Dear Prof. Dr. Rachel Franklin,

Please find attached to this letter our research paper

“A comparison of spatial and non-spatial methods in statistical modelling of NO2: prediction accuracy, uncertainty quantification, and model interpretation,

to be considered Geographical Analysis*.*

NO2 is a traffic-related air pollutant that negatively affect our health and the spatial prediction of it is an important subject in geographical analysis to inform decision-making and understand health-environment relationships. Recently, Machine Learning (ML) methods have shown to be powerful in capturing non-linear relationships between NO2 measurements and geospatial predictors for spatial prediction of NO2. but it is unclear if the spatial structure of NO2 is sufficiently captured in the response-covariates relationships and how prediction intervals and model interpretation derived from ML models compare with geostatistical models. In addition, it is commonly not evaluated how different models behave in different geographical areas.

In this study, we compared geostatistical methods with ML methods in the spatial prediction of NO2 and implemented two methods that to our knowledge have not been applied in air pollution mapping, one post-processes quantile random forest with L1-norm shrinkage (Lasso) regression and the other geostatistical stacked learning. We developed a comparison process that comprehensively compare the predicted mean, prediction intervals, and model interpretation of different spatial and non-spatial models. Spatial and non-spatial CV strategies are used to reveal prediction accuracy in different aspects. The two new methods considerably improve from the original (quantile) random forest and stacked learning methods, respectively.

With geospatial predictors and ground observations becoming increasingly available, many statistical methods have been developed in NO2 mapping, but a study that comprehensively dive into models with different structures and complexity to understand the strength and limitations of each is lacking in air pollution mapping. Our comparison study is important in understanding different model behaviours and pointing out good practices in spatial prediction of air pollution and future directions for improvements. We also paid full attention to computational efficiency in the methods applied, all the methods we applied are highly scalable and the Lasso post-processing further reduces model redundancy. We therefore strongly believe our study is highly relevant to *Geographical Analysis* .

Thank you very much for your consideration,

Dr. Meng Lu,

On behalf of Joaquin Cavieres and Dr. Paula Moraga