統計科學研究所
INSTITUTE OF STATISTICAL SCIENCE

SEMINAR STATAS

Seminar

Title：Adaptive Calibrations of Spatially Misaligned IoT Data

Speaker：Prof. ShengLi Tzeng

(Department of Applied Mathematics, National Sun Yat-sen University, Taiwan)

Time：10:30~12:00, Monday, September 4, 2023

Place：Auditorium, B1F, Institute of Statistical Science

Abstract

We consider a challenging problem of calibrating geo-referenced data that suffers from spatial misalignment due to the use of multiple instruments to measure the same variables. Some instruments are precise but expensive, while low-cost instruments are less accurate but more widespread. Data fusion techniques are often employed to combine these different sources and extract more information, but spatial misalignment hinders the direct application of usual fusion methods. Before data fusion, we need careful calibration for those untrustworthy observations. Otherwise, without reliable models, more data can introduce bias and noise.

To address this, we propose a strategy to calibrate fine particulate matter (PM2.5) data in Taiwan. We have two sources of PM2.5 concentration measurements from traditional monitoring stations and low-cost internet-of-things (IoT) devices called AirBoxes. AirBoxes are unreliable but easily deployed and form a large network. A one-size-fits-all calibration procedure for all AirBoxes does not work well, because the relationship between measurements from AirBoxes and traditional monitoring stations is not homogeneous in space and many outliers exist. We develop a fast, robust method to model the PM2.5 processes to account for the heterogeneity in the data. We also use a spatially varying coefficient regression framework to calibrate AirBox measurements. Our method, requiring no collocated data, gives spatially adaptive calibrations of AirBoxes, produces accurate PM2.5 concentration estimates at any location, and provides calibration formulas for new AirBoxes once they are added to the network. We illustrate our approach using hourly PM2.5 data in 2020. After the calibration, the results show that the PM2.5 prediction improves by about 38%–68% in root-mean-squared prediction error. This work provides a flexible and reliable solution to the problem of spatial misalignment with data from multiple sources.

Keywords: Geostatistics; Robust estimation; Spatially varying coefficient model; Heterogeneous variance; Misalignment.

※Online live streaming through Cisco Webex will be available.

※ The tea reception will be held at 10:10.