

## Special Section Guest Editorial: Integrating Remote Sensing, Machine Learning, and Data Science for Air Quality Management

Kaixu Bai,<sup>a</sup> Simone Lolli,<sup>b</sup> and Yuanjian Yang<sup>c</sup>

<sup>a</sup>East China Normal University, School of Geographic Sciences,

Key Laboratory of Geographic Information Science (Ministry of Education), Shanghai, China

<sup>b</sup>CNR-IMAA, Consiglio Nazionale delle Ricerche, Contrada S. Loja snc, Tito Scalo, Potenza, Italy

<sup>c</sup>Nanjing University of Information Science & Technology, School of Atmospheric Physics, Key Laboratory for Aerosol-Cloud-Precipitation of China Meteorological Administration, Nanjing, China

Monitoring air pollutant concentration at the near surface is of great significance for better assessing air pollution's impacts on air quality, health exposure risk, and even climate. Satellites are more promising than sparsely distributed ground monitors for regional and even global air pollutant concentration mapping given the vast spatial coverage merit. Remote sensing technology and its applications in air quality monitoring have been growing exponentially along with the presence of versatile satellite platforms and instruments, in particular, novel data processing and retrieval algorithms. To better monitor air pollutant concentration from satellites and to bridge knowledge gaps in understanding air pollutant impacts, recent applications have incorporated many state-of-the-art methods and techniques to help generate quality-enhanced air pollutant concentration datasets. New advancements in remote sensing, machine learning, and data science, such as hyperspectral remote sensing, multimodal data fusion, neural networks, machine learning, and artificial intelligence, have stepped in to facilitate satellite-based air quality monitoring and impact assessment.

Aiming at leveraging the state of the art in remote sensing, machine learning and data science to advance air quality monitoring and management, in particular triggering multidisciplinary applications and relevant research topics, we organized this special section and six papers were finally accepted for publication after rigorous peer-review procedures.

Yang et al. developed a LiDAR system for simultaneous measurement of atmospheric transmittance, turbulence, and wind (ACW-LiDAR) along a common path through an integrated optical and mechanical design. The experimental results indicate that the detection distance of atmospheric transmittance is greater than 10 km @ 1064 nm, with turbulence greater than 10 km @ 532 nm and 4 km for wind @ 1550 nm. The good agreements between measurements from ACW-LiDAR and ground meteorological automatic observation system well demonstrate the accuracy of this proposed system.

Bai et al. proposed a fusion algorithm to generate hourly 1-km PM<sub>2.5</sub> concentrations, which were then used in conjunction with hourly movement trajectory data points to better assess PM<sub>2.5</sub> exposures. The inter-comparison results indicate a better performance of exposure estimates considering individual movement than residence-based exposures, with the bias depending on activity types and distances from home to activity locations. The results highlight the critical importance of considering individual movement trajectory for exposure assessment, especially with long commuting distances.

Ma et al. developed nationwide machine-learned models to predict surface O<sub>3</sub>, NO<sub>2</sub>, and SO<sub>2</sub> concentration using the extreme gradient boosting method, and a multi-task learning model was then utilized to estimate the relative importance of influential factors on PM<sub>2.5</sub> and O<sub>3</sub> pollutions in the North China Plain. The modeling results indicate that both meteorological factors

and anthropogenic emissions contribute to  $PM_{2.5}$  and  $O_3$  pollutions, while  $NO_x$  played more important roles on  $O_3$  pollution whereas  $SO_2$  for  $PM_{2.5}$  pollution in the NCP.

Swetha et al. proposed a multimodal deep learning network ( $M^2$ -APNet) to predict major air pollutant concentration at a global scale from multimodal temporal satellite images, with convolutional neural network used to extract local features and a bidirectional long short-term memory to obtain longitudinal features from multimodal temporal data. The experimental results demonstrate the effectiveness of DEM modality over others in learning to predict major air pollutants and determining the air quality index.

Gu et al. developed a genetic algorithm-optimized backpropagation neural network (GA-BPNN) to better derive  $PM_{2.5}$  concentrations from satellite-based aerosol optical depth imageries. The validation results indicate better performance of GA-BPNN than conventional BPNN in terms of the modeling accuracy, especially for extreme values. 5-year long spatial distribution of  $PM_{2.5}$  concentrations in Dalian was finally mapped with GA-BPNN, and an evident decreasing trend was observed.

Mi et al. explored aerosol impacts on convective clouds in eastern China by making use of satellite remote sensing and machine learning methods. The results revealed that increased aerosol loading resulted in decreased cloud top brightness temperature, weakening precipitation intensity and mesoscale convective systems area in turn at a 90% confidence level. These results deepen our understanding of aerosol impacts on aerosol–cloud–precipitation interaction.

As editors, we are grateful for the dedication and passion of authors and referees contributing to this special section. We hope this special section sparks new ideas and triggers multidisciplinary collaboration to advance air quality monitoring and management by leveraging cutting-edge techniques in remote sensing, data science, and artificial intelligence.