Transactions of the Korean Nuclear Society Autumn Meeting
Changwon, Korea, October 30-31, 2025

Detection of Fault Signals in Nuclear Power Plant Measurements via LSTM Classifier

Yu Seob Kim, Seung Gyu Cho, Seung Jun Lee*
Ulsan National Institute of Science and Technology, 50, UNIST-gil, Ulsan, 44911
**Corresponding author: sjlee420@unist.ac.kr*

***Keywords :** Nuclear power plant, nuclear fatigue monitoring system, measurement signal, fault signal detection, long short-term memory

1. Introduction

In nuclear power plants, measurement instruments continuously generate real-time signals that indicate the operational state of individual components and systems. These signals play a pivotal role in plant operation and control, providing operators with essential information to ensure safe operation and structural integrity. Consequently, the accuracy and reliability of measurement signals are of critical importance in fatigue monitoring and structural integrity assessment.

To evaluate the fatigue life and structural integrity of major components and piping, the Nuclear Fatigue Monitoring System (NuFMS) utilizes measurement signals obtained from the Plant Monitoring System (PMS) to compute cumulative fatigue usage factors [1]. In actual operation, however, these signals are often contaminated by spurious signals-including noise, drift, and impulsive outliers-that compromise the accuracy of fatigue calculations, potentially leading to overly conservative predictions or misinterpretations of structural reliability.

Traditional fault signal detection has relied on statistical methods such as the z-score, which identifies deviations from the mean and standard deviation, and the Mahalanobis distance, which considers correlations among multiple variables. While effective under certain conditions, these methods are susceptible to high false alarm rates and degraded accuracy, particularly when measurement signals exhibit significant variability, nonlinear characteristics, or pronounced temporal correlations.

To overcome these challenges, this study introduces a Long Short-Term Memory (LSTM) network for detecting spurious signals in PMS data. LSTM networks effectively capture temporal dependencies across multiple time scales, allowing more robust recognition of underlying patterns compared with traditional statistical techniques. The proposed approach demonstrates improved reliability in spurious signal detection, thereby contributing to more accurate fatigue evaluations within NuFMS.

2. Proposed Condition Monitoring System

2.1 Overall Framework

The overall framework of the proposed condition monitoring system is illustrated in Fig. 1. In this study, we propose a condition monitoring system capable of fault signal detection. The system consists of four main processes: (i) data acquisition and preprocessing, (ii) interval segmentation and fault signal labeling, (iii) model training, and (iv) evaluation and validation. In the preprocessing stage, multi-sensor data are classified into non-fault and fault signals, followed by standardization procedures. The entire system employs an LSTM-based classifier to perform fault signal detection.

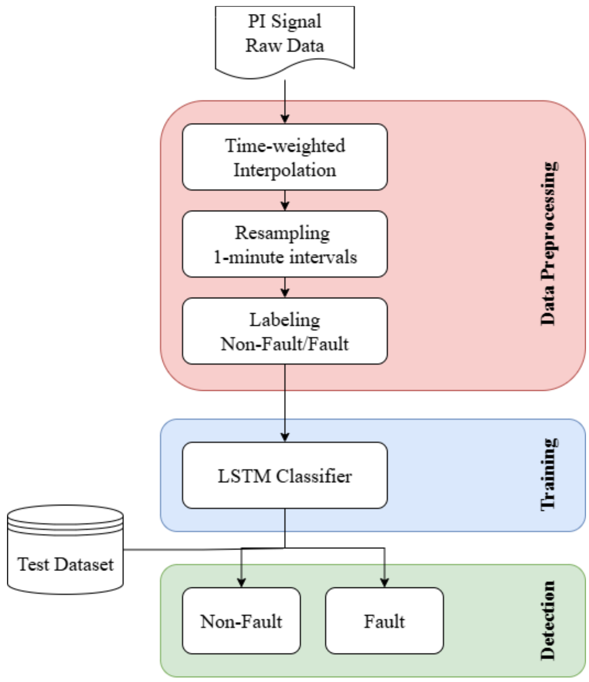

Fig. 1. Overview of the proposed condition monitoring system

3. Methodology

3.1 Data Acquisition and preprocessing

3.1.1 Data Acquisition

In this study, 22 key variables from OPR-1000 pressurized water reactor (PWR) were selected as the primary dataset. The raw data were obtained from the Plant Information (PI) system over a period of approximately 14 years, spanning from January 1, 2006, to December 31, 2019. Among these, the records from 2006 to 2016 were utilized for model training, while the records from 2017 to 2019 were employed for testing and validation. Each variable reflects the operational states of major systems and components in the nuclear

Table 1. Variable Description

Variable	Type	Unit	Measurement point	Description
NP1 ~ NP4	AI	%	Power channels	Nuclear Power
HL1 ~ HL3	AI	℃	Each loop Hot Leg	Hot Leg Temperature
CL1 ~ CL3	AI	℃	Each loop Cold Leg	Cold Leg Temperature
Srg_Temp	AI	℃	Pressurizer	Pressurizer Surge Line Temperature
PZR_W_Temp	AI	℃	Pressurizer	Pressurizer Water Temperature
PZR_S_Temp	AI	℃	Pressurizer	Pressurizer Steam Temperature
PZR_Pres1 ~ 5	AI	MPa	Pressurizer	Pressurizer Pressure
Spr1 ~ 2_Temp	AI	℃	Pressurizer spray	Pressurizer spray Temperature
Chrg, Let_Temp	AI	℃	Chemical Volume Control System	Charging, Letdown Temperature

power plant, and the corresponding signal types and measurement units are summarized in Table 1.

3.1.2 Data Characteristics

The measurements consist of both regularly recorded values at midnight each day and additional event-based entries triggered by changes in the monitored signals. As a result, the intervals between records are irregular, and it is uncommon for multiple variables to be recorded simultaneously within the same time step. These characteristics pose limitations in directly capturing the temporal correlations among variables.

3.1.3 Data Interpolation

To address the aforementioned issues, missing values in each variable were first imputed using linear interpolation. Subsequently, the entire dataset was resampled at one-minute intervals to normalize all variables onto a common time axis. Through this process, the irregular time series were transformed into uniform sequences, thereby enabling inter-variable comparisons and constructing input data suitable for LSTM-based learning.

3.1.4 Signal Fault Modes

Spurious signals encountered in nuclear power plant instrumentation are essentially similar to those observed in general industrial systems. Signal faults can arise from various causes, including: (i) hardware defects of the measuring instruments, (ii) adverse environmental conditions such as high temperature or radiation, (iii) electronic or mechanical failures of transducers during signal transmission, and (iv) malfunctions in communication networks. These faults may manifest in several characteristic forms, such as spikes, drifts, saturation, and stuck signals, as illustrated in Fig. 2 [2].

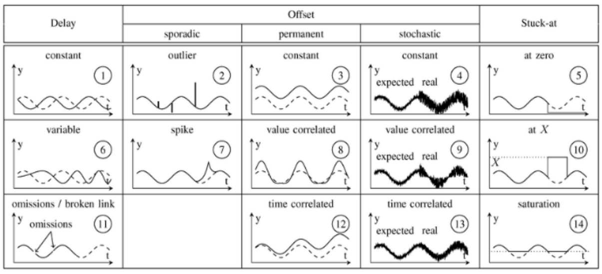

Fig. 2 Examples of Signal Fault Patterns by Measurement Fault Categories

Similar fault patterns were observed in the dataset used in this study. Specifically, impulse-type outliers, interval-based stuck signals, drifts, and loss of redundancy consistency among correlated sensors were identified, which are consistent with typical signal faults in industrial plants.

3.1.5 Dataset Partitioning

The dataset was divided by time period, with data from 2006–2016 used for training, and data from 2017–2019 employed for validation as an independent test set. Each year contained major operational events, such as unplanned trips and scheduled overhauls (OH). Based on these events, the spurious signal detection model was trained and evaluated in this study.

3.1.6 Data Normalization

To compensate for differences in the distribution of variables and to prevent certain variables from disproportionately influencing the model training, z-score normalization was applied. This normalization ensured that all input variables had zero mean and unit variance, thereby improving training stability and convergence speed. The normalization process is defined as:

$$x' = \frac{x - \mu}{\sigma} \tag{1}$$

where x denotes the original value, μ the mean, and σ the standard deviation of each variable.

3.2 Segmentation and Fault signal Labeling

3.2.1 Data Segmentation

In this study, the time series data were divided into smaller segments according to predefined criteria, and each segment was assessed for the presence of fault signals. To this end, two types of intervals were defined:

- Transient Interval: the period during which a signal exceeds a predefined threshold and remains unstable until it returns within the threshold range.

- Stable Interval: the period during which a signal fluctuates within the threshold range without exhibiting abrupt changes.

The thresholds were determined based on the distribution of short-term variations (Δx) observed under normal operating conditions for each variable. Specifically, the short-term fluctuation range in normal intervals was statistically quantified, and variable-specific thresholds were assigned to reflect individual characteristics of each signal.

In addition, the duration of each segment was limited to a maximum of 120 minutes. This constraint was imposed to ensure that when local fault signals (e.g., slowly drifting signals or anomalous values occurring within an otherwise normal trend) were present, the segmentation allowed such fault signals to be isolated without misclassifying the surrounding normal data as spurious signals. Consequently, even in cases where normal and fault data coexisted within the same interval, this approach enabled the effective extraction of local fault signals while minimizing false detections.

3.2.2 Fault signal Labeling

Based on the segmented intervals, a first-stage labeling process was performed using statistical techniques. The standard deviation and z-score methods were applied to each interval to detect values deviating from the normal distribution range.

However, such statistical approaches exhibited limitations, as they occasionally misclassified normal fluctuations under steady-state operating conditions as fault signals. In other words, physically valid variations were sometimes falsely labeled as fault signals simply because they exceeded mathematical thresholds.

To address this issue, a domain knowledge-based review procedure involving researchers and plant operators was incorporated following the initial labeling. Each interval was manually examined and validated to correct the false detections produced by the statistical methods. Through this process, the final ground truth dataset was constructed.

3.3 Model Training

In this study, an LSTM (Long Short-Term Memory)–based classification model was designed for spurious signal detection using time series data. The proposed model adopts a many-to-one architecture, in which each segmented sequence is provided as input, and the model is trained to determine whether the given interval contains fault data. The overall architecture of the proposed LSTM model is illustrated in Fig. 3 [3].

LSTM networks have a structural advantage in learning both long- and short-term dependencies in time series. This enables the model not only to capture instantaneous fault signals but also to reflect dynamic patterns accumulated over a period of time. Accordingly, the data of each interval are processed sequentially, and the final hidden state is used to output the classification result at the segment level.

The input data consist of 22 key variables that were preprocessed and normalized onto a common one-minute time axis, with each interval having a maximum length of 120 minutes. During training, each interval was transformed into a fixed-length sequence and fed into the model. The labels were assigned according to the presence or absence of fault signals in the corresponding interval (0: normal, 1: fault).

In particular, to address the severe class imbalance in the dataset, a higher weight was assigned to the fault signal class during the loss function calculation, as expressed in the following weighted cross-entropy formulation:

$$w = \frac{N_{neg}}{\max(N_{pos}, 1)} \tag{2}$$

$$L = \frac{1}{N}\sum_{i=1}^{N}(-(1-y_i)\log(1-p_i) - y_i w log(p_i)) \tag{3}$$

where N_{neg} is the total number of normal (negative) samples, N_{pos} is the total number of fault (positive) samples, respectively. This weighting ensures that the contribution of the minority (fault) class is amplified relative to the majority (normal) class, thereby encouraging the model to be more sensitive to fault signal detection.

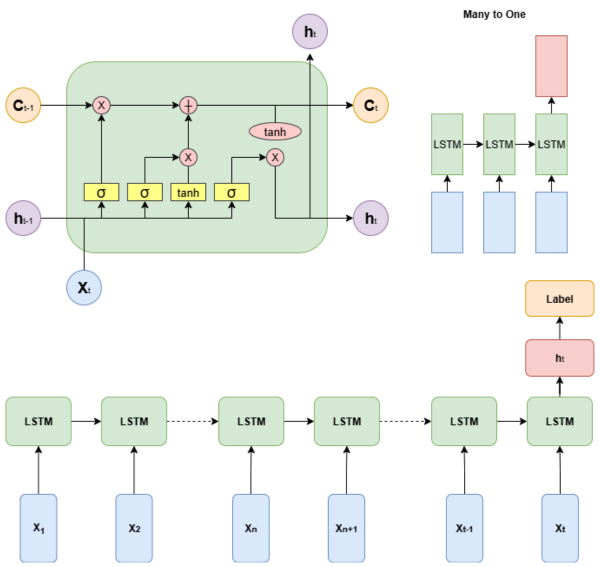

Fig.3 LSTM many-to-one structure

4. Results

In this study, the fault signal detection model was first applied to the NP among the key variables. The NP signal was chosen as the initial case study because it directly represents the overall operating condition of the plant, responds sensitively to system fault signals and sensor malfunctions, and is one of the core input parameters used in NuFMS fatigue evaluations.

4.1 Model performance Evaluation

The performance metrics of the LSTM-based fault signal detection model for the NP signal are summarized in Table 2. The overall accuracy reached 98.4%, indicating high predictive performance. For the normal class, the model achieved excellent results with a precision of 1.00, recall of 0.98, and F1-score of 0.99, demonstrating stable and consistent classification. In contrast, the performance for the fault class was more limited. The precision was only 0.28, meaning that a substantial number of normal signals were incorrectly classified as fault signals, while the recall was 0.905, showing that most actual fault signals were successfully detected. Consequently, the F1-score for the fault class was 0.43.

The confusion matrix (Fig. 4) provides further insight into the model's predictions. As shown, a portion of normal data was misclassified as fault signals (221 cases), while the majority of fault signals were correctly identified (86 cases). However, a small number of fault signals were incorrectly classified as normal (9 cases). This indicates that the model maintains high sensitivity in fault signal detection but still suffers from false alarms.

4.2 Discussion

The low precision for the fault class can be explained primarily by the severe class imbalance in the dataset—only 95 fault signals (about 0.7%) were included out of 14,537 test samples—as well as the influence of concurrent fault signals in other variables. In practice, when fault signals occurred simultaneously in multiple variables, the entire interval was labeled as fault. As a result, the NP signal was also judged as anomalous, even in cases where its own deviation was less distinct.

These results are summarized numerically in Table 2. The evaluation of the NP signal demonstrates the potential applicability of the proposed LSTM-based approach to NuFMS signal validation. However, further studies should extend this methodology to other key plant variables to assess its robustness and generalizability.

Table 2. Performance metrics of the LSTM classifier for NP

Class	Precision	Recall	F1-score	Segment
Non-Fault	1.000	0.980	0.990	14,442
Fault	0.280	0.905	0.428	95

※ Accuracy: 0.984

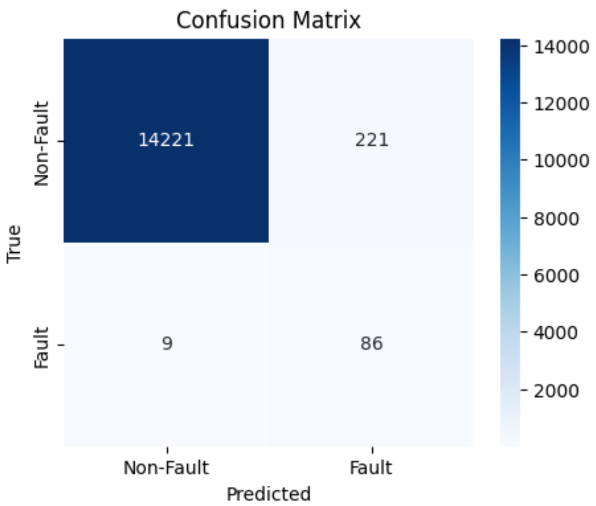

Fig.4 Confusion matrix of fault signal detection results for NP

5. Conclusion

This study applied an LSTM-based fault signal detection method to address the problem of spurious signals that undermine the reliability of measurement data in the NuFMS. To this end, major measurement variables from OPR-1000 pressurized water reactor (PWR) were collected, preprocessed, segmented, and labeled to construct a training dataset, and the nuclear power output signal (NP) was used for model training and evaluation.

The experimental results showed that the proposed LSTM model achieved a high recall of 0.905, successfully detecting the majority of fault signals. However, the precision was limited to 0.280, indicating that many normal signals were misclassified as fault signals. This low precision is interpreted as being caused primarily by the severe class imbalance in the test dataset, in which fault signals accounted for only about 0.7% of all samples, as well as the concurrent occurrence of spurious signals in other variables, which caused NP intervals to be classified as fault even when the NP signal itself did not show a distinct deviation.

Nevertheless, the study demonstrated the potential to automate spurious signal detection, which has traditionally relied on manual review. Since the NP signal is a core parameter directly used in NuFMS fatigue evaluations, the performance of the proposed model provides an important case study for assessing the feasibility of real-world application.

Future research directions include: (i) applying fault signal data augmentation techniques to alleviate class imbalance, (ii) developing a hybrid fault signal detection framework that integrates statistical detection methods, PCA-based dimensionality reduction, and probability density function-based thresholding [4], (iii) extending the validation to other key plant measurement variables beyond NP. (iv) exploring the use of autoencoder-based architectures, such as LSTM-AE, for not only detecting spurious signals but also reconstructing and restoring corrupted measurements.

Accordingly, the proposed approach is expected to substantially improve the reliability of NuFMS fatigue assessments by overcoming the inherent limitations of manual inspection and providing a systematic, automated validation process for PMS signals. In particular, the establishment of an automated signal validation framework ensures both consistency and efficiency, which in the long term can contribute not only to enhancing the accuracy of fatigue life evaluations but also to strengthening overall plant safety and facilitating regulatory compliance in nuclear power plant operations.

ACKNOWLEDGEMENTS

This work was supported by KOREA HYDRO & NUCLEAR POWER CO., LTD (No. 2024-Tech-09)

REFERENCES

[1] Korea Atomic Energy Research Institute.. "Development of fatigue monitoring system in nuclear power plants" (KAERI/RR-1307/93). (1994)

[2] Sebastian Zug. " Fault-Handling in Networked Sensor Systems ".

[3] Malhotra, P. et al., “Long-short-term memory networks for anomaly detection in time series,” 2015.

[4] Sojin Shin. "Multi-sensor data-based anomaly detection and diagnosis of a pumped storage hydropower plant" (2023)