Computer Science and Information Systems 20(2):631–656 https://doi.org/10.2298/CSIS220717018C

Machine Learning-based Intelligent Weather Modification Forecast in Smart City Potential Area

Zengyuan Chao

Weather Modification Center of Shijiazhuang Meteorological Bureau,
Shijiazhuang, China
czy93420@163.com

Abstract. It is necessary to improve the efficiency of meteorological service monitoring in smart cities and refine the prediction of extreme weather in smart cities continuously. Firstly, this paper discusses the weather prediction model of artificial influence under Machine Learning (ML) technology and the weather prediction model under the Decision Tree (DT) algorithm. Through ML technology, meteorological observation systems and meteorological data management platforms are developed. The DT algorithm receives and displays the real meteorological signals of extreme weather. Secondly, Artificial Intelligence (AI) technology stores and manages the data generated in the meteorological detection system. Finally, the lightning monitoring system is used to monitor the meteorological conditions of Shaanxi Province from September to December 2021. In addition, the different meteorological intelligent forecast performance of the intelligent forecast meteorological model is verified and analyzed through the national meteorological forecast results from 2018 to 2019. The results suggest that the ML algorithm can couple bad weather variation with the existing mesoscale regional prediction methods to improve the weather forecast accuracy; the AI system can analyze the laws of cloud layer variation along with the existing data and enhance the operational efficiency of urban weather modification. By comparison, the proposed model outperforms the traditional one by 35.26%, and the maximum, minimum, and average prediction errors are 5.95%, 0.59%, and 3.76%, respectively. This exploration has a specific practical value for improving smart city weather modification operation efficiency.

Keywords: Artificial intelligence; Machine learning; Weather modification operation; Intelligent forecast; Decision tree

1. Introduction

With a vast territory and complicated terrain, China still lacks effective meteorological monitoring systems and technologies. For example, some meteorological stations in China fail to receive radar data occasionally due to their geographical locations, thus interrupting normal weather forecasts. Compared with the Weather Surveillance Radar (WSR), the weather events positioning system is widely used because of its extensive coverage, low maintenance cost, and long-time continuous operation [1]. In particular, atmospheric convection is often responsible for lightning, precipitation, hail, gale, and tornado, and lightning is usually followed by precipitation. Atmospheric convection can be located by referring to the location of lightning. It has been argued that precipitation and lightning intensity have a specific correlation. Such correlation can be used to forecast the intensity of the weather conditions, which is of great significance to weather modification [2]. The lightning location data can be used to monitor convective weather and give early warning. Relevant research has shown that the echo intensity of radar well corresponds to ground flash frequencies. Positive and negative lightning frequency shows different characteristics with the storm's evolution. Generally, negative ground flash is closely related to the active period of convective weather development, and its spatial distribution corresponds to a vital updraft region and wind convergence area. In contrast, positive ground flash occurs in intense convective weather's initial and dissipation stages [3]. Lightning location is 30-60min ahead of radar echo, according to which the origin and intensity of the weather events can be calculated, and manual rain precipitation and hail prevention operations can be commanded [4]. Yet, the current weather events positioning system has shown some deficiencies. The occurrence, development, and decline of the weather events are judged by the data characteristics of lightning, providing the basis for the time selection of a weather modification. An important mechanism to improve weather modification efficiency is the real-time and accurate understanding of the distribution and transportation of precipitable water vapor in the atmosphere using lightning detection data [5]. The distribution and lightning of water vapor in weather events can be judged by its relation to lightning. The method of seeding catalysts, such as silver iodide or dry ice, is mostly used in weather modification in the convective weather process. Hence, properly-timed cloud seeding is much more cost-effective [6].

Lightning monitoring systems have been established in some areas without unified guidance on station site, equipment, communication, application, and forecast product release [7], and a full meteorological monitoring business system based on the existing system and resources will be constructed. Regarding the research of Artificial Intelligence (AI) technology in urban weather forecasting, Chang et al. (2020) used AI technology to obtain satellite image data. They verified the prediction of urban rain and flood by AI and hydraulics through 7·12 Beijing rainstorm data information [8]. Yang et al. (2019) proposed a short-term flood forecasting model based on AI technology to improve the parameter calibration, professional factors, and other problems of traditional flood forecasting models. They also processed the rainfall data of Xixian County from 2010 to 2018. It was found that the prediction duration of the model was 24 hours to 36 hours, with high prediction accuracy, meeting the requirements of flood forecasting [9].The main tasks of the lightning detection business include: (1) Strengthening the development and application research of lightning positioning technology; (2) Improving the lightning positioning system equipment; (3) Enhancing the comprehensive positioning technology of the national monitoring station network for meteorological business. Meanwhile, the performance of the meteorological monitoring station network should be evaluated, and the positioning accuracy and detection efficiency of the meteorological observation system should be improved. Thereupon, the foundation for the meteorological data application platform and shared resource databases are provided, applying the meteorological data widely while maximizing the benefits of the meteorological observation system.

The innovation of this paper is to use AI technology to establish a Local Area Network (LAN), which is used to transmit data and connect lightning detection stations. It can provide information connection for meteorological conditions in different regions. A Geographic Information System (GIS) collects and visualizes spatial data.

Multidimensional spatial data can be queried to realize data sampling. Besides, an intelligent weather forecast model for extreme weather is established through AI technology, Machine Learning (ML) technology, and the Decision Tree (DT) algorithm. The correlation between the meteorological forecast model and the monitoring and analysis system is proposed. Then, the weather forecast for extreme weather is accurate.

This exploration encompasses six sections. The first section is the introduction, explaining the importance and significance of researching intelligent meteorological forecasts, describing the problems in the current meteorological monitoring, and clarifying the research framework and main research contributions. The second section is the literature review, summarizing and analyzing the research of AI, ML, and Deep Learning (DL) methods in intelligent weather forecasting. It also defines the shortcomings of existing research and proposes the latest research algorithm. The third section describes the existing weather forecast models based on AI and ML algorithms, and the simulation parameters and environment are explained. The fourth section is the simulation analysis of the proposed model, which is compared with the other latest literature methods to verify the effectiveness of the proposed model. The fifth section is the discussion, which comparatively analyzes the results of the proposed model and previous research results. The sixth section is the conclusion, a summary pointing out the shortcomings and prospects.

2. Literature Review

2.1 AI Weather Forecast

AI refers to the simulation of human intelligence processes by machines, especially computer systems. There is a natural coupling relationship between the weather forecast, which needs massive amounts of miscellaneous data, and AI [10], which can deal with big data efficiently and infer from incomplete and uncertain information with insufficient spatial-temporal data density [11]. Additionally, AI can summarize expert knowledge and experience, improve the prediction level, and utilize an unusable understanding of statistical and numerical models. Mohammadi et al. (2018) improved the performance and efficiency of CMOS Ring Oscillator (RO) by multi-objective optimization and Particle Swarm Optimization [12]; Mohammadi et al. (2019) introduced an AIO method for modeling infinite impulse response system and evaluated the design and optimization of IIR digital filter. The results showed that the algorithm model outperformed other heuristic algorithms and Genetic Algorithms (GA) [13]; Shahraki et al. (2020) optimized the component values of analog active power filters based on multi-objective optimization. MOIPO outperformed other methods in minimizing quality factor deviation and cutting off frequency deviation [14]; Mohammadi et al. (2019) proposed an SI algorithm by introducing a novel index in IIR filter design, which was more suitable and reliable than the EC algorithm. Then, the performance of the proposed SI algorithm was analyzed from reliability, Mean Square Error (MSE), and IoS, which were better than other algorithms; Farzaneh et al. (2021)

believed that the AI method could help people plan urban construction, and AI-based modeling was widely used to predict building energy consumption [15].

2.2 ML Weather Forecast

ML, the core of AI research, primarily aims to obtain knowledge and information from input and historical data to solve more complex problems, reduce errors, and automatically learn and improve itself [16]. ML is also an interdisciplinary subject, studying how computers simulate human learning behavior to acquire and retain new knowledge and skills and reorganize the existing knowledge structure for self-improvement [17]. Kirkwood et al. (2021) believed that under ML technologies, AI could be applied to many fields, such as data mining, computer vision, natural language processing, biometrics, search engine, medical diagnosis, detection of credit card fraud, securities market analysis, DNA sequencing, speech and handwriting recognition, strategy games, and robots [18]. Hossain et al. (2020) suggested that meteorological information could be classified by DT learning in supervised learning according to the feature standard [19]. Xu et al. (2020) considered that the DT could be used to classify objects conforming to model input, which included decision points, state nodes, and result nodes. Precisely, each internal tree node (non-leaf node) and edge corresponded to an object attribute and its option. While each leaf node corresponded to the primary classification of objects [20]. Caron et al. (2020) believed that unsupervised learning could excavate the data structure and cluster input data according to feature similarities. Accordingly, the meteorological data under similar weather conditions in historical records were classified into clusters, such as gale and light rain. The current meteorological data are used for regression analysis to predict future weather events [21]. Bao et al. (2021) proposed a short-term power load forecasting model IGA-LS-SVM based on an AI algorithm. They analyzed weather, temperature, short-term power load, work, and holidays. The proposed model showed excellent predictability when associated with the existing Backpropagation algorithm, with a statistical significance of 0.8274, higher than other algorithms. Thus, the proposed IGA-LS-SVM was suitable for short-term power load forecasting [22]. Memiş et al. (2021) constructed a new classification algorithm using 18 learning real-world databases in ML, finding that FPFS-EC outperformed SVM, Fuzzy KNN, FSSC, FussCyier, HDFSSC, and BM-Fuzzy KNN in 13 of the 18 experimental datasets [23].

2.3 DL Weather Forecast

DL is a new field of ML and originates from Neural Networks (NNs) research. It can combine low-level features into more abstract high-level attribute classes or feature representations [24]. Traditionally, computer algorithms produce the desired results based on rules and data, while DL can generate appropriate rules and structures by simulating the complex human brain structure using massive amounts of training, which is especially suitable for problems that are difficult for traditional computers yet simple for human beings, such as Natural Language Understanding, handwriting recognition, and visual classification [25], as well as another extremely complex problem, weather forecast. The weather forecast system has inevitable delays, given extensive data acquisition, conversion, and analysis. DL can be combined with the NN to analyze images directly, which has the advantages of multiple factors combination and is convenient and efficient [26].

Yin et al. (2021) considered that the WSR detected weather information images through microwave signals and displayed the echoed images on the screen, on which blue to purple indicated the echo intensity from small to large (10~70dBz). Rainfall intensity, range, possibility, and movement could be judged from different colors, color block sizes, and changes [27]. Huang et al. (2021) argued that each NN layer could calculate the color depth, color area coverage, and regional change of the image, respectively, and predict the possibility, intensity, coverage, and duration of meteorological conditions [28]. Xia et al. (2020) proposed that a satellite cloud image was an image of cloud cover and ground surface features observed from top to bottom by meteorological satellites. The Convolutional Neural Network (CNN) could identify cloud species, genera, and precipitation conditions through satellite images' color, shape, structure, brightness, and texture. For example, the deeper the green in the infrared satellite cloud image was, the stronger the ground radiation was, and the better the weather was; an arc-shaped cloud line in the image indicated an actual arc-shaped cloud line [29]. Guo et al. (2020) suggested that the CNN could determine the location and intensity of weather events according to the meteorological characteristics of these satellite images and judge their future movement and evolution trend. Image recognition could predict natural disasters, such as tsunamis, hurricanes, and thunderstorms, through radar and satellite cloud images [30]. For instance, a typhoon is caused by high ground temperature and rising hot air flow that form a low-pressure center. With the change of air pressure and the earth's movement, the air flowing into the center of low pressure also rotates, forming a counter-clockwise air vortex. The cyclone will strengthen to form a typhoon if the temperature does not drop. At present, typhoon interference cloud clusters can be identified a few days earlier, so the moving direction of a typhoon can be judged based on cloud type, wind intensity, and humidity, and its path can be released and forecasted. The advantages of the above ML methods are outstanding, but they need to be operated in the existing set-up program, while reinforcement learning can automatically adjust the state to meet the operation requirements. Table 1 is the summary and analysis of the existing literature.

Based on the above research, most scholars have conducted weather forecasting models under the development of ML and DL. However, the training of weather forecast models is greatly limited by computer hardware. Currently, there is a large amount of data for training samples of the meteorological weather forecast. Regarding time scale, the influence of weather elements at different times in recent days on future weather should be considered. On the spatial scale, there are more than ten elements in the vertical direction. Each element has more than ten levels. The accuracy of longitude and latitude grid points that can be selected within the region is also high. The data to be trained far exceeds the capacity limit of the current Central Processing Unit (CPU) after merging. The training can only be conducted after the original data is cut and magnified. The massive data resources cannot be fully utilized. ML technology is used to reform the weather forecast system to compensate for this research's lack. A meteorological monitoring business system is built to integrate meteorological monitoring, early warning and forecasting, research and technology development, and lightning detection and protection technology services. In addition, this paper uses the DT algorithm to integrate the received thunderstorm flow meteorological signals to effectively simulate the formation of weather, which can predict extreme weather.

Table 1. The summary and analysis of the existing literature

Author (year)	Methods	Advantages	Disadvantages	Data	accuracy	Application
Mohammadi et al. (2018)	Multi-objective optimization	Better performance	Large dataset requirements	Target number	81.26%	RO
Mohammadi et al. (2019)	AIO	Data acquisition is relatively easy	It is difficult to build the model.	Pulse voltage	83.15%	Digital filter
Mohammadi et al. (2019)	SI algorithm	High robustness	Lack of stability	Pulse voltage	86.35%	IIR filter
Shahraki et al. (2020)	Multi-objective optimization	High robustness	Low accuracy	Frequency deviation	76.39%	Active power filter
Hossain et al. (2020)	DT	High accuracy	It is difficult to build the model.	Meteorological data	82.14%	Meteorological field
Xu et al. (2020)	DT + SVM	Data acquisition is relatively easy	Lack of stability	Meteorological data	80.13%	Meteorological field
Caron et al. (2020)	Unsupervised learning + clustering	High accuracy	Poor robustness	Meteorological data	83.15%	Meteorological field
Guo et al. (2020)	ML	High robustness	Low accuracy	Meteorological data	81.12%	Meteorological field
Farzaneh et al. (2021)	AI	High robustness	Low accuracy	No	76.38%	Building energy consumption
Kirkwood et al. (2021)	ML	Data acquisition is relatively easy	Lack of stability	No	No	Multiple areas
Bao et al. (2021)	AI algorithm	Better performance	Large dataset requirements	Power data	86.38%	Power load forecasting
Memiș et al. (2021)	Classification algorithm	The algorithm is easy to build.	Large dataset requirements	Multiple datasets	83.69%	Multiple areas
Yin et al. (2021)	Clustering algorithm	High accuracy	It is difficult to build the model.	Meteorological data	86.31%	Meteorological field
Huang et al. (2021)	ML	High accuracy	Lack of stability	Meteorological data	84.36%	Meteorological field

3. Intelligent Forecast Model of Weather Modification Operation

3.1 ML Weather Modification Forecast

The weather modification forecast system based on an ML algorithm includes meteorological monitoring, data storage, meteorological early warning, meteorological data sharing, weather modification enhancement, and hail suppression guidance product release. The key to a meteorological observation system is to strengthen the detection and construction of urban meteorological operations and promote the combination of district meteorological operations and weather modification scientific research and services. Finally, a meteorological monitoring business system is constructed, integrating meteorological monitoring, early warning and forecasting, research and technology development, and lightning detection and protection technology services [31]. Here, the construction demand of a regional meteorological monitoring system is analyzed explicitly from the aspects of early warning and forecasting comprehensive service based on the existing meteorological detection means and the accumulation of historical meteorological data. All kinds of data are comprehensively processed by developing a meteorological observation system and meteorological data management platform. Then, the technical means of weather modification, rain enhancement, and hail suppression are broadened based on meteorological monitoring data. Radar and satellite observation data and numerical model research are combined, and various applicable products have been developed for meteorological business and social services, thus providing visual analysis and auxiliary decision-making services for meteorological early warning and forecast and establishing a perfect region-wide weather modification forecast and early-warning business system.

The near forecast of weather modification and precipitation enhancement refers to the forecast of precipitation enhancement potential within zero to six hours. The comparison of real-time observation data is the primary decision-making basis for the forecast. The meteorological monitoring data network is used to obtain the original data of meteorological observation data in real-time and decompose the radar image combined with the observation data of radar data and satellite cloud image. The spatial-temporal changes in the weather are tracked based on the meteorological positioning, the development status is described through the exponential trend, and the early meteorological warning is realized. Figure 1 shows the structure of the meteorological early warning system. The meteorological information and the cloud images are analyzed comprehensively and used to understand the development of current weather events to facilitate weather modification. The best time for modification operation is selected according to the frequency of weather events in the cloud.

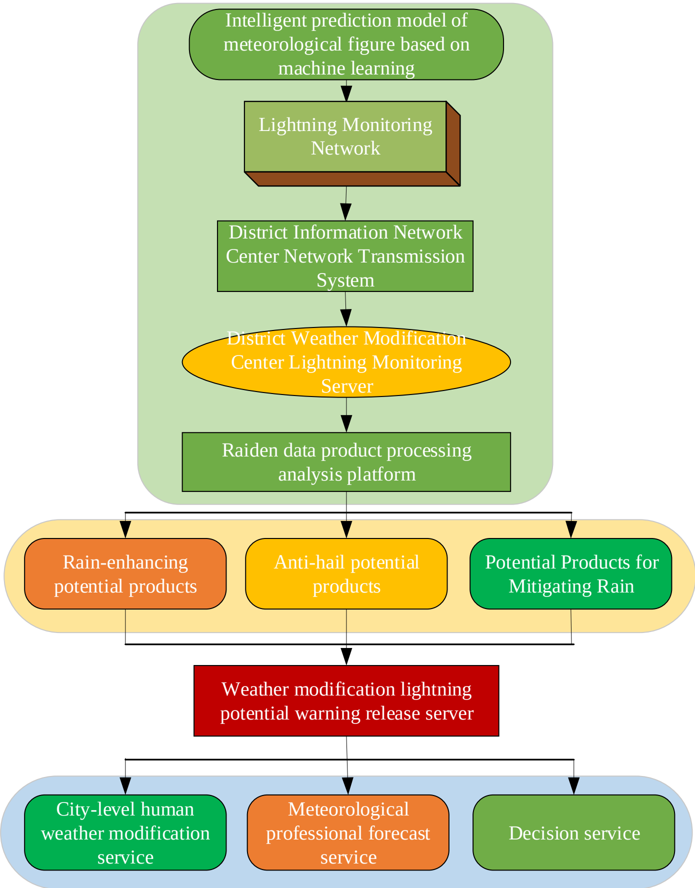

Fig. 1. Structure of meteorological early warning system

3.2 DT Weather Modification Detection

A DT is mainly used in meteorological detection algorithms. The algorithm can effectively simulate meteorological formation, generate the probability of location reliability in meteorological observation, and analyze the error of single and multi-meteorological station locations. The general process of the DT algorithm includes six aspects: collecting data, preparing data, analyzing data, constructing algorithms, testing algorithms, and applying algorithms. Any method can be used to obtain a dataset to collect data. Data often contains many errors and defects. In data preparation, feature selection is the most crucial step in data preprocessing. Attribute selection integrates subsets of data and eliminates worthless attributes. The strong correlation between attribute sets can simplify the model, reduce training time, and decrease the risk of overfitting. Analyzing the data requires selecting the appropriate statistical method to predict the results. If not, the database needs to be adjusted. Constructing the algorithm requires building the data structure of the tree and using the training data to establish the DT model. The testing algorithm evaluates the model performance using a test data set of the constructed DT model test sample. Finally, the DT can understand the internal relationship between data and has high robustness. This means that the model can be applied to new data for data prediction and analysis. The research and application of the DT algorithm suggest that there is a one-to-one correspondence between the received meteorological signal from the thunderstorm discharge and the algorithm calculation. The model output statistical method and linear stepwise regression method are used to analyze the statistical relationship between the lightning data and the nested grid model through the frequency of weather events and the main influence of meteorological parameters. The results can be employed to estimate the possibility of regional meteorological formation. Moreover, it is pointed out that extreme weather events are often formed from large-scale stratification instability and convergence provided by humidity and local wind field. The algorithm analyzes 274 meteorological signals received by four meteorological stations and then filters and compares them step by step. The crucial factors are selected from the four stations, and a Multiple Regression (MR) equation is established to calculate the real meteorological signal for receiving and displaying. The MR equation should use a variance ratio to test the significance of the regression equation. If the influence factor passes the significance test, it can enter the equation. Otherwise, it should not enter the regression equation. The criterion for excluding variables in the regression equation is also to use the variance ratio for the significance test. The test eliminates the variables that contribute the least to the sum of squares of partial regression, whether they are selected into or excluded from the regression equation. The selection and exclusion items that meet the conditions are eliminated. The stepwise regression method eliminates the factors that have little influence on the dependent variables, reducing the difficulty of analyzing problems. It also improves the calculation efficiency and the stability of the regression equation, with good prediction accuracy. The error is minimized through the accurate calculation of the algorithm. Meanwhile, an accurate forecast is conducted, and the research on the numerical model of the weather forecast is strengthened [32].

The DT method is one of the fast and most intuitive inductive learning methods. It has been widely used in expert systems, industrial control processes, financial insurance prediction, and medical diagnosis. It can integrate various data and forecast methods in the weather potential forecast on a platform. The DT method is applied in the meteorological monitoring system to compare the data of positioning instruments, conventional radiosonde, satellite cloud image, and T213 numerical forecast products and presents real and reliable meteorological formation data [33]. Specifically, the DT algorithm should be used in the meteorological monitoring system. First, in the DT generation phase, the algorithm judges the branches of the DT according to the root of the attribute parameter value of each node. It selects the path channel attribute to distinguish the sample parameters. The distance between the branch node tree and the root node is selected. The probability of a reliable position in generating meteorological observation can be determined according to the distance between nodes. Also, the error of single and multiple meteorological station positions can be analyzed. During this period, the branch judgment standard of the DT is selected according to the lightning detection algorithm as the tradeoff standard for using the data branch information value.

The DT algorithm can better integrate all data and forecast methods in the potential meteorological forecast on one platform. It is used to establish the meteorological monitoring and forecast system. The satellite cloud image and the real meteorological data are integrated and used for weather forecasts to improve the system's performance. The binary tree algorithm is used here, and the branch information is used as the judgment standard of DT branch selection. The binary DT is established through the established rules of attribute branching. Figure 2 presents the collection process of a meteorological signal. The specific path is to judge whether the selected path is a multi-possibility selection or whether the lightning signals collected by each station are inconsistent. The multi-station comparison method is employed to select the best path for the possibility of lightning formation. The collected data point signals (>ten-point) are counted. All the points are digitized and statistically analyzed to select the following path. Each attribute value's weight is found, and the best path is selected.

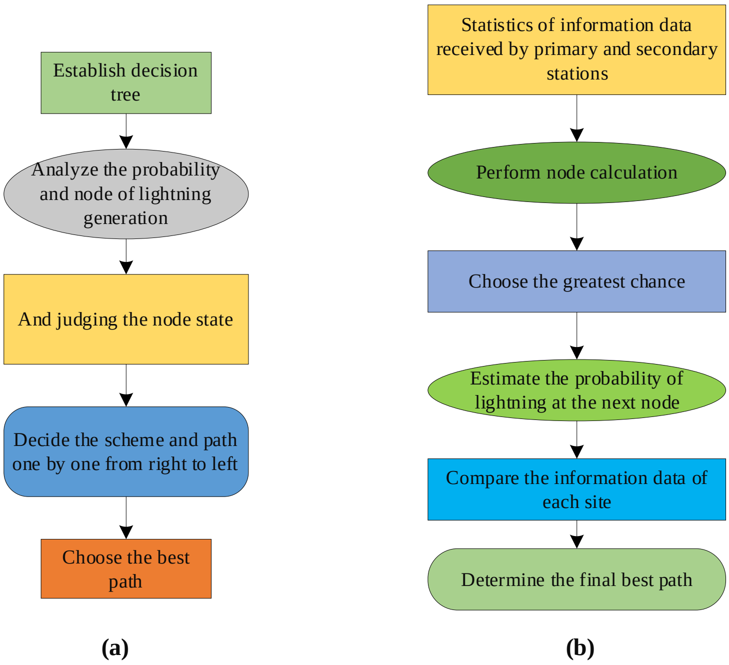

Fig. 2. Flowchart of meteorological signal collection under different conditions, (a) Flowchart of data collection when meteorological signals of each station are consistent;(b) Flowchart of data collection in case of inconsistent meteorological signals at each station

The local electronic map should be loaded into the lightning detection system. Mapinfo is used to add label layers, such as lightning locator and city names on the map, to assist in display and query. Then, the labeled map is loaded into the lightning detection system. The lightning detection system should include an essential layer for background display and an item layer for lightning positioning data. Figure 3 presents the process of loading electronic map data.

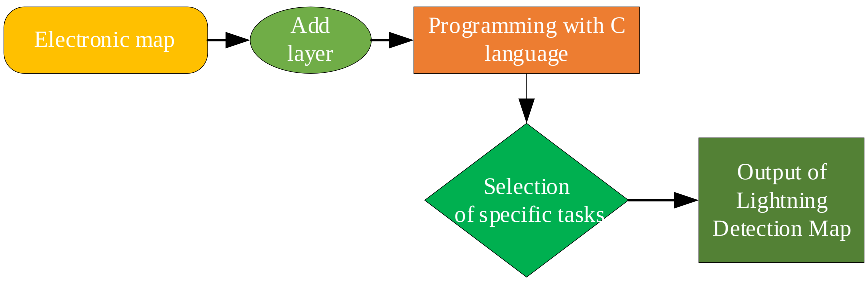

Fig. 3. Loading flow chart of electronic map data

The DT can be simplified using the impurity function φ (p1, p2, ..., p*j*) on each branch node. For any node *t*, the calculation of impurity *i*(*t*) reads:

$$i(t) = \varphi(p(1|t), p(2|t), \dots, p(j|t)) \tag{1}$$

p(*i*|*t*) represents the probability of the node *t* samples belonging to category *i*.

The impurity reduction after branch *s* divides the parent node *t* into child nodes t_R and t_L reads:

$$\Delta i(s,t) = i(t) - P_R i(t_R) - P_L i(t_L) \tag{2}$$

P_R and P_L represent the proportion of child nodes t_R and t_L samples in parent node *t* samples, respectively. The reduction in impurity characterizes the branching effect φ (s, t).

The branching criterion: the criterion with the largest impurity reduction is used for further branching, which is expressed by Eqs. (3) and (4).

$$\Delta i(s,t) = \max_{s_j \in S} \Delta i(s,t) \tag{3}$$

$$S = \{s_1, s_2, \text{K}\} \tag{4}$$

The criterion for branching termination: branching terminates when the maximum impurity reduction is less than the given threshold β.

The binary DT based on a growth-pruning method is used for error analysis of the lightning detection system. The data judgment accuracy of the original tree is not ideal because of redundant nodes. After continuous pruning and selection, the attribute of each node in the lightning detection DT is the threshold of the attribute of generating lightning weather. The contribution of different attributes greater than the threshold of the same node to the generation of lightning weather is different. The DT of lightning detection calculates the possible lightning data of each 4°×4° longitude and latitude grid point in Tibet (The geographic information map of Tibet is divided into multiple 4°×4° grid blocks). The maximum depth of the DT is five because the classification is based on the work area size and combines related references. The DT is used to analyze the possibility of lightning occurrence. The data analysis shows that the greater the value obtained is, the larger the possibility of lightning occurring in the detection point is; on the contrary, the smaller the probability is, the less likely lightning will appear in the detection point.

3.3 AI Meteorological Data Storage

AI is a branch of computational science, which is to study how to simulate human brain thinking through the computer to realize machine intelligence. AI involves a wide range of disciplines, including both natural and social disciplines, and is mainly realized through traditional programming technology and simulation. Firstly, manual programming, as an engineering method, is tedious and error-prone and is used in character recognition and AI chess players. Secondly, the simulation method is based on biological thinking and, thus, is more complex, including GA and Artificial Neural Network algorithms. The simulation method is more robust and can handle more complex problems [34].

The introduction of AI has excellent advantages in model correction, which can significantly improve the prediction ability of temperature, precipitation, and other meteorological elements. In addition, AI technology can also track the development of weather processes. AI technology can locate typhoons according to satellite cloud maps through image recognition to tracktyphoonsn in real time, quickly, and intelligently. On the one hand, the development process of typhoons can be accurately described. On the other hand, a good initial field for the model can be provided, which helps improve the ability of the model to predict typhoons. The meteorological monitoring system will produce a large amount of data, which needs to be effectively managed and stored. Through strong learning ability and fast reasoning speed, AI technology provides a database for multi-source multi-resolution Spatio-temporal observation data and multi-mode and multi-scheme prediction data in the meteorological field. The numerical weather model forecasts and simulates the weather based on the complex atmospheric motion process. AI algorithms are driven by big data. It can capture meteorological data's temporal and spatial characteristics and nonlinear relationships. SQL Server 2000 database stores lightning location data collected through the control and data operation statements. Data collection mainly includes: reading data text format and data entry. The data entry accuracy can be intuitively understood through feedback, including the database name, database authentication method, entry time, entry name, and the entry row number. The data is initially explored, and the model is constructed. Then, they are applied to all data through loop traversal.

Data query: multiple controls and SQL query statements are used to match the query information field. The queried contents are shown in the table. The lightning data query system can draw data output to give users an intuitive understanding of the lightning data, through which the strong center of weather events can be located. The queried lightning data are used to analyze the intensity center of lightning occurrence. Then, the lightning frequency, position of the intensity center, formation direction, distance from the detection point, and the intensity of the maximum charge center are obtained. Finally, the discharge area of the thunderstorm cloud center is obtained through the corresponding equation of the meteorological precipitation enhancement potential forecast.

Here, a smart city's intelligent weather prediction model is realized through ML technology, DT algorithm, and AI technology. The role of different research methods in this paper is shown in Table 2.

Table 2. Functions of research methods in this paper

Research method	Function
AI technology	Effectively store and manage data of meteorological monitoring system.
ML	The meteorological forecast system can be reformed. A meteorological monitoring business system can be established by integrating meteorological monitoring, early warning and forecasting, research and technology development, and lightning detection and protection technology services.
DT algorithm	Collect meteorological signals under different conditions.

3.4 System Parameters and Environment Settings

The network environment in the lightning monitoring system is based on the LAN and the General Packet Radio Service wireless network. The software architecture is established through the Client/Server technology network based on Windows 10 operating system. The lightning monitoring system is designed using the principle of function separation, independent work, non-interference, and scalability. The system includes a display unit for lightning data, a receiving unit for secondary station data, a multi-station positioning unit, and a data viewing unit. The receiving unit for the secondary station data can be expanded multiple times without changing the source program. It has strong adaptability to the increase or decrease in the number of secondary stations.

Background running environment: the hardware of the application server is HP PROLIANT DL580G7, equipped with 4 Xeon E7-4807 1.86GHzCPU, 16G DDR3 memory, 900G hard disk, configured with Windows 10 ServerR2 operating system, and JBoss and GeoServer. The database server is HP PROLIANT DL580G7, equipped with 4 Xeon E7-4807 1.86GHzCPU, 8G DDR3 Memory, and 900G hard disk, configured with Windows2008ServerR2 operating system, and PostgreSQL and SQLServer2005 [35].

3.5 Experimental Data

The data here comes from the National Climatic Data Center (NCDC). The lightning meteorological conditions from September to December 2021 in Shaanxi Province and the meteorological forecast results of Yanta District in Xi'an, Shaanxi Province, from 2018 to 2019 are retrieved through data. Table 3 reveals the data components of meteorological variables.

Table 3. Data composition of meteorological variables (data source: NCDC)

Name	Root Mean Square Error	Deviation	Correlation coefficient
Air temperature	0.88K	-0.13K	0.97
Specific humidityc	4.76%	4.76%	0.93
Wind speed	0.83m/S	-0.21%	0.82
Surface barometric pressure	3.74hPa	-0.38hPa	0.96
Hourly precipitation	0.94mm/H	-0.004 mm/H	0.72

Additionally, the number of input features and the length of output results of each sample must be the same to reflect that the intelligent prediction model of a smart city based on AI and ML does not need all training samples to maintain consistency. This is also the characteristic of the verification model proposed here, which is different from other ML algorithms. The number of training set data is 300, and the number of test set data is 100. Training data set samples are input in the training phase, n categories. The number of samples of each type and the number of samples of the maximum type m are recorded. The random list L of each category is used to modularize the number of category samples. Resampling is performed according to the modular operation results so that the number of samples of each class is m. Its index is stored in the corresponding L. Finally, all the L lists are combined to form a final list and output the training list. Secondly, the intelligent city prediction model based on AI and ML proposed here is developed for meteorological prediction, so its evaluation algorithm is more specific and rigorous than the general algorithm. The real-time test set of the algorithm is the real-time weather data collected every day during the test period and the results observed by multiple observers back-to-back, which cannot be used for model verification and evaluation in advance. In model training and testing, AI technology is used to store and manage the data of the meteorological monitoring system. The DT algorithm is used to collect meteorological signals under different conditions. This can effectively avoid the problem of model applicability caused by training data sets. Based on AI technology and ML, the overall consistency can be observed for a long time under different weather conditions in different regions. The model's applicability to different geographical locations is also tested while testing the consistency with manual observation.

4. Model Simulation and Comparative Analysis

4.1 Analysis of Meteorological Statistics

Here, the frequency of lightning occurrence in Yanta District, Xi'an City, Shaanxi Province, is statistically analyzed to conduct statistical analysis on the lightning monitoring system and verify its feasibility of the lightning monitoring system. The frequency change of a lightning occurrence in Yanta District, Xi'an City, Shaanxi Province is shown in Figure 4. The results reveal that the lightning monitoring system can calculate the lightning flash frequency in a cycle and draw it out through the broken line chart. It can well reflect the changing trend of the relevant data during this period and can intuitively show the monthly lightning frequency, which is conducive to intuitive and concise analysis. The frequency of lightning in September is significantly higher than that in other months, and that in December is only 0.2.

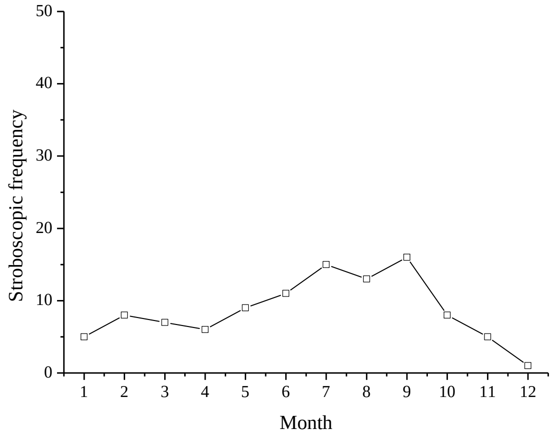

Fig. 4. Frequency change of a lightning occurrence in Yanta District, Xi’an City, Shaanxi Province

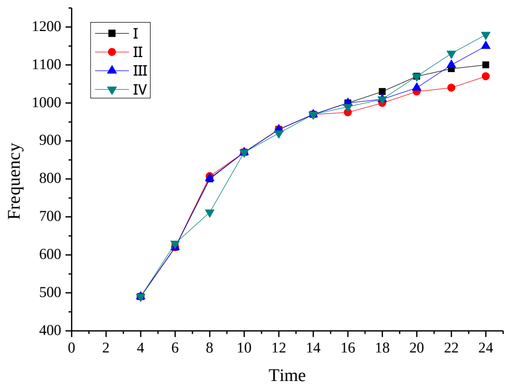

Fig. 5. Statistical chart of lightning frequency in different areas of a day\

In Figure 5, the lightning frequency in different areas (Xi'an: I, Hanzhong: II, Shangluo: III, and Weinan: IV) is calculated and plotted. It is found that the combination of lightning frequency and satellite cloud image information can be used to obtain the law of lightning occurrences in the future. The lightning frequency significantly increases with time, and the highest lightning frequency in IV is 1,163 times in 24 hours. Accordingly, the right time for weather-modification operations can be determined.

4.2 Performance analysis of intelligent forecasting

In Figure 6, two forecast dates in 2018 [36] and 2019 [37] are used as examples. The dataset of the four forecasting models is the meteorological forecast data of Yanta District, Xi’an City, Shaanxi Province, from 2018 to 2019. The different meteorological intelligent forecasting performances in the Yanta District of Xi’an City are predicted. Figure 6A shows the forecast results for 2018, revealing that under the training set and the test set, the forecast results of the model deviate greatly from the actual results. The main reason is the great fluctuation of China's meteorological situation in 2018 and the incomplete dataset, which leads to insufficient learning performance. In the meteorological intelligent forecast results in 2019 (Figure 6B), the model’s performance is consistent with the actual meteorological results in different sets, but there is a short-term fluctuation at 10-20 o'clock. Thus, the weather intelligent forecast system based on the ML algorithm performs well but relies on specific data.

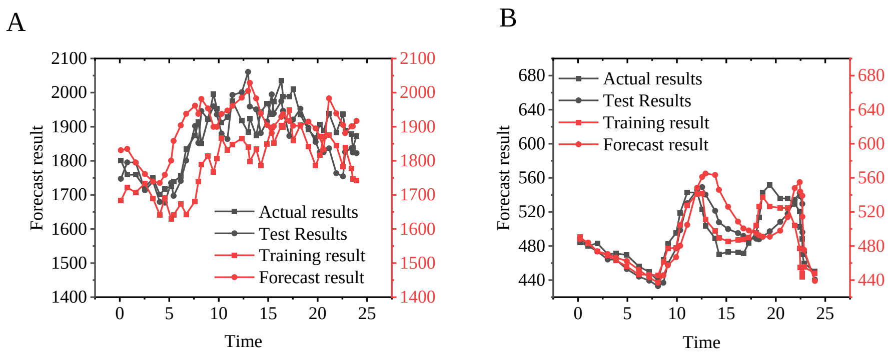

Fig. 6. Results of meteorological intelligent forecast performance

4.3 Comparative analysis of model performance

The performance of an intelligent city prediction model based on AI and ML is verified to analyze the performance results of the meteorological intelligent prediction model. The training data set is the weather forecast data of Xi’an on September 4, 2021. Figure 7 illustrates the performance results of the meteorological intelligent forecast model. The performance analysis of different methods suggests that the performance of the model is reduced by 20.35% without an AI management system, while the performance of the model without a DT algorithm is reduced by 15.63%. Without AI and DT algorithms, the model’s performance is reduced by 35.68%. The performance of the proposed model is improved by 35.26% compared with the traditional model. The performance of the latest literature models [38-41] is compared with that of the proposed model. Compared with the literature model in reference 23, the performance of the proposed model is slightly weaker, but it has a significant improvement compared with the model in other references. This further shows the effectiveness of the proposed weather intelligent forecast model, which has good forecast accuracy. The accuracy of the proposed model is not as good as that in reference 23, but the proposed model has much lower requirements for the computer configuration. Therefore, under the same equipment, the accuracy of the proposed model can exceed that in most literature. It is believed here that DL has a broader application for meteorological modeling in the future.

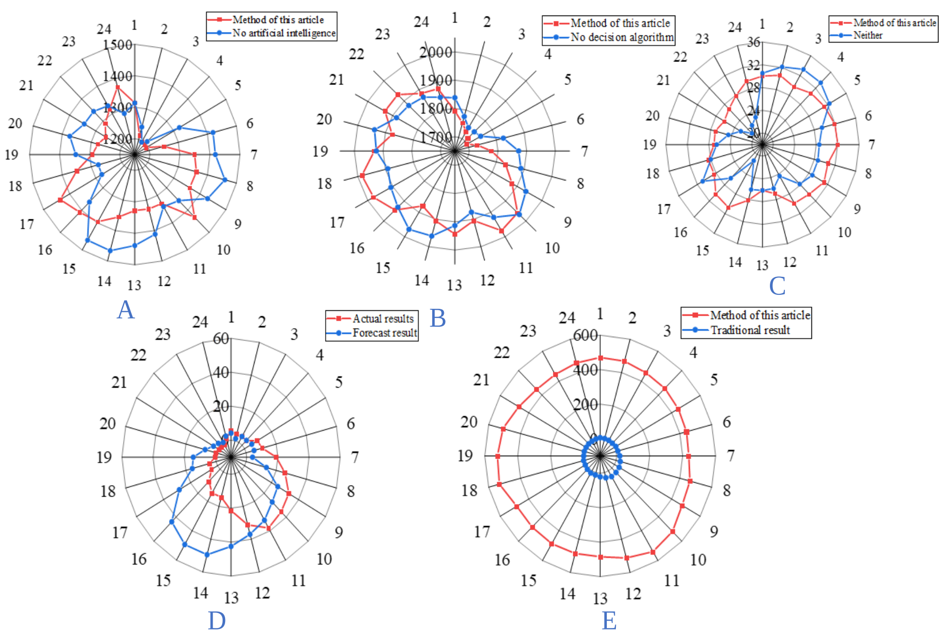

Fig. 7. Performance comparison analysis of different models((A) shows the performance comparison between the research method in this paper and the model without AI technology; (B) indicates the performance comparison between the research method in this paper and the model without DT algorithm; (C) represents the performance comparison between the research method in this paper and the model without AI technology and DT algorithm; (D) represents the comparison between the actual meteorological forecast results and the model prediction results; (E) shows the performance comparison between the research method in this paper and the traditional prediction model)

4.4 Performance Comparison of Different Models

Further, to compare the performance of the intelligent city prediction model based on AI and ML, the proposed model is compared with literature models in references 36-37 from Accuracy, Precision, Recall, and F1 score. The calculation reads:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} . \quad (5)$$

In Eq.(5), $TP + TN$ represents the number of samples predicted correctly, and $TP + TN + FP + FN$ stands for the total number of samples.

$$\Pr ecision = \frac{TP}{TP + FP} . \quad (6)$$

In Eq.(6), TP represents the number of samples correctly predicted as 1, and $TP + FP$ denotes the total number of samples predicted as 1.

$$\mathrm{Re}\,call = \frac{TP}{TP + FN} . \quad (7)$$

In Eq.(7), $TP + FN$ stands for the number of samples predicted to be 1 in the real case.

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} . \quad (8)$$

The F1 value is a harmonic average of model accuracy and recall.
The results are shown in Figure 8.

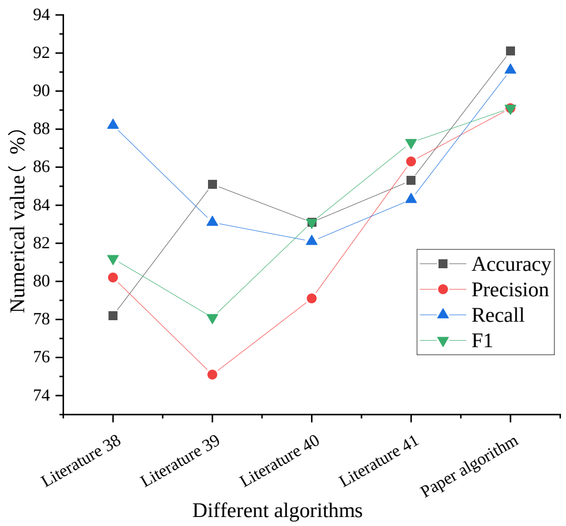

Fig. 8. Comparison of recognition accuracy of different models

Figure 8 tells that the recognition accuracy of the proposed intelligent weather forecast model has reached 92.1%, which is at least 15.1% higher than that of other models. At the same time, the Precision, Recall, and F1 score of the proposed model are also the highest, indicating that compared with other intelligent weather forecast models, the proposed model has better performance and higher prediction accuracy.

4.5 Analysis of actual forecasting performance

The meteorological forecast data of Xi'an from September 4 to September 30, 2021 is taken as the training data set to analyze the performance of the intelligent forecasting model of smart city integrated with AI and ML. Figure 9 demonstrates the forecast analysis results of the actual performance of the proposed model. It suggests that the accuracy of meteorological intelligent forecast results within one month is basically maintained at 96%, of which the maximum, minimum, and average forecast errors are 5.95%, 0.59%, and 3.76%, respectively. It proves that the proposed model has high forecast accuracy and good robustness, which can meet the urban daily weather forecast needs.

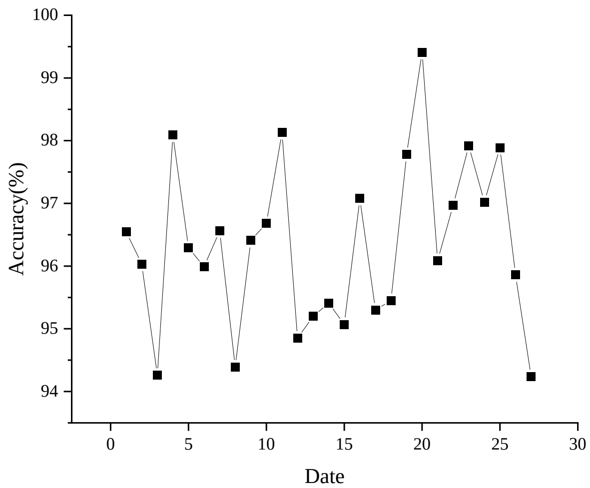

Fig. 9. Analysis results of actual forecast performance of the proposed model

Table 4. Analysis of meteorological early warning results

Thunderstorm Serial number	Lightning station Current location	Distance	Movement				Prediction accuracy
M0	143	148	0min	15min	30min	45 min	(%)
F0	132	165	2/3	44/147	NO	NO	91%
H0	145	165	NO	129/135	NO	NO	92.45%
D0	120	155	189/5	144/155	143/155	NO	91.35%
O0	95	99	235/7	118/134	115/134	NO	90.21%
P0	141	163	NO	NO	NO	NO	90.13%
MO	113	125	120/145	NO	NO	NO	92.54%

Table 4 displays the results of meteorological location early warning. The positions of seven meteorological images are analyzed, and the information of the lightning station determines the specific location information of these points. The actual data will be used to analyze the proposed model and compared with the existing data to obtain the final prediction result. From the meteorological early warning results in Table 4, the model has apparent early warning information within 0~30 minutes. After 45 minutes, the warning information disappears. Moreover, the model’s average prediction accuracy in seven images is 91%, proving the proposed model's effectiveness.

5. Discussion

The meteorological observation and early warning system study focuses on some of the leading image display core technologies in the development of GIS. The system adopts Mapinfo, MapX graphics development software, and a system application platform. Meanwhile, the enterprise user database platform based on SQL Server 2000 is designed and built. Many functions are realized, including the integrated analysis and retrieval of the monitoring attributes and spatial data of the actual lightning location data, the spatial distribution statistics of the lightning location data, and the visualization grid demonstration of charts. It facilitates the early warning and forecasting personnel of weather modification, precipitation enhancement, and hail suppression to provide services for the weather forecast and public users. Meteorological monitoring, early warning, and forecast are realized. The application of the ML algorithm improves the forecast accuracy of meteorological monitoring system data and realizes the data classification while simplifying data processing and improving the system calculation efficiency. In Reference [42], scholars also applied ML algorithms to construct intelligent weather forecasts and built an intelligent weather forecast system. The algorithm could help the relevant personnel intuitively analyze the weather changes to improve the forecast accuracy. This is consistent with the idea of this exploration.

Further, the relevant forecast factors are screened using various numerical products through the DT method. Then, the DT of the potential lightning forecast is established with the monitoring data of the lightning locator network, high altitude ground observation data, wind field data, and radar data of the Tibet Autonomous Region. The potential meteorological forecast with a spatial resolution of 1°*1° and temporal resolution of 1 hour is realized. The weather events positioning system is combined with the domestic mature satellite cloud images and weather detection system data, the occurrence and change process of meteorological phenomena are analyzed, and the temporary weather forecast is realized. In Reference [43][44], scholars also established a temporary early-warning system through an ML algorithm to help modify operations and predict weather phenomena, achieving high forecast accuracy. The results are consistent with the temporary weather early warning system proposed, proving that the proposed meteorological monitoring system is feasible, especially in the areas with imperfect meteorological monitoring systems, to improve the data processing technology of local meteorological monitoring, early warning, and forecasting products. Additionally, the proposed system can improve their fineness, visibility, evaluability, and forecast accuracy and help the weather modification operation of relevant personnel, such as precipitation enhancement and hail suppression.

6. Conclusion

Given the imperfection of meteorological monitoring, early warning, and forecasting in urban weather modification centers, targeted research is conducted here. Based on the GIS, the potential forecast model of meteorological precipitation enhancement monitoring is implemented together with a temporary early-warning system. Various data structures are analyzed in connection with the spatial data, and an intuitive and visual query analysis method is established to calculate weather formation in the spatial range. The calculation and forecast results are verified by real-time visual data regarding the thunderstorm weather. Through software development, the meteorological detection system has been put into trial operation. The main achievements of this paper are as follows.

1)After repeated operation and testing, the system is stable. Its function fully meets the design requirements of urban weather modification, precipitation enhancement, and hail suppression early warning and prediction, improving early warning and prediction in the region.

2)The GIS technology is applied to the data processing technology of lightning monitoring, early warning, and forecasting products, significantly improving the fineness, visualization, evaluability, and prediction accuracy. Mainly, the research involves the embedment of observation data of the new generation Doppler WSR, the surface precipitation data, and the early warning results into the GIS platform.

3)Under comparison analysis, the monitoring and forecasting results are more in line with the actual situation.

Still, there are some deficiencies. First, there is no perfect meteorological database. The existing data content of domestic scientific research institutions is only used for internal testing and is not fully open access. Moreover, NN or DL content can be used for effective data processing for the proposed model. In the follow-up study, these two aspects will be improved, and the model performance will be optimized.

The research contributions read: (1) The collected meteorological data are used for precipitation enhancement, hail suppression potential prediction, and meteorological early warning products. (2) A metropolitan LAN is established to transmit data and connect lightning detection stations based on regression algorithm, DT algorithm, and batch processing method. (3) Two main functional modules are evaluated from two aspects: loading a GIS electronic map and superimposing radar and cloud map and high-altitude data based on AI technology. A background database based on meteorological observation is established. It can collect and visualize GIS and query multidimensional spatial data to achieve data sampling. (4) The 24-hour meteorological monitoring technology is studied through comparison and superposition. Relevant forecast factors are screened to form potential forecast data for weather modification. The meteorological potential prediction products are built on the meteorological data analysis platform to monitor meteorological conditions, such as spatial location, distribution location, and time resolution.

Acknowledgment. The author acknowledge the help from the university colleagues.

References

1. Hu H., Lei T., Hu J., Zhang S., Kavan P. Disaster-mitigating and general innovative responses to climate disasters: Evidence from modern and historical China. International Journal of Disaster Risk Reduction, Vol. 28, 664-673. (2018)
2. Li R., Jin Y. The early-warning system based on a hybrid optimization algorithm and fuzzy synthetic evaluation model. Information Sciences, Vol. 435, 296-319. (2018)
3. Zhou W., Berrio J.S., Worrall S., Nebot E. Automated evaluation of semantic segmentation robustness for autonomous driving. IEEE Transactions on Intelligent Transportation Systems, Vol. 21, No. 5, 1951-1963. (2019)
4. Song T., Zhao H., Liu Z., Liu H., Hu Y., Sun D. Intelligent human hand gesture recognition by local–global fusing quality-aware features. Future Generation Computer Systems, Vol. 115, 298-303. (2020)
5. Marshall A., Ojiako U., Wang V., Lin F., Chipulu M. Forecasting unknown-unknowns by boosting the risk radar within the risk intelligent organization. International Journal of Forecasting, Vol. 35, No. 2, 644-658. (2019)
6. Wang B., Gu X., Ma L., Yan S. Temperature error correction based on BP neural network in the meteorological wireless sensor network. International Journal of Sensor Networks, Vol. 23, No. 4, 265-278. (2017)
7. Yogavani D., Prakash N.K. Implementation of wireless sensor network-based multi-core embedded system for the smart city. International Journal of Computer Technology and Applications (IJCTA), Vol. 10, No. 02, 119-123. (2017)
8. Chang D L, Yang S H, Hsieh S L, et al. Artificial intelligence methodologies applied to prompt pluvial flood estimation and prediction. Water, Vol. 12, No. 12, 3552. (2020)
9. Yang T, Sun F, Gentine P, et al. Evaluation and machine learning improvement of global hydrological model-based flood simulations. Environmental Research Letters, Vol. 14, No. 11, 114027. (2019)
10. Schwendicke F.a., Samek W., Krois J. Artificial intelligence in dentistry: chances and challenges. Journal of dental research, Vol. 99, No. 7, 769-774. (2020)
11. Schneider P., Walters W.P., Plowright A.T., Sieroka N., Listgarten J., Goodnow R.A., et al. Rethinking drug design in the artificial intelligence era. Nature Reviews Drug Discovery, Vol. 19, No. 5, 353-364. (2020)
12. Mohammadi A., Mohammadi M., Zahiri S.H. Design of optimal CMOS ring oscillator using an intelligent optimization tool. Soft Computing, Vol. 22, No. 24, 8151-8166. (2018)
13. Mohammadi A., Zahiri S.H., Razavi S.M. Infinite impulse response systems modeling by artificial intelligent optimization methods. Evolving Systems, Vol. 10, No. 2, 221-237. (2019)
14. Shahraki N.S., Zahiri S.H., editors. Multi-objective optimization algorithms in analog active filter design. 2020 8th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS); 1258-1261. (2020)
15. Farzaneh H., Malehmirchegini L., Bejan A., Afolabi T., Mulumba A., Daka P.P. Artificial Intelligence Evolution in Smart Buildings for Energy Efficiency. Applied Sciences, Vol. 11, No. 2, 763-773. (2021)
16. Lei Y., Yang B., Jiang X., Jia F., Li N., Nandi A.K. Applications of machine learning to machine fault diagnosis: A review and roadmap. Mechanical Systems and Signal Processing, Vol. 138, 106587-106593. (2020)
17. Dral P.O. Quantum chemistry in the age of machine learning. The journal of physical chemistry letters, Vol. 11, No. 6, 2336-2347. (2020)
18. Kirkwood C., Economou T., Odbert H., Pugeault N. A framework for probabilistic weather forecast post-processing across models and lead times using machine learning. Philosophical Transactions of the Royal Society A, Vol. 379, No. 2194, 20200099-20200106. (2021)
19. Hossain M.S., Mahmood H. Short-term photovoltaic power forecasting using an LSTM neural network and synthetic weather forecast. IEEE Access, Vol. 8, 172524-172533. (2020)

20. Xu W., Ning L., Luo Y. Wind speed forecast based on post-processing of numerical weather predictions using a gradient boosting decision tree algorithm. Atmosphere, Vol. 11, No. 7, 738-742. (2020)
21. Caron M., Misra I., Mairal J., Goyal P., Bojanowski P., Joulin A. Unsupervised learning of visual features by contrasting cluster assignments. arXiv preprint arXiv:200609882, 26-31. (2020)
22. Bao D L., Xin Y Z., Mei Z., Hui L., Guang-Qian L. Improved genetic algorithm-based research on optimization of least square support vector machines: an application of load forecasting. Soft Computing: 1-9. (2021)
23. Memiş S., Enginoğlu S., Erkan U. Numerical Data Classification via Distance-Based Similarity Measures of Fuzzy Parameterized Fuzzy Soft Matrices. IEEE Access, Vol. 9, 88583-88601. (2021)
24. Tian C., Fei L., Zheng W., Xu Y., Zuo W., Lin C.W. Deep learning on image denoising: An overview. Neural Networks, 124-131. (2020)
25. Bahri Y., Kadmon J., Pennington J., Schoenholz S.S., Sohl-Dickstein J., Ganguli S. Statistical mechanics of deep learning. Annual Review of Condensed Matter Physics, 224-233. (2020)
26. Minaee S., Boykov Y.Y., Porikli F., Plaza A.J., Kehtarnavaz N., Terzopoulos D. Image segmentation using deep learning: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 115-163. (2021)
27. Yin X., Hu Z., Zheng J., Li B., Zuo Y. Study on Radar Echo-Filling in an Occlusion Area by a Deep Learning Algorithm. Remote Sensing, Vol. 13, No.9, 1779-1783. (2021)
28. Huang C.J., Shen Y., Chen Y.H., Chen H.C. A novel hybrid deep neural network model for short-term electricity price forecasting. International Journal of Energy Research, Vol. 45, No. 2, 2511-2532. (2021)
29. Xia M., Tian N., Zhang Y., Xu Y., Zhang X. Dilated multi-scale cascade forest for satellite image classification. International Journal of Remote Sensing, Vol. 41, No. 20, 7779-800. (2020)
30. Guo Y., Cao X., Liu B., Gao M. Cloud Detection for Satellite Imagery Using Attention-Based U-Net Convolutional Neural Network. Symmetry, Vol. 12, No. 6, 1056-1063. (2020)
31. Khalyasmaa A.I., Eroshenko S.A., Tashchilin V.A., Ramachandran H., Piepur Chakravarthi T., Butusov D.N. Industry Experience of Developing Day-Ahead Photovoltaic Plant Forecasting System Based on Machine Learning. Remote Sensing, Vol. 12, No. 20, 3420-3426. (2020)
32. Benkercha R., Moulahoum S. Fault detection and diagnosis based on C4. 5 decision tree algorithm for the grid-connected PV system. Solar Energy, Vol. 173, 610-634. (2018)
33. Ali E.M., Ahmed M.M., Wulff S.S. Detection of critical safety events on freeways in clear and rainy weather using SHRP2 naturalistic driving data: Parametric and non-parametric techniques. Safety Science, Vol. 119, 141-149. (2019)
34. Landeras G., Bekoe E., Ampofo J., Logah F., Diop M., Cisse M., et al. New alternatives for reference evapotranspiration estimation in West Africa using limited weather data and ancillary data supply strategies. Theoretical and Applied Climatology, Vol. 132, No. 3, 701-716. (2018)
35. Bu J., Peng D.D., Xiao H., Yue Q., Han Y., Lin Y., et al. Analysis of meteorological conditions and prediction of epidemic trend of 2019-nCoV infection in 2020. MedRxiv, 1211-1234. (2020)
36. Işık E., Inallı M. Artificial neural networks and adaptive neuro-fuzzy inference systems approaches to forecast the meteorological data for HVAC: The case of cities for Turkey. Energy, Vol. 154, 7-16. (2018)
37. Fuente A., Meruane V., Meruane C. Hydrological early warning system based on a deep learning runoff model coupled with a meteorological forecast. Water, Vol. 11, No. 9, 1808. (2019)

38. Kang H., Hong T., Lee M. A new approach for developing a hybrid sun-tracking method of the intelligent photovoltaic blinds considering the weather condition using data mining technique. Energy and Buildings, Vol. 209, 109708-109712. (2020)
39. Li C., Xu P. Application on traffic flow prediction of machine learning in intelligent transportation. Neural Computing and Applications, 1-12. (2020)
40. Kuhn P., Wirtz M., Wilbert S., Bosch J., Wang G., Ramirez L., et al. Field validation and benchmarking of a cloud shadow speed sensor. Solar Energy, Vol. 173, 229-45. (2018)
41. Narote S.P., Bhujbal P.N., Narote A.S., Dhane D.M. A review of recent advances in lane detection and departure warning system. Pattern Recognition, Vol. 73, 216-34. (2018)
42. Pallonetto F., De Rosa M., Milano F., Finn D.P. Demand response algorithms for smart-grid ready residential buildings using machine learning models. Applied Energy, Vol. 239, 1265-82. (2019)
43. Barton C., Chettipally U., Zhou Y., Jiang Z., Lynn-Palevsky A., Le S., et al. Evaluation of a machine learning algorithm for up to 48-hour advance prediction of sepsis using six vital signs. Computers in biology and medicine, Vol. 109, 79-84. (2019)
44. Shen, C.-w., Min, C., & Wang, C.-c. (2019). Analyzing the trend of O2O commerce by bilingual text mining on social media. Computers in Human Behavior, Vol. 101, 474-483. https://doi.org/10.1016/j.chb.2018.09.031

Appendix.

Table 3. Variable information

Variables	Definition	Calculation
$p(i\|t)$	It indicates the proportion of samples with node number t belonging to category i.	It is calculated by Eq. (1).
s	Number of branches	Fixed value
t_R	Number of R branches	Fixed value
t_L	Number of L branches	Fixed value
P_R	The proportion of sample tR in the sample with node t.	It is calculated by Eq. (2).
P_L	The proportion of sample tL in the sample with node t.	It is calculated by Eq. (2).
β	Threshold size	Fixed value

Table 4. English abbreviations

Abbreviations	Full name	Meaning
CMOS	Complementary Metal Oxide Semiconductor	It refers to a technology used to manufacture large-scale integrated circuit (IC) chips or chips made with this technology. It is a read-write RAM chip on the computer mainboard.
RO	Reverse Osmosis	Osmosis is adopted to put clear water and saltwater in one tube. The middle is separated by a semi-permeable membrane, allowing water to pass through. Then, water flows from the place with low osmotic pressure to the place with high osmotic pressure.
AIO	All In One	It supports synchronous and asynchronous (callback-based) processing.
EC	Erasure Coding	When the data block or check block is lost or damaged, the system can be recovered according to the EC algorithm to protect data.
SIFT	Scale-invariant feature transform	It can detect and describe the local features in the image. It looks for the extreme points in the spatial scale and extracts their position, scale, and rotation invariants.
IoS	Indicator of Success	It is mainly the latest indicator to evaluate the difference between the existing and optimized models.
DNA	Deoxyribonucleic acid	It is a molecule with a double-stranded structure and is made up of deoxyribonucleotides.
SQL	Structured Query Language	It is a special-purpose programming language for managing relational database management systems.
DDR3	Double-Data-Rate Three Synchronous Dynamic Random Access Memory	It aims at the next-generation memory technology of Intel's new chip.
AI	Artificial Intelligence	It refers to the system and machine that can imitate human intelligence to perform tasks and improve itself iteratively based on the collected information.

Zeng-yuan Chao was born in Haidong, Qinghai,P.R. China, in 1978. He received the bachelor‘s degree from Chengdu University of Information Technology, P.R. China. Now, he works in Weather Modification Center of Shijiazhuang Meteorological Bureau.

His research interest include intelligent equipment , cloud precipitation and weather modification. E-mail: czy93420@163.com

Received: July 17, 2022; Accepted: November 10, 2022.