



## Long-term prediction of wind speed in La Serena City (Chile) using hybrid neural network-particle swarm algorithm

Juan A. Lazzús\* and Ignacio Salfate

Departamento de Física y Astronomía, Universidad de La Serena, Casilla 554, La Serena, Chile

\*jlazzus@dfuls.cl

### ABSTRACT

An artificial neural network was used for forecasting of long-term wind speed data (24 and 48 hours ahead) in La Serena City (Chile). In order to obtain a more effective correlation and prediction, a particle swarm algorithm was implemented to update the weights of the network. 43800 data points of wind speed were used (years 2003-2007), and the past values of wind speed, relative humidity, and air temperature were used as input parameters, considering that these meteorological parameters are more readily available around the globe. Several neural network architectures were studied, and the optimum architecture was determined by adding neurons in systematic form and evaluating the root mean square error (RMSE) during the learning process. The results show that the meteorological variables used as input parameters, have influential effects on the good training and predicting capabilities of the chosen network, and that the hybrid neural network can forecast the hourly wind speed with acceptable accuracy, such as:  $RMSE=0.81 [m \cdot s^{-1}]$ ,  $MSE=0.65 [m \cdot s^{-1}]^2$  and  $R^2=0.97$  for 24-hours-ahead wind speed prediction, and  $RMSE=0.78$ ,  $MSE=0.634 [m \cdot s^{-1}]^2$  and  $R^2=0.97$  for 48-hours-ahead wind speed prediction.

*Keywords: wind speed; time series forecasting; artificial neural network; particle swarm optimization; meteorological data.*

## Predicción a largo plazo de la velocidad de viento en la ciudad de La Serena (Chile) utilizando un algoritmo híbrido de red neuronal-enjambre de partículas

### RESUMEN

Una red neuronal artificial fue utilizada para la predicción de datos de la velocidad de viento a largo plazo (24 y 48 horas en adelante) en la Ciudad de La Serena (Chile). Para obtener una efectiva correlación y predicción, se implementó una optimización de enjambre de partículas para actualizar los pesos de la red. Se emplearon 43800 datos de velocidad de viento (años 2003-2007), y los valores pasados de velocidad del viento, humedad relativa y temperatura del aire fueron utilizados como parámetros de entrada, considerando que estos parámetros meteorológicos se encuentran fácilmente disponibles en todo el mundo. Se estudiaron varias arquitecturas de redes neuronales y la arquitectura óptima se determinó añadiendo neuronas de forma sistemática y evaluando la raíz del error cuadrático medio (RMSE) durante el proceso de aprendizaje. Los resultados muestran que las variables meteorológicas utilizadas como parámetros de entrada, tienen un efecto positivo sobre el correcto entrenamiento y capacidades predictivas de la red, y que la red neural híbrida puede pronosticar la velocidad del viento horaria con una precisión aceptable, como un  $RMSE=0.81 [m \cdot s^{-1}]$ ,  $MSE=0.65 [m \cdot s^{-1}]^2$  y  $R^2=0.97$  para la predicción de la velocidad del viento de 24 horas en adelante, y un  $RMSE=0.78$ ,  $MSE=0.634 [m \cdot s^{-1}]^2$  and  $R^2=0.97$  para la predicción de la velocidad del viento de 48 horas en adelante.

*Palabras clave: velocidad del viento; predicción de series de tiempo; redes neuronales artificiales; optimización de enjambre de partículas; datos meteorológicos.*

#### Record

Manuscript received: 30/04/2015

Accepted for publication: 17/02/2017

#### How to cite item

Lazzus, J. A. & Salfate, I. (2017). Long-term prediction of wind speed in La Serena City (Chile) using hybrid neural network-particle swarm algorithm. *Earth Sciences Research Journal*, 21(1). 29-35.

doi:<http://dx.doi.org/10.15446/esrj.v21n1.50337>

## 1. Introduction

Energy derived from wind has played a vital role in the history of mankind and is again receiving considerable attention because of its free and non-polluting character. With the development of wind energy technologies and the decrease of wind power production cost, wind power has rapidly developed around the world in recent years (Akdağ and Guler, 2011). As a clean energy source wind is considered an alternative to fossil fuels, which actually accelerate global warming. The first scientific research to utilize wind for generating electricity, was initiated by the Danish in the 1960s. The 1973 energetic crisis, forced many governments to realize the value of wind, as a renewable and independent energy source (Hanağasioglu, 1999). Electricity generation using wind energy has been well recognized as environmentally friendly, socially beneficial, and economically competitive for many applications (Monfared et al., 2009). Prediction of wind speed (WS) at the surface or near the surface, is essential in many areas of science and technology, e.g., wind energy generation, aviation, space vehicle launching, weather forecasting, and agro-meteorology (Kulkarni et al., 2008).

Wind field prediction at the level of wind farm, is still a challenging problem. Different methods have been developed (Kallos et al., 2007); and several studies have been performed to estimate the wind potential in different parts of the world (Çam and Yildiz, 2006). There are various strategies for wind speed prediction that can be classified into two categories: (1) statistical methods that can be subdivided into numerical weather prediction (NWP) and persistence and (2) artificial intelligence techniques that have subdivisions such as artificial neural networks (ANN) and fuzzy logic (Monfared et al., 2009).

With the developments made in chaos theory, researchers have looked for the determinism in various seemingly chaotic-looking fluctuations from different disciplines such as physics, chemistry, hydrology, atmospheric sciences, etc (Karunasinghe and Liong, 2006).

The time series prediction is one of the most important aspects in chaos theory. Time series contain much information about dynamic systems (Han and Wang, 2009). These systems are usually modeled by delay-differential equations. Some of them, for example, the Mackey–Glass equation (Mackey and Glass, 1977), the Ikeda equation (Ikeda, 1979), and equation for an electronic oscillator with delayed feedback (Chua et al., 1992), are standard examples of time-delay systems (Bezruchko et al., 2001).

The main problem of the time series study consist of predicting the next value of a series known up to a specific time, using the known past values of the series. In time series prediction, this is usually first embedded in a state space using delay coordinates:

$$\mathbf{x}(t) = [x(t), x(t+\tau), \dots, x(t+(d-1)\tau)] \quad (1)$$

where  $x(t)$  is the value of the time series at time  $t$ ,  $\tau$  a suitable time-delay and  $d$  the order of the embedding. This embedded vector is then used to predict the next value of the series  $x(t + \tau)$ . Therefore, the non-linear dependence of the level of a series on previous data points is of interest, partly because of the possibility of producing a chaotic time series. Note that short-term prediction for chaotic time series have been widely investigated by several techniques (Karunasinghe and Liong, 2006), however, the long-term prediction has not been widely studied in the literature.

In this work, chaotic time series data taken from the Mackey–Glass differential equation were used to develop a neural network. In order to still obtain a more effective correlation and prediction, particle swarm algorithm has been introduced to update the weights of all layers of the network. Next, this hybrid algorithm was used in the long-term prediction of the next 24 and 48 hours of the wind speed time series. To the best of the authors' knowledge, there is no application for the prediction of the wind speed that includes the long-term prediction, such as the one presented here.

## 2. Computational method

A feed-forward neural network was used to represent non-linear relationships among variables. This ANN was implemented by replacing standard back-propagation algorithm with particle swarm optimization (PSO).

PSO is a population-based optimization tool, where the system is initialized with a population of random particles and the algorithm searches for optima by updating generations (Eberhart and Kennedy, 1995). In each iteration, the velocity of each particle  $j$  is calculated according to the following formula (Lazzús, 2011):

$$\mathbf{v}_j^{k+1} = \omega \mathbf{v}_j^k + c_1 r_1 (\boldsymbol{\psi}_j^k - \mathbf{s}_j^k) + c_2 r_2 (\boldsymbol{\psi}_g^k - \mathbf{s}_j^k) \quad (2)$$

where  $\mathbf{s}$  and  $\mathbf{v}$  denote a particle position and its corresponding velocity in a search space, respectively.  $k$  is the current step number,  $\omega$  is the inertia weight,  $c_1$  and  $c_2$  are the acceleration constants, and  $r_1, r_2$  are elements from two random sequences in the range (0,1).  $\mathbf{s}_j^k$  is the current position of the particle,  $\boldsymbol{\psi}_j^k$  is the best one of the solutions that this particle has reached, and  $\boldsymbol{\psi}_g^k$  is the best solutions that all the particles have reached. In general, the value of each component in  $\mathbf{v}$  can be clamped to the range  $[-v_{max}, v_{max}]$  control excessive roaming of particles outside the search space (Kennedy et al., 2001). After calculating the velocity, the new position of each particle is:

$$\mathbf{s}_j^{k+1} = \mathbf{s}_j^k + \mathbf{v}_j^{k+1} \quad (3)$$

The total steps to calculate the output values, using the input values of the network were as follows (Lazzús et al., 2014):

$$p_i = \left( X_i - X_i^{min} \right) \frac{2}{X_i^{max} - X_i^{min}} - 1 \quad (4)$$

where  $X_i$  is the input variables  $i$ ,  $X_i^{min}$  and  $X_i^{max}$  are the smallest and largest value of the data, thus the input data are normalized using this equation. Next, the net inputs ( $N$ ) are calculated for the hidden neurons coming from the inputs neurons. For a hidden neuron:

$$N_i^h = \sum_j^n w_{i,j}^h p_i + b_{i,j}^h \quad (5)$$

where  $p_i$  is the vector of the inputs of the training,  $w_{i,j}^h$  is the weight of the connection among the input neurons with the hidden layer  $h$ , and the term  $b_{i,j}^h$  corresponds to the bias of the neuron of the hidden layer  $h$ , reached in its activation (Freeman and Skapura, 1991). The PSO algorithm is very different from any of the traditional methods of training (Lazzús, 2011). Each neuron contains a position and velocity. The position corresponds to the weight of a neuron ( $s_i^k \rightarrow w_{i,j}$ ) while the velocity is used to update the weight ( $v_j^{k+1} \rightarrow w_{i,j}^k$ ). Starting from these inputs, the outputs ( $y_j$ ) of the hidden neurons are calculated, using a transfer function  $f^h$  associated with the neurons of this layer (Freeman and Skapura, 1991).

$$y_j = f^h \left( \sum_i^n w_{i,j}^h p_i + b_{i,j}^h \right) \quad (6)$$

To minimize the error, the transfer function  $f$  should be differentiable. In the ANN, the hyperbolic tangent function (tansig) was used as

$$f(N_i) = \frac{e^{N_i} - e^{-N_i}}{e^{N_i} + e^{-N_i}} \quad (7)$$

All the neurons of the ANN have an associated activation value for a given input pattern; the algorithm continues finding the error that is presented for each neuron, except those of the input layer. After finding the output values,

the weights of all layers of the network are actualized  $w_{i,j} \rightarrow w'_{i,j}$  by PSO, using equations 2 and 3 (Lazzús *et al.*, 2014).

The velocity is used to control how much the position is updated. On each step, PSO compares each weight using the data set. The network with the highest fit ness is considered the global best. The other weights are updated based on the global best network rather than on their personal error or fitness (Pérez Ponce *et al.*, 2012; Lazzús *et al.*, 2014). In this article, we used the mean square error (MSE) to determine network fitness for the entire training set:

$$MSE = \frac{\sum_{i=1}^n (Y_i^{true} - Y_i^{calc})^2}{n} \quad (8)$$

where  $Y_i$  is the output value obtained from the normalized output ( $y_i$ ) of the network. This process was repeated for the total number of patterns to training. For a successful process the objective of the algorithm is to modernize all the weights by minimizing the total root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i^{true} - Y_i^{calc})^2}{n}} \quad (9)$$

$$\varepsilon = \min(RMSE) \quad (10)$$

Figure 1 presents a block diagram of the ANN+PSO algorithm developed in this study. In PSO, the inertial weight  $\omega$ , the constant  $c1$  and  $c2$ , the number of particles  $N_{part}$  and the maximum speed of particle summarizes the parameters to synchronize for their application in a given problem. An exhaustive trial-and-error procedure was applied for tuning the PSO parameters. Table 1 shows the selected parameters for this hybrid algorithm.

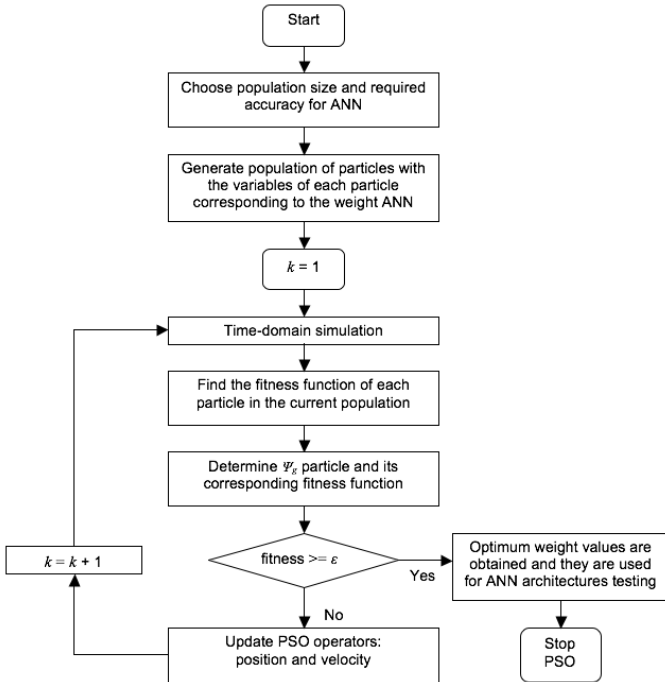


Figure 1. Flow diagram for training of the ANN using PSO algorithm

Table 1. Parameters used in the hybrid ANN+PSO algorithm.

Section	Parameter	Value
ANN	NN-type	feed-forward
	Number of hidden layers	1
	Transfer function	tansig (Eq. 7)
	Number of iterations	1500
	Normalization range	[-1,1]
	Weight range	[-100,100]
	Bias range	[-10,10]
PSO	Minimum error	1e-3
	Number of particles in swarm ( $N_{part}$ )	50
	Number of iterations ( $k_{max}$ )	1500
	Cognitive component ( $c_1$ )	1.494
	Social component ( $c_2$ )	1.494
	Maximum velocity ( $v_{max}$ )	12
	Minimum inertia weight ( $\omega_{min}$ )	0.5
	Maximum inertia weight ( $\omega_{max}$ )	0.7
Objective function	RMSE (Eq. 10)	

### 3. Simulations

#### 3.1. Mackey-Glass time series

To evaluate the capability of the proposed hybrid algorithm in the long-term prediction, the Mackey–Glass time series was used. Thus, a set of data points were generated from the Mackey–Glass time-delay differential equation (Mackey and Glass, 1977; Farmer, 1982) which is defined by:

$$\frac{dx}{dt} = \beta x(t) + \frac{\alpha x(t-\tau)}{1 + x(t-\tau)^{10}}, \quad (11)$$

where  $t$  is a variable,  $x$  is a function of  $t$ , and  $\tau$  is the time delay. The initial values of the time series are  $\alpha = 0.2$ ,  $\beta = 0.1$ , and  $x(0)=1.2$ . If  $\tau \geq 17$ , the time series show the chaotic behaviour (Farmer, 1982; Mirzaee, 2009).

The goal of the task is to use known values of the time series up to the point  $x=t$  to predict the value at some point in the future  $x=t+T$ . The standard method for this type of prediction is to create a mapping from  $d$  points of the time series spaced  $\Delta$  apart, that is  $(x(t), x(t-\Delta), \dots, x(t-(d-1)\Delta))$ , to a predicted future value  $x(t+T)$ .

In order to solve the Mackey–Glass equation, the fourth-order Runge–Kutta method was applied to find the numerical solution. The time series was obtained evaluating the solution of eq. (11) at each integer points. Step size of 0.1 was used to generate a time series, and  $x(t)$  is thus derived for  $0 \leq t \leq 1500$  with  $x(t)=0$  for  $t < 0$  in the integration. Four non consecutive points in the time series are given to generate each input vector  $X_i$  (where  $i=1, 2, \dots, n$ ) of the input matrix  $\mathbf{X}$ , as:

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} x(t) & x(t-\Delta) & \dots & x(t-(d-1)\Delta) \\ x(t+1) & x(t) & \dots & x(t) \\ \vdots & \vdots & \ddots & \vdots \\ x(t+n-1) & x(t+n) & \dots & x(t) \end{bmatrix} \quad (12)$$

A similar criterion was used to create the output matrix, as:

$$\mathbf{Y} = \begin{bmatrix} x(t+\Delta) \\ \vdots \\ x(t+n) \end{bmatrix} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} \quad (13)$$

Then, the ANN+PSO method was used in the long-term prediction. This hybrid algorithm was trained to predict the future value  $x(t+84)$  from the current value  $x(t)$  and the past values, using the standard form applied in the literature, for  $d=4$  and  $\Delta=T=6$  (Chng et al., 1996; Mirzaee, 2009).

$$x(t+84) = F[x(t), x(t-6), x(t-12), x(t-18)] \quad (14)$$

One thousand data points of the above format were collected. The first 500 were used for training while the others were used for testing the ANN+PSO method. Then, the following case was simulated with  $\alpha = 0.2$ ,  $\beta = 0.1$ ,  $x(0) = 1.2$  and  $\tau = 17$ . From this case, several network architectures were tested.

The most basic architecture normally used for the analysis of chaotic time series involves a neural network consisting of three or four layers (Lazzús et al., 2014). The input layer contains one neuron for each input parameter:  $x(t)$ ,  $x(t-6)$ ,  $x(t-12)$ , and  $x(t-16)$ . The output layer has one node generating the scaled estimated value of the chaotic time series  $x(t+84)$ . The number of hidden neurons needs to be sufficient to ensure that the information contained in the data utilized for training the network is adequately represented (Lazzús et al., 2014). There is no specific approach to determine the number of neurons of the hidden layer, many alternative combinations are possible. The optimum number of neurons was determined by adding neurons in systematic form and evaluating the MSE and RMSE of the sets during the learning process (Pérez Ponce et al., 2012; Lazzús et al., 2014). For our case the optimum architecture was 4-12-1.

The results obtained with the ANN+PSO method present a  $MSE=0.000063$  and  $RMSE=0.0079$  for training set, and  $MSE=0.000065$  and  $RMSE=0.0080$  for prediction set. These results show that the ANN+PSO model can be accurately trained and that the chosen architectures can predict the long-term  $x(t+84)$  with acceptable accuracy. Table 2 shows a comparison between some computational methods found in the literature (Martinetz et al., 1993; Whitehead and Choate, 1996; Bersini et al., 1997; Awad et al., 2009) and the result obtained with the ANN+PSO method. This comparison was made using the normalized root mean squared error (NRMSE), defined as:

$$NRMSE = \sqrt{\frac{\sum_{i=1}^n (y_i^{true} - y_i^{calc})^2}{\sum_{i=1}^n (y_i^{true} - \bar{y}_i)^2}} \quad (15)$$

**Table 2.** Comparison between computational methods found in the literature for the long-term prediction.

Method	NRMSE
Linear model	1.504
Cascade correlation NN	0.170
Fuzzy system [23]	0.103*
RBFNN <sup>a</sup> [24]	0.097*
Six order polynomial	0.085
Back-propagation NN	0.060
Genetic algorithm and RBFNN <sup>a</sup> [25]	0.050*
Neural gas [26]	0.050
This work case 1	0.038

\*long-term:  $x(t+85)$

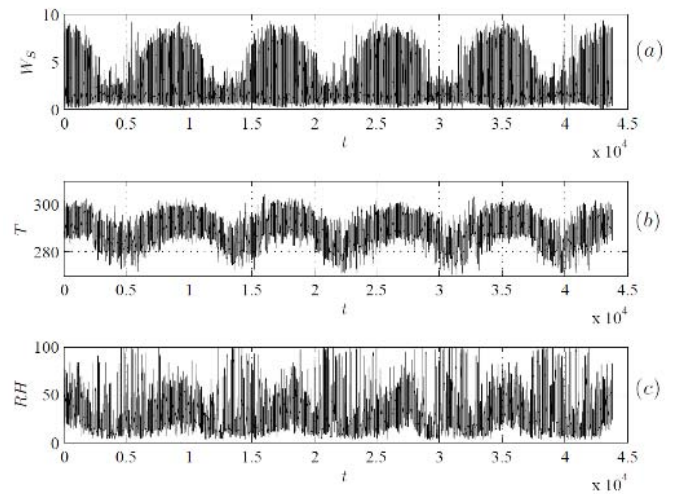
<sup>a</sup>Radial basis function neural network

### 3.2. Wind speed time series

Once the capability of the hybrid algorithm was proved, it was used to forecast the long-term of wind speed time series.

This study is based in data collected from a meteorological stations located on the semi-arid Norte Chico of Chile (29°54' S; 71°15' W; 10 m), located at south of the hyper-arid Atacama Desert. The region is characterized by complex topography with altitudes varying from sea level until 5000 m at the high Andes Cordillera. The climatic characteristics are influenced by the south-eastern Pacific subtropical anticyclone and the cold Humbolt Current which results in low precipitation rates (Kalthoff et al., 2002). The zone is one of the most sensitive areas in South America (Kalthoff et al., 2006), and recent studies of oceanic and atmospheric variability have confirmed the implications of the dynamics of *El Niño*-Southern Oscillation (ENSO) cycle on the local climate of this zone (Meinen and McPhaden, 2009).

43800 data points (years 2003-2007) were used. The future values of wind speed was predicted using the past values of the time series of wind speed  $WS(m \cdot s^{-1})$ , relative humidity  $RH(\%)$ , and air temperature  $T(K)$ . Figure 2 shows the time series of the selected meteorological data used. The data ranges and the properties of interest are listed in Table 3. As seen in this Table, hourly  $WS$  cover wide ranges, going from  $\approx 0$  to  $10 (m \cdot s^{-1})$ . Other wide range for the input parameters are:  $T$  from 270 to 305 (K), and  $RH$  from 2 to  $\approx 100 (\%)$ .



**Figure 2.** Time series of the meteorological data used in this study. (a) wind speed  $WS/m \cdot s^{-1}$ , (b) air temperature  $T/K$ , and (c) relative humidity  $RH/\%$ .

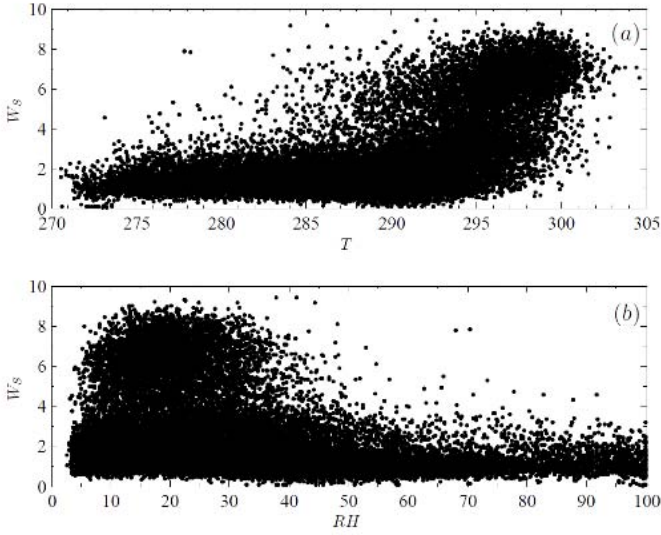
**Table 3** Summary of data used in this study.

Year	No. data points	$WS (m \cdot s^{-1})$	$T (K)$	$RH (\%)$
2003	8760	$\approx 0-10$	271-303	3- $\approx 100$
2004	8760	$\approx 0-10$	272-305	3- $\approx 100$
2005	8760	$\approx 0-10$	271-303	3- $\approx 100$
2006	8760	$\approx 0-9$	270-303	2- $\approx 100$
2007	8760	$\approx 0-10$	271-303	3- $\approx 100$

The influence of the selected meteorological data in this study ( $T$ ,  $RH$ , and  $WS$ ) over the climate of the semi-arid zone of the Atacama Desert has been revised and evaluated in other communications (Kalthoff et al., 2002; Kalthoff et al., 2006). Figure 3 shows  $WS$  as a function of the selected meteorological data:  $T$  and  $RH$ . Fig. 3a shows  $WS$  as a function of  $T$  with a coefficient of linear correlation ( $R^2$ ) of 0.5979. Fig. 3b shows  $WS$  as a function of  $RH$  with  $R^2$  of 0.2757. Note that the coefficient of linear correlation in this figure shows a non-linear relationship between

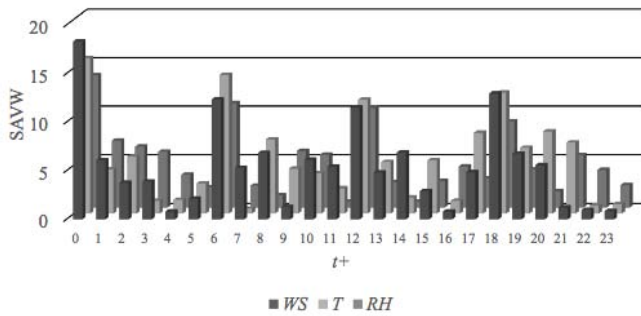
WS and the input parameters for this climate zone. Then, the relationship between WS and these meteorological data is highly non-linear, and consequently an ANN is the best alternative to model the hourly WS.

Two new cases were studied with this methodology, the long-term



**Figure 3.** Experimental data of WS from 2003 to 2007 as a function of the selected Meteorological parameters. a) WS vs T  $\rightarrow R^2=0.5979$ ; (b) WS vs RH  $\rightarrow R^2=0.2757$ .

prediction of the next 24 hours WS(t+24) and the long-term prediction of the next 48 hour WS(t+48). To select the best input parameters for solving the problem, the hourly data from the current value to 23 past hours ( $t-23, t-22, t-21, \dots, t$ ), were considered. Then, the sum of absolute values of weights (SAVW) was used (Lazzús, 2013). Figure 4 shows the input with more contribution for the prediction of future values of  $W_s$ . This Figure shows the great significance of the past values ( $t-18$ ), ( $t-12$ ), ( $t-6$ ), and the current value ( $t$ ) on the three meteorological data ( $W_s$ ,  $T$ , and  $RH$ ).



**Figure 4.** Influence of several points of time series on prediction of future values of wind speed. Bars represent the sum of absolute values of weights (SAVW) of the ANN+PSO.

Thus, the optimum input vector was:

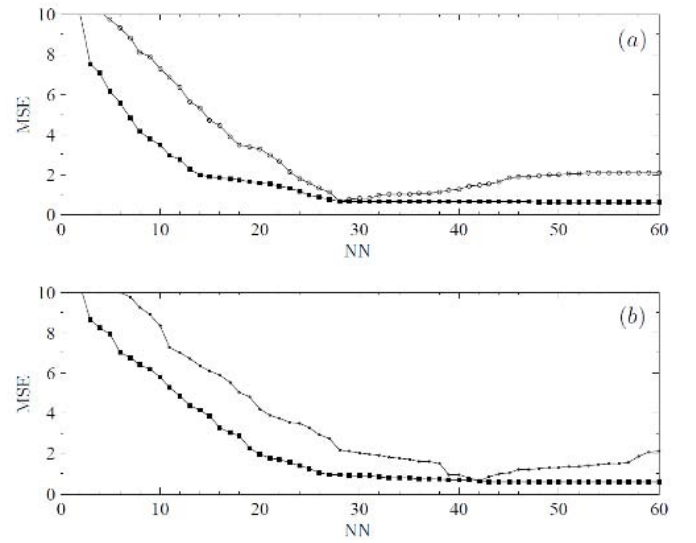
$$T(t+\tau) = F[W_s(t-18), W_s(t-12), W_s(t-6), W_s(t), T(t-18), T(t-12), T(t-6), T(t), RH(t-18), RH(t-12), RH(t-6), RH(t)] \quad (16)$$

where  $\tau$  is the future value to predict: 24 and 48 for the long-term prediction.

The leave-20%-out cross-validation method was used to estimate the predictive capabilities of the model. 34996 data points were used in the training set, and 8760 data points (not used in the training step) were used in the prediction set.

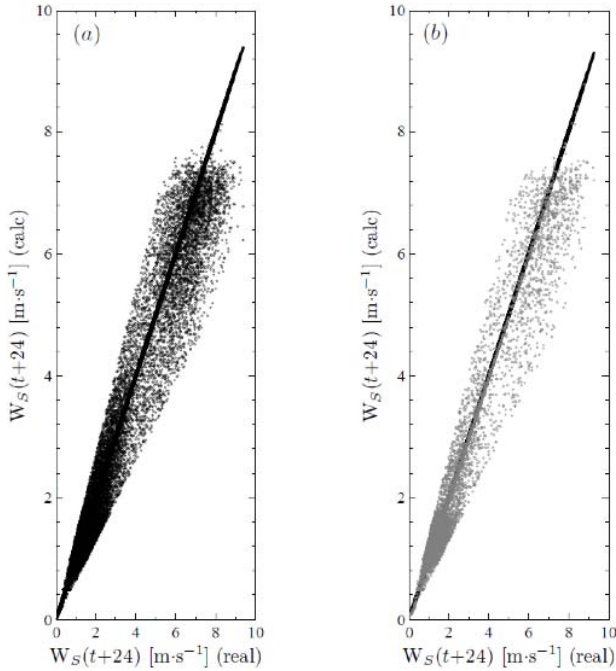
#### 4. Results and Discussion

Several network architectures were tested for the long-term wind speed prediction ( $T(t+24)$  and  $T(t+48)$ , separately). The optimum architecture was checked using the objective function (Eq. 10). Figure 5 shows MSE found in correlating the WS as function of the number of neurons in the hidden layer (NN). Fig. 5a shows the best topology found for the prediction of the WS for the next 24 hours  $T(t+24)$  with a network architecture 12-28-1. Fig. 5b shows the prediction of the  $W_s$  for the next 48 hours  $T(t+48)$  with an optimum network architecture of 12-42-1. Once the best architectures were determined, the optimum weights and biases required to carry out the estimate of future values of  $W_s$  were obtained.



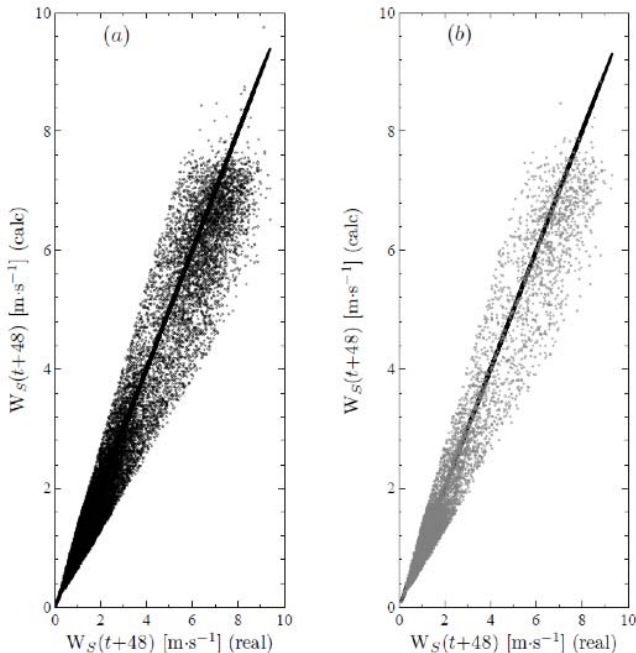
**Figure 5.** Deviations found in the correlation of wind speed as a function of the number of neurons in the hidden layer for: (a) long-term prediction WS(t+24), and (b) long-term prediction WS(t+48). In both graphics, training step (■) and prediction step (○).

Figure 6 shows a comparison between real data (black line) and calculated values (points) of long-term prediction of WS(t+24). Fig. 6a shows the forecasting of WS(t+24) with a correlation coefficient ( $R^2$ ) of 0.974 for the training set (years 2003-2006) with a slope of the curve ( $m$ ) of 0.965 (expected to be 1.0), and with RMSE=0.787 [ $m \cdot s^{-1}$ ] (MSE=0.620 [ $m \cdot s^{-1}$ ]<sup>2</sup> and MSEmax=10.028 [ $m \cdot s^{-1}$ ]<sup>2</sup>). Fig. 6b shows the forecasting of WS(t+24) with  $R^2=0.973$  during the prediction step (year 2007) with  $m=0.963$  (also expected to be 1.0), and with RMSE=0.807 [ $m \cdot s^{-1}$ ] (MSE=0.651 [ $m \cdot s^{-1}$ ]<sup>2</sup> and MSEmax=7.110 [ $m \cdot s^{-1}$ ]<sup>2</sup>).



**Figure 6.** Comparison between real and calculated values of  $WS(t+24)$  using the proposed ANN+PSO model. (a) Training set, period 2003-2006; and (b) Prediction set, period 2007.

Figure 7 shows a comparison between real data (black line) and calculated values (points) of long-term prediction of  $WS(t+48)$ . Fig. 6a shows the forecasting of  $WS(t+48)$  with  $R^2=0.970$  for the training set with  $m=0.963$ , and with  $RMSE=0.788$  [ $m \cdot s^{-1}$ ] ( $MSE=0.622$  [ $m \cdot s^{-1}$ ] $^2$  and  $MSE_{max}=9.386$  [ $m \cdot s^{-1}$ ] $^2$ ). Fig. 6b shows the forecasting of  $WS(t+48)$  with  $R^2=0.969$  for the prediction set with  $m=0.962$ , and with  $RMSE=0.796$  ( $MSE=0.634$  [ $m \cdot s^{-1}$ ] $^2$  and  $MSE_{max}=8.210$  [ $m \cdot s^{-1}$ ] $^2$ ).



**Figure 6.** Comparison between real and calculated values of  $WS(t+48)$  using the proposed ANN+PSO model. (a) Training set, period 2003-2006; and (b) Prediction set, period 2007.

A comparison was made with a neural network with standard back-propagation (BPNN) algorithm (Hagan and Menhaj, 1994), and similar architecture and database. This BPNN show results of MSE higher than 1 [ $m \cdot s^{-1}$ ] $^2$  and  $R^2$  lower 0.8 for the forecasting of  $WS(t+24)$  and  $WS(t+48)$ . And other comparison was made with a multiple linear regression (MLR) method, and similar database. The MLR method shows MSE higher than 4 [ $m \cdot s^{-1}$ ] $^2$  and  $R^2$  lower 0.7 for both cases. Note that, the predictions with the proposed ANN+ PSO method shows MSE a little higher than 0.6 [ $m \cdot s^{-1}$ ] $^2$  and  $R^2$  higher than 0.97. Table 4 summarizes the deviations obtained in the long-term prediction using the proposed method versus BPNN and MLR methods. These results show that the ANN+PSO can be accurately trained and that the chosen topologies can estimate the future values of WS with acceptable accuracy. These results represent a tremendous increase in accuracy for forecasting this important meteorological property and show that not only the optimum architecture obtained was crucial, also the appropriate selection of the independent parameters (T and RH). This is important because air temperature and relative humidity are commonly available parameters. Note that the coefficients of linear correlation of these parameters show a non-linear relationship with the WS for the climate of several geographic zones. Then, the relationship between WS and these meteorological data is highly non-linear, and consequently the ANN+PSO is a good tool for modeling WS for several applications.

**Table 4.** Summary of the deviations obtained with the ANN+PSO algorithm for the long-term prediction of wind speed..

Long-term	Set	MSE [ $m \cdot s^{-1}$ ] $^2$	RMSE [ $m \cdot s^{-1}$ ]	NRMSE	$R^2$
ANN+PSO					
$WS(t+24)$	Training	0.620	0.787	0.058	0.974
	Prediction	0.651	0.807	0.060	0.973
$WS(t+48)$	Training	0.622	0.788	0.053	0.970
	Prediction	0.634	0.796	0.059	0.969
BPNN					
$WS(t+24)$	Training	1.154	1.074	11.315	0.785
	Prediction	1.102	1.098	11.537	0.775
$WS(t+48)$	Training	1.212	1.097	11.654	0.763
	Prediction	1.236	1.105	11.852	0.757
MLR					
$WS(t+24)$	Correlation	4.154	2.038	21.984	0.675
	Prediction	4.251	2.076	20.989	0.671
$WS(t+48)$	Correlation	4.169	2.045	20.987	0.669
	Prediction	4.286	2.100	21.627	0.667

The results obtained by other models and other sites can be usefully compared with our results. Zhang et al. (2012) shows the performance analysis of four modified approaches for wind speed forecasting for four observation sites in Gansu (China), with MSE higher than 2 [ $m \cdot s^{-1}$ ] $^2$ . Liu et al. (2013) shows the forecasting models for wind speed using wavelet, wavelet packet, time series and artificial neural networks with MSE higher than 1 [ $m \cdot s^{-1}$ ] $^2$ . For Spain, the wind speed estimation was made using a multilayer perceptron with MSE higher than 1 [ $m \cdot s^{-1}$ ] $^2$  and  $R^2$  below 0.75 (Velo et al., 2014). Recently, Wang et al. (2014) shows the mean hourly wind speed prediction in the Hexi Corridor of China based on the seasonal adjustment method (SAM), exponential smoothing method (ESM), and radial basis function neural network (RBFN), with RSME higher than 0.7 [ $m \cdot s^{-1}$ ]. It must be mentioned that these results were obtained from different sites and based on different weather conditions, and the results cannot be compared directly with one another. However, results from the different methods show that the accuracy of ANN+PSO model employed in this study is good.

## 5. Conclusions

In this work, a neural network was used for the forecasting of long-term wind speed time series. In order to obtain a more effective correlation and prediction, particle swarm algorithm has been introduced to update the weights of all layers of the network. 43800 data points (years 2003-

2007) of wind speed were used. To distinguish between the different values of hourly data considered in this study, so that the network can discriminate and learn in optimum form, the following past values of the time series of meteorological data were used as input parameters: wind speed  $WS(m \cdot s^{-1})$ , relative humidity  $RH(\%)$ , and air temperature  $T(K)$ .

Based on the results and discussion presented in this study, the following main conclusions are obtained: i) The results show that the proposed ANN+PSO can be properly trained for predicting the hourly wind speed, with acceptable accuracy; ii) the meteorological variables used ( $WS$ ,  $T$ , and  $RH$ ), have influential effects, on the good training and predicting capabilities, of the chosen network; iii) The low deviations found with the proposed ANN+PSO method indicate that it can predict the future values of  $WS$  with better accuracy than other methods; and iv) The values obtained with the proposed method are believed to be sufficiently accurate for engineering calculations, among other uses.

#### Acknowledgements

The authors thank the Direction of Research and Development of the University of La Serena (DIDULS) through the research project PEQ16141, and the Department of Physics of the University of La Serena (DFULS) for the special support that made possible the preparation of this paper. Special acknowledgements to Dr. Pedro Vega (ULS) for the support that made possible the implementation of the meteorological station and to Mg. Julio Marín (ULS) for providing the meteorological data set.

#### References

- Akdağ, S.A. and Güler Ö. (2011). A Comparison of Wind Turbine Power Curve Models. *Energy Sources Part A*, 33, 2257–2263.
- Awad, M., Pomares, H., Rojas, I., Salameh, O. and Hamdon, M. (2009). Prediction of time series using RBF neural networks: a new approach of clustering. *International Arab Journal of Information Technology*, 6, 138–143.
- Bersini, H., Duchateau, A. and Bradshaw, N. (1997). Using incremental learning algorithms in the search for minimal effective fuzzy models. In: *Proceedings of 6th International Conference on Fuzzy Systems*, pp. 1417–1422.
- Bezruchko, B.P., Karavaev, A.S., Ponomarenko, V.I. and Prokhorov, M.D. (2001). Reconstruction of time-delay systems from chaotic time series. *Physical Review E*, 64, 056216.
- Çam, E. and Yildiz, O. (2006). Prediction of wind speed and power in the Central Anatolian region of Turkey by adaptive neuro-fuzzy inference systems (ANFIS). *Turkish Journal of Engineering and Environmental Sciences*, 30, 35–41.
- Chng, E.S., Chen, S. and Mulgrew, B. (1996). Gradient radial basis function networks for nonlinear and nonstationary time series prediction. *IEEE Transactions on Neural Networks*, 7, 190–194.
- Chua, L.O., Kocarev, L., Eckert, K. and Itoh, M. (1992). Experimental chaos synchronization in Chua's circuit. *International Journal of Bifurcation and Chaos*, 2, 705–708.
- Eberhart, R.C. and Kennedy, J. (1995). A new optimizer using particle swarm theory. In: *Proceedings of 6th International Symposium on Micro Machine and Human Science*, Nagoya, New York, NY, USA: IEEE, pp. 39–43.
- Farmer, J.D. (1982). Chaotic attractors of an infinite-dimensional dynamical system. *Physica D*, 4, 366–393.
- Freeman, J.A. and Skapura, D.M. (1991). *Neural networks: algorithms, applications and programming techniques*. Computation and Neural Systems Series. Massachusetts, USA: Addison-Wesley.
- Hagan, M.T. and Menhaj, M.B. (1994). Training feedforward networks with the Marquardt algorithm, *IEEE Transactions on Neural Networks*, 5, 989–993.
- Han, M. and Wang, Y. (2009). Analysis and modeling of multivariate chaotic time series based on neural network. *Expert Systems with Applications*, 36, 1280–1290.
- Hanağasioglu, M. (1999). Wind energy in Turkey. *Renewable Energy*, 16, 822–827.
- Ikeda, K. (1979). Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system. *Optics Communications*, 30, 257–261.
- Kallos, G., Galanis, G. and Katsafados, P. (2007). Local wind speed forecasting and applications to power prediction. *Geophysical Research*, 9, 93–99.
- Kalthoff, N., Bischoff-Gauß, I., Fiebig-Wittmaack, M., Fiedler, F., Thürauf, J., Novoa, E., Pizarro, C., Castillo, R., Gallardo, L., Rondanelli, R. and Kohler, M. (2002). Mesoscale wind regimes in Chile at 30° S. *Journal of Applied Meteorology*, 41, 953–970.
- Kalthoff, N., Fiebig-Wittmaack, M., Meißner, C., Kohler, M., Uriarte, M. and Bischoff-Gauß, I. (2006). The energy balance, evapotranspiration and nocturnal dew deposition of an arid valley in the Andes. *Journal of Arid Environments*, 65, 420–443.
- Karunasinghe, D.S.K. and Liong, S.Y. (2006). Chaotic time series prediction with a global model: artificial neural network. *Journal of Hydrology*, 323, 92–105.
- Kennedy, J., Eberhart, R.C. and Shi, Y. (2001). *Swarm Intelligence*. San Diego, CA, USA: Academic Press.
- Kulkarni, M.A., Patil, S., Rama, G.V. and Sen, P.N. (2008). Wind speed prediction using statistical regression and neural network. *Journal of Earth System Science*, 117, 457–463.
- Lazzús, J.A. (2011). Predicting natural and chaotic time series with a swarm-optimized neural network. *Chinese Physics Letters*, 28, 110504.
- Lazzús, J.A. (2013). Neural network-particle swarm modeling to predict thermal properties. *Mathematical and Computer Modelling*, 57, 2408–2418.
- Lazzús, J.A., Salfate, I. and Montecinos, S. (2014). Hybrid neural network-particle swarm algorithm to describe chaotic time series. *Neural Network World*, 24, 601–617.
- Liu, H., Tian, H., Pan, D. and Li, Y. (2013). Forecasting models for wind speed using wavelet, wavelet packet, time series and artificial neural networks. *Applied Energy*, 107, 191–208.
- Mackey, M.C. and Glass, L. (1977). Oscillation and chaos in physiological control systems. *Science*, 197, 287–289.
- Martinetz, T.M., Berkovich, S.G. and Schulten, K.J. (1993). Neural-gas network for vector quantization and its application to time-series prediction. *IEEE Transactions on Neural Networks*, 4, 558–569.
- Meinen, C.S. and McPhaden, M.J. (2000). Observations of warm water volume changes in the Equatorial Pacific and their relationship to El Niño and La Niña. *Journal of Climate*, 13, 3551–3559.
- Mirzaee, H. (2009). Linear combination rule in genetic algorithm for optimization of finite impulse response neural network to predict natural chaotic time series. *Chaos, Solitons & Fractals*, 41, 2681–2689.
- Monfared, M., Rastegar, H. and Madadi-Kojabadi, H. (2009). A new strategy for wind speed forecasting using artificial intelligent methods. *Renewable Energy*, 34, 845–848.
- Pérez Ponce, A.A., Lazzús, J.A. and Palma-Chilla, L. Hybrid neural network-particle swarm method to predict global radiation over the Norte Chico (Chile). *Journal of Renewable and Sustainable Energy*, 4, 023108.
- Velo, R., López, P. and Maseda, F. (2014). Wind speed estimation using multilayer perceptron. *Energy Conversion and Management*, 81, 1–9.
- Wang, J., Zhang, W., Wang, J., Han, T. and Kong, L. A novel hybrid approach for wind speed prediction. *Information Sciences*, 273, 304–318.
- Whitehead, B.A. and Choate, T.D. (1996). Cooperative-competitive genetic evolution of radial basis function centers and widths for time series prediction. *IEEE Transactions on Neural Networks*, 7, 869–880.
- Zhang, W., Wu, J., Wang, J., Zhao, W. and Shen, J. (2012). Performance analysis of four modified approaches for wind speed forecasting. *Applied Energy*, 99, 324–333.