DOI - 10.68648/ijte-q-2026-49-03-007

Identifying the Best Interpolation Technique for $PM_{2.5}$ Concentrations in Delhi

Sumit Suman
Assistant Professor
Dept. of Civil Engineering
Government Engineering College Khagaria
Khagaria, Bihar
✉ suman.sumit@geckhagaria.org.in

ABSTRACT

Continuous ambient air quality monitoring stations report $PM_{2.5}$ at a small number of discrete points, whereas exposure and health-impact studies require concentrations at arbitrary locations. This study identifies the interpolation technique that predicts $PM_{2.5}$ over Delhi most accurately and, unlike comparisons that adopt software defaults, optimises the governing parameters of every technique before comparing them. Four techniques available in ArcGIS Geostatistical Analyst were examined: inverse distance weighting (IDW), spline, kriging and cokriging. For each technique three parameters were varied systematically, and leave-one-out cross-validation across 24 monitoring stations gave the root mean square prediction error for every combination, in a spatial phase (10 dates) and a temporal phase (24 stations). Cokriging gave the lowest mean error in both phases, 26.21 and 26.03 µg/m3 respectively, improving on IDW by 6.9 and 4.3 per cent; wind speed, temperature and relative humidity were selected as the optimal secondary variables in both phases. The margin over the deterministic techniques narrows sharply in the temporal phase. Prediction surfaces generated with the optimised model reproduce the two annual pollution episodes reported for Delhi.

KEYWORDS *: $PM_{2.5}$, Spatial interpolation, Cokriging, Cross-validation, ArcGIS.*

NOTATIONS

$PM_{2.5}$	particulate matter of aerodynamic diameter < 2.5 µm (µg/m^3)
$Z_{(s0)}$	predicted value at an unsampled location
Z_i	measured value at the i^{th} sampled location
λ_i	weight assigned to the i^{th} measured value
d_i	distance between sampled and unsampled location
k	power function in IDW
γ(h)	semi-variance at lag distance h
n	number of measured values
P, M	predicted and measured concentration (µg/m^3)
S, M, T	sector type, neighbour count, technique parameter
ME, RMSE	mean and root mean square prediction error (µg/m^3)

INTRODUCTION

Air pollution causes a large share of premature mortality in urban areas; the World Health Organization attributes of the order of 0.8 million deaths a year worldwide to it. $PM_{2.5}$, particulate matter below 2.5 micrometres in aerodynamic diameter, is the most damaging of the regulated pollutants, since it remains suspended as an aerosol and penetrates deep into the respiratory tract. Delhi is among the worst affected cities in the world, and its $PM_{2.5}$ burden concentrates into two recurring episodes each dry season, the first coinciding with post-monsoon crop-residue burning in the upwind states [1].

Controlling emissions at source is the most economical strategy but is not always feasible, so there is growing demand for urban air quality assessment, which needs the concentration at any desired location so that it can be correlated with health outcomes there. In Delhi,

ambient $PM_{2.5}$ is measured at continuous ambient air quality monitoring (CAAQM) stations run by the Central Pollution Control Board (CPCB) and the Delhi Pollution Control Committee (DPCC). These are few and sit at discrete points, so concentrations elsewhere must be estimated by spatial interpolation, which rests on Tobler's first law: near things are more related than distant things [2].

Interpolation has been applied to Delhi before, but narrowly. Shukla et al. [3] mapped particulate matter across the city's supersites with ordinary kriging and IDW, reporting average errors of about 22 and 24 per cent. That study restricts itself to the primary pollutant and to a single family of techniques and, in common with the wider literature, runs each technique at the settings the software supplies. Li and Heap's review of comparative studies [4] found IDW, ordinary kriging and ordinary cokriging to be the most-used methods and identified data variation as a dominant control on their relative accuracy; yet the parameters governing how each method responds to that variation are seldom examined. Choi and Chong [5] adopt a power function of two for particulate matter on the authority of earlier work rather than testing it, and Hernández-Gress and Conchouso González [6], mapping pollutants over Mexico City, conclude that no single variogram model is universally optimal and that kriging accuracy shifts with hour and month.

Three questions therefore remain open: whether the defaults are appropriate for $PM_{2.5}$ in Delhi, whether meteorological covariates improve the prediction, and whether the ranking of techniques is stable between spatial and temporal interpolation. This study addresses all three by (i) optimising three governing parameters for each of IDW, spline, kriging and cokriging against the root mean square prediction error (RMSE); (ii) comparing the four optimised techniques in a spatial and a temporal phase using leave-one-out cross-validation; and (iii) generating prediction surfaces and time series for Delhi with the best technique.

LITERATURE REVIEW

Geostatistical interpolation is well established in meteorology. Ozturk and Kilic [7] mapped temperature and precipitation across the Aegean region of Turkey using ordinary kriging, taking geostatistical methods as a strong alternative to deterministic ones because they model the autocorrelation between measured points. Comparative studies nevertheless disagree on a single best technique. Bhunia et al. [8] compared five methods for soil organic carbon in West Bengal and found ordinary kriging superior on cross-validated RMSE, and Rabah et al. [9] ranked kriging first for water-level and chloride surfaces in the Gaza Strip. Yang et al. [10], in contrast, found IDW slightly better than the alternatives for daily rainfall over Greater Sydney, while Wei et al. [11], correcting geographically weighted regression residuals for $PM_{2.5}$ across China, found the completely regularised spline the most consistent. Simpson and Wu [12] found spline lowest in RMSE for lake sediment while kriging performed worst, and attributed that to kriging's assumptions of stationarity and normality not being met.

Read together with [4] and [6], the pattern is clear: the reported ranking is unstable, and much of that instability traces to parameters the studies did not vary. The case for meteorological covariates is separately well supported. Mangaraj et al. [13], using a dense sensor network over Delhi-NCR, found wind speed and boundary-layer height inversely correlated with surface $PM_{2.5}$ and concluded that local meteorology governs day-to-day concentrations. Cokriging is the natural way to bring such variables into an interpolator: it uses the autocorrelation of the primary variable together with its cross-correlation with up to three secondary variables, and cannot in principle do worse than kriging. Yet no study of Delhi has tested meteorological covariates inside a cokriging model.

METHODOLOGY

Study area and data

The study area is the National Capital Territory of Delhi. Data were obtained from 24 CPCB/DPCC CAAQM stations (Fig. 1) for $PM_{2.5}$ and five meteorological variables: wind speed (WS), wind direction (WD), relative humidity (RH), temperature (TEMP) and solar radiation (SR). Ten dates at ten-day intervals from 21 December 2019 to 21 March 2020 were used for the error analysis, and prediction surfaces generated at ten-day intervals from 1 October 2019 to 11 February

2020 so that both annual episodes were covered. All processing used ArcGIS 10.3 with the Geostatistical Analyst extension [14].

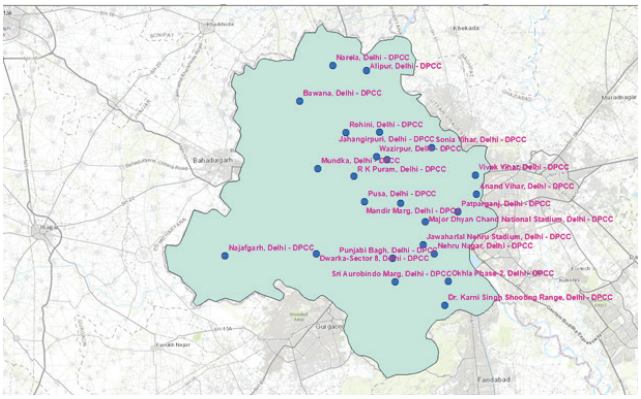


Fig. 1 The 24 CPCB/DPCC CAAQM stations used in the study, on the Delhi shapefile

Interpolation techniques

IDW predicts the value at an unsampled location as a distance-weighted mean of the measured values, the direct expression of Tobler's law:

$$Z(s_0) = \Sigma\, \lambda_i\, Z_i\,,\ \lambda_i = d_i^{-k} / \Sigma\, d_i^{-k} \qquad (1)$$

where Z is the $PM_{2.5}$ concentration, d the separation distance and k the power function; raising k reduces the influence of distant points, and k = 2 is the ArcGIS default [14]. Li and Heap [4] identify the power parameter as the main factor affecting IDW accuracy, which is why it is optimised here rather than assumed. Because the weights depend on distance alone, IDW makes no use of the structure of the data. Spline instead fits a surface through the measured points while minimising total curvature:

$$S(x, y) = T(x, y) + \Sigma\, \lambda_j\, R(r_j) \qquad (2)$$

with T and R set by the spline type: T1 (completely regularised) gives a smoother surface, T2 (with tension) one held closer to the sample range. Kriging is stochastic, predicting the value and quantifying its uncertainty; it uses the weighted-sum form of (1) but derives the weights from a semi-variogram fitted to the data, whose empirical semi-variance at lag h is

$$\gamma(h) = \tfrac{1}{2} \cdot \text{mean}\,[\,(Z_i - Z_j)^2\,] \qquad (3)$$

A stable, Gaussian, exponential or spherical model is fitted to the empirical semi-variogram, characterised by its nugget, sill and range (Fig. 2), and the weights follow from the resulting linear system. Because that system contains the covariances among the measured points themselves, kriging accounts for clustering in the network, which IDW cannot. Cokriging extends kriging by modelling $PM_{2.5}$ jointly with up to three secondary variables, so that both its autocorrelation and its cross-correlation with the covariates enter the weights.

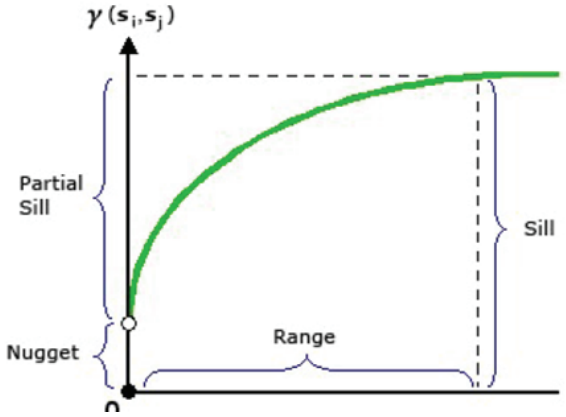


Fig. 2 Nugget, sill and range of a fitted semi-variogram

Parameters selected for optimisation

Three parameters were optimised per technique. Two are common to all four: the sector type S, controlling the shape of the search neighbourhood, and the neighbour count M, controlling its size; moving from S1 to S4 and M1 to M3 admits progressively more measured points into each estimate. The third is technique-specific: the power function k (1, 1.5, 2, 3, 4) for IDW; the spline type; the semi-variogram model for kriging; and for cokriging the choice of three secondary variables from WS, WD, RH, TEMP and SR. Levels are defined beneath Table 1.

Experimental set-up

Optimisation and comparison both rest on leave-one-out cross-validation. One parameter was varied at a time with the other two held fixed. For each setting one station was withheld, a surface was interpolated from the remaining 23, and the value at the withheld station was extracted from it. Repeating this for all 24 stations gives 24 independent prediction–measurement pairs, from which the mean error (ME) and the root mean square prediction error follow:

$$RMSE = [\ \Sigma\ (P - M)^2 / n\ ]^{½} \quad (4)$$

where P is the predicted value, M the measured value and n the number of stations. Plotting RMSE against the parameter level gives a minimum that identifies the optimum (Fig. 3). That setting was retained, the procedure repeated for the next parameter and then the next technique. The sequence was run twice: in the spatial phase the 24 stations on a single date form the validation set, repeated over ten dates; in the temporal phase the ten dates at a single station form it, repeated over 24 stations.

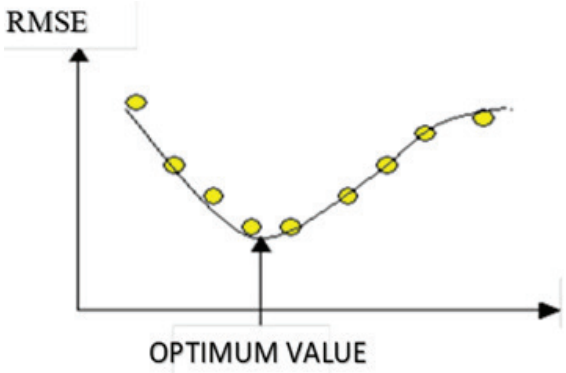


Fig. 3 Identifying the optimum level of a parameter from the RMSE minimum

RESULTS AND DISCUSSION

Optimum parameter values

Table 1 gives the parameter set that minimised RMSE for each technique in each phase. Two results stand out. First, the optimum neighbourhood is wide spatially (S3, M3) and narrow temporally (S1, M1): where the spatial field is the object of interest more stations stabilise the estimate, whereas with the location fixed, distant stations only add noise. Second, the same three secondary variables (wind speed, temperature and relative humidity) emerged as optimal for cokriging in both phases independently, which is physically reasonable and consistent with the meteorological control reported in [13]: wind governs dispersion, temperature and humidity govern boundary-layer stability and secondary aerosol formation.

The optimum power function for IDW was k = 2 spatially, confirming the default that [5] assumes, but the temporal phase gave k = 1 or k = 4 depending on the station, with no stable optimum. The default is therefore defensible for spatial work on $PM_{2.5}$ and not for temporal work. Only a parameter sweep can reveal that distinction, which is consistent with the hour- and month-dependence reported in [6].

Table 1. Optimum parameter values for each technique

Technique	Parameters optimised	Optimum (spatial)	Optimum (temporal)
IDW	Sector type, neighbours, power k	S1/S3, M3, k = 2	S1, M1, k = 1 or 4
Spline	Sector type, neighbours, spline type	S3, M3, T1	S1, M1, T1
Kriging	Sector type, neighbours, semi-variogram	S3/S4, M3, T1	S1, M1, T1
Cokriging	Secondary variables (3 of 5)	WS, TEMP, RH	WS, TEMP, RH

S1 = 1 sector; S2 = 4 sectors; S3 = 4 sectors, 45° offset; S4 = 8 sectors. M1 = 5/2, M2 = 10/6, M3 = 15/10 max/min neighbours. Spline: T1 = completely regularised, T2 = with tension. Kriging: T1 = stable, T2 = Gaussian, T3 = exponential, T4 = spherical

Spatial analysis

Table 2 lists the cross-validated RMSE of each optimised technique on the ten dates. The ranking is unambiguous: cokriging first, then kriging, then spline, with IDW last. A cokriging variant was best on nine dates out of ten, and averaged over the ten dates the best model reduced RMSE from 28.14 µg/m^3 for IDW to 26.21 µg/m^3, an improvement of 6.9 per cent. The ranking follows from what each technique uses: IDW derives its weights from distance alone; spline uses the concentrations but forces an exact fit; kriging derives weights from a fitted semi-variogram; cokriging adds the cross-correlation with meteorology. Each step up the ranking exploits additional structure.

Averaging conceals a difference between the two cokriging variants: the single-variable model has the marginally lower mean (26.21 against 26.27 µg/m^3), yet the three-variable model won on twice as many dates. The mean is dominated by 1 January 2020, the most variable date in the set, on which the simpler model did better (55.16 against 56.99 µg/m^3). Each extra covariate needs further autocorrelation parameters estimated,

and on such a day that added variance outweighs the information gained.

Table 2. Cross-validated RMSE (µg/m³) of each optimised technique for the spatial analysis

Date	IDW	Spline	Kriging	Cokriging (1 sec. var.)	Cokriging (3 sec. var.)
21 Dec 2019	31.45	31.65	31.07	29.98	**29.45**
1 Jan 2020	66.66	63.44	60.29	**55.16**	56.99
11 Jan 2020	39.56	40.17	**39.02**	**39.02**	**39.02**
21 Jan 2020	46.49	46.24	46.16	44.73	**43.85**
1 Feb 2020	22.31	21.25	20.73	19.75	**19.47**
11 Feb 2020	35.02	35.61	34.97	**34.86**	35.33
21 Feb 2020	8.23	8.03	7.97	7.93	**7.81**
1 Mar 2020	10.79	10.68	10.64	10.54	**10.24**
11 Mar 2020	5.75	5.76	5.74	5.52	**5.32**
21 Mar 2020	15.18	14.91	14.89	**14.62**	15.25
Mean RMSE	**28.14**	**27.77**	**27.15**	**26.21**	**26.27**
Mean ME	**+2.73**	**+1.84**	**+0.85**	**+0.57**	**+0.63**

Bold = lowest RMSE on that date

Temporal analysis

Table 3 gives the per-station results. Cokriging is again most accurate overall, 26.03 against 27.21 µg/m³ for IDW, but the margin narrows to 4.3 per cent and the ranking of the other three collapses: IDW, clearly worst spatially, is now second, and spline and kriging are separated by 0.02 µg/m³. Outright wins tell the same story: cokriging best at 9 stations, IDW at 8, spline at 6.

The explanation lies in what varies in each phase. Spatially, the field is $PM_{2.5}$ across Delhi on one day, and concentrations at different points in the city are broadly similar; the field is smooth and strongly autocorrelated, so a technique modelling that autocorrelation has much to exploit. Temporally, the location is fixed and the variation spans three months, over which $PM_{2.5}$ swings by more than an order of magnitude. Spatial autocorrelation is no longer the dominant structure, so cokriging's advantage is largely neutralised and the deterministic techniques close most of the gap. This is precisely the behaviour anticipated by [4], which identified data variation as a dominant impact factor and the resulting loss of accuracy as method-dependent. The choice of technique therefore matters far more for mapping than for gap-filling a station record.

Table 3. Cross-validated RMSE (µg/m³) of each optimised technique for the temporal analysis

Station	IDW	Spline	Kriging	Cokriging
Alipur	34.01	31.41	37.29	**31.36**
Anand Vihar	**12.97**	13.35	14.53	14.14
Ashok Vihar	7.19	**5.79**	7.73	7.18
Bawana	23.40	**16.89**	21.74	24.69
Dr. Karni Singh S.R.	20.44	16.38	22.82	**14.89**
Dwarka-Sector 8	**34.29**	51.60	37.95	39.00
Jahangirpuri	**20.19**	25.38	26.61	24.83
J. Nehru Stadium	17.48	17.91	**7.36**	10.18
Major Dhyan Chand Stadium	16.55	**6.89**	15.70	15.43
Mandir Marg	**20.75**	23.31	21.24	23.80
Mundka	49.41	**45.68**	49.82	51.40
Najafgarh	54.92	65.88	54.20	**52.38**
Narela	**12.14**	14.77	19.36	19.48
Nehru Nagar	58.63	55.27	56.00	**54.05**
Okhla Phase-2	24.93	**12.84**	16.92	13.33
Patparganj	17.58	20.54	17.41	**17.02**
Punjabi Bagh	**21.21**	28.28	27.11	21.75
Pusa	38.06	40.79	36.13	**32.48**
R K Puram	**46.29**	52.09	49.49	48.15
Rohini	25.57	**21.85**	23.50	24.94
Sonia Vihar	33.11	28.35	28.81	**25.90**
Sri Aurobindo Marg	34.73	38.85	36.57	**28.31**
Vivek Vihar	12.71	12.81	14.22	**10.99**
Wazirpur	**16.40**	17.27	21.11	19.13
Mean RMSE	**27.21**	**27.67**	**27.65**	**26.03**

Bold = lowest RMSE at that station.

Prediction surfaces

Surfaces were generated with the optimised cokriging model at ten-day intervals from 1 October 2019 to 11 February 2020. Predicted $PM_{2.5}$ ranged from 28.02–45.23 µg/m³ on 1 October 2019, the cleanest date, to 265.89–455.60 µg/m³ on 1 November 2019 and 256.50–493.63 µg/m³ on 1 January 2020 (Fig. 4). The clean-season surface is comparatively flat while both episode surfaces show a pronounced gradient, so the smoothness of the field is itself seasonal.

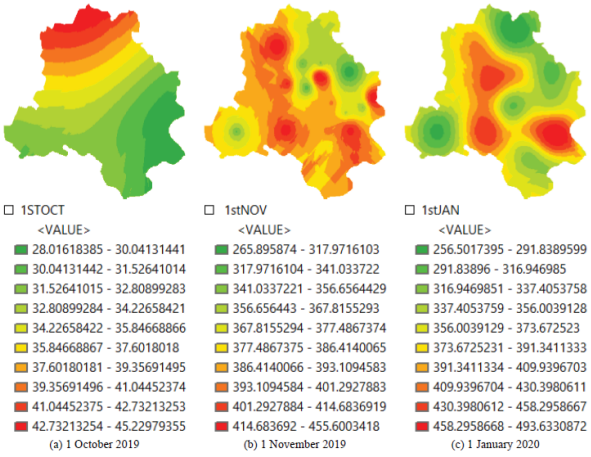
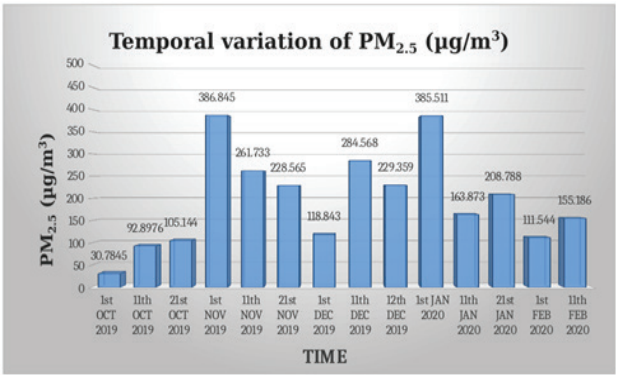


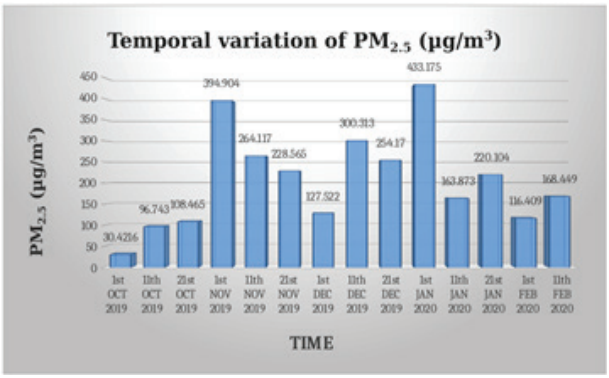
Fig. 4 Predicted $PM_{2.5}$ surfaces over Delhi from the optimised cokriging model (µg/m³); class breaks are given beneath each map

Temporal variation and episode reproduction

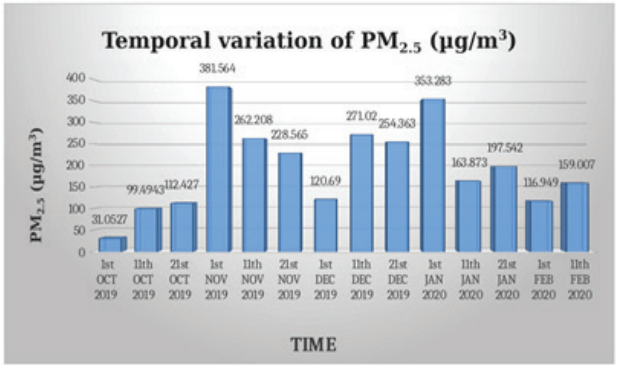
The same model was used to extract time series at three non-monitored locations (IIT Delhi, AIIMS and Connaught Place), the use case that motivates the study (Fig. 5). All three track one another closely and exceed the 24-hour national standard of 60 µg/m³ on every date after 1 October. Two peaks stand out: 1 November 2019 (382–395 µg/m³) and 1 January 2020 (353–433 µg/m³), matching the two dry-season episodes documented for the city. The first coincides with post-monsoon crop-residue burning upwind, whose contribution is itself contested: Kajino et al. [1] put it at 30-34 per cent of surface $PM_{2.5}$ over the Punjab-Haryana-Delhi region, rising to 50-56 per cent during plume transport events, whereas Mangaraj et al. [13] attribute only about 14 per cent of Delhi-NCR $PM_{2.5}$ to burning over Punjab. The second is attributed to local emissions and secondary formation under stable winter meteorology. That an independently optimised model reproduces both episodes, in the right months and with the right relative magnitudes, corroborates the predicted surfaces.



(a) IIT Delhi



(b) AIIMS



(c) Connaught Place

Fig. 5 Predicted $PM_{2.5}$ at three non-monitored locations, showing the two annual episodes

Systematic smoothing of extremes

The prediction–measurement pairs reveal a bias common to all four techniques. Fig. 6 plots predicted against measured $PM_{2.5}$ for the 24 stations; the fitted

line has a slope of 0.289, far below the 1:1 line, so high measured values are systematically under-predicted and low ones over-predicted. Every technique here estimates a value as a weighted combination of its neighbours, which can never exceed the range of those neighbours, so extremes are pulled towards the local mean.

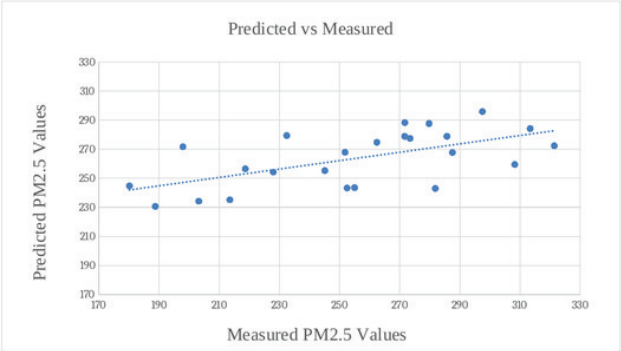


Fig. 6 Predicted versus measured $PM_{2.5}$ (µg/m³) for the 24 stations; fitted line y = 0.2887x + 189.98

Mean error across the network is nonetheless small, +0.57 µg/m³ for the best cokriging model against +2.73 µg/m³ for IDW (Table 2), because over-prediction at clean sites and under-prediction at polluted sites largely cancel. Interpolated surfaces are therefore reliable for city-wide averages but understate peak exposure at hotspots. This matters directly for the health-impact applications that motivate [3], and is a further reason to prefer cokriging, whose covariates reduce the effect.

CONCLUSIONS

Cokriging with wind speed, temperature and relative humidity as secondary variables is the most accurate of the four techniques for predicting $PM_{2.5}$ over Delhi, in both phases. The same three covariates were selected independently in each phase, so the result reflects the physics of the field rather than an artefact of one dataset. No previous study of Delhi had tested meteorological covariates inside a cokriging model.

Its advantage is conditional on what is being predicted. Spatially it is decisively better, with 6.9 per cent lower RMSE than IDW and the best result on nine of ten dates. Temporally, at a fixed location, the advantage narrows to 4.3 per cent and IDW becomes competitive. Reporting a single best interpolator without stating what is being interpolated is therefore misleading.

Optimising parameters before comparing techniques changes the comparison. The optimum search neighbourhood was wide spatially (S3, M3) and narrow temporally (S1, M1); k = 2 was confirmed for spatial interpolation of $PM_{2.5}$ but found unstable temporally. Rankings obtained at default settings measure the defaults as much as the techniques.

All four techniques smooth extremes, with a predicted-against-measured slope of 0.289, so interpolated surfaces are dependable for city-wide averages but understate hotspot peaks. The natural extension is to couple these predicted concentrations with health records at the same locations, so that the disease burden attributable to $PM_{2.5}$ can be quantified and mitigation planned accordingly.

ACKNOWLEDGEMENT

This work is carried out under the supervision of Prof. Mukesh Khare and Associate Professor, Dr. Sri Harsha Kota, Department of Civil Engineering, IIT Delhi. Their guidance throughout the work is gratefully acknowledged. The author also thanks the Central Pollution Control Board and the Delhi Pollution Control Committee for making the CAAQM data available online publicly.

REFERENCES

1. M. Kajino, K. Ishijima, J. Ching et al. (2025), "Impact of post-monsoon crop residue burning on $PM_{2.5}$ over northern India: optimizing emissions using a high-density in situ surface observation network", Atmospheric Chemistry and Physics, Vol. 25, Pp. 7137-7160. doi: 10.5194/acp-25-7137-2025

2. Waldo Tobler (2004), "On the First Law of Geography: A Reply", Annals of the Association of American Geographers, Vol. 94, No. 02, Pp. 304-310. doi: 10.1111/j.1467-8306.2004.09402009.x

3. Komal Shukla, Prashant Kumar, Gaurav S. Mann and Mukesh Khare (2020), "Mapping spatial distribution of particulate matter using Kriging and Inverse Distance Weighting at supersites of megacity Delhi", Sustainable Cities and Society, Vol. 54, Article 101997. doi: 10.1016/j.scs.2019.101997

4. Jin Li and Andrew D. Heap (2008), A Review of Spatial Interpolation Methods for Environmental Scientists,

Geoscience Australia, Record 2008/23, Canberra, 137 pp. GeoCat 68229, ISBN 978-1-921498-28-2.

5. Kanghyeok Choi and Kyusoo Chong (2022), "Modified Inverse Distance Weighting Interpolation for Particulate Matter Estimation and Mapping", Atmosphere, Vol. 13, No. 05, Article 846. doi: 10.3390/atmos13050846

6. Eva Selene Hernandez-Gress and David Conchouso Gonzalez (2026), "When Does Geostatistical Interpolation Work? Monthly and Hourly Sensitivity of Ordinary Kriging for Urban Air Pollutant Mapping in Mexico City", Algorithms, Vol. 19, No. 03, Article 213. doi: 10.3390/a19030213

7. Derya Ozturk and Fatmagul Kilic (2016), "Geostatistical Approach for Spatial Interpolation of Meteorological Data", Anais da Academia Brasileira de Ciencias, Vol. 88, No. 04, Pp. 2121-2136. doi: 10.1590/0001-3765201620150103

8. Gouri Sankar Bhunia, Pravat Kumar Shit and Ramkrishna Maiti (2018), "Comparison of GIS-based interpolation methods for spatial distribution of soil organic carbon (SOC)", Journal of the Saudi Society of Agricultural Sciences, Vol. 17, No. 02, Pp. 114-126. doi: 10.1016/j.jssas.2016.02.001

9. Fahid K. J. Rabah, Said M. Ghabayen and Ali A. Salha (2011), "Effect of GIS Interpolation Techniques on the Accuracy of the Spatial Representation of Groundwater Monitoring Data in Gaza Strip", Journal of Environmental Science and Technology, Vol. 04, No. 06, Pp. 579-589. doi: 10.3923/jest.2011.579.589

10. Xihua Yang, Xiaojin Xie, De Li Liu, Fei Ji and Lin Wang (2015), "Spatial Interpolation of Daily Rainfall Data for Local Climate Impact Assessment over Greater Sydney Region", Advances in Meteorology, Vol. 2015, Article ID 563629, 12 pp. doi: 10.1155/2015/563629

11. Peng Wei, Shaodong Xie, Liming Huang et al. (2023), "Spatial interpolation of regional $PM_{2.5}$ concentrations in China during COVID-19 incorporating multivariate data", Atmospheric Pollution Research, Vol. 14, No. 03, Article 101688. doi: 10.1016/j.apr.2023.101688

12. Greg Simpson and Yi Hwa Wu (2014), "Accuracy and Effort of Interpolation and Sampling: Can GIS Help Lower Field Costs?", ISPRS International Journal of Geo-Information, Vol. 03, No. 04, Pp. 1317-1333. doi: 10.3390/ijgi3041317

13. P. Mangaraj, Y. Matsumi, T. Nakayama et al. (2025), "Weak coupling of observed surface $PM_{2.5}$ in Delhi-NCR with rice crop residue burning in Punjab and Haryana", npj Climate and Atmospheric Science, Vol. 08, Article 18. doi: 10.1038/s41612-025-00901-8

14. ArcGIS 10.3 Help: Geostatistical Analyst, Environmental Systems Research Institute (Esri), Redlands, California, 2015. Available: https://desktop.arcgis.com/en/arcmap/10.3/