Long-Term Prediction, Chaos and Artificial Neural Networks. Where is the Meeting Point?

Pilar Gómez-Gil

***Abstract*—**This paper presents the advances of a research using a combination of recurrent and feed-forward neural networks for long term prediction of chaotic time series. It is known that point-to-point, long term prediction for chaotic time series is not possible; however, in this research we are looking for ways to build dynamical systems using artificial neural network, that contain the same characteristics as the unknown systems producing a time series. The created systems should be such that they have chaotic invariants similar to the ones presented in the time series. Being able to create such system, the bounds of the predictions may be defined. We present some general concepts related to chaos, dynamical systems and artificial neural networks and present one model that, if not able to predict yet, is able to autonomously oscillate in bounded limits and in a chaotic way. Some experimental results and future work is also presented.

***Index Terms*—** Chaotic signals, hybrid complex neural network (HCNN), Lyapunov exponents, recurrent neural networks.

I. Introduction

The possibility of accurate prediction is still impossible for some applications, as weather, stock values or some biological signals. This is because they behave chaotically. However, in many of these dynamical systems, it could be very useful to have a system able to predict if catastrophic events may occur.

Therefore the creation of a model with dynamical characteristics similar to the unknown model generating the time series produced by these systems could help in the prediction of such systems.

II. Deterministic Chaos and Long Term Prediction

A. *Some definitions*

Deterministic chaos is understood as an aperiodic behavior very sensitive to initial conditions. In the last years it has been found that many non-linear systems present a chaotic behavior. There is not a universally agreed definition of chaos. According to [1] some characteristics of deterministic chaos are:

- Chaotic trajectories are aperiodic and deterministic
- Chaotic systems are extremely dependant on initial conditions. Therefore small uncertainties in the initial state will grow exponentially fast average
- Chaotic behavior is bounded and presents strange attractors
- There is a particular pattern associated with chaotic behavior

B. *Non-Linear Dynamical Systems*

A dynamical system is one that changes with time. A non linear dynamical system can be represented by the equation:

$$\frac{d\mathbf{y}(t)}{dt} = \mathbf{F}(\mathbf{y}(t)), \quad \mathbf{y}(t) = [y_1(t), y_2(t), \dots y_d(t)], \qquad (1)$$
$$\mathbf{y}(0) = \mathbf{y}_0.$$

Vector **F** is nonlinear; d is the dimension of the system. This equation describes the motion of a point in a d-dimensional state space known as phase space.

A *trajectory* of the system is a curve of $\mathbf{y}(t)$ drawn in the phase space, showing the solution of (1) for a specific initial condition. A family of trajectories for different initial conditions is called a *phase portrait*. The behavior of a dynamical system can be visualized using a plot called *return map*, which is a representation of the state space of the system. A return map shows the relation between a given point in a time series with other points further away in time.

In the literature, many times nonlinear dynamical systems and chaotic systems are taken as the same. All chaotic systems are nonlinear, but not all nonlinear dynamical systems are chaotic.

C. *Chaotic Systems and Time Series*

Given a time series that represents a dynamical system, it is possible to find out the degree of nonlinearity in such system, using some numerical methods to calculate the invariants of the system [2]. The asymptotic behavior of a dynamical system when $t \rightarrow \infty$ is known as its *steady state*. A steady state y* is a set of values of the variables of a system for which the system does not change as time proceeds. That is:

$$\frac{d\mathbf{y}(t)}{dt} = 0. \qquad (2)$$

Some systems do not reach fixed points as $t \rightarrow \infty$, rather they oscillate. The solution to their differential equation is

Manuscript received June 30, 2006. This research was partially supported by the Department of Computer Science at Texas Tech University, the Dean Office of Research and Graduate Studies at the Universidad de las Americas, Puebla and the Mexican National Council of Science and Technology under grant No. 132900-A.

Pilar Gómez-Gil is with the Department of Computation, Electronics Physics and Innovation at Universidad de las Américas, Puebla. Phone: +52-222-229-2029 (e-mail: pgomezexttu@gmail.com).

periodic, known as a *limit cycle*. A *limit set* is defined as the set of points in the state space that a trajectory of the system repeatedly visits. The points of the limit cycle form the limit set.

An *attractor* is a set of points toward which a trajectory approaches after its transient state dies out [3]. Equilibrium points or fixed point and limit cycles are attractors. A *basin of attraction* is a set of initial conditions for which the system approaches the same attractor as $t \to \infty$.

When the steady state behavior of a system is bounded but not an equilibrium point or a limit cycle then the system is said to be chaotic. The geometrical object in state space to which chaotic trajectories are attracted is called a strange attractor. A strange attractor presents one or more injection regions, which are zones of the phase space with a very large number of trajectories crossing along them. Such closeness in the trajectories makes the system to "jump" from one trajectory to other, generating what is known as chaos.

Through the years, the existence of chaos has been characterized in time series using several methods, among others: analysis of Fourier spectra, fractal power spectra, entropy, fractal dimension, and calculation of Lyapunov exponents. However, several of these methods have proven not to be very efficient. Even though, the calculation of Lyapunov exponents has been a common way to determine if a time series resulting from an unknown dynamical system is chaotic.

D. Lyapunov Exponents

The *Lyapunov exponents* (LE) $\lambda_1, \lambda_2 \ldots \lambda_d$ are the average rates of expansion ($\lambda_i > 0$) or contraction ($\lambda_i < 0$) near a limit set of a dynamical system [4]. In other words, the LE are a quantitative measure of the divergence of several trajectories in the system. There is one LE for each direction in the *d*-dimensional phase space where the system is embedded. The variable *d* represents the Lyapunov dimension or phase space dimension. The LE's are *invariant measures*; that is, they do not change when the initial conditions of a trajectory are modified or if some perturbation occurs in the system. If at least one LE is positive, the system presents chaotic motion; hence it is very dependent on initial conditions. If that is the case, the magnitude of the LE reflects the time scale on which the dynamics of the system become unpredictable.

There are several algorithms developed to calculate LE from time series. Of special interest for our research is the one developed by Gencay and Decher [5] that estimates all the LEs of an unknown dynamical system using a technique based on a multivariate feed-forward neural network. The main idea of this algorithm is to estimate a function from the observations of the system that is topologically equivalent to the one describing the system, and then calculate the LE of that function.

E. Long-term prediction

Forecasting consists of the evaluation of future values of a time series, x_{T+h}, $h \geq 1$, based on the observations of its past values $x_1, x_2 \ldots x_T$ [6]. In general, prediction may be seen as a function approximation problem [7] which consists of defining a complicated function $f(\mathbf{y})$ using other function $\tilde{f}(\mathbf{y})$ that is a combination of simpler functions:

$$\tilde{f}(\mathbf{y}) = \sum_{i=1}^{N} a_i \varphi_i(\mathbf{y}). \tag{3}$$

If the bases functions $\varphi_i(\mathbf{y})$ are a linear combination of past outputs and/or inputs, the model is said to be *linear*. When the bases are nonlinear with respect to the past signal, the model is *non-linear*. For many years linear models have been the most popular tools to predict, but obviously they are not able to represent accurately non-linear dynamical systems.

A *predictor map* defined by (3) can be written as:

$$\mathbf{y}(t+1) = \mathbf{F}(\mathbf{y}(t), \mathbf{a}) \tag{4}$$

Single-step prediction takes place when several observations of past values are used by the predictor to calculate the next point. This is also called *next-point prediction* or *one-point prediction*. In this case the predictor map is applied once: $\mathbf{y}(t+1) = \mathbf{F}(\mathbf{y}(t), \mathbf{a})$ .This kind of prediction is normally carried out by *non-autonomous predictors*, where each prediction of time *t* requires as external inputs observations at times *t-1, t-2… t-d+1*. It is also possible to carry out single-step prediction using an *autonomous predictor*, which is able to fully represent the solution to the dynamic of the system, and therefore it only requires as external input the initial conditions of the system. In this case, from time *t=0,* the predictor will calculate the output at any time *t* without using any external inputs.

Point-by-point prediction takes place when, in an iterative way, the predictor calculates outputs at times *t, t+1, t+2* and so on. Here the predictor map is applied several times:

$$\begin{aligned}&\mathbf{y}(t+k) = \mathbf{F}(\mathbf{y}(t+k-1), \mathbf{a}) = \\ &\mathbf{F}(\mathbf{F}(\mathbf{y}(t+k-2), \mathbf{a}), \mathbf{a}) = \ldots = \mathbf{F}^k(\mathbf{y}(t), \mathbf{a})\end{aligned} \tag{5}$$

A non-autonomous predictor will require feedback from its own predictions to calculate new values when the value to be predicted is such that there are no more available observations. Point-by-point prediction is required for long-time prediction.

A non-linear model can not be accurately described by a transfer function. However, a model equivalent to linear case can be constructed for non-linear systems using the embedding theory of Takens [8]. First it is required to reconstruct an embedding space using the available time series, with characteristics equivalent to the original system. With this it is possible to define a map that transforms from the current reconstructed state of the trajectory to the next state.

Given:

$$\mathbf{y}(n) = (x(n), x(n+\tau), x(n+2\tau) \ldots x(n+(d-1)\tau) \tag{6}$$

where $x(n)$ is the time series, *d* is the embedding dimension, and τ the time lag, a map $\mathbf{F} : \Re^{\mathbf{d}} \to \Re^{\mathbf{d}}$, parameterized by $\mathbf{a} = (a_1, a_2, a_3 \ldots a_p)$, can be constructed such that

$$\mathbf{y}(n+1) = \mathbf{F}(\mathbf{y}(n), \mathbf{a}). \tag{7}$$

Data resulting from a non-linear dynamical system contain invariant information that is essential to describe the geometrical structure of its attractor. A way to construct a predictor map is calculating such invariants from the data and then imposing them as constraints on the calculation of parameters, in a way that the dynamical system defined by **F** has similar invariants to the unknown system.

It is important to remark that reliable point-to-point prediction of chaotic systems with unknown dynamics is impossible [9]. Indeed, there is not a theory for recognizing whether a constructed predictor has been able to truly identify the original system.

III. The Hybrid Complex Neural Network

A. *Some previous work*

There are several studies related to modeling and prediction of non-linear time series using neural networks. For example, Hayashi [10] analyzed the behavior of an oscillatory network with external inputs. His network is made of excitatory and inhibitory neural groups. Each excitatory cell is connected to an inhibitory cell and to other excitatory cells. Hayashi observed that, when the external inputs to the network were similar to a memory pattern, the network generated a limit cycle near such a pattern. For an input far from the memory patterns, a chaotic orbit was generated. Príncipe and Kuo [12] studied a dynamic modeling of chaotic time series using a recurrent neural network with a global feedback loop. Their network was trained using back- propagation through time. They proposed to use dynamic invariants as a measure of the success of the predictor, instead of a global error. Recurrent neural networks have shown to be crucial for activities involving non-linear dynamics and especially for chaos.

Logar [13] showed that a 3-node fully-connected recurrent neural network (RNN) is able to oscillate; hence it may capture the dynamics of sine waves and work as an autonomous predictor. Once trained, these kind of networks can accurately predict point-by-point fairly well during long periods of time, using no external inputs except the initial point of the signal. Figure 1 shows this kind of network, that we named *harmonic generator* [15]. Based on this oscillation ability, Oldham [14] developed a predictor model which consists of a fully-connected RNN pre-loaded with information about the Fourier components of the signal to be learned. Such a model is known as the complex network.

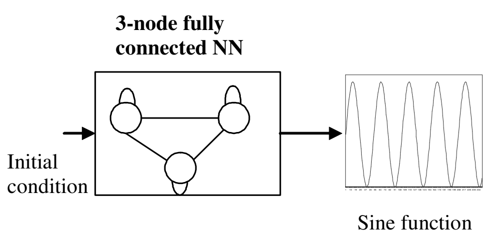


Fig. 1 A Harmonic Generator

B. *Architecture and Training of HCNN*

The "Hybrid Complex Neural Network" (HCNN) [16] is based on several harmonic generators, connected to other neurons using feed-forward and recurrent connections. Figure 2 shows a HCNN.

Training of the HCNN includes a mechanism to obtain information related to the chaotic dynamics of the training signal. This is done by introducing in the network information about Lyapunov exponents. To do so, we use a procedure similar to the one proposed at [5] which obtains an approximation of the Lyapunov exponents of a time series, based on a feed-forward neural network. Such neural network calculates a function that is topologically equivalent to the one describing the dynamical system. That network gets as input some past values of the series, generating the next value as output.

Training HCNN to produce a particular time series is done as follows:

1. Calculate the Fourier transform of a portion of the time series and obtain it fundamental frequency. The size of such portion should include at least one "pseudo" cycle of the series (is the time series is chaotic, there are not real cycles on it)
2. Using the fundamental frequency, train a number of harmonic generators to produce sine functions with frequencies multiples of such fundamental. The number of harmonic generators depends on the desired size of the HCNN
3. Embed the trained harmonic generator in a HCNN, which is made of n input neurons in the first layer, m harmonic generator in the second layer, r hidden nodes in the third layer and 1 node in the output layer. The topology includes feed forward connections between each layer. Only one node in each harmonic generator is connected to the next layer.
4. Keeping fixed the harmonic generators weights, train the feed-forward weights using the algorithm back propagation through time (BPTT) [17]. This training feeds input nodes with (p-1) values of the tine series and use p values to calculate the error to propagate. This step creates the part of the system that poses similar invariant characteristics to the expected dynamical system.
5. Train again the whole weights of the network using (BPTT).

The dynamics of each neuron in the network is given by:

$$\frac{dy_i}{dt} = -y_i + \sigma(x_i) + I_i \qquad (8)$$

where:

$$x_i = \sum_j w_{ji} y_j \qquad (9)$$

x_i represents the input to the i-th. neuron coming from other neurons, I_i is an external input to i-th. neuron, w_{ji} is the weight connecting neuron i to neuron j and $\sigma(x)$ is an arbitrary differentiable function, commonly a sigmoid. I_i is used only in the input layer of the HCNN.

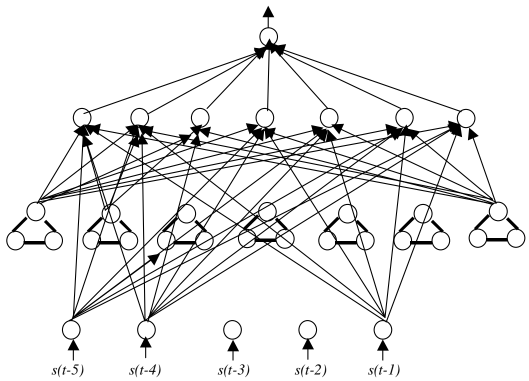


Fig. 2 An example of a Hybrid Complex Neural Network with 7 harmonic generators, 5 inputs and one output

IV. Results

The behavior of the Hybrid Complex Neural Network has been tested with several chaotic time series [16]. Here we present the results obtained when 512 points of an Electrocardiogram were fed to the network. Figure 3 shows such time series and Figure 4 its return map.

As a metric of the performance of the network we used the Maximum Square Error (MSE) as well as the maximum LE λ_1 of the original and predicted time series. LE was calculated using our own implementation of the method described at [17]. It is important to point out that, as with other numerical algorithms, this is strongly dependent of the size of data and accuracy of input parameters.

To calculate λ_1 for the time series used and predicted in this experiment we applied the following input parameters: Number of inputs points = 512 or 1,536 (size of the signal), embedded dimension = 3 (as suggested in [18]), time delay = 15 (calculated using auto-correlation function), time period of data = 0.028, maximum distance to look for neighbors = 0.008 and minimum distance = 1.0e-05. We obtained a Maximum LE of 3.23±0.027 for this time series. At the time when we developed these experiments there was not an agreement in the literature about what was the real maximum LE of an electrocardiogram. Values as low as 0.34±0.08 [18] or as large as 29.0 [19] can be found.

A HCNN with 5 input neurons, 7 harmonic generator, 7 hidden neurons and 1 output neuron (Fig. 2) was trained using 512 points of the ECG with 30,000 sweeps. A MSE of 2.5E-3 was obtained. Figure 5 shows the original and predicted signal obtained in this experiment. Figure 6 shows the predicted signal by itself.

We calculated the maximum LE for two different segments of the predicted signal: in the first 512 points and in the complete 2048 predicted points. This was because the prediction of the first 512 points was made using original data as input to the network, and the rest of the signal (1,536 points) was predicted using predicted data as input to the network. The maximum LE was 4.88±2.21 using the first 512 points, and 7.52±1.95 using the 2,048 points.

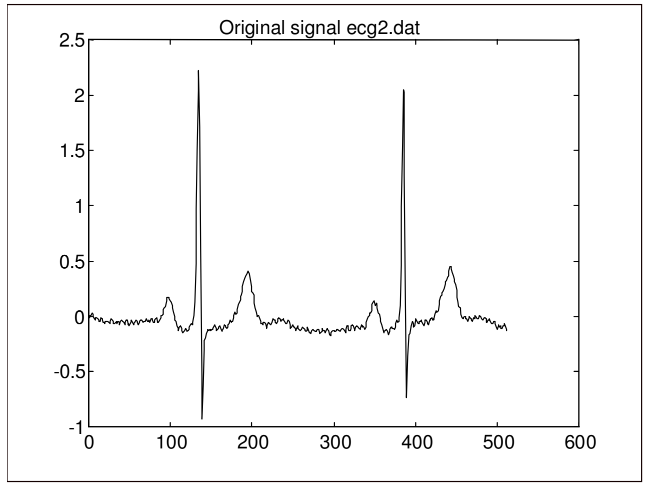


Fig. 3. The ECG used in this experiment

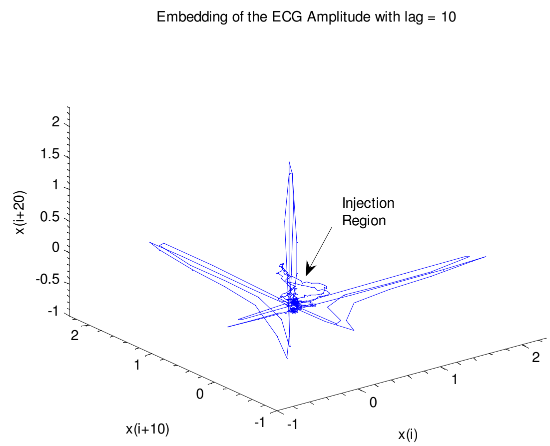


Figure 4. A Return map of the time series at figure 3.

Notice that the predicted output is enclosed to an upper value; that is, the network oscillates in an autonomous way, using their own predicted data and keeping their output with no divergence. This behavior is not seen in other linear predictors, or with predictors based on feed- forward neural networks, as shown at [16]. It is also noted that magnitude of the predicted signal is not as expected, especially in the peaks of the ECG. It can also be noticed that values obtained by topology B are noisier than the obtained by topology A.

Figure 7 shows a return map of the predicted signal, where compared to figure 4 the attractor is missing the peak due to the short peak values of the ECG in the predicted signal.

V. Conclusions and Future work

A model combining recurrent and feed forward network to predict chaotic time series, called the hybrid complex neural network, is presented in this research. The result obtained using a HCNN for prediction of electrocardiograms are not satisfactory yet, However, HCNN is able to oscillate in a bounded and autonomous way, and generates a signal with positive LE.

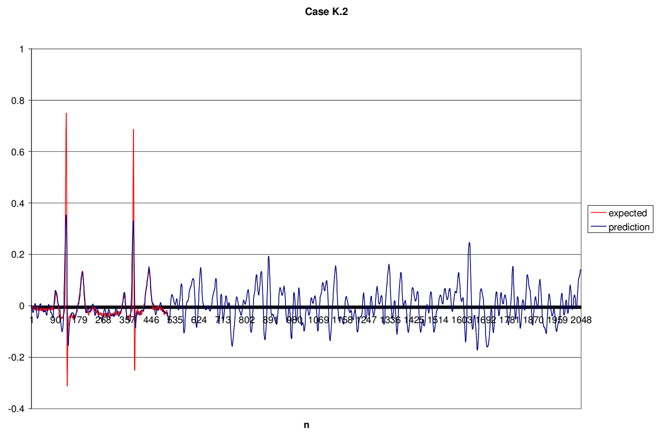


Figure 5. Original and predicted ECG

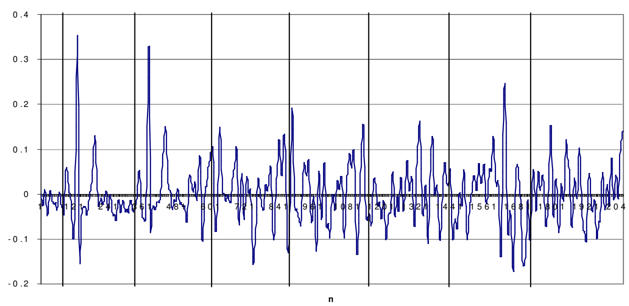


Figure 6. Predicted ECG using HCNN

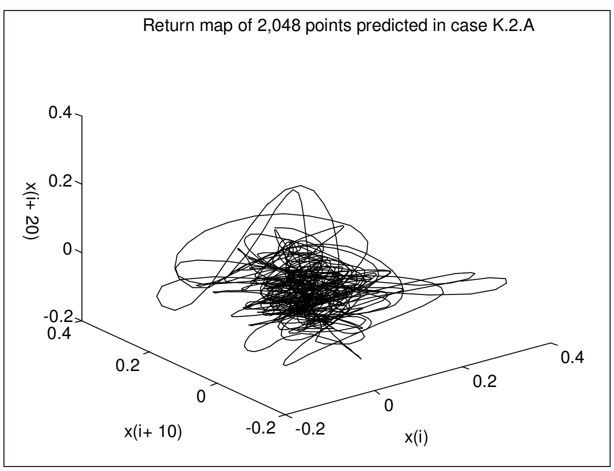


Figure 7. A Return Map of Predicted ECG

These results are encouraging to improve the model. Future work in this model may include the use of phase information during the training of the harmonic generators and the use of some variant of teacher forcing during training. Also the use of ensemble forecasting by training several HCNN in parallel that collaborate among them may contribute to improve the results.

ACKNOWLEDGMENT

The author thanks Dr. Brain Oldham for their invaluable help during the direction of part of this project.

REFERENCES

[1] D. Kaplan and R. J. Cohen. "Is Fibrillation Chaos?" *Circulation Research,* Vol. 67, No. 4, October 1990.

[2] S. Sharma. "An Exploratory Study of Chaos in Human – Machine System Dynamics" in *IEEE Transactions on Systems, Man and Cybernetics. Part A: Systems and humans.* Vol. 36, No. 2, March 2006

[3] L. Glass and M. C. Mackey. *From Clocks to Chaos*, Princeton University Press, Princeton, New Jersey, 1988.

[4] T.S. Parker and L.O. Chua. *Practical Numerical Algorithms for Chaotic Systems*, Springer-Verlag, New York, 1989.

[5] R. Gencay and W. D. Dechert. "An algorithm for the n Lyapunov exponents of an n-dimensional unknown dynamical system," *Physica D.* Vol. 59, pp. 142-157, 1992.

[6] P. J. Brockwell and R. A. Davis. *Time Series Theory and Methods, Second Edition.* Springer Editors, New York, 1991.

[7] J. Príncipe, L. Wang and Jyh-Ming Kuo. "Non-linear Dynamic Modeling with Neural Networks," *Proceedings of the First European Conference on Signal Analysis and Prediction*, 1997

[8] Takens, F. "Detecting Strange Attractors in Turbulence," Dynamical Systems and Turbulence, edited by D. A. Rand and L. S. Young, Springer-Verlag, Berlin, 1981

[9] L. Wang and D. Alkon. *Artificial Neural Networks: Oscillations, Chaos and Sequence Processing*, IEEE Computer Society Press, Washington DC, 1993.

[10] Y. Hayashi. "Oscillatory Neural Network and Learning of Continuously Transformed Patterns," *Neural Networks*, Vol. 7, No. 2, pp. 219-231, 1994

[11] R. Williams and D. Zipser. "Experimental Analysis of Real-time Recurrent Learning Algorithm," *Connection Sciences*, Vol. 1, No. 1, 1989

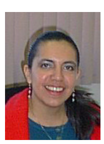
[12] J. Príncipe and J. Kuo. "Dynamic Modeling of Chaotic Time Series with Neural Networks," Advances in Neural Information Processing Systems 6, edited by T. Cowan and A. Koufmann, pp. 311-318, 1994

[13] A. Logar. *Recurrent Neural Networks and Time Series prediction,* Ph.D. Dissertation in Computer Science, Texas Tech University, Lubbock, TX, December 1992

[14] W. Oldham, W. and P. Gómez-Gil. "Modeling and Prediction of Time Series Using Recurrent Neural Networks: an Application to ECG," *Proceedings of the "Second Joint Mexico-US International Workshop on Neural Networks and Neurocontrol Sian Ka'an '97,"* Quintana Roo, México, August 1997.

[15] P. Gómez-Gil, M. Ramírez-Cortés. "Experiments with a Hybrid Complex Neural Network for long-term prediction of electrocardiogram" *Proceedings of the IEEE 2006 International World Congress on Computational Intelligence. IJCNN 2006.* IEEE Press. Canada 2006

[16] P. Gómez-Gil. *The Effect of Non-linear Dynamic Invariants in Recurrent Neural Networks for Prediction of Electrocardiograms.* Dissertation in Computer Science. Texas Tech University, December 1998. USA.



Pilar Gomez-Gil received the B.Sc. degree from the Universidad de las Americas (UDLA), at Puebla, Mexico in 1983, and the M.Sc. and Ph.D. degrees from Texas Tech University in 1991 and 1999, respectively, all in Computer Science.

Since 1985 she has been with the Department of Computer Science, Electronics, Physics and Innovation at UDLA. Her research interests include Artificial Neural Networks, Pattern Recognition and Forecasting. Currently she is involved with research related to prediction of electrocardiograms, recognition of antique handwritten documents, and identification of HLA sequences.

Dr. Gomez-Gil is a senior member of the IEEE, and founder member of the IEEE Computational Intelligence Society chapter Mexico, a member of the IEEE Society of Women Engineers, a member of the Honor Society of Computer Science, of the Honor Society of International scholars Phi Beta Delta, a member of the ACM and a member of the Mexican Association of Quality in Software Engineering