Prediction Intervals of Cooling System Temperatures Monitoring in Hydropower Generators Using Pretrained LSTM and Sobol Sensitivity Analysis

Chinachote Deevijit, Tanongkiat Kiatsiriroat, Thoranis Deethayat, and Attakorn Asanakham*

Energy Engineering Program, Faculty of Engineering and Graduate School, Chiang Mai University, 239 Huay Kaew Road, Muang District, Chiang Mai, 50200, Thailand
Email: chinachote.dee@egat.co.th (C.D.); tanongkiat_k@yahoo.com (T.K.); thoranisdee@gmail.com (T.D.); attakorn.asana@cmu.ac.th (A.A.)
*Corresponding author

Manuscript received December 10, 2025; accepted February 13, 2026; published April 30, 2026

***Abstract*—This study presents a data-driven framework that integrates a pretrained Long Short-Term Memory (LSTM) model with Sobol Global Sensitivity Analysis to forecast thermal behavior in hydropower generator cooling systems. The model was developed using real operational data from a hydropower plant in Thailand and demonstrated the capability to predict four key temperature variables: inlet air temperature, outlet air temperature, stator temperature, and outlet water temperature, under both normal and abnormal operating conditions. Monthly Prediction Intervals (PIs) were constructed using uncertainty sampling guided by Sobol analysis, capturing operational variability such as reductions in cooling water flow and elevated inlet water temperatures. Simulation results indicated that the combined approach effectively replicated thermal anomalies observed in historical data, particularly during the dry season. Each thermal variable exhibited distinct sensitivity profiles, especially stator temperature and both inlet and outlet air temperatures which responded critically on the heat exchange performance. These findings highlight the potential for implementing seasonally adaptive maintenance strategies and demonstrate the proposed framework's practical applicability for early anomaly detection and predictive maintenance planning in hydropower operations.**

***Keywords*—thermal forecasting, Long Short-Term Memory (LSTM) model, sobol global sensitivity analysis, hydropower generators**

I. Introduction

Maintaining thermal stability within the cooling systems of hydropower generators is critical for ensuring operational reliability and the long-term durability of generation equipment. Traditional monitoring and forecasting methods, such as thermocouple-based measurements and statistical regression techniques, often fail to capture nonlinear and time-dependent dynamics inherent in these systems. Consequently, their predictive effectiveness in complex, real operational environments remain limited.

In recent years, deep learning has emerged as a powerful tool for modeling dynamic processes in engineering, with Long Short-Term Memory (LSTM) networks gaining increasing attention for time-series applications. These models have demonstrated strong predictive performance across various domains, including short-term energy forecasting [1] and thermal monitoring in industrial equipment [2]. However, despite their accuracy, LSTM models are often criticized for their 'black-box' nature, which limits interpretability and makes it challenging to understand the influence of input features on the predicted outcomes. As a result, many recent studies have employed techniques such as Sobol Sensitivity Analysis [3–7] to assess the importance of individual input variables. Emerging research increasingly highlights the value of integrating deep learning with global sensitivity analysis to enhance model interpretability and robustness. In particular, Sobol Sensitivity Analysis has been widely applied in conjunction with neural networks to quantify the relative influence of input variables on predicted outputs. This combination has shown promising results across various engineering domains, including model stability assessment under parameter perturbations [3], feature selection for solar energy forecasting [4], and maintenance scheduling based on degradation simulations [5]. Structural engineering studies have also leveraged these techniques to identify key variables affecting bridge deterioration [6] and to estimate the wear life of casting molds under thermal stress [7]. These findings collectively highlight the transformative role of sensitivity analysis in engineering practice. By illuminating the contributions of individual input variables, such methods enhance system understanding and support more informed design, control, and maintenance strategies.

In this study, the researchers employed a pretrained Long Short-Term Memory (LSTM) model, developed using operational data from a hydropower plant in Thailand. The model utilized input variables such as electrical power output (P), cooling water flow rate (Q_w), inlet water temperature ($T_{w,in}$), reservoir level (H), and operation duration (M) to forecast key thermal outputs, including inlet air temperature ($T_{air,in}$), outlet air temperature ($T_{air,out}$), stator temperature (T_{stator}), and outlet water temperature ($T_{w,out}$). The model was further applied in the analysis of prediction uncertainty through Sobol Sensitivity Analysis, which was used to evaluate the variability of outputs under simulated scenarios, such as reduced coolant flow or elevated inlet water temperature. The approach aims to support predictive maintenance decision-making in hydropower systems by providing interpretable and reliable thermal forecasts.

II. Methods

This study comprises two main stages:

1) Employing a pretrained Long Short-Term Memory (LSTM) model to forecast the thermal behavior of a hydropower generator.
2) Applying Sobol Global Sensitivity Analysis to evaluate

the effects of input uncertainties and construct monthly Prediction Intervals (PIs) for use in predictive maintenance planning.

A. Pretrained LSTM Model

The LSTM model used in this study was developed to estimate key thermal indicators within the generator cooling system. It was trained using real operational data from the Sirikit Dam Hydropower Plant, covering three years: 2018 (training), 2020 (validation), and 2019 data, used exclusively for testing, contained known thermal anomalies. This multi-year approach ensures that the model captures seasonal variability relevant to hydropower operations.

The model accepts five mechanical input variables: electrical power output (P), time index (M), reservoir level (H), cooling water flow rate (Q_w), and inlet water temperature ($T_{w,in}$). It predicts four output variables that reflect thermal performance: inlet air temperature ($T_{air,in}$), outlet air temperature ($T_{air,out}$), stator temperature (T_{stator}), and outlet water temperature ($T_{w,out}$).

To maintain the integrity of the uncertainty analysis, the trained model is treated as a fixed "black box" during the simulation stage. The model demonstrates high predictive accuracy, with prediction errors for all four temperature outputs remaining below 1% compared to their respective ground truth values (see Table 1).

Table 1. Mean absolute prediction errors for thermal outputs (2019)

Variable	% Error
$T_{air,in}$	0.46%
$T_{air,out}$	0.09%
T_{stator}	0.93%
$T_{w,out}$	0.16%

The architecture of the LSTM model is shown in Fig. 1, in which five input variables are first processed through a single LSTM layer comprising 128 memory cells to capture temporal patterns in the input sequence. The final hidden state of the LSTM is then passed to a fully connected regression layer that simultaneously outputs four predicted temperature variables.

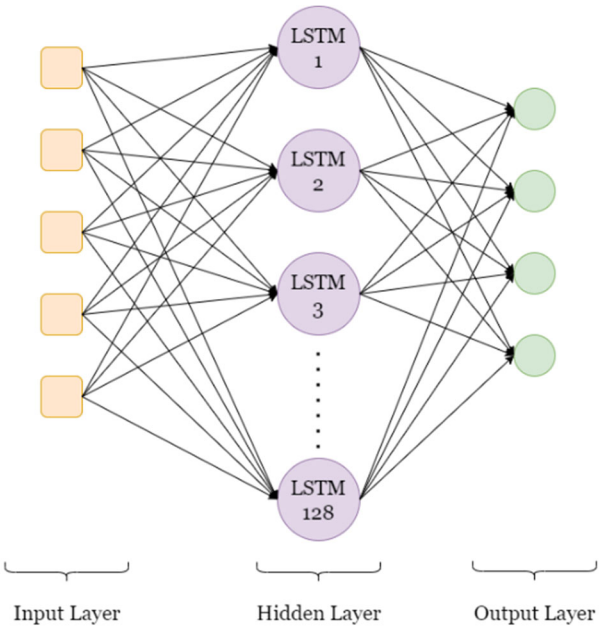

Fig. 1. Architecture of the pre-trained LSTM model used for thermal forecasting.

B. Sobol Sensitivity Analysis and Simulation

To evaluate the influence of input uncertainties on thermal behavior, the study employs Sobol Global Sensitivity Analysis, implemented using the SALib Python library [8]. This method allows for efficient quasi-random sampling in high-dimensional spaces and provides both first-order and total-order sensitivity indices.

The 2019 operational data was selected due to a real thermal deviation that occurred during live electricity production and was recorded by the plant's control system. This anomaly, being unplanned and naturally occurring, provides a robust test case to validate the framework under real-world uncertainty.

Input ranges for the four mechanical input parameters—electrical power output (P), cooling water flow rate (Q_w), inlet water temperature ($T_{w,in}$), and reservoir level (H)—were derived from the 2019 operational dataset. The simulation month (M) was not included in the sampling process but was treated as a fixed contextual input to account for seasonal variation.

For each calendar month, statistical bounds were defined as the mean ± one standard deviation (±1σ), assuming a normal distribution [9]. This assumption follows conventional engineering practice and was supported by exploratory analysis, which showed that most input variables exhibited near-Gaussian behavior under typical operating conditions.

The calculated monthly bounds were expressed in real engineering units (e.g., MW, L/min, °C, m.mSL) and are summarized in Table 2. These values form the basis for defining the uncertainty input space used in the Sobol simulations. In this context, each mean ± 1σ value serves as an upper and lower confidence bound for the respective input variable, defining the plausible operational range for uncertainty propagation.

Table 2. Monthly input parameter bounds used for sobol sampling (Mean ± 1σ)

Statistic	Month	P (MW)	Q_w (L/min)	$T_{w,in}$ (°C)	H (m.mSL)
Mean	1	89.10	5006.48	31.20	68.56
±1σ		±6.59	±2286.00	±0.79	±5.87
Mean	2	86.90	5091.15	31.20	65.08
±1σ		±6.59	±2286.00	±0.83	±5.43
Mean	3	82.07	5091.15	31.28	62.48
±1σ		±5.27	±2286.00	±0.95	±5.43
Mean	4	81.63	5260.48	31.48	59.22
±1σ		±6.15	±2201.33	±0.91	±4.78
Mean	5	79.87	2635.82	31.95	57.48
±1σ		±6.15	±1862.66	±1.22	±5.87
Mean	6	76.80	2889.82	32.62	55.09
±1σ		±5.71	±1862.66	±1.11	±7.17
Mean	7	75.92	3313.15	32.78	53.35
±1σ		±7.91	±2032.00	±1.22	±8.69
Mean	8	79.43	3651.81	32.27	59.43
±1σ		±9.23	±2116.66	±1.30	±6.09
Mean	9	81.19	3736.48	31.83	62.26
±1σ		±7.47	±1947.33	±0.95	±5.87
Mean	10	82.51	3905.81	31.99	62.26
±1σ		±6.59	±1693.33	±0.79	±4.56
Mean	11	81.19	3905.81	32.23	61.39
±1σ		±5.27	±1693.33	±1.03	±5.22
Mean	12	79.87	3905.81	31.52	61.17
±1σ		±5.27	±1862.66	±1.03	±5.43

Note: Each value represents the upper and lower bounds of the monthly confidence range.

Based on these bounds, 400 input combinations per month were generated using Sobol sequences and used to simulate the following operational scenarios:

- Normal Case: All inputs sampled within their monthly ±1σ bounds.
- Low-Flow Case: Q_w reduced by 30% from the lower statistical bound.
- High-Temperature Case: $T_{w,in}$ increased by 50% above the upper bound.
- Combined Case: Q_w reduced by 30% and $T_{w,in}$ increased by 50%, simulating a worst-case stress scenario.

Simulations were conducted specifically during May through July, a period typically associated with lower reservoir levels and higher generation demand, conditions that are prone to inducing thermal stress in the system. For each scenario, the LSTM model generates predicted outputs, which are statistically aggregated to compute the monthly mean and standard deviation. These are then used to construct Prediction Intervals (PIs) using the formula:

$$\text{PIs} = \text{mean} \pm \text{standard deviation}$$

This formulation assumes approximate normality in the outputs. These PIs offer a probabilistic envelope of expected system behavior, allowing operators to detect deviations indicative of potential thermal anomalies.

The overall process of the research framework is illustrated in Fig. 2, and the resulting prediction intervals are subsequently compared against actual thermal measurements from the 2019 dataset to assess the ability of the model to detect anomalies under real operating conditions.

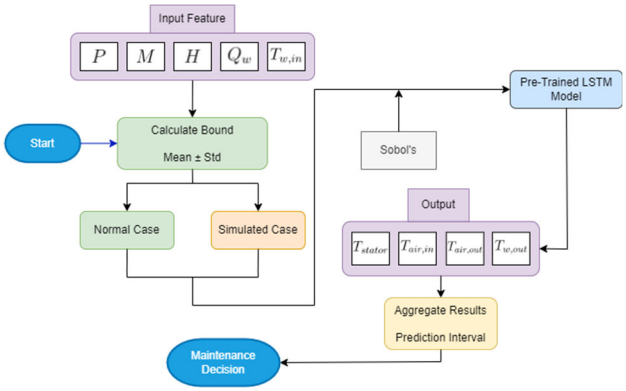

Fig. 2. Workflow of the proposed predictive maintenance framework, integrating LSTM-based temperature forecasting and Sobol-based uncertainty propagation.

III. Result

This section presents the monthly prediction results for four key thermal indicators $T_{air,in}$, $T_{air,out}$, T_{stator}, and $T_{w,out}$ obtained using the LSTM-based forecasting model and Sobol-driven uncertainty simulations. For each variable, the results include the baseline prediction of the model, ±1σ prediction intervals, and the outcomes under three operational scenarios (Low Flow, High Temperature, Combined), compared against actual 2019 measurements. The Prediction Intervals (PIs), constructed as mean ±1σ, provide a probabilistic confidence band that enables visual comparison between actual data and expected thermal behavior under uncertainty.

1) Inlet air temperature ($T_{air,in}$)

As shown in Fig. 3, the $T_{air,in}$ under normal operating conditions, remained within a range of 59–62 °C. In the Low Flow scenario, $T_{air,in}$ a minor increase in temperature, stayed within the ±1σ interval, indicating relatively low sensitivity to flow reduction alone. The High Temperature scenario led to a more pronounced rise in $T_{air,in}$ than upper bound. reflecting the direct thermal influence of elevated $T_{w,in}$. The Combined scenario resulted $T_{air,in}$ exceeding the upper prediction bound, which closely aligned with the peak observed in 2019.

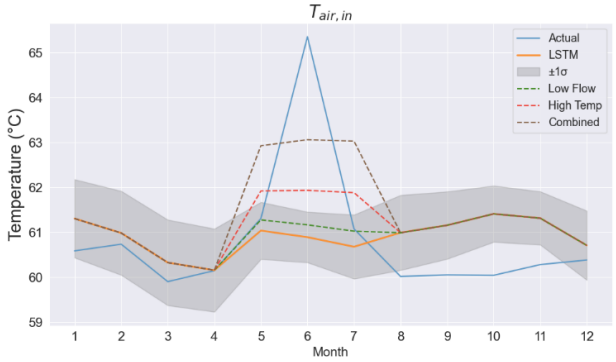

Fig. 3. The inlet air temperature ($T_{air,in}$) under both normal and simulated conditions.

2) Outlet air temperature ($T_{air,out}$)

As shown in Fig. 4, $T_{air,out}$ under normal conditions, ranged between 43–50 °C throughout the year. In the Low Flow scenario, the temperature increased slightly, approaching the upper bound, but remained within acceptable limits. The High Temperature scenario caused a more substantial rise in $T_{air,out}$, reflecting the direct thermal impact of elevated $T_{w,in}$. Under the Combined condition, $T_{air,out}$ rose significantly above the normal range, closely matching the anomaly observed in the 2019 data.

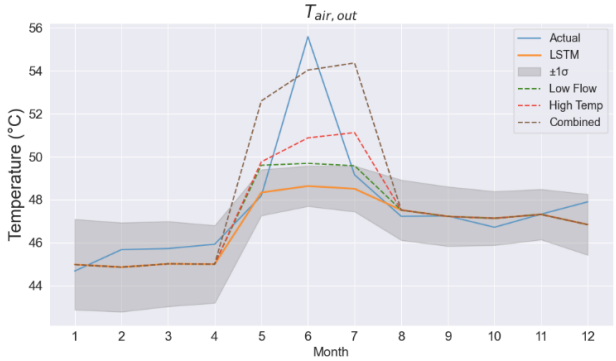

Fig. 4. The outlet air temperature ($T_{air,out}$) under both normal and simulated conditions.

3) Stator temperature (T_{stator})

As depicted in Fig. 5, T_{stator} ranged between 78–81 °C under baseline conditions. In the Low Flow scenario, a slight increase in temperature was observed, but T_{stator} remained well within the prediction intervals. In the High Temperature scenario, T_{stator} rose slightly above the upper bound of the prediction intervals but remained within a tolerable range. In the Combined scenario, T_{stator} increased significantly and

exceeded the upper bound, closely matching the actual anomaly recorded in June. This indicated that stator components were particularly vulnerable to overheating caused by both reduced coolant flow and elevated inlet water temperature.

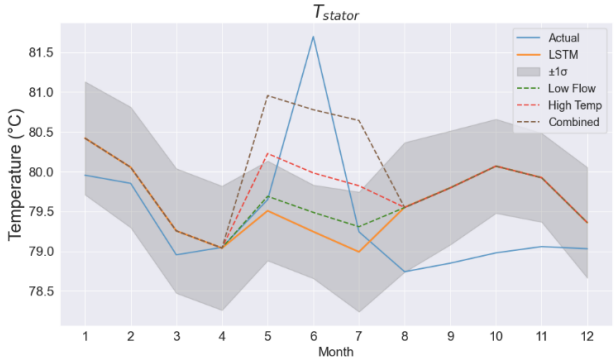

Fig. 5. The stator temperature (T_{stator}) under both normal and simulated conditions.

4) Outlet Water Temperature ($T_{w,out}$)

As shown in Fig. 6, $T_{w,out}$ under nominal conditions, ranged between 35–38 °C. In the Low Flow scenario, $T_{w,out}$ increased toward the upper bound but remained within the prediction intervals. The High Temperature case caused $T_{w,out}$ to rise further, approaching the actual values observed in 2019. In the Combined scenario, $T_{w,out}$ exhibited a noticeable increase, likely due to cumulative heat buildup within the cooling system.

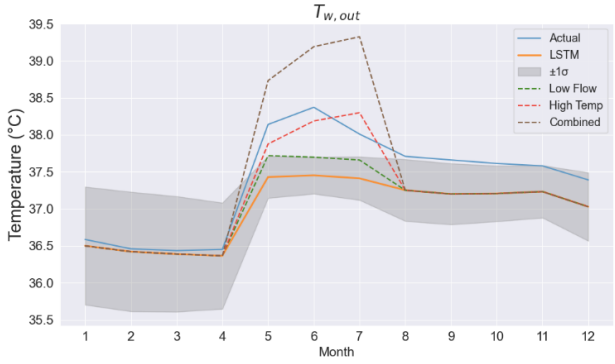

Fig. 6. The outlet water temperature ($T_{w,out}$) under both normal and simulated conditions.

Across all four thermal indicators, the Combined case consistently led to elevated values compared to baseline conditions. Moreover, the simulated responses closely matched the actual in 2019, confirming the model's capability to capture compounded thermal effects resulting from simultaneous reductions in cooling water flow and elevated inlet water temperature.

Among the outputs, $T_{air,in}$ and $T_{air,out}$ exhibited rapid and pronounced responses, indicating high sensitivity to both flow and thermal disturbances. In contrast, T_{stator} and $T_{w,out}$ demonstrated slower, cumulative temperature increases, reflecting the buildup of heat within the electrical and cooling systems.

These results validate the effectiveness of the proposed LSTM-Sobol framework not only in forecasting typical thermal behavior but also in detecting early signs of system stress through uncertainty-based scenario simulations. This capability provides valuable insights for condition-based monitoring and supports the development of predictive maintenance strategies to mitigate thermal risks in hydropower generator operations.

IV. Discussion

The results from this study highlighted the effectiveness of combining LSTM-based thermal forecasting with Sobol-driven uncertainty analysis to support predictive maintenance in hydropower generators. Among the simulated scenarios, the Combined case consistently reproduced abnormal temperature spikes observed in the actual 2019 data, particularly from May to July. This confirms that the proposed framework was capable of capturing thermal effects stemming from simultaneous disturbances in coolant flow rate and inlet water temperature.

Each thermal variable demonstrated distinct response characteristics. The inlet and outlet air temperatures ($T_{air,in}$ and $T_{air,out}$) exhibited high sensitivity to both flow and thermal perturbations, responding rapidly to changes in system inputs. In contrast, the stator temperature (T_{stator}) and the outlet water temperature ($T_{w,out}$) showed gradual increases, reflecting the accumulation of heat over time and underscoring their importance as indicators of long-term thermal risk. Sensitivity-based simulations such as these have proven valuable in dynamic systems for identifying critical variables and managing state transitions [10].

The integration of Sobol sensitivity sampling enabled the construction of Prediction Intervals (PIs) that account for variability in key operational parameters. Unlike traditional point forecasts, these intervals provided a probabilistic envelope that helps identify early indicators of thermal anomalies [11], particularly when actual measurements begin to exceed expected thresholds. This framework allowed maintenance teams to assess thermal behavior not just as static predictions but as dynamic bands of operational confidence.

Notably, the prediction intervals began to diverge from the baseline forecast before actual temperature anomalies appeared, offering early-warning potential. This trend was consistently observed across all four indicators—$T_{air,in}$, $T_{air,out}$, T_{stator}, and $T_{w,out}$—under the Combined simulation scenario. As illustrated in Figs. 3–6, the Combined case exceeded the ±1σ bounds one month prior to the peak measurements recorded in 2019, particularly between May and June. This early deviation highlights the model's ability to detect subtle shifts in system behavior before they fully manifest. Such capability is especially valuable in thermally sensitive components, where gradual heat buildup may precede visible failures. By identifying emerging anomalies earlier, the framework supports more responsive and preventative maintenance planning.

However, it is important to acknowledge the limitations of this approach. The simulations were conducted by varying only a defined subset of input features (P, M, H, Q_w, and $T_{w,in}$), while other unmeasured or unknown influences, such as ambient environmental factors or internal component degradation, were not explicitly considered. Furthermore, the input distributions were assumed to be Gaussian, which may not always reflect real-world variability. Future implementations could incorporate statistical validation

techniques to confirm the appropriateness of distributional assumptions. Additionally, some deviations in the actual 2019 measurements may not represent true thermal anomalies but could instead result from noisy data, sensor drift, or logging errors. These uncertainties highlight the need to integrate simulation frameworks with robust data validation in future applications.

Despite these limitations, the proposed model demonstrates strong potential for practical application in predictive maintenance strategies. By combining real-time monitoring with Prediction Intervals (PIs) derived from uncertainty-aware forecasting, the framework could enable plant operators to observe not only actual thermal trends but also assess proximity to operational thresholds. Deviations from the predicted range, particularly those that are persistent or compounding, may serve as early warning indicators for initiating maintenance actions. This approach supports the concept of condition-based scheduling and could be integrated into Supervisory Control and Data Acquisition (SCADA) systems or IoT-based monitoring platforms to enhance operational reliability and responsiveness in hydropower environments.

V. Conclusion and Future Work

This study proposed a practical framework that combined a pretrained LSTM model with Sobol Global Sensitivity Analysis to construct monthly Prediction Intervals (PIs) for forecasting thermal behavior in hydropower generator cooling systems. By simulating realistic operational uncertainties, particularly under reduced flow and elevated water temperature conditions, the framework effectively identified seasonal risks of overheating, especially during the dry-season months.

Results showed that air temperatures were highly sensitive to generated heat, while stator temperature rose significantly when exposed to both reduced cooling water flow and elevated inlet water temperature. In contrast, water-side variables exhibited delayed but meaningful responses, reinforcing their role in long-term thermal monitoring. The use of PIs, derived from variability in input parameters under an assumed Gaussian distribution, enabled early detection of anomalous trends and supported an uncertainty-aware perspective in assessing system behavior.

This approach was computationally efficient, required no retraining, and was suitable for integration with SCADA or IoT-based monitoring systems. Future work will focus on improving interpretability via SHAP analysis, validating distributional assumptions, and extending the framework to other critical subsystems.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Chinachote Deevijit: Writing—original draft, methodology, formal analysis, investigation, and validation; Tanongkiat Kiatsiriroat: Supervision; Thoranis Deethayat: Supervision; Attakorn Asanakham: Conceptualization, supervision, and writing—review & editing. All authors had approved the final version.

Acknowledgment

The authors expressed their gratitude to the Department of Mechanical Engineering, Faculty of Engineering, Chiang Mai University, and the Electricity Generating Authority of Thailand (EGAT) for their invaluable financial support of this research.

References

[1] L. Sun, T. Liu, Y. Xie, D. Zhang, and X. Xia, "Real-time power prediction approach for turbine using deep learning techniques," *Energy*, vol. 233, 121130, Oct. 2021. doi: 10.1016/J.ENERGY.2021.121130

[2] V. Jadhav, A. Deodhar, A. Gupta, and V. Runkana, "Physics informed neural network for health monitoring of an air preheater," in *Proc. PHM Society European Conference*, vol. 7, no. 1, 2022, pp. 219–230. doi: 10.36001/PHME.2022.V7I1.3343

[3] H. Shu and H. Zhu, "Sensitivity analysis of deep neural networks," in *Proc. AAAI Conf. Artificial Intelligence*, Jan. 2019. doi: 10.1609/aaai.v33i01.33014943

[4] Y. Sun, "Deep neural network regression and sobol sensitivity analysis for daily solar energy prediction given weather data," *Open Access Theses*, Purdue Univ., May 2018. https://docs.lib.purdue.edu/open_access_theses/1460. Accessed: May 06, 2025.

[5] X. Wang, D. Xu, N. Qu, T. Liu, F. Qu, and G. Zhang, "Predictive maintenance and sensitivity analysis for equipment with multiple quality states," *Math. Probl. Eng.*, vol. 2021, no. 1, 4914372, Jan. 2021. doi: 10.1155/2021/4914372

[6] P. Miao, "Prediction-based maintenance of existing bridges using neural network and sensitivity analysis," *Adv. Civ. Eng.*, vol. 2021, no. 1, 4598337, Jan. 2021. doi: 10.1155/2021/4598337

[7] B. Mrzygłód, M. Hawryluk, M. Janik, and I. Olejarczyk-Wożeńska, "Sensitivity analysis of the artificial neural networks in a system for durability prediction of forging tools to forgings made of C45 steel," *Int. J. Adv. Manuf. Technol.*, vol. 109, no. 5–6, pp. 1385–1395, Jul. 2020. doi: 10.1007/S00170-020-05641-Y

[8] J. Herman and W. Usher, "SALib: An open-source Python library for sensitivity analysis," *J. Open Source Softw.*, vol. 2, no. 9, p. 97, Jan. 2017. doi: 10.21105/JOSS.00097

[9] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, 5th ed. Hoboken, NJ: Wiley, 2010. https://archive.org/details/buku-stat-montgomery-5

[10] W. H. Ng, C. R. Myers, S. McArt, and S. P. Ellner, "A time for every purpose: Using time-dependent sensitivity analysis to help understand and manage dynamic ecological systems," *Am. Nat.*, vol. 202, no. 5, pp. 630–654, Nov. 2023. doi: 10.1086/726143

[11] T. Mancini, H. Calvo-Pardo, and J. Olmo, "Prediction intervals for deep neural networks," arXiv Preprint, 2010.04044, 2020. https://arxiv.org/pdf/2010.04044