Vratislava Mošová
Wavelets and prediction in time series

In: Jan Chleboun and Petr Přikryl and Karel Segeth and Jakub Šístek and Tomáš Vejchodský (eds.): Programs and Algorithms of Numerical Mathematics, Proceedings of Seminar. Dolní Maxov, June 8-13, 2014. Institute of Mathematics AS CR, Prague, 2015. pp. 156--162.

Persistent URL: http://dml.cz/dmlcz/702678

Programs and Algorithms of Numerical Matematics 17
J. Chleboun, P. Přikryl, K. Segeth, J. Šístek, T. Vejchodský (Eds.)
Institute of Mathematics AS CR, Prague 2015

WAVELETS AND PREDICTION IN TIME SERIES

Vratislava Mošová

Moravian College Olomouc
tř. Kosmonautů 1, 779 00 Olomouc
Czech Republic
vratislava.mosova@mvso.cz

Abstract

Wavelets (see [2, 3, 4]) are a recent mathematical tool that is applied in signal processing, numerical mathematics and statistics. The wavelet transform allows to follow data in the frequency as well as time domain, to compute efficiently the wavelet coefficients using fast algorithm, to separate approximations from details. Due to these properties, the wavelet transform is suitable for analyzing and forecasting in time series. In this paper, Box-Jenkins models (see [1, 5]) combined with wavelets are used to the prediction of a time series behavior. The described method is demonstrated on an example from practice in the conclusion.

1. Introduction

It is possible to get the first impression of a time series behavior from the line graph. However, the conclusions received are highly subjective. More accurate information can be provided for instance by the Box-Jenkins methodology. Box-Jenkins models use the fact that every time series $\{y_t \mid t = 1, \ldots, T\}$ is a realization of some stochastic process. Because such models are based on the stochastic nature of time series, correlations have important place in drawing them up. A prediction for the time series is then created on the basis of the mathematical model received. This paper deals with linking wavelets and Box-Jenkins models. The ability of wavelets to decorrelate data is then used to specify forecast in time series.

The contribution is divided into following parts: The description of standard Box-Jenkins models built is given in Section 2. Wavelets and their usage in forecasting time series are discussed in Section 3. The procedures described are presented in the example in Section 4.

2. Box-Jenkins models

The Box-Jenkins models are constructed for the stationary time series. It means that the mean value and the variance function are constant, the correlation and the covariation functions depend only on the time distance of random variables.

A special case of the stationary process is the series $\{a_t\}$ of uncorrelated random variables with constant mean value and constant variance function that is called the white noise. In what follows, suppose that every time series consists of an unsystematic component $\{a_t\}$ and of systematic components such as a trend, a seasonal component or a cyclical component.

The stationary Box-Jenkins process is denoted by ARMA(p, q). It is a process composed of an autoregressive process of order p and a process of moving averages of order q. The mathematical model of it is

$$y_t = \Phi_1 y_{t-1} + \cdots + \Phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}, \tag{1}$$

where $\Phi_1, \ldots, \Phi_p$ are parameters of the autoregressive part and $\theta_1, \ldots, \theta_q$ are parameters of the moving averages part of the model. This model can be rewritten using a backshift operator $B^i y_t = y_{t-i}$ in the form

$$\Phi_p(B) y_t = \theta_q(B) a_t, \tag{2}$$

where $\Phi_p(B) = (1 - \Phi_1 B - \cdots - \Phi_p B^p)$ and $\theta_q(B) = (1 - \theta_1 B - \cdots - \theta_q B^q)$.

The process AR(p) is stationary in case when roots of the polynomial $\Phi_p(B)$ lie outside the unit circle. The process MA(q) is invertible, when roots of the polynomial $\theta_q(B)$ lie outside the unit circle. But stationary models nearly absent in the economic practice. Fortunately, it is possible to convert a nonstationary model to a stationary one.

If d roots of the polynomial $\Phi_p(B)$ lie on the unit circle, the process is not stationary but it has a stochastic trend. Such process is denoted $I(d)$ and it is called the integrated process of order d. Its model has the form

$$(1 - B)^d y_t = a_t. \tag{3}$$

This model can be converted to a stationary one if d-times differentiation is applied to it. The combination of the stationary and the integrated process leads to the nonstationary process ARIMA(p, d, q),

$$\Phi_p(B)(1 - B)^d y_t = \theta_q(B) a_t. \tag{4}$$

When a seasonal oscillation with period s occurs in a time series, it is necessary to capture the dependence among the components of the original series and also the dependence among the components, which correspond to the different seasons. The seasonal model SARIMA$(p, d, q)(P, D, Q)$,

$$\Phi_P(B^s)\Phi_p(B)(1 - B)^d(1 - B^s)^D y_t = \theta_q(B)\theta_Q(B^s) a_t, \tag{5}$$

where P, D and Q are seasonal parameters of process, is used in this case. The left-hand side of (5) supplies the dependence inside the period and the right-hand side represents only the seasonal dependences.

Constructions of Box-Jenkins models are especially based on the information that is obtained from the correlograms, i.e. the graphs of values of the autocorrelation function ACF and the partial autocorrelation function PACF.

For stationary time series, the residual ACF is defined through autocorrelations with the delay k,

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \tag{6}$$

where $\gamma_k = E[(y_t - \mu)(y_{t-k} - \mu)]$. The residual ACF indicates the range of the linear dependence between y_t and y_{t-k}.

The partial autocorrelation with delay k is defined through partial regressive coefficients Φ_{kk} in the autoregression of order k

$$y_t = \Phi_{k1} y_{t-1} + \Phi_{k2} y_{t-2} + \cdots + \Phi_{kk} y_{t-k} + a_t, \tag{7}$$

where a_t is a value that is uncorrelated with $y_{t-1}, y_{t-2}, \ldots, y_{t-k}$. The function PACF gives the information cleaned from the influence of the variables $y_{t-1}, y_{t-2}, \ldots, y_{t-k}$.

First estimation properties of a given time series are based on the line graph, periodogram, ACF and PACF. Peaks in the periodogram indicate the presence of seasonal oscillations. It means that it is necessary to work with a seasonal model. Values greater than 1 in ACF and PACF mean that the series is not stationary. In this case, it is necessary to consider an integrated model. Removal of non-stationarity in the variance can be achieved by the Box-Cox transformation.

The model chosen has to be verified, i.e. monitored whether autocorrelation unsystematic components are zero by the Box-Pearson test and how good the received estimates of the parameters μ, ϕ, θ are by t-tests.

The model selected is the basis for the estimate of further development of the series. The calculation of the forecasted value y_{T+h} is done by means of the conditional mean value $E(y_{T+h} \mid y_{T-1}, y_{T-2}, \ldots)$.

3. Wavelet transform

The wavelet transform is a useful tool for detecting local properties and investigating nonstationary data. It is defined using wavelets, which form a basis in the space $L^2(R)$. Multiresolution analysis (MRA) is the most commonly used method to the construction of such basis.

During MRA, subspaces $V_j \subset L^2(R)$ are constructed with properties

1) $V_j \subset V_{j+1}$,
2) there exists $\varphi \in V_0$ such that $\{\varphi_{0,k}\}$, where $\varphi_{0,k}(x) = \varphi(x-k)$, is orthogonal and complete in $L^2(R)$,
3) $f(x) \in V_0$ if and only if $f(2^j x) \in V_j$,
4) $\cap_j V_j = \{0\}$,
5) $\overline{\cup_j V_j} = L^2(R)$.

When $\{V_j\}$ is MRA with scaling function φ, then there exists a scaling vector $\mathbf{u} = (\dots, u_{-1}, u_0, u_1, \dots)$ such that

$$\varphi(x) = \sqrt{2} \sum_{k \in Z} u_k \varphi(2x - k). \tag{8}$$

In this case, the associated wavelet ψ is defined by the formula

$$\psi(x) = \sqrt{2} \sum_{k \in Z} v_k \varphi(2x - k),\ v_k = (-1)^k u_{1-k}. \tag{9}$$

It follows from the MRA that there exists a subspace $W_j \subset L^2(R)$ such that

$$V_{j+1} = V_j \oplus W_j. \tag{10}$$

The subspaces V_j and W_j can be generated by means of dilations and traslations of the functions φ and ψ. It holds

$$W_j = \overline{\mathrm{span}\{\psi_{j,k}(x)\}}, \quad \text{where} \quad \psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k), \tag{11}$$

$$V_j = \overline{\mathrm{span}\{\varphi_{j,k}(x)\}}, \quad \text{where} \quad \varphi_{j,k}(x) = 2^{j/2} \varphi(2^j x - k). \tag{12}$$

Moreover, it can be seen that

$$V_{j+1} = V_J \oplus W_J \oplus W_{J+1} \oplus \dots \oplus W_j. \tag{13}$$

It means that it is possible to expand every function $f \in L^2(R)$ into the series

$$f(x) = \sum_{k \in Z} y_{J,k} \varphi_{J,k} + \sum_{j=J}^{\infty} \sum_{k \in Z} x_{j,k} \psi_{j,k}. \tag{14}$$

The scaling coefficients $y_{j,k}$ and the wavelet coefficients $x_{j,k}$ are calculated as inner products. It holds

$$y_{j,k} = \langle f, \varphi_{j,k} \rangle, \qquad x_{j,k} = \langle f, \psi_{j,k} \rangle. \tag{15}$$

It follows from (15), (11), (12), (8), (9) and (14) that

$$y_{j,k} = \frac{1}{\sqrt{2}} \sum_l u_l y_{j+1,2k+l}, \quad x_{j,k} = \frac{1}{\sqrt{2}} \sum_l (-1)^l u_{1-l} y_{j+1,2k+1}, \tag{16}$$

$$y_{l+1,k} = \frac{1}{\sqrt{2}} \sum_m u_{m-2k} y_{l,m} + \frac{1}{\sqrt{2}} \sum_m (-1)^m u_{1-m-2k} x_{l,m}. \tag{17}$$

Computation of the wavelet coefficients is realized by means of the Mallat algorithm. The relations (16) and (17) are the basis of this algorithm. First, approximations and details are computed from the data given (the decomposition phase). The approximations correspond to the trend and the details correspond to random components of the time series. The process can be repeated more times. The wavelet coefficients can be adapted or not. In the end, modified or original data are obtained from this set of values (the reconstruction phase).

In the following example, the wavelet transform is used to construct the prediction for the time series given. First, a decomposition into approximations and details is done. Then the proper ARIMA model and prediction are found for each of these parts. The resulting prediction is a sum of values from these two partial predictions.

4. Example

The monthly values of CPI inflation in the Czech Republic in the years 2004–2014 are given in Table 1. Find a suitable ARIMA model for this series from January 2004 to December 2012. Make a forecast for the rest of the series using the ARIMA model and then using ARIMA model modified by wavelets. Compare the results received to each other.

	1	2	3	4	5	6	7	8	9	10	11	12
2004	0.3	0.5	0.8	1.	1.2	1.4	1.7	2.	2.2	2.5	2.7	2.8
2005	2.8	2.7	2.6	2.6	2.5	2.4	2.2	2.1	2.	2.	1.9	1.9
2006	2.	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.7	2.6	2.5
2007	2.4	2.3	2.2	2.2	2.1	2.1	2.1	2.	2.	2.2	2.5	2.8
2008	3.4	3.9	4.3	4.7	5.	5.4	5.8	6.1	6.4	6.6	6.5	6.3
2009	5.9	5.4	5.	4.6	4.1	3.7	3.1	2.6	2.1	1.6	1.3	1.
2010	0.9	0.8	0.7	0.6	0.6	0.6	0.8	0.9	1.1	1.2	1.4	1.5
2011	1.6	1.7	1.7	1.8	1.8	1.9	1.9	1.9	1.8	1.9	1.9	1.9
2012	2.1	2.2	2.4	2.6	2.7	2.8	2.9	3.1	3.2	3.3	3.3	3.3
2013	3.2	3.	2.8	2.7	2.5	2.3	2.2	2.	1.8	1.6	1.5	1.4
2014	1.3	1.1	1.									

Table 1: Inflation 2000–2014

Solution. On the basis of ACF and PACF, the original time series were modeled through ARIMA(3,2,1) model.

Further, the decomposition of the time series to approximations and details by using the Daubechies wavelet Db3 was done. The first order extrapolation was used to expand the data beyond boundary. This allowed receiving such approximation coefficients that are close to the values of the original time series.

In the next step, appropriate ARIMA models were selected for the approximations and for the details separately. The approximation coefficients are not identical with the time series values, because the information is lost when wavelet decomposition is made. Moreover, a small change of range (e.g. a truncation of the time series or an extension of the data beyond boundary) may affect the shape of the ARIMA model. Therefore, Box-Jenkins models are different for the original data and for approximations.

The approximations were modeled with the help of ARIMA(2,2,1) process and the details were modeled as ARIMA(2,0,1) process in this example. The prognosis for the time series was obtained by adding up the forecasts for approximation and details. Note that it is possible to realize prediction using approximations only and ignore details, when the details are detected like random noise.

Choice of ARIMA models affects the shape of the predictions. Adequacy of the models is assessed by means of corresponding graphs ACF and PACF. Comparison of the predictions for the next 15 months is shown in Figure 1. Comparison of the values received is presented in Table 2.

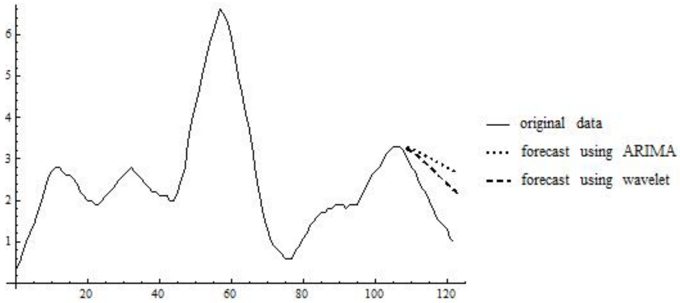

Figure 1: Comparison of predictions

	1	2	3	4	5	6	7	8
original	3.2	3.	2.8	2.7	2.5	2.3	2.2	2.
ARIMA	3.2789	3.24318	3.20814	3.16497	3.12164	3.07674	3.03056	2.98447
wavelets	3.24865	3.18837	3.12578	3.05159	2.97491	2.89764	2.81764	2.73695

	9	10	11	12	13	14	15
original	1.8	1.6	1.5	1.4	1.3	1.1	1.
ARIMA	2.93772	2.89095	2.84405	2.79705	2.75005	2.70301	2.65596
wavelets	2.65611	2.57463	2.49296	2.41125	2.32939	2.24748	2.16556

Table 2: Comparison of predictions

The root mean square error $RMSE = 1.07178$ in case of the ARIMA model and $RMSE = 0.78320$ in case of the model that uses the wavelet transform. It can be seen that the prognosis was improved by 36.8% when the wavelet modification was used.

5. Conclusion

The example has shown that the ARIMA model modification can lead to improved estimation of time series evolution. Note that wavelets can be used not only in forecasting non-stationary time series, but also to detect sudden changes, or to select cycles or fractal nature of time series.

6. Acknowledgement

This work was supported by project GAČR P403/12/1811: Unconventional managerials decision making methods development in enterprise economics and public economy.

References

[1] Artl, J., Artlová, M., and Rublíková, E.: *Analýza ekonomických časových řad s příklady*. VŠE, Praha, 2004.

[2] Najzar, K.: *Základy teorie waveletů*. Karolinum, Praha, 2004.

[3] Jansen, M. and Oonix P.: *Second generation wavelets and applications*. Springer-Verlag, London, 2005.

[4] Sebera, M. and Seberová, H.: Časové řady a wavelety. In: *XX. mezinárodní kolokvium o řízení osvojovacího procesu*, pp. 354–357. Vyškov, 2001.

[5] Siegel, A. F.: *Practical business statistics*. Elselvier, Oxford, 2012.