# TRAP: A Novel Road-level Spatial Interpolation to Improve Estimation Errors of Air Pollution

Hyesop Shin*[1]

[1]MRC/CSO Social and Public Health Sciences Unit, School of Health and Wellbeing, University of Glasgow, Clarice Pears Building, 90 Byres Road, Glasgow, G12 8TB

GISRUK 2023

**Summary**

Spatial interpolation techniques have been used in air pollution studies to generate area-level estimates. Despite the benefits of a mathematically sound concept, rapid implementation, and user-friendly software, interpolation suffers in areas with a low number of monitoring stations and when the built environment is ignored. The purpose of this study is to introduce TRAP, a nearly finished R package that is a new road-scale spatial interpolation method that uses road weighting. The $NO_2$ results in Seoul showed a small variation during the summer, but large daily variations during the winter. The road-overlaid outcomes gave improved results relative to the roadside measurements.

**KEYWORDS:** Spatial Interpolation, Air Pollution, TRAP, R Package, Road Weighting.

## 1 Introduction

Daily ambient air pollution has posed a major threat to population health. Despite efforts to legislate national pollution standards, daily $NO_2$ levels in cities such as Seoul or London have repeatedly exceeded WHO standards. Understanding the spatial and temporal aspects of air pollution, as well as their relationship to exposure, is critical for reducing further adverse health effects associated with air pollution.

Various Spatial Interpolation (SI) techniques have been used in air pollution studies to generate area-level estimates of pollutants, allowing researchers to further compute the potential exposure across different geographic scales. Despite criticisms, SI has remained a popular method due to its advantages of a mathematically sound concept, quick implementation speed, and useful software. However, when the number of air pollution monitoring stations is limited in comparison to the size of the city, estimation errors, known as small-scale variability, are likely to increase in monitor-sparse regions (Wu et al., 2019; Chen and Lin, 2022; Shiode and Shiode, 2011). Another drawback of SI methods is that they assume all locations exist in a two-dimensional space and that the distance

*hyesop.shin@glasgow.ac.uk

between monitoring stations is measured solely on the Euclidean distance, which excludes complex topographies and built environments (Chen and Lin, 2022; Li et al., 2015).

This study aims to examine small-scale variations that occur during SI prediction and to introduce a new road-scale spatial interpolation method that employs road weighting. This method, which is also computationally light and fast, is expected to help improve the prediction accuracy of road air quality. This presentation introduces the nearly finished development of the R package TRAP (Traffic-related Air Pollution).

## 2 Methods

### 2.1 Data Collection

This study selected Seoul as a case study and used $NO_2$ as the main source. Data were collected from 57 background stations (installed on the rooftops of district offices) and 19 roadside stations that were within 10km from the city boundary (see Figure 1). The data for the summer of 2013 were downloaded between the 1st of August and the 30th of September 2013, and between the 2nd of December 2013 and the 28th of February 2014. Units are measured in ppb (parts per billion). This study used 12-hour aggregations to reduce the short-term exposure estimation errors at hourly intervals.

Road layout was provided by the Korean Transport Database (KTDB). Seoul had 59,319 road segments in total, and the attributes of each segment included node ID, link ID, number of lanes, maximum speed, traffic light density, and road ranks. In Table 1, Road segments were grouped by their road ranks as a level to distinguish different pollution loading. Rural and county roads 105, 106, and 107 were excluded due to a small fraction existing in Seoul. The roads were regrouped as general roads (103, 104), and motorways (101, 102, 108) to distinguish the pollution loads by two road types.

| Rank | Name | Group |
|---|---|---|
| 101 | Highway | Highway |
| 102 | Urban highway | Highway |
| 103 | National road | General |
| 104 | Metropolitan road | General |
| 105 | Rural road (gov-supported) | - |
| 106 | Rural road | - |
| 107 | County road | - |
| 108 | Highway ramp | General |

Table 1: Road hierarchies and groups

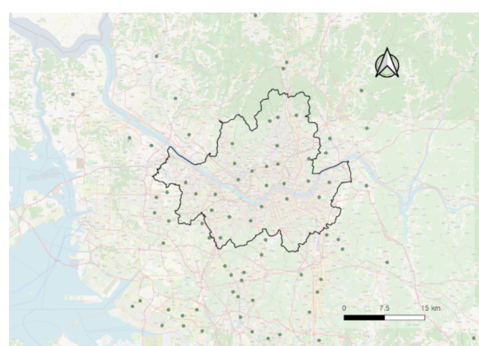

Figure 1: 57 Background pollution stations that are considered for spatial interpolation

## 2.2 Modelling Universal Kriging by considering Small-scale Variability

This study used Universal Kriging (UK), a non-stationary variant of Ordinary Kriging in which the mean varies deterministically in different locations (trend or drift), but only the variance remains constant (Kumar, 2007; Kim et al., 2014; Li et al., 2015). We aumotated the fitting of the conceptual semivariograms to empirical splines to krige the entire study extent, then manually adjusted nuggets and partial sills if artefacts (a.k.a. bull's eye effect) were observed. This study used variogram experiments to discover small-scale $NO_2$ variations under 5kms, and discovered that on some days the prediction was created with little or no spatial autocorrelation (which created a very smoothed outcome).

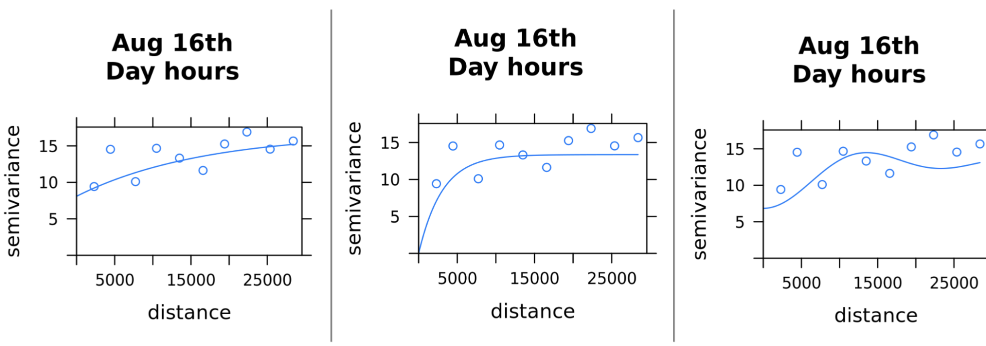


Figure 2: The example of small-scale variability of $NO_2$ on Aug.16th. Semivariogram examples of good-fit, overfit, and underfit. The middle semivariogram leads to bull's eye effects on the map while the right gives an overly smoothed outcome.

# 3 Adding a Road-Weight Function to Background Pollution Levels on Spatial Interpolation Predictions

As to improve the accuracy of $NO_2$ prediction, this study took roadside pollution data from roads and motorways respectively and measured the ratios (see Figure 3). The ratios were then applied universally to the road networks across Seoul.

# 4 Results

The average concentration of $NO_2$ in August was around 15ppb (see Figure 4). The distribution was consistent throughout the city. In comparison, $NO_2$ had a relatively higher average of 42ppb in February, with greater spatial variability and temporal oscillation. In the road-overlayed results, $NO_2$ levels on roads were approximately 80% and 30% higher than in the background areas during the summer and winter seasons, respectively (see Figure 5).

# 5 Conclusion

To better estimate a spatially gridded field for road air pollution, this study used Universal Kriging at 12-hour intervals and then added an extra road effect. The $NO_2$ results showed a small variation

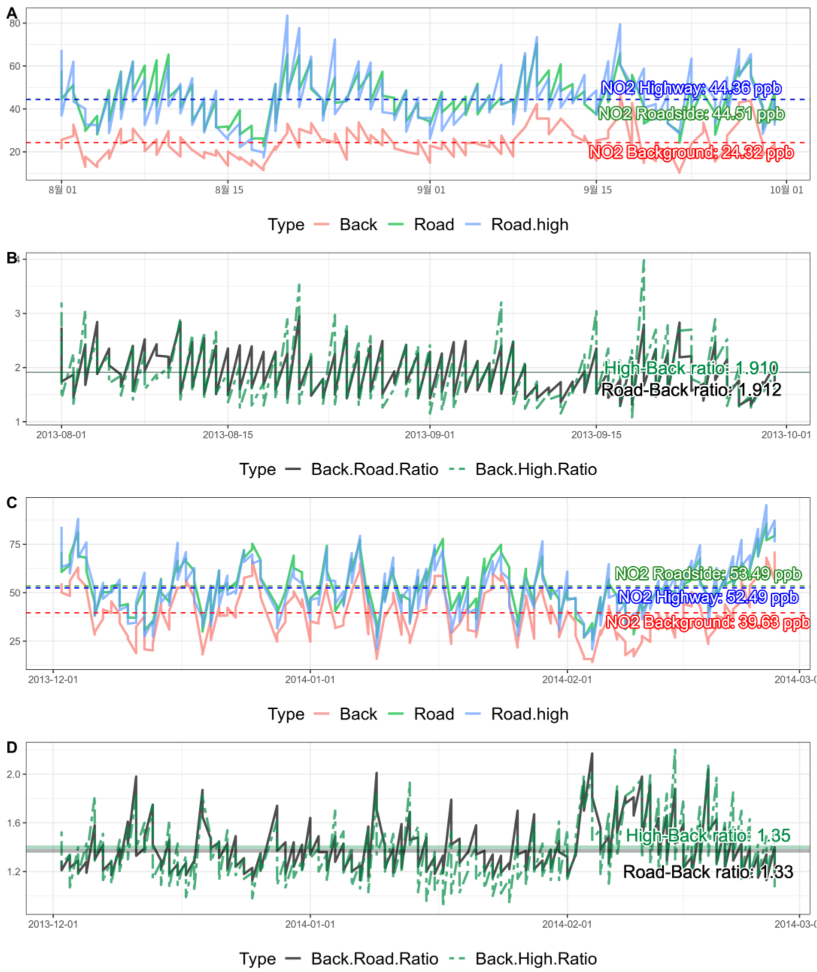


Figure 3: $NO_2$ ratio measured by every 12-hour average in late July-September 2013 and December-February 2013-2014. Graphs A and C are pollution levels of $NO_2$, and Graphs B and D are ratios between aggregated roadside and background stations, and between the aggregated urban highway and background stations.

during the summer, but large daily variations during the winter. Although statistical models cannot take into account atmospheric dispersion in the street canyon scale (Di Sabatino et al., 2008), the road-overlaid outcomes gave improved results relative to the roadside measurements which is promising. For future work, splitting the background-roadside ratio into districts might improve the model's accuracy.

## 6 Data and Code Availability

The data, codes, and figures are available on our GitHub repository: `https://github.com/dataandcrowd/GISRUK2023`.

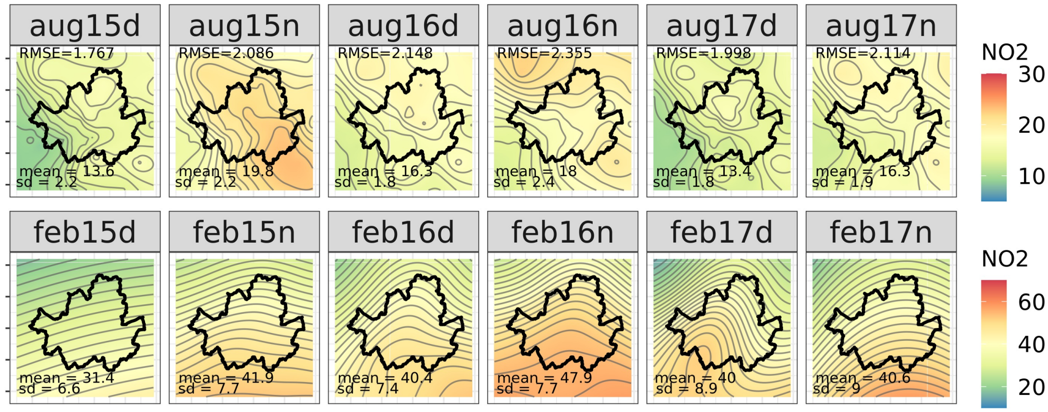


Figure 4: Interpolated $NO_2$ maps on August 15-17th by 12 hour intervals.

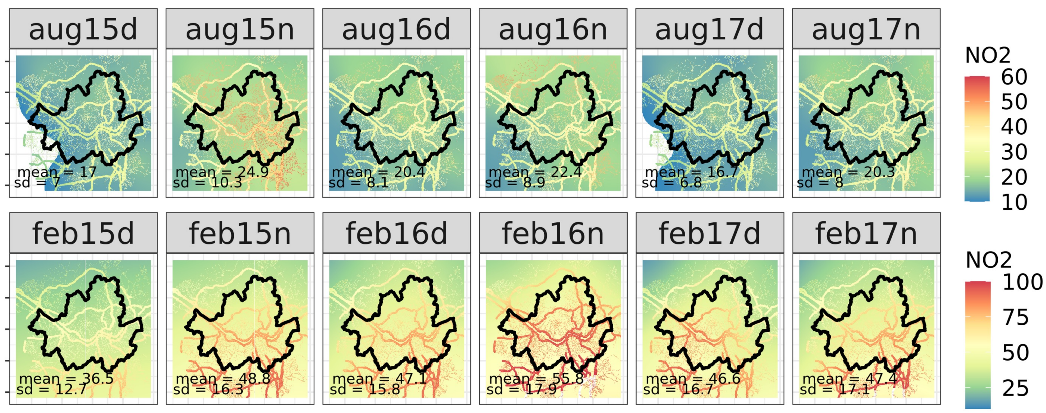


Figure 5: Time series outcomes of combined $NO_2$ interpolation and road effects in mid-February 2014

## 7 Biography

Hyesop Shin is a research associate at the University of Glasgow MRC/CSO Social and Public Health Sciences Unit, and has interests in agent-based modelling (ABM), urban air quality, traffic modelling, and physical activity.

## References

Chen, P.-C. & Lin, Y.-T. (2022). Exposure assessment of pm2. 5 using smart spatial interpolation on regulatory air quality stations with clustering of densely-deployed microsensors. *Environmental Pollution*, *292*, 118401.

Di Sabatino, S., Buccolieri, R., Pulvirenti, B., & Britter, R. E. (2008). Flow and pollutant dispersion in street canyons using FLUENT and ADMS-Urban. *Environmental Modeling & Assessment*, *13*(3), 369–381.

Kim, S.-Y., Yi, S.-J., Eum, Y. S., Choi, H.-J., Shin, H., Ryou, H. G., & Kim, H. (2014). Ordinary kriging approach to predicting long-term particulate matter concentrations in seven major Korean cities. *Environmental Health and Toxicology*, *29*, e2014012, `https://doi.org/10.5620/eht.e2014012`.

Kumar, V. (2007). Optimal contour mapping of groundwater levels using universal kriging - A case study. *Hydrological Sciences Journal*, *52*(5), 1038–1050, `https://doi.org/10.1623/hysj.52.5.1038`.

Li, L., Romary, T., & Caers, J. (2015). Universal kriging with training images. *Spatial Statistics*, *14*, 240–268.

Shiode, N. & Shiode, S. (2011). Street-level spatial interpolation using network-based idw and ordinary kriging. *Transactions in GIS*, *15*(4), 457–477.

Wu, C.-Y., Mossa, J., Mao, L., & Almulla, M. (2019). Comparison of different spatial interpolation methods for historical hydrographic data of the lowermost Mississippi River. *Annals of GIS*, *25*(2), 133–151, `https://doi.org/10.1080/19475683.2019.1588781`.