National Ground-Level NO2 Predictions via Satellite Imagery Driven Hybrid Neural Networks

Cao, Elton (School: Fairview High School)

Outdoor air pollution, specifically nitrogen dioxide (NO2), poses a global health risk. Land use regression (LUR) models are widely used to estimate ground-level NO2 concentrations by describing the satellite land use characteristics of a given location using buffer distance averages of variables. However, information may be leaked in this approach. Therefore, in this study, I leverage a convolutional neural network (CNN) architecture to directly pass pixel plots of satellite imagery for the prediction of U.S. national ground-level NO2. I designed CNN architectures of various complexity which inputs both image and numerical based data, testing both high and low resolution pixel plots. My resulting model accurately predicted NO2 concentrations at both daily ($R2 = 0.898$) and annual ($R2 = 0.964$) temporal scales, with coarse resolution imagery and simple CNN architectures displaying the best and most efficient performance. Furthermore, the CNN outperforms traditional buffer distance models, including random forest (RF) and neural network approaches. Additionally, with a novel graph neural network (GNN) based approach which leverages network interactions between monitoring sites, I developed a hybrid hierarchical GNN-CNN model which captures both short and long distance associations between monitors. The resulting hybrid model significantly improved prediction against new and unseen monitors. With the success of hybrid neural networks in this approach, satellite land use variables continue to be useful for the prediction of NO2. Using this computationally inexpensive model, I encourage the globalization of advanced LUR models as a low cost alternative to traditional NO2 monitoring.