Research Article

Multi-Step-Ahead Forecasting Model for Monthly Anchovy Catches Based on Wavelet Analysis

Nibaldo Rodríguez, Claudio Cubillos, and José-Miguel Rubio

School of Informatics Engineering, Pontificia Universidad Católica de Valparaíso, 2362807 Valparaíso, Chile

Correspondence should be addressed to Nibaldo Rodríguez; nibaldo.rodriguez@ucv.cl

Received 29 July 2013; Revised 29 December 2013; Accepted 12 January 2014; Published 24 February 2014

Academic Editor: Hector Pomares

Copyright © 2014 Nibaldo Rodríguez et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a $p$-step-ahead forecasting strategy based on two stages to improve pelagic fish-catch time-series modeling by considering annual and interannual fluctuations for northern Chile ($18^\circ$–$24^\circ$S). In the first stage, the stationary wavelet transform is used to separate the raw time series into an annual component and an interannual component, whereas the periodicities of each component are obtained using the Morlet wavelet power spectrum. In the second stage, a linear autoregressive model is constructed to predict each component and the unknown $p$-next values are forecasted by the addition of the two predicted components. We demonstrate the utility of the proposed forecasting model on monthly anchovy-catches time series for periods from January 1963 to December 2007. Empirical results obtained for 10-month-ahead forecasting showed the effectiveness of the proposed wavelet autoregressive strategy.

1. Introduction

During the last decades, both the direct method and the iterative method have been generally used in literature to implement $p$-step-ahead forecasting model [1–6]. The iterative method iterates $p$ times the same one-step-ahead forecasting model to obtain the $p$ predicted values whereas the direct method estimates a set of $p$ forecasting models, each returning a forecast for the $r$th value; $r \in \{1, 2, \ldots, p\}$. In other words, $p$ models are calibrated from the time series (one for each horizon). On the other hand, multi-step-ahead forecasting models play an important role in the management of marine resources. However, the multi-step-ahead forecasting task is difficult due to factors such as accumulation of errors, reduced accuracy, increased uncertainty, and nonstationarity [1–6].

In recent years, linear regression model [7–9] and artificial neuronal networks (ANN) [10, 11] have been proposed for pelagic fishes forecasting model. The disadvantage of models based on linear regressions is the supposition of stationarity and linearity of the time series of pelagic species catches. Although ANN allow modelling the nonlinear behaviour of time series, they also have some disadvantages such as the stagnancy of local minima due to the steepest descent learning method. A multilayer percpetron neural network model for forecasting the anchovy and sardines catches of northern Chile was proposed by [10, 11], which reported a coefficient of determination between 82% and 87%. Noteworthy is that the anchovy and sardine are important pelagic fisheries resources for economic development of northern Chile.

In this paper, a $p$-step-ahead forecasting strategy based on wavelet analysis and linear autoregressive (AR) model is proposed to improve prediction accuracy of monthly anchovy catches in the coastal zone of northern Chile. The advantage of wavelet analysis is its ability to separate low frequency components and high frequency components from a nonstationary time series. Besides, each component is more regular than the raw time series, which can help improve the forecasting performance [12, 13]. On the other hand, the wavelet analysis was also selected due to its successful use in electricity market [12, 13], smoothing methods [14–17], financial market [18–20], and ecological time series modeling [21, 22]. Moreover, variability analysis at different time scales based on the wavelet power spectrum has shown that climatic oscillations such as the El Niño-Southern Oscillation significantly affect the abundance of marine species [23, 24].
Finally, the proposed $p$-step-ahead forecasting strategy is based on two stages. In the first stage, the raw time series is decomposed using the 3-level stationary wavelet transform (SWT) to separate both the annual component (AC) and the interannual component (IAC). The periodicity of each component at different time scales is detected and quantified using the global Morlet wavelet power spectrum. In the second stage, both the AC and IAC are forecasted using a linear AR model and the unknown $p$-next values are forecasted by the algebraic sum of the two predicted values.

The rest of this paper is organized as follows. Section 2 briefly describes the wavelet analysis. The $p$-step-ahead forecasting strategy is presented in Section 3. The experimental results and performance evaluation are presented in Section 4 followed by the conclusions in Section 5.

2. Wavelet Analysis

In this section we briefly present the continuous wavelet transform and the discrete stationary wavelet transform.

2.1. Continuous Wavelet Transform. The continuous wavelet transform (CWT) decomposes a continuous time signal $x(t)$ into a family of “daughter-wavelet” functions $\psi_{s,b}(t)$, which are generated by the dilatation and the translation of a mother wavelet function. The CWT coefficients of a real signal $x(t)$ are obtained for the different scales and translations as follows [25, 26]:

$$W(s, b) = \int_{-\infty}^{\infty} x(t) \psi^*_s (t - b) \, dt,$$

where the $*$ denotes the complex conjugate, $s$ is a dilatation (scale) factor that controls the width of the wavelet function, and $b$ is a translation factor controlling its location. The scaled and translated daughter-wavelet function is obtained as

$$\psi_{s,b}(t) = \frac{1}{\sqrt{s}} \psi \left( \frac{t - b}{s} \right), \quad s \in \mathbb{R}^+, \quad b \in \mathbb{R},$$

where $\psi(t)$ is the mother wavelet function and the choice of the wavelet function is not arbitrary. There are several considerations when making the choice of a wavelet function, for example, real versus complex wavelets, continuous versus discrete wavelets, and orthogonal versus redundant wavelets. However, all the wavelets share a general feature. That is, fast oscillations have good time resolution but lower frequency resolution, whereas low oscillations have good frequency but poor time resolution. In this paper the CWT has been implemented using the Morlet mother wavelet function defined as

$$\psi(t) = \pi^{-1/4} \exp(i \omega_0 t) \exp \left( -\frac{t^2}{2} \right),$$

where $\omega_0$ defines the frequency center of the wavelet and here is set equal to $\omega_0 = 6$, as it yields the function to have zero mean and be localized in both time and frequency space, as well as providing a good balance between time and frequency. Moreover, for the Morlet wavelet, the scale $s$ is inversely related to the Fourier frequency; that is, $f = 1/s$, which simplifies the interpretation of the wavelet analysis and one may replace the scale $s$ by the Fourier period $1/f$ [25, 26].

For discrete time series $x_n = \{x(n), n = 0, 1, \ldots, N - 1\}$ with uniform time step $\delta t$, its CWT is obtained as a $J \times N$ matrix, whose $(j, k)$ elements are given by [27, 28]

$$W(j, k) = \sqrt{\delta t} \sum_{n=0}^{N-1} x(n) \psi^* \left( \frac{(n-k) \delta t}{s_j} \right), \quad j = 0, 1, \ldots, J - 1,$$

where $s_j$ represents a set of scales, $k = 0, 1, \ldots, N - 1$ denotes the shifting index, $J$ determines the largest scale, and $N$ is the number of data points in the time series.

The global wavelet power spectrum GWPS of a discrete time series $x_n$ is calculated as [27, 28]

$$\text{GWPS}(s_j) = \frac{1}{N} \sum_{n=0}^{N-1} |W(j, k)|^2, \quad j = 0, 1, \ldots, J - 1,$$

where GWPS represents the cumulate variance contributed by each time scale and $|W(j, k)|^2$ denotes the local variance distribution of the time series in the time-scale plane.

2.2. Stationary Wavelet Transform. The SWT was independently proposed by several researchers and it is known in literature under a variety of names, such as the nondecimated wavelet transform, invariant wavelet transform, and redundant wavelet transform. The key feature is that it gives a better approximation than the discrete wavelet transform (DWT) since it is redundant, linear, and shift invariant [14–17]. The SWT is similar to the DWT [25, 26] in that the high-pass and low-pass filters are applied to the input signal at each level, but the output signal is never decimated. Instead, the filters are upsampled at each level of decomposition [14].

Now, we consider the following discrete time series $x_0 = x_n$ with $N = 2^J$ for some integer $J$. At the first level of SWT, the input signal $a_0$ is convolved with a low-pass filter $h_0$ of length $r$ to obtain the approximation coefficient $a_1$ and with a high-pass filter $g_0$ of length $r$ to obtain the detail coefficient $d_1$. That is,

$$a_1(n) = \sum_k h_0(n-k) a_0(k), \quad d_1(n) = \sum_k g_0(n-k) a_0(k),$$

because no subsampling is performed, and $a_1$ and $d_1$ are of length $N$ instead of $N/2$ as in the DWT case. At the next level of the SWT, the approximation coefficient $a_1$ is split into two parts using the same previous scheme, but with a new pair of filter $h_1$ and $g_1$, which are obtained by inserting a zero value between the elements of the filters used in the previous step.
Input: A time series \( x = x(n), n = 0, 1, \ldots, N - 1 \) of \( N \) samples
Output: Annual and inter-annual component

1. \( J \leftarrow 3 \)
2. \( W_n \leftarrow d_{b2} \)
3. \([a d] \leftarrow \text{swt}(x, f, Wn)\)
4. \( \text{Aux} \leftarrow \text{zeros}(J, N) \)
5. \( x_a \leftarrow \text{iswt}(\text{Aux}, d, Wn) \)
6. \( x_{ia} \leftarrow \text{iswt}(a, \text{Aux}, Wn) \)
7. \( \text{return} x_a, x_{ia} \)

Algorithm 1: Component separation.

(i.e., \( h_0 \) and \( g_0 \)). The general process of the SWT is continued iteratively for \( j = 1, \ldots, J - 1 \) and is given as

\[
\begin{align*}
    a_{j+1}(n) &= \sum_k h_j(n-k) a_j(k), \\
    d_{j+1}(n) &= \sum_k g_j(n-k) a_j(k),
\end{align*}
\]

where \( h_j \) and \( g_j \) are obtained by the upsampling operator. The upsampling operator inserts \( 2^j - 1 \) zeros between the elements of the filters \( h_0 \) and \( g_0 \), respectively. The SWT is fully defined by the choice of a pair of filters (i.e., \( h_0 \) and \( g_0 \)) and the number of decomposition step \( J \).

3. Multi-Step-Ahead Forecasting Strategy

3.1. Wavelet Preprocessing. At this stage, the SWT is used to extract both the AC and IAC by using the Daubechies low-pass/high-pass filter with four coefficients (Db2) and three levels of decomposition. The periodic behavior of each component was obtained by using the GWPS with a 95% confidence level [27, 28], which can be seen in Figures 2, 3, and 4. Algorithm 1 explains the component separation process considering the Db2 wavelet filter and \( J \)-levels of decomposition. In this algorithm, line (3) performs the separation of components by using 3-level SWT with Db2 wavelet filter, whereas line (5) and line (6) implement the reconstruction of the AC and the IAC using the inverse 3-level SWT.

3.2. Wavelet Autoregressive Forecasting. The proposed \( p \)-step-ahead forecasting strategy is based on direct method. The direct method is to learn \( p \) single output models, each returning a direct prediction of the future value \( x(n+p) \). In order to predict the future value \( x(n+p) \) firstly we separate the original time series into two components by using Algorithm 1. The first component \( x_a \) presents the annual variabilities and is characterized by fast dynamics, while the second component \( x_{ia} \) indicates interannual fluctuation and is characterized by slow dynamics. Therefore the proposed forecasting model will be the coaddition of two predicted values given by

\[
\hat{x}(n+\tau) = \hat{x}_a(n+\tau) + \hat{x}_{ia}(n+\tau) + e(n),
\]

where \( \tau \) represents the forecasting horizon and \( e(n) \) denotes the \( n \)th value of the residual component.

The \( \hat{x}_a(n+\tau) \) and \( \hat{x}_{ia}(n+\tau) \) are estimated by using a linear autoregressive function with exogenous inputs given by the following two equations, respectively:

\[
\hat{x}_a(n+\tau) = F_{a,\tau} [x_a(n), x_{ia}(n)] = \sum_{i=0}^{m-1} \alpha_i x_a(n-i) + \sum_{i=0}^{m-1} \alpha_{msi} x_{ia}(n-i),
\]

where \([x_a(n-i)]\) represents the endogenous regressor vector, whereas \([x_{ia}(n-i)]\) plays the role of exogenous regressor vector and \( m \) is the size of a time window,

\[
\hat{x}_{ia}(n+\tau) = F_{ia,\tau} [x_{ia}(n), x_a(n)] = \sum_{i=0}^{m-1} \beta_i x_{ia}(n-i) + \sum_{i=0}^{m-1} \beta_{msi} x_a(n-i),
\]

where \([x_{ia}(n-i)]\) represents the endogenous regressor vector and the exogenous regressor vector is denoted by \([x_a(n-i)]\).

In order to estimate the parameters \([\alpha_i]\) and \([\beta_i]\) the linear least square method is used. Now suppose a set of \( T \) training input-output samples; then we can perform \( T \) equations of the following forms:

\[
\begin{align*}
    X_a &= Z x_a, \\
    X_{ia} &= Z x_{ia},
\end{align*}
\]

where \( Z \) is the regressor matrix of \( T \) rows and \( 2m \) columns and \( \alpha \) and \( \beta \) are vectors of \( 2m \) rows and one column. Finally, the \( \alpha \) and \( \beta \) values are calculated by using the following equations:

\[
\begin{align*}
    \alpha &= Z^\dagger X_a, \\
    \beta &= Z^\dagger X_{ia},
\end{align*}
\]

where \((\cdot)^\dagger\) denotes the Moore-Penrose pseudoinverse [29].

4. Measures of Accuracy Applied in the Models Performance

In this paper, three criteria of accuracy are used to evaluate the estimation capabilities during the test phase of the evaluated models. They are root mean square error (RMSE), mean
The time series data analyzed in this paper corresponds to the monthly fish catch of anchovy for periods from January 1963 to December 2007 (http://www.sernapesca.cl/). Besides, the original time series data is smoothed using the following operation: $x = \sqrt{x}$. The new smoothed original time series is shown in Figure 1. The first step was to implement the component separation process presented in Algorithm 1 using three different wavelet filters with three levels of decomposition to select the most suitable wavelet filter. The performance of Algorithm 1 was evaluated using the Daubechies filter, Symlets filter, and Coiflets filter. The first two filters are implemented using four and six coefficients denoted as Db2, Sym2, Db3, and Sym3, respectively, whereas the third filter is evaluated using six and twelve coefficients denoted as Coif1 and Coif2. The following step was to detect the most significant periodic fluctuations of each component based on the GWPS. The main goal of the detection phase was to find the significant peak power that explained the periodicities of the time series.

The GWPS for the AC and the IAC is presented in Figures 2, 3, and 4, respectively. In these figures, the peak power indicates the catches high activity. The black thick line designates the 95% confidence level against red noise and the values of the peak power obtained on the black thick line are significant [27, 28].

Figures 2(a) and 2(b) show the AC time series and the GWPS, respectively. From Figure 2(b) it can be observed that there are two peaks of significant power. The first peak has periodicities of 6 months, whereas the second peak has a period of 12 months. Therefore, the AC time series has a seasonal pattern. On the other hand, the AC time series obtained by using 4-level SWT with Db2 wavelet filter is presented in Figure 3(a), whereas Figure 3(b) shows that AC time series has a cycle of 20 months and another of 30 months, which leads to the conclusion that this new AC time series does not meet seasonal behavior, because it has interannual periods. Therefore, only 3-level SWT is used to evaluate the performance of the wavelet autoregressive (WAR) forecasting model proposed.

The AIC time series and GWPS are presented in Figures 4(a) and 4(b), respectively. From Figure 4(b) it can be observed that there are three peaks of significant power. The first peak has a period of 84 months and the second peak has a periodicity of 31 months, while the last peak has a periodic behavior of 152 months. Therefore, these three peaks represent significant interannual fluctuations and can be associated with the El Niño Southern Oscillation effect [30, 31].

Once both the AC and IAC were reconstructed by using the 3-level SWT with Db2 wavelet filter, each component was divided into two parts: a training data set ($T = 360$ observations, 2/3) and a test data set ($M = 180$ observations, 1/3). The training data set was firstly used to estimate the parameters of the WAR forecasting models, and the testing data set was used for computing the performance measures of the models and for validation purposes. The WAR forecasting model was calibrated with $2m = 2 \times 30$ previous months as input data due to the interannual variability of the original time series. In the calibration process overall parameters were estimated using the linear least squares method.

The performance measures of the WAR forecasting model as a function of the forecasting horizon using different wavelet filters with 3-level SWT are shown in Figure 5. From Figure 5 it is observed that the Coif2, Db3, and Sym3 gave a poor performance in comparison to Coif1, Sym3, and Db2. On the contrary, Db2 and Sym2 achieved the better performance. Therefore, the Db2 wavelet filter was used in all further analysis. From Figure 5 it can be seen that the performance measures significantly increase their values for a time horizon of more than 10-month-ahead forecasting.

Once both the AC and IAC have been separately predicted, their values must be added in order to obtain the $p$-step-ahead anchovy-catches forecasting model. In the following we present only 10-month-ahead forecasting results during the testing phase, whose results are illustrated in Figures 6, 7, and 8.

The performance evaluation of the $F_{\text{ax},10}[-]$ and $F_{\text{ai},10}[-]$ models is presented in Figures 6 and 7. As seen in Figures
Figure 2: (a) Annual component \( x_a \) from 3-level SWT and (b) GWPS.

Figure 3: (a) Annual component \( x_a \) from 4-level SWT and (b) GWPS.

Figure 4: (a) Interannual \( x_\text{ia} \) component from 3-level SWT and (b) GWPS.

Figure 5: Forecasting horizon for WAR forecasting.
Figure 6: 10-month-ahead WAR forecasting for annual component ($x_a$).

Figure 7: 10-month-ahead WAR forecasting for interannual component ($x_{ia}$).

Figure 8: 10-month-ahead WAR forecasting for observed anchovy catches versus estimated anchovy catches.

6(a) and 7(a) the actual value and predicted value are very close for the testing data with low performance measures. The model $F_{a,10}$ achieves a RMSE of 0.0093 and MAPE of 36%, whereas the $F_{ia,10}$ model obtained a RMSE of 0.0062 and MAPE 1.38%. On the other hand, from Figure 6(b) it is leveraged that over 80% of the predicted AC catches are acceptable with residual ranging from $-20\%$ to $20\%$ and Figure 7(b) illustrates that 96% of the predicted IAC catches values are within the range of 4%.

Figure 8(a) provides the observed monthly anchovy-catches data set versus forecasted anchovy catches, whose forecasting behavior is very accurate for testing data with a RMSE and MAPE of 0.0093 and 2.66%, respectively. On the other hand, from Figure 8(b) it can be observed that an
important fraction (over 95%) of the predicted values are acceptable with residuals ranging from -10% to 10%.

6. Conclusions

This paper presented a \( p \)-month-ahead forecasting strategy based on two stages to improve pelagic fish-catch time-series modeling. It is based on wavelet analysis and linear AR modeling. In the first stages, the two time series are constructed after 3-level stationary wavelet decomposition containing information about the original time series at annual and interannual time scale, whereas in the second stage each time series is modeled by using a linear AR model. As time series are not independent, each AR model included information of the other time series to improve the forecasting accuracy. The 10-month-ahead forecasting results show that the thirty previous months of each time series contain valuable information to explain the highest variance level of the monthly anchovy catches of northern Chile for periods from January 1963 to December 2007. Finally, the wavelet autoregressive forecasting strategy can be suitable as a very promising methodology for any other marine species of the fishing industry.

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by Grant CONICYT/ FONDECYT/Regular 1131105 and by the DI project of the Pontificia Universidad Católica de Valparaíso.

References


