Superensembling of Artificial Neural Network Models for Investigating the Effect of Polar Sea Ice on Indian summer Monsoon Rainfall

Shruthy Govindan
PG Scholar, CSE Department, Inderprastha
Engineering College
Ghaziabad, India

K. C. Tripathi, PhD
Associate Professor, CSE Department,
Inderprastha Engineering College
Ghaziabad, India

ABSTRACT

Multimodel ensembling is now a popular prediction among researchers of dynamical meteorology. The ensembling of model outputs obtained from more than one initializations reduces the errors in forecasting due to wrong initializations. Multi model ensembling reduces the errors on account of poor understanding of physical phenomenon. Apart from dynamical forecasting, statistical forecasting contributes significantly to the research of this area. It becomes a pertinent question if the ensembling reduces this error in the same light as the dynamical ensembling. To address this question we have considered Artificial Neural Network models which has become a popular statistical model among the researchers of this community. We define the Superensemble model as the model which is a model based on ensembles of many ANN models. The ensemble ANN model is a simple ensemble of more than one ANNs with different initializations. Various approaches to superensembling have been discussed. The superensemble ANN models have been used to study the impact of Antarctic Sea Ice Concentration on the Sea Surface Temperature in an Indian Ocean region. The superensemble forecast is compared with normal and ensemble forecast. The results give a mixed interpretation.

Keywords

Super ensemble neural networks, sea ice concentration, correlation, Indian summer monsoon rainfall

1. INTRODUCTION

Sea ice concentration is both an indicator and driver of high latitude climatic change with strong social and economic importance. Sea ice is simply frozen ocean water (D. N. Thomas and G. S. Dieckmann,2009). It forms, grows and melts in the ocean. Although it may not directly affect us, it is a key component of our planet because it influences climate and wildlife. Sea ice regulates the exchange between Earth's surface and the atmosphere, or the Earth's energy balance (Tripathi and Das, 2008). It is a key boundary condition for atmospheric models (Jaiser et.al., 2011). Polar sea ice affects the global climate variability through cryosphere-albedo feedback over a wide range of time scales Connections among observed climate variability, such as the Southern Oscillation, sea ice cover and Antarctic precipitation have been observed (Aagaard and Carmach, 2012). Several studies have been done to investigate the effect of the polar sea ice parameters such as extent and concentrations on the local and global atmospheric and oceanic phenomena.

India being primarily agricultural economy depends largely on agricultural yields. Further the farmers mostly rely on rain water for their crops. Hence advance prediction of

precipitation over the Indian Ocean region plays a crucial role in the economic development of Indian and neighboring countries. Indian summer monsoon, which is a part of the Asian monsoon system, is a regular annual phenomenon which brings heavy rainfall to India and adjacent countries during summer monsoon season (June to September; JJAS). It contributes about 70% - 90% of rainfall in most parts of country. Prediction of Rainfall is necessary for planning and management of water resources for the social and economic benefits as well as mitigation of loss of lives and property.

Investigation of long term impact of antarctic sea ice on Indian Summer Monsoon Rainfall (ISMR) is important, not only because of the potential importance of sea ice in global climate, but also for the practical purpose of exploring the Antarctic continent. Several studies have been done to investigate the effect of the Antarctic sea ice parameters such as Antarctic Sea Ice Extent and Antarctic Sea Ice Concentrations (AnSIC) on the local and global atmospheric and oceanic phenomena .The present author (Tripathi and Das, 2008) has demonstrated that AnSIC is a prospective predictor of southern Indian Ocean SST indices and also of All India Rainfall (Tripathi and Das, 2009).

Recently it has been found by Tripathi et.al., 2008 that Antarctic sea ice concentration has a possible relationship with the sea surface temperature. In our study, We investigated the possibility of prediction of effect of polar sea ice on indian summer monsoon rainfall in the Indian ocean using superensemble technique. We used super ensemble technique to predict the effect of polar sea ice on sea surface temperature in the Indian Ocean by training network with different random initializations like by varying number of hidden neurons in the hidden layer.

1.1 Artificial neural network and its role in forecasting

Artificial Neural Networks are statistical models which discover the underlying deterministic behavior of the dynamical system. The ANN model has found significant applications in meteorology and oceanography. ANNs provide a methodology for solving many types of non-linear problems that are difficult to solve by the traditional techniques. With ANNs, there exists the capability to extract the relationship between the inputs and outputs of a process, without the physics being explicitly provided . Thus, these properties of ANNs are well suited to the problem of weather forecasting under consideration.

Statistical prediction models such as the Artificial Neural Networks (ANN) tend to model the dynamical behavior of the system based on past information. Recently, ANN has been

used to investigate various aspects of Antarctic Ocean such as to predict 6-hourly AWS surface data (temperature, pressure) in West Antarctic (Reusch and Alley, 2002), Acoustic surveys for biomass estimation (Woodd-Walker et al., 2003), synoptic reconstructions of West Antarctic climate (Reusch et al., 2005), to evaluate the role of southern hemispheric tropical and polar influences during global climatic cycles and events (Crundwell et al., 2008), comparison of the accuracy of SST estimates from different techniques in the Antarctic ocean (Gupta and Malmgren, 2009), teleconnection between Antarctic Sea Ice Extent and Indian Summer Monsoon Rainfall (Prabhu et al. 2009).

1.2 Ensemble forecasting

Neural network ensemble is a learning technique where many neural networks are jointly used to solve a problem. Neural network ensemble helps in improving generalization and also removes the error related to initial conditions. The output of the neural network ensemble is considered to have high correlation and low root mean square error. Ensembles represent a natural extension of common neural network practice that undoubtedly owes its roots to earlier works using conventional forecasting tools. Generally to create such an ensemble, several networks are trained out of which best networks are chosen to create an ensemble. Construction of neural network ensemble involves two steps: training the individual networks and then combining their predictions. Individual networks can be trained using different ensemble techniques and these techniques are based on varying parameters related to design and training of ANN such as varying initial weights, varying network type, varying network architecture involving number of hidden neurons, number of hidden layer. To combine the networks, ensemble mean or weighted mean method is used so as each ensemble is assigned weights in such a way that root mean square error is minimized. Network is trained with different random initializations and 4 such models are selected for participating in the ensemble forecasting whose errors are better. The average output of the above 4 models is expected to give better results. We have modified the learning parameters of the network where a number of networks are built with different learning parameters, such as initial weights in an MLP and varying number of hidden neurons, etc.

An alternate baseline approach we investigated was the creation of a simple neural network ensemble where each network are trained independently and networks with good correlation are selected then ensemble mean of these networks are taken. Likewise hidden neurons are varied and then again same procedure applied .While theoretical results indicate that, if properly constructed and employed, neural network ensembles can generalize better than any individual model used separately, they do not provide general guidelines about the selection of different models in the ensemble.

1.3 Superensemble neural network

ANN superensembling is an extension of ensemble method which further improves the performance of a network by providing high correlation and low root mean square error. Several ensemble models are combined together to form a superensemble. We have used two aspect to create a superensemble that is the initial weights and the architecture of the ANN which essentially means the number of neurons in the hidden layer. Our baseline approach was the creation of a simple superensemble network using two ways first where output of various ensemble networks are combined during training phase to form a superensemble and second where mean of various ensembles are taken. We have also created a superensemble network where weighted mean concept is utilized and comparison between these results are done to see which network performs better .It has been found that weighted mean concept performs better than ensemble mean and gives a significant increase in correlation and root mean square error is decreased. For weighted concept we used network with 8 hidden neurons as it performed better then various weighted mean are statistically combined during training so that skill of ensemble of ensemble is factored into superensemble. In ANN ensemble, output of weighted ensemble mean is fed into the network as an input then the network is trained so that the output obtained gives better performance producing good correlation and root mean square error is reduced and then multiple linear regression is applied to establish a quantitative relationship between a group of predictor variables and output . We also created a ANN ensemble where output of various created ensembles is fed into the network as a input then the network is trained so that the provided output enhances the performance and gives better results in comparison to superensemble that we obtain by simply averaging all the various created ensembles.

The simplest and most common combination of the ensemble members is an ensemble mean (EM), which is calculated as a simple average of ensemble members:

$$EM = 1/N \sum_{i=1}^{N} Pi$$

where N is the total number of ensemble members and Pi is the ith ensemble member generated by the model number i.

For weighted ensemble mean (WEM) it is simple average of the random weights assigned to the networks :

$$WEM = 1/N \sum_{i=1}^{N} Wi$$

here each ensemble member is subscribed a weight, Wi, randomly.

1.4 ANN Ensemble

In ANN ensemble, output of weighted ensemble mean is fed into the network as an input then the network is trained so that the output obtained gives better performance producing good correlation and root mean square error is reduced . We also created a ANN ensemble where output of various created ensembles is fed into the network as a input then the network is trained so that the provided output enhances the performance and gives better results in comparison to superensemble that we obtain by simply averaging all the various created ensembles.

2. METHODOLOGY

2.1 Data

The monthly AnSIC data for 23 years (January 1982 to December 2004) downloaded from http://www.cdc.noaa.gov has been used. For the sea surface temperature, The extended reconstructed sea surface temperature from January 1871 to May 2004 for the IO region (in the neighborhood of IOD) has been used for the present analysis. We have used data for 22 years that is from January 1982 to December 2003 to analyze the effect of polar sea ice on SST which in turn has a very strong correlation with rainfall and thus can be used as a predictor for rainfall.

2.2 Correlation Analysis

The correlation analysis has been done between the time series of AnSIC and the ERSST, the ERSST time series being 'ahead' by 1 to 24 months. It has been found that the AnSIC has a very strong correlation (nearly 0.62) with the SST when

the former lags behind by 1 month. Hence, the AnSIC is used as the predictor for the Sea Surface Temperature in the region one month ahead. The anomalies of SST are calculated. The determination of predictors is an important step as this is a precursor to a good prediction model. Cross-correlation analysis was done for the determination of predictors and also to establish the basis for an attempt to design such a prediction model. Cross correlation value of a AnSIC series with a lagged series of ERSST gives an insight of the level of dependence of the future values in the series on the present value.

2.3 Normalization

The predictor and the predictand series has been normalized using the following scheme

$$XN=[(X-XMIN)/(XMAX-XMIN)]*0.6+0.2$$

Where X is the sea ice concentration, XMAX is the maximum concentration for the month, XMIN is the minimum concentration for the month, XN is the normalized concentration for the month. The factors of 0.6 and 0.2 are included so that the normalized lies within 0.8.

2.4 Partitioning of data

The data set comprises of 22-year monthly data (from January 1982 to December 2003). This implies 264 data points. Since we are predicting 1 months in advance, there are effectively 263 input-output relationships. The normalized data is partitioned into three datasets namely training, testing and validation set. The data set comprises of 22-year monthly data (from January 1982 to December 2003). This implies 264 data points. Since we are predicting 1 month in advance, there are effectively 263 input-output relationships. The test set is obtained by serially taking the last 22 points. The remaining 197 points are used for training and 44 points for validation.

2.5 Architecture of ANN Ensemble

Multilayer feed-forward neural network models are the most popular network paradigm for forecasting applications. The factors related to neural network model architecture include the number of input variables, the number of hidden layers and hidden nodes, the number of output nodes, the activation functions for hidden and output nodes, and the training algorithm and process. The activation functions used for all hidden nodes are the logistic function while the identity function is employed in the output layer.

The number of input nodes corresponds to the number of past lagged observations used to capture the underlying pattern here which is lag1. We used four levels of hidden nodes of 5, 6 ,7 and 8 to experiment within this study. In ensembling of models while prediction with dynamical approach the following procedure is adopted.

Ensemble of a single model is taken with several initial conditions. This takes care of the error introduced due to wrong initializations

Ensembles of various models obtained in step 1 are taken and second level of ensembling done .This takes care of error due to poor understanding of the model.

In context of ANN ensembling the step (i) corresponds to taking the ensemble forecast with fixed number of hidden neurons but with different initial weights. The step (ii) corresponds to varying the number of hidden neurons and repeating step (i). The step (i) takes into account the false minimum of error due to small perturbation that may have creeped in while feeding the input to ANN. Taking several initializations would cancel the effect of every wrong initial choice made or at least we may hope so. The second step ensures that the fractional degree of non-linearity that may be present in the actual dynamical system and which cannot be modeled with initial number of hidden layer neurons. Hence the two steps take care of wrong initialization and poor understanding of the problem. We call such ensembling as superensembling.

The design of various models proceeds by initially making an a ANN model with 5 hidden neurons. The RMSE and correlation of the response with desired response is obtained. Keeping the number of hidden neurons fixed, training is repeated with different initial weights. Best model based on correlation co-efficient and RMSE is retained. This is individual best model with 5 hidden layer neurons and is called ANN5. The process is repeated with 6 neurons and the best model is termed as ANN6. Similarly we obtain ANN7 and ANN8.

Apart from best individual model we retain three other better models with the same number of hidden layer neurons. Ensembling of all the four models is then done by averaging their outputs. The resulting model is termed as EANN5. Similarly we obtain EANN6, EANN7 and EANN8. Apart from averaging the outputs, ensemble models are created by taking the average of weights of better models. E.g. four better ANN models with 5 hidden layer neurons are taken and the corresponding weights averaged. This gives WANN5. Similarly we obtain WANN6, WANN7 and WANN8. Thus we have eight simple ensembles of ANNs viz EANN5, EANN6, EANN7, EANN8, WANN5, WANN6, WANN7 and WANN8. Using these simple ensembles we design superensemble models.

We have used four approaches to superensembling as follows:

By taking simple average of ensemble forecasts (called S_1)

By averaging the weights of various ensemble ANN models (EANN obtained in step i. The missing weights are set to 0 before taking the average of weights. This mode of superensembling is termed here as S_2.

The outputs of ensemble members of S_1 are fed as input to a second layer ANN keeping the targets unchanged. This is called S_3.

The outputs of ensemble members of S_2 are fed as input to a second layer ANN keeping the targets unchanged. This is called S_4.

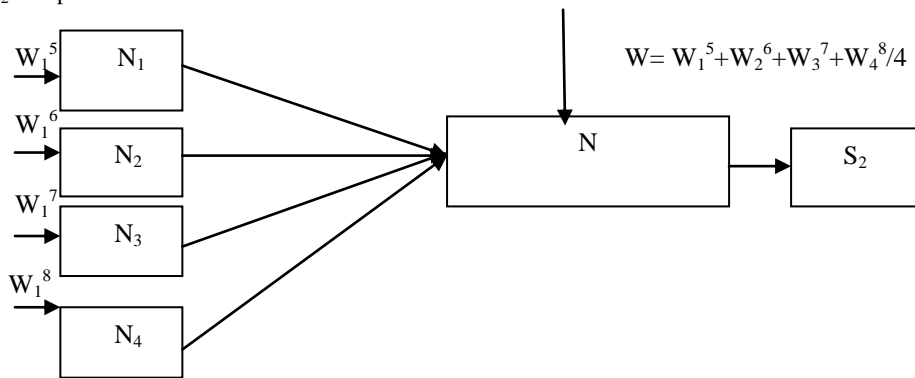a) E_i^j: i means the ensemble number and j specifies number of hidden neurons

S$_1$: Output

$$E= (E_1{}^5+ E_2{}^6+ E_3{}^7+ E_4{}^8)/4$$

$$S_1$$
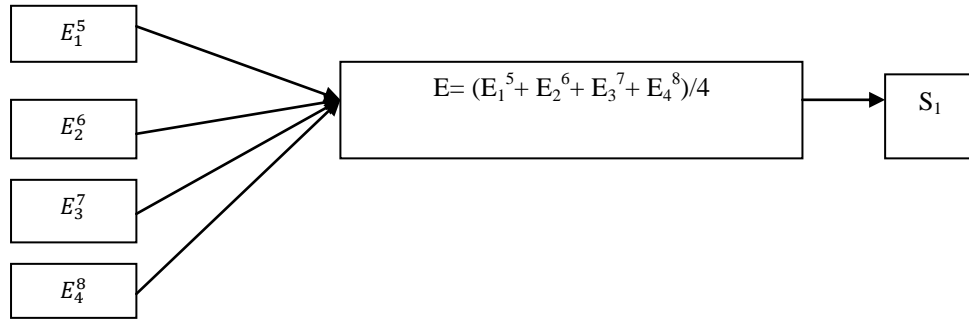
b) $\mathbf{W_i^j}$: W Specify weights assigned to networks, i means the ensemble number and j specifies number of hidden neurons (bold face specifies a vector quantity)
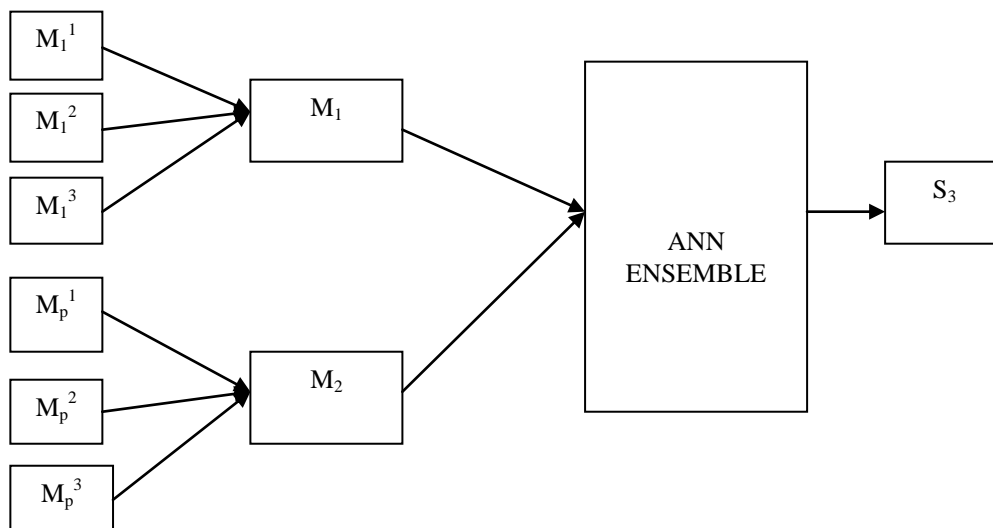
N_k: means the neural networks

S$_2$: Output

$W_1{}^5$ N_1

$W_1{}^6$ N_2

$W_1{}^7$ N_3

$W_1{}^8$ N_4

$$W= W_1{}^5+W_2{}^6+W_3{}^7+W_4{}^8/4$$

N

$$S_2$$

c) M_i^j: i means the number of hidden neurons and j means the number of initialization

$M_i= (M_i{}^1+M_i{}^2+M_i{}^3+M_i{}^4)/4$

Superensemble $M= M_1+M_2+M_3+M_4 = 1/4\sum_{i=1}{}^4M_i$

S$_3$:Output

$M_1{}^1$

$M_1{}^2$ M_1

$M_1{}^3$

$M_p{}^1$

$M_p{}^2$ M_2

$M_p{}^3$

ANN ENSEMBLE

$$S_3$$

d) N_i^j : i means the number of hidden neurons and j means the number of initialization

$N_i= (N_i{}^1+N_i{}^2+N_i{}^3+N_i{}^4)/4$

Superensemble $N= N_1+N_2+N_3+N_4 = 1/4\sum_{i=1}{}^4N_i$

$\mathbf{W_i^j}$: W Specify weights assigned to networks, i means the ensemble number and j specifies number of hidden neurons (bold face specifies a vector quantity)

S$_4$: Output

W_i : W Specify weights assigned to networks, i means the ensemble number (bold face specifies a vector quantity)
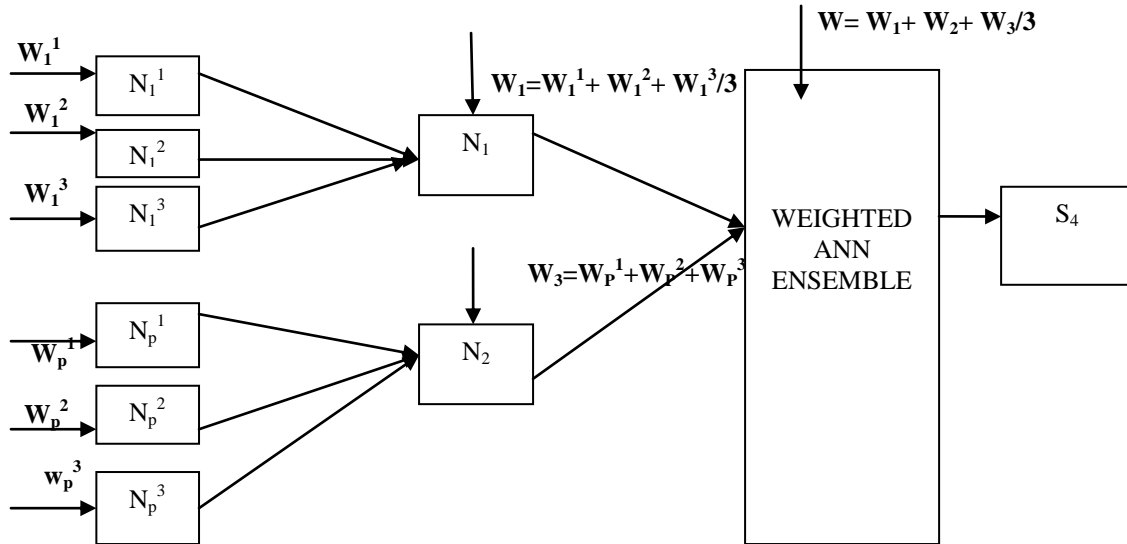

Fig.1 (a-d) shows the architecture of all the four steps

3. RESULTS AND DISCUSSIONS

Results presented below shows the performance measures of various superensemble models. The mode of superensembling is mentioned in short. We have also compared the performance of superensembling with simple ensemble forecasts as well as individual models. We have used correlation coefficient and RMS errors as the criteria for performance evaluation. The best results under all the categories are highlighted in red.

It can be seen that the best model is the Weighted ANN ensemble model. This is demonstrated by best in class correlation (0.75) and RMS error (0.0713). The best ensemble model, in the two categories of simple ensembling, also turns out to be the same with the same performance measures. However, the results show no significant improvement over the individual model ANN. Hence simple ensembling does not lead to significant improvement in the performance.

Looking at the superensembling results we see that the ensemble method do not lead to any significant improvement in performance.

Thus the results obtained indicate that the ANN ensemble gives better performance than its counterparts and can be effectively used in predicting the effect of polar sea ice on rainfall. However using the method of training ANN with outputs of ensemble ANN gives significantly better results (correlation coefficient=0.85 and RMSE=0.077).

Table 1. The performance measurement of ensemble averaging (superensemble forecast) using correlation coefficient and root mean square error and training ,testing and validation results are presented for these measurements.

	TRAIN	TEST	VALID
CC	0.7924	0.7290	0.8598
RMSE	0.0738	0.0845	0.0685

Table 2. The performance measurement of weighted ensemble (superensemble forecast) using correlation coefficient and root mean square error and training ,testing and validation results are presented for these measurements.

	TRAIN	TEST	VALID
CC	0.7924	0.8477	0.7819
RMSE	0.0757	0.0783	0.0702

Table 3. The performance measurement of ANN Ensemble using correlation coefficient and root mean square error and training ,testing and validation results are presented for these measurements.

|	TRAIN	TEST	VALID
CC	0.81	0.8589	0.8143
RMSE	0.0738	0.0772	0.0702

Table 4. The performance measurement of Weighted ANN Ensemble using correlation coefficient and root mean square error with 9 hidden neurons in the layer and training ,testing and validation results are presented for these measurements.

|	TRAIN	TEST	VALID
CC	0.8068	0.7529	0.8540
RMSE	**0.0727**	**0.0713**	**0.0673**

The results thus obtained show that ANN ensemble performs better than simple ensemble averaging and weighted ensemble mean better than ensemble averaging and Weighted ANN ensemble even performs better than ANN ensemble .
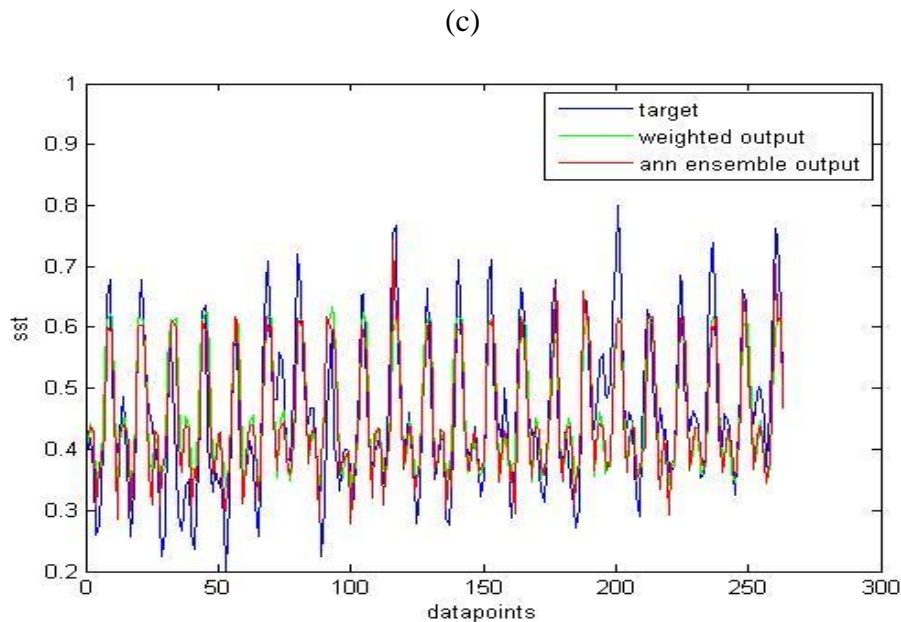
ANN ensemble shows better correlation and root mean square error is further minimized when we applied curve fitting techniques to the ANN ensemble output obtained.

(a)

(b)

(c)
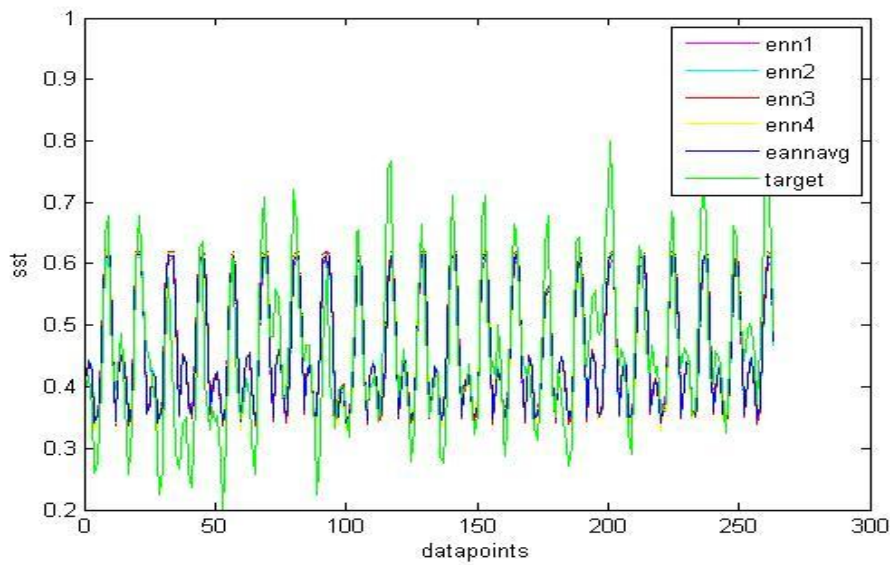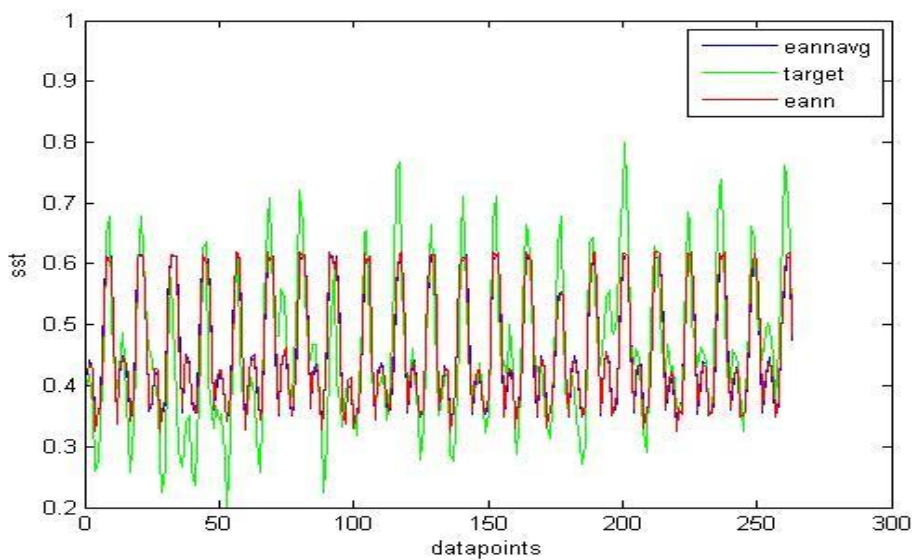
Figure 2. Comparison of various ensemble techniques (a) ensemble averaging output compared with target and individual ensembles output (b) ANN ensemble compared with target and ensemble averaging output (c) Weighted ensemble mean output compared with target and Weighted ANN ensemble

4. CONCLUSION

Antarctic Sea Ice Concentration, AnSIC, is a factor which significantly affects the amount of precipitation that the Indian subcontinent receives. The impact of AnSIC on the Indian monsoon is required to be investigated. This is achieved by mathematical modeling of the processes that connects the two. However, numerical prediction models require huge computational resources and manpower. This incurs a huge cost in such modeling. Statistical modeling is used since long to achieve the objective of the numerical weather prediction models at a low cost. After the advancements in the research of expert systems and other computational models such as the ANN, statistical prediction and modeling has gained significant attention of the researchers. The quality of forecasts by statistical and dynamical methods has been found to be of almost same accuracy. This is because both methods model the underlying deterministic laws that govern the actual system.

There are two sources of errors in prediction by dynamical systems. Firstly due to wrong initialization of the model and secondly due to wrong equations being incorporated in the model. Different researchers employ different equations for the modeling of same phenomena depending upon their understanding of the problem. Both types of errors are bound to creep in. The error due to wrong initialization cannot be eliminated because of finite capacity of recording instruments. Weather being a chaotic system deviates from that of the model fast because of this error. The error due to wrong equations cannot be eliminated because weather is not a deterministic system as of now. Hence these errors are bound to be present.

During the last decade researchers of the dynamical modeling community have devised the technique of Multi Model Ensemble, MME, to get reduce the errors due to the two sources. It is ensemble of ensembles of individual models initialized with slightly different initialization conditions. It has been found that MME forecast is better than individual forecasts. Since statistical and dynamical modeling work on

the same principle of modeling the underlying deterministic system and so far the quality of forecasts with the two types of models has been found to be similar, we are tempted to put up a question if the concept of MME is applicable in statistical modeling as well. If yes, what is the principle and how does the quality of forecasts compare. We have taken the ANN model to be our statistical model for the analysis. We call the two layer of ensemble of models a superensemble ANN model.

As a moderate first step we first wish to investigate if such a system would give better forecast than the individual model forecast. In the present study we have investigated the predictability of Indian summer monsoon rainfall using the AnSIC as predictor in order to see the effect of AnSIC on the Indian climate. Owing to the importance of MME forecast, we have designed a superensemble ANN architecture to carry out the predictability studies. Three approaches to superensembling were outlined and discussed. The results show no difference between the first two. Further, the results show no significant improvement over best individual forecast. However, the third approach to superensembling gives encouraging results. It was found that the forecast obtained by connecting two separate ANN layers is better than the individual forecast, their ensemble forecast and their superensemble forecast when first two approaches to superensembling were used. We have used the correlation coefficient and RMS error as the measures of quality of forecasts.

Although the improvements were slight but since they match with the improvements obtained by MME forecast, it can be prima facie concluded that superensembling of ANN models bears resemblance to the idea of MME forecast used by numerical models. However, further research is needed to consolidate our findings. We plan to carry out such research in future for time series forecasting of various meteorological parameters.

5. REFERENCES

[1] Aggarwal,Rashi., Kumar,Rajendra. (2013) 'A comprehensive review of Numerical Weather Prediction Models', *International Journal of Computer Applications*,Vol.47

[2] Basawaraj,Gadgay., Kulkarni, S., Chandrasekhar,B. (2012) ' Novel Ensemble Neural Network Models for better Prediction using Variable Input Approach', *International Journal of Computer Applications*,Vol.39-no.18.

[3] Mekanik, F. and Imteaz, M.A, (2012) 'A Multivariate Artificial Neural Network Approach for Rainfall Forecasting: Case Study of Victoria, Australia', *Proceedings of the World Congress on Engineering and Computer Science* Vol I,WCES .

[4] Gheyas, I. A. and Smith, L. S.(2012) 'A novel neural network ensemble architecture for time series forecasting', Neurocomputing, doi:10.1016/j.neucom.2011.08.005 .

[5] Shrivastava,Gyanesh., Karmakar,Sanjeev., Kowar, M.K., Guhathakurta, Pulak. (2012) 'Application of Artificial Neural Networks in Weather Forecasting: A Comprehensive Literature Review', *International Journal of Computer Applications* ,Vol 51– No.18.

[6] Hansen, L.K., Salamon, P.(1990) ' Neural network ensembles ', *IEEE Trans Pattern Anal* 12(10):993-1001.

[7] Hsieh, W.W.& Tang, B.(1998) 'Applying neural networks model to prediction and data analysis in meteorology and oceanography', *Bull. Amer. Meteorol Soc.*,79,pp.1855-1870.

[8] Maqsood,Imran., Khan,M.R. , Abaraham, A. (2004) ,'An ensemble of neural networks for weather forecasting', *Springer-Verlag, Neural Comput & Applic* ,Vol. 13 pp.112-122.

[9] Tripathi,K.C., Agarwal, Rashi., Hrisheekesha, P.N. (2013) , 'Global Prediction algorithms and predictability of anomalous points in a time series', *International journal of computers & technology*, Vol.10, no 9.

[10] Krishnamurti ,T.N and Kishtawal C.M., Larow,T. ,et, al.,(1994) , ' Improved weather and seasonal forecasts from multi-model superensemble ',Science ,pp.1548-1550

[11] Lynch ,P.(2008) ,' The Origins of Computer Weather Prediction and Climatic Modelling'. *Jounal on Computational Physics* ', 227, pp.3431–3444.

[12] McCulloh , W.S. & Pitts ,W.(1943), 'A logical calculus of the ideas immanent in neural nets', Bull Math. Biophys.,5(1943) ,pp.133-137

[13] Palmer, T.N. and others,(2004) ,' Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER)' ,*Bulletin of American Meteorological Society*,85,pp.853-872.

[14] Rai, S., Pandey , A. C. , Tripathi, K. C. and Dwivedi ,S. (2008), ' Predictive Skill of DEMETER models for wind prediction near Madagascar', *Indian Journal of Marine Sciences*, 37.

[15] Sahai, A. K. , Chattopadhyay ,R. and Goswami , B. N. (2008), 'A SST based large multi-model ensemble forecasting system for Indian summer monsoon rainfall', *Geophysical Research Letters*, 35, L19705, doi:10.1029/2008GL035461.

[16] Rai, Shailendra., Khare, N. and Pandey, A.C. (2008), 'Antarctica sea ice variability and southeast Indian Ocean SST: possible relationship', *Indian Journal of Marine Sciences*,Vol.37(1),pp.35-39.

[17] Bhowmik,S. K. R., Durai ,V.R(2012), 'Development of multimodel ensemble based district level medium range rainfall', *Indian Academy of Science*, 121, no. 2 , pp. 273-285.

[18] Yun, W.T., Stefanova ,L., Mitra, A.K., Dewar,W. and Krishnamurti ,T.N. (2004), 'Multi-Model Synthetic superensemble Algorithm for seasonal climate prediction using DEMETER Forecasts', *Tellus*, DEMETER .

[19] Krasnopolsky ,V.M. and Lin,Ying. (2012), 'A Neural Network Nonlinear Multimodel Ensemble to Improve Precipitation Forecasts over Continental US', *Advances in Meteorology*, Article ID 649450, 11 pages, 2012. doi:10.1155/2012/649450.

[20] Zhou ZH, Wu J, Tang W (2002), 'Ensembling neural networks: many could be better than all', *Artificial Intelligence* 137(1-2), pp. 239-263.

[21] Zurada, J.M.(2002), 'Introduction to artificial neural systems', *Jaico Publishing House, Mumbai, India*, pp. 30-400.