Multi Step Ahead Prediction of North and South Hemisphere Sun Spots Chaotic Time Series using Focused Time Lagged Recurrent Neural Network Model

[1]SANJAY L. BADJATE [2]SANJAY V. DUDUL
[1]Jawaharlal Darda Institute of Engineering & Technology M.I.D.C. Yavatmal (M.S.), India
[2] Applied Electronics Department Sant Gadgebaba University, Amravati University ,Amravati India
e-mail : [1]s_badjate@rediffmail.com, [2]dudulsv@rediffmail.com

Abstract:- Multi–Step ahead prediction of a chaotic time series is a difficult task that has attracted increasing interest in recent years. The interest in this work is the development of nonlinear neural network models for the purpose of building multi-step ahead prediction of North and South hemisphere sunspots chaotic time series. In the literature there is a wide range of different approaches but their success depends on the predicting performance of the individual methods. Also the most popular neural models are based on the statistical and traditional feed forward neural networks. But it is seen that this kind of neural model may present some disadvantages when long-term prediction is required. In this paper focused time lagged recurrent neural network (FTLRNN) model with gamma memory is developed not only for short-term but also for long-term prediction which allows to obtain better predictions of northern and southern chaotic time series in future. The authors experimented the performance of this FTLRNN model on predicting the dynamic behavior of typical northern and southern sunspots chaotic time series. Static MLP model is also attempted and compared against the proposed model on the performance measures like mean squared error (MSE), Normalized mean squared error (NMSE) and Correlation Coefficient (*r*) .The standard back propagation algorithm with momentum term has been used for both the models. The various parameters like number of hidden layers, number of processing elements in the hidden layer, step size, the different learning rules, the various transfer functions like tanh, sigmoid, linear-tanh and linear sigmoid, different error norms L_1,L_2 (Euclidean), L_3, L_4 ,L_5 and L_∞, and different combination of training and testing samples are exhaustively varied and experimented for obtaining the optimal values of performance measures. The obtained results indicates the superior performance of estimated dynamic FTLRNN based model with gamma memory over the static MLP NN in various performance metrics. In addition, the output of proposed FTLRNN neural network model with gamma memory closely follows the desired output for multi- step ahead prediction for all the chaotic time series considered in the study.

Keywords : Sunspots chaotic time series, multi- step prediction, Focused time lagged neural network (FTLRNN) , Multilayer perceptron (MLP), Self organizing feature map (SOFM).

1. Introduction

Predicting the future which has been the goal of many research activities in the last century is an important problem for human, arising from the fear of unknown phenomenon and calamities all around the infinitely large world with its many variables showing highly nonlinear and chaotic behavior. Chaotic time series have many applications in various fields of Science, e.g. astrophysics, fluid mechanics, medicine, stock market, weather, and is also useful in engineering such as speech coding [1], radar detection,modeling of electromagnetic wave propagation and scattering [2].The chaotic inter connected complex dynamical systems in nature are characterized by high sensitivity to initial conditions which results in long term unpredictability. The dynamical reconstruction seems to be extremely difficult, even in developing era of super computers, not because of computational complexity, but due to inaccessibility of perfect inputs and state variables. Many different methods have been developed to deal with chaotic time series prediction. Among them neural networks occupy an important place being adequately model the nonlinearity and nonstationarity.

Inspired from the structure of the human brain and the way it is supposed to operate, neural networks are parallel computational systems capable of solving number of complex problems in such a diverse areas as pattern recognition , computer vision , robotics , control and medical diagnosis, to name just few [3].Neural networks are an effective tool to perform any nonlinear input output mappings and prediction problem[4].. Predicting a chaotic time series using a neural network is of particular interest [5]. Not only it is an efficient method to reconstruct a dynamical system from an observed time series, but it also has many applications in engineering problems radar like noise cancellation [6] radar detection [7], demodulation of chaotic secure communication systems [8] and spread spectrum /code division multiple access (CDMA) systems [9,10] .It is already established that, under appropriate conditions, they are able to uniformly approximate any complex continuous function to any desired degree of accuracy [11]. Later, similar results were published independently in [12]. It is these fundamental results that allow us to employ neural network in time series prediction. Since neural networks models do not need any a priory assumption about the underlying statistical distribution of the series to be predicted, they are commonly classified as “data- driven” approach, to contrast them with the “model-driven” statistical methods. Neural networks are the instruments in broad sense can learn the complex nonlinear mappings from the set of observations [13]. The static MLP network has gained an immense popularity from numerous practical application published over the past decade, there seems to be substantial evidence that multilayer perceptron indeed possesses an impressive ability [14]. There have been some theoretical results that try to explain the reasons for the success [15] and [16]. Most applications are based on feed forward neural networks, such as the back propagation (BP) network [17] and Radial basis function (RBF) network [18-19]. It has also been shown that modeling capacity of feed forward neural networks can be improved if the iteration of the network is incorporated into the learning process [20]. Several methods with different performance measures have been attempted in the literature to predict the Chaotic time series. A new class of wavelet network is develop with a standard deviation of 0.0029 for short term ahead prediction of Mackey-Glass chaotic time series and annual sunspots for 1 step ahead prediction [21].By using recurrent predictor neural network for monthly sunspots chaotic time series for 6 months ahead prediction with EPA equals to 0.992 and ERMSE equals 4.419 , for 10 months ahead prediction with EPA of 0.980 and ERMSE of 7.050 , for 15 months ahead prediction of EPA 0.9222 ERMSE equals 13.658 and 20 months ahead prediction EPA of 0.866 and ERMSE of 16.79323 [22]. By using radial basis function with orthogonal least square Fuzzy model for monthly sunspots with prediction error +6 to -4 and for Mackey-Glass chaotic time series with RMSE of 0.0015 [23].It is also attempted with Hybrid network for Mackey-Glass time series with iterative prediction and NMSE of 0.053 [24].By using Elman neural network for yearly sun spots for 1 year ahead prediction with E_{rmse} equals 30.2931 and prediction accuracy of 0.9732[25] .

From the scrupulous review of the related research work, it is noticed that no simple model is available for long term prediction of sunspots chaotic time series so far and for the individual North and South hemisphere sunspots chaotic time series . It is necessary to develop a simple model that is able to perform short , medium and long term prediction of such individual southern and northern hemisphere chaotic time series with reasonable accuracy . In view of the remarkable ability of neural network in learning from the instances, it can prove as a potential candidate with a view to design a versatile predictor (forecaster) for the such chaotic time series.

The paper is organized as follows. First the optimal static NN based model on MLP is attempted to model the given system. Next on the same parameter the self organizing feature map and best dynamic focused time lagged NN model with built in gamma memory is estimated for prediction for all the short term and long term ahead prediction. Next the comparison between these models are carried out on the basis of the performance measures such as Mean Square Error (MSE), Normalized mean square error (NMSE) and Correlation coefficient (*r*) on testing as well as training data set for multi step ahead prediction (K=1,6,12,18,24 months ahead).

The various parameters like number of hidden layers, number of processing elements, step size, momentum value in hidden layer, in output layer the various transfer functions like tanh, sigmoid, linear-tan-h and linear sigmoid, different error norms L_1,L_2, L_3, L_4 ,L_5 and L_∞, Epochs variations and different combination of training and testing samples are exhaustively experimented for obtaining the proposed robust model for the multi step ahead prediction of the North and South hemisphere sunspots chaotic time series .

2 Static NN based model

Static NNs typically uses MLP as a backbone. They are layered feed forward networks typically trained with static back propagation. MLP solid based model has a solid foundation [26 -27]. The main reason for this is its ability to model simple as well as complex functional relationships. This has been proven through number of practical applications [28]. In [11] it is shown that all continuous functions can be approximated to any desired accuracy, in terms of the uniform norm, with a network of one hidden layer of sigmoid or (hyperbolic tangent) hidden units and a layer of linear or tan h output unit to include in the hidden layer. The paper does not explain how many units to include in the hidden layer. This is discussed in [29] and a significant result is derived approximation capabilities of two layer perception networks when the function to be approximated shows certain smoothness. The biggest advantage of using MLP NN for approximation of mapping from input to the output of the system resides in its simplicity and the fact that it is well suited for online implementation. The objective of training is then to determine a mapping from a set of training data to the set of possible weights so that the network will produce predictions y (t), which in some sense are close to the true outputs y (t). The prediction error approach is based on the introduction of measure of closeness in terms of mean square error (MSE) criteria:

$$V_N(\theta,Z^N)=\frac{1}{2N}\sum_{t=1}^{N}\left[y(t)-\hat{y}(t|\theta)\right]^T\left[y(t)-\hat{y}(t|\theta)\right] \text{ ------- (1)}$$
$$=\frac{1}{2N}\sum_{t=1}^{N}\varepsilon^2(t,\theta)$$

The weights are then found as:

$$\hat{\theta}=\arg\min_{\theta} V_N(\theta,Z^N)$$

by some kind of iterative minimization scheme:

$$\theta^{(i+1)}=\theta^{(i)}+\mu^{(i)}f^{(i)}$$

Where $\theta^{(i)}$ specifies the current iterate (number "i"), $f^{(i)}$ is the search direction and $\mu^{(i)}$ is the step size.

When NN has been trained, the next step is to evaluate it. This is done by standard method in statistics called independent validation [30]. It is never a good idea to assess the generalization properties of a NN based on training data alone. This method divides the available data sets into two sets namely training data set and testing data set. The training data set are next divided into two partitions: the first partition is used to update the weights in the network and the second partition is used to assess (or cross validate) the training performance. The testing data set are then used to assess how the network has generalized. The learning and generalization ability of the estimated NN based model is assessed on the basis of certain performance measures such as MSE, NMSE and the regression ability of the NN by visual inspection of the regression characteristics for different outputs of system under study.

3 FTLRNN Model

Time lagged recurrent networks (TLRNs) are MLPs extended with short term memory structures. Here, a "static" NN (e.g., MLP) is augmented with dynamic properties [14]. This, in turn, makes the network reactive to the temporal structure of information bearing signals. For a NN to be dynamic, it must be given memory. This memory may be classified into "short-term" and "long-term" memory. Long term memory is built into a NN through supervised learning, whereby the information content of the training data set is stored (in part or in full) in the synaptic weights of the network [31]. However, if the task at hand has a temporal dimension, some form of "short-term" memory is needed to make the network dynamic. One simple way of building short-term memory into the structure of a NN is through the use of time delays, which can be applied at the input layer of the network (focused). A short-term memory structure transforms a sequence of samples into a point in the reconstruction space [32].This memory structure is incorporated inside the learning machine. This means that instead of using a window over the input data, PEs created are dedicated to storing either the history of the input signal or the PE activations.

The input PEs of an MLP are replaced with a tap delay line, which is followed by the MLP NN. This topology is called the focused time-delay NN (TDNN). The focused topology only includes the memory kernels connected to the input layer. This way, only the past of the input is remembered. The delay line of the focused TDNN stores the past samples of the input. The combination of the tap delay line and the weights that connect the taps to the PEs of the first hidden layer are simply linear combiners followed by a static non-linearity. Typically, a gamma short-term memory mechanism is combined with nonlinear PEs in restricted topologies called focused. Basically, the first layer of the focused TDNN is a filtering layer, with as many adaptive filters as PEs in the first hidden layer. The outputs of the linear combiners are passed through a non linearity (of the hidden-layer PE) and are then further processed by the subsequent layers of the MLP for system identification, where the goal is to find the weights that produce a network output that best matches the present output of the system by combining the information of the present and a predefined number of past samples (given by the size of the tap delay line) [32]. Size of the memory layer depends on the number of past samples that are needed to describe the input characteristics in time. This number depends on the characteristics of the input and the task. This focused TDNN can still be trained with static back-propagation, provided that a desired signal is available at each time step. This is because the tap delay line at the input layer doesn't have any free parameters. So the only adaptive parameters are in the static feed forward path.

The memory PE receives in general many inputs, $x_1(n)$ and produces multiple outputs y = $[y_0(n), \ldots, y_D(n)]T$, which are delayed versions of $y_0(n)$ the combined input,

$$y_k(n) = g(y_{k-1}(n)) \quad y_0(n) = \sum_{j=1}^{P} x_j(n) \qquad \text{---(2)}$$

where, g(.) is a delay function.

These short-term memory structures can be studied by linear adaptive filter theory if g(.) is a linear operator. It is important to emphasize that the memory PE is a short-term memory mechanism, to make clear the distinction from the network weights, which represent the long-term memory of the network.

There are basically two types of memory mechanisms memory by delay and memory by feedback. We seek to find the most general linear delay operator (special case of the Auto Regressive Moving Average model) where the memory traces $y_k(n)$ would be recursively computed from the previous memory trace yk-1(n). This memory PE is the generalized feed forward memory PE. It can be shown that the defining relationship for the generalized feed forward memory PE is mentioned

$$g_k(n) = g(n) * g_{k-1}(n) \quad k \geq 1 \qquad \text{-(3)}$$

Where, * is the convolution operation, g(n) is a causal time function, and k is the tap index. Since this is a recursive equation, $g_0(n)$ should be assigned a value independently. This relationship means that the next memory trace is constructed from the previous memory trace by convolution with the same function g(n), the memory kernel yet unspecified. Different choices of g(n) will provide different choices for the projection space axes. When we apply the input x(n) to the generalized feed forward memory PE, the tap signals $y_k(n)$ become

$$y_k(n) = g(n) * y_{k-1}(n) \quad k \geq 1 \qquad \text{------(4)}$$

the convolution of yk – 1(n) with the memory kernel. For k=0 we have

$$y_0(n) = g_0(n) * x(n) \qquad \text{------(5)}$$

where, $g_0(n)$ may be specified separately. The projection x(n) of the input signal is obtained by linearly weighting the tap signals according to

$$x(n) = \sum_{k=0}^{D} w_k y_k(n) \qquad \text{-------(6)}$$

The most obvious choice for the basis is to use the past samples of the input signal x(n) directly, that is the kth tap signal becomes $y_k(n) = x(n-k)$. This choice corresponds to

$$g(n) = \delta(n-1) \qquad \text{------(7)}$$

In this case $g_0(n)$ is also a delta function δ(n) (delta function operator used in the tap delay line). The memory depth is strictly controlled by D, that is the memory traces store the past D samples of the input. The time delay NN uses exactly this choice of basis.

The gamma memory PE attenuates the signals at each tap because it is a cascade of leaky integrators with the same time constant gamma modal. The gamma memory PE is a special case of the generalized feed forward memory PE where,

$$g(n) = \mu(1-\mu)^n \quad n \geq 1 \qquad \text{---------(8)}$$

and $g_0(n) = \delta(n)$. The gamma memory is basically a cascade of low pass filters with the same time constant 1 - μ. The over all impulse response of the gamma memory is

$$g_p(n) = \binom{n-1}{p-1} \mu^P (1-\mu)^{n-p} \text{ , n} \geq \text{p} \quad \text{-----(9)}$$

Where, (:) is a binomial coefficient defined by

$$\binom{n}{p} = \frac{n(n-1)....(n-p+1)}{p!}$$

For integer values of n and p, the overall impulse response $g_p(n)$ for varying p represents a discrete version of the integrand of the gamma function, hence the name of the memory.

The gamma memory PE has a multiple pole that can be adaptively moved along the real Z-domain axis, that is the gamma memory can implement only low pass (0 < μ < 1) or high pass (1 < μ < 2) transfer functions. The high pass transfer function creates an extra ability to model fast-moving signals by alternating the signs of the samples in the gamma PE (the impulse response for 1 < μ < 2 has alternating signs). The depth in samples parameters (D) is used to compute the number of taps (T) contained within the memory structure(s) of the network.

4 Performance Measures

Three different types of statistical performance evaluation criteria were employed to evaluate the performance of these models developed in this study. These are as follows.

MSE : The mean square error is given by -

$$\text{MSE} = \frac{\sum_{j=0}^{p} \sum_{i=0}^{N} (d_{ij} - y_{ij})^2}{\text{N*P}} \quad \text{----- (10)}$$

Where P = number of output PEs (processing elements), N=number of exemplars in the data set, y_{ij} =network output for exemplar i at PEj, d_{ij}=desired output for exemplar i at PEj.

NMSE (Normalized Mean square Error)

The normalized mean square error is defined by the following formula :

Where P = Number of output PEs,

N = Number of exemplars in data set,

MSE = Mean square error,

d_{ij}= desired output for exemplar i at pej

$$NMSE = \frac{P * N * MSE}{\sum_{j=0}^{P} \left[\frac{\text{N} \sum_{i=0}^{N} dij^2 - \sum_{i=0}^{2} dij}{N} \right]} \quad \text{--------(11)}$$

Correlation Coefficient *(r)* : The mean square error (MSE) can be used to determine how well the network output fits the desired output, but it does not necessarily reflect whether the two sets of data move in the same direction. For instance by simply scaling the network output, we can change the MSE without changing the directionality of the data. The correlation coefficient solves this problem. By definition, the correlation coefficient between a network output x and a desired output d is.

$$r = \frac{\frac{\sum_{i} (xi - \bar{x})(di - \bar{d})}{N}}{\sqrt{\frac{\sum_i (d - \bar{d})}{N}} \sqrt{\frac{\sum_i (xi - x)}{N}}} \quad \text{------(12)}$$

where,

$$\bar{X} = \frac{1}{N} \sum^{N} Xi \quad \text{and} \quad \bar{d} = \frac{1}{N}$$

The correlation coefficient is confined to the range [-1, 1].

6 Sun Spot time series

A sun spot number is a good measure of solar activity which has a period of 11 years, so called solar cycle . The solar activity has a measure effect on earth, climate, space weather , satellites and space missions, thus is an important value to be predicted. But due to intrinsic complexity of time behavior and the lack of a quantitative theoretical model, the prediction of solar cycle is very difficult. Many prediction techniques have been examined on the yearly sunspots number time series as an indicator of solar activity . However, in more recent studies the international monthly sunspot time series, which has a better time resolution and accuracy , has been used . In particular , a nonlinear dynamics approach has been developed in [33] and prediction results are compared between several prediction techniques from both statistical and physical classes. There has been a lot of work on controversial issue of nonlinear characteristics of the solar activity [33-37]; and a several recent analysis have provided evidence for low dimensional deterministic nonlinear -chaotic behavior of the monthly smoothed sun spot time series [33,34,35] and has intense .The data considered the monthly variations from January 1749 to December 2006. The total samples are 3096 considered. The series is normalized in the range of -1 to +1. The monthly smoothed sunspot number time series is downloaded from the SIDC (World data center for the sun spot Index) [37].

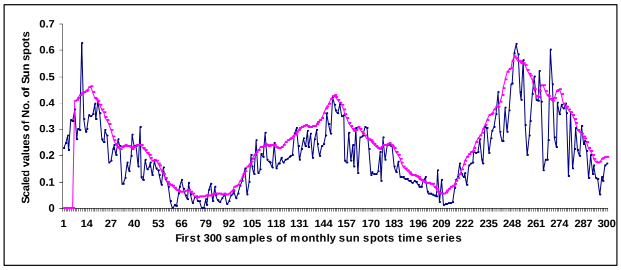

Fig.1 First three hundred samples of monthly sun spot time series .

7 Experimental Results

The choice of the number of hidden layers and the number of hidden units in each hidden layers is critical [38]. It has been established that a MLPNN that has only one hidden layer, with sufficient number of neurons, acts as a universal approximators of nonlinear mappings[39]. The tradeoff between accuracy and complexity of the model should be resolved accurately [40-41]. In practice, it is very difficult to determine a sufficient number of neurons necessary to achieve the desired degree of approximation accuracy. Frequently the number of units in the hidden layer is determined by trial and error .To determine the weight values, one must have a set of examples of how the output should relate to the inputs .The task of determining the weights from these examples is called training or learning, and is basically a conventional estimation problem. That is, the weights are estimated from the examples in such away that the network, according to metric, models the true relationship as accurately as possible. Since learning is a stochastic process, the learning curve may be drastically different from run to run .In order to compare the performance of a particular search methodology or effect of different parameters have on a system, it is needed to obtain the average learning curve over the number of runs so that the randomness can be averaged out. An exhaustive and careful experimentations has been carried to determine the configuration of the static MLP Model and the optimal proposed FTLRNN model for all the step (K=1, 6, 12, 18, 24) months ahead prediction. It is seen that the performance of this model is optimal on the test dataset for the following Wples = 10, No. of taps = 6, Tap Delay = 1. Trajectory Length = 50. All the possible variations for the model such as number of hidden layers, number of processing elements in each hidden layer, different transfer functions like tan h, linear tanh, sigmoid, linear sigmoid in output layer, different supervised learning rules like momentum ,step, conjugant gradient and quick propagation are investigated in simulation. The step size and momentum are gradually varied from 0.1 to 1 for static back propagation rule. After meticulous examination of the performance measures like MSE, NMSE, Correlation Coefficient (*r*), the optimum parameters are found and mentioned in the table 1 for 60% used as training samples, 25 % as testing samples and 15% cross validation samples.

Table 1: Parameters for the Neural network Models

Sr. no.	Parameters	Hidden Layer	Output Layer
1	Processing elements	15	1
2	Transfer function	tanh	Tanh
3	Learning rule	Momentum	Momentum
4	Step Size	1	0.1
5	Momentum	0.8	0.8

It is found that the performance of the selected model is optimal for 15 neurons in the hidden layer with regards to the MSE, NMSE, and the correlation coefficient (*r*) for the testing data sets. When we attempted to increase the number of hidden layer and the number of processing element in the hidden layer, the performance of the model is not to seen to improve significantly .On the contrary it takes too long time for training because of complexity of the model. As there is single input and single output for the given system, the number of input and output processing elements is chosen as one. Now the NN Model (1:15:1) is trained three times with different weight initialization with 1000 iterations of the static back propagation algorithm with momentum term for all the three models for all the 1, 6, 12, 18 and 24 months ahead predictions as shown in table 2.

Table 2 Performance of Neural Network Models for testing data set (for northern hemisphere sun pots)

K (months)	MLP Neural Network		FTLRNN		SOFM				
	MSE	NMSE	*r*	MSE	NMSE	*r*	MSE	NMSE	*r*
1	0.00252	0.03296	0.9765	0.00207	0.03124	0.98928	0.00211	0.3187	0.9812
6	0.00886	0.12354	0.93954	0.00437	0.06535	0.97497	0.00953	0.14241	0.93298
12	0.02661	0.39485	0.79227	0.01415	0.21006	0.91345	0.02847	0.42241	0.77782
18	0.04681	0.69153	0.4681	0.02173	0.32103	0.86821	0.07968	0.70850	0.58202
24	0.6454	0.95030	0.36815	0.02692	0.39647	0.83186	0.06715	0.98877	0.33654

Table 3 Performance of Neural Network Models for testing data set (for South hemisphere sun spots)

K (Months)	MLP Neural Network			FTLRNN			SOFM		
	MSE	NMSE	*r*	MSE	NMSE	*r*	MSE	NMSE	*r*
1	0.00410	0.08449	0.95903	0.00391	0.08301	0.95984	0.00432	0.09169	0.95361
6	0.00886	0.07109	0.90347	0.00548	0.11554	0.94075	0.00921	0.19407	0.90125
12	0.02149	0.45009	0.75559	0.01191	0.24945	0.87238	0.02201	0.46080	0.75047
18	0.03397	0.70810	0.56893	0.01711	0.35685	0.82576	0.03613	0.75325	0.54383
24	0.04547	0.94668	0.35761	0.02233	0.46493	0.76713	0.04640	0.96608	0.34182

From the table 2 and table 3 it is observed that FTLRNN model is able to predict the monthly southern and northern sunspots chaotic time series elegantly well as compared to multilayer perceptron (MLP) and self organizing feature map (SOFM) on testing data set with regards to MSE,NMSE and correlation coefficient (*r*). Also the graphs are plotted for the proposed FTLRNN model for 1, 6,12, 18 and 24 months ahead prediction for south and north direction monthly sunspots time series as shown in figure 11 to

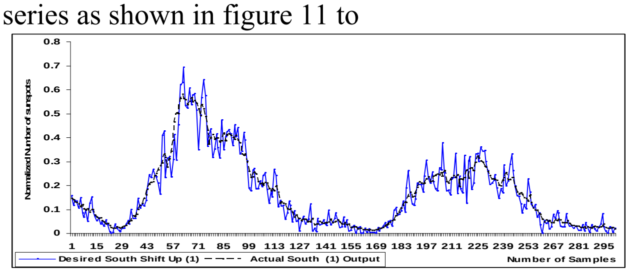

Fig 11 Desired output Vs Actual output for 1 month ahead for South direction Sun Spots for FTLRNN.

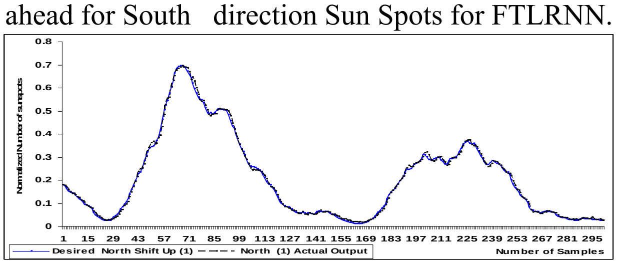

Fig 12 Desired output Vs Actual output for 1 month ahead for north direction Sun Spots for FTLRNN.

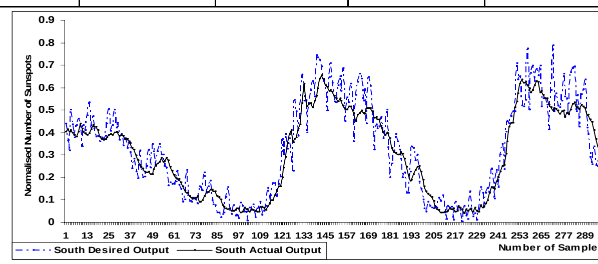

Fig 13Desired output Vs Actual output for 6 month ahead for South direction Sun Spots for FTLRNN

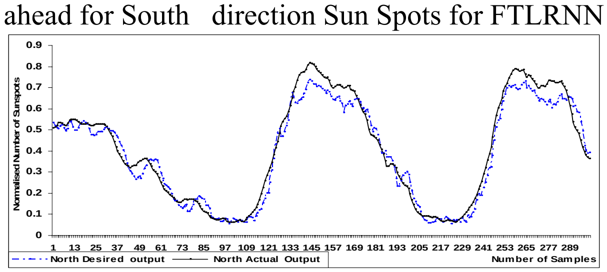

Fig 14 Desired output Vs Actual output for 6 month ahead for North direction Sun Spots for FTLRNN

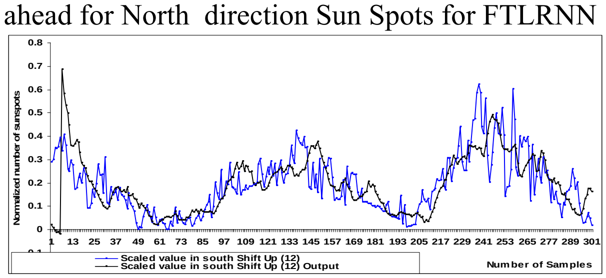

Fig 15 Desired output Vs Actual output for 12 months ahead for south direction Sun Spots for FTLRNN model.

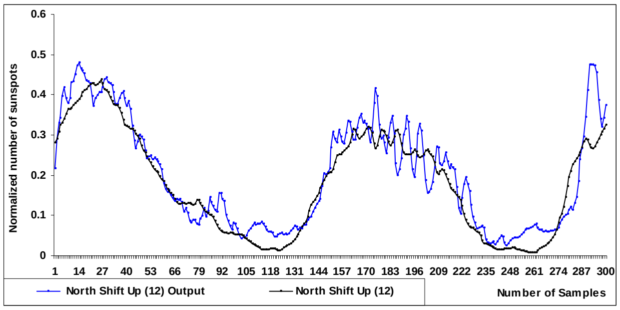

Fig.16 Desired output Vs Actual output for 12 months ahead for north direction Sun Spots for FTLRNN model.

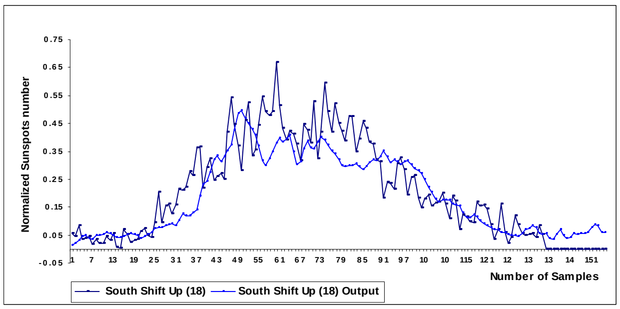

Fig.17 Desired output Vs Actual output for 18 months ahead for South direction Sun Spots for FTLRNN model

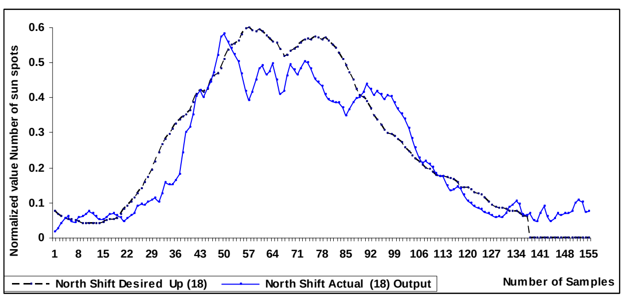

Fig.18 Desired output Vs Actual output for 18 months ahead for north direction Sun Spots for FTLRNN model

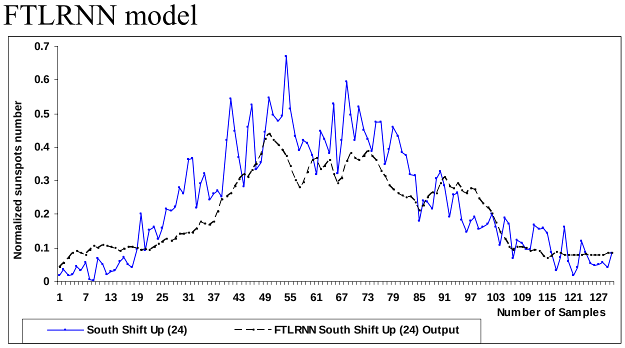

Fig.19 Desired output Vs Actual output for 24 months ahead for South direction Sun Spots for FTLRNN model

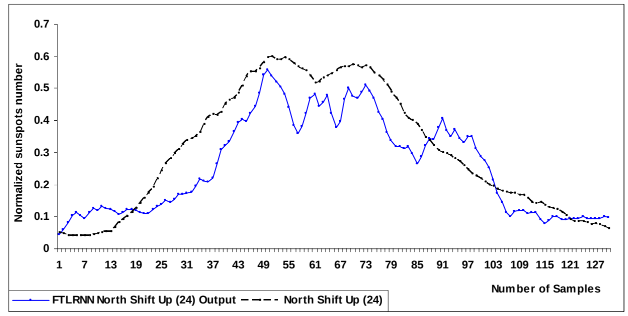

Fig.20 Desired output Vs Actual output for 24 months ahead for North direction Sun Spots for FTLRNN model.

8 Conclusion

It is seen that focused time lagged recurrent neural network model with gamma memory is able to predict the southern and northern sunspots chaotic time series quite well in comparison with the Multilayer preceptron (MLP) and self organizing feature map (SOFM) . Static NN configuration such as MLP NN based model and self organizing feature map (SOFM) network are failed to cope up with the underlying nonlinear dynamics of the sunspots chaotic time series. It is seen that MSE, NMSE of the proposed focused time lagged recurrent neural network (FTLRNN) dynamic model for testing data set as well as for training data set are significant better than those of static MLP NN and SOFM model. For the 12 ,18 and 24 months ahead prediction the value of MSE and NMSE for the proposed FTLRNN model is significantly improved .Also for the proposed FTLRNN model the output closely follows the desired output for all the months ahead prediction for northern and southern direction as shown in figure 12 to 21 as compared to the MLP and SOFM . It can be closely visually inspected from the figures 12,14,16,18 and 20 of desired output to actual output plot that for all the months ahead prediction for northern direction sunspots the output of FTLRNN model is closely follows the desired output than the Southern direction sunspots as in the figures 11,13,15,17,19 and the values of MSE, NMSE and correlation coefficient *r* are also better. In addition it is also observed that the correlation coefficient of this model for testing and training exemplars are much higher than MLP and self organizing feature map (SOFM) neural network. It is resulted from the experiments that the FTLRNN model learns the dynamics of monthly sunspot chaotic time series quite well as compared to Multilayer perceptron and self organizing feature map. On the contrary, it is observed that static MLP NN and self organizing feature map (SOFM) performs poorly bad, because on the one hand it yields much higher MSE and NMSE on testing data sets and on the other hand the correlation coefficient r for testing data set is far less than unity. Hence the focused time lagged recurrent neural network with gamma memory filter has out performed the static MLP based neural network and SOFM better for all the months ahead predictions for monthly north and south hemisphere chaotic time series .

9 References

[1] B. Townshend , 1992 : Nonlinear prediction of speech signals , in Nonlinear modeling and Forecasting , M. Casdagli and S. Eubank , Eds . Reading MA: Addison –Wesley pp 433-453.

[2] P.G. Cooper, M.N. Hays, and J.E. Whalen, “Neural networks for propagation modeling,”Electromagnetic Environmental Test facility, Fort Huachuca, AZ, Atlantic Res Corp, Rep .92-12-002, 1992

[3] Haykin, S., Neural Networks, 2006 –A comprehensive foundation, Second Edition, Pearson Education India, pp .443.

[4] Weigend A.S. and Greshenfeld N.A, 1993. Time series prediction: Forecasting the future and Understanding the past, Santa Fe Institute Studies in Science of complexity, Proceedings vol . XV, Addison –Wesley, Reading, MA.

[5] M.Casdagli,1989 .Nonlinear prediction of chaotic time series, Phys. D, vol 35, pp 335-356.

[6] Leung H. and Lo T., July 1993. Chaotic radar signal processing over the sea .IEEE J, Oceanic Eng., Vol.18 issue 3 pp 287-295.,ISSN no 0364-9059

[7] H. Leung , 1995. Applying Chaos to radar detection in an ocean environment An Experimental study, IEEE Journal of Oceanic Eng., vol. 20 pp56-64.

[8] K.M. Short, 1994 . Steps toward unmasking secure communications International Journal of Bifurcation Chaos, Vol 4, pp 959-977.

[9] G. Heidari- Bateni and C.D. McGillen, 1994. A chaotic direct sequence spread spectrum communication system , IEEE transactions communications, Vol 42. pp 1524-1527.

[10] Y. Fu and H. Leung, 2001. Narrowband interference cancellation in spread spectrum communication systems using chaos ., IEEE transactions circuits systems ,vol. 48 ,pp847-858.

[11] Cybenko. G., 1989. Approximations by superposition of a sigmoidal functions, Mathematics of Control, Signals and Systems, Vol 2 , no 4pp 303-314.

[12] K Hornik, M Stinchcombe & H White, Multilayer feed forward networks are universal approximators, Neural networks, vol 2,no 5, pp 359-366, 1989.

[13] Dudul Sanjay V., 2005. Prediction of a Lorenz chaotic attractor using two layer perceptron network , Elsevier Journal of Applied soft computing , 5(4) 4333-4355.

[14] Sanjay V Dudul , February2007. Identification of a Liquid of a Liquid Saturated Steam Heat Exchanger Using Focused Time Lagged Recurrent Neural Network Model, IETE Journal of Research Vol53 No1,pp69-82.

[15] Barron A. R., May 1993. Universal approximation bounds for superpositions of a sigmoidal function, IEEE Transactions on Information Theory, vol.39, issue 3, pp 930–945.,

[16] Judi sky, A., Hjalmarson, H., Benefits, A., Delyon, B., Ljung, L., Sjoberg, J., Zang, Q., 1995. “Nonlinear black-box models in system identification: Mathematical foundations”. Automatica, 31(12), pp. 1725-1750.

[17] R. and Cor M. Vanden Bleek 2000. attractors by neural network, Neural Computations., vol.12, pp 2355-2383 .

[18] Zeng Zhaocoui, Duan Yurong, 1999. Chaotic time series analysis based on radial basis function network, Chinease Journal of Chongqing University 22(6) 113-120.

[19] H. Leung, T. Lo and S.Wang, September 2001. Prediction of noisy chaotic time series using an optimal radial basis function neural network. IEEE Transactions on neural networks vol 12, no 5, pp1163-1172.

[20] Deco, G and Schurmann M,1994. Neural learning of chaotic system behavior, IEICE Transactions fundamentals E77(A),1840-1845.

[21] Billings Stephen A. and Hua- Liang Wei, July 2005. A new class of wavelet network for nonlinear system identification, IEEE

transactions on neural networks , Vol 16, No. 4, pp 862-874.

[22] Min Han , Jianhui Xi, Shiguo Xu and Fu-Liang Yin, December 2004. Prediction of chaotic time series based on the recurrent predictor neural network, IEEE transaction of signal processing , vol 52, no 12.

[23] Ali Gholipour, Babak N. Araobi and Caro Lucas, 2006. Predicting Chaotic time series using Neuraland neuro fuzzy models:A comparative study. Neural Processing Letters Learning and locally linear (2006) 24 :217-239, Springer.

[24] Hirotoka Inoue, Yashinobou Fukuango and Hiroyuki Narihinsa " Efficient Hybrid neural network for chaotic time series prediction" ICANN2001, LNCS 2130, pp 712-718, Springer.

[25] Min Han, Mingming Fan , Jianhui Xi Study of nonlinear multivariate time series prediction based on neural networks LNCS 3497 ,PP618-623, 2005.

[26] Quin S.Z., Su H T & McAvoy T J, 1992.Comparison of Four Neural Net Learning Methods for Dynamic System Identification, IEEE transactions on Neural networks vol 3, issue 1 pp 122-30. ,ISSN no. 1045-9227.

[27] Naraendra K S & Parthasarathy K, 1990. Identification and control of dynamic systems using neural networks, IEEE Trans on Neural Networks , vol 1 , no 1 , pp 4-27.

[28] Demuth & Beale M., Neural Network Tool Box for use with MATLAB ,2004. Users Guide , Version 4.0, The Math works, Inc Natick, MA.,www.mathworks.com.

[29] Franklin , G.F. Powell, J .D . And Workman M.L , 1998 .Digital control of dynamics systems , Addison –Weseley , Reading , MA ,3rd Edition.

[30] Fredric M Ham & I. Kostanic, 2002 Principles of Neuro computing for science and Engineering., Tata McGraw –Hill Edition.

[31] Vries, Bert De .& Principe Jose .C., 1992 The Gamma Model,. –A neural model for temporal processing. Neural networks, Vol. 5 pp 565-576., principe @brain .ee. ufl .edu.

[32] Principe J.C., Euliano N.R, Inc 2000 Neural and Adaptive Systems – Fundamental through Simulations, John Wiley and Sons.

[33] Sello S. , 2001. Solar cycle forecasting: a nonlinear dynamics approach , Astronomy and Astrophysics 377, pp 312—320.

[34] Lawrence ,J.K., Cadavid A.C. and Ruzumaikin, A.A., 1995 : Turbulant and chaotic dynamics underlying solar magnetic variability, Astrophysical Journal 455.

[35] Zhang Q, 1996 . A nonlinear prediction of the smoothed monthly sun spots numbers, Astronomy and Astrophysics 310 ,646.

[36] Schreiber, T, 1998 : Interdisciplinary application of nonlinear time series methods , Physical Reports 308

[37] http:// sidc .oma .be / index .php3.

[38] Huebner, U., Abraham, N. and Weiss C, 1989: Dimensions and Entropies of chaotic intensity pulsations in a single –mode far infrared NH3 Laser, Physical Review A (40) a, 6354-6360.

[39] I. V. Turchenko, June 2004,Simulation modeling of multi –parameter sensor signal identification using neural networks, Proceedings of the 2nd International conference on Intelligent systems , Sofia, Bulgaria , pp 48-53.

[40] Huang G., Yan Qiu Chen & Babri H.A.,2000 , Classification ability of single hidden layer feed forward neural networks, IEEE Transactions on neural networks , vol 11 no, pp 799-801.

[41] K W Lee & H.N. Lam ,1995. Optimal sizing of feed forward neural networks : Casestudies , proceedings of the 2nd New Zealand two stream International conference on Artificial neural networks and Expert systems pp 71-79.