Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Identifying pollution sources and predicting urban air quality using ensemble learning methods

Kunwar P. Singh [a,b,*], Shikha Gupta [a,b], Premanjali Rai [a,b]

[a] Academy of Scientific and Innovative Research, Council of Scientific & Industrial Research, New Delhi, India
[b] Environmental Chemistry Division, CSIR – Indian Institute of Toxicology Research, Post Box 80, Mahatma Gandhi Marg, Lucknow 226 001, India

HIGHLIGHTS

- Developed tree ensemble models for seasonal discrimination and air quality prediction.
- PCA used to identify air pollution sources; air quality indices used for health risk.
- Bagging and boosting algorithms enhanced predictive ability of ensemble models.
- Ensemble classification and regression models performed better than SVMs.
- Proposed models can be used as tools for air quality prediction and management.

GRAPHICAL ABSTRACT

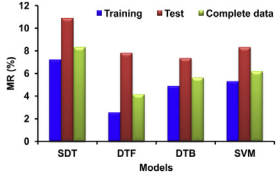
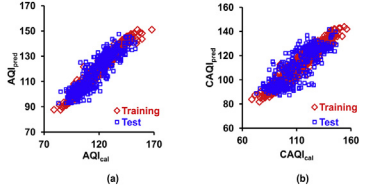


Figure shows misclassification rate in seasonal discrimination of air quality of Lucknow yielded by different models and suggest that the ensemble learning classification models (DTF and DTB) performed relatively better than SDT and SVM.



Figures show correlative distribution of calculated and model predicted values of (a) AQI, and (b) CAQI for Lucknow ambient air using DTB model.