

Article

Switch Cabinet Temperature Prediction Using a Fusion of CNN and LSTM Neural Networks

Xu Sun ^{1,2} , Yun Chen ^{1,2}, Jiang Wei ^{1,2}, Qi Liu ³, Hui Guo ^{1,2,*}  and Ruijian Cheng ^{1,2,*} ¹ Institute of Applied Physics, Jiangxi Academy of Sciences, Nanchang 330096, China² Jiangxi Key Laboratory of Advanced Copper-Based Materials, Institute of Applied Physics, Jiangxi Academy of Sciences, Nanchang 330096, China³ Institute of Energy Research, Jiangxi Academy of Sciences, Nanchang 330096, China

* Correspondence: guohui@jxas.ac.cn (H.G.); chengruijian@jxas.ac.cn (R.C.)

Abstract

This study investigates the effectiveness of a hybrid convolutional neural network (CNN) and long short-term memory (LSTM) model for predicting the temperature of switchgear within electrical power systems. Given the critical importance of temperature monitoring for operational safety and stability, this research integrates CNNs and LSTMs to leverage their respective strengths in spatial feature extraction and temporal data processing. Utilizing a dataset from 2020 comprising hourly data points along with comprehensive environmental and operational variables, the model aims to deliver precise temperature predictions. Initial results indicate a high level of accuracy, with the CNN-LSTM model achieving an R^2 score of 0.95 and a mean absolute error of 0.12 °C, highlighting its significant potential to enhance the monitoring and management of safety in power systems.

Keywords: CNN-LSTM mode; temperature monitoring; power system safety



Academic Editor: Friedhelm
Schwenker

Received: 24 June 2025

Revised: 28 September 2025

Accepted: 4 October 2025

Published: 21 October 2025

Citation: Sun, X.; Chen, Y.; Wei, J.; Liu, Q.; Guo, H.; Cheng, R. Switch Cabinet Temperature Prediction Using a Fusion of CNN and LSTM Neural Networks. *Appl. Syst. Innov.* **2025**, *8*, 157. <https://doi.org/10.3390/asi8050157>

Copyright: © 2025 by the authors. Published by MDPI on behalf of the International Institute of Knowledge Innovation and Invention. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The temperature monitoring of switchgear is crucial for ensuring the operational safety and stability of electrical power systems [1–3]. Switchgear refers to a comprehensive set of power distribution devices, integrating one or more switching mechanisms along with associated control, measurement, signaling, and protection components, all connected through electrical and mechanical interconnections within structural units [4]. These assemblies are essential for controlling, distributing, and safeguarding electrical systems during power generation, transmission, distribution, and transformation. Typical components within switchgear include circuit breakers, isolators, load switches, operating mechanisms, transformers, and various protective devices, and their applications span a range of environments, from power plants and substations to residential complexes and high-rise buildings [5]. Switchgear plays a critical role in distributing electrical energy and protecting circuits within the power system [6]. Therefore, malfunctions in switchgear not only disrupt normal power supply, affecting essential services and daily activities—particularly in critical facilities such as hospitals, banks, and theaters—but also pose serious safety risks. Faults can lead to electrical fires, often triggered by overloads, short circuits, or poor contacts, resulting in overheating of circuits and equipment. Moreover, failures in switchgear compromise overall system stability and may trigger cascading failures that extend outages and exacerbate the problem [7]. These malfunctions can also degrade the quality of electrical energy, leading to increased energy consumption and higher electricity

costs [8]. In this context, temperature is one of the most critical indicators of switchgear health, as abnormal temperature rise is closely associated with insulation degradation, contact resistance increase, and thermal stresses on components. Sustained overheating can accelerate material aging and eventually lead to catastrophic failures. While direct temperature measurements can be obtained through sensors such as infrared detectors or thermocouples, these approaches face limitations in practice: installation of sensors in high-voltage environments can be costly, intrusive, and sometimes infeasible, particularly for enclosed or compact switchgear structures. Furthermore, real-time measurement at all potential hot spots is rarely achievable. Consequently, temperature estimation based on correlated electrical and environmental parameters provides a more flexible and scalable approach. Accurate estimation enables continuous monitoring, predictive maintenance, and early fault detection, thereby enhancing both the stability and operability of power systems.

Accurate temperature prediction of electrical switchgear has been extensively studied due to its direct implications for operational safety and system reliability [9–11]. Traditional monitoring approaches, such as threshold alarms, manual inspections, and thermal imaging, have been widely applied in practice [12]. However, these methods are limited in their ability to process large-scale multivariate time-series data and often fail to capture nonlinear dependencies that characterize thermal variations in power systems. To address these challenges, machine learning models including support vector machines (SVMs), random forests (RFs), and artificial neural networks (ANNs) have been introduced [13]. While these approaches improve predictive accuracy compared to conventional statistical methods [14], their effectiveness remains constrained by insufficient feature extraction and difficulty in modeling long-term temporal dependencies. With the rapid advancement of deep learning, more powerful architectures have been explored. Convolutional Neural Networks (CNNs) are particularly effective in extracting hierarchical spatial features and have demonstrated strong performance in applications such as equipment fault detection and process monitoring [15,16]. Conversely, Long Short-Term Memory (LSTM) networks excel in sequence modeling by capturing both short- and long-term dependencies, overcoming the limitations of standard recurrent networks [17–19]. Studies in energy forecasting and thermal system analysis highlight the advantages of LSTMs in time-series prediction tasks. Nonetheless, CNNs alone often neglect sequential dependencies, while LSTMs alone may underperform in handling complex spatial correlations, suggesting the need for integrated frameworks. Recent research has increasingly focused on hybrid CNN-LSTM models, which leverage the spatial feature extraction of CNNs and the temporal learning capability of LSTMs. These hybrid approaches have achieved promising results in fields such as renewable energy forecasting, air quality prediction, and industrial process optimization [20–23].

Despite these technological advances, the use of hybrid models that combine CNNs and LSTMs for temperature prediction in power system switchgear remains relatively unexplored [24]. This hybrid model harnesses the spatial feature extraction capabilities of CNNs and the temporal data processing strengths of LSTMs, offering the potential for improved accuracy in temperature predictions. This study aims to assess the effectiveness of CNN-LSTM hybrid models in predicting switchgear temperatures within power systems, providing robust technical support for enhancing the safety and reliability of power system operations.

2. Methodology

2.1. Dataset and Switchgear

The dataset used in this study comprises hourly data collected throughout 2020, from 1 January to 31 December, amounting to 8784 data points. Each data point includes 11 features, resulting in a data matrix of dimensions (8784, 11). The recorded features are:

'Date', 'Time', 'Flue Gas Velocity (m/s)', 'O₂ Content (%)', 'Measured SO₂ Concentration (mg/Nm³)', 'Standardized SO₂ Concentration (mg/Nm³)', 'Particulate Matter Concentration (mg/Nm³)', 'Standardized Particulate Matter Concentration (mg/Nm³)', 'Flue Gas Flow (m³/h)', 'Inlet Temperature (°C)', and 'Outlet Temperature (°C)'. The primary variable of interest, 'Outlet Temperature (°C)', serves as the dependent variable in this temperature prediction study. This dataset provides a comprehensive view of the environmental and operational conditions affecting switchgear temperature, enabling the application of advanced predictive models.

The data was obtained from a 110 kV indoor high-voltage switchgear cabinet operating in a coal-fired power plant located in eastern China. The switchgear is part of the plant's flue gas desulfurization (FGD) system, and it manages the electrical power distribution for induced draft fans, circulating pumps, and auxiliary equipment. The cabinet model belongs to a widely used metal-enclosed switchgear series, rated for a nominal current of 3150 A and a short-time withstand current of 40 kA (3 s). The load profile during the observation period primarily reflected continuous base-load operation, with average load utilization around 70–85% of rated capacity. The inlet source refers to the flue gas channel upstream of the FGD absorber, while the outlet temperature represents the temperature of flue gas downstream of the absorber section. The cabinet is installed in a controlled indoor environment with ambient temperatures varying seasonally between 5 and 32 °C. These specifications are critical for interpreting the dataset, as they provide context for the thermal behavior of the switchgear and the applicability of the predictive models to other industrial scenarios. The testing process consists of several key steps. First, the dataset was divided into training (70%), validation (15%), and testing (15%) subsets to ensure that the models were evaluated on unseen data.

2.2. CNNs

Convolutional Neural Networks (CNNs) are a class of deep learning models specifically designed to handle high-dimensional data with grid-like topology, such as images, videos, and increasingly, other structured data types like time-series data and spectrograms [25]. The key innovation of CNNs lies in their ability to exploit spatially local correlations through convolutional operations, significantly reducing the number of trainable parameters compared to fully connected neural networks while maintaining high representational power.

The fundamental building block of a CNN is the convolutional layer, which applies a set of learnable kernels or filters to the input data. Each filter is typically small in spatial extent but spans the full depth of the input feature maps. Mathematically, for an input image $X \in \mathbb{R}^{H \times W \times D}$ with height H , width W , and depth D , and a filter $F \in \mathbb{R}^{k \times k \times D}$, the convolution operation is defined as

$$S(i, j) = (X * F)(i, j) = \sum_{m=1}^k \sum_{n=1}^k \sum_{d=1}^D X(i+m, j+n, d) \cdot F(m, n, d)$$

where $S(i, j)$ denotes the activation at location (i, j) in the resulting feature map, and $*$ represents the convolution operator. The resulting feature maps highlight salient local patterns such as edges, textures, or higher-level structures depending on the depth of the network.

To introduce nonlinearity and enhance representational capacity, CNNs apply activation functions such as the Rectified Linear Unit (ReLU) after convolutional operations. In addition, pooling layers (e.g., max pooling or average pooling) are often employed to down sample feature maps by summarizing activations within a local region. Pooling not only reduces computational cost and memory footprint but also introduces a degree of

translation invariance, which is desirable in many computer vision tasks. Deeper CNN architectures stack multiple convolutional and pooling layers, enabling hierarchical feature extraction where low-level features are captured in early layers, and high-level semantic concepts emerge in later layers. This spatial hierarchy learning is one of the most significant strengths of CNNs, allowing them to outperform traditional feature-engineering-based methods in visual recognition tasks.

In addition to convolution and pooling, modern CNN architectures incorporate batch normalization to stabilize training, dropout to prevent overfitting, and residual connections to enable training of extremely deep networks. Furthermore, CNNs are not limited to image recognition; they have been successfully applied to natural language processing via character-level or word-level embeddings, audio processing, medical image diagnosis, and even time-series forecasting where local temporal correlations can be treated analogously to spatial correlations.

2.3. LSTMs

Recurrent Neural Networks (RNNs) are designed for sequential data modeling, yet they often struggle with learning long-term dependencies due to the vanishing and exploding gradient problem [26,27]. Long Short-Term Memory (LSTM) networks were proposed as a solution to this limitation, introducing a memory cell mechanism that allows information to persist across longer time horizons. This architecture has since become a cornerstone for sequence modeling tasks, including natural language processing, speech recognition, and time-series forecasting. An LSTM unit consists of a cell state C_t , which acts as the memory of the network, and three types of gates—forget gate, input gate, and output gate—that regulate the flow of information [28,29]. Specifically, the forget gate determines how much of the past memory should be retained or discarded. The input gate controls how much of the new input information is written to the memory. The output gate decides how much of the memory is exposed to the next hidden state. Mathematically, the operations of an LSTM at time step t can be described as

$$\begin{aligned} f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\ C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t &= o_t * \tanh(C_t) \end{aligned}$$

where x_t is the input at time step t , h_t is the hidden state, C_t is the cell state, σ is the sigmoid activation function, and $*$ denotes the Hadamard (element-wise) product.

The gating mechanism ensures that gradients can propagate effectively over long sequences, thereby mitigating the vanishing gradient problem. By selectively updating and retaining memory, LSTMs capture both short-term patterns and long-term dependencies in sequential data. In practice, LSTMs have proven to be highly effective in diverse applications. In natural language processing, they are widely used for language modeling, machine translation, and text generation, where the meaning of words depends heavily on their context in long sequences. In time-series forecasting, LSTMs excel at modeling temporal dependencies in fields such as financial prediction, climate modeling, and energy demand forecasting. Additionally, LSTMs are frequently combined with CNNs to leverage both spatial and temporal features, such as in video classification or sensor-based human activity recognition.

Despite their strengths, LSTMs are computationally expensive due to the complex gating mechanisms and large parameter sets. Recent alternatives such as Gated Recurrent Units (GRUs) have been proposed to simplify the architecture while retaining comparable performance. More recently, Transformer-based models have shown superior scalability and performance in many sequence modeling tasks, though LSTMs remain an indispensable architecture in domains requiring efficient sequence learning with moderate data scales.

2.4. Model Training

In this study, three predictive models were implemented: a CNN model, an LSTM model, and a hybrid CNN–LSTM model. The CNN architecture consisted of two convolutional layers with 32 and 64 filters, respectively, each using a kernel size of 3 and ReLU activation. A max-pooling layer with pool size 2 followed each convolutional layer, and a dropout rate of 0.2 was applied to mitigate overfitting. The LSTM model comprised two stacked LSTM layers with 128 and 64 hidden units, respectively, followed by a fully connected dense layer with linear activation to produce the output. For the hybrid CNN–LSTM architecture, the first stage employed the same convolutional layers as in the standalone CNN to extract local spatial features from the input data. The resulting feature maps were then reshaped and fed into the LSTM layers, which modeled temporal dependencies across sequential inputs. This combination enables the model to capture both short-term correlations and long-term dependencies, improving overall predictive accuracy.

All models were trained using the Adam optimizer with an initial learning rate of 0.001 and mean squared error (MSE) as the loss function. The batch size was set to 64, and the maximum number of epochs was 200, with early stopping applied if validation loss did not improve for 20 consecutive epochs. Data were normalized using min–max scaling prior to training, and performance evaluation was carried out on an independent test set using R^2 , Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE) metrics. These implementation details ensure transparency and reproducibility of the experiments. All experiments were conducted on a workstation equipped with an NVIDIA RTX 3090 GPU and 64 GB RAM, ensuring efficient model training. Performance evaluation was carried out using four commonly adopted metrics: coefficient of determination (R^2), mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). This systematic process ensures that the testing framework is both rigorous and reproducible.

2.5. Presentation of the Testing Problem

The primary testing problem addressed in this study is the accurate prediction of switchgear outlet temperature under varying operational and environmental conditions. The challenge arises from the complex nonlinear relationships and strong temporal dependencies inherent in the data, as temperature dynamics are jointly influenced by multiple correlated factors such as inlet temperature, flue gas flow, and pollutant concentrations. Traditional approaches and single-model deep learning methods are insufficient for capturing both spatial correlations among input variables and long-term dependencies across time. Therefore, this study specifically designs a CNN–LSTM hybrid framework and evaluates its predictive performance against baseline CNN and LSTM models. The testing problem can thus be formally defined as a supervised regression task, where the goal is to minimize predictive error in outlet temperature forecasting while maintaining robust generalization ability across unseen data.

3. Results and Discussion

3.1. Correlation Analysis

Figure 1 illustrates the outlet temperature throughout 2020, indicating that the outlet temperature predominantly remains stable, fluctuating around 50 °C for much of the year. This stability suggests a consistent operational environment under normal working conditions. Notably, a significant temperature drop occurs between late April and early June, where temperatures decrease to nearly 15 °C. This sharp decline is attributed to a planned operational intervention and thus was excluded from the primary stability analysis to avoid skewing the results. Following this period, temperatures rapidly recover to normal levels, suggesting a possible correction or adjustment in the system, whether automated or manual. The rapid recovery implies that robust control mechanisms are in place to maintain the temperature within the desired range.

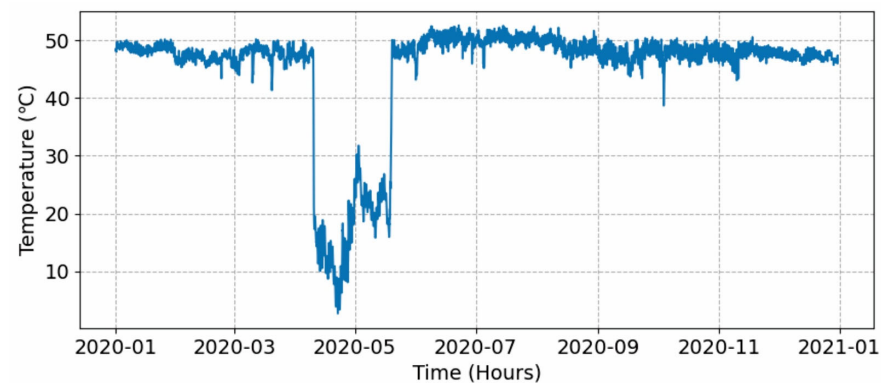


Figure 1. Outlet temperature during 2020.

Figure 2 presents a comprehensive view of the linear relationships among several variables pertinent to the operational conditions of electrical equipment [30]. A robust positive correlation of 0.95 exists between inlet and outlet temperatures. This strong correlation indicates that thermal conditions at the inlet significantly influence the temperatures observed at the outlet, reflecting a relatively stable thermal process within the system. Furthermore, the concentrations of particulate matter and sulfur dioxide exhibit a substantial correlation of 0.87. This relationship implies that conditions conducive to high particulate emissions are also favorable for increased sulfur dioxide concentrations, potentially due to common sources or environmental conditions affecting both parameters simultaneously. Conversely, the correlation matrix reveals strong negative correlations involving oxygen content. Specifically, oxygen content exhibits a strong inverse relationship with both inlet and outlet temperatures, with coefficients of -0.91 and -0.88 , respectively. These findings suggest that increased oxygen content is associated with lower temperatures at both the inlet and outlet, potentially attributable to the cooling effects or dilution of inlet gases. The correlation between oxygen content and the concentrations of sulfur dioxide and particulate matter, both at -0.79 , further supports the hypothesis that higher oxygen levels may enhance combustion completeness, thereby reducing these pollutants. Additionally, outlet temperature demonstrates strong positive correlations of 0.95 with flue gas flow and 0.87 with inlet temperature, reinforcing the notion that higher gas flows and elevated inlet temperatures predictably lead to higher outlet temperatures. This relationship underscores the thermodynamic consistency within the system, wherein input energy and mass flow significantly influence thermal output.

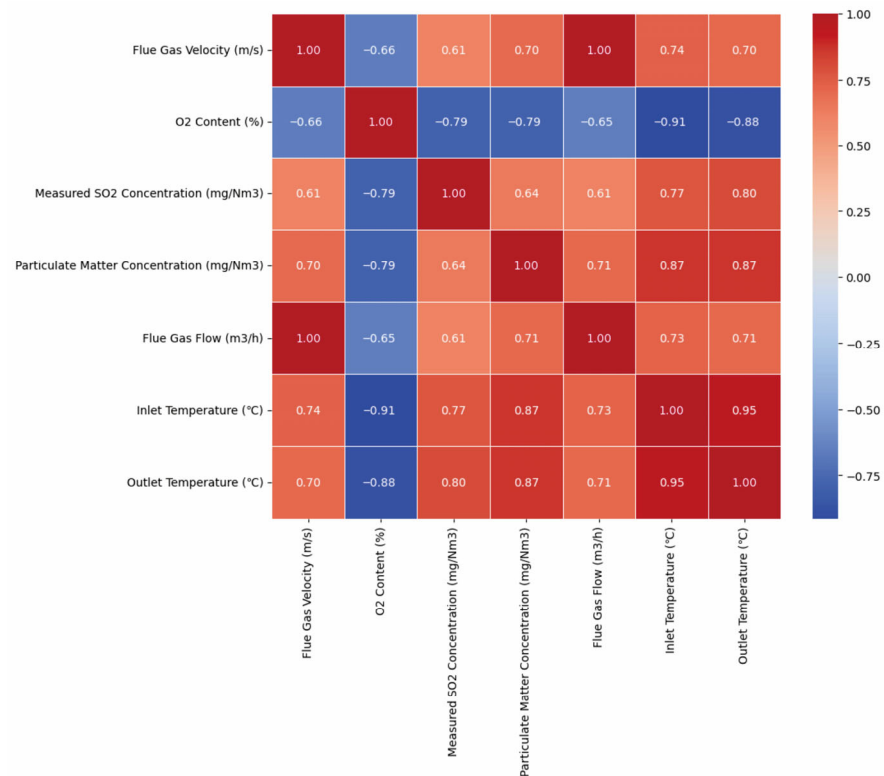


Figure 2. Pearson correlation coefficient matrix.

3.2. CNN Model Regression Results

Figure 3 depicting the performance of the CNN in predicting temperatures over time illustrates the model's robust capability to accurately replicate actual temperature fluctuations. This accuracy is reflected in the R^2 score of 0.81, indicating a strong linear correlation between predicted and actual temperatures. Such a score suggests that approximately 81% of the variance in the actual temperature data is explained by the predictions from the CNN model. Furthermore, the MAE of 0.22 °C indicates that the predictions deviate from the actual values by an average of only two-tenths of 1 °C, underscoring the model's precision.

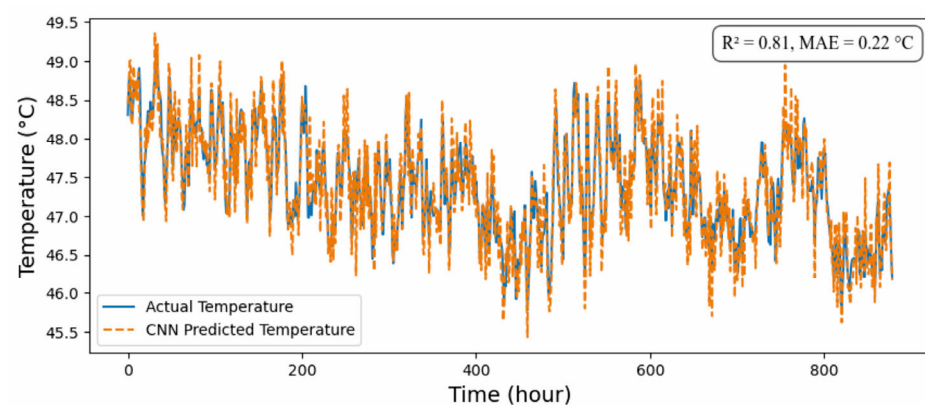


Figure 3. CNN model predicted outlet temperature.

3.3. LSTM Model Regression Results

Figure 4 presents the results from a Long Short-Term Memory (LSTM) network tasked with predicting temperatures, visualized by comparing actual and predicted values over a series of hours. The LSTM model achieves a commendable R^2 score of 0.92, indicating a high degree of linear correlation between the observed and predicted temperature values.

This score implies that 92% of the variance in the actual temperature is explained by the model's predictions. Furthermore, the Mean Absolute Error (MAE) of 0.15 °C highlights the model's precision, with an average prediction error of only 0.15 °C, demonstrating its efficacy in closely matching the actual temperature data points. The LSTM's ability to capture complex dependencies in temperature fluctuations over time is evident. The close alignment of the LSTM-predicted temperature (represented by the dotted red line) with the actual temperature (indicated by the solid blue line) across the graph indicates that the model effectively learns and replicates the temporal dynamics inherent in the dataset.

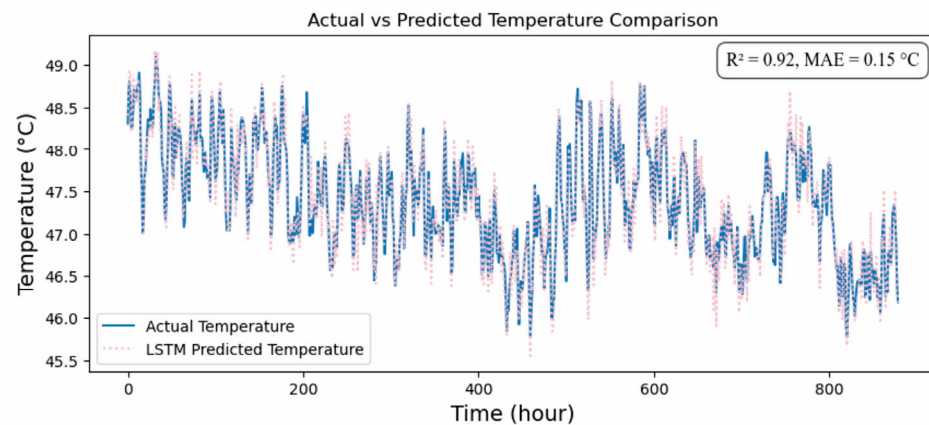


Figure 4. LSTM model predicted outlet temperature.

3.4. CNN-LSTM Regression Results

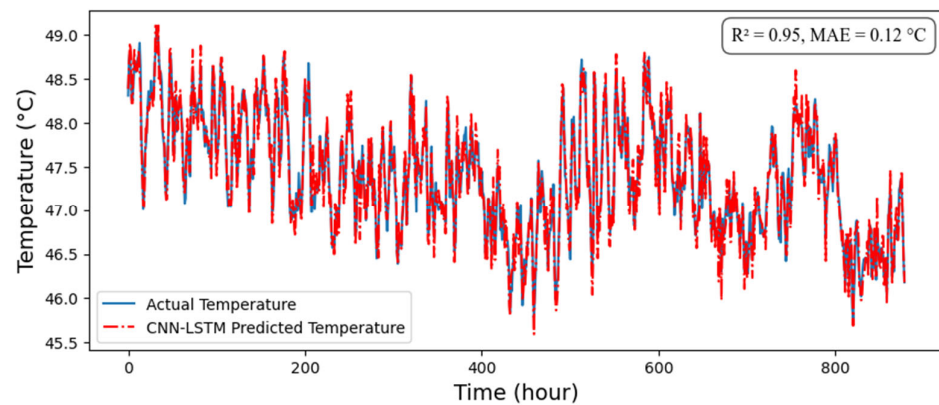
The combined CNN-LSTM model, as illustrated in the graph, demonstrates exceptional capability in predicting temperature variations over time, evidenced by a high R^2 score of 0.95 and a Mean Absolute Error (MAE) of 0.12 °C. The R^2 value indicates that 95% of the variance in the actual temperature readings is accounted for by the model's predictions, signifying strong predictive accuracy. The low MAE further reinforces the model's precision, suggesting that the average deviation of the predicted temperatures from the actual temperatures is minimal, at slightly greater than a tenth of 1 °C.

3.5. Presentation of the Testing Results and a Summary

Table 1 presents a comprehensive comparison of the three models. The CNN model achieved an R^2 of 0.81 with an MAE of 0.22 °C, while the LSTM model outperformed CNN with an R^2 of 0.92 and an MAE of 0.15 °C. The proposed CNN-LSTM hybrid model delivered the highest accuracy, achieving an R^2 of 0.95 and an MAE of 0.12 °C. Similar improvements were observed for RMSE and MAPE values, highlighting the robustness of the hybrid approach. Figure 5 further illustrates the residual distributions of the three models, where the CNN-LSTM shows the narrowest spread, indicating more reliable predictions. These findings confirm that the hybrid model successfully leverages the complementary strengths of CNNs in feature extraction and LSTMs in temporal modeling. In summary, the testing results demonstrate that the CNN-LSTM model provides best performance in temperature prediction for switchgear systems among three models, with potential for integration into real-time monitoring platforms to improve fault prevention and operational safety.

Table 1. The regression model and performance.

Model	R^2	MAE
CNN	0.81	0.22
LSTM	0.92	0.15
CNN-LSTM	0.95	0.12

**Figure 5.** CNN-LSTM model predicted outlet temperature.

4. Conclusions

This study investigated the application of hybrid CNN and LSTM models for the prediction of temperature fluctuations within switchgear systems. The findings demonstrate that the hybrid model exhibits a high R^2 score and a low mean absolute error, indicating its efficacy in accurately forecasting temperature variations. Such predictive accuracy is essential for enhancing the operational safety and efficiency of power systems. This hybrid methodology effectively captures the complex spatial and temporal patterns inherent in the temperature data, while also dynamically adjusting its predictive capability through an error-inverse weighting mechanism. The precision in temperature forecasting afforded by this model could significantly contribute to the prevention of system malfunctions and the maintenance of stable operations in critical infrastructure. Future research endeavors may aim to refine the model's architecture further, explore advanced techniques such as transfer learning, and evaluate its applicability in other domains of predictive maintenance within industrial contexts.

Author Contributions: Conceptualization and writing—original draft preparation X.S.; methodology, Y.C.; software, Q.L.; validation, J.W.; resources, H.G. and R.C.; data curation, Y.C.; writing—review and editing, H.G. and R.C.; visualization, X.S.; supervision, H.G.; project administration, R.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Jiangxi Key Laboratory of Advanced Copper-based Materials, grant number 2024SSY05021, Jiangxi Provincial Natural Science Foundation grant number 20232BAB214012, The emerging interdisciplinary cultivation Program of Jiangxi Academy of Sciences, grant number 2022YXXJC0102 and Ganpo Promising Talents Supporting Plan: Talent Development Project of Leading Academic and Technological Researchers in Key Disciplines, grant number 20232BCJ23099. All authors express gratitude to all individuals who provided valuable support throughout this study. Special acknowledgment is extended to colleagues at the Institute of Applied Physics, Jiangxi Academy of Sciences, for their expert contributions and dedication to the success of this project. Additionally, the financial backing from the Key Research and Development Program of Jiangxi Academy of Sciences, grant numbers 2022YBSG22022 and 2022YJC2014 was instrumental in facilitating data acquisition and the application of advanced analytical methodologies in this research.

Data Availability Statement: 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.

Acknowledgments: All authors contributed extensively to the work presented in this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Li, Y.; Cheng, W. A novel temperature monitoring system based on FBG for switchgear. In Proceedings of the 2010 Asia-Pacific Power and Energy Engineering Conference, Chengdu, China, 28–31 March 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 1–4.
- Shuxin, L.; Yundong, C.; Chunguang, H.; Xiaoming, L.; Jing, L. Development of on-line monitoring system of switchgear. In Proceedings of the 2011 1st International Conference on Electric Power Equipment-Switching Technology, Xi'an, China, 23–27 October 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 295–298.
- Yan, Y.; Qi, D.; Gu, H. A real-time IR-fusion switchgear contact monitoring system (SCMS). *IEEE Access* **2017**, *5*, 12114–12124.
- Flurschein, C. Switchgear. A review of progress. *Proc. IEE-Part A Power Eng.* **1956**, *103*, 239–262.
- Schaffer, J.S.; Hazel, T. Ensuring switchgear integrity in High-Power installations. *IEEE Trans. Ind. Appl.* **2014**, *51*, 2641–2650. [[CrossRef](#)]
- Smeets, R.; Van der Sluis, L.; Kapetanovic, M.; Peelo, D.F.; Janssen, A. *Switching in Electrical Transmission and Distribution Systems*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
- Gupta, M.; Elahi, B. Prioritizing Failure-Contributing Causes in Switchgear Systems. In Proceedings of the 2025 Annual Reliability and Maintainability Symposium (RAMS), Destin, FL, USA, 27–30 January 2025; IEEE: Piscataway, NJ, USA, 2025; pp. 1–6.
- Razip, H.D.; Khairudin, F.R.; Harun, M.N. The impact of under sized cable on MV vacuum metal clad switchgear during fault. In Proceedings of the 2012 IEEE International Power Engineering and Optimization Conference, Melaka, Malaysia, 6–7 June 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 312–316.
- Zeng, L.; Ge, X.; Zheng, W. DATA-DRIVEN transient temperature rise prediction model for medium-voltage switchgear. In Proceedings of the CIRED 2021-The 26th International Conference and Exhibition on Electricity Distribution, Online, 20–30 September 2021; IET: Hertfordshire, UK, 2021; pp. 162–166.
- Matin, M.; Dehghanian, A.; Zeinaddini, A.H.; Darijani, H. Interpretable machine learning modeling of temperature rise in a medium voltage switchgear using multiphysics CFD analysis. *Case Stud. Therm. Eng.* **2025**, *65*, 105585.
- Shen, G.; Zhang, L.; Zhu, J.; Xu, M.; Zhu, M.; Chong, P. Research on Temperature and Humidity Prediction Model of High Voltage Switchgear Based on SSA-BP Algorithm. In Proceedings of the 2022 4th International Conference on Smart Power & Internet Energy Systems (SPIES), Beijing, China, 9–12 December 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1059–1064.
- Alsumaidae, Y.A.M.; Yaw, C.T.; Koh, S.P.; Tiong, S.K.; Chen, C.P.; Ali, K. Review of medium-voltage switchgear fault detection in a condition-based monitoring system by using deep learning. *Energies* **2022**, *15*, 6762.
- Hoffmann, M.W.; Wildermuth, S.; Gitzel, R.; Boyaci, A.; Gebhardt, J.; Kaul, H.; Amihai, I.; Forg, B.; Suriyah, M.; Leibfried, T.; et al. Integration of novel sensors and machine learning for predictive maintenance in medium voltage switchgear to enable the energy and mobility revolutions. *Sensors* **2020**, *20*, 2099. [[CrossRef](#)] [[PubMed](#)]
- Heo, J.; Zhang, Y.; Bae, J.; Lim, S.; Shin, W.G.; Kim, S.B. Explainable Machine Learning-based Meta-modeling for Predicting Fire Damage to Motor Control Cabinets in a Switchgear Room. *Nucl. Eng. Technol.* **2025**, *58*, 103862.
- Sun, L.; Li, Y.; Gong, F.; Liu, G.; Yang, Z. Application of CNN in condition monitoring of high-voltage switchgear. In Proceedings of the International Conference on Physics, Photonics, and Optical Engineering (ICPPOE 2024), Singapore, 8–10 November 2025; SPIE: Bellingham, WA, USA, 2025; pp. 383–389.
- Ishaq, A.; Junaid, M.; Hussain, G.A.; Khan, S.U.; Chen, Y.; Yu, D. Partial Discharge Defect Classification in MV Switchgear by Using CWT and Deep Learning Approach. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 1–12. [[CrossRef](#)]
- Yan, J.; Wang, Y.; Wang, J.; Geng, Y.; Srinivasan, D. A novel transformer-LSTM-based multitask network for partial discharge condition assessment in gas-insulated switchgear. In Proceedings of the 2024 10th International Conference on Condition Monitoring and Diagnosis (CMD), Gangneung, Republic of Korea, 20–24 October 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 98–101.
- Santillan, P.B.; Magwili, G.V. Time-series prediction of busbar temperature rise based on LSTM neural network. In Proceedings of the 2023 IEEE 5th Eurasia Conference on IOT, Communication and Engineering (ECICE), Yunlin, Taiwan, 27–29 October 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 133–138.
- Moon, S.; Cho, H.; Koh, E.; Cho, Y.S.; Oh, H.L.; Kim, Y.; Kim, S.B. Remanufacturing decision-making for gas insulated switchgear with remaining useful life prediction. *Sustainability* **2022**, *14*, 12357. [[CrossRef](#)]
- Alsumaidae, Y.A.M.; Yaw, C.T.; Koh, S.P.; Tiong, S.K.; Chen, C.P.; Yusaf, T.; Abdalla, A.N.; Ali, K.; Raj, A.A. Detection of corona faults in switchgear by using 1D-CNN, LSTM, and 1D-CNN-LSTM methods. *Sensors* **2023**, *23*, 3108.

21. Alsumidaee, Y.A.M.; Paw, J.K.S.; Yaw, C.T.; Tiong, S.K.; Chen, C.P.; Yusaf, T.; Benedict, F.; Kadirgama, K.; Hong, T.C.; Abdalla, A.N. Fault detection for medium voltage switchgear using a deep learning hybrid 1D-CNN-LSTM model. *IEEE Access* **2023**, *11*, 97574–97589.
22. Alsumidaee, Y.A.M.; Yahya, M.M.; Yaseen, A.H. Reliable Mechanical Fault Diagnosis in Medium Voltage Electrical Switchgear Using a 1D-Convolutional Neural Network. *J. Eur. Syst. Autom.* **2024**, *57*, 1461.
23. Li, Z.; Qu, N.; Li, X.; Zuo, J.; Yin, Y. Partial discharge detection of insulated conductors based on CNN-LSTM of attention mechanisms. *J. Power Electron.* **2021**, *21*, 1030–1040. [[CrossRef](#)]
24. Zhou, N.; Xu, Y.; Cho, S.; Wee, C.T. A systematic review for switchgear asset management in power grids: Condition monitoring, health assessment, and maintenance strategy. *IEEE Trans. Power Deliv.* **2023**, *38*, 3296–3311. [[CrossRef](#)]
25. Purwono, P.; Ma'arif, A.; Rahmانيar, W.; Fathurrahman, H.I.K.; Frisky, A.Z.K.; Haq, Q.M.U. Understanding of convolutional neural network (cnn): A review. *Int. J. Robot. Control Syst.* **2022**, *2*, 739–748.
26. Van Houdt, G.; Mosquera, C.; Nápoles, G. A review on the long short-term memory model. *Artif. Intell. Rev.* **2020**, *53*, 5929–5955. [[CrossRef](#)]
27. Grossberg, S. Recurrent neural networks. *Scholarpedia* **2013**, *8*, 1888. [[CrossRef](#)]
28. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [[CrossRef](#)] [[PubMed](#)]
29. Zargar, S. Introduction to Sequence Learning Models: RNN, LSTM, GRU. Ph.D. Thesis, Department of Mechanical and Aerospace Engineering, North Carolina State University, Raleigh, NC, USA, 2021. Volume 37988518.
30. Schober, P.; Boer, C.; Schwarte, L.A. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth. Analg.* **2018**, *126*, 1763. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.