Research in Astron. Astrophys. **2011** Vol. **11** No. **4**, 491–496
http://www.raa-journal.org *http://www.iop.org/journals/raa*

Research in Astronomy and Astrophysics

Neural network prediction of solar cycle 24

A. Ajabshirizadeh[1], N. Masoumzadeh Jouzdani[1,3] and Shahram Abbassi[2,3]

[1] Department of Physics, University of Tabriz, Tabriz, Iran
[2] School of Physics, Damghan University, P. O. Box 36175–364 Damghan, Iran; *abbassi@mail.ipm.ir*
[3] School of Astronomy, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

Received 2010 October 13; accepted 2010 December 7

Abstract The ability to predict the future behavior of solar activity has become extremely import due to its effect on the environment near the Earth. Predictions of both the amplitude and timing of the next solar cycle will assist in estimating the various consequences of space weather. The level of solar activity is usually expressed by international sunspot number (R_z). Several prediction techniques have been applied and have achieved varying degrees of success in the domain of solar activity prediction. We predict a solar index (R_z) in solar cycle 24 by using a neural network method. The neural network technique is used to analyze the time series of solar activity. According to our predictions of yearly sunspot number, the maximum of cycle 24 will occur in the year 2013 and will have an annual mean sunspot number of 65. Finally, we discuss our results in order to compare them with other suggested predictions.

Key words: Sun: activity — sunspots — neural networks — prediction

1 INTRODUCTION

The successful prediction of a future event is arguably the most powerful way of confirming a scientific theory. Commonly in physics, a theory that describes a system in the natural world is regarded as correct and therefore useful if it can use the state of the system at one time to describe the state of the system at some other past or future time.

The prediction of solar activity over a few years is the oldest problem in solar physics, which arose as soon as the solar cycle itself was discovered. Unfortunately, this problem has not been solved, probably because the series of observational data available were not long enough for purely statistical analysis, and because we do not quite understand the physical nature of this phenomenon.

Most space weather phenomena are influenced by variations in solar activity. During the years of solar maximum, there are more solar flares causing a significant increase in solar cosmic ray intensity. The high-energy particles disrupt communication systems and shorten the lifetime of satellites. Coronal mass ejections and solar flares are the origin of shocks in the solar wind and cause geomagnetic disturbances in the earth's magnetosphere. The high rate of geomagnetic storms and sub-storms results in atmospheric heating and drag of Low Earth Orbit (LEO) satellites. Solar activity forecasting is especially useful to space mission centers since the orbital trajectory parameters of satellites are greatly affected by variations of solar activity. A dramatic solar weather event may not only disturb the orbits of satellites in the Earth's upper atmosphere, but also affect power grids on the ground, e.g. the power cuts in Quebec, Canada in 1989. The level of solar activity is usually expressed by the Zurich or International sunspot number.

Although this type of solar activity presents some clear periodicities, its prediction is quite difficult but not impossible, since the large range of forecasting methods used to predict the occurrence and amplitude of the solar cycle can be categorized into two types of models: statistical models and physical models. In statistical models, it is usual to represent the evolution of a physical system by using a time series. In contrast with a physical model, the statistical model only attempts to explain the system, and in particular the time series associated with it, in terms of itself, and perhaps in terms of correlations with other time series associated with the system. At this point, it is appropriate to address a common concern, which for obvious reasons is most usually expressed by physicists: what reasons are there for constructing a model that contains no physical understanding? Here are three reasons. Firstly, simply writing down the data set as a time series, together with organizing and examining it, is the first step in the scientific method: analyzing the sequence as a time series governed by a statistical model is a natural first step, until such time as a physical theory can be formulated. The second reason is that predictions from a statistical model might simply be useful in their own right. For example, in day to day life it makes no difference to most people whether the weather forecast was made from a statistical model or from a physical one. The final and most important reason is that it might be impossible for the physical system to be predicted from the basic physical principles governing it. This can be because the system is simply too complicated, which, for example, is the case for a system with plasma (Conway 1998).

One of the statistical models used for predicting the data is an artificial neural network method. The use of artificial neural networks has been recognized recently as a promising way of making predictions on temporal series with chaotic or irregular behavior (Weigend 1990). This technique has already been applied in the framework of solar-terrestrial physics for prediction of geomagnetic induced currents and storms (Lundstedt 1992) and as a way of recognizing a pattern during the onset of a new sunspot cycle (Koons 1990).

The aim of this paper is to predict the solar cycle. The structure of the paper is as follows. In Section 2, we provide a brief summary of the neural network methodology employed. In Section 3, we introduce the results of our network architecture to generate our best estimate of the behavior of cycle 24, and in Section 4, the conclusions and their comparison are presented.

2 ARTIFICIAL NEURAL NETWORK

An artificial neural network (ANN) is an information-processing system consisting of a large number of simple processing elements called neurons or units. A Neural Network (NN) system is characterized by (i) its pattern of connections between the neurons, (ii) its method of determining the weights in the connections (training or learning algorithm) and (iii) its activation function. In other words, ANNs are parallel computing systems that are widely used in prediction, pattern recognition and classification. Neural networks with a sufficient number of hidden units can approximate any nonlinear function to any degree of accuracy (Hornik et al. 1989).

There exist various types of NNs; however, for our predictions, we have used the most popular and simple NN, which is the Feed Forward Neural Network (FFNN) employing the Levenberg-Marquardt error learning algorithm (Levenberg 1944; Marquardt 1963). Although the learning algorithm using the back-propagation of errors (Rumelhart 1986) is more famous and usually used in the FFNN, it is also known as an algorithm with a very poor convergence rate. More significant improvement was possible by using various second order approaches such as Newton, conjugate gradient, or the Levenberg-Marquardt (LM) method. The LM algorithm is now considered as the most efficient. It combines the speed of the Newton algorithm with the stability of the steepest decent method (Hagen et al. 1994).

In an FFNN arrangement (Fig. 1), neurons (units) between layers are connected in a forward direction. Neurons in a given layer do not connect to each other and do not take inputs from subsequent layers. The input units send the signals to the hidden units, which then process the received information and pass the results to output units. The output units produce the final response to the input signals. A database of historical data describing the relationship between a set of inputs and known outputs is used to define the input and output units. Feed Forward networks often have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear activation functions allow the network to learn nonlinear and linear relationships between input and output vectors. The linear output layer lets the network produce values outside the range –1 to +1.

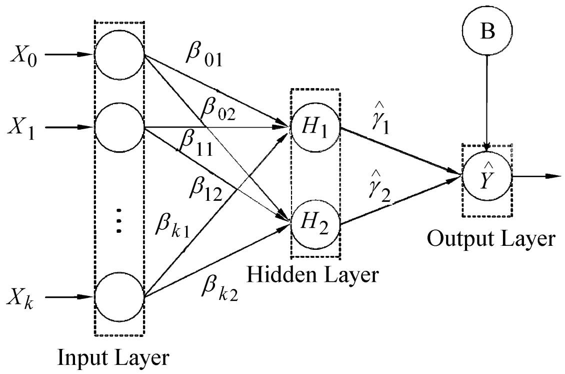

Fig. 1 An FFNN with one hidden layer and one output.

A typical FFNN is depicted as follows: An algebraic form of the neural network can be written

$$Y = y_0 + \sum_{j=1}^{h} y_j f_j(X, w_j), \tag{1}$$

where w_j is the vector of weights for the jth neuron, $X = (X_1, X_2, ..., X_k)$ is a vector of explanatory variables, and $f_j(X, w_j) = G(w_{0j} + X'w_j), j = 1, ..., h$ shows output of the hidden units. The function G is any activation function such as the hyperbolic tangent function $G(n) = (\mathrm{e}^n - \mathrm{e}^{-n})(\mathrm{e}^n + \mathrm{e}^{-n})^{-1}$ or logistic function $G(n) = (1 + \mathrm{e}^{-n})^{-1}$. An FFNN can be composed of more than one hidden layer with multiple outputs, which is similar to the system of nonlinear regression equations.

The FFNN is organized here with three layers: input, hidden, and output layers. The activation function in the first layer is log-sigmoid, and the output layer activation function is linear. Units between layers are connected by weights that are optimized for a minimum of the root mean square error (RMSE) between a known output and the predicted output. Training is the process by which the weights are adjusted according to the LM algorithm. A simplified definition of an NN is a computer program that has been trained to learn the relationship between a given set of inputs and a known output. In general, training an NN requires an optimum network architecture and sufficient historical information about the time series. The architecture of a feed-forward network is specified by the number of neurons used in the input, hidden and output layers of the network. The input layer needs a sufficient number of neurons so that the network has enough access to the recent history of the time series. The hidden layer of the network is responsible for the nonlinear processing capability of the network and as such needs to have sufficient neurons to represent the underlying complexity of the time series. We only consider networks with one output, which is required to produce a prediction a number of years ahead of the most recent input. NNs are trained until the RMSE between the output values predicted by the network and the target output values has reached a minimum. At this point, we say that the optimum result has been reached for the given situation. As applied to Sunspot Number (SSN) prediction, the RMSE was defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mathrm{SSN_{obs}} - \mathrm{SSN_{pred}})^2}, \tag{2}$$

where N is the number of training patterns, and $\mathrm{SSN_{obs}}$ and $\mathrm{SSN_{pred}}$ are the observed and predicted SSN values respectively. Generally, the time series is split into two data sets: a training set and a testing set. The training set is used to adjust the weights during training, while the testing set is used to verify the prediction performances of the network. Neural networks with large numbers of parameters are more at risk of overfitting. Overfitting is the problem of very bad predictions for the output of sample data in spite of having very good results for input sample data. Here, for overcoming this problem, we used an early stopping method which stops training when the validation set fails to reduce the validation sample's RMSE (Baum 1989).

Finally, before going on to present our results, we mention two different ways in which neural networks can be used to produce predictions. Firstly, in what we term "direct prediction," the network only relies on actual known data to generate any predictions. Consequently, the furthest ahead a prediction is obtainable would be limited by the last known data point in the time series plus the predict-ahead-time of the individual network. Alternatively, networks can be used to iteratively predict, sometimes called multistep prediction, in which the networks' predictions are subsequently fed back into the input layer as new data points. This potentially allows networks to predict arbitrarily far ahead; in practice, as predicted values make up more and more of the supposedly known input data, errors can be recursively compounded until no estimate of accuracy in the results can be calculated (Conway 1998).

3 FORECASTING SOLAR INDICES

Here we wish to predict up to 10 yr ahead, and consequently, we use yearly sunspot number since the use of monthly data would potentially require a large network. Furthermore, Hoyt et al. (1994) have shown that some of Wolf's reconstructed values were wrong, particularly for the early cycles 1–7. Thus, only the post–1850 data can be considered wholly reliable. It should be considered that cycle 23 began in 1996 May and reached its maximum in 2000 April, and now it is inferred to end in 2008 December (or probably later); therefore, its length should be 12.6 yr (or longer) (Li 2009). The yearly mean values of sunspot number were obtained electronically from the website: *ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/*.

Regarding time difference between data accessing, we choose various network architectures for our time series. After a large number of trial and error cases, we processed sunspot number time series with a neural network having a 128–42–1 structure, which means we used the sunspot values for the years 1882–2009 as the training set.

For the sunspot number R_z, we obtained 2013 as the year of the next maximum with a value of around 65. Regarding the accuracy of the year of maximum prediction, for the two cycles predicted with this network, in two cases, the date of maximum was predicted correctly. Using a comparison between the predicted value and observed value of Solar Cycles (SCs) 22 and 23, the uncertainty in the value of the sunspot maximum has been obtained to be ± 13. All of the predicted values for sunspot number have been added to Table 1. Also, comparison between the 1986–2009 observed sunspot number and the predictions of neural networks are shown in Figure 2 as well as the predicted shape and amplitude of SC 24 in terms of yearly sunspot numbers.

Table 1 Predicted Values of Sunspot Number from 2008 to 2018

Year	Predicted values	min RMSE	Year	Predicted values	min RMSE	Year	Predicted values	min RMSE
1986	12.68	0.060	1997	12.40	0.055	2008	14.46	0.039
1987	36.99	0.064	1998	66.74	0.053	2009	16.23	0.037
1988	86.17	0.057	1999	114.68	0.051	2010	17.91	0.049
1989	144.80	0.060	2000	132.90	0.056	2011	43.50	0.045
1990	135.97	0.055	2001	115.51	0.060	2012	57.64	0.047
1991	124.08	0.058	2002	104.19	0.055	2013	65.43	0.045
1992	92.14	0.053	2003	64.75	0.051	2014	56.74	0.042
1993	57.79	0.053	2004	42.20	0.051	2015	48.37	0.042
1994	38.55	0.051	2005	27.37	0.043	2016	18.58	0.041
1995	20.63	0.046	2006	19.94	0.039	2017	10.82	0.040
1996	11.22	0.052	2007	15.40	0.042	2018	14.17	0.041

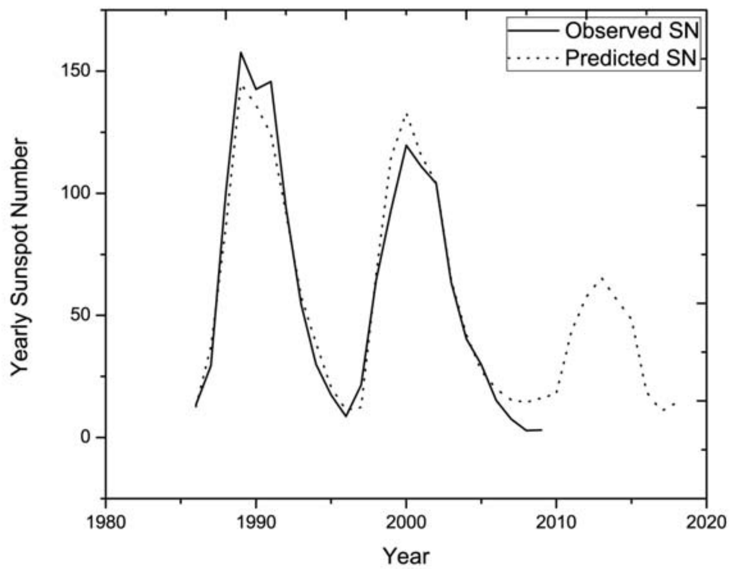

Fig. 2 Observed SCs 22 and 23 (*solid line*) and the predicted SCs 22, 23 and 24 (*dashed line*) in terms of yearly mean sunspot numbers.

4 DISCUSSION AND CONCLUSIONS

Our neural network method is based on one hidden layer. For reliable results, we use multi-step prediction having only one reasonable output. In terms of processing data, by changing the back-propagation algorithm to the the LM algorithm, our feed-forward neural network model becomes faster since the LM algorithm speeds up convergence while limiting memory requirements (Battiti 1992). We saw a similarity between the predicted SCs 24 and 20. We predict SC 24 will have a maximum of 65 ± 13 occurring in 2013. In general, our result is close to other predictions made for SC 24. For example, Li et al. (2005) obtained 2013 for the maximum of cycle 24 with a statistical method. Also, a recent article by Wang et al. (2009), using similar descending phases and a cycle grouping, predicted that peak amplitude for that monthly smoothed sunspot number in solar cycle 24 is near 100.2 ± 7.5, occurring in 2012. Furthermore, Chumak & Matveychuk (2010) predicted that the maximum amplitude of cycle 24 is 90 ± 20, which is in agreement with our results. Finally, our prediction fits well within the limits of other works as indicated in Pesnell (2008), where an average cycle was predicted using other methods such as statistical and precursor methods.

References

Battiti, R. 1992, Neural Computation, 4, 141

Baum, E. B., & Haussler, D. 1989, Neural Computation, 1, 151

Chumak, O. V., & Matveychuk, T. V. 2010, RAA (Research in Astronomy and Astrophysics), 10, 935

Conway, A. J. 1998, New Astronomy Reviews, 42, 343

Hagan, M. T., & Menhaj, M., 1994, IEEE Transaction on Neural Networks, 5, 989

Hornik, K., Stinchcombe, M., & White, H. 1989, Neural Networks, 2, 359

Hoyt, D. V., Schatten, K. H., & Nesme-Ribes, E. 1994, Geophys. Res. Lett., 21, 2067

Koons, H. C., & Gorney, D. J. 1990, EOS Transactions, 71, 677

Levenberg, K. 1944, Quarterly of Applied Mathematics, 2, 164

Li, K.-J. 2009, RAA (Research in Astronomy and Astrophysics), 9, 959

Li, K.-J., Gao, P.-X., & Su, T.-W. 2005, ChJAA (Chin. J. Astron. Astrophys.), 5, 539

Lundstedt, H. 1992, Planet. Space Sci., 40, 457

Marquardt, D. W. 1963, SIAM Journal on Applied Mathematics, 2, 431

Pesnell, W. D. 2008, Sol. Phys., 252, 209

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. 1986, Nature, 323, 533

Wang, J.-L., Zong, W.-G., Le, G.-M., Zhao, H.-J., Tang, Y.-Q., & Zhang, Y. 2009, RAA (Research in Astronomy and Astrophysics), 9, 133

Weigend, A. S., Huberman, B. A., & Rumelhart, D. E. 1990, Int. J. Neural Syst., 1, 193