

Assessment of numerous Data Assimilation Approaches in Regional Retrieval of Ground NO₂ Spatial Distribution via Remote Sensing products

*Hugo Wai Leung MAK^{1,2}, Jimmy Chi-Hung FUNG^{1,2}

1. Department of Mathematics, Hong Kong University of Science and Technology, 2. Division of Environment and Sustainability, Hong Kong University of Science and Technology

Combating with environmental pollution problems and reducing human exposure to unhealthy levels of airborne chemicals are important global missions nowadays, particularly in more developed cities in China. Existing monitoring network for ground pollutant measurements in most places of southern China are currently too sparse, therefore spatial variation and characteristics in rural regions or places with mixed land use may not be well detected and distinguished. Thus, the use of satellite remote sensing techniques play important roles in observing temporal changes of pollutants within troposphere through satellite scans. Based on tropospheric NO₂ column density inferred from Ozone Monitoring Instrument (OMI), together with certain meteorological quantities, ground NO₂ spatial distribution could be retrieved via data assimilation techniques, thus providing a possibility to detect changes of pollutants within a prescribed domain over long period.

In this research, we adopt the use of modified Berkeley High Resolution product (BEHR-HK v3.0C) and meteorological outputs from the Weather Research and Forecasting (WRF) model to describe tropospheric NO₂ column density within southern China, especially Greater Bay Area. The preliminary steps of such approach are mostly in line with NASA Standard Product, however it makes use of higher-resolution a priori NO₂ profiles and terrain patterns to reformulate the Air Mass Factor (AMF), thus doing a better job in capturing changes of NO₂ within lower troposphere, particularly in places with large spatial variabilities and terrain height differences like China.

Furthermore, based on understanding in atmospheric meteorology, we combine the use of updated tropospheric NO₂ column density and 8 meteorological factors from WRF (e.g. Planetary Boundary Layer Height (PBLH), atmospheric pressure and perturbation potential temperature), then retrieve the estimated spatial distribution of ground-level NO₂ concentration in southern China within 4 seasons of 2015, via combination of machine learning approaches and kernel-based probabilistic models. These data assimilation techniques include non-parametric models like Gaussian Process Regression (GPR), imposing of different statistical kernels (like Gaussian or Radial Basis Function (RBF)) or Support Vector Machine (SVM). Retrieval results are then validated by available measurement datasets obtained from ground monitoring stations. Respective accuracy and applicability of these models are assessed by same statistical metrics, which we conclude that GPR model reaches the best performance in predicting ground NO₂ concentration after a series of machine learning and training processes were applied, followed by Tree Ensemble model. This finding opens new window into long-term ground NO₂ retrieval and prediction, and dealing with environmental problems that require continuous monitoring and assessments.

Keywords: Satellite Remote Sensing Techniques, Tropospheric Column Density and Ground Retrieval, Inverse Approaches in Data Assimilation

