Evaluation of Geographic Information Systems-Based Spatial Interpolation Methods Using Ohio Indoor Radon Data

Ashok Kumar, Akhil Kadiyala* and Dipsikha Sarmah

Department of Civil Engineering, The University of Toledo, Toledo, OH, USA 43606

Abstract: This paper evaluates the performance of six different Geographic Information System based interpolation methods: inverse distance weighting (IDW), radial basis function (RBF), global polynomial interpolation, local polynomial interpolation, kriging, and cokriging, using the Ohio homes database developed between 1987 and 2011. The best performing interpolation method to be used in the prediction of radon gas concentrations in the unmeasured areas of Ohio, USA was determined by validating the model predictions with operational performance measures. Additionally, this study performed a zip code level-based analysis that provided a complete picture of the radon gas concentration distribution in Ohio.

The RBF method was identified to be the best performing method. While the RBF method performed significantly better than the IDW, it was statistically similar to the other interpolation methods. The RBF predicted radon gas concentration results indicated a significant increase in the number of zip codes that exceeded the United States Environmental Protection Agency and the World Health Organization action limits, thereby, indicating the need to mitigate the Ohio radon gas concentrations to safe levels in order to reduce the health effects. The approach demonstrated in this paper can be applied to other radon-affected areas around the world.

1. INTRODUCTION

Radon (^{222}Rn) gas is a naturally occurring radioactive gas (properties: colorless, odorless, and chemically inert), formed from radium (^{226}Ra), by the decay of uranium (^{238}U) in geological materials. 'Radon gas drifts upward through the ground to the surface of the soil and flows into the outdoor air or seeps into the buildings through the foundation cracks and other openings' [1]. While the geological factors primarily control the indoor radon gas levels, the physical characteristics of the houses (penetration factor, air exchange rate, depressurization effect) were observed to be important in the areas where there are no known or significant geological sources of radon [2]. Radon gas concentrations in homes can easily be monitored using devices such as the charcoal canister, the alpha track detector, the positive barrier, the electrostatic radon monitor, the scintillation counter, the ionization chamber, etc. The radon gas concentrations are generally expressed in alpha particles and the units for radon gas are "pico-curies per liter of air" (pCi/l). 'The United States Environmental Protection Agency's (USEPA's) action level for public safety is 4 pCi/l, where in immediate

measures must be taken to reduce the level to 2 pCi/l' [3]. However, the World Health Organization designated the indoor radon gas concentration action limit to 2.7 pCi/l [4].

'Radon gas exposure is the second major cause of lung cancer after cigarette smoking and accounts for 3 to 14% of the global lung cancer incidences' [4]. The cancerous risk due to radon gas exposure poses a major challenge to environmentalists and healthcare professionals, considering that people spend 90% of their time indoors [5] and a majority of the radon-induced cancerous incidences are a result of the prolonged exposure at lower to moderate radon gas concentrations [4]. It is practically not feasible to measure the radon gas concentrations in each and every house, and to mitigate the radon gas concentrations to safe limits, as it will be a time consuming task requiring large financial investments on a long-term basis. Under such circumstances, it is essential that one explores the available cost-effective alternatives that assist in determining the radon gas concentrations in unmeasured areas with a certain degree of confidence. Geographic Information Systems (GIS) based interpolation methods provide one such alternative, wherein a representative sample of the actual monitored radon gas concentrations for a designated area is adequate to generate a spatially interpolated surface of the radon gas concentrations that can be utilized in the prediction of radon gas concentrations in the unmeasured areas within the designated area.

*Address correspondence to this author at the Department of Civil Engineering, The University of Toledo, Toledo, OH, USA 43606; Tel: 419-530-8136; Fax: 419-530-8116; E-mail: akhilkadiyala@gmail.com

GIS offers researchers a variety of spatial interpolation methods: the inverse distance weighting (IDW), the radial basis function (RBF), the global polynomial interpolation (GPI), the local polynomial interpolation (LPI), the kriging, and the cokriging. Several studies demonstrated the applicability of GIS interpolation methods in estimating air quality data [6-11]. There are only a limited number of studies that investigated the use of GIS interpolation methods in examining radon gas data [12-16]. All the aforementioned studies either used a single interpolation method or a combination of two to three interpolation methods in examining their respective datasets, complimented by validation of the results with only limited number of statistical indicators. None of the prior studies compared and evaluated the performance of all six GIS interpolation methods using the radon gas data with a comprehensive set of operational performance measures and this research study aims to fill that knowledge gap. The homes database from the Ohio Radon Information Systems (ORIS) was used in this study to predict the radon gas concentrations and the performance of the six GIS interpolation methods were validated using operational performance measures recommended by Kadiyala and Kumar [17]. The identification of the best interpolation method using operational performance measures can help in not only making accurate predictions of the radon gas concentrations in unmeasured areas in Ohio, but also enables Ohio public health officials in implementing different control strategies to reduce the health impact.

2. METHODOLOGY

2.1. Study Area and Database Development

'The major sources of radon gas concentrations in the state of Ohio, USA are 'Ohio shale' and soil. Organic shale in Ohio is known to have elevated concentrations (10 ppm to 40 ppm) of uranium, five to 20 times the average levels in the earth's crust' [18, 19]. The USEPA estimates that radon gas exposure accounts for 900 (14%) cancer deaths per year in Ohio. The Air Pollution research group (APRG) of the Civil Engineering Department at the University of Toledo (UT) has been obtaining the data for radon gas concentration in homes in Ohio from various county health departments, commercial testing services, and university researchers since 1987. This enabled the APRG team at UT to compile the ORIS that included the homes database accumulated over 20 years. A detailed timeline of the development of the different components (homes, water, testers, and mitigation) of the ORIS and the corresponding statistics were discussed elsewhere [20].

Of the 1,862 zip codes in Ohio, radon gas concentrations were measured in 1,569 zip codes. Of the 1,569 zip codes, 861 zip codes had the number of radon gas concentration data points less than or equal to 20. This study adopted the use of 208,097 radon gas data points measured across the remaining 708 zip codes in Ohio, with each zip code having more than 20 measured radon gas concentration data points to account for the ample representation in each measured zip code. None of the prior studies on interpolation schemes used such a huge database with ample representation for each zip code. As the radon gas data are heavily skewed towards higher concentrations, the geometric mean (GM) of radon gas concentration data was used to input as a point source datum. Of the 708 zip codes, 22 zip codes were not represented in the Ohio zip code shape file obtained from the ESRI website [21]. The resulting 686 zip codes were thus used as the point source data inputs for the GIS interpolation methods.

'Uranium data for Ohio were extracted from the map published by Duval' [22]. The distribution map of uranium concentrations in Ohio state sediments and soils is available online [23]. 'A map of the Ohio's zip code areas was drawn to the same scale and overlaid on the uranium map, which resulted in obtaining the corresponding uranium concentrations for respective zip code areas. Each line of the uranium data file contained a zip code number, followed by three coded numbers that were representative of the modal, the maximum, and the minimum uranium concentrations, respectively' [15].

2.2. GIS Based Interpolation Methods

Of the six GIS based interpolation methods considered in this study, the IDW, the RBF, the GPI, and the LPI are referred as the deterministic interpolation methods, while the kriging and the cokriging are referred as the geostatistical interpolation methods. 'The deterministic interpolation methods generate a surface from the measured data points based on the degree of similarity or smoothing, while the geostatistical methods generate a surface using the statistical properties of the measured data points after consideration of the spatial configuration of the measured data points around the prediction location' [21]. Based on the data points used in generating a surface, one can further classify the deterministic interpolation methods as global and local interpolation methods. The global interpolation method uses all the measured sample data points in making the predictions, while the local interpolation method uses only a smaller sample of the measured data points from the neighborhood within a larger area. The GPI is categorized as the global interpolation method, while the IDW, the RBF, and the LPI are categorized as the local interpolation methods. Alternatively, one can also categorize the deterministic methods based on whether the surface generated passes through the monitored data points or not. When the surface passes through the measured points, the interpolation methods are referred as the exact interpolators and if not, they are referred as the inexact interpolators. The IDW and the RBF are the exact interpolators, while the GPI and the LPI are the inexact interpolators.

'The IDW method works on the principle of influential weightage that varies directly with the closeness of the measured values to the prediction location, i.e., the points closest to the prediction location are provided more weightage and the weights reduce as a function of distance from the prediction location' [21]. The RBF method is a special case of the splines that aims at fitting a surface through the measured sample data points with the minimum possible total curvature of surface. The GPI method fits a surface between the measured sample data points using a polynomial func-

tion. The LPI method also fits a surface between the measured sample data points using a polynomial function as is the case with the GPI method; the exception being that the LPI method fits multiple polynomials within the specified overlapping neighborhoods. The kriging method interpolates the value of a sample data point at a prediction location from the sample data point values at nearby measured locations. Kriging makes a prediction for an unknown value of a specific location using the fitted model obtained from quantification of the spatial data structure. The cokriging method interpolates the value of a sample data point at a prediction location from the observations of multivariable (two or more) data point values at nearby measured locations based on the auto-correlations and cross-correlations among the multivariable data. Additional details on the six interpolation methods inclusive of the mathematical expressions were documented elsewhere [21, 24].

2.3. Approach to Data Analysis

The Ohio zip code shape file obtained from the ESRI website consisted of 1,862 zip codes. The 686 zip codes (with radon gas concentration records greater than 20) were used as the input data to predict the radon gas concentrations in the remaining 1,176 zip codes of Ohio with the ArcGIS geostatistical analyst using the interpolation methods of the IDW, the RBF, the GPI, the LPI, the kriging, and the cokriging. The measured 686 zip codes were designated with the corresponding radon gas GM concentrations and zero values were assigned to the remaining 1,176 zip codes. The polygon features of the Ohio zip code shape file were then transformed into point features using the Data Management toolbox in ArcGIS. The point featured shape file was then divided into two: one shape file with 686 zip codes having radon gas GM concentration data and the second shape file with 1,176 zip codes having no radon gas GM concentration data. The Exploratory Spatial Data Analysis (ESDA) tools in the ArcGIS geostatistical analyst were used to assess the statistical properties of the radon gas GM data and to visualize the spatial distribution before generating the surfaces with the six interpolation methods. The Trend Analysis tool revealed the radon gas GM concentrations to be higher near Central Ohio and decreased towards the north and south directions.

The point shape file with 686 zip codes was divided to obtain 80% of the data that was used as training data for the generation of an output surface, and the remaining 20% data were used to validate the output surface by comparing the actual and predicted radon gas GM concentration values. The division of 686 zip codes into training (80%) and validation (20%) was based on the observation that the 80-20 combination provided the least root mean square error on performing the sensitivity analysis for the divisions of 90-10, 80-20, 70-30, and 60-40. Performance measures were then used to identify the best interpolation method on validating the actual and predicted radon gas GM concentration data values using the validation dataset and the identified best interpolation method generated surface was used to predict the radon gas GM concentrations in the remaining 1,176 Ohio zip codes.

2.4. Model Evaluation

Based on a comprehensive review of the literature on air quality model evaluation, Kadiyala and Kumar [17] enumerated a complete list of performance measures to validate indoor and atmospheric air quality models with emphasis on the modeling attributes of the extreme-end (i.e., peak-end/low-end) concentrations and the mid-range concentrations. 'The statistical performance measures, namely, correlation (CORR), slope of the regression line (m), ratio of the intercept of regression line to the average observed concentrations (b/C_0), normalized mean square error (NMSE), fractional bias (FB), fractional variance (FV), fraction of predictions within a factor of two of the observations (FA2), model comparison measure (MCM_2), geometric mean bias (MG), geometric mean variance (VG), revised index of agreement (IOA_r), scatter plots, quantile-quantile (Q-Q) plots, and bootstrap 95% confidence interval estimates over NMSE, FB, CORR, VG, and MG were recommended to be used for mid-range indoor air quality (IAQ) model validation. The study also recommended a set of primary performance measures (spearman rank correlation coefficient (ρ), m, b/C_0, FV, FA2, robust highest concentration ratio (RHC_{ratio}), MCM_2, MG, VG, scatter plots, Q-Q plots, bootstrap 95% confidence interval estimates over MG and VG) and a set of secondary performance measures (IOA_r, accuracy of paired peaks (A_p)) for IAQ model validation, when the emphasis was on the peak-end concentrations. With the exception of RHC_{ratio} and A_p, all other performance measures recommended for the peak-end IAQ model validation were recommended to be used for the low-end IAQ model validation' [17]. Additional details on the criteria used for selection and ranking of the operational model performance measures, based on the attributes of extreme-end and mid-range IAQ modeling are provided elsewhere [17]. This study validated the performance of the six interpolation methods using the criteria established by Kadiyala and Kumar [17, 25-27] and Kadiyala *et al.* [28].

Based on the recommendations made by Kadiyala and Kumar [17, 25-27] and Kadiyala *et al.* [28], an IAQ model is deemed acceptable from both extreme-end and mid-range modeling perspectives, if it meets the criteria of (i) $0.9 \leq$ CORR ≤ 1.0, (ii) $0.75 \leq$ m ≤ 1.25, (iii) $-25 \leq b/C_0$ (%) ≤ 25, (iv) $0 \leq$ NMSE ≤ 0.25, (v) $-0.25 \leq$ FB ≤ 0.25, (vi) $-0.5 \leq$ FV ≤ 0.5, (vii) $0.8 \leq$ FA2 ≤ 1.2, (viii) $0 \leq MCM_2 \leq 1.2$, (ix) $0.8 \leq$ MG ≤ 1.2, (x) $0.8 \leq$ VG ≤ 1.2, (xi) $0.8 \leq IOA_r \leq 1.0$, (xii) $0.9 \leq \rho \leq 1.0$, (xiii) $0.8 \leq RHC_{ratio} \leq 1.2$, and (xiv) $-15 \leq A_p \leq 15$. In addition to the above mentioned criteria, the degree of closeness of plotted points in the scatter and Q-Q graphical representations to the identity line can help determine the superiority of one model over the other, while the bootstrap estimates over NMSE, FB, CORR, VG, and MG provide the 95% confidence interval estimates. 'In this study, the performance measures of m, b/C_0, FV, ρ, RHC_{ratio}, MCM_2, IOA_r, and A_p were computed using Microsoft® Excel 2010; while, CORR, NMSE, FB, FA2, MG, VG, and bootstrap 95% confidence interval estimates over NMSE, r, FB, MG, and VG were computed using the BOOT v2.0. Graphical representations of the scatter plots and the Q-Q plots were obtained from using the MINITAB® 16 and MathWorks® MATLAB 2010b software, respectively' [17].

Table 1. Operational performance measures computed for the six interpolation methods.

Model(s)	M	SD	CORR	m	b/C$_o$(%)	NMSE	FB	FV	FA2	MCM$_2$	IOA$_r$	ρ	RHC$_{ratio}$	A$_p$	MG	VG
	< ------------------------------arithmetic values------------------------------>														Log Values	
Obs.	3.560	1.970	1.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	1.000	1.000	0.000	1.000	1.000
IDW	3.390	1.340	0.635	0.431	52.197	0.200	0.048	0.383	0.942	0.717	0.659	0.746	0.819	-68.523	0.987	1.140
RBF	3.430	1.370	0.666	0.462	50.331	0.180	0.036	0.362	0.942	0.695	0.558	0.764	0.842	-67.514	0.977	1.130
GPI	3.330	1.190	0.608	0.368	56.832	0.210	0.066	0.492	0.883	1.082	0.523	0.694	0.735	-62.419	1.068	1.820
LPI	3.360	1.290	0.648	0.423	52.238	0.190	0.056	0.419	0.942	0.746	0.486	0.721	0.787	-65.625	0.995	1.160
Kriging	3.410	1.240	0.676	0.426	53.398	0.180	0.041	0.454	0.927	0.723	0.564	0.742	0.789	-63.983	0.971	1.150
Cokriging	3.420	1.200	0.678	0.412	54.916	0.180	0.039	0.487	0.942	0.732	0.531	0.744	0.782	-64.037	0.965	1.150

Table 2. Bootstrap 95% confidence limit estimates over NMSE, FB, CORR, MG, and VG for individual interpolation methods comparison.

Model (s)	95% Confidence Limits (Lower Limit – Upper Limit)				
	NMSE	**FB**	**CORR**	**MG**	**VG**
IDW	0.114 – 0.280	-0.020 – 0.120	0.522 – 0.745	-0.067 – 0.053	0.099 – 0.170
RBF	0.101 – 0.259	-0.031 – 0.105	0.556 – 0.772	-0.077 – 0.041	0.092 – 0.161
GPI	0.130 – 0.291	-0.006 – 0.146	0.522 – 0.708	-0.074 – 0.215	-0.040 – 1.280
LPI	0.115 – 0.270	-0.013 – 0.129	0.555 – 0.740	-0.065 – 0.070	0.114 – 0.187
Kriging	0.101 – 0.251	-0.025 – 0.112	0.583 – 0.771	-0.087 – 0.040	0.100 – 0.170
Cokriging	0.102 – 0.250	-0.027 – 0.111	0.595 – 0.764	-0.094 – 0.037	0.101 – 0.174

3. RESULTS AND DISCUSSION

The following sections provide more information on the validation results obtained for the six GIS interpolation methods in addition to providing a zip code based analysis using the identified best interpolation method.

3.1. Evaluation of the Six Interpolation Methods

Table **1** presents a summary of the mean (M) statistic, the standard deviation (SD) statistic, and the operational performance measures recommended by Kadiyala and Kumar [17, 25-27] and Kadiyala *et al.* [28] for validating the six interpolation methods. An ideal interpolation method must meet all the stringent criteria recommended by Kadiyala and Kumar [17, 25-27] and Kadiyala *et al.* [28]. However, it is rare to find an air quality modeling technique that meets all the stringent guidelines. From Table **1**, one can note that none of the six interpolation methods considered in this study met the criteria for CORR, m, b/C$_0$(%), IOA$_r$, ρ, and A$_p$. All the six interpolation methods met the criteria for NMSE, FB, FV, FA2, MCM$_2$, RHC$_{ratio}$, and MG. With the exception of the GPI method, all the remaining interpolation methods met the VG acceptance criteria. The RBF method computed operational performance measures of CORR, m, b/C$_0$(%), NMSE, FB, FV, FA2, MCM$_2$, VG, ρ, and RHC$_{ratio}$ were much closer to the corresponding ideal values in com-

parison with the performance measures obtained using the other five interpolation methods.

Table **2** presents a summary of the bootstrap 95% confidence intervals for individual interpolation methods comparison over NMSE, FB, CORR, MG, and VG. An interpolation method is considered to exhibit a significant difference to zero property when there is no change in the sign between lower and upper confidence limits. From Table **2**, one can note that all the six interpolation methods exhibited the significant difference to zero property over NMSE and CORR. None of the interpolation methods exhibited the significant difference to zero property over FB and MG. With the exception of the GPI method, all the other interpolation methods were significantly different to zero when VG statistic was considered.

Table **3** presents a summary of the bootstrap 95% confidence intervals for between interpolation methods comparison over NMSE, FB, CORR, MG, and VG. The RBF method was significantly different to the IDW method with respect to all the considered statistics, i.e., NMSE, FB, CORR, MG, and VG. The combinations of IDW – GPI, IDW – Kriging, RBF – GPI, RBF – Kriging, RBF – Cokriging, GPI – LPI, and GPI – Kriging were observed to be statistically similar, irrespective of the statistic considered. The IDW – LPI methods combination exhibited the signifi-

Table 3. Bootstrap 95% confidence limit estimates over NMSE, FB, CORR, MG, and VG for between interpolation methods comparison.

Model (s)	95% Confidence Limits (Lower Limit – Upper Limit)				
	NMSE	FB	CORR	VG	MG
IDW – RBF	0.008 – 0.026	0.004 – 0.022	-0.050 – -0.011	0.004 – 0.013	0.003 – 0.019
IDW – GPI	-0.052 – 0.025	-0.054 – 0.021	-0.064 – 0.018	-1.137 – 0.166	-0.212 – 0.043
IDW – LPI	-0.013 – 0.022	-0.025 – 0.008	-0.049 – 0.021	-0.030 – -0.003	-0.029 – 0.011
IDW – Kriging	-0.003 – 0.045	-0.011 – 0.024	-0.095 – 0.008	-0.017 – 0.015	-0.005 – 0.034
IDW – Cokriging	-0.004 – 0.047	-0.012 – 0.028	-0.100 – 0.008	-0.022 – 0.016	0.002 – 0.041
RBF – GPI	-0.067 – 0.007	-0.065 – 0.007	-0.025 – 0.131	-0.157 – 1.146	-0.034 – 0.223
RBF – LPI	-0.004 – 0.028	0.005 – 0.038	-0.049 – 0.017	0.012 – 0.037	-0.002 – 0.043
RBF – Kriging	-0.017 – 0.026	-0.022 – 0.009	-0.057 – 0.030	-0.024 – 0.005	-0.011 – 0.022
RBF – Cokriging	-0.018 – 0.027	-0.024 – 0.013	-0.060 – 0.029	-0.029 – 0.007	-0.008 – 0.030
GPI – LPI	-0.014 – 0.047	-0.022 – 0.039	-0.098 – 0.032	-0.182 – 1.120	-0.056 – 0.204
GPI – Kriging	-0.001 – 0.066	-0.006 – 0.054	-0.130 – 0.010	-0.165 – 1.133	-0.028 – 0.226
GPI – Cokriging	0.006 – 0.045	-0.003 – 0.054	-0.121 - -0.005	-0.167 – 1.130	-0.021 – 0.231
LPI – Kriging	0.001 – 0.034	0.002 – 0.028	-0.073 – 0.013	0.005 – 0.026	0.012 – 0.039
LPI – Cokriging	0.003 – 0.031	0.003 – 0.029	-0.068 – 0.004	0.001 – 0.025	0.016 – 0.045
Kriging – Cokriging	-0.011 – 0.011	-0.006 – 0.009	-0.026 – 0.022	-0.008 – 0.004	0.001 – 0.011

cant difference to zero property only for the VG statistic, while the combinations of IDW – Cokriging and Kriging – Cokriging exhibited the significant difference to zero property only for the MG statistic. The RBF – LPI methods combination was statistically similar for the NMSE, CORR, and MG statistics, and were significantly different when the statistics considered were FB and VG. The GPI – Cokriging methods combination was statistically similar for the FB, VG, and MG statistics; the combination yielded significantly different results over NMSE and CORR. The LPI– Kriging and LPI – Cokriging methods combination exhibited the significant difference to zero property for all the statistics with the exception of CORR.

Figs. (**1** and **2**) present the scatter plots and the Q-Q plots obtained for the six interpolation methods. The higher degree of closeness of the plotted points to the 1:1 identity line (45° line) indicate the better performing interpolation method. Figs. (**1** and **2**) indicate that RBF performed better as the plotted points were much closer to the 1:1 identity line in comparison with other interpolation methods.

Based on the degree of closeness of the computed performance measures to the ideal values (Table **1**) and the plotted points to the 1:1 identity line (Figs. **1** and **2**), one can conclude that the RBF method performed the best in predicting the radon gas GM concentrations. The bootstrap 95% confidence interval estimates indicate that the RBF method performed significantly differently in comparison with the IDW method and performed similarly with the GPI, the kriging, and the cokriging methods. This suggests that the RBF method will always perform better than the IDW method; the same pattern cannot be expected in comparison with the GPI, the kriging, and the cokriging methods on using other radon datasets. Therefore, one can always prefer to use the RBF method over the IDW method in predicting the radon gas concentrations in other states and countries. Despite both the RBF and the IDW methods being exact interpolators, the superior performance of the RBF method over the IDW method may be attributed to the fact that the radon surface generated by the RBF method is not limited by the range (maximum to minimum) of the measured values used in training stage, i.e., the IDW method cannot predict values outside the range of values used in the training stage. There is a possibility that the GPI, the LPI, the kriging, and the cokriging methods outperform the RBF method with other radon datasets. Since RBF methods are good at generating smooth surfaces using large number of data points, they are recommended to be used in the places with ample representation of radon data points having the feature of gradual variations (such as in Ohio) and are inappropriate for the places having large variations within short distances. In view of the aforementioned model evaluation results and the adequate representativeness for radon concentrations in Ohio with gradual variations, one may adopt the use of RBF method in the prediction of radon gas GM concentrations for the unknown zip codes in Ohio.
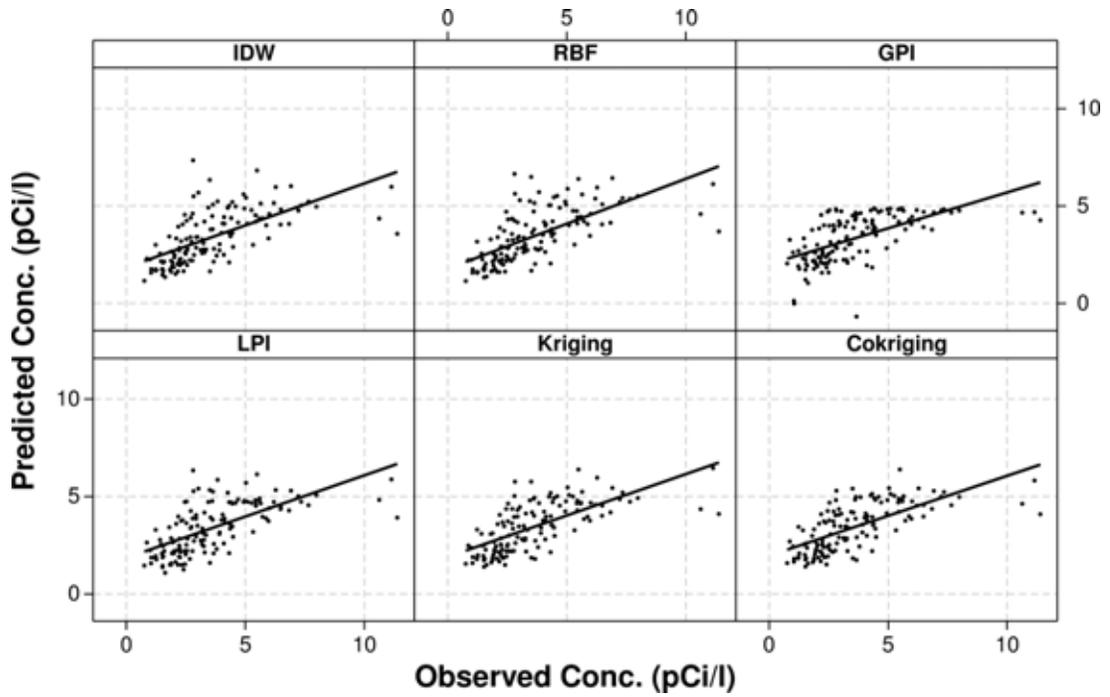
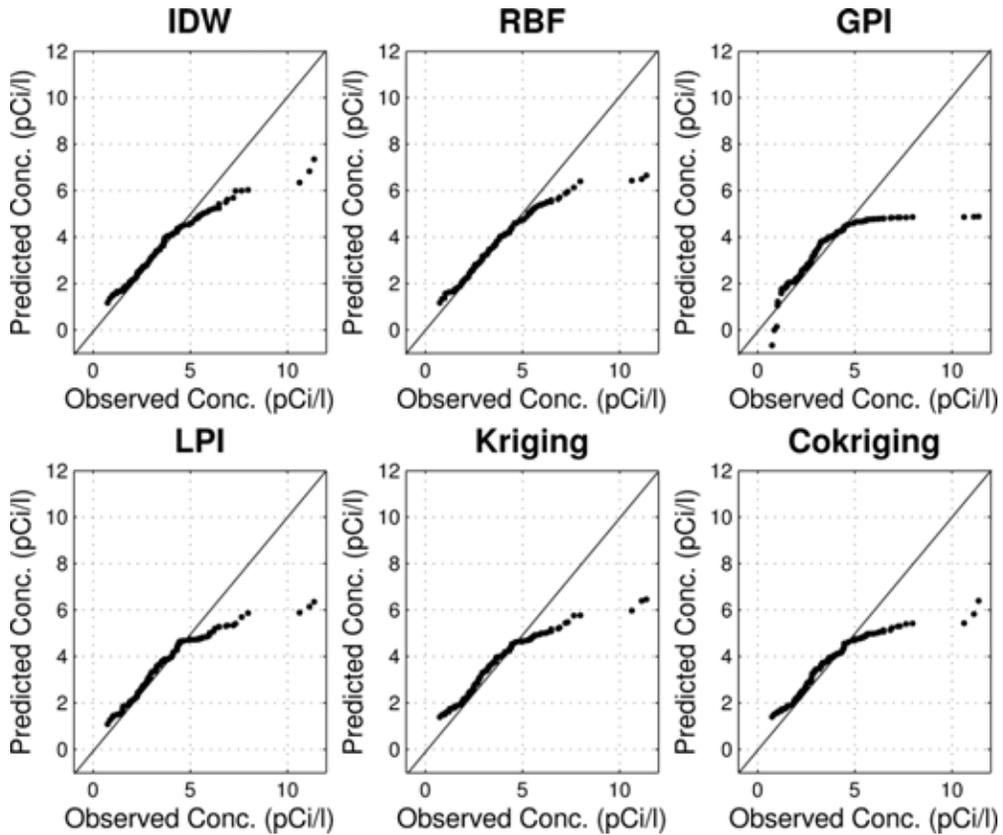**Fig. (1).** Scatter plots for the six interpolation methods.

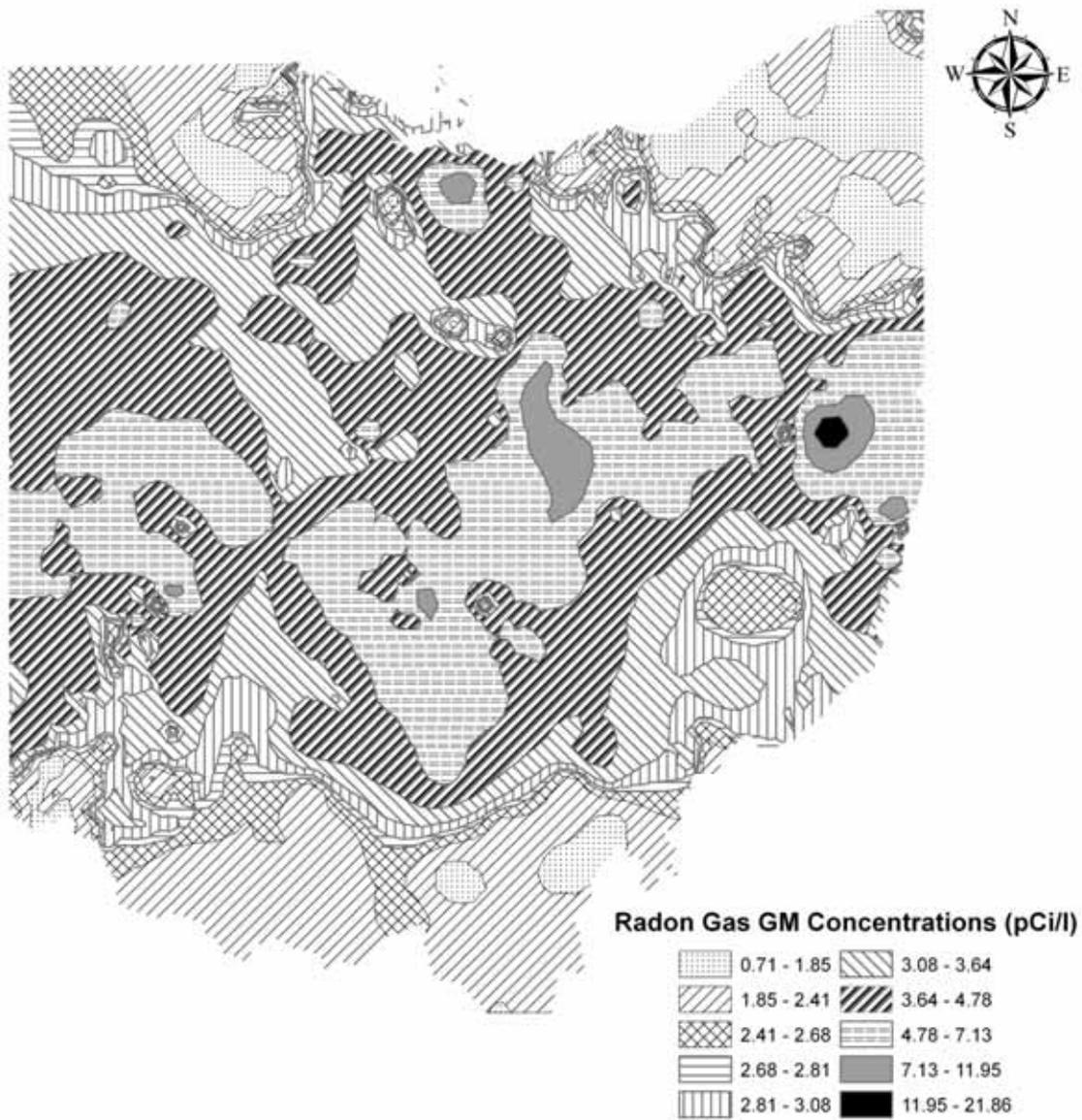**Fig. (2).** QQ plots for the six interpolation methods.

Radon Gas GM Concentrations (pCi/l)

0.71 - 1.85		3.08 - 3.64	
1.85 - 2.41		3.64 - 4.78	
2.41 - 2.68		4.78 - 7.13	
2.68 - 2.81		7.13 - 11.95	
2.81 - 3.08		11.95 - 21.86	

Fig. (3). Ohio radon gas GM concentration prediction map obtained using the RBF interpolation method.

Table 4. Statistics on radon gas GM concentrations in Ohio zip codes.

Criteria	Measured	RBF Predictions
No. of zip codes with radon gas GM concentrations > 2.7 pCi/l	574	1300
No. of zip codes with radon gas GM concentrations > 4 pCi/l	292	693
No. of zip codes with radon gas GM concentrations > 8 pCi/l	28	28
No. of zip codes with radon gas GM concentrations > 20 pCi/l	2	2

Fig. (**3**) presents the Ohio radon GM concentration prediction map obtained using the RBF method. Table **A-1** in the supplementary document provides a summary of the radon gas GM concentrations predicted for the unmeasured 1,176 zip codes using the RBF method.

3.2. Analysis Based on Zip Codes

This study analyzed the number of Ohio zip codes with radon gas GM concentrations greater than 2.7 pCi/l, 4 pCi/l, 8 pCi/l, and 20 pCi/l using the identified best interpolation method, i.e., the RBF method. Table **4** presents a summary

of the measured and the RBF (identified best interpolation method) predicted zip code results. These results demonstrate that there is a significant increase in the number of zip codes that exceeded the USEPA and WHO action limits of 4 pCi/l and 2.7 pCi/l, respectively. The zip code based analysis revealed 37.2% of the Ohio zip codes to have radon gas GM concentrations exceeding 4 pCi/l on using the RBF method against 18.6% of the zip codes based on the measured data. Similarly, 69.8% of the Ohio zip codes were predicted to have the radon gas GM concentrations exceeding the 2.7 pCi/l on using the RBF method against the 36.6% based on the measured data. There was no change in the number of zip codes having radon gas GM concentrations exceeding 8 pCi/l and 20 pCi/l. This indicates that majority of the unmeasured zip codes have the radon gas GM concentrations greater than the USEPA and WHO action limits and less than 8 pCi/l. This analysis suggests that more mitigation work is needed in the state of Ohio to lower the radon gas GM concentrations to the recommended action limits.

CONCLUSION

The performance of the six GIS based interpolation methods, namely, the IDW, the RBF, the GPI, the LPI, the kriging, and the cokriging were analyzed using the Ohio homes radon gas data. The prediction results were then evaluated using a comprehensive set of operational performance measures. The RBF method was identified to be the best performing interpolation method on the basis of the degree of closeness of computed operational performance measures to the corresponding ideal values. The bootstrap 95% confidence interval estimates indicate that the RBF method significantly outperformed the IDW method and was similar to the other interpolation methods. The identified best interpolation method, i.e., the RBF method, was used to evaluate the radon gas GM concentrations in Ohio on a zip code level basis. These results demonstrated an increase in the radon gas GM concentrations, thereby, indicating that more works needs to be done by radon planners in Ohio to reduce the health effects due to exposure to radon gas concentrations. The approach adopted in this study can be implemented world-wide to any radon affected areas.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   National Research Council, *Health Effects of Exposure to Radon: BEIR VI*. The National Academies Press: Washington, DC, 1999. Available at http://www.nap.edu/catalog.php?record_id=5499 (Accessed on February 25, 2014).

[2]   J.A. Harrell and A. Kumar, "Multivariate stepwise regression analysis of indoor radon data from Ohio, U.S.A.", *J. Off. Stats*, vol. 5, pp. 409-420, 1989.

[3]   A. Kumar, J.A. Harrell, and A. Heydinger, "Ohio goes online to combat indoor radon", *EM (by J. Air Waste Manage. Assoc.)*, vol. 2, pp. 10-12, 2001.

[4]   World Health Organization, *WHO handbook on radon: a public health perspective*. World Health Organization: Geneva, 2009. Available at http://www.ncbi.nlm.nih.gov/books/NBK143216/ (Accessed on: February 25, 2014).

[5]   N.E. Klepeis, W.C. Nelson, W.R. Ott, J.P. Robinson, A.M. Tsang, P. Switzer, J.V. Behar, S.C. Hern, and W.H. Engelmann, "The national human activity pattern survey (NHAPS): a resource for assessing exposure to environmental pollutants", *J. Expo. Anal. Env. Epid.*, vol. 11, pp. 231-252, 2001.

[6]   M. Jerrett, R.T. Burnett, P. Kanaroglou, J. Eyles, N. Finkelstein, C. Giovis, and J.R. Brook, "A GIS - environmental justice analysis of particulate air pollution in Hamilton, Canada", *Environ. Plann. A*, vol. 33, pp. 955-973, 2001.

[7]   D.W. Wong, L. Yuan, and S.A. Perlin, "Comparison of spatial interpolation methods for the estimation of air quality data", *J. Expo. Anal. Env. Epid.*, vol. 14, pp. 404-415, 2004.

[8]   L. Matejicek, P. Engst, and Z. Janour, "A GIS-based approach to spatio-temporal analysis of environmental pollution in urban areas: a case study of Prague's environment extended by LIDAR data", *Ecol. Model.*, vol. 199, pp. 261-277, 2006.

[9]   J. Wu, A.M. Winer, and R.J. Delfino, "Exposure assessment of particulate matter air pollution before, during, and after the 2003 Southern California wildfires", *Atmos. Environ.*, vol. 40, pp. 3333-3348, 2006.

[10]   C. de Fouquet, D. Gallois, and G. Perron, "Geostatistical characterization of the nitrogen dioxide concentration in an urban area part I: spatial variability and cartography of the annual concentration", *Atmos. Environ.*, vol. 41, pp. 6701-6714, 2007.

[11]   D.K. Jha, M. Sabesan, A. Das, N.V.V Kumar, and R. Kirubagaran, "Evaluation of interpolation technique for air quality parameters in Port Blair, India", *Universal J. Environ. Res. Technol.*, vol. 1, pp. 301-310, 2011.

[12]   H.C. Zhu, J.M. Charlet, and A. Poffijn, "Radon risk mapping in southern Belgium: an application of geostatistical and GIS techniques", *Sci. Total Environ.*, vol. 272, pp. 203-210, 2001.

[13]   A. Kumar, S. Maroju, and A. Bhat, "Application of ArcGIS geostatistical analyst for interpolating environmental data from observations", *Environ. Prog.*, vol. 26, pp. 220-225, 2007.

[14]   Kumar, A. Kadiyala, V. Devabhaktuni, A. Akkala, and D.V. Manthena, "Examination of Ohio indoor radon data", paper presented at *American Association of Radon Scientists and Technologists (AARST) Proceedings of 22nd International Radon Symposium*, Columbus, Ohio, 2010. Available at http://www.aarst.org/proceedings/2010/06_EXAMINATION_Of_OHIO_INDOOR_RADON_DATA.pdf [Accessed on: February 25, 2014].

[15]   D.V. Manthena, A. Kadiyala, and A. Kumar, "Interpolation of radon concentrations using GIS based kriging and cokriging techniques", *Environ. Prog. Sustain. Energy*, vol. 28, pp. 487-492, 2009.

[16]   D.V. Manthena, A. Kumar, and A. Kadiyala, "Development of an approach to predict radon gas concentrations for unmeasured zip codes using GIS based interpolation techniques", In *Advances in Mathematics Research*, Vol. 15 (A.R. Baswell, ed.), Nova Science Publishers Inc.: New York, pp. 47-76, 2012.

[17]   A. Kadiyala, and A. Kumar, *Guidelines for operational evaluation of air quality models: outdoor and indoor air quality models*. LAP

LAMBERT Academic Publishing GmbH & Co. KG: Saarbrücken, Germany, pp. 123, 2012.

[18]   J.A. Harrell, M.E. Belsito, and A. Kumar, "Radon hazards associated with outcrops of Ohio shale in Ohio", *Environ. Geol. Water Sci.*, vol. 18, pp. 17-26, 1991.

[19]   J.A. Harrell, J.P. McKenna, and A. Kumar, "*Geological controls on indoor radon in Ohio*", Ohio Division of Geological Survey Report of Investigations 144, pp. 36, 1993.

[20]   A. Kumar, C. Varadarajan, and A. Kadiyala, "Management of radon data in the state of Ohio, U.S.A.", *Open Environ. Biol. Monit. J.*, vol. 4, pp. 57-71, 2011.

[21]   ESRI, 2012. Available at http://www.esri.com [Accessed on: February 25, 2014].

[22]   J.S. Duval, "*Aerial radiometric contour maps of Ohio*", US Geological Survey Geophysical Investigations Map GP-968, 1985.

[23]   Ohio Radon Information System Website, 2012. Available at http://www.radon.utoledo.edu [Accessed on: February 25, 2014].

[24]   J. Li and A.D. Heap, *A review of spatial interpolation methods for environmental scientists.* Geoscience Australia: Canberra, Australia, 2008. Available at http://www.ga.gov.au/image_cache/GA125-26.pdf [Accessed on: February 25, 2014].

[25]   A. Kadiyala and A. Kumar, "Evaluation of indoor air quality models with the ranked statistical performance measures using available software", *Environ. Prog. Sustain. Energy*, vol. 31, pp. 170-175, 2012.

[26]   Kadiyala and A. Kumar, "Univariate time series prediction of air quality inside a public transportation bus using available software", *Environ. Prog. Sustain. Energy*, vol. 31, pp. 494-499, 2012.

[27]   Kadiyala and A. Kumar, "Multivariate time series models for prediction of air quality inside a public transportation bus using available software", *Environ. Prog. Sustain. Energy*, vol. 33, pp. 337-341, 2014.

[28]   Kadiyala, D. Kaur, and A. Kumar, "Development of hybrid genetic-algorithm-based neural networks using regression trees for modeling air quality inside a public transportation bus", *J. Air Waste Manage. Assoc.*, vol. 63, pp. 205-218, 2013.