ISSN 2959-6157

Exploring the Potential of Deep Learning Models in Weather Prediction: Case Studies on Sandy Weather, Wind Speed, and Rainfall

Jiawei Cao[1, *]

[1]Keystone Academy, Beijing, China
*Corresponding author: jiawei.cao@student.keystoneacademy.cn

Abstract:

Weather prediction is one of the major focuses of today's research, and researchers have found that traditional prediction models are limited in the ability of weather prediction, while deep learning has a relatively stronger performance in weather prediction, however, there are still blanks in the research on the accuracy and applicability of the models. Therefore, this study aims to explore the application of different deep-learning models in weather prediction and make recommendations accordingly. In this paper, deep learning models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory Networks (LSTMs) will be investigated for case studies in the prediction of three different kinds of weather, namely, sandy and dusty weather, wind speed, and rainfall, respectively, along with a summary of the strengths and weaknesses of the different models. The results show that deep learning models can accurately predict future weather conditions by summarizing and analyzing the temporal and spatial characteristics of the data, while the combination of different models can further improve the model performance according to their advantages. In summary, this study provides new ideas for the further development of weather prediction and provides an important reference for future related research.

Keywords: Weather prediction; deep learning; convolutional neural networks; recurrent neural networks; long short-term memory networks.

1. Introduction

Weather forecasting has become an essential part of human life and has a profound impact on daily manufacturing and living, while weather prediction, as a prerequisite for weather forecasting, has been an important research topic in the field of meteorology. However, the commonly used traditional weather prediction methods in this field usually produce unsatisfactory predictions with relatively low accuracy and precision, and predicting extreme weather is even more challenging. For example, rainfall can have a serious impact on crop growth, daily human activities, etc., but the complex formation process of rainfall weather makes it difficult for traditional models to accurately capture its patterns. The emergence of deep learning techniques, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), long short-term memory networks (LSTMs), and other models has brought new hope for improving the accuracy of weather prediction. At present, deep learning has been widely used in speech processing, image classification, text processing, and other fields with its strong data processing and feature extraction capabilities. In the field of weather prediction, deep learning can learn a large amount of existing data, summarize the features and laws, analyze the spatial and temporal characteristics of the data, and predict future weather conditions. For example, Zhang Weiming et al. successfully predicted the sand and dust weather in Inner Mongolia by using historical meteorological data using two deep learning models, long and short-term memory networks and convolutional neural networks [1]. Xia Houjie et al. similarly utilized a convolutional neural network and a long short-term memory network to construct a short-term heavy precipitation DCDL forecasting model, which improved the accuracy and hit rate in the validation [2]. Therefore, deep learning techniques can improve the accuracy of weather prediction, and these existing research results provide valuable experience for the application of deep learning in weather prediction. Among the existing research, Dong Runting et al. organized and analyzed the weather prediction applications combining spatio-temporal sequence prediction and numerical weather prediction models based on physical theories [3]. Zhou Kanghui et al. investigated the problems demonstrated by machine learning in strong convective monitoring and forecasting, such as sample imbalance and high requirements on data features and labeling [4]. However, existing studies have analyzed less about the applications of two deep learning models, CNNs and RNNs. Therefore, this paper aims to explore the application of CNNs and RNNs in weather prediction, and through the summary and analysis of the existing research, and analyzing the advantages and disadvantages of different models, provide the direction to build deep learning weather prediction models with a more reliable prediction approach.

2. Deep Learning Related Technologies

2.1 Convolutional Neural Networks (CNNs)

CNNs is a computational model that mimics the visual and nervous system functions of a biological brain and is widely used in image recognition, video recognition, language processing, and other application scenarios. CNNs are composed of three main parts: the convolutional layer, the pooling layer, and the fully connected layer. In the convolutional layer, the convolutional kernel will process each region of the input data in turn (the size of each processing region is decided by the user), and extract the data features of the local region by performing certain matrix operations between the specified region and the original data, and finally generate the feature map of the region. The pooling layer, on the other hand, slices the data, retains the most significant features, reduces the spatial size of the feature map, and reduces the model complexity. The pooling layer process can effectively reduce the amount of data and computation. The two commonly used pooling methods are maximum pooling and average pooling. Finally, in the fully connected layer, the features are summarized and the results are output by receiving the processed features and performing operations such as weight matrix, linear transformation, and activation function [5]. The overall workflow of convolutional neural network is to extract the features through multiple convolutional and pooling layers and finally output them through a fully connected layer as shown in Fig. 1.

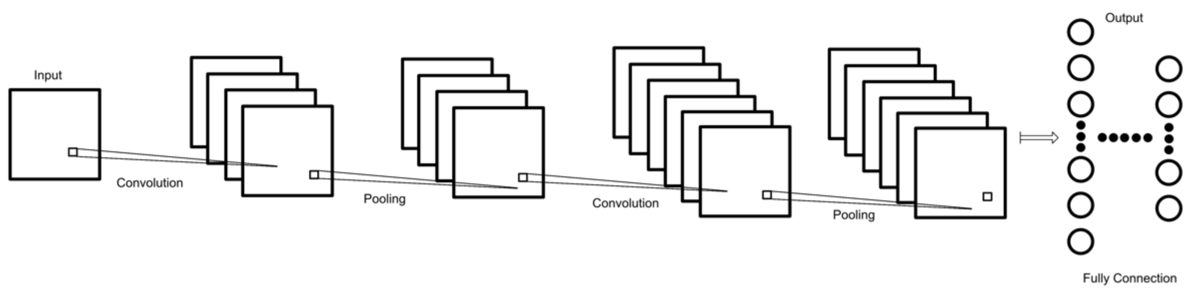

Fig. 1 Convolutional neural network workflow

2.2 Recurrent Neural Networks (RNNs)

RNNs are a special type of neural network model used to process data with sequential regularity, with common application scenarios such as language translation, speech recognition, etc. RNNs can create recurrent connections to allow data to be passed between different processing steps and capture the dependencies between data before and after the sequence. RNNs are mainly composed of an input layer, a hidden layer, and an output layer. The input layer receives the input sequence data and the individual elements and features in the sequence data are represented as a vector. In a rRNN, the hidden layer is a recurrent workflow, in each loop, the hidden layer receives the current vector of the input layer, processes its features, receives the hidden state from the previous loop, merges the two inputs, performs a linear combination, and outputs it through an activation function as a new hidden state. This step will be repeated until all the sequential data has been processed. By looping, the recurrent neural network can process sequential data with regularity and relationships. The output layer then receives the last hidden state and generates the final features and predictions [6].

2.3 Long Short-Term Memory Networks (LSTMs)

LSTMs belong to a class of variants of recurrent memory networks that are capable of processing sequential data with longer intervals or with long-term dependencies. Based on RNNS, LSTMs contain three gating units, forgetting gate, input gate, and output gate. The forgetting gate is used to determine the information that needs to be forgotten in the current time step, and the input gate is used to determine the information that needs to be saved in the current time step. Both the forgetting gate and input gate generate a value between 0 and 1 as a degree to perform the operation through a sigmoid function. The output gate is then used to determine whether the content of the current time step needs to be output or not [7].

3. Deep Learning Applications in Weather Prediction

3.1 Application of Deep Learning to Sand and Dust Weather Prediction

The use of two deep learning models, long short-term memory networks and convolutional neural networks, is an effective and commonly used predictive tool in sand and dust prediction. In the paper "Application Research of Recurrent Neural Network in Sand-Dust Storm Forecast in Inner Mongolia" by Zhang Weiming, the authors utilized a deep learning model, the long short-term memory network, to predict sand and dust weather in Inner Mongolia, and investigate the application of long short-term memory networks in the field of sand and dust prediction. This study was conducted by processing basic historical meteorological attributes (e.g., dust storm start/end time, magnitude, temperature, precipitation, barometric pressure, etc.), and processing imbalanced data by using downsampling and SMOTE algorithms. The authors used a neural network model structure with two layers of LSTM units and added a Batch-Normalization layer after the fully connected layer to enhance the generalization ability of the model. The model outputs the latest state after processing each input and finally outputs the result of the last step as a prediction of sand and dust weather. During the training process, the number of training rounds of the model was set to 20 through testing to achieve the highest accuracy and avoid overfitting. The model has a strong prediction ability, with a recall of more than 0.8 in predicting extraordinarily strong sand and dust storms and strong sand and dust storms, floating sand and dust and no dust storms, and an overall precision of more than 0.7, which verifies the effectiveness of LSTM neural networks in sand and dust weather prediction [1]. In "Spatiotemporal feature - based GCN - LSTM model for predicting sand - dust weather in northwest China" by Su Jia, Li Gaoya, and Zhang Xinsheng, the authors built a sand-dust weather prediction model by extracting features of cities in northwest China using graph convolutional network and long short-term memory network. The study used air quality data, climate data and geographic data, and processed the data through linear functions. The authors combined graph convolutional network and long short-term memory network in the model. Graph convolutional network was used to extract spatial features and construct neighbor relationships between nodes. The long and short-term memory network is then used to process the time series data and capture the relationship between the time series. Finally, the spatial and temporal features are integrated and output through a fully connected layer, and the model is used for sand and dust weather prediction. At last, the model was optimized by adjusting the parameters such as the number of layers and the number of neurons, setting the number of hidden layers to 2, the number of neurons in the graph convolution network to 128, and the number of neurons in the long and short-term memory network to 32 as optimal. The model is superior in predicting sand and dust weather, with a recall of 0.81 and a ROC-AUC of 0.92, which are the highest among the tested and compared models, and overall can correctly predict 70% of sand and dust weather [8]. Compared with Zhang Weiming's research model, this model is more superior in prediction results due to the consideration of spatial features based on the traditional recurrent neural network and the optimization of more model parameters. In summary, both long short-term memory networks and convolutional neural networks are capable of predicting dusty weather and producing more accurate data results. When building the model, according to the characteristics of the data, judging the spatial and temporal features, fusing multiple neural networks and adjusting the parameters in the test can make the model more effective and more accurate in predicting the sand and dust weather.

3.2 Application of Deep Learning to Wind Speed Prediction

Deep learning models such as convolutional neural networks and bidirectional long short-term memory networks usually have strong modeling and generalization capabilities, and play an important role in wind speed prediction application area. For example, in the study "A short-term wind speed prediction method utilizing novel hybrid deep learning algorithms to correct numerical weather forecasting" by Yan Han et al. weather forecasting" by Yan Han et al., the authors used a hybrid model, PCC-CEEMDAN-CNN-BLSTM-AMGS, within this model, the authors used two deep learning models, Convolutional Neural Network (CNN), Bidirectional Long Short-Term Memory Network (BLSTM), along with the use of other data processing methods to process and analyze the meteorological variables, such as wind speed, wind direction, pressure, temperature, relative humidity etc. After data preprocessing, a convolutional neural network is used to extract the spatial feature data generated by the CEEMDAN algorithm and pass it to a bidirectional long short-term memory network, which in turn captures the correlation of temporal features and further improves the accuracy of the wind speed prediction. After this, the model performance is further tuned by back propagation algorithm, hyperparameters and other methods. The model can predict wind speed effectively, and the prediction accuracy of the model is higher than that of the baseline model as judged by the evaluation metrics of mean square error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE) in the testing phase [9]. Therefore, with strong spatial analysis ability of convolutional neural network, it can effectively extract and capture the features of wind speed data, and with strong temporal analysis ability of bi-directional long short-term memory network, it can model and analyze the temporal data, and the combination of both of them can increase the accuracy and validity of the wind speed prediction model.

3.3 Application of Deep Learning to Rainfall Prediction

The accuracy of precipitation prediction models based on deep learning techniques has significant advantages. In the study "Observation Based Deep Learning Model for Short-duration Heavy Rain Nowcasting" published by Xia Houjie et al., the authors utilized a convolutional neural network and a long and short-term memory network to construct a short-term heavy precipitation DCDL forecasting model. This study used both ground meteorological data such as temperature and barometric pressure and radar reflectivity picture data. In this study, the convolutional neural network processes the spatial features of the radar time series data and the ground meteorological data through convolution and pooling operations, after which it is passed to the long and short-term memory network to process the time series information and capture the dependencies in the time series data. Finally, the result of precipitation probability is obtained by fully connected neural network with activation function. In the result analysis section, the authors compared two existing models, the SDL model and the CMA-SH3 model, with the DCDL forecast model, and the results show that the hit rate of the DCDL forecast model designed by the authors is improved by about 1%-15% compared with the other existing models, which proves that the use of simultaneous processing of spatial and temporal features in predicting the short-term precipitation can improve the accuracy of the model [2]. In a study by Chen Sheng et al. titled "Short-Term Precipitation Nowcasting Based on Multi-Scale Feature Deep Learning," the authors employed the U-Net deep learning architecture and convolutional neural networks to construct a model for forecasting short-term precipitation using radar data. Firstly, in this model, the authors built a "multi-scale feature fusion module," which combines multidimensional convolution kernels with depth wise separable convolution. That is, the incoming radar echo data undergo deep convolution, followed by multidimensional convolution on four sizes of kernels: 1×1, 3×3, 5×5, and 7×7. This process extracts features from the incoming data, and then the U-Net architecture is employed for image segmentation and feature extraction. The model can effectively reduce the computational cost and improve the feature extraction ability. During the model testing, the forecast results with POD and CSI scores are higher than the existing model by using three precipitation intensity conditions of 5 mm/h, 10 mm/h, and 25 mm/h, which proves that the form of combining multi-dimensional convolution and deep convolution can avoid feature loss and have stronger prediction ability [10]. In a study by Zhang Chi et al. titled "Using Deep Learning to Predict Daily Precipitation Distribution in the Northeast of the United States," the authors compared the weather prediction capabilities of three existing deep learning models: VGG, ResNet, and GoogleNet [11]. All three models are deep convolutional neural network architectures, VGG (Visual Geometry Group) is characterized by deeper model depth, improving the prediction performance of the model, ResNet (Residual Network) is characterized by the use of residual learning mechanism to increase the number of layers of the neural network, GoogleNet is characterized by the use of the Inception module, enhancing the ability to extract information at different scales and improving the performance. The study demonstrated that all three models can effectively forecast weak precipitation events. Among the RMSE, TS, and ETS prediction scores, VGG showed slightly lower prediction errors. However, all three models exhibited poor performance in predicting the timing of heavy precipitation events, indicating that the ability of convolutional neural networks to predict precipitation is relatively weaker when used individually. In the study "Deep Learning for Improving Numerical Weather Prediction of Heavy Rainfall" by Philipp Hess and Niklas Boers, a deep learning model based on the U-Net convolutional neural network architecture was employed for rainfall prediction. The research utilized atmospheric variable data and satellite observation data from the European Centre for Medium-Range Weather Forecasts. The data were partitioned based on time and space, and being separated into training, validation, and test sets for model evaluation. In this study, atmospheric variables and satellite observations are used as inputs to the U-Net model, which is combined into an input image containing 12 channels, and features are extracted through convolution and pooling, and finally decoded for output. In the assessment of the continuous prediction capability of the model, the authors' developed model shows a significant improvement in the prediction performance and is more advantageous in predicting heavy rainfall weather. In predicting heavy rainfall weather, the model showed a significant improvement in scores in predicting events above 90th and 99th percentile [12]. It demonstrates the effectiveness and superiority of U-Net convolutional neural network structure in rainfall prediction. In the paper "Rainfall Prediction Using Machine Learning & Deep Learning Techniques" by CMAK Zeelan Basha et al., the authors utilized multilayer perceptron (MLP) and Auto-Encoder neural networks to develop rainfall prediction models. They evaluated the performance of the models using root mean square error (RMSE) and mean square error (MSE) [13]. The model can predict rainfall weather to a certain extent, but there is still room for improvement. Therefore, deep learning techniques can effectively predict precipitation weather, and deep learning architectures such as convolutional neural networks and long short-term memory networks usually have better prediction capabilities and perform well, especially in handling multiple data with capturing spatial-temporal features.

4. Suggestions

Deep learning models such as convolutional neural networks, long short-term memory networks, and other deep learning models have excellent performance in the application scenario of weather prediction, possessing a certain degree of stability and reliability. When using deep learning technology to predict the weather, since deep learning models can analyze the spatial and temporal features of data comprehensively, the importance of spatial and temporal features should be considered when selecting data, and a suitable model architecture should be selected for each type of data, for instance, convolutional neural networks excel in handling spatially structured data, while long short-term memory networks demonstrate superior performance in processing time series data. By integrating the characteristics of multiple models and using them in combination, more accurate and reliable results can be obtained. Additionally, when selecting data, it's essential to choose highly correlated data and preprocess it appropriately, such as downsampling and data balancing, to enhance the model's predictive capability for complex weather conditions.

5. Conclusion

This study found that deep learning models, including CNNs, RNNs, and LSTMs, demonstrated high predictive capabilities in the application of weather forecasting. These deep learning models can accurately capture both temporal and spatial patterns in the data, extract features, and make precise predictions about future weather conditions. The complex weather prediction capabilities of deep learning models were particularly strong, highlighting the advantages of deep learning in weather forecasting. When predicting different weather phenomena, it is usually necessary to use different deep learning models according to the formation characteristics of the weather phenomena and the characteristics of the available data, for example, using CNNs to process spatial data, and utilizing LSTMs to process sequential data. For example, predicting sand and dust weather using a combination of LSTMs and CNNs can effectively improve the prediction accuracy. Therefore, adjusting model parameters, combining different model and data processing methods based on data characteristics, as well as combining deep learning models with traditional meteorological forecasting methods are effective approaches to improving the comprehensive capabilities of weather forecasting. By summarizing the application of deep learning in the field of weather forecasting, this study provides new methods and insights to enhance the accuracy and precision of weather predictions. It lays the foundation for the future development of weather forecasting technology and promotes technological advancements in this field. Finally, there is a relative lack of comprehensive comparison and evaluation of different models in the same task or under the same assessment conditions in this study; therefore, the evaluation of the strengths and weaknesses of different models is not completely accurate, and in the future, the same assessment criteria can be used to make comprehensive comparisons and summarize the patterns and results of different models.

References

[1] Zhang Weiming. Application Research of Recurrent Neural Network in Sand-Dust Storm Forecast in Inner Mongolia. Inner Mongolia University of Technology, 2020.

[2] Xia Houjie, Xiao An, Nie Daoyang. Observation Based Deep Learning Model for Short-duration Heavy Rain Nowcasting. Plateau Meteorology, 42(4): 1005-1017.

[3] Dong Runting, Wu Li, Wang Xiaoying, et al. Review of application analysis and research progress of deep learning in weather forecasting. Journal of Computer Applications, 2023, 43(6): 1958-1968.

[4] Zhou Kanghui, Zheng Yongguang, Han Lei, Dong Wansheng. Advances in Application of Machine Learning to Severe Convective Weather Monitoring and Forecasting. Meteor Mon, 47(3): 274-289.

[5] O'Shea Keiron, Nash Ryan. An Introduction to Convolutional Neural Networks. CoRR, 2015.

[6] Pascanu Razvan, et al. How to Construct Deep Recurrent Neural Networks. CoRR, 2014.

[7] Ghojogh, Benyamin, Ali Ghodsi. Recurrent Neural Networks and Long Short-Term Memory Networks: Tutorial and Survey. ArXiv.org, 2023.

[8] Su Jia, Li Gaoya, Zhang Xinsheng. Spatiotemporal feature-

based GCN-LSTM model for predicting sand-dust weather in northwest China. Journal of Arid Land Resources and Environment, 2024, 38(5): 111-120.

[9] Han Yan, et al. A Short-Term Wind Speed Prediction Method Utilizing Novel Hybrid Deep Learning Algorithms to Correct Numerical Weather Forecasting. Applied Energy, 2022, 312:118777–118777.

[10] Chen Sheng, Huang Qiqiao, Tan Jinkai, Liang Zhenqing, Wu Chong. Short-Term Precipitation Nowcasting Based on Multi-Scale Feature Deep Learning. Journal of Tropical Meteorology, 2023, 39(6): 799-806.

[11] Zhang Chi, Chen Guoxing, Yang Hongtao. Using Deep Learning to Predict Daily Precipitation Distribution in the Northeast of the United States. Transactions of Atmospheric Sciences, 2024, 47(01): 55-64.

[12] Hess, Philipp, Niklas Boers. Deep Learning for Improving Numerical Weather Prediction of Heavy Rainfall. Journal of Advances in Modeling Earth Systems, 2022, 14(3).

[13] Cmak Zeelan Basha, Nagulla Bhavana, Ponduru Bhavya, Sowmya V. Rainfall Prediction using Machine Learning & Deep Learning Techniques. International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, 92-97.