MAUSAM, **64**, 2 (April 2013), 221-230

551.509.333

Downscaling of MM5 model output using artificial neural network over western Himalaya

PIYUSH JOSHI and A. GANJU

Snow and Avalanche Study Establishment, Research and Development Centre

Him Parisar, Sector 37-A Chandigarh – 160 036, India

(*Received 21 June 2011, Modified 14 December 2011*)

e mail : piyush.joshi@sase.drdo.in

सार – पूर्व की ओर जाने वाली सिनॉप्टिक मौसम प्रणाली अर्थात पश्चिमी विक्षोभ (डब्ल्यू.डी.) की वजह से पश्चिमी हिमालय में हिम के रूप में बहुत अधिक मात्रा में वर्षण होता है। एम. एम. 5, संख्यात्मक मौसम पूर्वानुमान (एन. डब्ल्यू. पी.) मॉडल ग्रिड में वर्षण पूर्वानुमान उपलब्ध कराता है। एन.डब्ल्यू.पी. आउटपुट की सांख्यिकीय डाउनस्केलिंग से पूर्वानुमान की सटीकता में सुधार आ सकता है। पश्चिमी हिमालय में एम. एम. 5 मॉडल पूर्वानुमान को स्थान विशिष्ट वर्षण पूर्वानुमान में डाउनस्केल करने के लिए इस अध्ययन में एक अरैखिक पद्धति, कृत्रिम तंत्रिका संजाल (ए. ए. एन.) का उपयोग किया गया है। इस अध्ययन के लिए वर्ष 2003 से 2008 तक के शीतकालीन महीनों (नवम्बर से मार्च) के आँकड़ें लिए गए हैं। प्रशिक्षण के लिए 2003 से 2007 के आँकड़ें और वैधता प्रयोजन के लिए वर्ष 2007–2008 के शीतकालीन आँकड़ें लिए गए हैं। प्रशिक्षण की प्रक्रिया में आगत–निर्गत संबंध का पता लगाया गया और अंतिम भार मैट्रिक्स का आकलन किया गया।

ABSTRACT. Western Himalaya receives enormous amount of precipitation in the form of snow due to eastward moving synoptic weather system called western disturbance (WD). MM5 a numerical weather prediction (NWP) model, provides precipitation forecast over a grid. Statistical downscaling of NWP output can improve forecast accuracy. In this study artificial neural network (ANN), a non linear method is used to downscale MM5 model forecast to location specific precipitation forecast over western Himalaya. Data of winter months (November to March) from 2003 to 2008 are considered for the study. Data from 2003 to 2007 is used for training and data of winter 2007-2008 is used for validation purpose. In the training process the input-output relationship is extracted and final weight matrix are computed.

Key words – MM5 model, ANN model, Western Himalaya, Skill score, RMSE.

1. Introduction

Most natural systems are non-linear. The atmosphere, ocean and climate systems are presumed to be extremely high dimensional, deterministic dissipative dynamical systems. Their state at a future time is a function of their state at a previous time. The atmosphere is a continuous fluid described by partial differential equations and its dimensionality is infinite. Discretization of the continuum, necessary for numerical modeling, renders the dimensionality finite, but any reasonable discretization results in a high dimensionality.

Generally two methods are used to forecast weather (a) the empirical approach (b) the dynamical approach (Maqsood, 2004). The first approach is based upon the occurrence of analogues and is often referred as analogue forecasting. This approach is useful for predicting local scale weather, when a large data base is available. The second approach is based upon equations and forward simulations of the atmosphere and referred to as computer modeling. The dynamical approach is useful for modeling large scale weather phenomenon and may not predict short term weather efficiently. Most weather forecast systems are a combination of empirical and dynamical techniques.

Northwest India has complex mountain ranges. This region receives high amount of precipitation, in the form of snow, during winter months (December, January, February and March - DJFM). Winter precipitation in western Himalaya is mainly attributed to the passage of weather systems called western disturbances (WDs). These are eastward moving low pressure synoptic weather systems which originate over the Mediterranean Sea or Mid Atlantic Ocean and travel eastward over Iran, Afghanistan, Pakistan and India (Dimri *et al*., 2008).

Snow and Avalanche Study Establishment (SASE) have been using the MM5 model for precipitation forecast

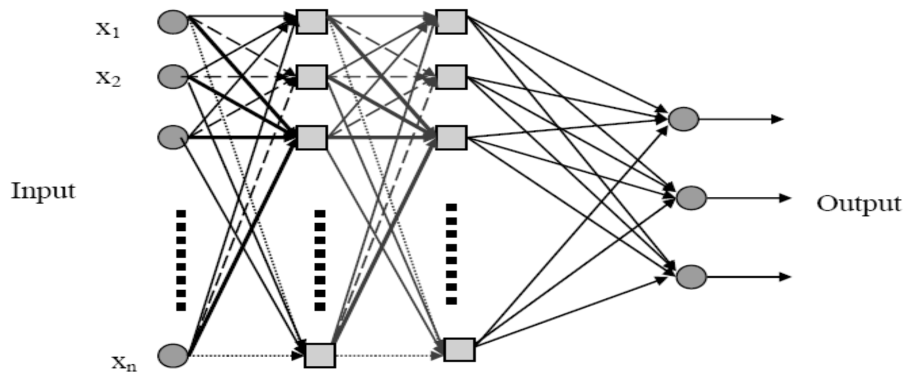

Fig. 1. Multi layered artificial neural network

at 10 km resolution over western Himalaya since 2002. The fifth generation Pennsylvania State University (PSU)/ National Centre for Atmospheric Research (NCAR) MM5 version 3.6 is a Limited Area Model (LAM) to simulate and predict mesoscale systems and regional atmospheric circulations (Anthes & Warner 1978). Its vertical coordinate system is terrain following sigma with options for non hydrostatics approximations. The complete MM5 modeling system consists of five modules: TERRAIN, REGRID, RAWINS/little_r, INTERPF and MM5. For the generation of terrain and land use, USGS 2' and 30" topography and land use data is used in MM5. Initial and lateral boundary conditions for the model are taken from National Centre for Medium Range Weather Forecasting (NCMRWF) T254 analysis and its forecast.

High resolution weather information from relatively coarse-resolution NWP models can be obtained by statistical downscaling of NWP model output. Statistical downscaling basically consists of development of statistical relationship between local variables (surface parameters) and large scale predictors (precipitation in present case) and application of such relationship to deduce local climate characteristics. Thus qualitative and quantitative location specific precipitation forecast can be improved after downscaling. In the present study Artificial Neural Network (ANN), a non-linear method is used for downscaling of MM5 model output to provide station specific precipitation forecast over western Himalaya.

Number of studies has been done using ANN approach in atmospheric science. Adya (1998) studied effectiveness of neural network in forecasting and prediction and found that ANN had potential for forecasting and prediction. Hall (1998) developed a neural network model using input from eta model and upper air soundings. The neural network produced a very good forecast of both the probability and amount of precipitation. Holger (2000) presented modeling issues of neural network models in the prediction of water resource variables. Naoya (2001) applied neural network method for the prediction of precipitation. Neural network was used to improve temperature forecasts produced by the Advanced Regional Prediction System and the improvement varied between 5% and 90% in terms of mean-squared error (Marzban, 2003). Coppola (2006) described an approach based on the application of ANN to a combination of satellite imageries and data from NWP model for real time rainfall estimation. Roebber (2006) used neural network for real time snow forecasting across the contiguous United States east of the Rockey Mountains. Roy Bhowmik *et al*., (2009) applied ANN to improve the forecast skill of maximum and minimum temperature over Delhi and found the method quite good for operational application.

2. Data and methodology

ANN is a computational structure modeled loosely on biological processes. Development of an ANN for any system involves topology of the network, a proper training algorithm and activation function [Chattopadhyay & Bandhyopadhyay (2007). In any ANN there is an input layer connected to output layer with different weights. In between there may be one or more hidden layers also connected by weights. This interconnected set of weights contains the knowledge generated by the ANN. Fig. 1 shows the architecture of a multi layered neural network.

In the present study a three layer network with one input layer, one hidden layer and one output layer is used to develop the model (Joshi and Dimri 2008). Back propagation learning algorithm is used to train the network. There are 11 nodes (number of input parameters)

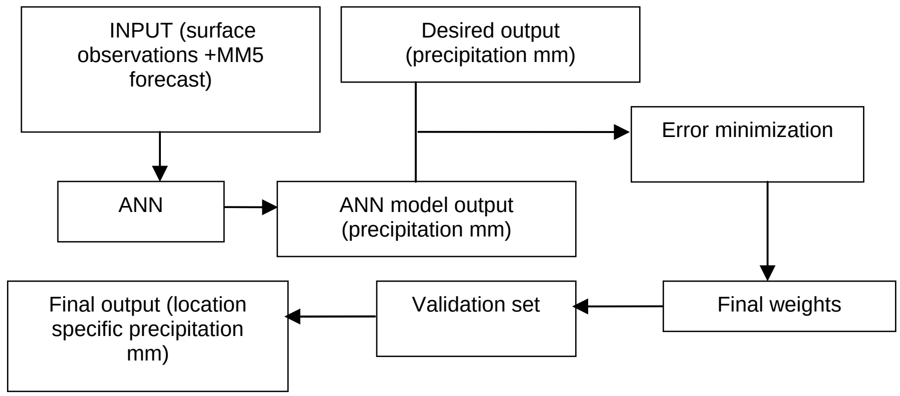

Fig. 2. ANN methodology in schematic form

in input layer, 7 nodes in hidden layer and one node (precipitation) in output layer. Table 1 shows network details of 11-7-1 network. The number of hidden layers is variable according to the problem. Generally one hidden layer is considered sufficient for most of the problems. Fig. 2 shows methodology in schematic form.

1.1. *Back propagation learning algorithm*

In ANN random weights are assigned between input-hidden and hidden output layers. If number of inputs, patterns and hidden nodes are denoted by n, m, and p respectively, the input at hidden layer is computed as:

$$Z_i = f\,(\sum W_{ij} * X_{kj}) \tag{1}$$

Final output O is computed with this input and hidden-out weights as;

$$O_k = f\,(\sum V_i * Z_i) \tag{2}$$

Where $i = 1, 2\text{-----}p$, $j = 1, 2, \text{------}n$, $k = 1, 2\text{-----}m$

f is a function called activation function. In this study the activation function is a sigmoid transfer function for both input-hidden layer and hidden output layer, given as:

$$f(x) = (1+e^{-x})^{-1} \tag{3}$$

Now the weights are adjusted in such a way that the error, the difference between the network output and desired output is reduced. The error is given by:

$$\text{Error} = 1/2\sum\,(\text{desired output-network output})^2$$

$$E = 1/2\sum\,(Y_k - O_k)^2 \tag{4}$$

TABLE 1

Three layer artificial neural network architecture

No of inputs	11
No of hidden layers	1
No of hidden nodes	7
No of outputs	1
Learning rate	0.3
Momentum	0.25
Activation function	Sigmoid transfer function $f(x) = (1+e^{-x})^{-1}$

The error term at output layer is computed as:

$$E_T = (Y_k - O_k) * O_k *(1 - O_k) \tag{5}$$

This error is propagated backward to the hidden layer and error term at hidden layer is computed as:

$$E_H = V_i * E_T * Z_i *(1 - Z_i) \tag{6}$$

These error terms are used for modification of the weights. The weights are adjusted so as to minimize the error functions. This technique is called gradient descent. Change in the weights is given as:

$$\Delta W_{ij} = \eta * E_H * X_{kj} + \alpha * \Delta W_{ij} \tag{7}$$

and,

$$\Delta V_i = \eta * E_T * Z_i + \alpha * \Delta V_i \tag{8}$$

Where η is the learning rate and α is the momentum.

TABLE 2

Principal observatories and data set considered for development and validation of ANN model

Station	Altitude (m)	Number of days (training set)	Number of days (validation set)
Kanzalwanwan	2440	432	136
Gulmarg	2800	432	136
Stage II	2650	435	136
Haddan Taj	3080	426	133
Banihal	3250	432	129
Pharkiyan	2960	432	136
Dras	3250	434	136

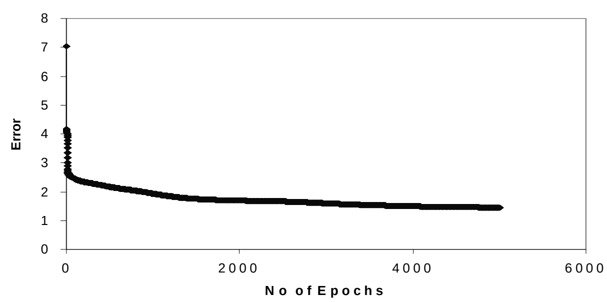

Fig. 3. Error minimization curve

These modified weights are used for the next pattern and error at each pattern is summed up to get the net error. The network is trained till the net error is minimized to a desired accuracy. Fig. 3 shows minimization of error with no. of epochs.

3. Study area and data set

Principal observatories of Snow and Avalanche Study Establishment (SASE), representing different road axes/sectors, [Kanzalwanwan (Bandipur-Gurez axis - BG), Haddantaj (Nogaon-Kaiyan axis - NK), Dras (Srinagar-Leh axis - SL), StageII (Chowkibal-Tangdhar axis - CT), Banihal (National Highway No. 1 A - NHIA), Pharkian (Keran sector), and Gulmarg] (Fig. 4) are considered for the study. The altitude of these locations varies from 2440 m to 3250 m (Table 2). These observatories represent the climatic and weather conditions of the region and are also nodal centers for assessing/disseminating avalanche hazards in time.

TABLE 3

Input parameters considered for the study

S. No.	Parameter	Time of observation
1	Maximum temperature (T_x)	1730 hrs (previous day)
2	Minimum temperature (T_n)	0830 hrs
3	Dry bulb temperature (T)	0830 hrs
4	Av wind speed (ws_{avg})	Average of last 24 hrs
5	Presssure	Change in last 24 hrs
6	Cloud amount (cla)	0830 hrs
7-11	MM5 output	Precipitation forecast for day 1 to day 5

Further, these sectors are considered as they represent different geographical and climatic conditions of the Jammu and Kashmir (J&K).

The data set consists of data of November, December, January, February and March months from 2003 to 2008. There are 432-435 data points for training set and 129-136 data points for validation set given in Table 2. Six surface parameters including maximum temperature, minimum temperature, dry bulb temperature, average wind, pressure change, cloud amount and MM5 output have been considered as input parameters for the study described in Table 3. Around 80% data is used to train the network and trained network is applied to the rest 20 % data for validation of the network.

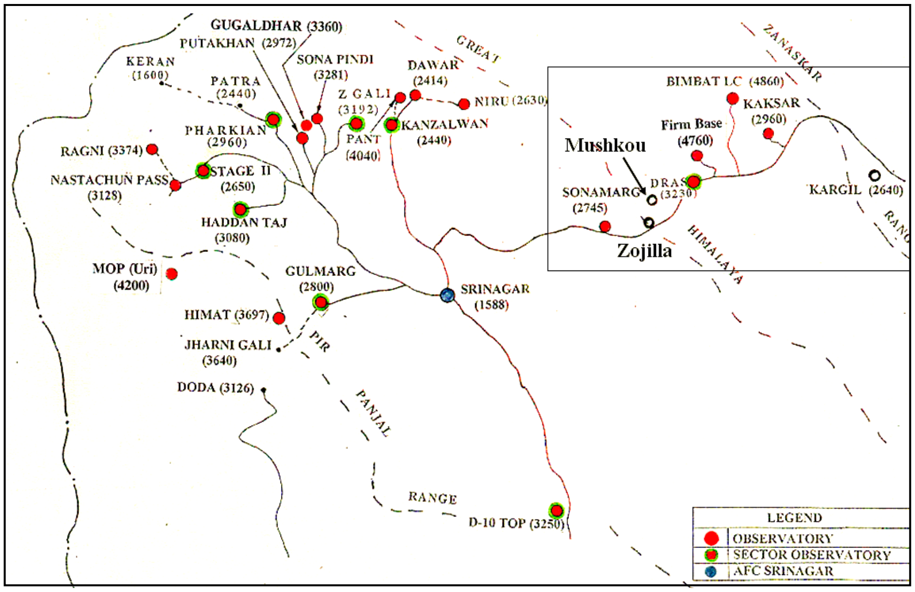

Fig. 4. Observatory network in J & K

TABLE 4

Verification measures for occurrence/non-occurrence of precipitation

S. No.	Station	POD		FAR		C-Non		CSI		BIAS		PC (%)	
		MM5	ANN	MM5	ANN	MM5	ANN	MM5	ANN	MM5	ANN	MM5	ANN
1	Kanzalwanwan	.78	.72	.32	.21	.81	.90	.57	.60	1.15	.91	80	84
2	Gulmarg	.44	.89	.62	.55	.82	.73	.26	.43	1.19	1.96	74	76
3	Stage II	.56	.67	.14	.30	.94	.80	.52	.52	.65	.95	78	74
4	Haddan Taj	.54	.93	.31	.46	.89	.64	.43	.51	.78	1.73	78	73
5	Banihal	.47	.67	.53	.58	.93	.88	.30	.34	1.00	1.60	88	85
6	Pharkiyan	.60	.81	.36	.40	.82	.72	.44	.53	.94	1.34	74	75
7	Dras	.47	.42	.33	.38	.91	.90	.38	.33	.71	.68	79	76

Parameters are normalized to values between 0.1 and 0.9 to ensure that the model has similar sensitivity to changes in various parameters. For a variable x with maximum value, x_{max} and minimum value, x_{min} the normalized value x_A is calculated as:

$$x_A = 0.1 + 0.8\,(x - x_{min}) / (x_{max} - x_{min}) \qquad (9)$$

4. Results and discussion

Six surface parameters and MM5 model output (precipitation) are considered for the study. Corresponding to these parameters ANN model is trained for the prediction of precipitation in next 24 hours. The precipitation amount recorded in 24 hours is in terms of water equivalent in mm. Precipitation is classified into five categories: one for no precipitation case *i.e.*, <= 2.5mm (no precipitation) and four for precipitation: 2.5 to 12.0 mm (cat I), 12.1 to 24.0 mm (cat II), 24.1 to 48.0 mm (cat III) and => 48.1 mm (cat IV). To verify the forecast probability of occurrence/non-occurrence of precipitation and probability of occurrence in certain category are assessed.

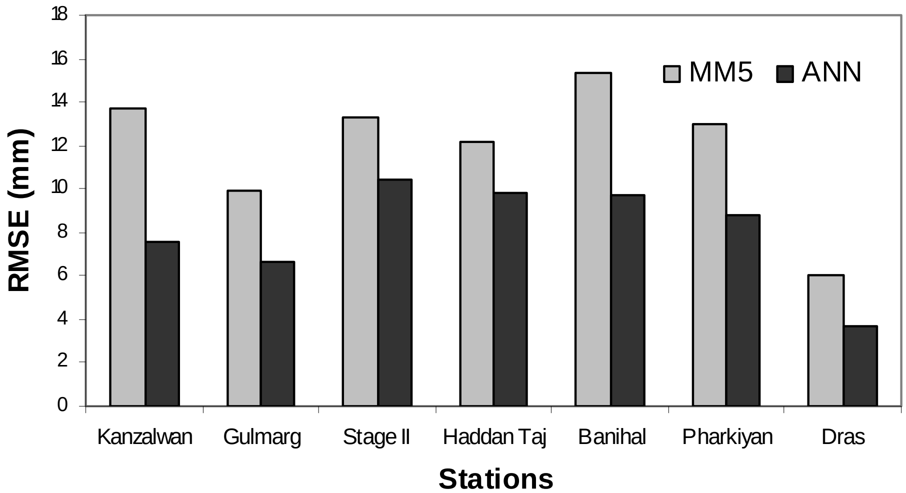

Fig. 5 (i). Root mean square error for training set

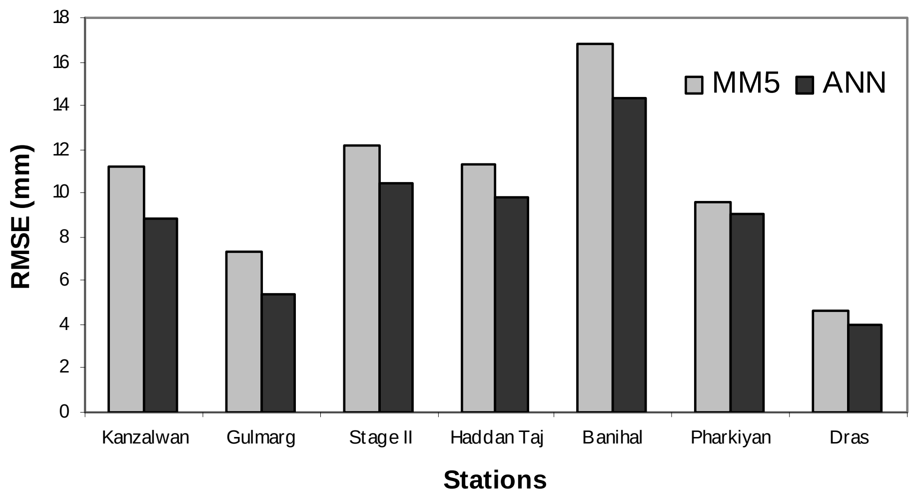

Fig. 5 (ii). Root mean square error for validation set

TABLE 5

Verification measures for quantitative categorical precipitation forecast

Measures		Kanzalwan		Gulmarg		Stage II		Haddan Taj		Banihal		Pharkiyan		Dras	
		MM5	ANN	MM5	ANN	MM5	ANN	MM5	ANN	MM5	ANN	MM5	ANN	MM5	ANN
CSI	Cat I	.52	.47	.60	.67	.42	.55	.56	.57	0	.17	.37	.56	.67	.69
	Cat II	.15	.21	.25	0	.08	.25	.15	.19	0	0	.07	.23	.50	0
	Cat III	.15	.33	0	0	.12	.38	0	.08	.50	0	.12	.08	0	0
	Cat IV	.17	.25	0	0	0	0	0	0	0	0	.25	0	0	0
HSS		.21	.31	.19	.07	.05	.29	.11	.18	.07	.11	.12	.20	.32	.12
PC (%)		47	52	58	67	41	58	44	51	14	27	36	50	67	56
% for non Occurrence Prediction		88	86	86	96	75	77	81	95	93	95	79	88	82	80

4.1. *Performance of probability of occurrence/non-occurrence of precipitation*

To verify the model for occurrence/non-occurrence case probability of detection (POD), false alarm rate (FAR), critical success index (CSI), model bias are computed for both MM5 and ANN model using 2 × 2 contingency tables (Wilks, 1995) (Appendix 'A') given in Table 4.

POD is fraction of those occasions when the forecast event occurred on which it was also forecast. The values of POD for MM5 are in the range 0.44 to 0.78. After downscaling values are in the range 0.42 to 0.93. There is improvement in POD after use of ANN except for 2 stations. But the difference is not much and values are acceptable as these values are for a particular location. Skill scores for non occurrence of precipitation (C-NON) is in the range 0.81 to 0.94 and 0.64 to 0.90 for MM5 and ANN model which are higher than POD indicating that the number of non-precipitating days is generally higher than number of days with precipitation.

FAR is proportion of forecast events that fail to materialize. The FAR is generally higher for MM5 Model than that for ANN. The CSI is also improved after downscaling.

Bias though not an accuracy measure, but tells about over forecasting and under forecasting. Model bias for ANN is slightly higher at some stations compared to MM5. This shows that after downscaling, model has a tendency to slightly over forecast the precipitation for a particular location.

The overall performance, measured by percent correct (PC) for MM5 and ANN model is 74-80% and 74-85% respectively. Thus, after applying ANN for downscaling of MM5 precipitation forecast, the forecast accuracy is quite good.

4.2. *Performance of probability of quantitative precipitation forecast*

To verify quantitative precipitation forecast root mean square error (RMSE) is calculated corresponding to training and test set for MM5 and ANN model respectively. RMSE varies from 6 to 15.38 and 4.62 to 12.15 for MM5 training set and test set respectively. For ANN training and test set these values are in the range from 3.73 to 9.76 and 3.94 to 10.43. Figs. 5(i&ii) show the results for training and test set for all stations. RMSE is reduced using ANN model for all the stations.

Quantitative precipitation in different categories is verified by calculating CSI and Heidke skill score (HSS) using 4 × 4 contingency table (Appendix 'B'). These skill scores for MM5 and ANN model are given in Table 5.

The CSI is the hit rate for the quantity being forecast, after removing correct 'no' forecasts from consideration. CSI is higher in category I as compared to other categories for all stations. This shows that the model could capture the precipitation amount in lower precipitation categories efficiently. Generally precipitation occurs in I category.

HSS ranges from 0.05 to 0.32 and 0.07 to 0.31 for MM5 and ANN model respectively for all stations. After downscaling, HSS is improved except at Gulmarg and Dras, where the value is slightly decreased. There are no negative scores for HSS showing that the forecast is better than the reference forecast.

The overall performance in different categories of precipitation varies from 14% to 67% for MM5 model and 27% to 67% for ANN model. Thus for all stations, forecast is improved after applying ANN model.

5. Conclusion

In present study ANN model for location specific precipitation forecast is developed using MM5 NWP model output and six surface meteorological parameters. Three layer feed forward network with a non-linear differentiable sigmoid transfer function is used. RMSE for quantitative forecast and different skill scores are computed for both MM5 and ANN models. Results show that after post processing of the MM5 model output RMSE is reduced significantly and skill scores are improved. There are some limitations in this technique as there is no definite rule to decide the number of hidden nodes, learning rate and momentum.

There are a large number of non-precipitating days for all stations, which makes the training data set biased towards no precipitation case. A large data set will give better results; however ANN can be used as an effective tool for downscaling of NWP model outputs. This improved location specific precipitation forecast would definitely be helpful for avalanche forecasting in western Himalayan region.

Besides precipitation, location specific forecast of temperature and other surface parameters are also very important. NWP models predict these parameters over a grid. To downscale NWP model forecast of these parameters at a station ANN model can be used. At present ANN is applied to generate only 24 hour precipitation forecast. It is possible to extend the forecast lead time up to 72 hours using ANN model.

Acknowledgements

We acknowledge all the technical assistance of Hazard Assessment and Forecasting Division (HAFD) for their sincere efforts in fieldwork.

References

Adya, M., 1998, "How effective are neural networks at forecasting and prediction? A review and evaluation", *Journal of Forecasting*, **17**, 481-495.

Anthes, R. A. and Warner, T. T., 1978, "Development of hydrodynamic models suitable for air pollution and other meso-meteorological studies", *Mon. Wea. Rev*., **106**, 1045-1078.

Chattopadhyay, S. and Bandyopadhyay, G., 2007, "Artificial neural network with back propagation learning to predict mean monthly total ozone in Arosa, Switzerland", *International Journal of Remote Sensing*, **28**, 4471-4482.

Coppola, E., 2006, "Validation of improved Taman neural network for operational satellite-derived rainfall estimation in Africa", *Journal of Applied meteorology and Climatology*, 1557-1572.

Dimri, A. P., Joshi, P. and Ganju, A., 2008, "Precipitation forecast over western Himalaya using k-nearest neighbour method", *International Journal of Climatology*, **28**, 14, 1921-1931.

Hall, T., 1998, "Precipitation forecasting using a neural network", *Weather and Forecasting*, **14**, 338-345.

Holger, R., 2000, "Neural networks for the prediction and forecasting of water resource variables : A review of modeling issues and applications", *Environment Modelling and Software*, **15**, 101-124.

Joshi, P. and Dimri, A. P., 2008, "Artificial neural network based precipitation forecast over the western Himalaya", Proceedings NSSW 2008, SASE Chandigarh., 2008, 251.

Marzban, C., 2003, "Neural networks for post processing model output: ARPS", *Monthly Weather Review*, **1**, 103-111.

Maqsood, I., 2004, "An ensemble of neural network for weather forecasting", *Neural Comput & Applic*., **13**, 112-122.

Naoya, M., 2001, "Prediction of precipitation by a neural network method", *Journal of Natural Disaster Science*, **23**, 1, 23-33.

Roebber, Paul J., 2006, "Real time forecasting of snowfall using a neural network", *Weather and Forecasting*, **22**, 678-684.

Roy Bhowmik, S. K., Nath, Sankar, Mitra, A. K. and Hatwar, H. R., 2009, "Application of Neural Network Technique to improve the location specific forecast of temperature over Delhi from MM5 model", *Mausam*, **60**, 1, 11-24.

Wilks, D. S., 1995, "Statistical methods in the atmospheric sciences an introduction", Academic Press. San Diego.

Appendix 'A'

2 × 2 contingency table:

		Forecast	
		Yes	No
Observed	Yes	A	B
	No	C	D

The values in the contingency table are defined as follows,

1. When an event is predicted to occur (forecast occurrence) and in reality it does occur (observed occurrence) then it is classified as A, otherwise (observed non-occurrence) it is classified as C.

2. When an event is predicted not to occur (forecast non-occurrence) and in reality it does occur (observed occurrence) then it is classified as B, otherwise (observed non-occurrence) it is classified as D.

3. A + B: Total number of cases of occurrence of precipitation as observed.

4. C + D: Total number of cases of non-occurrence of precipitation as observed.
5. A + B + C + D : Total number of forecasts.

Probability of detection: $\mathrm{POD} = \dfrac{A}{A+B}$

False Alarm Rate: $\mathrm{FAR} = \dfrac{C}{C+A}$

Correct non-occurrence: $\mathrm{C\text{-}NON} = \dfrac{D}{D+C}$

Critical success index: $\mathrm{CSI} = \dfrac{A}{A+B+C}$

Bias for occurrence: $\mathrm{BIAS} = \dfrac{A+C}{A+B}$

Percentage correct $\mathrm{PC} = \dfrac{A+D}{A+B+C+D} 100\%$

For a best/perfect forecast series:

B = 0 and C = 0 and hence

POD = 1, FAR = 0, MR = 0, C-NON = 0, Bias = 1, CSI = 1, TSS = 1, HSS = 1, PC = 100%

Appendix 'B'

4 × 4 contingency table (categorical verification of forecasts)

		Forecast				
		I	II	III	IV	Total
Observed	I	a	b	c	d	J
	II	e	f	g	h	K
	III	i	j	k	l	L
	IV	m	n	o	p	M
	Total	N	O	P	Q	T

(*Category I* : 2.5 to 12.0 mm; *Category II* : 12.1 to 24.0 mm; *Category III* : 24.1 to 48.0 mm and *Category IV* : ≥48.1 mm)

Total number of observed events in *category I* is:
J = a + b + c + d

Total number of forecast events in *category I* is:
N = a + e + i + m

In the similar way O, K, P, L, Q and M are computed. Then the total numbers of events are:
T= J + K + L + M = N + O + P + Q

Percentage correct: $\mathrm{PC} = \dfrac{a+f+k+p}{T} 100\%$

Critical success index: $\mathrm{CSI} = \dfrac{a}{J+N-a}, \dfrac{f}{K+O-f}, \dfrac{k}{L+P-k}, \dfrac{P}{M+Q-p}$

Heidke skill score: $\mathrm{HSS} = \dfrac{a+f+k+p-\dfrac{JN+KO+LP+MQ}{T}}{T-\dfrac{JN+KO+LP+MQ}{T}}$