MAUSAM, **64**, 2 (April 2013), 231-250

551.58 : 519.65

Climatology at any point : A neural network solution

PULAK GUHATHAKURTA, AJIT TYAGI* and B. MUKHOPADHYAY*

India Meteorological Department, Shivajinagar, Pune, India

* *India Meteorological Department, New Delhi, India*

(*Received 20 July 2011, Modified 22 February 2012*)

e mail : pguhathakurta@rediffmail.com

सार – सभी उपयोगकर्ताओं, योजना बनाने वालों, आपदा प्रबंधन कार्मिकों, पर्यटन आदि द्वारा तापमान, अधिकतम तापमान, न्यूनतम तापमान, वायुमंडलीय दाब, वर्षा आदि जैसे मौसम प्राचलों की जलवायु विज्ञान पर सूचनाओं की उन्नत जानकारी की अत्याधिक माँग रही है। किसी स्थान विशेष में वेधशाला के अभाव और कभी–कभी दीर्घ अवधि के पहले के आँकड़ों की अनुपलब्धता के कारण मौसम विज्ञान समुदाय उस स्थान विशेष के लिए अपेक्षित सूचनाओं को उपलब्ध नहीं करा पाता है। इस शोध पत्र में स्थानिक अंतर्वेशन के क्षेत्र में न्यूरल नेटवर्क के तुलनात्मक नए अनुप्रयोग बताए गए हैं। पूरे बारह महीनों के अधिकतम और न्यूनतम दोनों तापमानों के लिए न्यूरल नेटवर्क अंतर्वेशन निदर्श विकसित किए गए हैं। यह मॉडल उस स्थान विशेष पर सामान्य अधिकतम और न्यूनतम तापमानों को तैयार करने के लिए सूचनाओं के रूप में अक्षांश, देशान्तर और उन्नयन जैसी भौगोलिक सूचनाओं का उपयोग करता है। अंतर्वेशन के लिए स्थानिक मॉडलों के निष्पादनों की तुलना अन्य सामान्यतः प्रयुक्त पद्धति के साथ की गई है।

ABSTRACT. Advance knowledge of information on climatology of meteorological parameters like temperature, maximum temperature, minimum temperature, atmospheric pressure, rainfall etc are of great demands from all the users, planners, disaster managements personals, tourism etc. The information is required at the place concerned but this cannot be fulfilled by the meteorological community due to absent of observatory at that place and also some time absent of past data of long period. The present paper has described a comparatively new application of the neural network in the field of spatial interpolation. Neural network interpolation models are developed for both maximum and minimum temperatures for all the twelve months. The model utilizes geographical information like latitude, longitude and elevation as inputs to generate normal maximum and minimum temperatures at a place. The performances of the models are compared with the other commonly used method for spatial interpolation.

Key words – Spatial interpolation, Neural network, Climate normal, Temperature.

1. Introduction

The influence of climate on our lives is all pervading as it is an important factor in agriculture, commerce, industry, transport, tourism and almost all factors of our daily life. Advance knowledge of information on climatology of meteorological parameters like temperature, maximum temperature, minimum temperature, atmospheric pressure, rainfall etc are of great demands from all the users, planners, disaster managements personals, tourism etc. Climatological normals are useful way to describe the average weather of a place and they serve as a bench mark for comparing current as forecast climatic condition of the future. However due to several constraints it is not possible to have a meteorological observatory at every place. The latest climate normal for the period 1971-2000 of India Meteorological Department has climate normal of several meteorological parameters of around 410 observatory stations.

Interpolation belongs to a very useful branch of mathematics, called approximation theory. Statistical interpolation considers fields having random fluctuations. Various interpolation methods have been invented to solve interpolation problems in practical applications. Cressie (1991) and Meyers (1994) have summarized many traditional interpolation methods. Commonly used interpolation methods for meteorological applications include nearest-station assignment, inverse-distance weighting, inverse-distance-square weighting, Thiessen-polygon method, orthogonal-polynomial approximation, Lagrange method, interpolation by splines, kriging, radial basis functions and interpolation by empirical orthogonal functions. Though these methods vary in complexity, all use information present in known sample locations to interpolate values for other locations. The accuracy with which a method can generate interior estimates depends on the complexity that underlies the spatial structure of the field. Each method has its merits and is applicable according to temporal length scale, spatial length scale, stationarity, and variability of the field under consideration. Several interpolation schemes are available for spatial analyses which help to generate gridded data set or to get values at desired point (New *et al*. 2000, Mitra *et al.* 2003, Yatagai *et al*. 2005, Rajeevan *et al*. 2006, Xie *et al*. 2007). One of the major limitations of these schemes is ignoring the geo-position importance of the points.

The strengths and weaknesses of existing methods for spatial interpolation and that describes efforts to downscale climate data generated by GCM simulations are available in various literatures (Giorgi and Mearns 1991; Hewitson and Crane 1996; Joubert and Hewitson 1997; Wilby and Wigley 1997; Wilby *et al*. 1998). The techniques used to downscale output from GCMs can be divided between two general categories; process-based and empirical approaches. Process-based approaches embed higher-resolution physical models within GCM grids (Dickinson *et al.* 1989; Giorgi *et al*. 1994; Russo and Zack 1997). These methods are computationally expensive and therefore may not be practical.

Empirical methods of downscaling generally use transfer functions to relate local conditions to large-scale climate features. The accuracy of these transfer functions is not perfect. Statistical approaches are valid only within the range of the sample data (Katz 1977). The relation between synoptic-scale features and local climatological fields often is highly nonlinear and changes with atmospheric circulation. These difficulties, along with spatial dependencies (*e.g*., spatial autocorrelation and anisotropy) that are inherent to many climatological fields, create the need for complex mathematical specifications and estimation techniques (Bogardi *et al*. 1993; Matyasovsky *et al*. 1994). Finally, it remains uncertain whether these empirical approaches can be applied to the non-stationarity of climate change (Palutikof *et al*. 1997).

The spatial interpolation of temperature is invariably influenced by elevation. Elevation emerged as the strongest covariate for estimating both daily maximum and minimum temperature (Jarvis and Stuart 2001). According to them this relationship tended to be stronger for maximum than minimum temperature. Elevation has guided the spatial interpolation of temperature in numerous other studies as well (Price *et al*. 2000; Johnson *et al*. 2000; Kurtzman and Kadmon 1999). In addition to the principal influence of elevation, Jarvis and Stuart (2001) show that northing (north map coordinate), directional distances to the coast, and urbanization emerge as some of the most important covariates. Typically, more variables were required to interpolate daily minimum temperature than maximum temperature; however, even in the case of minimum temperature the addition of more than three covariates has a minimal effect on interpolation accuracy (Jarvis and Stuart 2001). Temperature also depends on the urbanization and the degree of urbanization was a particularly important covariate for minimum temperature (Jarvis and Stuart 2001). By addressing elevation, Choi *et al*. (2003) showed a 30% reduction in minimum temperature interpolation errors by accounting for urbanization. Nonetheless, elevation exerted the primary influence.

Artificial Neural Networks (ANN) technique may be capable to address many of the difficulties described above and therefore may be a useful tool to downscale GCM output. Hornik *et al*. (1989) described ANNs as universal approximators. ANNs can approximate nonlinear relations and their derivatives without knowing the true nonlinear function; therefore ANNs can make accurate predictions for highly nonlinear systems (Werbos 1974; Rumelhart *et al*. 1986; Fischer and Gopal 1994; Gopal and Scuderi 1995).

Gardner and Dorling (1998) have used ANNs to analyze climate data. The effect of atmospheric circulation on local precipitation is investigated using ANNs (Hewitson and Crane 1994; Cavazos 1997; Crane and Hewitson 1998). McGinnis (1997) uses ANNs to develop transfer functions that predict snowfall from grid scale information generated by GCMs. Zhang and Scofield (1994) use an ANN to estimate convective rainfall and recognize cloud merger from satellite data.

Neural network technique can be considered as a tool for spatial interpolation which can particularly adapt at handling massive amounts of data, dealing with complex nonlinear relationships, coping with non-normal and inter-correlated inputs, and allowing incorporation of additional data and expert knowledge about a particular geographical domain within the estimation process (Bishop, 1995). Rigol *et al*. (2001), Snell *et al*. (2000) have used neural network technique for spatial interpolation of surface temperature and down scaling the GCM outputs. In India an attempt has already been made by Karmakar *et al*. (2009) for spatial interpolation of rainfall variables using artificial neural network for the Chattisgarh region.

The climatological information is always available at uneven spaced intervals as the information is made from the average of past data of observations of climate parameters taken at different observatories. The observatories are not situated at evenly space of points. Spatial interpolation techniques are used to generate climate information at evenly space points from the unevenly space points. This is called generation of gridded data. In this regard well known interpolation techniques like inverse distance weighting, Kriging etc.

as well as neural network are used. However the gridded data are generally used for getting general pattern of climate variable.

Our main purpose in this paper is to have climate information at any point, *i.e.*, unevenly space point. Though theoretically inverse distance weighting, kriging and other methods can be used to have information at ungaged point or at a point where there is no observation or observatory, but neither of these methods so far has been applied in such a case due to limited accuracy. In other word these spatial interpolation techniques can not generate information at any point with reasonable accuracy but can be applicable to generate gridded data set which gives climate features over larger area. Even ANN is used in this category. But in ANN there is flexibility of including more parameters as inputs. Temperature is invariably influenced by elevation, latitude etc. In this paper we have used latitude, longitude and elevation as inputs and temperature as output. From the known set of input output pattern, we have extracted the relationship between input and output by using standard training algorithm of ANN.

In the present case we have used monthly climatological normal temperature both maximum and minimum of observatory stations of India Meteorological Department. The neural network models are developed using the geographical information (latitude, longitude and elevation) of these stations as input and temperature as input. The models so developed are able to generate normal maximum temperature and minimum temperature at any point with reasonable accuracy by providing latitude, longitude and elevation of the point as inputs to the model. We have also compared the results with the other standard method used in spatial interpolation like inverse distance weighting.

Inverse distance weighting and Kriging are the most widely used methods in spatial interpolation of climate variables. But the spatial interpolation for these methods is generally used to generate the value at evenly spaced points i.e. at grid points. The values generated at evenly spaced points are used for getting spatial patterns by contour analysis. Even artificial network can also be used to generate values at evenly spaced points. However in many practical cases users from different sectors like industry, tourism, health and agriculture require the climate information at uneven spaced points *i.e.*, at their point of interest. So far standard interpolation techniques have limited application to address these issues. The present artificial network interpolation model which uses only geo coordinates *i.e.*, latitude, longitude and attitude as inputs addresses these problems. We have selected four high altitudes and four low altitude stations for comparisons of the results from neural network model with inverse distance weighting and Kriging methods. Cross validation technique is used for getting estimations from Inverse distance weighting and Kriging methods for these eight stations.

2. Data

As a mandate of World Meteorological Organization, India Meteorological Department is preparing climate normal for all meteorological parameters in every 10 years. The recent update is the 30 year climate normal based on the data for the period 1971-2000. These normal is available for 413 surface observatories of the country. We have taken monthly mean maximum and minimum temperature of the available stations from India Meteorological Department, Pune.

3. Methodology

Multi layer feed-forward networks form an important class of neural networks. Typically the network consists of a set of sensory units or input nodes, that constitute the input layer, one or more hidden layers of neurons or computation nodes, and an output layer. Multi layer Perceptron (MLP) neural networks with sufficiently many nonlinear units in a single hidden unit layer have been established as universal function approximators. The advantages of the MLP are:

(*i*) Hidden unit outputs (basis functions) change adaptively during training, making it unnecessary for the user to choose them beforehand.

(*ii*) The number of free parameters in the MLP can be unambiguously increased in small increments by simply increasing the number of hidden units.

(*iii*) The basis functions are bounded making overflow errors and round-off errors unlikely.

The MLP is a feed-forward network consisting of units arranged in layers with only forward connections to units in subsequent layers. The connections have weights associated with them. Each signal traveling along a link is multiplied by its weight. The input layer, being the first layer, has input units that distribute the inputs to units in subsequent layers. In the following (hidden) layer, each unit sums its inputs and adds a threshold to it and nonlinearly transforms the sum (called the net function) to produce the unit output (called the activation). The output layer units often have linear activations, so that output activations equal net function values.

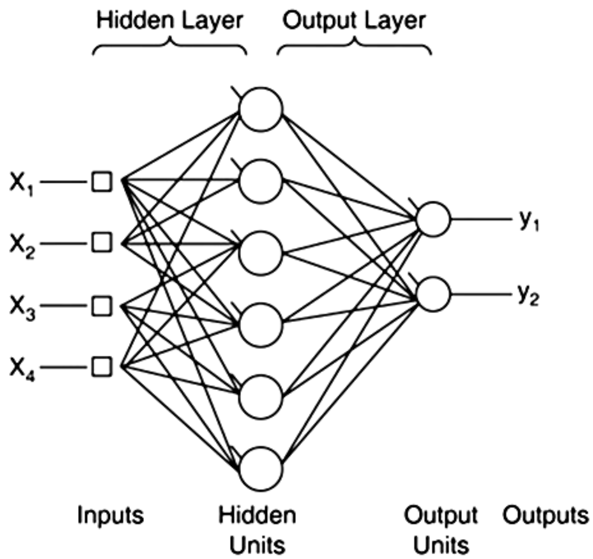

Fig. 1. A three layer (one input layer, one hidden layer and one output layer) neural network model

The layers between the input and the output layers are called hidden layers, and the units in the hidden layers are called hidden units. The network shown below has one hidden layer.

The training data set consists of N_v training patterns $[(x_p, t_p)]$, where p is the pattern number. The input vector x_p and the desired output vector t_p have dimensions N and M, respectively. y_p is the network output vector for the p^{th} pattern. For the j^{th} hidden unit, the net input $NET_p(j)$ and the output activation $O_p(j)$ for the p^{th} training pattern are:

$$NET_p(j) = \sum_{i=1}^{N+1} w(j,i).x_p(i), \qquad 1 \le i \le N_h$$

$$O_p(j) = f\left[NETp(j)\right]$$

where $w(j, i)$ denotes the weight connecting the i^{th} input unit to the j^{th} hidden unit and $X_p(N+1) = 1$.

The kth output for the pth training pattern is y_{pk} and is given by

$$y_{pk} = \sum_{i=1}^{N+1} w_{io}(k,i).x_p(i) + \sum_{j=1}^{Nh} w_{ho}(k,j).O_p(j), \qquad 1 \le k \le M$$

where $w_{io}(k,i)$ denotes the output weight connecting the ith input unit to the kth output unit and $w_{ho}(k,j)$ denotes the output weight connecting the jth hidden unit to the kth output unit. Note that we use bypass weights connecting the inputs and the outputs. This results in fewer hidden units being necessary. Also, this means that no hidden units are required for the linear network case. Note that the network has linear output activation, as support vector machines (SVM).

The overall performance of an MLP neural network, measured as mean square error (MSE), can be written as

$$E = \sum_{k=1}^{M} E(k) = \frac{1}{Nv}\sum_{p=1}^{Nv} E_p$$

where E_p is the sum of squared errors for the pth pattern and $E(k)$ is the MSE error for the kth output unit defined as

$$E_p = \sum_{k=1}^{M}\left[t_{pk} - y_{pk}\right]^2$$

$$E(k) = \frac{1}{Nv}\sum_{p=1}^{Nv}\left[t_{pk} - y_{\partial pk}\right]^2$$

where t_{pk} denotes the kth element of the pth desired output vector.

We have used a three layer (one input layer, one hidden layer and one output layer) neural network model (Fig. 1). The input layer has three input nodes *viz*., latitude, longitude and elevation while the output layer has only one node, the temperature. There are 413 stations for which climatological normal for maximum temperature is available and there are 410 stations for which climatological normal for minimum temperature is available. 80% of the stations are used as training the data set and the remaining 20% are used as testing data set.

The training algorithm used Output Weight Optimization - Hidden Weight Optimization. The Output Weight Optimization - Hidden Weight Optimization (OWO-HWO) method was introduced by Chen *et al.* (1999) and then modified by Yu and Manry (2002) to train the neural network. They have shown that the OWO-HWO method is superior in terms of convergence to standard OWO-BP (output weight optimization-back propagation) which uses OWO to update output weights and back propagation to update hidden weights.

The models are developed for the maximum temperature as well as minimum temperature and for all the twelve months separately. Initially 10 hidden nodes are taken for developing the model so that there are twenty four separate neural network interpolation models.

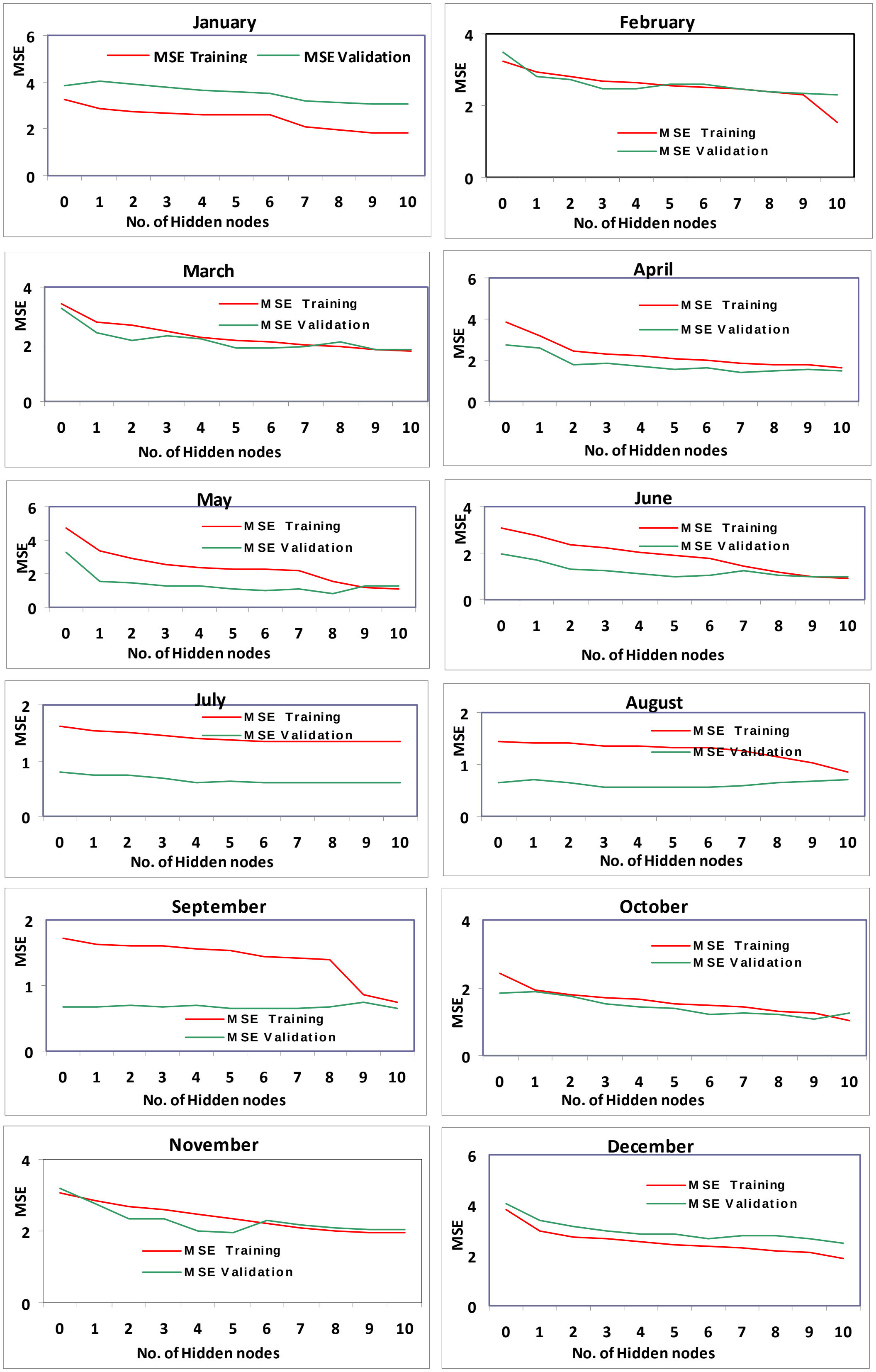

Fig. 2. Performances of the NN models for minimum temperature in pruning for different hidden nodes in training and validation sets

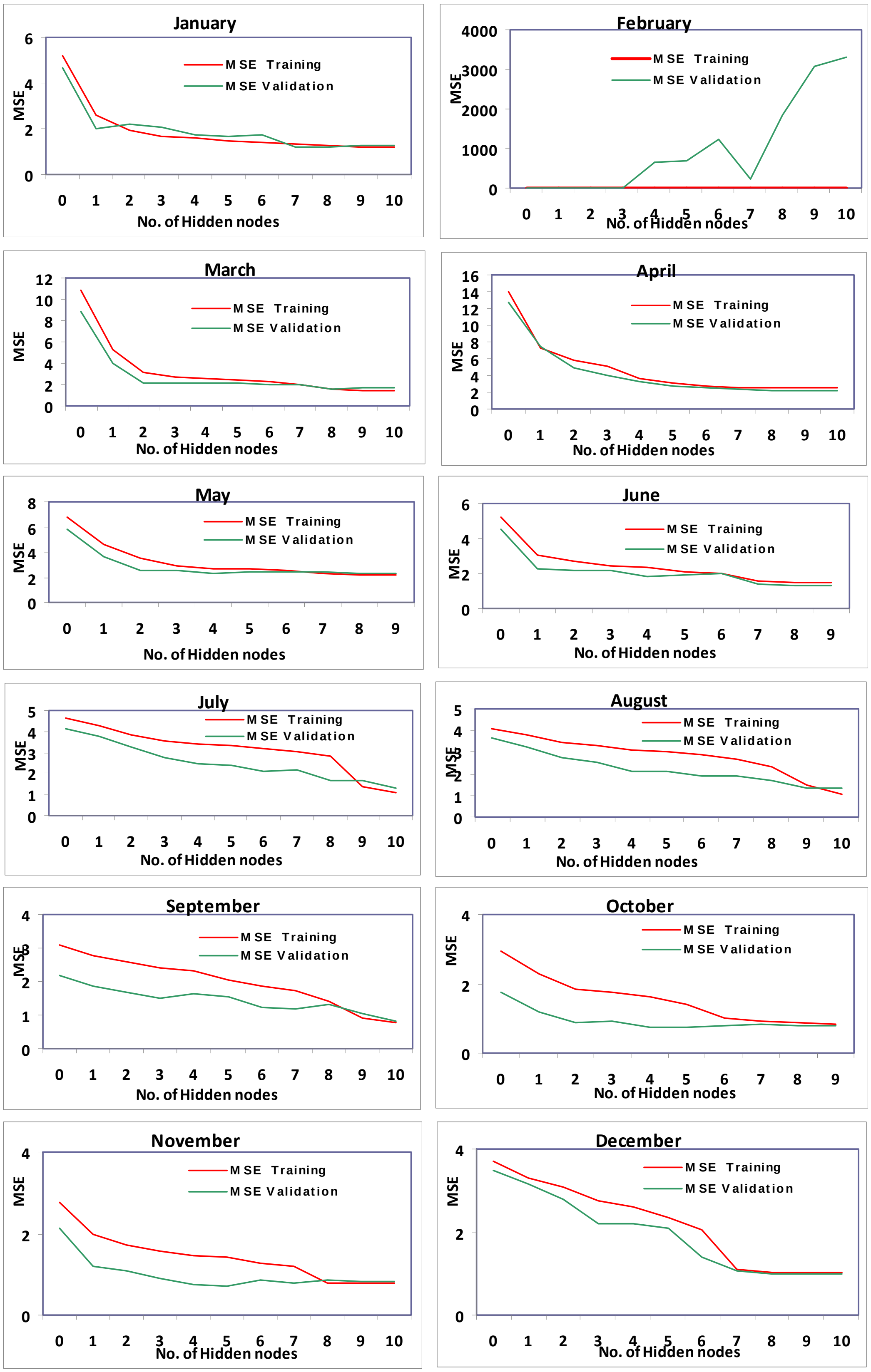

Fig. 3. Performances of the NN models for maximum temperature in pruning for different hidden nodes in training and validation sets

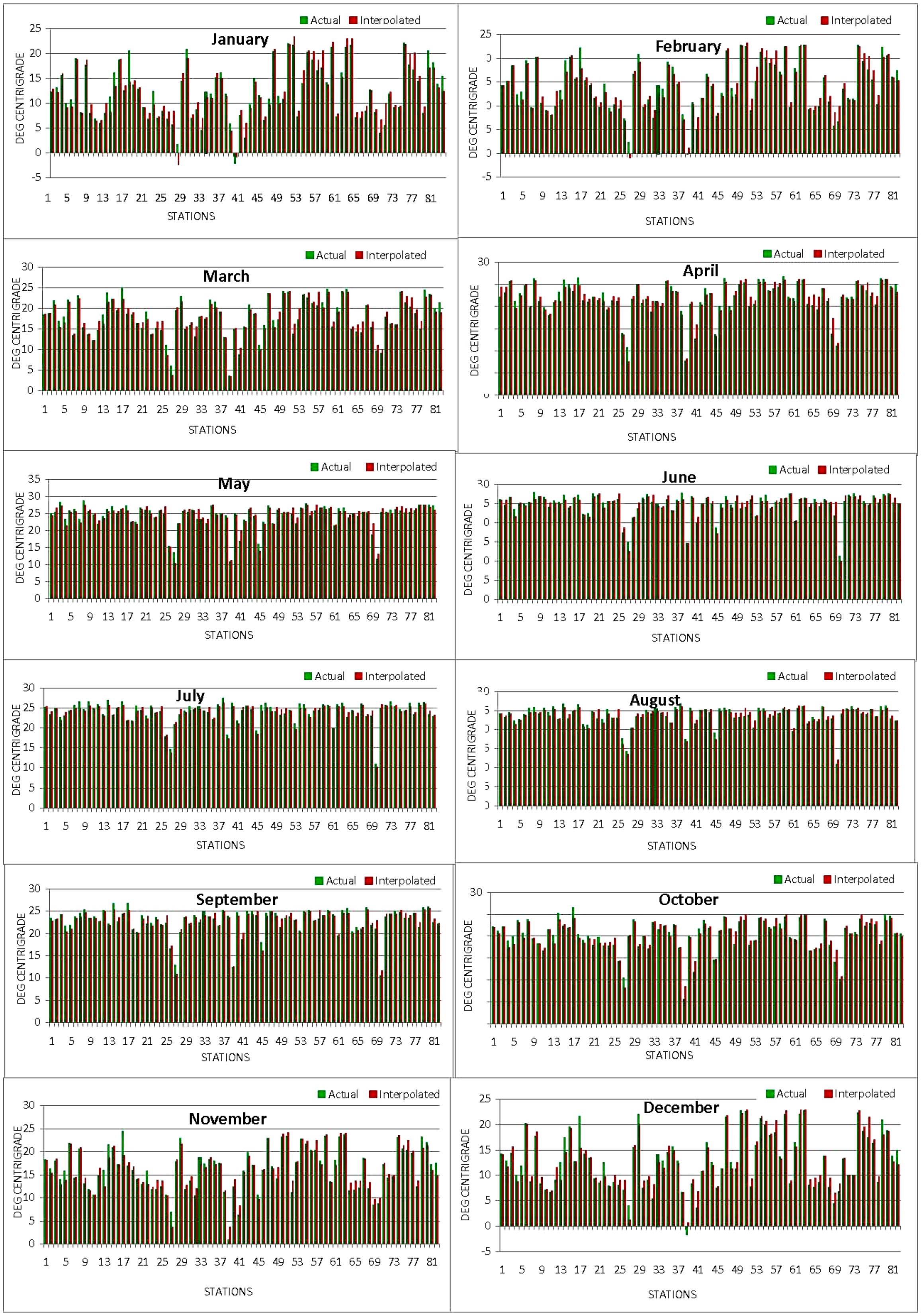

Fig. 4. Performances of the NN interpolation models for minimum temperature for the twelve months for the validation stations

4. Results and discussions

10 hidden nodes are taken for developing the models initially. The pruning method which is generally applied to select optimum number of hidden nodes is used for the final number of hidden nodes. Fig. 2 and Fig. 3 show the pruning results showing the performances of the models for varying hidden nodes for twelve months and for the minimum and maximum temperature respectively. Final numbers of hidden nodes for each month are selected based on the performance of the models both in the training period and validation period. Table 1 shows the hidden nodes for each month for the final models. Ones we have developed 24 models for each of the twelve months (12 models for twelve months in case of minimum temperature and 12 models for twelve months in case of maximum temperature), the models can be used to give climatological information *i.e.*, normal maximum and minimum temperatures at any geographical point of India provided we have geographical information (*i.e.*, latitude, longitude and elevation which has to be provided as inputs to the model) of that point.

TABLE 1

Numbers of hidden nodes in the final NN models for minimum and maximum temperatures

Months	No. of hidden nodes for minimum temperature NN model	No. of hidden nodes for maximum temperature NN model
January	9	9
February	10	3
March	10	8
April	10	9
May	10	9
June	9	10
July	10	10
August	10	10
September	10	10
October	9	8
November	9	10
December	10	8

4.1. *Performances of the minimum temperature models during the training and validation periods*

As mentioned earlier there are 410 stations for which monthly climatological normal of minimum temperatures are available. 80% of these stations *i.e.*, 328 stations are chosen randomly to use as developing the models while remaining 82 stations are used to validate the models. Fig. 4 gives the comparison of interpolated minimum temperature values obtained from the NN model for the 82 stations which are not used for training the model. It may be mentioned that we have 12 separate models for each of the twelve months. Performances of the models are encouraging for all the months and it is able to produce the lowest temperature values and also highest temperature values quite accurately.

The performances of the models have been further verified by computing Root Mean Square Error (RMSE) of the models during both training set and validation set of stations (Table 2). Table 2 also shows the standard deviation of the minimum temperature. The RMSE for the validation set is less than even one third of standard deviation for all the months. To see the performances of the models for each of the stations (validation set of 82 stations) we have categorized the mean absolute deviation of the interpolated value obtained from the NN model from the actual as (*i*) within 1 (*ii*) within 2 (*iii*) within 3 and (*iv*) within 4 and (*v*) greater than 4. Table 3 shows the percentage of stations falling in each category for the minimum temperature. Number of stations coming in the 1st category varies from 46% (January) to maximum 83% (September) while the number of stations coming in 2nd category varies from 80% to 100%. Only 3.7% and 1.2% of stations (for the month of January and February) are coming in the last category *i.e.*, where deviation is more than 4.

4.2. *Performances of the maximum temperature models during the training and validation periods*

As mentioned earlier there are 413 stations for which monthly climatological normal of minimum temperatures are available. 80% of these stations *i.e.*, 330 stations are chosen randomly to use as developing the models while remaining 83 stations are used to validate the models. Fig. 5 gives the comparison of interpolated maximum temperature values obtained from the NN model for the 83 stations which are not used for training the model. It may be mentioned that as in the case of minimum temperature we have 12 separate models for each of the twelve months. Performances of the models are encouraging for all the months and it is able to produce the highest temperature values and also lowest temperature values quite accurately.

The performances of the models have been further verified by computing Root Mean Square Error (RMSE) of the models during both training set and validation set of

TABLE 2

RMSE of NN models for minimum temperature and maximum temperature during training and validation periods and its comparison with standard deviation

Months	Minimum Temperature (deg centigrade)			Maximum Temperature (deg centigrade)		
	Standard Deviation	RMSE		Standard Deviation	RMSE	
		Training Period	Validation Period		Training Period	Validation Period
January	5.79	1.35	1.75	5.30	1.09	1.12
February	5.54	1.24	1.52	5.31	1.58	1.43
March	4.85	1.33	1.34	5.04	1.28	1.27
April	4.22	1.28	1.23	4.87	1.58	1.45
May	3.92	1.05	1.11	4.98	1.50	1.50
June	3.41	0.99	0.98	4.23	1.22	1.15
July	2.87	1.16	0.78	3.21	1.05	1.15
August	2.83	0.93	0.84	3.02	1.02	1.15
September	3.15	0.86	0.81	2.91	0.89	0.89
October	3.89	1.13	1.04	3.24	0.94	0.88
November	4.83	1.39	1.43	3.57	0.88	0.90
December	5.61	1.36	1.58	4.46	1.02	0.99

TABLE 3

Performances of the NN interpolation models for minimum temperature for the twelve months: Absolute deviation from actual in different categories for the validation stations (82 stations)

Months	No of station (%) for which Absolute Deviation from Actual				
	$\leq$ 1 deg C	$\leq$ 2 deg C	$\leq$ 3 deg C	$\leq$ 4 deg C	> 4 deg C
January	46.3	79.3	91.5	96.3	3.7
February	53.7	80.5	93.9	98.8	1.2
March	51.2	82.9	98.8	100.0	0.0
April	68.3	89.0	95.1	100.0	0.0
May	69.5	91.5	96.3	100.0	0.0
June	72.0	97.6	98.8	100.0	0.0
July	75.6	100.0	100.0	100.0	0.0
August	80.5	96.3	100.0	100.0	0.0
September	82.9	97.6	100.0	100.0	0.0
October	69.5	91.5	100.0	100.0	0.0
November	61.0	84.1	95.1	98.8	1.2
December	54.9	82.9	93.9	97.6	2.4

stations. Last three columns of Table 2 show the standard deviation of the minimum temperature, RMSE of the maximum temperature models for the training and validation sets of stations. The RMSE for the validation set is less than even one third of standard deviation for all the months except for July and August where it is less

TABLE 4

Performances of the NN interpolation models for maximum temperature for the twelve months: Absolute deviation from actual in different categories for the validation stations (83 stations)

Months	No of station (%) for which Absolute Deviation from Actual within				
	$\leq$ 1 deg C	$\leq$ 2 deg C	$\leq$ 3 deg C	$\leq$ 4 deg C	> 4 deg C
January	69.9	92.8	97.6	100.0	0.0
February	59.0	85.5	96.4	98.8	1.2
March	73.5	90.4	94.0	97.6	2.4
April	58.3	85.7	95.2	96.4	3.6
May	57.8	85.5	95.2	96.4	3.7
June	62.7	91.6	98.8	100.0	0.0
July	77.1	94.0	96.4	98.8	1.2
August	68.7	92.8	97.6	98.8	1.2
September	78.3	94.0	98.8	100.0	0.0
October	86.7	95.2	98.8	100.0	0.0
November	83.1	94.0	97.6	100.0	0.0
December	78.3	92.8	98.8	100.0	0.0

than half of the standard deviation. To see the performances of the models for each of the stations (validation set of 83 stations) we have categorized the mean absolute deviation of the interpolated value obtained from the NN model from the actual as (*i*) within 1 (*ii*) within 2 (*iii*) within 3 and (*iv*) within 4 and (*v*) greater than 4. Table 4 shows the percentage of stations falling in each category for the maximum temperature. Number of stations coming in the 1st category varies from 58% (January) to maximum 87% (September) while the number of stations coming in 2nd category varies from 86% to 95%. For 94% to 99% stations the mean absolute deviations are within 3 deg C for all the months.

4.3. *Comparison of the results with results obtained by inverse distance weighting*

A general form of finding an interpolated value *u* at a given point *x* based on samples $u_i = u(x_i)$ for $i = 0, 1,..., N$ using IDW is an interpolating function:

$$u(x) = \sum_{i=0}^{N} \frac{w_i(x)u_i}{\sum_{j=0}^{N} w_j(x)},$$

where

$$w_i(x) = \frac{1}{d(x, x_i)^p}$$

is a simple IDW weighting function, as defined by Shepard (1968), *x* denotes an interpolated (arbitrary) point, x_i is an interpolating (known) point, *d* is a given distance (metric operator) from the known point x_i to the unknown point *x*, *N* is the total number of known points used in interpolation and *p* is a positive real number, called the power parameter.

TABLE 5

Performances of the IDW (Shepard's) method for maximum and minimum temperatures for the twelve months

Months	RMSE in deg Centigrade	
	Max	Min
Jan	2.10	2.14
Feb	2.21	2.09
Mar	2.37	2.03
Apr	2.47	1.94
May	2.44	1.93
Jun	2.27	1.87
Jul	2.17	1.82
Aug	2.13	1.81
Sep	2.13	1.81
Oct	2.09	1.82
Nov	2.07	1.96
Dec	2.02	2.11

Fig. 5. Performances of the NN interpolation models for maximum temperature for the twelve months for the validation stations

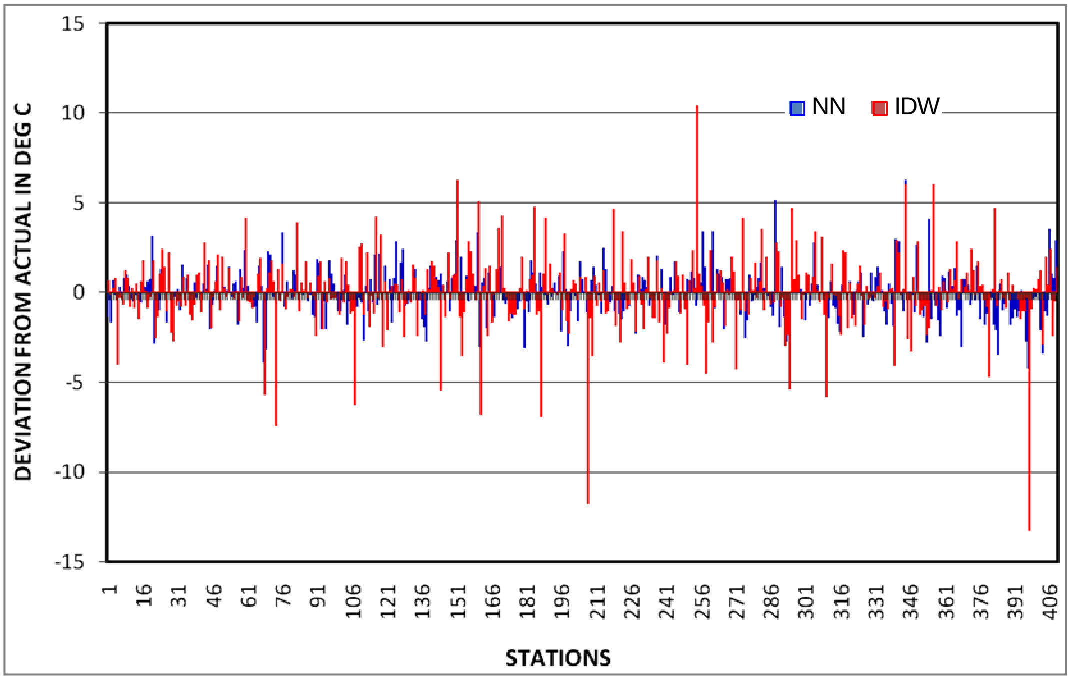

Fig. 6. Comparison of the NN interpolation model (NN) and Inverse Distance Weighting model (IDW) for minimum temperature for January

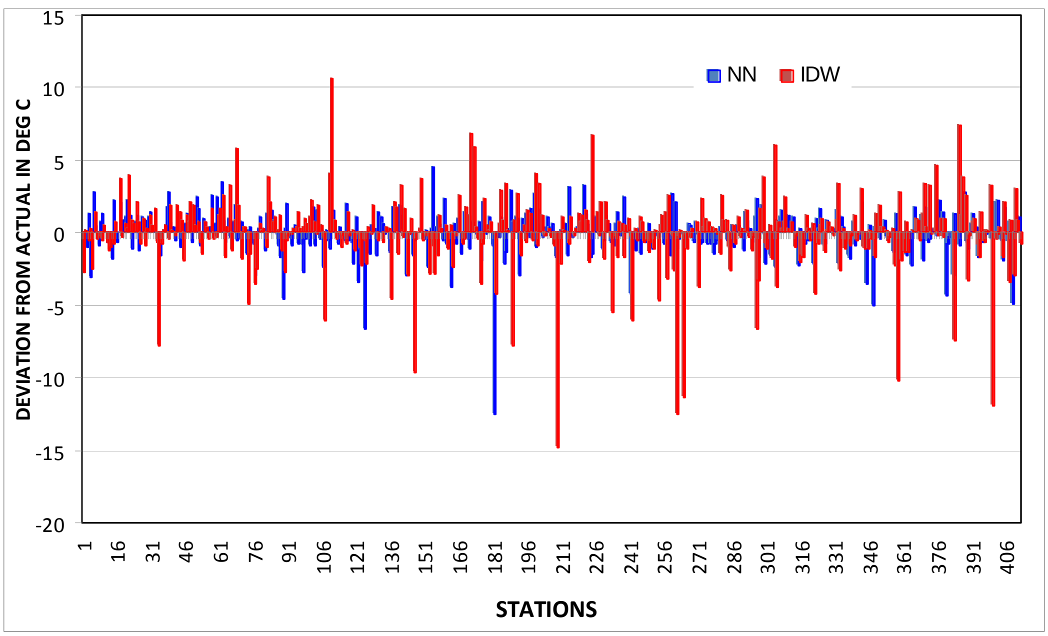

Fig. 7. Comparison of the NN interpolation model (NN) and Inverse Distance Weighting model (IDW) for maximum temperature for May

This method is commonly known as Shepard method and is mostly used for generating gridded data set for climatological parameters (Piper and Stewart, 1996; New *et al*. 2000; Kiktev *et al*. 2003; Caesar *et al*. 2006; Rajeevan *et al*. 2006; Srivastava *et al*. 2009 etc).

In order to compare the results with the results obtained by NN models for both the training set and validation set of stations, we have used Shepard's method to generate values at each station instead of grid point. In this process we have used the remaining stations for

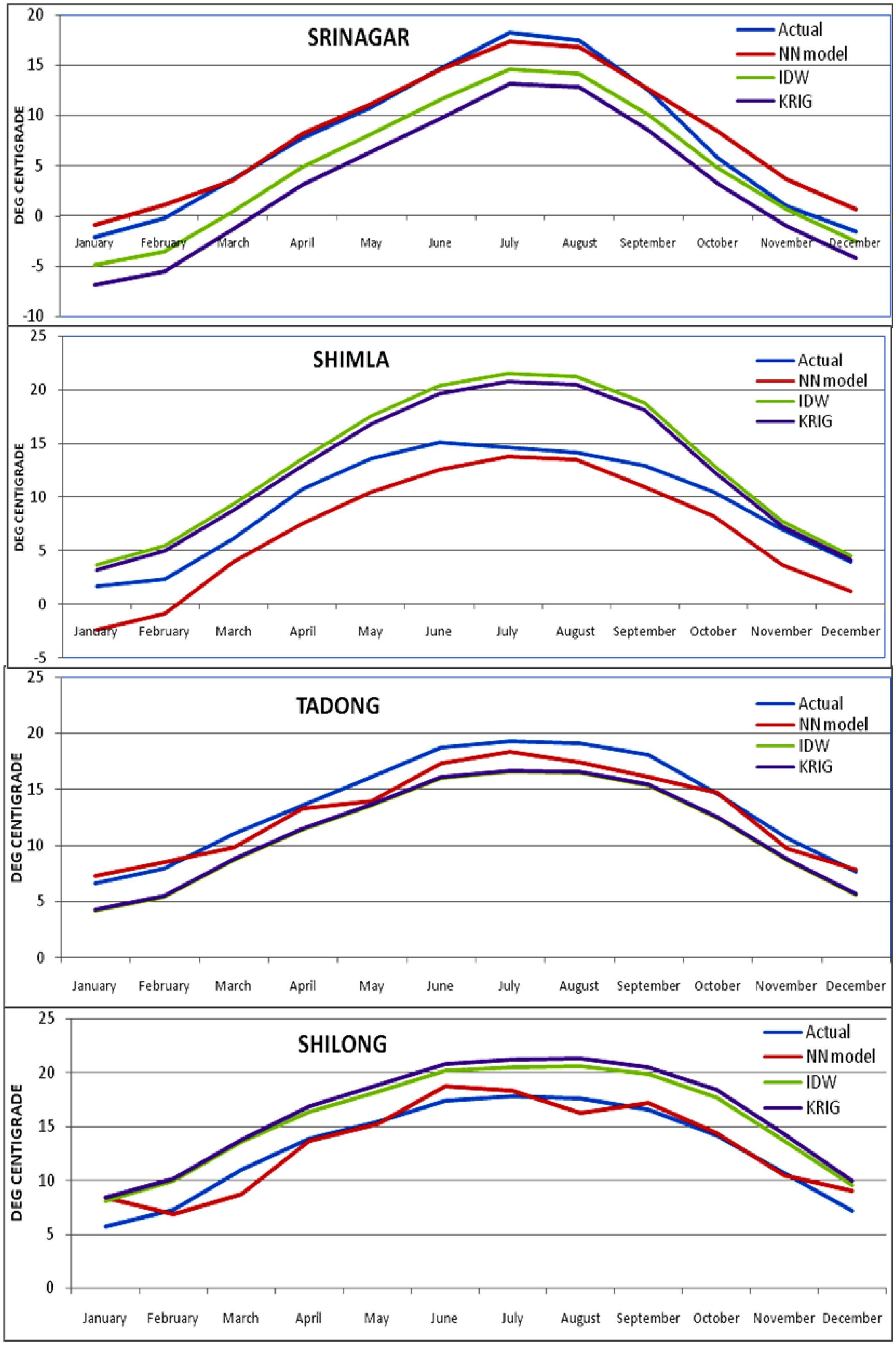

Fig. 8. Comparison of the performances of NN interpolation model (NN) along with IDW and Kriging methods for minimum temperature for all the twelve months for four high elevation stations

search and IDW. Table 5 shows the Root Mean Square Error (RMSE) for both maximum and minimum temperatures and for the twelve months. The RMSE values are very high compare to the RMSE values presented in third, fourth and sixth and seventh column of Table 2. In Fig. 6 and Fig. 7 we have presented the deviation of temperature values obtained from neural network model and inverse distance weighting model for the two extreme months *i.e*., minimum temperature for January month and maximum temperature for the month of May. For many stations, the deviation is more than 5 deg centigrade for IDW method for both the months while for NN models only one station in January and two stations in May have deviation more than 5 deg centigrade. All these results clearly show that NN model considered here for spatial interpolation is superior than the inverse distance method which normally used for generating gridded data of climate.

Since we have used station's latitude, longitude and elevation as inputs to produce the temperature (output) in the Neural Network model we have tried with the simple regression giving same parameters as input and output. For example RMSE in regression model for May maximum temperature is 3.9 which is even higher than the Shepard's interpolation result. The results are very poor as the regression method can explore the nonlinearity between inputs and output which was explained by the NN models.

4.4. *Comparison of the results with results obtained by Kriging method*

Kriging is an interpolation method named after a South African mining engineer named D. G. Krige (1951) who developed the technique in an attempt to more accurately predict ore reserves. It was further modified by Georges Matheron (1963) as the "theory of regionalized variables". Over the past several decades kriging has become a fundamental tool in the field of geostatistics.

Kriging is based on the assumption that the parameter being interpolated can be treated as a regionalized variable. A regionalized variable is intermediate between a truly random variable and a completely deterministic variable in that it varies in a continuous manner from one location to the next and therefore points that are near each other have a certain degree of spatial correlation, but points that are widely separated are statistically independent (Davis, 1986). Kriging is a set of linear regression routines which minimize estimation variance from a predefined covariance model.

The first step in ordinary kriging is to construct a variogram from the scatter point set to be interpolated. A variogram consists of two parts: an experimental variogram and a model variogram. Suppose that the value to be interpolated is referred to as *f*. The experimental variogram is found by calculating the variance (*g*) of each point in the set with respect to each of the other points and plotting the variances *versus* distance (*h*) between the points. Several formulas can be used to compute the variance, but it is typically computed as one half the differences in f squared. Once the experimental variogram is computed, the next step is to define a model variogram. A model variogram is a simple mathematical function that models the trend in the experimental variogram.

The shape of the variogram indicates that at small separation distances, the variance in f is small. In other words, points that are close together have similar f values. After a certain level of separation, the variance in the f values becomes somewhat random and the model variogram flattens out to a value corresponding to the average variance.

Once the model variogram is constructed, it is used to compute the weights used in kriging. The basic equation used in ordinary kriging is as follows:

$$F(x, y) = \sum_{i=1}^{n} w_i f_i$$

where n is the number of scatter points in the set, f_i are the values of the scatter points, and w_i are weights assigned to each scatter point. This equation is essentially the same as the equation used for inverse distance weighted interpolation except that rather than using weights based on an arbitrary function of distance, the weights used in kriging are based on the model variogram.

In addition to modified Spepard method (IDW), we have also used Kriging method to have an estimation of interpolation value at each of the point. Here also as in the IDW method, we have considered all the points excluding the point for which interpolation is obtained (cross validation). The overall performances are comparable with the IDW method but are not better than the neural network model.

4.5. *Validation and comparison of the Neural Network interpolation model for high elevation stations*

As we know the rotation and spinning of the earth cause the variation of solar radiation over earth's surface. Thus temperature has both latitudinal and longitudinal variation. Also there is high relationship between the

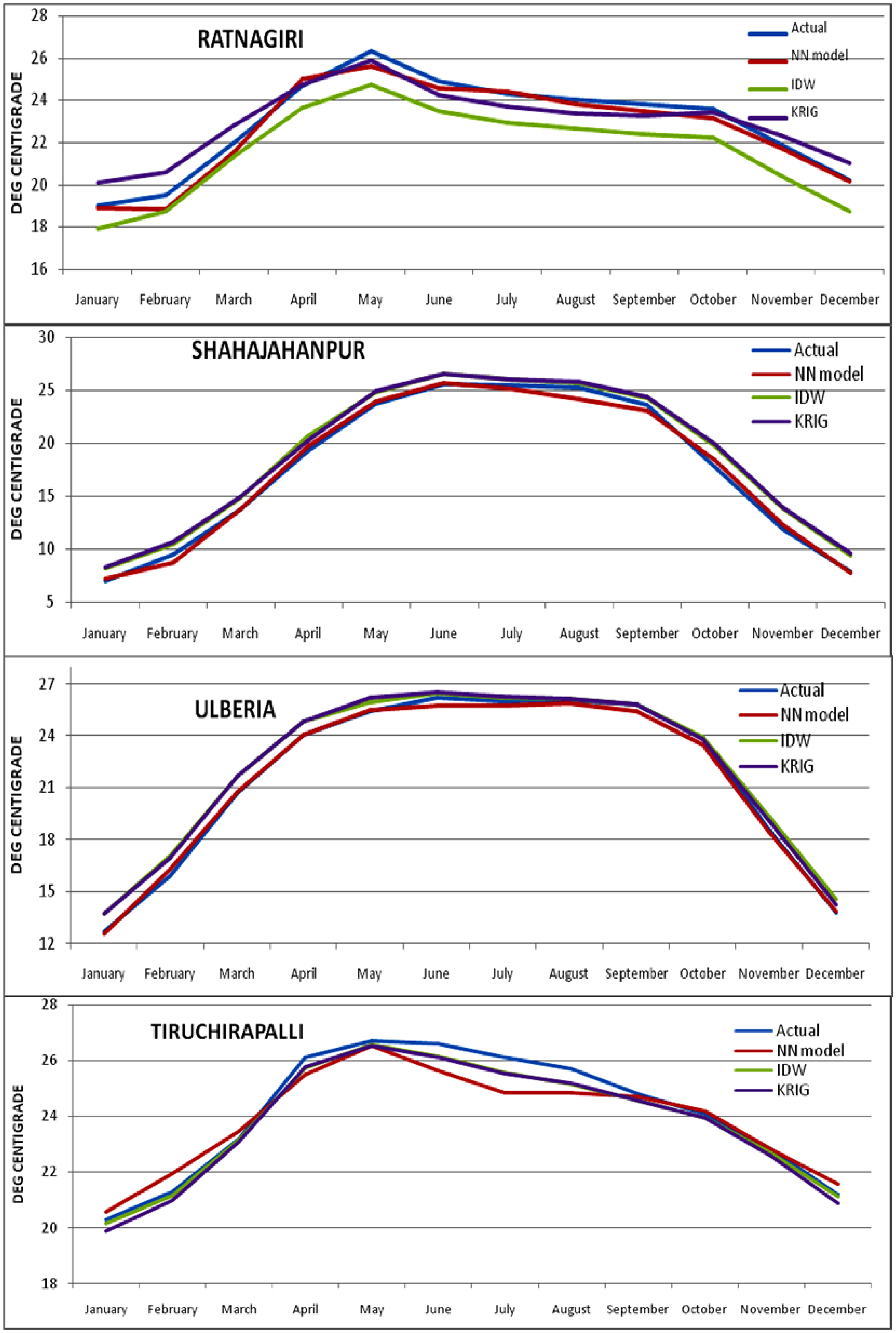

Fig. 9. Comparison of the performances of NN interpolation model (NN) along with IDW and Kriging methods for minimum temperature for all the twelve months for four low elevation stations

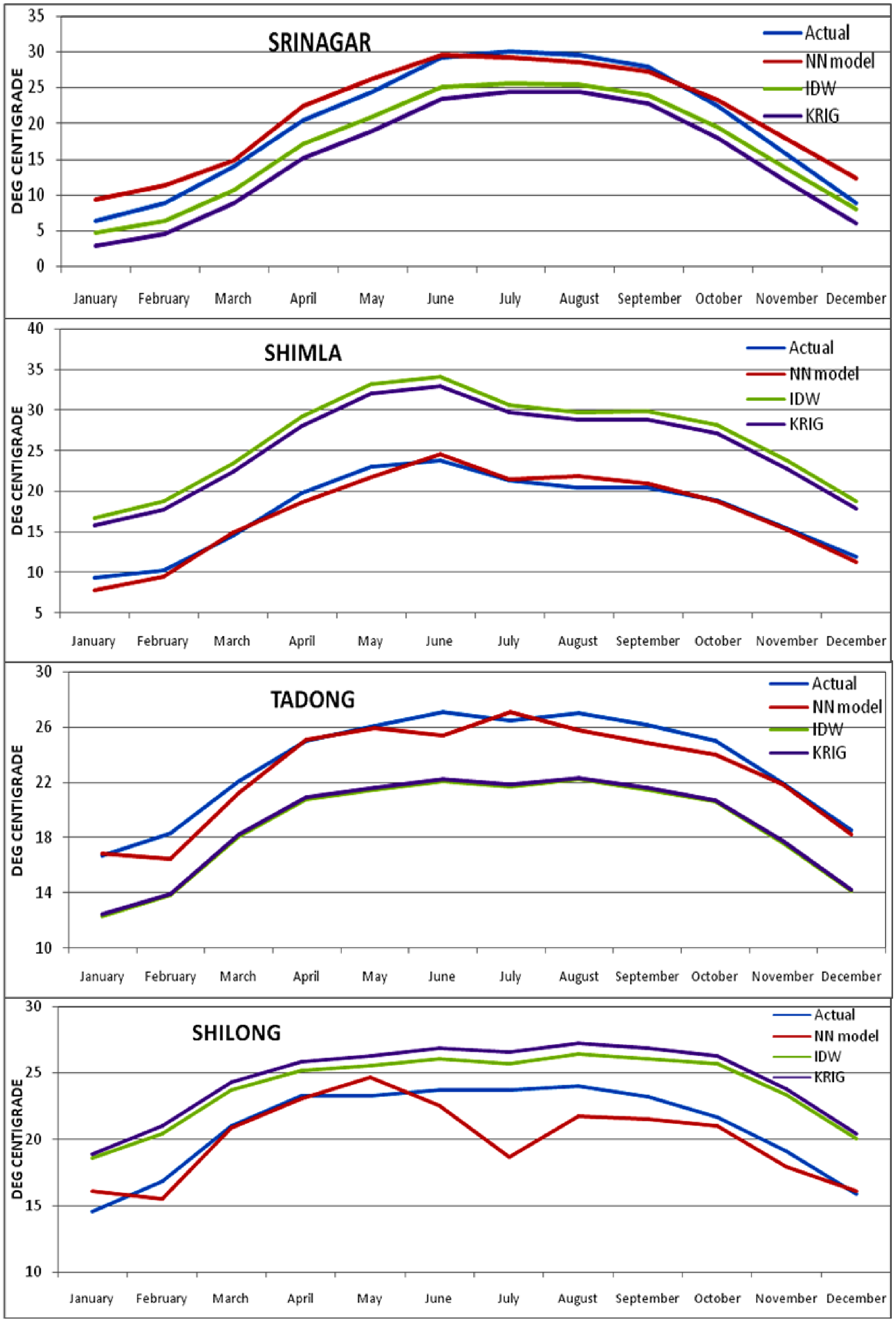

Fig. 10. Comparison of the performances of NN interpolation model (NN) along with IDW and Kriging methods for maximum temperature for all the twelve months for four high elevation stations

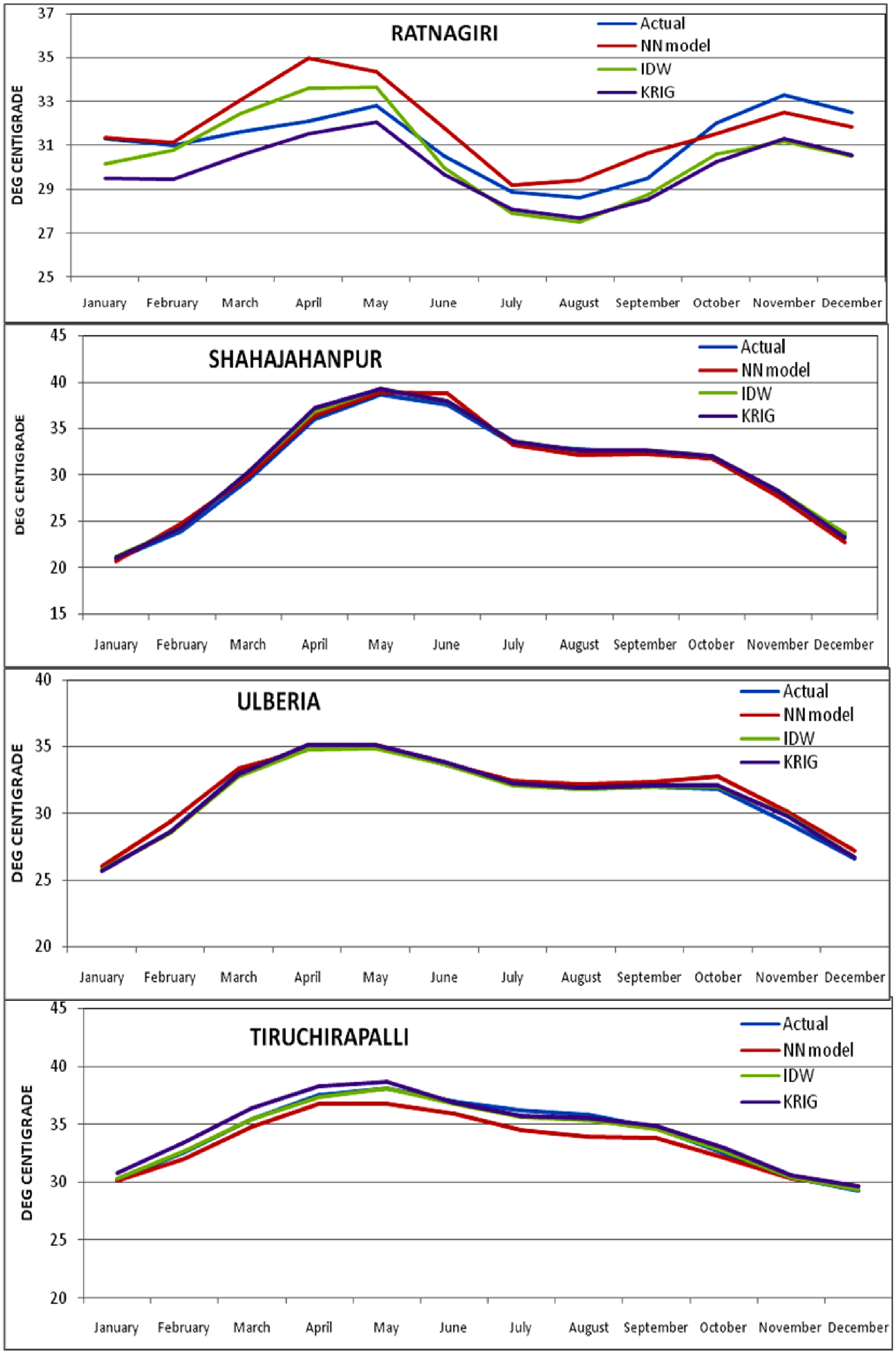

Fig. 11. Comparison of the performances of NN interpolation model (NN) along with IDW and Kriging methods for maximum temperature for all the twelve months for four low elevation stations

maximum and minimum temperatures with elevation. Our training and validation sets of stations thus represent all types of stations. To validate the neural network interpolation model and compare with inverse distance weighting and Kriging methods for high elevation stations, we have taken four stations *viz*., Srinagar (elevation 1587m), Shimla (elevation 2202m), Tadong (elevation 1322m) and Shillong (elevation 1598m) from the validation set of stations. As stated earlier cross validation method is used to generate the interpolate values at these four points for with inverse distance weighting and Kriging methods. Fig. 8 and Fig. 10 show the performances of the neural network interpolation models for these four stations for the twelve months for minimum and maximum temperatures respectively. Model values are very close to the actual values for Srinagar and Tadong and close to Shilong for minimum temperature. For Shimla the model values are lower side of the actual values for all the months. However the monthly values are very close to the actual values for all the four stations in the case of maximum temperature except for the month of July for the station Shillong. For all the four stations both IDW and Kriging methods values are almost showing similar results with differences from actual values for than the differences between neural network values and actual for all the months and for both maximum and minimum temperatures. However for Tadong though the values from both IDW and Kriging methods are closer to the actual but even neural network values are closer for all the months for minimum.

4.6. *Validation and comparison of the Neural Network interpolation model for low elevation stations*

For comparison of the performances of neural network interpolation model with the two most widely used spatial interpolation schemes and validation of the models, we have selected four low elevation stations *viz*., Ratnagiri (elevation 67m) of Maharashtra, Shahjanpur (elevation 155m) of Uttar Pradesh, Uluberia (elevation 5m) of West Bengal and Tiruchirapalli (elevation 88m) of Tamilnadu from the validation set of stations. All the four low elevation stations are taken from west, north, east and southern parts of the country in order to have a good representation.

Fig. 9 and Fig. 11 show the performances of the neural network interpolation models for these four stations for the twelve months for minimum and maximum temperatures respectively. The results are also compared with the values from inverse distance and kriging methods. Model values are very close to the actual values very close to the actual values for all the four stations and for both minimum and maximum temperatures. Except for Ratnagiri, values from all the models *i.e.* neural network, inverse distance weighting and kriging methods are very close to the actual for both maximum and minimum temperatures. In Ratnagiri, for minimum temperature, both neural network and kriging methods give very close results to the actual while inverse distance weighting differs much from the actual. Differences of values from actual in maximum temperature for Ratnagiri are more for the month of March, April and May for neural network model than the other two models.

5. Conclusions

To get the local effects of climate information, several techniques have been developed to downscale climate data generated by GCMs. Artificial Neural Network techniques have already been used to interpolate temperature data from a grid structure to interior points with a high degree of accuracy. Information on climate normal at a place is widely required by all the sectors of our life. This information cannot be given due to unavailability of observatory at the place demanded by the users. Standard spatial interpolation techniques are used for generating interpolation for evenly spaced points i.e. grid points and are generally used to get spatial pattern of the climate variables. Our main purpose in this paper is to have climate information at any point, *i.e*., unevenly space point. Though theoretically inverse distance weighting, kriging and other methods can be used to have information at ungaged point or at a point where there is no observation or observatory, but neither of these methods so far have been applied in such a case due to limited accuracy. In other word these spatial interpolation techniques can not generate information at any point with reasonable accuracy but can be applicable to generate gridded data set which gives climate features over larger area. Even ANN is used in this category. But in ANN there is flexibility of including more parameters as inputs. Since temperature is invariably influenced by elevation, latitude etc we have used latitude, longitude and elevation as inputs and temperature as output. From the known set of input output pattern, we have extracted the relationship between input and output by using standard training algorithm of ANN (mentioned in the text).

The present model which uses Neural Network technique with training algorithm used Output Weight Optimization - Hidden Weight Optimization is able to give normal maximum and minimum temperature at a place where there is no observatory. The model performs significantly better than the standard interpolation techniques used for generating gridded data set from the station data. The inclusion of elevation data in the neural network model has improved the performance of the model for both maximum and minimum temperatures.

This can be clearly seen from Figs. 8 and 10 showing the performances for high elevation station. It may possible to improve the accuracy of the model by including more variables like northing (north map coordinate), directional distances to the coast, and urbanization in the input parameters.

Acknowledgements

The authors are thankful to the ADGM(R), India Meteorological Department, Pune for providing the latest climate normal (1971-2000) and also other facilities for carrying out the work. The authors also express their thanks to the referee for his suggestions which lead to better representation of the results.

References

Bishop, C. M., 1995, "Neural Networks for Pattern Recognition", Oxford, Clarendon Press.

Bogardi, I., Matyasovsky, I., Bardossy, A. and. Duckstein, L., 1993, "Application of a space–time stochastic model for daily precipitation using atmospheric circulation patterns", *J. Geophys. Res.*, **98** (D9), 16 653-16 667.

Caesar, J., Alexandar, L. and Vose, R., 2006, "Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set", *J. of Geophys. Res*., 111, DO5101, doi:10.1029/2005JD006280.

Cavazos, T., 1997, "Downscaling large-scale circulation to local winter rainfall in north-eastern Mexico", *Int. J. Climatol.,* **17,** 1069-1082.

Chen, H. H., Manry, M. T. and Chandrasekaran, H., 1999, "A neural network training algorithm utilizing multiple sets of linear equations", *Neurocomputing*, **25**, 1-3, 55-72.

Choi, J., Chung, U. and Yun, J. I., 2003, "Urban-effect correction to improve accuracy of spatially interpolated temperature estimates in Korea", *J. Appl. Meteor.*, **42,** 1711-1719.

Crane, R. G. and Hewitson, B., 1998, "Doubled CO_2 precipitation changes for the susquehanna basin: Down-scaling from the genesis general circulation model", *Int. J. Climatol.,* **18,** 65-76.

Cressie, N., 1991, "Statistics for spatial data". Wiley, New York, p900.

Davis, J. C., 1986, "Statistics and Data Analysis in Geology", 2nd edition, Wiley, New York.

Dickinson, R., Errico, R., Giorgi, F. and Bates, G., 1989, "A regional climate model for the western United States". *Climatic Change*, **15**, 383-422.

Fischer, M. M. and Gopal, S., 1994, "Artificial neural networks: A new approach to modeling interregional telecommunication flows", *J. Reg. Sci.*, **34**, 503-527.

Gardner, M. W. and Dorling, S. R., 1998, "Artificial neural networks (the multilayer perceptron)-A review of applications in the atmospheric sciences", *Atmos. Environ.*, **32**, 14/15, 2627-2636.

Giorgi, F. and. Mearns, L. O., 1991, "Approaches to the simulation of regional climate change: A review", *Rev. Geophys.,* **29**, 191-216.

Giorgi, F., Brodeur, C. S. and Bates, G., 1994, "Regional climate change scenarios over the United States produced with a nested regional climate model", *J. Climate*, **7**, 375-399.

Gopal, S. and Scuderi, L., 1995, "Application of artificial neural networks in climatology: A case study of sunspot prediction and solar climate trends", *Geogr. Anal.*, **27**, 1, 42-59.

Hewitson, B. and Crane, R. G., 1996, "Climate downscaling: Techniques and application", *Climate Res.*, **7**, 85-95.

Hewitson, B. and Crane, R. G., 1994, "Precipitation controls in southern Mexico", *Neural Nets: Applications in Geography*, B. Hewitson and R. Crane, Eds., Kluwer Academic Publishers, 121-143.

Hornik, K., Stinchcombe, M. and White, H., 1989, "Multilayer feedforward networks are universal approximators", *Neural Networks*, **2**, 359-366.

Jarvis, C. H. and Stuart, N., 2001, "A comparison between strategies for interpolating maximum and minimum daily air temperatures", *Journal of Applied Meteorology*, **40**, 1060-1074, b. "The interaction between guiding variable and interpolation method", *Journal of Applied Meteorology*, **40**, 1075-1084.

Johnson, G. L., Daly, C., Taylor, G. H. and Hanson, C. L., 2000, "Spatial variability and interpolation of stochastic weather simulation model parameters", *J. Appl. Meteor.*, **39**, 778-796.

Joubert, A. and Hewitson, B., 1997, "Simulating present and future climates of southern Africa using general circulation models", *Prog. Phys. Geogr.*, **21**, 51-76.

Karmakar, S., Kowar, M. K. and Guhathakurta, P., 2009, "Spatial Interpolation of Rainfall Variables using Artificial Neural Network", Proc. of ACM International Conference on Advancement of Computing, Communication and Control, ISBN ACM 978-1-60558-351-8, ACM New York, NY, USA, 547-552.

Katz, R., 1977, "Assessing the impact of climatic change on food production", *Climatic Change*, **1**, 85-96.

Kiktev, D., Sexton, D. M. H., Alexander, L. and Folland, C. K., 2003, "Comparison of modeled and observed trends in indices of daily climate extremes", *J. Climate*, **16**, 3560-3571.

Krige, D. G., 1951, "A statistical approach to some mine valuations problems at the Witwatersrand", *J. Chemical, Metallurgical and Mining Society of South Africa*, **52**, 119-139.

Kurtzman, D. and Kadmon, R., 1999, "Mapping of temperature variables in Israel: A comparison of different interpolation methods", *Climate Res.*, **13**, 33-43.

Matyasovsky, I., Bogardi, I., Bardossy, A. and Duckstein, L., 1994, "Local temperature estimation under climatic change", *Theor. Appl. Climatol.*, **50**, 1–2, 1-13.

Matheron, G., 1963, "Principles of geostatistics", *Economic Geology*, **58**, 1246-1266.

McGinnis, D. L., 1997, "Estimating climate-change impacts on Colorado plateau snowpack using downscaling methods", *Prof. Geogr.*, **49**, 1, 117-125.

Meyers, D. E., 1994, “Spatial interpolation: An overview”, *Geoderma*, **62**, 17-28.

Mitra, A. K., Das Gupta, M., Singh, S. V. and Krishnamurti, T. N., 2003, “Daily rainfall for the Indian monsoon region from merged satellite and rain gauge values: Large scale analysis from real time data”, *J. Hyrdrometeorology*, **4**, 769-781.

New, M. G., Hulme M. and Jones, P. D., 2000, “Representing twentieth-century space time climate variability, part II: Development of 1901-96 monthly grids of terrestrial surface climate”, *J. Climate*, **13**, 2217-2238.

Palutikof, J. P., Winkler, J. A., Goodess, C. M. and Andresen, J. A., 1997, “The simulation of daily temperature time series from GCM output. Part I: Comparison of model data with observations”, *J Climate*, **10**, 2497-2513.

Piper, S. C. and Stewart, E. F., 1996, “A gridded global data set of daily temperature and precipitation for terrestrial biospheric modeling”, *Global Biogeochem. Cycles*, **10**, 752-782.

Price, D. T., McKenney, D. W., Nalder, I. A., Hutchinson, M. F. and Kesteven, J. L., 2000, “A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data”, *Agric. For. Meteor.*, **101**, 81-94.

Rajeevan, M., Bhate, J., Kale, J. D. and Lal, B., 2006, “High resolution daily gridded rainfall data for the Indian region: Analysis of break and active monsoon spells”, *Current Science*, **91**, 3, 296- 306.

Rumelhart, D. E., Hinton, G. E. and Williams, R. J., 1986, “Learning representations by back-propagating errors”, *Nature*, **323**, 533-536.

Rigol, J. P., Jarvis, C. H. and Stuart, N., 2001, “Artificial neural networks as a tool for spatial interpolation”, *International Journal of Geographical Information Science*, **15**, 4, 323-343.

Russo, J., and Zack, J., 1997, “Downscaling GCM output with a mesoscale model”, *J. Environ. Manage.*, **49**, 19-29.

Shepard, D., 1968, “A two-dimensional interpolation function for irregularly spaced data”, Proc. 1968 ACM Nat. Conf, 517-524.

Snell, S. E., Gopal, S. and Kaufmann, R. K., 2000, “Spatial interpolation of surface air temperatures using artificial neural networks: evaluating their use for downscaling GCMs”, *J. Climate*, **13**, 886-895.

Srivastava, A. K., Rajeevan, M. and Kshirsagar, S. R., 2009, “Development of a high resolution daily gridded temperature data set (1969–2005) for the Indian region”, *Atmospheric Science Letters*, **10**, 4, 249-254.

Wilby, R. L. and Wigley, T., 1997, “Downscaling general circulation model output: A review of methods and limitations”, *Prog. Phys. Geogr.*, **21**, 530-548.

Wilby, R. L., Wigley, T., Conway, D., Jones, P., Hewitson, B., Main, J. and Wilks, D., 1998, “Statistical downscaling of general circulation model output: A comparison of methods”, *Water Resour. Res.*, **34**, 11, 2995-3008.

Werbos, P., 1974, “Beyond regression: New tools for prediction and analysis in the behavioral sciences”, Ph.D. dissertation, Dept. of Applied Mathematics, Harvard University, p453. [Available from Harvard University Archives, Pusey Library, Harvard University, Cambridge, MA 01238].

Xie, P., Yatagai, A., Chen, M., Hayasaka, T., Fukushima, Y., Liu, C. and Yang, S., 2007, “A gauge-based analysis of daily precipitation over East Asia”, *Journal of Hydrometeorology*, **8**, 607-627.

Yatagai, A., Xie, P. and Kitoh, A, 2005, “Utilization of a new gauge-based daily precipitation dataset over monsoon Asia for validation of the daily precipitation climatology simulated by the MRI/JMA 20 km mesh AGCM”, SOLA, 1, 193- 196, doi:10.2151/ sola.2005-050.

Yu, C. and Manry, M. T., 2002, “A modified hidden weight optimization algorithm for feed-forward neural networks”, *Thirty-Sixth Asilomar Conference on Signals, Systems & Computers*, **1**, 1034-1038.

Zhang, M. and Scofield, R. A., 1994, “Artificial neural network techniques for estimating heavy convective rainfall and recognizing cloud mergers from satellite data”, *Int. J. Remote Sens.*, **15**, 6, 3241-3261.