URSI GASS 2021, Rome, Italy, 28 August - 4 September 2021

The More, the Better? Forecasting Lightning from Multi-Source Meteorological Data via Deep Neural Networks

Yangli-ao Geng(1), Dong Zheng(2), Qingyong Li*(1), Wen Yao(2), Tianyang Lin(1), Liangtao Xu(2), Weitao Lyu(2), Yijun Zhang(3)

(1) Beijing Key Lab of Traffic Data Analysis and Mining, Beijing Jiaotong University, Beijing, China

(2) State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing, China

(3) Department of Atmospheric and Oceanic Sciences, Fudan University, Shanghai, China

Lightning forecasting usually requires comprehensive analysis from a variety of meteorological data. Recent decades have witnessed the advance of weather observation and simulation technologies, triggering an explosion of meteorological data which are collected from multiple sources (e.g. radar, automatic stations and numerical weather prediction) and usually characterized by SpatioTemporal (ST) structure. As a result, how to adequately exploit these multi-source ST data emerges as a promising but challenging topic for Lightning forecasting. To address this issue, we propose a data-driven lightning forecasting framework (referred to as LightNet+) based on deep neural networks DNNs. Our framework design enables LightNet+ to make forecasts by mining complementary information distributed in multiple data sources, which may be heterogeneous in spatial (continuous vs. discrete) and temporal domains (observation for the past vs. simulation for the future).

More specifically, our data consist of three sources: weather simulation data by the Weather Research and Forecasting (WRF) model, LIGhtning observation (LIG) data and Automatic Weather Station (AWS) data. In each forecasting case, as illustrated in Figure 1a, we use the WRF simulation data for the next six hours ($\mathcal{W}$), the LIG data ($\mathcal{L}$) and the AWS data ($\mathcal{A}$) for the past three hours as inputs, and aim to forecast lighting for the next six hours ($\hat{\mathcal{L}}$), with a spatial resolution of 4km$\times$4km. Among these inputs, LIG and AWS can provide accurate location information of recent thunderstorms, while WRF can display relatively reliable trend information of thunderstorms for long-term forecasting. They demonstrate complementary advantages for lightning forecasting.

As shown in Figure 1b, LightNet+ consists of four major modules: WRF encoder, historical lightning (LIG) encoder, AWS encoder and forecast (FORE) decoder. Among them, the three encoders aim to extract forecast-related information from $\mathcal{W}$, $\mathcal{L}$ and $\mathcal{A}$ respectively; FORE decoder seeks to "decode" lightning forecasts $\hat{\mathcal{L}}$ from the information extracted by the encoders. All the four modules are constructed based on a special DNN model named convolution long-short term memory [1], which is well-known for handling ST data.

Our experimental results demonstrate: (i) LightNet+ achieves about three times improvement in equitable threat score compared with three established lightning schemes; (ii) the more data sources are fed into LightNet+, the higher forecasting quality it achieves. Moreover, LightNet+ is flexible for the number of input data sources and can be extended to other weather forecasting tasks confronting the challenge of multi-source ST data.

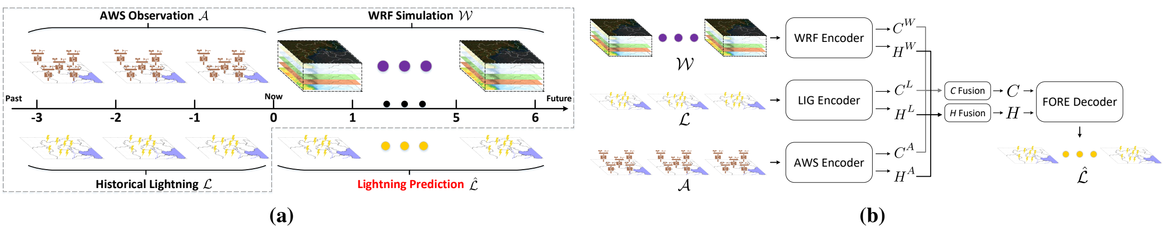

Figure 1. (a) The spatiotemporal strcuture of data in one forecasting case. (b) The architecture of LightNet+, which receives $\mathcal{W}$, $\mathcal{L}$ and $\mathcal{A}$ as input and output $\hat{\mathcal{L}}$.

References

[1] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-Chun Woo. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In *Proceedings of Advances in Neural Information Processing Systems*, pages 802–810, 2015.