WAVELET-NETWORK BASED ON L_1-NORM MINIMISATION FOR LEARNING CHAOTIC TIME SERIES

V. Alarcon-Aquino[1], E. S. Garcia-Treviño[1], R. Rosas-Romero[1], J. F. Ramirez-Cruz[2], L. G. Guerrero-Ojeda[1], & J. Rodriguez-Asomoza[1]

[1]Department of Electrical and Electronic Engineering
Communications and Signal Processing Group, CENTIA
Universidad de las Americas, Puebla
72820 Cholula, Puebla MEXICO
[2]Department of Computer Science
Instituto Tecnológico de Apizaco
Tlaxcala MEXICO

Received: April 19th, 2003. Accepted :September 30th, 2005

ABSTRACT

This paper presents a wavelet-neural network based on the L_1-norm minimisation for learning chaotic time series. The proposed approach, which is based on multi-resolution analysis, uses wavelets as activation functions in the hidden layer of the wavelet-network. We propose using the L_1-norm, as opposed to the L_2-norm, due to the well-known fact that the L_1-norm is superior to the L_2-norm criterion when the signal has heavy tailed distributions or outliers. A comparison of the proposed approach with previous reported schemes using a time series benchmark is presented. Simulation results show that the proposed wavelet-network based on the L_1-norm performs better than the standard back-propagation network and the wavelet-network based on the traditional L_2-norm when applied to synthetic data.

RESUMEN

En este artículo se presenta una red neuronal-wavelet basada en la minimización de la norma L_1 para aprendizaje de series de tiempo caóticas. El método propuesto, el cuál se basa en un análisis multi-resolución, utiliza wavelets como funciones de activación en la capa oculta de la red neuronal-wavelet. Se propone utilizar la norma L_1, en lugar de la tradicional norma L_2, debido a que la norma L_1 es superior a la norma L_2 cuando la señal tiene distribuciones sesgadas o de colas pesadas. Se presenta una comparación del método propuesto con esquemas reportados previamente utilizando series de tiempo caóticas conocidas. Los resultados de simulación revelan que la red neuronal-wavelet basada en la norma L_1 tiene una mejor eficiencia que la red neuronal con propagación hacia atrás y la red neuronal-wavelet basada en la tradicional norma L_2 cuando se aplica a datos sintéticos.

KEYWORDS: Wavelet-networks, Wavelets, Multi-resolution Analysis, Learning Chaotic Time Series.

1. INTRODUCTION

Despite the potential of neural networks, there are several problems that remain to be solved. First, it is well known that the back-propagation algorithm for training the multi-layer perceptron, also known as back-propagation network (BPN), suffers for having slow convergence, and the convergence is not normally guaranteed [1 & 9]. Second, the number of hidden units, which determines the structure of the network, is chosen empirically by trial and error. Furthermore, the activation functions are global functions which do not allow local learning or manipulation of the network [1&3]. As a result, there have been some efforts to improve the performance of neural networks (see e.g. [1, 3, 7, 8, 10-14]). The results reported in these works show the advantages offered by combination of neural networks with wavelet representation when compared with general statistical methods and the back-propagation network. Examples of wavelet-networks (WN) applied to time series prediction are presented in [10-14]. Wavelet-networks are a class of neural networks that employ wavelets as activation functions. These have been recently proposed as an alternative approach to the traditional neural networks with sigmoidal activation functions (see e.g., [3, 7, 14]). In [7], a $(1+\frac{1}{2})$-layer neural network based on wavelets is introduced. However, this approach is still using an algorithm of back-propagation type for the learning of the network. In [3], the off-line learning algorithm is based on the traditional L_2-norm, which is basically a non-iterative Least Squares problem and it can be solved using the Moore-Penrose pseudo inverse rule [9]. The L_2-norm may be inaccurate especially where the measurements contain large errors [6]. It is well known that the L_1-norm outperforms the L_2-norm when the signal has heavy tailed distributions or outliers. Generally speaking, the L_1-norm is superior to the L_p-norm (with $p > 1$ i.e., $p = 2, \infty$) criteria if the error distribution has long tails [4&6].

In the work reported in this paper an algorithm is proposed for learning chaotic times series based on wavelet-networks and the L_1-norm minimisation. The results reported show that wavelet-networks have better approximation properties than the back-propagation network when applied to synthetic data. This is due to the fact that wavelets, in addition to forming an orthogonal basis, have the capability to explicitly represent the behaviour of a function with different resolutions of input variables [3]. The rest of this paper is organised as follows. Section 2 presents a review of wavelet theory. Section 3, introduces the wavelet-network based on the L_1-norm minimisation algorithm, and its implementation through linear programming and the dual simplex method. Section 4 presents a brief description of chaotic time series. In Section 5, a comparison of different approaches reported in the literature and the wavelet-network based on the L_1-norm minimisation using synthetic data is presented. Finally, Section 6 presents the conclusions drawn from previous sections and suggests some future work.

2. A REVIEW OF WAVELET THEORY

Wavelet transforms involve representing a general function in terms of simple, fixed building blocks at different scales and positions. These building blocks are generated from a single fixed function called mother wavelet by translation and dilation operations. The continuous wavelet transform considers a family [2 & 5]

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{x-b}{a}\right), \tag{1}$$

where $a \in \Re^+$, $b \in \Re$ with, $a \neq 0$ and $\psi(\cdot)$ satisfies the admissibility condition. For discrete wavelets the scale (or dilation) and translation parameters in Eq. (1) are chosen such that at level m the wavelet $a_0^m \psi\left(a_0^{-m} x\right)$ is a_0^m times the width of $\psi(x)$. That is, the scale parameter $\{a = a_0^m : m \in Z\}$ and the translation parameter $\{b = kb_0 a_0^m : m, k \in Z\}$. This family of wavelets is thus given by [2]

$$\psi_{m,k}(x) = a_0^{-m/2}\psi(a_0^{-m}x - kb_0), \tag{2}$$

so the discrete version of the wavelet transform is

$$d_{m,k} = \left\langle g(x), \psi_{m,k}(x)\right\rangle = a_0^{-m/2}\int_{-\infty}^{+\infty} g(x)\psi(a_0^{-m}x - kb_0)dx, \tag{3}$$

where $\langle \cdot,\cdot \rangle$ denotes the L^2-inner product.

To recover $g(x)$ from the coefficients $\{d_{m,k}\}$, the following stability condition should exist [2],

$$A\|g(x)\|^2 \leq \sum_{m\in Z}\sum_{k\in Z}\left|\langle g(x),\psi_{m,k}(x)\rangle\right|^2 \leq B\|g(x)\|^2, \tag{4}$$

with A>0 and $B<\infty$ for all signals $g(x)$ in $L^2(\Re)$ denoting the frame bounds. These frame bounds can be computed from a_0, b_0 and $\psi(x)$ [2]. The reconstruction formula is thus given by

$$g(x) \approx \frac{2}{A+B}\sum_{m\in Z}\sum_{k\in Z}\langle g(x),\psi_{m,k}(x)\rangle\psi_{m,k}(x). \tag{5}$$

Note, that the closer A and B, the more accurate the reconstruction. When A=B=1, the family of wavelets then forms an orthonormal basis [2].

2.1 Orthonormal Bases and Multi-resolution Analysis

The mother wavelet function $\psi(x)$, scaling a_0 and translation b_0 parameters are specifically chosen such that $\psi_{m,k}(x)$ constitute orthonormal bases for $L^2(\Re)$ [2&5]. To form orthonormal bases with good time-frequency localisation properties, the time-scale parameters (b,a) are sampled on a so-called *dyadic grid* in the time-scale plane, namely, $a_0=2$ and $b_0=1$ [2 & 5]. Thus, from Eq. (2) substituting these values, we have a family of orthonormal bases

$$\psi_{m,k}(x) = 2^{-m/2}\psi(2^{-m}x-k). \tag{6}$$

Using Eq. (3), the orthonormal wavelet transform is thus given by

$$d_{m.,k} = \langle g(x),\psi_{m,k}(x)\rangle = 2^{-m/2}\int_{-\infty}^{+\infty} g(x)\psi_{m,k}(2^{-m}x-k)dx \tag{7}$$

and the reconstruction formula is obtained from Eq. (5). A formal approach to construct orthonormal bases is provided by multi-resolution analysis (MRA) [5]. The idea of MRA is to write a function $g(x)$ as a limit of successive approximations, each of which is a smoother version of $g(x)$. The successive approximations thus correspond to different resolutions [5].

2.2 Discrete Wavelet Transform: Decomposition and Reconstruction

Since the idea of multi-resolution analysis is to write a signal $g(x)$ as a limit of successive approximations, the differences between two successive smooth approximations at resolution 2^{m-1} and 2 give the detail signal at resolution 2^m. In other words, after choosing an initial resolution L, any signal $g(x)\in L^2(\Re)$ can be expressed as [2&5]:

$$g(x) = \sum_{k\in Z} c_{L,k}\phi_{L,k}(x) + \sum_{m=L}^{\infty}\sum_{k\in Z} d_{m,k}\psi_{m,k}(x), \tag{8}$$

where the detail or wavelet coefficients $\{d_{m,k}\}$ are given by Eq. (7), while the approximation or scaling coefficients $\{c_{m,k}\}$ are defined by

$$c_{m.,k} = 2^{-m/2}\int_{-\infty}^{+\infty} g(x)\phi_{m,k}(2^{-m}x-k)dx\,, \tag{9}$$

where $\phi_{m,k}(x)$ denotes the scaling function. Equations (7) and (9) express that a signal $g(x)$ is decomposed in details $\{d_{m,k}\}$ and approximations $\{c_{m,k}\}$ to form a multi-resolution analysis of the signal [5]. Equation (8) is used for the wavelet-network proposed in the following section.

3. PROPOSED WAVELET-NETWORK: L_1 - NORM MINIMISATION

Assume that the training data T for approximation is a set of observations as follows:

$$T = \{(x_i, y_i) : x_i \in \Re, y_i \in \Re, i = 1,2,...,M\}\,,$$

where x_i denotes the input, y_i represents the output and *M* is the number of observations. The algorithm for the L_1-norm may be formulated as a linear programming problem and the idea is to find a design vector $\boldsymbol{\omega} \in \Re^N$ that minimises the energy function $\boldsymbol{J(x)}$, viz.

$$J(x) = \sum_{i=1}^{M} \left|r_i(x)\right|,$$

Where

$$r_i(x) = y_i - \sum_{j=1}^{N} \theta_j(x_i)\omega_j\,, (i = 1,2,...,M)\,. \tag{10}$$

The model represented by Eq. (10) is also called least absolute deviations optimisation model, where $\boldsymbol{\theta} \in \Re^{M\times N}$ ($M \geq N$), $y \in \Re^M$ and $\boldsymbol{\omega} \in R^N$ [6]. The matrix $\boldsymbol{\theta}$ denotes the wavelet and scaling functions, $\boldsymbol{\omega}$ are the weights to be estimated and $\boldsymbol{r} \in \Re^{\boldsymbol{M}}$ is the unknown error vector that account for model errors. Minimisation of the L_1-norm is complicated by the fact that the function $\|\boldsymbol{\theta\omega} - \boldsymbol{y}\|_{\boldsymbol{p}}$ is not differentiable for this value of *p* [4]. However, it may be formulated as a linear programming (LP) problem. Let the matrix $\boldsymbol{\theta}$ comprises the basis functions, that is, wavelet $\psi_{m,k}(x)$ and scaling $\phi_{m,k}(x)$ functions $\boldsymbol{m,k} \in Z$, where *m* and *k* denote the dilation and translation parameters respectively. Following the theory of MRA proposed by Mallat [5], any function $g(x) \in L^2(\Re)$ may be decomposed as in Eq. (8). Equation (8) can be represented in a more compact form, $g(x) = \sum_j \theta_{m,j}(x_i)\omega_{m,j}$ where $\omega_{m,j} \in \left[c_{L,k} d_{m,k}\right]$ denotes the weights to be estimated and $\theta_{m,j} \in \left[\phi_{L,k}\psi_{m,k}\right]$ are the basis functions. Figure 1 shows the structure of a typical wavelet-network for learning chaotic time series.

The scaling functions, $\phi - \boldsymbol{nodes}$, capture the lowest frequency components, and are analogous to the bias value that is present in neural networks. The details of the signal are then captured by adding wavelet functions, $\psi - \boldsymbol{nodes}$. This process of adding wavelet functions constitutes the hierarchical multi-resolution learning [3]. The weights, $\omega_{m,j}$, are then estimated using the L_1-norm minimisation. The minimisation problem

$$\text{minimise}\sum_i \left|y_i - \sum_j \theta_{m,j}(x_i)\omega_{m,j}\right| \tag{11}$$

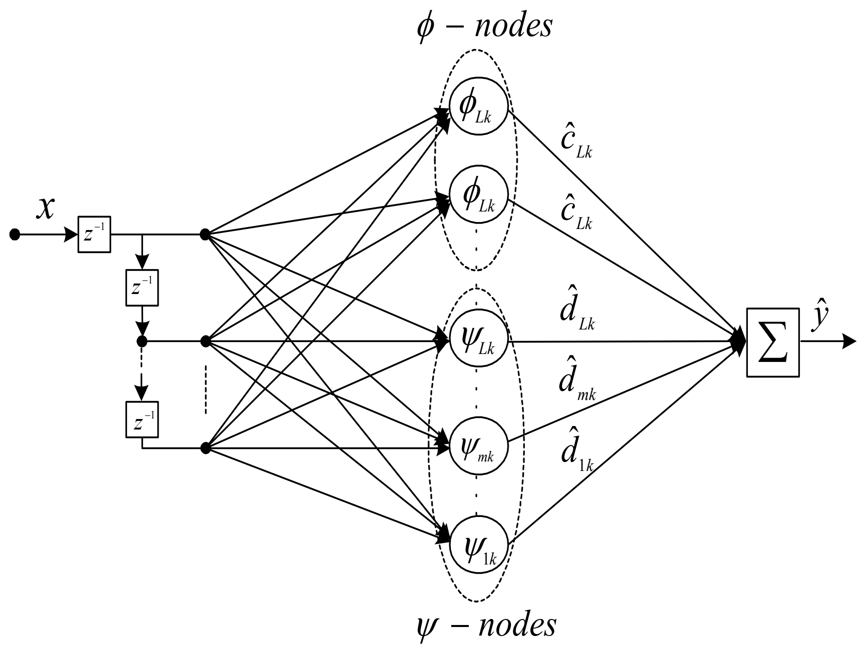


Figure 1. Structure of a wavelet-neural network with tapped delay line (z^{-1} : delay). The hidden layer comprises scaling and wavelet functions, and $\hat{c}$ and $\hat{d}$ denote the estimated scaling and wavelet coefficients respectively

may be converted into a LP problem as follows [6]:

$$\text{minimise} \sum_i \; r_{m,i} \tag{12}$$

$$\text{subject to } r_{m,i} \geq y_i - \sum_j \theta_{m,j}(x_i)\omega_{m,j}, (i = 1,2,\ldots,M),$$

$$r_{m,i} \geq \sum_j \theta_{m,j}(x_i)\omega_{m,j} - y_i ,$$

where $r_{m,i} = y_i - \sum_j \theta_{m,j}(x_i)\omega_{m,j}$ denotes the absolute error with which $\omega^*_{m,j} - optimal$ fail to satisfy the exact solution. In order to solve this problem, Eq. (12) may be written as the standard form of a LP problem by adding some slack variables. Then, the optimal solution of $r^*_{m,i}$ and $\omega^*_{m,j}$ would be [6]

$$r^*_{m,i} = \max\begin{Bmatrix} y_i - \sum_j \theta_{m,j}(x_i)\omega^*_{m,j} \\ \sum_j \theta_{m,j}(x_i)\omega^*_{m,j} - y_i \end{Bmatrix} = \left| y_i - \sum_j \theta_{m,j}(x_i)\omega^*_{m,j} \right| \tag{13}$$

The on-line version of this algorithm is obtained by adding more variables in Eq. (12), viz.

$$\text{minimise} \sum_{i=1}^{M+1} r_{m,i} \tag{14}$$

$$\text{subject to } r_{m,i} \geq y_i - \sum_j \; \theta_{m,j}(x_i)\omega_{m,j}, (i = 1,2,\ldots,M+1) r_{m,i} \geq \sum_j \; \theta_{m,j}(x_i)\omega_{m,j} - y_i$$

and addition of new basis functions introduce a new variable $\omega_{m,N+1}$, thus previous solutions may be used to update wavelet and scaling coefficients. When M is becoming large Eq. (14) may be converted to its dual LP problem (after adding some slack variables), and it may be solved through the dual simplex algorithm [6], viz.

$$\text{maximise} \sum_i v_{m,i} y_i$$

$$\text{subject to} - \sum_i \theta_{m,j}(x_i) + \sum_i \theta_{m,j}(x_i) v_{m,i} = 0, (j = 1,2,...,N) \;\; 0 \le v_{m,i} \le 2, (i = 1,2,...,M) \tag{15}$$

The advantage of using this dual version is that it allows having a sensitivity analysis, that is, previous solutions of the original problem may be used to solve a new problem. Equations (14) and (15) are then used in Section 5 to evaluate the proposed wavelet-network based on the L_1-norm minimisation for learning chaotic time series.

4. DESCRIPTION OF CHAOTIC TIME SERIES

In this section, we briefly review the chaotic time series used in this work. Chaos is the mathematical term for the behaviour of a system that is inherently unpredictable. Unpredictable phenomena are readily apparent in all areas of life [15]. Many systems in the natural world are now known to exhibit chaos or non-linear behaviour, the complexity of which is so great that they were previously considered random. One might argue that the many factors that influence this kind of systems are the reason for this unpredictability. But chaos can occur in systems that have few degrees of freedom as well. The critical ingredient in many chaotic systems is what mathematicians call sensitive dependence to initial conditions. If one makes even slightest change in the initial configuration of the system, the resulting behaviour may be dramatically different [15]. The chaotic time series are generated by a logistic map function. This function was explored by ecologists and biologists who used it to model population dynamics. It was popularized by Robert May in 1976 as an example of a very simple non-linear equation being able to produce very complex dynamics. The mathematical equation that describes the logistic map function is given by

$$x(i+1) = \alpha x(i)(1 - x(i)), \tag{16}$$

where $\alpha = 4$, and the initial condition $x(0) = 0.000104$. It is worth noting that for the logistic map function, the specific parameters are selected because they are the more representative values for chaotic behaviour, and also note that it is the most commonly used in the literature (see e.g., [1, 3, 8, 15, 16]).

5. SIMULATION RESULTS

This section presents a comparison between the L_2-norm reported in [3] and the L_1-norm for the wavelet-network described in Section 3. The orthogonal Battle-Lemarié wavelet is used as basis functions (see Figure 2). We select the Battle-Lemarie wavelet due to the well-known fact that it has a high degree of smoothness. This important characteristic allows a better frequency localisation and thus a better approximation of more complicated functions [2]. Furthermore, a comparison has been done with other reported approaches such as BPN [9] and wavelet-networks [7 & 8]. It is worth mentioning that the approach reported in [7] needs to estimate not only the weights but also the translation and dilation parameters using an algorithm of back-propagation type, whereas the approach presented here the weights are estimated only, and the dilation and translation parameters follow a regular grid in the input space.

The performance of the different learning algorithms is evaluated using a time series generated by the logistic map function defined in Eq. (16) which contains periodic and aperiodic cycles. The first 300 points are used for training, and generalisation is done with the next 200 points. The commonly used notation for BPN $1 \times 6 \times 1$ denotes the number of inputs, hidden units and outputs, respectively. The learning rate for the BPN was set to 0.1 and the momentum term was set to 0.07. The sigmoid function is considered as a basis function. For the wavelet-networks two levels of resolution are considered. Since the resolution levels depends on the wavelet itself, the levels of resolutions are chosen according to global and local approximation errors for the training data. New $\psi - nodes$, added to the wavelet-network, are trained to minimise the error from the coarser resolutions. This procedure is continued until the wavelet-network satisfies the performance requirements. The computational complexity of this procedure depends on the total number of training data. Simulation results show that the first two resolution levels reduce considerably the global and local approximation errors. Note that the weights of $\phi - nodes$ and $\psi - nodes$ at other resolution levels are unaffected by the addition of new $\psi - nodes$ allowing independent training of each node [3]. Figure 3 shows the prediction of the logistic map function using the proposed wavelet-network based on the L_1-norm minimisation, and the mean square error (MSE) at each multi-resolution stage. This figure shows that the proposed approach has a good prediction and generalisation performance on unseen data. Note that the optimal wavelet-network may be designed by removing nodes with small contribution to the approximation of the unknown function. This is due to the fact that the orthonormality allows removal of nodes without retraining [3].

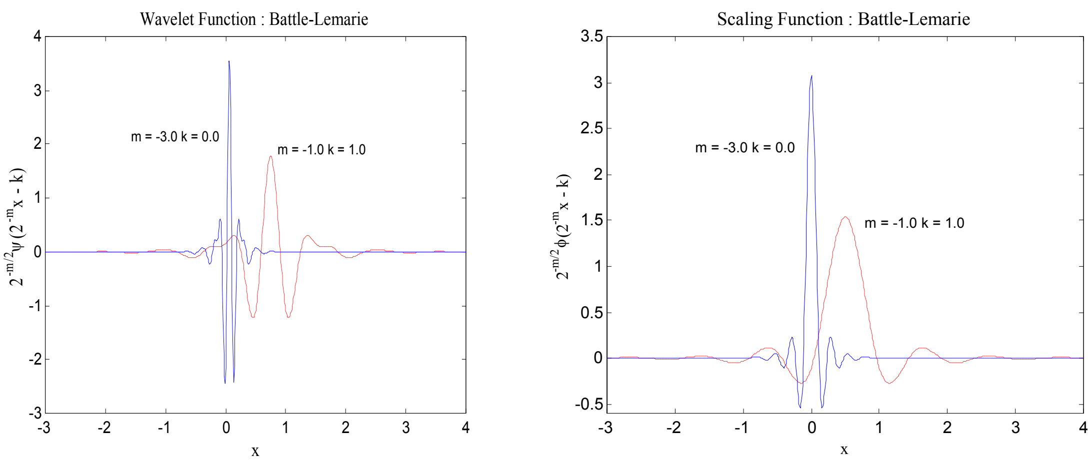
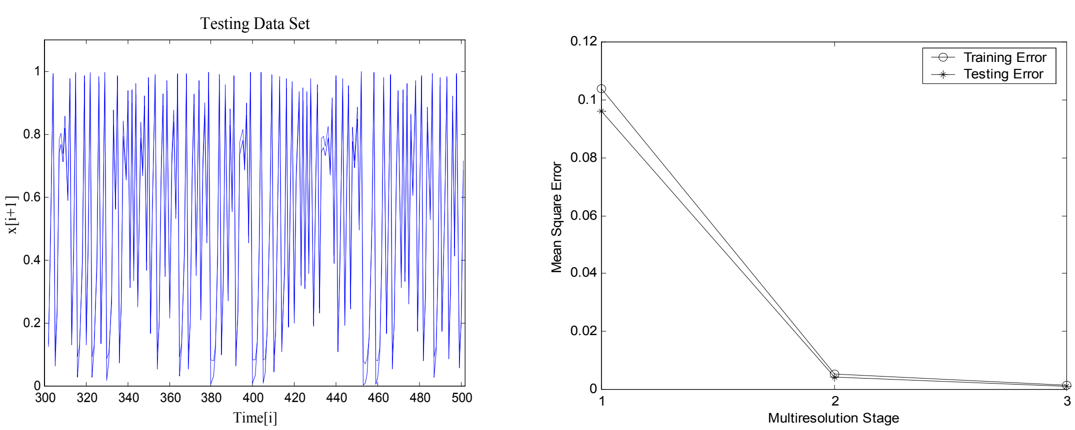


Figure 2. Example of the Battle-Lemarié wavelet and scaling function with different dilations and translations parameters



(a) (b)

Figure 3. (a) Prediction of the testing data set using the wavelet-network based on the L_1-norm with two resolution levels ($m = 0, -1$). Actual testing data set (solid line), prediction network output (dotted line). (b)Training error and testing error for each multi-resolution stage.

Table I shows the performance of different approaches, and the approach proposed in this paper. It can be seen that wavelet-networks outperform BPN, and this is due to the fact that neural networks based on conventional single-resolution schemes cannot learn complicated time series, and consequently the training process often converge very slowly and the trained network may not generalise well. This is also due to the fact that wavelets, in addition to forming an orthogonal basis, have the capability to explicitly represent the behaviour of a function a different resolutions of input variables. That is, a wavelet-network is first trained to learn the mapping at the coarsest resolution level and then trained to include elements of the mapping at higher resolutions until the desired level of generalisation has been reached [3]. This hierarchical multi-resolution training can result in a more meaningful interpretation of the resulting mapping and adaptation of networks that are more efficient than conventional methods [14]. Furthermore, note that the proposed approach performs better than previous reported wavelet-networks schemes [3, 7, 8] for the testing data set. The wavelet-network based on the L_1-norm performs better than the L_2-norm. This is due to the well-known fact that the L_1-norm has advantages in situations where the data are erratic (presence of outliers, long-tailed error distributions or the error distribution is not well known) [4]. The wavelet-network based on the L_1-norm eliminates the need for the user to guess the size of the network. That is, the network begins learning by having only scaling functions and new wavelet nodes are added until the optimality criterion or tolerance is satisfied [3]. Furthermore, the learning is faster, due to the fact that only a single layer (either scaling or wavelet nodes) is trained at any given time and there is no need to propagate errors backwards as in BPN.

Table I
Performance Comparison

MSE			
Approaches	Model	Training Set	Testing Set
BPN[9]	1-6-1, 5000 Iter.	0.005295	0.006491
WN [7], [8]	1-6-1, 5000 Iter.	0.001112	0.001073
WN-L_2 norm [3]	2 ϕ –nodes, 4 ψ -nodes	0.001443	0.001201
Proposed WN-L_1 norm	2 ϕ -nodes, 4 ψ -nodes	0.001245	0.000974

6. CONCLUSIONS

An approach for learning chaotic time series based on wavelet-networks and the L_1-norm minimisation has been proposed. This approach has been evaluated using a chaotic time series generated by a logistic map function. The results show that the wavelet-network proposed in this paper performs better than previous reported approaches [3, 7-9] when applied to synthetic data. Note that wavelet-networks have better approximation properties than the back-propagation network. This is due to the fact that wavelets, in addition to forming an orthogonal basis, have the capability to explicitly represent the behaviour of a function a different resolutions of input variables. The wavelet-network based on the L_1-norm minimisation eliminates the need to select the number of hidden units before the training phase begins. This hierarchical nature enables each wavelet to model the residual error from the previous approximation. That is, a wavelet-network can first be trained to learn the mapping at the coarsest resolution level and then trained to include elements of the mapping at higher resolutions until the desired level of generalisation is reached [3]. This hierarchical multi-resolution training can result in a more meaningful interpretation of the resulting mapping and adaptation of networks that are more efficient than conventional methods [14].

The wavelet selected as basis function for the studied approach plays an important role in the approximation process. The selection of the wavelet depends on the type of function to be approximated. Based on the analyzed chaotic time series, a wavelet with good frequency localization like the Battle-Lemarié showed better approximation results when tested with the logistic map function. This is due to the fact that the Battle-Lemarié wavelet has high degree of smoothness. This important characteristic allows a better approximation of more complicated functions. Further work needs to be done to evaluate the performance of the proposed approach under different scenarios of data streams. At present, the wavelet-network based on the L_1-norm minimisation is being extended to recurrent neural networks and other types of wavelets.

ACKNOWLEDGEMENTS

This work was carried out during the first author stay at Imperial College London, UK, and was supported by the National Council for Science and Technology (CONACYT), MEXICO.

REFERENCES

[1] Safavi, A. A., Romagnoli, J. A. Application of Wave-Nets to Modelling and Optimisation of a Multidimensional Chemical Process. IEEE International Conference on Neural Networks, Vol. 4, pp. 1724-7128, (1995)

[2] Daubechies, I. Ten Lectures on Wavelets. New York: SIAM, (1992).

[3] Bakshi, B. R. and Stephanopoulus, G. Wavelets as Basis Functions for Localised Learning in a Multi-resolution Hierarchy, in Proceedings of International Joint Conference on Neural Networks, pp. II/140-II/145, Baltimore. (1992)

[4] Bloomfield, P. and Steiger, W. L., Least Absolute Deviations:Theory, Applications, and Algorithms, Edited by P. Huber and M. Rosenblatt, Progress in Probability and Statistics, Birkhauser, (1983)

[5] Mallat, S. A Theory for Multi-resolution Signal Decomposition: The Wavelet Representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, No. 7, pp. 674-693, (1989)

[6] Chvatal, V. Linear Programming, W.H. Freeman and Company, (1983).

[7] Zhang, Q. and Benveniste, A. Wavelet Networks, IEEE Transactions on Neural Networks, Vol. 3, No. 6, pp. 889-898 (1992)

[8] Chen, Z., Feng, T. J. and Meng, M. C., The Application of Wavelet Neural Network in Time Series Prediction and System Modelling Based on Multi-resolution Learning, IEEE SMC'99 Conference (Systems, Man. and Cybernetics), Vol. 1, pp. 425-430 (1999)

[9] Haykin, S., Neural Networks: A Comprehensive Foundation, MacMillan, (1994)

[10] Kristea, P., Tuduse, R., Cristea, A., Time Series Prediction with Wavelet Neural Networks, 5th Seminar on Neural Network Applications in Electrical Engineering Faculty of Electrical Engineering. University of Belgrade, Yugoslavia. (2000)

[11] Rying, E. A., Bilbro, G. L. and Lu, J.-C., Focused Local Learning With Wavelet Neural Networks, IEEE Transactions on Neural Networks, Vol. 13, No. 2, March (2002).

[12] Galvão, R. K. H. and Yoneyama, T., A Competitive Wavelet Network for Signal Clustering, IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, Vol. 34, No. 2, April (2004)

[13] Zhiguo Z., Ye S., Adaptive Wavelet Neural Network for prediction of Hourly Nox and No2 Concentrations, Proceedings of the 2004 Winter Simulation Conference Simulation Center, Harbin Institute of Technology, CHINA, (2004).

[14] Sitharama Iyengar, S., Cho, E. C., Phoha, Vir V., Foundations of Wavelet Networks and Applications, Chapman & Hall, CRC USA, (2002).

[15] Devaney, R. L., Chaotic Explosions in Simple Dynamical Systems, The Ubiquity of Chaos, Edited by Saul Krasner. American Association for the Advancement of Science. Washington D.C., USA, (1990).

[16] Pritchard, J., The Chaos CookBook: A Practical Programming Guide, Part of Reed International Books. Oxford, UK, (1992).

Authror's Biography