# COMBINING WAVELET AND KALMAN FILTERS FOR FINANCIAL TIME SERIES FORECASTING

Autores: Fabiano Guasti Lima e Alexandre Assaf Neto
Publicação: Asian Economic and Financial Review

## ABSTRACT

Among the numerous techniques that played the role of forecasting a future value, there are several different previously verified models found in literature and, more recently, the time series separation filters appeared as a complementary alternative to the current forecasting techniques. Therefore, the general purpose of this study is to carry out a comparative analysis of the combined use of wavelet and Kalman filters along with forecasting models for financial time series in order to verify which of them produces the best forecast. After testing the junction of techniques for a high volatility time series, such as IBOVESPA, the results first indicate the use of the Kalman filter and next the use of *wavelets* with recurring neural networks, with error of 0.72% measured by MAPE. Therefore, this paper is a contribution to the area by creating a way of reducing errors in forecasting and, consequently, developing better risk management in investment positions in the financial market.

**KEY WORDS**: Wavelets, Kalman, Forecasting, Financial Time Series

## 1. INTRODUCTION

The current financial crises occurring due to the financial market instability led to the scenery of increasing relevance of the risk management in the investment strategies. Thus, the market itself became fickler and the worry of the investors concerning unpleasant surprises requires the information to be better prepared for dealing with the market adversities.

The financial market movements seem to be clearer in relation to the turbulences caused by the information published daily. Along with those sudden market behaviors, the companies which operate together in these markets are seeking advances in their internal management, mainly in the risk mitigation. The need of financial projections concerning this market volatility, mainly the short term ones, is essential for the management of their positions in the cash market and future exchanges through, principally, the use of the forecasting models of future values.

There are several models to forecast time series in the literature, from the most simple and easiest to comprehend ones to the most complex ones which include different parameters such as the ARIMA models and the GARCH family ones. And the fact of making use of more complex statistical models does not necessarily mean an improvement in the result forecast since the series present noise along the time.

The use of the time series decomposition through *wavelets*, aiming the analysis of the time series, appeared as alternative for the noise reduction in the time series. The combination of this methodology with the traditional forecasting models were used by Granger (1992, p. 3), Tak (1995, p. 43), Ariño (1995), Ukil & Zivanovic (2001, p. 103), Ma, Wong & Sankar (2004, p. 5824) and Aminghafari (2007, p. 715).

In Brazil, the works of time series forecast of Chiann (1997, p. 32), Homsy, Portugal & Araújo (2000, p. 10), Zandonade & Morettin (2003, p. 205), Lima (2004, p. 133) and Rocha (2008, p. 120) applying several methodologies of time series forecast and the use of filters as the *wavelets* are distinct.

The *wavelets* are functions which consist in fractioning the original time series in two subseries, one concerning the high frequency and the other one the low frequency aiming to reduce the noise effect in the forecasts (Gençay, Selçuk & Whitcher, 2002). The use of this series filtering process brought significant improvement in the forecasting models according to what can be seen in the works mentioned before which used only the traditional forecasting models.

In contrast to this static approach of the forecasting models, the Dynamic Linear Models (DLM) appeared, introduced by Kalman (1960), which are formulated with the characteristic of incorporating changes in the parameters as the evolutions in the time series. The increase in the number of observations of the series is interpreted, therefore, as additional information to the current information set, causing the parameters to present a dynamic evolution, impeding any static quantification of the subjacent relations to the global behavior of the series.

The operationalization of the DLM is obtained by adopting the state space model and making using of the Kalman filter for the sequential update of the unobservable components. The representation in state space is done by a system of two dynamic equations which describes the way through which the observations are generated according to the state vector and its dynamic evolution. The Kalman filter consists basically in an algorithm that supplies updated estimates of the state vector every single time moment.

Recently, new works are making use of the Kalman filter methodology to make the forecasts such as, Aiube (2005, p. 108) and Corsini & Ribeiro (2008, p. 11) and with the combined use of Kalman and *wavelets* as in Postalcioglu, Erkan & Bolat (2005, p. 951).

The fact is that both the theories and the modelings start from their premises, adopt their analysis methods and obtain their results. Such a fact reports the dimension that there is no absolute sureness in favor of either theory. Neither is there a way of combining those methods and filters in a single forecast scheme for decision-taking.

The several applications which can be done with the junction of the filtering techniques with the forecast models make this field of study to be one of the most dynamic one in the finance study. Therefore, the continuity of the studies concerning the theme is justified. Besides, since it is a relatively new subject, the exploration tends to bring new results.

Following this path of study, this work appears and presents the following research question: does the combination of filters in financial time series improve its forecasting capability?

The general goal of this study is to perform a comparative analysis of the combined use of the *wavelet* and *Kalman* filters, along with the forecasting models for the most renowned financial time series, which are the GARCH models and the neural networks ones, in order to check which one produces the best future forecast.

Therefore, this work aims to analyze the effect of the Double application of a space and state filter, associated with a high and low-frequency decomposition criteria in different occurrence orders, besides comparing the quality of the forecasts done separately by each hypothesis.

## 2. LITERATURE REVIEW

The financial time series differ from the other time series because they present characteristics peculiar to their elements. Those characteristics, according to Enders (2004, p. 10), are that these series are not serially correlated, but dependent.

Besides, according to the author, such series show the presence of volatility clusters in their return logs. Such a fact allows these volatility groups to be defined in several ways, but they are not clearly observable, in the analysis of these series in which you desire to model the phenomena that creates them, so that forecast can be done afterwards.

In this analysis, the variance measure of a log return, for a given period of time, depends on the past return logs, besides other elements unknown up to that time, so that its conditional variance does not coincide with the total variance of the series (called unconditional variance).

Enders (2004, p. 32) states that the models of the ARCH and GARCH family are the most common ones to be used in the modeling of the conditioned variance, even though other models can be adjusted.

The neural network models, or more precisely, artificial neural networks, are models of parallel processing distributed, formed by units of simple adjustment, which have natural tendency to store experimental knowledge and make it available for the use (Haykin, 2001, p. 28).

According to Oliveira (2003, p. 89), the reason for the use of a neural network is very simple and direct since it is possible to find an approach for modeling that improves the forecasts for financial time series, which are highly non-linear data, with a low amount of parameters, with the readiness of estimating them.

In the neural networks, what gives its power of prediction is exactly its parallel processing, that is, in which just one entry observed of the series is given and, with just one exit, which is obtained by the ponderation of the entry neurons and in the hidden layer, forms the process in a parallel way in such a way that it improves the network performance. (Zhang, Patuwo & Hu, 1998, p. 47).

The algorithm starts from an arbitrary configuration for the synaptic weight of neurons. In response to the statistical variations, the weights are adjusted in a continuous way in time. The calculations of those adjustments are completed within an interval of time convergence which is the sampling period. This process is known as adaptive filter (Haykin, 2001, p. 145).

This filtering process is the one which will constitute the refeeding bonds around the neurons, constituting an element of essential importance in the process of forecasting the time series through neural networks.

The denomination ‘filter’ comes from the communication engineering field and it means using a mechanism which enables the passage of components with frequencies in a given frequency band (Morettin & Toloi, 2004, p. 441).

This research understands that a filter works, in fact, as a process of data transmission which went through a “cleaning” process. This process happens by means of mathematical transformations in the time series which allow this purification of the series elements.

Also according to the authors, the main reason for the use of this kind of analysis in the treatment of time series is the fact that the spectrum provides a quite simple description of the effect of a linear transformation in the stationary process.

There are several kinds of filters which can be used in the treatment of time series. In this work, two filters will be used: the *wavelet* and the Kalman.

The filtering process through *wavelets* has the purpose of making the separation of the data of the original series in other two subseries by their frequency components. The *wavelets* are mathematical functions that increase the time interval of the data enabling each component to be allocated in its own scale (Misiti *et al.*, 2007, p. 4-3).

Polikar (1999, p. 7) states that this filtering process to allocate each component of the time series in its proper scale refers to the identification of the correspondent coefficients to each scale, whether high or low frequency, forming their new subseries. The inverse filtering, also called inverse-transformed, consists in applying the inverse filters in the decomposted signal and has the power to rebuild the original signal, rejoining both frequency bands.

The use of the *wavelets*, according to Gençay, Selçuk & Whitcher (2002, p. 10), is valid thanks to its capacity of decomposing a time series in scales which refer to the frequency domain for the time domain. What it is sought, therefore, is that, from a financial time series, their high and low-frequency representative subseries can be obtained in an attempt to soothe the variation effects of the financial market through the application of a filter by a wavelet function.

Gençay, Selçuk & Whitcher (2002, p. 133) say that a financial time series can be decomposed by a wavelet analysis, by a sequence of projections of father and mother wavelets, from the functions $\Phi$ and $\Psi$. The mother wavelet works as a window of finite covering which pursues the time series. The captivation of the high and low frequency points takes place by the translation and dilatation of the shape of the wavelet.

This representation for time series $y_t$ can be given from the functions previously defined by:

$$y_t = \sum_k a_{j,k}\Phi_{j,k}(t) + \sum_k d_{j,k}\Psi_{j,k}(t) + \sum_k d_{j-1,k}\Psi_{j-1,k}(t) + \ldots + \sum_k d_{1,k}\Psi_{1,k}(t) \qquad (1)$$

where $j$ is the number of the components and $k$ ranges from 1 to the number of quotients of the specific component. The quotients $a_{j,k}, d_{j,k}, \ldots, d_{1,k}$ are the coefficients of the transformed of wavelets given by the projections

$$a_{j,k} = \int \Phi_{j,k}(t).y_t dt \text{ , called approximation part} \quad (2)$$

and

$$d_{j,k} = \int \Psi_{j,k}(t).y_t dt \text{ , called detail part} \quad (3)$$

Thus, the explicit goal for the use of the wavelets, as described by Donoho & Tohnstone (1994, p. 439), is the noise reduction, also known as *denoising*, that the analysis of the wavelets performs in the choice of the coefficients which must be kept in order to preserve the information and consistency of the data of the original financial time series.

The Kalman filter, on the other hand, was introduced by Rudolph Emil Kalman and appears in the literature in 1960 when the author described his algorithm for the application solution of the filter of discrete data (Grewal & Andrews, 2008, p. 21).

According the same authors, the Kalman filter is a set of mathematical equations developed in the shape of a computational algorithm that form an iterative process, developed to perform future forecasts and estimate model variances for time series.

This means that, through the filtering process, by the Kalman filter, you start from an observable variable (a financial time series) and you can estimate another non-observable variable, called state variable, being able to estimate the past, present and future states by the value forecast.

This estimation, states Harvey (2001, p. 24), of the parameters which are unknown, happens by the maximization process of verosimilarity through forecast error decomposition, as already commented in this work.

Oliveira (2007, p. 74) mentions that the Kalman filter is extremely useful and has good results found in the literature due to its optimality and structure in providing formulations of easy implementation and processing in real time.

The structuring of the model in the shape of state spaces is made from a time series $\{y_t\}_{t=1}^{n} = \{y_1, y_2, \ldots, y_t, \ldots, y_n\}$ with $n$ elements. Such variables are called observable variables and represent a *n x 1* vector and associate with the state variables $x_t$ by a Markov[1] process creating an equation called measuring or observation equation:

$$y_t = A_t x_t + \varepsilon_t \quad (4)$$

With $t = 1, \ldots, T$ , being $A_t$ is a matrix $n \times m$, $\varepsilon_t$ a vector serially non-correlated with the zero average and covariance matrix $M_t$ and $x_t$ is a vector $m \times 1$ which contains the non-observable state variables.

As seen in the equation (4) above, the calculation of a state of a linear dynamic system in the time *t*, $x_t$ is calculated recursively starting from prior estimations of the state in the time *t-1*, $x_{t-1}$ and the new data provided in the entrance $y_t$ being unnecessary the storage of all the previous data to estimate the current state of the system.

It is understood as state of a system, a column vector $m \times 1$ containing variables which are of interest to the analyst. To do so, the Kalman filter is used, several times, combined with neural

[1] For more information on the Markov processes, see appendix B.

networks, in which these variables described as of the analyst´s interest are given by the weights of the neural network, aiming to find the best estimates for these variables.

Therefore, the applicable equations to the Kalman filter can be described:

- The Kalman gain

$$K_t = \frac{P_{t|t-1} A_t^T}{A_t P_{t|t-1} A_t^T + V_t} \quad (5)$$

- updating equation

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t \left[ y_t - A_t \hat{x}_{t|t-1} \right] \quad (6)$$

$$P_{t|t} = P_{t|t-1} - K_t A_t P_{t|t-1} \quad (7)$$

- forecasting equation

$$\hat{x}_{t+1|t} = B_{t+1|t} \hat{x}_{t|t} \quad (8)$$

$$P_{t+1|t} = B_{t+1|t} P_{t|t} B_{t+1|t}^T + Q_t \quad (9)$$

Thus, the equations of the Kalman filter keep themselves connected one to the other through the estimates of the state vector $\hat{x}$ and by the matrix of the error state correlation $P_t$. The updating equations make the correction of the $y_t$ for each step *t*, while the forecast equations do the future estimation for the instant *t+1*, a step ahead, before the next measure is made available in the system. Such process is repeated recursively up to the state convergence.

Consequently, if $\hat{x}_t$ is given as the current state of the system, then $\hat{x}_{t|t-1}$ refers to the state estimative for the step *t* given the knowledge of the variables in the prior step *t-1*, while $\hat{x}_{t+1|t}$ refers to the state estimative in the instant *t+1* given the information in the prior instant *t*.

The description of the Kalman filter is relied on the idea that both the noise of the measuring and transition equations follows a normal distribution. In other statistical words, it would be enough to say that the first two moments are enough to describe the states of the system, being $\hat{x}_t = E\left[x_t\right]$ and $P_t = E\left[\left(x_t - \hat{x}_t\right)\left(x_t - \hat{x}_t\right)^T\right]$. The estimator is many times said to be excellent for minimizing the error variance.

This iteration principle, described above, is the core of the Kalman filter, in which each new observation, in a given time instant, is made possible by the system, not only the state space vector but also the matrix of the state covariance are updated.

The combined use of filters is one of the main contributions of this work in the sense of checking the contribution that this combined use has to reduce the forecast quality measures for financial time series.

Concerning the forecast combined models, also called hybrid models, Souza (2008, p. 4) studied such forecast models of short, medium and long-term time series, confronting linear and non-linear models. For non-linear hybrid models, he considered the use of neural networks with radial basis function– RNs-RBF, with training based on the extended Kalman filter, that is, in the training phase; he used data filtered by the Kalman algorithm. For the linear models, he considered the Box, Jenkins & Reinsel model (1994, p. 33).

In the same line of the neural network use for forecasting, Oliveira (2007, p. 8) showed better performance for forecasting comparing the ARIMA-GARCH models, *feedforward* neural networks and neural networks trained with the Kalman algorithm and applied to the bonds of the finance, food, industry and service sectors.

Singhal & Wu (1989, p. 1188) have perhaps been the first ones to demonstrate the performance improvement of a supervised neural network which made use of the extended Kalman filter.

These authors demonstrated that their algorithm, although of great computational effort, converged with less iteration than the traditional methods of retroprogation. After this work, several other authors made simplifications and improvements in the algorithm and diversified its use in problems related to the fields of engineering, health and transportation such as the articles of Shah & Palmieri (1990, p. 42), Williams (1992, p. 244) and Puskorus & Feldcamp (1994, p. 288). In Brazil, it was also used a lot in recent researches such as Oliveira (2007, p. 80) and Pereira (2009, p. 103).

Concerning the use of the Kalman and the wavelet filters for the filtering process, Postalcioglu, Erikan & Bolat (2005, p. 951) assure that the Kalman filter removes disturbance or flaws of a time series (or a signal), using the initialization and transmission of the statistics of error covariance. They also comment that the application of the Kalman filter becomes impractical in models of great scale, such as demonstrated for the oscillator system. The alternative for this kind of system is the wavelet filter. In their research, they used the *Coiflet 2* wavelet, with 9 levels of decomposition and showed that the response of the wavelet filter is better when compared to the result of the Kalman filter concerning the noise suppression filtering called *denoising*. Even though the author has not done this work aiming to forecast data, it is clear the use of the improvement that the techniques can bring in case of noise suppression in time series.

Tak (1995, p. 43) made short-term forecasts for *Standard and Poor's* 500 (S&P 500) with daily data of the financial time series during the period from May, 1980 to December, 1990, based on the segregation theory of the time series decomposed by wavelets.

The author adopted two sublevels for the transformed application of wavelet, and compared the forecasts of the ARIMA models of feedforward neural networks. He sought to certify whether or not the decomposition through wavelets would bring improvement in the quality of the forecasts comparatively to the traditional models without decomposition.

Ma, Wong & Sankar (2004, p. 5824), on the other hand, also made forecasts with the ARIMA-GARCH models and used the wavelet decomposition for the S&P100 and obtained 3% forecast error reduction with the use of Haar´s primary wavelets. The results of Tak (1995, p. 44) were relatively worse with the use of Morlet´s wavelets and the Mexican hat and proved that the filtering through wavelets reduced the MAPE in just 9%.

In Brazil, the works of forecast of time series of Chiann (1997, p. 32), Homsy, Portugal & Araújo (2000, p. 10), Zandonade & Morettin (2003, p. 205), Lima (2004, p. 133), and Rocha (2008, p. 120) which used the wavelet decomposition aiming the improvement in the quality of the treatment of time series stand out. Of the authors mentioned above, Chiann (1997, p. 31) and Homsy, Portugal & Araujo (2000, p. 10) worked with wavelet analysis in non-financial time series.

Lima (2004, p. 133) applied a wavelet decomposition method for the behavior of the Ibovespa along with the ARIMA-GARCH models and recurrent neural networks and reached a 3%-error reduction for forecasts 21 steps ahead, with distinction to the neural networks and Deubechies 1 wavelets.

In the same line of thought, Rocha (2008, p. 121), mentioning Lima (2004, p.155), got good results combining forecasting models of exponential soothing and the ARIMA model reducing the error in 7.08% measured by the MAPE. He used the Deubechies 2 wavelet function.

In the several works analyzed, it can be seen that the use of filters presented great interdisciplinary application, which, in part, explains the strong presence of wavelets in the works researched in the last years. These conclusions enrich the relevance of the adoption of such filter tools to the forecast of financial series, applied to the Brazilian market. They also corroborated for the idea that it intends to obtain an improvement in the forecast quality, providing the investor of data so that he can take his decision and positioning in the future market.

## 3. METHODOLOGY

Concerning its objectives, it is an exploratory and descriptive research. It is exploratory because, according to Beuren (2006), it seeks to develop an overview of the great themes mentioned in the work in order to make it clear for the formulation of the research hypothesis. It is also descriptive since it seeks to describe the main characteristics present in a time series with grouping of volatility, statistical measures of the normality and non-linearity behavior that allow building and describing the model which rules its behavior.

For the consecution of the research, we chose the use of software, in which each one had its own role. The use of more than one kind of software was necessary since there has not yet been, in the academic community, software which fits all the steps described in this research.

At first, a series of return logs is formed and the descriptive statistics is analyzed aiming to check the stylized facts in the financial time series, such as asymmetry, kurtosis, volatility cluster grouping, normality, etc.

Then, the normality test is applied and the series stationary is tested, since its data must be stationary. In case this hypothesis is violated, succeeding differences in the time series are applied to make it stationary and, once more, the test is applied up to the moment it is checked.

From the stationary series, the presence of serial autocorrelations in the data is verified by applying the BDS independence test. In case it is independent, the time series can represent a behavior of random or chaotic stroll. These steps will not be verified since they are not the object of this research.

When the dependent behavior of the time series is obtained, the type of the dependence, whether linear or non-linear, is checked through the McLeod-Li test. If the series present linear behavior, linear models can be extracted. Otherwise, it is necessary to detect the non-linearity level, if in the average, in the variance or total, in which the best models are the neural networks ones.

The next step would then be to make the forecasts with the ARIMA-GARCH models and the recurrent neural networks for the series of pure log returns, that is, without any filtering process in the data. In the case of the neural networks, the recurrent networks proposed by Williams & Zipser (1989) were used. Concerning the data inserted in the recurrent neural network, the normalization process proposed by Azoff (1994) was used.

$$Z_t = \frac{R_t - \min\{R_t\}_{t=1}^{N}}{\max\{R_t\}_{t=1}^{N} - \min\{R_t\}_{t=1}^{N}} \in [0;1] \tag{10}$$

From the log returns, the extended Kalman filter is applied, since the series non-linearity presence has already been verified, and the forecasts must be made getting to the series below:

$$\{R_t\}_{t=1}^{n+t} = \left\{ R_1, R_2, \ldots, R_t, \ldots, R_n, \underbrace{EKFR_{n+1}, EKFR_{n+2}, \ldots, EKFR_{n+t}}_{previsões} \right\} \tag{11}$$

To make this type of forecast, first of all, the application of the Kalman filter on the data series is adopted. Afterwards, the wavelet filter is applied on the series filtered by the Kalman.

After the decomposition transformed by the wavelets is applied, the forecast is made making use of the recurrent neural networks. In this phase, as in the forecast made only with the neural networks, the wavelet must be chosen.

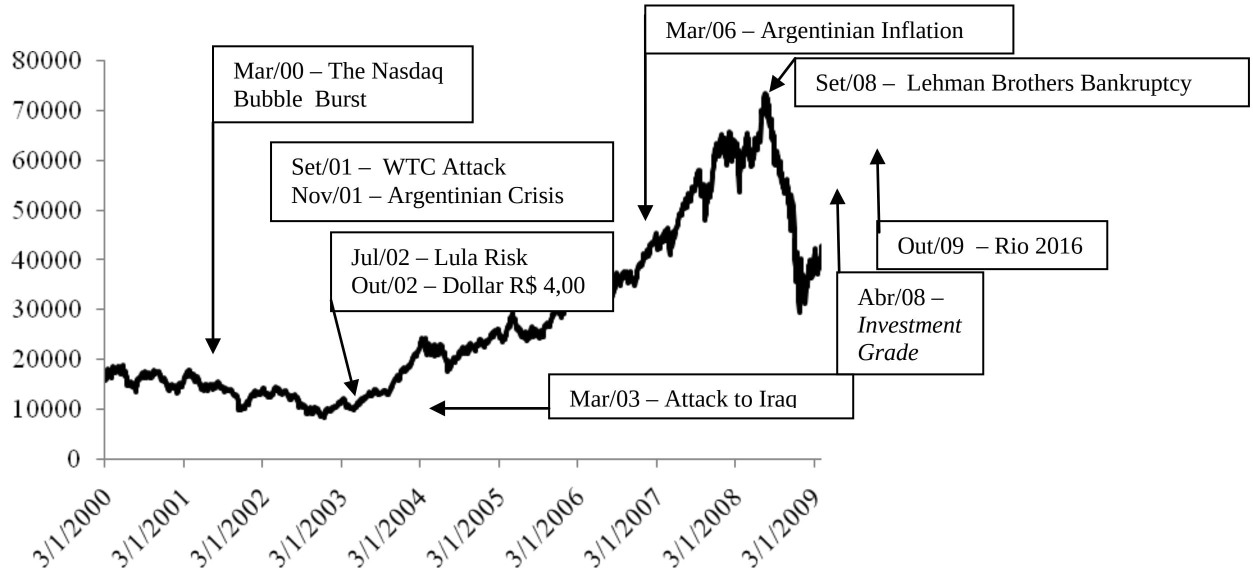
In this research, the reverse application was also considered, that is, decompose first by using the wavelet filter and then the Kalman filter.

The accuracy statistics of the forecasts are calculated according to Gooijer & Hyndman (2006, p. 457) which raised the main error measures in the works of time series forecasting. The term “accuracy” designates “how well-adjusted” the model is, that is, “how much” the model is able to reproduce the data already known.

As there is a time series with *n* periods of time, consequently there will be *n* error terms and, then, it is possible to calculate, for this research, the MAPE (*Mean Absolute Percentage Error*), Pearson´s linear correlation coefficient and the TIC – *Theil´s Inequality Coefficient*. This coefficient will always be between zero and one, being that zero indicates a perfect adjustment.

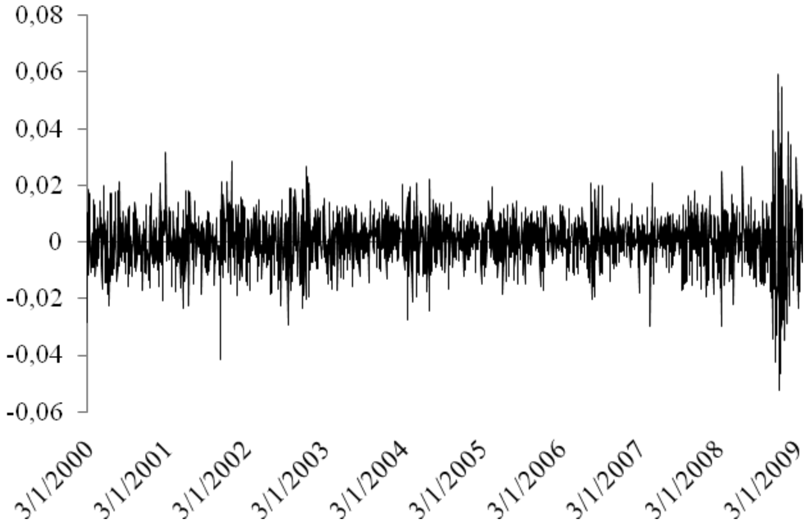
The analysis described previously was applied in the IBOVESPA series, taken from 01/03/2000, when it represented 16,930 points, up to 12/30/2009, when it presented 68,588 points. During this period, the valorization was of 305.13%. The exchange rate quotes were considered of closing in the daily period summing up 2,477 observations.

Picture 1 highlights the IBOVESPA with its exchange rate quote and some fact prominences worldwide weighty which had some influence in the index volatility behavior.




**Picture 1**
Nominal time series of the daily IBOVESPA
Source: BM&FBOVESPA, Sistema Enfoque

As seen, the IBOVESPA series is strongly volatile and that, constantly, suffers the influence of exogen variables such as crisis, rumors, which are not directly linked to its structure, but that influence in the oscillation of its returns. Picture 2, below, highlights the log returns.




**Picture 2**
Time series of log returns of the daily IBOVESPA

It is observed that the daily returns oscillates around zero presenting a variability, which depends on time, called volatility, with periods of high and low variability and days in which the return is abnormal, called *outlier*. Besides, several volatility clusters can be noticed, which take place due to the uncertainties of the market caused by economic and social phenomena as pointed out in Picture 1.

The statistics test indicates the JB value = 1241.254 with a p-value equal to zero. The trust level adopted was of 95%, being its probability of significance below 5%, indicating the rejection of the nule hypothesis which reveals that the series does not follow a normal distribution. It is emphasized that the same hypothesis would also be rejected to the significance level of 1%.

The descriptive statistics for the IBOVESPA series indicate for an average next to zero, which agrees with the classical finance theory that the return mean of a financial asset is always zero. The unconditional standard deviation features the mean oscillations of log returns. The excess of kurtosis with values above 3 is one of the main factors which can have led to the rejection of the nule hypothesis of normality.

We chose to apply the Augmented Dickey-Fuller test (ADF) since it is the most indicated and used one in the literature. As seen, the test checks if the series is stationary in the level or if it is necessary to make differences between it so that it becomes stationary.

The ADF test presented indicates that the p-value is below 5%, revealing the rejection of the nule hypothesis. Therefore, the log return series of the IBOVESPA is stationary.

The BDS indicates that, in every dimension, the log returns of the IBOVESPA do not follow an independent and identically distributed behavior, as observed in table 5.2, the low p-values, in which the nule hypothesis is rejected.

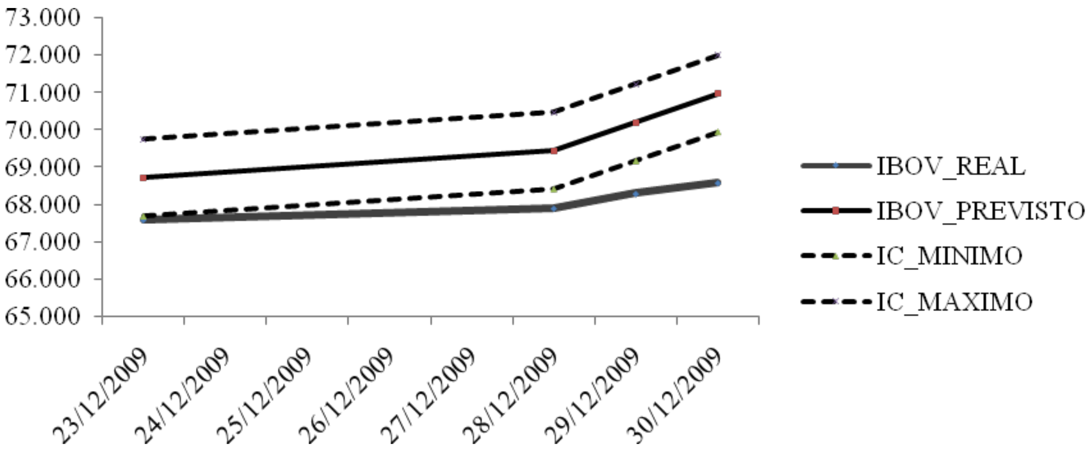
Thus, the series does not have an independent and identically distributed behavior. That means that there is time dependence between the log returns. In other words, the future returns are influenced by past returns. Concerning the fact of not being identically distributed, it refers to the fact they have time intervals with different probability distributions for the log returns. This feature shows the presence of non-linearity of the log returns.

The McLeod-Li test, for 5 discrepancies, rejected the nule hypothesis due to the fact that the p-value is below 5%. This indicates that the log return time series of the IBOVESPA has a non-linear behavior. The log return series raised to the power of 2 presents strong autocorrelation, which gives evidences that the generalized conditional heterocedasticidity autoregressive model can be used to improve the modeling of the series.

From now on, the non-linearity of the log returns is known. This fact discards any models other than the ones of non-linear characteristics. Therefore, the GARCH models and the recurrent time neural networks are applied for making the forecasts.

To perform the static forecasts with 4 steps ahead, first the ARIMA-GARCH conditional volatility model was estimated. The estimated model according to the analysis of the functions of the autocorrelation and partial autocorrelation was an AR(1)-GARCH(1,1) model. The selection of the order p, q of the model was done minimizing the information criterion of the AIC (*Akaike Information Criteria*).

Picture 3 illustrates the values forecasted, the real values for a static forecast of 4 steps ahead, as well as the confidence interval (CI) of 95% for the values forecasted. It is observed that the confidence interval is 100% out of the real values of the IBOVESPA.




**Picture3**
Graphic of the real and 4-step- ahead-forecasted IBOVESPA with static forecast for the ARIMA-GARCH model

It is worth pointing out that, in the graphics that follow, even though there are 8 dates, the forecast was of 4 steps ahead. The date of the forecast values was kept only for teaching matters, since there were holidays and bridges of periods without stock exchange from 12/24/2009 to 12/27/2009, returning to have stock exchange on 12/28/2009.

The low predictive capacity of this model, already identified above, presents the indicators of the data forecast analysis in table 1.

**Table 1**
Accuracy statistics of the AR(1)-GARCH(1,1) model for the IBOVESPA

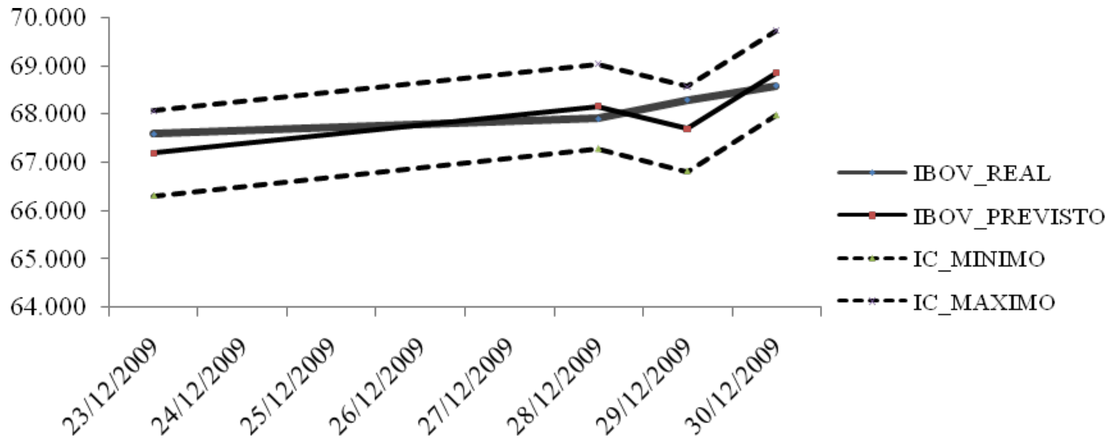
| Accuracy statistics of the predictive model | Values |
|---|---|
| TIC | 0.013089 |
| MAPE | 2.56% |
| Correlation | 0.227729 |

It is observed that the correlation between the real values and the forecast ones happens by the tendency movement verified in the real series and in the forecasted series. The TIC value shows a good adjustment, but a relatively high error.

Due to the low efficiency of the model of the GARCH family, the non-parametric models of the neural networks were also used. The neural networks represent an escape when it is unable to create models which are adequate to the reality of the data. This difficulty is found in the financial time series due to the continuous change in the volatility for short periods of time.

The recurrent neural network was used here, in which there is a connection of network refeeding between the processors in the same layer and also in different layers. The neural network adjustment had two phases: the first one is the training phase, in which 2,300 of the 2,477 values available were used, which corresponds to 92.85% of the data.  The second phase is the network performance, in which the other data for the test were used, for later use of the network. The network contained a neuron in the entrance layer, four neurons in the intermediate layer and one in the exit layer. The activation function used was the logistical function with 200 training periods.

The results found are in Picture 4.




**Picture 4**
Graphic of the real and 4-step- ahead-forecasted IBOVESPA with static forecast for the Recurrent Neural Networks

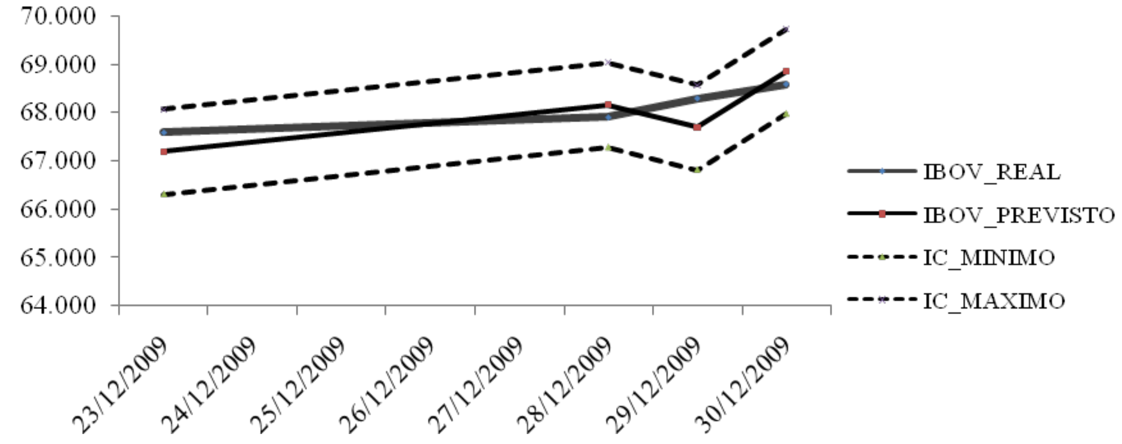
It can be observed that the adjustment and forecast quality improve sensibly with the use of the recurrent neural networks. The confidence interval comprises almost all the real values; just one was left out. Table 2 presents the indicators of forecasting analysis for the recurrent neural networks.

**Table 2**
Accuracy statistics with the use of recurrent neural networks for the IBOVESPA

| Accuracy statistics of the predictive model | Values |
|---|---|
| TIC | 0.005389 |
| MAPE | 0.84% |
| Correlation | 0.544006 |

Afterwards, it is believed that the combined use of the neural networks with forecasting filters tends to improve, since the filters would have the capacity to dilute the turbulence of the returns.

The first filter applied was the *wavelets,* as studied in Lima (2004). The results can be seen in Picture 5.  The wavelet chosen was the Deubechies 1 wavelet, since it is a primary wavelet.




**Picture5**
Graphic of the real and 4-step- ahead-forecasted IBOVESPA with static forecast for the Recurrent Neural Networks with wavelet filter

The results point to an improvement in the quality of the values forecasted and consequently the improvement in the forecasting indicators as it can be seen in table 3.

**Table 3**
Accuracy statistics with the use of recurrent neural networks for the IBOVESPA with the wavelets filter.

| Accuracy statistics of the predictive model | Values |
|---|---|
| TIC | 0.004417 |
| MAPE | 0.83% |
| Correlation | 0.705138 |

It can be seen that the forecasting adjustment was better, as given by the smallest TIC. The mean error of the forecasts reduced 0.01% and the correlation between the forecasted and real values increased significantly.

Thus, the process of the use of the wavelet filter helped with the error reduction in the forecasts. Another decomposition process of the time series for forecasting use is through the structural models. The model applied herein is the one of state spaces whose advance happened from Harvey´s (2001) work, in which his theory was described.

A state space model is built, based on the independence of the future event of the process in relation to its past state, since there is the present state. That is, the information about the past is inserted in the state of the process. The state of a process means finding the smallest number of independent variables so that, from knowing those variables at the initial instant, the behavior of the system for future instants is determined.

The representation of this state is done by two dynamic equations, in which one of them is the equation that indicates the values observed of the process obtained in function of the state vector and, the other one is the transition equation which will indicate the dynamic evolution of the not-observed state vector. (Souza, 1989, p. 54).

As seen, these models suppose that the typical movements stand being decomposed in non-observable portions such as, for example, tendency, seasonality, cyclical part and random part (error). The contribution of this process is that each component can be interpreted directly, due to the way the model is estimated.

The tool used for the estimation and forecast by structural models is the Kalman filter, which estimates forecasting equations and the updating of the positions in every instant of time.

It is worth pointing out that, according to Harvey (2001, p. 231), the variables do not necessarily represent measures of physical quantity, besides not being the only one. As the estimated state variables are independent, they can not be expressed as algebraic functions of other state variables.

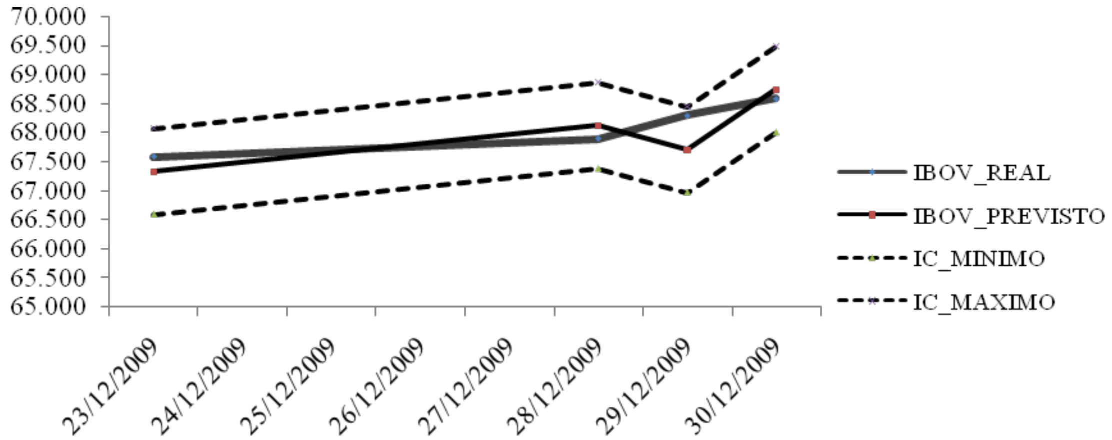
In table 4 are the results obtained according to the use of the structural models with the Kalman filter.

**Table4**
Accuracy statistics with the Kalman filter for the IBOVESPA

| Accuracy statistics of the predictive model | Values |
|---|---|
| TIC | 0.005505 |
| MAPE | 0.86% |
| Correlation | 0.598134 |

It can be observed that the forecasting adjustment by the Kalman filter were not higher than the ones obtained by the use of neural networks and wavelets. The error forecasts increased and the correlation with real values decreased, being even more efficient than the use of just neural networks.

Picture 6, which follows, illustrates the forecasts made.




**Picture 6**
Graphic of the real and 4-step- ahead-forecasted IBOVESPA with the static forecasting for the Kalman filter

The separated use of the Kalman and wavelet filters identified that the wavelets with the use of recurrent neural networks presented better adjustment and better quality in the static forecasts, even though it is highlighted that the Kalman filter was more efficient than the econometric models of the GARCH family.

Now, there is the following question: what would happen if the filtering techniques were combined? That is, if the Kalman filter was used and the wavelets were then used in its filter to make the forecasting with recurrent neural networks, since the networks improve the non-linear adjustment of the data or whether the reverse would bring better results.

The results of this combination are seen below.

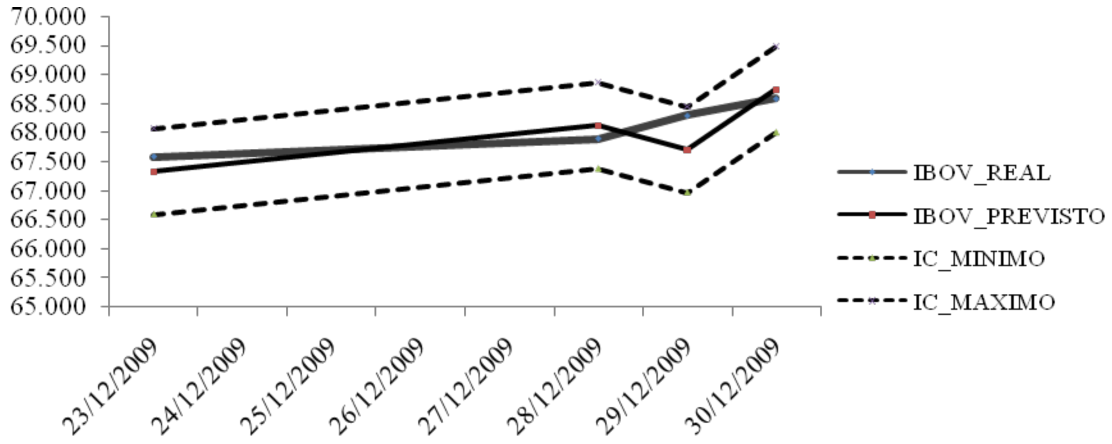
Table 5 illustrates the combined application of the 4-step ahead static forecasts, making use of Kalman first algorithm and, on it, the wavelet filter with decomposition in one level is applied.

**Table 5**
Accuracy statistics combining recurrent neural networks first with the Kalman filter and then wavelets for the IBOVESPA

| Accuracy statistics of the predictive model | Values |
|---|---|
| TIC | 0.004546 |
| MAPE | 0.72% |
| Correlation | 0.671659 |

The adjustment was better obtained, as well as the error statistics of the forecasts. The correlation is statistically more efficient than all the models verified up to here.

The forecasting graphic is shown in the Picture 7 below:




**Picture 7**
Real and 4-step ahead forecast IBOVESPA with static forecasting through the use of recurrent neural networks with the Kalman´s filter and then the wavelets.

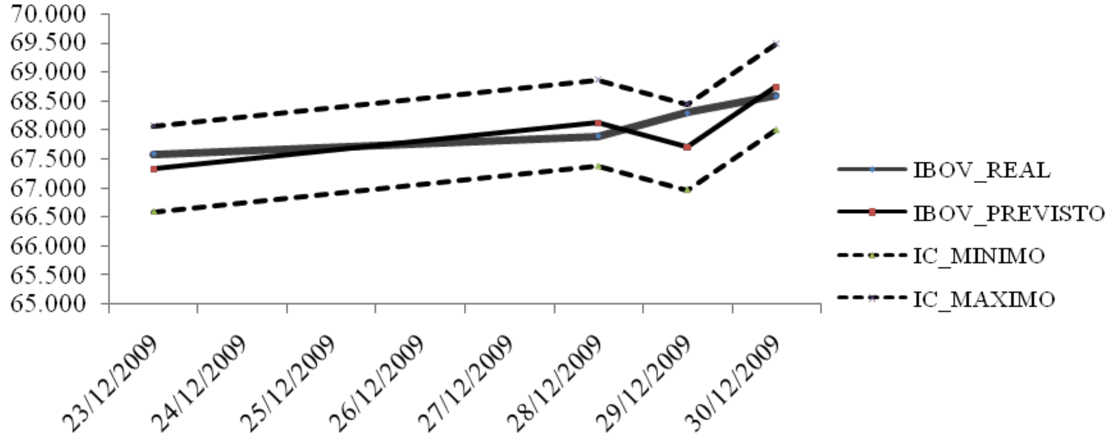
Then, the application of the combined use of the techniques to the reverse, that is, first the wavelet decomposition in a level on each of the forecasts with the Kalman filter was applied. The results are presented in table 6, below:

**Table 6**
Accuracy statistics combining recurrent neural networks with the wavelet filter first, and the Kalman filter afterwards for the IBOVEPA

| Accuracy statistics of the predictive model | Values |
|---|---|
| TIC | 0.005321 |
| MAPE | 0.93% |
| Correlation | 0.563688 |

The reverse of the combined technique of the use of filters does not improve the quality and adjustments of the forecasts as seen in the statistics above.

The forecasting graphic is shown in the Picture 8 below:




**Picture 8**
Real and 4-step ahead forecast IBOVESPA with static forecasting through the use of recurrent neural networks with the wavelet filter and the Kalman filter afterwards.

Summarizing the whole analysis process, it can be noticed that the combined use of the filtering techniques brought relative benefit in the forecasting adjustment and in its quality. Table 7, which follows, summarizes all the measures, remembering that the 21-day IBOVESPA volatility was of 12.71%.

**Table 7**
Summary of the forecasting statistics for the IBOVESPA

| Measure | ARIMA GARCH | RN-RECORR | RN-REC_WAV | Kalman´s Filter | FK_WAV | WAV_FK |
|---|---|---|---|---|---|---|
| TIC | 0,013089 | 0,005389 | 0,004417 | 0,005505 | 0,004546 | 0,005321 |
| MAPE | 2,56% | 0,84% | 0,83% | 0,86% | 0,72% | 0,93% |
| CORREL | 0,227729 | 0,54406 | 0,705138 | 0,598134 | 0,671659 | 0,563688 |

Therefore, the use of the Kalman filter, with afterdecomposition of wavelets and forecast with neural networks brought more benefits than all the other model combinations.

## CONCLUSION

Results with the use of just one filter had already been present in the literature, which attested the improvement in error reduction. And concerning the questions sought, support was found in the literature which had already been pointing to the need of improvement in the volatility filters together with the techniques already developed and tested.

The *background* raised pointed to the existence of two main filters: the Wavelets and Kalman. Such filters were applied in several series with the most distinct objectives: from tests in graphic images to even forecasting data such as sea currents, vehicle flow as well as in the construction of forecasting models of time series. And the main conclusion of the works reviewed was that the use of filters brings some improvement in the resulting intended.

Concerning the main results of this research, we came to check that the use of the filtering techniques can, indeed, reduce the error in the forecasts only in the IBOVESPA series which presented high volatility in the period.

In the case, the wavelets were the ones that could reduce the error with the use of recurrent neural networks. Such a result has the support in the existent and reviewed literature in this research. Thus, the first one of the specific objectives which were set was accomplished.

In these conditions, the forecasting models of time series with volatility filters suffer influence in the quality of the data used for the forecast. One of the key elements in this process can have been the strong influence of the economic crises in the period considered, as distinguished for the IBOVESPA. Such windows of study can be incorporated for the purpose of revealing whether this behavior would be present for series with more refined behaviors or that did not suffer pressure of macroeconomic variables, as well as the influence of the kinds of wavelets.

Finally, it is believed, in a very positive way, that the path to be followed by the forecasting quantitative models is really the combined use of modeling techniques, noise reduction, data segmentation, so that together each one can give their contribution for such a volatile and uncertain financial market. Thus, the knowledge of the peculiarities of the models and techniques will allow the researches to advance. Anyway, it keeps being a field to be followed and refined again in researches in the field of quantitative methods applied to finances.

## LITERATURE REFERENCES
## REFERÊNCIAS BIBLIOGRÁFICAS

Aiube, Fernando Antonio Lucena (2005). Modelagem dos preços futuros de commodities: abordagem pelo filtro de partículas. Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro. Tese de Doutorado.

Aminghafari, Mina, & Poggi, Jean-Michel (2007). Forecasting time series using wavelets. *International Journal of Wavelets Multiresolution and Information Processing*, 5, 5:709-724.

Azoff, E. Michael (1994). *Neural network time series forecasting of financial markets*. John Wiley & Sons, Inc., Hoboken, NJ, 212 p.

Ariño, Miguel A. (1995). Time series forecasts with *wavelet*: an application to car sales in the spanish market. ISDS – Institute of Statistics & Decision Sciences, Duke University. Discussion Paper.

Beuren, Ilse K. (2006). *Como elaborar trabalhos monográficos em contabilidade.* Atlas, São Paulo, 3 ed., 195 p.

Box, George E. P. Jenkins, Gwilym. M., & Reinsel, Gregory C. (1994). *Time series analysis: forecasting and control.* Prentice Hall, New York, 3. ed., 592 p.

Chiann, Chang (1997). Análise de ondaletas em séries temporais. IME – Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo. Tese de Doutorado.

Corsini, Flávio Pinheiro, & Ribeiro, Celma de Oliveira (2008). Dinâmica e previsão de preços de commodities agrícolas com o filtro de Kalman. Anais do XXVIII Encontro Nacional de Engenharia de Produção, Rio de Janeiro.

Donoho, David L., & Johnstone, Iain M. (1994). Ideal spatial adaptation by *wavelets* shrinkage. *Biometrika*, 81, 3:425-455.

Enders, Walter (2004). *Applied econometric time series.* John Wiley & Sons, Inc., Hoboken, NJ, 2. ed., 480 p.

Gençay, Ramazan, Selçuk, Faruk, & Whitcher, Brandon J. (2002). *An introduction to wavelets and other filtering methods in finance and economics.* Academic Press, San Diego, CA, 359 p.

Gooijer, Jan G. de, Hyndman, Rob J. (2006). 25 years of time series forecasting. *International Journal of Forecasting*, 22:443-473.

Granger, Clive W. J. (1992). Forecasting stock market prices: lessons for forecasters. *International Journal of Forecasting*, 8:3-13.

Grewal, Mohinder S., & Andrews, Angus P. (2008). *Kalman Filtering: theory and practice using MATLAB.* John Wiley & Sons, 3. ed., 575 p.

Harvey, Andrew C. (2001). *Forecasting, structural time series models and Kalman filter*. Cambridge University Press, Cambridge, 554 p.

Haykin, Simon S. (2001). *Redes neurais: princípios e prática*. Bookman, Porto Alegre, 2. ed., 900 p.

Homsy, Guilherme V., Portugal, Marcelo S., & Araújo, Jorge P. (2000). Ondaletas e previsões de séries de tempo: uma análise empírica. Anais do XXII Encontro Brasileiro de Econometria, Campinas.

Kalman, Rudolph Emil (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME – Journal of Basic Engeneering*, 82:35-45.

Lima, Fabiano Guasti (2004). Um método de análise e previsão de sucessões cronológicas unidimensionais lineares e não-lineares. Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo. Tese de Doutorado.

Ma, Irwin, Wong, Tony, Sankar, Thiagas. (2004). Forecasting the volatility of a financial index by wavelet transform and evolutionary algorithm. *IEEE International Conference on Systems, Man & Cibernetics*, 6:5824-5829.

Misiti, Michel, Misiti, Yves, Oppenheim, Georges, & Poggi, Jean-Michel (2007). *Wavelets and their applications*. ISTE Ltd., London, 330 p.

Morettin, Pedro Alberto, & Toloi, Clélia Maria de C. (2004). *Análise de séries temporais*. Edgard Blucher, São Paulo, 535 p.

Oliveira, Mauri Aparecido de (2007). Aplicação de redes neurais artificiais na análise de séries temporais econômico-financeiras. Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo. Tese de Doutorado.

Polikar, Robi (1999). The wavelet tutorial. Disponível em http://users.rowan.edu/~polikar/WAVELETS/WTtutorial.html. Acesso em: 10/01/2010.

Postalcioglu, Seda, Erkan, Kadir, & Bolat, Emine Dogru (2005). Comparison of Kalman filter and wavelets filter for denoising. *IEEE Neural Networks and Brain*, 3:951-953.

Puskorius, Gintaras Vincent, & Feldkamp, Lee Albert (1994). Neuro control of nonlinear dynamical systems with Kalman filter trained recurrent networks. *IEEE Transactions on Neural Networks*, 5:271-297.

Rocha, Vanessa Bueno da (2008). Uma abordagem de *wavelets* aplicada à combinação de previsões: uma análise teórica e experimental. Departamento de Matemática, Universidade Federal do Paraná. Dissertação de Mestrado.

Shah, Samir, & Palmieri, Francesco (1990). A fast, local algorithm for training feedforward neural networks, Proceedings of International Joint Conference on Neural Networks, San Diego, CA (USA), 3: 41-46.

Singhal, Sharad, & Wu, Lance (1989). Training feed-forward networks with extended Kalman filter. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1187-1190.

Souza, Reinaldo Castro (1989). Modelos estruturais para previsão de séries temporais: abordagens clássica e bayesiana. 17º Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro.

Souza, Rodrigo Clemente Thom de (2008). Previsão de séries temporais utilizando rede neural treinada por filtro de Kalman e evolução diferencial. Pontifícia Universidade Católica do Paraná, Curitiba. Dissertação de Mestrado em Engenharia de Produção.

Tak, Bongnam (1995). A new method for forecasting stock prices using artificial neural network and ondaleta theory. Universidade de Pensilvânia, Estados Unidos. Tese de Doutorado.

Ukil, Abhisek, Zivanovic, Rastko (2001). Adjusted Haar wavelet for application in the power systems disturbance analysis. *Digital Signal Processing*, 18:103-115.

Williams, Ronald J. (1992). Training recurrent networks using the extended Kalman filter. Proceedings of International Joint Conference on Neural Networks, Baltimore, MD (USA), 4: 241-247.

Williams, Ronald J., & Zipser, David (1989). A learning algorithm for continually running fully recurrent neural networks. *Neural Computation*, 1: 270-280.

Zhang, Guoqiang, Patuwo, B. Eddy, & Hu, Michael Y. (1998). Forecasting with artificial neural networks: the state of the art. *International Journal of Forecasting*, Kent, OH, (USA), 14:35–62.

Zandonade, Eliana, & Morettin, Pedro Alberto (2003). Wavelets in State Space Models. *Applied Stochastic Models in Business and Industry*, 15:199-219.