

REVIEW

Open Access

Modeling of wireless networks using multivariate time models

Edilberto Rozal^{1*}, Evaldo Pelaes², Joaquim Queiroz³ and Camil Salame^{3,4}

Abstract

The literature analysis of propagation models has investigated different statistical prediction methods to identify appropriate techniques for this purpose. This article presents the results of propagation channel modeling, based on multivariate time series models using data collected in measurement campaigns and the main characteristics of urbanization in the city of Belém-PA. Transfer function models were used to evaluate the relationship between received power signal and other variables, such as the height of buildings, the distance between buildings, and the distance to the radio base station. A multivariate model was designed in which the contributions due to the height of the buildings and the distance between buildings had a significant effect on the received power signal. The utilized sample failed to identify the contribution of distance to the source for the received power signal. The result obtained with the proposed model appeared to be accurate for the samples used in the study.

Keywords: autocorrelation function, ARIMA model, multivariate time model, partial autocorrelation function, transfer function, wireless

Review

Introduction

Currently, there is a wide range of channel models, with theoretical fundamentals and experimental for predictions for mitigation of path in mobile communication systems, consisting of one of the most important steps in the planning of mobile cellular systems. A correct prediction enables the designer of mobile systems to have the ability to predict the minimum power that is required to radiate from a transmitter to meet acceptable quality cover in a predetermined area, which is extremely important for improving the technique of frequency reuse and implementation of projects with shared bandwidth.

These models differ in their applicability on different types of terrain and different environmental conditions. Thus, there is a suitable model for all situations [1]. In real cases, the land on which the propagation occurs has varied topography, vegetation, and buildings randomly distributed. Statistical methods [2] consider a statistical modeling of signal fading and interprets the mobile radio signal as a random variable whose

probability distribution is to be determined. Although the propagation loss calculation can be performed, yet with limited accuracy, using techniques such as ray tracing, since the intrinsic limitation of this approach is the high processing time required to evaluate all routes ray [3,4]. There has been proposed a technique to improve performance computing, enabling to optimize the simulation models based on the propagation of 3D ray tracing techniques (*Ray-tracing 3D*) [5].

The most commonly used methods for the calculation of coverage are empirical or semi-empirical [6,7]. The empirical models are proposed from extensive collection of information signal coverage of a given region and the use of interpolation techniques. Resulting in expressions that allow to calculate the average attenuation of the signal path in the area remains question. In [8], analyses of the information available about various propagation models for both indoor and outdoor environments have been performed, where the main characteristics of the radio channel, such as fading and path loss, are discussed.

The models of Perez-Vega and Zamanillo [9] and Hata [10] are the empirical models. On the other hand, the propagation models are also derived semi-empirical from collecting information to cover a given area and the

* Correspondence: edilibe@ufpa.br

¹Department of Mathematics, Federal University of Pará (UFPA), Castanhal, Pará, Brazil

Full list of author information is available at the end of the article

classical electromagnetic theory, considering phenomena such as reflection structures, refraction, and diffraction at the edges of obstacles. Belonging to the class of models of Ibrahim–Parsons [11], Walfisch–Ikegami [12], and others.

In order to make the expressions obtained by Okumura et al. [13] suitable for computer implementation, Hata developed analytical expressions, based on the technique of Okumura, that allow to calculate the average attenuation path, for urban areas, suburban, and open (rural). The equations proposed by Hata are limited to certain ranges of input parameters (distances up to 100 Km) and are applicable only to land almost flat and are valid for the frequencies of 150 to 1500 MHz. This model has been proven to be accurate and is used by computer simulation tools.

The propagation model for urban environments formulated by Ibrahim–Parsons takes urbanization levels into account, use of the terrain, and the variation in height between the mobile station (EM) and the radio base station (RBS). These empirical characteristics were extracted from measurements taken in the city of London, on frequencies between 168 and 900 MHz. This model was studied in urban areas without undulations. It is used for distances between antennas smaller than 10 km and receiving antenna height of less than 3 m.

The model of Walfisch–Ikegami has its formulation based on the characteristics of urban regions, such as density and average height of buildings, and the width of the streets. This model is effective in cases where the height of the antennas RBS is smaller than the average height of buildings, a situation where there is considerable guidance signal RF along the routes considered. This model predicts two different situations for calculating the average attenuation path between RBS and the mobile: the line-of-sight.

To determine which model is most appropriate for a given region, measurements can be performed in the area of interest and used to evaluate the performance of each model against the measurements. However, determination of the best locations for the RBS is very important for obtaining adequate data rates and for providing a better estimation of coverage without needing to conduct a series of signal propagation measurements, which are very expensive and time-consuming. It is therefore important to develop effective propagation models for mobile communications in order to provide guidelines for efficient estimations in mobile communication systems.

This article presents a time series model for characterizing the received power signal (dBm) at the Gentil Bittencourt Avenue, located in the urban area of the city of Belem, Pará, Brazil. This study addresses the possible links between received power signal and the influence of

the height of the buildings, as well as the distance between buildings, the transmitter and receiver units. Transfer function models were used to assess the effects of the received power on the time series and to evaluate the relationship between the height of buildings and other variables of interest.

Materials and methods

Measurement environment

With the objective of observing the behavior of the received power (dBm), measurements were performed involving one street in the urban area of Belém-PA, Brazil. The acquisition of verticalization and measurements of the tested buildings and residences, which summed to a total of approximately 254 points (including residences and buildings), was performed using AUTOCADMAP and ORTOFOTO, obtained in cooperation with the Companhia de Desenvolvimento e Administração da Área Metropolitana de Belém—PA, CODEM.

The Gentil Bittencourt Avenue, involved in the measurement campaign, is located in the large urban center of the city of Belém. This pathway predominates a high level of vertical integration, exceeding 150 buildings; some even reach 80 m in height. The study of the street represents a rich area for theoretical analysis of cellular coverage-based propagation models, empirical and semi-empirical, given the size of the analyzed area (macrocell). Figure 1 represents a part of the Gentil Bittencourt Avenue, where the data collection took place.

Setup of measurement

The set of equipment used in the setup of measurements is given by a transmission system and a receiving system. The transmission system consists of a transmitter antenna used by the local operator of the model 739632 manufactured by KathereinTM with dual polarization ($\pm 45^\circ$) and operates on a scale of 880; a 960 MHz with a gain of 14.5 dBi of vertical polarization, positioned at a height of 35 m above the ground. The actually radiated power by the transmitting antenna was 42.29 dBm.

During the measurements, a CW signal on the frequency of 890.4 MHz was transmitted. To eliminate the interference, all channels using the same frequency or frequencies adjacent to the one used in the measurements were off the plant operator's site. The receiving system is the model E7474A TDMA produced by Agilent. The receiving antenna used in the measurements was a monopole model TPM 8003A produced by PlusTM which operate within a range of 825; 896 MHz with gain of 3 dBi. This was assembled on a car at a height of 1.5 m above the ground and the incoming signal was issued for a laptop that had a PCMCIA card installed, this card was the interface between the acquisition and storage system. Besides the acquisition



Figure 1 View of a part of the Gentil Bittencourt Avenue, typically urban environment containing buildings up to 80 m height (font Google earth, September 2012).

of the received power, the test motion system uses a GPS to give the position of geographic information of all measurements. Due to the environment of measurement, it has a high planting density of mango trees, at many points occurred loss of satellite visibility, causing GPS not providing an accurate measurement of geographical position. In order to avoid this loss of position, a chronometer was used to measure the time that the vehicle (with constant speed of 20 km/h) used to travel the entire length of the studied street. A subsequent treatment, performed in Matlab 7.1, converted the time of distance.

Time series

A time series is a set of statistics, usually collected at regular intervals. Time series data occur naturally in many application areas, such as economics, finance, environmental, and medicine. The methods of time series analysis pre-date those for general stochastic processes and Markov Chains. The aims of time series analysis are to describe and summarize time series data, fit low-dimensional models, and make forecasts [14].

We write our real-valued series of observations as $\dots X_{-2}, X_{-1}, X_{-0}, X_1, X_2, \dots$, a doubly infinite sequence of real-valued random variables indexed by integer numbers.

One simple method of describing a series is that of classical decomposition. The notion is that the series can be decomposed into four elements:

- Trend (T_t)—long-term movements in the mean;
- Seasonal effects (I_t)—cyclical fluctuations related to the calendar;

- Cycles (C_t)—other cyclical fluctuations (such as a business cycles);
- Residuals (E_t)—other random or systematic fluctuations.

The idea is to create separate models for these four elements and then combine them, either additively:

$$X_t = T_t + I_t + C_t - E_t \quad (1)$$

or multiplicatively:

$$X_t = T_t \cdot I_t \cdot C_t \cdot E_t \quad (2)$$

ARIMA models

Box et al. [15] first introduced ARIMA models, the term deriving from autoregressive integrated and moving average.

A key concept underlying time series processes is that of stationarity. A time series is stationarity when it has the following three characteristics:

- Exhibits mean reversion in that it fluctuates around a constant long-run mean;
- Has a finite variance that is time-invariant;
- Has a theoretical correlogram that diminishes as the lag length increases.

The autoregressive process of order p is denoted AR (p), and defined by

$$Y_t = \sum_{i=1}^p \varphi_i Y_{t-i} + e_t \quad (3)$$

where $\varphi_1, \dots, \varphi_p$ are fixed constants. Y_t is expressed linearly in terms of current and previous values of a

white noise series $\{e_t\}$. This noise series is constructed from the forecasting errors; $\{e_t\}$ is a sequence of independent (or uncorrelated) random variables with mean 0 and variance σ^2 .

Using the lag operator L (the lag operator L has the property: $L^n Y_t = Y_{t-n}$) we can write the AR(p) model as

$$Y_t \left(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \right) = e_t, \quad (4)$$

$$\Phi(L) Y_t = e_t. \quad (5)$$

where $\Phi(L) Y_t$ is a polynomial function of Y_t .

The moving average process of order q is denoted MA (q) and defined by

$$Y_t = e_t + \sum_{i=1}^q \theta_i e_{t-i}, \quad (6)$$

where θ_1, θ_q are fixed constants, $\theta_0 = 1$, and $\{e_t\}$ are the sequence of independent (or uncorrelated) random variables with mean 0 and variance σ^2 .

Or, using the lag operator

$$Y_t = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_p L^q) u_t, \quad (7)$$

$$Y_t = \Theta(L) u_t \quad (8)$$

The combination of the two processes to give a new series of models called ARMA(p, q) models is defined by

$$Y_t = \sum_{i=1}^p \phi_i Y_{t-i} + e_t + \sum_{i=1}^q \theta_i e_{t-i}. \quad (9)$$

where again $\{e_t\}$ is white noise, $\{\phi_i / i = 1, 2, \dots, p\}$ are the coefficients of AR model and $\{\theta_i / i = 1, 2, \dots, q\}$ are the coefficients of MA model.

Using the lag operator

$$Y_t \left(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \right) = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_p L^q) e_t, \quad (10)$$

$$\Phi(L) Y_t = \Theta(L) e_t. \quad (11)$$

There is an assumption for ARMA process that the time series for analysis should be stationary, that is, the mean of the time series and the covariance among its observations are not time-varied. This process is stationary for appropriate ϕ and θ .

According to the target model, the process is non-stationary, so the series should be transformed to a stationary process by the model construction. This can be often achieved by a differentiation process. The first-order differencing of the original time series is defined as

$$\Delta Y_t = Y_t - Y_{t-1} = Y_t - B Y_t.$$

For the high-order differentiation, we have

$$\Delta^d Y_t = (1 - B)^d Y_t. \quad (13)$$

If we ever find that the differenced process is a stationary process, then we can look for an ARMA model of that. The process $\{Y_t\}$ is said to be an ARIMA(p, d, q). If $X_t = \Delta^d Y_t$ is an ARIMA (p, q) process.

After the d th-order differentiations of Y_t in Equation (10), the ARIMA (p, d, q) can be constructed as

$$\Phi(L) Y_t^d = \Theta(L) e_t. \quad (14)$$

A time series (TS) may be defined as a set of observations Y_t as a function of time [16]. The principal tools utilized for analysis of a time series are the autocorrelation and partial autocorrelation functions (PACFs).

The autocorrelation function (ACF) represents a simple correlation between Y_t and Y_{t-k} as a function of the lag k . The ACF of TS $\{Y_t\}$ may be defined as [16]

$$\rho = \frac{\sum_{t=0}^{N-k-1} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=0}^{N-1} (Y_t - \bar{Y})^2}. \quad (15)$$

where N represents the length of the TS and \bar{Y} is the expected value from the observations, calculated for the time variation (*delay*) k . The autocorrelation coefficient (ρ) of a TS varies between -1 and 1 .

The PACF represents the correlation between Y_t and Y_{t-k} as a function of the lag k , filtering the effect of the other lags on Y_t and Y_{t-k} . The PACF is defined as the sequence of correlations between $(Y_t$ and $Y_{t-1})$, $(Y_t$ and $Y_{t-2})$, $(Y_t$ and $Y_{t-3})$, and so on, because the effects of prior lag on t remain constant. The PACF is calculated as the coefficient value ϕ_{kk} in the equation

$$Y_t = \phi_{k1} Y_{t-1} + \phi_{k2} Y_{t-2} + \phi_{k3} Y_{t-3} + \dots + \phi_{kk} Y_{t-k} + e_t. \quad (16)$$

Transfer function model

Transfer function model is different from the ARIMA model. The ARIMA model is univariate time series model, but transfer function is multivariate time series model. This means that ARIMA model relates the series only to its past. Besides the past series, transfer function model also relates the series to other time series. Transfer function models can be used to model single-output and multiple-output systems [16]. In the case of single-output model, only one equation is required to describe the system. It is referred to as a single-equation transfer function model. A multiple-output transfer function model is referred to as a multi-equation transfer

function model or a simultaneous transfer function model (see [17-20]).

The models of transfer functions use predicted values of explaining variables so as to produce predictions for the dependent variable. A general model of the transfer function, with an explaining variable and k lags, is expressed as

$$Y_t = \mu + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + \beta_k X_{t-k} + Z_t \quad (17)$$

Where:

- Y_t is the dependent variable of the temporal series;
 - μ is the constant term;
 - β_0 is the parameter associated with the present value of the explaining variable;
 - X_t is the present value of the explaining variable;
 - β_1, \dots, β_k are parameters associated with the discrepant values of the explaining variable.
- Such values are called *impulse response weight* or *transfer function weight*;
- Z_t is the temporal series of auto-correlated errors.

Transfer function models are used in series related to one or more entry series. For example, the daily consumption of electricity may be related to certain environmental variables, such as external temperature and relative humidity.

Assume that X_t and Y_t are properly transformed series such that both are stationary. In a linear system with simple input and output, the series of X_t input and Y_t output are related through a linear filter as

$$Y_t = v(B)X_t + N_t, \quad (18)$$

where $v(B) = \sum_{j=-\infty}^{\infty} v_j B^j$ is referred to as a filter transfer function by Box et al. and N_t is a noise series of the system that is independent of the input series X_t .

The coefficients in the transfer function model (18) are frequently called impulse response weights. The transfer function model is considered stable if the sequence of those impulse response weights is finite (absolutely addable), i.e., $\sum |v_j| < \infty$. Therefore, in a stable system, a limited entry always produces a limited exit. In a casual model, the system does not respond the entry series until it has really been applied into the system. In other words, the exit is affected by the entries in the system only in terms of past or present values. A casual model is also called realizable model, as it seems to be all those real physical systems. In practice, one considers frequently only the following casual stable model:

$$Y_t = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots + N_t = v(B)X_t + N_t \quad (19)$$

where $v(B) = \sum_{j=-\infty}^{\infty} v_j B^j$, $\sum |v_j| < \infty$ and X_t and N_t are independent.

The objective of the modeling of transfer function is to identify and to estimate the transfer function $v(B)$ and the noise model for N_t on the basis of the available information of the entry series X_t and of the response series Y_t . The greatest difficulty is that the information about X_t and Y_t is finite, and the transfer function in (19) contains a infinite number of coefficients. So, as to relieve that difficulty, the transfer function $v(B)$ is represented the following rational way

$$v(B) = \frac{w_s(B)B^b}{\delta_r(B)}, \quad (20)$$

where $w_s(B) = w_0 - w_1 B - \dots - w_s B^s$, $\delta_r(B) = 1 - \delta_1 B - \dots - \delta_r B^r$, and b is a discrepancy parameter that represents the lag of present time which occurs before the impulse of the entry variable may produce an effect over the exit variable. For a stable system, one assumes that the roots of $\delta_r(B) = 0$ are out of the unitary circle [14]. Once $w_s(B)$, $\delta_r(B)$, and b are found out, the v_j weights of the impulse response may be obtained by being equal the coefficients of B^j in both sides of the following

$$\delta_r(B)v(B) = w_s(B)B^b \quad (21)$$

In practice, the values of r and s in system (21) rarely exceed 2. Some transfer function models may be seen in Wei [16]. Such models may be used in the identification of the transfer function parameters. The analysis of such models indicate that the peak occurrences suggest parameters in the numbering item of the transfer function, very similar to models of moving averages, and the occurrence with a behavior of exponential decaying may indicate the existence of parameters in the transfer function denominator, which is similar to auto-regressive models.

The cross-correlation function (CCF) is a useful measure for direction and intensity between two random variables. For two stochastic processes X_t and Y_t for $t = 0, \pm 1, \pm 2, \dots$, one may say that X_t and Y_t are stationary together with each other if both of them are univariate stationary processes and if the cross correlation between X_t and Y_t , $\text{Cov}(X_t, Y_t)$ is a function only of time difference ($s - t$). In such cases, the cross covariance function between X_t and Y_t is [16]

$$\gamma_{xy}(k) = E \left[(X_t - \mu_x) (Y_t - \mu_y) \right] \quad (22)$$

where E represents the (average) mathematical expectance.

For $k = 0, \pm 1, \pm 2, \dots$, with standardization, we have got the following CCF:

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x \sigma_y} \quad (23)$$

For $k = 0, \pm 1, \pm 2, \dots$, where σ_x and σ_y are the deviations standard of X_t and Y_t , respectively. It is important to realize that the cross covariance function $\gamma_{xy}(k)$ and the CCFs $\rho_{xy}(k)$ are the generalizations of the autocovariance and of the autocorrelation functions because $\gamma_{xx}(k) = \gamma_x(k)$ and $\rho_{xx}(k) = \rho_x(k)$. However, unlike correlation function, the CCF is not asymmetrical, that is, $\rho_{xy}(k) \neq \rho_{xy}(-k)$. Instead, we have got

$$\begin{aligned} \gamma_{xy}(k) &= E(X_t - \mu_x)(Y_{t+k} - \mu_y) \\ &= E(Y_{t+k} - \mu_y)(X_t - \mu_x) = \gamma_{yx}(-k) \end{aligned} \quad (24)$$

which drives to

$$\rho_{xy}(k) = \rho_{yx}(-k) \quad (25)$$

Thus, the CCF measures not only the intensity of an association, but also its direction. So, the relation between the X_t and Y_t series is important to examine the CCF, $\rho_{xy}(k)$, for both lags, the negative and the positive ones, $k > 0$ and $k < 0$. The diagram of the CCF is known as cross correlogram.

The CCF, $\rho_{xy}(k)$, is defined only when X_t and Y_t are bivariate processes and at the same time stationary. Thus, in the application of transfer functions, one must at first evaluate if the X_t and Y_t processes are stationary at the same time.

In the general transfer function model

$$Y_t = v(B)X_t + N_t \quad (26)$$

One may assume that the X_t entry series follows an ARIMA process

$$\Phi(L)Y_t^d = \Theta(L)e_{t,t}$$

where a_t is a white noise, given by

$$\alpha_t = \frac{\varphi_x(B)}{\theta_x(B)} X_t \quad (27)$$

This series is frequently called prewhitened entry series. Applying the same prewhitening transfer in the response series Y_t , one may obtain the filtered response series

$$\beta_t = \frac{\varphi_x(B)}{\theta_x(B)} Y_t \quad (28)$$

As $e_t = \theta_x^{-1}(B)\varphi_x(B)N_t$ the transfer function model—changing (27) and (28) into (29), we have got

$$\beta_t = v(B)\alpha_t + e_t \quad (29)$$

The impulse response weights v_k may, consequently, be found out as

$$v_k = \frac{\sigma_\beta}{\sigma_\alpha} \rho_{\alpha\beta}(k) \quad (30)$$

Thus, the transfer function $v(B)$ is obtained according to the following steps:

Prewhitening time for entry series

$$\varphi_x(B)X_t = \theta_x(B)\alpha_t \quad (31)$$

in other words

$$\alpha_t = \frac{\varphi_x(B)}{\theta_x(B)} X_t \quad (32)$$

where a_t is a white noise series, which represents the random parts of all data, with zero average and variance σ_a^2 .

Calculating the filtered exit series

In other words, transforming the exit series Y_t using a prewhitened model to generate the series.

$$\beta_t = \frac{\varphi_x(B)}{\theta_x(B)} Y_t \quad (33)$$

Calculating the sampling CCF $\hat{\rho}_{\alpha\beta}(k)$ between α and β to estimate v_k .

Identifying b , $\delta_r(B) = (1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r)$ and $w_s(B) = (w_0 - w_1 B - \dots - w_s B^s)$ by the agreement of the standard of \hat{v}_k . Once b , r , and s are chosen, preliminary estimates \hat{w}_k and $\hat{\delta}_j$ may be found out from their relation to \hat{v}_k as it is shown in Equation (21). So, a preliminary estimate of the transfer function \hat{v}_k will be

$$\hat{v}(B) = \frac{\hat{w}_s(B)B^b}{\hat{\delta}_r} (B) \quad (34)$$

Once the preliminary model of the transfer function is obtained, one may calculate the series of noise estimated

$$\hat{N}_t = Y_t - \hat{v}(B)X_t \quad (35)$$

$$Y_t - \frac{\hat{w}_s(B)}{\delta_r(B)} B^b X_t \quad (36)$$

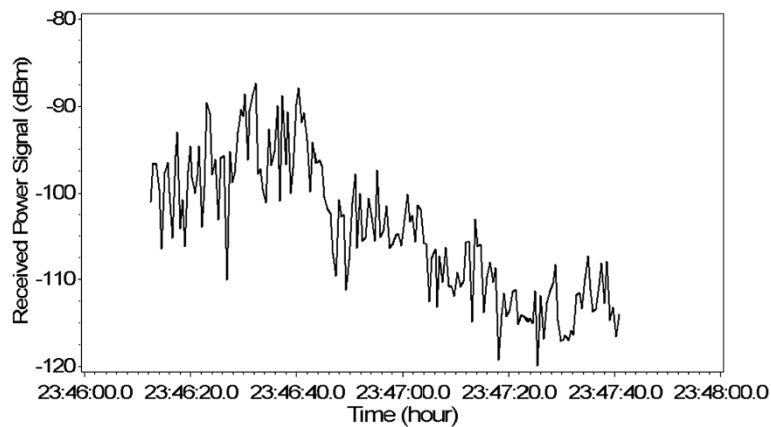


Figure 2 Received power signal (dBm) along The Gentil Bittencourt Avenue.

The appropriate model for the noise may be identified by examining its ACF and PACF or by another identification tool of time univariate series. So

$$\varphi(B)n_t = \theta(B)a_t \tag{37}$$

Combining (34) and (35) we have got the following model of transfer function

$$Y_t = \frac{w(B)}{\delta(B)}X_{t-b} + \frac{\theta(B)}{\varphi(B)}a_t \tag{38}$$

After the identification of a model of preliminary transfer function, as shown in Equation (24), it is necessary to estimate the general parameters $\delta = (\delta_1, \dots, \delta_r)'$, $w = (w_0, w_1, \dots, w_{Ls})'$, $\varphi = (\varphi_1, \dots, \varphi_p)'$, $\theta = (\theta_1, \dots, \theta_{Lq})'$, and σ_a^2 . In general, such estimates are carried out interactively with the help of computer programmes.

Before the model may be used for prediction, control, or other purposes, it is necessary to test the appropriateness of it. In the transfer function model, it is assumed that a_t are independent, white noises of the entry series

X_t and, so, they are also independent from the prewhitened entry series \hat{a}_t . Thus, in the diagnosis test of a transfer function model, one must analyze the residuals \hat{a}_t of the noise model as well as the residuals a_t of the prewhitened entry model so as to verify if the considerations keep the same, that is to say, if the residuals are independent. In this case, the cross correlation test is carried out, in which for a proper mode the sampling CCF $\hat{\rho}_{\hat{a}a}(k)$, between \hat{a}_t and a_t , must not show standards and must be within its two standard-errors, $2(n-k)^{-1/2}$; in other words, the noise series a_t and the X_t entry series must be independent.

Analysis and discussion of the results

Most time series analyses and forecasting procedures assume that observations of the time series use equally spaced time or distance intervals and that processes should be stationary.

Figures 2, 3, 4, and 5 present the series of received power signal (dBm) (Y_t), distance between the transmitter

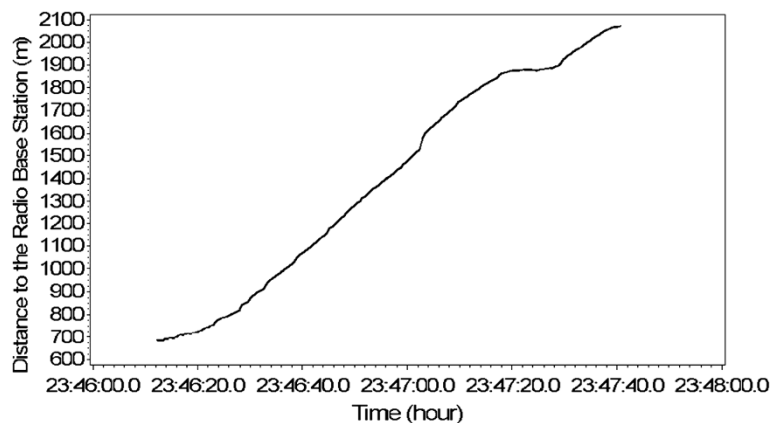


Figure 3 Distance to the RBS (m).

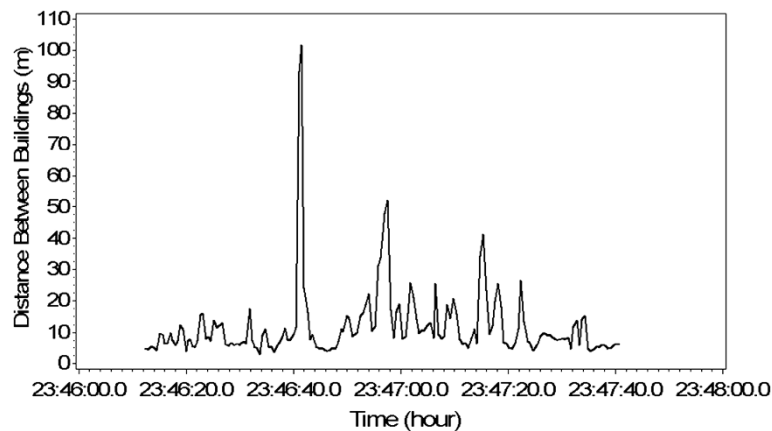


Figure 4 Distance between buildings (m).

and receiver, (X_{1t}), distance between buildings (X_{2t}), and height of buildings (X_{3t}), respectively, collected along the Gentil Bittencourt Avenue. The variable of received power signal (dBm) was used as the dependent variable (Y_t). Selection of the received power signal data as the dependent variable was made in order to provide information on the influence of varying distance between the transmitter and receiver (X_{1t}), distance between buildings (X_{2t}) and height of buildings (X_{3t}), considered in this case as independent variables.

In Figures 2 and 3, a trend was observed. There is an indication of a negative correlation between the series of received power signal (dBm) (Y_t) and the distance between the transmitter and receiver (X_{1t}), i.e., a decrease in received power signal (dBm) (Y_t) is accompanied by an increase in the distance between the transmitter and receiver (X_{1t}). To make the series stationary, differentiations were performed.

First, in the construction of a transfer function model, a set of input data must be pre-adjusted. In the pre-

adjustment of the data series, the response series and the cross-correlation analysis, the following procedures were used: (i) adjustment of an ARIMA model to the input series so that the residual of the model is white noise, (ii) filtration of the response series with the same model used in the input series, (iii) realization of a cross correlation between the filtered response series and the filtered input series to determine the relationship between the two; and (iv) interpretation of the cross-correlation graph. Autoregressive indicators suggest terms in the denominator, and indicators of moving averages suggest terms in the numerator.

Overall, several models are adjusted until they find one that fits more properly to the data series. The models are to be adjusted based on the analysis of the ACFs and PACF.

In Figures 2 and 3, one can observe the presence of tendency. There is an evidence of a negative correlation between the grades received signal strength (dBm) (Y_t) and distance between the transmitter and receiver, (X_{1t}),

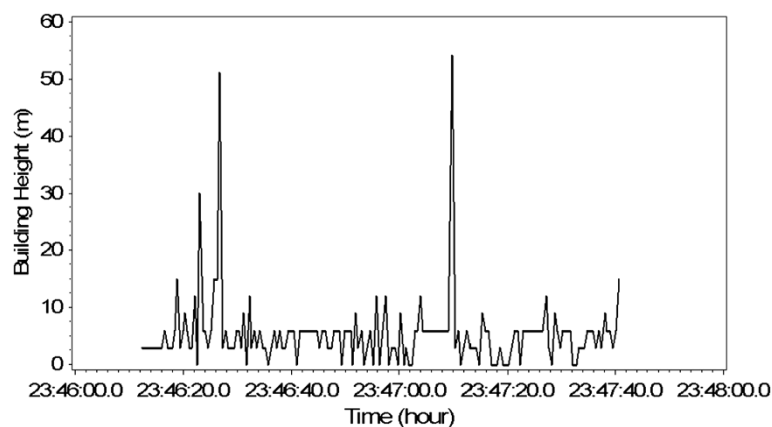


Figure 5 Height of the buildings (m).

Table 1 ARIMA models adjusted for the input series

Series (variable)	Adjusted model	Model
Received signal strength (dBm)	$Y_t = Y_{t-1} - 0.534_{(p<0.0001)} a_{t-2} + a_{3t}$	ARIMA(0,1,1)
Distance to the RBS (m)	$(1 - 0.967_{(p<0.0001)} B)(1 - B)X_{1t} = (1 - 0.309_{(p<0.0001)} B - 0.420_{(p<0.0001)} B^2)a_{1t}$	ARIMA(1,1,2)
Distance between Buildings (m)	$(1 - 0.478_{(p<0.0001)} B)(1 - B)X_{2t} = (1 + 0.916_{(p<0.0001)} B + 0.551_{(p<0.0001)} B^2)a_{2t}$	ARIMA(1,1,2)
Height of the buildings (m)	$X_{3t} = 5.382_{(p<0.0001)} + 0.893_{(p<0.0001)} a_{3t-1} + a_{3t}$	ARIMA(0,0,1)

where t is the time index, Y_t , X_{1t} , X_{2t} , and X_{3t} are the variables, a_1 , a_{1t} , a_{2t} , and a_{3t} are random errors, P , is p -value.

or, a decrease in the received signal strength (dBm) (Y_t) is accompanied by an increase in the distance between transmitter and receiver, (X_{1t}). To make the series stationary same differentiations were performed. Observe that the series (X_{2t}) (Figure 4) was also differentiated.

For the analysis and interpretation of data and calculation of parameters of the model, a computer program using the SAS software was designed [21], through which the subroutine of proc ARIMA made the adjustment of ARIMA models. This adjustment, which is accomplished in an iterative manner, consists of three steps. The first is the identification of the model, where the observed series is transformed into a series stationary. The second step is to estimate the model, in which the orders p and q are selected and the corresponding parameters are estimated. The third step is to forecast, in which the estimated model is used to predict future values of the time series considered.

Table 1 presents the models adjusted for the input series (p indicates the significance of the estimate). The autocorrelations of the residues until lag 12 are presented in Table 2. It may be noted that the descriptive levels indicate no autocorrelation, i.e., a good fit of the model.

The correlation of influence of the explanatory variables is verified. X_{1t} , X_{2t} , and X_{3t} the variable response Y_t , were constructed and analyzed the functions of cross-autocorrelations between these variables. Figures 6, 7, and 8 show the graphs of the cross-correlation between the response series Y_t and the filtered series X_{1t} , X_{2t} , and X_{3t} .

The results for the cross correlations between the received power signal and the distance to the RBS are obtained. Figure 6 shows that none of the lags presented significant spikes. Therefore, it was concluded that, for

Table 2 Residual autocorrelations for the entry series

Series	χ^2	$P_r > \chi^2$	Cross correlations					
X_{1t}	1.02	0.7972	0.017	0.002	-0.059	0.000	-0.011	0.000
	1.91	0.9928	0.016	0.008	-0.020	0.047	-0.015	-0.013
X_{2t}	0.14	0.9865	0.004	-0.002	0.002	0.004	0.010	-0.020
	0.53	1.0000	0.020	0.007	0.025	-0.010	-0.011	0.013
X_{3t}	4.65	0.4597	0.033	0.066	0.102	0.018	-0.043	0.002
	16.58	0.1208	0.026	0.084	0.000	0.151	0.091	-0.077

the analyzed sample, the correlation between Y and X_{1t} showed no significant correlation, being characterized as white noise. Thus, the X_{1t} variable was not included in the model.

The results coming from the cross correlations between Y_t and X_{2t} (Figure 7) show the following features.

- There is a significant negative lag spike. This is a case with a model with feedback.
- There is a lag spike in lag 15, which indicates that in this model there discrepancy.
- After the first spike, there is only another lag spike at lag 3, indicating a parameter in the numerator of the transfer function.
- After the first spike, the graph decays exponentially with apparently sinusoidal pattern to zero, indicating the possibility of two or more parameters in the denominator for the model.

The results coming from the cross correlations between Y_t and X_{3t} (Figure 8) show the following features:

- No positive lag has significant spike.
- There are spikes in the negative lags -14 , -6 e -5 (model with feedback), which indicates that in this model there is a lag.
- After the first spike, it indicates the possibility of the numerator of the model parameters. The cross-correlations exhibit an exponential decay in a sinusoidal pattern. This indicates the possibility of two or more parameters in the denominator transfer function model.

The transfer function model adjusted for potency (Y_t) has included explanatory variables X_{2t} (distance between buildings) and X_{3t} (height of buildings) and according to the analyzes of the cross-correlations the following model transfer function was specified:

- For the variable X_{2t} , a discrepancy of two lags and a parameter in the denominator is specified.
- For the variable X_{3t} , a discrepancy of two lags and two parameters in the denominator is specified.
- After the identification and estimation of transfer function model an appropriate process modeling

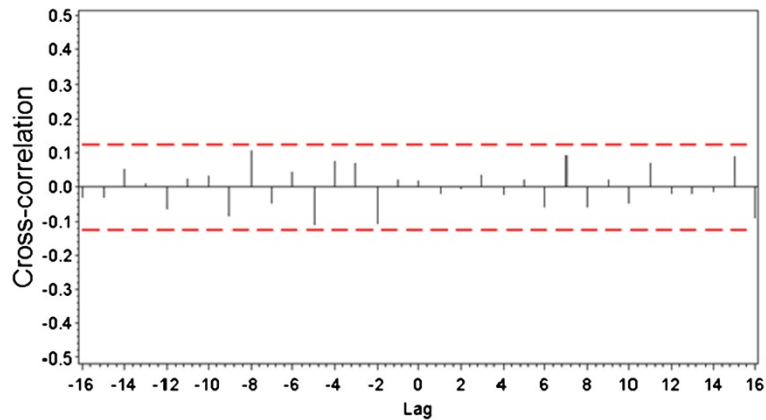


Figure 6 Cross correlation between the received power signal and the distance to the radio base station.

(waste) to the model was necessary to estimate appropriate (model) error.

In this method, we found the following model transfer function:

$$\begin{aligned}
 Y_t = & \frac{w_{02}}{(1 - \delta_3 B^3)} X_{2t-2} \\
 & + \frac{w_{03}}{(1 - \delta_1 B - \delta_5 B^5)} X_{3t-2} \\
 & + \frac{e_t}{(1 - \varphi_2 B^2)(1 - \varphi_4 B^4)} \quad (39)
 \end{aligned}$$

The CCF was used to assess the relationships between the input and output variables and define the order and parameters of the transfer function. The estimated parameters of the transfer function and noise model are made iteratively using various language programs SAS [21] through the procedure of proc ARIMA. Based on the information about the cross-correlations, the transfer

function model shown in Equation (40) was applied. It was necessary to make corrections in the model residuals until it was possible to obtain a better adjustment of the model to the data, with the residuals showing white noise behavior.

The residual analysis shown in Table 3, represented by the first 12 autocorrelations of the shown residuals, indicates a good fit to the modeled data. It can be observed that the residual autocorrelations, especially in the early delays, are virtually zero, indicating the presence of white noise, i.e., the model captured the information contained in the signal, leaving out only the random part.

Equation (40) shows the model fitted for the received power signal (dBm) (Y_t), which included the explanatory variables X_{2t} (distance between buildings) and X_{3t} (height of the buildings). It can be observed that both input variables showed a displacement of two lags in relation to the response variable. This may be due to the nature of the input variables; both were

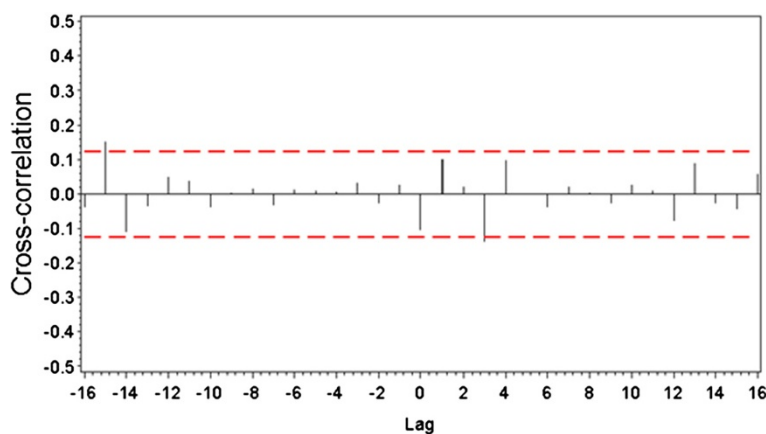


Figure 7 Cross correlation between the received power signal and the distance between buildings.

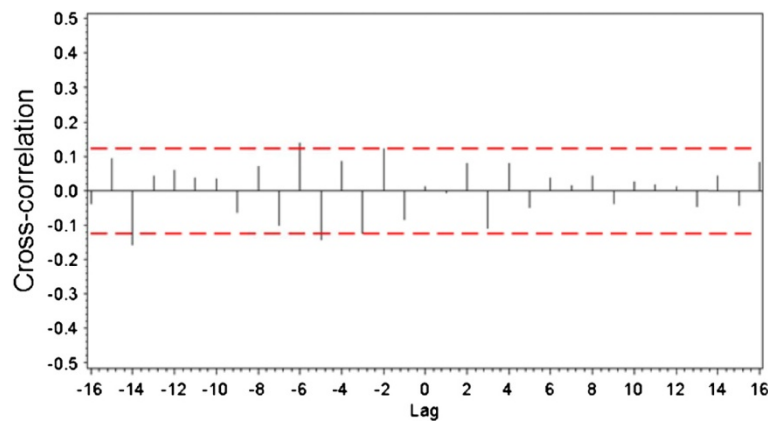


Figure 8 Cross correlation between the received power signal and the building height.

related to the environment where the measurements were taken.

$$\begin{aligned}
 Y_t = & \frac{-0.008}{(1 - 0.766B^3)} X_{2t-2} \\
 & + \frac{0.019}{(1 + 0.558B - 0.586B^5)} X_{3t-2} \\
 & + \frac{e_t}{(1 + 0.512B^2)(1 + 0.290B^4)} \quad (40)
 \end{aligned}$$

The evaluation of the adjustment for the transfer function may be observed by means of the cross correlations among the residuals and the entry variables. Table 4 shows the first 11 correlations (including the zero lag) for the 2 entry variables. One may observe that the cross correlations are statistically non-significant. Although for the variable height of the buildings, in the lags 2 and 9, the correlation has been a little bigger than the one for the other lags, even so, one may observe that the values have been, in general, low. Thus, it is acceptable to consider that the cross correlations of the three entry variables with the residuals from the model have been non-significant, indicating that the transfer function model offers a reasonable adjustment to the data collected. Figure 9 shows the response graph for the model and the observed values for the response variable. The yellow-shaded area indicates 95% confidence intervals. A good fit between the observed signal and the fit signal was obtained; with these data, predictions of the strength of the received signal can be made.

Figure 10 shows the variation of received signal strength (measured and simulated) and theoretical

models of Ibrahim–Parsons [11], Okumura et al. [13] and Walfisch–Ikegami [12] as a function of distance from the transmitter antenna along the Gentil Bittencourt Avenue in the urban area of Belém-PA.

Each propagation model’s approach varies in relation to the classification of the analyzed environment. In the case of locality analysis, simulations of all models were made considering the characteristics of urban environment involved, whereas this environment predominate residences and buildings with an average height varying from 5 to 80 m, respectively.

In order to make a more careful study of performances of each model, statistical analyses of the measures were made for each data file, aiming to measure the deviations between simulated and measured values. With this, information was available to provide subsidy to state what the best model for characterizing the propagation environment for cellular mobile communication of the studied street. Table 5 shows the average and standard deviation for each model in dB compared to the values of received signal power obtained in the field.

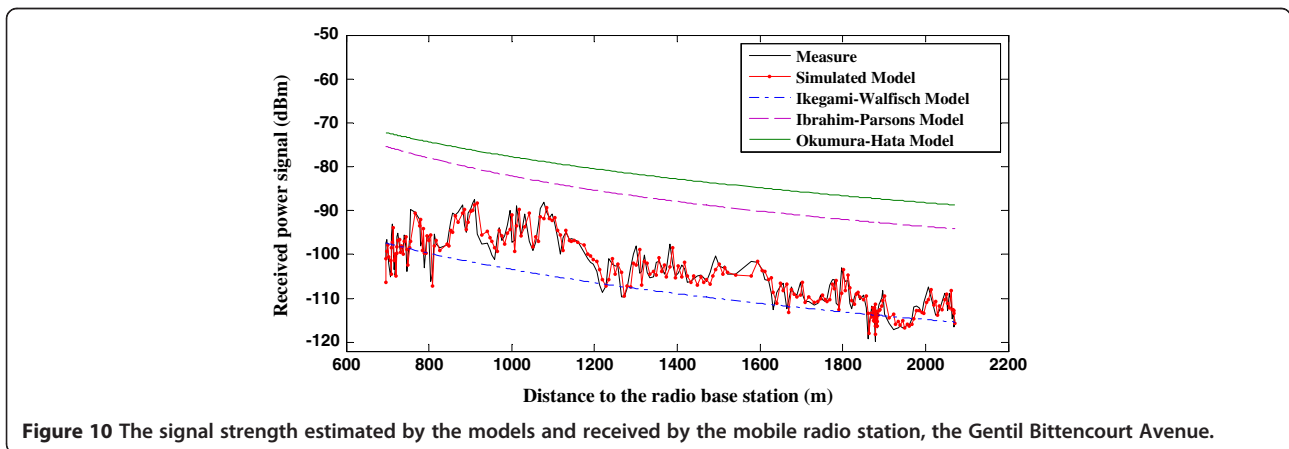
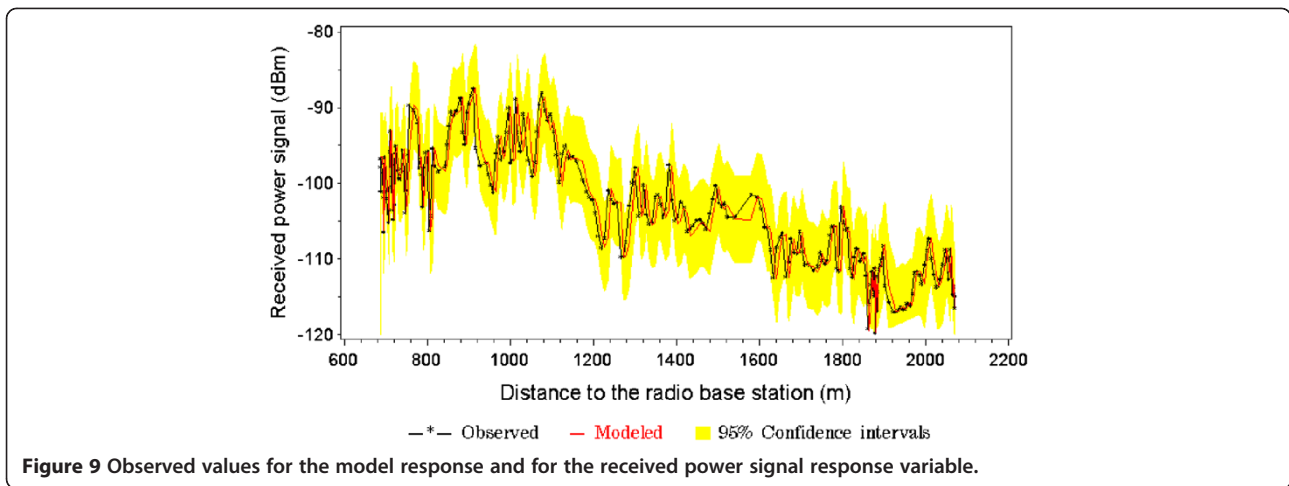
And by the analysis of graphs of the results shown in Table 5, one can deduce that the proposed model is resulting in lower average deviation when compared with field measurements. For this model, the mean square error compared to the theoretical power level is 2.5 dB. In which the maximum acceptable deviation in the signal level that was received by the mobile in

Table 3 Analysis of the residuals for model (40)

Up to delay	$P_r > \chi^2$	Cross correlations						
6	0.5112	-0.034	-0.004	-0.088	-0.025	-0.040	-0.043	
12	0.3712	0.073	-0.074	-0.035	-0.100	-0.083	0.010	

Table 4 Cross-autocorrelations of the residuals of the response series and the input series

Series	χ^2	$P_r > \chi^2$	Cross correlations						
Y_r, X_{2t}	0.14	0.9865	0.004	-0.002	0.002	0.004	0.010	-0.020	
	0.53	1.0000	0.020	0.007	0.025	-0.010	-0.011	0.013	
Y_r, X_{3t}	4.65	0.4597	0.033	0.066	0.102	0.018	-0.043	0.002	
	16.58	0.1208	0.026	0.084	0.000	0.151	0.091	-0.077	



relation to the prediction is 8 dB [7]. Note that the average and standard deviation calculated for the proposed model showed values very close to those obtained for the data collected in the field (measured values).

However, from the theoretical models, nearest measured value is Walfisch–Ikegami model, with a mean square error of 8.34 dB. Note that this model could also be used in the prediction of the spread in the analyzed location. Ibrahim–Parsons and Okumura et al.’s models, which showed the worst results, should undergo their

adjustment coefficients. Possibly the analyzed urban environment for obtaining these models does not present many similarities with that found in the studied region in this study.

Conclusions

In this article, it was shown that how the fitting of signal strength data series used in multivariate time series may present advantages compared with other signal propagation models. The mean squared error obtained with this methodology was on the order of 2.58 dB, much better than the traditional models used. Moreover, this model makes it possible to verify the relationship between the response series and the other series. It thus shows the need to use samples containing information on more streets in order to verify whether there are significant correlations between the received signal and the distance to the base station.

Abbreviations

ACF: Autocorrelation function; AR: Autoregressive; ARIMA(p, d, q): Autoregressive integrated moving average process; χ^2 : Chi-square test; C_i: Cycles; CCF: Cross correlation function; E_i: Residuals; EM: Mobile station;

Table 5 Comparison between the mean square error, the standard deviation, and the mean for the four models, the measured value for Avenida Gentil Bittencourt

	Mean squared error (dBm)	Mean (dBm)	Standard deviation (dBm)
Measured	–	–104.2997	7.8006
Proposed model	2.5821	–104.3195	7.6118
Walfisch–Ikegami model	8.3439	–97.3022	5.8662
Ibrahim–Parsons model	28.7999	–75.8739	6.1063
Okumura et al.’s model	34.9879	–69.6186	5.3701

ERB: Radio base; I: Integrated; I_s: Seasonal effects; MA: Moving average; PACF: Partial autocorrelation function; Pr: P-value; TS: Time series; T_i: Trend.

Competing interests

The authors declare that they have no competing interests.

Acknowledgment

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

Author details

¹Department of Mathematics, Federal University of Pará (UFPA), Castanhal, Pará, Brazil. ²Electrical Engineering Department, Federal University of Pará (UFPA), Belém, Pará, Brazil. ³Department of Statistics (ICEN), Federal University of Pará (UFPA), Belém, Pará, Brazil. ⁴Statistics and Mathematics, Federal University of Pará (UFPA), Belém, Pará, Brazil.

Received: 9 May 2012 Accepted: 31 October 2012

Published: 27 November 2012

References

1. MVSN Prasad, Path loss deduced from VHF and UHF measurements over indian subcontinent and model comparison. *IEEE Trans. Broadcast* **52**(3), 290–298 (2006)
2. MD Yacoub, *Foundations of Mobile Radio Engineering* (CRC Press, New York, 1993)
3. Z Chen, A Delis, HL Bertoni, Radio-wave propagation predictions using ray-tracing techniques on a network of workstations (NOW). *J. Parallel Distrib. Comput* **64**, 1127–1156 (2004)
4. V Degli-Esposti, D Guiducci, A de'Marsi, P Azzi, F Fuschini, An advanced field prediction model including diffuse scattering. *IEEE Trans. Antennas Propagat* **52**(7), 1717–1728 (2004)
5. AM Cavalcante, MJ Sousa, CSJ Sales, JCWA Costa, CRL Francês, GPS Cavalcante, *Computational parallelization strategy applied in full 3D ray-tracing*, *International Microwave and Optoelectronic Conference - IMOC'2005* (Brasília, 2005), pp. 1–4
6. MF Iskander, Z Yun, Propagation prediction models for wireless communications systems. *IEEE Trans. Microwave Theory Tech.* **50**, 662–673 (2002)
7. TS Rappaport, *Wireless Communications Principles and Practice* (Prentice Hall (Englewood Cliffs, NJ), 2002)
8. T Sarkar, J Zhong, K Kyungjung, A Medouri, M Salazar- Palma, A Survey of Various Propagation Models for Mobile Communication. *IEEE Antenna Propagat. Mag.* **45**(3), 51–82 (2003)
9. C Perez-Vega, JM Zamanillo, Path loss model for broadcasting applications and outdoor communication systems in the VHF and UHF Bands. *IEEE Trans. Broadcast. (USA)* **48**, 91–96 (2002)
10. M Hata, Empirical formula for propagation loss in land mobile radio service. *IEEE Trans. Veh. Technol.* **VT-29**(3), 317–325 (1980)
11. MF Ibrahim, JD Parsons, Signal Strength Prediction in Built-up Areas. Part 1: Median Signal Strength. *IEE Proc.* **130**(5), 377–384 (1983)
12. AP Garcia, H Ortega, A Navarro, AH Rodriguez, Effect of terrain on electromagnetic propagation in urban environments on the Andean region, using the COST231-Walfisch-Ikegami model and GIS planning tools *Antennas and Propagation*. *IEEE Antenna Propagat. Mag.* **1**, 270–275 (2003)
13. T Okumura, E Ohmori, K Fukuda, Field strength and its variability in VHF and UHF land mobile service. *Rev. Electr. Commun. Lab.* **16**(9–10), 825–873 (1968)
14. PA Morettin, *CMC ToIoi, Análise de Séries Temporais* (Edgard Blücher, São Paulo, 2004)
15. GEP Box, GM Jenkins, GC Reinsel, *Times Series Analysis: Forecasting and Control*, 3rd edn. (Holden-Day, San Francisco, 1994)
16. WS Wei, *Time Series Analysis: Univariate and Multivariate*, 2nd edn. (Addison-Wesley, Reading, MA, 2005)
17. JD Cryer, KS Chan, *Time Series Analysis. With Application in R*, 2nd edn. (Springer, New York, 2008)
18. ZL Dupe, *Prediction Nino3.4 SST Anomaly Using Simple Harmonic Model* (Paper presented at the Second International Conference on Science and Technology for the Assessment of Global Climate Change and Its Impacts, Jakarta, 1999)
19. L Lon-Mu et al., *Time Series Analysis and Forecast*, 2nd edn. (Scientific Computing Associates Corporation, Villa Park, IL, 2006), pp. 1–565
20. D West, S Dellana, J Jarretti, Transfer function modeling of process with dynamic inputs. *J. Qual. Technol.* **34**(3), 315–327 (2002)
21. SAS Institute Inc, *SAS/ETS User's Guide, Version 9* (SAS Institute Inc, Cary, 2003)

doi:10.1186/1687-6180-2012-248

Cite this article as: Rozal et al.: Modeling of wireless networks using multivariate time models. *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:248.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com