Wavelet Denoising Based Multivariate Polynomial For Anchovy Catches Forecasting

Nibaldo Rodriguez
Pontificia Universidad Católica de Valparaíso, Escuela Ingeniería Informática,
Valparaíso, Chile
FirstName.Name@ucv.cl

and

Guillermo Cabrera
Pontificia Universidad Católica de Valparaíso, Escuela Ingeniería Informática,
Valparaíso, Chile
FirstName.Name@ucv.cl

Abstract

In this paprer, a multivariate polynomial (MP) combined with denoising techniques is proposed to forecast 1-month ahead monthly anchovy catches in the north area of Chile. The anchovy catches data is denoised by using discrete stationary wavelet transform and then appropriate is used as inputs to the MP. The MP's parameters are estimated using the penalized least square (LS) method and the performance evaluation of the proposed forecaster showed that a 98% of the explained variance was captured with a reduced parsimony.

Keywords: Forecasting, multivariate polynomial, wavelet transform.

Resumen

En este artículo, un polinomio multivariado (PM) combinado con técnicas de eliminación de ruido is propuesto para pronosticar capturas de anchovetas mensuales en el área norte de Chile. La data de capturas de anchovetas mensuales es suavizada utilizando la transformada wavelet estacionaria y dicha data es la entrada al PM. Los parámetros del PM son estimados usando el método de los mínimos cuadrados penalizados y la evaluación de rendimiento del pronosticador propuesto muestra que un 98% de la varianza explicada fue capturada con una reducida parsimonia.

Palabras claves: Pronóstico, polinomio multivariado, transformada wavelet.

1 Introduction

Linear regression has been successful in describing and forecasting the fishery dynamics of a wide variety of species [1, 2], but this technique is inefficient for capturing nonlinear phenomena. The forecasting of monthly anchovy catches in the north area of Chile is a basic topic, because it play a central role in management of stock. In fisheries management policy the main goal is to establish the future catch per unit of effort (CPUE) values in a concrete are during a know period keeping the stock replacements. To achieve this aim lineal regression methodology has been successful in describing and forecasting the fishery dynamics of a wide variety of species [1, 2]. However, this technique is inefficient for capturing both nonstationary and nonlinearities phenomena in anchovy catch forecasting time series. Recently there has been an increased interest in combining nonlinear techniques and wavelet theory to model complex relationship in nonstationary time series. Nonlinear model based on Neural networks have been used for forecasting model due to their ability to approximate a wide range of unknown nonlinear functions [3]. Gutierrez *et. al.* [4], propose a forecasting model of monthly anchovy catches based on a sigmoidal neural network, whose architecture is composed of an input layer of 6 nodes, two hidden layers having 15 nodes each layer, and a linear output layer of a single node. Some disadvantages of this architecture is its high parsimony as well as computational time cost during the estimation of linear and nonlinear weights. As shown in [4], when applying the Levenberg Marquardt (LM) algorithm, the forecasting model achieves a determination coefficient of 82%. A better result of the determination coefficient can be achieved if sigmoidal neural network is substituted by a reduced multivariate polynomial (MP) combined with wavelet denoising techniques based on translation-invariant wavelet transform. Coifman and Donoho [6] introduced translation-invariant wavelet denoising algorithm based on the idea of cycle spinning, which is equivalent to denoising using the discrete stationary wavelet transform (SWT) [7, 8]. Beside, Coifman and Donoho showed that SWT denoising achieves better root mean squared error than traditional descrete wavelet transform denoising. Therefore, we employ the SWT for denoising monthly anchovy catches data.

This paper evaluates the performance of a multivariate polynomial based on stationary wavelet denoising for one-step monthly anchovy catches forecasting and the parameters of the proposed MP model are estimated using the penalized least square (LS) method.

The rest of the paper is organized as follows: the multivariate polynomial based forecasting model and estimating algorithm are presented in Section 2. Section 3 presents some experiments and results related to forecasting monthly anchovy catches on the north area of Chile. Finally, the conclusions are drawn in the last section.

2 Stationary wavelet transform based Denoising

A signal can be represented at multiple resolutions by decomposing the signal on a family of wavelets and scaling functions [6, 7, 8]. The scaled (approximation) 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 singal 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 to 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 tacked by mean of the un-decimated stationary wavelet transform. The USWT 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 upsampled 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 h filter to obtain the approximation coefficients $a(n)$ and with g filter to obtain the detail coefficients $d(n)$, so that

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

$$d(n) = \sum_k g(n-k)x(k) \tag{3b}$$

Because no subsampling is performed, $a(n)$ and $d(n)$ are of length N instead of $N/2$ as in the DWT case. The output of the SWT $[a(n)\ d(n)]$ is then the approximation coefficients and the detail coefficients. In this paper, the $h(n)$ and $g(n)$ filters are based on Haar wavelet filters, which are given as [11]

$$h = \left[\frac{1}{\sqrt{2}}\ \frac{1}{\sqrt{2}}\right] \tag{4a}$$

$$g = \left[\frac{-1}{\sqrt{2}}\ \frac{1}{\sqrt{2}}\right] \tag{4b}$$

2.1 Wavelet Denoising Algorithm

The wavelet denoising algorithm is based on three stages: (i) the stationary wavelet transform of time series $x(n)$; (ii) thresholding the wavelet coefficients; (iii) the inverse stationary wavelet transform of the thresholding wavelet coefficients to obtain the denoised time series $s(n)$.

According to the original hard thresholding rule with the universal threshold, the wavelet coefficients of $x(n)$ are thresholded by the threshold value given by [9]

$$T = \sigma \sqrt{2log(N)} \tag{5a}$$

$$\sigma = \frac{median(d(n))}{0.6745} \tag{5b}$$

where N is the length of time series $d(n)$ and σ is the noise level.

3 Multivariate polynomial based forecasting

The forecasted signal $s(n)$ is approximated using a reduced multivariate polynomial model, which is obtained as [12]

$$y(u) = \alpha_0 + \sum_{k=1}^{r} \sum_{j=1}^{l} \alpha_{kj} u_j^k + \sum_{j=1}^{r} \alpha_{rl+j} (u_1 + \ldots + u_l)^j + \sum_{j=2}^{r} (\alpha_j^T \cdot u)(u_1 + \ldots + u_l)^{j-1} \tag{6}$$

where r is the order of approximation, u denotes the regressor vector $(u_1, u_2, \ldots, u_l)$ containing l lagged values, and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_K)$ is the parameters vector to be estimated. The K value is the total number of terms of the proposed polynomial and is defined as

$$K = 1 + r + l(2r - 1) \tag{7}$$

In order to estimate the parameters $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_K)$ of the multivariate polynomial forecaster the least squares method is used.

3.1 Least Squares Based Estimation

In order to estimate the linear parameters $\{\alpha_i\}$, the least square algorithm is proposed. Now suppose a set of training input-output samples, denoted as $\{u_{i,k}, s_i, i = 1, \ldots, N_s, k = 1, \ldots, l\ )$, where s_i represent the desired anchovy catches data and u_i denoted the input data, which is obtained as

$$u_i = [s_i(t-1) s_i(t-2) \cdots s_i(t-l)] \tag{8}$$

Then we can perform N_s equations of the form of (6) as follows

$$y = \alpha u \tag{9}$$

The optimal values of the linear parameters α are obtained using the following error objective function

$$E(u, \alpha) = \sum_{i=1}^{N_s} (s_i - y_i)^2 + b \parallel \alpha \parallel^2 \tag{10}$$

where $\parallel \cdot \parallel$ denotes the eucludian norm and b is a regulation constant. Minimizing the error objective function results is

$$\alpha = (H^T H + bI)^{\dagger} H^T y \tag{11}$$

where $H \in R^{N_s \times l}$, $y \in R^{N_s \times 1}$, I is the $(l \times l)$ identity matrix and $(\cdot)^{\dagger}$ is the Moore-Penrose generalized inverse [5].

4 Experiments and Results

The monthly anchovy catches data was conformed by historical data from January 1963 to December 2005, divided into two data subsets as shown in Fig.1. In the first subset, the data from January 1963 to December 1995 was chosen for the training phase (parameters estimation), while the remaining was used for the validation phase. The forecasting process starts by applying the wavelet denoising algorithm and normalization step to the anchovy catches data. Then, the LS learning algorithms is performed to adapts the MP's parameters with denoised historical data. The proposed forecasting structure is based on model order r, regulation constant b and l lagged value. Empirically, we found that a good starting point will be $r = 3$, $b = 0.001$. While maintaining theses values, the lagged values order can be varied within $[1, 12]$ using parsimony principle for best forecasting model. Therefore, for the monthly anchovy catches time series, 12 MP architecture are used to analyze. The MP architecture according to mean squared error (MSE) is given in

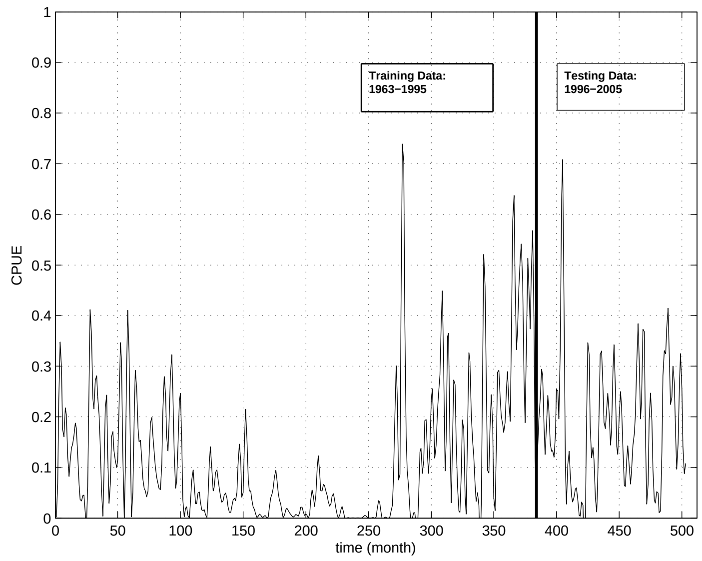

Figure 1: Anchovy catch per unit of effort

Lagged Inputs	MSE	R-Squared
$u_1, u_2, ..., u_{12}$	0.00064	96.59
$u_1, u_2, ..., u_{11}$	0.00061	96.49
$u_1, u_2, ..., u_{10}$	0.00057	96.74
$u_1, u_2, ..., u_9$	**0.00035**	**98.04**
$u_1, u_2, ..., u_8$	0.00038	98.07
$u_1, u_2, ..., u_7$	0.00050	97.36
$u_1, u_2, ..., u_6$	0.00082	95.48
$u_1, u_2, ..., u_5$	0.00115	93.57
u_1, u_2, u_3, u_4	0.00167	91.50
u_1, u_2, u_3	0.00246	87.04
$u_1, u_2,$	0.00449	77.10
u_1	0.00726	61.23

Table 1: Statistic of the forecasting model with LS

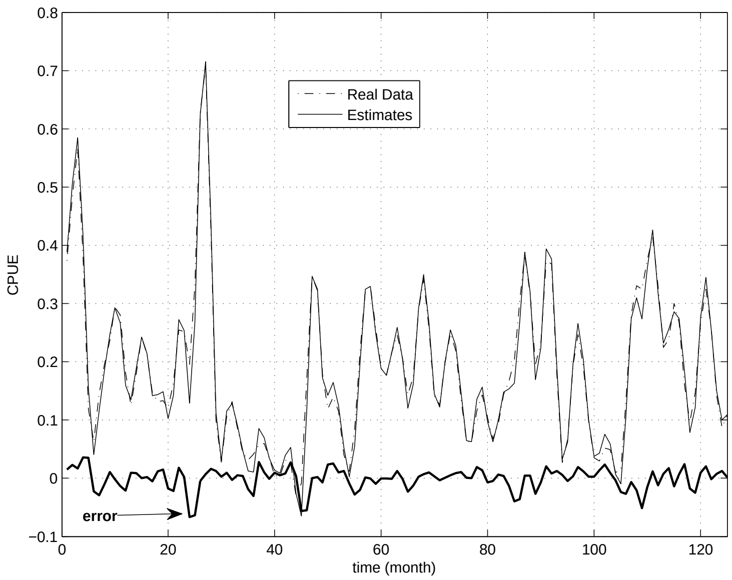

Figure 2: Observed anchovy catches vs estimated anchovy catches with test monthly data

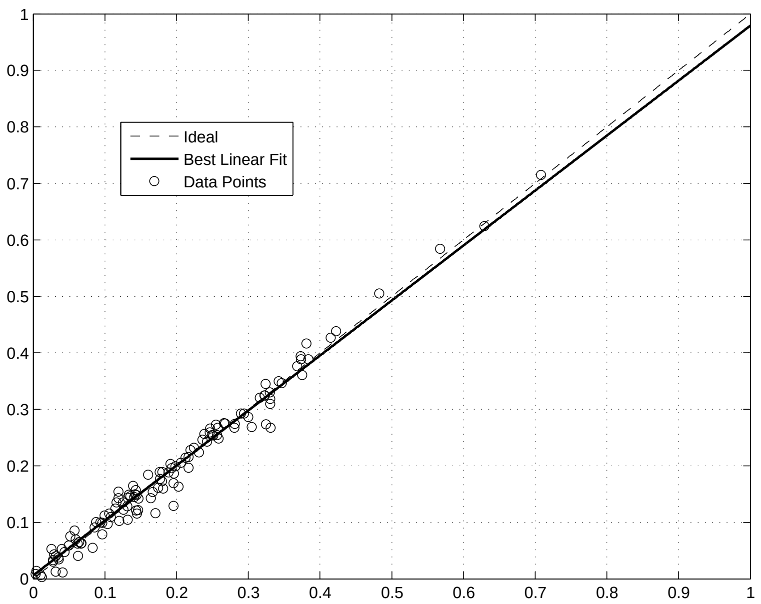

Figure 3: Observed anchovy catches

Table 1. From Table 1, we conclude that best multivariate polynomial forecasting model is based on 9 lagged values and achieves a determination coefficient ($R - squared$) of 98% using smaller number of parameters.

Fig.2 describes the performance evaluation of the validation phase with testing data for the MP(3,9) forecasting model. This plot shows the original time series and the one predicted by proposed forecasting model. The difference is so tiny that it is impossible to tell one from another by eye inspection. That is why you probably see only the MP(3,9) forecasting curve. Besides, from Fig.2 it can be observed the estimated error by the forecasting model

Fig.3 illustrates the determination coefficient estimation of the validation phase between the observed and estimated anchovy catches data with the best forecasting model. Please note that the forecasting model shown that a 98% of the explained variance was captured by the proposed MP(3,9) model.

5 Conclusions

In this paper, one-step-ahead forecasting of monthly anchovy catches based on stationary wavelet denoising technique and multivariate polynomial with penalized least square method was presented. The proposed forecasting model can predict the future CPUE value based on previous values $s(t-1), s(t-2), .., s(t-9)$ and the results found shown that a 98% of the explained variance was captured with a reduced parsimony.

References

[1] Stergiou KI., Prediction of the Mullidae fishery in the easterm Mediterranean 24 months in advance, *Fish. Res.*, vol. 9, pp. 67-74, 1996

[2] Stergiou KI, Christou ED., Modelling and forecasting annual fisheries catches: comparison of regression, univariate and multivariate time series methods, *Fish. Res.*, vol. 25, pp. 105-138, 1996

[3] Hornik K., Stinchcombe M.,White H., Multilayer feedforward networks are universal approximators, *Neural Network*, vol. 2, no. 5, pp.359-366, 1989.

[4] Gutierrez J.C., Silva C., Yaez E., Rodriguez N., Pulido I., Monthly catch forecasting of anchovy engraulis ringens in the north area of Chile: Nonlinear univariate approach, *Fisheries Research*, vol. 86, pages 188-200, 2007

[5] D. Serre, Matrices: Theory and applications.*New York: Springer-Verlag* , 2002

[6] Coifman, R. R. and Donoho, D. L.: Translation-invariant denoising, Wavelets and Statistics, Springer Lecture Notes in Statistics 103, pp. 125-150, Springer-Verlag, 1995.

[7] Nason, G. and Silverman, B.: The stationary wavelet transform and some statistical applications, Wavelets and Statistics, Springer Lecture Notes in Statistics 103, pp. 281-300, Springer-Verlag, 1995.

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

[9] Donoho, D. L. and Johnstone, I. M.: Ideal spatial adaptation by wavelet shrink-age, Biometrika, vol. 81, no. 3, pp.425-455,1994.

[10] Kennedy, J., Eberhart, R.C.: Particle Swarm Optimization. *Proceedings of IEEE International Conference on Neural Networks*, NJ: Piscataway, pages 1942-1948, 1995.

[11] Percival, D. and Walden, A.: Wavelet Methods for Time Series Analysis, Cambridge University Press, Cambridge UK, 2000

[12] Toh, K. -A., Yau, W. -Y., Jiang, X.: A reduced multivariate polynomial model for multimodal biometrics and classifiers fusion, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 2, pp. 224-233, 2004