EGU2020-22302
https://doi.org/10.5194/egusphere-egu2020-22302
EGU General Assembly 2020

Nowcasting Lightning Occurrence Using Machine Learning Techniques: The Challenge of Identifying Outliers

Amirhossein Mostajabi[1], Declan Finney[2], Marcos Rubinstein[3], and Farhad Rachidi[1]
[1]Electromagnetic Compatibility Laboratory, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland
[2]Institute for Climate and Atmospheric Science, University of Leeds, Leeds, UK
[3]Institute for Information and Communication Technologies, University of Applied Sciences of Western Switzerland (HES-SO), Yverdon-les-Bains, Switzerland

Lightning is formed in the atmosphere through the combination of complex dynamic and microphysical processes. Lightning can have a considerable influence on the environment and on the economy since it causes energy supply outages, forest fires, damages, injury and death of humans and livestock worldwide. Therefore, it is of great importance to be able to predict lightning incidence in order to protect people and installations. Despite numerous attempts to solve the important problem of lightning prediction (e.g., [1]–[3]), the complex processes and large number of parameters involved in the problem lend themselves to the potential application of a machine learning (ML) approach.

We recently proposed a ML-based lightning early-warning system with promising performance [4]. The proposed ML model is trained to nowcast lightning incidence during any one of three consecutive 10-minute time intervals and within a circular area of 30 km radius around a meteorological station. The system uses the real-time measured values of four meteorological parameters that are relevant to the mechanisms of electric charge generation in thunderstorms, namely the air pressure at station level (QFE), the air temperature 2 m above ground, the relative humidity, and the wind speed. The proposed algorithm was implemented using the data from 12 meteorological stations in Switzerland between 2006-2017 with a granularity of ten minutes. The stations were selected to be well distributed among different ranges of altitude and terrain topographies.

The algorithm requires the filtering out of a portion of the data which are identified as outliers. However, the process of the automatic identification of outliers is a challenging task which could also affect the model's performance. In this presentation, we discuss this problem and present approaches that can be used to optimize the process.

References

[1] D. Aranguren, J. Montanya, G. Solá, V. March, D. Romero, and H. Torres, "On the lightning hazard warning using electrostatic field: Analysis of summer thunderstorms in Spain," J.

Electrostat., vol. 67, no. 2–3, pp. 507–512, May 2009.

[2] G. N. Seroka, R. E. Orville, and C. Schumacher, “Radar Nowcasting of Total Lightning over the Kennedy Space Center,” Weather Forecast., vol. 27, no. 1, pp. 189–204, Feb. 2012.

[3] Q. Meng, W. Yao, and L. Xu, “Development of Lightning Nowcasting and Warning Technique and Its Application,” Adv. Meteorol., vol. 2019, pp. 1–9, Jan. 2019.

[4] A. Mostajabi, D. L. Finney, M. Rubinstein, and F. Rachidi, “Nowcasting lightning occurrence from commonly available meteorological parameters using machine learning techniques,” npj Clim. Atmos. Sci., vol. 2, no. 1, p. 41, 2019.