Progress In Electromagnetics Research M, Vol. 14, 135–145, 2010

A COMPARISON OF SPATIAL INTERPOLATION METHODS FOR ESTIMATION OF AVERAGE ELECTROMAGNETIC FIELD MAGNITUDE

M. Azpurua and K. Dos Ramos

Instituto de Ingeniería
Venezuela

Abstract—Several georeferenced measurements of electric field were done in a pilot area of Caracas, Venezuela, to verify that the magnitude of radio frequency electromagnetic fields is below the human exposure limits, recommended by the International Commission on Non-Ionizing Radiation Protection. The collected data were analyzed using geographical information systems, with the objective of using interpolation techniques to estimate the average electromagnetic field magnitude, to obtain a continuous dataset that could be represented over a map of the entire pilot area. This paper reviews the three methods of interpolation used: SPLINE, Inverse Distance Weighting (IDW) and KRIGING. A statistical assessment of the resultant continuous surfaces indicates that there is substantial difference between the estimating ability of the three interpolation methods and IDW performing better overall.

1. INTRODUCTION

There is no single preferred method for data interpolation. Aspects of the algorithm selection criteria need to be based on the actual data, the level of accuracy required, and the time and/or computer resources available. In the absence of criteria for selecting among the available techniques, this paper compares three spatial interpolation techniques — SPLINE, Inverse Distance Weighting (IDW), and KRIGING — with the goal of determining which method creates the best representation of reality for measured electric field intensity. The benefits and limitations of these commonly used interpolation methods are discussed in this paper.

Received 31 August 2010, Accepted 17 September 2010, Scheduled 26 September 2010
Corresponding author: M. Azpurua (bazpurua@fii.gob.ve).

The data used in this assessment are average electric field intensity, measured over a frequency range of 100 kHz to 6 GHz using a isotropic E-field probe, taking 35 samples per second during a 6 minutes period in each of the 206 measurements points. It was considered a hybrid data acquisition pattern random — systematic of uniform distribution. Therefore, a matrix of data points spaced approximately 100 m from each other is fixed over the 2.64 km^2 pilot area.

Selecting an appropriate spatial interpolation method is fundamental to surface analysis since different methods of interpolation can result in different surfaces and ultimately different results. Statistical techniques are used to evaluate the three interpolation methods against independently collected data [1].

2. SPATIAL INTERPOLATION

Interpolation is a method or mathematical function that estimates the values at locations where no measured values are available. Spatial interpolation assumes the attribute data are continuous over space. This allows the estimation of the attribute at any location within the data boundary. Another assumption is the attribute is spatially dependent, indicating the values closer together are more likely to be similar than the values farther apart. These assumptions allow for the spatial interpolation methods to be formulated. The goal of spatial interpolation is to create a surface that is intended to best represent empirical reality thus the method selected must be assessed for accuracy [2].

3. SPATIAL INTERPOLATION METHODS

The techniques assessed here include the deterministic interpolation methods of SPLINE and Inverse Distance Weighting (IDW) and the stochastic method of KRIGING. Each method selected requires that the exact data values for the sample points are included in the final output surface. These spatial interpolation methods have various decision parameters. The selected techniques, SPLINE, IDW and KRIGING, are not all of the interpolation methods used in spatial analysis; but are the most common techniques that are available in GIS software.

3.1. SPLINES

SPLINES method estimates values using a mathematical function that minimizes the total surface curvature, resulting in a smooth surface that passes exactly through the sampled points. While there are more entry points specified, the greater the influence of distant points and the smoother the surface. Advantages of splining functions are that they can generate sufficiently accurate surfaces from only a few sampled points and they retain small features. A disadvantage is that they may have different minimum and maximum values than the data set and the functions are sensitive to outliers due to the inclusion of the original data values at the sample points. This is true for all exact interpolators, which are commonly used in GIS, but can present more serious problems for SPLINE since it operates best for gently varying surfaces, i.e., those having a low variance [3]. For the mathematical formulation of the SPLINES interpolation method, it can be consulted [3].

3.2. Inverse Distance Weighting (IDW)

Inverse Distance Weighting is based on the assumption that the nearby values contribute more to the interpolated values than distant observations. In other words, for this method the influence of a known data point is inversely related to the distance from the unknown location that is being estimated. The advantage of IDW is that it is intuitive and efficient. This interpolation works best with evenly distributed points. Similar to the SPLINE functions, IDW is sensitive to outliers. Furthermore, unevenly distributed data clusters results in introduced errors [4].

The simplest form of IDW interpolation is called, Shepard method [4] and it uses weight function w_i given by (1):

$$w_i = \frac{h_i^{-p}}{\sum\limits_{j=0}^{n} h_j^{-p}} \tag{1}$$

where p is an arbitrary positive real number called the power parameter (typically $p = 2$) and h_j are the distances from the dispersion points to the interpolation point, given by:

$$h_i = \sqrt{(x - x_i)^2 + (y - y_i)^2} \tag{2}$$

where (x, y) are the coordinates of the interpolation point and (x_i, y_i) are the coordinates of each dispersion point. The weight function varies with a value of unity at the dispersion point to a value close to zero as the distance to the dispersion point increase. The weight functions are normalized as a sum of the weights of the unit. Then, the interpolated value of the electric field $E(x, y)$ is given by:

$$E(x, y) = \sum_{j=0}^{n} w_j E(x_j, y_j) \tag{3}$$

In order to improve the computational time is possible to set bounds to the dispersion points that contribute to the calculation of the interpolated value, to all those dispersion points within a given search radius centered on the interpolated point. For the particular application developed in this work, it was determined that the most appropriate search radius was 500 m, so that the computation times were manageable.

3.3. KRIGING

Similar to IDW, KRIGING uses a weighting, which assigns more influence to the nearest data points in the interpolation of values for unknown locations. KRIGING, however, is not deterministic but extends the proximity weighting approach of IDW to include random components where exact point location is not known by the function. KRIGING depends on spatial and statistical relationships to calculate the surface. The two-step process of KRIGING begins with semivariance estimations and then performs the interpolation. Some advantages of this method are the incorporation of variable interdependence and the available error surface output. A disadvantage is that it requires substantially more computing and modeling time, and KRIGING requires more input from the user [1].

KRIGING belongs to the family of linear least squares estimation algorithms. The aim of KRIGING is to estimate the value of an unknown real-valued function, E, at a point, (x, y), given the values of the function at some other points, $\{(x_1, y_1), (x_2, y_2), \ldots, (x_3, y_3)\}$. A KRIGING estimator is said to be linear because the predicted value $E(x, y)$ is a linear combination that may be written as:

$$E(x, y) = \sum_{j=0}^{n} \lambda_j E(x_j, y_j) \tag{4}$$

The weights λ_i are solutions of a system of linear equations which is obtained by assuming that E is a sample-path of a random process $F(x, y)$, and that the error of prediction

$$\varepsilon(x, y) = F(x, y) - \sum_{j=0}^{n} \lambda_j E(x_j, y_j) \tag{5}$$

is to be minimized in some sense. For instance, the so-called simple kriging assumption is that the mean and the covariance of $F(x, y)$ is known and then, the KRIGING predictor is the one that minimizes the variance of the prediction error [1].

4. METHODOLOGY

4.1. Study Site

The measurements were taken in a pilot zone of Caracas, Venezuela (Figure 1) with an approximate area of $2.64\,\mathrm{km}^2$, which represents 0.609% of the total geographical area of the Caracas. A total of 206 measurements points, were selected over the pilot zone. This area is characterized by a dynamic economic-business activity which is evident given the presence of shopping centers, office buildings of the national telephone operators and other telecommunications companies. On the other hand, this area also boasts a large number of hospitals and schools, which is of interest to know the impact of electromagnetic fields.

Figure 1. Study site.

4.2. Data Collection Process

During the process of data collection, measurements were performed according to the following considerations:

- Time considerations: Measurements were taken over a period of 30 days, from February 15th to March 15th of 2010. Measurements were performed only on working days (from Mondays to Fridays). Each measurement was taken between 8:00 am and 5:00 pm. Previous investigations suggest this is the more suitable time interval to take this kind of measurements [5].
- Geographical considerations: For each measurement point geographical coordinates were taken using a GPS navigation unit.
- Measurements considerations: Each measurement was carried out averaging the magnitude of the electric field intensity, measured 35 times per second, using an isotropic E-field probe during six (6) minutes, as indicated by ICNIRP [6]. The instrument measurement bandwidth was 100 kHz–6 GHz. To minimize the interference caused by the operators, all the cellphones and personal radiation sources were disconnected. Finally, it is assumed and procured that measurements are taken in the far field zone.

4.3. Data Processing and Representation Tools

The spatial analysis and the geographical data representation carried out in this work, was mainly supported by the usage of two informatics tools: gvSIG 1.9 and Past 2.02. GvSIG is a Geographic Information System (GIS), that is, a desktop application designed for capturing, storing, handling, analyzing and deploying any kind of referenced geographic information in order to solve complex problems, while Past is a data analysis package that includes common statistical, plotting and modeling functions. Both, gvSIG and Past are free software tools distributed under the GNU/GPL license.

4.4. Assessment Methods

Different measures of fit may be used to determine how well an interpolated map represents the observed data. With most methods, some measure may be constructed of the closeness of the interpolated values $E(x, y)$ to the values E_i observed at control sites $\emptyset_i$. In this work, it was calculated the mean absolute error "MAE", the mean squared error "MSE" and the Euclidean distance "D", between a set of control points (on which measurements of electric field intensity were done, but not included in the interpolation process) and the interpolated results. The number of control points taken was approximately 10% of the total number of measurement points, which are twenty (20), randomly distributed over the study site perimeter taking into account that the control point shall not be located closer to 20 m from a previous measured point. The equations used in those calculations were:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|E(x,y) - E_i| \tag{6}$$

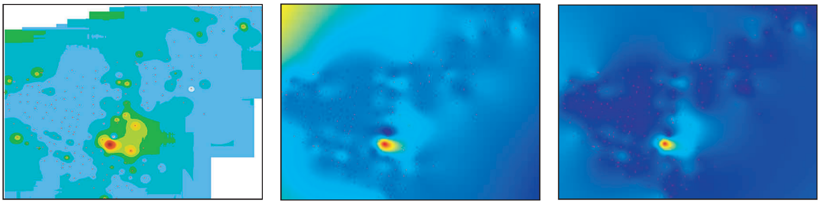
$$MSE = \frac{1}{n}\sum_{i=1}^{n}(E(x,y) - E_i)^2 \tag{7}$$

$$D = \sqrt{\sum_{i=1}^{n}(E(x,y) - E_i)^2} \tag{8}$$

Additionally, a month later, 10% of the measurements were repeated under similar conditions to the original ones, following the same measurement protocol, to ensure the results are valid over time without suffering significant changes. It was found that the maximum difference in the electric field intensity measured for one point, was about 0.5 V/m. Therefore, it was fixed as a criterion, that only the interpolated results with an error lower than 0.5 V/m will be accepted as valid.

5. RESULTS

The interpolation process was carried out using the methods: IDW, SPLINES and KRIGING. The results of each interpolation process were represented over the study site map. The mapping result of the



(a) (b) (c)

Figure 2. Interpolated maps of the *E*-field intensity average magnitude. (a) IDW, (b) SPLINES, (c) KRIGING.

Table 1. Results of the measures of fit applied to the interpolation methods.

MEASURES OF FIT	MAE (V/m)	MSE $(V/m)^2$	D (V/m)	Max Error (V/m)	Min Error (V/m)
IDW	0.17	0.05	1.01	0.55	0.008
KRINGING	0.74	0.73	3.81	1.66	0.111
SPLINE	0.89	1.11	4.71	1.98	0.034

E-field average magnitude is shown in Figure 2, in which the red dots are the dispersion points. The E-field intensity average magnitude is represented using a color scale where the lower values are represented using blue, and the higher values using red. The interpolated values are valid for the points inside the perimeter of the study site, even when the algorithms that calculate the raster layer, considers all pixels within rectangle defined by the base map.

The measures of fit carried out, yielded the results shown in Table 1. It shows that within the interpolation methods used, the IDW method is the one that best estimated the measurement results of the electric field average magnitude.

The results show that the average magnitude of the electromagnetic fields is mainly determined by the distance between the sources and the observation point. Furthermore, the results also show that the obstacles represented by the typical buildings and similar have no important influence on the magnitude of the average electric field intensity.

Also it's important to compute the number of control points whose interpolated results are within the acceptance margin. For the IDW method only two (2) of twenty (20) control points are lightly outside the acceptance interval. Similarly, only 9 and 5 of the 20 control points met the acceptance criteria for KRIGING and SPLINES interpolation methods, respectively. This fact it's shown in Figure 3 using a scatter plot. Figure 3 also shows that KRIGING and SPLINES methods have a very poor relationship with the actual E-field average magnitude, due to the large dispersion observed around the perfect prediction line.

When the error distribution is represented in histograms (Figure 4), it could be seen that the IDW interpolation method has a superior performance than KRIGING and SPLINES methods. Additionally, for each method, the errors are concentrated around its mean value, and the error density diminishes as it gets apart of its mean value.

Finally, we elaborated a raster layer of the average electric field strength using the IDW method and this layer is superimposed on a base map of the pilot area in order to obtain the representation shown in Figure 5.

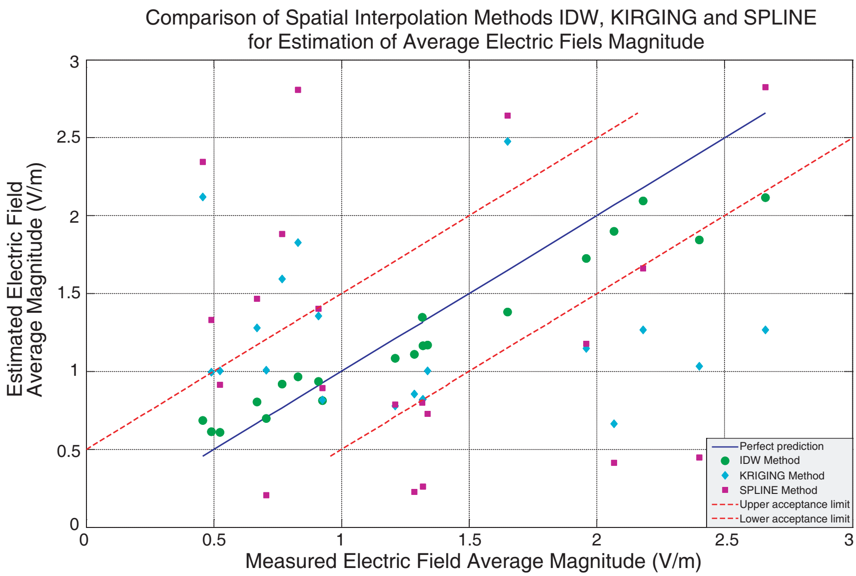
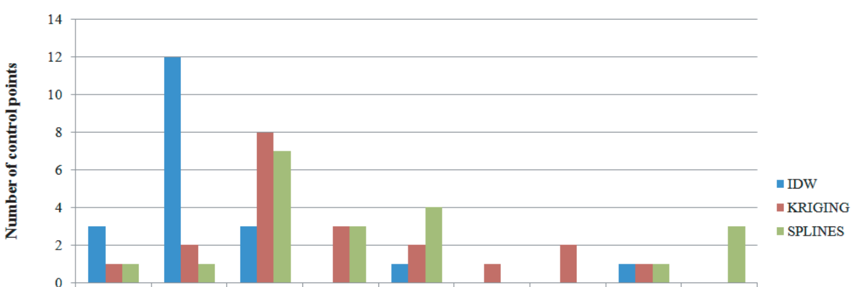


Figure 3. Scatter plot for the comparison of the results obtained using the interpolation methods IDW, KRIGING and SPLINES for the calculation of the average electric field intensity magnitude.



	[0-0.25)	[0.25-0.5)	[0.5-0.75)	[0.75-1.00)	[1.00-1.25)	[1.25-1.5)	[1.5-1.75)	[1.75,2)	[2,∞)
IDW	3	12	3	0	1	0	0	1	0
KRIGING	1	2	8	3	2	1	2	1	0
SPLINES	1	1	7	3	4	0	0	1	3

Error bins

Figure 4. Error distribution in the calculation of the E-field intensity average magnitude using the interpolation methods IDW, KRIGING and SPLINES.

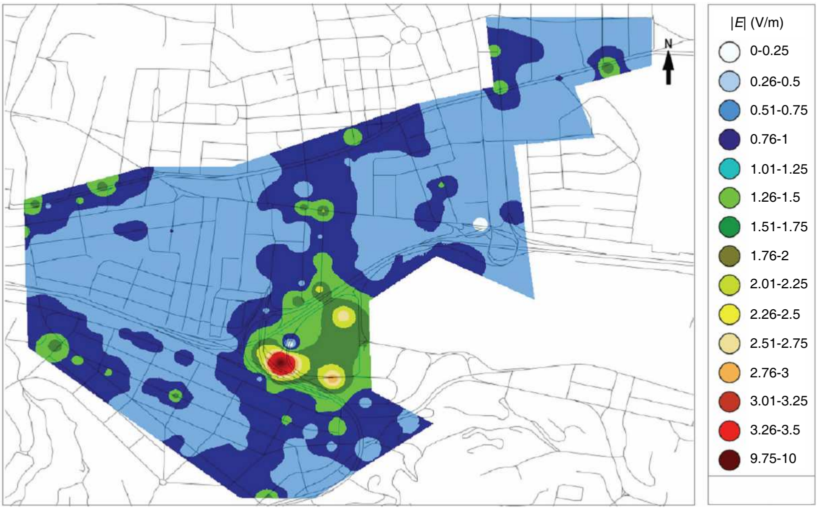


Figure 5. Geographical representation of the average electric field intensity magnitude.

6. CONCLUSION

This study has shown that IDW interpolation method is most likely to produce the best estimation of a continuous surface of the average magnitude of electric field intensity. The IDW method exactness was superior to the one shown by the SPLINES and KRIGING techniques. This could be justified not only because the magnitude of the radiated electromagnetic fields in free space decreases with the inverse of distance (free space loss) but also because in a urban environment, the ground irregularity and clutter, increase the effect of absorption, shadowing, scattering, divergence, and defocusing of the diffracted waves, contributing to the attenuation of the electromagnetic fields. That is why attenuation and, in consequence, the electromagnetic field average magnitude, at a defined height, can be described on a statistical basis considering all those effects combined together. Therefore, for the far field region, the expected field strength varies as $1/D^n$, where D in the distance from source, and n is a dimensionless number, greater than 1, that typically varies from 1, 3 for open country areas to about 2, 8 for heavily built-up urban areas [7].

In our test case, an $n = 2$, which correspond with the power parameter of the IDW interpolation method, was determined as the more suitable, and it is coherent with the expected $n = 2.2$ proposed in [8] for the perdition of propagation of interference from industrial radio frequency equipment. Hence, the average magnitude of the electromagnetic fields is mainly determined by the distance between the sources and the observation point; and the statistically determined power parameter that resumes the effects of absorption, shadowing, scattering, divergence, and defocusing of the diffracted waves, showing that for our particular case, the environmental electromagnetic field distribution produce little impact over the human exposure to electromagnetic fields. Furthermore, the results also show that specific obstacles represented by the typical buildings and similar have no important influence on the magnitude of the average electric field intensity.

Nevertheless, it should be noted that IDW interpolation method should be tuned adjusting the power parameter and the search radius to improve accuracy, in each specific case. It also should be considered the incorporation of an altimetric model of the radiation sources and measurement points, as a way to improve the interpolation results.

REFERENCES

1. Legendre, P. and L. Legendre, *Numerical Ecology*, 2nd Edition, Elsevier, Canada, 1998.
2. Anderson, S., "An evaluation of spatial interpolation methods on air temperature in phoenix," Department of Geography, Arizona State University Tempe, AZ 85287-0104, 2001.
3. Magnus, E. and M. Clyde, *Control Theoretic Splines: Optimal Control, Statistics, and Path Planning*, Princeton University Press, 2010.
4. Shepard, D., "A two-dimensional interpolation function for irregularly-spaced data," *ACM Annual Conference/Annual Meeting*, 517–524, 1968.
5. García, E., J. Cepeda, B. Melcón, G. Búrdalo, M. de Barrios, M. Fuentes, and A. Fernández, "Mapa de radiaciones no ionizantes en la ciudad de león," Universidad de León, Spain, 2008.
6. IRPA/ICNIRP, "Guidelines for limiting exposure to time varying electric, magnetic and electromagnetic fields (up to 300 GHz)," *Health Physics*, Vol. 74, No. 4, 1998.
7. CCIR Report 239-7, "Propagation statistics required for broadcasting services using the frequency range 30 to 1000 MHz," 1990.
8. CISPR 11, "Industrial, scientific and medical equipment — Radio-frequency disturbance characteristics — Limits and methods of measurement," Edition 5.1 2010-05, Annex D.