Review Paper:

Predictability of El Nino Southern Oscillation using Artificial Neural Networks

Sarkar Partha Pratim*, Janardhan Prashanth and Roy Parthajit
Department of Civil Engineering, National Institute of Technology, Silchar, Assam 788010, INDIA
*Parthasarkar.nits@gmail.com

Abstract

El Nino- Southern Oscillation (ENSO) is an interactive phenomenon between the ocean and atmosphere across the tropical Pacific region which results in changes in sea surface temperatures (SSTs). It occurs every 2 to 8 years and has adverse effects on the climate across the globe. ENSO has large-scale social and economic impacts. It directly affects the infrastructure, health, agriculture, and trade and energy sectors. Hence, prediction of ENSO conditions is crucial to both the researchers and the general populaces. Therefore, a lot of research work has been done for its precise prediction. Artificial Neural Networks (ANNs) that are functionally based on the human brain have been popular and efficient in climate change studies recently.

As ENSO has several nonlinear features and is a complex natural phenomenon with various salient features and factors contributing to it. Researchers have started using ANNs as an alternative to other traditional methods to model this event. In this study, the physics of ENSO and the substantial work done by researchers to predict the SST anomalies (one of the major factors in this event) using different artificial neural network techniques is studied in detail and the key findings are highlighted which further emphasizes the importance of artificial intelligence in the domain of climate change.

Keywords: Artificial neural networks, El Nino Southern oscillation, Sea surface temperatures, Modeling, Nonlinear models, Prediction.

Introduction

El Nino Southern Oscillation (ENSO) occurs naturally and has devastating and appalling impacts socially, economically as well as environmentally. These climatic disturbances occurred in every 2 to 8 years. El Nino and La Nina are the 2 phases of ENSO. The former occurs when there is warming of ocean surface and the sea surface temperatures (SSTs) becomes higher than normal while the latter is just the reverse of El Nino phase which results when SSTs fall down from the average SSTs. The southern oscillation is a consistent periodic variation of atmospheric pressure across the equatorial Pacific Ocean and is an atmospheric component of ENSO. El Nino and the SST variations from normal surface temperatures are very closely linked.

ENSO is certainly a predictable event but it is the degree of its precise prediction, which is of utmost importance to us. A lot of research work has been done in this field for its successful prediction. The interest in ENSO grew up a notch after a very high strong and heavily observed El Nino event was reported in 1982-83. It had cataclysmic effects in many South American countries, parts of Asia and Australia.[10] The 1982-83 El Nino certainly sparked an outburst in research interest.

It was during the mid-1980s that prediction and seasonal experimental forecasting of ENSO gathered momentum. The prediction models could be classified into 3 types namely Statistical models, Coupled Physical models and Hybrid models.[24,25] Barnston et al[5] observed that models based on statistical analysis produced similar capabilities as compared to other models in predicting SST aberrations across the equatorial Pacific region. Most of these kind of models predicted results for 6 months in future and were mostly linear regression models which depended on operations (matrix) to improve the correlation of observed and forecasted data.[4,17,21] On the other hand, ENSO certainly has features, which are nonlinear.

During the El Nino and La Nina phases of ENSO, the SST and sea surface wind aberration patterns clearly do not match with one another[32,65] and this observed non-linearity has been elucidated by An and Jin[1]. To understand and model these nonlinear irregularities of ENSO, artificial neural network (ANN) models which mirror the working of a human brain (biological neural network) and primitively from the domain of computational intelligence, have been used. These models were developed to perform nonlinear regression[53,65] and nonlinear CC[11,19].

This study addresses the work done by researchers in the prediction of SST anomalies using artificial neural network models (Nonlinear models). This study also explores the recent developments that have taken place in the non-linear domain and deliberates future scopes and improvements.

ENSO: The physics and indices involved

ENSO- The processes concerned: The abnormal rise of the SSTs, which leads to the ENSO phenomenon, lasts from a few months up to a year or sometimes even more. It is represented by the following events:

1) The western tropical Pacific region and Indonesia is covered with severe high pressure and the eastern tropical Pacific region is covered by low air pressure which is known as the negative phase of the Southern Oscillation[7,23].
2) The Easterly winds known as trade winds get weakened across the equatorial Pacific, resulting in build-up of warm waters on the eastern part of the equatorial Pacific[66]. This is also called as the disruption of Walker Circulation.
3) The equatorial regions to the east of 160ºE receive heavy precipitation[16,44].
4) The Hadley circulation is enhanced in the Pacific region[46].
5) The Aleutian low is dislodged southwards during the Northern Hemisphere winter season[7,8,61].

The oceanographic studies[20,28,42,66] discovered that during a normal condition, low pressure exists over Darwin and a high pressure exists over Tahiti, which results in the air circulation from east to west. These winds are the trade winds which bring warmer surface waters towards the western side and because of which Australia and western Pacific get precipitation.

However, when this pressure weakens, the trade winds are weakened and can no longer force the warmer surface waters towards the west. This causes severe drought conditions in Australia, while across the eastern part due to heavy precipitation, floods occur at the western beachfront of equatorial South America. It can be stated that the SSTs and atmospheric pressure highly relate to each other. The intensification and decline of Easterly trade winds are function of changes in the pressure tilt across the tropical Pacific.

The ENSO observations emphasizes on Sea surface temperature anomalies (SSTA) in four different geographical locations of the equatorial Pacific.

The ENSO indices: The monitoring of the tropical Pacific SST aberrations is done with the help of few indices averaged over a given region and are described below. The inconsistencies are generally estimated with respect to an average of thirty years.

Niño 1+2 (0º-10º S, 90º W-80º W): It is the tiniest and the most eastern of the Niño SST regions. It is located at the beachfront South America. At this location, the nearby populaces first perceived El Niño. It shows the biggest deviation of all the other such indices.

Niño 3 (5 º N-5 º S, 150 º W-90 º W): At one time, this location was considered a key location for observing and forecasting ENSO events, however later researchers discovered that the prime regions for coupled sea air communications for ENSO are located more towards the west[57].

Niño 3.4 (5 º N-5 º S, 170 º W-120 º W): It generally employs a 5-month running average. The two phases are identified when the SSTs in the region Niño 3.4 exceed by 0.4 ^{0}C for an El Nino event and are reduced by 0.4 ^{0}C for a La Nina event for a period of 6 months or more.

Niño 4 (5 º N-5 º S, 160 º E-150 º W): This region records SST irregularities in the central Pacific. This area shows lesser change as compared to other mentioned Niño region.

ONI (Oceanic Niño Index): It is also a key indicator to monitor these events. The ONI employs the data from Niño 3.4 region. The ONI usually employs a 3-month running average and if the anomalies are to exceed by +0.5 ^{0}C or -0.5 ^{0}C for a period of 5 continuous months, then such an event can be termed as certain El Nino or La Nina.

To describe the remarkable feature of every ENSO (both phases included) event, Trenberth et al[58] presented the Trans-Niño Index (TNI). They proposed that to minutely understand features of every El Niño or La Niña, SST irregularities between the Niño 1+2 and Niño 4 regions have to studied along with Niño 3.4 region.

In this way, TNI measures the rise in SST irregularities between the central and eastern tropical Pacific. At the point when the SST angle plays a major role (Niño 4 region is characterized by positive anomalies and Niño 1+2 area by the negative anomalies), researchers have coined this occasion as a "central Pacific El Niño " or "El Niño Modoki."

The difference between them is the fact that in El Niño Modoki, there is strong anomalous warming of the central tropical region in the Pacific Ocean. The cooling takes place at both eastern and western tropical regions of the Pacific Ocean. These regional gradients give result to 2 walker circulation cells as compared to one in case of El Niño.

The El Niño Modoki is also characterized by a wet zone across the central Pacific. The locations of different Nino indices (regions) across the equatorial pacific sector are shown in fig. 1.

ENSO Prediction: The Neural network models

Artificial neural networks (ANNs) are the foundations of AI and are capable of solving complex tasks which are very difficult by human standards or traditional statistical methods. ANNs are based on the working of the human brain and the neurons and which are the functional units of the brain and serve as the basic units of an artificial neural network.

These neurons function parallel to one other and are organized in layers. The first layer that serves as the input layer is fed the raw information, which is then processed by the subsequent layers and the final layers, produces the output. Interestingly, a neural network can adjust itself to the changes in the surroundings and therefore results in better learning of a system.

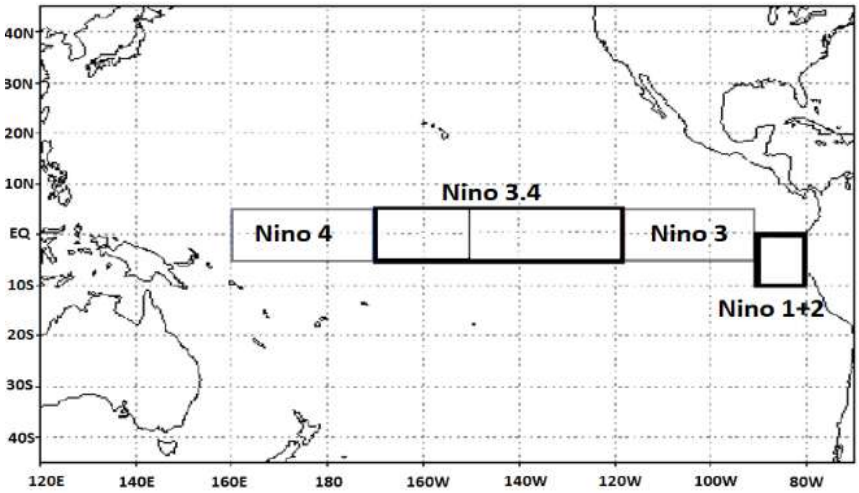

Fig. 1: Locations of different Nino regions across the Pacific

Mathematically the working a neuron can be explained with the help of the equations given below:

$$u_m = \sum_{i=1}^{n} w_{mi}\, x_i \quad (1)$$

and

$$y_m = \varphi(u_m + b_m) \quad (2)$$

where $x_1, x_2, \ldots, x_i$ are the initial input signals; $w_{m1}, w_{m2}, \ldots, w_{mi}$ are the corresponding weights of neuron M; u_m is the linear combiner of the weights with their input signals; b_m is the bias of the network; $\varphi(.)$ is known as the activation function and y_m is the final output of the neuron, M. The bias is a constant that adjusts the final output u_m so that the overall model fits the given data.

The function of the activation function is to analyze and decide when a neuron is to be initialized by adjusting the synaptic strengths corresponding to that particular the neuron and by adding the bias to it. The non-linearity is brought in the neuron with the introduction of activation function. The NNs can be broadly classified into three fundamental network architectures:

1. Single-Layer Feedforward networks (Single-Layer Perceptron): It is the simplest type of NN architecture. The input layer of signal nodes projects the output layer of computation nodes. It is non-cyclic in nature.

2. Multilayer Feedforward networks (Multi-Layer Perceptron): There is a presence of one or more layers of hidden neurons in this type of NN architectures. The hidden layers compute the weighted inputs and then convert them by employing suitable activation function to finally obtain the desired output. The number of hidden layers in the NN model largely depends on the complexity of the problem. The input layer is the first layer fed with input vectors which creates the input signals which are then applied to the second layer neurons which becomes the first hidden layer. The second layer output is used as an input for the third layer, and the process goes on in a similar way for the remainder of the network.

3. Recurrent networks: These NN architectures are similar to feed forward networks, the contrasting feature being the presence of at least one feedback loop. Generally, there are two types of feedback, feedback generated from hidden neurons as well as from the output neurons and is moved backwards and self-feedback where the output of the neuron is brought back into its own input.

The domains that employ utilize artificial neural networks (ANNs) are advanced and are spread continuously in the recent times as computer technology and artificial intelligence have developed tremendously. ANN as such has been widely employed and put in work by researchers taking in account of growing concerns of frequent El Nino Southern oscillation events and with it the related climatic abnormalities. The work of Tangang et al[53] can be considered pivotal in this field. The SSTA in the tropical Pacific was predicted by using NN models in the Nino 3.4 region. Empirical orthogonal function (EOF) analysis was used in this study.

Wind stress data transformed using EOFs across the tropical Pacific (20°S-20°N, 120°E-70°W) for 4 seasons and the Nino 3.4 SSTA data was used as the inputs. For training the NN models, the data from the period of 1952-1981 was used and for forecast validation, data from the period 1982-1992 was employed. A simple three-layered neural network model (feed forward) as the one shown in the fig. 2 was used.

In the hidden layer, hyperbolic tangent activation function was employed and a linear transformation of the output was done. Extended empirical orthogonal functions (EEOF) were used to compress. The predictor field was compressed by extended empirical orthogonal functions (EEOF) which resulted in smaller NNs producing identical or comparatively better results than the larger NN models; the output of the network is given by:

$$z = \sum_j \widetilde{w}_j \; y_j + b,$$

where y_j is the outcome of the j^{th} hidden neuron.

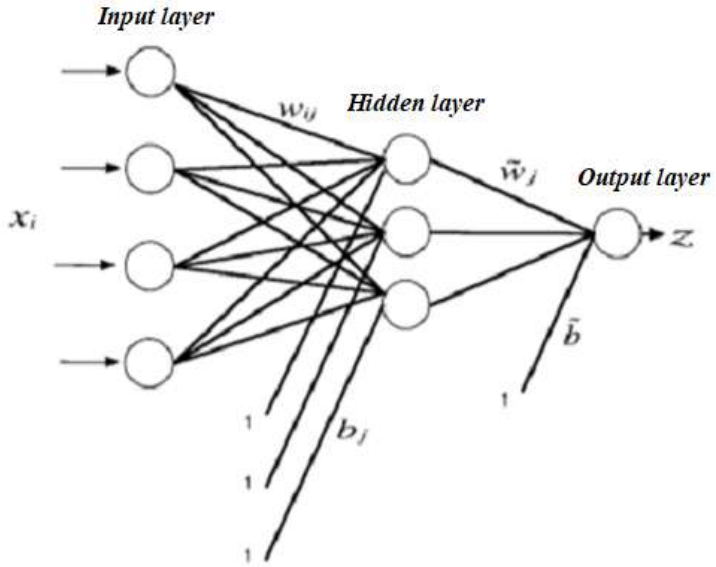

Fig. 2: A 3-layer feed-forward NN model having an input layer, a hidden layer and the output layer

$$y_j = \tan h \left(\sum_i w_{ij} x_i + b_j \right)$$

whereas the terms w, x and b bear the same descriptions as described earlier. The optimization or the network training was done using the back-propagation method[48,50]. The steepest descent algorithm was used in this study to minimize the cost function[22] which was of the form as shown:[12]

$$C = \frac{a_1}{2} \sum (z - z_a)^2 + a_2 \sum w^2 + a_3 \sum \tilde{w}^2$$

where C denotes the sum of squared error and the two terms on the left hand side are weight and weight adjusting terms respectively. This NN model attained forecasting capabilities similar to the other ENSO models of that time at 6–month lead-time. The obtained individual forecasts were then compared with the ensemble forecasts, which were obtained by calculating the mean of six separate forecasts. For example, for the January-February- March (JFM) period of 1982 at 3-month time ahead, the ensemble forecasts were obtained through averaging of the particular forecasts of the same period at 3-month time ahead.

A similar method was used to create ensemble forecasts for those at 4, 5, 6, 7, and 8-month leads. This averaging method increased the skills of mid-range periods like from 5-10 months but it did prove to be a disadvantage for lesser or very shorter duration periods or periods larger than 1 year. The correlation of 0.66 and the root-mean-square error (rmse) of 0.83 for the individual forecasts was achieved at 6-month lead. A better correlation of 0.71 and an rmse of 0.73 were achieved by the ensemble forecasts. The results from the study clearly in fig. 3 show the effectiveness of NN models in forecasting of ENSO even at longer time durations (more than 6 months to a year). It was also found that the underlying relationship between Nino 3.4 SSTA and the wind stress became progressively nonlinear at these longer leads. The interpretation of NN outcomes was done in presence of certain theories e.g. the part played by the Rossby waves in sparking the commencement of an ENSO event and the part played by delayed-oscillator theory in the growth and ceasing of an ENSO event.

Tangang et al[53,54] did a sequential work. They created two types of NN models for forecasting the SST anomalies over standard equatorial Pacific regions namely Nino 3, 3.4, 3.5, 4, P2, P4 and P5. The new region Nino 3.5 (120°W to 180° and 10°S to 5°N) was proposed by Trenberth and Hoar[56]. This region includes parts of the Nino 3 and Nino 4 regions and extended farther into the southern hemisphere.

The first model used SLP field as the predictors and the other used wind stress field. The neural network model design is the same as given by Tangang et al.[53] To reduce noisy data from those of the earlier models, ensemble averaging was done over 20 forecasts with randomly initialized weights. The west central equatorial region was the best forecasted but the eastern boundary regions not so well. The models with SLP gave better results. 9, 12 and 15 months ahead forecasted results gave correlational skills of 0.78, 0.80 and 0.75 shown in fig. 4. The NN results were verified with samples that were not included in training of the network model. From these above-mentioned works, it can be concluded that NNs had an advantage over other traditional models when it comes to forecasting for longer durations in the future.

The cause behind this would be the increase in non-linearity. Tangang et al[55] used ANNs to the three regions namely Nino 3, Nino 3.5 and Nino 4 which independently represented the different parts the equatorial Pacific Ocean (central, eastern and western) respectively as compared to employing NNs to only one region mentioned in the previous work. The SST data sets of Tangang et al[53] were employed in this study as

well. The extended empirical orthogonal functions (EEOF) of SLP field was used which helped in reducing the size of NNs. The introduction of new regions showed that the surface wind and sea level pressure anomalies were responsible for the major changes in SST with respect to ENSO, originating from the region near the equator in Indian Ocean and moved slowly towards the Pacific region in the east[3].

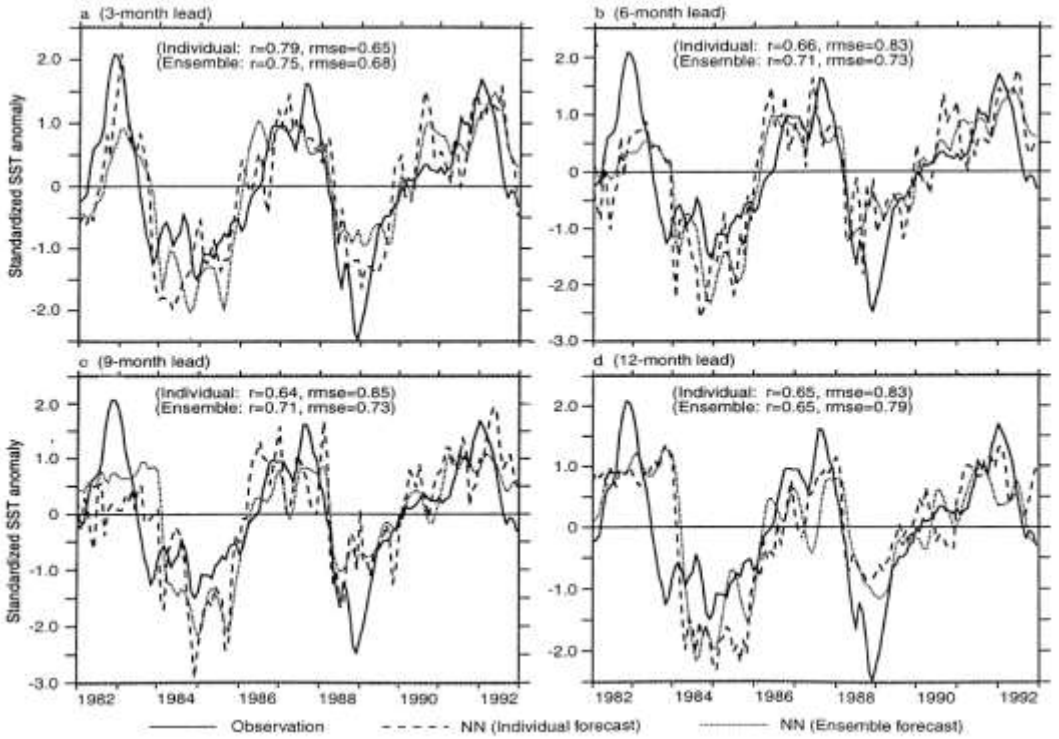

Fig. 3: NN forecasts of Standardized SST anomalies at future leads of 3, 6, 9 and 12 months[53]

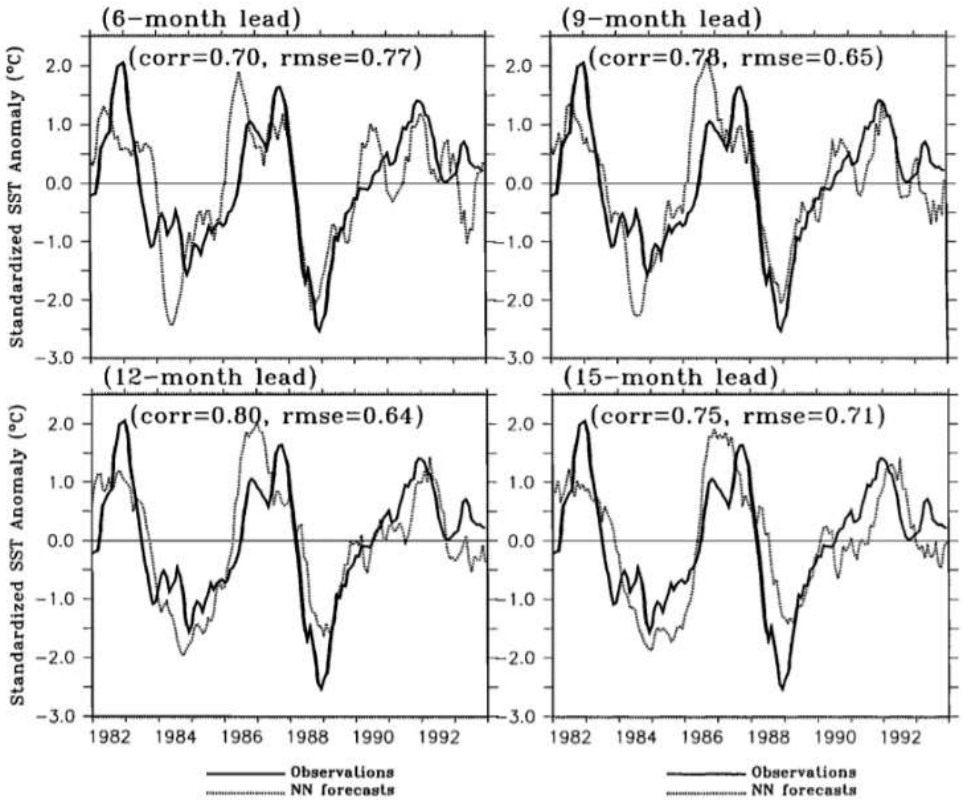

Fig. 4: The NN forecasted Standardized SST anomalies at time ahead of 3, 6, 9, and 12 months with the correlation coefficient (r) for the period of 1983-92 [54]

The NN model was similar to the one used by Tangang et al.[53] The Nino 4 was the best predicted region with a good competence of 1-year time ahead forecast. The forecast comparison of Niño 4 region using NNs with the observed irregularities is shown in the figure 4 with being the statistical parameter of measure being the correlation coefficient (r). Spectral analysis was applied to NNs to obtain crucial input features and nonlinear characteristics.

The results showed greater nonlinear behavior with increase in lead time and hence it can be stated that the relation between the SST and the atmospheric field became more and more nonlinear as well with increase in time duration. The regions to the east of Nino 3.5 and Nino 4 highlighted this trait more and more.

Tang et al[52] performed a performance and results oriented comparative study between neural network (NN), linear regression (LR) and canonical correlational models (CCA). The sea level pressure (SLP) data[63] and the sea surface temperature (SST) data across the tropical Pacific were used in this study.[47,50] EOF was used for reducing the data size. The overall forecasting skills of the three models were compared along with the persistence are shown in fig. 5.

The similar characteristic of NN was once again highlighted that it predicted better results at longer durations (6 months and longer). Despite its non-linearity, NN did not provide any improvement over LR and CCA models. It was then concluded that this lack of improvement may be due to several contributing factors, one being the tropical Pacific region which is mostly linear and can be a key factor for the so obtained NN results.

Penland and Sardeshmukh[41] observed in their study that SST data from the tropical Pacific are comparable to a linear model for a time ahead of up to a year. It was also suggested that the data records were either very insufficient or may be were not large enough. NNs have better generalizing capability but they also require substantial amount of data for proper learning of a system. NNs require data of better quality and less noise. One more possibility was finding a better way to develop NN models. The framing of a sound NN model requires more analyzing than creating a LR or a CCA model. The neural networks till the early 2000s were relatively new method in climate change studies.

Baawain et al[2] developed ANN models to forecast the start of an ENSO event with the help of two indices namely, SOI and Niño3, to simulate ENSO. These indices were used one at a time. Two NN models were used for forecasting ENSO using monthly averaged data. The first NN model used SOI as the output to predict ENSO up to 12 months ahead. The NN inputs were air speed and temperature parameters along with SST anomalies for the 2 selected locations in the Pacific Ocean.

The other NN model used Niño3 as the indicator representing SST anomalies as an output for four locations having the same previous inputs in the Pacific Ocean. For the above-mentioned NN model monthly average data, range from a period of 1994 to 2004 was used. The NN models were multi linear fee forward NNs which were trained using the error back propagation algorithm. The training and testing data were in the ration of 3 to 1. Log activation function and a linear scaling function[45] were applied to the input and output layers of the model.

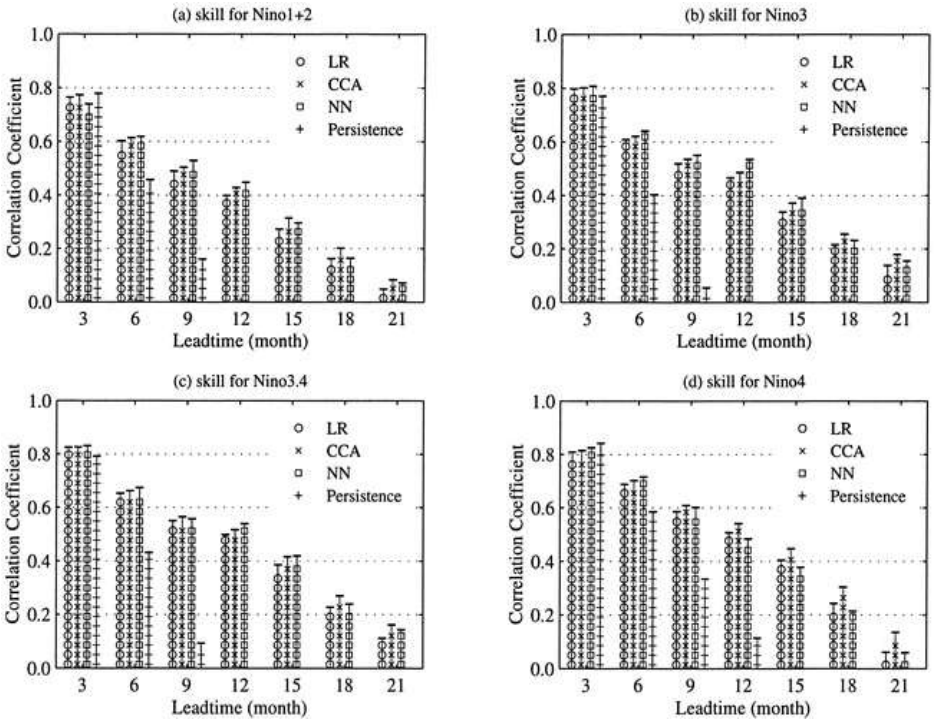

Fig. 5: The prediction capabilities of the four different models and persistence for the 4 regions at different lead times[52]

The optimization of the network was by varying the activation function in the hidden layers as well as the number of iterations. The models did fare well in prediction of onset of ENSO up to 12 months in advance. For SOI and Niño3, coefficient of correlation values of 0.7 and 0.8 were obtained for 12-month time ahead forecasting. These 2 NN models were then compared among themselves and were found to be hugely consistent with 75% compliance in their prediction ability. Fig. 6 shows the actual observed and predicted indicators, SOI and Niño 3 respectively while the table 1 shows the various corresponding NN architectures involved in developing the models.

Wu et al[65] introduced a neural network that was feed forward in nature to predicting SST aberrations in the Pacific Ocean. They used SST and SLP anomalies as the inputs to their model to forecast five leading principal SST components at lead times from a period of 3 to 15 months. The principal component mentioned in this study is nothing but EOF of the SSTs. It is done to reduce the data set to manageable size of prominent modes[18] and is linearly uncorrelated. They made a comparison with LR models and observed that NN models showed higher correlational capabilities and considerable low rmse over most study areas, especially across eastern equatorial Pacific and the western Pacific.

The seasonal and decadal aberrations were also examined in this study. A standard feed forward NN built this research using Bayesian regularization[9]. For training of the NN model Bayesian regularization was employed where the optimization of weight parameter was done using a Bayesian approach[26].

The data sets were then divided into 2 parts, 85 % of the data set was used in training and remaining for an over fitting test. The training set of 85 % was used to train both NN model and a LR model. The NN model was only accepted if it showed greater correlational skills and lesser MSE than the LR model.

Table 1
The NN architectures used in prediction of SOI and Niño3[2]

SST Forecast in months	Correlation Tripathi et al	Correlation coefficients for the different sites					
		Site A	Site B	Site C	Site D	Site E	Site F
1	0.97	0.99	0.96	0.90	0.97	0.99	0.96
2	0.95	0.96	0.94	0.80	0.91	0.98	0.88
3	0.59	0.96	0.94	0.79	0.92	0.98	0.89
4	0.76	0.95	0.94	0.82	0.91	0.98	0.89
5	0.90	0.95	0.94	0.85	0.92	0.98	0.90
6	0.92	0.96	0.94	0.85	0.91	0.98	0.87
7	0.77	0.96	0.94	0.83	0.91	0.98	0.87
8	0.89	0.96	0.94	0.83	0.91	0.98	0.87
9	0.87	0.96	0.95	0.81	0.91	0.98	0.87
10	0.69	0.95	0.94	0.83	0.91	0.98	0.88
11	0.76	0.95	0.94	0.82	0.92	0.98	0.88
12	0.99	0.96	0.94	0.81	0.92	0.98	0.88

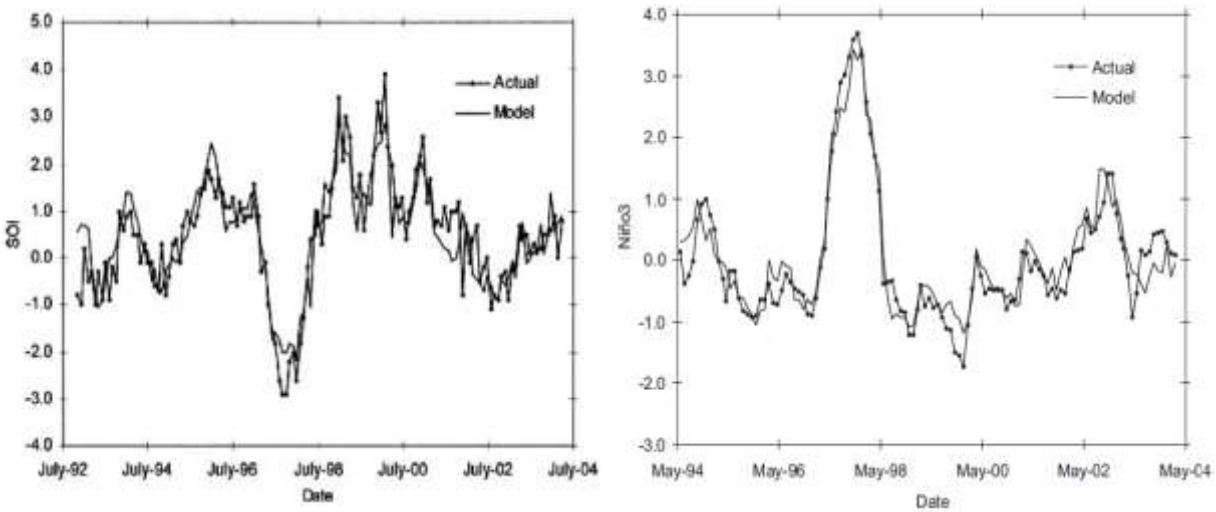

Fig. 6: Actual and predicted SOI and Niño3 respectively with no lead time[2]

This whole process was continued till 30 NN models showed the above mentioned characteristic. The final forecast was made from the average of predictions of these 30 NN models. As there were 30 models and 10 portions were kept for validation purposes, a total of 300 such models were used for the entire prediction. This process was continued until all SST PCs at 3, 6, 9, 12 and 15 months were predicted. Moreover, 8 different iterations for each such model were done as the number of hidden neurons varied from 1 to 8.

As it is evident from the previous studies that the non-linearity between the observed and forecasted varies or grows as the prediction duration increases, the nonlinear (NL) models showed better forecasting skills than their counterparts over the far western Pacific (west of 155◦E) and the eastern equatorial Pacific. Their correlation skills were improved by 0.10-0.14. It can be concluded that nonlinear models have advantage over LR models at longer lead-time i.e. greater than 6 months and during the boreal summer and fall seasons. Fig. 7 compares the skills between the two models at different lead time.

Tripathi et al[59] employed neural networks across the Indian Ocean (27º to 35ºS and 96º to 104º E) to predict SST anomalies. In the study, twelve NNs were created for every month in a year and the network training was performed on the mean values across the region. A feed forward NN combined with logistic activation function was employed in the study. The Reynolds' reconstructed SST data set ranging from a period of 1950 to 2001 was used which consisted of 624 data points i.e. for the 52 years for every 12 months. From the 624 data points, 12 series of corresponding 52 points were obtained in a way that each series served as the data set of a particular month.

Hence, separate time series was formulated for each and every month. Then the average aberrations for each month were calculated. The entire time series data set of 52 months were then into three parts, namely training, testing and validation. For every monthly SST anomaly prediction, a separate NN structure was created as creating a single ANN model to predict the entire anomalies for all the 12 months would increases the number of free parameters enormously and increase the complexity as well.

The output obtained from each NN was the predicted aberrations for the corresponding month. The NN models therefore had one neuron as their outputs. The number of neurons would vary depending upon the number of predictors selected for example from one (for December) to four (for June). For achieving consistency, lag correction was also done. The predictor months were chosen on the basis of their correlational skills. If a correlational coefficient of 0.5 was achieved, then that month was selected as predictor otherwise not.

The number of predictors at max was kept at four. It was observed that the best correlation was achieved with past data of the same region and not with the bordering grid points. The cross validation of the network was given more importance as compared to the other studies mentioned in this study. The training of the data sets was stopped after modification of free parameters for every 10 cycles. Then the performance of the NN was tested and training was resumed. With decline in performance the training was stopped as it meant that the model was over fitting the data sets. The results obtained were then compared with a Linear Multivariate Regression (LMR) model.

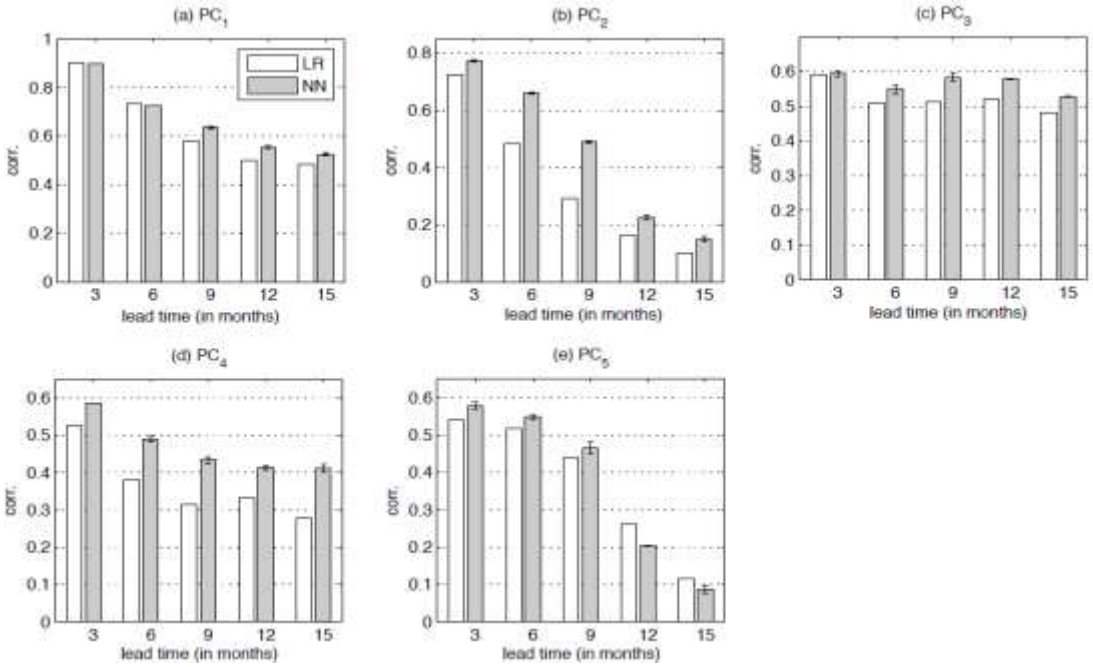

Fig. 7: Skill comparison of NN model and LR model in predicting the 5 principal SST components. The error bar shows ±1 standard error of the prediction capability from the 30-member NNs average[65]

It was observed that LMR model failed to make any forecast when there was reliance of present anomalies on the past anomalies showing nonlinear behavior. During linear dependency, NN and LMR models showed similar forecasting skills. NN models showed better trend predicting ability and their corresponding forecasts were better than the average forecast for every month. From the fig. 8 above, this behavior can be understood properly. For the months of Sept, Oct and Nov, it is evident from the figure that both LMR and NN models show lesser than normal forecasting skills, which is due to the fact that there is a linear relationship between past and present aberrations. This brings us to another important aspect of ANNs that they are better suited for nonlinearity.

Two separate nonlinear methods, support vector regression (SVR) and Bayesian NN (BNN) were used by Martinez and Heish[30] alongside a LR model to forecast tropical Pacific SSTs. They forecasted at lead times of 3 to 15 months. The predictors were SLPs and SSTs. They studied 2 data sets 1950-2005 and from 1980 -2005 with the latter having another component known as the warm water volume (WWV), which is volume integral of water above the 20^0 isotherm between 5°N–5°S, 120°E–80°W[31]. The sole reason of using two data sets of SST and SLP data was prior to 1980, the data set did not contain WWV which was an extra predictor used in this study.

The LR model was achieving similar sort of results as well. SVR can be termed as a supervised learning model[35] which is associated with learning algorithms that helps in regression analysis. The SVR had two fundamental advantages over the BNN model i.e. it did not have multiple minima in optimization process and a robust error norm to outliers in the data. But then again in spite of these advantages, it did not improve on the results obtained by BNN significantly. The reason would be that BNN uses probability distribution over the network weights. It can naturally address the issues of over fitting and model selection sometimes without even needing a separate cross validation set.[43] The data was then divided into 10 parts and 9 parts were used to construct the model and the remaining one was used for testing. The 9 parts were further divided into 2 parts for training (two thirds) and cross validation (one third).

The numbers of hidden layers were randomly picked for which the error function had to be optimized. Since the error function contained many local minimum, it led to many possible different models. The model with least Mean Square Error (MSE) was kept. SVR provided a good solution in this matter as it avoids multiple local minima and uses a robust error norm,

$$E_{\in}(y-y') = \begin{cases} 0, & if\ |y-y'| < \in, \\ |y-y'| - \in, & otherwise \end{cases}$$

whereas y and y' are the observed and predicted outputs and for outliers. From the fig. 9 and 10, it can be seen that the nonlinear models have performed better than the LR models and evidently the models with WWV predictor comparatively produced even better skills. It was hence observed that predictors such as WWV can definitely characterize an ENSO event. Its inclusion along with other traditional or common predictors increased the prediction capability of the model.

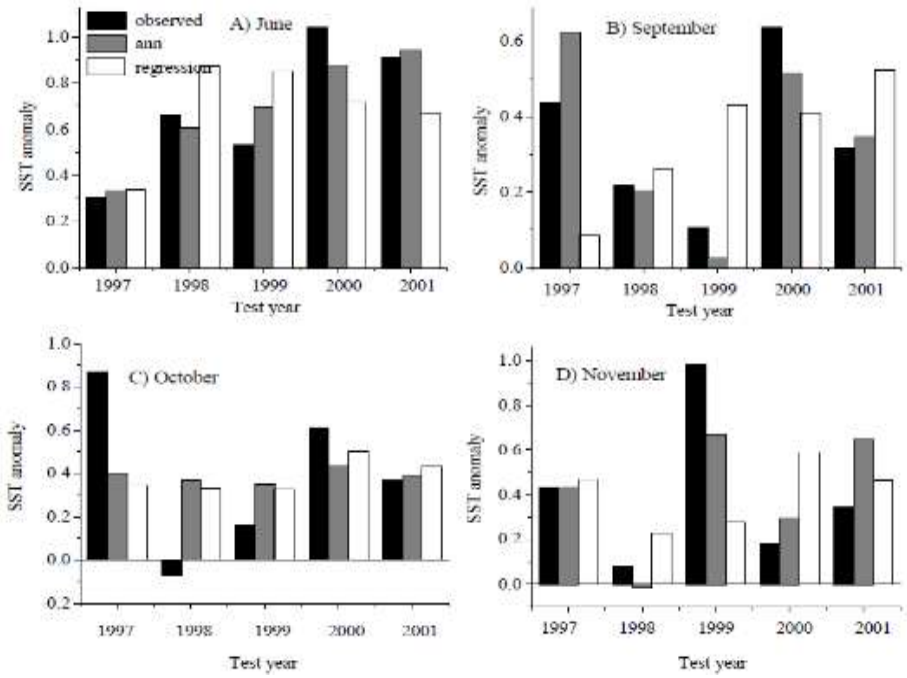

Fig. 8: Predictive skills comparison between NN and LR models for the months of (A) September, (B) October, (C) and (D) November[59]

Patil et al[37] exhibited the advantage of non-linear autoregressive (NAR) neural network models to forecast SST aberrations on a monthly scale for one to twelve months in the future at six different locations based on 61-year data records. The data records ranged from Jan 1945 to Dec 2005. Thus for the 61 years, there were 732 values (61 years and their corresponding 12 months). The initial study was done with a traditional feed forward back propagation NN (FFBP)[68]. To improve on its skills another network NAR network was employed.

The study regions were coded as AS (19ºN - 20ºN and 68ºE), BOB (18ºN - 19ºN and 90ºE), EEIO (1ºS - 1ºN and 90ºE), SOUTHIO (9ºS -11ºS and 95ºE - 98ºE), THERMO (16ºS - 14ºS and 56ºE - 58ºE), and WEIO (1ºS - 1ºN and 65ºE). These are all located in the north Indian Ocean. NAR is a special kind of recurrent network having feedback protrusions shown in the figure 11. It uses a static back propagation while training.

The training was done using Levenberg-Marquardt algorithm[6]. The feed forward NN includes sigmoid activation function in the hidden layers. The output layer of neurons is transformed using a linear activation function. The number of input neurons was kept at 24 and the output of the network was one.

The NAR network produced good predictions over the different locations used in this study and was superior to the results obtained by FFBP NN architecture. The statistical parameters like correlation coefficients (CC), MSE, mean absolute error (MAE) and Nash Sutcliffe error (NSE) were used. The NAR network attained correlation above 0.90 between predicted and observed values. MSE and MAE were observed to be less than 0.23◦C and 0.38◦C respectively and NSE coefficient of the order of 80%. The results of this study and a comparison with a previous study are shown in the table 2.

Mohongo and Deo[29] forecasted SST anomalies in the region around Indian Ocean using NAR network using exogenous input (NARX) on a monthly and seasonal scale. This model is based on the principle of forecasting based on time series.[67] This present study employs a multi layered NARX network. The governing equation for the NARX model is given below, the model can predict the series *y(t)* based on the past values, d of time series *y(t)* and *x(t)* :

$$y(t) = f((x)(t-1),.....,x(t-d), y(t-1) ,...y(t-d))$$

The prior values of x(t) and y(t) time series are stored by the past values d given in the above equation which are called tapped delay lines (TDLs). The input neurons were set at four which was equal to the number of predictors. For example, the four predictors for the month of Feb of a certain year would be Jan of the same year and Dec, Nov and Oct of last year. The study locations were coded EAF, a region located along the eastern African shore (6-7ºS, 39-40ºE) and EQT, a location that decisively lied around Indian Ocean Dipole's western pole (0-1ºS, 59-60ºE). The data ranged for a period of 142 years from Jan 1870 to Dec 2011. Hence, a total of 1704 data points were obtained (142 years and the corresponding 12 months).

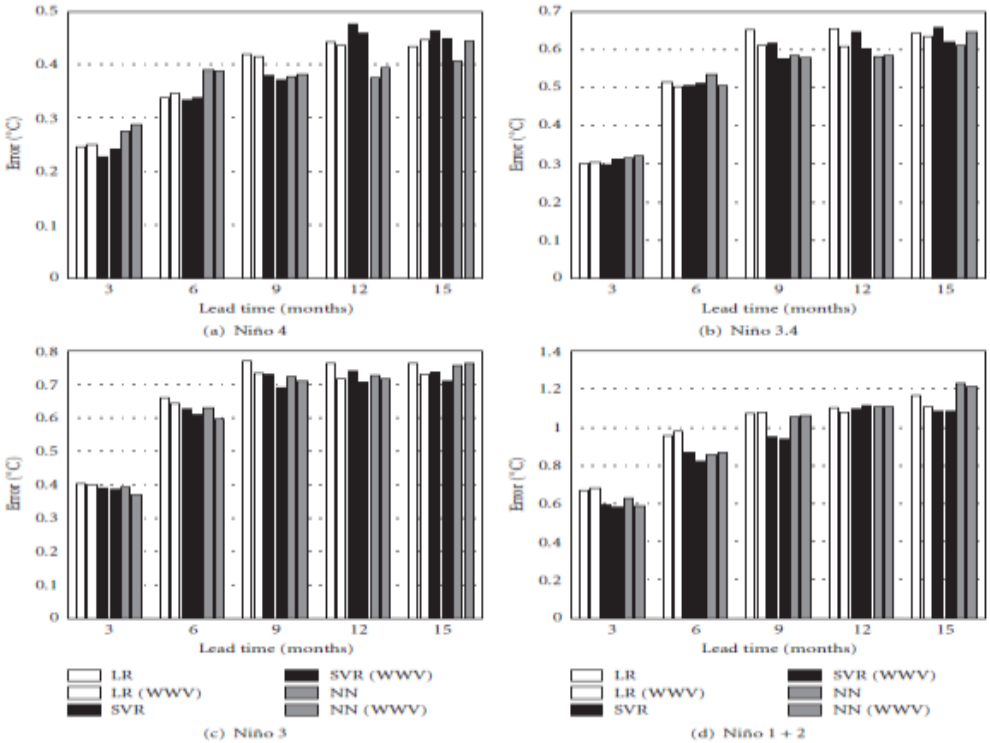

Fig. 9: RMSE values obtained by the different models during the predictions for the Nino 4, Nino 3.4, Nino 3, and Nino 1 + 2 regions[30]

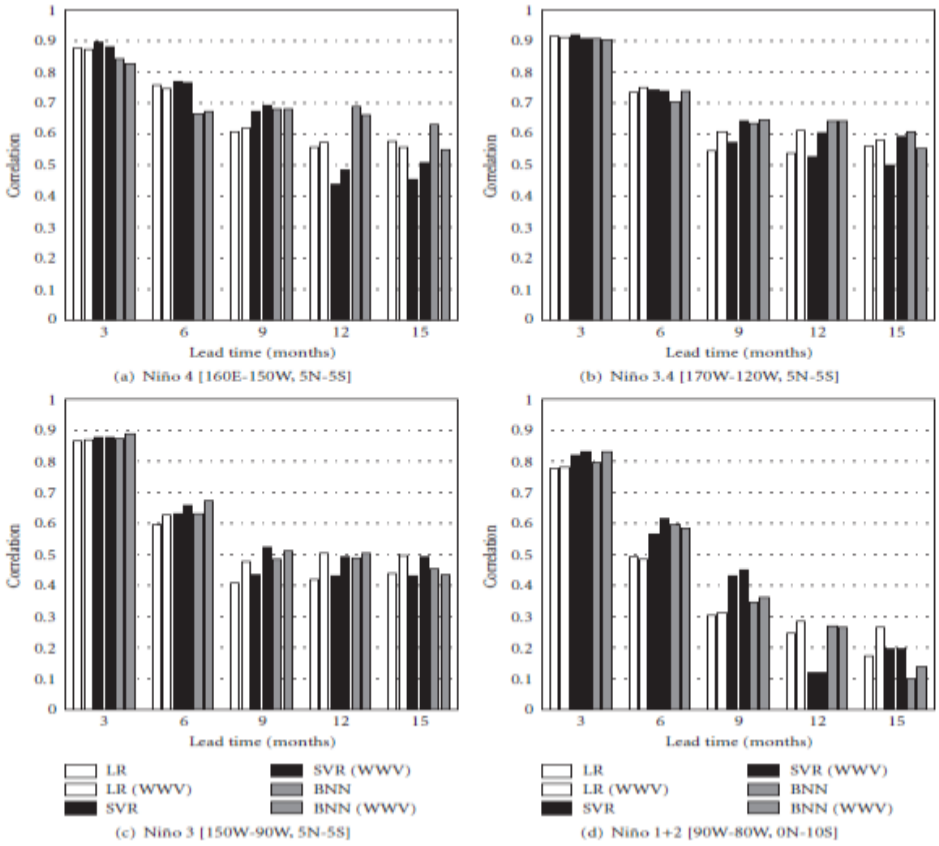

Fig. 10: Correlation values obtained by the different models during the predictions for the Nino 4, Nino 3.4, Nino 3, and Nino 1 + 2 regions[30]

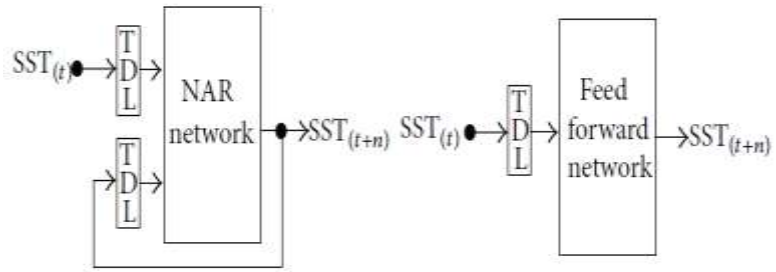

Fig. 11: NAR architecture for SST prediction[37]

Table 2
Skill comparison in terms of Correlation coefficient at different locations (CC)[37]

Lead time (month)	1	2	3	4	5	6	7	8	9	10	11	12
Number of hidden Neurons (SOI)	9	4	12	14	16	4	12	10	10	6	6	16
Number of hidden Neurons (Niño3)	12	12	16	5	6	6	13	15	8	7	5	16
Hidden layer activation function (SOI)	Log	tanh	Log	Log	Log	tanh	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian	Log
Hidden layer activation function (Niño3)	Log	Log	Log	tanh	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian	tanh	Log
Number of epochs (SOI)	3100	500	15000	4250	6600	2500	775	460	500	575	875	4260
Number of epochs (Niño3)	1200	1500	6400	300	115	125	275	280	225	135	530	160

To achieve necessary consistency, the normalization scheme of Tripathi et al[59] was employed in this study and the obtained SST aberrations were mapped onto the range 0.2 - 0.8. This process is supported by the works of Maier and Dandy[27]. The results of this model were compared with three other NN namely, Feed forward NN, Radial Basis function NN (RBFNN) and Generalized Regression NN (GRNN) and also with a linear system, the Autoregressive Integrated Moving Average with Exogenous Input (ARIMAX) model. The architecture RBF NN is shown in fig. 12. It is a three layered NN. The first layer compiles all the input parameters and hence requires more neurons than a standard feed forward NN.

The hidden layer employs RBFS as their activation function. They perform the non-linear transformations of the inputs and they have only one layer of hidden neurons. Unlike FFNN and GRNN, RBFNN involves an unsupervised training component.[13]

A GRNN (Fig.13) is a variation of RBFNN. For every epoch, one hidden neuron is centered and these units are called kernels. These are nothing but Gaussian functions. The output, which is a special linear layer, is then obtained in a slightly different manner as compared to other NN models[60]. These minute changes improve the local approximation capability of the neural network. Eight different performance evaluators were used to assess the skill of the models and are shown in the table 3. The testing period of the models ranged from 1981 to 2010. From the table above, it was observed that NARX has outperformed other NN models and ARIMAX. Therefore, it can be stated that it NARX model held good for both monthly and seasonal predictions.

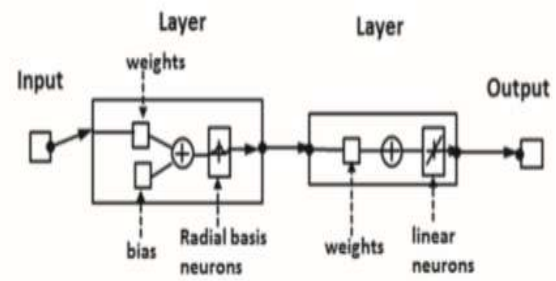

Fig. 12: The architecture of a standard Radial Basis function NN[29]

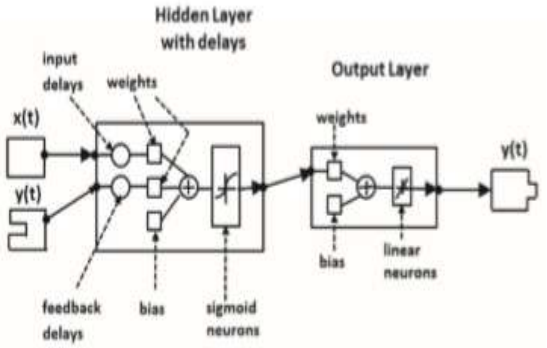

Fig. 13: Diagrammatic representation of a typical Generalized Regression NN[29]

Table 3
kill comparison of the ANNs and ARIMAX model using various statistical performance indicators on the average monthly and seasonal SST anomalies respectively during the testing period of 1981-2010[29]

	MODEL	R	r^2	CE	MaxAE (ºC)	SDD (ºC)	RMSE (ºC)	SDE (ºC)	MAE (ºC)	MSRE (ºC)	MAPE (ºC)
MONTHLY	NARX	0.90	0.81	0.79	0.30	0.27	0.12	0.12	0.10	3.33	127.13
	FFNN	0.82	0.68	0.63	0.41	0.27	0.16	0.16	0.13	14.36	193.74
	RBFNN	0.80	0.64	0.60	0.46	0.27	0.17	0.17	0.13	75.05	194.21
	GRNN	0.80	0.64	0.60	0.46	0.27	0.17	0.17	0.13	81.89	197.34
	ARIMAX	0.80	0.64	0.59	0.45	0.27	0.17	0.17	0.14	50.26	185.40
SEASONAL	NARX	0.98	0.95	0.95	0.13	0.25	0.05	0.05	0.04	1.42	60.36
	FFNN	0.97	0.93	0.93	0.19	0.25	0.07	0.07	0.05	2.68	71.64
	RBFNN	0.96	0.92	0.90	0.19	0.25	0.08	0.07	0.07	6.52	92.43
	GRNN	0.96	0.92	0.91	0.20	0.25	0.08	0.07	0.06	6.78	95.60
	ARIMAX	0.96	0.93	0.91	0.19	0.25	0.07	0.07	0.06	5.09	90.31

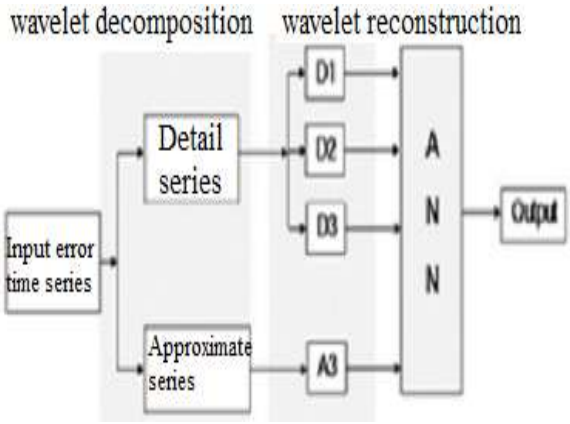

Fig. 14: Functioning of the Wavelet Neural Network[38]

Patil et al[38] introduced wavelet neural network to forecast SST anomalies. The study was carried out to predict daily, monthly and seasonal SST anomalies at different locations in the Indian Ocean region. An error time series was generated by calculating the variation between obtained and observed SST anomalies and subsequent time ahead errors were predicted. This predicted error was then introduced to the numerical estimate and SST aberrations were calculated at 5 times ahead in future. The performance was assessed using various statistical performance indicators. WNN is a hybrid network in which the inputs have to be preprocessed using wavelet transform.

Deka et al[14] and Dixit et al[15] highlighted various applications of WNN and Shoaib et al[49] used these NNs for run off predictions. The pre-processing involves removal of unwanted singnals with the help of the filters (high and low pass) and then feeding the broken down components into the network as inputs.

Wavelet transform (WT) is a cyclic function for breaking down the series of data as shown in fig. 14. It segregates important features (useful data) of a series from its surroundings. The advantage is that it helps in retaining time dependent information. The filtering depends on the width of the WT which can be adjusted for analyzing low frequency data. These transforms are created from a mother wavelet φ by either decreasing or increasing the scale parameter *s* or by changing the form or position using translation parameter *u*. These two parameters actuate the resolution or width of filtered part. The mother wavelet $\varphi(x)$ is given by:

$$\varphi_{s,u}(x) = \frac{1}{\sqrt{s}}\varphi\frac{(x-u)}{s}$$

where $\varphi_{s,\,u}(x)$ is a function of translation parameter, *u* and scale parameter S are given by:

$$S = (x_1 w_1 + x_2 w_2 + \cdots) + \emptyset,$$

where *x*, *w* and $\emptyset$ represent the input, weight and bias respectively. The mother wavelet helps in mining hidden information in the signal through compression known as high pass and helps in approximating information through dilation known as low pass. These 2 phases are analyzed separately. These decompositions are again filtered through various levels until the network achieves desired results.

In this work, the wavelet defined above is itself used as an activation function. The transformation was achieved during the training of the model as the test data was not included during the for actual application. The Discrete Wavelength Transform (DWT) decomposes the initial signals up to three levels of sub signals and thus three detail signals and one approximate signal are acquired.

This study was performed with seven different WTs namely Haar wavelet, Symlet wavelet, Daubechies wavelet, Coiflet wavelet and Meyer wavelet with discrete approximation[34]. The study area locations were the same as in Patil et al.[37] It was observed that Meyer wavelet with discrete approximation based on a particular mother WT showed better results than the other wavelets. The NAR network was then set for the final computation (WNN being a hybrid network).

The data set was divided into parts for training and testing (70:30). Daily, weekly and monthly prediction of SST anomalies was made. At every time scale, SST anomalies were forecasted over five lead times ahead. The results obtained by the seven WTs were assessed using various statistical indicators of r, RMSE, and MAE. The results obtained by the seven WTs at site AS in terms of r and MAE are given in the fig. 15. The concept of combining neural and numerical methods showed in this work resulted in precise daily, weekly and monthly SST predictions over five lead times ahead at the selected sites.

Patil et al[38,39] and compared *in situ* instrumentally observed data and tested the results at various sites in the Indian Ocean. A numerical method was employed to investigate the precision of location specific information with *in situ* data records. Then NNs were used to predict SST anomalies a few times ahead in future using reanalysis data. Reanalysis, numerical, and *in situ* SST data were used in this study.

This observational data ranged from a period of Jan 2002 to June 2015, spanning around 13.5 years. The Regional Ocean Modeling System (ROMS) which supplied the numerical observations of daily SSTs, is an equation based climatic ocean model employed for various kind of coastal operations[33,62]. These observations ranged from January 2013 to May 2015 for a period of 29 months. Based on the availability of the data, 6 locations were chosen. These sites were located across the Indian Ocean.

The period from 10 Dec, 2014 to May 27, 2015 was used for testing to assess the precision of such numerical model over a lead time of 5 days in future. This constituted 20% of the total data records. The statistical parameter r was high and proved to be a good prediction but RMSE and MSE were not as low as expected and hence, it left scope for further improvements.

Therefore, WNNs were employed at these 6 locations to predict daily values of SST 5 days ahead in time. The WNN was trained by NOAA OI version 2 reanalysis data and the testing was done as per the observations made by buoys under RAMA project. A comparison study was performed to assess the results obtained by WNN. The comparisons were made with a standard three-layered FFBP neural network and with a persistent model (PER). The PER model at every next step takes the predicted value equivalent to the present value i.e. SST (t+Δt) =SST(t) where the present surface temperature value is SST (t) and SST (t+Δt) is the future value at (t+Δt).

The variations in the temperatures were found to be very less in comparison to originally observed values. For building the model the average SST anomalies were favored and not their absolute values. As a result, the re-analysis data were converted into abnormalities by subtracting the calculated average values from their absolute values. The decomposition of abnormalities at every site was achieved with a wavelet transform at three levels to produce detailed and approximate components.

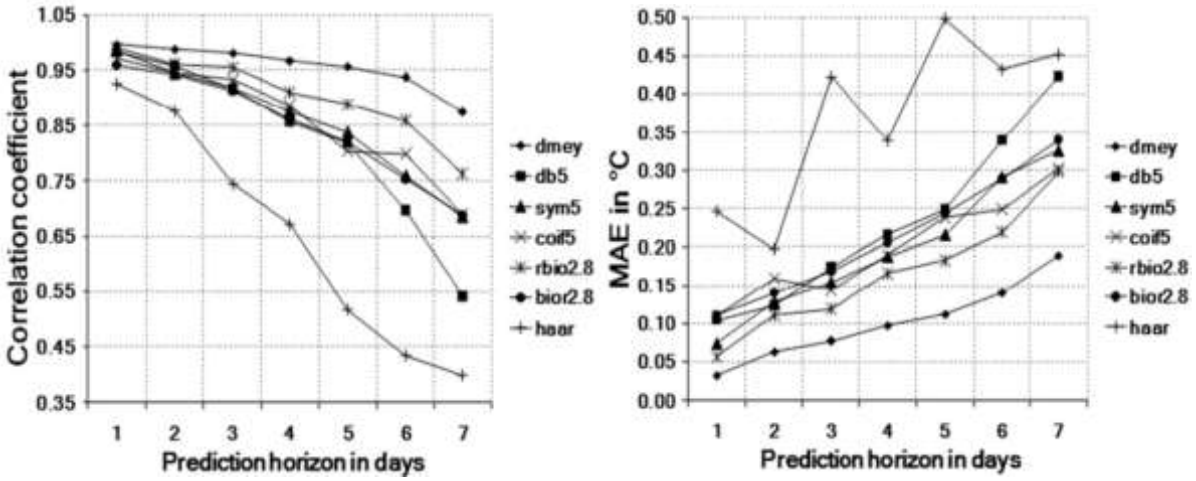

Fig. 15: Performance of the 7 WTs at site AS (19N to 20N and 68E) in terms of correlation coefficient (r) and MAE.[38]

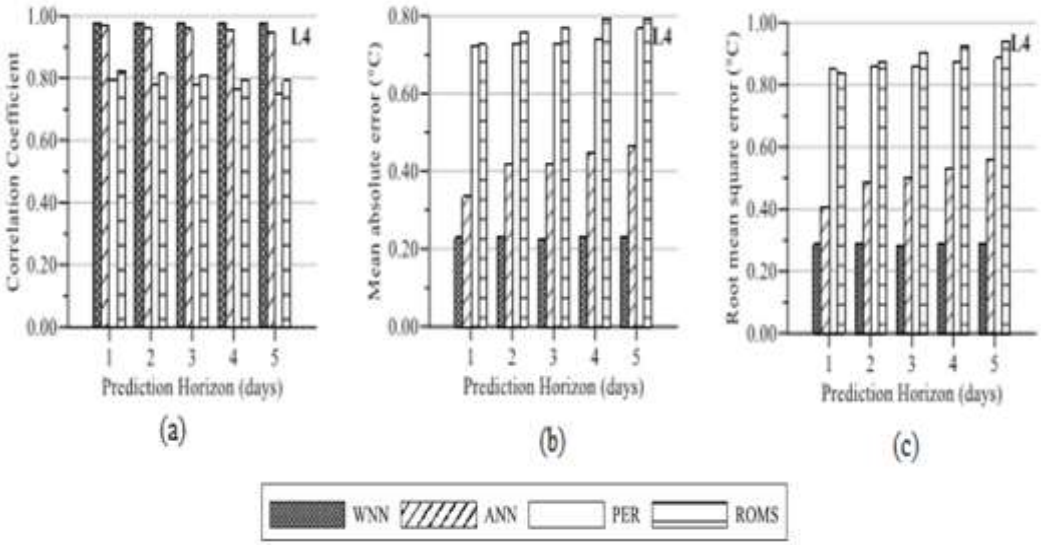

Fig. 16: Skill comparison between WNN, ANN, PER and numerical ROMS based SST prediction in terms of r, MAE and RMSE at site L4 (time lead of 5 days in future).[39]

These filtered components were then fed to NN for training to produce time ahead forecasting of SSTA. The ANN part of this WNN is similar to the architecture of NN used by the authors. The skill comparison between the different models was further analyzed in detail with the help of statistical performance measures namely r, RMSE and MAE. Figure 16 highlights the capabilities of the different models used in this study at location L4 over the time horizon of 5 days. The predictions at other sites presented identical results.

Conclusion and Discussion

ENSO is a natural climate phenomenon which has impacts globally and occurs regularly. It has caused droughts, floods, wildfires, and dust and snow storms. It has resulted in huge losses in terms of lives and economy of many countries which are directly affected by this event. ENSO occurs every 2 to 8 years; hence its precise prediction becomes very important. As discussed in the study earlier, ENSO has several nonlinear features which make it very difficult to predict with high degree of accuracy. Over the years, researchers and government bodies have toiled a lot to accumulate and analyze data concerning this climatic event.

Artificial Neural Networks (ANNs) have proven to a big help in analyzing and prediction of ENSO. As ANNs are highly nonlinear and the non-linearity is spread throughout the network, it comparatively gives better results than other traditional methods or models. As the computation in ANNs is very robust, the decline in performance is not sudden. With adequate amount of data, ANNs shows good generalization capacity.

The works of various researchers mentioned in the study clearly show viability and relevance of NNs in precise prediction of ENSO. NNs along with its advantages have certain disadvantages as well. A network that is smaller than required does not learn the data samples properly and larger than required mostly over fits the data. There is no fixed method to attain compatibility; it requires great deal of trials and analyzing of the results. Much emphasis has to be given in pre-processing of the data before it is fed into the network which reduces the risk of learning undesirable noisy features. The results of all the mentioned works showed the capability of NNs to predict SSTA with lesser errors as compared to the other standard methods (mostly numerical ones).

Most of the above-mentioned studies are location specific and it can be aforementioned that ANNs do not have the capability to include space and time variability of SST together for any specified area. In spite of these few disadvantages ANNs have proven to be very successful among the researchers in climate change studies due to the fact that they provide uniformity in data processing. They employ similar notations in all domains of application which makes it straight forward and more accessible to the research community.

Conclusion

For monitoring specific locations to predict the onset of ENSO events, NNs are much more suited and preferred to other techniques. ANNs provided better forecasts at longer time leads and with new upcoming models like WNN which considers both physics based and data driven methods together in prediction of SST anomalies. It can be stated that ANNs are certainly very efficient. They can certainly adapt and learn unknown and uncertain systems with ENSO being one of them.

References

1. An S.I. and Jin F.F., Collective role of thermo cline and zonal advective feedbacks in the ENSO mode, *J. Clim*., **14,** 3421–3432 **(2001)**

2. Baawain M.S., El Niño southern-oscillation prediction using southern oscillation index and Niño3 as onset indicators: Application of artificial neural networks, *J. Environmental Engineering and Science*, **4(2),** 113-121 **(2005)**

3. Barnett T.P., Prediction of the El Niño of 1982–83, *Mon. Wea. Rev.*, **112**, 1403–1407 **(1984)**

4. Barnston A.G. and Ropelewski C.F., Prediction of ENSO episodes using canonical correlation analysis, *J. Climate*, **5**, 1316–1345 **(1992)**

5. Barnston A., Van Den Dool H.M., Zebiak S.E., Barnett T.P., Ji M., Rodenhuis D.R., Cane M.A., Leetmaa A., Graham N.E., Ropelewski C.R., Kousky V.E., O'Lenic E.A. and Livezey R., Long-lead seasonal forecasts: Where do we stand?, *Bull. A.n. Meteorol. Soc.*, **75**, 2097-2114 **(1994)**

6. Basterrech S., Mohammed S., Rubino G. and Soliman M., Levenberg—Marquardt Training Algorithms for Random Neural Networks, *The Computer Journal*, **54(1),** 125–135 **(2011)**

7. Bjerknes J., Atmospheric teleconnections from the equatorial Pacific, *Mon. Wea. Rev*., **97**, 163-172 **(1969)**

8. Bjerknes J., Large-scale atmospheric response to the 1964-65 Pacific equatorial warming, *J. Phys. Oceanography*, **2,** 212-217 **(1972)**

9. Bishop C.M., Neural networks for pattern recognition, Clarendon Pr. Oxford **(1995)**

10. Cane M.A., Tropical Pacific ENSO models: ENSO as a mode of the coupled system, Climate system modeling, 787 **(1992)**

11. Cannon A.J. and Hsieh W.W., Robust nonlinear canonical correlation analysis: application to seasonal climate forecasting. Nonlinear Processes in Geophysics, *European Geosciences Union*, **15(1),** 221-232 **(2008)**

12. Chauvin Y., A back propagation algorithm with optimal use of Hidden Units, Advances in Neural Information processing systems, Vol.1, Morgan Kaufmann Publishers Inc., San Francisco, 46-55 **(1989)**

13. Chen F., Cowan C.F.N. and Grant P.M., Orthogonal least squares learning algorithm for radial basis function networks, IEEE Transactions on Neural Networks, **2(2),** 302-309 **(1991)**

14. Deka P.C. and Prahlada R., Discrete wavelet neural network approach in significant wave height forecasting for multistep lead time, *Ocean Eng*., https://doi: 10.1016/j.oceaneng.2012.01.017, **43**, 32–42 **(2012)**

15. Dixit P., Londhe S. and Dandawate Y., Removing prediction lag in wave height forecasting using neurowavelet modelling technique, *Ocean Eng*., https://doi: 10.1016/j.oceaneng.2014. 10.009, **93**, 74-83 **(2015)**

16. Flohn H. and Fleer H., Climate teleconnections with the equatorial Pacific and the role of ocean/atmosphere coupling, *Atmosphere*, **13**, 96-107 **(1975)**

17. Graham N.E., Michaelsen J. and Barnett T.P., An investigation of the El Nino-Southern oscillation cycle with statistical models, 2, Model results, *J. Geophys. Res.*, **92**, 14271–14289 **(1987)**

18. Hannachi A., Jolliffe I.T. and Stephenson D.B., Empirical orthogonal functions and related techniques in atmospheric science: A review, *Int. J. Climatol*., **27**, 1119–1152 **(2007)**

19. Hsieh W.W., Nonlinear canonical correlation analysis by neural networks, *Neural Networks*, **13**, 1095–1105 **(2000)**

20. Hurlburt H.E., Kindle J.C. and O'Brien J.J., A numerical simulation of the onset of El Niño, *Journal of Physical Oceanography*, **6(5),** 621-631 **(1976)**

21. Inoue M. and O'Brien J.J., A forecasting model for the onset of El Nino, *Mon. Wea. Rev.*, **112**, 2326–2337 **(1984)**

22. Jacobs R.A., Increased rates of convergence through learning rate adaptation, *Neural Networks*, **1(4)**, 295-307 **(1988)**

23. Julian P.R. and Chervin R.M., A study of the Southern Oscillation and Walker Circulation phenomena, *Mon. Wea. Rev*., **106**, 1433-1451 **(1978)**

24. Latif M., Barnett T.P., Cane M.A., Flügel M., Graham N.E., Von Storch H., Xu J.S. and Zebiak S.E., A review of ENSO prediction studies, *Clin. Dyn.*, **9**, 167-179 **(1994)**

25. Latif M., Anderson D., Barnett T., Cane M., Kleeman R., Leetmaa A., O'Brien J., Rosati A. and Schneider E., A review of the predictability and prediction of ENSO, *J. Geophys. Res*., **103**, 14375–14393 **(1998)**

26. MacKay D., Bayesian interpolation, *Neural Computation*, **4(3)**, 415-447 **(1992)**

27. Maier H.R. and Dandy G.C., Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications, *Environmental Modelling and Software*, **15(1),** 101-124 **(2000)**

28. McCreary J., Eastern Tropical Ocean Response to Changing Wind Systems: with Application to El Niño, *Journal of physical Oceanography*, **6**, 632-645 **(1976)**

29. Mahongo S.B. and Deo M.C., Using artificial neural networks to forecast monthly and seasonal sea surface temperature anomalies in the Western Indian Ocean, *The International Journal of Ocean and Climate Systems*, **4(2),** 133–150 **(2013)**

30. Martinez S.A. and Hsieh W.W., Forecasts of tropical Pacific sea surface temperatures by neural networks and support vector regression, *International Journal of Oceanography*, **2009**, 1-13 **(2009)**

31. Meinen C.S. and McPhaden M.J., Observations of warm water volume changes in the equatorial Pacific and their relationship to El Ni˜no and La Ni˜na, *Journal of Climate*, **13(20),** 3551–3559 **(2000)**

32. Monahan A.H., Nonlinear principal component analysis: Tropical Indo–Pacific Sea surface temperature and sea level pressure, *J. Climate*, **14**, 219–233 **(2001)**

33. Moore A.M., Arango H.G., Di Lorenzo E., Cornuelle B.D., Miller A.J. and Neilson D.J., A comprehensive ocean prediction and analysis system based on the tangent linear and adjoint of a regional ocean model, *Ocean Model*, **7(1),** 227–258 **(2004)**

34. Moraud E.M., Wavelet networks, School of Informatics, University of Edinburgh, 8 [Available online at http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/AV0809/martinmoraud.pdf.] **(2009)**

35. Muhammad I. and Yan Z., Supervised Machine Learning Approaches: A Survey, *ICTACT Journal on Soft Computing*, **5(3),** 249-268 **(2015)**

36. Nash J.E. and Sutcliffe J.V., River flow forecasting through conceptual models, Part 1, A discussion of principles, *Journal of Hydrology*, **10**, 282-290 **(1970)**

37. Patil K., Deo M.C., Ghosh S. and Ravichandran M., Predicting Sea Surface Temperatures in the North Indian Ocean with Nonlinear Autoregressive Neural Networks, *International Journal of Oceanography*, https://doi:10.1155/2013/302479, **2013**, 1–11 **(2013)**

38. Patil K., Deo M.C. and Ravichandran M., Prediction of Sea Surface Temperature by Combining Numerical and Neural Techniques, *Journal of Atmospheric and Oceanic Technology*, doi:10.1175/JTECH-D-15-0213.1, **33(8),** 1715–1726 **(2016)**

39. Patil K. and Deo M.C., Prediction of Daily Sea Surface Temperature Using Efficient Neural Networks, *Ocean Dynamics* https://doi: 10.1007/s10236-017-1032-9, **67**, 357–368 **(2017)**

40. Peddada S.D. and Haseman J.K., Analysis of nonlinear regression models: a cautionary note, *Dose-response: an International Journal*, **3(4),** 342-352 **(2005)**

41. Penland C. and Sardeshmukh P.D., The optimal growth of tropical sees surface temperature anomalies, *J. Climate*, **8**, 1999–2024 **(1995)**

42. Philander S.G.H. and Pacanowski R.C., Oceanic response to cross-equatorial winds (with application tocoastal upwelling in low latitudes), *Tellus*, **33,** 201-210 **(1981)**

43. Pourret O., Naim P. and Marcot B., Bayesian Networks, Wiley, New York **(2008)**

44. Quinn and Burt W., Use of the Southern Oscillation in weather prediction, *J. Appl. Meteor*., **11**, 616-628 **(1972)**

45. Ramachandran P., Zoph B. and Le Q.V., Searching for Activation Functions, ArXiv, http: //arxiv.org/abs/1710.05941 **(2017)**

46. Reiter E.A., Long-term wind variability in the tropical pacific; its possible causes and effects, *Mon. Wea. Rev*., **106**, 324-330 **(1978)**

47. Reynolds R.W. and Smith T.M., Improved global sea surface temperature analyses using optimum interpolation, *J. Climate*, **7**, 929–948 **(1994)**

48. Rumelhart D.E., Hinton G.E. and Williams R.J., Learning internal representations by error propagation, in Parallel Distributed Processing, Rumelhartand D.E. and McClelland J.L., eds., MIT Press, Cambridge, 318-362 **(1986)**

49. Shoaib M., Shamseldin A.Y. and Melville B.W., Comparative study of different wavelet based neural network models for rainfall–runoff modeling, *J. Hydrol.*, https://doi: 10.1016/j.jhydrol. 2014.04.055, **515**, 47–58 **(2014)**

50. Smith A.E., Control chart representation and analysis via back propagation neural networks, Proceedings of the 1992 International Fuzzy Systems and Intelligent Control Conference **(1992)**

51. Smith T.M., Barnston A.G., Ji M. and Chelliah M., The impact of Pacific Ocean subsurface data on operational prediction of tropical Pacific SST at the NCEP, *Wea. Forecasting*, **10,** 708–714 **(1995)**

52. Tang B., Hsieh W.W., Monavan A.H. and Tangang F.T., Skill Comparisons between Neural Networks and Canonical Correlation Analysis in Predicting the Equatorial Pacific Sea Surface Temperatures, *Journal of Climate*, **13**, 287-293 **(2000)**

53. Tangang F.T., Hsieh W.W. and Tang B., Forecasting the equatorial Pacific sea surface temperatures by neural network models, *Climate Dyn*., **13**, 135–147 **(1997)**

54. Tangang F.T., Tang B., Monahan A.H. and Hsieh W.W., Forecasting ENSO events—A neural network-extended EOF approach, *J. Climate*, **11**, 29–41 **(1998)**

55. Tangang F.T., Hsieh W.W. and Tang B., Forecasting the regional sea surface temperatures of the tropical Pacific by neural network models, with wind stress and sea level pressure as predictors, *J. Geophys. Res*., **103**, 7511–7522 **(1998)**

56. Trenberth K. and Hoar T.J., The 1990-1995 E1 Nino Southern Oscillation event: Longest on record, *Geophys Res. Left*, **23**, 57-60 **(1996)**

57. Trenberth K.E., the definition of El Nino, *Bulletin of the American Meteorological Society*, **78(12),** 2771-2777 **(1997)**

58. Trenberth K.E. and Stepaniak D.P., Indices of El Niño Evolution, *Journal of Climate*, **14(8),** 1697-1701 **(2001)**

59. Tripathi K.C., Das M.L. and Sahai A.K., Predictability of sea surface temperature anomalies in the Indian Ocean using artificial neural networks, *Indian Journal of Marine Sciences*, **35(3),** 210–220 **(2006)**

60. Wasserman P.D., Advanced Methods in Neural Computing, Van Nostrand Reinhold, New York, 255 **(1993)**

61. White W.B. and Walker A.E., Meridional atmospheric teleconnections over the North Pacific from 1950 to 1972, *Mon. Wea. Rev*., **101**, 817-822 **(1973)**

62. Wilkin J.L., Arango H.G., Haidvogel D.B., Lichtenwalner C., Glenn S.M. and Hedström K.S., A regional ocean modeling system for the Long-term Ecosystem Observatory, *Journal of Geophysical Research: Oceans*, **110(C6),** 1-13 **(2005)**

63. Woodruff S.D., Slutz R.J., Jenne R.L. and Steurer P.M., A Comprehensive Ocean–Atmosphere Data Set, *Bull. Amer. Meteor. Soc*., **68**, 1239–1250 **(1987)**

64. Wu A. and Hsieh W.W., Nonlinear canonical correlation analysis of the tropical Pacific wind stress and sea surface temperature, *Climate Dynamics*, **19,** 713-722 **(2002)**

65. Wu A., Hsieh W.W. and Tang B., Neural network forecasts of the tropical Pacific sea surface temperatures, *Neural Networks*, https://doi: 10.1016/j.neunet.2006.01.004, **19(2)**, 145–154 **(2006)**

66. Wyrtki K., El Niño-the dynamic response of the equatorial Pacific Ocean to atmospheric forcing, *J. Phys. Oceanography*, **5**, 572-584 **(1975)**

67. Xie H., Tang H. and Liao Y.H., Time Series Prediction Based on NARX Neural Networks: An Advanced Approach, Paper presented at the 2009 International Conference on Machine Learning and Cybernetics, 12–15 July 2009, IEEE, Baoding, China, 626 **(2009)**

68. Yu X., Efe M.O. and Kaynak O., A general backpropagation algorithm for feedforward neural networks learning, *IEEE Transactions on Neural Networks*, **13(1),** 251–254 **(2002)**.

(Received 27th January 2020, accepted 25th April 2020)