electronics

Article

Ice Cover Prediction for Transmission Lines Based on Feature Extraction and an Improved Transformer Scheme

Hongchang Ke [1,†], Hongbin Sun [2,*,†], Huiling Zhao [1,†] and Tong Wu [2,†]

[1] School of Computer Technology and Engineering, Changchun Institute of Technology, Changchun 130012, China; rj_khc@ccit.edu.cn (H.K.); rj_zhl@ccit.edu.cn (H.Z.)
[2] Intelligent Distribution Network Measurement Control and Safe Operation Technology of National Development and Reform Commission, Changchun Institute of Technology, Changchun 130012, China; wutong@stu.ccit.edu.cn
* Correspondence: dq_shb@ccit.edu.cn
† These authors contributed equally to this work.

Abstract: Frequent and severe icing on transmission lines poses a serious threat to the stability and safe operation of the power system. Meteorological data, inherently stochastic and uncertain, requires effective preprocessing and feature extraction to ensure accurate and efficient prediction of transmission line icing thickness. We address this challenge by leveraging the meteorological features of icing phenomena and propose a novel feature preprocessing method that integrates Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and spectral clustering. This method effectively preprocesses raw time series data, extracts key features, and constructs a similarity matrix and feature vector. The resulting feature vector serves as a new data representation, facilitating cluster analysis to isolate meteorological and icing-related features specific to transmission lines. Subsequently, we introduce an enhanced Transformer model for predicting transmission line icing thickness. The proposed model leverages the extracted meteorological and icing features by independently embedding variable tokens for each input feature. This approach improves the model's prediction accuracy under multiple feature inputs, leading to more effective learning. The experimental results demonstrate that the performance of the proposed prediction algorithm is better than the three baseline algorithms (hybrid CEEMDAN and LSTM (CEEMDAN-LSTM), hybrid CEEMDAN, spectral clustering, and LSTM (CEEMDAN-SP-LSTM), and hybrid CEEMDAN, spectral clustering, and Transformer (CEEMDAN-SP-Transformer)) under multiple feature inputs and different parameter settings.

Keywords: ice cover prediction; feature extraction; spectral clustering; Transformer

check for updates

Citation: Ke, H.; Sun, H.; Zhao, H.; Wu, T. Ice Cover Prediction for Transmission Lines Based on Feature Extraction and an Improved Transformer Scheme. *Electronics* **2024**, *13*, 2339. https://doi.org/10.3390/electronics13122339

Academic Editor: Prasan Kumar Sahoo

Received: 24 May 2024
Revised: 9 June 2024
Accepted: 13 June 2024
Published: 14 June 2024

Copyright: © 2024 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

The increasing frequency of extreme weather events globally presents a significant challenge to power grids. Icing disasters, a common and severe threat to transmission and distribution lines, endanger the safety and stability of these critical infrastructures [1–3]. Icing can result in catastrophic consequences, including pole and tower collapse, broken wires and strands due to uneven icing or de-icing, insulator flashover, and equipment damage from line movement (dancing) [4–6]. The transmission line icing disasters caused by micro-meteorology and micro-terrain on transmission lines are as follows: severe icing on transmission lines leads to disasters such as pole collapse and tower collapse; uneven icing or de icing of the line leads to wire and strand breakage disasters, insulator string icing flashover, and equipment damage caused by the swinging of transmission lines [7,8]. Accordingly, research in this field holds profound significance for preventing and mitigating power system accidents caused by icing, ensuring reliable and safe power supply, reducing natural disaster risks, and maintaining the smooth operation of the social and economic sectors [9–11].

Recent research on transmission line icing has focused on several key areas: understanding the formation mechanisms, identifying influencing factors, developing monitoring technologies, creating prediction models, and implementing ice melting technologies. For the prediction and evaluation of icing, research explores image recognition, segmentation, degree evaluation, pattern clustering, type recognition, and prediction—primarily using machine learning methods. Zhao et al. proposed a multi-classification support vector prediction model, which was an icing thickness prediction model based on multi-classification support vector machine (SVM). The model was based on actual operating environment measured optical cable icing datasets and simulation experiments [12]. Chen et al. implemented a novel hybrid model PR-KELM, which combined principal component analysis, ReliefF, and kernel-based Extreme Learning Machine (ELM) methods to predict the degree of icing on transmission lines [13]. Wang et al. proposed a hybrid model that combines the Random Forest (RF) and Chaotic Grey Wolf Optimization Extreme Learning Machine (CGWO-ELM) algorithms for predicting short-term ice thickness [14].

Current research recognizes the critical role of preprocessing time series data due to inherent variability and uncertainty. Removing noise points and outliers demonstrably improves prediction accuracy. Feature extraction, a cornerstone of machine learning and deep learning, optimizes the learning process in two key ways. First, it reduces the risk of overfitting by lowering the data dimensionality and feature count. Second, it streamlines the learning process by eliminating irrelevant features.

Empirical mode decomposition (EMD) is an adaptive time series analysis method that excels at capturing local signal features by decomposing them into intrinsic mode functions (IMFs). EMD finds application in machine learning and deep learning prediction tasks, particularly in areas like stock market forecasting and short-term load prediction. Commonly used modal decomposition methods include EMD, Ensemble Empirical Mode Decomposition (EEMD), and Complete Ensemble EMD with Adaptive Noise (CEEMDAN). These methods share the ability to capture nonlinear data features, a challenge for traditional linear approaches. Additionally, EMD's advantage lies in its lack of reliance on pre-defined basis functions, allowing for broader applicability across various fields and data types. As a data preprocessing step, these methods pave the way for enhanced performance in subsequent machine learning or deep learning models. By combining these decomposition techniques with machine learning or deep learning models, the extracted features can serve as potent inputs, ultimately improving the model's ability to make predictions on complex data. Existing research demonstrates the successful integration of EMD, EEMD, and CEEMDAN with machine learning for time series forecasting in areas like carbon pricing and wind power prediction. Wang et al. implemented an innovative nonlinear ensemble paradigm for carbon price forecasting, integrating CEEMDAN, Sample Entropy (SE), long short-term memory (LSTM), and Random Forest (RF) to address the challenges posed by non-stationary and nonlinear carbon price series [15]. Ding et al. introduced a hybrid forecasting model that combines Complementary Ensemble Empirical Mode Decomposition (CEEMD) with the Whale Optimization Algorithm (WOA)-Kernel Extreme Learning Machine (KELM) to address the challenge of predicting short-term wind power, which is characterized by its intermittent and fluctuating nature [16]. Liao et al. constructed a hybrid long-term runoff forecasting framework that integrated EEMD with an Artificial Neural Network (ANN) and was enhanced by a multicore parallel algorithm, addressing the challenges of model construction, prediction accuracy, and computational efficiency in runoff forecasting [17]. However, there is relatively little research on prediction for icing on thickness prediction. Li et al. introduced an EEMD feature-extraction method that utilized time–frequency analysis to address the challenge of accurately predicting transmission line icing levels [18]. However, directly using the components of modal decomposition as inputs for machine learning or deep learning models was not feasible for prediction of icing cover on transmission lines.

Leveraging advancements in signal processing and data analysis techniques, spectral clustering has emerged as a promising tool for addressing these challenges [19]. This powerful

clustering method, rooted in graph theory, excels at identifying clusters within complex data structures [20–22]. By mapping data points into a lower dimensional space, spectral clustering simplifies and streamlines clustering tasks. In the context of transmission line icing prediction, spectral clustering plays a crucial role in polymerizing the IMF components into clusters. This process reduces input redundancy and mitigates model training delays, which could arise from using multiple unclustered IMFs. It is particularly valuable for identifying and analyzing the characteristics of transmission line icing signals, as the icing process itself is accompanied by complex physical changes reflected in the signal's nonlinear and non-stationary nature. There are currently some existing works on the combination of clustering and EMD for prediction, such as passive islanding detection [23] and interpretable building energy consumption forecasting [24], but they are not suitable for predicting icing on transmission lines.

In recent years, more advancements in computer hardware and software have propelled deep learning into the forefront of prediction tasks, particularly short-term forecasting. Deep learning's efficacy stems from its use of multi-layered neural networks that construct complex, end-to-end representations that map inputs directly to outputs. Convolutional neural networks (CNNs), long short-term memory (LSTM) networks, Gated Recurrent Units (GRUs), and Transformers are some of the most prevalent deep learning methods employed for prediction [25–27]. There are some related works on deep learning-based prediction on ice cover for transmission lines. Wang et al. revealed a beta variational graph attention autoencoder (β-VGATAE) for unsupervised wind turbine blade icing detection, leveraging the spatial structure of sensor data through a graph attention network to overcome the limitations of traditional supervised learning methods [28]. Han et al. introduced an ISSA-CNN-LSTM model for predicting transmission line icing thickness, utilizing an improved sparrow search algorithm, convolutional neural network, and long short-term memory to address the challenge of accurate ice thickness forecasting in power systems [29]. In terms of the damage caused by electric power line icing, Li et al. introduced a CNN and Spark large data platform-based ice data preprocessing framework and comprehensively analyzed the impact of meteorological data on ice cover thickness [30]. Wen et al. presented a Gated Temporal Convolutional Network (GTCN) enhanced by a multi-source information fusion scheme to predict transmission line icing tension, addressing the complexity introduced by various meteorological factors and line attributes [31]. A Transformer is a deep learning architecture that adopts a self-attention mechanism instead of the traditional recurrent neural network (RNN), which is suitable for various data-sequence-processing tasks and has achieved significant success in various fields, such as the Transformer Encoder-based Attention Network (TEANet) framework for stock movement prediction [32] and spatial–temporal traffic prediction [33]. However, there is almost no research on the application of Transformers for the icing on transmission lines.

In summary, this work proposes a prediction scheme for ice covering on transmission lines based on feature extraction and an improved Transformer. The designed scheme first implements the feature-extraction method based on CEEMDAN and spectral clustering and then constructs an improved Transformer-based ice-thickness-prediction model. To the best of our knowledge, this is the first prediction method that combines CEEMDAN, spectral clustering, and the Transformer model. In addition, we improved the Transformer model based on the icing features of transmission lines. The main contributions of this work are as follows:

- Firstly, we employed CEEMDAN to decompose the raw time series data, simultaneously extracting key features and reducing dimensionality through similarity and Laplacian matrices. This decomposition effectively isolates critical features within transmission line icing signals.
- Next, spectral clustering is applied to the decomposed IMFs, mapping them into a lower dimensional space. This simplifies the clustering components, improving their interpretability and suitability for machine learning and deep learning algorithms.

Notably, this work represents the first to combine CEEMDAN, spectral clustering, and their application to transmission line icing prediction.

- Furthermore, leveraging the extracted meteorological and icing features, we design a Transformer-based ice-prediction model specifically tailored to transmission lines. This model incorporates independent variable token embeddings for each input feature, enhancing prediction accuracy under multiple feature inputs and promoting more effective model learning.
- Finally, we comprehensively created experimental settings and performed prediction comparison results analysis and evaluation results analysis to verify the convergence and advantage of the proposed algorithm compared with three benchmarks: hybrid CEEMDAN and LSTM (CEEMDAN-LSTM), hybrid CEEMDAN, spectral clustering, and LSTM (CEEMDAN-SP-LSTM), and hybrid CEEMDAN, spectral clustering, and Transformer (CEEMDAN-SP-Transformer).

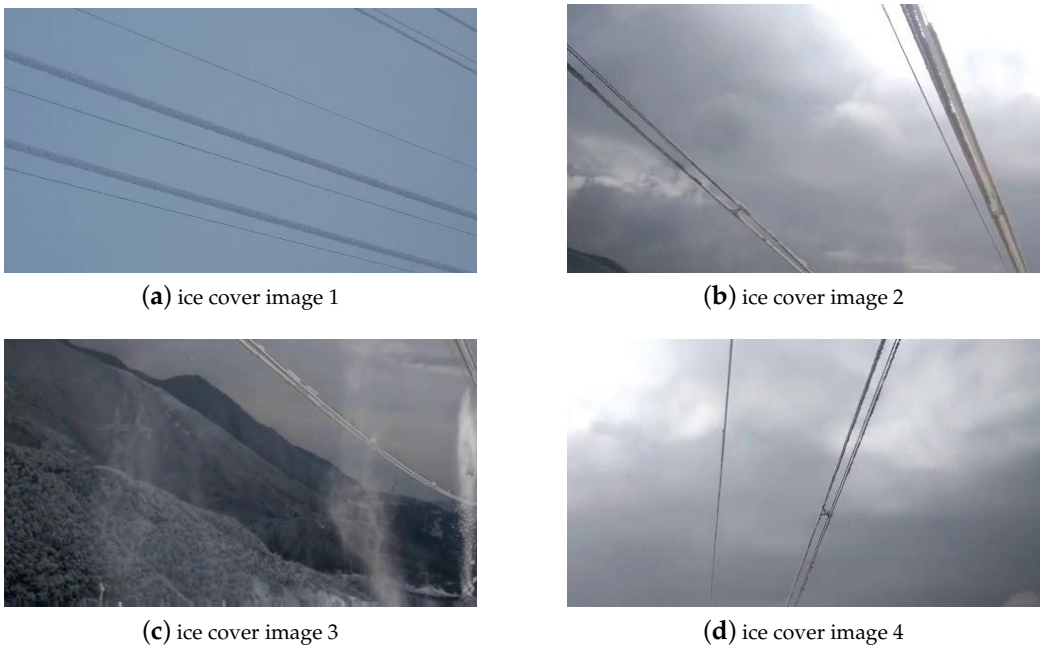
The remainder of this paper is organized as follows. Section 2 elaborates on the data preprocessing scheme based on the hybrid CEEMDAN and spectral clustering. In Section 3, we construct the prediction model based on the improved Transformer. In Section 4, the extensive experimental results are given. Finally, Section 5 concludes our paper.

2. Data Preprocessing Scheme Based on Hybrid CEEMDAN and Spectral Clustering

In this section, we describe the mode decomposition methods and spectral clustering algorithm, then propose the data preprocessing algorithm combined with CEEMDAN and spectral clustering.

2.1. Time Series Analysis of Input Data

Figure 1 shows parts of on-site images for icing on transmission lines, which usually occurs in environments with high temperatures, rain, snow, and large temperature differences.



(**a**) ice cover image 1 (**b**) ice cover image 2

(**c**) ice cover image 3 (**d**) ice cover image 4

Figure 1. Samples of ice cover on transmission lines.

In terms of the transmission lines, time series data play a critical role in analyzing and understanding icing conditions. By continuously monitoring and recording specific parameters like temperature, humidity, wind speed, and precipitation, we can gain valuable insights into the formation and melting processes of ice accumulation. This information is crucial for predicting the potential impact of icing on transmission lines and implementing preventative measures to ensure safe operation. These time series datasets are typically

collected using various sensors and monitoring equipment strategically deployed along the transmission lines. These devices capture real-time parameter readings and store the data for further processing and analysis. By analyzing these time series trends, we can not only understand the current operating state of the lines, but also identify potential issues and predict future changes. This allows for proactive measures to be taken, minimizing the risk of icing-related disruptions to power transmission. For the convenience of calculation, a time series is generally defined as follows:

$$X = \{(x_1), (x_2), ..., (x_n)\}, \tag{1}$$

Time series mining can uncover valuable information and meaningful hidden distribution patterns. A time series is a purely random sequence that evolves over time, reflecting the overall dynamic characteristics and regularity of things. The values of series are not correlated with each other. Time series are nonlinear, with mutation and uncertainty. In time series data mining, it is necessary to effectively analyze the characteristics of time series and adopt appropriate methods to process different types of time series, in order to mine more valuable information. Then, based on the vibration or electrical signals emitted by the sensors installed on the transmission line, the data are preprocessed to remove noise and normalize the signal. The research adopts a dataset that includes real-time monitoring data of transmission lines in high-altitude and cold regions, including, but not limited to parameters such as line temperature, humidity, and wind speed. These data are collected through sensor networks and obtained in real time through data interfaces. Normalize the collected data using the following normalization equation:

$$x_{norm} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \tag{2}$$

where x_{norm} stands for the normalized data series. For convenience, we define the time series of input data as $x(t)$ in time slot t.

2.2. Data Decomposition by CEEMDAN

Firstly, we used CEEMDAN to decompose the preprocessed data signal into a finite number of intrinsic mode functions (IMFs). Furthermore, we analyzed these IMFs and extracted the ones most relevant to ice cover characteristics (based on frequency content, energy, etc.). The empirical mode decomposition (EMD) algorithm is a classic modal decomposition method commonly used to process nonlinear and non-stationary time series data. The EEMD and CEEMDAN algorithms incorporate Gaussian white noise on the basis of classical modal decomposition methods to alleviate the modal aliasing problem of EMD algorithms, but there is inevitably a problem of residual white noise. Fully adaptive noise set empirical mode decomposition can solve such problems.

We added different Gaussian white noise $n^i(t)$ to the original time sequence $x(t)$. The mode decomposition by CEEMDAN is as follows: Step 1: Calculate the corresponding first-order modal decomposition signals $x^i(t) = x(t) + n^i(t), i = 1, 2, \cdots, I$

Step 2: Perform EMD on $x^i(t)$, and after decomposition, obtain the mean of $x_i(t)$ to obtain the first component $c_1(t)$:

$$c_1(t) = \frac{1}{I}\sum_{i=i}^{I} c^i(t). \tag{3}$$

Gain the first residue term $r_1(t)$:

$$r_1(t) = x(t) - c_1(t). \tag{4}$$

Step 3: Add white noise to the residue term $r_1(t)$, and perform EMD. Take the mean of the results to obtain the second component $c_2(t)$:

$$c_2(t) = \frac{1}{I}\sum_{i=1}^{I} EMD_1(r_1(t) + EMD_1(n^i(t))). \tag{5}$$

Gain the second residue term $r_2(t)$:

$$r_2(t) = x(t) - c_2(t). \tag{6}$$

Step 4: Loop step 2 and step 3 until the margin cannot be further decomposed. The implementation steps of the CEEMDAN algorithm can be summarized as follows:

$$x(t) = \frac{1}{I}\sum_{i=1}^{I} c_i(t) + r(t), \tag{7}$$

where $c_i(t)$ is the i-th component and $r(t)$ is the residue term.

2.3. Spectral Clustering

Spectral clustering is a powerful clustering method based on graph theory, particularly suitable for identifying clusters in complex data structures. It makes clustering tasks simpler and more intuitive by mapping data points to a low-dimensional space. The extracted features are used as the input for the spectral clustering algorithm, and then, the features are divided into several groups, each representing a specific ice accumulation state or condition. The spectral clustering algorithm constructs a similarity matrix between the data points and transforms it into a low-dimensional-space problem, thereby achieving effective data segmentation.

At present, spectral clustering algorithms are mostly based on the Affinity Matrix, which can be defined as follows:

$$W_{ij} = \exp\left(-\frac{d(v_i, v_j)}{2\sigma^2}\right), \tag{8}$$

where v_i, v_j represent the data samples. $d(v_i, v_j)$ stands for the distance between v_i and v_j. σ is the scale parameter, and W changes with σ. We normalized the scale matrix W_{ij} to W_{ij}^{norm}:

$$W_{i,j}^{norm} = D_{i,j}^{1/2} W_{i,j} D_{i,j}^{1/2}, \tag{9}$$

where $D_{i,j}$ is equivalent to standardizing the Laplace matrix $W_{i,j}$ once, which is defined as $D_{i,j} = \sum_{i \in I, j \in J} W_{i,j}$.

Furthermore, we calculated the minimum first k eigenvalues, then found the corresponding eigenvectors, normalized them, and obtained the final feature matrix, then performed a traditional clustering on the feature matrix:

Step 1: Construct the similarity matrix S of the sample based on the generation method of the input similarity matrix;

Step 2: Construct the adjacency matrix and degree matrix based on the similarity matrix;

Step 3: Calculate the Laplace matrix $D_{i,j}$;

Step 4: Construct a standardized Laplace matrix $W_{i,j}^{norm}$;

Step 5: Calculate the eigenvectors f corresponding to the smallest k_1 eigenvalues;

Step 6: Normalize the matrices composed of their corresponding eigenvectors f by rows, and finally, form an $n \times k_1$-dimensional feature matrix F;

Step 7: Take each row in F as a k_1-dimensional sample, consisting of n samples, and cluster them using the input clustering method with a clustering dimension of k_2;

Step 8: Obtain cluster partition $C = (c_1, c_2, \cdots, c_{k_2})$.

2.4. Data Preprocessing Scheme Based on CEEMDAN and Spectral Clustering

We usually integrate the spectral clustering algorithm and the CEEMDAN algorithm into CEEMDAN-SP for extracting icing features of transmission lines. This integration method can fully utilize the advantages of two algorithms, namely the ability of CEEMDAN to decompose complex signals into simpler components and the ability of spectral clustering to group data points (or signal components) according to similarity, thereby enhancing the detection and analysis of transmission line icing. As shown in Figure 2, the main steps of CEEMDAN-SP for transmission line icing feature extraction are as follows: The raw input data (temperature, humidity, and wind speed) are first preprocessed. Then, EMD of the data is performed based on CEEMDAN, and the noise mode function is calculated for each IMF created during EMD. Then, the similarity matrix is constructed based on the decomposed data, and a Laplace matrix is constructed. Furthermore, the eigenvalues and eigenvectors of the Laplace matrix are computed. The feature vectors are selected to retain based on the noise mode function. The feature vectors are combined into a new data representation. The new data sequences are clustered. Finally, the algorithm outputs the clustering results.

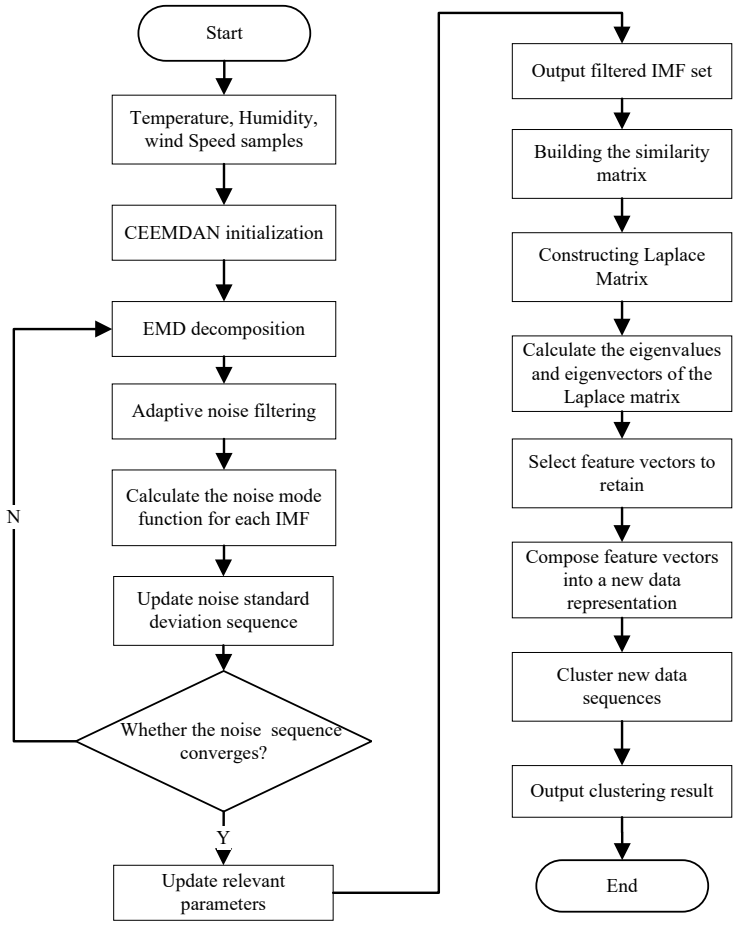


Figure 2. The data preprocessing method based on the hybrid CEEMDAN and spectral clustering.

3. Prediction Model Based on Improved Transformer

3.1. Transformer Model

A Transformer is a deep learning architecture that employs the self-attention mechanism rather than traditional recurrent neural networks (RNNs) for a wide range of sequential data processing tasks and has achieved significant success in a variety of domains. The core of the Transformer is the self-attention mechanism, which excels at taking global contextual information into account. The self-attention mechanism outperforms traditional methods in capturing the internal correlations of data or features, and it better solves the

long-range dependency problem. The Transformer consists of an encoder stack and a Decoder stack. Both the encoder and Decoder stacks are composed of N identical encoder and Decoder layers, respectively. Each encoder and Decoder layer is a tandem combination of two core modules: self-attention and feedforward network.

In terms of time series processing, in time slot t, the Transformer can be represented as:

$$X(t) = Attention(Q, K, V) + H(t-1), \tag{10}$$

where Q is the query matrix, K the key matrix, and V the value matrix, $Xt-1$ represents the output at the previous moment $(t-1)$, and $Attention$ is the attention system, which is defined as:

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}}V), \tag{11}$$

where QK^T is the dot product of the query matrix and key matrix. d_k stands for the input dimensions for elements of the queries and keys.

3.2. Improved Transformer Model

For time series prediction, the attention mechanism does not require a causality model and, thus, requires masking of the attention blocks of the traditional Transformer model. Therefore, we captured the correlation between multivariate variables using the sparse self-attention mechanism by setting the Transformer encoder to sparse self-attention to reduce the time series complexity and encode the sequence representations using a feedforward network by modeling each time series as a variable token. The specific model is shown in Figure 3: each variable of the time series is first embedded independently as a variable token; the correlation between the multiple variables is captured by the sparse self-attention mechanism, and the sequences in each block are processed independently by the shared feedforward network. This processing solves the problem of embedding multiple variables as a single time marker in the traditional Transformer architecture. This inverse operation allows the embedded markers to better capture the global features of the time series and to better utilize the correlation between multiple variables.

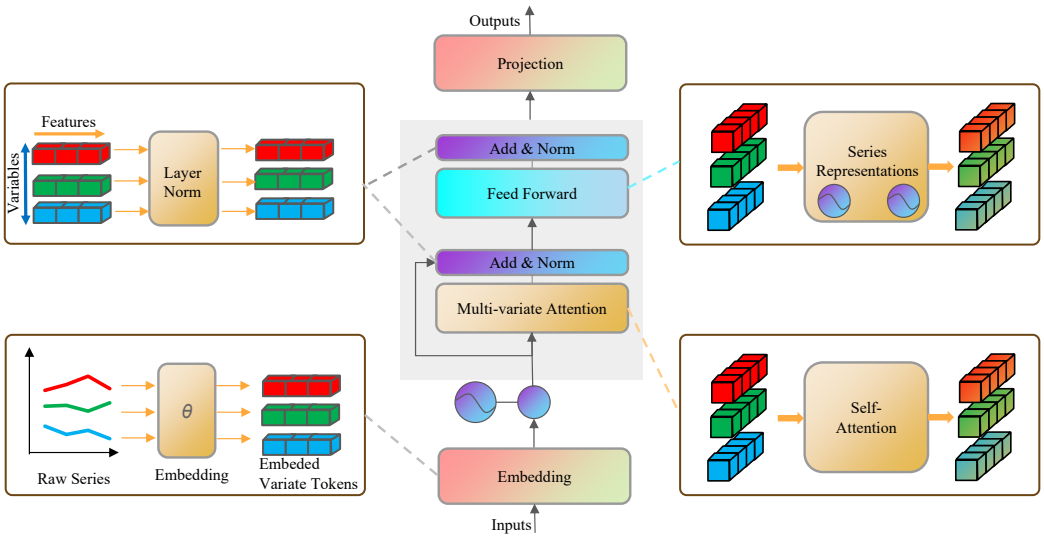


Figure 3. Structure of the proposed model.

3.3. Hybrid Feature Extraction and Transformer-Scheme-Based Ice Cover Prediction for Transmission Lines

Based on the proposed hybrid CEEMDAN, spectral clustering, and Transformer models, we designed the ice cover prediction for transmission lines based on improved feature extraction and the Transformer scheme. As shown in Figure 4, the prediction model consists of five stages. The first stage involves collecting raw data, including temperature, humidity, wind

speed, and ice thickness. In the second stage, the raw data, respectively, undergo anomaly processing to remove and correct some anomalous data. Then, in the third stage, CEEMDAN is used to decompose the data into multiple IMFs, and then, the IMFs with different features are clustered based on the spectral clustering to obtain the final feature components. In the fourth stage, we utilized the proposed Transformer-based deep learning model for multiple rounds of training and testing to learn the prediction model, and finally, in the fifth stage, we gained the prediction results of ice cover thickness for transmission lines.

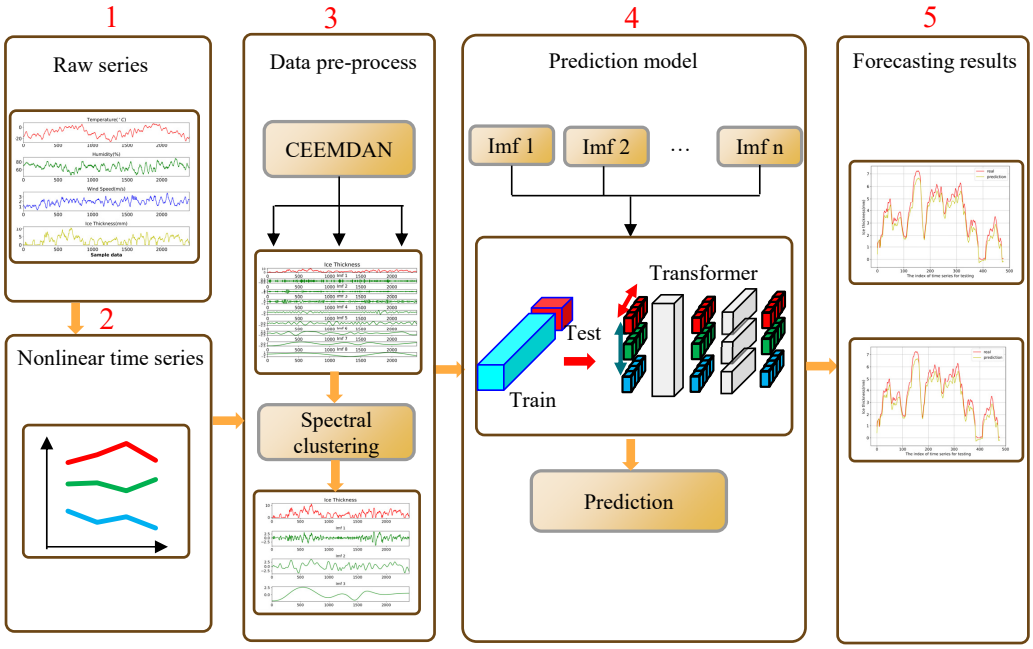


Figure 4. Structure of the ice-cover-prediction scheme.

4. Result Analysis

4.1. Experimental Data

To verify the proposed method for predicting ice thickness on transmission lines, this study selected as analysis objects two 220 kV and one 110 kV lines in a provincial power grid in northern China during the period of November to March in 2020 and 2021. The data were obtained from the ice-monitoring terminals every three hours, including environmental temperature, relative humidity, wind speed, and ice thickness data. It is worth noting that, in addition to temperature, humidity, and wind speed, wind direction is also an important factor that affects the thickness of ice cover on transmission lines. However, due to the maintenance of the data-collection device, we have not obtained complete data on the wind direction. The data depicting the relationship between the thickness of ice accumulation on the transmission line and meteorological parameters are shown in Figure 5. To train and test the proposed algorithm, we selected 2400 sample points per group as a dataset, and a total of 10 datasets were selected for cross-training and validation, then the average value was taken. We divided the dataset into the following proportions: 80% for the training set and 20% for the test set.

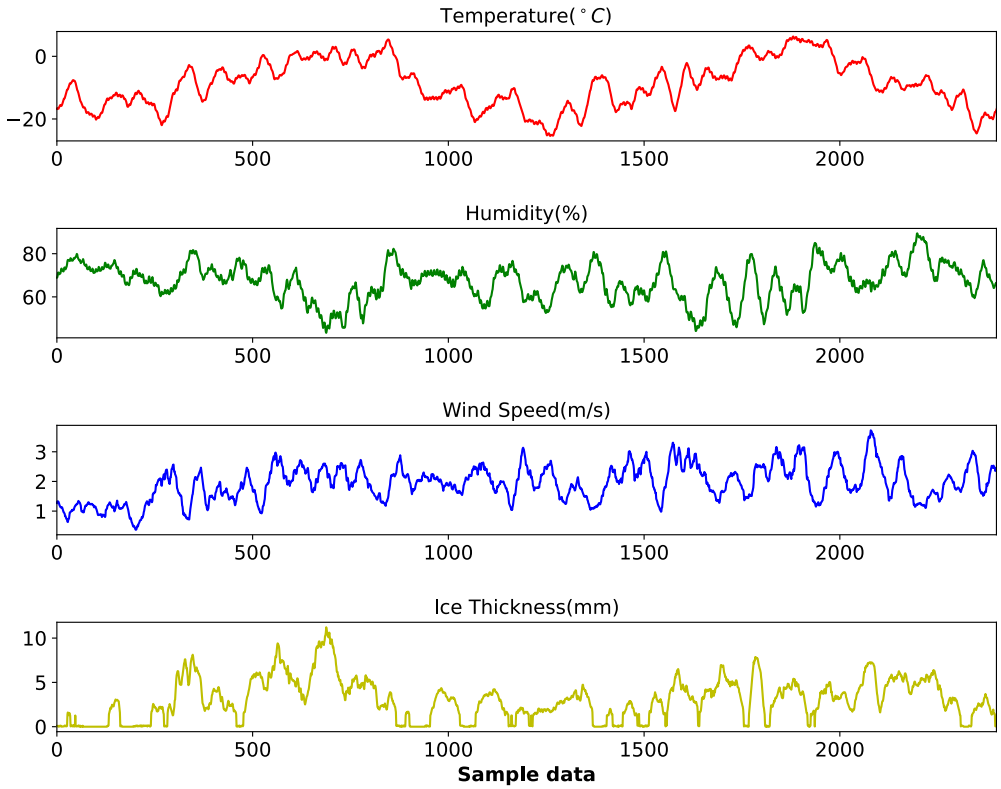


Figure 5. The data of meteorological parameters and ice thickness.

4.2. Experimental Settings

As depicted in Figure 5, temperatures, humidity, and wind speeds were all in an unstable state. We trained and predicted on 20 sets of samples for the ice-coverage-prediction models based on CEEMDAN-LSTM, CEEMDAN-SP-LSTM, CEEMDAN-SP-Transformer, and the proposed model, to determine their accuracy. A detailed description of the comparison results between the benchmark prediction algorithms will be illustrated in the subsequent paragraphs. Meanwhile, we chose the mean-squared error (MSE), root-mean-squared error (RMSE), and mean absolute error (MAE) as the evaluation metrics. Assume the predicted value and true values are $\hat{\mathbf{y}} = \{\hat{y}_1, \hat{y}_2, ..., \hat{y}_n\}$, $\mathbf{y} = \{y_1, y_2, ..., y_n\}$, respectively. Then, the metric functions are defined as:

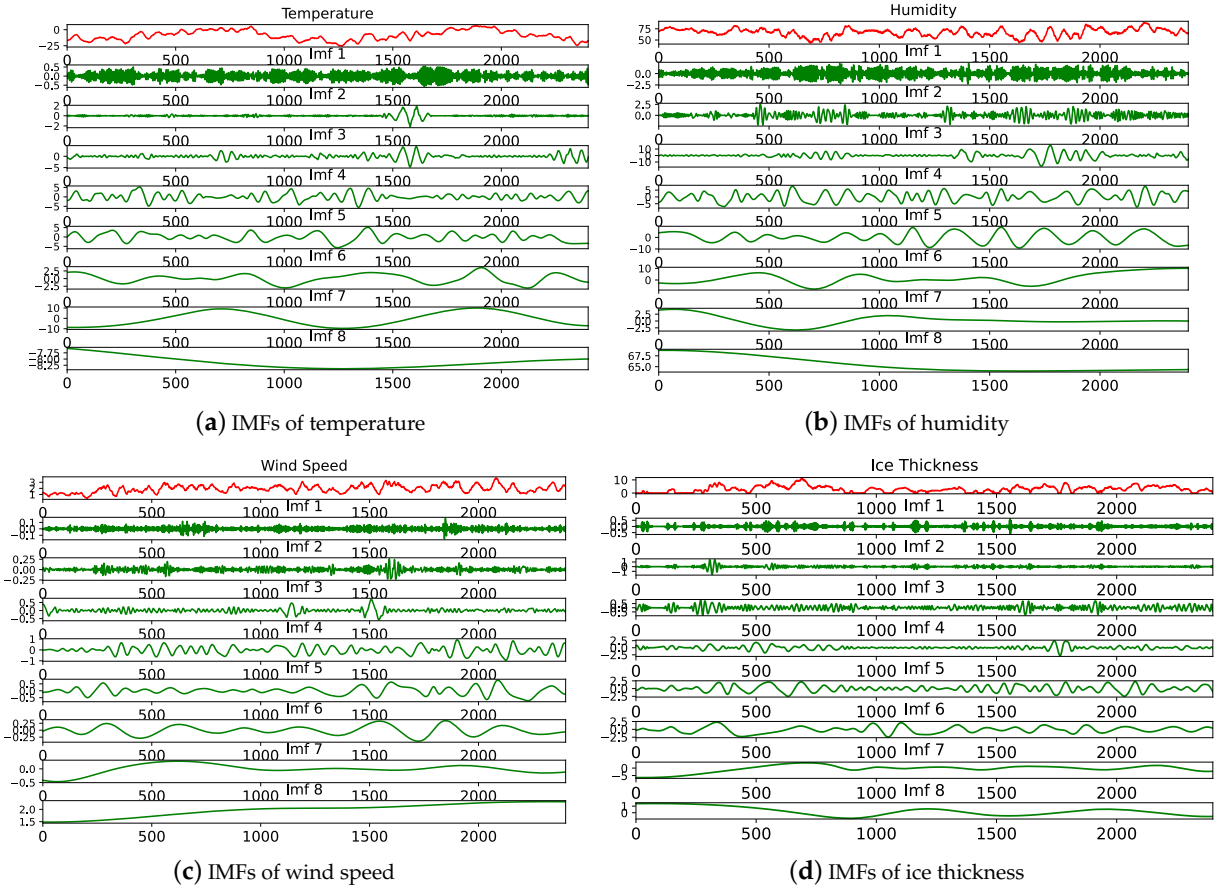
$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \tag{12}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}, \tag{13}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \mid y_i - \hat{y}_i \mid . \tag{14}$$

4.3. Data Decomposition and Clustering

Given that the original meteorological data contain exceptional noise due to numerous outliers, this study first employs normalization procedures on the raw data to enhance the model's recognition capabilities. Subsequently, CEEMDAN decomposition was conducted, resulting in the model being decomposed into different sub-time series, with a series of IMF components generated at various scales, all aimed at constructing a stable sequence. Then, noisy data were added layer by layer to enhance the generalizability of the decomposition. Figure 6 presents the results obtained after decomposing the data using CEEMDAN.



(**a**) IMFs of temperature

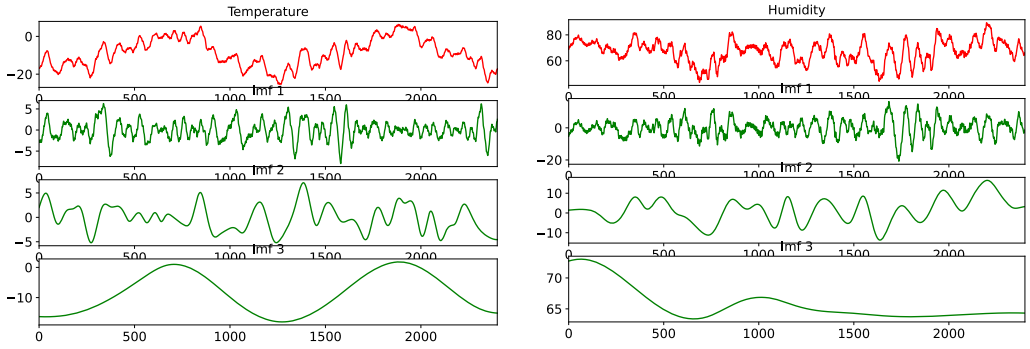
(**b**) IMFs of humidity

(**c**) IMFs of wind speed

(**d**) IMFs of ice thickness

Figure 6. IMF components of data based on CEEMDAN.

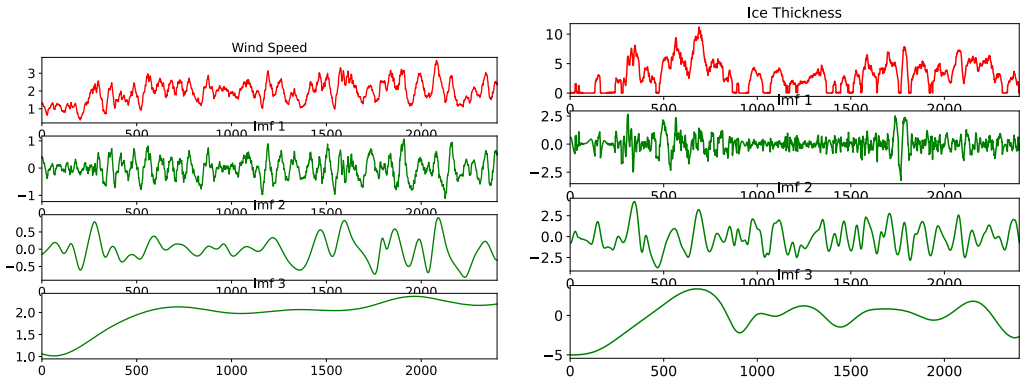
As previously stated in Section 2, spectral clustering can reduce the complexity of time series by effectively performing feature clustering. To enhance the accuracy and efficiency of prediction, cluster sub-sequences with similar eigenvalues according to spectral clustering theory. The curve depicted in Figure 7 presents the clustering results for each weather data characteristic. As shown in Figure 7a–d, it can be seen that the eight sub-sequences derived from all weather data features cluster into three clusters.



(**a**) Clusters of IMFs for temperature

(**b**) Clusters of IMFs for humidity

Figure 7. *Cont.*

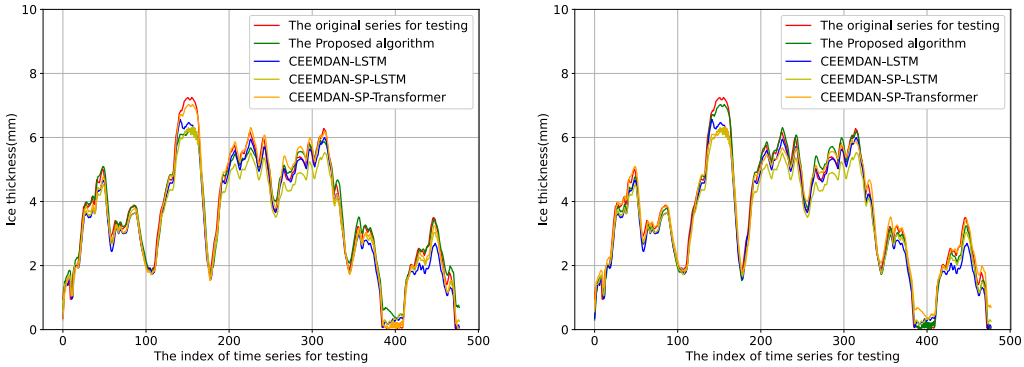


(**c**) Clusters of IMFs for wind speed (**d**) Clusters of IMFs for ice thickness

Figure 7. Clusters of IMFs based on spectral clustering.

4.4. Forecast Comparison Results of the Benchmark Algorithms

To verify the merits and advantages of the proposed algorithm, we introduced three benchmark algorithms, which are CEEMDAN-LSTM, CEEMDAN-SP-LSTM, and CEEMDAN-SP-Transformer. We present the comparison of prediction performance, convergence, and metric indication results of four algorithms under different learning rates. Figure 8 shows the comparison on different benchmark algorithms for ice thickness prediction. The red curve represents the true time series of the ice thickness.



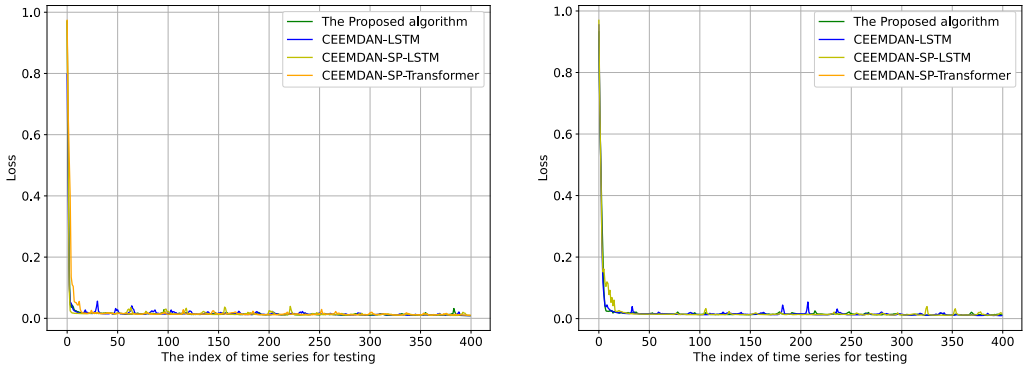
(**a**) Learning rate of $\alpha = 0.0001$ (**b**) Learning rate of $\alpha = 0.0005$

Figure 8. Comparison of different benchmark algorithms for ice thickness prediction.

As shown in Figure 8a, when the learning rate $\alpha = 0.0001$, the green curve (the proposed algorithm) illustrates the best performance, whose predicted result is closest to the true ice thickness. Although the orange curve (CEEMDAN-SP-Transformer) shows good performance, the effect is slightly inferior because it does not establish tokens based on each input time series like the proposed model. This once again confirms that the proposed model is more suitable for time series prediction. As shown in Figure 8b, when the learning rate $\alpha = 0.0005$, the performance of the proposed prediction scheme is better than the other three algorithms.

4.5. Evaluation Result Analysis of the Benchmark Algorithms

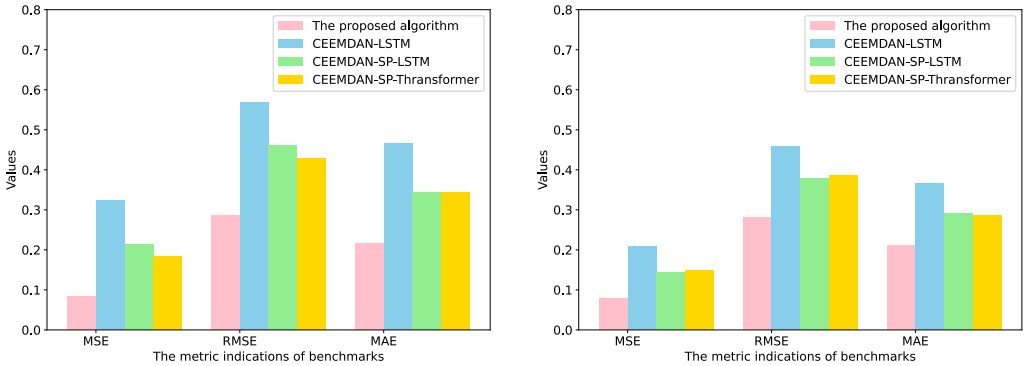
Figure 9 reveals the comparison of the different benchmark algorithms for loss. As shown in Figure 9a,b, when $\alpha = 0.0001$ and $\alpha = 0.0005$, all learning algorithms can converge, but the convergence of the proposed algorithm is the best and most stable.



(**a**) Learning rate of $\alpha = 0.0001$ (**b**) Learning rate of $\alpha = 0.0005$

Figure 9. Comparison of different benchmark algorithms for loss.

Figure 10 reveals the comparison of the different benchmark algorithms for metric indications. As described in Section 4.2, the MSE, RMSE, and MAE were selected as evaluation metrics for the predication algorithms. As shown in Figure 10a,b, when $\alpha = 0.0001$ and $\alpha = 0.0005$, regardless of the MSE, RMSE, and MAE, the performance indicators of the proposed model were superior to the other three baseline algorithms.



(**a**) Learning rate of $\alpha = 0.0001$ (**b**) Learning rate of $\alpha = 0.0005$

Figure 10. Comparison of different benchmark algorithms for metric indications.

5. Conclusions

In this paper, we propose a prediction scheme for ice covering transmission lines based on feature extraction and an improved Transformer. The proposed model can accurately predict the thickness of icing on transmission lines, which plays a crucial role in the safe and stable operation of the power grid. This model can provide technical support for intelligent detection in the power grid. The conclusions of this paper are as follows:

Firstly, the designed preprocessing method for meteorological and icing features integrates CEEMDAN and spectral clustering. The integration of the spectral clustering technique with the CEEMDAN algorithm offers a novel approach to identifying the features of icing on transmission lines. A key aspect of this hybrid scheme is the application of spectral clustering to accurately differentiate between the states of icing and non-icing on transmission lines. This distinction is vital as it aids in the timely issuance of warnings and in the implementation of preventive measures against the potential damages that can be inflicted by icing.

Secondly, the improved Transformer-based ice-thickness-prediction model for transmission lines is proposed. This designed model is suitable for time series features and provides the assurance of the accuracy of the prediction results with multiple feature inputs. In the experimental results, with inclusive parameter settings, the convergence of the proposed learning model was verified and compared with the other three baseline

algorithms. The performance of the proposed learning model was best in terms of forecast accuracy for icing thickness and metric indications for the algorithms during the integral time period in the training and test stages.

However, there are certain limitations to our work. We did not consider wind direction as a factor affecting ice cover prediction. In future work, we will collect wind direction data as part of the time series input of the proposed prediction model, which can better reflect the impact of vertical wind or win parallel to the transmission lines on ice cover thickness.

Author Contributions: H.K.: idea and physical model building, writing and modification of the manuscript. H.S.: writing—original draft, supervision, English proofing, and submission. H.Z.: experiments' execution and parts of the writing. T.W.: experiments' execution and parts of the writing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Jilin Province Scientific and Technological Planning Project of China (YDZJ202201ZYTS642, 20210101415JC).

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare that no conflicts of interest exit in the submission of this manuscript, and the manuscript has been approved by all authors for publication. We would like to declare on that the work described was original research that has not been published previously, and is not under consideration for publication elsewhere, in whole or in part. All the authors listed have approved the manuscript that is enclosed.

References

1. Qiao, X.; Zhang, Z.; Jiang, X.; Sundararajan, R.; Ma, X.; Li, X. AC failure voltage of iced and contaminated composite insulators in different natural environments. *Int. J. Electr. Power Energy Syst.* **2020**, *120*, 105993. [CrossRef]
2. Xu, F.; Li, D.; Gao, P.; Zang, W.; Duan, Z.; Ou, J. Numerical simulation of two-dimensional transmission line icing and analysis of factors that influence icing. *J. Fluids Struct.* **2023**, *118*, 103858. [CrossRef]
3. Yang, L.; Chen, J.; Hao, Y.; Li, L.; Lin, X.; Yu, L.; Li, Y.; Yuan, Z. Experimental Study on Ultrasonic Detection Method of Ice Thickness for 10 kV Overhead Transmission Lines. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–10. [CrossRef]
4. Long, X.; Gu, X.; Lu, C.; Li, Z.; Ma, Y.; Jian, Z. Prediction of the jump height of transmission lines after ice-shedding based on XGBoost and Bayesian optimization. *Cold Reg. Sci. Technol.* **2023**, *213*, 103928. [CrossRef]
5. Yang, L.; Chen, Y.; Hao, Y.; Li, L.; Li, H.; Huang, Z. Detection Method for Equivalent Ice Thickness of 500-kV Overhead Lines Based on Axial Tension Measurement and Its Application. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–11. [CrossRef]
6. Zhou, F.; Zhu, J.; An, N.; Wang, C.; Liu, J.; Long, L. The anti-icing and deicing robot system for electricity transmission line based on external excitation resonant. *IEEJ Trans. Electr. Electron. Eng.* **2020**, *15*, 593–600. [CrossRef]
7. Liu, C.; He, Q.; Lu, Y. Transmission line PSOEM-LSSVM icing prediction model. *J. Electr. Power Sci. Technol.* **2020**, *35*, 131–137.
8. Veerakumar, R.; Gao, L.; Liu, Y.; Hu, H. Dynamic ice accretion process and its effects on the aerodynamic drag characteristics of a power transmission cable model. *Cold Reg. Sci. Technol.* **2020**, *169*, 102908. [CrossRef]
9. Zhu, T.; Yuan, Y.; Xiang, H.; Liu, G.; Dai, X.; Song, L.; Liao, R. A composite pore-structured superhydrophobic aluminum surface for durable anti-icing. *J. Mater. Res. Technol.* **2023**, *27*, 8151–8163. [CrossRef]
10. Potapov, A.; Jäger, C.; Henning, T. Ice coverage of dust grains in cold astrophysical environments. *Phys. Rev. Lett.* **2020**, *124*, 221103. [CrossRef] [PubMed]
11. Han, B.; Ming, Z.; Zhao, Y.; Wen, T.; Xie, M. Comprehensive risk assessment of transmission lines affected by multi-meteorological disasters based on fuzzy analytic hierarchy process. *Int. J. Electr. Power Energy Syst.* **2021**, *133*, 107190. [CrossRef]
12. Zhao, Y.; Li, X.; Tian, R.; Feng, X.; Wu, J.; Hao, J.; Dou, W. Prediction of ice thickness of Optical Fiber Composite Overhead Ground Wire (OPGW) based on multi-class support vector machine. In Proceedings of the 14th International Photonics and Optoelectronics Meetings (POEM 2022), Wuhan, China, 18–20 December 2022; Volume 12614, pp. 67–69.
13. Chen, Y.; Fan, J.; Deng, Z.; Du, B.; Huang, X.; Gui, Q. PR-KELM: Icing level prediction for transmission lines in smart grid. *Future Gener. Comput. Syst.* **2020**, *102*, 75–83. [CrossRef]
14. Wang, W.; Zhao, D.; Fan, L.; Jia, Y. Study on icing prediction of power transmission lines based on ensemble empirical mode decomposition and feature selection optimized extreme learning machine. *Energies* **2019**, *12*, 2163. [CrossRef]
15. Wang, J.; Sun, X.; Cheng, Q.; Cui, Q. An innovative random forest-based nonlinear ensemble paradigm of improved feature extraction and deep learning for carbon price forecasting. *Sci. Total Environ.* **2021**, *762*, 143099. [CrossRef] [PubMed]
16. Ding, Y.; Chen, Z.; Zhang, H.; Wang, X.; Guo, Y. A short-term wind power prediction model based on CEEMD and WOA-KELM. *Renew. Energy* **2022**, *189*, 188–198. [CrossRef]
17. Liao, S.; Wang, H.; Liu, B.; Ma, X.; Zhou, B.; Su, H. Runoff Forecast Model Based on an EEMD-ANN and Meteorological Factors Using a Multicore Parallel Algorithm. *Water Resour. Manag.* **2023**, *37*, 1539–1555. [CrossRef]

18. Li, H.; Chen, Y.; Zhang, G.; Li, J.; Zhang, N.; Du, B.; Liu, H.; Xiong, N. Transmission line ice coating prediction model based on EEMD feature extraction. *IEEE Access* **2019**, *7*, 40695–40706. [CrossRef]
19. Langone, R.; Alzate, C.; De Ketelaere, B.; Vlasselaer, J.; Meert, W.; Suykens, J.A. LS-SVM based spectral clustering and regression for predicting maintenance of industrial machines. *Eng. Appl. Artif. Intell.* **2015**, *37*, 268–278. [CrossRef]
20. El Hajjar, S.; Dornaika, F.; Abdallah, F. One-step multi-view spectral clustering with cluster label correlation graph. *Inf. Sci.* **2022**, *592*, 97–111. [CrossRef]
21. Yang, X.; Zhu, M.; Cai, Y.; Wang, Z.; Nie, F. Fast spectral clustering with self-adapted bipartite graph learning. *Inf. Sci.* **2023**, *644*, 118810. [CrossRef]
22. Berahmand, K.; Mohammadi, M.; Faroughi, A.; Mohammadiani, R.P. A novel method of spectral clustering in attributed networks by constructing parameter-free affinity matrix. *Clust. Comput.* **2022**, *25*, 869–888. [CrossRef]
23. Thomas, S.R.; Kurupath, V.; Nair, U. A passive islanding detection method based on K-means clustering and EMD of reactive power signal. *Sustain. Energy, Grids Netw.* **2020**, *23*, 100377. [CrossRef]
24. Zheng, P.; Zhou, H.; Liu, J.; Nakanishi, Y. Interpretable building energy consumption forecasting using spectral clustering algorithm and temporal fusion transformers architecture. *Appl. Energy* **2023**, *349*, 121607. [CrossRef]
25. Shakiba, F.M.; Azizi, S.M.; Zhou, M.; Abusorrah, A. Application of machine learning methods in fault detection and classification of power transmission lines: A survey. *Artif. Intell. Rev.* **2023**, *56*, 5799–5836. [CrossRef]
26. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1–11.
27. Liu, Y.; Hu, T.; Zhang, H.; Wu, H.; Wang, S.; Ma, L.; Long, M. itransformer: Inverted transformers are effective for time series forecasting. *arXiv* **2023**, arXiv:2310.06625.
28. Wang, L.; He, Y.; Shao, K.; Xing, Z.; Zhou, Y. An Unsupervised Approach to Wind Turbine Blade Icing Detection Based on Beta Variational Graph Attention Autoencoder. *IEEE Trans. Instrum. Meas.* **2023**. [CrossRef]
29. Han, Z.; Lv, H.; Liang, Z.; Yi, J. Transmission line icing thickness prediction model based on ISSA-CNN-LSTM. *J. Phys. Conf. Ser.* **2023**, *2588*, 012020. [CrossRef]
30. Li, L.; Luo, D.; Yao, W. Analysis of transmission line icing prediction based on CNN and data mining technology. *Soft Comput.* **2022**, *26*, 7865–7870. [CrossRef]
31. Wen, Y.; Wu, J.; Huang, H.; He, J.; Liao, Y.; Li, R. Multi-source Information Fusion with Gated Temporal Convolutional Network for Transmission Line Icing Tension Prediction. In Proceedings of the 2021 16th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), Chengdu, China, 26–28 November 2021; pp. 687–691.
32. Zhang, Q.; Qin, C.; Zhang, Y.; Bao, F.; Zhang, C.; Liu, P. Transformer-based attention network for stock movement prediction. *Expert Syst. Appl.* **2022**, *202*, 117239. [CrossRef]
33. Ye, X.; Fang, S.; Sun, F.; Zhang, C.; Xiang, S. Meta graph transformer: A novel framework for spatial–temporal traffic prediction. *Neurocomputing* **2022**, *491*, 544–563. [CrossRef]

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.