Spatially Adaptive Calibrations of AirBox $PM_{2.5}$ Data

ShengLi Tzeng[a], Chi-Wei Lai[b], Hsin-Cheng Huang[c]

[a]*Department of Applied Mathematics, National Sun Yat-sen University, Taiwan, R.O.C.*
[b]*Institute of Statistics, National Tsing Hua University, Taiwan, R.O.C.*
[c]*Institute of Statistical Science, Academia Sinica, Taiwan, R.O.C.*

Abstract

Two networks are available to monitor $PM_{2.5}$ in Taiwan, including the Taiwan Air Quality Monitoring Network (TAQMN) and the AirBox network. The TAQMN, managed by Taiwan's Environmental Protection Administration (EPA), provides high-quality $PM_{2.5}$ measurements at 77 monitoring stations. More recently, the AirBox network was launched, consisting of low-cost, small internet-of-things (IoT) microsensors (i.e., AirBoxes) at thousands of locations. While the AirBox network provides broad spatial coverage, its measurements are not reliable and require calibrations. However, applying a universal calibration procedure to all AirBoxes does not work well because the calibration curves vary with several factors, including the chemical compositions of $PM_{2.5}$, which are not homogeneous in space. Therefore, different calibrations are needed at different locations with different local environments. Unfortunately, most AirBoxes are not close to EPA stations, making the calibration task challenging. In this article, we propose a spatial model with spatially varying coefficients to account for heteroscedasticity in the data. Our method gives adaptive calibrations of AirBoxes according to their local conditions and provides accurate $PM_{2.5}$ concentrations at any location in Taiwan, incorporating two types of measurements. In addition, the proposed method automatically calibrates measurements from a new AirBox once it is added to the network. We illustrate our approach using hourly $PM_{2.5}$ data in the year 2020. After the calibration, the results show that the $PM_{2.5}$ prediction improves about 37% to 67% in root mean-squared prediction error for matching EPA data. In particular, once the calibration curves are established, we can obtain reliable $PM_{2.5}$ values at any location in Taiwan, even if we ignore EPA data.

Keywords: Heterogeneous variance; kriging; microsensor; monitoring station; robust estimation; spatially varying coefficient model.
