Geophysical Research Abstracts
Vol. 21, EGU2019-10543, 2019
EGU General Assembly 2019

Prediction of daily maximum ozone threshold exceedances by preprocessing and ensemble artificial intelligence techniques

Bing Gong (1), Martin G. Schultz (2), and Felix Kleinert (3)

(1) Forschungszentrum Jülich, Supercomputing Center, Jülich, Germany (b.gong@fz-juelich.de), (2) Forschungszentrum Jülich, Supercomputing Center, Jülich, Germany (m.schultz@fz-juelich.de), (3) Forschungszentrum Jülich, Supercomputing Center, Jülich, Germany (f.kleinert@fz-juelich.de)

The objective of this research is to forecast local daily maximum ozone threshold exceedances. We utilised synthetic minority over-sampling and under-sampling as preprocessing step to address the imbalanced data issues. Meanwhile, some ensemble machine learning algorithms were proposed to compare with traditional single algorithms, resulting in significant improvement of the classifier accuracy. A novel approach by adding a distance -based regional data set, which includes the neighbor monitoring stations meteorological and air quality variables, could effectively capture the ozone transport characteristics, therefore, further enhancing the prediction performance. The results demonstrate that such combination of preprocessing technologies and ensemble algorithms can effectively and accurately forecast ozone threshold exceedances. The relevance of individual variables for ozone prediction were estimated through tree-based algorithms, which confirm the hypothesis that the regional data could benefit better ozone prediction in a data-driven way. This presentation draws from a published study on air pollution in the Hong Kong area. These results will be compared to ozone predictions at German stations and expanded to include more advanced deep learning algorithms for further improvement of the model performance.