Evaluation of nonlinear forecasts for radioelectric spectrum

Luis Pedraza [#*1], Cesar Hernández [#*2], Ingrid Paez [*3]
[#] Faculty of Technology, Universidad Distrital Francisco José de Caldas,
Carrera 7 No. 40B – 53, Bogota, Colombia
[1] lfpedrazam@udistrital.edu.co
[2] cahernandezs@udistrital.edu.co
[*] Faculty of Engineering, Industrial and Systems Engineering Department,
Universidad Nacional de Colombia, Carrera 45 N° 26-85, Bogota, Colombia
[1] lufpedrazama@unal.edu.co
[2] ceahernandezsu@unal.edu.co
[3] ippaezp@unal.edu.co

Abstract—This paper presents the development and evaluation of three nonlinear models to forecast the power reception of different channels of the global system for mobile communications (GSM) in order to analyze spatial opportunity to reuse frequencies by secondary users (SUs) in a cognitive radio (CR) network. Markov, empirical mode decomposition-support vector regression (EMD-SVR) and wavelet neural models were utilized to forecast the channel occupancy status. Results were evaluated using the criteria of availability and occupancy times of channels, different types of mean error, and observation time. This study forecasts not only the reception power but also the occupancy and availability time of channels to determine the accuracy percentage that can have of the channel use time for the primary users (PUs) and SUs in CR systems. The analysis of the models presents the wavelet neural model as the one with the best behavior in the variables evaluated. Thus, this study suggests that this model represents a promising alternative to CR systems.

Keyword-Radioelectric Spectrum, Cognitive Radio, Time Series, Forecast Models, Nonlinear Models.

I. INTRODUCTION

Radioelectric spectrum occupancy has been widely studied due to its importance for the construction of new spectrum assigning policies in emerging technologies as well as in monitoring activities both in licensed and unlicensed bands. Real measurements for spectrum use within a determined band enable the corresponding authorities to guarantee that licensees meet local and regional spectrum regulations [1]. 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. 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 forecast of communications. This could contribute to spectral efficiency improvement efforts [2-4]. 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 [5]. Besides, SUs can select the channels with higher probability of vacancy in multi-channel wideband sensing scenarios by means of prediction information [6], and also PUs occupancy models can be used as empty channel indicators replacing the spectrum sensing procedures [7].

Various initiatives have been proposed to model radio spectrum [2, 8-14]. One of the fundamental differences between the typical proposals and the one presented here lies in the fact that in those the time series of the duty cycle of different channels is modeled, whereas in our proposal the time series of received power in three GSM channels are modeled by means of different occupancy levels. To do this, three models have been employed. In the first place, the Markov model, which has been used especially to forecast binary channel occupation [10, 11] in wireless networks. For example in [15], 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), an appropriate method to forecast non-stationary signals, and empirical mode decomposition (EMD), was used to analyze both stationary and non-stationary signals. In [12] 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 [16-20].

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)) [21-23].

The remainder of this paper is then divided in the following way: section II describes the characteristics of the models employed; in section III the statistical analysis of the spectrum measures and the models' design is carried out; section IV presents comparative results of the models studied; and section V shows the conclusions drawn in this paper.

II. THEORY AND BACKGROUND

The models that will be described in this part start from the assumption of non-linear time series. This is particularly true for short-term analyses.

A. Markov Model

The parameters of the model are unknown and must be settle 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 [24].

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 ... m_N}, where *x* is the function *x*: Ω→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 [24].The process *x* is said to be a Markov chain if it satisfies Markov's property [24]

$$P\left(x_{k+1} = m_{x_{k+1}} \mid x_0 = m_0, \dots, x_k = m_k\right) = P\left(x_{k+1} = m_{x_{k+1}} \mid x_k = m_k\right) \quad (1)$$

$$\forall k \geq 1 \;\; y m_0, m_1, \dots, m_k \in M$$

The initial distribution of *x* is defined by X = (X_m: m∈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 [24],

$$\pi_{ji} = P(x_{k+1} = j \mid x_k = i), \;\; i, j \in M \quad (2)$$

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.

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 [25].

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 [26].

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 [24]

$$J_k^{(sr)} := \sum_{l=1}^{k} \langle x_{l-1}, e_r \rangle \langle x_l, e_s \rangle \quad (3)$$

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 by [24]

$$O_k^{(r)} := \sum_{l=1}^{k} \langle x_{l-1}, e_r \rangle \tag{4}$$

An auxiliary process is also needed to estimate the vectors$\pi, \alpha, \gamma y \xi$. This has the following form [24]:

$$T_k^{(r)}(g) := \sum_{l=1}^{k} \langle x_{l-1}, e_r \rangle g(y_l) \tag{5}$$

where g is a function.

The representation in discrete time of the reception power process is obtained from [24]

$$P_{r_{k+1}} = \alpha(x_k) P_{r_k} + \gamma(x_k) + \xi(x_k) w_{k+1} \tag{6}$$

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 [24]. 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 $\hat{\theta} = \{\hat{\pi}_{ji}, \hat{\gamma}_i, \hat{\alpha}_i, \hat{\xi}_i \;\; 1 \le i, j \le n\}$. 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 [24, 27].

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

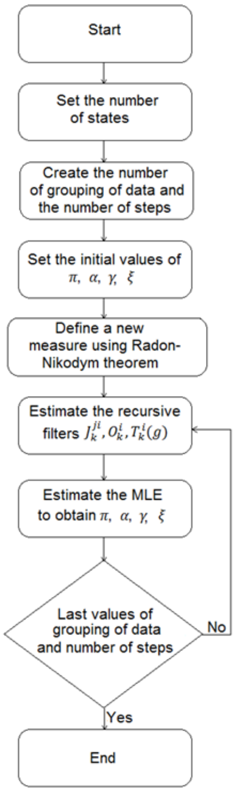

Fig. 1. Flow chart of hidden Markov model

B. 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 [28]:

- 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 EMD algorithm shown below [29]:

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 2 with i=i+1 else the decomposition is finished and $r_i(t)$ is the residue.

At the end of the algorithm we obtain

$$x(t) = \sum_{i=1}^{n} imf_i(t) + r_n(t) \tag{7}$$

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

2) 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 [30]:

$$f(x) = \langle w, \Phi(x)\rangle + b \tag{8}$$

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 solution to the convex optimization problem [31]:

$$\min_{w,\xi_i,\xi_i^*} L = C\sum_{i=1}^{N}(\xi_i + \xi_i^*) + \frac{1}{2}\|w\|^2 \tag{9}$$

$$\begin{cases} y_i - \langle w, \Phi(x_i)\rangle - b \le \varepsilon + \xi_i \\ \langle w, \Phi(x_i)\rangle + b - y_i \le \varepsilon + \xi_i^* \\ \quad \xi_i, \xi_i^* \qquad\qquad \ge 0 \end{cases}$$

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_iy ξ_i^* called "slack variables", which characterize the deviation of training samples outside the ε-margin [31]. In formula (9) 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 [30]. 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 [30, 31]:

$$\max_{a_i,a_i^*} L = -\frac{1}{2}\sum_{i,j=1}^{N}(a_i^* - a_i)(a_j^* - a_j)\left(\Phi(x_i), \Phi(x_j)\right) - \sum_{i=1}^{N}(a_i^*(y_i - \varepsilon) - a_i(y_i + \varepsilon)) \tag{10}$$

Subject to

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

where the variables α_iy 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 [30]:

$$w = \sum_{i=1}^{N}(a_i - a_i^*)\Phi(x_i) \tag{11}$$

$$f(x) = \sum_{i=1}^{N}(a_i - a_i^*)\langle\Phi(x_i), \Phi(x)\rangle + b \tag{12}$$

In (12) the scalar product in the feature space ‹φ(x_i),φ(x)› 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[30]. The most employed kernel function is the Gaussian RBF with a width of σ[30].

$$k(x, x_i) = exp\left\{-\frac{|x-x_i|^2}{\sigma^2}\right\} \tag{13}$$

3) Regression support vector in time series forecast

In the modeling and forecast of nonlinear time series, the phase space reconstruction (PSR) is essential [32]. 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 [33]. 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) [33]. Given a time series $\{x_t\}_{t=1}^{N}$, reconstruct the feature vector [33]:

$$X_t = \left(x_t, x_{t-1}, \dots, x_{t-(m-1)}\right) \tag{14}$$

(m is embedding dimension, delay time sets in 1; $t = m$, m +1,…, 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 [33]:

$$Y_t = x_{t+1} = f\left(x_t, x_{t-1}, \dots, x_{t-(m-1)}\right) = f(X_t) \tag{15}$$

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 [34]:

$$Y_N = \hat{x}_{N+1} = \sum_{i=1}^{N-m}(a_i - a_i^*)\, K(X_N, x_i) + b \tag{16}$$

where $X_N=(x_N, x_{N-1}, \dots, x_{N-(m-1)})$

4) Forecast model EMD-SVR

The EMD-SVR model, as it can be seen in Fig. 2, uses the EMD algorithm to decompose data series $\{x_1, \dots, x_l\}$ in a finite set of IMFs. Then, the forecasts on these IMFs are made using SVR model to obtain the forecasted value $\widehat{imf_i}(l+1)$; and finally, according to formula (7), the forecasted value $\hat{x}(l+1)$ can be found by the sum of previously forecasted results [12].

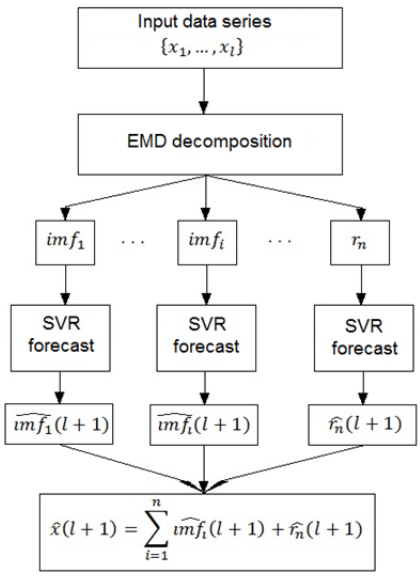

Fig. 2.EMD-SVR flow chart [12]

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 raw 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[12].

C. Wavelet Neural Model

In [35] the idea of "wavelet networks" is proposed, the theory combines both wavelets and neural nets.

1) 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"[35, 36].

A wavelet is a "small wave" function, usually noted as *ψ(·)*. 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. The function *ψ(·)* is generally referred to as mother wavelet. A wavelets family can be generated from translation and expansion of this mother wavelet [17].

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 [37]. Wavelet coefficients may be expressed as [38]:

$$W_{\Phi}[j_0, k] = \frac{1}{\sqrt{M}} \sum_n f[n] \Phi_{jo,k}[n] \tag{17}$$

$$W_{\psi}[j, k] = \frac{1}{\sqrt{M}} \sum_n f[n] \psi_{j,k}[n] j \geq j_0 \tag{18}$$

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 (17) represent the approximate coefficients, whereas the ones in (18) represent the detailed coefficients.

2) 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 is modified by the fact that neurons may die and new synapses may grow [17].

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; 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. Fig. 3 shows a scheme used to represent a neuron [39].

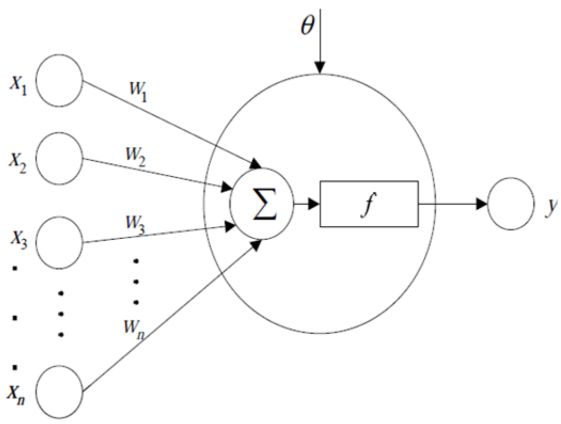

Fig. 3. Model of a neuron [39]

In Fig, 3, the output of the neuron is given by [39]:

$$y = f[(\sum_{i=1}^{n} w_i x_i - \theta)], i = 1,2,3 \dots n \tag{19}$$

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

3) 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 backpropagation 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 Fig. 4, the wavelet neural network model developed is presented.

The neural model employed in this research makes use of a multilayer backpropagation neural network, whose output error is propagated backwards to adjust the weights that involve the error minimization. The backpropagation networks learn with the descending gradient method, which defines the way to train the output nodes in a multilayer network [40].

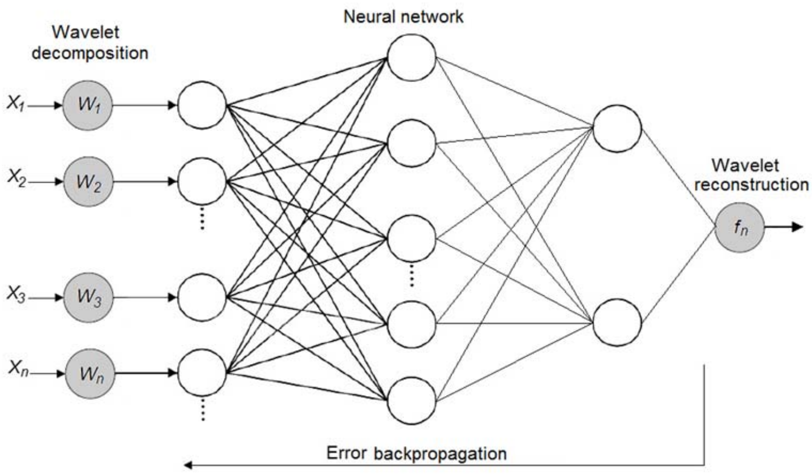

Fig. 4. Wavelet neural network

III. CASE STUDY AND EXPERIMENT PROCEDURE

The decision to carry out this study was made during the spectrum measures campaign held in Bogota-Colombia, where we obtained the measurements employed in this work from a spectrum occupancy study previously carried out [1, 41]. 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 correspond to a week, from December 23 to 29, 2012. In some studies [42], it has been indicated that a reasonable option for obtaining 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. While a 24 hours measurement period could be thought as adequate in order to properly characterize the activity of determined spectrum bands [43], 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 [43]. Thus, the data gathered along the 7 days were used to train the models previously presented and to particularly forecast the reception power data of 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 duty cycles of 60 channels in the GSM band. From these, three channels with different occupancy levels (high, medium and low), were chosen. Fig. 5 presents results of power measures 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 at selected channels are depicted in Fig. 6. Threshold (λ) used, which for this event is of -89dBm, was obtained from (20) with a probability of false alarm (Pfa) of 1% [44]:

$$P_{fa} = \frac{\Gamma\left(m, \frac{\lambda}{2}\right)}{\Gamma(m)} \tag{20}$$

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

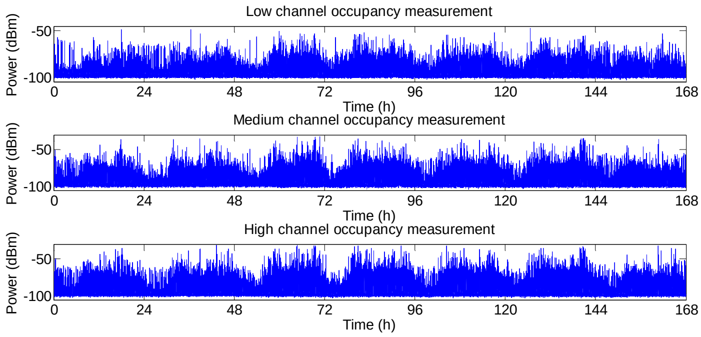

Fig. 5. Power measurements for three GSM band downlink channels [45]

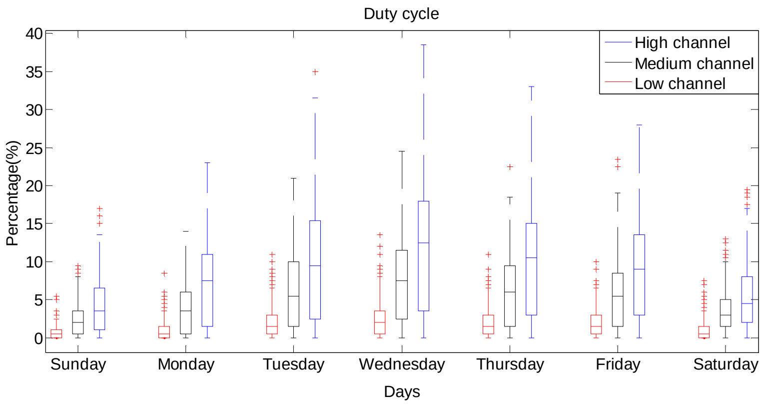

Fig. 6. Duty cycles for three GSM band downlink channels [45]

Fig. 7, 8 and 9 present histograms corresponding to distribution opportunities 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.

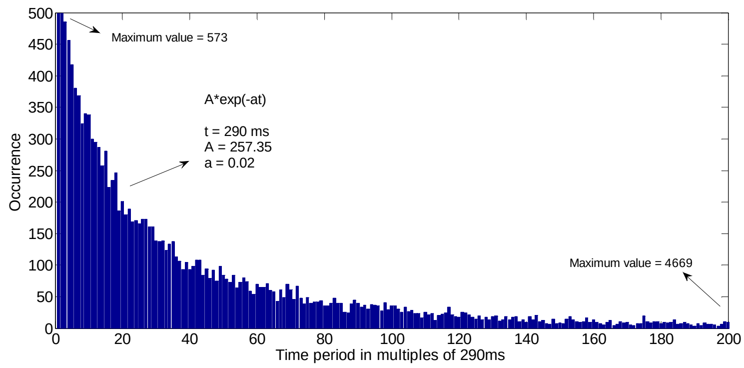

Fig. 7. Time period opportunities distribution for low channel

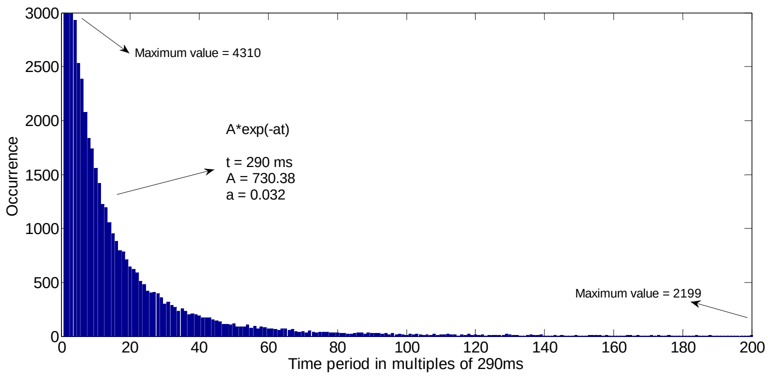

Fig. 8. Time period opportunities distribution for medium channel

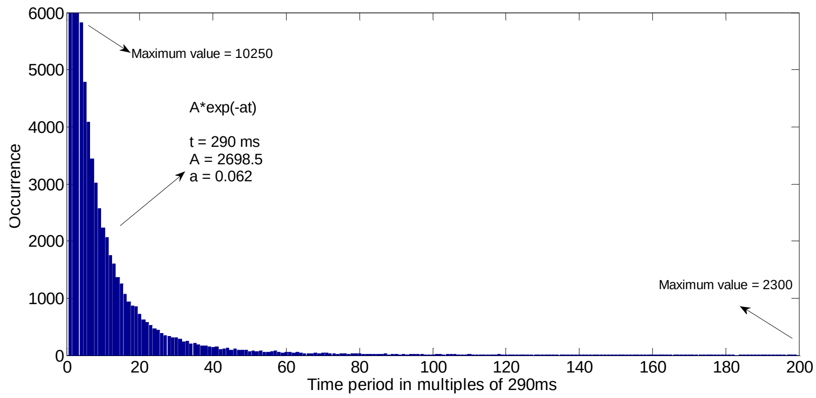

Fig. 9. Time period opportunities distribution for high channel

In the following we proceeded to analyze the time series of measured channels during a week, which were equivalent to 1062514 samples. To do this, autocorrelation function (ACF) is initially presented, as it is observed in Fig. 10. ACF diagrams for the three channels present a form which is alternately positive and negative, decaying to zero, therefore there correlation [45].

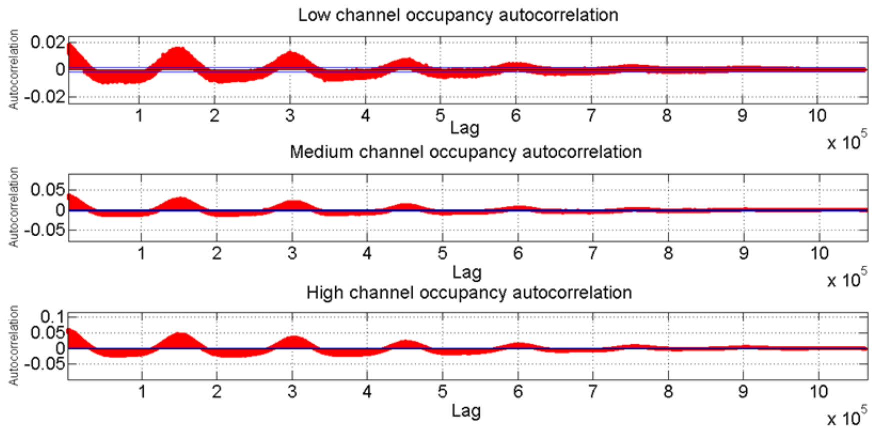

Fig. 10. Autocorrelation for three GSM band downlink channels [45]

A. Design of the Markov Model

The design of this model is based on the flow chart of Fig. 1. Since the estimations of the parameters are calculated through the EM algorithm, 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 [24]:

$$E[y_{k+1}|F_k] = E[\alpha(x_k)y_k + \gamma(x_k) + \xi(x_k)w_{k+1}|F_k] = \langle\alpha, \Pi\hat{x}_k\rangle y_k + \langle\gamma, \Pi\hat{x}_k\rangle \quad (21)$$

where $\hat{x}_k = E[x_k|F_k]$. The number of states was chosen according to the minor akaike information criterion, in this case it is of three states, compared with the values of two and four states. Fig. 11, 12 and 13 depicts 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.

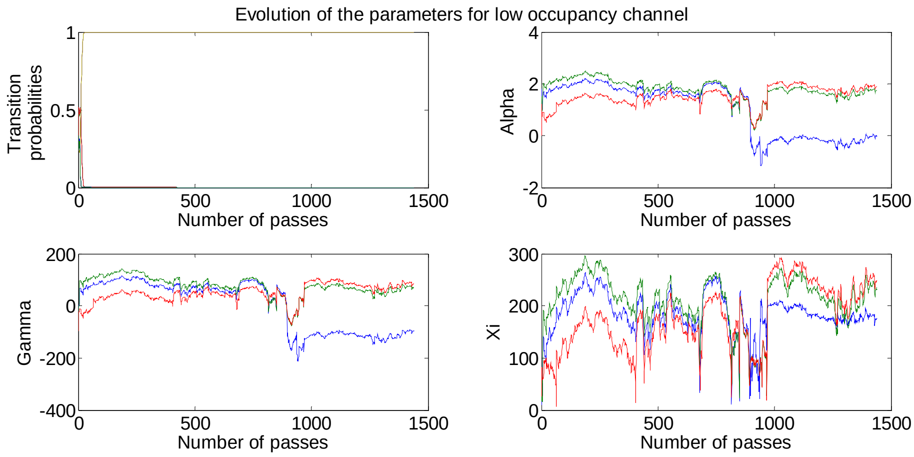

Fig. 11. Evolution of the parameters α, γ, ξ and the transition probability of the low occupancy channel

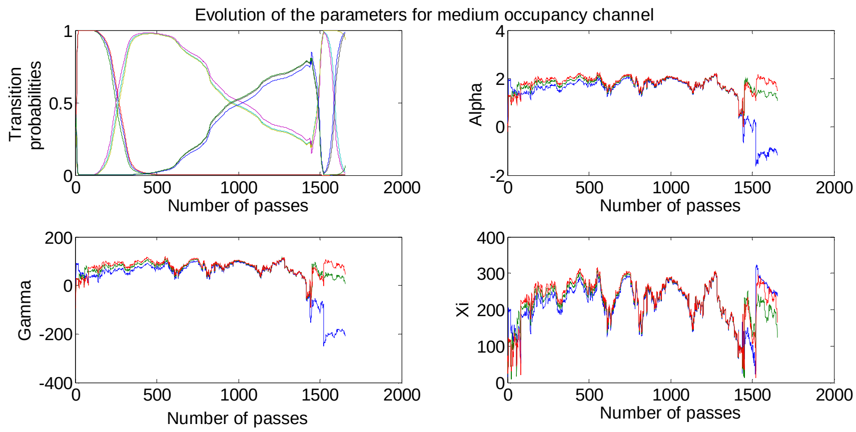

Fig. 12. Evolution of the parameters α, γ, ξ and the transition probability of the medium occupancy channel

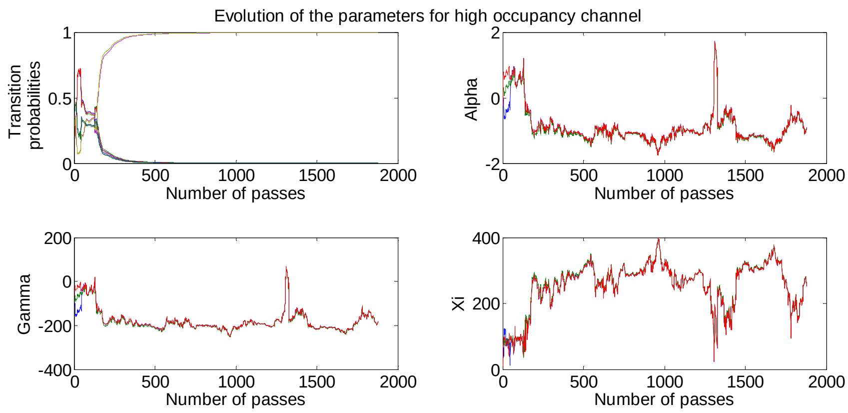

Fig. 13. Evolution of the parameters α, γ, ξ and the transition probability of the high occupancy channel

B. Design of the EMD-SVR Model

The EMD-SVR model has a higher time-consuming in processing than the other 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 Fig. 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 Fig. 14, 15 and 16:

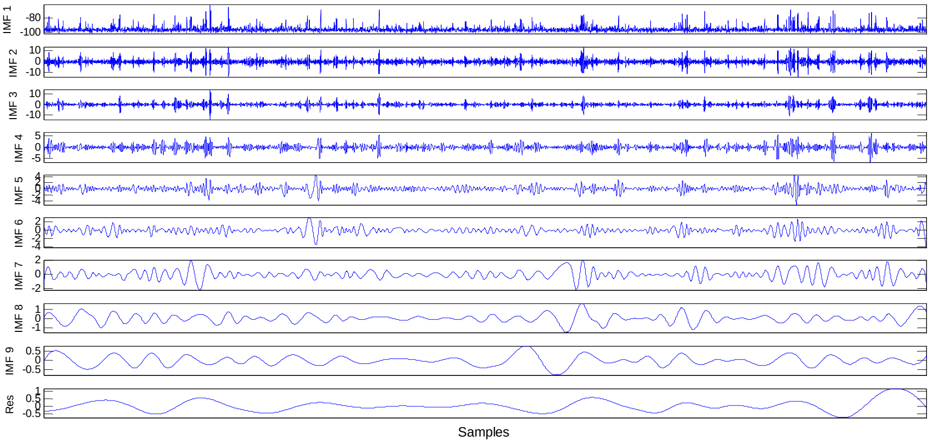

Fig. 14.EMD data results for the low occupancy channel

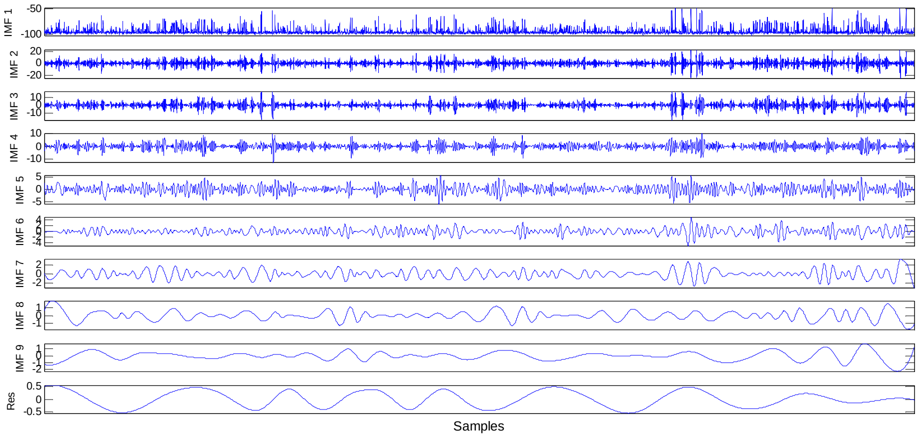

Fig. 15.EMD data results for the medium occupancy channel

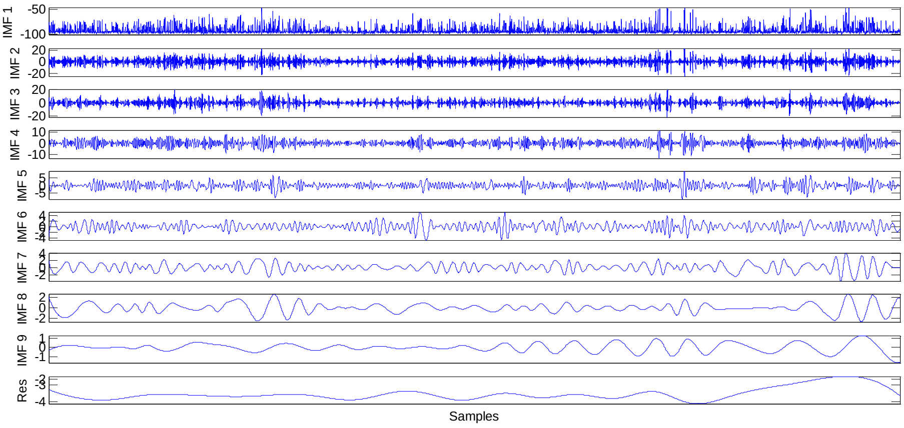

Fig. 16.EMD data results for the high occupancy channel

- A normalized processing of data series was carried out, in order to improving 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 is reconstructed and data corresponding to one hour are forecast.

C. 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 mothers wavelet *Daubechies, Coiflets*and*Symlets* [46]. The result is two levels containing in total four coefficients.

The multilayer backpropagation neural network used in this study is depicted in Fig. 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.

IV. RESULTS AND DISCUSSION

Fig. 17 shows forecasts from the Markov model obtained from Fig, 1, from the EMD-SVR model based on Fig. 2, and from the wavelet neural model presented in Fig. 4. These forecasts were contrasted with the measured data for the powers of Friday from 5 to 6 pm.

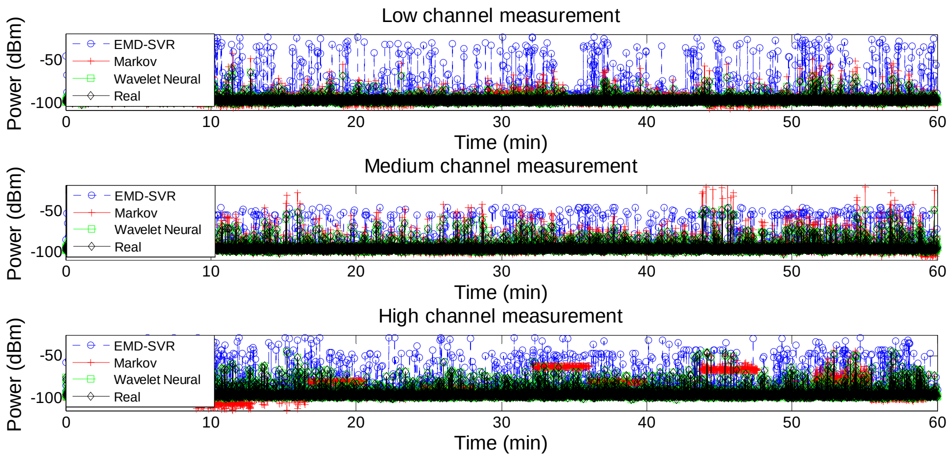

Fig. 17. Time series measured and forecasted for low, medium and high GSM channels using Markov, EMD-SVR and wavelet neural models

Availability and occupation times for the channels measured and forecasted are depicted in Fig. 18 and 19. 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 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 I. The above was calculated for the three GSM channels selected.

TABLE I. 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**	**Low occupancy 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

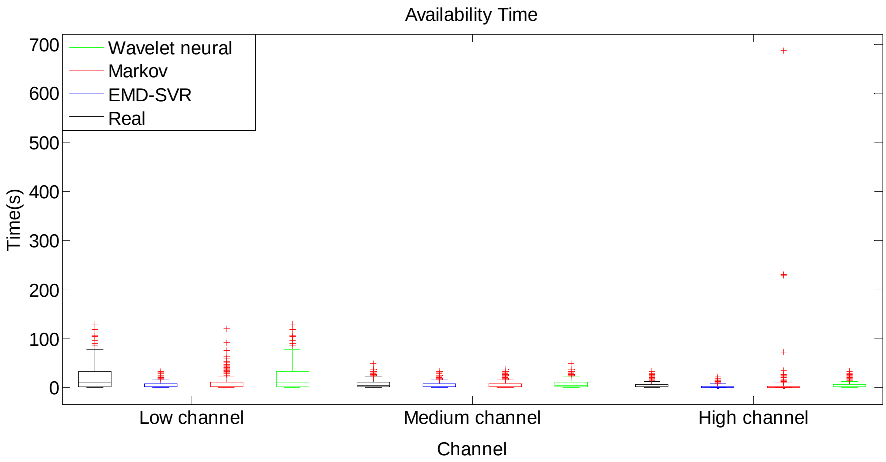

Fig. 18. Availability time of the forecasted channels

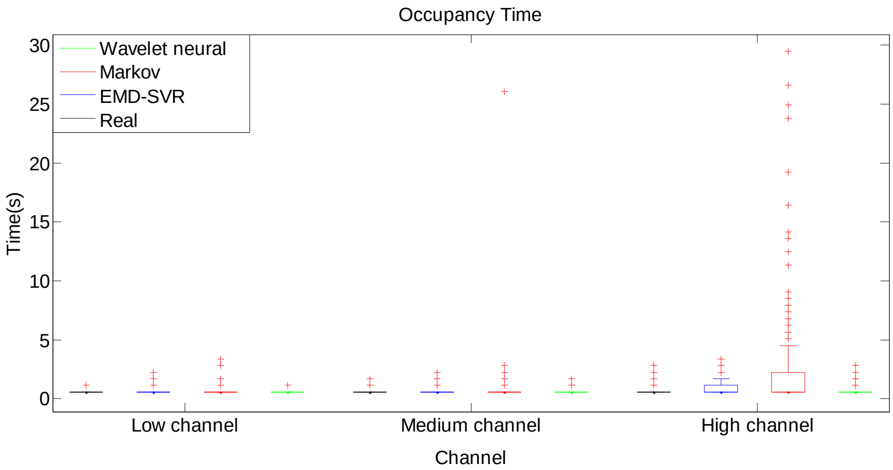

Fig. 19. Occupancy time of the forecasted channels

Tables II, III and IV displays the errors between real and forecasted 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 EMD-SVR and Markov 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.

TABLEII.SMAPE comparison among the forecasted models

Channel occupancy	**SMAPE-EMD-SVR**	**SMAPE-Markov**	**SMAPE-Wavelet neural**
Low	-0.0681	-0.0231	-0.0017
Medium	-0.0654	-0.02	-0.0020
High	-0.0991	-0.1201	-0.0019

TABLE III.MAPE comparison among the forecasted models

Channel occupancy	**MAPE- EMD-SVR**	**MAPE- Markov**	**MAPE- Wavelet neural**
Low	0.0556	0.0227	0.00089
Medium	0.0598	0.0189	0.0011
High	0.0890	0.1117	0.0010

TABLE IV. MAE comparison among the forecasted models

Channel occupancy	MAE- EMD-SVR	MAE- Markov	MAE- Wavelet neural
Low	5.296	2.1336	0.0866
Medium	5.411	1.6016	0.1
High	8.022	4.3067	0.1005

Fig. 20 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 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, wavelet neural method reduces forecast error by as much as 99.62% when compared with Markov model and as much as 99.8% when compared with EMD-SVR model.

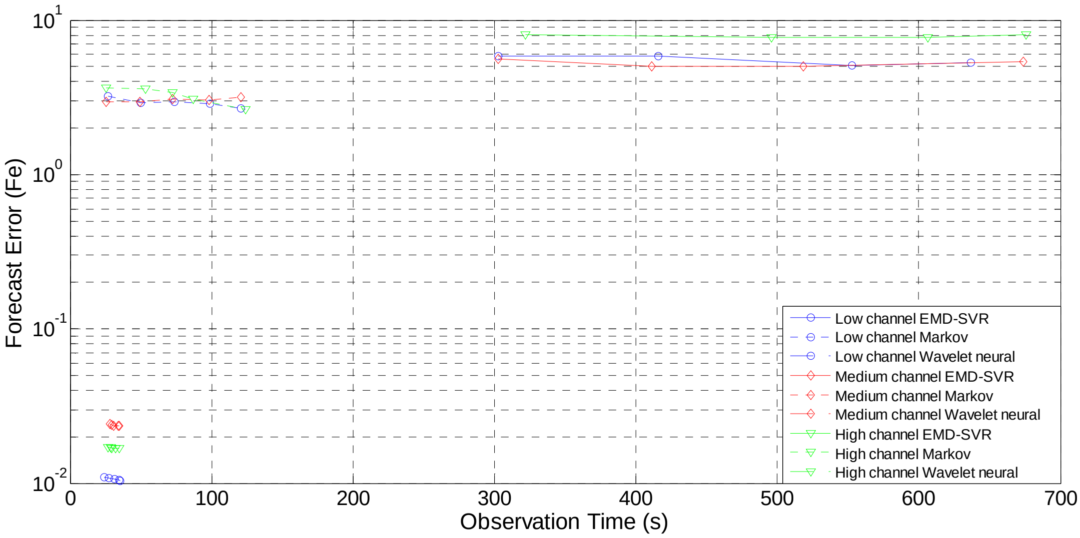

Fig. 20. Forecast error vs. Observation time

V. CONCLUSIONS

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 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 wavelet neural gave the most optimal results. Other important aspect is that forecast error and observation time decrease 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 systems of CR.

ACKNOWLEDGEMENT

The work was supported by the Administrative Department of Science, Technology and Innovation (COLCIENCIAS).

REFERENCES

[1] L. Pedraza, F. Forero, and I. Paez, "Evaluación de ocupación del espectro radioeléctrico en Bogotá-Colombia," Ingeniería y Ciencia, vol. 10,no. 19, pp. 127-143, 2014.
[2] A. Gorcin, et al., "An autoregressive approach for spectrum occupancy modeling and prediction based on synchronous measurements," in Int. Symp. on Personal Indoor and Mobile Radio Commun., Toronto, 2011, pp. 705-709.
[3] L. Pedraza, Redes inalámbricas mesh "Caso de estudio: ciudad Bolívar", 1st ed. Bogota: Universidad Distrital Francisco José de Caldas, 2012.
[4] L. Pedraza, et al., Ocupación espectral y modelo de radio cognitiva para la ciudad de Bogotá, 1st ed. Bogota: Universidad Distrital Francisco José de Caldas, 2016.
[5] M. Wellens, J. Riihijarvi, and P. Mahonen, "Empirical time and frequency domain models of spectrum use," Physical Communication, vol. 2,no. 1-2, pp. 10-32, 2009.
[6] C. Song, D. Chen, and Q. Zhang, "Understand the predictability of wireless spectrum: a large-scale empirical study," in International Conference on Communications, Cape Town, 2010, pp. 1-5.

[7] Z. Sun, J. N. Laneman, and G. J. Bradford, "Sequence detection algorithms for dynamic spectrum access networks," in IEEE International Symposium on New Frontiers in Dynamic Spectrum, Singapore, 2010, pp. 1-9.
[8] M. Lopez-Benitez and F. Casadevall, "Time-dimension models of spectrum usage for the analysis, design, and simulation of cognitive radio networks," IEEE Transaction on Vehicular Technology, vol. 62,no. 5, pp. 2091-2104, 2013.
[9] S. Yarkan and H. Arslan, "Binary time series approach to spectrum prediction for cognitive radio," in Vehicular Technology Conf., Dublin, 2007, pp. 1563-1567.
[10] T. Black, B. Kerans, and A. Kerans, "Implementation of hidden markov model spectrum prediction algorithm," in Int. Symp. on Commun. and Inform. Technologies, Gold Coast, 2012, pp. 280-283.
[11] L. Yang, et al., "Spectrum usage prediction based on high-order markov model for cognitive radio networks," in Int. Conf. on Comput. and Inform. Technology, Bradford, 2010, pp. 2784-2788.
[12] C. Yu, Y. He, and T. Quan, "Frequency Spectrum Prediction Method Based on EMD and SVR," in Intelligent Systems Design and Applications, Kaohsiung, 2008, pp. 39-44.
[13] Z. Wang and S. Salous, "Time series arima model of spectrum occupancy for cognitive radio," in Seminar on Cognitive Radio and Software Defined Radios: Technologies and Techniques, London, 2008, pp. 1-4.
[14] Y. Chen and H. S. Oh, "A Survey of Measurement-based Spectrum Occupancy Modeling for Cognitive Radios," IEEE Communications Surveys & Tutorials, pp. 1-36, 2014.
[15] M. Lopez-Benitez and F. Casadevall, "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, vol. 60,no. 6, pp. 2519-2533, 2011.
[16] U. Lotric and A. Dobnikar, "Neural Networks with Wavelet Based Denoising Layers for Time Series Prediction," Neural Computing and Applications, vol. 14,pp. 11-17, 2005.
[17] D. Veitch, "Wavelet Neural Networks and their application in the study of dynamical systems," Master thesis, Department of Mathematics, University of York, York, 2005.
[18] G. Capizzi, C. Napoli, and F. Bonanno, "Innovative Second-Generation Wavelets Construction With Recurrent Neural Networks for Solar Radiation Forecasting," IEEE Transactions on neural networks and learning systems, vol. 23,no. 11, pp. 1805-1815, 2012.
[19] K. Minu, M. Lineesh, and C. Jessy, "Wavelet Neural Networks for Nonlinear Time Series Analysis," Applied Mathematical Sciences, vol. 4,no. 50, pp. 2485-2595, 2010.
[20] J. Božić and D. Babić, "EUR/RSD Exchange Rate Forecasting Using Hybrid Wavelet-Neural Model: A CASE STUDY," Computer Science and Information Systems, vol. 12,no. 2, pp. 487-508, 2015.
[21] R. Hyndman, et al., Forecasting with Exponential Smoothing: The State Space Approach, 1st ed. Berlin: Springer-Verlag Berlin Heidelberg, 2008.
[22] C. Stolojescu-Crisan, "Data mining based wireless network traffic forecasting," in Int. Symp. on Electron. and Telecommun., Timis, 2012, pp. 115-118.
[23] R. J. Hyndman, "Another look at forecast-accuracy metrics for intermittent demand," Foresight: Int. J. Appl. Forecast., no. 4, pp. 43-46, 2006.
[24] C. Erlwein, "Applications of hidden Markov models in financial modelling," Doctor of Philosophy, Department of Mathematical Sciences, Brunel University, London, 2008.
[25] M. Zakai, "On the optimal filtering of diffusion processes," Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, vol. 11,no. 3, pp. 230-243, 1969.
[26] R. Elliott, L. Aggoun, and J. Moore, Hidden Markov Models vol. 29. New York: Springer-Verlag, 1995.
[27] G. McLachlan and T. Krishnan, The EM Algorithm and Extensions, 2nd ed. New Jersey: Wiley, 2008.
[28] N. E. Huang, et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," in Proceedings of the Royal Society, London, 1998, pp. 903-995.
[29] I. Magrin and R. Baraniuk, "Empirical Mode Decomposition based time-frequency attributes," in SEG Meeting, Houston, 1999.
[30] D. Li and Y. Cao, "SOFM based support vector regression model for prediction and its application in power system transient stability forecasting," in International Power Engineering Conference, Singapore, 2005, pp. 765-770.
[31] F. Dufrenois and D. Hamad, "Fuzzy weighted support vector regression for multiple linear model estimation : application to object tracking in image sequences," in International Joint Conference on Neural Networks, Orlando, 2007, pp. 1289-1294.
[32] Y. Xie, et al., "Phase-space reconstruction of ECoG time sequences and extraction of nonlinear characteristic quantities," Acta Physica Sinica, vol. 51,no. 2, pp. 205-214, 2002.
[33] L. Shukuan, et al., "Phase Space Reconstruction of Nonlinear Time Series Based on Kernel Method," in The Sixth World Congress on Intelligent Control and Automation, Dalian, 2006, pp. 4364-4368.
[34] K. Wei, Y. Li, and P. Zhang, "Analysis and Application of Time Series Forecasting Model via Support Vector Machines," Systems Engineering and Electronics, vol. 27,no. 3, pp. 529-532, 2005.
[35] Q. Zhang and A. Benveniste, "Wavelet networks," IEEE Trans. Neural Netw., vol. 3,no. 6, pp. 889–898, 1992.
[36] I. Daubechies, Ten Lectures on Wavelets, 8th ed. Philadelphia: SIAM, 2004.
[37] C. M. Akujuobi, et al., "Wavelet-based differential nonlinearity testing of mixed signal system ADCs," in IEEE Proceedings in SoutheastCon, Richmond, 2007, pp. 76-81.
[38] L. Chun-Lin, "A tutorial of the wavelet transform," National Taiwan University 2010.
[39] E. Frimpong and P. Okyere, "Monthly energy consumption forecasting using wavelet analysis and radial basis function neural network," Journal of Science and Technology, vol. 30,no. 2, pp. 157-163, 2010.
[40] J. Li, et al., "Brief Introduction of Back Propagation (BP) Neural Network Algorithm and Its Improvement," in Advances in Computer Science and Information Engineering. vol. 2, ed New York: Springer, 2012, pp. 553-558.
[41] L. Pedraza, F. Forero, and I. Paez, "Metropolitan Spectrum Survey in Bogota Colombia," in IEEE International Conference on Advanced Information Networking and Applications Workshops, Barcelona, 2013, pp. 548-553.
[42] M. López and F. Casadevall, "Methodological aspects of spectrum occupancy evaluation in the context of cognitive radio," European Transactions on Telecommunications, vol. 21,no. 8, pp. 680-693, 2010.
[43] ITU-R, "Report ITU-R SM.2256, spectrum occupancy measurements and evaluation," Geneva 2012.
[44] F. Digham, M. S. Alouini, and M. Simon, "On the Energy Detection of Unknown Signals Over Fading Channels," IEEE Transactions on Communications, vol. 55,no. 1, pp. 21-24, 2007.
[45] L. Pedraza, C. Hernandez, and E. Rodriguez, "Modeling of GSM Spectrum Based on Seasonal ARIMA model," in 6th IEEE Latin-American Conference on Communications, Cartagena de Indias, 2014, pp. 1-4.

AUTHOR PROFILE

Luis Pedraza is an associate professor at the Universidad Distrital Francisco José de Caldas. His research interests include wireless networks.

Cesar Hernández is an associate professor at the Universidad Distrital Francisco José de Caldas. His research interests include telecommunications.

Ingrid Paez is full time professor at the Industrial and Systems Engineering Department, Universidad Nacional de Colombia, Bogota. Her main research interest is the area of wirelesscommunication systems.