Evaluation of Interpolation Techniques for Air Quality Monitoring using Statistical Error Metrics – A Review

[1]Goutham Priya M,
[1]Research Scholar,
IRS, College of Engineering,
Chennai, India, 600 025

[2] Jayalakshmi S
[2]Assistant Professor,
IRS, College of Engineering,
Chennai, India, 600 025

***Abstract* - Air pollution is the gravest crisis of the modern times all over the planet because of towering level of urbanization and industrialization. It has both global and local impacts. The discharge of such air pollutants in heavy concentrations over the city causes health hazards to the people. The monitoring of the pollutants has become a priority in urban areas. Environmental measurements of air pollutants are often taken at specific location with ground based instruments. Whilst these measurements offer helpful information about the pollutants around these stations, but little about the conditions further afield. The need for air pollutant information at unsampled locations arises when producing a spatially continuous air pollution map for a whole area. This need can be accomplished by spatial interpolation with which predictions are made at unsampled locations on the basis of the information from the nearest available measured sampling stations. This paper addresses the evaluation of interpolation techniques using statistical error metrics. The interpolation methods can be grouped as geostatistical, deterministic and interpolation with barriers. Statistical error metrics are employed for cross validation of the interpolation methods. The error metrics can be classified as dimensional and non-dimensional statistics. Dimensional statistics includes scale dependent metrics, percentage error metrics, relative error metrics and scale free error metrics whereas non-dimensional statistics includes correlation coefficient and index of agreement. GIS based interpolation techniques and statistical error metrics can be used in combination to determine the air quality in remote areas**.

Keywords: Air Pollution, Spatial Interpolation, Statistical Error Metrics, GIS.

I. INTRODUCTION

Currently, air pollution poses a major threat due to its deteriorating effects on the health of mankind resulting in lung and heart related diseases on exposure. Air Pollution may be defined as presence or introduction of toxic and detrimental substances in air, which are caused by growing industrialization, consuming electricity, developing transportation etc. Owing to the increase in the world population the exploitation of the natural resources also increases resulting in pollution.

Studies on air pollution are being through lately focussing on its causes and effects but the data available are limited and not extensive. It is not obligatory that an air pollutant measured at a meticulous location is been emitted from the exact location, it may also have moved in from a different location due to wind direction and wind velocity. Therefore an extensive data is required for the study which in turn calls for a large number of air quality monitoring stations which is impossible to establish for a developing Country like India due to its economic condition. Therefore the data from the existing stations has to be either interpolated or extrapolated to get the desired extensive data.

Interpolation is a mathematical method of estimating new data within a range of distinct set of known data whereas; extrapolation is also a mathematical method of estimating new data beyond a range of original observation data. Both of these methods have its own advantages and disadvantages but of the two methods, the method of interpolation is preferred. This is because there is a greater likelihood of getting a legitimate estimate from interpolation, while in extrapolation an assumption is made that the same trend is followed beyond the original data which may not be true in certain cases.

Interpolation by conventional method is laborious and time consuming. Therefore, GIS comes in picture, where in the available interpolation methods are: Inverse Distance Weighting (IDW), Kriging, Global Polynomial Interpolation (GPI), Local Polynomial Interpolation (LPI), Radial Basis Function (RBF), Diffusion interpolation with barriers (DI), Kernel Interpolation with Barriers (KI). This paper gives an outline and compares the different interpolation techniques which can be used for interpolating air quality parameters and also studies the Error Statistics like MAE (Mean Average Error), MAPE (Mean Average Percentage Error), RMSE (Root Mean Square Error), Index of Agreement (d) etc. to identify the best interpolation technique for any air pollutant.

II. INTERPOLATION METHODS – AN APPROACH

Spatial continuous data engage a striking role in expansion, risk evaluation, and management in environmental society. For mountainous and marine regions, these data are usually not available and are very expensive to obtain. Environmental data like air pollutant data collected from the field sampling surveys are typically point data. Nevertheless, for assessing air pollution, spatial continuous data are required for the region of interest for justified interpretations [1].

Spatial Interpolation can be categorised in several ways:

1. Global or Local Interpolation
2. Exact or approximate Interpolation
3. Gradual or abrupt Interpolation
4. Deterministic or geostatistical Interpolation

Since, in this study, as we are dealing with the accuracy of the interpolation techniques through error metrics, we focus more on the fourth category: Deterministic or Geostatistical Interpolation.

A. DETERMINISTIC INTERPOLATION

Deterministic Interpolation provides no estimation of errors with predicted values. It includes Inverse Distance Weighting (IDW), Radial Basis Function (RBF), Global Polynomial Interpolation (GPI) and Local Polynomial Interpolation (LPI). IDW is based on a hypothesis that the quantity of influence in close proximity sample points ought to be larger than the effect of more remote points. It means that the points neighbouring to the prediction location are assumed to have superior influence on the predicted value than those further away. The general equation is given in (1).

$$Z\,(s_o) = \sum \lambda_i Z\,(s_i) \tag{1}$$

Where,

$Z\,(s_o)$ – predicted value at location s_o

λ_i – Weights assigned

$Z\,(s_i)$ – Observed value at location s_i

In air pollution modelling or in air quality assessment studies, IDW is considered to be popular alternative to other interpolation techniques at various scales. The methodology usually employs a cross validation approach in which the data sets are divided into groups, out of which a particular set of data is kept as test data and the remaining sets as training data. The test data are predicted with the help of training data. This method is followed until all the data has been estimated. This cross validation process can be conceded in any number of iterations. IDW proved to be the finest method over kriging for the evaluation of the air quality parameters NO_2 (Nitrogen dioxide), SO_2 (Sulphur dioxide), SPM (Suspended Particulate Matter) in Port Blair and PM_{10} in central region of Thailand and therefore can be relied upon for air pollution studies to acquire spatially continuous data [2] [3].

Apart from air pollution studies IDW can also be employed to study meteorological parameters like temperature, Precipitation etc. IDW with other interpolation techniques such as spline with the regression models like stepwise and forced entry can also be compared using one-left-out technique [4]. The technique was performed on 4 temperature variables: mean daily temperature of the coldest month (January), mean daily temperature of the warmest month (August), the lowest mean monthly minimum temperature (January) and the highest mean monthly maximum temperature (June). The Mean daily temperature was accurately estimated by the local interpolation methods whereas the mean monthly temperature was accurately estimated by the regression models. When IDW is combined with the regression models the accuracy of the data is increased by 5%. Therefore combination of interpolation techniques can be used to estimate air pollutants and other parameters for better accuracy [5]. But when comparing 3D-shape function based spatiotemporal interpolation method to IDW, the former leads the show because of dividing the domain data into a number of tetrahedrons with the sampled input data as vertices [6]. Even though IDW is the most extensively used type of interpolation it has its own limitation mainly due to the fact that the affiliation between two observed values is not simply a function of distance and in many cases the distance relationship is not constant throughout [7].

RBF is a suitable estimator for irregularly distributed data. Hence they are employed in studies which involved multivariate data interpolation. They are global, exact and deterministic interpolation. It is based on an assumption that the estimated value can be approximated to several degree of meticulousness by summing up of standard mathematically distinct values [8]. The equation of RBF is given in (2).

$$y\,(x) = \sum w_i\, \varphi\,(\,\|\, x - c_i \,\|\,) \tag{2}$$

Where,

y – Estimation function

φ – Radially symmetrical Function

c_i – Centres (data points)

w_i – Coefficients

The advantage of RBF is that it requires only a few observed points but comparatively well distributed centres for the shape parameter in order to produce a high-quality estimate of the pollutant. RBF can be explored with leave-one-out cross validation process by means of a huge amount of geospatiotemporal data to access the trend of $PM_{2.5}$ for the US in 2009 [9]. RBF is found to be dominant in obtaining a spatially continuous data and also addresses the computational issues while handling big data. Moreover, RBF when comprehended with ANN provided by MATLAB tools provides good results on classified evaluation of real effects of pollutants on human health [10]. In general, RBF does not execute well for an observable with rapid changes in data within a small distance.

GPI is a deterministic approach of interpolation that fits a plane, typically polynomial, through the measured data points. It is usually used to fit a surface where the region varies slowly over the area of interest and therefore can be applied to air pollution studies. Nevertheless, it should be noted that the more multifarious the polynomial, the more complicated it is to attribute substantial significance to it. Moreover, it is highly vulnerable to outliers in particular at the boundary. On the grounds of air pollution modelling many studies has not been done. Of a few GPI has been identified as the best interpolation technique for the evaluation of RSPM in Tamil Nadu with the lowest RMSE of 21.528 by using 20% of the data as test data and the remaining 80% as training data and executing a number of iterations [11].

LPI method uses a particular sample of recognized points from the entire dataset and a polynomial equation to assess unidentified values. It fits many polynomials, each within specified overlapping neighbourhoods. These neighbourhoods overlap, and the value used for each prediction is the value of the fitted polynomial at the centre of the neighbourhood. There are two main disadvantage of using LPI. Firstly, the prediction standard errors, which, indicates uncertainty with the predicted value and secondly, the spatial condition number (SCN), which is a measure of stability of the accuracy of the predicted values [12]. A small change in SCN will have drastic effect in the solution vector. LPI is usually applied when the sample data are equally spaced and the values are normally distributed. Also, LPI is found to be more suitable than GPI [13].

B. *GEOSTATISTICAL INTERPOLATION*

Geostatistical Interpolation provides estimation of errors with predicted values. Kriging is a geostatistical interpolation technique that considers equally the distance and the degree of variation involving known data points while estimating values in unknown areas. A kriged approximate is a biased linear blend of the known sample values about the point to be estimated. It assumes that spatial variation consists of following three components, viz., spatially correlated component representing the variation of regionalized variable, a drift or structure, representing a trend and a random error term. It uses semi variance to determine spatial dependence. Equation 3 gives the general equation for kriging.

$$Z(s_o) = \sum \lambda_i Z(s_i) \quad (3)$$

Where,

$Z(s_o)$ – value being predicted for the target location s_o

λ_i – Weights assigned to each measured point

$Z(s_i)$ – Observed value at location s_i

It may seem as though equation (3) resembles the equation (1) of IDW. In IDW, the weight, λ_i, depends exclusively on the distance to the prediction location. However, with the kriging method, the weights are based not only on the distance between the measured points and the prediction location but also in general spatial arrangement of the measured points. The different types of kriging are formulated in the table I

Table I
Different Types of Kriging

S.NO	TYPE OF KRIGING	MODEL USED
1	Ordinary Kriging	$Z(s) = \mu + \varepsilon(s)$
2	Simple Kriging	$Z(s) = \mu + \varepsilon(s)$
3	Universal Kriging	$Z(s) = \mu(s) + \varepsilon(s)$
4	Indicator Kriging	$I(s) = \mu + \varepsilon(s)$
5	Probability Kriging	$Z(s) = \mu_2 + \varepsilon_2(s)$
6	Disjunctive Kriging	$f(Z(s)) = \mu_1 + \varepsilon(s)$

Source:www.gisresources.com

Kriging method of interpolation plays a significant role in determining the data at unsampled stations. It can also be used for optimizing Air Quality Monitoring Network (AQMN) with Geographic Information System (GIS) [14]. The missing or omitted data can be filled with the help of the Kriging. Among the various types of kriging models Ordinary Cokriging (OCK) outperforms the other models because it uses an additional parameter, DEM for interpolating air pollutants as well as meteorological parameters. When the study deviates to study the performance comparison between OK and OCK in terms of correlation, if the correlation is small OK performs better and as the correlation increases the results turns the other way. So it was concluded that OCK cannot always outperform OK in all aspects [15].

Kriging can also be incorporated in studies involving Univariate and Multivariate interpolation methods, where in two methods were used from multivariate geostatistical interpolation, simple Kriging with varying local means and linear regression. In Multivariate methods two secondary information were used exhaustively. The elevation of the study area derived from DEM and distance to a regional rainfall maximum. It was concluded that Theisson Polygon Method created highest errors whereas OK created the lowest. Therefore OK is considered as the best method beyond linear regression to interpolate rainfall [16]. Interpolation studies can also be conducted on ground water availability due to severe water shortages and dramatic decline in groundwater levels resulting in deterioration of the Oasis. A similar study was executed in the Minqin Oasis in Northwest China, where ground water data was available from 48 observation wells for 22 years (1981 – 2002). Three interpolation methods IDW, Radial Basis Function (RBF) and Kriging were used for the same which established that Kriging to be the best method for interpolating ground water [17].

Furthermore, kriging has been identified as the preferable interpolation technique for mapping traffic-related air pollution in Ijebu-Ode, Nigeria [18]. The major pollutants mapped using the techniques were NO_2, SO_2 and CO and spatio-temporal analysis of urban air quality was carried out. 3D spatial interpolation was done to understand the distribution of pollutants in space to perform more complex environmental analysis using processed DEM data. The DEM was created with photogrammetric tools by processing the image and data acquired [19]. From the studies discussed it can be noted that OK is the frequent type of kriging employed for the interpolation of air pollutant as well as meteorological parameters.

C. *INTERPOLATION WITH BARRIERS*

Interpolation with barriers incorporates diffusion interpolation with barriers (DI) and Kernel Interpolation with barriers (KI). DI depicts how particles diffuse inside a particular barrier. It refers to the gradual flow of particles. If barriers are not present then the output becomes similar to kernel interpolation with Gaussian kernel. It predicts on an automatically generated grid unlike other techniques which uses triangles with variable sizes. Barriers can be additive, cumulative or flow.

KI is a variation of first order LPI in which the volatility in the calculation is prohibited by means of regression coefficients similar to ridge regression. It also uses the shortest distance between the two points for prediction on the sides of the barrier which are connected by straight lines. The kernels used are Exponential, Gaussian, Quadratic, Epanechnikov, Polynomial of Order 5, and Constant. It is mostly applicable in hydrological and meteorological applications.

III ACCURACY METRICS

The cross validation of the interpolation techniques can be executed through Statistics by employing Accuracy Metrics. The Schematic flow diagram is given in Figure.1. Overall in Statistics, to check the accuracy of the actual and interpolated values in GIS interpolation, there are two methods – Dimensional and Non Dimensional Statistics.

A. *DIMENSIONAL STATISTICS*

There are four types of error metrics namely scale dependent metrics, percentage error metrics, relative error metrics and scale-free error metrics. Scale Dependent Metrics measures the average magnitude of the error in a set of predictions without considering its direction. The most commonly used scale dependent metrics are Mean Absolute Error (MAE) (Equation 4), Geometric Mean Absolute Error (GMAE), Mean Square Error (MSE) and Root Mean Square Error (RMSE) (Equation 5). From interpretation point of view, MAE is used to validate any type of GIS modelling.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\, \mathrm{O_i} - \mathrm{E_i} \,| \tag{4}$$

Where, || indicates –ve signs are ruled out.

O_i – Observed Values

E_i – Estimated Values

n – Number of values

MAE cannot be compared across series of data as it is scale dependent. Whereas, GMAE is an intermittent method [20] of accuracy assessment as division by zero may occur. Therefore it is less used in studies concerning interpolation. However, RMSE is more powerful and useful when large errors are particularly undesirable, since it is been squared even before it is averaged. It gives relatively high weight to normal errors. It also increases with the variance of the frequency distribution of error magnitudes [21].

$$\mathrm{RMSE} = \sqrt{[\frac{1}{n}\ \sum_{i=1}^{n} |\, \mathrm{O_i} - \mathrm{E_i} \,|^2]} \tag{5}$$

If MAE = RMSE, all the errors are of same magnitude. If, RMSE ≥ MAE, the greater the difference between them the greater the variance in the individual errors in the sample. While comparing RMSE and MAE, it can be observed that the value of RMSE is always greater than MAE and their difference does not increase monotonically. This is due to the fact that the differing in the error-magnitude variances associated with the set of errors is a constant [22]. Therefore it was concluded that RMSE is not a good indicator for an average model performance and MAE was used instead. But in some cases RMSE is found to be more appropriate when the error distribution is Gaussian and satisfies the triangle inequality requirement [23].

To evaluate the predictive availability of the interpolation techniques, viz, IDW, OK, UK, RMSE has been used with which IDW was selected as the best interpolation techniques with a RMSE value 0.683 – 0.703 which was low when compared to OK and UK [2]. Both RMSE and MAE were used in combination to assess the accuracy of the interpolation techniques employed to study the air pollutants at Port Blair [3]. IDW proved to be the finest technique as it bears the lowest values of MAE and RMSE. Therefore, the lower the values of MAE and RMSE the more reliability of a particular interpolation increases.

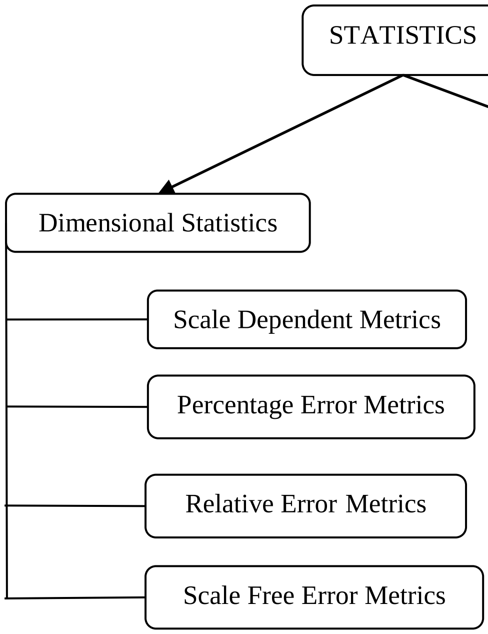

Figure. 1 Accuracy Metrics - An Overview

Percentage Error Metrics are scale independent and quantitative forecasting method since it produces a measure of relative overall fit. It is mostly used to compare the forecasting between any two methods in percentage which is easy to interpret. The most commonly used metric is Mean Absolute Percentage Error (MAPE) (Equation 6) also called as Mean Absolute percentage Deviation (MAPD).

$$\text{MAPE}= \left\{\frac{1}{n}\sum_{i=1}^{n} \frac{O_i - E_i}{O_i}\right\}*100 \tag{6}$$

It is Scale sensitive - cannot be used with low volume of data because O_i in the denominator, if less/small, MAPE gives extreme values which is worthless. MAPE can be used to validate accuracy of the optimization of Air Quality Monitoring Network (AQMN) [14]. The value of MAPE should be lesser indicating that low percentage of errors in the estimated data. But, eventually, there is a disadvantage of having indefinite results due to division by zero for intermittent data [20].

Relative Error Metrics are recommended when the task involves calibrating a model for a set of time series. It is also an alternative to percentage error metric for scale independent measurements. The most commonly used metrics are Geometric Mean Relative Absolute Error (GMRAE) and Mean Relative Absolute Error (MRAE). Since they are scale independent they can be used for assessing forecast accuracy across multiple series.

Scale-free Error Metrics is a general method applicable for the measurement of forecast accuracy without the problems seen in other methods. It can be used to compare results from a single series or multiple series. The commonly used metric is Mean Absolute Scaled Error (MASE) is given in equation (7).

$$\text{MASE} = \frac{1}{n}\sum \frac{e_t}{\left[\frac{1}{n-1}\sum(O_i - E_i)\right]} \tag{7}$$

B. NON DIMENSIONAL STATISTICS

In statistics, correlation coefficient (r^2) can be used to determine the relationship between two or more variables. The value ranges from -1 to +1. If the value ranges close to +1 then there is a positive correlation between the variables, viz if one variable increases other variable also increases. If the value ranges close to -1 then there is a negative correlation between the variables, viz, if one variable increases other variable decreases. If the coefficient is 0 then there is no relationship among the variable of study.

Correlation coefficient was used to determine the best interpolation technique in Middle Black Sea region [24] in which it was found that all the interpolation techniques considered were appropriate for some months since the coefficient among the measured and the estimated temperatures varied between 0.80 and 0.95. Moreover, correlation coefficient can also be used to study the correlation between the estimated values from two different interpolation techniques. The results of ozone estimation from both IDW and kriging had similar high correlation [5] indicating that both the interpolation techniques can be employed to acquire the data at unsampled locations. It can also be used in combination with other accuracy metrics like RMSE to validate the performance of any model. Such a study was done in Minqin Oasis in Northwest China to identify the areas with good source of ground water due to water shortages [17].

Index of correlation (d) is used to determine the relationship between the actual and interpolated data. It is a dimensionless parameter and the value ranges from 0 to 1 indicating better agreement. IOA can be expressed as in equation (8).

$$IOA = 1- \left[\frac{\sum_{i=1}^{n}(O_i - E_i)^2}{\sum_{i=1}^{n}(|O_i - \mu| + |E_i - \mu|)^2}\right] \tag{8}$$

Where ,

μ - Mean of the observed values

IOA is mostly used in studies of social research wherein logically there is great requirement of an index of agreement for each fact [25]. In air pollution studies, two spatial interpolation algorithms can be compared to identify the best technique by cross validation procedure using the index of agreement. The interpolation maps can be integrated with gradient and directional gradient maps that may serve as aides in the definition of critical sampling points [26].

IV. SUMMARY

Choosing a best type of interpolation technique for a study area depends on several factors. There is no universal technique that is appropriate for all exertion. Each technique works in a different way, except for the most part utilizes the perception of spatial auto-correlation; near points are more similar than points far away. It depends on the character of the variable and on the time-scale on which the variable is represented. For any study, it is mandatory to execute the different interpolation techniques and compare the results to determine the best technique. The eminence of sample point can affect the choice of interpolation technique as well. If the sample points are feebly scattered, the surface might not correspond to the actual terrain very well.

The real-world acquaintance of the subject will primarily influence which interpolation technique to use. If there are few sample points, Kriging can be used where more sample points can be added in areas where the topography changes unexpectedly or recurrently. If the sample points are closely located and have extreme differences in values then IDW would be the best option. IDW is a good interpolator for phenomena whose distribution is strongly correlated with distance, such as air, water and noise. The potential techniques that are employed regularly for the studies involving soil/geology modelling are universal Kriging and linear regression models in combination with Kriging. When a smooth surface is required like in temperature data, spline is the best option. Also there is much research need to be done on the interpolators such as LPI, GPI, Kernel and Diffusion to identify their unique characteristics. Geostatistical interpolation techniques are recommended to provide the regional variability of the data which involves variogram analysis, where secondary variables can be incorporated in order to improve the estimation and also will provide interpolation error estimation. It also preserves the original variability of the data.

V. CONCLUSION

Visiting each and every location in the area of study to measure the concentration of an observable fact such as an air pollutant is typically complicated or expensive. Instead, disseminated sample point locations can be selected and a predicted value can be assigned to all other locations. Estimation of pollution fields, especially, air pollution has become the area of interest because of its ill-effects to human health.

In this paper, the various aspects of interpolation techniques (IDW, RBF, GPI, LPI, Kriging, DI and KI), which can be employed for air pollution estimation has been discussed. From this, a particular interpolation technique that outperforms the other technique can be identified by assessing the accuracy between the observed and the estimated values. This can be accomplished by error metrics (MAE, RMSE, MAPE, MASE etc) through statistics. The study concluded that, for a study area, there is no pre determined optimal interpolation scheme [27]. It relies on the spatial arrangement of monitoring networks. Therefore, for an area of interest, to identify the best interpolation technique, the various techniques can be studied and compared using statistical error metrics. Also, apart from air pollution spatial interpolation can also be applied for other environmental and meteorological parameters like soil, rainfall, water quality, temperature, etc. The methodology proves to be exceptionally accommodating for decision makers to wrap up any results without even have to rely on the monitoring stations.

REFERENCES

[1] J Li and A D Heap,'A Review of Spatial Interpolation Methods for Environmental Scientists', *Geoscience Australia*, Canberra, 2008.

[2] Phatarapon Vorapracha, Pongtep Phonprasert, Suparada Khanaruksombat and nuchanaporn Pijarn, 2015, 'A Comparison of Spatial Interpolation Methods for predicting concentrations of Particle Pollution (PM_{10})', *International Journal of Chemical, Environmental & Biological Sciences*, Vol.3, Issue.4, ISSN: 2320 - 4087, pp 302 - 306.

[3] Dilip Kumar Jha, M.Sabesan, Anup Das, N.V.Vinithkumar, and R.Kirubagaran, 2011, 'Evaluation of Interpolation Technique for Air Quality Parameters in Port Blair, India", *Universal Journal of Environmental Reseearch and Technology*, Vol.1, Issue.3, ISSN: 2249 0256, pp 301-310.

[4] Daniel Kurtzman and Ronen Kadmon, 1999, 'Mapping of temperature variables in Israel: a comparison of different interpolation methods', *CLIMATE RESEARCH*, Vol. 13, pp 33–43.

[5] D Rojas Avallaneda, 2007, 'Spatial interpolation Techniques for estimating the levels of pollutant concentrations in the atmosphere', *Revista Mexicana De Fisica*, Vol.53, Issue.6, pp 447 – 454.

[6] L.Li, X.Zhang and R.Piltner, 2008, 'An Application of the Shape Function Based Spatiotemporal Interpolation Method to Ozone and Population-Based Environmental Exposure in the Contiguous U.S', *Journal of Environmental Informatics*, Vol.12, Isuue.2, pp 120-128.

[7] Lu.G.Y and Wong D.W, 2008, 'An adaptive inverse-distance weighting spatial interpolation technique', *Computers & Geosciences*, Vol.34, No.9, pp 1044-1055.

[8] Hardy R.L, 1978, 'The application of Multiquadratic Equations and Point Mass Anomaly Models to Crustal Movement Studies, *National Oceanic and Atmospheric Administration*.

[9] Travis Losser, Lixin Li and Reinhard Piltner, 2014, 'A Spatiotemporal Interpolation using Radial Basis Functions for Geospatiotemporal Big Data', *Fifith International Conference on Computing for Geospatial Research and Application*, IEEE, pp 17-24.

[10] Liu Jie, Yang Peng, L V Wensheng and LIU Agudamu, 2014, ' Comprehensive Assessment Grade of Air Pollutants based on Human Health Risk and ANN Method', *International Symposium on safety science and technology, Science Direct*, Procedia Engineering, pp 715-720.

[11] Goutham Priya M, Jayalakshmi S, Samundeeswari R, 2018, 'A Study on Comparison of Interpolation Techniques for Air Pollution Modelling', *Indian Journal of Scientific Research*, Vol.17, Issue.2, pp 58-63.

[12] A. Schaum,1993, 'Theory and Design of Local Interpolators', *CVGIP: Graphical Models and Image Processing*, Vol. 55, No. 6, November, pp. 464-481.

[13] Humair Ahmed, Muhammad Tariq Siddique, Muhammad Iqbal and Fayyaz Hussain, 2017, 'Comparative study of interpolation methods for mapping soil pH in the apple orchards of Murree, Pakistan, *Soil & Environment*, VOl.36, No.1, pp. 70-76.

[14] Shareef M, Husain T and Alharbi B, 2016, 'Optimization of Air Quality Monitoring Network Using GIS Based Interpolation Techniques', *Journal of Environmental Protection*, Vol.7, pp 895 – 911.

[15] Xian Luo and Youpeng Xu, Yi Shi, 2011,' Comparison of Interpolation Methods for Spatial Precipitation under Diverse Orographic Effects', *IEEE*.

[16] Alan Mair and Ali Fares, 2011, 'Comparison of Rainfall Interpolation Methods in a Mountainous Region of a Tropical Island', *Journal of Hydrologic Engineering*, Vol.16, pp 371 – 383.

[17] Yue Sun, Shaozhong Kang, Fusheng Li and Lu Zhang, 2009, 'comparison of interpolation methods for depth to ground water and its temporal and spatial variations in the Minqin oasis of northwest china', *Environment Modelling and Software*, Vol.24, Issue.10, pp 1163 – 1170.

[18] Oludare Hakeem Adedeji, Olasumbo Oluwafunmilayo and Tope-Ajayi Opeyemi Oluwaseun, 2016, 'Mapping of Traffic-Related Air Pollution Using GIS Techniques in Ijebu-Ode, Nigeria', *Indonesian Journal of Geography*, Vol.48, Issue.1, pp 73-83.

[19] Shujun Song, 2008, 'A GIS-based approach to Spatio-temporal Analysis of Urban air quality in Chengdu Plain', The international Archives of the Photogrammetry, *Remote Sensing and Spatial Information Sciences*', Vol.XXXVII, pp 1447-1450.

[20] Syntetos.A.A and Boylan.J.E, 2005, 'The Accuracy of intermittent demands estimates', *International Journal of Forecasting*, Vol.21, pp 303 – 314.

[21] Rob J Hyndman, 2006, 'Another Look into Forecast Accuracy Metrics for Intermittent Demand', *Foresight*, Issue.4, pp. 43-46.

[22] Cort J Willmort and Kenji Matsuura, 2005, 'Advantages of Mean Absolute Error (MAE) over Root Mean Square Error (RMSE) in Assessing Average Model Performance', *Climate Research*, Vol.30, pp. 79-82.

[23] T.Chai and R.R.Draxler, 2014, 'RMSE or MAE?-Arguments against avoiding RMSE in literature', *Geoscientific Model Development*, Vol.7, pp. 1247-1250.

[24] Mustafa GÜLER and Tekin KARA, 2014, 'Comparison of Different Interpolation Techniques for Modelling Temperatures in Middle Black Sea Region', *Journal of Agricultural Faculty of Gaziosmanpasa University*, Vol.31, Issue.2, ISSN:1300-2910, pp 61 – 71.

[25] W.S.Robinson, 1957, 'The Statistical Measure of Agreement', *American Sociological Review*, Vol.22, No.1, pp. 17-25.

[26] Libardo Antonio Londoño-Ciro, Julio Eduardo Cañón-Barriga, 2015, 'Imputation of spatial air quality data using gis-spline and the index of agreement in sparse urban monitoring networks' *Revista Facultad de Ingeniería*, Vol.76.

[27] Lamiaa Khazaz, Hassane Jarar Oulidi, Saida El Moutaki, Abdessamad Ghafiri, 2015 "Comparing and Evaluating Probabilistic and deterministic Spatial Interpolation Methods for Ground Water Level of Haouz in Morocco", *Journal of Geographic Information System*, Vol.7, pp 631-642.