2015; Cancemi and Lo Frano, 2023).

Among the several available machine learning (ML) training models (to be chosen based on the data and the problem being solved), the three main types are: a) supervised learning, b) unsupervised learning, and c)

* Corresponding author.
E-mail address: salvatore.cancemi@phd.unipi.it (S.A. Cancemi).

https://doi.org/10.1016/j.nucengdes.2023.112502
Received 20 March 2023; Received in revised form 21 June 2023; Accepted 15 July 2023
Available online 19 July 2023

reinforcement learning.

Supervised learning (SL) is a type of machine learning training that involves providing labelled training data to a machine learning algorithm. The algorithm learns to recognize patterns and makes predictions based on the labelled. A supervised learning model could be trained on historical equipment data, such as operation conditions, failure modes, and maintenance records. The input variables might be temperature, pressure, vibration levels, operating hours, etc., and the output variable would be the occurrence of failure. This trained model can then be used to predict when a machine/item is likely to fail based on the provided input variables, allowing maintenance to be scheduled proactively in order to prevent the unexpected downtime.

Unsupervised learning (UL) is a type of machine learning training that involves providing unlabelled data to a machine learning algorithm. The algorithm learns to identify patterns and structure in the data without any prior knowledge of what does the data represent. UL can be used for anomaly detection (AD) in machine health data. A model could be trained on normal operational data from machinery. Once the model has learned this 'normal' state, it can then monitor new data for any deviations from it. Such deviations (anomalies) could be early indicators of a potential failure, enabling the proactive maintenance.

Reinforcement learning (RL) is a machine learning technique that focuses on training an agent to make decisions that maximize a reward within a specific environment. The agent repetitively learns through a process of trial and error, analysing the resulting outcomes and adjusting its behaviour accordingly. RL could be used to optimize maintenance schedules. In this scenario, the RL agent's environment is the state of the machinery. Actions could be to maintain or not maintain each equipment/item, and the reward could be inversely related to maintenance costs and the number of machine failures. Over time, the RL agent could learn an optimal policy for maintaining each piece of equipment/item to minimize both the overall costs and the machine failures.

The nuclear industry still employs PM and RM maintenance strategies, particularly for Class I components (ASME III, 1980), due to safety and cost concerns. However, both often result in premature replacement of functional components. On the other hand, PM relies on historical data and degradation curves, but the lack of failure data limits the effectiveness of AI approaches. To address these limitations, this study proposes an innovative methodology that utilizes the unsupervised approach and digital twin (DT) concept that provide synthetic anomaly dataset. The digital twin is represented by pressurized water reactor (PWR) 2-loop simulator PCTRAN® (Cancemi and Lo Frano, 2022; Cancemi and Lo Frano, 2022; International Atomic Energy Agency, 2019; U.S. NRC, 2021). Digital twin is a virtual representation of a physical object, system, or process. It is a computerized model that is designed to simulate the behavior, performance, and characteristics of a real-world object or system in a digital environment. DT combined with unsupervised machine learning can enhance predictive maintenance in several ways:

- Lack of Failure Data: Unlike supervised learning that requires failure examples, unsupervised learning learns from normal operation data and flags anomalies, solving the issue of infrequent failure events.
- Real-Time Anomaly Detection: By applying unsupervised learning to digital twin data, potential problems can be identified in real-time, enabling quicker maintenance responses.
- Handling Complex Relationships: Unsupervised learning can better understand complex variable relationships in the system compared to rule-based systems.
- Continuous Learning: Unlike traditional methods needing manual reprogramming for changes, unsupervised learning algorithms adapt over time, improving system understanding as more data is processed.

This study deals with two distinct scenarios of anomalies, wherein the cooling system in the hot leg and cold leg fail, respectively. The percentage of pipe failure fraction varies for each of the scenarios. This work is not aimed to describe a loss of cooling accident (LOCA) event, which is ruled by 10CFR50.46 (U.S. NRC, 2023) and whose several aspects are analyzed extensively and in great detail in studies found in the open scientific literature (U.S.NRC, 2023; Lewis, 1977; International Atomic Energy Agency, 2016; González-González et al., 2023; Veshchunov et al., 2022). The aim of this study is so to forecast potential anomaly patterns beforehand and assess the efficacy of the suggested methodology by contrasting it against safety systems such as reactor protection system (RPS) and engineered safety features (ESF).

2. State of the art

The unsupervised approach is based on autoencoder. Autoencoder is a neural network used for anomaly detection that can learn complex patterns in the data. It is trained on normal data, and if the reconstruction error of new input data is above the threshold, it is considered an anomaly. Autoencoders can be trained in an unsupervised manner and used for equipment failure prediction. Moreover, it is able to reveal potential issues, even if the specific failure mode is not yet known or detected by a sensor monitoring system. The study (Wang and Takehisa, 2014) is one of the first applications of autoencoders in AD. The authors show that autoencoders, through their capability to perform nonlinear reduction in dimensionality, can effectively detect anomalies in a spaceship health monitoring dataset. Further, they compared the performance of autoencoders with Principal Component Analysis (PCA). Their findings indicated that autoencoders performed more effectively than PCA in identifying anomalies, underscoring the benefits of employing non-linear approaches for such tasks. Chen J. et all. (Chen et al., 2017) present an ensemble approach based on autoencoder. The authors use multiple autoencoders, each trained on a different subset of the feature space. This method has demonstrated enhancement in both robustness and accuracy of anomaly detection, especially in high-dimensional datasets. Zhou and Paffenroth (Zhou and Paffenroth, 2017) introduce an autoencoder model for anomaly detection in high-dimensional data. The methodology is based on convolutional architecture, particularly effective for image data. The model outperforms traditional methods on several image datasets. The study (Zong et al., 2018) propose an innovative model by means Deep Autoencoding Gaussian Mixture Model (DAGMM) for anomaly detection. This model combines the advantages of deep autoencoders and Gaussian Mixture Models, providing a powerful tool for capturing complex data distributions. The authors demonstrate that DAGMM outperforms conventional approach. The effectiveness of the model is proven through several datasets, especially in cases which the anomalies are sensitive. Meng-Die Wang et al. (Wang and Takehisa, 2014) proposed a different type of neural network algorithm to detect anomaly pattern in an accident scenario. The authors adopted a particular algorithm called Long Short-Term Memory (LSTM), which is a type of recurrent neural network (RNN) architecture mainly used for time-series data, such as natural language processing and speech recognition. This study focused on 13 anomaly events, each of which event comprises record and states from 26 sensors (6 physical parameter) data for feeding the LSTM. However, this method seems limited in the comparison of the results from analyzed events of the safety systems implemented in NPPs, as the RPS and EFS. This limitation is due to different set-up of input signals. To overcome this problem, the approach used in this study is based on different neural architecture (autoencoder) which employs the same input signals of RPS and EFS systems. Furthermore, LSTM architecture can be resource-intensive and require more time to train due to their complex structure that handles long-term dependencies.

The proposed methodology involves using an autoencoder neural network for anomaly detection, which is linked to a synthetic dataset generated via digital twin of NPP (PCTRAN). This approach compensates for the lack of failure data from nuclear plants (extremely rare), by digitizing the NPP. The integration of the autoencoder with PCTRAN enables a direct comparison with the existing RPS and EFS monitoring systems in NPP. Digital twins can be used for simulation and testing. This means that different scenarios can be simulated to investigate how the system would respond. In this study LOCA event is simulated but other events can be studied.

In this study, eight LOCA were simulated, consisting of four incidents in the cold leg and four in the hot leg, each of which characterized by different pipe break size. In section 2 a description of the PCTRAN model is provided. The methodology based on undercomplete autoencoder is extensively described in section 3. The validation of the proposed methodology was carried out with reference to the study section 5 of Santus et al. (Santus et al., 2020), who used a resonant test rig to perform fatigue tests on corroded drill pipe connections and pipe bodies with varying sizes and steel grades, and a monitoring system consisting of vertical and horizontal laser displacement and strain gauge sensors.

3. Safety system in nuclear power plant

The safety of a nuclear power plant is maintained by means of a variety of complex systems; among them, the most important are the Reactor Protection System (RPS) and the Engineered Safety Features (ESF). The RPS has the dual objectives to prevent hazardous reactor operation and to safeguard against the discharge of radioactive materials. These are accomplished by triggering a reactor trip when safe operational limits are exceeded and activating ESF in the event of an accident. Safe operational limits are established based on the Final Safety Analysis Report of the plant and continuously monitored by local sensors (International Atomic Energy Agency, 2021; U.S. NRC, 2023; U. S. NRC, 2012). When a process signal surpasses a specified set-point, the analog signal is transformed into a digital output by a bistable and monitored by the trip logic matrix. As a consequence of that, the logic matrices determine whether to initiate a reactor trip or an engineered safety features activation. The RPS consists of two separate and independent analog and logic circuit trains. If an analog circuit detects an unsafe condition, it sends signals to the protection system logic cabinets, where the relevant logic contacts are opened. Then, the logic matrices check whether the coincidence for a reactor trip function has been met. If so, the protection system triggers the reactor trip breakers, halting power to the control rod drive mechanisms and causing the injection of control rods into the reactor core. If an accident occurs, the protection system activates the necessary safety equipment. Moreover, the logic trains automatically enable or disable permissive, which are interlocks of protection-grade (U.S. NRC, 2012). The ESF main function is to detect accident-related parameters and trigger equipment that can mitigate the consequences of accidents. It includes activating equipment that removes core decay heat, provides long-term core cooling, terminates steam line breaks, and protects the containment building (last barrier against fission product leakage). The ESF receives inputs from different sources, such as pressurizer pressure, containment pressure, containment radiation, steam generator pressure, steam generator level, 4160 Vac ESF bus voltage, and refuelling water tank (RWT) level. The first four input parameters can be used to detect a loss of coolant accident or steam line break. The steam generator level input triggers the activation of the auxiliary feedwater system (AFW), while the 4160 Vac ESF bus voltage input is responsible for sequencing loads onto the diesel generator during accident conditions. Lastly, the RWT level input prompts the emergency core cooling equipment to switch to long-term core cooling mode of operation. The input signals received from the ESF are used to actuate eight distinct systems, which include the safety injection actuation signal (SIAS), containment spray actuation signal (CSAS), containment isolation signal (CIS), recirculation actuation signal (RAS), containment radiation signal (CRS), steam generator isolation signal (SGIS), auxiliary feedwater actuation signal (AFAS), and emergency diesel generator (EDG) sequencing signal (U.S. NRC, 2012).

4. Nuclear power plant simulator

PCTRAN® is a software for simulating nuclear power plants. It was developed by MST Inc. PCTRAN (International Atomic Energy Agency, 2019) simulator version 6.0.4, can be obtained from the IAEA website. It is built on a standard 2-loop PWR 1800MWth design with inverted U-bend steam generators (SGs) and a dry containment system (looking like Westinghouse, AREVA or Korean Advanced PWR's). Fig. 1 shows the PCTRAN® graphical interface of the plant which includes in a single display all the main synoptic systems.

The components enclosed in red box are numbered in order to uniquely identify them and help the user in understanding how they behave when the plant operating conditions change. This is allowed because the mimic system mimic is interactive. The list of the plant components is provided in Table 1. Moreover, as indicated in (International Atomic Energy Agency, 2019), A and B are used to distinguish the components belonging to the system circuits, facilitating this way the elements loops' detection (Cancemi and Lo Frano, 2023; International Atomic Energy Agency, 2019; Cliff Po, 2009).

The simulator allows to perform, on the personal computer, simulations of transient and accident analysis. The components and features of the plant are represented as graphical elements, organized in a systematic manner on the interfaces. This allows the user to interact with the simulation software by directly changing the setting of the graphical elements.

The MS Access database stores essential plant data, such as geometry, physical parameters, and trip set-points, etc., as well as the initial conditions, such as reactor thermal power, core pressure, core temperature, and fuel-life cycle (International Atomic Energy Agency, 2019). The database also contains various malfunctions, such as LOCA, steam line break, fuel failure during power operation, and anticipated transient without scram. The RPS and EFS are represented at the bottom of the graphical interface (Cheng et al., 2012; Qi et al., 2022). In Table 2 the RPS and Emergency Core Cooling System (ECCS) setpoints implemented in PCTRAN are shown. The PCTRAN model has undergone rigorous benchmarking and verification processes. The study (International Atomic Energy Agency, 2019) presents the validation of the PCTRAN that was carried out referring to the Three Mile Island (TMI) nuclear accident, while the recent study (Qi et al., 2022) considered for validation the Fukushima accident scenario and two other representative accident conditions. Qi et al. (2022) successfully demonstrated the control logic and the transient response of the accidents.

5. Methodology

The methodology proposed in this study is based on undercomplete autoencoder (AE) (Sakurada and Yairi, 2014). An undercomplete autoencoder (AE) is a neural network architecture employed for unsupervised learning tasks, including identifying anomalies. The fundamental concept behind this type of AE is to create a condensed representation of the input data by first mapping it to a lower dimensional space and then mapping it back to its original, higher dimensional space. The encoder segment is characterized by several completely interconnected layers that downsize the dimensionality of the input information. Each of these layers contains specific weights and biases, which are tweaked to decrease the reconstruction error. The output of the encoder is a latent space representation, which is generally significantly smaller compared to the original input data. On the other hand, the decoder comprises a sequence of fully interconnected layers, designed to map the latent space back into the format of the original input data. During the training phase, tuning is made to the weights and biases of decoder with the aim of achieving optimal accuracy in replicating the original input data (Torabi et al., 2023). The encoder function can be mathematically represented by the following equation:

$$h = f(W_e \bullet X + b_e) \quad (1)$$

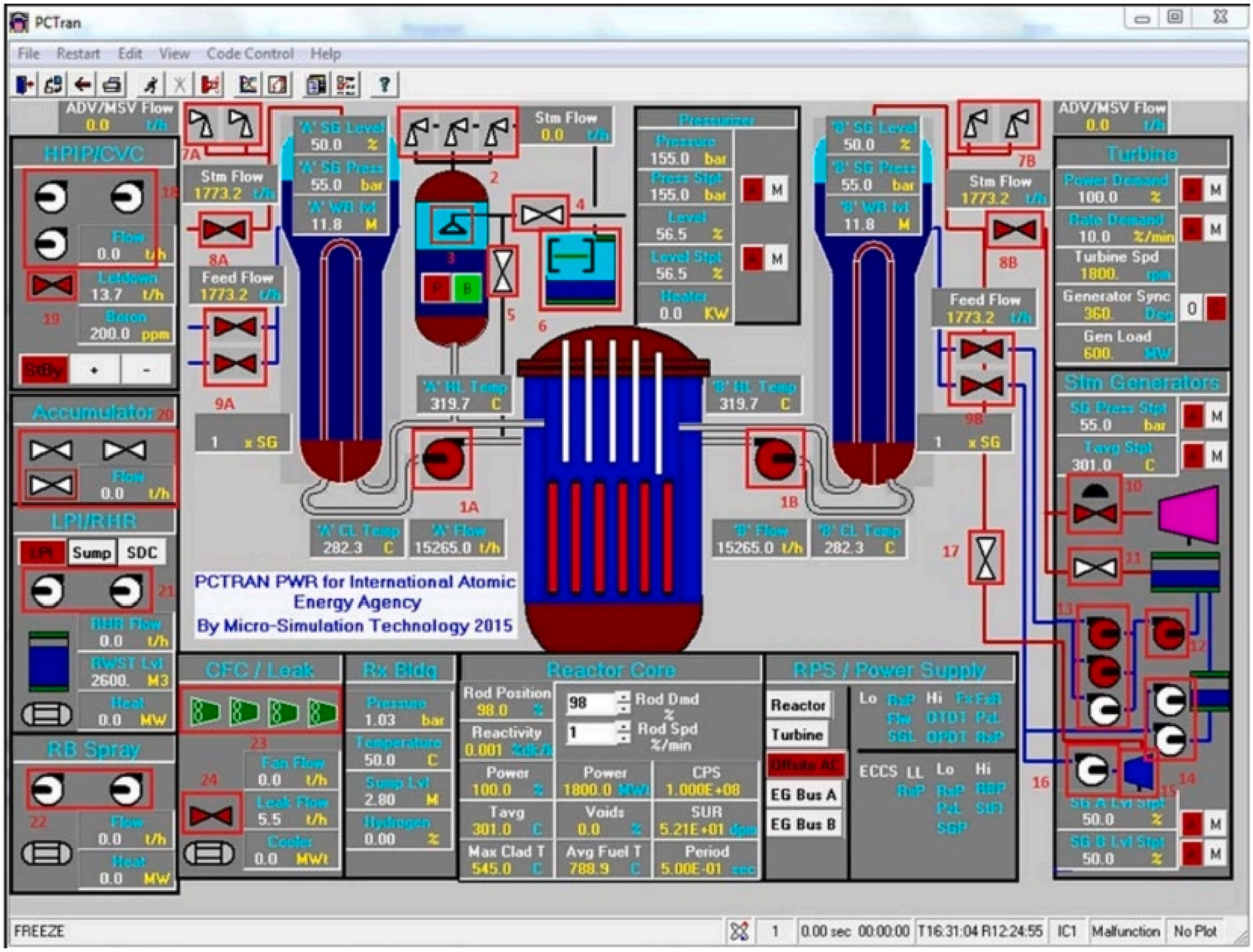

Fig. 1. PCTRAN® plant mimic with the main system (from (International Atomic Energy Agency, 2019).

Table 1
List of Components.

Number	Component	Number	Component
1A, 1B	Reactor Coolant Pumps (RCPs)	13	Feedwater pumps
2	Pilot-Operated Relief Valves (PORV)	14	Motor-driven auxiliary feedwater pumps
3	Pressurizer spray nozzle	15	Turbine of the auxiliary feedwater pump
4	Pressurizer auxiliary spray valve (from CVCS)	16	Turbine-driven auxiliary feedwater pump
5	Pressurizer spray valve (from a cold leg)	17	Valve for steam to the turbine of the auxiliary feedwater pump
6	Pressurizer relief tank	18	High pressure injection (HPI) pumps/Charging pump
7A, 7B	Atmospheric Dump Valves (ADVs) or Main Steam Safety Valves (MSSVs)	19	Letdown valve
8A, 8B	Main Steam Isolation Valves (MSIVs)	20	Accumulators
9A, 9B	Feed Water Isolation Valves (FWIVs)	21	Low pressure injection (LPI) pumps
10	Turbine governor valve	22	Reactor building spray pumps
11	Turbine bypass valve	23	Containment cooling fans
12	Condensate pumps	24	Containment vent valve

Table 2
Trip set-point of RPS and ECCS system.

System	No ID	Description	Value	Unit system
RPS	1	High pressure reactor scram setpoint (bar)	165.7	bar
RPS	2	Low pressure reactor scram setpoint (bar)	132.2	bar
RPS	3	Low SG narrow range scram setpoint (%)	17	bar
RPS	4	High neutron flux reactor scram setpoint (fraction of full power)	1.18	[–]
RPS	5	Low core flow reactor scram setpoint (fraction of full flow)	0.87	[–]
RPS	6	High-high SG level turbine trip setpoint (%)	82	%
ECCS	7	HPI automatic start setpoint (bar)	129.69	bar
ECCS	8	Accumulator initiation pressure setpoint (bar)	43	bar
ECCS	9	LPI system initiation pressure (bar)	11.36	bar
ECCS	10	High RB press for SI initiation (bar)	1.3	bar
ECCS	11	Simultaneous low PZR level with RCS pressure for Safety Injection initiation (fraction of full)	0.15	[–]
ECCS	12	Simultaneous low RCS pressure with Pressurizer level for Safety Injection initiation (bar) 128 Low SG press for SI initiation (bar)	128	bar
ECCS	13	Low SG press for Safety Injection initiation (bar)	38	bar

where X is the input vector data, W_e is the weight matrix for the encoder, b_e is the bias vector for the encoder, f is the activation function used by the encoder, and h is the compressed representation of the input data.Similarly, the decoder function can be mathematically represented as (2):

$$X' = g(W_d \bullet h + b_d) \tag{2}$$

where X' is the reconstructed output, W_d is the weight matrix for the decoder, b_d is the bias vector for the decoder, g is the activation function used by the decoder, and h is the compressed representation of the input data.

Anomalies can be detected by evaluating the magnitude of the mean square error (MSE), by training an AE to minimize the reconstruction error. This error is the criterion used to measure how well the AE has learned the relationships between the features of the input set. If the AE can reproduce the input with good accuracy in the output, the MSE will be small. If not, the AE will produce a larger error. It is possible to identify anomalies by comparing the reconstructed MSE to a predetermined threshold. The MSE is given by:

$$MSE = \frac{1}{N} \times \sum (X - X')^2 \tag{3}$$

Where N is the number of samples in the input data, X is vector of observed values and X' is a vector of predicted values. In order to increase the soundness of model even mean absolute error (MAE) is used.

The lack of failure data presents a significant challenge for developing an accurate anomaly detection methodology, especially in industries such as nuclear power where safety and reliability are critical. To tackle this challenge a synthetic dataset by PCTRAN software is provided.

The dataset contains 8 test cases of LOCA, with four incidents taking place in the cold leg and the other four in the hot leg. Each scenario is distinguished by a different pipe break size. The set-up of all test cases is collected in the Table 3. The delay time refers to the duration it takes for a malfunction to be activated after it has been implemented. In all test cases, the coolant leak is triggered exactly 1000 s after the simulator has been started. The ramp time is the time required to achieve the maximum break size is 2200 s. For sudden changes recorded by the sensors, the traditional monitoring platforms continue to be effective. Based on the test-case analysed, the transient lasts between ~2058 and ~3389 s. In this study, the undercomplete autoencoder consists of three layers, characterized by 10 nodes in the input layer, 4 nodes in the hidden layer, and 10 nodes in the output layer. The model is trained for 100 epochs with a batch size of 10, and 5% of the data is set aside for validation after each epoch. The choice of a proper activation function needs often trials and error approach. The experimental-based activation function is the one that may offers the optimal performance for a particular task. In this study, several trials have been carried out by considering different activation functions for the autoencoder algorithm. Evaluating based on anomaly detection instances, it was found that the ReLU activation function proved to be the most effective for the model proposed (Dubey et al., 2022; LeCun et al., 2012).

Table 3
Set-up of PCTRAN implemented malfunction.

Test Case No.	Event Description	Delay Time [s]	Ramp Time [s]	Pipe Break Size [%]	Nominal Pipe Dimension [cm^2]
1	Loss of Coolant Accident (Cold Leg)	1000	2200	2	100
2	Loss of Coolant Accident (Cold Leg)	1000	2200	5	100
3	Loss of Coolant Accident (Cold Leg)	1000	2200	7	100
4	Loss of Coolant Accident (Cold Leg)	1000	2200	10	100
5	Loss of Coolant Accident (Hot Leg)	1000	2200	2	100
6	Loss of Coolant Accident (Hot Leg)	1000	2200	5	100
7	Loss of Coolant Accident (Hot Leg)	1000	2200	7	100
8	Loss of Coolant Accident (Hot Leg)	1000	2200	10	100

The architecture of the autoencoder is determined by monitoring the model performance based on metrics such as reconstruction error. The weights and biases of the encoder and decoder are adjusted to minimize this reconstruction error using Adam's algorithm (Kingma and Ba, 2017). During the training phase, only the data related to healthy conditions are provided to the model, and no abnormal occurrences are included. The model is trained using the first 975 s of the dataset. As a result, the model can identify anomalies after being trained. However, since the model has only learned event records corresponding to a healthy condition of the component, when the entire dataset, including anomalous events, is supplied to the model, it produces a large reconstruction error. This reconstruction error can be considered as a measure of the "anomaly" of the input signal (Cancemi and Lo Frano, 2022; Cancemi and Lo Frano, 2022).

The threshold is calculated based on the distribution of MSE and MAE. The MSE and MAE distribution of test case 1 are shown in Fig. 2a and Fig. 2b. The learning curve of test case 2 is represented in Fig. 2c. The suggested threshold automatically calculated for test case 1 is k = 0.23 and k = 0.19 for MSE and MAE respectively. The threshold evaluation workflow is applied uniformly across all the test cases. Their initial conditions are as follows: the power of NPP is 100%, the pressure of the Reactor Cooling System (RCS) is 155 bar, the average temperature of the RCS is 306.9° C, and the pressure of the Steam Generator is 70 bar. The PCTRAN simulator can graph 93 variables, which may not always correspond to actual signals from a monitoring system but provides a general overview of the transient. Since the aim of this study is to present a genuine data acquisition system, the input signals utilized by the RPS and EFS systems, which were described in section 3, have been chosen. In addition, Gaussian noise is applied to signal to simulate real platform acquisition. The selected input signals are pressure of RCS (P), level narrow range of both steam generators (NSGA; NSGB), power nuclear flux (PWNT), flow reactor coolant loop A and B (WRCA; WRCB), pressure of reactor building (PRB) and radiation monitoring reactor

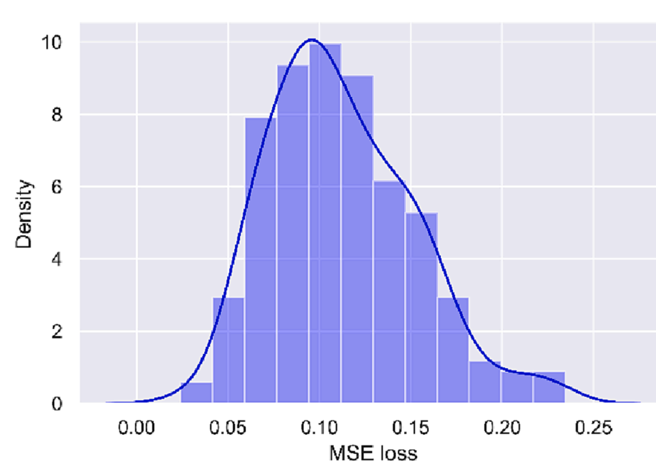

Fig. 2a. Distribution of MSE loss of training set for Test-Case 1.

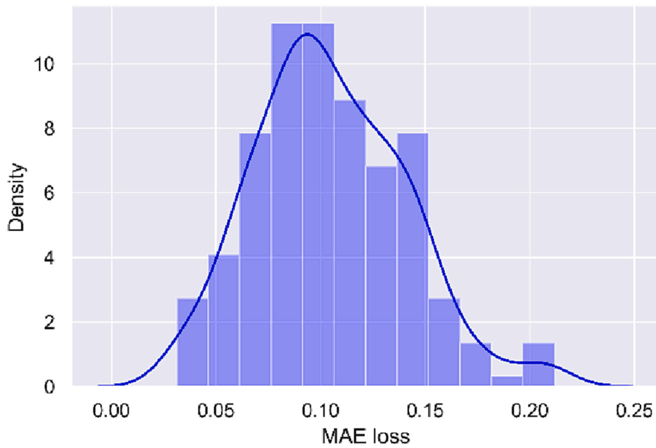

Fig. 2b. Distribution of MAE loss of training set for Test-Case 1.

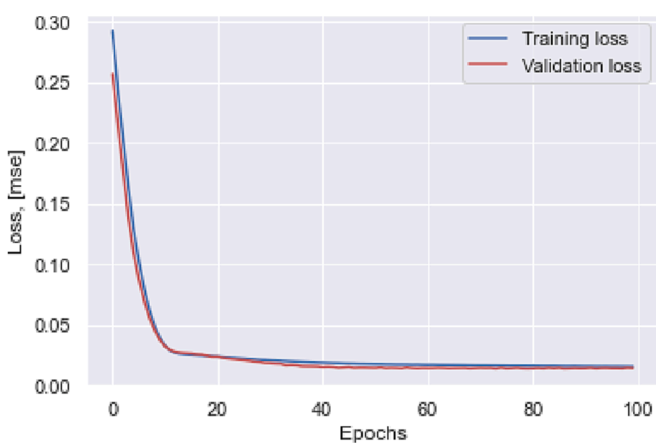

Fig. 2c. Learning curve for Test-Case 2.

building (RM1). The safety system integrated into the simulator was activated based on the trip set-point specified in Table 2. The proposed methodology is summarized in Fig. 4. The superscript in the output vector represents the reconstructed signals after compression into hidden layer (See Fig. 3).

6. Results and discussion

Predictive maintenance is safety-enhancing strategy and a cost-effective approach as well which can be implemented in nuclear industry. Neural networks algorithms can be used for the detection of potential equipment failures before they occur, allowing maintenance teams to take preventative action and avoid costly downtime and safety concerns. The possibility to anticipate anomalous behaviour can be beneficial, particularly in the frame of long-term operation, for preventing it from escalating into a severe accident. The results of all test case are shown in Figs. 4-11 and collected in Table 4. By analysing them, the predictive capacity (or effectiveness) in detecting abnormal events of the developed model is evident. The neural model is able to detect anomaly before the intervention of RPS and ECCS for all the eight test cases. The cold/hot leg pipe break malfunction is set up at the 1000-second. For the purpose of comparison, the analysis includes the reactor scram time and ECCS activation, provided in relation to the time of each event's initiation.

The results obtained taken into account the MSE metric are shown from Figs. 4–11 and describes in the following. In test case 1, the pipe break size is 2 cm^2 as indicated in Table 3. The model successfully detects the initial anomaly at 1095 s, the time instant that precedes the reactor scrams by 2205 s and the intervention of the high-pressure safety injection (HPSI) by 2211 s. In the test case 2, where the break size is 5 cm^2, the reactor scram take place at 2248 s, followed by HPSI intervention at 2254 s. However, the model is able to detect the initial anomaly only after a delay of 65 s from the predetermined malfunction setup, at 1065 s. Despite this delay, the model detects the anomaly much earlier, (1183 s) than the reactor scram and 1189 s prior HPSI activation. Analysing the results of test cases 3–4, it is possible to say that the neural model is able to detect the anomalies 944.5 s and 731 s respectively before the reactor scram, and at 950 s and 954.5 s respectively, earlier

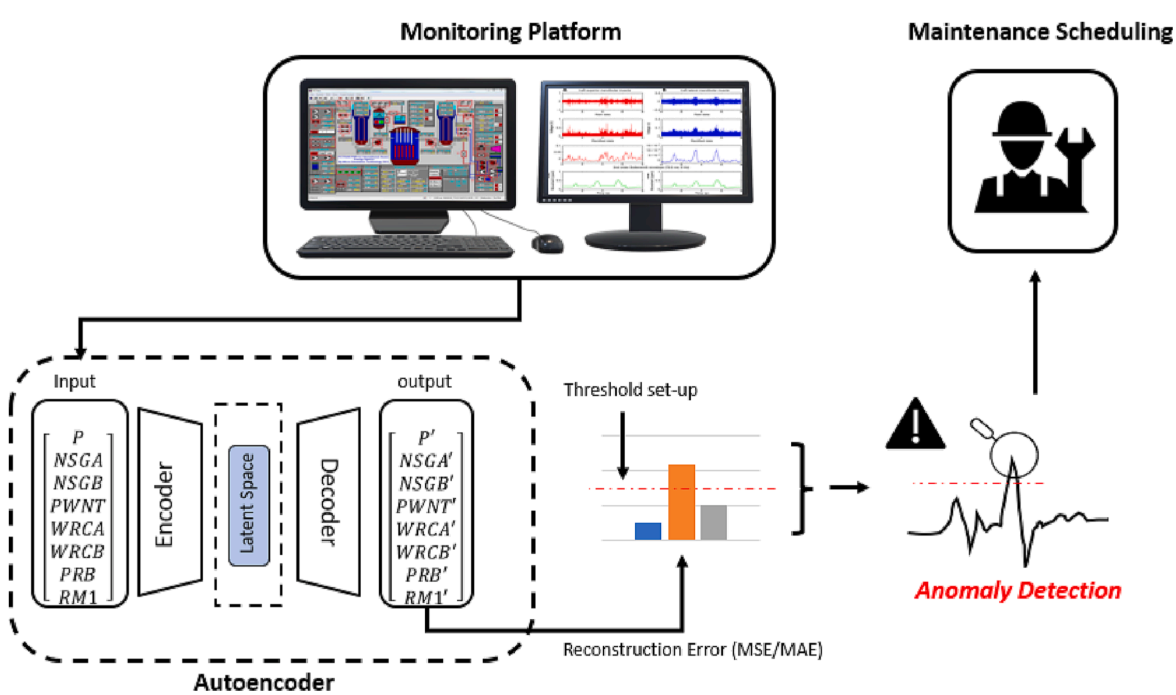

Fig. 3. Methodology Workflow.

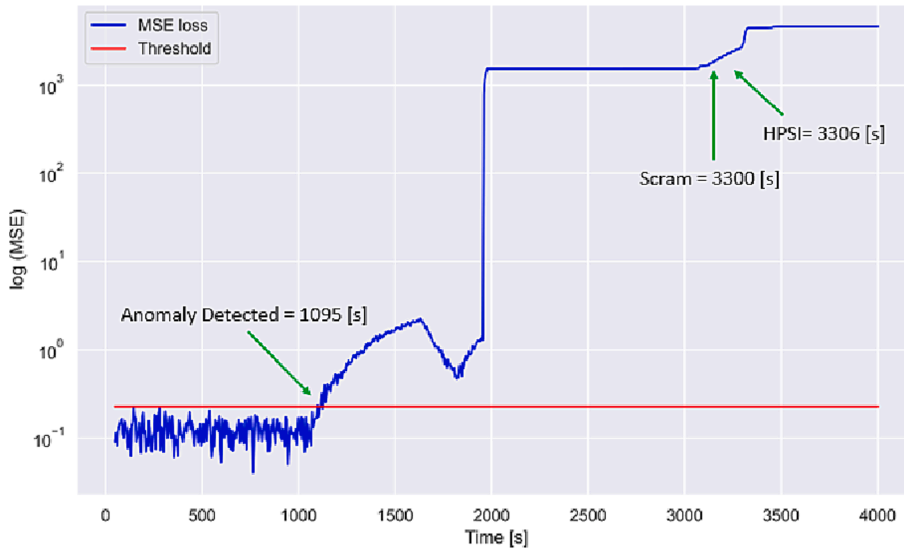

Fig. 4. Test case 1.

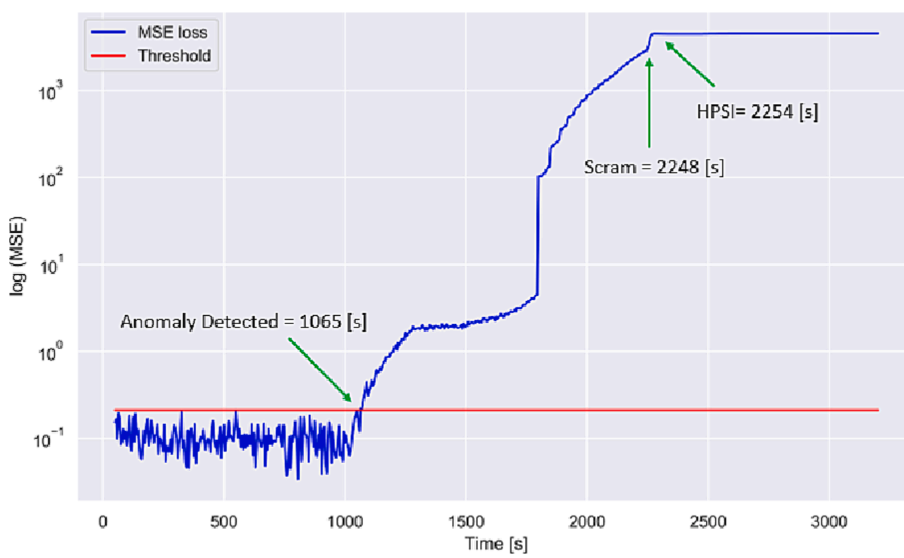

Fig. 5. Test case 2.

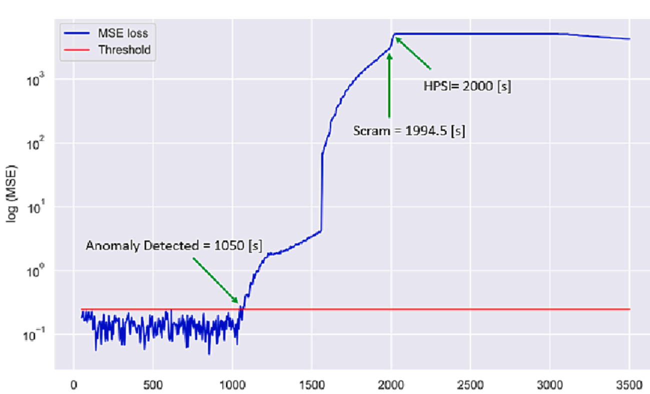

Fig. 6. Test case 3.

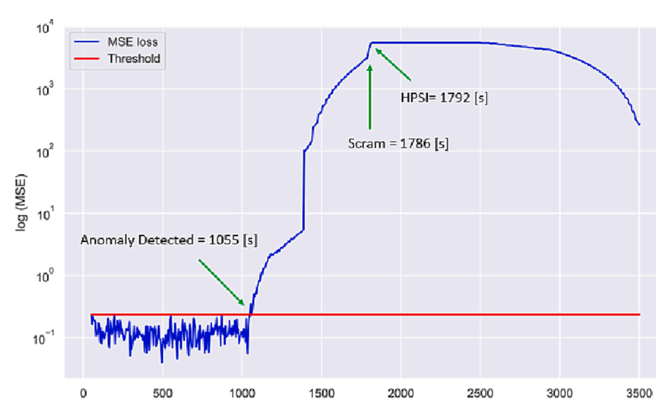

Fig. 7. Test case 4.

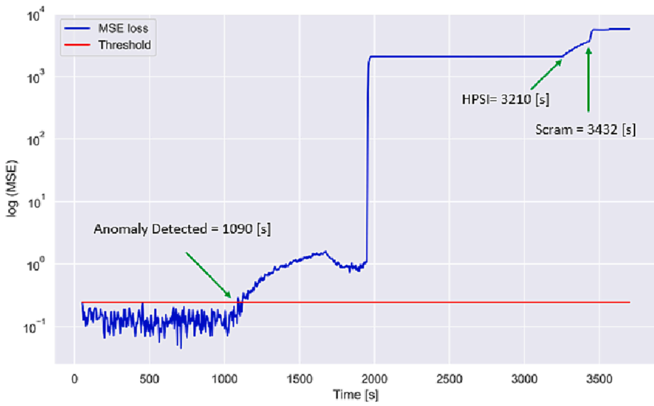

Fig. 8. Test case 5.

than HPSI activation.

With regards to test cases 1–4, the reactor scram occurs as the reactor protection system (RPS) reaches its low-pressure setpoint of 132.2 bar. Additionally, the HPSI is triggered by the automatic start setpoint of the high-pressure injection (HPI) system at 129.69 bar (see Table 2). Test cases 5–8 show a cooling loss in the hot leg of the reactor due to a pipe break failure. The break area for each of these test cases are as follows: 2 cm^2, 5 cm^2, 7 cm^2, and 10 cm^2 respectively. Also in these cases, the model proves to be performing and efficient in detecting anomalies before the RPS and ECCS systems intervention. In test case 5, the model detects the first anomaly at 1090 s, which is significantly earlier than when HPSI and the reactor scram switch activates (at 3210 and 3432 s respectively). The model predicts the anomaly 2120 s before the HPSI intervention and 2342 s the reactor scram. The HPSI is triggered due to high reactor building pressure setpoint reaching 1.30 bar. In test cases 6, 7, and 8, the neural model flags anomalies at 1075, 1065, and 1050 s, respectively. The maximum delay in prediction is observed in the test case 1, where the model detects the anomaly 95 s after the malfunction occurred. However, in all three cases, the model predicts the anomaly well before the safety system intervention, with respective lead times of 1172, 928, and 735 s. The best model performance is observed in test case 1, where the model flags the anomaly 36 min and 45 s before the reactor scram. The anomaly threshold detection differs at least 10 s between the MSE and MAE metrics respectively (test case-1-2-3-5).

7. Model validation

Model validation is an essential step in any machine learning project. The goal of model validation is to estimate the generalization error of the model, which is the error rate that the model is likely to achieve when applied to new, unseen data. To this end, the neural model presented in this manuscript is validated based on the study (Santus et al., 2020), which deals with a resonant test rig designed for fatigue testing of tubular components and describes the rig's operational principle and control strategy. Santus et al. tested corroded drill pipe connections and drill pipe bodies of varying sizes and steel grades, subjecting all specimens to a relatively high stress amplitude under fatigue loading. Results showed that fatigue initiation occurred in the central region of the pipe specimens where stress amplitude was highest and away from the tool joints of connection specimens, mainly due to differences in outer diameters, despite the stress concentration due to the threaded connection itself. Fatigue cracks initiated at pitting corrosion sites, which reduced the potential fatigue strength of drill pipes. The control of bending vibration in that test rig is achieved through the use of strain gauges and laser sensors. In order to verify the methodology, the neural model is utilized to investigate data from the experimental study (Santus et al., 2020). The input signals consist of readings from laser sensors positioned horizontally and vertically along the pipe. Strain gauge sensors are not taken into account due to their accidental failure in several test cases. The model was validated based on 19 test cases. The results of all

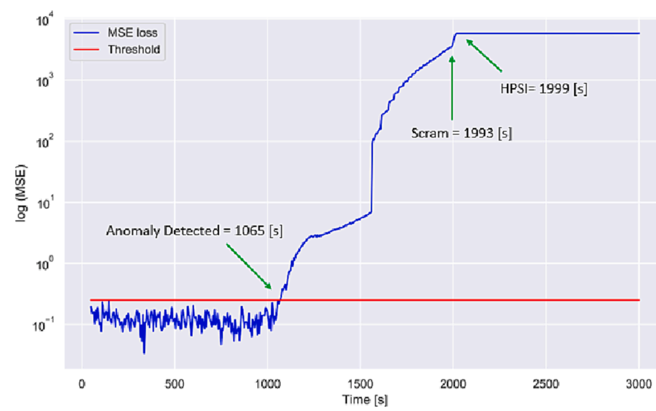

Fig. 10. Test case 7.

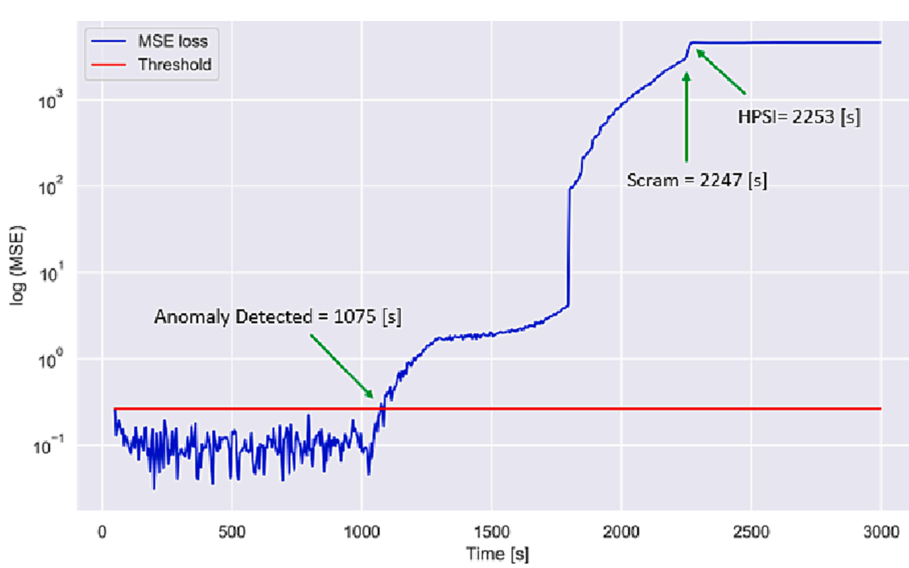

Fig. 9. Test case 6.

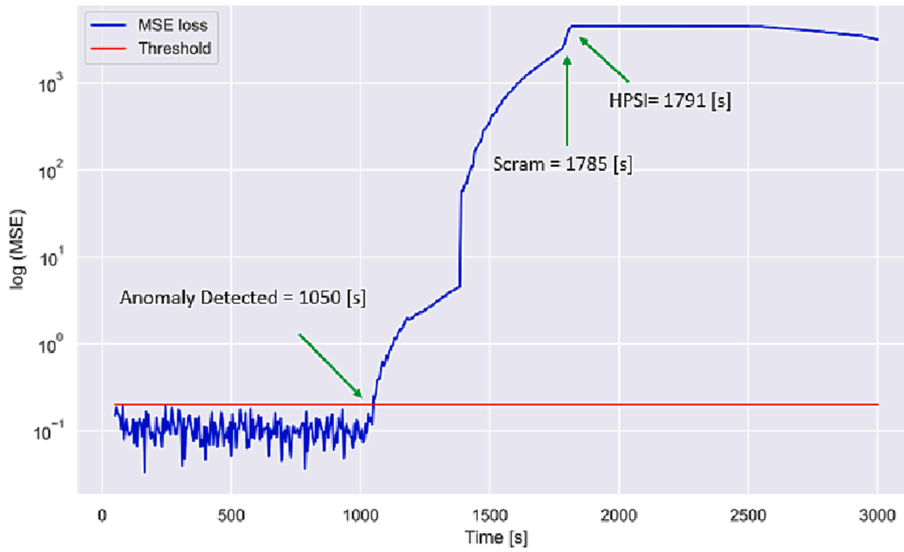

Fig. 11. Test case 8.

Table 4
Results of test case: anomaly detection and reactor scram and ECCs times [s].

No. Test Case	Anomaly detected (MSE) [s]	Anomaly detected (MAE) [s]	Reactor Scram [s]	Trip Setpoint No. ID	ECCS [s]	Trip Setpoint No. ID
1	1095	1105	3300	2	3306	7
2	1065	1055	2248	2	2254	7
3	1050	1040	1994.5	2	2000	7
4	1055	1055	1786	2	1792	7
5	1090	1100	3432	2	3210	10
6	1075	1075	2247	2	2253	7
7	1065	1065	1993	2	1999	7
8	1050	1055	1785	2	1791	7

Table 5
Summary of the validation results.

No. Test Case	Material	Failure Time [s]	1st Anomaly detected [s]	Residual Life [s] (Failure Time – 1st Anomaly detected)
1A	S140	9772	5709	4063
2A	S140	2801	2534	267
3A	S140	3758	3597	161
4A	S140	4684	4340	344
5A	S150	4532	4020	512
6A	S150	3014	254	471
7A	S150	3792	3045	747
8A	S150	5970	5025	945
9A	S150	3909	3029	880
10A	S150	3737	3469	268
11A	S150	4618	4062	556
12A	S150	4382	3783	599
13A	S140	9674	8137	1537
14A	Z140	4699	3135	1564
15A	Z140	4471	2984	1487
16A	Z140	4699	3135	1564
17A	S140	9937	8444	1493
18A	UD165	4317	3567	750
19A	UD165	5333	4002	1331

tests are summarized in the Table 5. To this regard, in Fig. 12 and Fig. 13 the record of displacement signals and results of test case 1A are presented. The algorithm exhibited good performance and generalization, accurately predicting failures in 100% of the analysed cases. Furthermore, in 74% of these cases, the model predicted a pipe break at least 500 s in advance. The highest level of performance was observed in test cases 1A, 13A, 14A, 15A, 16A, 17A, and 19A, with failure predictions ranging between 68 and 22 min (or 102,284 and 33,892 cycle) before the event.

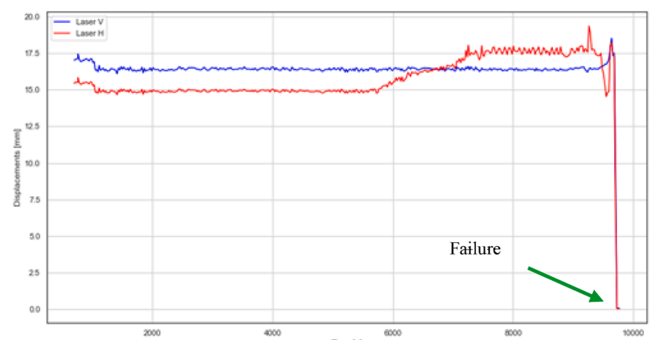

Fig. 12. Vertical and Horizontal Laser Displacements recorded during a test.

8. Conclusion

This study has proposed an innovative methodology based on unsupervised neural network to predict potential anomalies in the cooling system of a PWR: a breaching in the hot leg and cold leg was considered purposely. The size of pipe break was varied so to consider multiple scenarios.

The neural network has been fed by synthetic dataset provide by PCTRAN simulator and the efficiency in event identification was studied. Particularly the abnormal event was detected if the differences between the acquired and predicted readings exceed the pre-set threshold. The main conclusions are as follow:

- The neural model successfully detects anomalies in 100% of test cases before reactor scram and ECCS activation.
- The first anomaly is detected by the model within just ~1 min of its implementation in the reactor simulator.
- Anomalies are detected up to 38 min before the safety system activation.
- Training the unsupervised model requires a large amount of data.
- Lack of data can be overcome by using a digital twin of the reactor.

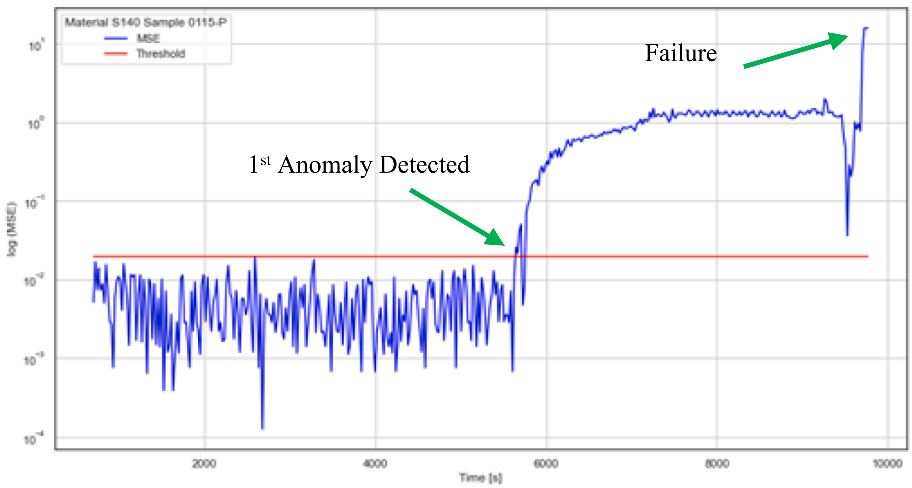

Fig. 13. Results of test case 1A.

- The neural model is validated based on experimental campaign of (Santus et al., 2020).
- The model performs well for slow deviations from the nominal pattern, but its performance decreases when deviations are impulsive in nature (such as large pipe breaks).

DT technology provides a powerful tool for predictive maintenance, especially when combined with machine learning methods like autoencoders, there are some potential differences and challenges when transitioning from synthetic to real-world data:

- Data Quality and Completeness: DT data is typically clean and complete because it's generated under controlled conditions. Real-world data, on the other hand, can be noisy, have missing values, or be affected by outliers. This may require additional preprocessing steps or robust methods to handle such inconsistencies in the real-world data.
- Variability and Uncertainty: DT data is generated based on models and simulations, which might not fully capture the inherent variability and uncertainty of real-world operations. Unexpected events, human errors, and other factors not included in the model can create discrepancies between the digital twin and the actual system.
- Computational Considerations: While digital twin data allows for robust testing and validation without affecting real-world operations, the computational cost of creating and maintaining a digital twin might be high. In addition, the model trained on digital twin data would need to be efficient enough to run in a real-time or near-real-time setting for practical use in predictive maintenance.
- Model Generalization: A model trained on synthetic data might not generalize well to real-world data due to differences in distributions or dynamics of the data. It's essential to validate and fine-tune the model with real-world data to ensure its performance.

It is beneficial to incorporate real-world data into the training process as much as possible and validate the model's performance on real-world data to ensure its applicability and effectiveness in a real-world setting.

CRediT authorship contribution statement

S.A. Cancemi: Writing – review & editing, Software. **R. Lo Frano:** Conceptualization, Methodology, Writing – original draft, Funding acquisition, Writing – review & editing. **C. Santus** Writing – review & editing. **T. Inoue** Revision, Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

We would like to express our gratitude to Alessandro Burchianti of ACTA Srl, Rosignano (LI), Italy, for carrying out the crucial testing processes essential to this study.

References

ASME III, Division 1 – subsection NB-3232, ASME, Boiler and Pressure Vessel Committee, 1980.
Cancemi S.A., Lo Frano R., (2022), Unsupervised Machine Learning approach for NPP LTO Program. In: 31st International Conference Nuclear Energy for New Europe (NENE), September 12-15, 2022, Portoroz, Slovenia.
Cancemi S.A., Lo Frano R., (2022). Anomalies detection in structures, system and components for supporting nuclear long term operation program. In: Proceedings of the 2022 29th International Conference on Nuclear Engineering. Volume 15: Student Paper Competition. Virtual, Online. August 8–12, 2022. V015T16A110. ASME. 10.1115/ICONE29-93193.
Cancemi S.A., Lo Frano R., (2023). Unsupervised Neural Network Approach for Identifying Anomalous Events in Nuclear Power Plants. Proceeding of ICONE 30, May 21-26, 2023. Kyoto (J)[In press].
Chen, J., Sathe, S., Aggarwal, C., Turaga, D., 2017. Outlier detection with autoencoder ensembles. In: Chawla, N., Wang, W. (Eds.), Proceedings of the 2017 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, Philadelphia, PA, pp. 90–98.
Cheng, Y.-H., Shih, C., Chiang, S.C., Weng, T.L., 2012. Introducing PCTRAN as an evaluation tool for nuclear power plant emergency responses. Ann. Nucl. Energy 40 (1), 122–129. https://www.sciencedirect.com/science/article/pii/S0306454911004233.
Cliff Po, L.-C., 2009. Personal Computer Transient Analyzer for a Two-loop PWR and TRIGA Reactor. International Atomic Energy Commission Workshop on Nuclear Power Plant Simulator for Education Trieste, Italy.
Deloitte. (2015). Predictive Maintenance: How Data Can Increase Production.
Dubey, S.R., Singh, S.K., Chaudhuri, B.B., 2022. Activation functions in deep learning: A comprehensive survey and benchmark. Neurocomputing 503, 92–108.
González-González, R., Montero-Mayorga, J., García-Melara, J., 2023. Investigation of small break LOCA with failure of the emergency core cooling system accident management in VVER-1000 reactor type. Ann. Nucl. Energy 183, 109632. https://doi.org/10.1016/j.anucene.2022.109632.
International Atomic Energy Agency, 2018. Maintenance Optimization Programme for Nuclear Power Plants. IAEA.
International Atomic Energy Agency, 2019. PCTRAN Generic Pressurized Water Reactor Simulator Exercise Handbook, IAEA-TCS-68. IAEA, Vienna.

International Atomic Energy Agency. (2016). Safety Standards for Protecting People and the Environment Specific Safety Requirements, Safety of Nuclear Power Plants: Design (No. SSR-2/1 (Rev. 1)). Vienna: IAEA. ISBN 978-92-0-109315-8.
International Atomic Energy Agency (2021). Format and Content of the Safety Analysis Report for Nuclear Power Plants, IAEA Safety Standards Series No. SSG-61, IAEA, Vienna.
Kingma D. P., Ba J. (2017). Adam: A Method for Stochastic Optimization. arXiv, doi: 10.48550/arXiv.1412.6980.
LeCun, Y.A., Bottou, L., Orr, G.B., Müller, KR. (2012). Efficient BackProp. In: Montavon, G., Orr, G.B., Müller, KR. (eds) Neural Networks: Tricks of the Trade. Lecture Notes in Computer Science, vol 7700. Springer, Berlin, Heidelberg. 10.1007/978-3-642-35289-8_3.
Lewis, E.E., 1977. Nuclear Power Reactor Safety. John Wiley & Sons Inc, New York.
PwC. (2017). Predictive Maintenance 4.0: Predictive Maintenance in the Industrial Internet of Things.
Qi, B., Xiao, X., Liang, J., Po, L.-C., Zhang, L., Tong, J., 2022. An open time-series simulated dataset covering various accidents for nuclear power plants. Sci. Data 9 (1). https://doi.org/10.1038/s41597-022-01879-1.
Sakurada M., Yairi T., (2014). Anomaly detection using autoencoders with nonlinear dimensionality reduction. In: MLSDA'14: Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis, Dec. Pages 4-11, 10.1145/2689746.2689747.
Santus C., Burchianti A., Inoue T., Ishiguro H. (2020). Fatigue resonant tests on S140 and S150 grade corroded drill pipe connections and pipe bodies, Int. J. Pressure Vessels Pip., 184, 104107., 10.1016/j.ijpvp.2020.104107.
Torabi, H., Mirtaheri, S.L., Greco, S., 2023. Practical autoencoder based anomaly detection by using vector reconstruction error. Cybersecurity 6, 1. https://doi.org/10.1186/s42400-022-00134-9.
U.S.NRC. (2012). Westinghouse Technology Systems Manual Section 12.1 Reactor Protection System.
U.S.NRC (2012). Westinghouse Technology Systems Manual, Section 12.3: Reactor Protection System – Engineered Safety Features Actuation Signals.
U.S. NRC. 52.157 Contents of applications; technical information in final safety analysis report. (n.d.). https://www.nrc.gov/reading-rm/doc-collections/cfr/part052/part052-0157.html. Accessed 10 Jan. 2023.
U.S.NRC. 50.46 Acceptance Criteria for Emergency Core Cooling Systems for Light-Water Nuclear Power Reactor. https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-0046.html. Accessed 10 Jan. 2023.
U.S.NRC (2021). Report TLR/RES-DE-REB-2021-01. The State of Technology of Application of Digital Twins.
Veshchunov, M.S., Boldyrev, A.V., Shestak, V.E., 2022. LOCA advanced failure criteria: Implementation in the mechanistic fuel performance code and practical applications. Nucl. Eng. Design 381, 112039. https://doi.org/10.1016/j.nucengdes.2022.112039.
Wang, M.-D., Takehisa, Y., 2014. Anomaly detection using autoencoders with nonlinear dimensionality reduction. In: Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis (MLSDA'14). Association for Computing Machinery, New York, NY, USA, pp. 4–11. https://doi.org/10.1145/2689746.2689747.
Zhou, C., Paffenroth, R.C., 2017. Anomaly Detection with Robust Deep Autoencoders. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '17). Association for Computing Machinery, New York, NY, USA, pp. 665–674. https://doi.org/10.1145/3097983.3098052.
Zong, B., Song, Q., Min, M.R., Cheng, W., Lumezanu, C., Cho, D., Chen, H., 2018. Deep autoencoding gaussian mixture model for unsupervised anomaly detection. Proceedings of the International Conference on Learning Representations.