Commun. Comput. Phys. Vol. **31**, No. 1, pp. 131-153
doi: 10.4208/cicp.OA-2020-0158 January 2022

A Deep Spatio-Temporal Forecasting Model for Multi-Site Weather Prediction Post-Processing

Wenjia Kong[1], Haochen Li[3], Chen Yu[2], Jiangjiang Xia[4], Yanyan Kang[5] and Pingwen Zhang[2,*]

[1] *Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China.*
[2] *School of Mathematical Sciences, Peking University, Beijing 100871, China.*
[3] *School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China.*
[4] *Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China.*
[5] *Beijing Weather Forecast Center, Beijing 100089, China.*

Received 10 August 2020; Accepted (in revised version) 3 August 2021

Abstract. In this paper, we propose a deep spatio-temporal forecasting model (DeepSTF) for multi-site weather prediction post-processing by using both temporal and spatial information. In our proposed framework, the spatio-temporal information is modeled by a CNN (convolutional neural network) module and an encoder-decoder structure with the attention mechanism. The novelty of our work lies in that our model takes full account of temporal and spatial characteristics and obtain forecasts of multiple meteorological stations simultaneously by using the same framework. We apply the DeepSTF model to short-term weather prediction at 226 meteorological stations in Beijing. It significantly improves the short-term forecasts compared to other widely-used benchmark models including the Model Output Statistics method. In order to evaluate the uncertainty of the model parameters, we estimate the confidence intervals by bootstrapping. The results show that the prediction accuracy of the DeepSTF model has strong stability. Finally, we evaluate the impact of seasonal changes and topographical differences on the accuracy of the model predictions. The results indicate that our proposed model has high prediction accuracy.

AMS subject classifications: 62P12, 86A10, 93B15, 97M10

Key words: Weather forecasting, post-processing, spatio-temporal modeling, deep learning.

*Corresponding author. *Email addresses:* `pzhang@pku.edu.cn` (P. Zhang), `wjkong@pku.edu.cn` (W. Kong), `lihaochen_bjut@sina.com` (H. Li), `yuchen1995@pku.edu.cn` (C. Yu), `xiajj@tea.ac.cn` (J. Xia), `yanyan051022@163.com` (Y. Kang)

1 Introduction

Weather forecasting has always been a matter of general concern. Accurate weather forecasts can reduce the adverse effects caused by extreme weather, reduce economic losses, and have an important impact on various industries such as tourism and transportation. Nowadays, as supercomputers gradually enter a period of rapid development, numerical weather prediction (NWP) has become a major technical method and research direction in the field of weather forecasting. The idea of numerical weather prediction was first proposed by Bjerknes [3] in the early 20th century and achieved rapid development afterwards [6, 32]. Nevertheless, NWP forecasts often carry significant systematic bias. Hence, post-processing has become standard practice since at least Glahn et al. (1972) [19], in which Glahn et al. demonstrated a version of model output statistics (MOS) that improves the raw NWP forecast accuracy. The MOS method has been widely used since it was proposed [1, 10, 18, 21, 39, 44]. Besides MOS, other statistical algorithms are commonly adopted in post-processing of weather prediction, such as Kalman filtering [13, 14, 28, 43], the analog ensemble [12], anomaly numerical-correction with observations [29] and Markov Chain models [5, 37].

In addition to traditional statistical methods, machine learning and artificial neural networks [25] are gradually being widely used in weather forecasting [4, 7, 20, 31, 40, 41, 45, 46]. Haochen Li et al. [26] proposed a model output machine learning (MOML) method for grid temperature forecasting. The results showed a better performance than the ECMWF (European Centre for Medium-range Weather Forecasts) model without post-processing and the traditional post-processing methods MOS, especially for winter. Huan Zheng et al. [48] used k-means algorithm to divide the samples into several categories based on the similarity of weather in historical days and proposed a new extreme gradient boosting (XGBoost) model for short-term wind power forecasting. Rasp et al. [30] established a fully-connected neural network to predict the 2-m temperature in Germany. The results showed that the neural network approach significantly outperforms traditional statistical methods. Zaytar et al. [47] established the seq2seq model based on LSTM (Long Short-Term Memory), and made end-to-end [25] predictions on the temperature, humidity, and wind speed of 9 cities in Morocco. The experimental results were better than traditional statistical methods. In the prediction of extreme weather forecasts, Ashesh Chattopadhyay et al. [8] proposed CapsNets to predict the geographic area of extreme surface temperature in North America. The results show that the multivariate data-driven framework is expected to achieve accurate extreme weather predictions.

In the field of weather forecasting, it is often necessary to consider the impact of both temporal and spatial characteristics. Xingjian Shi et al. [33] proposed the convolutional LSTM (ConvLSTM) and used it to build an end-to-end trainable model for the precipitation nowcasting problem. Experiments showed that the ConvLSTM network can capture spatio-temporal correlations. Due to the shortcoming of ConvLSTM to model the dynamic changes of clouds, Xingjian Shi introduced a new precipitation prediction model named TrajGRU (Trajectory Gate Recurrent Unit) that can actively learn the location-