PRECIPITATION ESTIMATION FROM RADAR AND RADIOMETRIC OBSERVATIONS FROM TRMM DATA USING ARTIFICIAL NEURAL NETWORKS

Rajesh Kumar [1], **I. M. L. Das** [1, 2], **A. K. Mittal** [1,2,3]
R. M. Gairola [4], **Samir Pokhrel** [5], **Vijay K. Agarwal** [6]

[1] M N Saha Centre of Space Studies, Institute of Interdisciplinary Studies, University of Allahabad, Allahabad 211 002, India, rky78@rediffmail.com
[2] Department of Physics, University of Allahabad, Allahabad 211 002, India, drimldas@yahoo.com
[3] "As (2) above, but Email":--- a_k_mittal@indiatimes.com
[4] Meteorology and Oceanography Group, Space Application Centre, Indian Space Research Organization, Ahmedabad 380 015, India, rmgairola@yahoo.com
[5] "As (4) above, but Email":--- pokhrelsamir@yahoo.com
[6] "As (4) above, but Email":---vkagarwal6@yahoo.com

ABSTRACT

Artificial Neural Network (ANN) technique has been used for the estimation of precipitation, mainly from passive and active microwave measurements from space. ANN has been used to estimate precipitation using TRMM Microwave Imager (TMI) onboard Tropical Rainfall Measuring Mission (TRMM) satellite. A precipitation algorithm designed to generate rainfall estimates using a combination of TMI and TRMM Precipitation Radar (PR) data has been developed. The inputs for the ANN are brightness temperatures (BT) from TMI and the output is PR-rainfall. The networks are trained and cross validated. Once the training is complete, the independent data sets (which were not included in the training) were used to test the performance of the network. Instantaneous precipitation estimates demonstrates correlations of around 0.82 to 0.97 with independent test data sets for the Indian regions. Multiple regression (MR) technique is also used to estimate precipitation using the same database and the results are compared with those obtained from ANN. The model developed can be used for the estimation of precipitation at high spatial and temporal resolutions on instantaneous basis.

1. INTRODUCTION

Rainfall is a highly discontinuous process both in space and time. The spaceborne measurement and monitoring of rainfall is a topic of major interest since it influences the global hydrological cycle and the nature of climate variability. Assessment of precipitation contributes to improved weather forecasting, in small and large spatial scales, and a study of global rainfall leads to better understanding of global climate variability. Quantitative assessment of precipitation is needed to improve understanding of the behavior of global energy and circulation patterns. Since about 70% of the earth is occupied by water, land based techniques for precipitation estimates (e.g. Rain gauges) are not sufficient for global rainfall estimation. Therefore a feasible method for estimating global rainfall is space based remote sensing. In the remote sensing scenario, an orbiting satellite records radiant energy at various wavelengths ranging from visible, infrared to microwave and various algorithms are used to estimate rainrates from the recorded radiant energies. During the last few years, various satellite missions have been dedicated to the measurements of global precipitation which employ some kinds of microwave radiometer like High resolution Microwave Multifrequency Radiometer (HMMR). Various algorithms have been developed that estimate rainfall from microwave data as reviewed by [1] and [2]. Several of them utilize BT values obtained by SSM/I instrument at various frequencies and polarizations, and attempt to solve the inverse problem. Recent research has shown that ANN techniques can be successfully used for the precipitation estimation from radiometric measurements [3]. ANN is a non-parametric method for representing the complex relationship between radiometric measurements and radar rainfall.

2. METHODOLOGY

a. Neural Networks

In recent years, ANN's have become extremely popular for prediction and forecasting as it provides a convenient and powerful means of performing nonlinear classification and regression [4]. It is composed of individual processing elements called units or nodes. The nodes are connected by links or weights. An ANN can have multiple layers of nodes interconnected with other nodes in the same or different layers. The layers are classified as input layers, hidden layers and output layers. The input nodes are connected by the output nodes through hidden nodes. The hidden nodes capture the non-linearity in the mapping between input and output information. The inputs are processed using interconnecting weights by weighted summation functions to produce a sum that is passed through a transfer function and the output of this transfer function is the output of the node. ANN learning can be supervised or unsupervised. In the supervised learning scheme, the ANN training is done by adjusting the input weights on each node in such a manner that the output of the network is consistent with the desired output. In unsupervised learning scheme, no external criteria are used by the network output for ANN training. The most popular supervised learning scheme is backpropagation. A description of ANN architectures, learning schemes and other specifics is given in [5].

b. Architecture and Inputs

The ANN architecture used in this study consists of an input layer, a hidden layer, and an output layer (as shown in Fig. 1).The four inputs of ANN are BT's corresponding to 10.7GHz (H), 19.4GHz (H), 21.3GHz (V) and 85.5GHz (H) channels of TMI and the output is PR rainfall. Vertical and horizontal channels corresponding to 10.7GHz, 19.4GHz and 85.5GHz are highly correlated (as correlation coefficient is nearly equal to 1) with each other as well as with rainrate, so only horizontal polarization corresponding to these channels are used. 37GHz channel is also not included in the present study because of low correlation with rain rate. The hidden layer has 16 nodes. For this study, TMI BT's (level 1B11), and PR rainfall (level 2A25) data are used. The transfer function used is sigmoid function. All the inputs and outputs have to be scaled in such a way as to be within the limits of the activation functions used in the output layer [6]. For example, as the outputs of the logistic transfer function are between 0 and 1, the data are scaled in the range 0.2 to 0.8. If the values are scaled to the extreme limits of the transfer function, the size of the weight updates is extremely small and flat spots in training are likely to occur. Here the database is generated from the collocation of TMI and PR, which share a common swath of about 220km on the surface. Two days collocated data base is used for the present study (13^{th} Jul. 1999 & 26^{th} Oct. 1999). There are about 12130 collocated TMI and PR data points. Here the rain-rate ranges from 00 to 56.7 mm/hr. This dynamic range of rainfall is quite sufficient for ANN training. Data before introducing to ANN is preprocessed to avoid the unwanted behavior of ANN [7]. Data points between 20mm/hr and 56.7mm/hr were small in number, so some additional copies of these data (3735 data points) were included in original data set so that ANN can pay equal attention towards the large values also. Once the preprocessing of the data is done, the original data is randomly divided into three parts training (9865), validation (3000) and testing (3000) [8]. About two-third of the data is used for the ANN training. Once the internal weights are fixed, the ANN performance is checked with the independent test data sets which were not included in the training.

3. RESULTS

Best result for the estimation is obtained with ANN training for 10000 iterations; learning rate parameter h ranges from 0.001 to 0.05 and the momentum parameter μ ranges from 0.01 to 0.03. The initial weights are randomized in the range [-1.0, +1.0]. The time evolution of the error during the training phase is shown in the Fig.2. From the Fig. 2, it is clear that the error (cost function) decreases substantially after few hundreds iterations and remains constant after about 4000 iterations, still more iterations were performed to ensure that the ANN had learned the large values as well as the small values. At this point the ANN is considered to be trained. In Fig.2, black line shows the absolute error with number of iterations for the training data and red line for validation data. Small part of the TRMM data set in the Northern Indian Ocean region during October 26, 1999, just 3 days before the Orissa super cyclone, have been taken for the training and test purposes, with the corresponding track nos. 10994, 11004 and 11005. The observed versus ANN retrieved rainfall for both training and test data sets are shown in Fig. 3 and 4. Significantly high correlations of 0.969 and 0.967 are achieved for both training and test data sets.

Multiple regression techniques for rainfall estimation have been studied by several workers [9] & [10]. We also performed multiple regressions using the above set of database. The correlation coefficient varies from 0.937 to 0.940 and RMSE from 5.708 to 5.596 for the same data sets used for the training and test. Based on these experiments, it is clear that the ANN performs better than the multiple regressions. The performance of the network was checked with other independent data sets for the Indian region whose error statistics along with training and test data sets are given in Table I Feed Forward Backpropagation Networks (FBBN) used for the present study are not only the network structures which may be used for this kind of problem e.g. [11] used a modified version of the counter propagation network. Therefore an improvement in the overall approach would include training the ANN to recognize the dependency of BT vectors on not only rainfall but variables such as water vapor, true surface emissivity, ice shape and size distribution, to name a few.

4. CONCLUSION

A combine TMI-PR precipitation estimation algorithm has been developed and validated against independent PR data. The ANN technique performed consistently better than MR in terms of correlations and RMSE. From this study it is clear that ANN model appears to be a useful tool for the rainfall estimation. But, one should be cautious in choosing the appropriate model inputs for such a complex problem, which reduces the size of the network and consequently reduces the training time and increases the generalization capability of the network for a given data set.

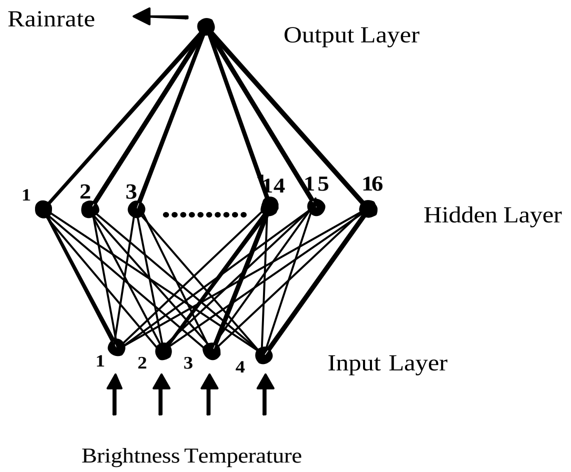

Fig. 1. A simple ANN architecture showing the input and output connections through the hidden units.

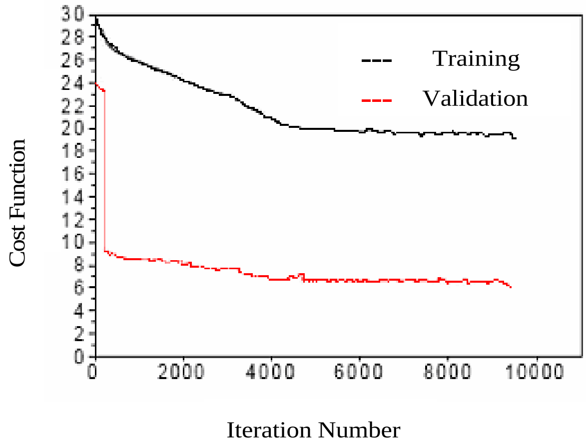

Fig. 2. ANN error distribution for training.

Table I. Error statistics for the different regions for trained ANN.

Statistics		Training Set	Test Set	Set I	Set II	Set III
Correlation Coefficient		0.969	0.967	0.825	0.826	0.925
Mean	Observed (mm/hr)	12.131	11.735	4.985	4.833	6.325
	Calculated (mm/hr)	12.049	11.823	5.411	5.292	4.060
S. D.	Observed	16.466	16.345	6.044	5.881	4.459
	Calculated	16.112	16.118	8.246	7.246	6.528
Bias(mm/hr)		-0.081	0.088	0.426	0.458	-2.264
RMSE(mm/hr)		4.023	4.201	4.737	4.114	3.717

Here, Set I – 13th Jul 1999 (9352), Set II – 26th Oct 1999 (11004), Set III – 26th Oct 1999 (11005)

ACKNOWLEDGEMENTS

The authors would like to thank SAC/ISRO for providing financial support through a sponsored research grant for this work. The data of TRMM used in the study from NASA/GSFC is thankfully acknowledged.

REFERENCES

[1] E. C. Barrett, J. Dodge, H. M. Goodman, J. Janowiak, C. Kidd, and E. A. Smith, "The first wet-net precipitation intercomparison project (PIP-1)", *Remote Sens. Rev.*, vol. 11, p.p. 49-60, 1994.
[2] E. E. Ebert, " Results of the 3rd algorithm intercomparison project (GPCP)", *Melbourne, Australia, BMRC Res. Rep. 55, Bureau Meteorol*., p. 199, 1996.
[3] D. Tsintikidis, J. L. Haferman, E. N. Anagnostou, W. F. Krajewski, and T.F. Smith, "A Neural Network Approach to Estimating Rainfall from Spaceborne Microwave Data", *IEEE Trans. Geosci. Remote Sensing*, vol. 35, No. 5, September 1997.
[4] J. A. Freeman, and D. M. Skpura, "*Neural Networks: Algorithms, Applications, and Programming Techniques. Addison-Wesley*, 401 pp.
[5] V. B. Rao, and H. V. Rao, "C++ *Neural Networks and fuzzy logic, New York: MIS*, ch. 5, pp. 67-92, 1993.
[6] Minns, A. W., and Hall, M. J., "Artificial neural networks as rainfall-runoff models", *Hydrological Sciences Journal*, vol. 41(3), pp. 399-417, 1996.
[7] Kaastra, I., and Boyd, M. S., "Forecasting futures trading volume using neural networks", *The Journal of Future*

Markets, vol. 15 (8), pp. 953-970, 1995.
[8] Burden, F. R., Brereton, R.G., and Walsh, P.T., “Cross-validatory selection of test and validation sets in multivariate calibration and neural networks as applied to spectroscopy”, *Analyst*, vol. 122(10), pp. 1015-1022, 1997.
[9] Bauer, P. and P. Schluessel, “Rainfall, total water, ice water and water vapor over sea from polarized microwave simulations and special sensor microwave/imager data”, *J. Geophys. Res.*, vol. 98, pp. 20,737, 1993.
[10] Ferraro, R. R. and G. F. Marks, The development of SSM/I rain-rate retrieval algorithms using ground-based radar measurements, *J. Atmos. Oceanici Technol.* Vol. 12, p.p. 755, 1995.
[11] Hsu, K., H. Gupta, X. Gao, and S. Sorooshian, “Estimation of physical variables from multichannel remotely sensed imagery using a neural network: Application of rainfall estimation. *Water Resources Res.*, vol. 35, 1605-1618, 1999.

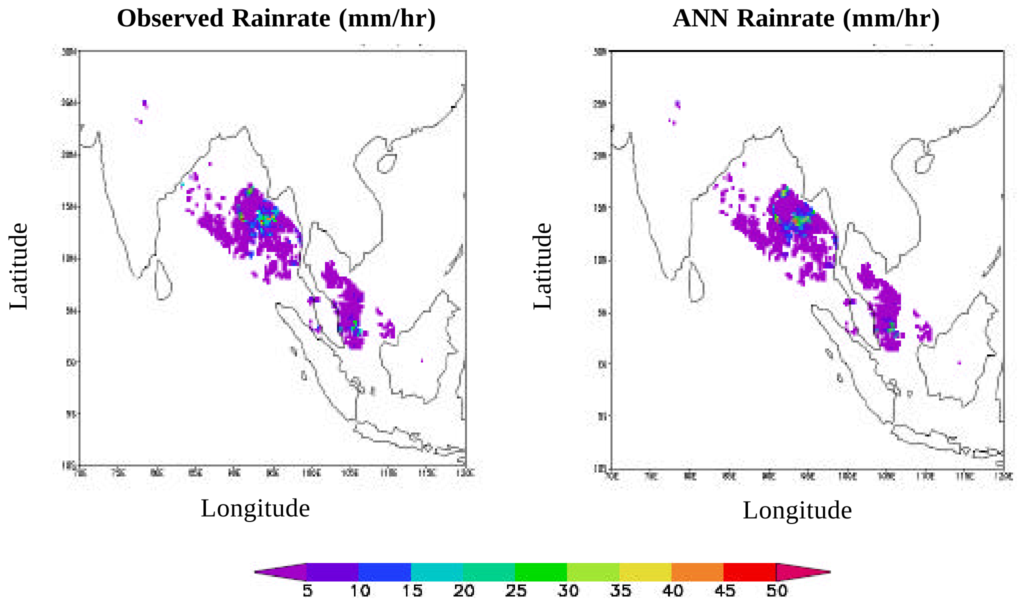

Fig. 3. Observed and ANN rain rate for the geographical area used for the network training.

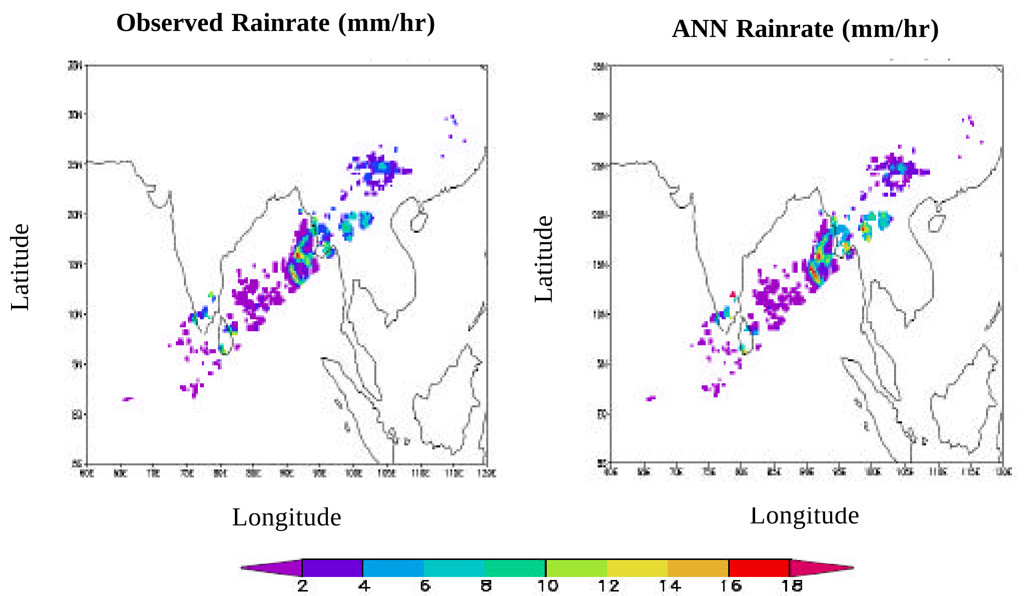

Fig. 4. Observed and ANN rain rate for the geographical area used for testing (Set I) the performance of the ANN.