ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 9(48), DOI: 10.17485/ijst/2016/v9i48/99370, December 2016* ISSN (Online) : 0974-5645

Study of Models to Forecast the Radio-electric Spectrum Occupancy

Luis F. Pedraza[1*], Cesar A. Hernandez[1] and E. Rodriguez-Colina[2]

[1]Universidad Distrital Francisco José de Caldas, Bogota, Colombia. and Universidad Nacional de Colombia, Industrial and Systems Engineering Department, Bogota, Colombia; lfpedrazam@udistrital.edu.co, cahernandezs@udistrital.edu.co, lufpedrazama@unal.edu.co
[2]Universidad Autónoma Metropolitana Iztapalapa, Electrical Engineering Department, Mexico City, 14387 Ciudad de, Mexico; erod@xanum.uam.mx.

Abstract

Objectives: The analysis of spatial opportunities to reuse frequencies by secondary users (SU) in a Cognitive Radio (CR) network is the main objective of this work. **Methods/Statistical Analysis:** Here lineal and no-lineal models are developed and evaluated to forecast the received power of different channels base on measurements performed in Bogota Colombia for the global system for mobile communication (GSM) bands. Seasonal autoregressive integrated moving average (SARIMA), generalized autoregressive conditional heteroskedastic (GARCH), Markov, empirical mode decomposition-support vector regression (EMD-SVR) and wavelet neural models were utilized for the forecasting of the channel occupancy. **Findings:** The analysis performed shows that the wavelet neural model presents less error for the received power forecast compared to the other models analyzed and developed in this work. This is in particular for the different types of mean absolute error evaluated. In addition, the accuracy percentage reach by the wavelet neural model is greater than 95% for the forecast of the availability and occupancy times of channels. **Application/Improvements:** Accordingly, this research indicates that wavelet neural model depicts a hopefuloption to CR systems, in order to forecast received power for the detection of spectral opportunities.

Keywords: Forecast Models, Primary User, Radio-electric Spectrum, Secondary User, Time Series

1. Introduction

Radio-electric spectrum occupancy is widely studied due to its importance for the construction of new spectrum assigning policies in emerging technologies, as well as in monitoring activities in both licensed and unlicensed bands. Real measurements for spectrum use within a determined band, allows the corresponding authorities to guarantee that licenses meet local and regional spectrum regulations[1,2]. On the other hand, precise parameter estimates like time quantity and geographical region where the different spectrum band is actually used bring useful information to determine spectral opportunities for variant technologies within a domain[3]. In this paper such technologies correspond to GSM technology variant in time domain.

The spectrum sensing in CR provides the necessary information about the status of the wireless channels, modeling and prediction of communications activity. This could contribute to spectral efficiency improvement efforts[4-11]. The prediction information of the channel status can be used by SUs to decide the sensing periods and channel occupancy duration for a single channel sensing scenario[12]. Besides, based on prediction information, SUs can select the channels with higher probability of vacancy in multi-channel wideband sensing scenarios[13], and also primary users (PUs) occupancy models can be used as empty channel indicators replacing the spectrum sensing procedures[4,14].

Different initiatives for radio-electric spectrum channel modeling have been proposed[4,15-22]. The difference between those proposals and the present work lies in that, in the first ones, the duty cycle time series is modeled for different types of channels, whereas in our proposal

* *Author for correspondence*

the time series of received power in three GSM channels is modeled by means of different occupancy levels. For this purpose, two linear and three nonlinear models have been employed. The following linear models were used initially: SARIMA model, which is adequate for analyzing time series with seasonality, in accordance with different studies[23-25]. In the second place, GARCH model, which had been applied to traffic modeling and forecast for different communications networks, was also used[26-28].

The three nonlinear models employed were, in the first place, the Markov model, which has been used especially to forecast binary channel occupation[17,18] in wireless networks. For example in[29], a discrete Markov chain was implemented to model duty cycles of channels with different wireless technologies. Second, the EMD-SVR model, combining Support Vector Regression (SVR), which is an appropriate method to forecast non-stationary signals, and Empirical Mode Decomposition (EMD), used to analyze both stationary and non-stationary signals. In[19] the EMD-SVR algorithm showed good results at forecasting the signal of a radar frequency monitoring system. Finally, the wavelet neural model was considered on the basis of the highly accurate results it has proven in forecasting different types of time series[30-34].

The analysis of the results obtained in the forecasts by the models is based on the following variables: availability time of the channel, occupancy time of the channel, observation time and error criteria analysis (symmetrical mean absolute percentage error (SMAPE), mean absolute percentage error (MAPE) and mean absolute error (MAE))[35-37].

The remainder of this paper is then divided into the following sections: section 2 describes the characteristics of the models employed; in section 3 the statistical analysis of the spectrum measures and the models' design is carried out; Section 4 presents comparative results of the models studied; and section 5 shows the conclusions drawn in this paper.

2. Models for Forecasting Spectrum Occupancy

GARCH and SARIMA are deployed in a time series that is assumed as linear and with a known statistical distribution. As it is presented below, this is partially met in long-term analysis of time series measurements. Then, the models that will be described start from the assumption of non-linear time series[38]. This is particularly true for short-term analyses.

2.1 Seasonal Autoregressive Integrated Moving Average Model

In general, if a time series displays potential seasonality indexed by *s*, then using a multiplied seasonal ARIMA (p,d,q) (P,D,Q)s model is beneficial, where *d* is the level of non-seasonal differencing, *p* is the autoregressive (AR) non-seasonal order, *q* is the moving average (MA) non-seasonal order, *P* is the number of seasonal autoregressive terms, *D* is the number of seasonal differences, and *Q* is the number of seasonal moving average terms. The seasonal autoregressive integrated moving average model[39] is given in the equation (1),

$$\Theta_p(B)\Phi_p(B^s)\nabla^d \nabla_s^D x_t = \theta_q(B)\Theta_Q(B^s)e_t \quad (1)$$

where B is the backward shift operator, x_t is the observed time series of load at time t, e_t is the independent, identical, normally distributed error (random shock) at period t; $\nabla_s^D x_t = (1-B^s)^D x_t$, $\Phi_p(B^s)$ and $\Theta_Q(B^s)$ are the seasonal AR(p) and MA(q) operators, respectively, which are defined[39] in equations (2) and (3),

$$\Phi_p(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_p B^{Ps} \quad (2)$$

$$\Theta_Q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \cdots - \Theta_Q B^{Qs} \quad (3)$$

where $\Phi_1, \Phi_2, \ldots, \Phi_p$ are the parameters of the seasonal AR(p) model, and $\Theta_1, \Theta_2, \ldots, \Theta_Q$ are the parameters of the seasonal MA(q)[25].

The Box-Jenkins methodology consists of four iterative steps[40]:

Step 1: Identification. This step focuses on the selection of d, D, p, P, q and Q. The number of order can be identified by observing the sample autocorrelations (ACF) and the sample partial autocorrelations (PACF).

Step 2: Estimation. The historical data is used to estimate the parameters of the tentative model in Step 1.

Step 3: Diagnostic checking. Diagnostic test is used to check the adequacy of the tentative model.

Step 4: Forecasting. The final model in Step 3 is used to forecast the values[40,41].

2.2 Generalized Autoregressive Conditional Heteroskedastic Model

An important number of models, most of which have the property that conditional variance depends on past, have been proposed for capturing special data characteristics. Models commonly used are those with autoregressive conditional heteroskedastic (ARCH) introduced in[42] and generalized ARCH (GARCH) given by[43]. Modeling ARCH-GARCH considers conditional error variance, as a compression function of the past of the series.

ARCH modeling usually requires a great amount of lags q, and therefore a great number of parameters. This might yield a model with a great number of parameters, which is opposed to the parsimony principle. This fact often entails difficulties when using the model to describe data in an adequate way[42]. On the contrary, a GARCH model uses an inferior quantity of parameters, which makes it preferable to an ARCH model[44-46]. In this paper, the GARCH model with order $p \geq 0$ and $q \geq 0$ for the discrete-time stochastic process r_t is expressed in equations (4) and (5)

$$r_t = \mu + \varepsilon_t, \qquad \varepsilon_t \sim N(0,1) \tag{4}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i r_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{5}$$

Where ε_t is an independent and identically distributed process with a zero mean and one standard deviation, μ is the mean constant offset, σ_t^2 is variance, and α_0 is the constant in the conditional variance. Unknown parameters for model are α_0, α_i and β_j for some positive integer p, q.

Just as in an ARIMA model, ACF and PACF are useful for p and q order identification in a GARCH (p,q) process[45].

2.3 Markovmodel

The parameters of the model are unknown and must be settle by taking observable data into consideration. The main idea behind a Hidden Markov Model (HMM) is that the latent state of the system, together with other non-observable information, are hidden as part of an observation process affected by some "noise". This hidden information is assumed to keep track of the dynamics of the finite-state Markov chain in discrete or continuous time[47].

1. Markov chain: Let (Ω, F, P) be a probability space and let $(X_k)_{k\in N}$ be a sequence of random variables in the state-space set $M = \{m_1, m_2 \ldots m_N\}$, where x is the function x: $\Omega \rightarrow M$ and N is the set of natural numbers.

The process x is said to be a Markov chain if it satisfies Markov's property[47].

$$P(x_{(k+1)} = m(x_{(k+1))} \mid x_0 = m_0,\ldots, x_k = m_k) = P(x_{(k+1)} = m(x_{(k+1)}) \mid x_k = m_k) \tag{6}$$

$$\forall k \geq 1 \quad ym_0, m_1, \ldots, m_k \in M$$

The initial distribution of x is defined by $X = (X_m: m \in M)$, $X_m = P(x = m) = P(\{w: x(w) = m\})$. Furthermore, the Markov chain $(X_k)_{k\in N}$ is characterized by its transition probability matrix Π. For a particular element π_{ji} of the transition probability matrix we have[47]:

$$\pi ji = P(x(k+1) = j \mid xk = i),\ i, j \in M \tag{7}$$

2 Hidden Markov model: In a HMM a Markov chain is embedded in a stochastic process, which is a series of observations. The Markov chain itself is not observable; this lies "hidden" in the observation. The aim is to estimate the underlying Markov chain, this is, filter the sequence $\{x_k\}$ out of the observations[47].

3. Change of probability measure: A summary of a change of probability measure technique for the filtering process is shown below. The change of the measure technique has been widely used in filtering applications. It was introduced for the stochastic filtering in[48].

This new "ideal" probability measure is equivalent to the real world measure, which is the measure under which we have the observation process. Changing the real measurement for an ideal one leads to easier ways of calculating filters such as the results of the Fubini-type, which can be used instead of direct calculations that require hard semimartingale methods.

4. Changes of the measure techniques: The theory of the evolution of the measures is based on the equivalence of two probability measures linked through the Radon-Nikodym theorem[49].

5. Adaptive and recursive filters: The aim is to attain adaptive and recursive filters for the Markov chain. Adaptive filters enable coefficients to adjust to the current situation of

the series. This adjustment is achieved with the help of a recursive algorithm within the filter. Consequently, a "self-tuning" model is generated, which adjusts itself to changes in the time series data. In a recursive filter the filter's previous output values are used as input for the calculations.

The number of jumps of a Markov chain from a state r to a state s is considered in the time k defined by [47]

$$J_k^{(sr)} := \sum_{l=1}^{k} \langle x_{l-1}, e_r \rangle \langle x_l, e_s \rangle \quad (8)$$

Where e_r and e_s are unitary vectors.

Secondly, consider the occupation time, given by the length of time x spent in state r up to time k. This is given

$$O_k^{(r)} := \sum_{l=1}^{k} \langle x_{l-1}, e_r \rangle \quad (9)$$

An auxiliary process is also needed to estimate the vectors $\pi, \alpha, \gamma y \xi$. This has the following form[47]:

$$T_k^{(r)}(g) := \sum_{l=1}^{k} \langle x_{l-1}, e_r \rangle g(y_l) \quad (10)$$

where g is a function.

The representation in discrete time of the reception power process is obtained from[47].

$$R_{Y_{k+1}} = \alpha(x_k) R_{Y_k} + \gamma(x_k) + \xi(x_k) w_{k+1} \quad (11)$$

Where $\{x_k\}$ is a discrete time of the Markov chain, and $\{w_k\}$ is the sequence of standard normal random variables, independent and identically distributed. For this discrete-time version the optimal parameters can be derived by means of filtering techniques[47]. The optimal parameter estimation is obtained through the technique of maximum likelihood estimation (MLE).

6. Expectation maximization algorithm. This algorithm is employed to derive estimations from the optimal parameters to the adjusted model parameter $\hat{\theta} = \{\hat{\pi}_{ji}, \hat{\gamma}_i, \hat{\alpha}_i, \hat{\xi}_i \ 1 \leq i, j \leq n\}$. The Expectation Maximization (EM) algorithm is an iterative procedure to find the MLE in problems with incomplete data, in which calculating MLE may result difficult due to the missing values or the optimization of the likelihood function is analytically intractable[47,50].

Figure 1 shows the HMM algorithm used to forecast the received power in the GSM channels.

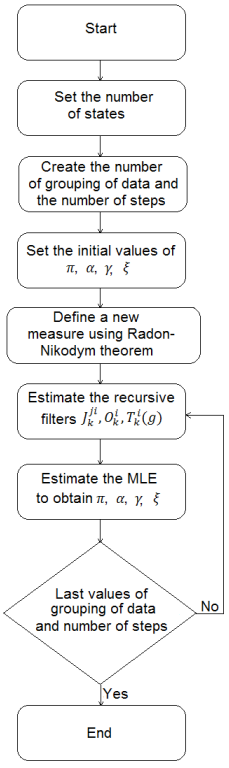

Figure 1. Flow chart of hidden Markov model.

2.4 Empirical Mode Decomposition and Support Vector Regression model

Down below are described the EMD and SVR methods, which were used together to forecast reception power of the GSM channels.

1. Empirical mode decomposition. A principle of EMD is to decompose a signal x(t) into a sum of functions that satisfies two conditions[19,51]:

- The number of extreme and the number of zero crossings must be either equal or differ at most by one.
- The mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

These functions are known as intrinsic mode functions (IMF) and are represented as $imf_i(t)$. IMFs are obtained using the Algorithm 1 shown below[52]:

(1) **Initialize**: $r_0(t)=x(t)$, i=1

(2) Extract the i-th IMF:

(a)Initialize: $h_0(t)= r_{i-1}(t)$, j=1

(b)Extract the local minima and maxima of $h_{j-1}(t)$

(c)Interpolate the local maxima and the local minima by a cubic spline to form upper and lower envelopes of $h_{j-1}(t)$

(d)Calculate the mean $m_{j-1}(t)$ of the lower and upper envelopes

(e)$h_j(t)= h_{j-1}(t)-m_{j-1}(t)$

(f)**If** stopping, the criterion is satisfied, then set $imf_i(t)=h_j(t)$ **else** go to (b) with j=j+1

(3) $r_i(t)= r_{i-1}(t)-imf_i(t)$

(4) **If**$r_i(t)$ still has 2 extreme at least, then go to 2with i=i+1 **else** the decomposition is finished and$r_i(t)$ is the residue.

Algorithm 1. Algorithm for the EMD[52].

At the end of the algorithm we obtain.

$$x(t) = \sum_{i=1}^{n} imf_i(t) + r_n(t) \tag{12}$$

where$r_n(t)$ is the residue of the decomposition, which can be the mean tendency or a constant.

1. Support vector regression. Consider a training data set $\{(x_i, y_i)\}_{i=1}^{N}$, where each $x_i \in R^n$ denotes an input value and an objective value $y_i \in R$. The generic SVR constitutes a linear function[53,54]:

$$f(x) = \langle w, \phi(x) \rangle + b \tag{13}$$

Where $\varphi(\cdot)$ is a nonlinear mapping from R^n to a higher dimensional space called feature space. The regression vector w ($w \subset R^n$) and the bias term b ($b \in R$) provide a solution to the convex optimization problem[54,55]:

$$\min_{w,\, \xi_i, \xi_i^*} L = C \sum_{i=1}^{N} (\xi_i + \xi_i^*) + \frac{1}{2}\|w\|^2 \tag{14}$$

$$\begin{cases} y_i - \langle w, \phi(x_i) \rangle - b \leq \varepsilon + \xi_i \\ \langle w, \phi(x_i) \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

where the parameter ε adjusts the size of the regression approximation error to control the support vector number and the generalization capacity. The larger is the value of ε, the lower is the accuracy. The presence of errors in the data set is measured by means of other internal parameters ξ_i y ξ_i^* called "slack variables", which characterize the deviation of training samples outside the ε-margin[19,55]. In equation (14) C is a constant that determines sanctions to estimation errors. A considerable C assigns significant sanctions to errors, in such a way that the regression is trained to minimize errors with a lower generalization, while a smaller C assigns an inferior number of sanctions to errors[19,53]. In standard SVR the values of C and ε must be specified beforehand.

The previous optimization problem can be easily solved using this double formulation[19,53,55]:

$$\max_{a_i, a_i^*} L = -\frac{1}{2}\sum_{i,j=1}^{N} (a_i^* - a_i)(a_j^* - a_j)\langle \phi(x_i), \phi(x_j) \rangle - \sum_{i=1}^{N} (a_i^*(y_i - \varepsilon) - a_i(y_i + \varepsilon)) \tag{15}$$

Subject to:

$$\sum_{i=1}^{N} (a_i - a_i^*) = 0, \qquad a_i, a_i^* \in [0, C]$$

where the variables α_i y α_i^* are determined by quadratic programming techniques. Then, the solution of vector *w* and the SVR regression function are obtained from the following expressions[53,56]:

$$w = \sum_{i=1}^{N} (a_i - a_i^*)\phi(x_i) \tag{16}$$

$$f(x) = \sum_{i=1}^{N} (a_i - a_i^*)\langle \phi(x_i), \phi(x) \rangle + b \tag{17}$$

In equation (17) the scalar product in the feature space $\langle \varphi(x_i), \varphi(x) \rangle$ can be replaced by a kernel function $k(x_i, x)$. Kernel functions enable the dot product to be performed in high-dimensional feature space using low-dimensional space data input without knowing the transformation φ[19,53]. The most commonly employed kernel function is the Gaussian RBF with a width of σ[19,53].

$$k(x, x_i) = exp\left\{-\frac{|x - x_i|^2}{\sigma^2}\right\} \tag{18}$$

2. Regression support vector in time series forecast. In the modeling and forecast of nonlinear time series, the phase space reconstruction (PSR) is essential[19,57]. In general terms, the dimension of phase space of nonlinear time series is likely to be very high, even infinite, which in most of the cases is not known. So, the information hidden in the time series can be exposed only when time series is expanded to multi-dimensional space[19,58]. Therefore, PSR permits to make forecasts, short term, of forward behavior of a time series, using information based only in previous values.

The traditional PSR often adopts a method called Coordinate Delay (CD)[58]. Given a time series$\{x_t\}_{t=1}^{N}$, reconstruct the feature vector[58]:

$$X_t = (x_t, x_{t-1}, \ldots, x_{t-(m-1)}) \tag{19}$$

(m is embedding dimension, delay time sets in 1; $t = m, m+1, \ldots, N-1$) modeling time series, consists in finding a function *f*: $R^m \rightarrow R$ between the self-correlative

input X_t and the output Y_t, such that[58]:

$$Y_t = x_{t+1} = f(x_t, x_{t-1}, \cdots, x_{t-(m-1)}) = f(X_t) \quad (20)$$

When applying SVR to process the training data set $\{(X_t, Y_t)\}_{i=m}^{N}$ a time series forecast model can be created in the following way[59]:

$$Y_N = \hat{x}_{N+1} = \sum_{i=1}^{N-m} (a_i - a_i^*) K(X_N, x_i) + b \quad (21)$$

Where $X_N=(x_N, x_{N-1}, \ldots, x_{N-(m-1)})$

3. Forecast model EMD-SVR. The EMD-SVR model, as it can be seen in Figure 2, uses the EMD algorithm to decompose data series $\{x_1,\ldots..,x_l\}$ in a finite set of IMFs. Then, the forecasts on these IMFs are made using SVR model to obtain the forecast value $\widetilde{imf}_i(l+1)$; and finally, according to equation (12), the forecast value $\hat{x}(l+1)$ can be found by the sum of previously forecast results[13,19].

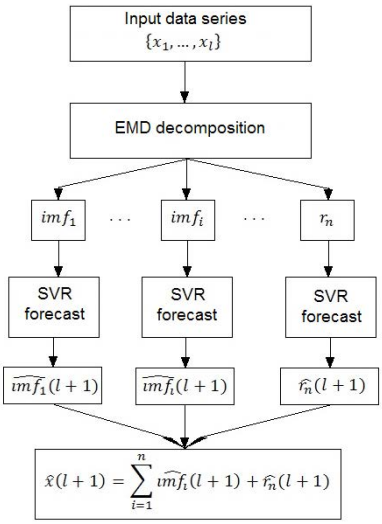

Figure 2. EMD-SVR flow chart[19].

Through EMD the different information features of raw data can be displayed in several scales, reason for which this method may better capture the local fluctuations of this type of data. What is more, having each IMF similar frequency characteristics, simpler frequency components and strong regularity, EMD has the capacity to reduce the complexity of SVR modeling and improving SVR forecast and accuracy[13,19].

2.5 Wavelet Neural Model

In[60] the idea of "wavelet networks" is proposed, the theory combines both wavelets and neural nets.

4. Wavelet. The wavelets are a kind of functions used to locate a given function in the position and in the scale. Wavelets are the base of the wavelet transform, which "divides data from functions or operators into different frequency components, and then each component is studied with a resolution e to its scale"[60-62]. A wavelet is a "small wave" function, usually noted as $\psi(\cdot)$. A small wave grows and decreases in a finite period of time opposed to a "large wave", such as the sine wave, which rises and decreases several times within an infinite period of time[62]. The function $\psi(\cdot)$ is generally referred to as mother wavelet. A wavelets family can be generated from translation and expansion of this mother wavelet[31].

Discrete Wavelet Transform (DWT) uses mother wavelets such as *Haar, Daubechies, Coefiman*, among others. Through DWT, a signal in different frequency bands with different resolutions to decompose the signal into high scale is analyzed, which are low-frequency components also called approximate coefficients; and low scale, which are high-frequency components called detailed coefficients. This way, wavelet transform is an implementation of a filter bank that decomposes a signal into multiple signals[63,64]. Wavelet coefficients may be expressed as[65]:

$$W_\varphi[j_0, k] = \frac{1}{\sqrt{M}} \sum f[n] \varphi_{j_0,k}[n] \quad (22)$$

$$W_\psi[j, k] = \frac{1}{\sqrt{M}} \sum_n f[n] \psi_{j,k}[n] j \geq j_0 \quad (23)$$

Where f[n] is the sample projection within time domain, $\varphi_{jo,k}$ is the scale function and $\psi_{j,k}$ is the translation function, which are discrete functions defined between [0,M-1], for a total of M dots. Coefficients in equation (22) represent the approximate coefficients, whereas in equation (23) they represent the detailed coefficients.

5. Neural network. The neural network has a series of external inputs and outputs that take or supply information to the surrounding environment. Inter-neural connections receive the name of synapses which have associated weights. These weights are used to store the knowledge received from the environment. Learning is achieved by adjusting these weights according to a learning algorithm. It is also possible for neurons to

evolve by modifying their own topology, which occurs by the fact that neurons may die and new synapses may grow[31,66].

Generally, a number of input/destinations are required in order to train a network. One neuron receives numerical information through a number of input nodes, internally processes it, and generates a response. Processing is usually carried out in two stages: firstly, input values are linearly combined; and secondly, the result is used as an argument of a nonlinear activation function. The combination uses the weights attributed to each connection, and a constant term. Figure 3 shows a scheme used to represent a neuron[64].

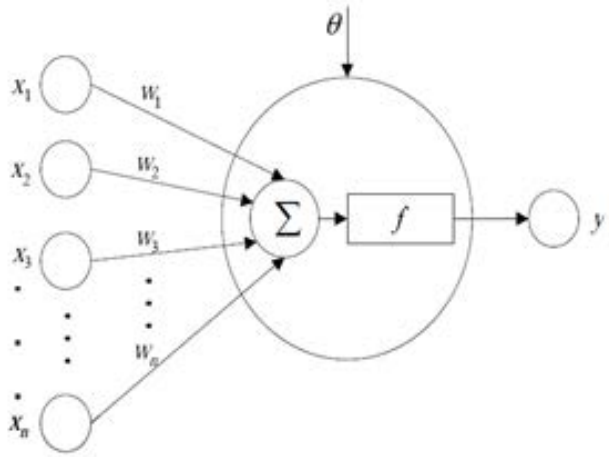

Figure 3. Model of a neuron[64].

In Figure 3, the output of the neuron is given by[64]:

$$y = f\left[\left(\sum_{i=1}^{n} w_i x_i - \theta\right)\right], i = 1,2,3 \ldots n \quad (24)$$

where x_i is the neuron input, w_i is the weight, θ is the offset and f is the activation function[64].

6. Wavelet neural network. Wavelet neural networks combine the wavelets theory with the neural networks. In the model proposed in this paper, the processes corresponding to wavelet and the neural networks are carried out separately. The input signal is decomposed using a mother wavelet, then the coefficients are sent to the input of the multilayer back propagation neural network, and finally the output of the neural network is reconstructed by means of the wavelet analysis for obtaining the power forecast of the GSM channels. In Figure 4, the wavelet neural network model developed is presented.

The neural model employed in this research makes use of a multilayer back propagation neural network, whose output error is propagated backwards to adjust the weights that involve the error minimization. The back propagation networks learn with the descending gradient method, which defines the way to train the output nodes in a multilayer network[67].

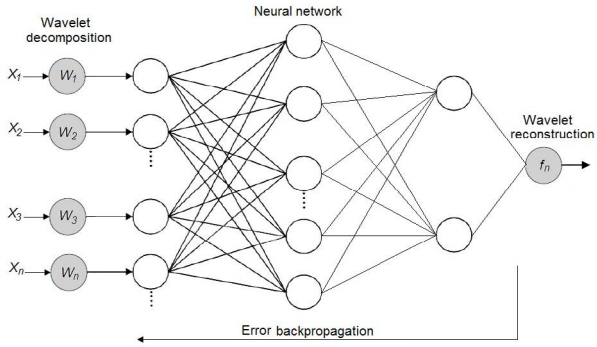

Figure 4. Wavelet neural network.

3. Case Study and Experiment Procedure

The decision to undertake this study was made during the spectrum measurements campaign held in Bogota-Colombia where we obtained the measurements employed in the present work from a spectrum occupancy study previously carried out[1,3]. The band analyzed was the GSM 850 MHz, as it is a band constantly used and viable for analysis in time function with conventional equipment, like a spectrum analyzer. Measurements used in this study correspond to a week, from December 23 to 29, 2012. In some studies[68], it has been indicated that a reasonable option to obtain representative data without any a priori information about a band is to consider measurement periods of at least 24 hours in order to avoid under or overestimating frequency bands occupancy with some temporary patterns. Although a 24 hours measurement period could be thought as sufficient in order to properly characterize the activity of determined spectrum bands[69], in this research 7 days were analyzed, including patterns for workdays and weekends. Additionally, this time period is sufficient to measure occupancy in mobile networks with low use, as indicated in[69]. Thus, the data gathered

along the 7 days were used to train the models previously presented and to particularly forecast the reception power data for Friday from 5 pm to 6 pm. This period was chosen since during this time an increased use of the channels by PUs was perceived.

The channels to be modeled were selected after measuring the duty cycles of 60 channels at GSM band. From these, three channels with different occupancy levels (high, medium and low), were chosen. Figure 5 presents results of power measure for three downlink channels during a week. Spectrum analyzer configuration for this band was the following: a resolution bandwidth of 100kHz with a sweep time of 290ms, which guarantees GSM signal detection with a bandwidth of 200kHz. Daily duty cycles from Pus in selected channels are shown in Figure 6. The threshold (λ) used, which for this event is of -89dBm, was obtained from equation (25) with a probability of false alarm (Pfa) of 1%[70]:

$$P_{fa} = \frac{\Gamma\left(m, \frac{\lambda}{2}\right)}{\Gamma(m)} \quad (25)$$

where Γ(.) y Γ(. , .) are complete and incomplete gamma functions, respectively, and m is the product of time times bandwidth.

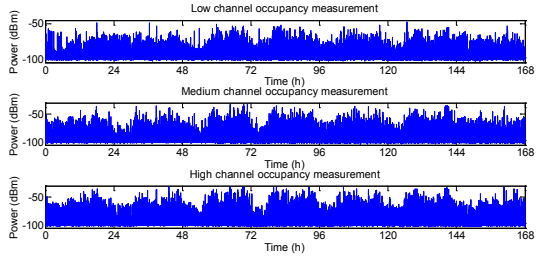

Figure 5. Power measurements for three GSM band downlink channels[41].

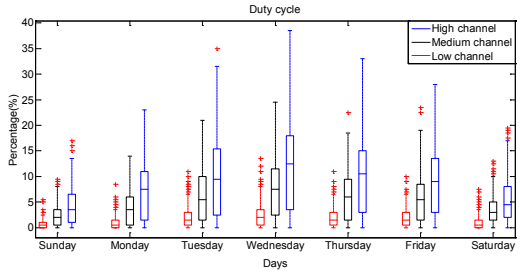

Figure 6. Duty cycles for three GSM band downlink channels[41].

Figures 7, 8 and 9 present histograms corresponding to opportunities distribution during time periods of GSM band channels; it is noted that such opportunities have an exponential behavior, whose approximate equations are exhibited in each figure. When channel occupancy increased occurrence was present especially during shorter time periods. In low, medium and high occupancy channels, total times of opportunities were approximately 84 hours, 81 hours and 78 hours, respectively, which indicates a relatively low occupancy.

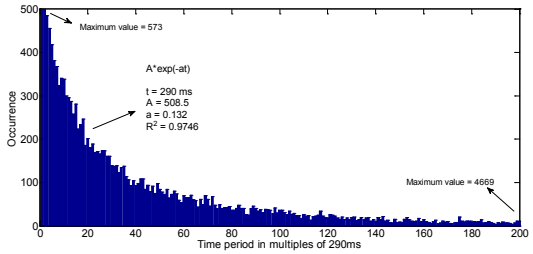

Figure 7. Time period opportunities distribution for low channel.

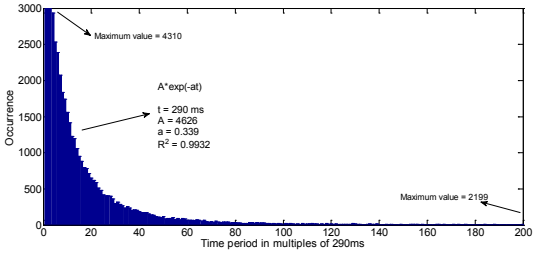

Figure 8. Time period opportunities distribution for medium channel.

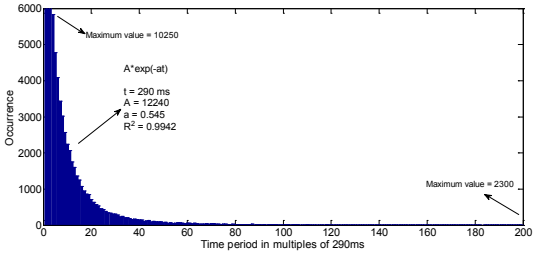

Figure 9. Time period opportunities distribution for high channel.

Following, we proceeded to analyze the time series of measured channels during a week, which were equivalent to 1062514 samples. To do this, ACF is initially presented,

as it is observed in Figure 10. ACF diagrams for the three channels present a form which is alternately positive and negative, decaying to zero, for which they could be, considered as short term correlations[41].

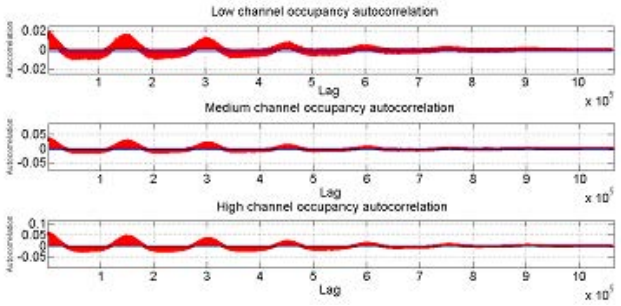

Figure 10. Autocorrelation for three GSM band downlink channels[41].

When analyzing channels stationarity of Figure 10, it is observed that the mean and variance are constant and similar to each other, on each one of the days from Monday to Friday. Therefore measurements on weekend are not taken into consideration when training the analyzed models, because the mean and variance are not similar and change in significant way regarding to measurements from Monday to Friday.

3.1 Design of SARIMA Model

In Figure 11 the trend and seasonality are presented in occupancy level for the three channels. Seasonality had a period of 24 hours, practically without trend and with stationary components, which makes possible the use of a SARIMA model to forecast behavior of the GSM channels[41].

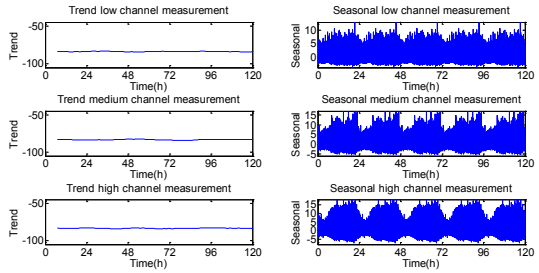

Figure 11. Seasonality and trend components of the GSM channels[41].

Delay difference *s*, which for this event is selected as five (Δ_5), was equivalent to the number of days of the week in which signal was stationary[19]. Applying the augmented Dickey–Fuller test[71], in the series of three channels between Mondays to Friday, the null hypothesis of unit root is rejected, which indicates stationarity. In order to find the parameters of SARIMA (p, d, q) (P, D, Q)s model, ACF and PACF were calculated for Δ_5 of respective channels, as it is shown in Figure 12.

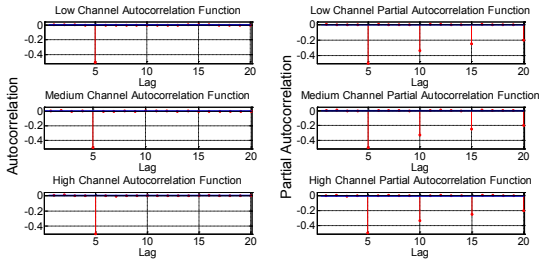

Figure 12. Simple and partial autocorrelation for GSM channels[41].

Using Box-Jenkins methodology[39], Figure 12 shows that PACF of Δ_5 decays to zero with a seasonal pattern and crosses confidence level initially in lag 5 for negative side. This suggests that a term non-seasonal AR(1) could be used, and a seasonal MA(5) could be added.

In order to avoid forecast overestimation (small variance and big errors), the Akaike information criterion (AIC)[72] was selected to evaluate different reasonable combinations, as it is observed in Table 1. Thus, the selected models were: SARIMA$(1,0,5)(1,0,1)_5$, SARIMA$(1,0,5)(0,0,1)_5$ and SARIMA$(1,0,5)(0,0,1)_5$, for occupancy levels of low, medium and high channels, respectively, and the characteristic equations, in the same order are:

$$(1-0.0135B)(1-0.55B^5)(1-B)(1-B^5)x_t = (1-0.997B^5)(1-0.546B^5)e_t \quad (26)$$

$$(1-0.0192B)(1+0.996B^5)(1-B)(1-B^5)x_t = (1+0.0085B^5)e_t \quad (27)$$

$$(1-0.0199B)(1-0.016B^5)(1-B)(1-B^5)x_t = (1-0.994B^5)e_t \quad (28)$$

Table 1. AIC values for different SARIMA models

p	d	q	P	D	Q	Low occupancy channel AIC	Medium occupancy channel AIC	High occupancy channel AIC
1	0	5	0	0	1	-8.24	-30.6	-50.82
1	0	5	1	0	0	-8.3	-32.7	-51.7
1	1	5	0	0	1	-14.1	-46.9	-76.2
1	0	5	1	0	1	-8.19	-32.6	-50.9

3.2 Design of GARCH Model

When analyzing in detail the large amount of acquired information, the existence of standard deviation was observed; therefore the GARCH model was used to forecast the behavior of measured series. Stochastic models ARIMA and SARIMA are methods for univariate modeling. The main difference among former models and GARCH model lies in the constant variance assumption.

Even though for the developed model there is stationarity in original signal from Monday to Friday, for this case the fifth difference is developed because there is a greater degree of stationarity. In Figure 13 the difference for each channel is presented. Channel measurements are converted into returns by logarithmic transformation. The logarithmic returns are defined in equation (29),

$$r_t = \frac{\ln P_t}{P_{t-1}} \quad (29)$$

where, P_t is power value in time t and P_{t-1} is power value in time t-1.

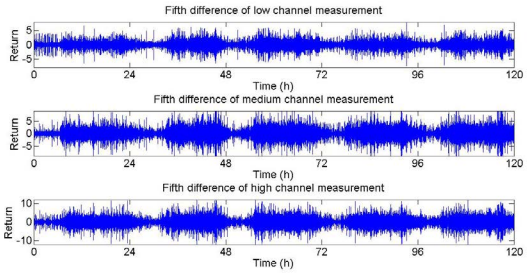

Figure 13. Fifth difference of measured powers in channels of GSM band.

Now we present a formal statistical test in order to found the existence of ARCH effects in the data and correlation. H = 0 implies that there exist no significant correlation as well as H = 1 indicates that there occurs a significant correlation. In Tables 2 and 3, all the *p* values show that *Ljung-Box-Pierce Q-Test* and *Engle ARCH test* in lag 10, 15 and 20 are significant, revealing the existence of ARCH effects (heteroskedasticity), indicating that GARCH modeling is appropriate.

Table 2. Ljung-Box-Pierce Q-Test for autocorrelation: (at 95% confidence) for GSM channels

Lag	H	p value	Low channel statistical test	Medium channel statistical test	High channel statistical test	Critical value
10	1	0	725124	731923	731240	18.3
15	1	0	725136	731956	731266	24.99
20	1	0	725138	731996	731313	31.41

Table 3. Engle ARCH test for heteroskedasticity: (at 95% confidence) for GSM channels

Lag	H	p value	Low channel statistical test	Medium channel statistical test	High channel statistical test	Critical value
10	1	0	574940	578554	576595	18.3
15	1	0	578008	581225	579079	24.99
20	1	0	578710	581829	579500	31.41

Dependence in data $x_1,\ldots,x_n$ was determined by computing correlations. This was done by representing the ACF.

If the time series is the result of a fully random phenomenon, the autocorrelation should be close to zero for all the time-lag separations. Otherwise, one or more of the autocorrelations will be considerably different from zero. Other useful way to examine dependencies of the series is to revise the PACF, where the reliance of intermediate elements (those within the lag) is eliminated[41]. In Figure 14, graphs of ACF, PACF and ACF of square returns present the existence of correlation in data of channels occupancy.

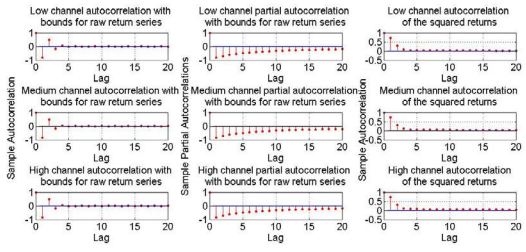

Figure 14. Correlation graphs for GSM band channels.

Table 4. GARCH model comparison for low channel

Model	AIC	BIC	Standard error	Log likelihood	SMAPE	MAPE	MAE
GARCH(0,1)	201838	201873	$7.8x10^{-4}$	96127.5	0.0249	0.0253	2.3606
GARCH(1,1)	192263	192309	$7.82x10^{-4}$	96127.5	0.0249	0.0253	2.3604
GARCH(0,2)	192622	192649	$7.8x10^{-4}$	96127.5	0.0248	0.0252	2.3492
GARCH(1,2)	192265	192299	0.0016	96127.5	0.0244	0.0248	2.3075
GARCH(2,1)	191587	191621	$7.33x10^{-4}$	96127.5	0.0251	0.0255	2.3792
GARCH(2,2)	191581	191622	0.0034	96127.5	0.0243	0.0247	2.3060

Table 5. GARCH model comparison for medium channel

Model	AIC	BIC	Standard error	Log likelihood	SMAPE	MAPE	MAE
GARCH(0,1)	876834	876854	$7.6x10^{-4}$	422041	0.0374	0.0393	3.4198
GARCH(1,1)	844089	844117	$6.6x10^{-4}$	422041	0.0427	0.0440	3.8676
GARCH(0,2)	844984	845012	$6.6x10^{-4}$	422041	0.0375	0.0395	3.4385
GARCH(1,2)	844091	844125	0.0012	422041	0.0411	0.0429	3.7699
GARCH(2,1)	843470	843504	$6.0x10^{-4}$	422041	0.0410	0.0427	3.7531
GARCH(2,2)	843472	843513	$5.0x10^{-4}$	422041	0.0434	0.0452	3.9895

Table 6. GARCH model comparison for high channel

Model	AIC	BIC	Standard error	Log likelihood	SMAPE	MAPE	MAE
GARCH(0,1)	122311	122311	$7.8x10^{-4}$	608609	0.0514	0.0542	4.656
GARCH(1,1)	121722	121725	$6.6x10^{-4}$	608609	0.055	0.058	5.013
GARCH(0,2)	122030	122033	$6.7x10^{-4}$	608609	0.053	0.055	4.795
GARCH(1,2)	121722	121726	$5.3x10^{-4}$	608609	0.056	0.059	5.127
GARCH(2,1)	121430	121434	$6.5x10^{-4}$	608609	0.054	0.057	4.922
GARCH(2,2)	121431	121435	$5.4x10^{-4}$	608609	0.062	0.067	5.939

Below, in Tables 4, 5 and 6, the evaluation and selection of the GARCH model for each channel was performed.

The GARCH model selection for each channel was done by fulfilling $\alpha_i + \beta_i < 1$ criterion, so the model is stationary and then taking into account the more proximate values to zero of MAE, MAPE and SMAPE from Tables 4, 5 and 6. Therefore, the selected models for low, medium and high channel are GARCH(2,2), GARCH(0,2) and GARCH(0,1), respectively.

Parameters for low channel model were estimated and are presented in Table 7. GARCH(2,2), where $\alpha_1 + \alpha_2 + \beta_1 + \beta_2 < 1$ is fulfilled.

Table 7. Parameters estimation for low channel model

Parameter	Estimated value	Standard error	t value
μ	- 96.112	0.0019308	-49778.3308
α_0	0.003516	0.00041447	8.4833
GARCH(1)	0.098255	0.19212	0.5114
GARCH(2)	0.90062	0.19201	4.6905
ARCH(1)	0.00029573	0.00018772	1.5753
ARCH(2)	0	0.00020886	0

Thus, the model according to Table 7 is presented in equations (30) and (31),

$$r_t = -96.112 + \epsilon_t \quad (30)$$

$$\sigma_t^2 = 0.003516 + 0.098255\sigma_{t-1}^2 + 0.90062\sigma_{t-2}^2 + 0.00029573\epsilon_{t-1}^2 \quad (31)$$

For medium channel, GARCH (0,2), model values presented in Table 8 are estimated.

Table 8. Parameters estimation for medium channel model

Parameter	Estimated value	Standard error	t value
μ	-95.061	0.0024331	-39069.8019
α_0	5	0.012924	386.8834
ARCH(1)	0.085692	0.0010392	82.4572
ARCH(2)	0.088298	0.0010582	83.4378

Therefore, equations (32) and (33) are obtained,

$$r_t = -95.061 + \epsilon_t \quad (32)$$

$$\sigma_t^2 = 5 + 0.085692\epsilon_{t-1}^2 + 0.088298\epsilon_{t-2}^2 \quad (33)$$

For high channel, GARCH(0,1), the following parameters were obtained, as shown in Table 9.

Table 9. Parameters estimation for high channel model

Parameter	Estimated value	Standard error	t value
μ	-94.585	0.0026236	-36051.8702
α_0	5	0.015341	325.9324
ARCH(1)	0.86058	0.0044771	192.2169

Then the model is described in equations (34) and (35),

$$r_t = -94.585 + \epsilon_t \quad (34)$$

$$\sigma_t^2 = 5 + 0.86058\epsilon_{t-1}^2 \quad (35)$$

ARCH-GARCH model analysis is based on evaluation of standardized residuals[46]. One assumption with GARCH model is that for an advantageous model, residuals should follow a white noise process, this is, it is expected that residuals be at random, independent and identically distributed, following a normal distribution[41]. Figure 15 presents the relationship among innovations (residuals) derivate from adjusted model, the corresponding conditional standard deviations and returns. Figure 15 show that both innovations and returns exhibit variations. In the following we intend to find out if by performing GARCH the autocorrelation of the standardized innovations disappears, which would indicate the effectiveness of GARCH model.

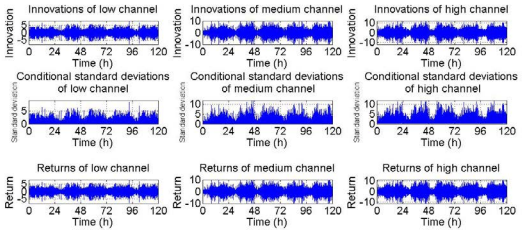

Figure 15. Innovations, conditional standard deviations and returns of GSM channels.

Figure 16 corresponds to the autocorrelation of the squared standardized innovations, in which correlation was not observed.

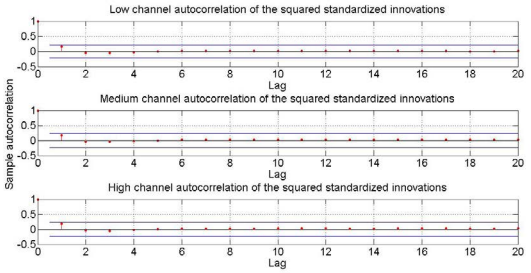

Figure 16. Autocorrelation of the squared standardized innovations of GSM channels.

In Tables 10 and 11, results of *Ljung-Box-Pierce Q-Test*

Table 10. Ljung-Box-Pierce Q-Test in standardized innovations for GSM channels

Lag	H	Low channel p value	Medium channel p value	High channel p value	Low channel statistical test	Medium channel statistical test	High channel statistical test	Critical value
10	0	0.424	0.402	0.701	25787	26701	33455	18.3
15	0	0.7014	0.6883	0.8236	26447	28617	37143	24.99
20	0	0.947	0.876	0.9355	26945	30313	40772	31.41

Table 11. Engle ARCH test in standardized innovations for GSM channels

Lag	H	Low channel p value	Medium channel p value	High channel p value	Low channel statistical test	Medium channel statistical test	High channel statistical test	Critical value
10	0	0.539	0.479	0.6212	26930	27093	33757	18.3
15	0	0.776	0.7144	0.7697	27432	28443	36248	24.99
20	0	0.908	0.863	0.8841	27792	29443	38240	31.41

and *Engle ARCH test* for later analysis are presented using standardized innovations. These tests indicate no correlation presence or ARCH effects. We have GARCH effects and also correlation between innovations that disappear after treating the data. Therefore, the GARCH model is a proper model to explain the variances of the three channels.

Normality verification was performed by analyzing histograms of residuals and normal probability graph, as shown in Figure 17. The histograms of the three channels show that the residuals are normally distributed. In turn, the probability graph confirms that residuals respond to a normal distribution, since most of data are spread along the straight line.

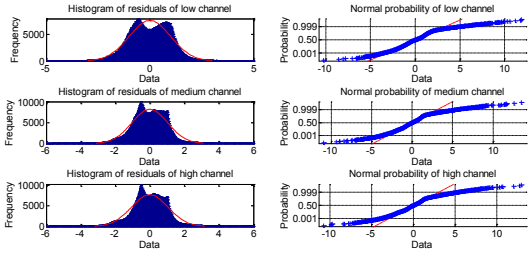

Figure 17. Histogram of residuals and normal probability for GSM channels.

3.3 Design of the Markov Model

The design of this model is based on the flow chart of Figure 1. Since the estimations of the parameters are calculated through the Algorithm 1, initial values are selected for the implementation. These values must be reasonable for the algorithm to obtain the local maxima. Initial values to the algorithm can be found using the method of least squares in the first dots of the data. The estimation of the resulting parameters work as approximations to the initial values of the parameters $\alpha, \gamma y \xi$ which are:

$\alpha = 1.53$, $\gamma = -96.3192$ and $\xi = 3.2551$, $\alpha = 0.09$, $\gamma = -81.8678$ and $\xi = 6.7551$, $\alpha = 0.05$, $\gamma = -94.8265$ and $\xi = 8.7551$,

for the low, medium, and high occupancy channels, respectively. Initial values for the transition probability matrix Π are set in 1/N, where N indicates the number of states in which the implementation is defined. Reception power value to be forecasted can be calculated by[46]:

$$E[y_{k+1}|F_k] = E[a(x_k)y_k + \gamma(x_k) + \xi(x_k)w_{k+1}|F_k] = (\alpha, \Pi\hat{x}_k)y_k + (\gamma, \Pi\hat{x}_k) \quad (36)$$

where $\hat{x}_k = E[x_k|F_k]$. The number of states was chosen according to the minor AIC, in this case it is of three states, compared with the values of two and four states. Figures 18, 19 and 20 depict the evolution of parameters α, γ, ξ and the transition probability after 1440, 1654 and 1879 passes for the channels GSM with low, medium and high occupancy, respectively.

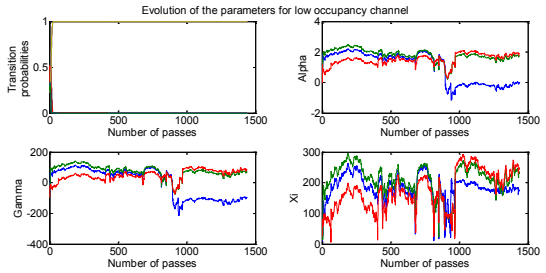

Figure 18. Evolution of the parameters α, γ, ξ and the transition probability of the low occupancy channel.

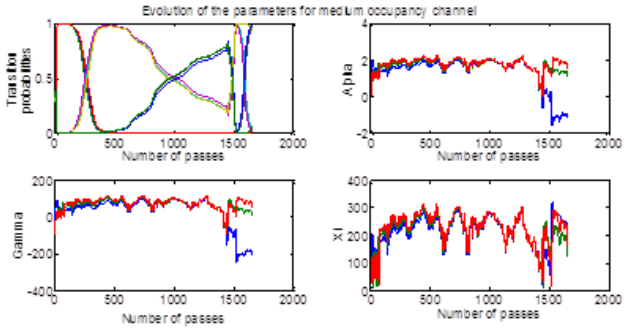

Figure 19. Evolution of the parameters α, γ, ξ and the transition probability of the medium occupancy channel.

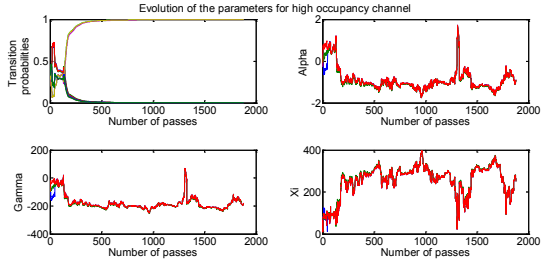

Figure 20. Evolution of the parameters α, γ, ξ and the transition probability of the high occupancy channel.

3.4 Design of the EMD-SVR Model

The processing of the EMD-SVR model turned out to be the more time-consuming of the models presented, reason for which the computational resources used in the simulation were insufficient to process all the input data

(one week); therefore, the model had to be trained with 152000 data, which approximately corresponded to a day of measurements. With these data, the following 6351 values equivalent to Friday from 5 to 6 pm were forecasted and the results were then validated. The procedure carried out for the development of EMD-SVR model displayed in Figure 2 may be summarized in the following steps:

EMD algorithm was executed. In this step 10 data of the time series were obtained (9 IMF plus 1 residue) as shown in Figures 21, 22 and 23:

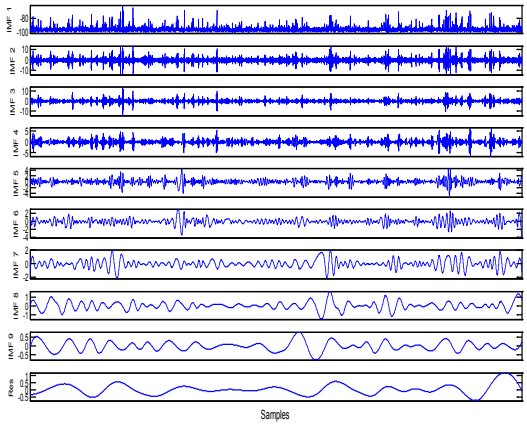

Figure 21. EMD data results for the low occupancy channel.

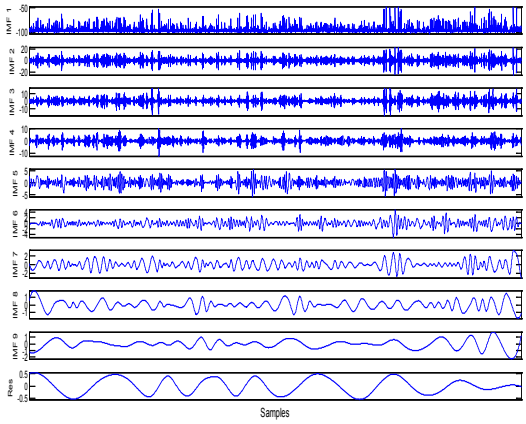

Figure 22. EMD data results for the medium occupancy channel.

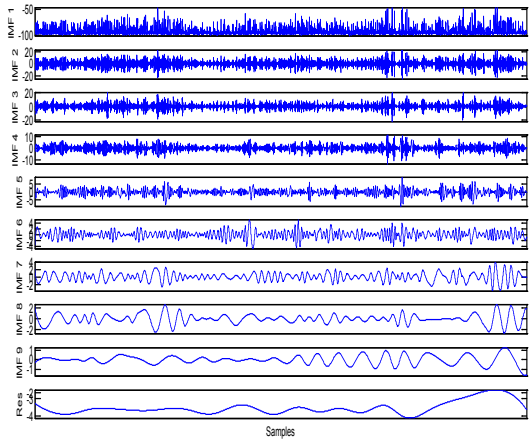

Figure 23. EMD data results for the high occupancy channel.

- A normalized processing of data series was carried out in order to improve the model accuracy.
- Data was divided into two groups. The former group of 152000 data was used as training data set and the latter 6351 data were the test data set.
- The SVR model for each series of the branching based on the training data set was created, and then it was reconstructed and data corresponding to one hour were forecast.

3.5 Design of the Wavelet Neural Model

The input signal to the model, corresponding to the received power from the GSM channels is decomposed using the mother wavelet *Discrete Meyer* (dmey), whose results presented a lower error when compared with mother's wavelet *Daubechies, Coiflets*and*Symlets*[73]. The result was two levels containing in total four coefficients.

The multilayer back propagation neural network used in this study is depicted in Figure 4. It contains two inputs, two outputs and two hidden layers. The network was trained with the former 714952 data from the input signal and the number of training patterns was increased until the error decreased and became relatively constant, condition that was achieved for 1000 training patterns. Finally, the output of the neural network was

reconstructed by means of a wavelet analysis in order to obtain the forecasted power.

4. Results and Discussion

In Figure 24 an example is displayed where there is application of the designed time series models. Here, the user of CR detects the power from a primary Base Station (BS) and can shift through the cell coverage as indicated by the direction of the arrows. The CR user can forecast the power level it will receive from the primary BS.

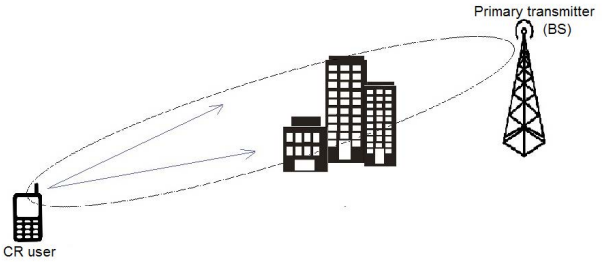

Figure 24. Example of application to forecast the received power from the BS.

In the example of Figure 24, in order to analyze the time series models, the forecast of the power is performed by the CR user, making a comparison with the spectrum analyzer in which the measurements were made. However, this depends on the architecture of CR deployed in the environment. Due to the processor and power consumption being more limited in the CR user's computer, it is recommended to use an infrastructure architecture where the forecast is carried out by the CR BS, which is provided with a better processor than the CR user, and has no limitations on power consumption. However, there is a time period between data capture in the environment and the processing, which adds a delay to the response. This must not be ignored, but the forecast helps to reduce the negative impact of the delayed response.

Figure 25 shows forecasts from the SARIMA model obtained from equations (26), (27) and (28); from the GARCH model based on equations (30), (31), (32), (33), (34) and (35), from the Markov model obtained from Figure 1; from the EMD-SVR model based on Figure 2; and from the wavelet neural model presented in Figure 4. These forecasts were contrasted with the power measured data for Friday from 5 pm to 6 pm. This period was chosen since during this time an increased use of the channels by Pus was perceived.

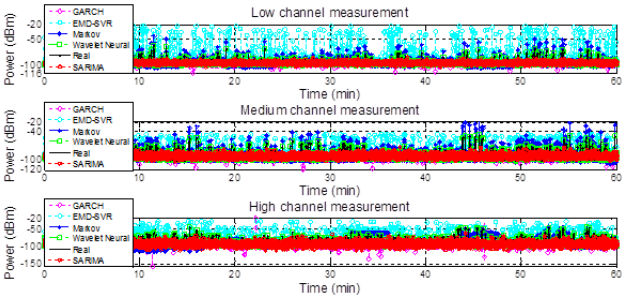

Figure 25. Time series measured and forecasted for low, medium and high GSM channels using SARIMA, GARCH, Markov, EMD-SVR and wavelet neural models.

Availability and occupation times for the channels measured and forecasted are depicted in Figures 26 and 27. Availability time makes possible to analyze the accuracy with which SUs may use the available time of GSM channels in a CR system. Likewise, occupation time helps to examine the accuracy within the time in which PUs make use of GSM channels. Average accuracy obtained between real and forecasted data, for occupancy time as well as for availability time are presented in Table 12. The calculated for the three GSM channels selected.

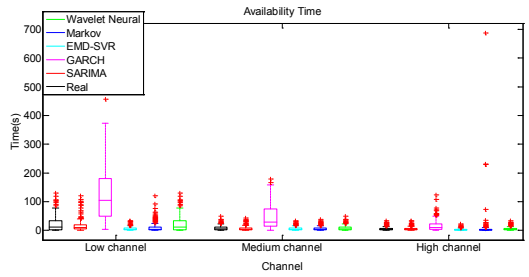

Figure 26. Availability time of forecasted channels.

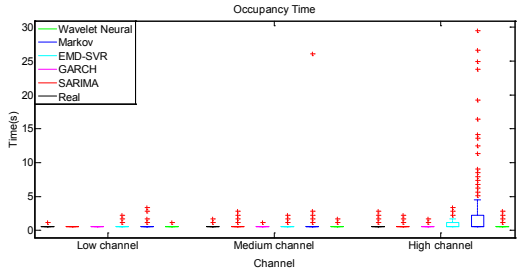

Figure 27. Occupancy time of forecasted channels.

Tables 13, 14 and 15 display the errors between real and forecast data of the models. This analysis included different variables to estimate errors such as SMAPE, MAPE and MAE. Wavelet neural model offers lower error values than the other models in all the criteria analyzed. The reduction of the error with the wavelet neural model becomes from 10 to 80 times with regards to the other models.

Figure 28 shows the evaluation of the performance facing the forecast. This analysis was carried out in a computer with a dual core processor of 2.4GHz and a RAM memory of 4GB. In the best of the cases: the SARIMA model diminishes error up to 7.8% in disadvantage of a 158% in the observation time for the medium occupancy channel. The GARCH model wanes error about 12.8%with an increase in the observation time of 169% for the high occupancy channel. Also, the Markov model reduces error in 27% with an increment in the observation time of 391% for the high occupancy channel. In the same way, in the EMD-SVR model forecast error decreases by as much as 12.1% at the expense of an increment in the observation time of 28% for the low occupancy channel. Then, in the wavelet neural method the total of the error is reduced in 5.45% in detriment of a 47.5% in the observation time for the low occupancy channel. Finally, the wavelet neural method reduces forecast error by as much as 99.6% when compared with Markov and GARCH models, and as much as 99.8% when compared with SARIMA and EMD-SVR models.

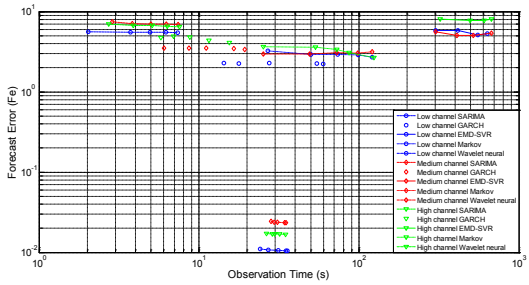

Figure 28. Forecast error vs. Observation time.

Table 12. Accuracy percentage of availability and occupancy times

	Percentage of availability time accuracy			Percentage of occupation time accuracy		
	Low occupancy channel	Medium occupancy channel	High occupancy-channel	Lowo-ccupancy-channel	Medium occupancy-channel	High occupancy-channel
MARKOV model	31	41	32	79	46	60
EMD-SVR model	30	42	44	81	80	62
Wavelet neural model	100	97	99.8	100	95.1	99.9
SARIM Amodel	82	54	60	58	77	78
GARCH model	31	30	43	44	46.6	44.2

Table 13. SMAPE comparison among the forecasted models

Channel-occupancy	EMD-SVR	Markov	Wavelet neural	SARIMA	GARCH
Low	0.0681	0.0231	0.0017	0.0170	0.0243
Medium	0.0654	0.02	0.0020	0.0470	0.0375
High	0.0991	0.1201	0.0019	0.0488	0.0514

Table 14. MAPE comparison among the forecasted models

Channel occupancy	EMD-SVR	Markov	Wavelet neural	SARIMA	GARCH
Low	0.0556	0.0227	0.00089	0.0172	0.0247
Medium	0.0598	0.0189	0.0011	0.0466	0.0395
High	0.0890	0.1117	0.0010	0.0497	0.0542

Table 15. MAE comparison among the forecasted models

Channel occupancy	EMD-SVR	Markov	Wavelet neural	SARIMA	GARCH
Low	5.296	2.1336	0.0866	1.6042	2.306
Medium	5.411	1.6016	0.1	4.2987	3.4385
High	8.022	4.3067	0.1005	4.4195	4.6565

5. Conclusions

Two lineal models (SARIMA and GARCH) and three nonlinear models (Markov, EMD-SVR and wavelet neural) were evaluated in order to forecast the reception power in GSM band channels. The model wavelet neural proved more advantageous in a CR system over the SARIMA, GARCH, EMD-SVR and Markov models, as it presents very high accuracy in availability and occupancy times, with which the use of spectrum efficiency is improved and the interference level and collisions between PUs and SUs will be reduced. In addition, the different types of errors analyzed help demonstrate that the wavelet neural gave the most optimal results. Other important aspect is that even though SARIMA and GARCH models presented lower observation time than wavelet neural model, the forecast error is much lower by using the aforementioned model. As observed, at the best of the cases, observation times of less than 29 seconds for the three GSM channels are achievable, which makes feasible its use in practical CR systems.

For a CR system, the forecast developed in the GSM band could help to improve the use of spectral efficiency, since it would allow CR users to share channels and avoid collisions with PUs in the found opportunities.

The time series models forecast not only the reception power; but the occupation and availability times for GSM channels. It would also be feasible to use the training data from one day for the forecast of a CR user's received power from a primary BS.

The significance of the forecast of received power is that CR users can save energy in the process of detecting the spectrum, take advantage of spectral opportunities, thereby increasing the rate of successful transmission and transmission opportunities, reduce the time to find an available channel, and adjust transmission power levels to protect against collisions and interference with the PUs.

6. Acknowledgments

The work was supported by the Administrative Department of Science, Technology and Innovation (COLCIENCIAS) of Colombia, under the support program 528.

7. References

1. Pedraza L, Forero F, Paez I. Evaluación de Ocupación Del Espectro Radioeléctrico En Bogotá-Colombia, Ingeniería y Ciencia. 2014; 10(19):127-43.
2. Vijayakumar N, Saravana T. Spectral Efficient Cognitive Radio Transmission using Full Duplex Communication, Indian Journal of Science and Technology. 2015; 8(35). DOI: 10.17485/ijst/2015/v8i35/80091.
3. Pedraza L, Forero F, Paez I. Metropolitan Spectrum Survey in Bogota Colombia. IEEE International Conference on Advanced Information Networking and Applications Workshops; Barcelona. 2013.
4. Gorcin A, Celebi H, Qaraqe K, Arslan H. An Autoregressive Approach for Spectrum Occupancy Modeling and Prediction based on Synchronous Measurements. IntSymp on Personal Indoor and Mobile Radio Commun; Toronto. 2011.
5. Pedraza L. Redes Inalámbricas Mesh Caso De Estudio: Ciudad Bolívar 1st Edition, Bogota: Universidad Distrital Francisco José de Caldas; 2012.
6. Pedraza L, Hernandez C, Galeano K, Rodriguez E, Paez I. Ocupación Espectral Y Modelo De Radio Cognitiva Para Bogotá. 1st Edition, Bogota: Universidad Distrital Francisco José de Caldas; 2016.
7. Poongodi K, Singh HK, Kumar D. Co-Operation Based Resource Selection in Cognitive Radio Network via Potential Games, Indian Journal of Science and Technology. 2015; 8(S2). DOI: 10.17485/ijst/2015/v8iS2/58727.
8. Nanthini SB, Hemalatha M, Manivannan D, Devasena L. Attacks in Cognitive Radio Networks (CRN) - A Survey, Indian Journal of Science and Technology. 2014; 7(4). DOI: 10.17485/ijst/2014/v7i4/48646.
9. Patil V, Patil S.A Survey on Spectrum Sensing Algorithms for Cognitive Radio. International Conference on Advances in Human Machine Interaction; Bangalore: India, 2016.
10. Vijayakumar P, Malarvizhi S. Reconfigurable Filter Bank Multicarrier Modulation for Cognitive Radio Spectrum Sharing - A SDR Implementation, Indian Journal of Science and Technology. 2016; 9(6). DOI: 10.17485/ijst/2016/v9i6/80403.
11. Nandhakumar P, Arunkumar. Analysis of OFDM System with Energy Detection Spectrum Sensing, Indian Journal of Science and Technology. 2016; 9(16). DOI: 10.17485/ijst/2016/v9i16/90230.
12. Wellens M, Riihijarvi J, Mahonen P. Empirical Time and Frequency Domain Models of Spectrum use, Physical Communication. 2009; 2(1-2):10-32.
13. Song C, Chen D, Zhang Q. Understand the Predictability of Wireless Spectrum: A Large-Scale Empirical Study. International Conference on Communications; Cape Town. China, 2010.
14. Sun Z, Laneman JN, Bradford GJ. Sequence Detection Algorithms for Dynamic Spectrum Access Networks. IEEE International Symposium on New Frontiers in Dynamic Spectrum; Singapore. 2010.
15. Lopez-Benitez M, Casadevall F. Time-Dimension Models of Spectrum usage for the Analysis, Design, and Simulation of Cognitive Radio Networks, IEEE Transaction on Vehicular Technology. 2013; 62(5):2091-104.

16. Yarkan S, Arslan H. Binary Time Series Approach to Spectrum Prediction for Cognitive Radio. Vehicular Technology Conf; Dublin. 2007.
17. Black T, Kerans B, Kerans A. Implementation of Hidden Markov Model Spectrum Prediction Algorithm. International Symposium on Communication and Information Technologies; Gold Coast. Australia, 2012.
18. Yang L, Dong Y, Zhang H, Zhao H, Shi H, Zhao X. Spectrum usage Prediction based on High-Order Markov Model for Cognitive Radio Networks. Int. Conference on Computer and Information Technology; Bradford. 2010.
19. Yu C, He Y, Quan T. Frequency Spectrum Prediction Method Based on EMD and SVR. Intelligent Systems Design and Applications; Kaohsiung. 2008.
20. Wang Z, SalousS .Time Series Arima Model of Spectrum Occupancy for Cognitive Radio. Seminar on Cognitive Radio and Software Defined Radios: Technologies and Techniques; London. 2008.
21. Chen Y, Oh HS. A Survey of Measurement-based Spectrum Occupancy Modeling for Cognitive Radios, IEEE Communications Surveys and Tutorials. 2014:1-36.
22. Lopez M, Casadevall F. Space-Dimension Models of Spectrum usage for Cognitive Radio Networks, IEEE Transactions on Vehicular Technology. 2016.
23. Shu Y, Yu M, Liu J, Yang O. Wireless Traffic Modeling and Prediction using Seasonal ARIMA Models. IEEE International Conference on Communication Anchorage; China, 2003.
24. Wang W, Niu Z. Time Series Analysis of NASDAQ Composite based on Seasonal ARIMA Model. International Conference on Manage and Service Science; Wuhan. China, 2009.
25. Tran VG, Debusschere V, Bacha S. Hourly Server Workload Forecasting up to 168 Hours Ahead using Seasonal ARIMA Model. IEEE International Conference on Indian Technology; Athens. 2012.
26. Anand NC, Scoglio C, Natarajan B. GARCH - Non-Linear Time Series Model for Traffic Modeling and Prediction. IEEE Network Operations and Management Symposium; Salvador. Brazil, 2008.
27. Tran Q, Ma Z, Li H, Hao L, Trinh Q. A Multiplicative Seasonal ARIMA/GARCH Model in EVN Traffic Prediction, International Journal of Communications, Network and System Sciences. 2015; 8(4):43-9.
28. Zhanga Y, Fayb D, Kilmartina L, Mooreb A. A Garch-based Adaptive Playout Delay Algorithm for VoIP, Computer Networks. 2010; 54(17):3108–22.
29. Lopez-Benitez M, Casadevall F. Empirical Time-Dimension Model of Spectrum Use Based on a Discrete-Time Markov Chain with Deterministic and Stochastic Duty Cycle Models, IEEE Transactions on Vehicular Technology. 2011; 60(6):2519-33.
30. Lotric U, Dobnikar A. Neural Networks with Wavelet Based Denoising Layers for Time Series Prediction, Neural Computing and Applications. 2005; 14:11-7.
31. Veitch D. Wavelet Neural Networks and their Application in the Study of Dynamical Systems. York: University of York. 2005.
32. Capizzi G, Napoli C, Bonanno F. Innovative Second-Generation Wavelets Construction with Recurrent Neural Networks for Solar Radiation Forecasting, IEEE Transactions on neural networks and learning systems. 2012; 23(11):1805-15.
33. Minu K, Lineesh M, Jessy C. Wavelet Neural Networks for Nonlinear Time Series Analysis, Applied Mathematical Sciences. 2010; 4(50):2485-595.
34. Božić J, Babić D. EUR/RSD Exchange Rate Forecasting Using Hybrid Wavelet-Neural Model: A Case Study, Computer Science and Information Systems. 2015; 12(2):487-508.
35. Hyndman R, Koehler AB, Ord JK, Snyder RD. Forecasting with Exponential Smoothing: The State Space Approach. 1st Edition, Berlin: Springer-Verlag Berlin Heidelberg. 2008.
36. Stolojescu-Crisan C. Data Mining based Wireless Network Traffic Forecasting. International Symposium on Electronics and Telecommunication; Timis. Romania, 2012.
37. Hyndman RJ. Another Look at Forecast-Accuracy Metrics for Intermittent Demand, Foresight: International Journal of Applied Forecast. 2006; 5(4):43-6.
38. Pedraza L, Hernandez C, Paez I. Evaluation of Nonlinear Forecasts for Radioelectric Spectrum, International Journal of Engineering and Technology. 2016; 8(3):1611-26.
39. Box G, Jenkins G, Reinsel C. Time Series Analysis: Forecasting and Control. 4th Edition, 2008.
40. Permanasari AE, Hidayah I, Bustoni I. SARIMA (Seasonal ARIMA) Implementation on Time Series to Forecast the Number of Malaria Incidence. International Conference on Information Technology and Electrical Engineering; Yogyakarta. 2013.
41. Pedraza L, Hernandez C, Rodriguez E. Modeling of GSM Spectrum Based on Seasonal ARIMA model. 6th IEEE Latin-American Conference on Communications; Cartagena de Indias. 2014.
42. Engle R. The Use of ARCH/GARCH Models in Applied Econometrics, Journal of Economic Perspectives. 1982; 15(4):157-68.
43. Bollersleva T. Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics. 1986; 31(3):307-27.
44. Chinomona A. Time Series Modelling with Application to South African Inflation Data. Pietermaritzburg: University of KwaZulu-Natal; 2009.
45. Edward N. Modelling and Forecasting Using Time Series Garch Models: An Application of Tanzania Inflation Rate Data. Tanzania: University of Dar es-Salaam; 2011.
46. Talke I. Modelling Volatility in Time Series Data. Pietermaritzburg: University of KwaZulu-Natal. 2003.
47. Erlwein C. Applications of Hidden Markov Models in Financial Modelling. London: Brunel University; 2008.
48. Zakai M. On the Optimal Filtering of Diffusion Processes. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete. 1969; 11(3):230-43.
49. Elliott R, Aggoun, Moore J. Hidden Markov Models. New

York: Springer-Verlag. 1995.
50. McLachlan G, Krishnan T. The EM Algorithm and Extensions. 2nd Edition, New Jersey. Wiley; 2008.
51. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q. The Empirical mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis. Proceedings of the Royal Society; London.1998:1-96.
52. Magrin I, Baraniuk R. Empirical Mode Decomposition based Time-Frequency Attributes, SEG Meeting; Houston.1999; 30(6):74-86.
53. Li D, Cao Y. SOFM based Support Vector Regression Model for Prediction and its Application in Power System Transient Stability Forecasting. International Power Engineering Conference; Singapore. 2005.
54. Wang W, Zhao H, Li Q, Liu Z.A Novel Hybrid Intelligent Model for Financial Time Series Forecasting and Its Application. Conference on Business Intelligence and Financial Engineering; Beijing. 2009.
55. 55. Dufrenois F, Hamad D. Fuzzy Weighted Support Vector Regression for Multiple Linear Model Estimation: Application to Object Tracking in Image Sequences. International Joint Conference on Neural Networks; Orlando. 2007.
56. Dufrenois F, Colliez J, Hamad D. Crisp Weighted Support Vector Regression for Robust Single Model Estimation: Application to Object Tracking in Image Sequences. Conference on Computer Vision and Pattern Recognition; Minneapolis. 2007.
57. Xie Y, Xu J, Yang H, Hu S. Phase-Space Reconstruction of ECoG Time Sequences and Extraction of Nonlinear Characteristic Quantities, Acta Physica Sinica. 2002; 51(2):205-14.
58. Shukuan L, Jianzhong Q, Guoren W, Shaomin Z, Lijia Z. Phase Space Reconstruction of Nonlinear Time Series Based on Kernel Method. The Sixth World Congress on Intelligent Control and Automation.Dalian.Shenyang. 2006.
59. Wei K, Li Y, Zhang P. Analysis and Application of Time Series Forecasting Model via Support Vector Machines, Systems Engineering and Electronics. 2005; 27(3):529-32.
60. Zhang Q, Benveniste A. Wavelet Networks, IEEE Trans Neural Netw. 1992; 3(6):889–98.
61. Daubechies I. Ten Lectures on Wavelets. 8th Edition, Philadelphia: SIAM; 2004.
62. Ashikin A, Ghazali R, Mat M. The Wavelet Multilayer Perception for the Prediction of Earthquake Time Series Data. International Conference on Information Integration and Web-based Applications and Services - iiWAS 11; Vietnam. 2011.
63. Akujuobi CM, Awada E, Sadiku M, Ali W. Wavelet-based Differential Nonlinearity Testing of Mixed Signal System ADCs. IEEE Proceedings in South East Conference Richmond. 2007.
64. Frimpong E, Okyere P. Monthly Energy Consumption Forecasting using Wavelet Analysis and Radial basis Function Neural Network, Journal of Science and Technology. 2010; 30(2):157-63.
65. Chun-Lin L. A tutorial of the Wavelet Transform. National Taiwan University, 2010.
66. Tummala P, Srinivasu P, Avadhani P, Murthy Y. Comparison of Variable Learning Rate and Levenberg-Marquardt Back-Propagation Training Algorithms for Detecting Attacks in Intrusion Detection System, International Journal on Computer Science and Engineering. 2011; 3(11):3572-82.
67. Li J, Cheng J, Shi J, Huang F. Brief Introduction of Back Propagation (BP) Neural Network Algorithm and Its Improvement, Advances in Computer Science and Information Engineering. New York: Springer; 2012. p. 553-8.
68. Lopez M, Casadevall F. Methodological Aspects of Spectrum Occupancy Evaluation in the Context of Cognitive Radio, European Transactions on Telecommunications. 2010; 21(8):680-93.
69. ITU-R. Report ITU-R SM.2256, Spectrum Occupancy Measurements and Evaluation. Date accessed: 24/04/2013. http://www.itu.int/pub/R-REP-SM.2256-2012.
70. Digham F, Alouini MS, Simon M. On the Energy Detection of Unknown Signals over Fading Channels, IEEE Transactions on Communications. 2007; 55(1):21-4.
71. Dickey DA, Fuller WA. Distribution of Estimators for Autoregressive Time Series with A Unit Root, J of the Amer. Statistical Assoc. 1979; 74(366):427-31.
72. Bozdogan H. Model Selection and Akaike's Information Criterion (AIC): The General Theory and its Analytical Extensions, Psychometrika. 1987; 52(3):345-70.
73. Debnath L, Shah F. Wavelet Transforms and Their Applications. 2nd edition, New York: Birkhäuser; 2014.