EGU24-8449, updated on 23 Jul 2024
https://doi.org/10.5194/egusphere-egu24-8449
EGU General Assembly 2024

Radiation fog nowcasting with XGBoost using station and satellite data

Michaela Schütz, Jörg Bendix, and Boris Thies
University of Marburg, Geography, Germany (vorndram@staff.uni-marburg.de)

The research project "FOrecasting radiation foG by combining station and satellite data using Machine Learning (FOG-ML)" represents a comprehensive effort to advance radiation fog prediction using machine learning (ML) techniques, with focus on the XGBoost algorithm. The nowcasting period is up to four hours into the future.

The initial phase of the project involved developing a robust classification-based model that could accurately forecast the occurrence of radiation fog, a challenging meteorological phenomenon. Radiation fog is particularly difficult to predict because it depends on a complex interplay of factors such as ground cooling, humidity, and minimal cloud cover. It often forms rapidly and in local areas. This required careful analysis of the chronological order of the data and consideration of autocorrelation to increase the effectiveness of model training.

Building upon this foundation, the next two phases concentrated on improving the model's forecasting performance for visibility classes (step 2) and for absolute visibility values (step 3). The main focus was then on a nowcasting period of up to two hours. This nowcasting period is critical in fog prediction as it directly impacts transportation planning and safety. The use of ground-level observations in step 2 and integration of satellite data in step 3 provided a rich dataset that allowed for more nuanced model training and validation.

In the latest phase of research, satellite data has been incorporated to further refine the prediction model, especially regarding the fog formation and dissipation. Satellite imagery provides additional variables of atmospheric data that are not readily available from ground-based observations. This integration aims to address one of the inherent limitations in fog forecasting methods, particularly in areas where ground-based observations are sparse.

Throughout the different stages, the project emphasized the need for thorough data processing and validation. This included the implementation of cross-validation techniques to assess the generalizability of the models and the use of various metrics to gauge their predictive power. This has also included the incorporation of trend information, which has proven to be crucial for forecasting with XGBoost. Our research has also shown that not only the overall performance, but also the performance of the transitions (fog formation and resolution) should be analyzed to get a complete picture of the model performance. This finding was consistent throughout the entire study, regardless of classification-based forecast or regression-based forecast.

We have been able to significantly improve the performance of our nowcasting model with each step. We will be presenting the key findings and latest results from this research at EGU24.

All results from step 1 can be found in “Current Training and Validation Weaknesses in Classification-Based Radiation Fog Nowcast Using Machine Learning Algorithms” from Vorndran et al. 2022. All results from step 2 can be found in “Improving classification-based nowcasting of radiation fog with machine learning based on filtered and preprocessed temporal data” from Schütz et al. 2023.