Efficient Calibration of Numerical Model Output Using Hierarchical Dynamic Models

Numerical air quality models, such as the Community Multiscale Air Quality (CMAQ) system, play a critical role in characterizing pollution levels at fine spatial and temporal scales. Nevertheless, numerical model outputs may systematically overestimate or underestimate pollutants concentrations due to various reasons. In this work, we propose an hierarchal dynamic model that can be implemented to calibrate grid-level CMAQ outputs using point-level observations from sparsely located monitoring stations. Under a Bayesian framework, our model presents a flexible quantification of uncertainties by considering deep hierarchies for key parameters, which can also be used to describe the dynamic nature of data structural changes. In addition, we adopt several newly-emerged techniques, including triangulations of research domain, tapering-based Gaussian kernel, sparse Gaussian graphical model, variational Bayes and ensemble Kalman smoother, that significantly speed up the whole calibration procedure. Our approach is illustrated using daily PM2.5 datasets of China's Beijing-Tianjin-Hebei region, which contains 68 monitoring stations and is covered by 2499 CMAQ 9-km grids. In contrast to existing methods, our model gives more accurate calibration results in most of the grids with higher computation efficiencies. This allows us to provide an effective calibration tool for large-scale numerical model outputs and generate better high-resolution maps of pollutants.

Key words: CMAQ; PM2.5; calibration; hierarchical dynamic model; variational Bayes;