2006 International Joint Conference on Neural Networks
Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada
July 16-21, 2006

Solar Activity Forecasting by Incorporating Prior Knowledge from Nonlinear Dynamics into Neural Networks

Salvatore Marra and Francesco C. Morabito, *Senior Member, IEEE*

***Abstract*— This paper presents an efficient approach for the prediction of sunspot-related time series commonly used for monitoring solar activity, namely the Yearly Sunspot Number and the *R*12 Index. The method consists in matching a "de-rectification" procedure of sunspot data with the use of nonlinear dynamics tools in order to design neural network based predictors with close-to-optimal performance. In fact, whereas the "de-rectification" process allows to obtain time series that can be modeled by neural structures much better than the original datasets, the incorporation of the prior knowledge extracted by using nonlinear dynamics into neural networks generates models able to fully capture the chaotic dynamics of solar activity. The proposed approach produces prediction results that outperform the most accurate methods existing in literature both for short and medium-term forecasting horizons.**

I. Introduction

NOWADAYS, it is well-known that high-frequency radio communications, electric power transmission lines, space activities concerning operations of low-Earth orbiting satellites, long term climate variations, weather and a number of ionospheric parameters are affected by the emission of solar particles and electromagnetic radiations reaching the Earth. Consequently, forecasting the future behavior of the solar activity, that is strongly related to the number of dark spots observed on the sun (Fig. 1), is a topic of growing interest in the scientific community. The monthly and yearly sunspot numbers as well as the related time series are the most suitable indexes that are used to characterize the level of solar activity.

In this work, we use data provided by the *Sunspot Index Data Center* of the *Federation of Astronomical and Geophysical Data Analysis Services* [1]. In the last years, several studies have revealed the chaotic nature of sunspot data [2], [3] and many prediction methods have been applied to the related time series. In most cases, the used predictors are based on learning machines, particularly Artificial Neural Networks (ANNs) [4]-[7]. This is surely due to their well-known characteristics of adaptability and nonlinear universal mapping approximation. Solar activity, in fact, is very difficult to predict using standard models due to high frequency content, noise contamination, high dispersion level, high variability in phase and amplitude [3].

S. Marra and F. C. Morabito are with the Department of Informatics, Mathematics, Electronics and Transportation, University *Mediterranea* of Reggio Calabria, 89055 ITALY (Phone +39-0965-875224; fax: +39-0965-875220; e-mail: salvomarra@ing.unirc.it; morabito@unirc.it).

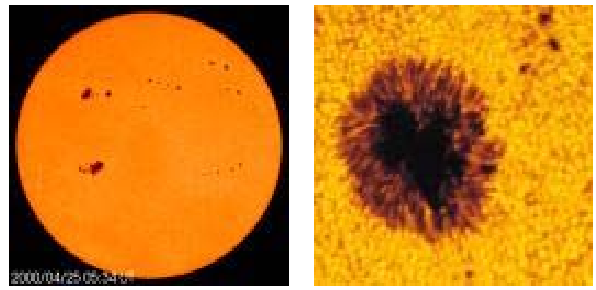

Fig. 1. Dark spots observed on the sun.

We can cite the work presented in [4], or the method in [7], that combines fossil and sunspot data to train a neural system to forecast the Yearly Sunspot Number. However, it is important to point out that interesting results have been also obtained using other prediction approaches [8], [9].

This paper describes a new methodology to forecast sunspot-related time series by exploiting Time Delay Neural Networks as prediction systems. Firstly, we "de-rectify" the original time series to obtain new data sets that can be modeled by neural architectures better than the original ones. According to one of our previous works [10], we use tools of the nonlinear dynamics theory to extract useful indications in order to generate the most suitable training set and design the network structure with close-to-optimal performance. Subsequently, after a pre-processing phase of training data and targets, we appropriately initialize the network weights and train the models. The proposed approach is fast, efficient, and allows to build neural architectures for sunspot-related time series that outperform the existing forecasting methods both for short and medium-term prediction horizons. The organization of the paper is as follows. The next Section describes the sunspot data considered in this work. In Section 3, an overview on time series prediction approaches is presented. The description of the proposed methodology is provided in Section 4. Section 5 is dedicated to the experimental results obtained and Section 6 gives the concluding remarks.

II. Sunspot Data

Solar activity is regularly monitored by many world observatories and research centers, that are able to provide the relative number of dark spots observed on the sun day after day. These data are recorded to give the so-called sunspot-related time series. In this work we consider the

Yearly Sunspot Number, shown in Fig. 2, that depicts the yearly number of dark spots from 1700 to 2004, and the *R*12 *Index* (shown in Fig. 9 later in this paper), also called the International Monthly Smoothed Sunspot Number, defined as follows:

$$R12_t = (MSN_{t-6}/2 + MSN_{t-5} + \ldots + MSN_{t+5} + MSN_{t+6}/2)/12, \quad (1)$$

where MSN_t is the number of dark spots observed during the month *t*. Most of the world observatories exploits this time series in order to forecast a part of solar activity. For example, the Paris-Meudon's observatory currently uses an ANN-based predictor system, developed by Fessant [5], for long-term predictions of the *R*12 Index. The temporal data we take into account start from January 1849 and end in March 2005 but, obviously, they continue up to now.

III. Review of Time Series Prediction Approaches

The prediction of a time series consists, given a set of past observations (x_t , x_{t-1}, ..., x_{t-n}), in finding the future value x_{t+m}, with m>0. The forecast $\tilde{x}_{t+m}$ is generally computed on the past history as following:

$$\tilde{x}_{t+m} = f(x_t, x_{t-1}, \ldots, x_{t-n}). \quad (2)$$

ANNs, that are well-known as universal function approximators, are extensively employed to estimate the unknown function *f*. In order to process temporal patterns, an ANN must contain memory. If the neural model is a Multi Layer Perceptron (MLP), the simplest way to build memory into the network is to feed it with a tapped delay line which stores past values of the input. This kind of architecture, shown in Fig. 3(a), is called Time Delay Neural Network (TDNN). However, the main problem is the choice of the optimal number of "taps" for the delay line. In fact, if the "time window" is too narrow, the network could not capture the information necessary for an adequate modelling, while if the window is too wide, some inputs may act as noise [11]. Recurrent Neural Networks (RNNs) do not need a time window to store the past values of a time series, since they possess an internal memory and, for this reason, seem the most suitable architectures to process temporal patterns. The Back Propagation Through Time algorithm (BPTT) is the most used one to train this kind of neural models. Unfortunately, unlike standard back-propagation algorithms, the BPTT is not guaranteed to converge to a local error minimum and it can be used with small networks sizes in the order of 3 to 20 units, because larger networks may require many hours of computation on current hardware [12]. So, it is not wrong to affirm that both feed-forward and recurrent models present advantages and disadvantages. RNNs, (Fig. 3(b)) are very useful when an unknown number of past values is necessary to correctly predict the next values, but when the time series under study presents chaotic behavior, TDNNs designed by incorporating the prior knowledge extracted from temporal data by using nonlinear dynamics tools can provide excellent performance [10]. In fact, given a chaotic time series (x_i , x_{i-1}, ... ,x_{i-n}), re-constructing the state space from which the time series originated and using the embedded data vectors $\boldsymbol{x_i}$, e. g.

$$\boldsymbol{x_i} = [x_i, x_{i-\tau}, x_{i-2\tau}, \ldots, x_{i-(d-1)\tau}], \quad (3)$$

as inputs of a predictor network can allow one to adequately capture the dynamics of the underlying system. In (3), τ is the "time delay", an integer indicating when the values x_i and $x_{i-\tau}$ are least correlated, and *d* is the "minimum embedding dimension", that is the number of variables necessary to reconstruct the dynamics. Both these parameters have to be correctly determined to obtain a good prediction accuracy.

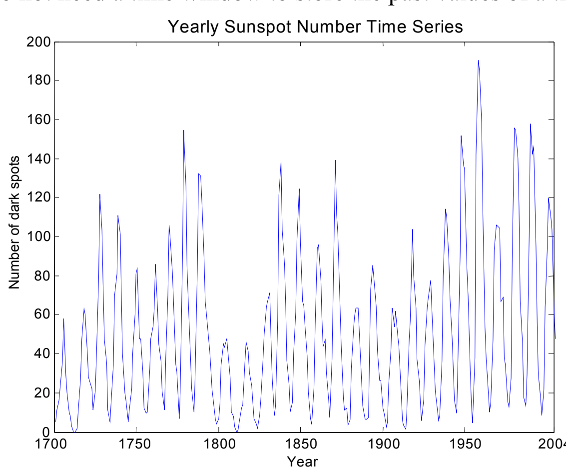

Fig. 2. Yearly Sunspot Number from 1700 to 2004.

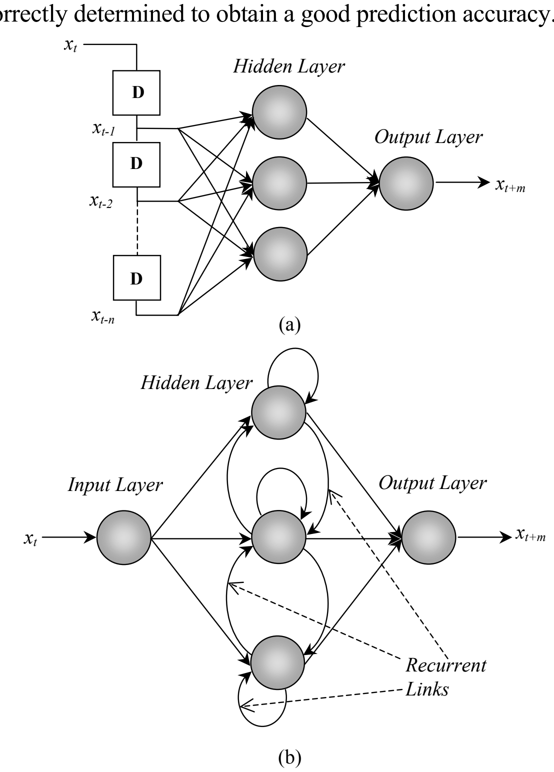

Fig. 3. (a) Time Delay Neural Network; (b) Recurrent Neural Network.

It is important to point out, indeed, that interesting results in time series prediction can be also obtained using local models, generally based on nearest-neighbor methods [13], clustering algorithms [9], and fuzzy inference systems [14].

IV. Incorporating Prior Knowledge from Nonlinear Dynamics into NNs for Sunspot Time Series Prediction

The methodology proposed in this work for sunspot data prediction has its strength in the efficient matching between the "de-rectification" process of the time series and the incorporation of the prior knowledge that can be extracted by nonlinear dynamics tools into TDNNs. In fact, we use a more natural representation of the sunspot time series, obtained by "de-rectifying" the temporal data in order that the sign of the signal is switched at every cycle minimum, as shown in Fig. 4 for the Yearly Sunspot Number. As suggested in [7], this is well motivated since the approximate 11 year solar cycle actually consists of a 22 year magnetic cycle that flips polarity every 11 years. In this way, the new time series appears much like a sum of sinusoids, whose oscillating pattern can be "learnt" much better than the original time series. Taking into account the chaotic behavior of sun activity and one of our previous works [10], we design the forecasting TDNNs with only one hidden layer and a number of taps in the input delay line at least equal to d-1, with d the minimum embedding dimension provided by Cao's method [15], an efficient tool that, unlike other classical methods, to compute the embedding dimension needs only the time delay τ, which can be determined by the first minimum of the Average Mutual Information function (AMI). The number of hidden units is fixed using the simple heuristic rule that the number of total parameters can be approximately equal to 1/5 of the training examples. After that, we normalize the training set and targets so that they have zero mean and unity standard deviation, initialize the weights with the Nguyen-Widrow's method [16] and train the network using the Levenberg

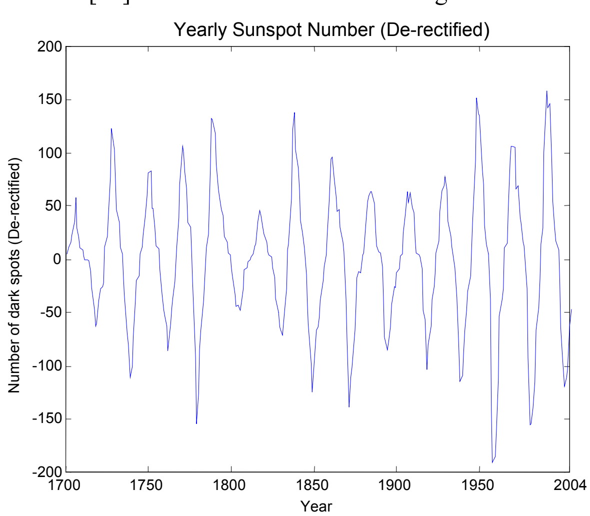

Fig. 4. De-rectification of the Yearly Sunspot Number.

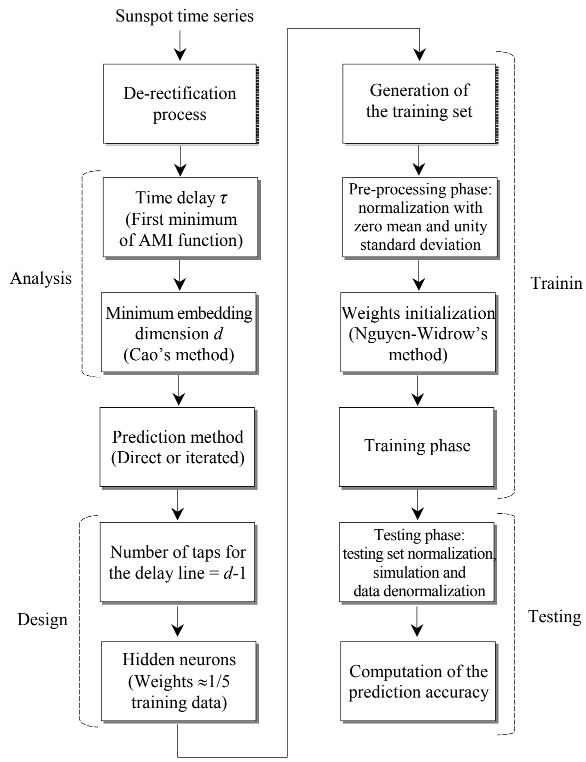

Fig. 5. Flow chart of the proposed methodology for sunspot data prediction.

Marquardt algorithm with Bayesian regularization [17]. Fig. 5 in detail describes the proposed method to design sunspot data predictors that will be explained as follows. After the de-rectification process, the time series under study is examined with AMI function and Cao's method. In the case of multi-step ahead predictions, we must select the forecasting method. In fact, multi-step ahead predictions can be obtained by means of the "direct method", consisting in training the model to directly predict the desired value, or the "iterated method", that recursively uses a single-step predictor until the required prediction horizon is reached. Subsequently, we can design the TDNN with the indications given by the analysis, generate the learning set, whose training sequences correspond to the embedded data vectors, normalize the data, initialize the network weights, train the network and, finally, verify the prediction accuracy.

To improve the performance of the model, we can increase the number of taps of the delay line and repeat the design procedure. However, increasing too much the number of taps does not always lead to better performance. This can be well explained taking into consideration the nature of sunspot data: in the case of chaotic time series, in fact, all relevant information about a next event is contained in a number of past values related to the dimension of the attractor from which the time series originated [10], [18], and too many inputs may act as noise.

V. EXPERIMENTAL RESULTS

Sunspot-related time series are often used to test the quality of a prediction method, so we can compare our approach with numerous results found in literature. MATLAB® is the exploited computing environment. According to the common practice, the prediction accuracy is evaluated in terms of the Normalized Mean Squared Error (NMSE), also called by some authors the Average Relative Variance (ARV):

$$\text{NMSE} = \frac{1}{\sigma^2 N} \sum_{i=1}^{N} (x_i - \tilde{x}_i)^2 \qquad (4)$$

where:
x_i = actual value of the i^{th} point of the series of length N;
$\tilde{x}_i$ = predicted value;
σ^2 = variance of the true time series in the interval N.

A. Yearly Sunspot Number

The Yearly Sunspot Number is the first time series we take into account. Data from 1700 to 1920 are generally used for the training phase, and the performance of the model is evaluated on three testing sets: from 1921 to 1955 (Test1), from 1956 to 1979 (Test2), and from 1980 to 1994 (Test3). After the de-rectification process, we use AMI function and Cao's method to determine τ, d and to generate the embedded data vectors. As depicted in Fig. 6, the minimum embedding dimension is 7, so we have trained TDNNs with a minimum number of 7 inputs (6 taps in the delay line) obtaining very interesting results with 8 inputs (7 taps in the delay line) and 5 neurons in the hidden layer for a total number of weights approximately equal to 1/5 of the training set. The Levenberg-Marquardt algorithm with Bayesian regularization allows to reach optimal performance in terms of speed convergence and generalization ability, but for this first time series we have found excellent results by exploiting the back propagation algorithm with variable learning rate and momentum[1].

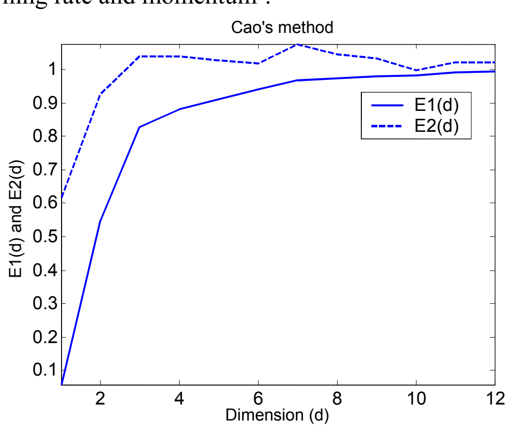

Fig. 6. Cao's method for the Yearly Sunspot Number.

[1]The MATLAB m-file to test the predictors designed in this work as well as the network weights and biases can be found in the "download section" of the page: http://neurolab.ing.unirc.it/salvatore.marra

TABLE I
PERFORMANCE COMPARISON FOR THE SINGLE-STEP PREDICTION OF THE YEARLY SUNSPOT NUMBER

Design Methods	Param.	NMSE Test1 (1921-1955)	NMSE Test2 (1956-1979)	NMSE Test3 (1980-1994)
WNet [20]	113	0.086	0.350	0.313
SSNet [22]	N/A	0.077	N/A	N/A
DRNN [21]	30	0.091	0.273	N/A
COMM [7]	N/A	0.065	0.240	0.188
Hybrid NN [24]	N/A	0.085	0.157	N/A
ScaleNet [23]	N/A	0.057	0.130	N/A
RNN+BPTT [11]	155	0.084	0.300	N/A
RNN+CBPTT [11]	15	0.092	0.251	N/A
HCA [9]	N/A	0.045	0.071	N/A
VGBP [19]	11	0.033	0.052	0.033
Proposed	51	0.022	0.065	0.027

This is not surprising since it is known that any training method is problem dependent and cannot pretend to be completely universal. Table I summarizes the achieved results for the single-step prediction obtained by the TDNNs designed and compares them with the best ones of many previous works on this time series. For each design method we give the number of free parameters used and the lowest NMSE (see [7], [11] and [19] for the main comparison results). From the results in Table I it is possible to observe that our approach outperforms feed-forward MLPs such as the Weight Elimination Feed Forward Network (WNet) [20], fully recurrent neural networks as the Dynamical Recurrent Neural Network (DRNN) [21], the Soft Weight Sharing Network (SSNet) [22], the Scale Neural Network (ScaleNet) [23], the Hybrid NN [24], the Committee Prediction method (COMM) presented in [7], the Hybrid Clustering Algorithm (HCA) [9], the recurrent networks trained with the BPTT algorithm and one of its variant, the Constructive BPTT (CBPTT) [11]. Only the Violation Guided Back Propagation technique (VGBP) [19] is able to reach a better performance on the Test 2. Fig. 7 shows the global aspect of the results on the three tests and for the period between 1995 and 2004, for which we obtain a NMSE of 0.045 (there are no comparison results for this test at moment). As regards the multi-step prediction task, by iteratively using the designed single-step predictor we have obtained the results gathered in Table II.

TABLE II
NMSES OBTAINED FOR MULTI STEP PREDICTIONS OF THE YEARLY SUNSPOT NUMBER CUMULATED TEST SET (1921-1979)

Steps ahead	Design Method: Proposed	RNN+EBPTT	RNN+CBPTT	RNN+BPTT
2	0.12	0.53	0.69	0.88
3	0.18	0.79	0.99	1.14
4	0.23	0.80	1.17	1.22
5	0.28	0.88	0.99	1.01
6	0.27	0.84	1.01	1.02

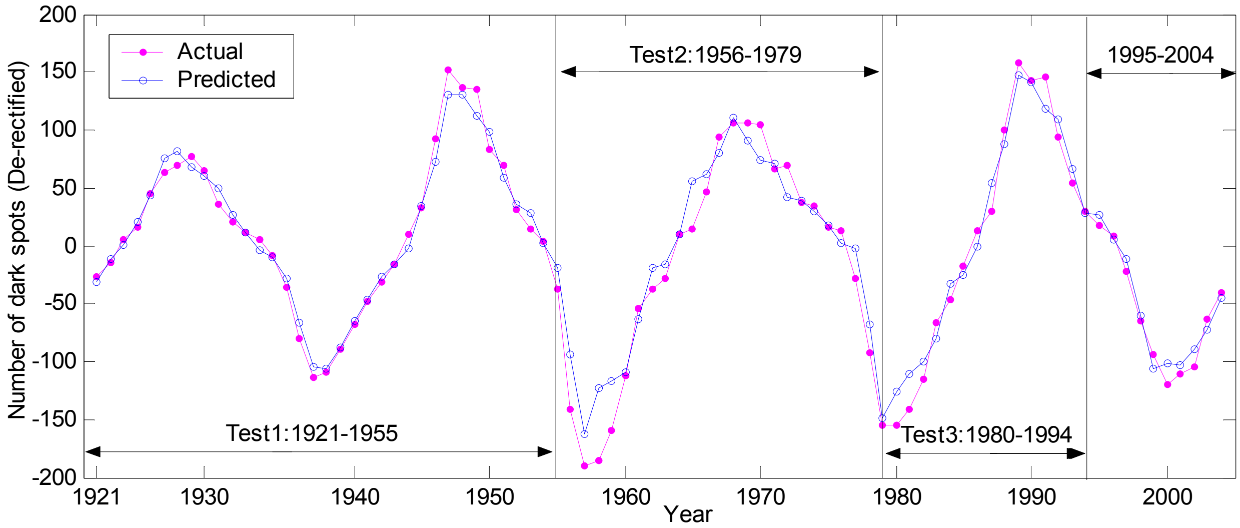

Fig. 7. Single-step prediction of the Yearly Sunspot Number test sets

In order to make comparisons with the most recent results published in literature about a medium-term horizon prediction of the Yearly Sunspot Number [26], we have computed the NMSEs on the cumulated test set involving Test1 and Test2. As one can see from Table II, in all the multi-step ahead predictions the TDNN designed presents the best performance.

Furthermore, for a correct comparison with the 4-steps ahead forecasting model proposed in [24], we use points from 1700 to 1955 as training set and data covering the period between 1956 and 1997 as testing set. The achieved forecasting accuracy is reported in Table III, whereas Fig. 8 graphically illustrates the considered multi-step prediction.

TABLE III
PERFORMANCE COMPARISON FOR THE 4-STEPS AHEAD PREDICTION OF THE YEARLY SUNSPOT NUMBER FROM 1956 TO 1997

Training set	Test set	Design method	NMSE
1700-1955	1956-1997	Network Committee	0.336
		Proposed	0.137

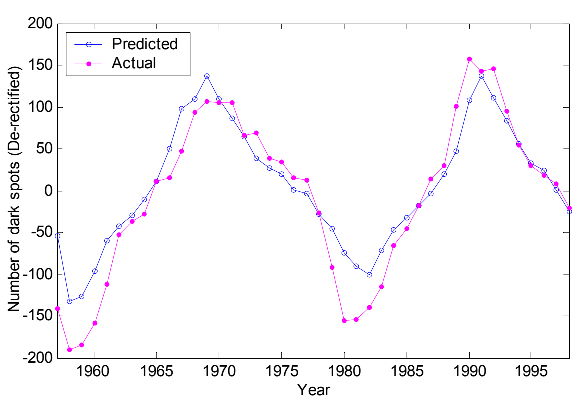

Fig. 8. 4-steps ahead prediction of the Yearly Sunspot Number from 1956 to 1997.

B. R12 Index

The International Monthly Smoothed Sunspot Number, also called *R*12 Index, is a particular filtered signal of the monthly sunspot number, that is the mean number of dark spots observed month by month. Fig. 9 shows both the time series from the January 1945 – January 2005 period, approximately. Various studies have revealed the intrinsic nonlinear dynamics of this time series [2], [3], [25]. In particular, on the basis of the nonlinear dynamic systems theory, Zhang [2] estimated accurately the fractal dimension $D = 2.8 \pm 0.1$, and the largest Lyapunov exponent $\lambda_1 = 0.023 \pm 0.004$ bits/month for the 1850-1992 period. The important information that can be obtained from this data is the existence of an upper limit for reliable predictions: 3.6 ± 0.6 years, that is 43 ± 7 months, approximately.

According to the procedure described in [8], we train TDNNs by using a learning set ranging from January 1849 to August 1991 and test the predictor models exploiting September 1991 - November 1995 data. The forecast of interest is 6-months ahead.

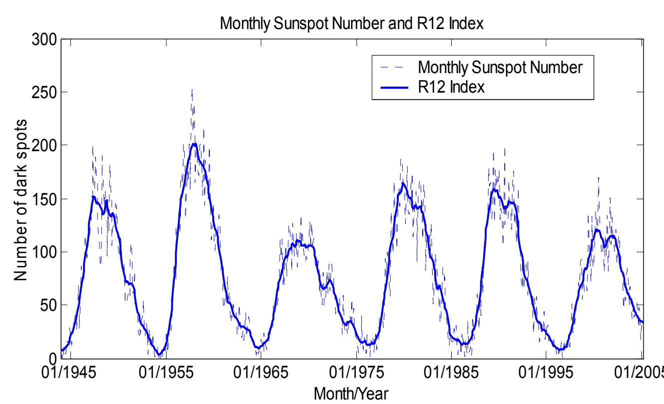

Fig. 9. Monthly sunspot number and *R*12 Index from January 1945 to January 2005.

For a further evaluation of the proposed method, we also verify the prediction accuracy on a larger testing set covering the period between September 1991 and March 2005. Firstly, we "de-rectify" the time series as previously explained for the Yearly Sunspot Number. The mutual information analysis gives the first minimum at 12, so we use this value as time delay whereas the Cao's method, depicted in Fig. 10, indicates a minimum embedding dimension of 6. Consequently, we have trained TDNNs with a minimum number of 5 "taps" in the input delay line obtaining optimal results with 8 "taps" and 5 neurons in the hidden layer. In this case, we use the direct method to forecast the $R12$ Index 6-months ahead. After generating the embedded data vectors that constitute the training set, we carry out their normalization with zero mean and unity standard deviation, train the network using Bayesian regularization and test the prediction accuracy. The achieved results are summarized in Table IV. We think that it is also important to consider the Strong Error Percentage (SEP), which represents the percentage of examples of the testing set for which the distance between the actual value and the predicted one is more than a fixed threshold, for example 20 ($R12$ values range from 0 to 201). This statistic is useful since it indicates whether there are many small errors or a few important ones.

For purpose of comparison, results of significant published prediction approaches have been considered [8]. From Table IV it is possible to observe that our TDNN outperform the recent Conditional Distribution Discrimination Tree method (CDDT) [8], but also the 43-12-1 MLP designed by Fessant [5], the McNish-Lincoln method [27] and the Stewart-Ostrow technique [28]. It is important to point out that the testing set of the last three approaches covers the period between September 1991 and November 1994 [5]. As regards the first two testing sets, we achieve not only the lowest NMSE, but also a SEP(20) of 0%, that is no 6-months ahead prediction error is greater than 20, whereas only the 8.5% of 163 predictions on the last testing set presents an error greater than 20. Fig. 11 depicts the 6-months ahead prediction carried out on part of the training set and the largest testing set (September 1991 - March 2005).

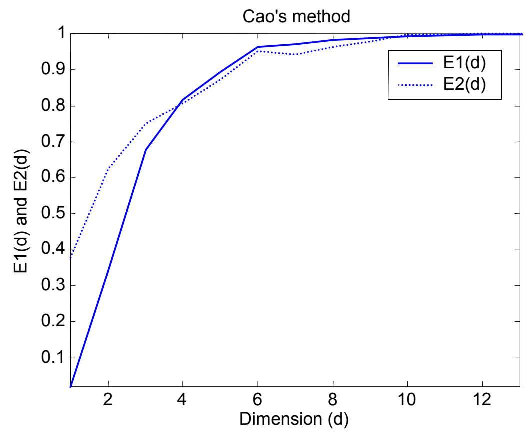

Fig. 10. Cao's method for the $R12$ Index.

TABLE IV
PERFORMANCE COMPARISON FOR THE 6-MONTHS AHEAD PREDICTION OF THE $R12$ INDEX ON DIFFERENT TESTING SETS

Training Set	Testing Set	Design Method	NMSE	SEP(20)
Jan. 1849- Aug. 1991	Sep. 1991- Nov. 1994 (39 points)	McNish - Lincoln	0.37	N/A
		43-12-1 MLP	0.028	N/A
		Stewart - Ostrow	0.026	N/A
		Proposed	0.021	0%
	Sep. 1991- Nov. 1995 (51 points)	CDDT	0.021	N/A
		Proposed	0.014	0%
	Sep. 1991- Mar. 2005 (163 points)	Proposed	0.026	8.5%

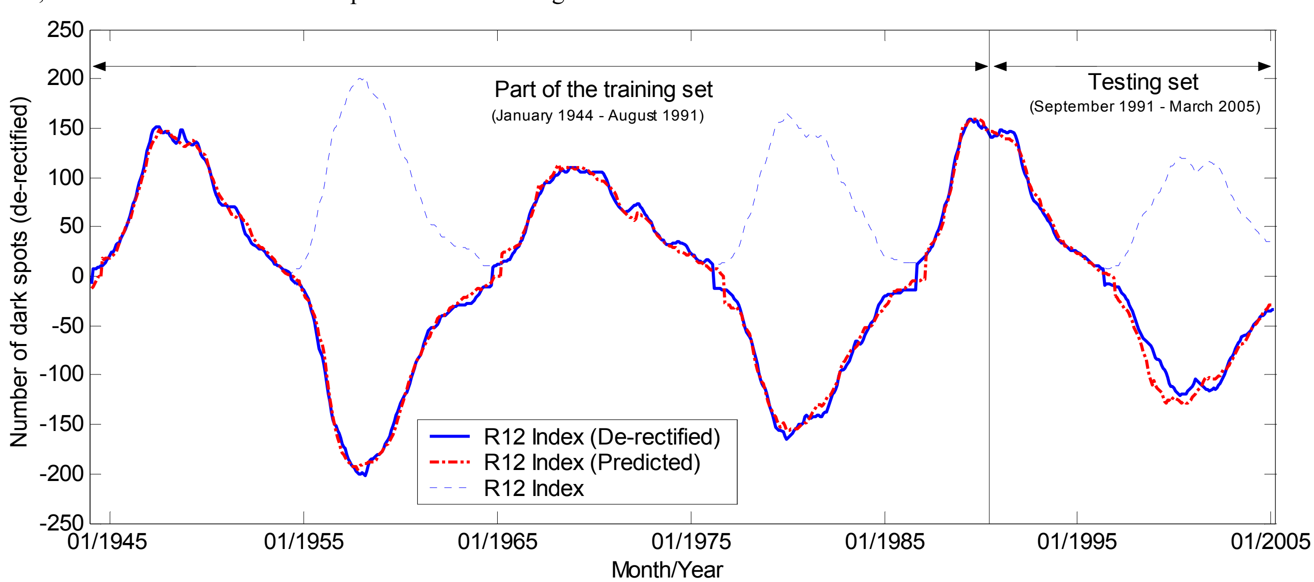

Fig. 11. 6-months ahead prediction of the $R12$ Index.

VI. Conclusion

In this paper, an efficient methodology for predicting the future behavior of solar activity has been presented. The idea to match the "de-rectification" process of sunspot data with the incorporation of the prior knowledge extracted by nonlinear dynamics tools into Time Delay Neural Networks has proved to be the best neural architecture to forecast this kind of temporal patterns. In particular, the indications provided by the Average Mutual Information function and the Cao's method for the de-rectified *R*12 Index data have been of fundamental importance to generate not only a suitable training set for the predictor model, but also a neural architecture with close-to-optimal performance.

To our knowledge, the results outperform other sunspot forecasting methods. Our method also proved effective out to the "maximum theoretical forecast horizon" of 4 years for sunspot data. Finally, we would like to point out the great importance of sunspot time series predictions, remembering that *Skylab*, the space laboratory launched by the USA in the 1973, was brought to a premature demise in 1979 due to improperly forecasting increased atmospheric drag associated to a sunspot maximum [7].

Acknowledgment

The authors are grateful to one of the referees whose interest and suggestions will be priceless for future works. It is a pleasure to work with these fine people.

References

[1] R. A. M. Van der Linden & the SIDC team. *Online catalogue of the sunspot index*. Available: http://sidc.oma.be/html/sunspot.html

[2] Q. Zhang, "A nonlinear prediction of the smoothed monthly sunspot numbers," *Astronomy and Astrophysics*, vol. 310, pp. 646-650, 1996.

[3] S. Sello, "Solar cycle forecasting: a nonlinear dynamics approach," *Astronomy and Astrophysics*, vol. 377, pp. 312-320, 2001.

[4] S. Bengio, F. Fessant and D. Collobert, "On the prediction of solar activity using different neural networks models," *Annales Geophysicae*, vol. 14, pp.20-26, 1995.

[5] F. Fessant, C. Pierret and P. Lantos, "Comparison of neural network and Mc-Nish and Lincoln methods for the prediction of the smoothed sunspot index," *Solar Physics*, vol. 168, pp. 423-433, 1996.

[6] Lj. R., Cander and X. Lamming, "Neural networks in ionospheric prediction and short-term forecasting," *Proceedings of the IEEE 10th International Conference on Antennas and Propagation*, vol. 2, pp. 14-17, 1997.

[7] E. A. Wan, "Combining fossils and sunspots: committee predictions," *Proceedings of the IEEE International Conference on Neural Networks*, vol. 4, pp. 2176-2180, 1997.

[8] M. G. Genet and A. Pétrowsky, "Sunspot number prediction by a conditional distribution discrimination tree," *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 1, pp. 814-819, 2003.

[9] A. Sfetsos and C. Siriopoulos "Time series forecasting with a hybrid clustering scheme and pattern recognition," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans,* vol. 344, no. 3, pp. 399- 405, 2004.

[10] S. Marra, F. C. Morabito and M. Versaci, "Design of neural predictors using tools of chaos theory and Bayesian learning," *Proceedings of the IEEE International Joint Conference on Neural Networks,* pp. 2222-2227, 2005.

[11] R. Boné, M. Crucianu and J. P. Asselin de Beauville, "An algorithm for the addition of time-delayed connections to recurrent neural networks," *Proceedings of the ESANN'2000*, Bruges, Belgium, pp. 293-298, 2000.

[12] H. Jaeger, "A tutorial on training recurrent neural networks covering BPPT, RTRL, EKF and the "echo state network" approach", *GMD Report 159*, German National Research Center for Information Technology, 2002.

[13] J. Vesanto, "Using the SOM and local models in time series prediction," in *Proceedings of the Workshop on Self-Organizing Maps (WSOM'97)*, Espoo, Finland, pp. 209-214, 1997.

[14] S. H. Ling, H. K. Lam, F. H. F. Leung and Y. S. Lee, "A genetic algorithm based fuzzy tuned neural network," *Proceedings of the 12th IEEE International Conference on Fuzzy Systems*, vol. 1, pp. 220-225, 2003.

[15] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series", *Physica D*, vol. 110, pp. 43-50, 1997.

[16] D. Nguyen and B. Widrow, "Improving the learning speed of two layer neural networks by choosing initial values of the adaptive weights," *Proceedings of the International Joint Conference on Neural Networks*, vol. 3, pp. 21-26, 1990.

[17] D. J. MacKay, "Bayesian interpolation", *Neural Computation,* vol. 4, no. 2, pp. 415-447, 1992.

[18] I. De Falco, A. Iazzetta, P. Natale and E. Tarantino, "Evolutionary neural networks for nonlinear dynamcis modelling," *Lecture Notes in Computer Science*, vol. 1498, pp. 593-602, 1998.

[19] B. W. Wah and M. L. Qian, "Time-series predictions using constrained formulations for neural-network training and cross validation", in *Proceedings of the Int'l Conference on Intelligent Information Processing, 16th IFIP World Computer Congress*, pp. 220-226, 2000.

[20] A. S. Weigend, B. A. Huberman, D. E. Rumelhart, "Predicting sunspot and exchange rates with connectionist networks," in *Nonlinear Modeling and Forecasting*, M. Casdagli and S. G. Eubank, Eds. Reading, MA: Addison Wesley, 1992, pp. 395-432.

[21] A. Aussem, "Dynamical recurrent neural networks towards prediction and modeling of dynamical systems," *Neurocomputing,* vol. 28, pp. 207-232, 1999.

[22] S. Nowland, G. Hinton, "Simplifying neural networks by soft weight sharing," *Neural Computation*, vol. 4, no. 4 pp. 473-493, 1992.

[23] A. B. Geva, "ScaleNet - Multiscale neural-network architecture for time series prediction," *IEEE Transactions on Neural Networks*, vol. 9, no. 5, pp. 1471-1482, 1998.

[24] G. M. Cloarec and J. Ringwood, "Incorporation of statistical methods in multi-step neural network prediction models," *Proceedings of the IEEE International Joint Conference on Neural Networks,* vol. 3, pp. 2513-2518, 1998.

[25] M. Han, J. Xi and S. Xu, "Chaotic system identification based on Kalman Filter," *Proceedings of the IEEE International Joint Conference on Neural Networks*, pp. 675-680, 2002.

[26] R. Boné and M. Crucianu, "An evaluation of constructive algorithms for recurrent networks on multi-step ahead prediction", *International Conference on Neural Information Processing (ICONIP'02)*, Singapore, vol. 2, pp. 547-551, 2002.

[27] A. G. McNish and J. V. Lincoln, "Prediction of sunspot numbers," *Eos, Transactions American Geophysical Union*, vol. 30, pp. 673, 1949.

[28] F. G. Stewart and S. M. Ostrow, *Journal des telecommunications*, vol. 37, pp. 228, 1970.