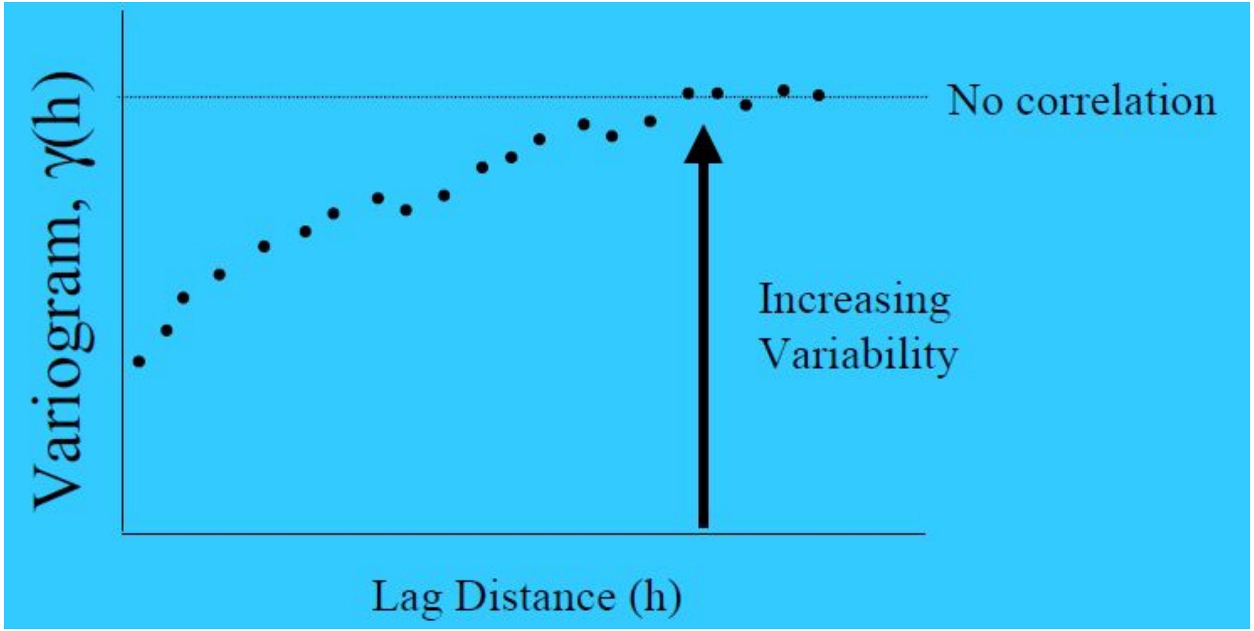
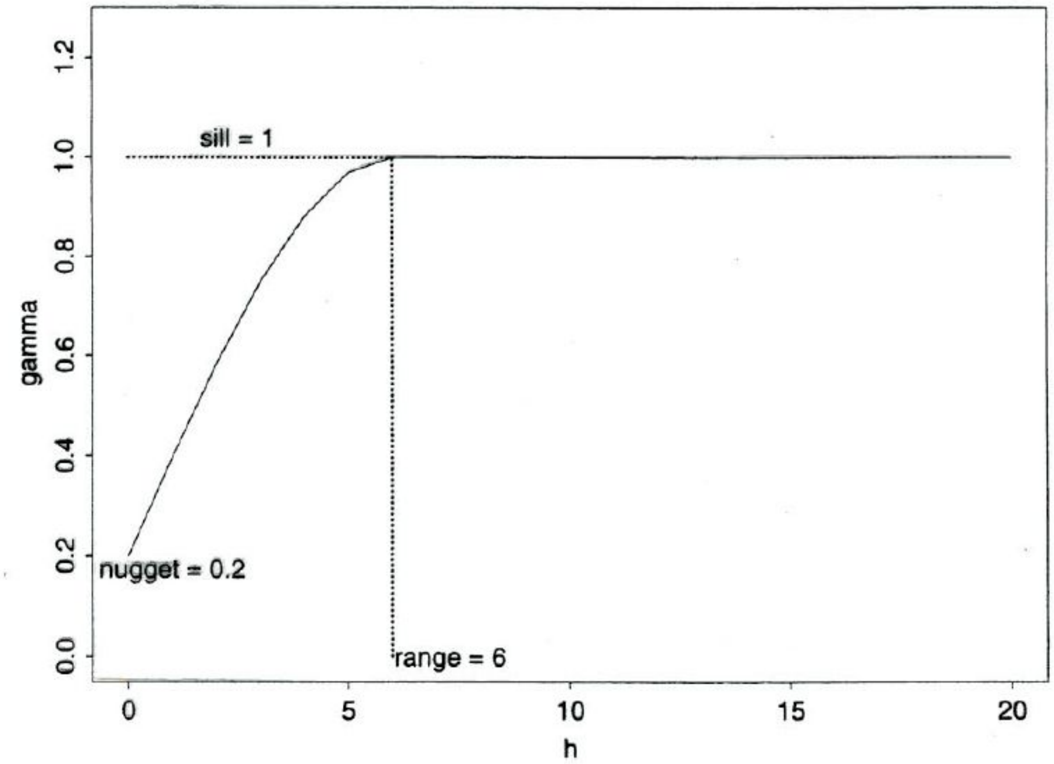
Variogram & Kriging

Prakkash Manohar

Variogram



$$\gamma(h) = \frac{1}{2|N(h)|} \sum_{N(h)} (z_i - z_j)^2$$



A generic variogram showing the *sill* and *range* parameters along with a *nugget effect*.

Variogram Assumptions

1. Normality
2. Stationarity
3. No Trend in the data

Covariance Function

- Model the covariance using a continuous function
- Provides variogram values for all distances (regardless of how small they are)

Spherical Models

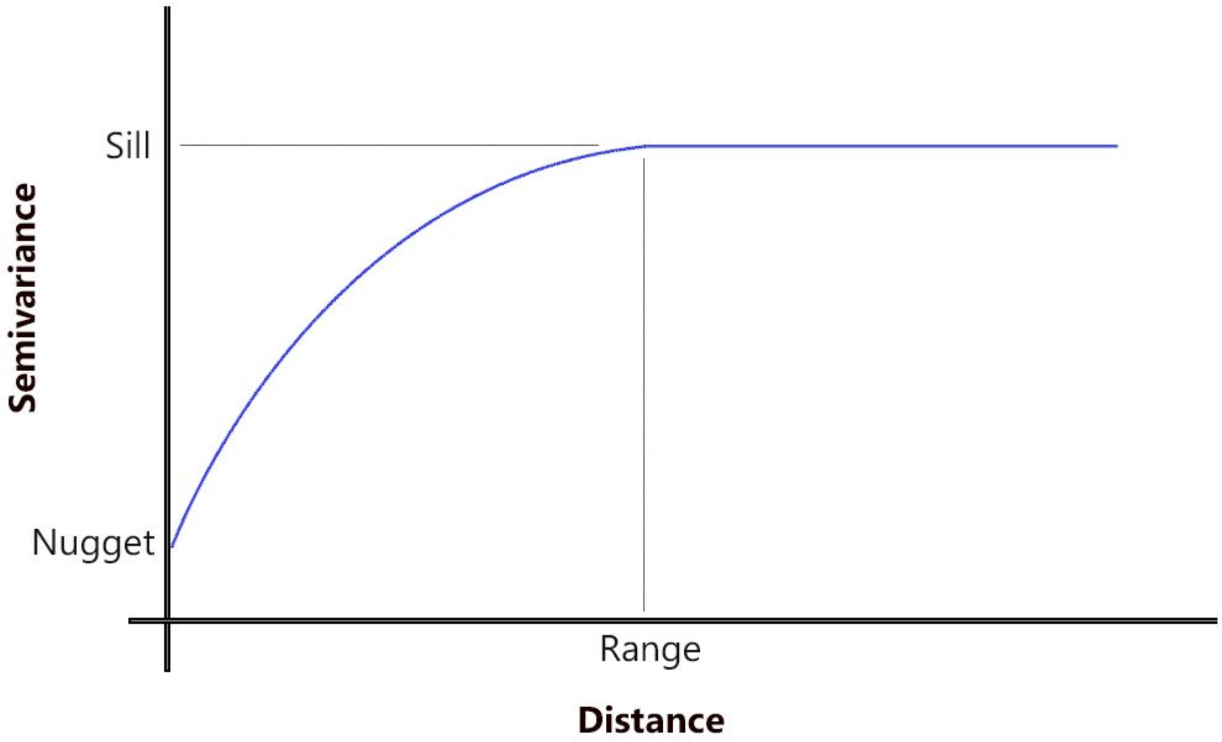
- One of the most common models used in variogram modelling.
- Modified quadratic equation where spatial dependence flattens out as the sill and range.

$$\gamma(h) = \gamma_0 + s\left[1.5\left(\frac{h}{a}\right) - 0.5\left(\frac{h}{a}\right)^3\right], \quad h \leq a$$

$$= \gamma_0 + s, \quad h > a.$$

$$\text{where} : h = \text{offset}, a = \text{range},$$

$$\gamma_0 + s = \text{sill}, \gamma_0 = \text{nugget}$$


Sill
Semivariance
Nugget
Range
Distance

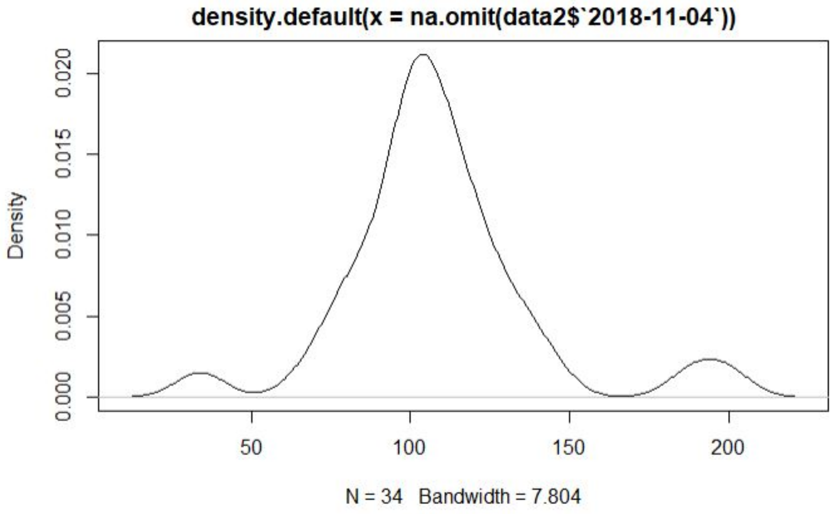
Diwali Air Pollutant Concentration Data

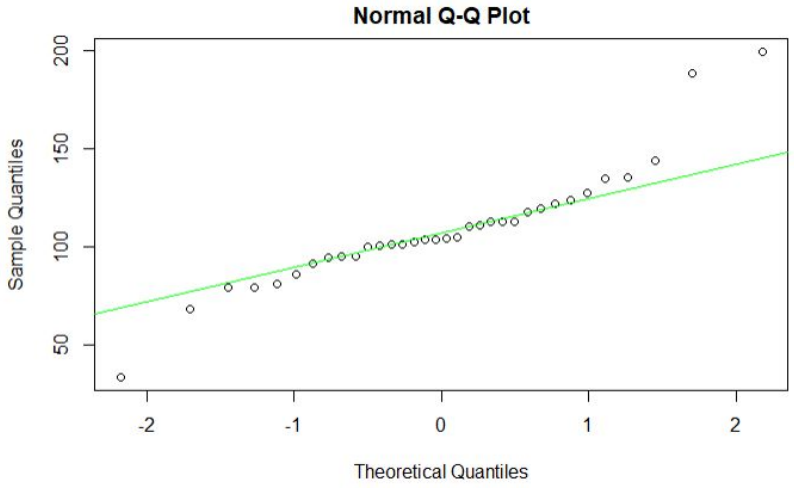
Dataset Source: CPCB

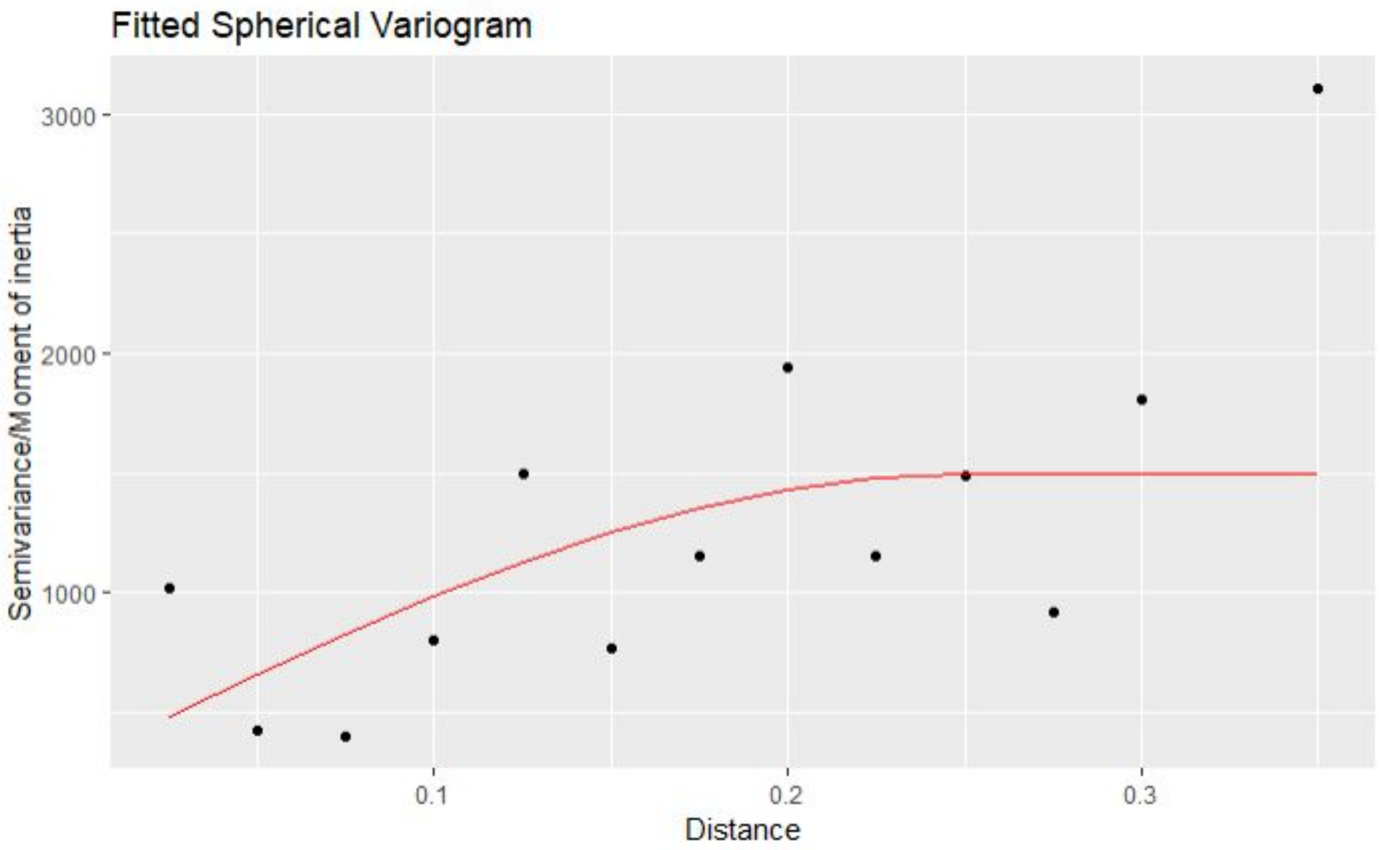
36 localities of Delhi * 13 pollutants *($PM_{2.5}$, CO, NO, NO_2, NO_x, Ozone, SO_2, PM_{10}, NH_3, CH_4, CO_2, SPM, Black Carbon)* * 8 days *(4th Nov, 2018 to 11th Nov, 2018)*

1. Data Cleaning and Structuring
2. Variogram for $PM_{2.5}$ concentration
3. Fitting a Spherical Variogram
4. Ordinary Kriging for $PM_{2.5}$ concentration

Variogram for $PM_{2.5}$ concentration (on 4th Nov’18)






Fitted Spherical Variogram
3000
2000
1000
Semivariance/Moment of inertia
0.1
0.2
0.3
Distance

The results for variogram fitting are:

Range: $a = 0.25$

Nugget: $C_0 = 300$

Sill: $C_0 + C_1 = 1500$

Ordinary Kriging (B.L.U.E)

$$\hat{V}(x_0) = \sum_{i=1}^{n} w_i V(x_i) \qquad R(x_0) = \sum_{i=1}^{n} w_i V(x_i) - V(x_0) \qquad \sum_{i=1}^{n} w_i = 1$$

$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \tilde{C}_{ij} - 2\sum_{i=1}^{n} w_i \tilde{C}_{i0}$$

results in

$$w = C^{-1} D$$

$$\underbrace{\begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{bmatrix}}_{(n+1)\times 1} = \underbrace{\begin{bmatrix} \tilde{C}_{11} & \cdots & \tilde{C}_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \tilde{C}_{n1} & \cdots & \tilde{C}_{nn} & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}^{-1}}_{(n+1)\times(n+1)} \underbrace{\begin{bmatrix} \tilde{C}_{10} \\ \vdots \\ \tilde{C}_{n0} \\ 1 \end{bmatrix}}_{(n+1)\times 1}$$

IDW Interpolation

$$\hat{V}(x_0) = \frac{\sum_{i=1}^{n} w_i V(x_i)}{\sum_{i=1}^{n} w_i} \qquad where, w_i = \frac{1}{d(x_0, x_i)^2}$$

- In the general case, the power 2 can be replaced with any power depending on the kind of interpolation values required.
- Here, 2 is taken as the usual/default value.

Ordinary Kriging & IDW

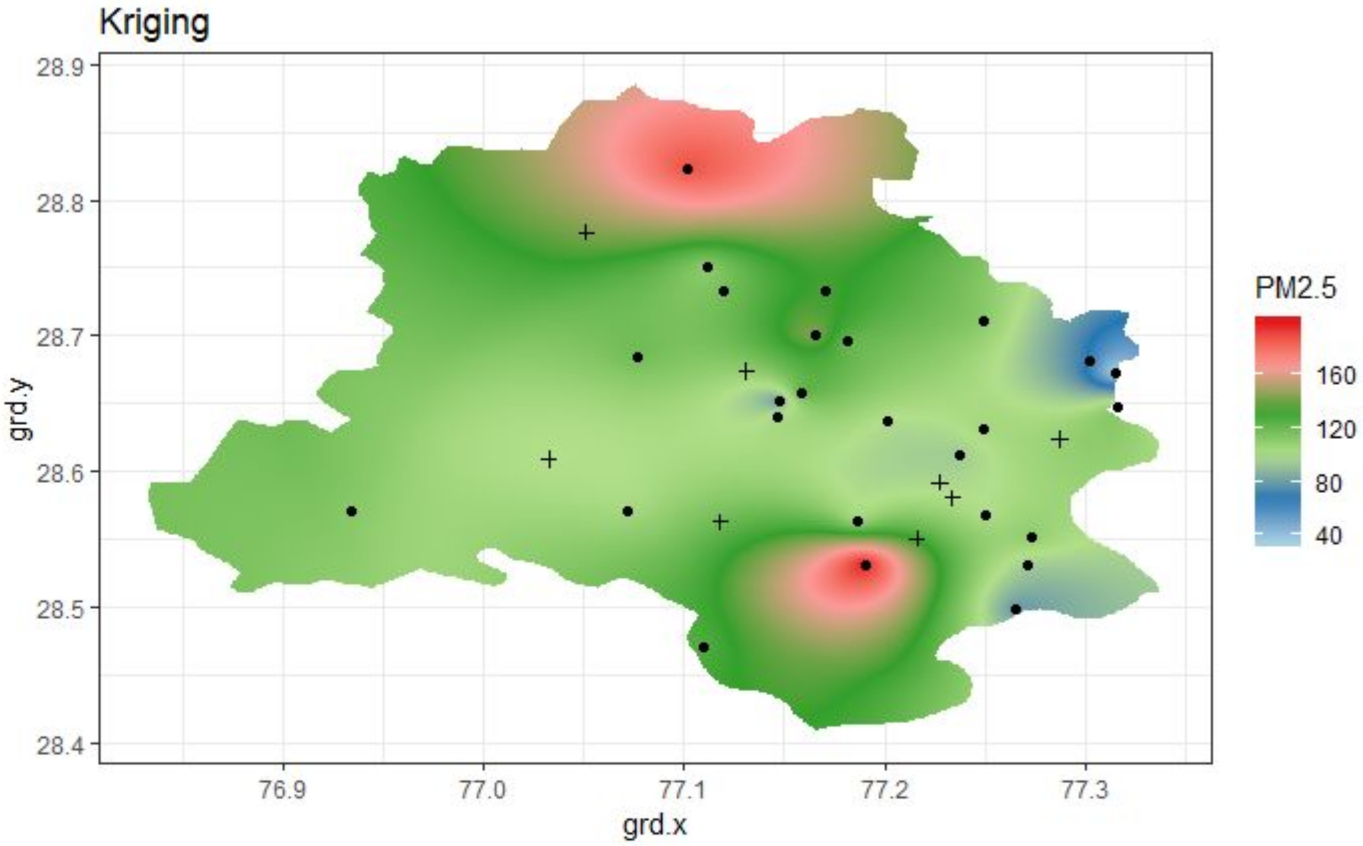
26 locations for building the kriging and the IDW models.

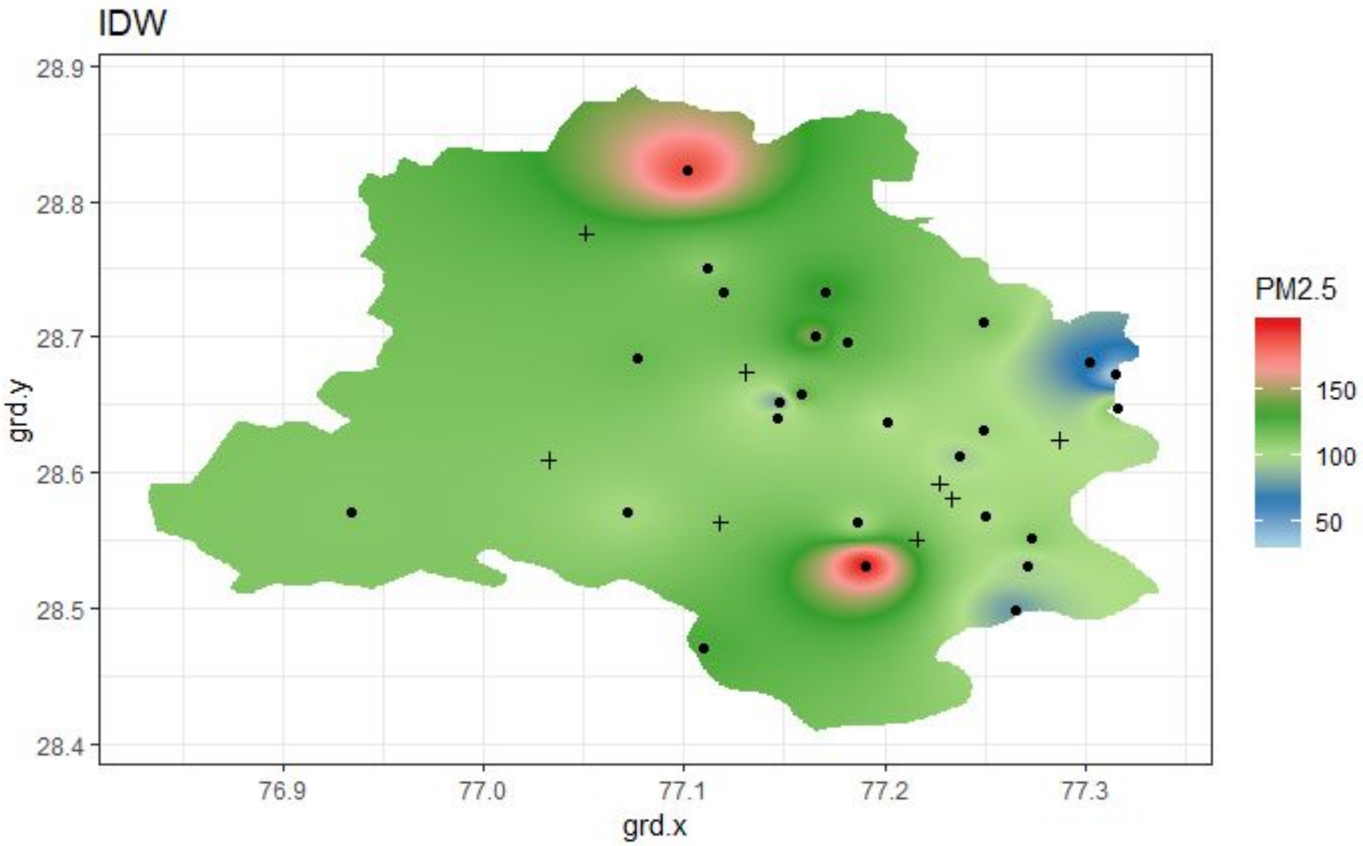
8 locations for prediction and cross validation using the built model.

Interpolation Results & Error Metrics

	X	Y	PM2.5	Predicted_PM2.5_Kriging	Predicted_PM2.5_IDW
1	77.05107	28.77620	134.79	143.05	127.02
2	77.11801	28.56278	102.81	119.65	114.05
3	77.23383	28.58028	95.53	101.76	104.72
4	77.22731	28.59182	79.18	98.41	103.42
5	77.03254	28.60909	100.79	100.97	112.08
6	77.28721	28.62375	91.58	108.03	99.13
7	77.13102	28.67404	101.39	110.57	110.79
8	77.21594	28.55042	85.91	131.35	122.68

	Ordinary Kriging	IDW Interpolation
Mean Absolute Error (MAE)	15.23	14.68
Root Mean Square Error (RMSE)	19.94	17.62
Mean Absolute Percentage Error (MAPE)	17%	16%


Kriging
28.9
28.8
28.7
28.6
28.5
28.4
grd.y
76.9
77.0
77.1
77.2
77.3
grd.x
PM2.5
160
120
80
40


IDW
PM2.5
150
100
50
28.9
28.8
28.7
28.6
28.5
28.4
grd.y
76.9
77.0
77.1
77.2
77.3
grd.x

Thank You!