

Article

Research on Lightning Prediction Based on GCN-LSTM Model Integrating Spatiotemporal Features

Wei Zhou, Wenqiang Wang and Xupeng Wang 

School of Information and Control Engineering, Qingdao University of Technology, 777 Jialingjiang Road, Huangdao District, Qingdao 266520, China; zhouwei@qut.edu.cn (W.Z.); wangwenqiang@stu.qut.edu.cn (W.W.)

* Correspondence: wangxupeng@qut.edu.cn

Abstract: To overcome the limitations of spatiotemporal feature extraction that are inherent in conventional lightning warning algorithms relying solely on temporal analysis, we propose a novel prediction framework integrating a Graph Convolutional Network (GCN), Long Short-Term Memory (LSTM) architecture, and a multi-head attention mechanism. The methodology innovatively constructs station adjacency matrices based on geographical distances between meteorological monitoring stations in Qingdao, Shandong Province, China, where GCN layers capture inter-station spatial dependencies while LSTM units extract localized temporal dynamics. A dedicated multi-head attention module was developed to enable adaptive fusion of global spatiotemporal patterns, significantly enhancing lightning warning level prediction accuracy at target locations. The GCN-LSTM model achieved 93% accuracy, 59% precision, 64% recall, and a 59% F1 score. Experimental evaluation on operational meteorological data demonstrated the model's superior performance: it achieved statistically significant accuracy improvements of 6% ($p = 0.019$), 3% ($p = 0.026$), and 2% ($p = 0.03$) over conventional LSTM, TGCN, and CNN-RNN baselines, respectively. Comprehensive assessments through precision–recall analysis, confusion matrix decomposition, and spatial generalizability tests confirmed the framework's robustness. The key theoretical advancement introduced by this study lies in the synergistic coupling of graph-based spatial modeling with deep temporal sequence learning, augmented by attention-driven feature fusion—an architectural innovation addressing critical gaps in existing single-modality approaches. This methodology establishes a new paradigm for extreme weather prediction with direct applications in lightning hazard mitigation.

Keywords: graph convolutional network; long short-term memory network; attention mechanism; lightning warning; spatiotemporal characteristics



Academic Editor: Vijay Tallapragada

Received: 3 March 2025

Revised: 3 April 2025

Accepted: 9 April 2025

Published: 11 April 2025

Citation: Zhou, W.; Wang, W.; Wang, X. Research on Lightning Prediction Based on GCN-LSTM Model Integrating Spatiotemporal Features. *Atmosphere* **2025**, *16*, 447. <https://doi.org/10.3390/atmos16040447>

Copyright: © 2025 by the authors. 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

Lightning is a strong local convective weather phenomenon [1]. It typically occurs between clouds (intra-cloud, or IC, lightning) or between clouds and the ground (cloud-to-ground, or CG, lightning) and is recognized as one of the ten most severe natural disasters worldwide. Lightning disasters are often accompanied by heavy rain, severe storms, hail, and other intense convective weather, generating powerful currents, extreme temperatures, and intense electromagnetic radiation [2–4]. These effects can cause significant damage to buildings, distribution systems, and electronic devices and can even trigger forest fires and explosions in oil fields and chemical plants, leading to substantial economic losses and casualties. With rapid social and economic development relying on information technology, lightning disasters have become a major public hazard in the information age [5–7].

However, it is difficult to predict lightning events and prevent them in advance because of suddenness and randomness of lightning events, often resulting in unavoidable economic losses and threats to life. Nevertheless, as people's understanding of the mechanisms of lightning deepens, an increasing number of proactive lightning protection technologies have been applied in real scenarios [8,9]. They can provide early warning solutions before lightning strikes, reducing or even eliminating economic and life losses. The study and application of proactive lightning protection measures are providing more effective solutions to mitigate the impacts of lightning hazards.

Devices commonly used for detecting lightning include atmospheric electric field instruments, meteorological radar, and lightning locators [10]. Existing research primarily focuses on utilizing traditional meteorological observation data and physical models to predict lightning occurrences [11–14]. However, it is important to note that the generation of lightning involves complex nonlinear processes such as charge accumulation, convective activity, and breakdown discharge. These physical phenomena exhibit irregularity in their spatial distribution and display instantaneous, sudden temporal characteristics. Current physical models often require simplifications of atmospheric dynamics and charge distribution to simulate these complex processes, which limits their ability to accurately capture changes in localized meteorological phenomena. Statistical methods [15] primarily predict events by summarizing physical laws from historical observational data, and they also require an in-depth understanding of the subject matter being statistically forecast. However, lightning events are highly random and uncertain, and relying solely on historical data often fails to achieve a high accuracy rate. In addition, traditional machine learning methods [16,17] rely on a deep understanding of the data and domain for model training, requiring manual intervention in selecting features. However, the importance of features in lightning activity changes over time, and traditional machine learning methods cannot adaptively handle these important features dynamically. Manual preprocessing and cleaning of lightning data are not only time-consuming but also prone to human errors that can compromise data quality. It was not until the emergence of deep learning that it became possible to automatically learn and extract features through a multi-layered structure, without the need for manual design or selection of features. This reduces the reliance on expert knowledge and enables direct learning of complex and abstract feature representations from raw data, overcoming the drawbacks of manual feature selection inherent in traditional machine learning. However, current research predominantly utilizes a single model (such as LSTM) for lightning prediction and fail to adequately extract spatial features from the data. Moreover, most studies focus solely on data from a single monitoring station during model training, neglecting the influence of surrounding stations, thereby limiting the model's performance in spatial dimensions. In order to address these challenges, we propose a spatiotemporal feature extraction model and introduce an attention mechanism to enhance the model's ability to perceive global features, enabling it to dynamically select the most relevant information at each time step.

In summary, our contributions are as follows:

- Extracting spatial features from non-Euclidean lightning data through graph convolution operations, overcoming the limitations of traditional models in handling spatial data from stations;
- Innovatively combining GCN and LSTM to effectively enhance the model's spatial feature extraction capabilities and time series modeling abilities;
- Introducing an attention mechanism that processes different parts of the sequence at varying levels, thereby comprehensively extracting the global features of the data.

The remainder of this paper is organized as follows: Section 2 reviews related work. Section 3 discusses relevant technologies. Section 4 describes the data sources. Section 5

defines the problem and the system model. In Section 6, we present the experimental results and analysis. Finally, Section 7 concludes the paper.

2. Related Work

The non-stationarity and high complexity of lightning data render traditional prediction methods inadequate for accurately capturing the long-term dependencies associated with lightning occurrences and the intricate interactions among various meteorological elements. Consequently, these methods fail to extract advantageous features from non-stationary data, which results in limited accuracy of prediction models. This limitation persisted until the advent of deep learning techniques [18], which are adept at managing complex nonlinear relationships, discerning dependencies in extensive time series, and extracting valuable features from high-dimensional data.

Riyang Bao et al. [19] proposed an integration of the LSTM network with the ordinary kriging algorithm for inputting the electric field time series data measured by an atmospheric electric field instrument into the LSTM network for training; they employed ordinary kriging interpolation to interpolate data from the network site, thereby obtaining the electric potential distribution, enabling relatively accurate predictions of the areas where lightning strikes may occur. In addition, Riyang Bao et al. [20] proposed a method for spatiotemporal localization of lightning by integrating the enhanced ResNet50 model with an MLP neural network. A sparse autoencoder was employed to extract time series data features from multiple electric field measurement points to construct visual images, and then a multilayer perceptron was used to determine the precise locations of lightning occurrences. Xu Yang et al. [21] proposed a lightning prediction method based on Convolutional Neural Networks (CNNs) and Bi-directional Long Short-Term Memory (Bi-LSTM) networks, where atmospheric electric field signals are first decomposed using BEADS to obtain denoised signals, which are subsequently processed by a CNN to extract spatial features. An atmospheric electric field signal prediction model was then established utilizing Bi-LSTM. The experimental results indicated that the model significantly enhanced lightning prediction accuracy. Tao Guo et al. [22] utilized atmospheric electric field data in conjunction with lightning localization data to label positive and negative samples for assessing the occurrence of lightning during atmospheric electric field fluctuations. An LSTM network was employed to analyze and predict these time series data. Experimental results demonstrated that this direct prediction method achieved high accuracy within a short time frame. Mingyue Lu et al. [23] proposed a Residual Network (ResNet) for lightning monitoring with multi-source spatial data based on deep learning and compared its performance with that of GoogLeNet and DenseNet. The proposed system achieved significant output by applying stepwise sensitivity analysis and one-factor sensitivity analysis. Ling Fan et al. [24] introduced a 3D U-Net-based Light3DUnet model to simulate cloud-to-ground (CG) and intra-cloud (IC) lightning activities for ground-based proximity forecasting. Yang et al. [25] proposed a thunderstorm identification method based on an amalgamation of the shrapnel distribution region's area and meteorological radar reflectivity. The method was tested on 312 thunderstorms during 17 weather processes in Nanjing, and the identification parameters were optimized to enhance efficiency in thunderstorm identification for forthcoming forecasts. John et al. [26] developed 'LightningCast', a CNN-based system for forecasting approaching thunderstorms. The system was trained on geosynchronous satellite data and deployed for comprehensive lightning activity analysis. Its primary feature is the transformation of voluminous satellite imagery data into precise predictions, demonstrating robust forecasting capabilities. Zhou H et al. [27] proposed an encoder–decoder architecture system, LightningNet, based on multi-source observation data; LightningNet is capable of predicting the approach of lightning by integrating six

infrared bands' brightness temperatures, synthetic reflectance, and ground flash densities from the Himawari-8 satellite as predictive factors. Test results demonstrated that LightningNet exhibited strong performance in lightning forecasting within the first hour, particularly when utilizing data from all three sources, outperforming models based on single- or dual-source data significantly.

Although the aforementioned studies have made certain progress in the field of lightning prediction, most are based on a singular prediction model, which may not optimally utilize the spatial and temporal dimensions inherent in the data. Moreover, even though some models employ multi-source data, they often exhibit limited flexibility and efficiency in the extraction and utilization of features, thereby hindering adaptive recognition and exploitation of crucial spatiotemporal features in lightning prediction.

3. Methods

To systematically address the challenge of lightning prediction, our approach combines spatiotemporal data processing with a deep learning architecture. As shown in Figure 1, from data acquisition to prediction generation, there are three consecutive stages: data preprocessing, model training, and prediction generation. This structured workflow ensures an efficient fusion of spatial and temporal features essential for accurate forecasting, while the modular design allows adaptation to different meteorological scenarios.

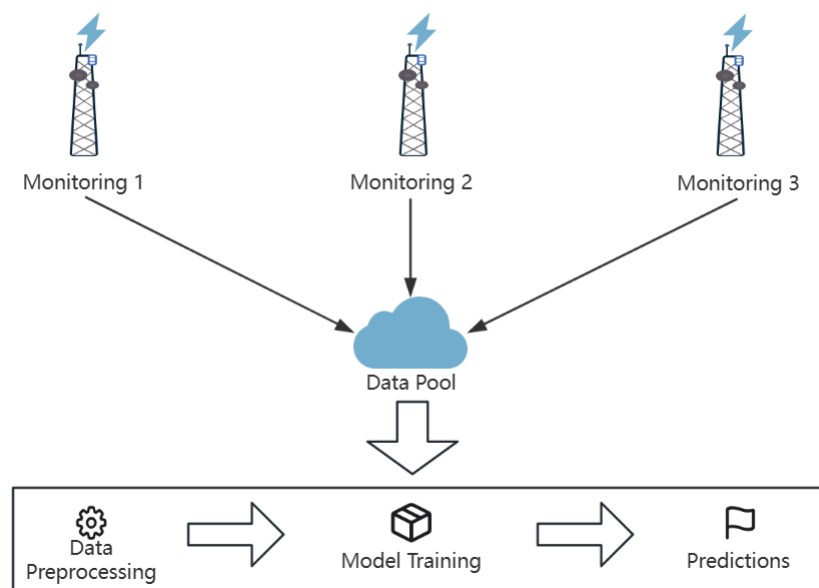


Figure 1. Method workflow diagram.

3.1. GCN Model

A GCN is a kind of convolutional neural network suitable for processing graph-structured data [28–30]. The core purpose of GCNs is to use graph convolution to extract spatial features of graph data with non-Euclidean structures. For a graph $G = (V, E, A)$, the input signal X and output signal Y , the processing f performed by the GCN network can be defined as

$$f(X, A) = Y \quad (1)$$

where V denotes the number of nodes of the graph, while E denotes the set of edges of the graph, A is the adjacency matrix of the graph, and the element A_{ij} in the matrix A denotes the connectivity between nodes V_i and V_j in the graph G . The forward propagation formula for the GCN is notated as

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^l W^l\right) \quad (2)$$

where $\tilde{A}$ is the adjacency matrix after joining the closed loop; $\tilde{D}$ is the diagonal matrix; H^l denotes the output value of first layer; X is the identity matrix, $H^0 = X$; $\sigma(\cdot)$ denotes the activation function; and W^l denotes the parameter value of the first layer.

3.2. LSTM Model

A Bi-LSTM network is a variant of RNN architecture that comprises a backward and a forward LSTM network, both connected to the same input layer [31,32]. By integrating the information from both directions, Bi-LSTM can more comprehensively explain and represent the sequence data. The structure of the Bi-LSTM network model is shown in Figure 2. The output at time t of the Bi-LSTM is calculated as follows:

$$\vec{h}_t = \overrightarrow{LSTM}\left(\vec{h}_{t-1}, X_t, \vec{C}_{t-1}\right), t \in [1, T] \quad (3)$$

$$\overleftarrow{h}_t = \overleftarrow{LSTM}\left(\overleftarrow{h}_{t-1}, X_t, \overleftarrow{C}_{t-1}\right), t \in [1, T] \quad (4)$$

$$h_t = \sigma\left(W_h \left[\vec{h}_t, \overleftarrow{h}_t\right] + b_h\right) \quad (5)$$

where $\vec{h}_{t-1}$ and $\vec{C}_{t-1}$, respectively, represent the hidden state and cell state of the input point at moment $t-1$ in the forward LSTM layer, while $\overleftarrow{h}_{t-1}$ and $\overleftarrow{C}_{t-1}$, respectively, represent the hidden state and cell state of the input point at moment t in the reverse LSTM layer. W_h and b_h are the weight matrix and bias term, respectively.

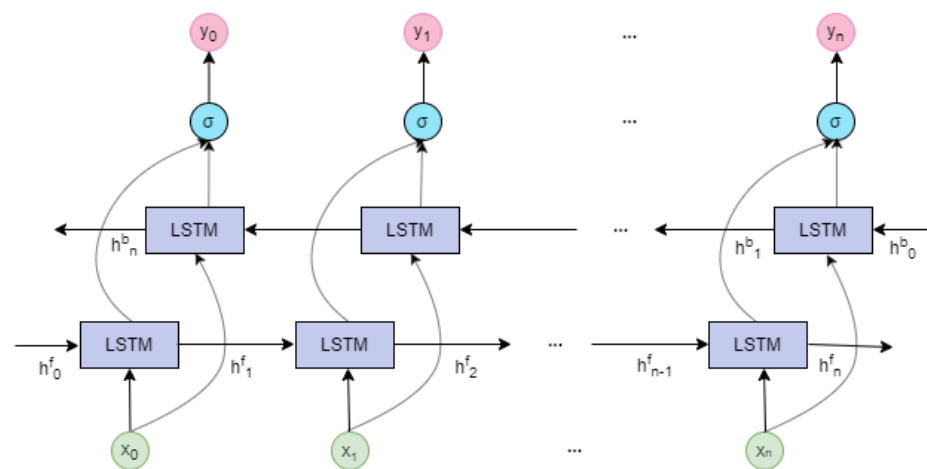


Figure 2. Bi-LSTM network structure.

3.3. Multi-Head Attention

An attention mechanism is a mechanism used for sequence modeling and feature extraction. A multi-head attention mechanism [33,34] can learn richer and more diverse data representations by integrating multiple parallel single-head self-attention modules. The structure of a multi-head attention mechanism is shown in Figure 3. The implementation steps are as follows:

$$s_i = F(Q, k_i) \quad (6)$$

$$\alpha_i = \text{softmax}(s_i) = \frac{\exp(s_i)}{\sum_{j=1}^N \exp(s_j)} \quad (7)$$

$$\text{Attention} = ((K, V), Q) = \sum_{i=1}^N \alpha_i \mu_i \quad (8)$$

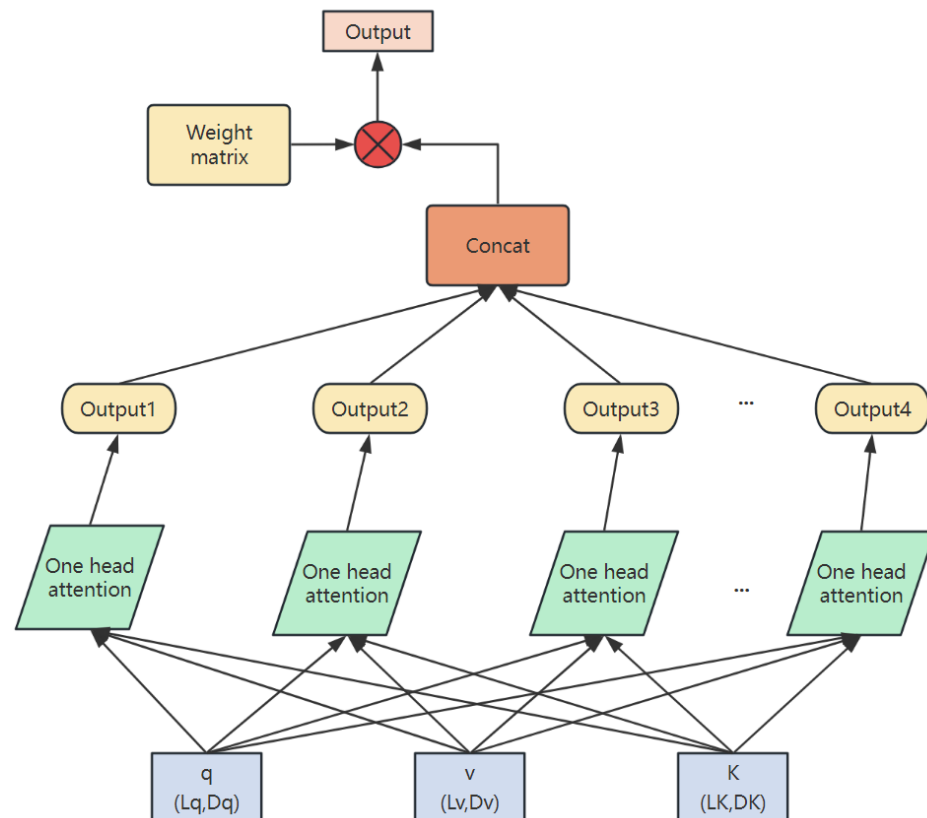


Figure 3. Multi-head attention mechanism.

First, calculate the similarity s_i between the query and the key. Then, use the softmax function to numerically transform the similarity s_i into weight coefficients α_i . Finally, use these coefficients α_i to calculate a weighted sum of the values.

3.4. GCN–LSTM–Attention

Lightning is highly localized in three-dimensional space and is produced by the accumulation of electrical charges within a cloud or between the cloud and the ground. Therefore, the intensity of the field strength varies with distance at the time of lightning generation. Constructing an adjacency matrix based on the distance between the detection sites can precisely represent the weight of the field strength influenced by the distance factor. Meanwhile, the randomness and short duration of lightning often result in high frequency during data collection. Therefore, to capture the characteristic relationships between each data element in the sequence, the LSTM network with a long-term memory function is used. In addition, due to the unsteady nature of lightning data, measures were needed to enable the model to concentrate on significant features. In this study, an attention mechanism is introduced to enhance the model's ability to perceive global features in order to predict the probability of lightning and classify its intensity levels. The model mainly includes three modules: a spatial feature extraction module, a temporal feature extraction module, and a multi-head attention module. The overall structure of the model is shown in Figure 4.

Spatial feature extraction module. The module consists of two layers of GCNConv designed to build an adjacency matrix based on the geographical location of the monitoring site, thus assigning a connection weight matrix to each data point. This distance-based adjacency matrix provides the GCN module with an intuitive representation of the spatial relationship between sites, which helps the model capture spatial features better.

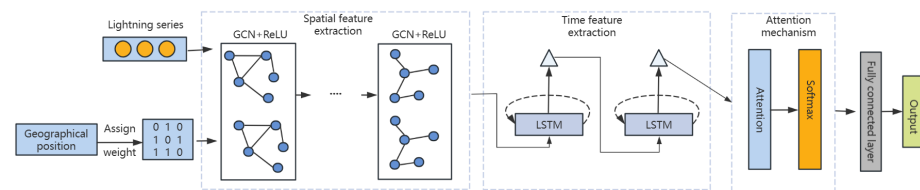


Figure 4. GCN-LSTM-Attention structure.

To construct the GCN, the adjacency matrix and data features are input into the network, and information on nodes and their adjacent nodes is processed by the two-layer GCNConv, so as to effectively extract the spatial features of the data. ReLU functions are used to introduce nonlinearity and enhance the expressiveness and generalization ability of the model. The feature extraction calculation process of the spatial feature extraction module is shown below:

$$\tilde{A} = A + I \quad (9)$$

$$\hat{A} = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} \quad (10)$$

$$D = \sum_j \tilde{A}_{ij} \quad (11)$$

$$H^{(i+1)} = \sigma(\hat{A}H^{(i)}W^{(i)}) \quad (12)$$

where d_{ij} denotes the Euclidean distance between nodes i and j ; Lat_i and Lat_j denote the latitudes of node i and node j , respectively; c and b denote the differences in longitude and latitude of node i and node j , respectively, with $c = Long_j - Long_i$ and $b = Lat_j - Lat_i$; R represents the radius of the earth; $w \in [0, 1]$ represents the weight coefficient; a_{ij} represents the influence of node i on node j 's degree of influence; $A \in R^{N \times N}$, which consists of the elements a_{ij} ; A is the original adjacency matrix, which represents the connectivity between all nodes in the graph; and I represents the unit matrix, which is used to add the closed loop in A .

Temporal feature extraction module. The module consists of two layers of Bi-LSTM. Since the output of the GCN is a 2D eigenmatrix and the LSTM requires 3D input, we use the view method in the PyTorch framework (v1.13.1) to convert the $N \times D$ 2D output of the GCN into $1 \times N \times D$ 3D data. Bi-LSTM enhances the model's understanding of time series data by considering past and future information to more fully capture the dynamic changes and long-term dependencies of the data.

Multi-head attention module. The module is utilized to handle the high non-linearity in thunderstorm data by focusing on different parts of the input sequence, thus extracting features at various levels. For each attention head, the input feature matrix $x \in R^{n \times d}$ is used. Initially, the scores are computed by taking the dot product of X and the weight vector $\omega \in R^d$ of that head: $scores = \omega \times X$. Subsequently, the softmax function is applied to normalize these scores: $weights = \text{softmax}(scores)$. Finally, the weighted sum is calculated to produce the output of the head: $output = \sum(weights \times X)$. After processing by all heads, the outputs are concatenated to form a comprehensive feature representation.

3.5. Indicators for Model Evaluation

To validate the GCN-LSTM model for lightning prediction, this study employs precision, recall, and the F1 score to assess the overall predictive performance of the model and analyzes the predictive effectiveness of the model across each category via a confusion matrix heatmap.

Precision represents the proportion of samples that the model predicts to be positive that are actually positive. The calculation formula is

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

where TP represents the number of samples that are actually positive and correctly predicted to be positive; FP represents the number of samples that are actually negative but are incorrectly predicted to be positive.

Recall represents the proportion of samples that are actually positive that the model correctly predicts to be positive. The calculation formula is

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

where FN represents the number of samples that are actually positive classes but are incorrectly predicted to be negative classes.

The $F1$ score is the harmonic average of the accuracy rate and the recall rate, designed as a balanced reflection of both accuracy and recall. The calculation formula is

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (15)$$

Confusion Matrix: The distribution between the predicted results of the model and the actual labels is displayed in the form of a matrix, helping to analyze the prediction performance of the model in different categories.

4. Case Study

4.1. Study Area

Qingdao is located on the eastern coast of Shandong Province, with geographical coordinates roughly between $119^{\circ}30'$ and $120^{\circ}40'$ east longitude and $35^{\circ}35'$ and $37^{\circ}09'$ north latitude. It is situated on the shore of Jiaozhou Bay, bordered by the Yellow Sea to the east, Jiaozhou Peninsula to the north, Weifang to the west, and Rizhao City to the south, establishing it as a crucial seaside and port city. Qingdao's geographic location is characterized by a unique maritime climate. In the summer months of July and August, due to the oceanic climate, the warm and humid airflow easily forms convective clouds when meeting the cold air on land, thus leading to the occurrence of thunder and lightning. The geography of the study area is shown as a map in Figure 5.

4.2. Data Sources

In this study, the data set used was derived from three monitoring stations affiliated with the Qingdao Meteorological Bureau in Shandong Province, China, from 8 August 2018 to 13 August 2018, with one-minute sampling intervals. The main data type is atmospheric electric field intensity, recorded in time series format, and each time point contains key variables such as electric field intensity and average electric field intensity. The data collection equipment is a high-precision atmospheric electric field instrument, which can monitor and record the changes in the atmospheric electric field in real time.

The electric field data from the monitoring stations, as shown in Figure 6, display monitoring values captured at a high sampling frequency in a single day. It is evident that lightning activity is marked by frequent alternations between positive and negative field strengths, with oscillation trends demonstrating periodicity throughout the day. This suggests that lightning activity is likely closely associated with movement of and changes in of thunderclouds, which produce significant fluctuations in the atmospheric electric

field. These fluctuations reflect the dynamic charge movements and interactions within the thunderclouds, thereby demonstrating significant instability and nonlinearity. Therefore, more accurate predictive models are required to reasonably forecast extreme thunderstorm weather and mitigate the potential negative impacts it may cause.

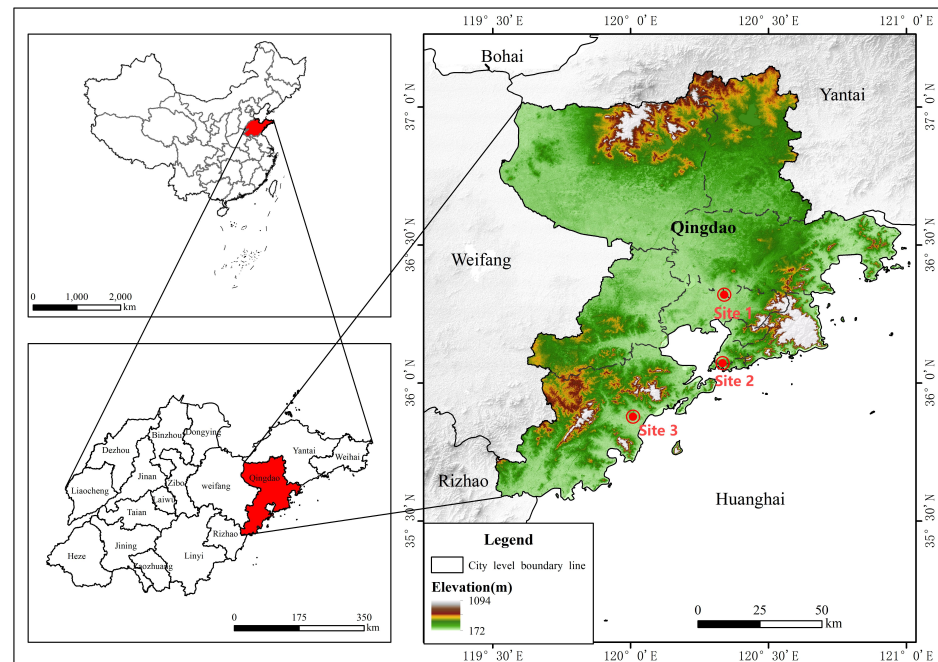


Figure 5. Geographical overview map of the study area.

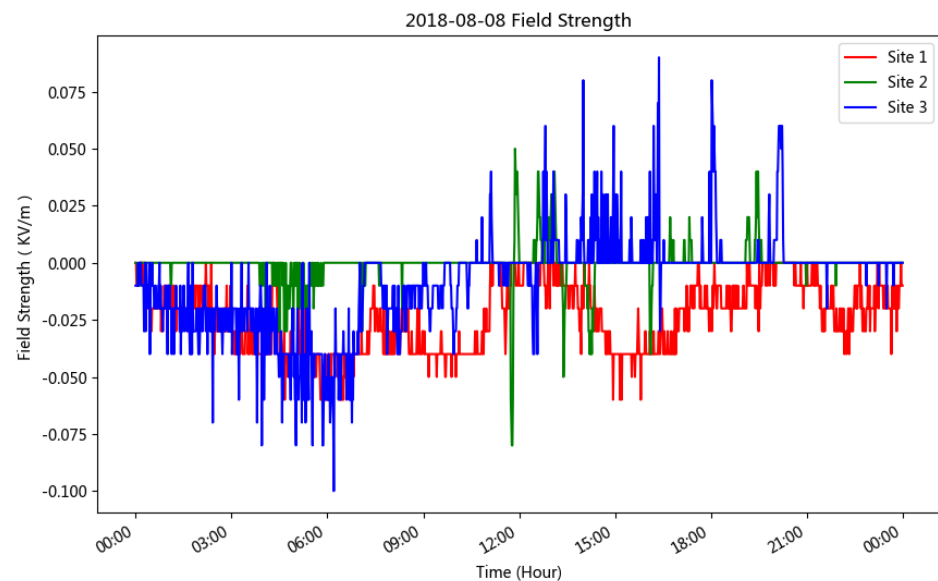


Figure 6. Field Strength on 8 August 2018.

5. Experiment

5.1. Data Processing

In this study, in order to ensure the integrity and accuracy of the data, we carefully pre-processed the original data. Firstly, we used the linear interpolation method to fill in the small number of values that were missing from the data due to equipment failure or external factors. Specifically, taking time as the horizontal axis and using several valid data points before and after the missing data points as references, each missing value was

estimated by constructing a linear equation to ensure the continuity of the data. For the detected outliers, we made corrections according to the actual distribution characteristics of the data and the actual situation of the monitoring site. For example, if the electric field strength data of a monitoring site deviates significantly from the normal range of maximum or minimum value, we made reasonable corrections to the outlier by analyzing the data from surrounding sites during the same period and combining the meteorological conditions and equipment operating status at that time, so as to ensure the authenticity and reliability of the data. In addition, in order to eliminate the dimensional differences between different eigenvalues and help the model to learn the underlying relationships in the data more efficiently, we carried out linear normalization of the data. The Min–Max scaling method was used to map each eigenvalue to the interval [0, 1].

After preprocessing, the dataset was partitioned into a training set and a test set. Specifically, 80% of the dataset was designated for the training set to ensure that the model underwent sufficient training and could accurately capture the data patterns. The remaining 20% of the dataset was used as a test set to evaluate the model's generalization ability and performance on unseen data. Through this preprocessing and data splitting strategy, we aimed to maximize the model's prediction accuracy and ensure its stability and reliability in practical applications.

5.2. Choice of Loss Function

The intense, transient, localized, and random characteristics of thunderstorms result in an extremely irregular distribution of collected data samples. The distribution of site data samples is shown in Table 1.

Table 1. Sample quantity distribution of lightning activity levels (Levels 0–3) across meteorological stations.

Site	Level 0	Level 1	Level 2	Level 3
Site 1	7757	390	343	150
Site 2	8274	327	28	11
Site 3	8258	348	27	7

The general loss function allocates equal attention to all sample categories, resulting in a disproportionate contribution of easy-to-categorize samples to the total loss, and fails to efficiently compute the minority samples. Therefore, this paper employs the Focal Loss function, which is explicitly designed to tackle class imbalance. By diminishing the influence of the majority class and augmenting the model's capacity to learn from the minority class, the Focal Loss function facilitates improved performance. The formula is as follows:

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (16)$$

Here, p_t represents the predicted probability that a sample belongs to the true class; α_t is the balancing factor, which aims to control the weight of each sample class to ensure that the class distribution becomes more balanced; and γ is the modulating factor, designed to reduce the loss of easily classified samples, thereby allowing the model to focus more on those hard-to-classify samples.

5.3. Hyperparameter Setting

Hyperparameters have a significant impact on the performance of a deep learning model. Generally, hyperparameters are not directly adjustable via conventional optimization methods such as gradient descent. Instead, they necessitate iterative experimentation and the application of expert knowledge to continually optimize and identify the optimal

combination. For the GCN-LSTM model structure proposed in this paper, we employed the generalization error derived from the test set as the evaluation metric. Through an iterative process, we meticulously fine-tuned several key hyperparameters (such as learning rate, hidden layer size, batch size, and optimizer) to find the optimal combination. The final hyperparameter settings are shown in Table 2.

Table 2. Hyperparameter settings.

Hyperparameter	Value
optimizer	Adam
Learning rate	0.001
batch_size	64
Epoch	50
Dropout	0.2
α_t	1
γ	3.8

5.4. Experimental Process

From a single lightning weather event, the formation of lightning and the movement of thunderstorm clouds make lightning activity exhibit clear spatial characteristics. Owing to the diurnal variations in lightning weather, the occurrence time, electric field strength, and average electric field strength show significant temporal characteristics. Therefore, lightning weather contains both spatial and temporal features. The prediction of lightning weather can thus be transformed into a spatiotemporal sequence prediction problem. In this study, taking Qingdao as the study area, the problem is defined as follows:

Definition 1 (Constructing the Lightning Weather Network Graph). Assume that there are N atmospheric electric field monitoring stations deployed in the city, and represent the lightning weather network graph as $G = (V, E, A)$, where $A \in R$. Each node records features such as electric field strength and average electric field strength. E represents the edge set, where an edge is established between two nodes if there is a relationship between them. $A \in R^{N \times N}$ is the weighted adjacency matrix of graph G , which represents the strength of the relationship between two nodes. In order to more intuitively understand the spatial relationship between the sites, we drew the adjacency matrix heatmap based on the distance between the sites, as shown in Figure 7. The distance between different weather stations is indicated by the depth of the color: the darker the color, the greater the distance. The calculation process for adjacency matrix construction is shown below:

$$d_{ij} = 2R \times \arcsin \sqrt{\sin^2 \frac{c}{2} + \cos(\text{Lat}_i) \cos(\text{Lat}_j) \sin^2 \frac{b}{2}} \quad (17)$$

$$a_{ij} = \begin{cases} w \times \frac{1}{d_{ij}}, & i \neq j \\ 0, & i = j \end{cases} \quad (18)$$

Definition 2 (Lightning Level Probability Prediction). Each station collects target feature data at regular intervals, considering a monitoring period containing T time intervals. The lightning data collected at all stations can be represented as a matrix $Y \in R^{N \times T}$. Due to the numerous features of the lightning data, F is used to represent the number of features collected at each station. The feature data collected at all stations form a three-dimensional dataset $X \in R^{N \times T \times F}$, where $X_{i,j,k}$ represents the value of feature k collected at time j from monitoring station i . Specifically, the model utilizes the historical data $X_t = (X_t^1, X_t^2, \dots, X_t^N) \in R^N$, $X_t = \{X_{i,j,k} | 1 \leq i \leq N, t - T + 1 \leq j \leq t, 1 \leq k \leq F\}$ collected from each station over the past

T time steps to predict the lightning level $y_{i,t+1}$ at the next time step for station i , which can be expressed as

$$y_{i,t+1} = F(X_t) \quad (19)$$

where F represents the model that predicts $y_{i,t+1}$ based on the data X_t .

Adjacency Matrix Heatmap Based on Inter-Station Distances

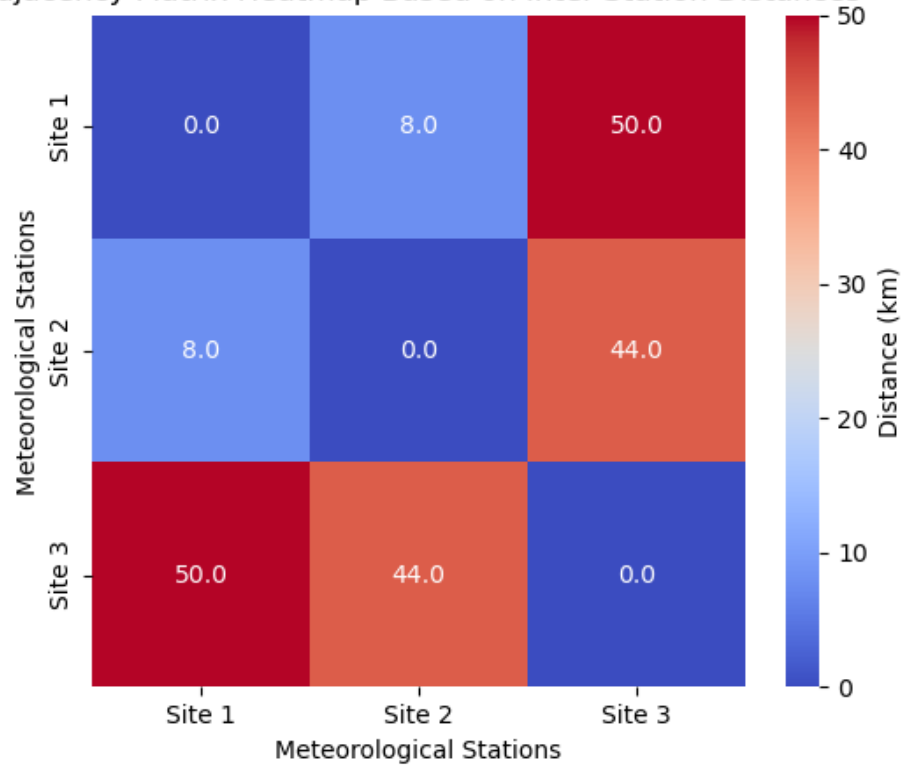


Figure 7. Adjacency matrix heatmap.

The settings of relevant parameters of the model are shown in Table 3.

Table 3. Model parameters.

Model Layer	Hidden Unit	Output Feature
GCN_1	256	256
GCN_2	512	512
LSTM_1	256	512
LSTM_2	512	1024
Attention Heads	1024	4096
FC	-	4

6. Results and Discussion

6.1. Classification Comparison

In order to display the predictive performance of each model more intuitively, this paper uses the confusion matrix heatmap for visual display, as shown in Figure 8. For the case of no lightning (Level 0), the TGCN and GCN-LSTM models had little difference in prediction effectiveness, correctly predicting 4659 and 4625 samples, respectively. Similarly, the LSTM and CNN-RNN models also showed slight differences in validity, correctly predicting 4546 and 4565 samples, respectively. For Level 1, the CNN-RNN model outperformed the other models, correctly predicting 126 samples—11, 22, and 13 more than the LSTM, TGCN, and GCN-LSTM models, respectively. For Level 2, the CNN-RNN and

GCN-LSTM models predicted eleven samples each—eight and six more than the LSTM and TGCN models, respectively. At Level 3, the GCN-LSTM model had the best predictive performance, correctly identifying four samples. For the occurrence of lightning phenomena (Levels 1–3), although the other three models (LSTM, TGCN, CNN-RNN) showed certain forecasting ability at some levels, they were ineffective at forecasting the most severe (Level 3) lightning activity. In contrast, the GCN-LSTM model had obvious advantages in dealing with complex temporal and spatial characteristics and could predict the occurrence of lightning events more accurately on the basis of comprehensive consideration of temporal and spatial characteristics.

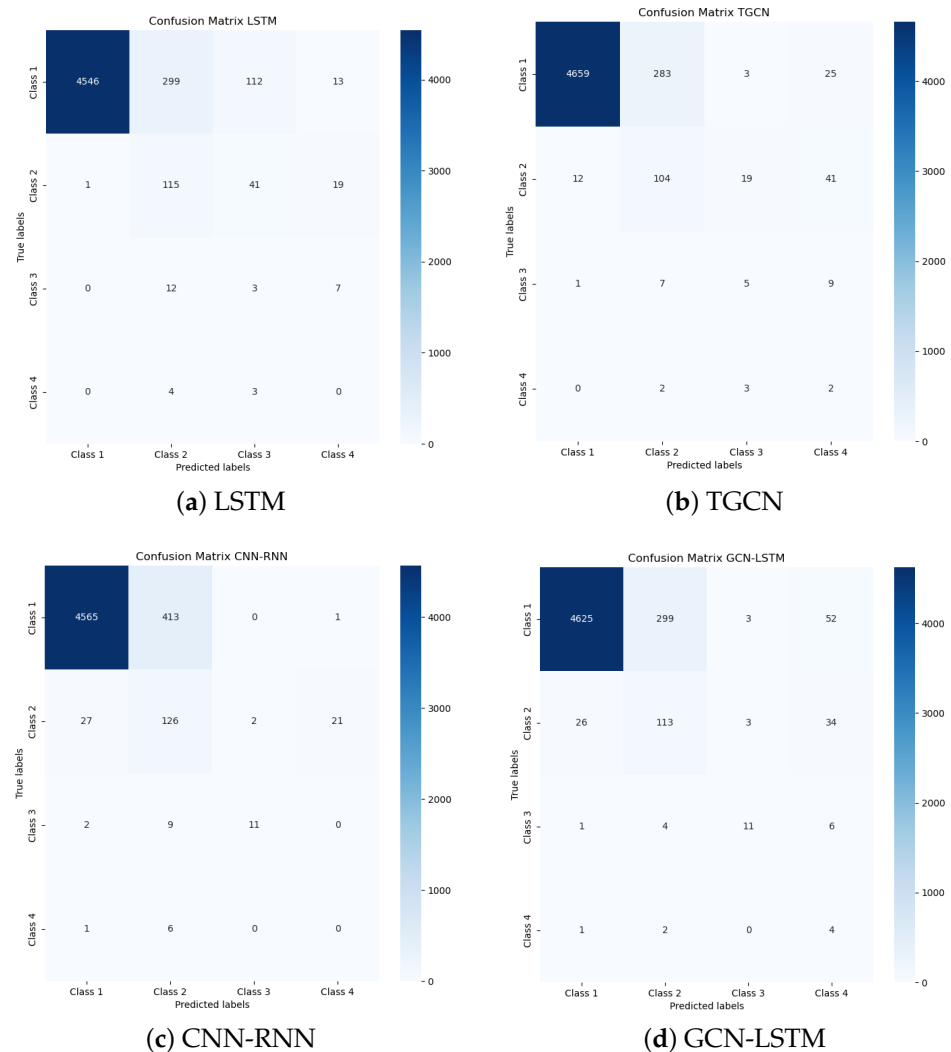


Figure 8. Confusion matrix heatmap.

6.2. Performance Comparison

In order to verify the validity and accuracy of GCN-LSTM coupling model for thunderstorm weather prediction more intuitively, two single models and one coupled model were established at the same time for experimental comparison. The specific results of model evaluation are shown in Table 4. Figure 9 shows the performance of different models as measured by four indicators—accuracy, accuracy, recall, and F1 score—in the form of radar charts. It can be seen from the table that a single model has certain limitations in predicting performance. Although LSTM models perform well in time series modeling, they are limited in capturing spatial features, resulting in poor overall predictive performance compared to coupled models. In contrast, the GCN-LSTM coupling model achieves

significantly improved prediction performance by combining the spatial feature extraction capability of GCN and the time series modeling capability of LSTM. Specifically, the GCN-LSTM model was improved in accuracy by 3% compared to the single TGCN model, 6% compared to the single LSTM model, and 2% compared to the coupled CNN-RNN model. In terms of F1 scores, the GCN-LSTM model was also improved by 19%, 2% and 3% over the single TGCN, LSTM and coupled CNN-RNN models, respectively. These results show that the GCN-LSTM model has obvious advantages in handling lightning prediction tasks. The reason is that the GCN module can effectively capture spatial correlations between meteorological stations, the LSTM module can capture long-term dependencies of time series data, and the attention mechanism can dynamically focus on key features of the data to improve the sensitivity of the model to important information. This combination enables the GCN-LSTM model to explain and predict the temporal and spatial characteristics of lightning activity more comprehensively.

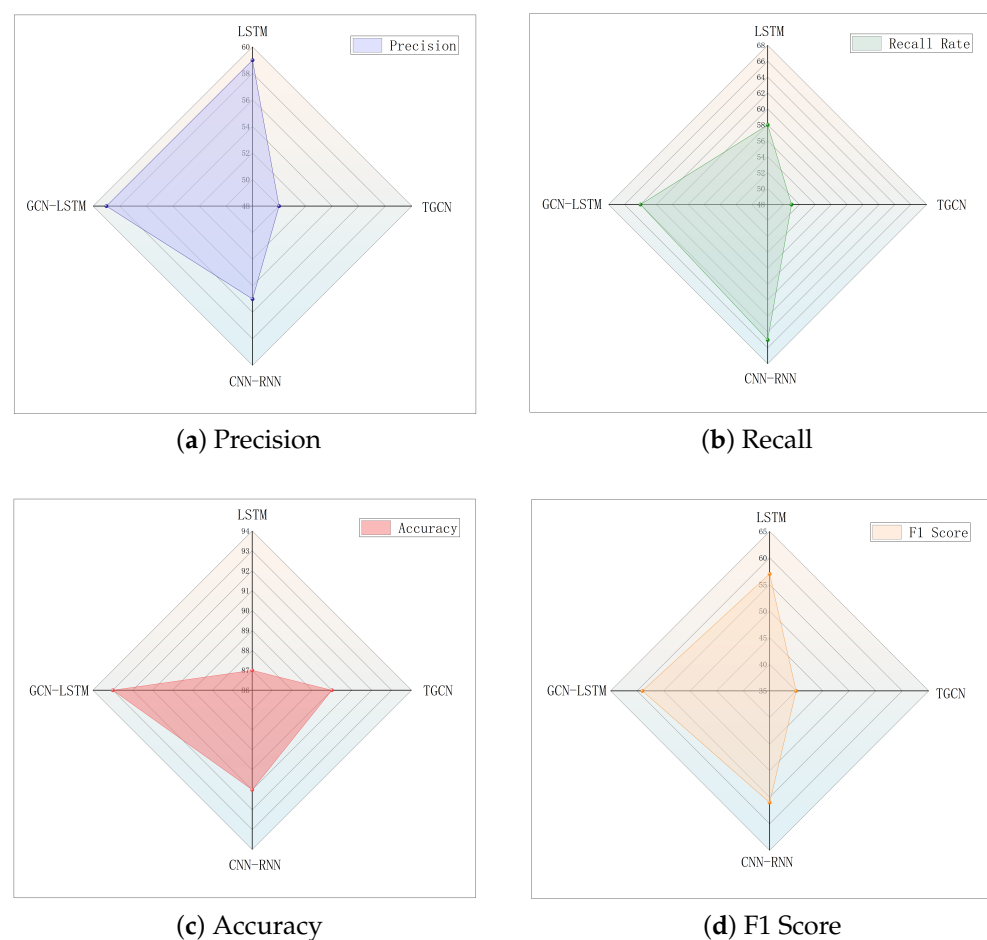


Figure 9. Comparison of performance metrics among different models. Each panel illustrates a different performance metric.

Table 4. Comparison of evaluation indexes of each model.

Model	Accuracy	Precision	Recall Rate	F1 Score
TGCN	90%	50%	51%	40%
LSTM	87%	59%	58%	57%
CNN-RNN	91%	55%	65%	56%
GCN-LSTM	93%	59%	64%	59%

In practical applications, this performance improvement is of great significance. More accurate lightning prediction can help meteorological departments to issue warnings in a more timely manner, giving the public and relevant departments more time to take protective measures, so as to effectively reduce the economic losses and casualties caused by lightning disasters. For example, predicting lightning activity in advance can guide the power department to arrange line inspection and maintenance reasonably to prevent a widespread power outage from being caused by lightning strikes. It can also help aviation departments optimize flight plans and reduce flight delays and safety incidents caused by lightning activity.

7. Conclusions

Lightning is often accompanied by heavy rain, hail, and other strong convective weather; thus, accurate prediction of lightning conditions enables timely and effective preventive measures. This study thoroughly analyzes the stochastic and transient characteristics of lightning occurrence, leveraging deep learning principles. It integrates a Graph Convolutional Network (GCN) with a Long Short-Term Memory (LSTM) network, yielding a proposed GCN-LSTM model enhanced by an attention mechanism. This model fully utilizes the spatial extraction capabilities of the GCN for graph structural data features and the temporal extraction capabilities of the LSTM network for sequential data features. It incorporates an attention mechanism that dynamically adjusts the model, effectively addressing data imbalance issues and enhancing focus on critical sample classes. The experiments demonstrate that the proposed GCN-LSTM model achieves a lightning prediction accuracy of 93%, with precision and recall rates of 59% and 64% respectively, outperforming baseline models by 2–6% in accuracy. These results prove the effectiveness and reliability of the integrated model in giving early warnings of lightning activity, providing new ideas and methods for lightning prediction. The model can help farmers and foresters take precautions against lightning-induced losses, such as protecting crops and preventing forest fires. Moreover, the model has good portability and adaptability and can be adapted and applied to other regions with different climatic conditions and lightning patterns. For future research, further exploration can be performed by combining real-time meteorological data such as weather radar and satellite data, as well as incorporating physical models or integrated learning methods, which may produce better performance.

Author Contributions: Conceptualization, W.W. under the guidance of W.Z.; methodology, W.W.; software, Windows; validation, W.W. and X.W.; formal analysis, W.W.; investigation, W.W.; resources, X.W.; writing—original draft preparation, W.W.; writing—review and editing, X.W.; visualization, W.W.; supervision, W.Z. and X.W.; project administration, W.Z.; All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Natural Science Foundation of China (61502262 and 42201506).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to acknowledge the support and encouragement of their colleagues during this study.

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

References

1. Yair, Y. Lightning hazards to human societies in a changing climate. *Environ. Res. Lett.* **2018**, *13*, 123002. [\[CrossRef\]](#)
2. Bala, K.; Choubey, D.K.; Paul, S.; Ghosh Nee, M.L. Classification Techniques for Thunderstorms and Lightning Prediction: A Survey. In *Soft Computing-Based Nonlinear Control Systems Design*; Smith, J., Ed.; IGI Global: Hershey, PA, USA, 2018; pp. 123–145. [\[CrossRef\]](#)
3. Knowles, S.; Devine, J.B. Climate change and amplified representations of natural hazards in institutional cultures. *Oxf. Res. Encycl. Environ. Sci.* **2018**, 9780190228620, 62. [\[CrossRef\]](#)
4. Allen, J.T.; Allen, E.R. A review of severe thunderstorms in Australia. *Atmos. Res.* **2016**, *178–179*, 347–366. [\[CrossRef\]](#)
5. Zhang, W.; Meng, Q.; Ma, M.; Zhang, Y. Lightning casualties and damages in China from 1997 to 2009. *Nat. Hazards* **2011**, *57*, 465–476. [\[CrossRef\]](#)
6. Yin, Q.; Liu, H.; Fan, X.; Zhang, Y.; Chen, L. Lightning fatalities in China, 2009–2018. *J. Agric. Meteorol.* **2021**, *77*, 150–159.
7. Gomes, C.; Doljinsuren, M.; Myagmar, D. Lightning incidents in Mongolia. *Geomat. Nat. Hazards Risk* **2015**, *6*, 686–701.
8. Villamil, D.E.; Rojas, H.E.; Santamaria, F.; Diaz, W. Lightning Risk and Disaster Risk Management at the Beginning of the 2020s. In Proceedings of the 2021 35th International Conference on Lightning Protection (ICLP) and XVI International Symposium on Lightning Protection (SIPDA), Colombo, Sri Lanka, 20–26 September 2021; pp. 1–5.
9. Yuan, Z. Study on the Causes of Rural Lightning Disaster and Countermeasures of Lightning Protection and Disaster Reduction. *J. Atmos. Sci. Res.* **2021**, *4*, 22–26. [\[CrossRef\]](#)
10. Hayward, L.; Whitworth, M.; Pepin, N.; Dorling, S. A comprehensive review of datasets and methodologies employed to produce thunderstorm climatologies. *Nat. Hazards Earth Syst. Sci.* **2020**, *99*, 2463–2482. [\[CrossRef\]](#)
11. Lorenc, A.C. Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.* **2010**, *112*, 474. [\[CrossRef\]](#)
12. Droegeleier, K.K. The numerical prediction of thunderstorms: challenges, potential benefits and results from real-time operational tests. *WMO Bull.* **1997**, *46*, 324–336.
13. Simon, T.; Mayr, G.J.; Umlauf, N.; Zeileis, A. NWP-based lightning prediction using flexible count data regression. *Adv. Stat. Climatol. Meteorol. Oceanogr.* **2019**, *5*, 1–16. [\[CrossRef\]](#)
14. Bala, K.; Choubey, D.K.; Paul, S. Soft computing and data mining techniques for thunderstorms and lightning prediction: A survey. In Proceedings of the 2017 International Conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 20–22 April 2017.
15. Mazany, R.A.; Businger, S.; Gutman, S.I.; Roeder, W. A lightning prediction index that utilizes GPS integrated precipitable water vapor. *Weather Forecast.* **2002**, *17*, 1034–1047. [\[CrossRef\]](#)
16. Bogdan, B.; Ustrnul, Z. Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives. *Atmosphere* **2022**, *13*, 180. [\[CrossRef\]](#)
17. Coughlan, R.; Di Giuseppe, F.; Vitolo, C.; Barnard, C.; Lopez, P.; Drusch, M. Using machine learning to predict fire-ignition occurrences from lightning forecasts. *Meteorol. Appl.* **2021**, *28*, e1973. [\[CrossRef\]](#)
18. Ren, X.; Li, X.; Ren, K.; Song, J.; Xu, Z.; Deng, K.; Wang, X. Deep learning-based weather prediction: A survey. *Big Data Res.* **2021**, *23*, 100178. [\[CrossRef\]](#)
19. Bao, R.; He, Z.; Zhang, Z. Application of lightning spatio-temporal localization method based on deep LSTM and interpolation. *Measurement* **2022**, *189*, 110549. [\[CrossRef\]](#)
20. Bao, R.; Zhang, Y.; Ma, B.J.; Zhang, Z.; He, Z. An Artificial Neural Network for Lightning Prediction Based on Atmospheric Electric Field Observations. *Remote Sens.* **2022**, *14*, 4131. [\[CrossRef\]](#)
21. Yang, X.; Xing, H.; Ji, X. Thunderstorm prediction method based on CNN-BiLSTM using BEADS. In Proceedings of the 2021 IEEE 15th International Conference on Electronic Measurement and Instruments (ICEMI), Nanjing, China, 29–31 October 2021; pp. 193–199.
22. Guo, T.; Liu, R.; Yang, H.; Shi, L.; Li, F.; Zhang, L.; Chen, Y.; Liu, Z.; Luo, F. Predict atmosphere electric field value with the LSTM neural network. In Proceedings of the 2017 International Conference on Computer Systems, Electronics and Control (ICCSEC), Dalian, China, 25–27 December 2017; pp. 263–266.
23. Lu, M.; Zhang, Y.; Chen, M.; Yu, M.; Wang, M. Monitoring Lightning Location Based on Deep Learning Combined with Multisource Spatial Data. *Remote Sens.* **2022**, *14*, 2200. [\[CrossRef\]](#)
24. Fan, L.; Zhou, C. Cloud-to-Ground and Intra-Cloud Nowcasting Lightning Using a Semantic Segmentation Deep Learning Network. *Remote Sens.* **2023**, *15*, 4981. [\[CrossRef\]](#)
25. Yang, B.; Gao, X.; Han, Y.; Zhang, Y.; Gao, T. A Thunderstorm Identification Method Combining the Area of Graupel Distribution Region and Weather Radar Reflectivity. *Earth Space Sci.* **2020**, *7*, e2019EA000733. [\[CrossRef\]](#)
26. Cintineo, J.L.; Pavolonis, M.J.; Sieglaff, J.M. ProbSevere LightningCast: A Deep-Learning Model for Satellite-Based Lightning Nowcasting. *Weather Forecast.* **2022**, *37*, 1239–1257. [\[CrossRef\]](#)
27. Zhou, K.; Zheng, Y.; Dong, W.; Wang, T. A Deep Learning Network for Cloud-to-Ground Lightning Nowcasting with Multisource Data. *J. Atmos. Ocean. Technol.* **2020**, *37*, 927–942. [\[CrossRef\]](#)

28. Cao, P.; Zhu, Z.; Wang, Z.; Zhu, Y.; Niu, Q. Applications of graph convolutional networks in computer vision. *Neural Comput. Appl.* **2022**, *34*, 13387–13405. [[CrossRef](#)]
29. Bhatti, U.A.; Tang, H.; Wu, G.; Marjan, S.; Hussain, A. Deep learning with graph convolutional networks: An overview and latest applications in computational intelligence. *Int. J. Intell. Syst.* **2023**, *1*, 8342104. [[CrossRef](#)]
30. Zhang, S.; Tong, H.; Xu, J.; Maciejewski, R. Graph convolutional networks: A comprehensive review. *Comput. Soc. Netw.* **2019**, *6*, 1–23. [[CrossRef](#)]
31. Siami-Namini, S.; Tavakoli, N.; Namin, A.S. A Comparison of ARIMA and LSTM in Forecasting Time Series. In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018.
32. Lindemann, B.; Müller, T.; Vietz, H.; Jazdi, N.; Weyrich, M. A survey on long short-term memory networks for time series prediction. *Procedia CIRP* **2021**, *99*, 650–655. [[CrossRef](#)]
33. Niu, Z.; Zhong, G.; Yu, H. A review on the attention mechanism of deep learning. *Neurocomputing* **2021**, *452*, 48–62. [[CrossRef](#)]
34. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *arXiv* **2017**, arXiv:1706.03762.

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.