

NONLINEAR AND DISCONTINUITIES MODELING OF TIME SERIES USING ARTIFICIAL NEURAL NETWORK WITH RADIAL BASIS FUNCTION

Alfonso TIERRA¹

DOI: 10.21163/GT_2016.112.10

ABSTRACT:

At present, the Global Navigation Satellite System (GNSS) is used to obtain one Geodetic Reference System by achieving a network of a continuous GNSS monitoring. Due to a variety of factors, the time series may expose data gaps and discontinuities that change its trend. This study proposes to use radial basis function of a neural network with the main objective to generate a single modeling that resolves such difficulties. Examinations were realized using data of the time series of the ellipsoidal height of the station CONZ continuous monitoring located in Concepcion-Chile, between GPS the weeks 1170 and 1833. This station has been chosen due to the circumstances that their time series reflect non-linear behavior, the appearance of data gaps, and it demonstrates also changes in its trend as result of an intense earthquake. The Artificial Neural Network (ANN) technique has been involved with the purpose to realize a single model. The structure of the ANN that has been obtained after training was Radial Basis Function (RBF). In that specific RBF the input vector has been in the week GPS and the output vector has been the ellipsoidal height. Finally, the time series was modeled for the determination of the capacity of a potential generalization in such possible prediction. The result of such prediction yields a mean arithmetic 0.01mm, and standard deviation 2.5 mm. This leads to the conclusion that it may be possible to use in any time series one single neural model only to obtain a reasonable prediction.

Key-words: GPS, Time series, Radial basis function network, Discontinuity, Gap, Nonlinear.

1. INTRODUCTION

Actually, the Global Navigation Satellite System (GNSS) and especially the Global Positioning System (GPS) is used to definition of the System Geodetic System (SGR) by implementation of Continuously Operation Reference Station (CORS) network. The observations from GNSS can be used to obtain the time series such as geodetic coordinates and velocities of one station, of which series have a nonlinear behavior. The time series modeling have been carried out by different methods and applied in various fields of science

Besides the nonlinear behavior of the time series of the geodetic coordinates, satellite signals may get temporarily lost creating gaps in GPS data bases. Additionally, some CORS stations are sometimes located in places where there is geodynamic activity in form of earthquakes, volcanoes, mass wasting, besides other. The occurrence of any of these events may result in discontinuities changing the trend of the time series.

¹ *Universidad de las Fuerzas Armadas, Grupo Geoespacial, 171-5-231B Sangolquí, Ecuador, artierra@espe.edu.ec.*

Modeling of times series has been studied by many authors, like by Altamini, Collilieux and Boucher (2006), as well as Altamini, Collilieux and Metivier (2011) who actualized the ITRF datum definition based on times series of station positions and Earth Orientation Parameters, and recently, the Post-Seismic Deformation (PSD) correction has been used in his equations to actualizing the ITRF2014. Wang, Wang and Robersts (2007) proposed in a study to establish a dynamic geodetic Datum considering the velocities with non-linear variations and using different interpolation methods to improve the local datum precision.

Amiri-Simkooei and Asgari (2011) performed studies of total electron contents. Güral, Erdoğan and Tiryakioglu (2013) realized an analysis of GNSS reference stations by time series using AR and ARMA techniques. Also, the ARMA technique; Li et al (2000) in GPS time series to the crustal deformation has been applied. Mémin et al (2014) investigated non-tidal oceanic and atmospheric loading, while Li et al (2014) used artificial neural network to build up regional crustal movement velocity field. Bogusz and Klos (2015) studied a noise analysis of coordinates times series. Roggero (2012) using statistic techniques in GNSS coordinates times series to detecting, estimate and remove the level shifts, to predict the station coordinates in data gaps were been realized.

In the present paper, modeling of the time series of the height ellipsoidal in the continuous monitoring station CONZ located in Concepción-Chile will be demonstrated. This specific station was chosen because it presents gaps in GPS data, nonlinear behavior, and also two discontinuities that change the trend of the series, one of them being the result of a local earthquake. The presences of such conditions are particularly useful to demonstrate the possibility of resolving such issues, with a single model with the use of an artificial neural network trained. The modeling of the time series was structured by an artificial neural network of the type Radial Basis Function (RBF) and for its learning supervised training has been used.

The next sections describe ANN structure, time series data of the ellipsoidal height in the station CONZ, the training using radial basis functions network, followed by the results, discussion and subsequently the conclusions.

2. ARTIFICIAL NEURAL NETWORK (ANN)

ANN is able to provide efficient and elegant solutions to many problems in different issues, and therefore it has been used with encouraging results in a wide range of applications such as modeling, pattern recognition, signal processing, environmental control, robotics and others.

The Artificial Neural Network -ANN was conceived from the observation of the human brain behavior in comparison with digital computers (Haykin, 1999). An ANN is the computational system that imitates abilities of the biological nervous system by means of the use of simple number of artificial interconnected neurons (also called nodes or processing elements), which enables information to be processed in parallel (Potukuchi & Wexler, 1997). **Fig. 1** reflects a similitude between a biological and artificial neuron. Therefore, the dendrites in the biological neuron are the inputs in the artificial neuron; the cell body is the process the integration the inputs and activation function, while the axon is the output. The connection between neurons is known as synaptic weights.

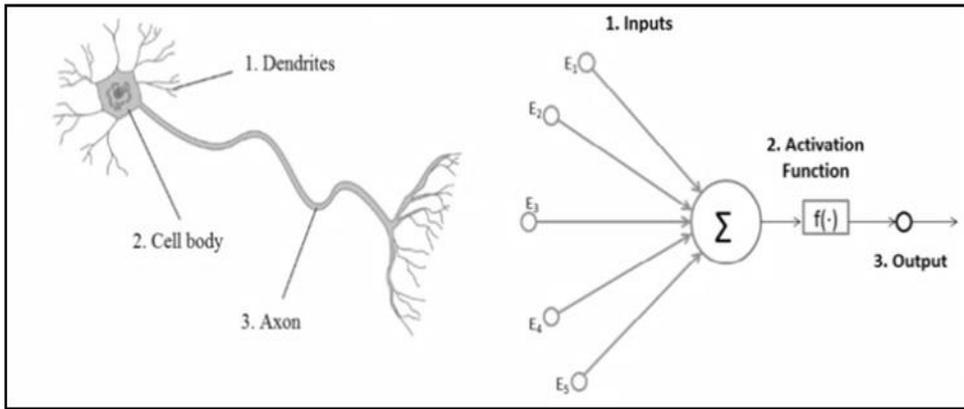


Fig. 1. Biological and neuron artificial
(Source: Tierra & Romero, 2014).

Fig. 2 depicts typical ANN with three distinctive layers, which are the input layer with (r) input signals that are connected to the net, the hidden layer where are of the (s) neurons and the last layer, which is called the output layer that contains the (k) neurons that provides output signal computed by the network to the environment external. ANN can be structured with different architectures, depending on the number of layers to be used, such as the manner of neurons arrangement, the transfer functions to be used and the form used to connect neurons (Loesch & Sari, 1996). One of them is known as the Radial Basis Function Network (RBF) and that is which has been used in the present study.

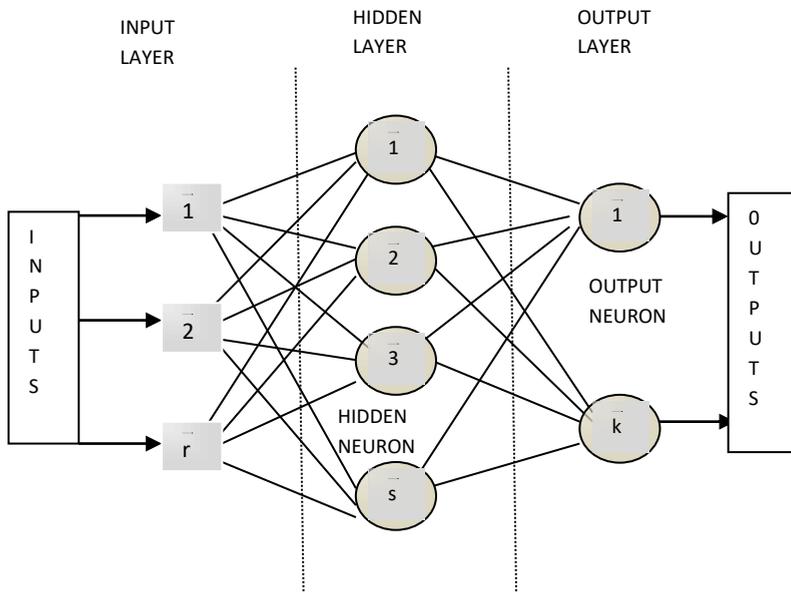


Fig. 2. Typical artificial neural network with the three distinctive layers described in the text.

2.1 Radial Basis Functions Network

The RBF define a non-linear relation between the input and output variables. This net in its basic form consists of the three layers with forward connections. The first layer is inputs that connect the network to its environment. The hidden layer uses neurons with activation functions of Radial Basis (RBF), applies a nonlinear transformation from the input space to the hidden space. The neurons of the output layer are used to combine linearly the outputs of the hidden neurons (Haykin, 1999; He & Xu, 2007; Dreyfus, 2005). **Fig. 2** illustrates a typical RBF structure.

Fig. 3 demonstrates a radial basis network with X^R inputs ($x_1, x_2, x_3, \dots, x_r$), a hidden radial basis layer of S^H neurons, an output linear layer of S^P neurons, and Y^T output ($y_1, y_2, y_3, \dots, y_t$) of the net. Furthermore, it contains the b_1, b_2 BIAS and synaptic weights W^H (w_1, w_2, \dots, w_h) and W^T (w_1, w_2, \dots, w_t) of the hidden and output layer respectively.

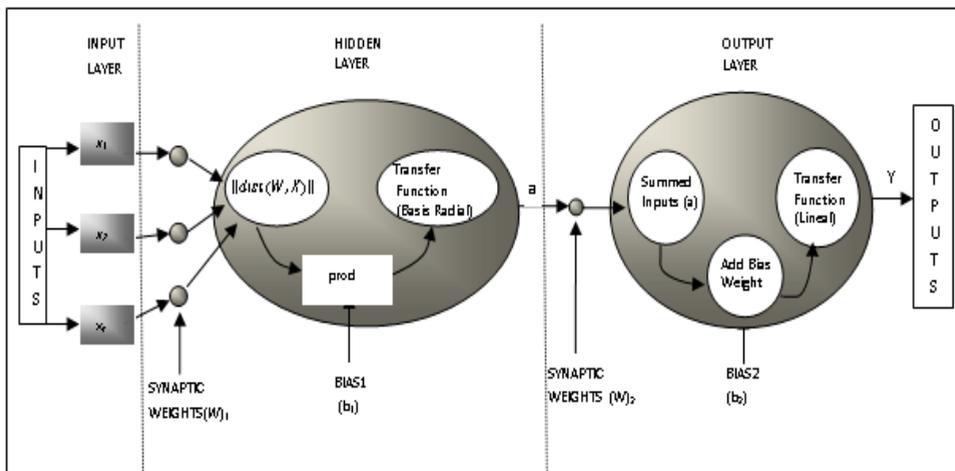


Fig. 3. Radial basis function network architecture with three layers.

In the hidden layer, the X^R inputs are connected with hidden layer neurons. Here, the distance $\| \text{dist}(W, X) \|$ between the weights W_1 and vector X^R is determined and produces a vector with S^H elements. Afterwards, the bias b_1 is combined, through the product, with the output of the distance, which generates element-by-element multiplication. This result is used in the transfer function of the basis radial type in each neurons and produce the outputs A^H (a_1, a_2, \dots, a_h) of the hidden layer. There are several forms of radial basis activation functions, of which the most common one is the Gaussian function, defined by (He & Xu, 2007; Tierra, De Freitas & Guevara, 2009):

$$a_h = G_h(X) = \exp\left(-\frac{\|X - u_h\|^2}{2\sigma_h^2}\right) \tag{1}$$

where σ is the width parameter, u is the vector determining the center of the basis function G and X is the dimensional input vector.

The outputs of A^H are connected with the output layer neurons and they are combined with the weights W_2 , added with the bias b_2 that has weight 1, and the output vector Y is obtained, therefore:

$$Y_t = \sum_{i=1}^H w_i a_i + b_2 \quad (2)$$

where each w_i is the output weight that corresponds to the connection between a hidden and an output unit.

For the training of the RBF supervised learning has been used, which starts with the consideration by learning with a teacher, who has obviously knowledge of the studied environment. This state of knowledge may be represented by a set of input-output examples. Therefore, the values of inputs signals and the corresponding output signals are known.

In the instructed or supervised learning, the error in the output of a neuron t (difference between the output Y_t and desired outputs T_t) in the first step (Haykin, 1999). Tierra, Dalazoana and De Freitas (2008) are defined by:

$$e_t = T_t - Y_t \quad (3)$$

The total error E_{TE} , eq. (04), for all output is:

$$E_{TE} = \frac{1}{2} \sum_{i=1}^T e_i^2 \quad (4)$$

where T is the number of neurons in the output layer.

The training process is made until the cost function MSE (Mean Square Error) is a minimum, given by eq. (05):

$$MSE = \frac{1}{M} \sum_{i=1}^M e_i^2 \quad (5)$$

where M is the total number of examples.

3. TIME SERIES DATA OF THE ELLIPSOIDAL HEIGHT

Tests were realized using time series data of the ellipsoidal height of the station CONZ of continuous monitoring located in Concepcion-Chile, between GPS week 1170 and 1833. The station data are available in <ftp://ftp.sirgas.org/pub/gps/SIRGAS/> and consists in the weekly coordinates (X,Y,Z) aligned to IGS05 reference frame, which later were transformed to geodetic coordinates (latitude, longitude, ellipsoidal height).

This station has been chosen due to the specific circumstances of having time series, which reflect no-linear behavior, appearance of data gaps and additionally it demonstrates trend changes, as illustrated in **Fig. 4**.

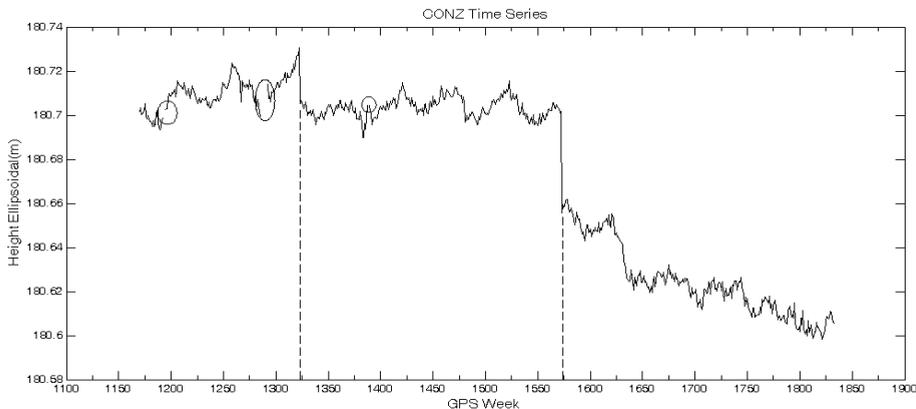


Fig. 4 Time series data of ellipsoidal heights. Ellipses represent data gaps of two, six, and one GPS week, in dashed lines have occurred trend changes. In the GPS week 1572 occurred the earthquake in Chile of 8.8 Mw and 30.1 km deep.

4. RBF NETWORK TRAINING

In order to perform the learning process of the RBF, 663 data of the time series have been used, of which all were divided into three groups. The first group includes 550 data and has been used for the training phase. The second group includes 70 data, which were used to validate the training itself. Finally, the third group included 35 data, which were used as test data for determine capacity in the prediction of the RBF. The validation and test data were chosen randomly. Some 22 data are missing or lost due to the presence of the gaps. The data groups are illustrated in **Fig. 5**.

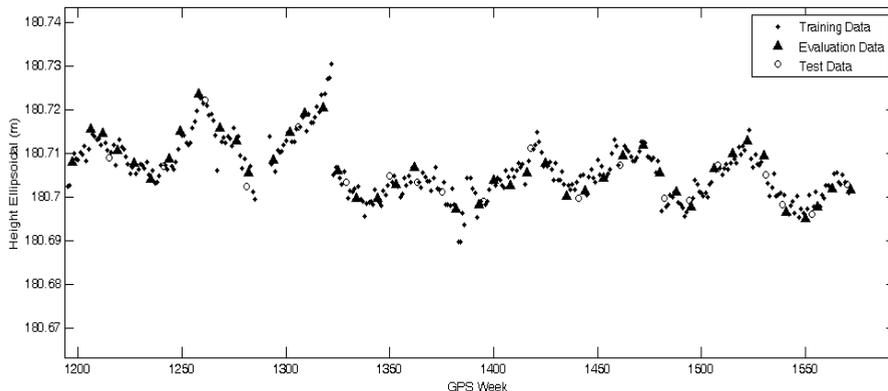


Fig. 5 Times series data of ellipsoidal height. Training data are the points, evaluation data are the triangle and test data are the circle.

The trained RBF [1 142 1] indicates that the net structure has one variable in the input layer that has been in the GPS week (**Fig. 6**). Some 142 neurons in the hidden layer were needed to perform learning. The output layer has been one neuron, of which the output of the net has been represented by the height of the ellipsoidal.

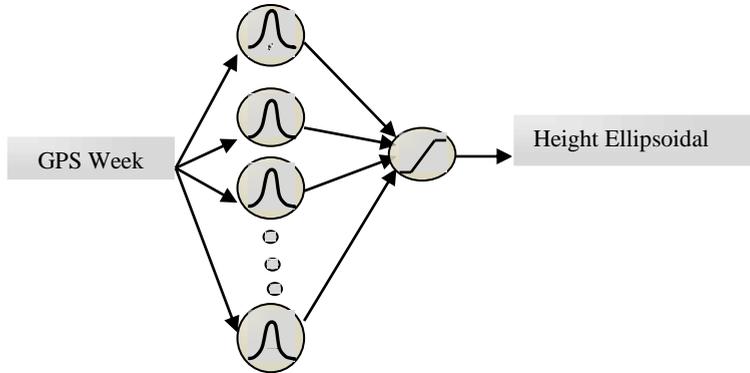


Fig. 6. Radial basis function network structure [1 142 1].

5. RESULTS AND DISCUSSION

The tests data input (GPS week) were used in the net trained RBF in order to predict the ellipsoidal heights. The obtained results were compared with the known ellipsoidal height of the tests data and their differences were calculated. These differences had a mean arithmetic value of 0.1 mm and a standard deviation of 2.4 mm, while all differences were less than 5 mm. **Fig. 7** illustrates the differences obtained between the trained RBF with the test data.

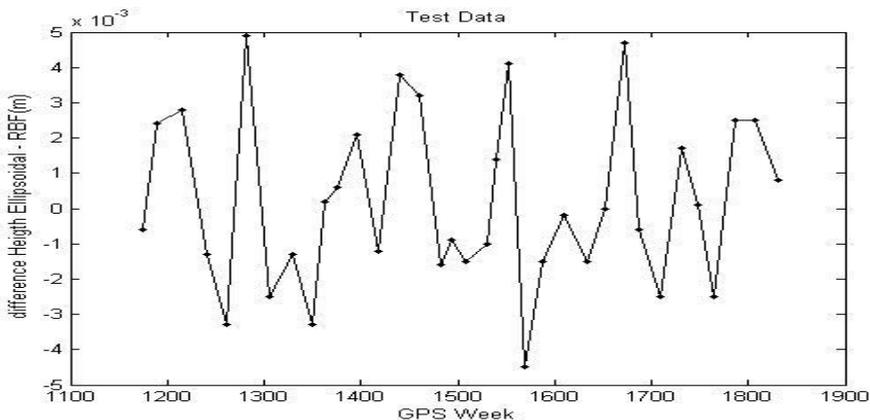


Fig. 7. Difference of ellipsoidal heights with RBF in function with time expressed in GPS weeks.

Finally, this RBF network trained has been used in the reconstruction of data gaps, for modeling of nonlinear time series as well as does obtain a continuation of the trends where discontinuities occurred (**Fig. 8**). To perform such aim, the 633 GPS week has been used as input vector due to the fact that it obtained the ellipsoidal respective. These results were than compared with the known ellipsoidal height and their differences were calculated in order to determine the capacity of generalization of the trained RBF.

The obtained differences had a mean arithmetic value of 0.01mm, a standard deviation value of 2.5 mm with extreme values of 12mm (maximum) and -23mm (minimum). Approximately 99.2% of the points have less than 1 cm, and 97% have less than 5 mm of difference.

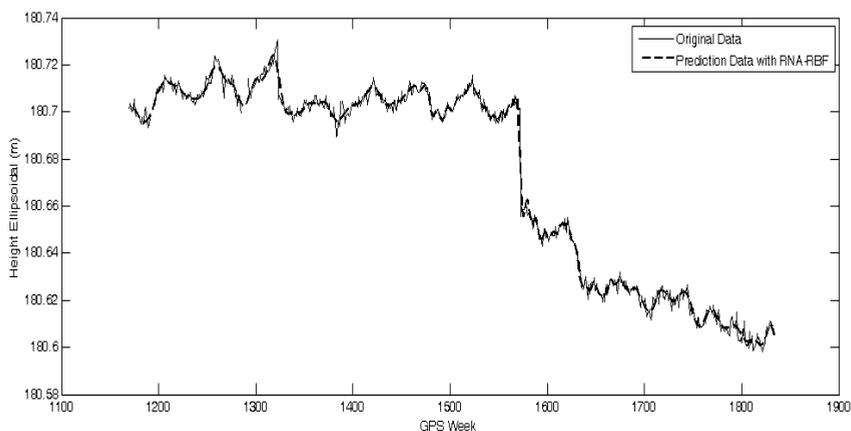


Fig. 8 Predicting times series with RBF network trained. The original data of the time series are represented by continuous line, and dashed lines indicated the time series predicted with the trained RBF.

A regression analysis in the post training between the network response and the corresponding targets has been performed. **Fig. 9** illustrates the graphical output provided by regression. The correlation coefficient (R-value) between output and target were calculated. All R, including the training, validation, test and all network are relatively close to one, indicating an excellent performance.

The differences obtained of the all series between output and target in the figure 12 has been plotted. In comparison with all data of all series, output as well as target studied here, the most fundamental differences were obtained between the GPS weeks of 1322-1323 and 1572-1573. In these weeks occurred the trend changes of the series. As it is unknown to date what generated the very first change, the second has been much clearer due to the presence of a local earthquake in Concepción-Chile (**Fig. 10**).

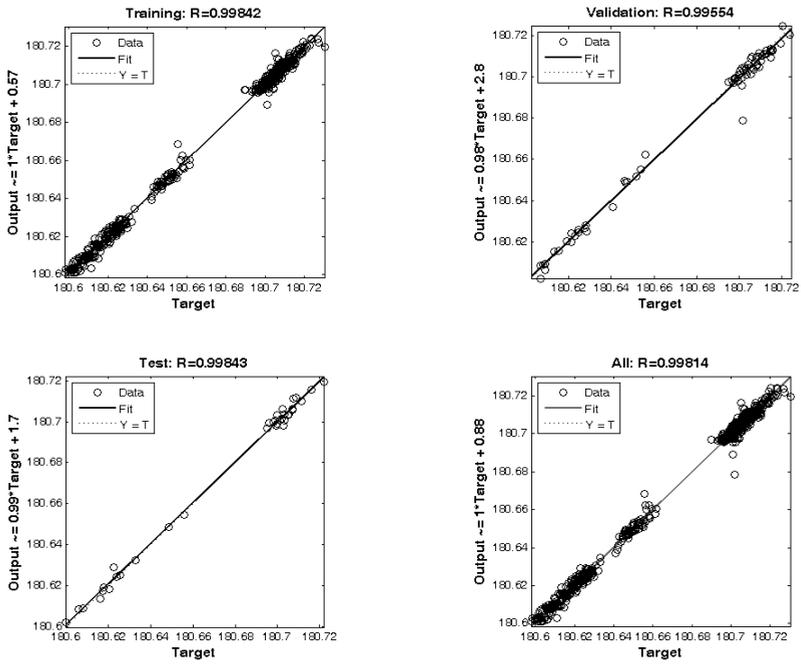


Fig. 9. Graphical output provided by regression, correlation coefficients. The best linear fit is indicated by a solid line, while the circles are the data.

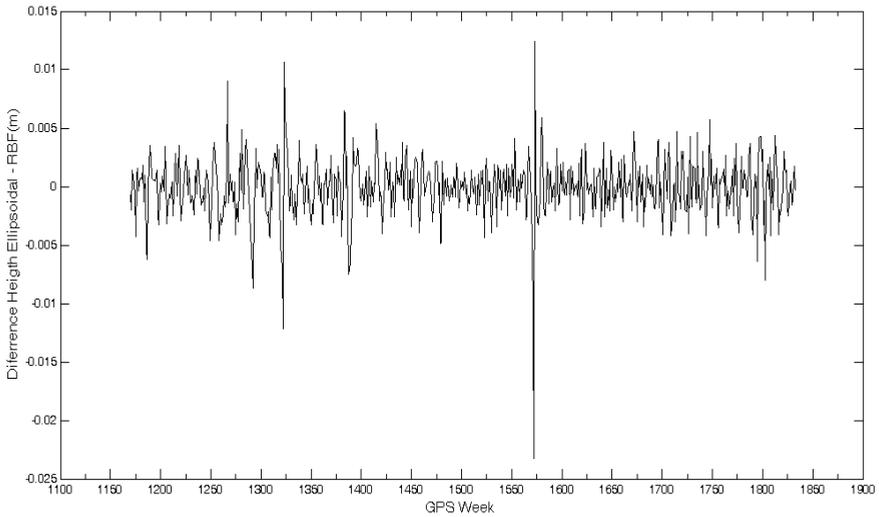


Fig. 10 Differences in the time series between height Ellipsoidal original and predicted.

In the first change trend was obtained a difference of -12 mm for the 1322 GPS week in the prediction, and 11 mm for the 1323 GPS week. In the second change trend a

difference of -23mm for the 1572 GPS week, and 12 mm for the 1573 GPS week were obtained. In both cases the modeling with the RBF trained follows the changes of the trends presents in the time series, as well as the RBF was able to rebuild the data gaps.

In a future work, this methodology will be used to develop a velocities model from a continuous monitoring GNSS network. This model will be applied in areas of high geodynamic activity due to natural events that cause discontinuities and change in the trend of the time series. In this way, it will facilitate to having a single model, avoiding to having models in each trend change.

6. CONCLUSIONS

With the results obtained the conclusion are summarized as follows:

- The correlation coefficients R were relatively close to one, indicating that a single trained RBF is able to be used in prediction of expected data with great reliability in nonlinear time series as well as in the presence of discontinuities and also where data gaps appear.
- The RBF network trained is able to follow the changes of the trends present in the time series.
- The most visible differences obtained with the RBF are located in the discontinuities where their trend changes are occurring.
- The application of the artificial neural network techniques indicates a potentially useful tool to model nonlinear time series.

REFERENCES

- Altamini, Z., Collilieux, X. & Boucher, C. (2006) Accuracy assessment of the ITRF datum definition. In: Xu P, Liu J, Dermanis A (eds) VI Hotine-Marussi symposium on theoretical and computational geodesy. *International Association of Geodesy Symposia*, 132, 101-110.
- Altamini, Z., Collilieux, X. & Metivier, L. (2011) ITRF08: An improved solution of the international terrestrial reference frame. *Journal of Geodesy*, 85, 457-473.
- Amiri-Simkooei, A. & Asgari, J. (2011) Harmonic analysis of total electron contents time series: Methodology and results. *GPS Solut*, 16, 77-88. doi: 10.1007/s10291-011-0208-x.
- Bogusz, J. & Klos, A. (2015) On the significance of periodic signals in noise analysis of GPS station coordinates time series. *GPS Solut*, 478. doi: 10.1007/s10291-015-0478-9.
- Dreyfus, G. (2005) *Neural Networks: Methodology and Applications*. Third Edition. Springer-Verlag, Berlin-Heidelberg, Germany.
- Gülal, E., Erdoğan, H. & Tiryakioglu, I. (2013) Research on the stability analysis of GNSS reference stations network by time series analysis. *Digital Signal Processing*, 23(6), 1945-1957. doi:10.1016/j.dsp.2013.06.014.
- Haykin, S. (1999) *Neural Networks: A Comprehensive Foundation*, Second Edition. Prentice Hall, New Yersey.
- He, X. & Xu, Sh. (2007) *Process Neural Networks: Theory and Applications*. Springer, China.

- Mémin, A., Watson, C., Haigh, I., MacPherson, L. & Tregoning, P. (2014) Non-linear motions of Australian geodetic stations induced by non-tidal ocean loading and the passage of tropical cyclones. *Journal of Geodesy*, 88(10), 927-940. doi:10.1007/s00190-014-0734-8.
- Li, C., Li, J. R., Huang, R. & Tan, L. (2014) Building up regional crustal movement velocity field with BP neural network base on Euler vector. *Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomatics and Information Science of Wuhan University*, 39 (3), 362-366.
- Li, J., Miyashita, K., Kato, T. & Miyazaki, S. (2000) GPS times series modeling by autoregressive moving average method: Application to the crustal deformation in central Japan. *Earth Planets Space*, 52, 155-162.
- Loesch, C. & Sari, S. (1996) *Redes Neurais Artificiais: Fundamentos e Modelos (Artificial Neural Networks: Foundations and Models)*. Editora da FURB, Blumenau, 166pp.
- Potukuchi, S. & Wexler, A. (1997) Predicting vapor pressures using neural network. *Atmospheric Environment*, 31(4), 741-753.
- Roggero, M. (2012) *Discontinuity detection and removal from data time series*. In: VII Hotine-Marussi Symposium on Mathematical Geodesy, Roma, 6-10 giugno 2009, 135-140.
- Tierra, A. & Romero, R. (2014) Planes coordinates transformation between PSAD56 to SIRGAS using a Multilayer Artificial Neural Network. *Geodesy and Cartography*, 63(2), 199-209. doi: 10.2478/geocart-2014-0014.
- Tierra, A., De Freitas, S. & Guevara, P. (2009) *Using Artificial Neural Network to Transformation of Coordinates from PSAD56 to SIRGAS95*. International Association of Geodesy Symposia: Geodetic Reference Frames. Springer-Verlag, Berlin-Germany, 134, 173-178. doi: 10.1007/978-3-642-00860-3.
- Tierra, A., Dalazoana, R. & De Freitas, S. (2008) Using artificial neural network to improve the transformation of coordinates between classical geodetic reference frames. *Computers & Geoscience*, 34(3), 181-189.
- Wang, J., Wang J. & Robersts, C. (2007) Investigations into a Dynamic Geocentric Datum. Observing our Changing Earth. *International Association of Geodesy*, 133, 11-19.