University of Tabriz

Journal of Civil and Environmental Engineering
Volume 51 (2022), Issue 4 (Winter), 141-152

EXTENDED ABSTRACT

Spatial Distribution of The Carbon Monoxid Using Common and Modern Interpolation Methods (Case study of Tehran)

Navid Hooshangi[a,*], Navid Mahdizadeh Gharakhanlou[b]

[a] *Department of Surveying Engineering, College of Earth Sciences Engineering, Arak University of Technology, Arak 3818146763, Iran*
[b] *Geospatial Information Science Division, Faculty of Geodesy and Geomatics Engineering, and Center of Excellence in Geo-Information Technology, K.N. Toosi University of Technology, Tehran 1996715433, Iran*

Received: 10 March 2020; **Accepted:** 20 December 2020

Keywords:
Carbon monoxide, spatial interpolation, common interpolation methods, Artificial neural network, Fuzzy Takagi-Sugeno.

1. Introduction

Air pollution is a major threat to public health, especially in the metropolises (Kalo et al., 2020). Due to the disadvantages of air pollution, understanding the various aspects of this issue is of great importance. Producing accurate air pollution maps plays an important role in managing and quantifying existing and future health risks (Alimissis et al., 2018).

Estimating the spatial distribution of air pollution continuously over a wide geographical area, especially in areas that have not been measured is a major concern in health studies (Masroor et al., 2020). Although spatial interpolation methods have been widely used in various applications to estimate unknown values in unsampled locations, many fundamental problems remain unresolved (Kalo et al., 2020). The superior methods extracted in previous research show that the results obtained in one phenomenon or one area are not extendable to all phenomena and places. Therefore, the evaluation and selection of interpolation techniques play an important role in the spatial zoning of CO pollution. Based on the results presented by García-Santos et al. (2020) and given reviewing the methods used in previous research, Inverse Distance Weight (IDW), Kriging (simple, ordinary, and universal), and Radial Base Function (RBF) methods were selected as common and classical methods of evaluation. New interpolation methods including artificial neural networks (ANN) and fuzzy-based methods have been developed in various fields. Alimissis et al. (2018) expressed the ability of ANN in predicting the pollutants of nitrogen dioxide, nitrogen monoxide, carbon monoxide, sulfur, and ozone. ANN and linear interpolations have also been used for daily nitrous oxide measurements (Bigaignon et al., 2020). In performed research, only the temporal forecast of air pollution in each station is considered and no spatial zoning is done. Tutmez and Hatipoglu (2010) compared the Takagi-Sugeno fuzzy method with fuzzy clustering and Universal Kriging in nitrate modeling so that their study demonstrated the superiority of fuzzy methods.

Since the spatial distribution of air pollutants is one of the major concerns of Tehran and authorities, the main objective of this research is to evaluate the capability of some proposed methods' functionality (e.g., ANN and Fuzzy Sugeno by Fuzzy C-means Clustering) along with the common interpolation methods (e.g., IDW, RBF and Simple Kriging (SK), Ordinary Kriging (OK) and Universal Kriging (OK)) in estimating carbon monoxide gas pollution.

* Corresponding Author
E-mail addresses: hooshangi@arakut.ac.ir (Navid Hooshangi), navid_mehdizade@email.kntu.ac.ir (Navid Mahdizadeh Gharakhanlou).

2. Methodology

In this study, CO data were collected in all air-pollution measuring stations in Tehran (34 stations) in 1397. The present research is carried out in three main stages: 1) Exploratory Spatial Data Analysis (ESDA), 2) Applying various interpolation methods, and 3) Investigation of results using the RMSE (Root Mean Square Error) and the MAE (Mean Absolute Error).

2.1. Exploratory Spatial Data Analysis (ESDA)

The first step is ESDA in order to search and prepare the data. In this phase, the sources of errors are analyzed using statistical analysis, trend analysis, as well as data distribution (Esri, 2020). Some interpolation methods such as Kriging need assuming the data distribution normal; therefore, the data distribution must be investigated before the execution in such interpolation methods. Data-trend (data's overall behavior regardless of small changes) is among the influential parameters in interpolation that needs to be examined at the ESDA stage.

2.2. Applying proposed interpolation methods as well as the common ones

In this phase, various and common interpolation methods (e.g., IDW, RBF, SK, OK and UK) in addition to the proposed interpolation methods (ANN and Fuzzy Takagi-Sugeno by Fuzzy C-means Clustering) were applied to the air pollution data. The IDW method proposes a linear combination of existing data and only takes into account distances, regardless of the position and arrangement of points (García-Santos et al., 2020). The Kriging method uses autocorrelation and statistical relationships between the measured points, so it is appropriate for phenomena having spatial correlations in terms of distance and direction (Belkhiri et al., 2020). The Kriging method is divided into different types, including simple, ordinary, and Universal. The Radial Basis Function (RBF) method assumes that there is no abrupt change in the surface. This method works by minimizing the overall curvature of the surfaces so that fits a mathematical surface (Ding et al., 2018). Neural network and fuzzy methods are new methods used in the field of interpolation. The artificial neural network is one of the computational methods that extracts the knowledge and the rule behind the information by processing the experimental data without considering the physics of the problem. One of the main categories of fuzzy inference systems is the Takagi-Sugeno fuzzy system (Ma et al., 2012). The fuzzy system visualizes input variables' space over output variables' space by using the concept of linguistic variables and fuzzy decision-making process. One of the advantages of this method is its simple implementation of different types of data (Hooshangi and Alesheikh, 2015).

2.3. Investigation of results using RMSE and MAE

In the third step, the accuracy of the discussed methods was investigated using RMSE and MAE. The purpose of this phase is to validate the quality and accuracy of the methods. In this research, cross-validation was used due to the lack of data.

3. Results and discussion

SPSS software and frequency analysis were used to identify outliers in the data. For each month, the type of distribution and trend analysis were studied. Due to the inaccurate density of stations, normal distribution was not observed in almost all months of the year. For ESDA and implementation of common interpolation methods, ArcGIS software was used and for the implementation of the Takagi-Sugeno fuzzy method as well as the ANN, MATLAB software was used. RMSE and MAE values were calculated for each of the interpolation methods. Table 1 and Table 2 present the results of different interpolation methods based on the RMSE and MAE evaluation parameters.

Table 1. Summary of interpolation results by optimizing each parameter based on RMSE (ppm unit)

	Farvardin	Ordibehesht	Khordad	Tir	Mordad	Shahrivar	Mehr	Aban	Azar	Dey	Bahman	Esfand
IDW	0.562	0.472	1.071	0.827	0.964	0.909	1.334	0.817	0.936	0.917	0.921	0.762
RBF	0.450	0.467	1.032	0.732	0.904	0.874	1.102	0.784	0.836	0.873	0.863	0.721
SK	0.527	0.445	1.056	0.755	0.933	0.812	1.227	0.757	0.746	0.962	0.915	0.741
OK	0.568	0.442	0.970	0.757	0.912	0.852	1.351	0.737	0.720	0.896	0.915	0.760
UK	0.581	0.440	1.054	0.724	0.914	0.857	1.234	0.781	0.762	0.962	0.913	0.712
NN	0.539	0.586	1.152	0.834	0.913	0.889	1.279	0.794	0.845	0.945	0.952	0.713
FS	0.349	0.442	1.092	0.696	0.910	0.786	1.082	0.792	0.729	0.899	0.918	0.662

Table 2. Summary of interpolation results by optimizing each parameter based on MAE (ppm unit)

	Farvardin	Ordibehesht	Khordad	Tir	Mordad	Shahrivar	Mehr	Aban	Azar	Dey	Bahman	Esfand
IDW	0.254	0.284	0.428	0.331	0.453	0.328	0.463	0.323	0.421	0.436	0.451	0.354
RBF	0.229	0.289	0.462	0.272	0.407	0.336	0.383	0.278	0.387	0.357	0.324	0.363
SK	0.272	0.273	0.438	0.247	0.437	0.344	0.429	0.291	0.364	0.462	0.415	0.355
OK	0.232	0.276	0.368	0.261	0.423	0.333	0.438	0.273	0.323	0.367	0.367	0.347
UK	0.248	0.262	0.442	0.274	0.441	0.343	0.484	0.296	0.373	0.433	0.423	0.339
NN	0.259	0.272	0.491	0.312	0.462	0.328	0.442	0.332	0.428	0.447	0.442	0.352
FS	0.205	0.293	0.378	0.244	0.436	0.319	0.355	0.287	0.374	0.368	0.426	0.333

As given in Tables 1 and 2, the errors driven from interpolation methods are high (at least RMSE = 0.349 and MAE = 0.205). The reasons can be due to the low number of pollution detection stations, the inappropriate dispersion of stations, as well as the lack of station information in most months of 1397. Among the existing methods, the fuzzy Sugeno, RBF, Ordinary Kriging, and Universal Kriging were selected as the optimal methods in the calculation of 5 months, 3 months, 3 months, and 1 month, respectively. The main reason for the fuzzy method's result in this study was the independence of the fuzzy method's function from the normal distribution of data. On the other hand, fuzzy methods not only are suitable methods for modeling complex systems, but they are also more flexible. Based on RMSE and MAE, IDW, Simple Kriging, and neural network methods were not selected as suitable methods in different months. It is due to the low number of data. Regarding the low number of data, the neural network cannot be trained well on the basis of the differences between stations. In the IDW method also, the low density of stations actually decreases the intensity of the spatial correlation rule.

4. Conclusions

Accurate spatial distribution of air pollution plays an important role in managing air pollution reduction. In this study, the capability of common interpolation methods along with some proposed interpolation methods in modeling the amount of air pollution was investigated in Tehran. This study indicated that the estimated fuzzy method performed better than the other methods according to RMSE and MBE evaluation parameters. This expresses the high capability of this method in the interpolation. The reason is that the fuzzy Sugeno method initially separates the data with a fuzzy classification and then applies interpolation on points. IDW and ANN methods were not selected as appropriate methods in different months. In contrast, the Kriging method, which uses statistical data, and the RBF method, had better results. The results of this research may not be expandable in the zoning of other geographical phenomena; however, applying and investigating these interpolation methods for different phenomena as well as different regions are highly recommended.

5. References

Alimissis A, Philippopoulos K, Tzanis CG, Deligiorgi D, "Spatial estimation of urban air pollution with the use of artificial neural network models", Atmospheric Environment, 2018, 191, 205-213.

Belkhiri L, Tiri A, Mouni L, "Spatial distribution of the groundwater quality using kriging and Co-kriging interpolations", Groundwater for Sustainable Development, 2020, 11, 100473.

Bigaignon L, Fieuzal R, Delon C, Tallec T, "Combination of two methodologies, artificial neural network and linear interpolation, to gap-fill daily nitrous oxide flux measurements", Agricultural and Forest Meteorology, 2020, 291, 108037.

Ding Q, Wang Y, Zhuang D, "Comparison of the common spatial interpolation methods used to analyze potentially toxic elements surrounding mining regions", Journal of Environmental Management, 2018, 212, 23-31.

Esri, Understanding interpolation analysis. Esri, http://resources.arcgis.com/en/ help/main/10.1/index.html #//009z0000006w000000, 2020.

García-Santos G, Scheiber M, Pilz J, "Spatial interpolation methods to predict airborne pesticide drift deposits on soils using knapsack sprayers", Chemosphere, 2020, 258, 127231.

Hooshangi N, Alesheikh AA, "Evaluation of ANN, ANFIS and fuzzy systems in estimation of solar radiation in Iran", Journal of Geomatics Science and Technology, 2015, 4, 187-200.

Kalo M, Zhou X, Li L, Tong W, Piltner R, "Chapter 8- Sensing air quality: Spatiotemporal interpolation and visualization of real-time air pollution data for the contiguous United States", LI, L., ZHOU, X. & TONG, W. (eds.) Spatiotemporal Analysis of Air Pollution and Its Application in Public Health, Elsevier,2020.

Ma S-Q, Chen F-C, Wang Q, Zhao Z-Q, "Sugeno Type Fuzzy Complex-Value Integral and Its Application in Classification", Procedia Engineering, 2012, 29, 4140-4151.

Masroor K, Fanaei F, Yousefi S, Raeesi M, Abbaslou H, Shahsavani A, Hadei M, “Spatial modelling of PM2. 5 concentrations in Tehran using Kriging and inverse distance weighting (IDW) methods”, Journal of Air Pollution and Health, 2020, 5, 89-96.

Tutmez B, Hatipoglu Z, “Comparing two data driven interpolation methods for modeling nitrate distribution in aquifer”, Ecological Informatics, 2010, 5, 311-315.