Polibits
ISSN: 1870-9044
polibits@nlp.cic.ipn.mx
Instituto Politécnico Nacional
México

Rodriguez, Nibaldo; Bravo, Gabriel; Barba, Lida
Haar Wavelet Neural Network for Multi-step-ahead Anchovy Catches Forecasting
Polibits, vol. 50, 2014, pp. 49-53
Instituto Politécnico Nacional
Distrito Federal, México

Haar Wavelet Neural Network for Multi-step-ahead Anchovy Catches Forecasting

Nibaldo Rodriguez, Gabriel Bravo, and Lida Barba

Abstract—**This paper proposes a hybrid multi-step-ahead forecasting model based on two stages to improve pelagic fish-catch time-series modeling. In the first stage, the Fourier power spectrum is used to analyze variations within a time series at multiple periodicities, while the stationary wavelet transform is used to extract a high frequency (HF) component of annual periodicity and a low frequency (LF) component of inter-annual periodicity. In the second stage, both the HF and LF components are the inputs into a single-hidden neural network model to predict the original non-stationary time series. We demonstrate the utility of the proposed forecasting model on monthly anchovy catches time-series of the coastal zone of northern Chile (18^oS-24^oS) for periods from January 1963 to December 2008. Empirical results obtained for 7-month ahead forecasting showed the effectiveness of the proposed hybrid forecasting strategy.**

Index Terms—**Neural network, wavelet analysis, forecasting model.**

I. Introduction

In order to develop sustainable exploitation policies, forecasting the stock and catches of pelagic species off northern Chile is one of the main goals of the fishery industry and the government. However, fluctuations in the environmental variables complicate this task. To the best of our knowledge, few publications exist on forecasting models for pelagic species. In recent years, linear regression models [1], [2] and artificial neuronal networks (ANN) [3], [4] have been proposed for forecasting models. 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 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 multilayer perceptron (MLP) combined with an autoregressive integrated moving average (ARIMA) model.This forecaster obtained a coefficient of determination R^2 of 82%, which improved slightly when combining the MLP model with the ARIMA model, reaching an R^2 of 87%.

In this paper, the proposed forecasting model is based on single-hidden neural network combined with Haar stationary wavelet transform (SWT). The stationary wavelet decomposition was selected due to its popularity in hydrological [5], [6], electricity market [7], financial market [8] and smoothing methods [9]–[11]. This SWT technique is based on the discreet wavelet transform (DWT) or the stationary wavelet transform (SWT) [12]. The advantage of these wavelet transforms in non-stationary time series analysis is their capacity to separate low frequency (LF) from high frequency (HF) components. On the one hand, the LF component reveals long-term trends, while the HF component describes short-term fluctuations in the time series. Being able to separate these components is a key advantage in proposed forecasting strategies since the behavior of each frequency component is more regular than the raw time series.

In this paper, Haar stationary wavelet decomposition is applied to build a hybrid multi-step-ahead forecasting model to achieve more accurate models than conventional single-hidden neuronal network. Our proposed multi-step-ahead anchovy catches forecasting model is based on two phase. In the first phase,the Fourier power spectrum is used to analyze variations within a time series at multiple periodicities, while the stationary wavelet transform is used to extract a high frequency (HF) component of annual periodicity and a low frequency (LF) component of inter-annual periodicity. In the second stage, both the HF and LF components are the inputs into a single-hidden neural network model with N_i input nodes, N_h hidden nodes and two output nodes to predict the original non-stationary time series.

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.

II. Proposed forecasting model

This section presents the proposed forecasting model for one-month-ahead anchovy catches in northern Chile, which is based on the Haar stationary wavelet transform and single-hidden neural network model.

A. Stationary wavelet decomposition

A signal $x(n)$ can be represented at multiple resolutions by decomposing the signal on a family of wavelets and

Manuscript received on August 7, 2014, accepted for publication on September 22, 2014, published on November 15, 2014.

Nibaldo Rodriguez (corresponding author) is with the Pontificia Universidad Católica de Valparaíso, Av. Brasil 2241, Chile (e-mail: nibaldo.rodriguez@ucv.cl).

Gabriel Bravo is with the Universidad San Sebastián, Concepción, Chile (e-mail: gabo.bravoro@hotmail.com).

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

scaling functions [9]–[11]. The approximation (scaled) signals are computed by projecting the original signal on a set of orthogonal scaling functions of the form:

$$\phi_{jk}(t) = \sqrt{2^{-j}}\phi(2^{-j}t - k) \tag{1}$$

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

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

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:

$$a_1(n) = \sum_k h_1(n-k)x(k) \tag{3a}$$

$$d_1(n) = \sum_k g_1(n-k)x(k), \tag{3b}$$

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:

$$a_{j+1}(n) = \sum_k h_{j+1}(n-k)a_j(k) \tag{4a}$$

$$d_{j+1}(n) = \sum_k g_{j+1}(n-k)a_j(k), \tag{4b}$$

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 and g_j; respectively.

Therefore, the output of the SWT is then the approximation coefficients a_J and the detail coefficients $d_1, d_2, ..., d_J$, whereas the original signal $x(n)$ is represented as a superposition of the form:

$$x(n) = a_J(n) + \sum_{j=1}^{J} d_j(n) \tag{5}$$

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 haar wavelet filters [12].

B. Neural network forecasting model

In order to predict the future value $\hat{x}(n-h)$, we can separate the original time series $x(n)$ into two components by using Haar stationary wavelet decomposition. The first extracted component x_H of the time series is characterized by slow dynamics, whereas the second component x_H is characterized by fast dynamics. Therefore, in our forecasting model a time series is considered as nonlinear function of several past observations of the components x_L and x_H as follows:

$$\hat{x}(n+h) = f\big(x_L(n-m), \ldots x_H(n-m)\big); \tag{6}$$

the h value represents forecasting horizon and m denotes lagged values of both the LF and HF components.

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

$$y_k(n) = \sum_{j=1}^{N_h} b_j \phi_j(u_i, v_j), k = 1, 2 \tag{7a}$$

$$\hat{x}(n+h) = y_1(n) + y_2(n), \tag{7b}$$

where N_h is the number of hidden nodes, $u = [u_1, u_2, \ldots u_{2m}]$ denotes the input regression vector containing $2m$ lagged values, $[b_1, \ldots b_{N_h}]$ represents the linear output parameters, $[v_1, v_2, \ldots v_{N_h}]$ denotes the nonlinear parameters, and $\phi_j(\cdot)$ are hidden activation functions, which are derived as

$$\phi_j(u_i) = \phi\Big(\sum_{i=1}^{m} v_{j,i}u_i\Big) \tag{8a}$$

$$\phi(u) = \frac{1}{1 + exp(-u)}. \tag{8b}$$

In order to estimate both the linear and nonlinear parameters of the MLP, we use the Levenberg-Marquardt (LM) algorithm [13]. The LM algorithm adapts the $\theta = [b_1, \ldots b_{N_h}, v_{j,1}, \ldots v_{j,m}]$ parameters of the neuro-forecaster minimizing mean square error, which is defined as:

$$E(\theta) = \frac{1}{2}\sum_{i=1}^{N_s}\Big(x(n+h) - \hat{x}(n+h))\Big)^2. \tag{9}$$

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

$$\theta = \theta + \Delta\theta \tag{10a}$$

$$\Delta\theta = (\Upsilon\Upsilon^T + \mu I)^{-1}\Upsilon^T E, \tag{10b}$$

where Υ represents the Jacobian matrix of the error vector evaluated in θ_i and the error vector $e(\theta_i) = d_i - y_i$ is the

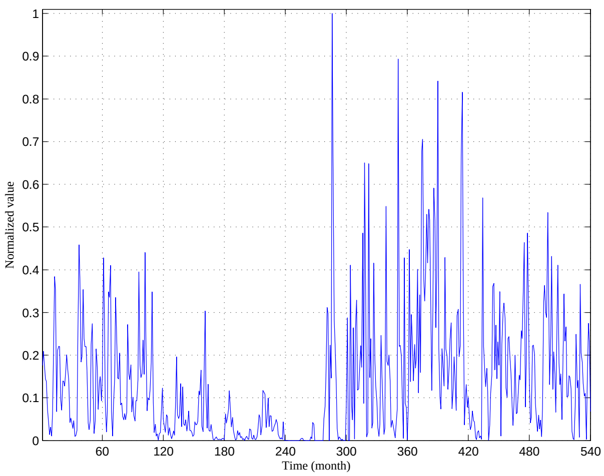

Fig. 1. Monthly anchovy catches time series

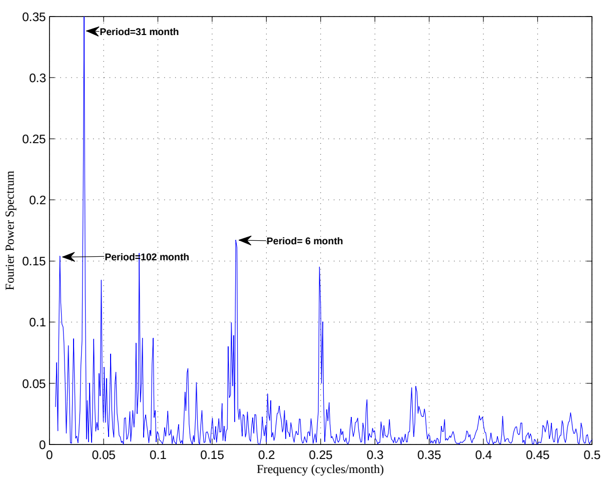

Fig. 2. Fourier power spectrum of catches time series

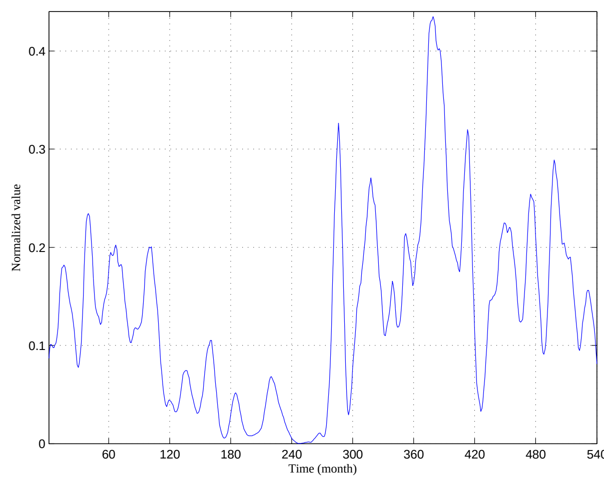

Fig. 3. Low frequency anchovy catches time series

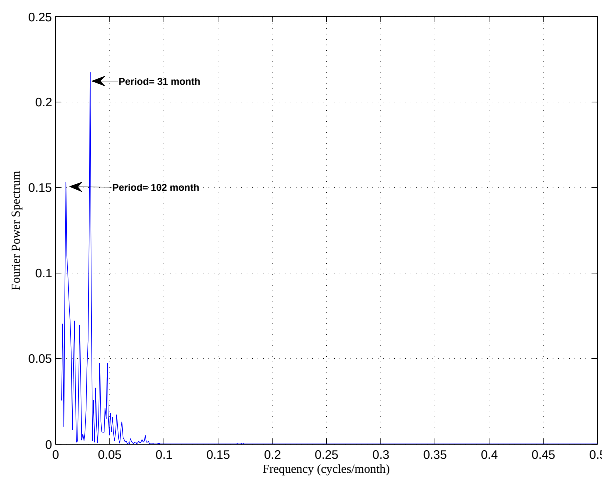

Fig. 4. Fourier power spectrum of LF time series

error of the MLP neural network for i patter, 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 hybrid model for 7-month-ahead anchovy catches forecasting. The data set used corresponded to anchovy landings off northern Chile. These samples were collected monthly from 1 January 1963 to 30 December 2008 by the National Fishery Service of Chile (www.sernapesca.cl). The raw anchovy 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 70% of the time series were chosen for the training 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 the Figures 1 and 2; respectively. From Figure 2 it can be observed that there are two peaks of significant power. The first peak has an inter-annual periodicities of 31 months, whereas the second peak has a intra-annual periodicities of 6 months. 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 31 months. Both the LF and HF times series are presented in Figures 3 and 5; respectively, whereas the power spectrum of both time series are illustrated in Figure 4 and 6; respectively.

In this study, three criteria of forecasting accuracy called root mean squares error (RMSE), mean absolute percentage error (MAPE) and ralative error (RE) were used to evaluate the forecasting capabilities of the proposed hybrid forecasting models, which are defined as

$$RMSE = \sqrt{MSE}, \tag{11}$$

$$MAPE(\%) = \frac{1}{N_s}\sum_{i=1}^{N_s}\left|\frac{A_i - F_i}{F_i}\right| \times 100, \tag{12}$$

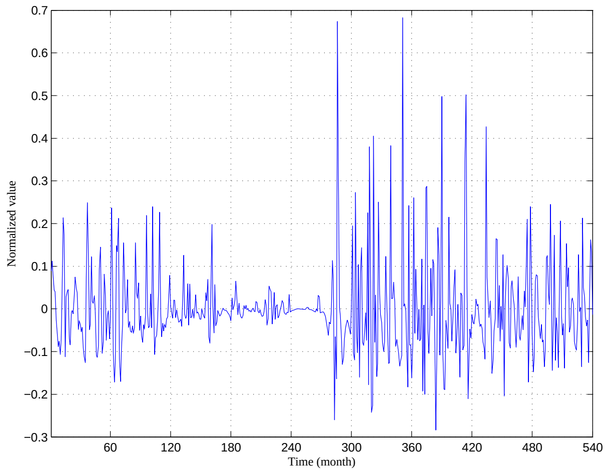

Fig. 5. High frequency anchovy catches time series

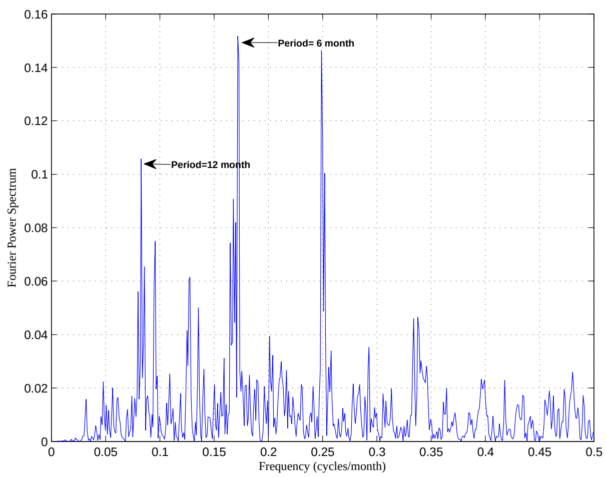

Fig. 6. Fourier power spectrum of HF time series

$$RE(\%) = \frac{A_i - F_i}{F_i} \times 100, \tag{13}$$

where A_i is the actual value at time i, F_i is the forecasted value at time i, and N_s is the number of forecasts.

The MLP neural network was calibrated using 31 previous months as input data plus one bias unit due to the periodicity of 31 months of the raw time series (see Figure 2). Finding the optimal number of hidden nodes is a complex problem, but in all our experiments, the number of hidden nodes is set as $\sqrt{N_i + N_o} = \sqrt{62+2}$ (number of input nodes and output nodes). In the training process, overall weights were initialized by a Gaussian random process with a normal distribution $N(0, 1)$ and the stopping criterion was a maximum number of iterations set at 200. Due to the random initialization of the weights, we used 50 runs to find the best MLP neural network with a low prediction error.

The Figures 7 and 8 shows the results of testing data for 50 run and 200 iterations, whose best result was achieved in the run 49. After the training-testing process, the architecture

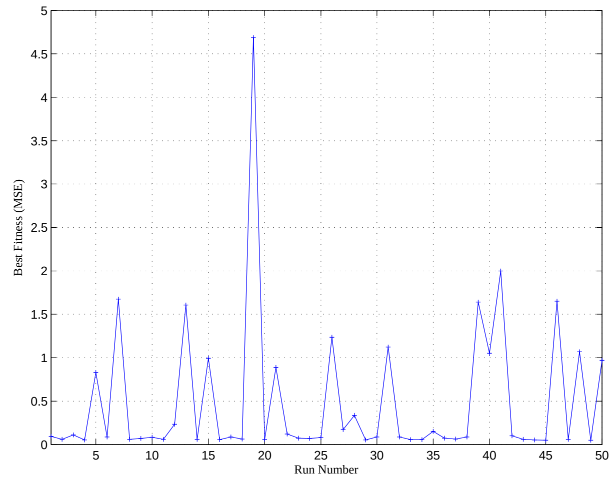

Fig. 7. Run versus MSE

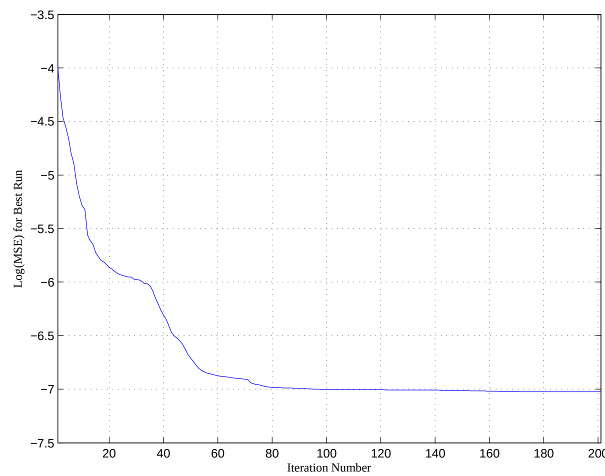

Fig. 8. Iteration number for Best Run

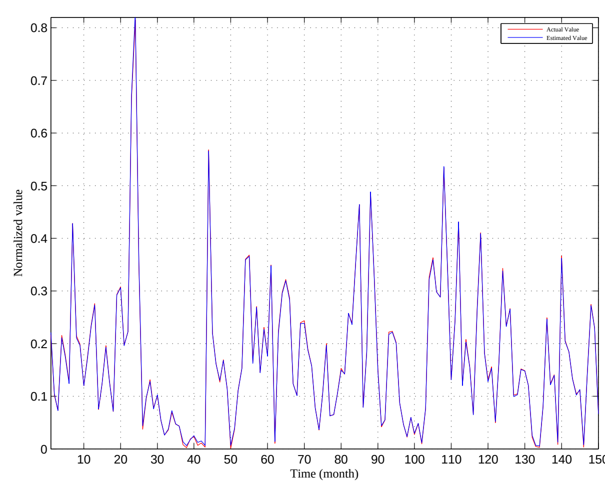

Fig. 9. Seven-month-ahead MLP(31,8,2) forecasting for test data set

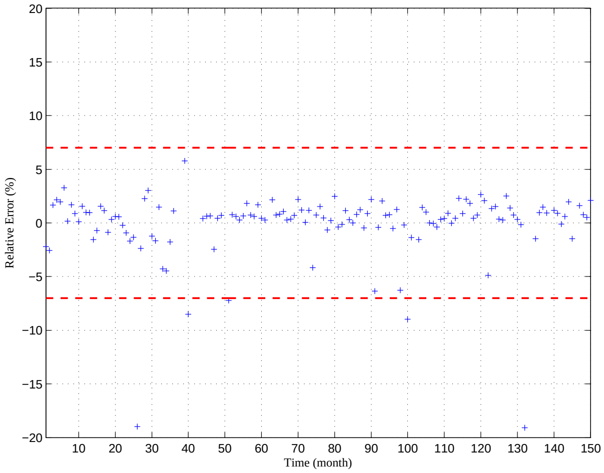

Fig. 10. Retative error for Seven-month-ahead MLP(31,8,2) forecasting during test data set

was calibrated with 31 input nodes, 8 hidden nodes, and two output node; this is denoted as MLP(31,8,2).

Figures 9 and 10 show the results obtained with the MLP(31,8,2) forecasting model during the testing phase. Fig. 9 provides data on observed monthly anchovy catches versus forecasted catches; this forecasting behavior is very accurate for testing data with a MAPE of 10.87% and a RMSE of 0.0028. On the other hand, from Figure 10 it can be observed that an important fraction (over 90%) of the predicted catches values are acceptable with residuals ranging from 7% to -7%.

IV. Conclusions

In this paper was proposed a 7-month-ahead anchovy catches forecasting strategy to improve prediction accuracy based on Haar stationary wavelet decomposition combined with a single-hidden neural network model. The reason of the improvement in forecasting accuracy was due to use Haar 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 a MAPE of 10.87% and a RMSE of 0.0028. Besides, proposed forecasting results showed that the 31 previous month contain valuable information to explicate a highest variance level for anchovy catches forecasting. Finally, hybrid forecasting strategy 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 VRIEA of the Pontificia Universidad Católica de Valparaíso.

References

[1] K. Stergiou, “Prediction of the mullidae fishery in the easterm mediterranean 24 months in advance,” *Fisheries Research*, vol. 9, pp. 67–74, 1996.

[2] K. Stergiou and E. Christou, “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. Gutierrez, C. Silva, E. Yañez, N. Rodriguez, and I. Pulido, “Monthly catch forecasting of anchovy engraulis ringens in the north area of chile: Nonlinear univariate approach,” *Fisheries Research*, vol. 86, pp. 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.

[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] Z. Bai-Ling, C. Richard, A. J. Marwan, D. Dominik, and F. Barry, “Multiresolution forecasting for futures trading using wavelet decompositions,” *IEEE Trans. on neural networks*, vol. 12, no.4, 2001.

[9] R. 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] 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.