ISSN 2395-8618

Bi-variate Wavelet Autoregressive Model for Multi-step-ahead Forecasting of Fish Catches

Nibaldo Rodriguez and Lida Barba

***Abstract*—This paper proposes a hybrid multi-step-ahead forecasting model based on two stages to improve monthly pelagic fish-catch time-series modeling. In the first stage, the stationary wavelet transform is used to separate the raw time series into a high frequency (HF) component and a low frequency (LF) component, whereas the periodicities of each time series is obtained by using the Fourier power spectrum. In the second stage, both the HF and LF components are the inputs into a bi-variate autoregressive model to predict the original time series. We demonstrate the utility of the proposed forecasting model on monthly sardines catches time-series of the coastal zone of Chile for periods from January 1949 to December 2011. Empirical results obtained for 12-month ahead forecasting showed the effectiveness of the proposed hybrid forecasting strategy.**

***Index Terms*—Wavelet analysis, bi-variate regression, forecasting model.**

I. Introduction

MULTI-STEP-AHEAD forecasting of pelagic species time series is one of the main goals of the fishery industry and the government. To the best of our knowledge, very publications exist on one-step-ahead forecasting models for fisheries time series based on both autoregressive integrated moving average (ARIMA) models [1], [2] and multilayer perceptron (MLP) neural network models [3], [4]. On the one hand, the disadvantage of models based on linear regression is the supposition of stationarity of the fishes catches time series. However, the fisheries time series are non-stationary due to climatic fluctuations. On the other hand, although MLP neural networks allow modeling the non-linear behavior of a time series, they also have some disadvantages such as slow convergence speed and the stagnancy of local minima due to the steepest descent learning method. To improve the convergence speed and forecasting precision of anchovy catches off northern Chile, Gutierrez [3] proposed a hybrid model based on a MLP neural network combined with an ARIMA model, whose model gave an explained variance of 87%.

In this paper, a multi-step-ahead forecasting model of monthly fishes catches is proposed to achieve a more accurate model than a MLP neural network model. Our proposed forecasting model is based on two phase. In the first phase, the haar stationary wavelet transform (SWT) is used to extract a high frequency (HF) component of intra-annual periodicity and a low frequency (LF) component of inter-annual periodicity. The wavelet decomposition was selected due to its popularity in hydrological [5], [6], electricity market [7], financial market [8] and smoothing methods [9], [10], [11]. In the second stage, both the HF and LF components are the inputs into a bi-variate autoregressive (BAR) model to predict the original time series. Besides, the proposed BAR model is compared with a MLP neural network model with N_i input nodes, N_h hidden nodes and two output nodes.

This paper is organized as follows. In the next section, we present hybrid multi-step-ahead forecasting model. The simulation results are presented in Section 3 followed by conclusions in Section 4.

Manuscript received on May 28, 2015, accepted for publication on July 30, 2015, published on October 15, 2015.

Nibaldo Rodriguez is with the School of Computer Engineering at the Pontificia Universidad Católica de Valparaíso, Av. Brasil 2241, Chile (e-mail: nibaldo.rodriguez@ucv.cl).

Lida Barba is with the School of Computer Engineering at the Universidad Nacional de Chimborazo, Av. Antonio Jose de Sucre, Riobamba, Ecuador (e-mail: lbarba@unach.edu.ec).

II. Proposed Multi-step-ahead Forecasting

In order to predict the future values of time series $x(n)$, we can separate the raw time series $x(n)$ into two components by using Haar SWT. The first extracted component x_H of the time series is characterized by fast dynamics, whereas the second component x_L is characterized by low dynamics. Therefore, in our forecasting model a time series is considered as a functional relationship of several past observations of the components x_L and x_H as follows:

$$\hat{x}(n+h) = f(x_L(n-m), x_H(n-m));$$

the h value represents forecasting horizon and $i = 1, 2, \ldots m$ denotes lagged values of both the LF and HF components. Besides, the functional relationship $f(\cdot)$ in this paper is estimated by using a BAR model and a MLP neural network model. The following three subsections present the SWT, BAR forecasting model and MLP forecasting model.

A. Stationary Wavelet Transform

Let $x(n)$ denote the value of a time series at time n, then $x(n)$ can be represented at multiple resolutions by decomposing the signal on a family of wavelets and scaling functions [9], [10], [11]. The approximation (scaled) signals are computed by projecting the original signal on a set of orthogonal scaling functions of the form:

$$\phi_{jk}(t) = \sqrt{2^{-j}}\phi(2^{-j}t - k),$$

ISSN 2395-8618

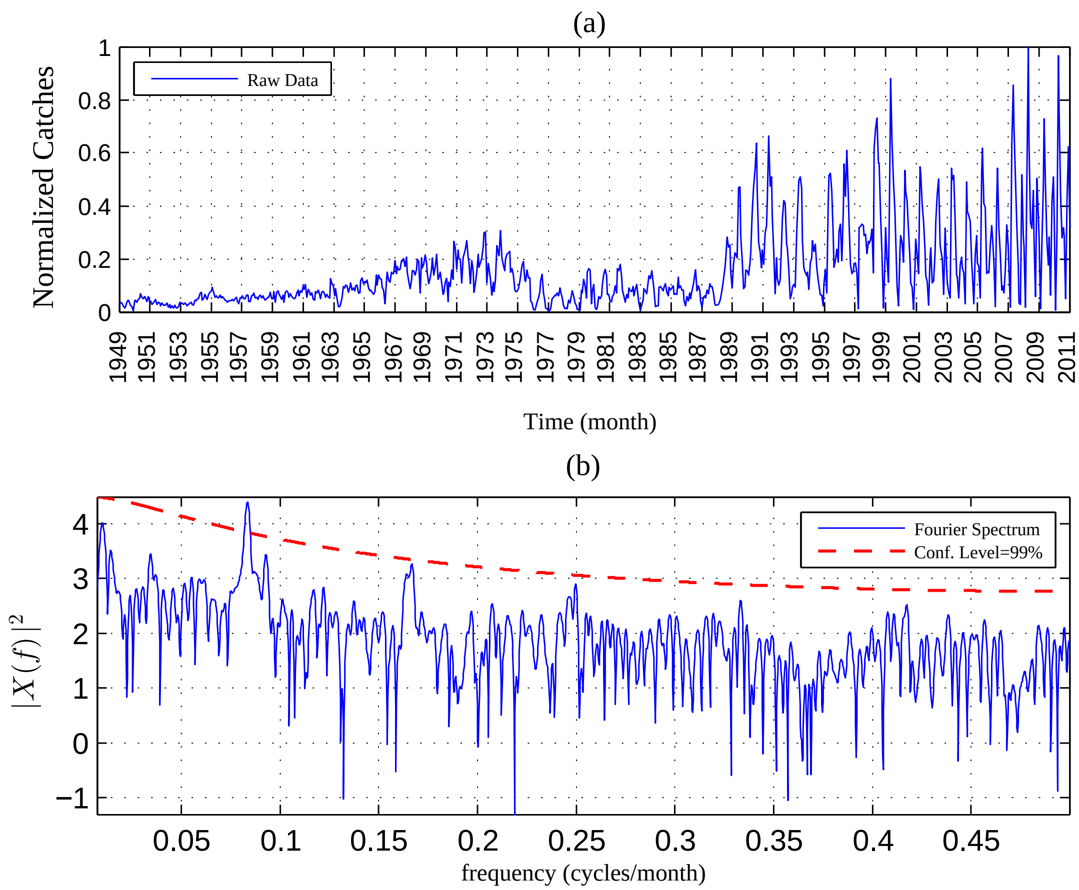

Fig. 1. Monthly sardines catches

or equivalently by filtering the signal using a low pass filter of length r, $h = [h_1, h_2, ..., h_r]$, derived from the scaling functions. On the other hand, the detail signals are computed by projecting the signal on a set of wavelet basis functions of the form

$$\psi_{jk}(t) = \sqrt{2^{-j}}\psi(2^{-j}t - k),$$

or equivalently by filtering the signal using a high pass filter of length r, $g = [g_1, g_2, ..., g_r]$, derived from the wavelet basis functions. Finally, repeating the decomposing process on any scale J, the original signal can be represented as the sum of all detail coefficients and the last approximation coefficient. In time series analysis, discrete wavelet transform (DWT) often suffers from a lack of translation invariance. This problem can be tackled by means of the un-decimated stationary wavelet transform (SWT). The SWT is similar to the DWT 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 up-sampled at each level.

Consider the following discrete signal $x(n)$ of length N where $N = 2^J$ for some integer J. At the first level of SWT, the input signal $x(n)$ is convolved with the $h_1(n)$ filter to obtain the approximation coefficients $a_1(n)$ and with the $g_1(n)$ filter to obtain the detail coefficients $d_1(n)$, so that:

$$a_1(n) = \sum_k h_1(n-k)x(k),$$
$$d_1(n) = \sum_k g_1(n-k)x(k),$$

because no sub-sampling is performed, $a_1(n)$ and $d_1(n)$ are of length N instead of $N/2$ as in the DWT case. At the next level of the SWT, $a_1(n)$ is split into two parts by using the same scheme, but with modified filters h_2 and g_2 obtained by dyadically up-sampling h_1 and g_1.

The general process of the SWT is continued recursively for $j = 1, ..., J$ and is given as:

$$a_{j+1}(n) = \sum_k h_{j+1}(n-k)a_j(k)$$
$$d_{j+1}(n) = \sum_k g_{j+1}(n-k)a_j(k)$$

where h_{j+1} and g_{j+1} are obtained by the up-sampling operator inserts a zero between every adjacent pair of elements of h_j

ISSN 2395-8618

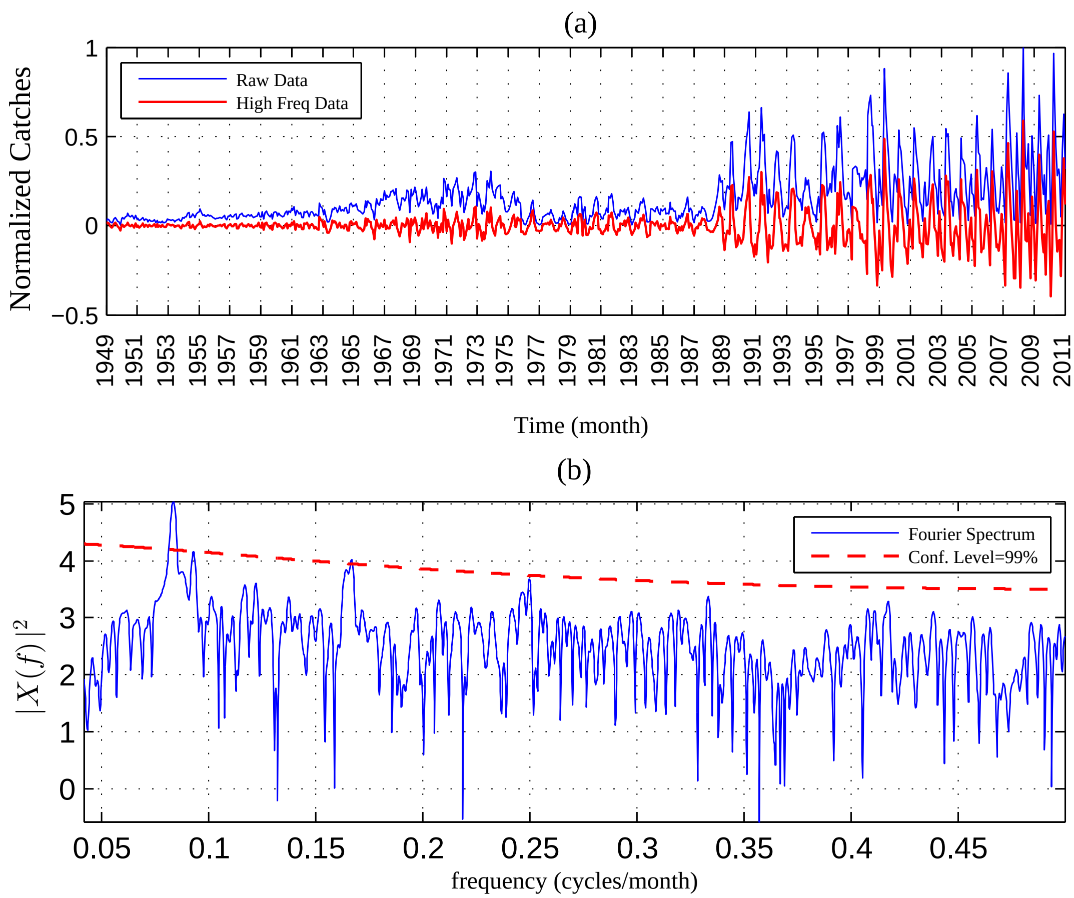

Fig. 2. High frequency sardines catches

and g_j; respectively. Therefore, the output of the SWT is then the approximation coefficients a_J and the detail coefficients $d_1, d_2, ..., d_J$. The wavelet decomposition method is fully defined by the choice of a pair of low and high pass filters and the number of decomposition steps J. Hence, in this study we choose a pair of Daubechies Db2 filters (has two wavelet and scaling coefficients) [12].

B. Bi-variate Forecasting Model

A bi-variate wavelet autoregressive (BWAR) model is used to estimate the function $\hat{f}(\cdot)$, which is given as

$$\begin{aligned} U &= ZA, \\ u_{i,1} &= X_H(i+h), \\ u_{i,2} &= X_L(i+h), \\ z_{i,j} &= X_H(i-j), j = 0, \dots, m-1, \\ z_{i,m+j} &= X_L(i-j), j = 0, \dots, m-1, \end{aligned}$$

where U is the matrix dependent variables of M rows by 2 columns, M is the set of input-output samples, Z is the regressor matrix of M rows by $2m$ columns and A is the parameters matrix of $2m$ rows by 2 columns. In order to estimate the parameters A the linear least squares method is used, which is given as

$$A = Z^{\dagger}U;$$

$(\cdot)^{\dagger}$ denotes the Moore-Penrose pseudoinverse [13].

C. Neural Network Forecasting Model

A single-hidden neural network with two output nodes is used to estimate the function $\hat{f}(\cdot)$, which is defined as

$$\begin{aligned} u_k(n) &= \sum_{j=1}^{N_h} b_j \phi_j(z_i, v_j), k = 1, 2, \\ \hat{x}(n+h) &= u_1(n) + u_2(n), \end{aligned}$$

where N_h is the number of hidden nodes, $z = [z_1, z_2, \dots z_{2m}]$ denotes the input regression vector containing $2m$ lagged values, $[b_1, \dots b_{N_h}]$ represents the linear output parameters, $[v_{j,1}, v_{j,2}, \dots v_{j,2m}]$ denotes the nonlinear parameters, and

ISSN 2395-8618

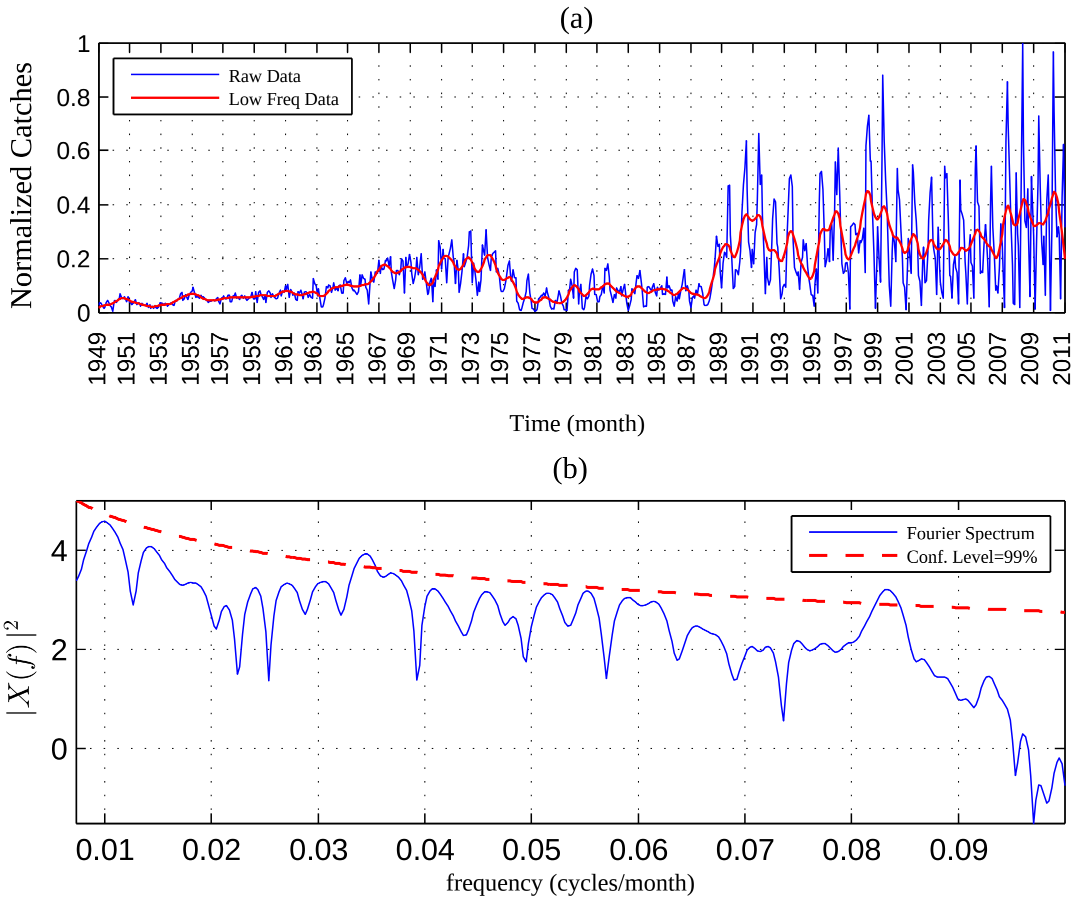

Fig. 3. Low frequency sardines catches

$\phi_j(\cdot)$ are hidden activation functions, which are derived as:

$$\phi_j(z_i) = \phi\Big(\sum_{i=1}^{2m} v_{j,i} z_i\Big),$$
$$\phi(z) = \frac{1}{1+exp(-z)}.$$

In order to estimate both the linear and nonlinear parameters of the MLP, we use the Levenberg-Marquardt (LM) algorithm [14]. The LM algorithm adapts the $\theta = [b_1, \ldots b_{N_h}, v_{j,1}, \ldots v_{j,2m}]$ parameters of the neuro-forecaster minimizing mean square error, which is defined as:

$$E(\theta) = \frac{1}{2}\sum_{i=1}^{M}\Big(e(\theta_i)\Big)^2$$

Finally, the LM algorithm adapts the parameter θ according to the following equations:

$$\theta = \theta + \Delta\theta,$$
$$\Delta\theta = (\Upsilon\Upsilon^T + \mu I)^{-1}\Upsilon^T e,$$

where Υ represents the Jacobian matrix of the error vector evaluated in θ_i and the error vector $e(\theta_i) = x(n+h) - \hat{x}(n+h)$ is the error of the MLP neural network for i pattern, I denotes the identity matrix and the parameter μ is increased or decreased at each step of the LM algorithm.

III. Experiments and Results

In this section, we apply the proposed BWAR model for 12-month-ahead sardines catches forecasting. The data set used corresponded to landing of sardines in the south of Chile. These samples were collected monthly from 1 January 1949 to 31 December 2011 by the National Fishery Service of Chile (www.sernapesca.cl). The raw sardines data set have been normalized to the range from 0 to 1 by simply dividing the real value by the maximum of the appropriate set. On the other hand, the original data set was also divided into two subsets. In the first subset the 85% of the time series were chosen for the calibration phase (parameters estimation), whereas the remaining data set were used for the testing phase.

The normalized raw time series and the Fourier power spectrum are present in Figures 1(a) and 1(b); respectively. The red thick line in Figure 1(b) designates the confidence level against red noise spectrum. From Figure 1(b) it can be observed that there are one peaks of significant power, which

ISSN 2395-8618

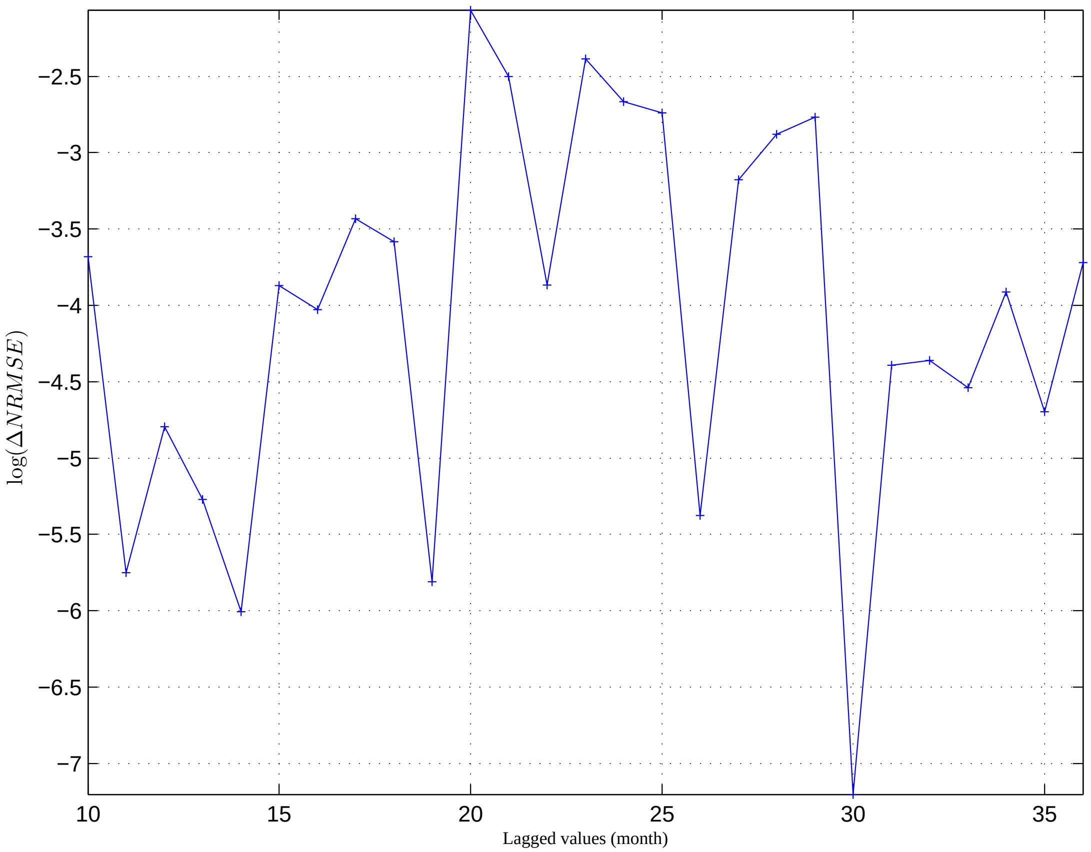

Fig. 4. NRMSE versus Lagged values

has an annual periodicities of 12 months ($freq = 0.083$). After we applied the Fourier power spectrum to the raw time series, we decided to use 3-level wavelet decomposition due to the significative peak of 12 months. Both the HF and LF components (times series) are presented in Figures 2(a) and 3(a); respectively, whereas the power spectrum of both time series are illustrated in Figure 2(b) and 3(b); respectively.

In this study, three criteria of forecasting accuracy called normalised root mean squares error (NRMSE), modified Nash-Sutcliffe efficiency coefficient (MNSE) and coefficient of determination (R2) were used to evaluate the forecasting capabilities of the proposed hybrid forecasting models, which are defined as

$$NRMSE = \sqrt{\frac{\sum_{i=1}^{L}\left(x(i)-\hat{x}(i)\right)^2}{\sum_{i=1}^{L}\left(x(i)-\bar{x}\right)^2}}$$

$$MNSE = 1 - \frac{\sum_{i=1}^{L}\left|x(i)-\hat{x}(i)\right|}{\sum_{i=1}^{L}\left|x(i)-\bar{x}\right|}$$

$$R2 = 1 - \frac{\sum_{i=1}^{L}\left(x(i)-\hat{x}(i)\right)^2}{\sum_{i=1}^{L}\left(x(i)-\bar{x}\right)^2},$$

where $x(i)$ is the actual value at time i, $\hat{x}(i)$ is the forecasted value at time i, $\bar{x}$ is the mean of observed data and L is the number of forecasts.

Find the order of the bi-variate autoregressive model is a complex task, but here we will use the following metric to evaluate different lagged values, which is given as

$$\Delta(NRMSE) = NRMSE(m) - NRMSE(m-1),$$

where m denotes values with m=2,…,36 months.

Figure 4 shows the results of testing data for lagged values between 10 and 36 months due to significant periods of the low frequency component, whose best result was achieved with $m = 30$ months, whereas Figures 5 show the results obtained with the best BWAR(30) forecasting model during the testing phase. Figure 5(a) provides data on observed monthly sardines catches versus forecasted catches; this forecasting behavior is very accurate for testing data with a NRMSE of 11% and a MNSE of 98%. On the other hand, from Figure 5(b) it can be observed a good fit to a linear curve with a coefficient of determination of 98%.

Once evaluated the BWAR(30) forecasting model perform calibration of the neural network with $N_i = 60$ input nodes, $N_h = \sqrt{N_i + N_o} = \sqrt{60+2} = 8$ hidden nodes and $N_o = 2$ output nodes. In the training process, overall weights were initialized by a Gaussian random process with

ISSN 2395-8618

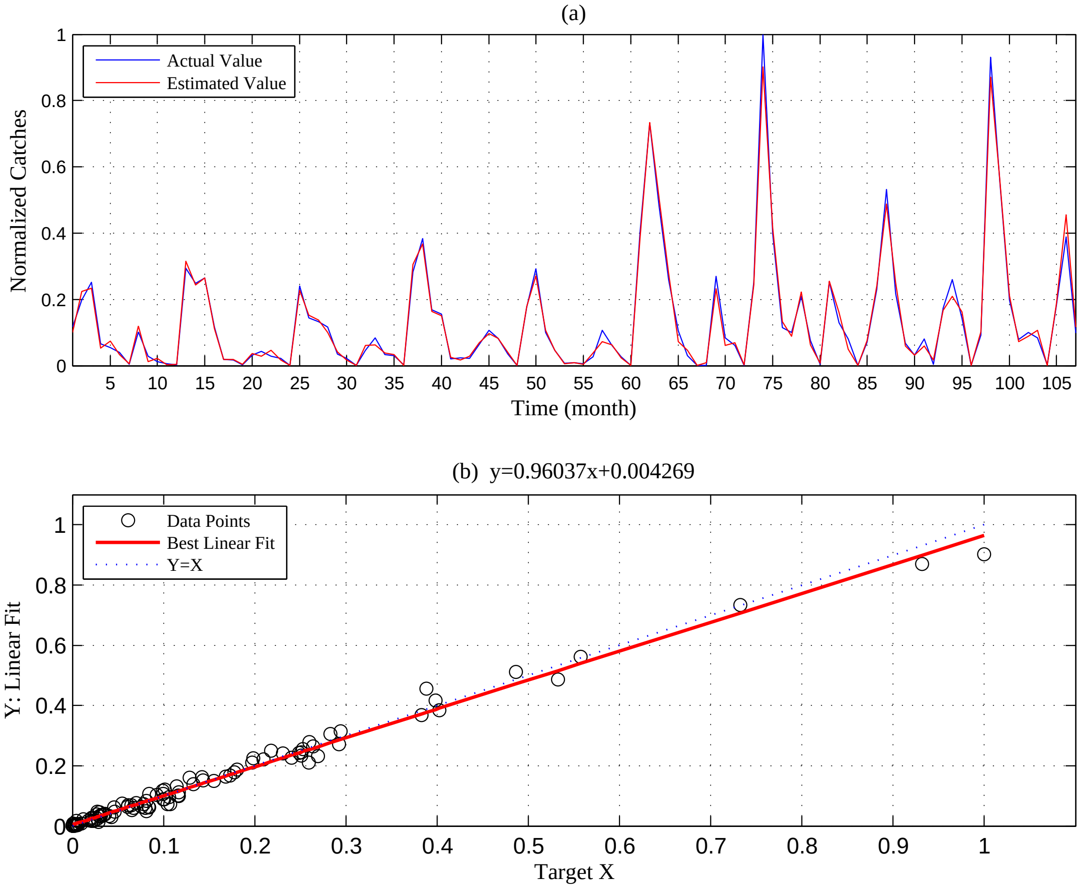

Fig. 5. Twelve-month-ahead BWAR forecasting for test data set

a normal distribution $N(0,1)$ and the stopping criterion was a maximum number of iterations set at 500. Due to the random initialization of the weights, we used 30 runs to find the best MLP neural network with a low prediction error. Figures 6(a) and 6(b) show the results obtained with the MLP(60,8,2) forecasting model during the testing phase. Figure 6(a) illustrates the observed data set versus forecasted data set, which obtains a NRMSE and a MNSE of 35% and 68%; respectively. On the other hand, Figure 6(b) shows the scatter curve between observed values and forecasted values with a R2 of 89%.

IV. Conclusions

In this paper was proposed a multi-step-ahead forecasting model to improve prediction accuracy based on Haar stationary wavelet decomposition combined with a bi-variate autoregressive model. The reason of the improvement in forecasting accuracy was due to use Daubechies SWT to separate both the LF and HF components of the raw time series, since the behavior of each component is more smoothing than raw data set. It was show that the proposed hybrid forecasting model achieves 11% and 98% of NRMSE and MNSE; respectively. Besides, the experimental results demonstrated a better performance of the proposed model when compared with a MLP neural network prediction model. Finally, hybrid forecasting model can be suitable as a very promising methodology to any other pelagic species.

Acknowledgment

This research was partially supported by the Chilean National Science Fund through the project Fondecyt-Regular 1131105 and by the project DI-Regular $037.442/2015$ of the Pontificia Universidad Católica de Valparaíso.

References

[1] S. KI., "Prediction of the mullidae fishery in the easterm mediterranean 24 months in advance," *Fisheries Research*, vol. 9, pp. 67–74, 1996.
[2] S. K.I. and C. E.D., "Modelling and forecasting annual fisheries catches: comparison of regression, univariate and multivariate time series methods," *Fisheries Research*, vol. 25, pp. 105–138, 1996.
[3] J. C. Gutierrez, S. C., Y. E., R. N., and P. I., "Monthly catch forecasting of anchovy engraulis ringens in the north area of chile: Nonlinear univariate approach," *Fisheries Research*, vol. 86, no. 188-200, 2007.
[4] S. P. Garcia, L. B. DeLancey, J. S. Almeida, and R. W. Chapman, "Ecoforecasting in real time for commercial fisheries: The Atlantic white shrimp as a case study," *Marine Biology*, vol. 152, pp. 15–24, 2007.
[5] J. F. Adamowski, "Development of a short-term river flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis," *Journal of Hydrology*, vol. 353, no. 3-4, pp. 247–266, 2008.

ISSN 2395-8618

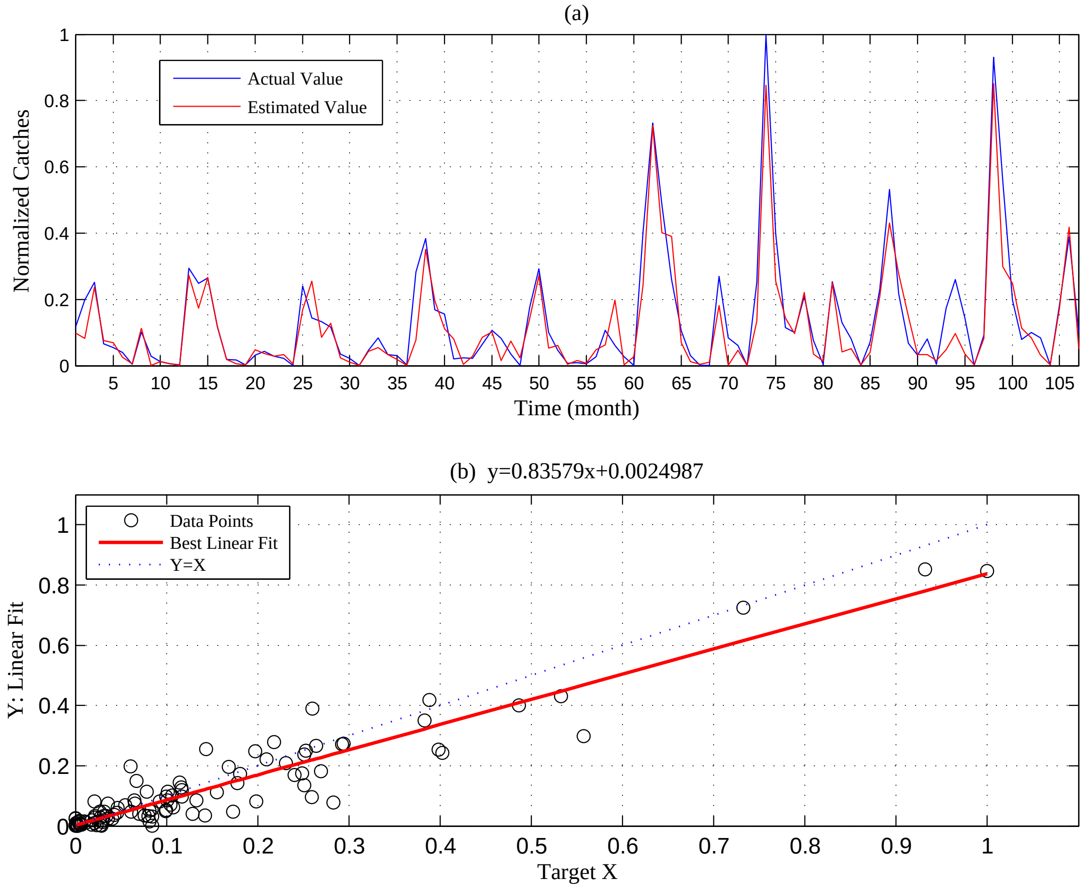

Fig. 6. Twelve-month-ahead WMLP forecasting for test data set

[6] O. Kisi, "Stream flow forecasting using neuro-wavelet technique," *Hydrological Processes*, vol. 22, no. 20, pp. 4142–4152, 2008.

[7] N. Amjady and F. Keyniaa, "Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method," *International Journal of Electrical Power Energy Systems*, vol. 30, pp. 533–546, 2008.

[8] B.-L. Z., R. C., M. A. J., D. D., and B. F., "Multiresolution forecasting for futures trading using wavelet decompositions," *IEEE Trans. on neural networks*, vol. 12, no. 4, pp. 765–775, 2001.

[9] R. Coifman and D. L. Donoho, "Translation-invariant denoising, wavelets and statistics," *Springer Lecture Notes in Statistics*, vol. 103, pp. 125–150, 1995.

[10] G. Nason and B. Silverman, "The stationary wavelet transform and some statistical applications, wavelets and statistics," *Springer Lecture Notes in Statistics*, vol. 103, pp. 281–300, 1995.

[11] J.-C. Pesquet, H. Krim, and H. Carfantan, "Time-invariant orthonormal wavelet representations," *IEEE Trans. on Signal Processing*, vol. 44, no. 8, pp. 1964–1970, 1996.

[12] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*. Cambridge, England: Cambridge University Press, 2000.

[13] D. Serre, *Matrices: Theory and Applications*. Springer,NewYork, NY, USA, 2002.

[14] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the marquardt algorithm," *IEEE transactions on neural networks*, vol. 5, no. 6, pp. 989–993, 1996.