Transactions of the SMiRT 27
Yokohama, Japan, March 3-8, 2024
Division XII

FALSE SENSOR-DATA DETECTION STRATEGY FOR POST-HAZARD CONDITION MONITORING OF NUCLEAR SYSTEMS USING STATISTICAL APPROACHES AND LONG SHORT-TERM MEMORY

Joomyung Lee[1], Harleen Kaur Sandhu[2], Saran Srikanth Bodda[3], Abhinav Gupta[4], and Nam Dinh[5]

[1]Postdoctoral Research Scholar, Department of Nuclear Engineering, NCSU, USA
[2]Postdoctoral Research Scholar, CNEFS, NCSU, USA (hksandhu@ncsu.edu)
[3]Research Faculty, CNEFS, NCSU, USA
[4]Director, CNEFS, NCSU, USA
[5]Professor, Department of Nuclear Engineering, NCSU, USA

ABSTRACT

The upcoming generation of nuclear power plants is integrating an autonomous OnLine Monitoring (OLM) system for safe and efficient operation. The OLM system's accuracy relies on high-quality data from the nuclear facility. Erroneous sensor data can adversely affect the OLM system's performance, leading to undetected issues such as degraded locations during events like earthquakes. This study introduces a False Signal Detection and Correction Model (FSDCM) with two components: one for detecting false sensor data through statistical correlation analysis and another for correction using deep learning algorithms. The FSDCM ensures accurate sensor data by overwriting the acquired false data with that from previously trained deep learning algorithms. A case study on a nuclear piping-equipment system validates the FSDCM's effectiveness in detecting and correcting false sensor data, thereby enhancing data reliability for the OLM system. This approach improves the plant's operational condition recognition, contributing to secure and efficient operations.

INTRODUCTION

Proactively identifying equipment malfunctions and degradation in structures, systems, and components (SSCs) of nuclear power plants is vital for maintaining uninterrupted daily operations (Sandhu et al., 2022b). Usually, vibrational health monitoring, as a form of condition monitoring, is implemented to assess the present state of the nuclear power plant and its safety-related systems. Optimally placed sensors in a nuclear plant gather data for effective operation monitoring. Precise data from the vibration monitoring system is crucial for accurate diagnoses. Therefore, the detection of anomalous sensor data is important for the safety and security of nuclear facilities. The operator or the condition monitoring system should not only perceive the suspicious data but also rectify the data by replacing it with reliable data.

Next-generation nuclear power plant operations depend on the research of the Digital Instrumentation and Control (DIC) framework, which incorporates the OnLine Monitoring (OLM) system. The U.S. Nuclear Regulatory Commission's (NRC) Office of Nuclear Regulatory Research (RES) collaborates with diverse research institutions in the development of the DIC system. This collaboration highlights the integration of advanced tools and methodologies, with a focus on digital data, OLM, wireless technologies, and the formulation of guidelines to include success criteria for the DIC applications (Sandhu et al., 2023b; USNRC, 2022). For researching the development and performance of Online Monitoring (OLM) systems, a significant challenge arises due to the inherent vulnerability of the analysis functions within contemporary systems to false sensor data. Inaccurate sensor data poses a significant risk to safety in nuclear power plants, as it not only introduces potential cybersecurity issues but also a possibility for deliberate manipulation of the data by malicious hackers. This highlights the importance of robust frameworks to ensure the integrity and reliability of the information collected from sensors in critical infrastructure systems.

Identifying false sensor data is difficult, particularly when signal data is not immediately interpretable. Common sensor-related measurement errors include random noise and systematic errors, some of which can be easily noticed by skilled operators. For instance, if an illogical data trend appears from a single thermocouple among numerous sensors in a system, operators may question the sensor's reliability. Similarly, during a seismic event, if the measured acceleration-time series signal remains constant, operators may suspect a malfunction based on their engineering intuition. However, distinguishing between actual and false sensor data becomes complex when raw sensor data from nuclear piping-equipment systems, represented as acceleration-time series signals during an earthquake, appears arbitrary. Even experienced operators may struggle to differentiate between genuine and erroneous data in acceleration-time series signals contaminated by randomly distributed errors. Detecting false sensor data in such scenarios requires in-depth analysis, utilizing statistical approaches like correlating sensor signals with physical variables.

Detecting damage in vital SSCs of a nuclear facility is crucial for predictive maintenance, and it relies on high-quality sensor data. A recent review (Yi et al., 2017) discusses frameworks for validating sensors in structural health monitoring, covering methodologies like statistical analysis, probability quantification, optimization, regression, and data denoising. Other research (Li et al., 2022) focuses on anomaly detection in time-series sensor data using methods like Pearson and Spearman correlation coefficients, spatial correlation, and temporal correlation graphs with Graph Neural Network (GNN). A recent study (Yu et al., 2022) uses a sequential probability ratio test to capture sensor failure in a nuclear power plant, defining failure as minute drifts in signal data. However, these studies mainly focus on the detection and isolation of false sensor data without regenerating correct synthetic data. In post-hazard scenarios for nuclear facilities, correlation analysis becomes essential for detecting false data in sensor responses from piping-equipment systems (Sandhu, 2021). While past studies on non-nuclear structural applications show promising results for detecting sensor malfunctions, the dynamic sensor response from nuclear systems, characterized by various low- and high-frequency vibration modes, requires further investigation (Sandhu et al., 2022b).

This paper aims to develop a novel algorithm for false sensor data detection and correction. The novel False Signal Detection and Correction Model (FSDCM) proposed in this research consists of two functional components: 1) detection of false sensor data based on statistical correlation analysis, and 2) correction of false data by implementing and designing deep learning algorithms. The performance of the proposed FSDCM is evaluated for the monitoring of nuclear piping-equipment systems against seismic hazards. It focuses on three aspects: the target system, event types, and manipulation of sensor data errors. The erroneous sensor responses can be comprehended by a comparative numerical analysis since, within a system, the temporal changes in physical variables must exhibit a consistent dynamic pattern. After filtering out the false data, the FSDCM can overwrite the inferred sensor data by replacing it with data obtained from previously trained deep learning algorithms. Hence, the Online Monitoring (OLM) system can reliably operate with corrected sensor data.

For the case study, false sensor data acquired from accelerometers on a nuclear piping-equipment system is considered for vibration condition monitoring against seismic hazards. The functionality of piping-equipment systems within nuclear power plants is crucial for the safe transport of coolants, such as water, gas, or molten sodium, and the interconnections of these systems are essential for the overall safety of the facility. A finite element model of the piping-equipment system is created, and simulations are conducted to collect sensor data. To introduce error mode manipulation, random signals are injected into the system. The proposed False Sensor Data Correction Method (FSDCM) offers several innovations, including the ability to identify false data even in scenarios with randomly distributed errors, the capacity to rectify inaccurate data using AI-generated sensor data, and applicability to support the Online Monitoring (OLM) system.

This study implements a correlation analysis between various sensor data from a nuclear piping-equipment system and develops a deep learning algorithm utilizing recurrent neural networks (RNNs). The subsequent sections present the model's technical aspects and an evaluation of its performance.

OVERVIEW OF STATISTICAL AND AI APPROACHES

The correlation coefficient quantifies the statistical connection between two variables through a numerical value ranging from -1 to 1. It's crucial to note that this measure doesn't imply causation, and drawing cause-and-effect conclusions based solely on the numerical value is inappropriate. As explained in a previously published book (Evans, 1996), the degree of association is verbally characterized using the coefficient, and illustrated in Table 1. In this study, three types of coefficients—Pearson correlation coefficient (PCC), Spearman correlation coefficient (SCC), and Kendall correlation coefficient (KCC)—are employed for time-dependent correlation analysis.

Table 1: Interpret Correlation Coefficients

Coefficient	**Verbal Description**
0 - 0.19	Very Weak
0.2 - 0.39	Weak
0.4 – 0.59	Moderate
0.6 – 0.79	Strong
0.8 – 1	Very Strong

In line with the objective of detecting erroneous sensor behavior, the correlation coefficients between two variables are computed with a temporal consideration. This involves revealing the correlations between two dynamically changing variables by monitoring the temporal shifts in correlation coefficients. For the specific case study outlined in the paper, all three correlation coefficients—Pearson, Spearman, and Kendall—are employed to identify the relationship between sensors and identify instances of inaccurate sensor data.

Machine Learning (ML) relies on Artificial Neural Networks (ANNs), computational models inspired by the human brain. Deep Neural Networks (DNNs) excel in extracting complex features due to multiple hidden layers. DNNs, with non-linear functions, are proficient in tasks like object and speech recognition. Recurrent Neural Networks (RNNs) process sequential data, retaining information through a feedback loop, facilitating dynamic pattern identification. Challenges like the vanishing gradient problem in deep learning are mitigated by integrating Long Short-Term Memory (LSTM) units with activation functions like ReLU. The LSTM module in the RNN determines information incorporation and retention. In this study, an RNN with an LSTM module predicts acceleration-time series sensor data contaminated by false signals.

PROPOSED APPROACH WORKFLOW

The False Signal Detection and Correction Model (FSDCM) comprises two primary functions: a detecting function and a correcting function. To achieve the model's goals, the correction of false signals follows their detection. The operational steps of FSDCM involve data collection, detecting false data, distinguishing false data from true data, and substituting removed sensor data with corrected data. To formulate the functions of the model, the workflow of FSDCM is outlined in Figure 1. The steps for the workflow comprise:

- *Simulation data generation*: A finite element (FE) model is created to obtain the system's sensor responses due to seismic events. The simulation data against a variety of earthquake loads would be used to develop and test the model. For this research, ANSYS is used to design the FE model of the piping system and simulate the earthquake responses.

- *Time-dependent correlation analysis*: Based on the simulation data, false data sets are introduced by manually injecting errors. Correlation analysis between temporal data sets, including true and manipulated data, is used to identify how the correlation changes with time to detect the error injection.

- *Machine learning-based correction model development*: The main objective of the correcting function is to infer removed false sensor data from the true sensor data. Therefore, long short-term memory (LSTM), a type of recurrent neural network (RNN), is used as the machine learning model to deal with sequential simulation data for data training, validation, and testing of the model.

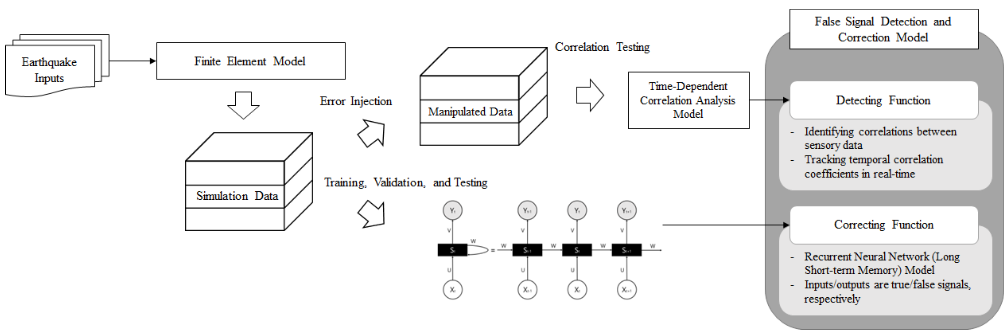


Figure 1. Proposed workflow

The proposed framework is developed using a three-dimensional piping system, as shown in Figure 2. An FE model is created to simulate the system's response to 10 earthquake loads (Sandhu et al., 2023a). Triaxial sensor responses are collected at nodes 2 to 10. This study considers the degradation of pipe walls over time due to flow-accelerated corrosion and erosion. The FE model is updated to reflect degradation at specific locations—nodes 1, 3, 5, 7, or 9—by reducing pipe wall thickness. Degradation severity is classified as Minor, Moderate, or Severe, with the uncertainty introduced through the following uniform distribution: (i) Minor: $[20\% - 30\%]$, (ii) Moderate: $[45\% - 55\%]$, and (iii) Severe: $[70\% - 80\%]$ (Sandhu et al., 2022a,c). Monte-Carlo simulation is used to generate random values, and to produce sensor responses for various locations and degradation severity levels (Bodda, 2018; Sandhu, 2015; Sandhu et al., 2019).

Two instances of false data injection are examined in this study: one involves the manipulation of data from a singular sensor, specifically sensor 3 data, while the other entails the manipulation of data from multiple sensors, namely sensor 3 and sensor 7 data. The details regarding the nature of data manipulation are elucidated in the following section.

DEVELOPMENT OF DETECTION STRATEGY

Correlation coefficient consistency test

The principal hypothesis under investigation is that physical symptoms within a given system should manifest consistent responses in the context of a particular event. To substantiate this hypothesis, a consistency test is conducted. In the course of this study, 10 earthquake records and their corresponding

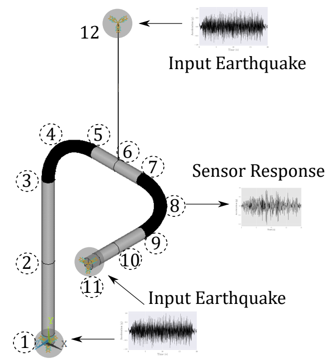


Figure 2. Simple Piping System

simulations are executed to obtain sensor responses. The Pearson correlation coefficients (PCCs), Spearman correlation coefficients (SCCs), and Kendall correlation coefficients (KCCs) computed between variables derived from the case of #2-earthquake load and its simulation are presented in Figure 3, Figure 4, and Figure 5.

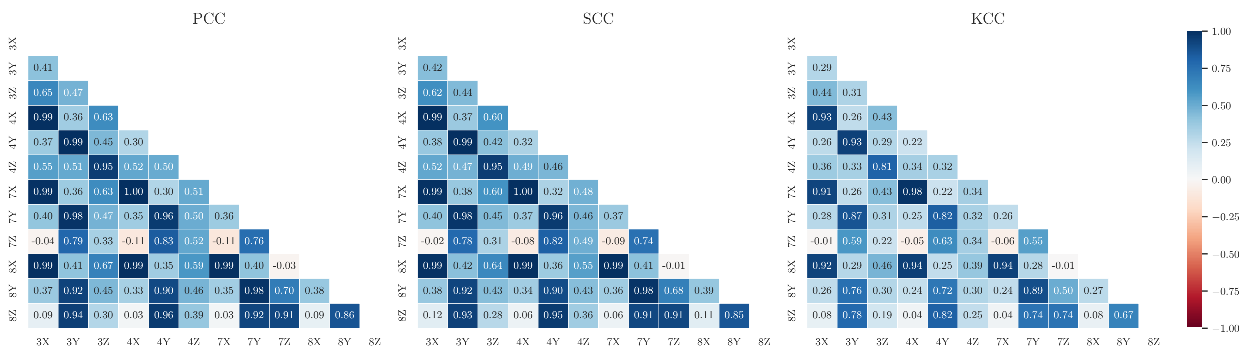


Figure 3. Correlation coefficients between sensor data

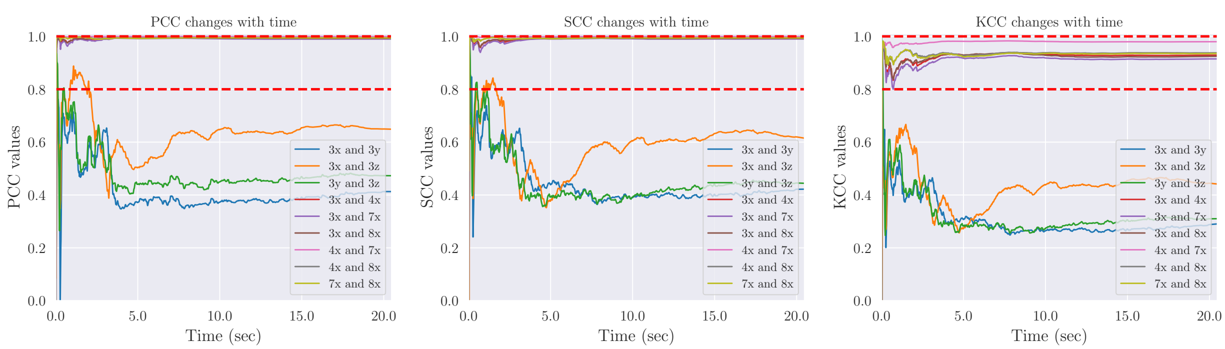


Figure 4. Time-dependent correlation coefficients analysis between sensor data

Strong positive correlations exist between responses in the same direction, notably seen in PCC and SCC analyses. Time-dependent correlation analysis reveals rapid convergence of correlations within the sensor responses acquired in the same directions and fluctuations between different direction responses over time. While KCCs between the same directions are relatively smaller, they still indicate robust correlations, as represented by the red dashed lines in Figure 4. Figure 5 confirms consistency between the same direction responses across all cases, even though some cases show lower KCC values. In conclusion, the main hypothesis is validated as responses to seismic events consistently maintain coherence within a system for vibrations in the same direction.

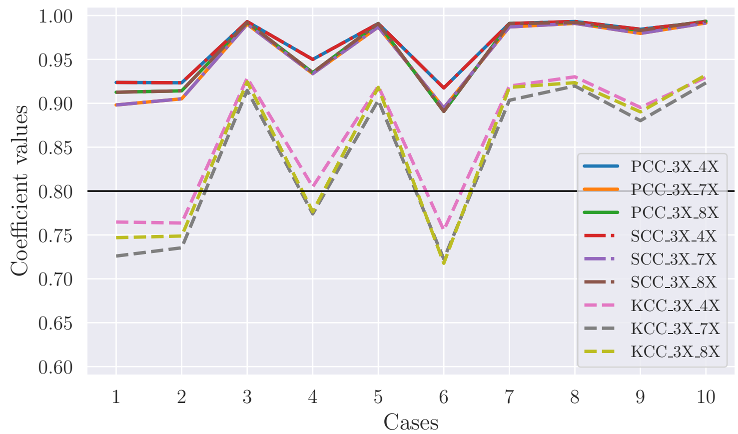


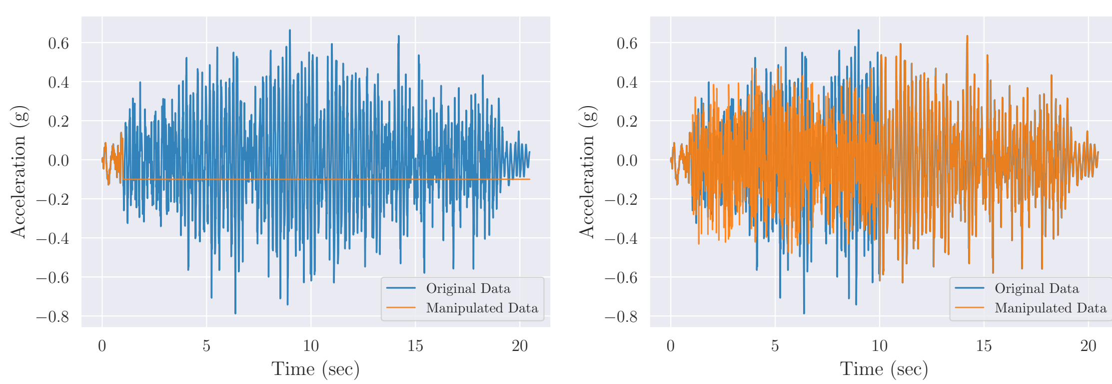
Figure 5. Correlation coefficient analysis for all the cases

Data manipulation

The primary challenge in false sensor detection arises from the engineer's reliance on intuition to differentiate suspicious signals from genuine data, particularly due to the transient nature induced by seismic events. Two potential behaviors of sensor malfunction are contemplated:

- Case 1: fixation at a constant value (Figure 6a), and,
- Case 2: random signal injection using normal distributions (Figure 6b).

The time of error occurrence and the duration of the error serve as independent variables employed to generate false sensor signal data.



(a) Fixation at Constant Value

(b) Random Signal Injection

Figure 6. Malfunction at sensor 3

Time-dependent correlation analysis for model performance

A time-dependent correlation model is formulated to identify the sensor generating false data and identify the timing of the erroneous sensor behavior. The results from case 1 and case 2 are illustrated in Figure 7a and Figure 7b respectively. The Pearson, Spearman, and Kendall correlation coefficients between the false signal and the true signal exhibit a gradual decrease over time. Specifically, in case 1, the correlation coefficients (3X-4X, 3X-7X, and 3X-8X) associated with sensor 3 transition from the "very strong correlated" range to the "very weak correlated" range. This shift indicates that the response from sensor 3 becomes unrelated to the responses from other sensors. The decline in correlation coefficients commences concurrently with the error occurrence (1 second for sensor 3 and 3 seconds for sensor 7). Therefore, the analysis model provides insights into when the false sensor data begins. Similar to case 1, three correlation coefficients between false and true signals in case 2 begin decreasing at the error injection time. Although they tend to slowly recover after the random signal injection ends, the restored correlations are limited and not fully recovered. In both cases, all three correlation coefficients contribute to the model's efficacy in identifying the false sensor number and the time of error occurrence.

DEVELOPMENT OF DATA CORRECTION STRATEGY

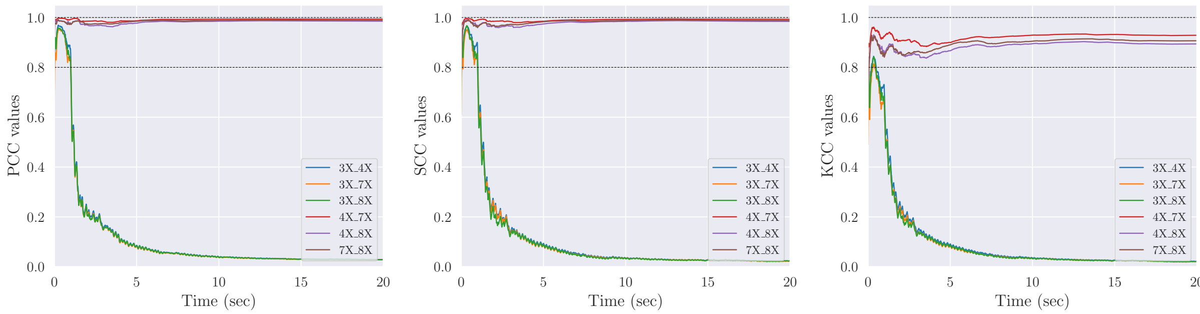
An additional innovative aspect of the Fault-Sensor Data Correction Module (FSDCM) is its integration of a Recurrent Neural Network (RNN) model to address data discrepancies. The model employs an inference process to identify false sensor data by comparing it with the authentic signal data. The RNN model is constructed, training of the model is conducted using past sensor data, and the performance of the trained model is demonstrated using false sensor data. To correct previously detected false signals, the input involves true sensor data, and the output comprises false sensor data. The number of input/output features depends on the specific case study, considering the three directions (X, Y, and Z) for each sensor data. This study employs cross-validation, excluding the test set from training, to prevent overfitting. Overfitting hinders model generalization on new data. Using a three-way hold-out method, the dataset is divided into training (70%), validation (20%), and test (10%) sets. The training set updates LSTM model parameters, the validation set validates during training, and the test set evaluates final performance. The Reduce Learning Rates on Plateau (RLRP) strategy adjusts the learning rate to enhance model performance, where the learning rate governs model parameter updates.

KEY RESULTS

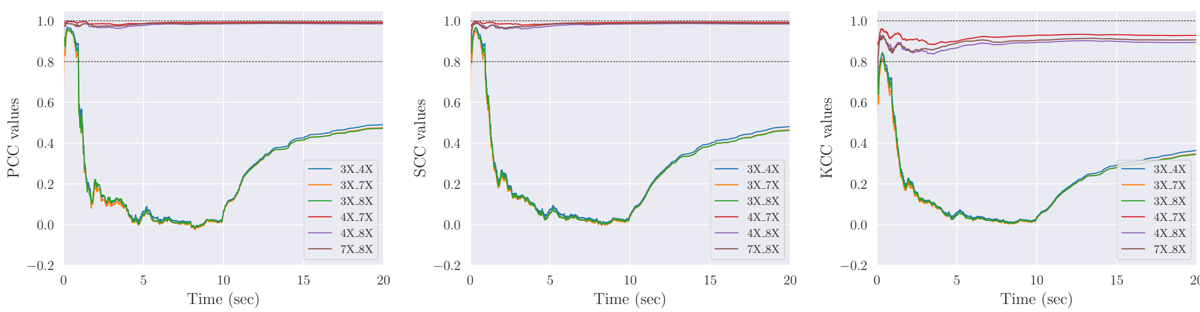
From the outcomes of cases 1 and 2, it is evident that data from sensor 3 (or sensors 3 and 7, depending on the example case study) is unreliable and requires replacement with dependable data. The LSTM model is employed to predict the acceleration-time series sensor data from sensor 3 (or sensors 3 and 7, depending on the example case study) using data from other sensors. Figure 8 and Figure 9 depict a close alignment between the temporal pattern of the inferred output and the actual data, signifying a successful match.

CONCLUSIONS

It is essential to integrate false sensor data detection and correction into advanced Online Monitoring (OLM) systems to prevent malfunctions or cyberattacks. This study introduces the Fault-Sensor Data Correction Module (FSDCM), demonstrating its effectiveness in detecting and rectifying false sensor data. The model identifies sensors responsible for generating erroneous signals and determines the timing of these signal initiations. This is achieved through time-dependent correlation coefficient analysis. FSDCM utilizes a machine learning paradigm, specifically the Long Short-Term Memory (LSTM) model, to generate substitute sensor data for false signals. To enhance model efficacy, strategies include the three-way holdout approach for dataset partitioning and the Reduce Learning Rates on Plateau (RLRP) for adaptive regulation during training. Empirical investigations confirm the proficiency of the trained LSTM model. Even in scenarios with multiple sensor failures, false signals can be replaced with reliable data from the LSTM model.



(a) Case 1



(b) Case 2

Figure 7. Time-dependent correlation analysis model performance

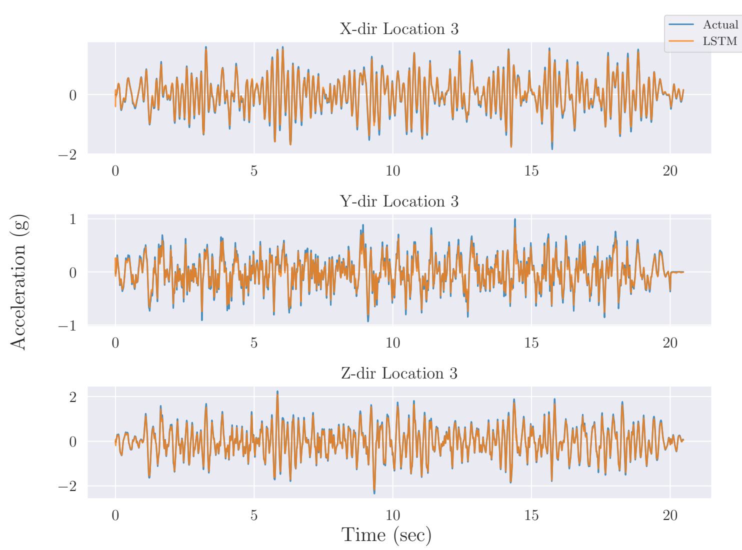


Figure 8. Sensor 3 prediction

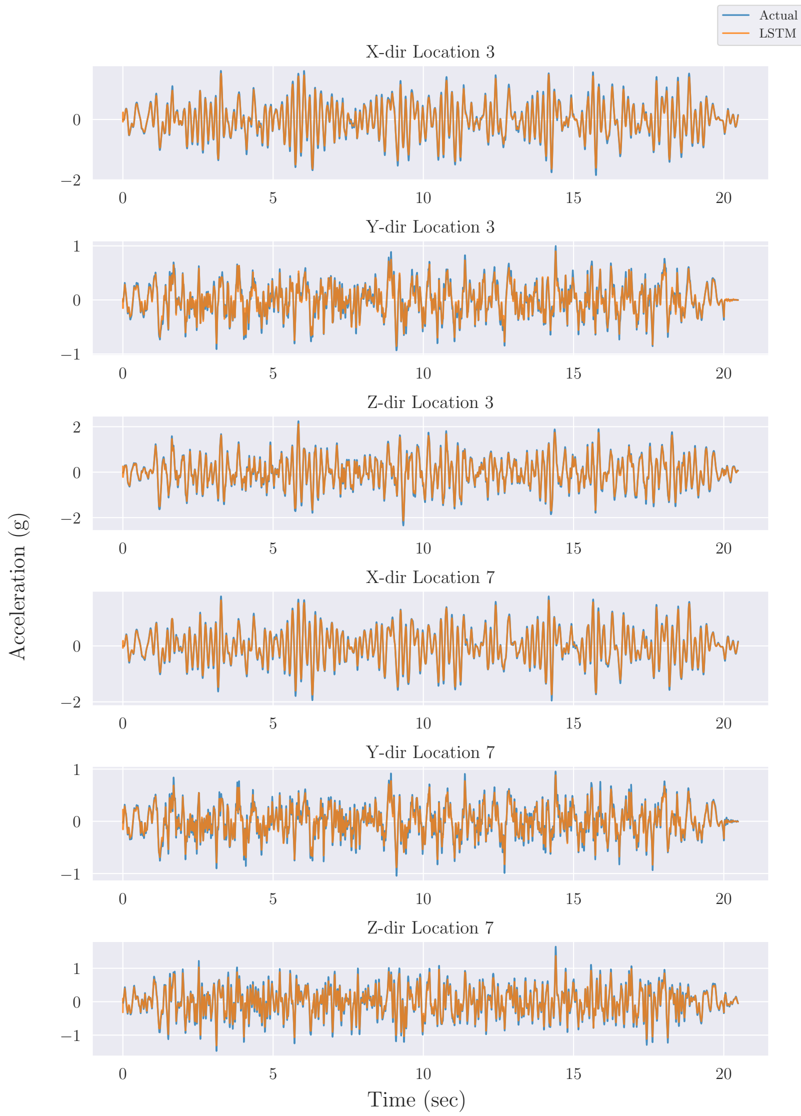


Figure 9. Sensor 3 and 7 inference

REFERENCES

Bodda, S. (2018). *Multi-Hazard Risk Assessment of a Flood Defense Structure*. Master's thesis, North Carolina State University.

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences.* Thomson Brooks/Cole Publishing Co.

Li, H., Wang, X., Yang, Z., Ali, S., Tong, N. and Baseer, S. (2022). "Correlation-Based Anomaly Detection Method for Multi-sensor System." *Computational Intelligence and Neuroscience*, 2022, pp. 1–13, ISSN 1687-5273, 1687-5265, doi:10.1155/2022/4756480.

Sandhu, H. K. (2015). *Flooding Fragility of Concrete Gravity Dam-Foundation System*. Master's thesis, North Carolina State University.

Sandhu, H. K. (2021). *Artificial Intelligence Based Condition Monitoring of Nuclear Piping-Equipment Systems*. Ph.D. thesis, North Carolina State University.

Sandhu, H. K., Bodda, S., Sauers, S. and Gupta, A. (2023a). "Condition Monitoring of Nuclear Equipment-piping Systems Subjected to Normal Operating Loads Using Deep Neural Networks." *Journal of Pressure Vessel Technology*, pp. 1–35.

Sandhu, H. K., Bodda, S. S. and Gupta, A. (2022a). "Deep Learning Framework for Post-Hazard Condition Monitoring of Nuclear Safety Systems." In *International Workshop on Structural Health Monitoring*, IWSHM.

Sandhu, H. K., Bodda, S. S. and Gupta, A. (2022b). "Post-hazard condition assessment of nuclear piping-equipment systems: Novel approach to feature extraction and deep learning." *International Journal of Pressure Vessels and Piping*, p. 104849.

Sandhu, H. K., Bodda, S. S. and Gupta, A. (2023b). "A Future with Machine Learning: Review of Condition Assessment of Structures and Mechanical Systems in Nuclear Facilities." *Energies*, 16(6), ISSN 1996-1073, doi:10.3390/en16062628.

Sandhu, H. K., Bodda, S. S., Sauers, S. and Gupta, A. (2022c). "Condition Monitoring of Nuclear Equipment-Piping Using Deep Learning." In *International Conference on Structural Mechanics in Reactor Technology*, IASMiRT.

Sandhu, H. K., Patel, P., Bodda, S. S. and Gupta, A. (2019). "External Multi-Hazard Probabilistic Risk Assessment Methodology and Applications: A review of the State-of-the-Art." In *International Conference on Structural Mechanics in Reactor Technology*, IASMiRT.

USNRC (2022). "Digital Instrumentation and Controls Research." https://www.nrc.gov/about-nrc/regulatory/research/digital.html, accessed: Aug 18, 2023.

Yi, T.-H., Huang, H.-B. and Li, H.-N. (2017). "Development of sensor validation methodologies for structural health monitoring: A comprehensive review." *Measurement*, 109, pp. 200–214, ISSN 0263-2241, doi:https://doi.org/10.1016/j.measurement.2017.05.064.

Yu, F., Ma, X., Miao, Z. and Zhang, M. (2022). "On Line Monitoring of Sensor Failure in Nuclear Power Plant Based on Sequential Probability Ratio Test." In *International Conference on Nuclear Engineering*, volume 86359, p. V001T01A005, American Society of Mechanical Engineers.