

Article

Not peer-reviewed version

Ensemble Learning for Spatial Modeling of Icing Fields from Multi-Source Remote Sensing Data

[Shaohui Zhou](#) , [Zhiqiu Gao](#) ^{*} , Bo Gong , Hourong Zhang , Haipeng Zhang , Jinqiang He , Xingya Xi

Posted Date: 30 April 2025

doi: 10.20944/preprints202504.2609.v1

Keywords: icing grid; EBKI; stacking model; remote sensing data fusion; SHAP model interpretation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Ensemble Learning for Spatial Modeling of Icing Fields from Multi-Source Remote Sensing Data

Shaohui Zhou ¹, Zhiqiu Gao ^{1,*}, Bo Gong ², Hourong Zhang ², Haipeng Zhang ², Jinqiang He ² and Xingya Xi ³

¹ Collaborative Innovation Centre on Forecast and Evaluation of Meteorological Disasters, Key Laboratory for Aerosol-Cloud-Precipitation of China Meteorological Administration, School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing, 210044, China

² China Southern Power Grid Co., Ltd. Electric Power Research Institute, Guangzhou, 510700, China

³ School of Atmospheric Sciences, Sun Yat-sen University, and Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, 519082, China

* Correspondence: zgao@nuist.edu.cn

Abstract: Accurate real-time icing grid fields are critical for preventing ice-related disasters during winter and protecting property. These fields are essential both for mapping ice distribution and for predicting icing using physical models combined with numerical weather prediction systems. However, developing precise real-time icing grids is challenging due to the uneven distribution of monitoring stations, data confidentiality restrictions, and the limitations of existing interpolation methods. In this study, we propose a new approach for constructing real-time icing grid fields using 1,339 online terminal monitoring datasets provided by the China Southern Power Grid Research Institute Co., Ltd. (CSPGRI) during the winter of 2023. Our method integrates static geographic information, dynamic meteorological factors, and ice_kriging values derived from parameter-optimized Empirical Bayesian Kriging Interpolation (EBKI) to create a spatiotemporally matched, multi-source fused icing thickness grid dataset. We applied five machine learning algorithms—Random Forest, XGBoost, LightGBM, Stacking, and Convolutional Neural Network Transformers (CNNT)—and evaluated their performance using six metrics: R, RMSE, CSI, MAR, FAR, and fbias, on both validation and testing sets. The stacking model performed best, achieving an R value of 0.634 (0.893), RMSE of 3.424 mm (2.834 mm), CSI of 0.514 (0.774), MAR of 0.309 (0.091), FAR of 0.332 (0.161), and fbias of 1.034 (1.084), respectively, when comparing predicted icing values with actual measurements on pylons. Additionally, we employed the SHAP model to provide a physical interpretation of the stacking model, confirming the independence of selected features. Meteorological factors such as relative humidity (RH), 10-meter wind speed (WS_{10}), 2-meter temperature (T_2), and precipitation (PRE) demonstrated a range of positive and negative contributions consistent with the observed growth of icing. Thus, our multi-source remote sensing data fusion approach, combined with the stacking model, offers a highly accurate and interpretable solution for generating real-time icing grid fields.

Keywords: icing grid; EBKI; stacking model; remote sensing data fusion; SHAP model interpretation

1. Introduction

Global climate change has significantly increased the frequency and severity of extreme weather events, intensifying the impact of meteorological disasters on infrastructure [1,2]. Since the first recorded case of icing on power transmission lines in 1932, many regions worldwide have faced severe icing disasters. One of the most notable examples is the 1998 ice storm in Canada, which caused direct economic losses of \$1 billion and indirect losses of \$30 billion to the power system [3]. These extreme weather events increasingly threaten power systems, transportation, and agricultural production, particularly in cold regions and areas prone to winter storms.

In China, icing disasters are especially frequent in the southern regions. Since 2004, extreme cold waves have become more common, with major events in 2008 and 2016 causing significant snow and ice damage to power infrastructure, transportation, and agriculture [4]. For example, the 2008 ice storm in southern China caused severe damage to the power grid, toppling 8,709 transmission towers and breaking more than 27,000 transmission lines, leading to direct economic losses exceeding \$18.7 billion [5,6]. Southern China's unique geographical and climatic conditions, particularly in the Yunnan-Guizhou Plateau, further elevate the risk of icing disasters. Factors such as wind speed, liquid water content, and the orientation of transmission lines significantly influence the icing process [7]. As China's power grid rapidly expands, the risk of meteorological disasters affecting its infrastructure has also increased, with icing posing a major threat to the system's safe and stable operation. Therefore, it is crucial to develop accurate real-time icing grids and reliable ice thickness predictions to safeguard the grid. To monitor icing on transmission lines, the China Southern Power Grid Company (CSPGC) has installed over 3,100 icing monitoring devices across more than 1,100 lines, forming a monitoring network in regions prone to moderate or severe icing. These devices use conductor tension data to provide real-time icing values, offering essential information for disaster prevention and mitigation. However, the spatial distribution of these monitoring devices is uneven, with a higher concentration in eastern Yunnan, Guizhou, and the northern regions of Guangxi and Guangdong. As a result, the response to icing disasters across southern China is incomplete and fragmented. To address this, combining real-time icing data from monitoring devices with real-time meteorological data is essential for constructing accurate icing grid fields and generating reliable icing distribution maps [8].

Physical models, such as the Makkonen and Jones models, are commonly used in conjunction with Numerical Weather Prediction (NWP) models for icing forecasts [9–11]. However, the initial grid for icing forecasts is typically based on data from previous forecast cycles, leading to cumulative errors over time. By integrating real-time icing grid data at the start of the forecast, these inaccuracies can be significantly reduced, improving the reliability of icing predictions.

The complexity of transmission line corridors in the China Southern Power Grid region, characterized by diverse terrain and environmental conditions, poses further challenges for icing grid interpolation techniques. Although research on icing grid fields is limited due to data confidentiality, established interpolation methods used for air pollutants (such as $PM_{2.5}$ and O_3) have been adapted for icing grids. These methods include: physical interpolation techniques (e.g., inverse distance weighting, spline interpolation, ordinary Kriging, and Empirical Bayesian Kriging (EBKI)), physical models combined with interpolation (e.g., Lenhard [22], Goodwin, Makkonen [23,24], and Jones models [25]), and machine learning-based statistical models (e.g., multivariate nonlinear regression, artificial intelligence models) [12–14]. However, the 'black-box' nature of many machine learning models limits their interpretability, making it difficult to explain the influence of input features on predictions. To address this, the SHapley Additive exPlanation (SHAP) algorithm, based on game theory, has been developed. SHAP quantifies the impact of input features on model outputs, offering clearer insights into machine learning predictions [15].

Transmission line icing is influenced by various factors, including atmospheric circulation, micro-meteorological conditions, local topography, and transmission line characteristics [7]. Atmospheric circulation data, often derived from latitude-longitude grids, typically lacks the resolution needed for detailed icing grid analysis. Therefore, this study explores the creation of icing grid fields using multi-source data fusion and ensemble machine learning models. Our study contributes in the following ways: 1) we combine static geographic data, high-resolution dynamic meteorological factors, and ice_kriging values derived from parameter-optimized EBKI to spatiotemporally match real icing data from transmission line towers; 2) we apply the Bayesian Optimization Algorithm (BOA) to optimize three machine learning models—Random Forest, XGBoost, and LightGBM—then combine them using linear regression to form a stacking model. A Convolutional Neural Network Transformer (CNNT) serves as a fifth model, with all models undergoing ten-fold cross-validation; 3) we evaluate six icing error metrics to compare and identify

the best ice_kriging error correction scheme for the four southern provinces (Yunnan, Guizhou, Guangxi, Guangdong); 4) finally, we apply the SHAP model to provide a reasonable physical interpretation of the most significant feature factors in the best-performing model.

2. Materials and Methods

2.1. Geographic and Climatic Features of the YGGG Region

This study focuses on the Yunnan, Guizhou, Guangxi, and Guangdong (YGGG) region in southern China, which covers an area of approximately 975,000 km² (97.53 °E to 117.30 °E, 20.22 °N to 29.23 °N; see Figure 1). In recent years, severe low-temperature events, including rain, snow, and ice disasters, have led to frequent icing incidents on power transmission lines in this region. These incidents pose a serious threat to the stable operation of the power supply, as well as public safety and property security. Every year, from late December to early January, cold air masses from the north frequently move southward. As these air masses are blocked by the Yunnan-Guizhou Plateau and the Nanling Mountains, they contribute to the formation of the "Yunnan-Guizhou quasi-stationary front" and the "South China quasi-stationary front." These atmospheric conditions lead to significant icing on power transmission lines, particularly in areas such as Guizhou, northeastern Yunnan, northern Guangdong, and the mountainous regions of northern Guangxi, threatening the safety and reliability of the national power grid. The Yunnan-Guizhou region encompasses a variety of climate zones, and the effects of micro-topography further exacerbate local climate variability, making icing events more unpredictable and hazardous in these areas.

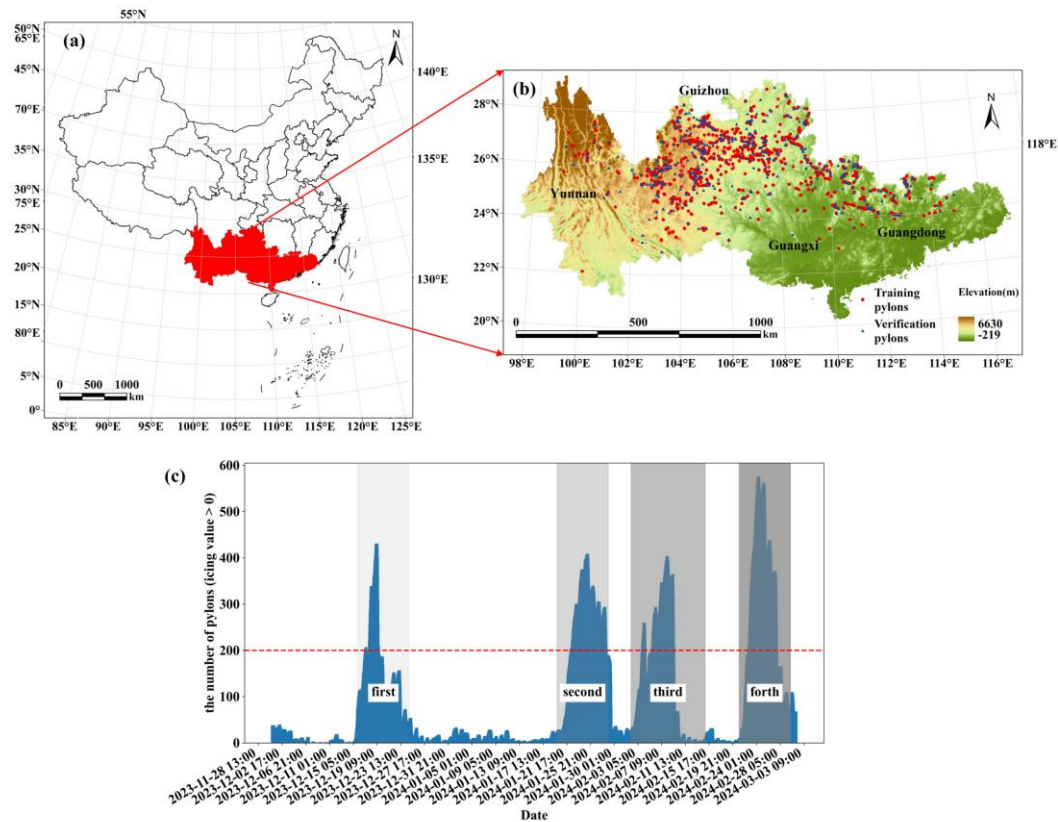


Figure 1. Study area and locations of icing monitoring pylons. (a) The locations of the four southern provinces (Yunnan, Guizhou, Guangxi, and Guangdong) in China. (b) Location and elevation of icing monitoring pylons (divided into training and verification pylons) in the four southern provinces. Changes over time in the number of pylons (icing value > 0). The gray shaded areas labeled 'first,' 'second,' 'third,' and 'forth' represent the four cold wave events.

2.2. Data Source and Data Preprocessing

2.2.1. Microtopography, Icing Grade, NDVI, LULC, and Elevation Data

The 90-meter grid Digital Elevation Model (DEM) data used in this study were sourced from the Resource and Environment Data Cloud Platform (<http://www.resdc.cn/>). Microtopography in the study area was identified and extracted using GIS analysis techniques, combined with the 90-meter DEM data [16,17]. Based on the classification by the China Southern Power Grid Research Institute Co., Ltd. (CSPGRI), nine types of microtopographies prone to icing were identified: ridge, valley, saddle, windward slope, leeward slope, moisture ascent area, ridge-moisture area, valley-moisture area, and saddle-moisture area (Figure 2b). Each of these microtopography types was assigned a numerical label, with ridge as 1, valley as 2, and so on up to 9.

CSPGRI carried out extensive surveys focusing on geomorphological features, climatic conditions, and icing monitoring. Drawing on these surveys, operational experience, and statistical analysis of icing recurrence intervals, they developed a map depicting icing severity (Figure 2a) for the YGGG region. This map highlights areas at risk for a potential 1-in-50-year icing event, which is critical for understanding long-term icing patterns. Elevation data for the mountainous regions were also extracted from the 90-meter DEM dataset. Using ArcGIS Pro 3.0, a focal statistics averaging method was applied to the DEM data within a 300×300 grid. After averaging the data three times, the final mountainous elevation data (Figure 2d) were obtained by subtracting these averaged results from the original DEM data, isolating the majority of the mountainous terrain. For vegetation analysis, the monthly Normalized Difference Vegetation Index (NDVI) data from December 2023 to February 2024 were sourced from the MODIS MOD13A3 product, with a 1-kilometer spatial resolution (Figure 2c). These NDVI datasets were accessed via the NASA Earthdata platform (<https://search.earthdata.nasa.gov/search>) on April 28, 2025.

Additionally, the land use and land cover (LULC) data were derived from the GlobeLand30 dataset, which offers a 30-meter spatial resolution (Figure 2d). This dataset classifies land cover into categories such as cultivated land, forest, grassland, shrubland, wetland, water bodies, artificial surfaces, bare land, and glaciers/permanent snow, represented by numerical codes [10, 20, 30, 40, 50, 60, 80, 90, 100], respectively. The GlobeLand30 V2020 dataset, developed by the Aerospace Information Research Institute of the Chinese Academy of Sciences, has been thoroughly validated. Using over 230,000 sample points and the Landscape Shape Index model, the dataset achieved an overall accuracy of 85.72% and a Kappa coefficient of 0.82. The dataset was downloaded from <https://www.webmap.cn/commres.do?method=globeIndex>, accessed on April 28, 2025. All the aforementioned datasets underwent quality control and normalization procedures. Their spatial resolutions were standardized to 90 meters to ensure consistency and accuracy for further analysis.

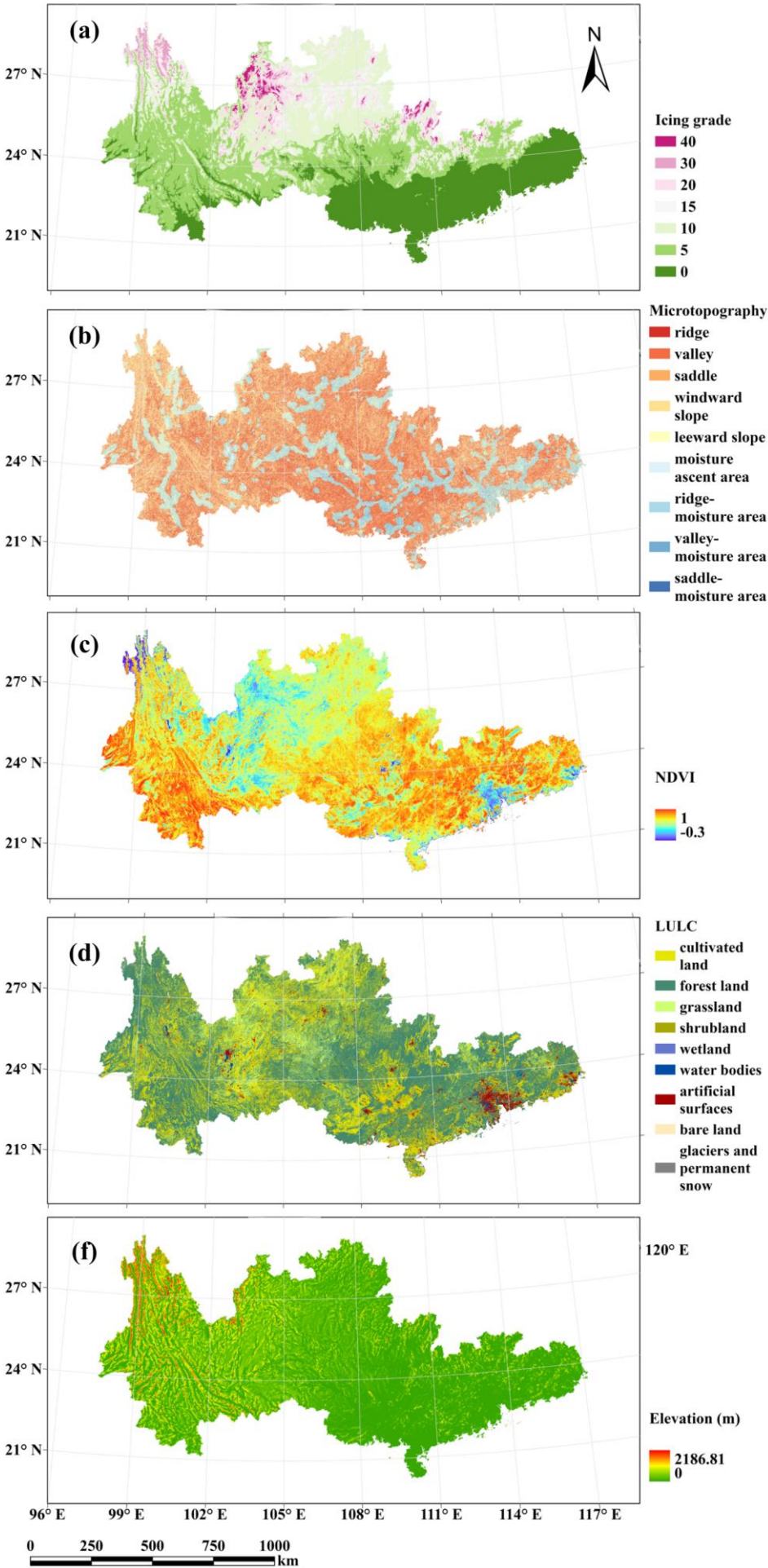


Figure 2. Icing grade (a), microtopography (b), NDVI (c), LULC (d), and mountainous elevation data (f).

2.2.2. Meteorological and Icing-Monitored Data

The observed data used in this study were obtained from the China Meteorological Administration's land data assimilation system (CLDAS-V2.0) real-time product dataset. According to the documentation available on the official website (https://data.cma.cn/data/cdcdetail/dataCode/NAFP_CLDAS2.0_RT.html), this dataset is constructed through the integration of multiple data sources, including ground-based observations and satellite data. It is further refined using advanced techniques such as multi-grid variational assimilation, physical inversion, and terrain correction. Compared to other products, this dataset offers superior quality, featuring higher spatial and temporal resolutions. The key variables used from this dataset include 2-meter air temperature, 2-meter relative humidity, 10-meter wind speed, surface pressure, and precipitation. These variables are processed by the China Meteorological Public Service Center to a latitude-longitude grid format, covering a geographical range from 15 °N to 32.97 °N and 94 °E to 120.97 °E, with a spatial resolution of 0.03 ° × 0.03 ° (3 km by 3 km) and a temporal resolution of 1 hour.

Through collaboration with the CSPGRI, we collected cumulative icing data from 1,339 icing pylons during the period from December 17, 2023, at 9:00 a.m. to February 27, 2024, at 1:00 p.m., encompassing a total of 413 time points. Each time point included icing data from more than 200 pylons (icing values greater than zero) (Figure 1c). These data were divided chronologically into two sets: the first 289 time points were used for training and validation, while the remaining 124 time points were reserved exclusively for testing. Of the initial 289 time points, 937 pylons were used for training, and 402 pylons were set aside for validation (Figure 1b). For testing, the 937 pylons were used across the final 124 time points. To interpolate the real-time icing values from the pylons, we employed the Empirical Bayesian Kriging Interpolation (EBKI) method, a widely used technique for converting point data into gridded formats. Using this method, we first interpolated the icing data onto a 540-meter by 540-meter grid, which was then resampled to a finer 90-meter by 90-meter grid. Finally, bilinear interpolation was applied to align the grid values with the pylons used in validation and testing. These kriging grid values (ice_kriging) were subsequently corrected using machine learning models to enhance accuracy.

2.2.3. Evaluation Indicators

Various commonly used evaluation metrics are applied to assess the predictive performance of models. In this study, we utilize the following key indicators: the correlation coefficient (R; Eq. (1)) and root mean square error (RMSE; Eq. (2)). Additionally, to compare the performance of different machine learning methods and multivariate nonlinear regression models for gridded interpolation of icing values, we employ several categorical performance indices: the missing alarm rate (MAR; Eq. (3)), frequency bias (fbias; Eq. (4)), false alarm rate (FAR; Eq. (5)), and critical success index (CSI; Eq. (6)). The threshold for determining the presence of icing is set at 2 mm. The formulas for calculating these error metrics are as follows:

$$R = \frac{\sum_i^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (1)$$

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

$$MAR = \frac{M}{H + M} \quad (3)$$

$$fbias = \frac{H + F}{H + M} \quad (4)$$

$$FAR = \frac{F}{H + F} \quad (5)$$

$$CSI = \frac{H}{H + M + F} \quad (6)$$

In these formulas, n represents the number of samples, $\hat{y}_i$ represents the i -th predicted value, and y_i represents the i -th actual value. For the categorical performance metrics: H (Hit) indicates an event recorded by both the icing-monitored pylons and the proposed icing gridded product; M (Miss) refers to an event identified only by the icing-monitored pylons; F (False alarm) is an event detected only by the proposed icing gridded product.

The evaluation metrics are defined as follows: Missing Alarm Rate (MAR) calculates how often the proposed product fails to estimate the icing intensity observed by the icing-monitored pylons; Frequency Bias (fbias) compares the number of icing events detected by the proposed product to those registered by the pylons. An $fbias > 1$ indicates that the icing intensity is overestimated, while $fbias < 1$ indicates underestimation; False Alarm Rate (FAR) measures the fraction of events incorrectly identified by the proposed product; Critical Success Index (CSI) combines MAR and FAR to assess the product's overall ability to accurately detect various icing intensities. The optimal values for fbias and CSI are 1, while for MAR and FAR, the ideal value is 0.

2.2.4. Proposed Icing Gridded Algorithms

This study utilized five machine learning algorithms across five experiments, with the objective of improving the accuracy of real-time icing grid fields. The overall approach begins by identifying the static and dynamic factors that influence the real-time icing grid. The real-time icing values of the pylons are then calculated and interpolated using empirical Bayesian kriging to form a 90 m by 90 m icing grid. These kriging icing values are subsequently interpolated back to the geographical coordinates of the pylons through bilinear interpolation. Next, a machine learning model is built to establish a statistical relationship between the static and dynamic factors and the real-time icing values, with the overarching goal of correcting the kriging icing values to produce more accurate real-time icing grids. The technical workflow, as illustrated in Figure 3, can be divided into three key steps:

The first step involves data fusion, cleaning, and standardization. A total of six static factors were selected: micro-topography data, the 2022 ice distribution grade map for a 50-year return period, monthly Normalized Difference Vegetation Index (NDVI) data, land use/land cover data (LULC), 90 m grid terrain elevation data (DEM), and mountainous elevation data (DEM_moun). In addition, eight dynamic factors were identified: 10-meter wind speed (WS_{10}), latitude, longitude, surface pressure (PRS), relative humidity (RH), 2-meter temperature (T_2), hourly precipitation (PRE), and kriging icing values (ice_kriging). All 14 factors were standardized before being input into the machine learning models.

The second step focuses on optimizing the machine learning models using the Bayesian Optimization Algorithm (BOA) and performing cross-validation. The dataset consists of cumulative icing values from 1,339 pylons, along with meteorological and geographical data for 413 time points collected between December 17, 2023, and February 27, 2024. The dataset was split chronologically: 70 % of the time points (289) were used for training and validation, and the remaining 30 % (124) were used for testing. Of the training/validation data, 70 % of the pylons (937) were randomly selected for training, while the remaining 30 % (402) were used for validation. This division allows for an assessment of the model's ability to correct kriging icing values on unknown pylons at the same time point and on the same pylons at unknown time points. The dimensions of the training, validation, and testing sets were $268,192 \times 14$, $114,733 \times 14$, and $115,082 \times 14$, respectively.

Machine learning algorithms such as Random Forest (RF), XGBoost, LightGBM, stacking models, and Convolutional Neural Network Transformers (CNNT) have been widely used in various fields. For example, RF, XGBoost, and LightGBM have been applied to correct 10-meter wind speed forecasts from the WRF model in southern China [18], while stacking models have been used to analyze the influence of meteorological factors on ozone formation in Hangzhou [15]. CNNT models have also been employed for denoising fluorescence microscopy images [19]. In this study, these machine learning algorithms are primarily used to correct kriging icing values, aiming to produce a more accurate and realistic icing grid. The BOA algorithm is employed to fine-tune the hyperparameters of all models except for CNNT, resulting in optimal model configurations.

The third and final step is model evaluation and error analysis. The trained machine learning models were applied to both the validation and testing sets to correct the kriging icing values, and their performance was evaluated using standard error metrics. The primary objective is to identify the best model for correcting icing grid values on unknown pylons at the same time point, as well as for the same pylons at unknown time points. Although the available data covers a limited time period, the 7:3 split for time and pylons within the same year provides a relatively robust framework for developing accurate icing grid correction models.

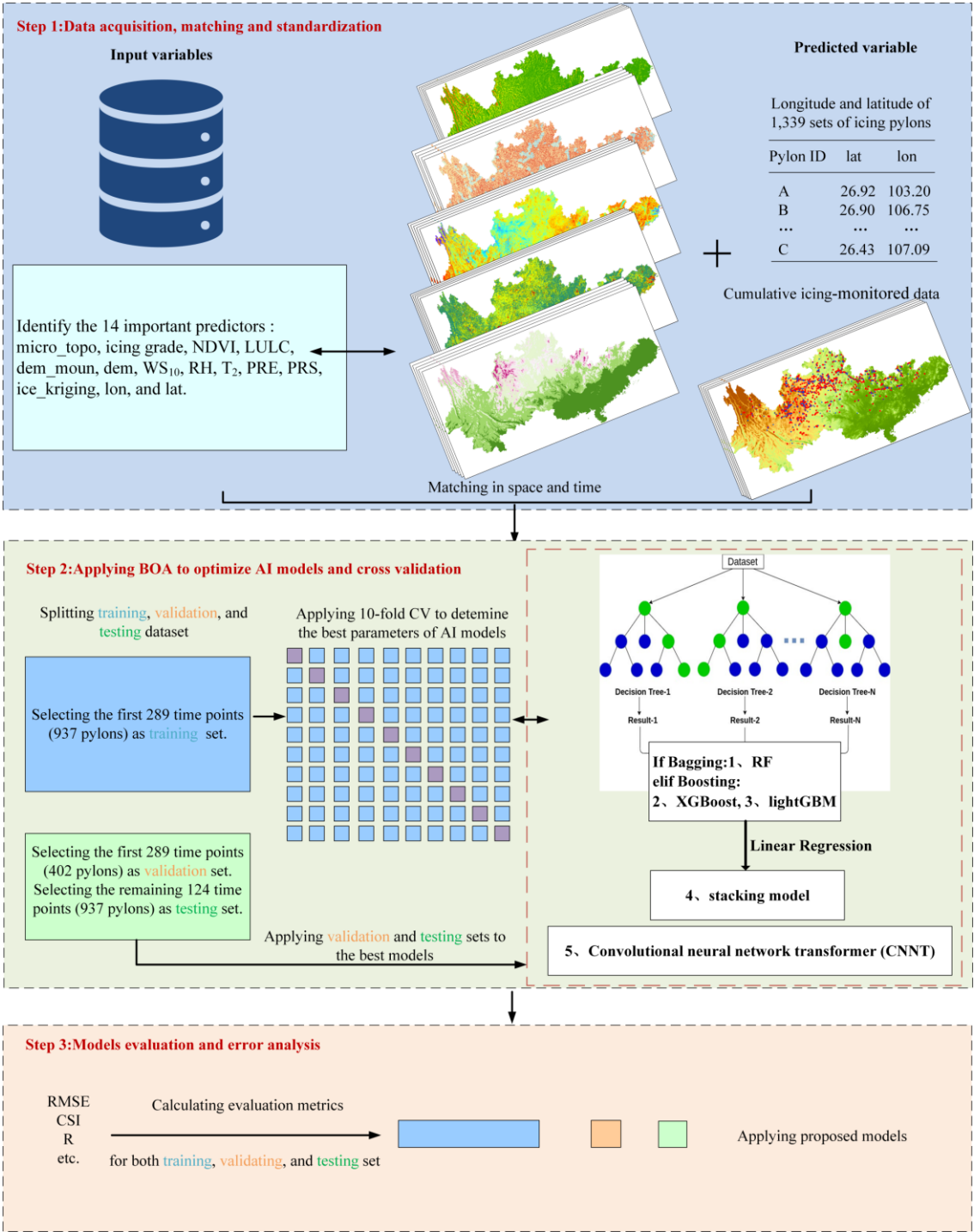


Figure 3. Flowchart of correcting kriging icing values using five machine learning models.

3. Results

3.1. Optimal Parameters Selection for EBKI

EBKI relies on three main search parameters: Radius, Maximum Neighboring Features (NBR_MAX), and Minimum Neighboring Features (NBR_MIN). In this study, we conducted grid-based optimization for these parameters, varying the radius between 1 and 5, NBR_MAX between 1 and 15, and NBR_MIN between 1 and NBR_MAX. To evaluate the model's performance, we selected a severe ice accretion event on February 24, 2024, at 09:00, during which more than 500 pylons experienced ice accretion greater than 0 mm. We recorded the real-time ice accretion values for 937 pylons used in training, processed them using EBKI, and bilinearly interpolated them to the

coordinates of 402 validation pylons to generate the corresponding ice_kriging values. These interpolated values were compared to the actual monitored ice accretion values of the validation pylons, and the model's accuracy was assessed using two key metrics: RMSE and CSI.

To determine the optimal parameter settings, we normalized the RMSE values from all experiments to a [0,1] range, and their reciprocals were calculated to align with the interpretation of performance (lower RMSE is better). Similarly, the CSI values were normalized to the same range, where higher values indicate better performance. The final evaluation score was computed as a weighted sum of the normalized RMSE and CSI values, with each metric assigned a weight of 0.5. This approach ensured that optimal parameter settings simultaneously minimized RMSE while maximizing CSI. The results of the experiment showed that varying the radius had little impact on the evaluation scores for the other parameters. However, as shown in Figure 4, the CSI and RMSE values generally improved (i.e., CSI increased and RMSE decreased) as the NBR_MAX and NBR_MIN values increased. For instance, with a radius of 5, the optimal performance was achieved when NBR_MAX was set to 5 and NBR_MIN to 1. Under these conditions, the highest evaluation score was 0.752, with corresponding RMSE and CSI values of 3.90 mm and 0.67, respectively.

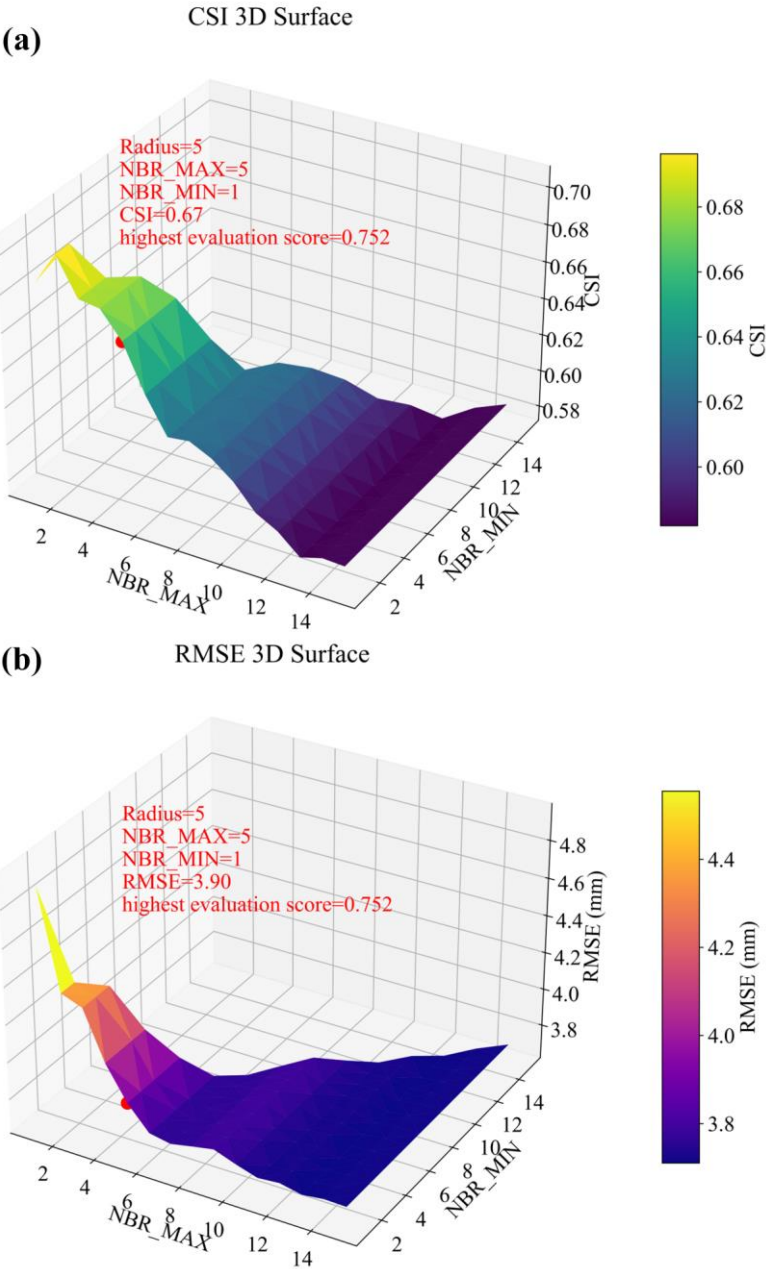


Figure 4. 3D variation of CSI (a) and RMSE (b) values with NBR_MAX, NBR_MIN, and Radius set to 5.

3.2. Evaluation of the Proposed Models on the Training Set, Validation Set, and Testing Set

After parameter optimization, the EBKI model provides relatively accurate predictions of icing gridded values across the training, validation, and testing sets. Tables 1 and 2 summarize the six error evaluation metrics (R, RMSE, CSI, MAR, FAR, and fbias) for EBKI and the five machine learning models applied to these datasets. The results indicate that the stacking model outperforms the others in both the validation and testing sets, showing superior performance in all error evaluation metrics. For the stacking model, the metrics on the validation set (and testing set) are as follows: R: 0.634 (0.893), RMSE: 3.424 mm (2.834 mm), CSI: 0.514 (0.774), MAR: 0.309 (0.091), FAR: 0.332 (0.161), and fbias: 1.034 (1.084).

In contrast, the Convolutional Neural Network Transformer (CNNT) model shows signs of overfitting. Although adjustments were made, the CNNT model performed worse than EBKI in the validation set for predicting icing gridded values (ice_kriging). While R and CSI slightly improve in the testing set, increasing from 0.851 to 0.859 and from 0.668 to 0.713, respectively, the RMSE also rises from 3.315 mm to 3.469 mm, indicating poorer overall performance. Among the models, CNNT performs the worst, whereas the other four machine learning models display reduced errors and improved CSI scores compared to EBKI.

The improvements in error evaluation metrics on the testing set compared to the validation set reveal important insights. Regardless of whether the common EBKI interpolation method or machine learning approaches are used, interpolating icing values from known pylons to unknown pylons (validation set) is particularly challenging, as evidenced by a maximum CSI score of 0.514. However, interpolating icing values from the same batch of training pylons to unknown time points (testing set) results in a much higher maximum CSI score of 0.774. This suggests that predicting icing values for unknown pylons is more difficult than for unknown time points (Figure 5).

The Taylor distribution map in Figure 6 further illustrates the models' performance. The CNNT model performs the worst in the validation set, while the stacking model consistently performs the best across both validation and testing sets, as indicated by the blue star. Not only does the stacking model achieve the highest R and RMSE performance, but its standard deviation is also closer to the observed values, represented by the black dashed line.

Table 1. Evaluation table of R, RMSE, and CSI for EBKI and five artificial intelligence models in icing gridded interpolation for training set, validation set, and testing set.

Model	training set			validation set			testing set		
	R	RMSE (mm)	CSI	R	RMSE (mm)	CSI	R	RMSE (mm)	CSI
EBKI	0.834	2.224	0.580	0.574	3.602	0.477	0.851	3.315	0.668
lightGBM	0.984	0.708	0.903	0.630	3.435	0.504	0.883	2.949	0.774
XGBoost	0.986	0.659	0.920	0.618	3.477	0.501	0.887	3.016	0.758
RF	0.988	0.609	0.969	0.630	3.417	0.511	0.883	2.951	0.767
stacking	0.987	0.661	0.933	0.634	3.424	0.514	0.893	2.834	0.774
CNNT	0.965	1.069	0.821	0.559	3.775	0.453	0.859	3.469	0.713

Table 2. Evaluation table of MAR, FAR, and fbias for EBKI and five artificial intelligence models in icing gridded interpolation for training set, validation set, and testing set.

Model	training set	validation set	testing set
-------	--------------	----------------	-------------

	MAR	FAR	fbias	MAR	FAR	fbias	MAR	FAR	fbias
EBKI	0.205	0.318	1.166	0.295	0.404	1.183	0.080	0.291	1.297
lightGBM	0.074	0.026	0.951	0.324	0.336	1.017	0.096	0.157	1.072
XGBoost	0.059	0.023	0.963	0.328	0.337	1.014	0.107	0.166	1.07
RF	0.020	0.012	0.993	0.313	0.334	1.032	0.094	0.166	1.086
stacking	0.048	0.022	0.974	0.309	0.332	1.034	0.091	0.161	1.084
CNNT	0.128	0.067	0.934	0.405	0.346	0.910	0.170	0.165	0.994

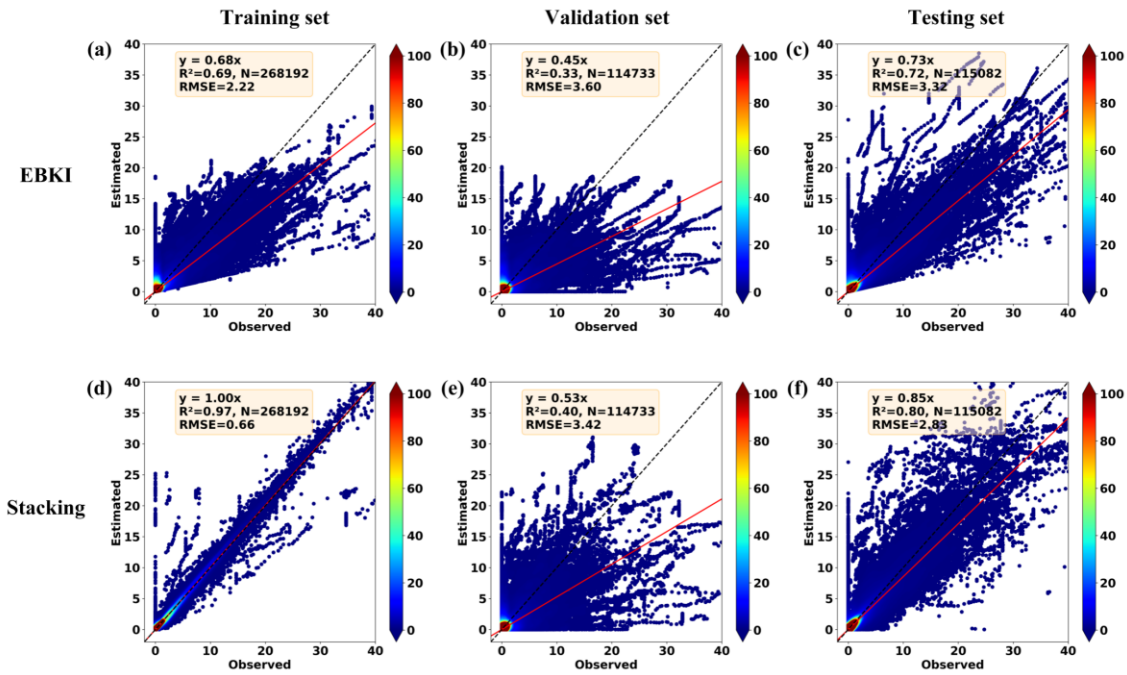


Figure 5. The scatter density map of EBKI and the stacking model for the training set (a, d), validation set (b, e), and testing set (c, f).

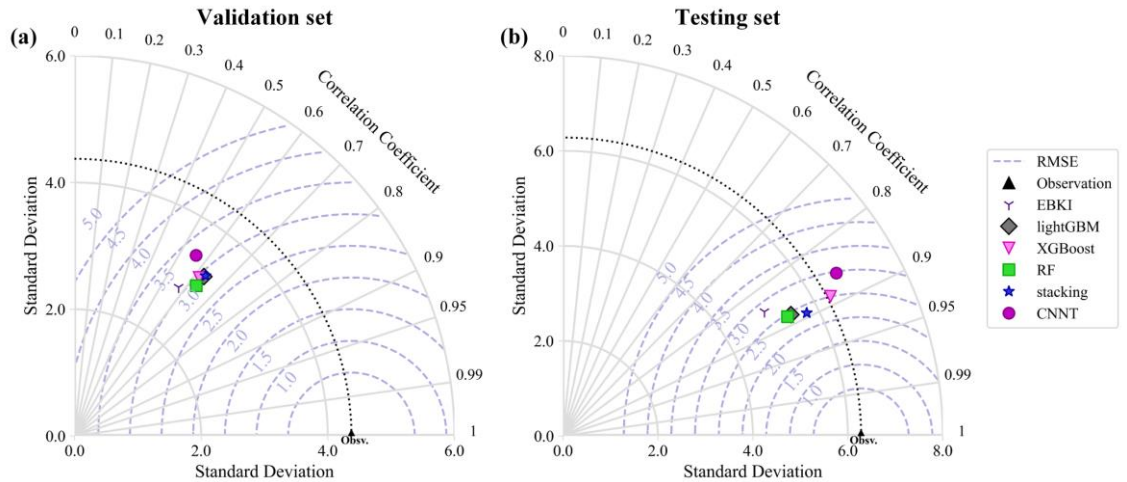


Figure 6. Taylor distribution map of icing gridded interpolation for EBKI and five artificial intelligence models in the validation (a) and testing (b) sets.

3.3. Spatial-Temporal Variations in the Best Model

The stacking model demonstrated the best performance in both the validation and testing sets, as evidenced by the comparison of evaluation metrics across the different models. Next, we further analyze the spatiotemporal variations of the stacking model. Figure 7 presents the evaluation metrics—CSI, MAR, FAR, and RMSE—over time for both the validation and testing sets, along with a bar chart showing the number of pylons with an icing value of 0. The correlation coefficient (R) between the number of pylons with an icing value of 0 and the metrics CSI, MAR, FAR, and RMSE is also calculated.

The results show that the yellow line, representing MAR, indicates that the stacking model mainly reduces MAR at each time point. On the other hand, the orange line, representing FAR, shows little change between the stacking model and EBKI. As a result, the stacking model consistently achieves higher CSI values compared to EBKI at each time point in both the validation and testing sets. Additionally, RMSE decreases for the stacking model at each time point compared to EBKI. In the validation set, the correlation coefficients (R) between the number of pylons with an icing value of 0 and the metrics CSI, MAR, FAR, and RMSE for EBKI (stacking) are: -0.32 (-0.41), 0.26 (0.25), 0.31 (0.54), and -0.51 (-0.51). In the testing set, the corresponding R values for EBKI (stacking) are: -0.75 (-0.07), 0.42 (-0.16), 0.31 (0.12), and -0.24 (-0.26).

These results indicate that in the validation set, there is a strong correlation between the number of pylons with an icing value of 0 and the evaluation metrics for both EBKI and the stacking model. However, in the testing set, the correlation between the stacking model metrics and the number of pylons with an icing value of 0 is significantly weaker than that of EBKI. This suggests that in the validation set, as the number of pylons with an icing value of 0 increases, the icing event intensity decreases, making it more challenging for both EBKI and the stacking model to interpolate icing values for unknown pylons. However, when all pylons are included in the interpolation process, the stacking model's performance is largely unaffected by the number of pylons with an icing value of 0. This finding suggests that to improve future icing interpolation studies, it will be critical to increase the number of pylons used for icing monitoring, leading to more accurate grid-based icing interpolation results.

Lastly, when examining the four cold wave periods, the RMSE values progressively increase with each cold wave phase. This indicates that during these alternating cold wave events, the weather becomes more unpredictable, leading to greater uncertainty and making it more difficult to capture changes in icing values on the pylons.

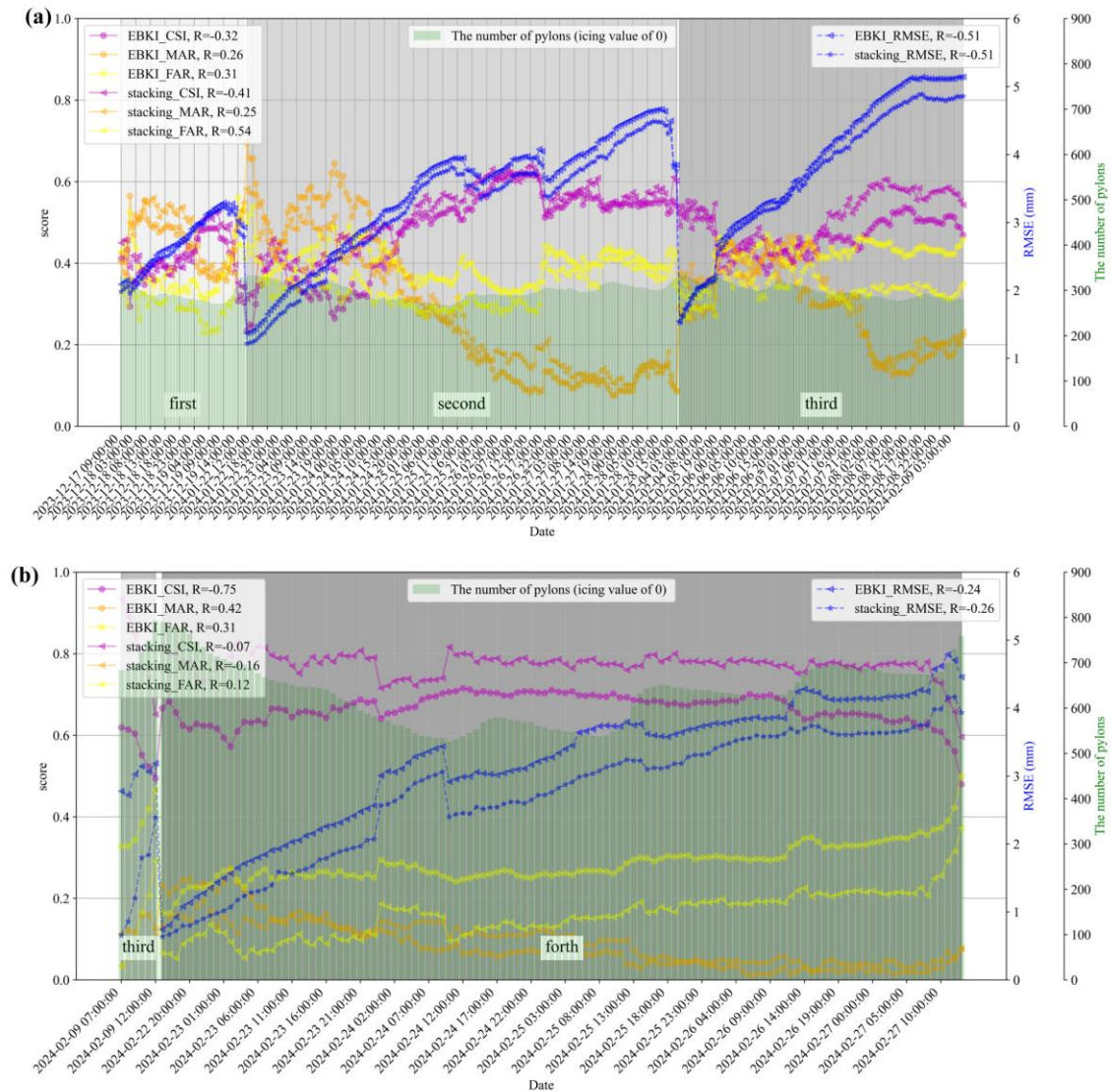


Figure 7. Scatter line plots of CSI, MAR, FAR, and RMSE over time for the EBKI and stacking model in the validation set (a) (testing set (b)); Bar plot of the number of pylons (icing value of 0) over time; R represents the correlation coefficient between the number of pylons (icing value of 0) and CSI, MAR, FAR, and RMSE; The gray shaded areas labeled 'first,' 'second,' 'third,' and 'forth' represent the four cold wave events.

Figure 8 illustrates the spatial distribution of CSI and RMSE for both the EBKI and stacking models on the validation and testing sets. In the validation set, the CSI and RMSE distribution maps show that, following parameter optimization of EBKI, most pylons in the border area between Yunnan and Guizhou, central-eastern Guizhou, and northern Guangxi achieve CSI scores above 0.5 and RMSE values below 5.4 mm. In Guizhou, approximately half of the pylons exhibit RMSE values below 1.2 mm. After further correction of the ice_kriging values using the stacking model, there is an improvement in CSI scores, particularly in northern Guizhou, where they increase from below 0.1 to above 0.1 in the validation set.

Although the stacking model only shows modest improvements over the parameter-optimized EBKI in the validation set, its performance is significantly better on the testing set. As shown in Figure 8e–h, the stacking model notably enhances both CSI and RMSE scores in the testing set, with CSI scores increasing to above 0.9 and RMSE values decreasing to below 0.6 mm in the border area between Yunnan and Guizhou.

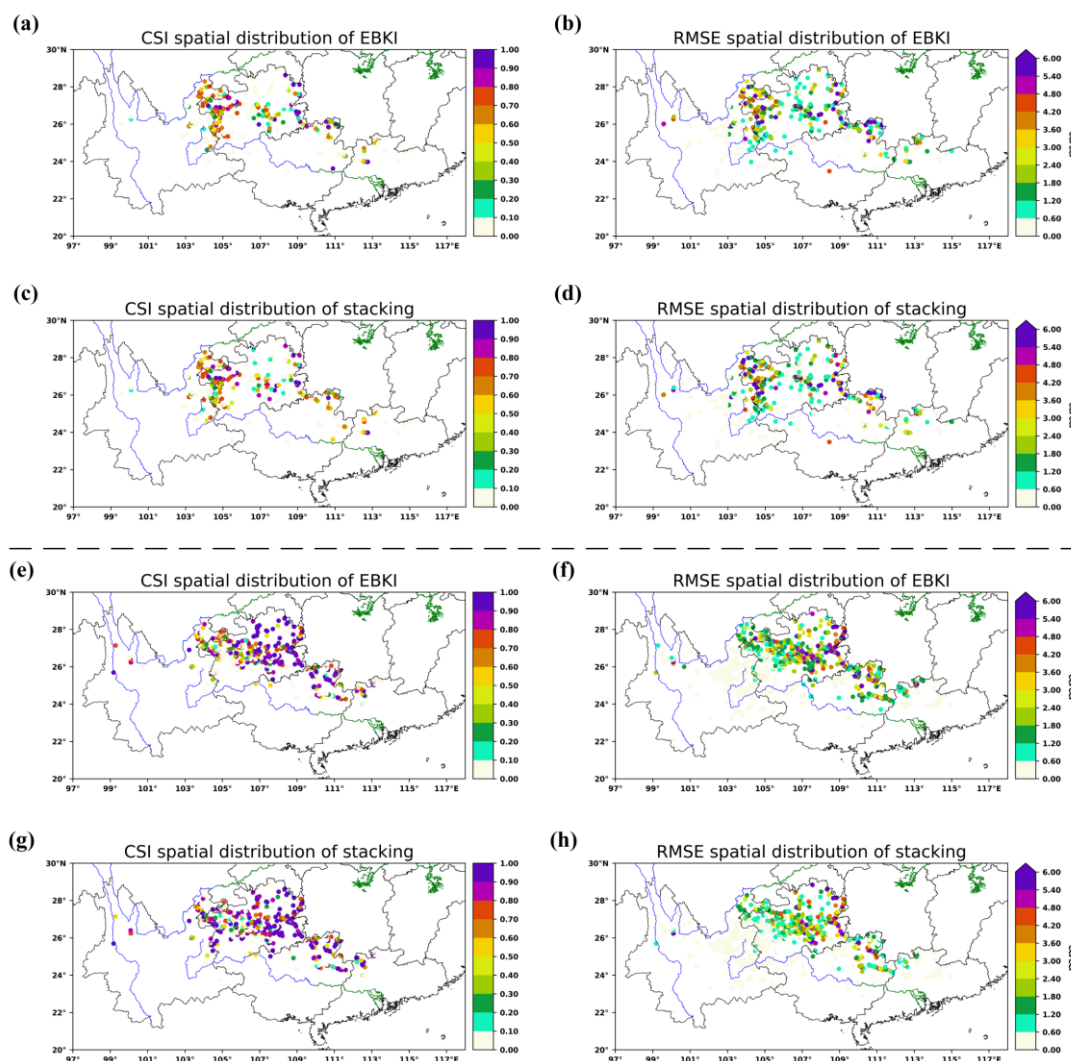


Figure 8. The spatial distribution of CSI and RMSE for the EBKI and stacking model on the validation (a, b, c, d) and testing (e, f, g, h) sets.

4. Discussion

4.1. Interpretation of Feature Factors in SHAP Model

All global features are input variables for the stacking model, and their importance is visualized using a summary plot based on SHAP values (Figure 9). Figure 9a shows the feature importance ranking, where the most influential features in descending order are: ice_kriging, dem_moun, lon, dem, lat, NDVI, PRS, T₂, micro_topo, icing grade, LULC, RH, WS₁₀, and PRE. Figure 9b represents the magnitude of the features, with red shading indicating larger values and blue shading indicating smaller values. The ranking from top to bottom reflects the decreasing importance of these features. The plot also illustrates the impact of each feature on the model's predictions, represented by changes in the Shapley values. For example, considering the most important feature, ice_kriging, the redder the color, the larger the ice_kriging value, and the corresponding Shapley value increases, staying above zero. This indicates that as the ice_kriging value increases, its positive contribution to the model's prediction (tendency for overestimation) grows. Conversely, when the color turns blue (smaller ice_kriging values), the Shapley values gradually decrease from positive to negative, meaning the feature's contribution shifts from overestimation to underestimation.

The plot also shows how data points are spread along the y-axis to avoid overlapping. For example, the blue region for ice_kriging shows significantly more dispersion along the y-axis compared to the red region, indicating that the dataset contains many samples with lower ice_kriging values. This distribution aligns closely with the training set data, where 211,054 samples had ice_kriging values below 2 mm, while only 57,138 samples had values of 2 mm or more. This demonstrates that there are significantly more samples without icing than those with icing in the dataset.

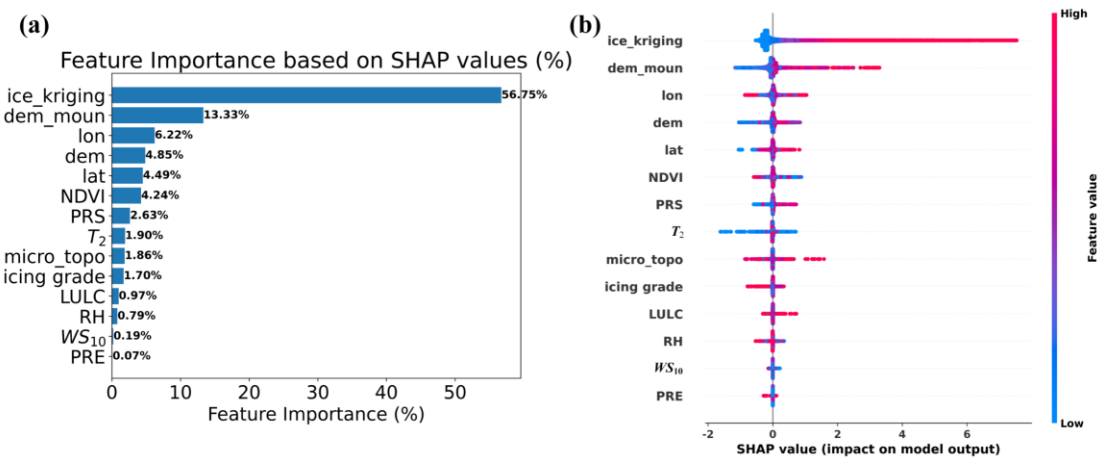


Figure 9. Summary plot of feature importance ranking (a) and the positive and negative impact (b) of features in SHAP.

Local features refer to a subset of input variables in the stacking model, and their effects can be interpreted by analyzing feature dependencies and interactions, as illustrated in Figure 10. Interaction refers to the phenomenon where the influence of one variable depends on the levels of other variables [20]. The SHAP values in this study show that the interaction between features is zero, indicating that the variables in the stacking model are independent, with no significant interaction effects. This absence of interaction can also be seen in the feature dependence plots displayed in Figure 10b, c, d, e, and f.

In these feature dependence plots, which display the behavior of variables such as dem_moun, relative humidity (RH), 10-meter wind speed (WS₁₀), 2-meter temperature (T₂), and precipitation (PRE), it is evident that when the ice_kriging value is very low (i.e., when the EBKI interpolated icing value is below 2 mm), the contribution to predicting ice cover remains nearly zero, regardless of how the individual meteorological factors change. Even when the ice_kriging value increases beyond 2 mm, indicating more significant ice accretion, the contributions to the prediction fluctuate between overestimation and underestimation, but remain independent of changes in the individual meteorological factors.

There are, however, specific ranges where changes in meteorological factors begin to influence SHAP values, reflecting their role in predicting icing. For example, relative humidity (RH) starts affecting predictions when it exceeds 60 %. Wind speed (WS₁₀) impacts predictions when speeds are below 5 meters per second. Temperature (T₂) plays a role when it is between -10 °C and 5 °C. Precipitation (PRE) has an effect when the value is below 5 mm.

These ranges align closely with the conditions typically associated with icing. Research by Fikke et al. (2006) [21] suggests that rime icing tends to occur under conditions where precipitation is less than 0.05 mm, ground temperatures are between -10 °C and 1 °C, relative humidity is above 90 %, and wind speeds are below 6 meters per second. This consistency implies that the changes in SHAP values are influenced by these independent meteorological factors within the range where icing is likely to occur.

Regarding the static features, such as micro_topo and LULC data, the impact of saddle-moisture areas, water bodies, and artificial surfaces on SHAP values is minimal. This is likely due to the small number of samples in these categories in the training set, with only 578 samples for saddle-moisture areas, 289 samples for water bodies, and 867 samples for artificial surfaces. In contrast, other factors within these datasets contribute both positively and negatively to the predictions of icing.

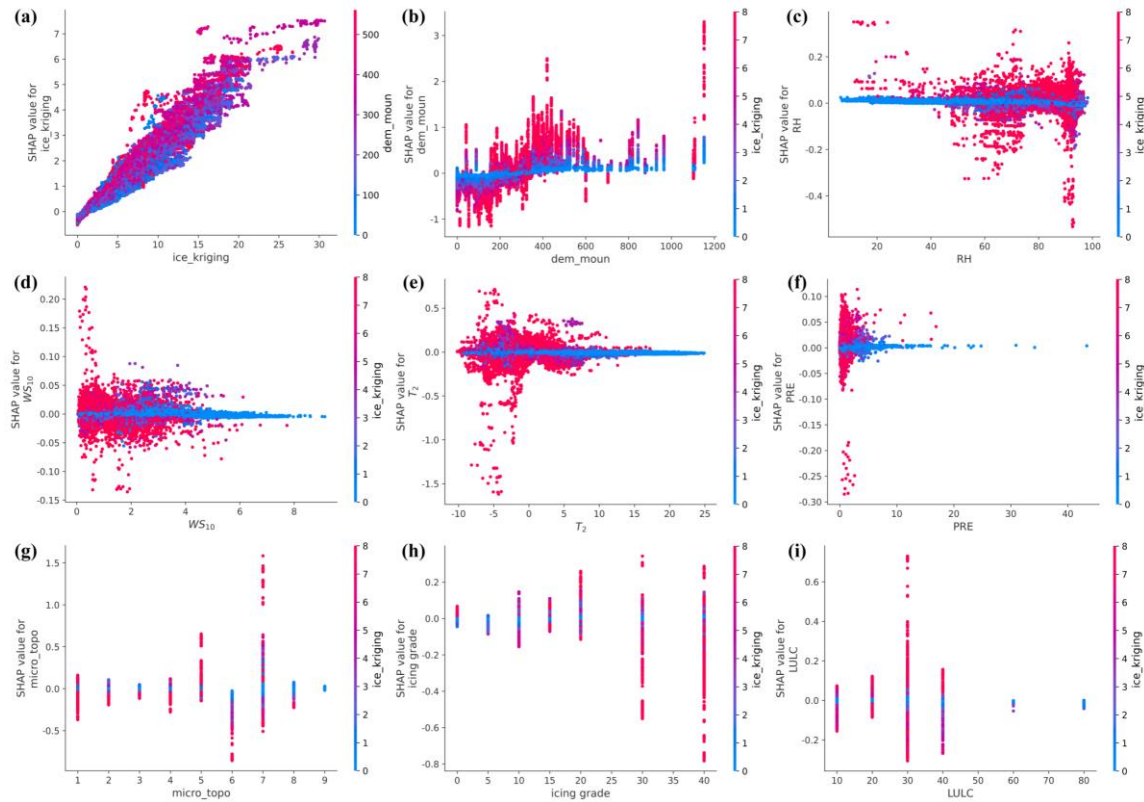


Figure 10. Feature dependence plots of the interactions between ice_kriging (a) and dem_moun; Feature dependence plots of the interactions between dem_moun (b), RH (c), WS₁₀ (d), T₂ (e), PRE (f), micro_topo (g), icing grade (h), LULC (i), and ice_kriging. Ridge, valley, saddle, windward slope, leeward slope, moisture ascent area, ridge-moisture area, valley-moisture area, and saddle-moisture area of the microtopography are replaced by numbers 1-9, respectively; Cultivated land, forest land, grassland, shrubland, wetland, water bodies, artificial surfaces, bare land, and glaciers and permanent snow in the LULC are represented by the numbers [10, 20, 30, 40, 50, 60, 80, 90, 100], respectively.

5. Conclusions

We have utilized a dataset comprising 1,339 cumulative ice thickness measurements from the online ice monitoring system provided by CSPGRI, covering the period from December 2023 to February 2024. By integrating static geographic information—such as NDVI, LULC, and surface elevation derived from satellite-based remote sensing products—with dynamic meteorological elements and ice_kriging values obtained from the parameter-optimized Empirical Bayesian Kriging Interpolation (EBKI), we created a spatiotemporally matched, multi-source fusion dataset for ice thickness grids. These static factors, derived from remote sensing sources, offer critical environmental context that enhances the spatial representativeness and physical interpretability of the model inputs. We applied five machine learning algorithms and six evaluation metrics to identify the optimal stacking model. Additionally, the SHAP algorithm was used to examine the influence of various static and dynamic factors, as well as the ice_kriging values derived from EBKI, on the accuracy of the ice thickness grid predictions.

The optimized EBKI model demonstrated good predictive capability, with the best-performing parameters being a radius of 5, NBR_MAX of 5, and NBR_MIN of 1. In the validation and testing sets, the EBKI model achieved correlation coefficients (R) of 0.574 and 0.851, RMSE values of 3.602 mm and 3.315 mm, and CSI values of 0.477 and 0.668, respectively. These results indicate that the parameter-optimized EBKI model can produce reasonably accurate ice thickness predictions (referred to as ice_kriging), forming a strong foundation for the icing-grid model.

Among the machine learning models tested, the stacking model was the most effective in correcting the ice_kriging values, as demonstrated by its performance in both the validation and testing sets. The stacking model achieved correlation coefficients of 0.634 and 0.893, RMSE values of 3.424 mm and 2.834 mm, CSI values of 0.514 and 0.774, and fbias values of 1.034 and 1.084 for the validation and testing sets, respectively. These results highlight the stacking model's superiority in accurately predicting real-time icing conditions.

Through our analysis of the relationship between the number of pylons with an icing value of 0 and the evaluation metrics, it became clear that increasing the number of icing monitoring stations is critical for improving the CSI score of the icing-grid model. A greater number of monitoring stations would provide more data points, which would enhance the model's ability to interpolate icing values across a wider area, particularly in regions with sparse coverage.

The SHAP model offered valuable insights into the role of various factors in the performance of the stacking model. Meteorological variables such as relative humidity (RH), 10-meter wind speed (WS₁₀), 2-meter temperature (T₂), and precipitation (PRE) were shown to have significant and independent contributions to the accuracy of the predictions, particularly in terms of correcting the ice_kriging values. These findings underscore the importance of meteorological factors in the development of an accurate and physically interpretable icing-grid model.

In conclusion, the stacking model, supported by the SHAP analysis, represents an optimal approach for improving the accuracy of ice thickness grid predictions. The incorporation of remote sensing-derived static features such as NDVI, LULC, and terrain elevation has proven to be a valuable complement to ground-based and meteorological data, further enhancing the model's spatial generalization. Future work will focus on extending this model to a 90 m by 90 m grid across four southern provinces, enabling the production of real-time icing grid products that can support more accurate and timely forecasting of icing events.

Author Contributions: Conceptualization, S.Z.; methodology, S.Z., and Z.G.; software, S.Z.; validation, S.Z., and Z.G.; formal analysis, S.Z.; investigation, B.G. and H.Z.; resources, H.Z. and J.H.; data curation, B.G. and H.Z.; writing—original draft preparation, S.Z.; writing—review and editing, S.Z., and Z.G.; visualization, S.Z. and X.X.; supervision, S.Z.; project administration, S.Z., B.G. and H.Z.; funding acquisition, Z.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the China Southern Power Grid Research Institute in November 2023, the Second Batch of Open Competitive Negotiation Procurement Projects for Production Technology and Support Services (grant no. CG1500062001647685).

Data Availability Statement: The 90-meter grid Digital Elevation Model (DEM) data are available from the Resource and Environment Data Cloud Platform (<http://www.resdc.cn/>, accessed on 28 April 2025). The monthly Normalized Difference Vegetation Index (NDVI) data from December 2023 to February 2024 were obtained from the MODIS MOD13A3 product at a 1-kilometer spatial resolution via the NASA Earthdata platform (<https://search.earthdata.nasa.gov/search>, accessed on 28 April 2025). The land use and land cover (LULC) data with a 30-meter resolution were sourced from the Globeland30 V2020 dataset, developed by the Aerospace Information Research Institute, Chinese Academy of Sciences (<https://www.webmap.cn/commres.do?method=globeIndex>, accessed on 28 April 2025). Meteorological Administration's land data assimilation system (CLDAS-V2.0) real-time product dataset, accessible at https://data.cma.cn/data/cdcdetail/dataCode/NAFP_CLDAS2.0_RT.html (accessed on 28 April 2025).

Acknowledgments: The authors sincerely thank the China Southern Power Grid Research Institute for providing the icing data from the online icing monitoring system, as well as the 2022 ice distribution grade map for a 50-year return period. We are grateful to the anonymous reviewers for their useful comments and suggestions.

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

References

1. Lu, Z.; Han, Y.; Liu, Y. Occurrence of Warm Freezing Rain: Observation and Modeling Study. *JGR Atmospheres* 2022, 127, e2021JD036242, doi:10.1029/2021JD036242.
2. Shen, H.; Wan, B.; Zhou, S.; Kang, J.; Chen, H.; Gao, Z. The Synoptic Characteristics of Icing Events on Transmission Lines in Southern China. *Atmosphere* 2023, 14, doi:10.3390/atmos14121789.
3. Ma, T.; Niu, D.; Fu, M. Icing Forecasting for Power Transmission Lines Based on a Wavelet Support Vector Machine Optimized by a Quantum Fireworks Algorithm. *Applied Sciences* 2016, 6, 54, doi:10.3390/app6020054.
4. Liao, Z.; Zhai, P.; Chen, Y.; Lu, H. Differing Mechanisms for the 2008 and 2016 Wintertime Cold Events in Southern China. *Intl Journal of Climatology* 2020, 40, 4944–4955, doi:10.1002/joc.6498.
5. Zhou, B.; Gu, L.; Ding, Y.; Shao, L.; Wu, Z.; Yang, X.; Li, C.; Li, Z.; Wang, X.; Cao, Y.; et al. The Great 2008 Chinese Ice Storm: Its Socioeconomic–Ecological Impact and Sustainability Lessons Learned. *Bull. Amer. Meteor. Soc.* 2011, 92, 47–60, doi:10.1175/2010BAMS2857.1.
6. Zhao, L.; Ma, Q.; Yang, G.; Wang, X.; Zhao, L.; Yang, X.; et al. Disasters and Its Impact of a Severe Snow Storm and Freezing Rain over Southern China in January 2008. *Climatic and Environmental Research* 2008, 13, 556–566, doi:10.3878/j.issn.1006-9585.2008.04.20.
7. Wang, L.; Chen, Z.; Zhang, W.; Lu, Z.; Cheng, Y.; Qu, X.; Gul, C.; Yang, Y. The Causes and Forecasting of Icing Events on Power Transmission Lines in Southern China: A Review and Perspective. *Atmosphere* 2023, 14, 1815, doi:10.3390/atmos14121815.
8. Yang, J.; Zhu, K.; Liu, B.; Li, X.; Chen, Q.; Yin, Q.; Si, J.; Gao, Z. Method for the Development of Ice Thickness Distribution Maps for Power Transmission Infrastructures in China. *J. Cold Reg. Eng.* 2015, 29, 06014004, doi:10.1061/(ASCE)CR.1943-5495.0000086.
9. Musilek, P.; Arnold, D.; Lozowski, E.P. An Ice Accretion Forecasting System (IAFS) for Power Transmission Lines Using Numerical Weather Prediction. *SOLA* 2009, 5, 25–28, doi:10.2151/sola.2009-007.
10. Podolskiy, E.A.; Nygaard, B.E.K.; Nishimura, K.; Makkonen, L.; Lozowski, E.P. Study of Unusual Atmospheric Icing at Mount Zao, Japan, Using the Weather Research and Forecasting Model. *J. Geophys. Res.* 2012, 117, 2011JD017042, doi:10.1029/2011JD017042.
11. Wang, Q.; Zhou, S.; Zhang, H.; Su, H.; Zheng, W. Prediction of Conductor Icing Thickness Based on Random Forest and WRF Models. In *Proceedings of the 2021 International Conference on Intelligent Computing, Automation and Applications (ICAA)*; 2021; pp. 959–962.
12. Pan, Z.; Zhang, W. Research on Automatic Drawing Method of Power Grid Icing Area Distribution Map Based on ArcGIS. *Electric Power Information and Communication Technology* 2018, 16, 44–48. <https://doi.org/10.16543/j.2095-641x.electric.power.ict.2018.10.008> (in Chinese)
13. Zhang, H.; Wang, Z.; Zhang, W. Exploring Spatiotemporal Patterns of PM_{2.5} in China Based on Ground-Level Observations for 190 Cities. *Environmental Pollution* 2016, 216, 559–567, doi:10.1016/j.envpol.2016.06.009.
14. Mao, F.; Hong, J.; Min, Q.; Gong, W.; Zang, L.; Yin, J. Estimating Hourly Full-Coverage PM_{2.5} over China Based on TOA Reflectance Data from the Fengyun-4A Satellite. *Environmental Pollution* 2021, 270, 116119, doi:10.1016/j.envpol.2020.116119.
15. Zhang, L.; Wang, L.; Ji, D.; Xia, Z.; Nan, P.; Zhang, J.; Li, K.; Qi, B.; Du, R.; Sun, Y.; et al. Explainable Ensemble Machine Learning Revealing the Effect of Meteorology and Sources on Ozone Formation in Megacity Hangzhou, China. *Science of The Total Environment* 2024, 922, 171295, doi:10.1016/j.scitotenv.2024.171295.

16. Minár, J.; Evans, I.S. Elementary Forms for Land Surface Segmentation: The Theoretical Basis of Terrain Analysis and Geomorphological Mapping. *Geomorphology* 2008, 95, 236–259, doi:10.1016/j.geomorph.2007.06.003.
17. Wu, J.; Wen, Y.; Zhang, Q.; et al. GIS based classification and extraction algorithm of ice-prone micro-terrain and its 3D application. *High Voltage Technology* 2023, 49, 1-5. <https://doi.org/10.13336/J.1003-6520.hve.20230088> (in Chinese)
18. Zhou, S.; Gao, C.Y.; Duan, Z.; Xi, X.; Li, Y. A Robust Error Correction Method for Numerical Weather Prediction Wind Speed Based on Bayesian Optimization, Variational Mode Decomposition, Principal Component Analysis, and Random Forest: VMD-PCA-RF (Version 1.0.0). *Geosci. Model Dev.* 2023, 16, 6247–6266, doi:10.5194/gmd-16-6247-2023.
19. Rehman, A.; Zhovmer, A.; Sato, R.; Mukouyama, Y.; Chen, J.; Rissone, A.; Puertollano, R.; Liu, J.; Vishwasrao, H.D.; Shroff, H.; et al. Convolutional Neural Network Transformer (CNNT) for Fluorescence Microscopy Image Denoising with Improved Generalization and Fast Adaptation. *Sci Rep* 2024, 14, 18184, doi:10.1038/s41598-024-68918-2.
20. He, Z.; Yang, Y.; Fang, R.; Zhou, S.; Zhao, W.; Bai, Y.; Li, J.; Wang, B. Integration of Shapley Additive Explanations with Random Forest Model for Quantitative Precipitation Estimation of Mesoscale Convective Systems. *Front. Environ. Sci.* 2023, 10, 1057081, doi:10.3389/fenvs.2022.1057081.
21. Fikke, S.; Ronsten, G.; Heimo, A.; Kunz, S.; Ostrozlik, M.; Persson, P.-E.; Sabata, J.; Wareing, B.; Wichura, B.; Chum, J.; et al. COST-727: Atmospheric Icing on Structures Measurements and Data Collection on Icing: State of the Art. Publication of MeteoSwiss 2006, 110..
22. Lenhard, R.W. An Indirect Method for Estimating the Weight of Glaze on Wires. *Bulletin of the American Meteorological Society* 1955, 36, 1–5, doi:10.1175/1520-0477-36.1.1.
23. Makkonen, L. Modeling Power Line Icing in Freezing Precipitation. *Atmospheric Research* 1998, 46, 131–142, doi:10.1016/S0169-8095(97)00056-2.
24. Makkonen, L. Models for the Growth of Rime, Glaze, Icicles and Wet Snow on Structures. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 2000, 358, 2913–2939, doi:10.1098/rsta.2000.0690.
25. Jones, K.F. A Simple Model for Freezing Rain Ice Loads. *Atmospheric Research* 1998, 46, 87–97, doi:10.1016/S0169-8095(97)00053-7.

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.