International Journal of Advanced Engineering Research and Science (IJAERS)
Peer-Reviewed Journal
ISSN: 2349-6495(P) | 2456-1908(O)
Vol-8, Issue-9; Sep, 2021
Journal Home Page Available: https://ijaers.com/
Article DOI: https://dx.doi.org/10.22161/ijaers.89.19

Using Geostatistics to Map Received Power in Wireless Communication Networks

Edilberto Rozal[1], Evaldo Pelaes[2]

[1]Department of Mathematics, Federal University of Pará (UFPA), Castanhal, Pará, Brazil
[2]Electrical Engineering Department, Federal University of Pará (UFPA), Belém, Pará, Brazil

Received: 18 Aug 2021,

Received in revised form: 14 Sep 2021,

Accepted: 21 Sep 2021,

Available online: 29 Sep 2021

©2021 The Author(s). Published by AI Publication. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

***Keywords*—Anisotropy, Geostatistics, Kriging, Semivariogram, Wireless.**

Abstract*—*The literature analysis of propagation models has investigated different prediction methods to identify appropriate techniques for this purpose. The predictive algorithms of these models usually deal with large amounts of data, requires a sophisticated computer processing and knowledge, sometimes detailed of the topography of the terrain. For being based on measurements performed at specific locations, empirical models tend not to provide very reliable results when applied to regions that differ significantly from the original region This article proposes a method based on measured data that incorporates the effects of neighborhood on the calculation of received power (dBm) and uses the theory of geostatistics to estimate the extent of the spatial correlation between measurements of samples in the region of interest. The results show that it is possible to identify the vectors with better reception of the signal emitted by the base transceiver station by the spatial perspective of received power measurements (dBm) and to identify homogeneous zones and those zones where the service operator may or may not favor the user.

I. INTRODUCTION

Currently, a wide variety of communication channel models exist with theoretical and experimental foundations to predict path attenuation in mobile communications systems, and their development is one of the most important steps in mobile communication planning. A correct estimate enables the designer of mobile systems to predict the minimum power required to radiate from a transmitter to supply a predetermined area with acceptable coverage quality, which is of fundamental importance for the improvement of the frequency reuse technique and to implement projects with shared bandwidth (Liaskos et al., 2018) [1].

These models differ in their applicability in different types of terrain and different environmental conditions. Therefore, no model is appropriate for all situations. The land on which propagation occurs has varied topography, vegetation and buildings that are randomly distributed; however the propagation loss can be calculated.

To determine which model is most appropriate for a given region, measurement campaigns can be performed in the area of interest to evaluate the performance of each model. This evaluation involves comparing the statistical errors of each model in relation to measured values quantitatively. Through these statistical parameters, a table comparing the models considered can be constructed, which allows a statistical analysis to determine which model best fits the aforementioned study region. Classic statistics is traditionally used to develop propagation models (Haneda et al., 2016; Salous, 2013; Shu Sun et al., 2014) [2-4]. Thus, it is assumed that realizations of random variables are mutually independent. However, there are several phenomena that involve scenarios that show spatial dependence.

The propagation models are generally based on deterministic models [1,2], and modified, based on the results obtained from measurement campaigns in one or more regions [1]. The models obtained are given through the abacus, as model of Okumura (A.Mawjoud, 2013) [5], for example, or expressions which provide the median attenuation, like the models of Okumura-Hata(Arthur et al., 2019; Gao et al., 2020) [6,7], Ibrahim-Parsons (Rozal et al., 2012) [8], Walfisch-Bertoni (Neto et al., 2003) [9], Ikegami-Walfisch (Alqudah, 2013; Cheerla et al., 2018) [10,11], Blomquist and Ladell (Loo et al., 2017) [12] and Lee (Wang et al., 2016) [13].

The predictive algorithms of these models usually deal with large volume of data, requires a sophisticated computer processing and knowledge, sometimes detailed, the topography of the terrain. For being based on measurements taken at specific locations, the empirical models tend not to provide very reliable results when applied to regions that differ significantly from the original region [3].

Therefore, classical statistics is often used when disregarding the possible correlation between neighboring samples; thus the relationships that may exist between the sample units are not explored satisfactorily. One of the methods that incorporates neighborhood effects in the calculation of the received power (dBm) is the methodology proposed in this study, namely, received power mapping in wireless communications networks by spatial inference using the Kriging process developed by Matheron (Matheron et al., 2019) [14]. Based on this methodology the spatial plan of received power measures (dBm) can be identified and, the vectors of better signal reception emitted by the BTS (base transceiver station) can be identified by the gradient of lines of iso-values; moreover, homogeneous zones can be identified as well as those where users may or may not be favored by the service operator. Thus, estimates of statistics, graphs, dispersion and surface maps that spatially describe the behavior of the power variable of the received signal (dBm) were obtained.

II. RELATED WORKS

Currently, many researchers employ geostatistical interpolation techniques for coverage prediction, based mainly on Kriging techniques. This Module includes a set of procedures necessary for geostatistical techniques (exploratory analysis, semivariogram generation and modeling and interpolation by kriging), aiming at the 2D analysis of spatially distributed data regarding the interpolation of surfaces generated from the georeferenced samples obtained from the received power. In [15], Konak (Konak, 2010) estimated signal propagation losses in wireless LANs using Ordinary Kriging (OK). In [16], Phillips et al. (Phillips et al., 2012) used OK on a 2.5 GHz WiMax network to produce radio environment maps that are more accurate and informative than deterministic propagation models. Kolay et al.(Kolyaie et al., 2011; Kolyaie & Yaghooti, 2011)[17,18] used drive-tests to collect signal strength measurements and compared the performance of empirical and spatial interpolation techniques. Mezhoud et al.(Mezhoud et al., 2020) [19] proposed an approach for coverage prediction based on the hybridization of the interpolation technique by OK and a Neural Network with MLP-NN architecture, this methodology was motivated by the lack of quality of the MLP-NN test database, which satisfactorily enriched the network's training dataset. Faruk et al.(Faruk et al., 2019) [20] evaluated and analyzed the efficiencies of empirical, heuristic and geospatial methods for predicting signal fading in the very high frequency (VHF) and ultra-high frequency (UHF) bands in typically urban environments. Path loss models based on artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS) and Kriging techniques were developed. Sato et al.(Sato et al., 2021) [21] proposed a technique that interpolates the representative map of the mobile radio signal in the spatial domain and in the frequency domain.

III. AREA OF STUDY

3.1. Area of Study

Belém, capital of the state of Pará, belonging to the metropolitan mesoregions of Belém. With an area of approximately 1064,918 km², located in northern Brazil, with a latitude of -01° 27' 21" and longitude of -48° 30' 16", altitude of 10 meters and distance of 2.146 miles of Brasília. The city is the capital of the Metropolitan Region of Belém, as with 2.100.319 inhabitants, is the 2nd most populous region, 12ª of the country 177ª of the world, well as being the largest urban agglomeration in the region. The city of Belém, considered the largest of the equator line, is also classified as a capital with the best quality of life in Northern Brazil. Fig. 1 shows an aerial view of the large urban center of the state capital bathed by the bay of Guajará.

Fig. 1: Partial view of the large urban center of Belém/PA (souce: https://cityofmangotrees.wordpress.com/2015/04/15/feliz-lusitania/, september 2021).

To observe the behavior of the received power (dBm), a measurement campaign involving 11 streets in the urban area of Belém - Pará, Brazil was performed. The acquisition of the verticalization and parcel measurements of buildings and residences, which resulted in a total of about 1800 points (from houses and buildings), was made by AUTOCADMAP and ORTOFOTO provided by the Company Development and Administration of the Metropolitan Area of Belem – CODEM.

The neighborhoods involved in the measurement campaign are located in the central region of Belém. The neighborhoods identified as Nazaré, Batista Campos and Umarizal present a high degree of vertical integration as a whole with more than 190 buildings and some that reach 70 meters high. The Alcindo Cacela, Conselheiro Furtado, Governor José Malcher, Magalhães Barata and Nazaré Avenues show a predominance of buildings of all sizes. These avenues have the greatest concentration of commercial buildings. In addition, there are many mango trees along the Magalhães Barata and Nazaré avenues.

3.2. Materials and Methods

3.2.1. Measurement Setup

The equipment used in the measurement setup included a transmission system and a receiving system. The transmission system consisted of a transmitting antenna positioned at a height of 35 m above the ground and operated by the local operator (Oi Celular). It was the 739632model produced by KathereinTM with dual polarization (± 45º) and operates within the range from 825 to 896 MHz with a gain of 15 dBi in vertical polarization.

The receiving system is the E7474A TDMA model produced by Agilent. The receiving antenna used in the measurements was the TPM 8003A monopole model produced by PlusTM, which operates in the range from 825 to 896 MHz with a gain of 3 dBi. It was mounted on a car and the received signal was collected by a laptop that had a PCMCIA card installed, which was the interface between the acquisition and storage system. In addition to the acquisition of the received power, the movement test system uses GPS coordination to determine the geographic position information of all measures.

3.2.2. Geostatistics

Geostatistics is used in spatial interpolation and quantification of uncertainty for variables that exhibit spatial continuity, i.e., that can be measured anywhere in the region of interest. It uses traditional statistical concepts such as random variables (RVs), cumulative distribution functions (CDFs), probability density functions (PDFs), expected value, and variance. In geostatistics, the RV represented by Z (u), where u is the vector of location coordinates, is related to a location in space. In this case, the main statistics are defined below (Gooverts, 1984; Isaaks, 1990) [22, 23].

The cumulative distribution function (CDF) provides the probability that the RV Z is less than or equal to a given z value, usually called the cut value.

$$F(u:z) = Pr\,o\,b\{Z(u) \leq z\} \quad (1)$$

The probability density function (PDF) is derived from the CDF, assuming it is differentiable, i.e.:

$$f(u:z) = F'(u:z)\lim_{dz\to 0}\frac{F(u;z+dz)-F(u;z)}{dz} \quad (2)$$

When the CDF is performed for a specific set of information, for example, (n) consisting of n neighboring data values Z(u) = z (u), a = 1,..., n, the notation "conditional to n" refers to the conditional cumulative distribution function (CCDF), which is defined as follows:

$$F(u:z|(n)) = Pr\,o\,b\{Z(u) \leq z|(n)\} \quad (3)$$

The expected value, $E\{Z\}$, is the weighted average of n possible outcomes where each outcome is weighted by its probability of occurrence. In the continuous case and assumingthat the integrals exist, the expected value is defined as follows:

$$E\{Z\} = m = \int_{-\infty}^{+\infty} z.\,dF(u;z)$$

$$= \int_{-\infty}^{+\infty} z.\,f(u;z)dz$$

$$\approx \sum_{k=1}^{K} z'_k[F(u;z_{k+1}) - F(u;z_k)] \quad (4)$$

With $z'_k \in]z_k, z_{k+1}]$

Where $F(u;z)$ and $f(u;z)$ are the CDF and PDF, respectively. The integral $\int_{-\infty}^{+\infty} z.\,dF(u;z)$ was approximated by *K* classes with frequencies $[F(u;z_{k+1}) -$

$F(u; z_k)$]and z'_k is a value in the *k-th*class, for example, the center of the class.

The variance$Var\{Z\}$Var{Z}, defined as the squared expected deviation V.A. Z in relation to its average in the continuous case, is written as follows:

$$Var\{Z\} = \sigma^2 = \int_{-\infty}^{+\infty} (z-m)^2 dF(u;z)$$

$$= \int_{-\infty}^{+\infty} (z-m)^2 f(u;z)dz \; (z-m)^2 f(u;z)dz \quad (5)$$

In many situations, it is desirable to know the pattern of dependence of one variable X in relation to another Y. The joint distribution of the results of a pair of random variables X and $Y$is characterized by the joint CDF (or bivariate), which is defined as follows:

$$F_{xy}(x,y) = Prob\{X \leq x, \; and \; Y \leq y\} \quad (6)$$

Which is estimated, in practice, by the proportion of data pairs jointly below the respective values (cut value) *x* and *y*. This distribution can be shown in a scatter diagram where each pair of data (x_i,y_i) is plotted as a point. The degree of dependence between two variables X and Ycan be characterized by the dispersion around the 45º line in the scatter diagram, as shown in Fig. 2.

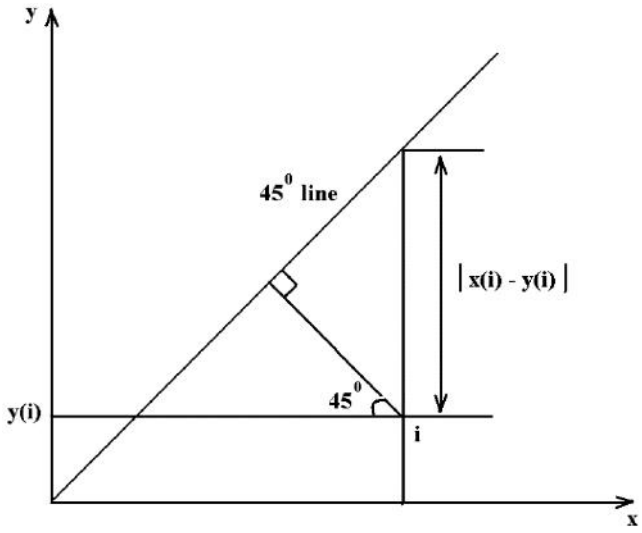

Fig. 2: Pair (x_i ,y_i) on a scattergram

The moment of inertia of the scatter diagram around the 45º line, called the "semivariogram" of the set of pairs (x_i,y_i) , is defined as follows:

$$\gamma_{XY} = \frac{1}{N}\sum_{i=1}^{N} d_i^2 = \frac{1}{2N}\sum_{i=1}^{N}(x_i - y_i)^2 \quad (7)$$

The higher the value of the semivariogram, the greater the dispersion and the less closely related are the two variables X and Y.

The centered covariance (on the average, *m*), or simply covariance, is given by the following:

$$\text{Cov}\{X,Y\} = \sigma_{XY} = E\{XY\} = E\{[X - m_x].[Y - my]\}$$

$$= E\{XY\} - m_x.my \quad (8)$$

The standard covariance between two RVs X and Y is known as the correlation coefficient, i.e.,

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{Cov\{X,Y\}}{\sqrt{Var\{X\}.Var\{Y\}}} \in [-1,+1] \quad (9)$$

The experimental relationship between the semivariogram and the covariance can be obtained by developing equation (7), which gives the following:

$$\gamma_{X'Y'} = 1 - \rho_{X'Y'} \in [0,2] \gamma_{X'Y'} = 1 - \rho_{X'Y'} \in [0,2] \quad (10)$$

Where $X'$and Y' are standardized variables$X' = (X - m_X)/\sigma_X$ and $Y' = (Y - m_Y)/\sigma_Y$.

The RVs X and Y can represent the same property measured in two different space locations, which are characteristic of regionalized variables at *x*and *x* + *h* that are separated by a vector *h* (called the lag or distance between locations), and $X = Z(x)$, $Y = Z(x + \boldsymbol{h})$.

In this case, the γ_{XY} semivariogram and the ρ_{XY} correlation measure the degree of variability or similarity between the two RVs X and Y.

This case is of particular interest in problems of spatial interpolation where an area (map) with a particular property, $Z(\boldsymbol{u})$, $\boldsymbol{u} \in \boldsymbol{A}, \boldsymbol{A}$area, must be identified from n samples of $Z(\boldsymbol{u})$. The combination of all n($\boldsymbol{h}$) data pairs of $Z(\boldsymbol{u})$over the same area with such pairs separated by approximately the same vector $\boldsymbol{h}$ (in length and direction) allows the characteristic (or experimental) semivariogram of the spatial variability in A to be estimated:

$$\gamma(h) = \frac{1}{2N(\boldsymbol{h})}\sum_{\alpha=1}^{N(\boldsymbol{h})}[z(\boldsymbol{u}_\alpha) - z(\boldsymbol{u}_\alpha + \boldsymbol{h})]^2 \quad (11)$$

The semivariogram characterizes the degree of spatial dependence between two random variables $Z(\boldsymbol{u})$ and $Z(\boldsymbol{u} + \boldsymbol{h})$ separated by vector $\boldsymbol{h}$.

With a single sample, all that is known of a random function $Z(\boldsymbol{u})$is a single point. Then, if the values for the non-sampled locations must be estimated, the restriction that the regionalized variable is statistically stationary must be introduced. To summarize, the hypothesis of stationarity establishes that the first two moments (mean and variance) of the difference $[Z(\boldsymbol{u}) - Z(\boldsymbol{u} + \boldsymbol{h})]$are independent of the $\boldsymbol{u}$ location and are only a function of the vector$\boldsymbol{h}$.

When the semivariogram graph is the same for any direction of $\boldsymbol{h}$, it is called isotropic, and it represents a much simpler situation than when it is anisotropic. In the latter case, the semivariogram should be transformed before being used. Therefore, it is advisable to examine semivariograms forseveral directions to evaluate the existence of anisotropy.

The semivariogram is the preferred tool for statistical inference because it has some advantages over covariance (Matheron et al., 2019) [14]. For a continuous function to be chosen with semivariograms, it is necessary to satisfy the defined positive property. In practice, linear combinations of basic models that are valid, i.e., permissible, are used. One of the most used basic models in geostatistics is the spherical model, which is given below:

$$\gamma_{(h)}\begin{cases}0, & |h|=0\\ C\left[\frac{3}{2}\left(\frac{|h|}{a}\right)-\frac{1}{2}\left(\frac{|h|}{a}\right)^3\right] & 0<|h|\le a\\ C & |h|>a\end{cases} \quad (12)$$

The C and components are called level and range, respectively. The level, also known as the "sill", represents the variability of the semivariogram until its stabilization. The range (or variogram amplitude) is the observed distance to the level where the variability stabilizes. It indicates the distance at which samples are spatially correlated, as shown in Fig. 3.

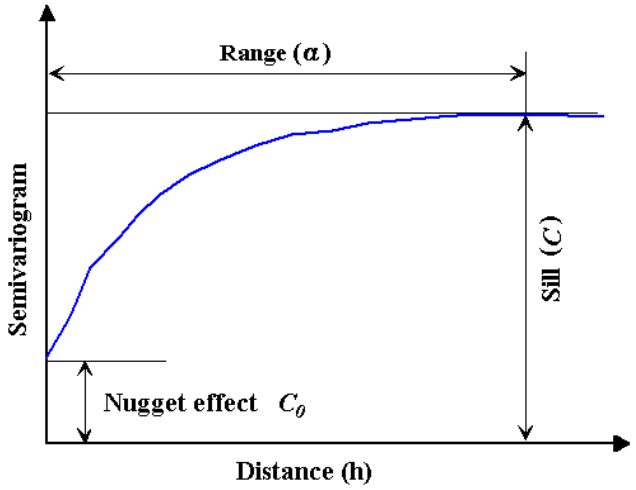

Fig. 3: Parameters of the semivariogram

After the model of spatial dependence between two random variables Z($\boldsymbol{u}$) and Z($\boldsymbol{u}+\boldsymbol{h}$)is established, the problem of estimating an unknown value $Z(\boldsymbol{u})$from the values available can be addressed. The goal is not only to find an estimate$Z^*(\boldsymbol{u})$ of the unknown value, but also model the uncertainty of this estimate. The uncertainty depends on the available information: the observed $z_{(u)}$'s values and\the established model of spatial dependence.

Kriging is a generic name adopted in geostatistics for a family of algorithms of least-squares regression based on the linear regression estimator $Z^*(\boldsymbol{u})$, which is given by the following:

$$Z^*(\boldsymbol{u})-m(\boldsymbol{u})=\sum_{\alpha=1}^{n(u)}\lambda_\alpha(\boldsymbol{u})[Z(\boldsymbol{u}_\alpha)-m(\boldsymbol{u}_\alpha)] \quad (13)$$

Where $\lambda_\alpha(\boldsymbol{u})$is the weight assigned to each observed value of $Z(\boldsymbol{u}_\alpha)$ located within a certain neighborhood $W(\boldsymbol{u})$ centered at $\boldsymbol{u}$. The $m(\boldsymbol{u})$ weights are chosen to minimize the estimation or error variance $\sigma_E^2(u)=Var[Z^*(\boldsymbol{u})-Z(\boldsymbol{u})]$ under the non-biased condition of the estimator.

The ordinary kriging (OK) considers the local variation of the average restricted to the domain of stationarity of the average to the local neighborhood $W(\boldsymbol{u})$centered on the location **u** to be estimated. In this case, the common average (stationary) $m(\boldsymbol{u}_\alpha)$ in equation 13 is considered. The unknown average $m(\boldsymbol{u}_\alpha)$can be eliminated by considering the sum of the weights ($\lambda_\alpha(\boldsymbol{u})$) of the Kriging equal to 1, i.e.,

$$Z_{KO}^*(u)=\sum_{\alpha=1}^{n(u)}\lambda_\alpha^{KO}(u)Z(u_\alpha) \quad \text{with} \quad \sum_{\alpha=1}^{n(u)}\lambda_\alpha^{KO}(u)=1 \quad (14)$$

The minimization of error variance ($Var[Z^*(\boldsymbol{u})-Z(\boldsymbol{u})]$) under the condition $\sum_{\alpha=1}^{n(u)}\lambda_\alpha^{KO}(u)=1$, allows the weights $\lambda_\alpha$to be determined from the following system of equations, called the ordinary Kriging system (normal equations with constraints):

$$\begin{cases}\sum_{\beta=1}^{n}\lambda_\beta^{KO}(\boldsymbol{u})C(\boldsymbol{u}_\beta-\boldsymbol{u}_\alpha)+\mu(\boldsymbol{u})=C(\boldsymbol{u}-\boldsymbol{u}_\alpha),\\ \sum_{\beta=1}^{n}\lambda_\beta^{KO}(\boldsymbol{u})=1\end{cases} \quad (15)$$

Where $C(\boldsymbol{u}_\beta-\boldsymbol{u}_\alpha)$ and$C(\boldsymbol{u}-\boldsymbol{u}_\alpha)$ are, respectively, the covariance among points $\boldsymbol{u}_\beta$ and $\boldsymbol{u}_\alpha$ and $\boldsymbol{u}$ and $\boldsymbol{u}_\alpha$ and $\mu(\boldsymbol{u})$ is the Lagrange parameter associated with the restriction $\sum_{\beta=1}^{KO}\lambda_\beta(u)=1$.

Unlike more traditional linear estimators, Kriging uses a weighting system that considers a spatial correlation model specific to the variable in study area A. Kriging provides not only a least squares estimate of the variable under study but also the associated variance error.

IV. ANALYSIS AND DISCUSSION OF RESULTS

Fig. 4 shows the data distributed in the study area. There are eleven streets in which measurements of the received power (dBm) were taken by the mobile station over fairly short distances (under 5 m) between measurements.

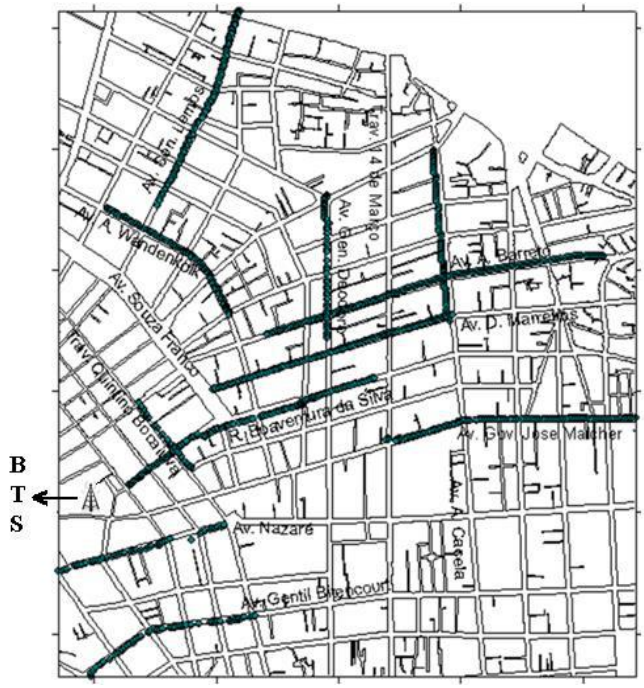

Fig. 4: Sampling points for power measurement in the study area [8]

Fig. 5 shows clearly that the distribution of the data evidences the slight asymmetry to the right, which indicates the presence of high values of received power (dBm). However, the values near the mean (-90.28) and median (-94.28) indicate that the distribution approaches normal values.

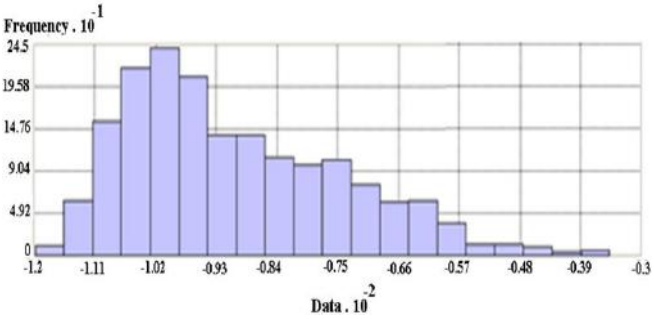

Fig. 5: Histogram for the received signal power (dBm)

To conduct the analysis and diagnosis of the effects of spatial autocorrelation samples, was used ARCGIS geostatistical module program(Johnston et al., 2001) [24] called ArcMap which is the application that is used for handling / generation of digital maps.

Fig. 6 presents a QQ-plot (quantil-quantil plot), which compares the received power distribution (dBm) with a standard normal distribution, note that there is a reasonable approximation of the distributions.

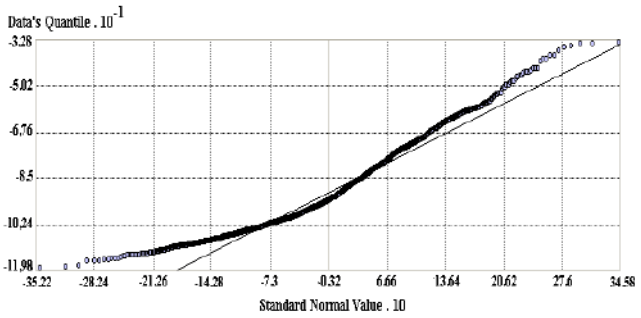

Fig. 6: QQ-plot for the power of the received signal (dBm)

The use of a geostatistical method requires that the data be spatially stationary (Pyrcz & Deutsch, 2014; Shiquan Sun et al., 2020; Tobler, 1989) [25-27]. It was observed that the distribution of the received power signal (dBm) shows higher peaks in certain directions, probably due to the short distance between the base station and the mobile station. This observation clarifies the presence of a spatial trend in the data. In this case, this trend should be removed and used for the geostatistical analysis because it is free of trends and therefore stationary. After the geostatistical analysis, the trend should be added to the results so that the predictions yield more accurate results. A first-order surface was used in this case to remove the trend, as shown in Fig. 7.

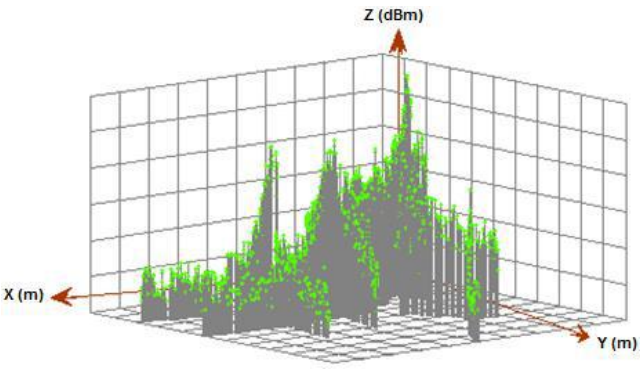

Fig. 7: Spatial distribution of power data of the received signal (dBm)

For the use of Kriging, foremost, an analysis is made by means of the spatial dependence of the semivariogram. Fig. 8 shows that the experimental points in the isotropic case semivariogram, where the samples show a range of spatial dependence around 30 meters. The semivariographic model in this case is given by:

$$\gamma(h) = 12.63C_0 + 194.3sph(335.06)(16)$$

Where: $C_0$is the nugget effect and $sph(335.06)$ is a spherical model for $h = 335.06$.

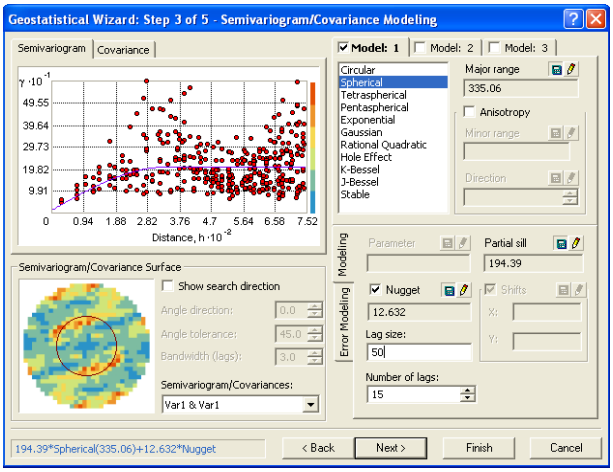

Fig. 8: Omnidirectional Semivariographic Model adjusted to the power of the received signal (dBm)

The semivariographic analysis showed the presence of an anisotropy from the southwest (SW) direction towards the northeast (NE) (Chilès & Delfiner, 2012) [28]. Because the trend was removed, the directional components of the spatial autocorrelation occur on a small scale, which will be included in the semivariographic modeling. The received power variable (dBm) was an ellipse of anisotropy with a major axis in the 57° direction (angle relative to geographic north) and minor axis in the 145° direction. Figs. 9 and 10 shows the semivariogram in both directions of 57° and 145°, respectively. The adjusted model, in this case, was represented by a spherical,

$$\gamma(h)_{57^{\circ}} = 29.19 + 189.06 \rightleftarrows Sph\left[\left(\frac{|h|}{749.17}\right) + \left(\frac{|h|}{288.32}\right)^{3}\right] (17)$$

$$\gamma(h)_{145^{\circ}} = 28.76 + 189.06 \rightleftarrows Sph\left[\left(\frac{|h|}{283.93}\right) + \left(\frac{|h|}{749.17}\right)^{3}\right] \quad (18)$$

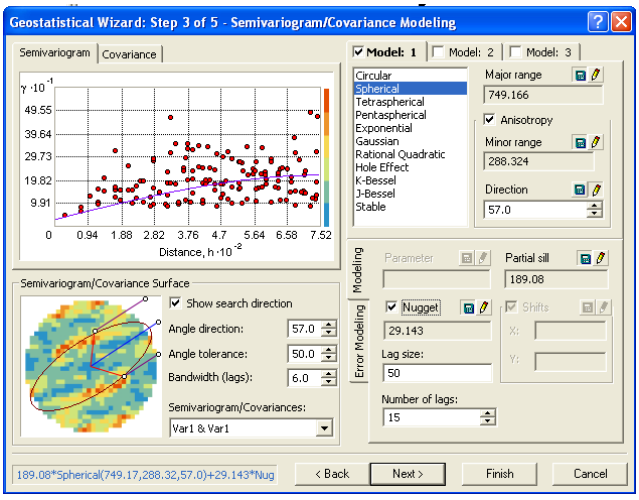

Fig. 9 :Directional Semivariographic Model in the 57º direction adjusted for the power of the received signal (dBm)

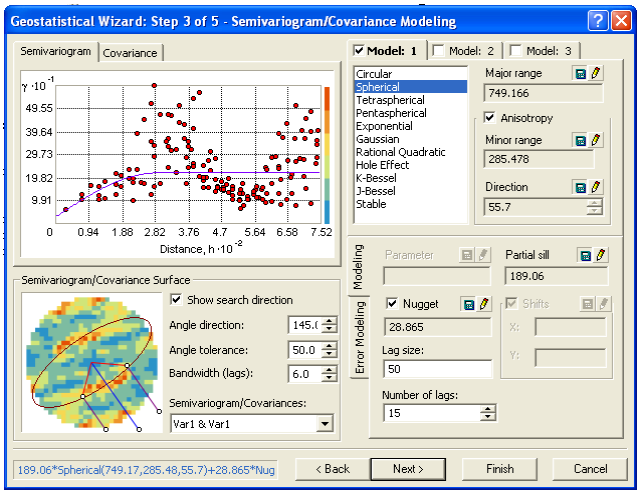

Fig. 10 : Directional Semivariographic Model in the 145º direction adjusted for the power of the received signal (dBm)

The nugget effect indicates that, around 13.38%=(29.195/(29.195+189.06))*100 of the total variability of the samples is due to the random component, and the remaining 86.62% is explained by the spatial autocorrelation component of the existing residues. The maximum and minimum ranges around 749.17 m and 288.32 m, respectively, indicate that there is an ellipse of influence of spatial contagion with maximumand minimum rays equal to the ranges; if these effects become negligible, the small spatial scale is determined from these limits. Thus, there is evidence that users who receive a signal with similar quality tend to be located close to each other.

Based on the framework defined by the variographic model, spatial inference was performed through the Kriging process to obtain a map of spatial distribution. The result of the mapping by Kriging for the received power (dBm) is shown in Fig. 11.

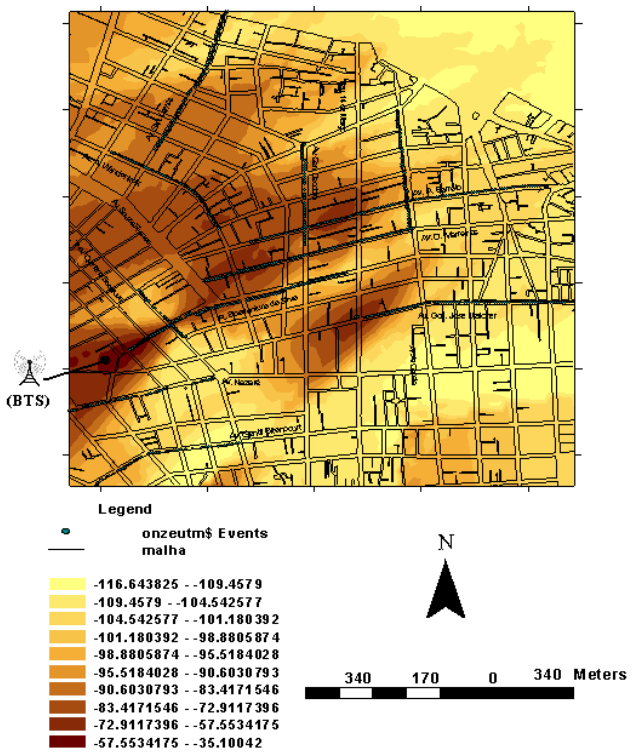

Fig. 11 : Spatial map of the received power (dBm)

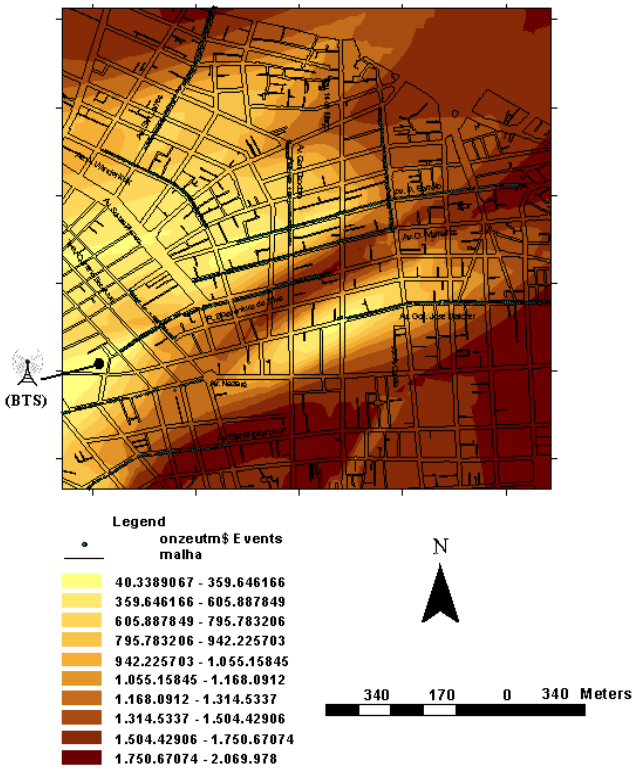

Fig. 12: Spatial map of the distance to the BTS

The map in Fig. 11 shows the spatial distribution using color levels to provide information about the distribution pattern of received power (dBm). The distribution of values shows the regions with higher levels of received power (dBm) in brown and the areas with lower signal strength in yellow. The highest levels of power are observed mainly in three regions: on the Governor Jose Malcher Avenue corner with Trav. 14 de Março; a large region that starts near the Trav. 14 de Março corner with Antonio Barreto Avenue and extends toward the Visconde de Souza Franco Avenue and Boaventura da Silva street; and another small region near the Dr. Moraes street with Boaventura da Silva street. The BTS is located near this area.

Using the same methodology adopted for the received power variable (dBm), a map of the spatial distribution of the color levels that provides information about the spatial distribution pattern of the distance from the mobile station to the transmitting antenna is shown Fig. 12. The potential of the applied methodology can be observed when comparing the maps showing the spatial distribution of the received power (dBm) by the receiving unit and the distance between the transmitter and receiver antennas. As anticipated, lower power levels are observed at greater distances from the base station.

V. COMPARISON OF RESULTS

Fig. 13 shows the variation of the received signal strength (simulated and measured) and theoretical models Okumura-Hata [6] and Ibrahim-Parsons [8] as a function of distance from the transmitter antenna along the eleven paths studied.

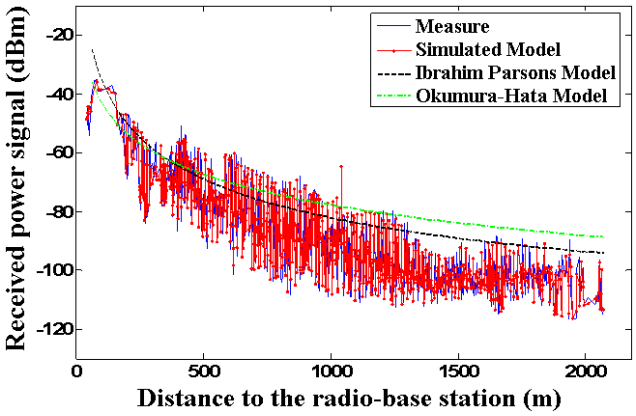

Fig. 13: The signal strength received by the mobile station and estimated by theoretical models and simulated

Each type of propagating approached showed some random variations regarding the classification of the analyzed environment. In the case of region analysis, simulations of all models were performed considering the involved characteristics of urban environment, whereas this environment predominate residences and buildings with an average height ranging from 3m to 100m, respectively. The parameters used in the analysis of models had the following values:

The receiving antenna height: $h_r = 1.5$ m.

The receiving antenna height: $h_t = 35$ m.

Operating frequency: f = 877.44 MHz

In order to perform a study more insightful for each model, were made for each data file, statistical analysis of the measures, aiming to measure the deviations between the simulated and measured values. This made it possible to obtain information to provide subsidy to state what the best model for characterizing the propagation environment for mobile cellular paths studied. Table 1 shows the average and standard deviation in dB for each model in relation to the values of received signal strength from the field.

Table 1:Comparison between the three theoretical models and the measured value for the paths involved in the measurement campaign

Measure/ Model	Mean squared error (dB)	Mean (dBm)	Standard deviation (dBm)
Measured	-	-89.7875	15.7682
Proposed model	0.37	-89.1993	16.1729
Ibrahim-Parsons	13.23	-79.8353	11.0324
Okumura et al.	16.54	-76.2427	8.8034

Through the analysis of graphs and the results shown in Table 1, one can deduce that the proposed model is resulting in lower average deviation compared with the field measurements. For this model, the mean square error with respect to the level of theoretical power is 0.37 dB Whereas the maximum acceptable deviation in the signal level received by the mobile in relation to the prediction, is 8 dB Note that the average and standard deviation calculated for the proposed model showed values very similar to those obtained for the data collected in the field (measured values).

However, among the theoretical models, nearest measured values is the template of Ibrahim-Parsons, with a mean square error of 13.23 dB The model of Okumura-Hata, that had the worst outcome, one should go through their adjustment coefficients. Possibly the urban environment analyzed for obtaining this model does not show many similarities with that found in the region studied in this study.

VI. CONCLUSIONS

In view of the above, it can be concluded that the spatial inference allows the regions where the levels of received power (dBm) are either intense or not to be identified, which demonstrates the negative effects to the subscribers who are in regions with low signal levels. In addition, this methodology allows all parts of the region of interest to be assessed individually based on their geographical coordinates and not just a generic statement of values, as in traditional propagation models. From the spatial distribution map of the received power (dBm), it is possible to identify the areas that are over- or underestimated in terms of signal reception, which can result in increased investment by the local operator to those regions where the signal is weak. Moreover, the spatial mapping of the received power (dBm) can also help in planning and developing wireless communications networks because iso-value maps can be used to identify neighborhoods that benefit from having high received power (dBm) in a given city.

The model proposed showed an good result with mean square error in order of 0.37 dB in relation to the measured signal, considering the data of the eleven paths of measuring campaign; whereas for the models of Ibrahim Parsons and Okumura-Hata this error was on the order of 13.23 and 16.54 dB, respectively. This performance is due to the fact that the geostatistical model considered the georeferenced data, enabling the identification of the interaction effects in this same space, using a kriging process.Therefore, the spatial estimation techniques used for wireless communications networks should be applied to other scenarios to estimate the signal strength along all avenues of a given city.

ACKNOWLEDGEMENTS

The authors thank the Capes and Project Ericsson/Oi Celular/UFPA for the experimental setup and the data used.

REFERENCES

[1] Liaskos, C., Nie, S., Tsioliaridou, A., Pitsillides, A., Ioannidis, S., & Akyildiz, I. (2018). A New Wireless Communication Paradigm through Software-Controlled Metasurfaces. *IEEE Communications Magazine*, *56*(9), 162–169. https://doi.org/10.1109/MCOM.2018.1700659

[2] Haneda, K., Zhang, J., Tan, L., Liu, G., Zheng, Y., Asplund, H., Li, J., Wang, Y., Steer, D., Li, C., Balercia, T., Lee, S., Kim, Y., Ghosh, A., Thomas, T., Nakamura, T., Kakishima, Y., Imai, T., Papadopoulos, H., … Ghosh, A. (2016). 5G 3GPP-like channel models for outdoor urban microcellular and macrocellular environments. *IEEE Vehicular Technology Conference*, *2016-July*. https://doi.org/10.1109/VTCSpring.2016.7503971

[3] Salous, S. (2013). *Measurement and Channel Modelling Radio Propagation Measurement and Channel Modelling*. John Wiley & Sons Ltd.

[4] Sun, Shu, Rappaport, T. S., Heath, R. W., Nix, A., & Rangan, S. (2014). MIMO for millimeter-wave wireless

communications: Beamforming, spatial multiplexing, or both? *IEEE Communications Magazine*, *52*(12), 110–121. https://doi.org/10.1109/MCOM.2014.6979962

[5] A.Mawjoud, S. (2013). Comparison of Propagation Model Accuracy for Long Term Evolution (LTE) Cellular Network. *International Journal of Computer Applications*, *79*(11), 41–45. https://doi.org/10.5120/13789-1931

[6] Arthur, J. K., Amartey, A. T., & Brown-Acquaye, W. (2019). Adaptation of the Okumura-Hata Model to the Environment of Accra. *2019 International Conference on Communications, Signal Processing and Networks, ICCSPN 2019*, 1–6. https://doi.org/10.1109/ICCSPN46366.2019.9150198

[7] Gao, R., Zhao, Y., Wang, Y., & Yan, T. (2020). An Improved Propagation Prediction Model Based on Okumura-Hata Model. In *Advances in Intelligent Systems and Computing* (Vol. 905). https://doi.org/10.1007/978-3-030-14680-1_101

[8] Rozal, E., Pelaes, E., Queiroz, J., & Salame, C. (2012). Modeling of wireless networks using multivariate time models. *Eurasip Journal on Advances in Signal Processing*, *2012*(1), 1–13. https://doi.org/10.1186/1687-6180-2012-248

[9] Neto, A. P., Rozal, E. O., & Pelaes, E. G. (2003). Bidimensional Statistics Analysis. *Proceedings of the 2003 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference - IMOC 2003. (Cat. No.03TH8678)*, *2*, 801–806. https://doi.org/10.1109/IMOC.2003.1242682

[10] Alqudah, Y. A. (2013). On the performance of Cost 231 Walfisch Ikegami model in deployed 3.5 GHz network. *2013 The International Conference on Technological Advances in Electrical, Electronics and Computer Engineering, TAEECE 2013*, 524–527. https://doi.org/10.1109/TAEECE.2013.6557329

[11] Cheerla, S., Ratnam, D. V., & Borra, H. S. (2018). Neural network-based path loss model for cellular mobile networks at 800 and 1800 MHz bands. *AEU - International Journal of Electronics and Communications*, *94*(July), 179–186. https://doi.org/10.1016/j.aeue.2018.07.007

[12] Loo, Z. Bin, Chong, P. K., Lee, K. Y., & Yap, W. S. (2017). Improved path loss simulation incorporating three-dimensional terrain model using parallel coprocessors. *Wireless Communications and Mobile Computing*, *2017*. https://doi.org/10.1155/2017/5492691

[13] Wang, T., Ai, B., He, R., & Zhong, Z. (2016). Local mean power estimation over fading channels. *IEEE Vehicular Technology Conference*, *2016-July*, 16–20. https://doi.org/10.1109/VTCSpring.2016.7504323

[14] Matheron, G., Pawlowsky-Glahn, V., & Serra, J. (2019). *Matheron's Theory of Regionalised Variables*. Oxford University Press. https://doi.org/10.1093/oso/9780198835660.001.0001

[15] Konak, A. (2010). Estimating path loss in wireless local area networks using ordinary kriging. *Proceedings - Winter Simulation Conference*, *Badman 2006*, 2888–2896. https://doi.org/10.1109/WSC.2010.5678983

[16] Phillips, C., Ton, M., Sicker, D., & Grunwald, D. (2012). Practical radio environment mapping with geostatistics. *2012 IEEE International Symposium on Dynamic Spectrum Access Networks, DYSPAN 2012*, 422–433. https://doi.org/10.1109/DYSPAN.2012.6478166

[17] Kolyaie, S., & Yaghooti, M. (2011). Evaluation of Geostatistical Analysis Capability in Wireless Signal Propagation Modeling. Geocomputation 2011, 76–83.

[18] Kolyaie, S., Yaghooti, M., & Majidi, G. (2011). Analysis and simulation of wireless signal propagation. *Archives of Photogrammetry, Cartography and Remote Sensing*, *22*(1), 261–270.

[19] Mezhoud, N., Oussalah, M., Zaatri, A., & Hammoudi, Z. (2020). Hybrid Kriging and multilayer perceptron neural network technique for coverage prediction in cellular networks. *International Journal of Parallel, Emergent and Distributed Systems*, *35*(6), 682–706. https://doi.org/10.1080/17445760.2020.1805609

[20] Faruk, N., Popoola, S. I., Surajudeen-Bakinde, N. T., Oloyede, A. A., Abdulkarim, A., Olawoyin, L. A., Ali, M., Calafate, C. T., & Atayero, A. A. (2019). Path Loss Predictions in the VHF and UHF Bands within Urban Environments: Experimental Investigation of Empirical, Heuristics and Geospatial Models. *IEEE Access*, *7*, 77293–77307. https://doi.org/10.1109/ACCESS.2019.2921411

[21] Sato, K., Suto, K., Inage, K., Adachi, K., & Fujii, T. (2021). Space-frequency-interpolated radio map. *IEEE Transactions on Vehicular Technology*, *70*(1), 714–725. https://doi.org/10.1109/TVT.2021.3049894

[22] Gooverts, P. (1984). Geostatistics for Natural Resources Characterization. In *Geostatistics for Natural Resources Characterization*. https://doi.org/10.1007/978-94-009-3701-7

[23] Isaaks, E. H. (1990). Applied geostatistics. *Choice Reviews Online*, *28*(01), 28-0304-28–0304. https://doi.org/10.5860/choice.28-0304

[24] Johnston, K., Ver Hoef, J. M., Krivoruchko, K., Lucas, N., & Magri, A. (2001). *ArcGIS 9 Geostatistical Analyst Tutorial*.

[25] Pyrcz, M. J., & Deutsch, C. V. (2014). *Geostatistical Reservoir Modeling* (2nd ed., Vol. 148). Oxford University Press.

[26] Sun, Shiquan, Zhu, J., & Zhou, X. (2020). Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies. *Nature Methods*, *17*(2), 193–200. https://doi.org/10.1038/s41592-019-0701-7

[27] Tobler, W. R. (1989). Frame independent spatial analysis. *The Accuracy of Spatial Databases*, *December*, 115–122. https://doi.org/10.1201/b12612-33

[28] Chile` S, J.-P., & Delfiner, P. (2012). *Geostatistics: Modeling Spatial Uncertainty* (2nd ed., Vol. 148). John Wiley & Sons, Inc.