AIRPOCALYPSE RED ALERT

Visualizing the Spatial Interpolation of Air Pollution Data in Beijing

SMU SINGAPORE MANAGEMENT UNIVERSITY

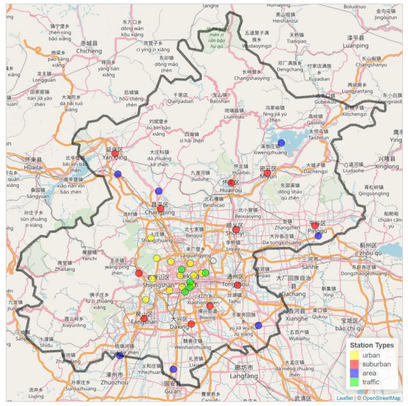
Ayushi Jaiswal | Daxaniie D/o Selvaraj | Soh Hui Shan

Guided by Prof. Kam Tin Seong

Introduction and Motivation

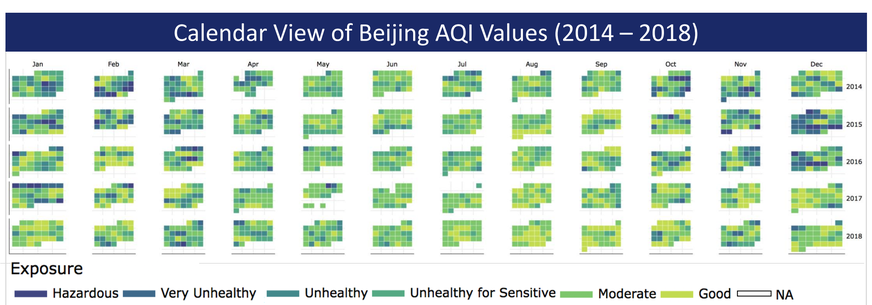
In China, air pollution was previously estimated to contribute to over 1 million deaths annually. The capital city Beijing has suffered serious air pollution for many years. One of the many serious events was in December 2015 when Beijing issued its first red alert as a severe smog engulfed the city for weeks elevating public awareness to unprecedented levels and prompting the government to roll out emergency measures. This project aims to visualise the changing patterns of air pollution across Beijing over time. Spatial interpolation is performed using 3 models to estimate pollutant concentrations at unknown locations.

About Our Data



Historical Beijing Air Quality data is obtained for 5 years from 2014 to 2018. Hourly data from 35 monitoring stations were collected which includes Air Quality Index (AQI) as well as concentration of six pollutants: PM2.5, PM10, CO, NO, O3 and SO2. The stations are categorised into traffic, urban, suburban and area stations. The density of the stations are higher in the city centre.

Data Exploration



Calendar view is used to visualise the trends of the average daily AQI values over 5 years. The colour indicates the levels of severity.

Tools Used

R, Shiny, Excel, Tableau, Python, Selenium, rmarkdown

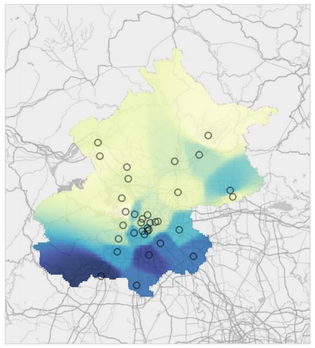
Methodology: Spatial Interpolation

Three spatial interpolation models are built to interpolate unknown values - Inverse Distance Weighted (IDW), Ordinary Kriging (OK) and Thin Plate Spline (TPS). The app is designed for different users like socio-economist, geo-statisticians and climatologist respectively.

Inverse Distance Weighted

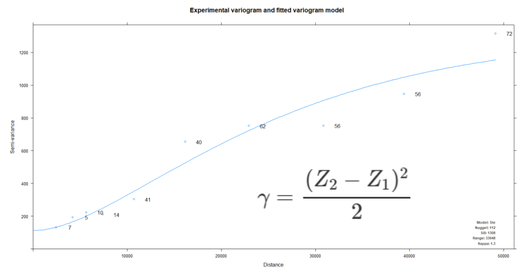
Inverse Distance Weighted technique (IDW) computes an average value for unsampled locations using values from nearby weighted locations.

$$\hat{Z}_j = \frac{\sum_i Z_i / d_{ij}^n}{\sum_i 1 / d_{ij}^n}$$

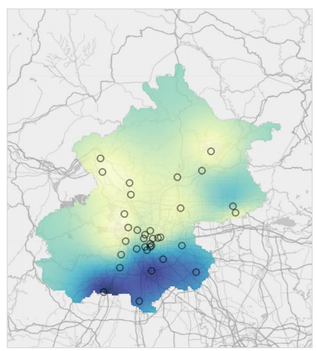


Ordinary Kriging

Kriging uses variogram model to compute weights of neighboring points based on distribution of values. It lets the localized pattern produced by the sample points define the weights for interpolation of unknown values.



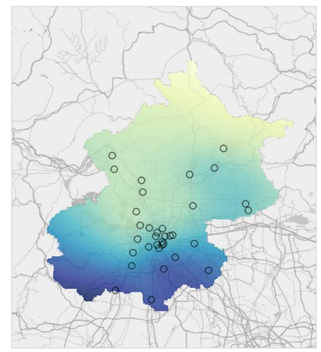
Variogram plots the semi-variance (γ) of the attribute values with increasing distance between station pairs.



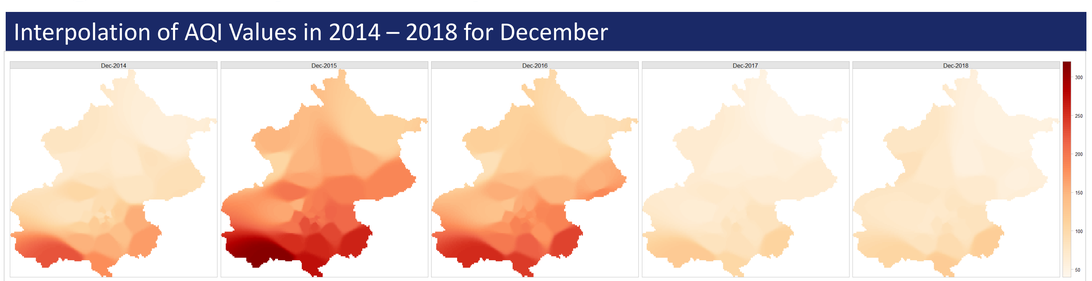
Thin Plate Spline

Thin Plate Splines (TPS) spline-based technique for data interpolation and smoothing. TPS fits a mapping function f(x) between corresponding point-sets y_i and x_i that minimizes the following function.

$$E_{\text{tps}}(f) = \sum_{i=1}^{K} \|y_i - f(x_i)\|^2$$



Insights from Spatial Interpolation



R Packages Used:
gstat, sp, sf, fields, leaflet, rgdal, raster, rasterVis, automap, tidyverse, eesim, dygraphs, plotly

External File:
Beijing spatial shapefile

Conclusion

Leveraging on the interactivity of Shiny, our app exposes the input parameters for model calibration, hence incorporating flexibility and transparency for different users, which could remain black-box otherwise. The severity of Air Pollution Red Alert in December 2015 can be examined visually and the improvement can be seen in the Air Quality Index based on measures taken up by the government over time.