

Research Article

Simulation of Daily Rainfall from Concurrent Meteorological Parameters over Core Monsoon Region of India: A Novel Approach

Utpal Misra ^{1,2}, Atri Deshamukhya,² Sanjay Sharma ³ and Srimanta Pal⁴

¹Patkai Christian College, Dimapur 797103, India

²Assam University, Silchar 788011, India

³Kohima Science College, Jotsoma, Kohima 797002, India

⁴Indian Statistical Institute, Kolkata, India

Correspondence should be addressed to Sanjay Sharma; sanjay_sharma11@hotmail.com

Received 13 January 2018; Revised 30 April 2018; Accepted 27 May 2018; Published 25 July 2018

Academic Editor: Stefania Bonafoni

Copyright © 2018 Utpal Misra et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the present work, daily rainfall is simulated over the core monsoon region of India by using a feedforward multilayer perceptron (MLP) model. Daily rainfall is found to be optimally dependent on four concurrent meteorological parameters, namely, geopotential height, specific humidity, zonal, and meridional wind at 1000 mb, 925 mb, 850 mb, and 700 mb pressure levels during 00, 06, 12, and 18 Greenwich Mean Time (GMT). The architecture of the optimized feedforward MLP model consists of 64 nodes in the input layer, 10 nodes in the hidden layer, and 1 node in the output layer. The results from the model are compared with the 3B42 (version 7) rainfall product. In terms of root mean square error (rmse) and correlation coefficient (cc), the model is performing better compared to the satellite-derived 3B42 rainfall product, whereas in terms of bias, the performance of the 3B42 product is better compared to the model. The weight matrices of the feedforward MLP model are estimated at a particular location (22.5°N, 82.5°E). These weight matrices are able to simulate daily rainfall at neighbourhood locations also with reasonably good accuracy with cc in the range of 0.41 to 0.55. The performance of the model improves in case of an aerial average of daily rainfall with significantly enhanced cc (0.72). The model is able to capture monthly and intraseasonal variation of rainfall with reasonably good accuracy, with cc of 0.88 and 0.68, respectively. The simulation model has a limitation that it is not able to simulate extreme high rainfall events (>60 mm/day). Overall, the developed model is performing reasonably well. This approach has a potential to be used as a rain parameterization scheme in the dynamical atmospheric and coupled models to simulate daily rainfall. Nevertheless, the present approach can also be used for multistep prediction of rainfall.

1. Introduction

Rainfall is a core component of the hydrological cycle. The knowledge of rainfall at different time scales is of paramount importance for many scientific and operational applications. Over continental regions, operational measurement of rainfall in general, is carried out by ground-based rain gauges. Over the ocean, relatively few rain gauges are available at buoy platforms. Due to its nature of point observation and limited access, rainfall measurement from rain gauges is supplemented by radars [1], satellites [2], rain parameterization schemes of physical models [3], and empirical statistical models [4]. Although the merged rainfall

product is a new trend where precipitation estimation is based on the combined use of satellite, rain gauge, and the numerical weather prediction (NWP) model rain output, the fundamental issue of accuracy of rain product of each component is still a topic of active research. It is further pointed out that at finer scales, merged rainfall product performs better compared to model output, whereas model rainfall products are better at larger spatial (>10⁶ km²) and temporal (monthly and longer) averaged scales [5].

The nature of height profiles of temperature and humidity are important parameters for the development of moist convective systems in the atmosphere, which culminate in rainfall [6]. Many researchers have tried to find

a possible association of different meteorological parameters with rainfall, such as temperature [7]; geopotential height at 850 mb [8]; relative humidity at 700 mb; specific humidity; perceptible water, zonal, and vertical components of wind velocity [9]; meridional wind [10]; 850 mb zonal wind gradient anomaly [11]; wet day frequency; minimum and maximum temperature; and cloud cover [12].

In recent years, artificial neural networks (ANNs), under nonparametric category, have found their ways to solve many problems related to rainfall such as rainfall forecasting [13–20], rainfall-runoff model [21–23], rainfall estimation by radars [1, 24, 25] and satellites [26–31], and temporal and spatial rainfall disaggregation [32, 33]. The advantage of an ANN approach is that it can be used to develop a functional relationship, including a nonlinear relationship, amongst the various parameters of the process under study even in the absence of full understanding of its mathematical model [34]. The ANN techniques such as feedforward, feedback, and competitive neural networks are extensively used for rainfall forecasting at different time scales such as, yearly [35–37], monthly/seasonal [13, 14, 38–44], and weekly/daily basis [45–47]. By using the soft computing techniques, the rainfall forecasting can be categorized in two groups: either by using historical time series rainfall data [48] or by using historical time series data of meteorological variables [45]. Over the Indian region, ANN techniques have been utilized extensively for forecasting of long range Indian summer monsoon rainfall with time lagged climatic indices—rainfall relationship approach [49] or from the stochastic time series data of rainfall [35, 41]. The simulation of daily rainfall by using height profiles of concurrent meteorological parameters is still to be seen over the Indian monsoon region in particular and over the globe in general. This aspect needs an attention, and the present work is an effort in this direction. The main objective of the study is to identify suitable concurrent meteorological parameters with respect to pressure depth in the atmosphere and to simulate daily rainfall by using these identified meteorological parameters. For this purpose, a feedforward multilayer perceptron (MLP) model is considered. The simulation is carried out over the core monsoon region of India. The manuscript consists of five sections. The description of the study region and the characteristics of meteorological parameters over the study region are provided in Section 2. In Section 3, the methodology of the proposed work is outlined. The results are provided in Section 4. Section 5 consists of summary, discussion, and conclusion.

2. Study Region and Characteristics of Meteorological Parameters over the Region

2.1. Study Region. The study has been carried out within the Indian core monsoon region as shown in Figure 1. The simulation study is carried out on a grid point 22.5°N, 82.5°E, and results are tested at four neighbourhood locations (22.5°N, 77.5°E; 22.5°N, 80.0°E; 22.5°N, 85.0°E; 22.5°N, 87.5°E). The study region consists of the central India plateau region coupled with the Chota Nagpur plateau and the plain

of the West Bengal. Over the study region, terrain elevation is varying from mean sea level to around 2 km of height.

2.2. Characteristics of Rainfall over the Region. In order to have a preliminary information about the number of rainy days during the study period (1953–2004) at 22.5°N, 82.5°E geolocation, gridded daily rainfall data [50] is utilized. The whole rainfall data set is divided into two parts, namely, 1953–1994 (for training and testing of a proposed simulation model) and 1995–2004 (for validation of the model). The training (testing) period consists of total 11,505 (3835) days, out of which around 25% (26%) are rainy days. Similarly, validation period consists of a total 3653 days, out of which 30% are rainy days.

During 1953–1994, the climatology of monthly mean rainfall is shown in Figure 2. The annual mean rainfall is 1490 mm. The location receives maximum rainfall during the southwest monsoon season (June–Sep) with maximum mean rainfall in the month of July (450 mm) followed by August (415 mm). In the months of June and September, the mean rainfall is around 200 mm. There is a significant standard deviation in the rainfall during the monsoon season, varying in the range 70–190 mm with the maximum value in the month of June (190 mm). The significant standard deviation is attributed to the various climatic indices, namely, El Nino Southern Oscillation (ENSO), Indian Ocean Dipole (IOD), and so on, which affects the onset of monsoon and received rainfall.

Further, the interannual variation of the southwest monsoon rainfall during 1953–1994 is shown in Figure 3. Rainfall is varying in the range of 600 mm (in the year 1979) to 2600 mm (in the year 1961). A trend line is fitted. The equation of the best fit line is provided in the figure panel. The negative slope (−8.60) of the best fit line suggests an overall decreasing trend of the annual rainfall. In order to further explore the trend of the annual rainfall pattern, the interannual variation of the number of rainy days in different rainfall categories, namely, $0 < \text{rainfall (R)} \leq 20$ mm/day, $20 < \text{R} \leq 40$ mm/day, $40 < \text{R} \leq 60$ mm/day, and $\text{R} > 60$ mm/day are shown in Figures 4(a)–4(d), respectively. The equation of the trend line is provided in the respective figure panel. It is observed that characteristic of interannual trend is different for these categories. In the range of $0 < \text{R} \leq 20$ mm/day, it has an increasing trend (slope: +0.30). Whereas for the category $20 < \text{R} \leq 40$ mm/day and $\text{R} > 60$ mm/day, it has a decreasing trend, and for the range $40 < \text{R} \leq 60$ mm/day, apparently there is no trend. It is interesting to mention that in the year 1961, the year of maximum rainfall of 2600 mm (Figure 3), the rainfall in the range of $40 < \text{R} \leq 60$ mm/day and $\text{R} > 60$ mm/day had significant contribution in the overall rain accumulation. Overall, the result suggests that over the grid point, the number of rainy days with $0 < \text{R} \leq 20$ mm/day are increasing, whereas the number of days with $\text{R} > 20$ mm/day are decreasing. Further, during the training and testing (1953–1994) and validation (1995–2004) periods, the percentage occurrence of rainy days for various categories is provided in Table 1. During training, testing, and validation periods, the

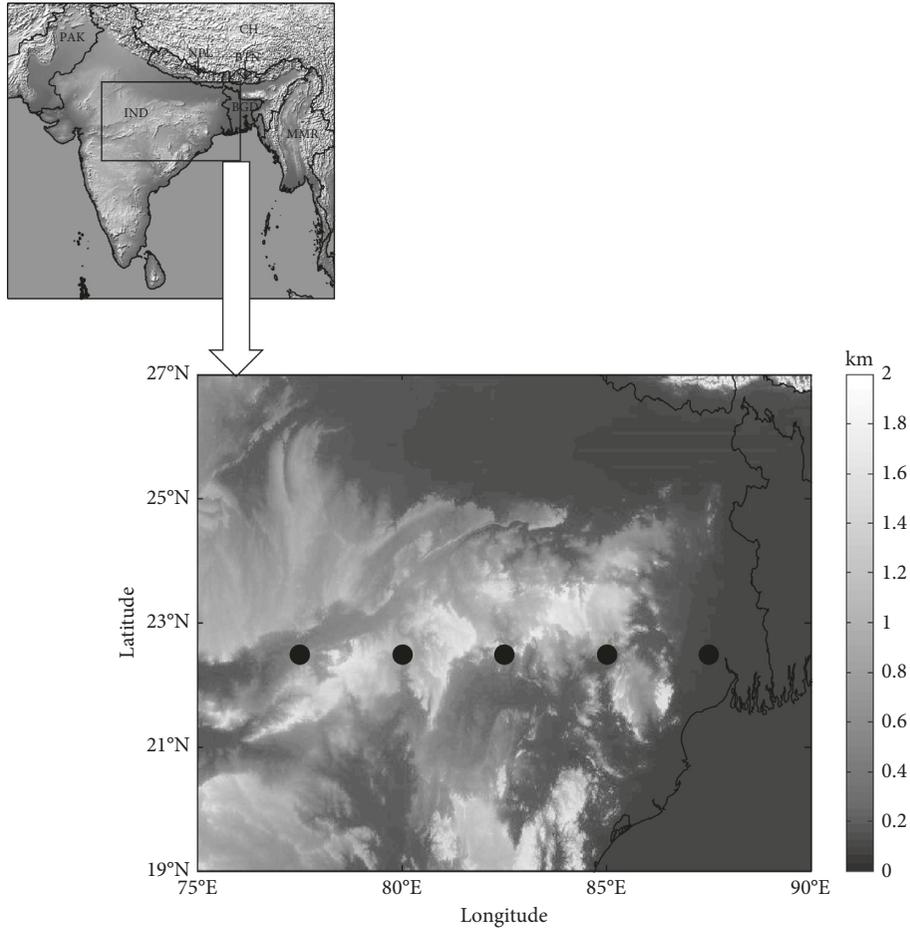


FIGURE 1: Geographical map of the study region. IND: India; CH: China; BTN: Bhutan; MMR: Myanmar; PAK: Pakistan; NPL: Nepal; BGD: Bangladesh.

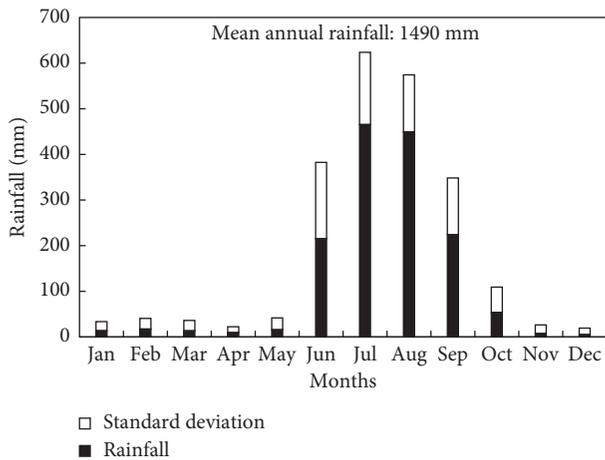


FIGURE 2: Climatology of the monthly variation of rainfall at 22.5°N, 82.5°E.

majority of the rainy days (around 24%, 25%, and 30% resp.) are in the category of $0 < R \leq 60$ mm/day and around 1.19%, 1.25%, and 0.71% are in the category of $R > 60$ mm/day. The result suggests that rainfall characteristics are similar in nature during these periods.

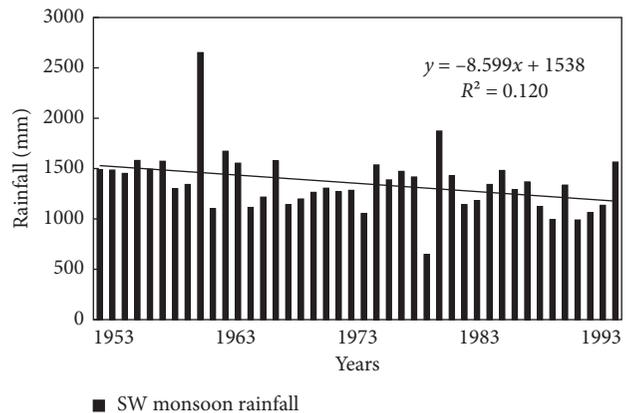


FIGURE 3: Interannual variation of rainfall at 22.5°N, 82.5°E, during 1953–1994.

Further, it is to be mentioned that the temporal and spatial inhomogeneity in the meteorological data may be present due to various factors such as due to the method of data collection, change in topography, and change in global or regional climatic indices. Therefore, it is important to test the homogeneity of the data set. The homogeneity test can be done in two ways: absolute method and relative method [51].

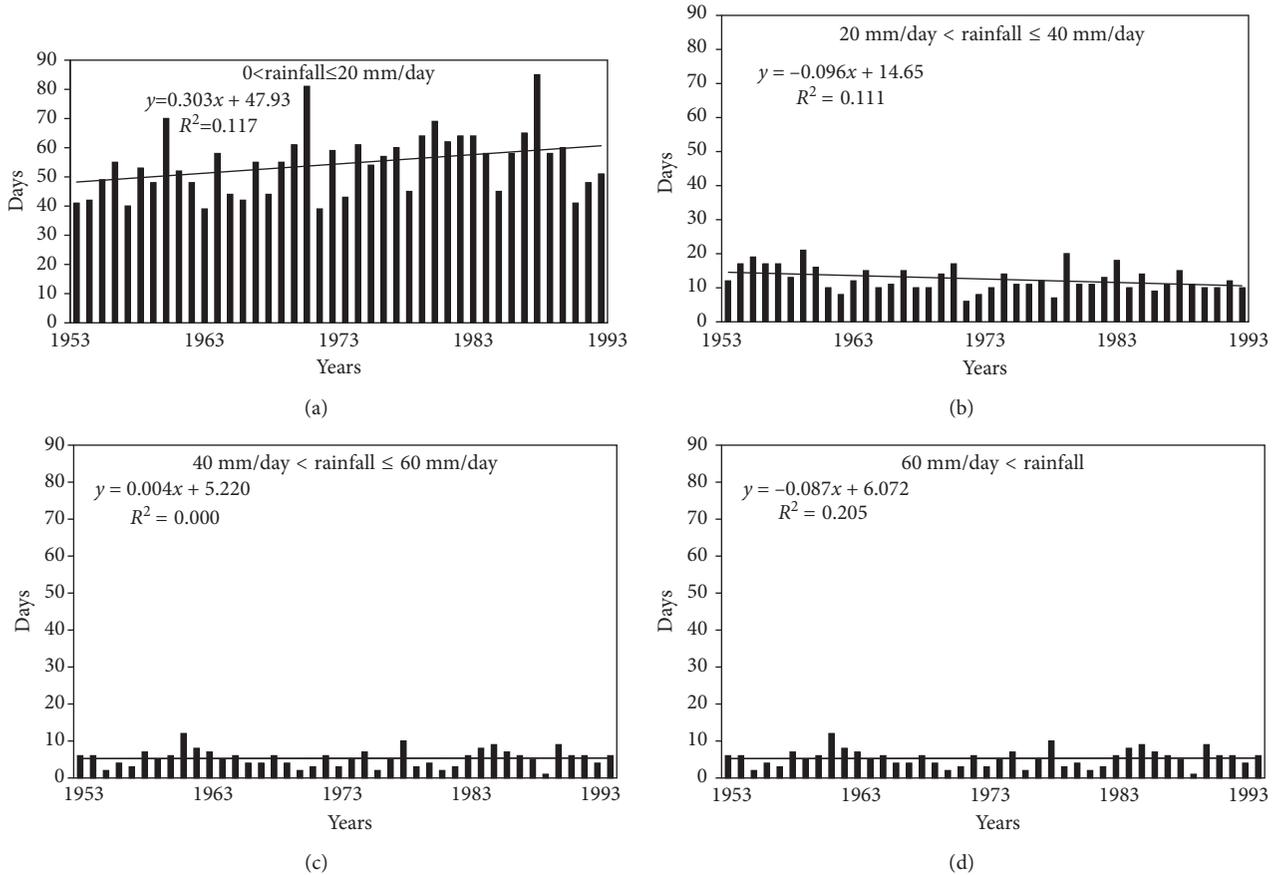


FIGURE 4: Interannual variation of number of days with rainfall. (a) $0 < \text{rainfall} \leq 20 \text{ mm/day}$. (b) $20 < \text{rainfall} \leq 40 \text{ mm/day}$. (c) $40 < \text{rainfall} \leq 60 \text{ mm/day}$. (d) $\text{Rainfall} > 60 \text{ mm/day}$.

TABLE 1: Characteristics of rainy days during the model training, testing, and validation periods.

Range (mm/day)	Number of Rainy days (%)		
	During 1953–1994		During 1995–2004
	Training*	Testing*	Validation*
>0 to ≤ 20	2121 (18.44%)	745 (19.43%)	907 (24.83%)
>20 to ≤ 40	435 (3.78%)	144 (3.75%)	134 (3.67%)
>40 to ≤ 60	182 (1.58%)	59 (1.54%)	44 (1.21%)
>60 to ≤ 80	70 (0.61%)	30 (0.78%)	16 (0.43%)
>80 to ≤ 100	28 (0.24%)	7 (0.18)	5 (0.14%)
>100	39 (0.34%)	11 (0.29%)	5 (0.14%)

*Training: total number of days: 11,505; *testing: total number of days: 3835; *validation: total number of days: 3653.

For absolute method, the test is applied for each station separately, and in relative method, neighbouring stations are also taken into account. Both the methods have limitations and advantages. As the utilized daily rainfall data is at a single location, the absolute method is considered for the homogeneity test. We tested the homogeneity of daily rainfall data for different months using three tests namely, Pettitt's test [52], standard normal homogeneity test (SNHT) [53], and the Buishand test [54] at 95% ($\alpha = 0.05$) confidence level. The results are presented in Table 2. If the p value is more than 0.05, the data is considered as homogeneous, otherwise

TABLE 2: Homogeneity test of daily rain data at the training location (1953–94).

Month	p value (Pettitt's test)	p value (SNHT test)	p value (Buishand's test)
Jan	0.02	0.33	0.17
Feb	0.03	0.4	0.42
Mar	0.09	0.19	0.02
Apr	0.02	0.32	0.04
May	0.0011	0.42	0.25
Jun	0.0001	0.03	0.34
Jul	0.47	0.01	0.01
Aug	0.08	0.12	0.09
Sep	0.01	0.11	0.17
Oct	0.08	0.19	0.03
Nov	0.29	0.19	0.06
Dec	0.55	0.48	0.23

$\alpha = 0.05$ (95% confidence level). If p value is more than 0.05, the data cannot be considered as inhomogeneous.

inhomogeneous. From the table, it is observed that out of three tests, two tests show the inhomogeneous result for the months of April, June, and July; one test shows the inhomogeneous result for the months of January, February, March, May, September, and October. On the other hand, daily rainfall data for the months of August, November, and December show homogeneous results for the entire three

tests. The inhomogeneity in the rainfall data in the months of June and July during the monsoon season is attributed to the variability in the onset of the monsoon.

2.3. Characteristics of Meteorological Parameters over the Location. Similar to rainfall, the characteristics of meteorological parameters are also studied at the same location. The meteorological parameters are considered from the NCEP/NCAR reanalyzed data set [55]. It is a gridded data set with a grid size of $2.5^\circ \times 2.5^\circ$. The data set consists of observational as well as NWP output meteorological parameters. The meteorological parameters at each grid point are categorized as “A”, “B”, “C”, and “D” [55]. The class “A” indicates that the gridded meteorological parameters are primarily the observed values; therefore, the parameters under this class are the most consistent. The parameters under class “B” are basically derived with the help of observed and model values. The parameters under “C” are fully derived from the model, whereas those under the class “D” are climatological data.

At the outset, four meteorological parameters of level “A” (atmospheric temperature, geopotential height, u -wind, and v -wind) and two parameters of level B (specific humidity and pressure vertical velocity (PVV) (ω)) are considered at eight pressure levels (1000, 925, 850, 700, 600, 500, 400, and 300 mb) during the four times a day (00, 06, 12, and 18 GMT). The climatology of mean monthly variation of the six meteorological parameters, namely, atmospheric temperature, geopotential height, specific humidity, u -wind (zonal wind), v -wind (meridional wind), and PVV (ω) at eight pressure levels are shown in Figures 5(a)–5(f), respectively. For atmospheric temperature, the maximum value is in the month of May (premonsoon) at 1000–700 mb pressure levels and at 600–300 mb pressure levels, its maximum value is in the month of July and August (monsoon). The higher value of temperature at higher pressure levels during the southwest monsoon is attributed to the release of latent heat by virtue of phase conversion of water in the convective clouds present during the southwest monsoon season. The characteristics of the monthly variation of geopotential height (Figure 5(b)) show that over 1000 to 600 mb, the minimum geopotential height is during the monsoon season. On the other hand, at 400 and 300 mb, this feature is reversed; that is, the minimum height is observed during the winter (Jan). In case of specific humidity, it is always maximum during the monsoon months (July and August), with overall lesser values towards the lower pressure levels (or upper heights) (Figure 5(c)). The velocity vectors are characterized in terms of magnitude and direction. As far as sign convention of u -wind vector (zonal wind) is concerned, +ve sign indicates westerly wind, that is, wind flow from west to east and –ve sign indicates easterly flow, that is, wind flow from east to west. For u -wind (Figure 5(d)), it is observed that from the premonsoon to monsoon seasons, its values at lower pressure levels of 600, 500, 400, and 300 mb changes from westerly to easterly and again becomes westerly during postmonsoon season (Nov-Dec). During the premonsoon season, the maximum value

of westerly at 300 mb (around 30 m/sec) is attributed to the presence of the jet stream. Its value at higher pressure levels (1000, 925, and 850 mb) is though predominantly westerly, but it has a lesser magnitude varying in the range of 0–5 m/sec. For v -wind (meridional wind), vector +ve sign indicates southerly flow, that is, flow from south to north and –ve sign indicates northerly flow, that is, wind flow from north to south. Its seasonal variation at various pressure levels has different characteristics (Figure 5(e)). At 300 mb, wind vector is always southerly with varying magnitude, having a minimum value during the monsoon and the maximum value during the postmonsoon. It is further observed that, except at 300 mb, v -wind vector at all pressure levels is predominantly northerly and has minimum magnitude during the monsoon season. It has significant value during the premonsoon followed by a postmonsoon season. Similarly, sign convention of PVV vector (ω) is such that +ve sign denotes downward motion and –ve sign denotes upward motion. For ω also, at each pressure level, there is seasonal variation in its characteristic. At 600, 500, 400, and 300 mb pressure levels, ω value has downward air motion during the premonsoon and postmonsoon seasons, whereas it becomes upwards during the monsoon season (Figure 5(f)). On the other hand, at 700, 850, 925, and 1000 mb, the air flow is in the upward direction during the premonsoon which changes to downwards direction in the monsoon and the postmonsoon seasons. Overall, systematic variation is observed in the monthly variation of these meteorological parameters at various pressure levels as well as of ground rainfall. The results suggest that these meteorological parameters at the location may be utilized to simulate rainfall.

3. Methodology

The relative role of the considered meteorological parameters for the formation of precipitating system and subsequent rainfall is uncertain. On the other hand, an MLP network has capability to map nonlinear relation between input and output parameters. It is assumed that, if rainfall characteristics with respect to meteorological parameters at a place for a certain period are known, then the architecture of an MLP model can be designed and trained to generate corresponding weight matrices, and thereafter, rainfall can be simulated from the concurrent meteorological parameters by using the generated weight matrices. It is to be mentioned that in the present work, we have not tried a forecasting (of rainfall) approach, where generally a time series data of either rainfall or meteorological parameters are used as inputs and time-lagged rainfall data as an output for the training of a selected neural network architecture. When a time series or temporal sequence data are used in that case, a feedback neural network is a better option. For the present study, the approach is different in the sense that there is a simulation of rainfall with the help of concurrent meteorological parameters, where daily rainfall is a cumulative effect of the favorable properties of height profiles of the meteorological parameters at four times in a day. In this article, nature of the proposed objective falls under the pattern recognition of a static input and output data. For the

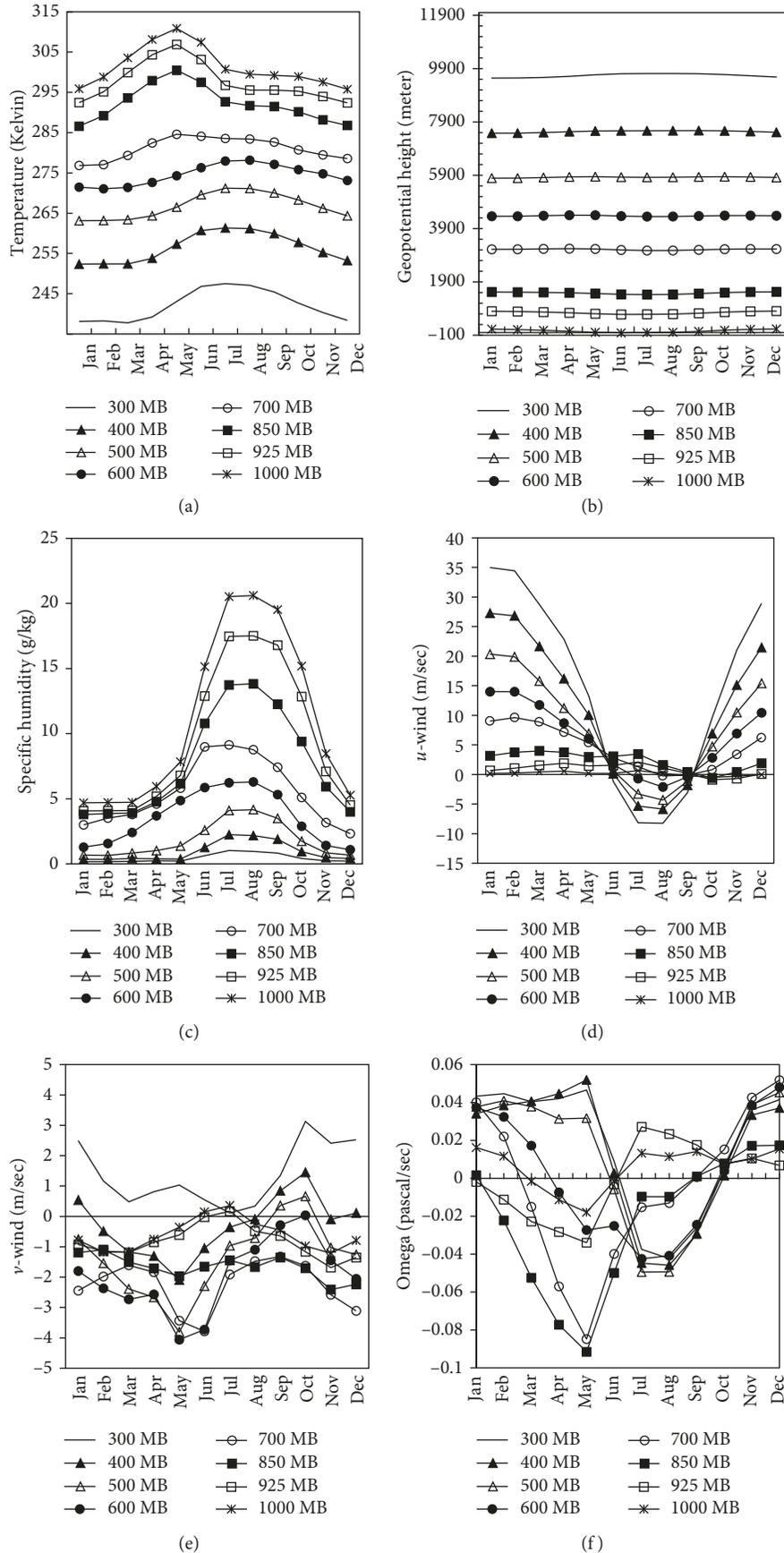


FIGURE 5: Climatology of monthly variation of (a) atmospheric temperature, (b) geopotential height, (c) specific humidity, (d) zonal wind (u -wind), (e) meridional wind (v -wind), and (f) pressure vertical velocity (ω).

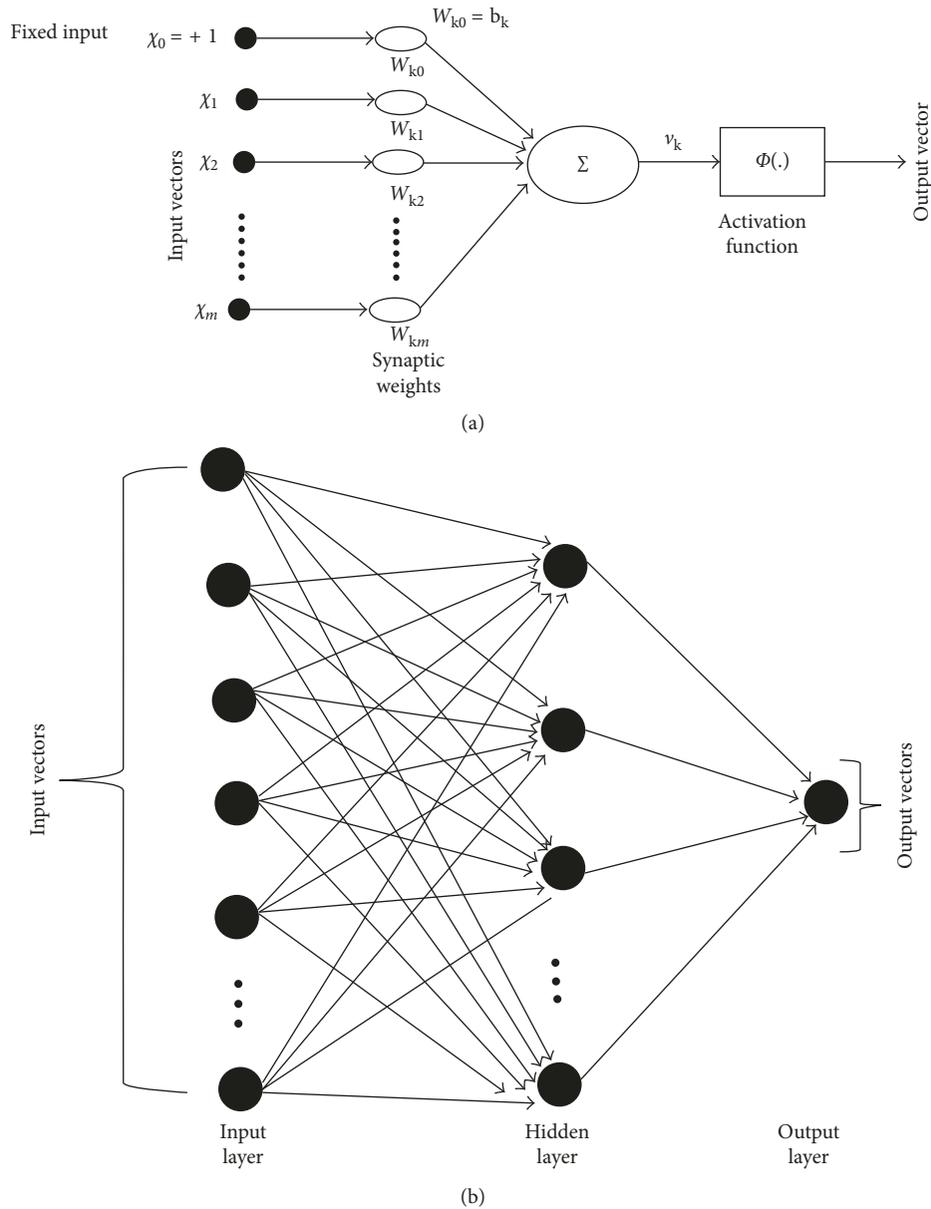


FIGURE 6: Schematic diagram of (a) an elementary neuron model with multiple inputs and (b) an architecture of multilayer perceptron model.

objectives of this nature, feedforward networks are reasonably suitable. Therefore, in the present study, a feedforward MLP network is considered to simulate daily rainfall. This model consists of an input and output layers and in between one or more hidden layer(s). The nodes between the two adjacent layers are interconnected, the value of the parameters in the subsequent layers are a weighted sum of the input values. The hidden layer helps to capture the nonlinear relation between the input and output parameters. A model of a neuron and an architecture of an MLP network are shown in Figures 6(a) and 6(b), respectively. The learning of an MLP is carried out primarily in the four steps, namely, (i) normalization of input and output vectors, (ii) estimation of activation function for hidden and output layers, (iii) estimation of

error for output and hidden layers, and (iv) adjustment of weights by many iterations, till the prescribed limit is achieved [56].

The training of the feedforward MLP is carried out on MATLAB platform. To simulate daily rainfall by using a feedforward MLP model, working methodology is planned into four steps, namely, (i) data preparation, (ii) identification of optimum depth of pressure level to be considered for the model, (iii) selection of a suitable number of input concurrent meteorological parameters for the model, and (iv) development of optimum weight matrices of an MLP model at 22.5°N, 82.5°E location. Thereafter, by using the developed weight matrices, daily rainfall is also simulated at neighboring locations. The detail is provided in the following subsections.

TABLE 3: Total number of vectors for the input layer of feedforward MLP.

Vectors	Parameter	Number of pressure levels (8)				Time (4)		Total vectors (A × B)
		(1000, 925, 850, 700, 600, 500, 400, and 300 mb)				(00, 06, 12, and 18 GMT)		
		(A)				(B)		
X1-X32	AT	8				4		32
X33-X64	GPH	8				4		32
X65-X96	SH	8				4		32
X97-X128	<i>u</i> -wind	8				4		32
X129-X160	<i>v</i> -wind	8				4		32
X161-X192	PVV	8				4		32
Total vectors								192

AT: atmospheric temperature; GPH: geopotential height; SH: specific humidity; *u*-wind: zonal wind; *v*-wind: meridional wind; PVV: pressure vertical wind.

3.1. Data Preparation for Training, Testing, and Validation of Feedforward MLP Model. From NCEP/NCAR reanalyzed data set [55], six meteorological parameters, namely, atmospheric temperature (AT), geopotential height (GPH), specific humidity (SH), *u*-wind, *v*-wind, and PVV (ω) are considered at eight pressure levels, (1000, 925, 850, 700, 600, 500, 400, and 300 mb) and at four times a day (at 00, 06, 12, and 18 GMT) at the geolocation 22.5°N, 82.5°E for the period 1953–2004. There are total 192 input vectors as provided in Table 3. For the same geolocation, daily rainfall data are considered from daily gridded rainfall data set [50]. Meteorological parameters are considered as input and daily rainfall is considered as output. The data set is separated into two groups, namely, (i) training and testing data set (1953 to 1994) and (ii) validation data set (1995 to 2004). The data set (both input and output) for the period 1953–1994 are randomized and then normalized within 0 and 1. All the input and output data sets are normalized between 0 and 1, by using the following formula:

$$X_N = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}, \quad (1)$$

where X_{\max} and X_{\min} are the maximum and minimum value of each input and output parameter, respectively, X_i is the specific value of each parameter for which we want to get the normalized value, and X_N is the normalized value that corresponds to X_i value. Therefore, in the data file, there is always a positive value between 0 and 1. Since actual minimum rainfall is not negative, model output of the rainfall is not negative.

The data sets for the period 1953–1994 are further divided into two parts: one part consists of 80% data points which are to be utilized for training purpose and the remaining 20% are to be utilized for testing purpose. The data set for the period 1995–2004 (validation data set) is also normalized between 0 and 1 without randomization.

3.2. Sensitivity Analysis of the Optimum Depth of Pressure Level. For this purpose, at the outset, six meteorological parameters at eight pressure levels, and at four times a day, are considered (Table 3). In order to find the optimum depth of the pressure level, a detailed sensitivity analysis is carried out for different slabs of the pressure level. These slabs are (1) at 1000 mb, (2) from 1000 to 925 mb, (3) from 1000 to 850 mb, (4) from 1000 to 700 mb, (5) from 1000 to 600 mb,

(6) from 1000 to 500 mb, (7) from 1000 to 400 mb, and (8) from 1000 to 300 mb. For slabs 1, 2, 3, 4, 5, 6, 7, and 8, the total number of meteorological parameters are 24, 48, 72, 96, 120, 144, 168, and 192, respectively (Table 4). A feedforward MLP model is trained and tested for the period 1953–1994 at 22.5°N, 82.5°E. For training purpose, the number of hidden layers is considered in the ratio of 10:1 with respect to total number of nodes in the input layer. The model is validated for the period 1995–2004 for each considered pressure depth. The simulated daily rainfall is compared with the observed daily rainfall (gridded rainfall data) in terms of root mean square error (rmse), bias, and correlation coefficient (cc) using the following relations:

$$\begin{aligned} \text{rmse} &= \sqrt{\left(\frac{1}{n} \sum_{i=1}^n (S_i - O_i)^2\right)}, \\ \text{bias} &= \frac{1}{n} \sum_{i=1}^n (S_i - O_i), \\ \text{cc} &= \frac{\sum_{i=1}^n (S_i - \bar{S})(O_i - \bar{O})}{n \sqrt{\sum_{i=1}^n (S_i - \bar{S})^2 \sum_{i=1}^n (O_i - \bar{O})^2}} \end{aligned} \quad (2)$$

where S_i is the model simulated rainfall, O_i is the observed rainfall, \bar{S} and \bar{O} are their arithmetic mean, and n is the number of data points.

The results are provided in Table 4. It is observed that for the slab with pressure levels from 1000 to 700 mb, the rmse is minimum and cc is maximum for both testing and validation data sets (testing: rmse: 11.82 mm/day and cc: 0.51. validation: rmse: 9.78 mm/day and cc: 0.48). Therefore, an atmospheric slab from 1000 to 700 mb (four levels) is considered as an optimum depth of pressure level to simulate the daily rainfall. Therefore, at this stage, the six meteorological parameters at four pressure levels and at four times a day constitute total 96 inputs to an MLP model.

3.3. Sensitivity Analysis for an Optimum Number of the Input Meteorological Parameters. Once the optimum depth of pressure level (1000–700 mb) is identified, the next step is to identify suitable meteorological parameters for the MLP model. A detailed sensitivity analysis is carried out in different combinations of meteorological parameters. For this purpose, twelve different combinations are tried as shown in

TABLE 4: Sensitivity analysis of MLP model with respect to input meteorological parameters at different pressure levels between 1000 and 300 mb.

Pressure level slabs (mb)	Number of meteorological parameters	Error statistics of simulated daily rainfall with respect to the observed daily rainfall					
		Testing data set			Validation data set		
		rmse (mm/day)	bias (mm/day)	cc	rmse (mm/day)	bias (mm/day)	cc
1000	24	12.00	2.46	0.46	12.08	-0.99	0.39
1000-925	48	11.88	2.37	0.49	14.45	-1.48	0.40
1000-850	72	11.87	2.37	0.50	9.79	1.49	0.47
1000-700	96	11.82	2.34	0.51	9.78	1.57	0.48
1000-600	120	12.09	2.55	0.48	10.06	1.61	0.43
1000-500	144	11.99	2.64	0.52	10.48	2.35	0.46
1000-400	168	12.15	2.75	0.48	9.99	0.53	0.42
1000-300	192	12.01	2.64	0.45	10.13	1.17	0.42

mb: millibar.

TABLE 5: Sensitivity analysis of ANN model for different combination of input meteorological parameters between the pressure levels 1000 and 700 mb.

Name of the meteorological parameters	Number of meteorological parameters	Error statistics of simulated daily rainfall with respect to the observed daily rainfall		
		rmse (mm/day)	bias (mm/day)	cc
Without AT	80	10.66	-1.4	0.48
Without GPH	80	21.47	-6.77	0.47
Without ω	80	10.97	-0.36	0.45
Without SH	80	12.36	-2.25	0.46
Without u -W	80	25.23	-3.49	0.34
Without v -W	80	11.16	-0.4	0.46
Without AT and ω	64	9.63	-0.9	0.49
Without AT and v -W	64	18.15	-2.13	0.32
Without ω and v -W	64	11.29	1.44	0.28
Without AT and SH	64	12.42	-2.35	0.46
Without ω and SH	64	16.16	-3.7	0.35

AT: atmospheric temperature; GPH: geopotential height; SH: specific humidity; ω vertical wind velocity; v -W: meridional wind; u -W: zonal wind.

Table 5. For each combination, the total number of input meteorological parameters to an MLP model is shown in the same table. An optimized MLP model is trained and tested for the period 1953-1994 at 22.5°N, 82.5°E and the MLP model is validated for the period 1995-2004 for each considered combinations. The error statistics between the simulated daily rainfall and observed daily rainfall is calculated for each case during the validation period, and it is presented in Table 5. From the error statistics, it is found that amongst all the cases, the minimum rmse and maximum cc (rmse: 9.63 mm/day, bias: -0.9 mm/day, cc: 0.49) are observed when, out of the six parameters, atmospheric temperature and PVV (ω) are excluded. Further, it is also observed that as an individual parameter, u -wind is the most crucial parameter, that is, without it, the rmse is maximized. Eventually, optimum number of input meteorological parameters at each pressure level is found to be four. These four parameters (geopotential height, specific humidity, u -wind, and v -wind) at four pressure levels (1000, 925, 850, and 700 mb) at four times a day (00, 06, 12, and 18 GMT) constitute total 64 inputs (Table 6) to a feedforward MLP network to develop its optimized weight matrices. Further, an analysis has also been carried out to check the consistency of the selected meteorological parameters. The selected

meteorological parameters from the NCEP/NCAR reanalyzed data set are compared with NOAA 20th century reanalysis data set [57] in terms of cc, and the results are presented in Table 7. It is observed that the calculated cc for individual meteorological parameters between these two data sets are pressure level-dependent. Overall, the maximum cc is observed for geopotential height, which is varying in the range 0.94-0.99 with the maximum cc at lower heights (1000, 925, and 850 mb level). It is followed by specific humidity which has a cc in the range 0.94-0.87 with higher cc at lower heights (1000 and 925 mb). For the wind parameters, v -wind (meridional wind) has the highest correlation (0.71-0.76) at lower heights (1000-850 mb), whereas in contrast, u -wind (zonal wind) has a maximum correlation (0.83) at a higher height (700 mb). Overall, it is observed that for the selected meteorological parameters, cc between these two data sets is significantly higher, which suggests that the considered parameters from NCEP/NCAR reanalyzed data set are consistent.

3.4. Estimation of Optimum Weight Matrices for the MLP Model. After identification of the optimum depth of pressure level and suitable meteorological parameters with the

TABLE 6: Finally selected meteorological parameters as an input to the MLP model.

Pressure level (mb)	Meteorological parameter	Time (GMT)	Number of inputs
1000	GPH, SH, u -W, v -W	00, 06, 12, 18	16
925	GPH, SH, u -W, v -W	00, 06, 12, 18	16
850	GPH, SH, u -W, v -W	00, 06, 12, 18	16
700	GPH, SH, u -W, v -W	00, 06, 12, 18	16

mb: millibar; GPH: geopotential height; SH: specific humidity; u -W: zonal wind; v -W: meridional wind; GMT: Greenwich Mean Time.

TABLE 7: Comparison of the selected input parameter (NCEP/N-CAR reanalysis data set) with the NOAA 20th century reanalysis data set.

Parameter	1000 mb	925 mb	850 mb	700 mb
	cc	cc	cc	cc
GPH	0.99	0.99	0.99	0.94
SH	0.94	0.94	0.93	0.87
u -W	0.63	0.76	0.76	0.83
v -W	0.71	0.74	0.76	0.65

mb: millibar; GPH: geopotential height; SH: specific humidity; u -W: zonal wind; v -W: meridional wind.

number of nodes in the hidden layer with a ratio of 10:1 with respect to number of input nodes, further experiments are conducted to find the optimum number of nodes in the hidden layer by trial and error method. The error statistics is provided in Table 8. It is observed that rmse and bias is least with 10 nodes in the hidden layer. Therefore, the optimized MLP network is a three-layer network, consisting of 64 nodes in the input layer, 10 nodes in the hidden layer, and one node in the output layer (64:10:1). The positive bias and negative bias imply that the simulated daily rainfall is over- and underestimated, respectively, compared to observed daily rainfall. In order to have a better visualization of the characteristics of the input meteorological parameters on a rainy day, a case study is provided for the simulation of rainfall on 17th August, 1995. The values of the input meteorological parameters are provided in Table 9. Here, the geopotential height and specific humidity are relatively higher in the evening hours, whereas the wind vectors are relatively less in the evening hours. On this particular day, the simulated rainfall and observed rainfall are found to be 10.82 mm/day and 9.6 mm/day. The simulated rainfall is closely matching the observed rainfall.

4. Results

The result section consists of the following three subsections: (i) simulation of daily rainfall and its year-wise comparison with observed and satellite derived daily rainfall at the training location during the validation period (1998–2004), (ii) simulation of daily rainfall at each neighborhood location and their aerial average within the core monsoon region during the validation period (1998–2004) and its year-wise comparison with observed and satellite derived rainfall, (iii) assessment of the simulated rainfall with respect to

TABLE 8: Selection of number of neuron of hidden layer for three-layer MLP model having 64 input nodes and 1 node in the output layer.

Number of nodes in hidden layer	Error statistics of simulated daily rainfall with respect to the observed daily rainfall		
	rmse (mm/day)	bias (mm/day)	cc
6	9.63	-0.9	0.49
50	10.60	-1.1	0.48
40	10.05	0.98	0.46
30	9.88	0.91	0.48
20	9.78	0.92	0.48
10	9.53	0.82	0.51
5	9.64	-0.88	0.48

TABLE 9: The meteorological parameters along with observed and simulated rainfall, on 17th August 1995.

8/17/1995	Observed rainfall = 9.6 mm		Simulated rainfall = 10.82 mm	
	0.0 GMT	6.00 GMT	12.00 GMT	18.00 GMT
	GPH700	3096	3109	3089
GPH850	1451	1465	1443	1467
GPH925	715	728	705	730
GPH1000	33	43	21	49
SH700	0.00808	0.00818	0.00835	0.00912
SH850	0.01403	0.0128	0.013	0.01421
SH925	0.01672	0.01757	0.01754	0.01738
SH1000	0.01917	0.02161	0.02132	0.02017
U700	-1.7	-3.5	-2.8	1.4
U850	-0.49998	-3.7	-3	1
U925	1.4	-2.2	-2.3	-0.79999
U1000	1.9	-1.1	-1.6	-0.99998
V700	-4.8	-5.3	-3.7	-2.6
V850	-4.9	-4.5	-1.2	-0.99998
V925	-3.1	-4.8	-1.4	-1.9
V1000	-0.79999	-3.8	-2	-1.6

GPH is in meter; SH is in g/kg; u and v are in meter/sec.

observed rainfall in terms of interannual variation of monthly rainfall and intraseasonal variation of the monsoon rainfall.

4.1. Simulation of Daily Rainfall and Its Year-Wise Comparison with the Observed and Satellite-Derived Daily Rainfall. The temporal variation of the simulated and observed daily rainfall during 1998–2004 is provided in Figure 7(a). It consists of total 2557 number of days. The daily rainfall is plotted against the Julian days. The overall error statistics of the simulated rainfall with respect to observed value is provided in the figure panel. It is observed that the simulated daily rainfall is able to capture the overall features of the monsoon rainfall as well as dry season reasonably well. The overall rmse, bias, and cc during this period are found to be 9.44 mm/day, 0.90 mm/day, and 0.52, respectively. The temporal variation shows that the MLP model could simulate the maximum rainfall up to 60 mm/day, which further suggests that the developed model is not able to capture the extreme rainfall events (>60 mm/day).

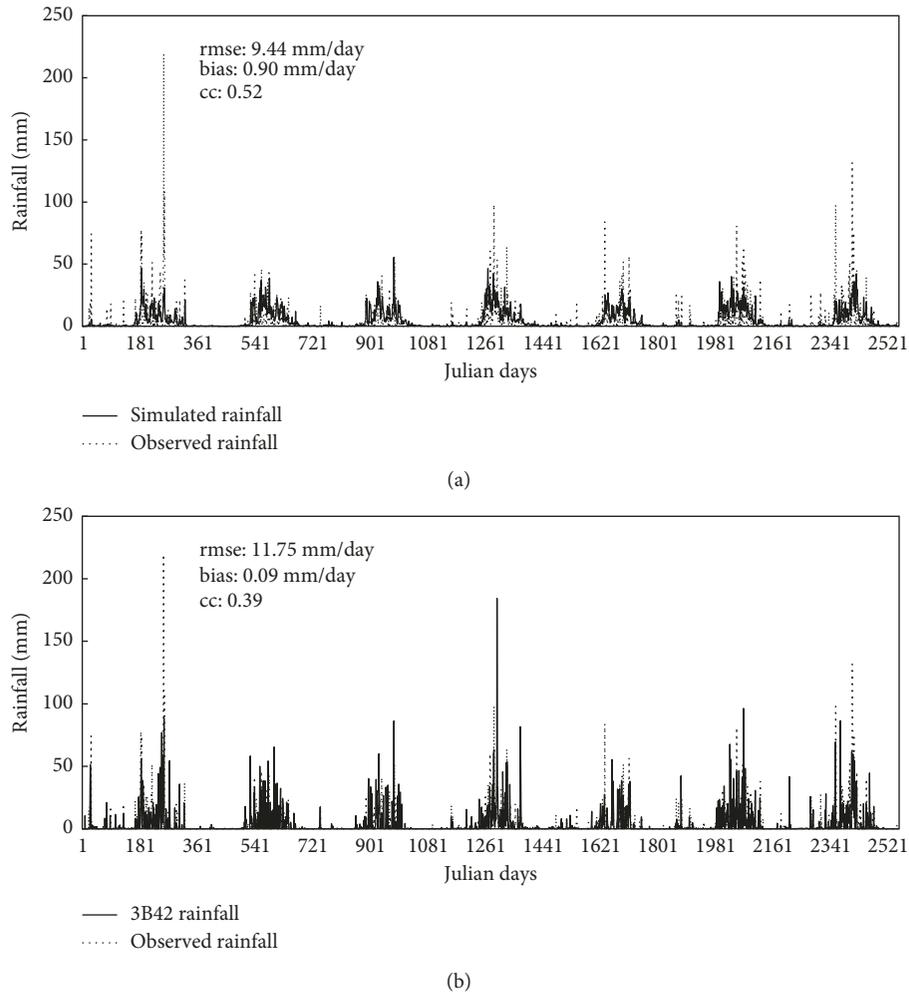


FIGURE 7: Variation of (a) simulated and observed daily rainfall and (b) 3B42 and observed daily rainfall at 22.5°N , 82.5°E during 1998–2004.

To assess the relative performance of the simulation model, Tropical Rainfall Measuring Mission (TRMM) Multisatellites Precipitation Analysis (TMPA) data product 3B42 (version 7) [58] is considered. It is a multisatellite rainfall product. It uses an optimal combination 2B-31 (TRMM product) 2A-12 (TRMM product) along with the Special Sensor Microwave Imager (SSM/I), Advanced Microwave Scanning Radiometer (AMSR), and Advance Microwave Sounding Unit (AMSU) rainfall data products to adjust infrared rainfall estimates from a geostationary satellite. This data product utilized TRMM Microwave Imager and TRMM-precipitation Radar as a calibration source to the Infrared (IR) estimated rainfall from a geostationary satellite [58]. For each day, it provides 3 hours of average rainfall during 0000 to 2100 UTC. The pixel resolution of the data product is 25 km. The temporal variation of the 3B42 and observed daily rainfall during 1998 to 2004 is provided in Figure 7(b). The daily rainfall is plotted against the Julian days. The overall error statistics is provided in the figure panel. It is observed that 3B42 rainfall is also able to capture the overall features of the monsoon rainfall and dry season reasonably well. It is observed that rmse (cc) of simulated daily rainfall is lesser (more) compared to 3B42 rainfall

product, whereas the bias for simulated daily rainfall is more compared to the 3B42 rainfall product. It thereby suggests that in terms of rmse and cc, the performance of the rainfall simulation model is relatively better compared to the 3B42 rainfall product.

Further, year-wise (during 1998–2004) error statistics for simulated versus observed daily rainfall at training location is provided in Table 10. There is interannual variation in the error statistics. The maximum rmse is observed in the year 1998, followed by 2004 with $-ve$ bias. This result is consistent with Figure 7(a), which shows that in both the years, extreme rainfall events are more, particularly in the year 1998 compared to other years. In 1998 and 2004, the observed daily rainfall is found up to around 225 and 125 mm/day, respectively (Figure 7(a)). It is attributed to the fact that the MLP model is not able to capture the extreme rainfall events. The model captured up to 50 mm/day in the year 2000. In terms of rmse (6.59 mm/day), the optimum simulation is observed in the year 1999, which is attributed to the absence of extremely high and low daily rainfall events. Overall, the cc is varying in the range of 0.50–0.62. Further, a comparative study has also been carried out between the 3B42 versus observed daily rainfall. The result is provided in Table 10.

TABLE 10: Error statistics of model simulated and 3B42 daily rainfall with respect to observed daily rainfall at the training location (22.5°N, 82.5°E) during the validation period.

Years	Simulated versus observed rainfall			3B42 versus observed rainfall		
	rmse (mm/day)	bias (mm/day)	cc	rmse (mm/day)	bias (mm/day)	cc
1998–2004	9.44	0.9	0.52	11.75	0.09	0.39
1998	14.47	−0.53	0.53	16.05	−0.34	0.40
1999	6.59	1.70	0.61	9.82	0.68	0.34
2000	6.64	2.29	0.57	8.93	0.89	0.33
2001	8.69	0.95	0.62	12.09	0.14	0.55
2002	7.05	1.03	0.54	8.72	−0.02	0.36
2003	8.40	1.11	0.55	12.11	−0.02	0.25
2004	11.45	−0.26	0.50	12.80	−0.67	0.43

TABLE 11: Error statistics of model simulated and 3B42 daily rainfall with respect to observed daily rainfall at different locations within the core monsoon region (1998–2004).

Location	Simulated versus observed rainfall			3B42 versus observed rainfall		
	rmse (mm/day)	bias (mm/day)	cc	rmse (mm/day)	bias (mm/day)	cc
22.5°N, 77.5°E	11.08	1.36	0.41	12.93	0.23	0.35
22.5°N, 80.0°E	8.52	−0.66	0.50	11.59	0.31	0.34
22.5°N, 82.5°E	9.44	0.90	0.52	11.75	0.09	0.39
22.5°N, 85.0°E	6.87	0.03	0.55	12.09	0.75	0.21
22.5°N, 87.5°E	9.73	−1.13	0.41	15.4	0.03	0.14

Similar to simulated versus observed rainfall statistics, the maximum rmse is observed in the year 1998, followed by 2004. Unlike the simulated rainfall, in terms of rmse (8.72 mm/day), the optimal estimation is observed in the year 2002. Overall result suggests that in terms of rmse and cc, the MLP model is performing better compared to 3B42 daily rainfall product, whereas in terms of bias 3B42 product is relatively better.

4.2. Simulation of Daily Rainfall at Other Locations within the Core Monsoon Region and Their Aerial Average. In order to assess the robustness of the estimated weight matrices of the MLP network, the same weight matrices are utilized to simulate daily rainfall at other neighbourhood locations also. The selected locations are 22.5°N–77.5°E; 22.5°N–80°E; 22.5°N–85°E; and 22.5°N–87.5°E. The error statistics of the simulated rainfall versus observed daily rainfall at different locations is provided in Table 11. For the simulated rainfall, the error statistics is reasonably consistent with the training location, where rmse, bias, and cc are varying in the range of 6.87 to 11.08 mm/day, −1.13 to 1.36 mm/day, and 0.41 to 0.55, respectively. For the 3B42 versus observed daily rainfall also, the error statistics is reasonably consistent, where rmse, bias, and cc are varying in the range of 11.59 to 15.40 mm/day, 0.03 to 0.75 mm/day, and 0.14 to 0.39, respectively. Overall, it is observed that at each location, the rmse (cc) for the simulated rainfall is less (more) compared to the 3B42 rainfall.

Thereafter, an areal average of the simulated and 3B42 rainfall within the core monsoon region is calculated and compared with the observed daily rainfall. The comparison of the temporal variation of an aerial average of simulated and 3B42 daily rainfall with the observed rainfall from 1998 to 2004 is shown in Figures 8(a) and 8(b), respectively. The daily rainfall is plotted against the Julian days. The error

statistics is provided in the respective figure panels. It is observed that the simulated rainfall is able to capture the average features of the monsoon rainfall as well as dry season reasonably well. It is observed that compared to point observations, performance of the MLP model improves significantly by virtue of enhanced cc. It is also observed that the performance of the MLP model is reasonably better compared to 3B42 rainfall product. Similar to the analysis at the training location, the year wise error statistics (during 1998–2004) of the aerial average of simulated and 3B42 daily rainfall with respect to the observed rainfall are provided in Table 12. In both the cases, there is interannual variation in the error statistics. For simulated versus observed daily rainfall, rmse, bias and cc are varying in the range of 3.49 to 6.8 mm/day, −0.53 to 0.89 mm/day and 0.62 to 0.77, respectively, with a minimum value of rmse and bias in 2000 and 2002–2003, respectively, and maximum value of cc in 2001. Further, for 3B42 versus observed daily rainfall rmse, bias, and cc are varying in the range of 5.05 to 8.26 mm/day, −0.05 to 0.79 mm/day, and 0.42 to 0.60, respectively, with a minimum value of rmse and bias in 2000 and 2002, respectively, and the maximum value of cc in 2004. The results are similar in nature with the error statistics of training location (Table 10). The maximum value of rmse (6.82 mm/day) and bias (−0.53 mm/day) of simulated rainfall in 1998 are attributed to the occurrence of extremely high rainfall events in that particular year, which recorded 1698 mm of rainfall, an amount, well above the average rainfall. The higher rmse and −ve bias in the simulated rainfall in the year 1998 further suggests that though the model could simulate normal rainfall with reasonably good accuracy, a simulation of extreme high rainfall events is still an issue. It is observed that with reference to the observed aerial average daily rainfall, the simulated rainfall has smaller rmse and higher cc compared to 3B42 daily rainfall. Overall

TABLE 12: Error statistics of aerial average model simulated and 3B42 daily rainfall with respect to observed daily rainfall within the core monsoon region (1998–2004).

Years	Simulated versus observed rainfall			3B42 versus observed rainfall		
	rmse (mm/day)	bias (mm/day)	cc	rmse (mm/day)	bias (mm/day)	cc
1998	6.82	-0.53	0.62	7.68	0.23	0.55
1999	4.89	-0.16	0.74	8.26	0.79	0.44
2000	3.49	0.89	0.76	5.05	0.37	0.56
2001	4.47	0.2	0.77	6.35	-0.05	0.55
2002	4.64	0.01	0.69	6.2	0.03	0.50
2003	4.13	0.00	0.75	7.04	0.23	0.42
2004	4.18	0.28	0.73	5.49	0.24	0.60

TABLE 13: Rainfall (R) characteristics at different rainfall ranges, 1995–2004.

Year	Number of rainy days for various range							Number of total rainy days	Annual rainfall (mm)	Monsoon rainfall (mm)
	R1	R2	R3	R4	R5	R6	R7			
1995	50	13	7	1	3	6	0	80	993	841
1996	30	13	8	7	2	2	5	67	1359	1306
1997	47	12	10	10	4	2	3	88	1379	1088
1998	42	21	8	4	3	3	6	87	1698	1372
1999	75	20	9	5	3	0	0	112	1053	946
2000	73	17	3	3	1	1	0	98	745	712
2001	105	27	11	5	1	2	4	155	1587	1402
2002	97	18	5	3	1	2	1	127	1032	903
2003	112	23	14	5	2	1	2	159	1509	1238
2004	94	18	11	5	3	2	5	138	1641	1429

mm: millimeter; R: rainfall; R1: $0 < R < 10$ (mm/day); R2: $10 \leq R < 20$ (mm/day); R3: $20 \leq R < 30$ (mm/day); R4: $30 \leq R < 40$ (mm/day); R5: $40 \leq R < 50$ (mm/day); R6: $50 \leq R < 60$ (mm/day); R7: $R \geq 60$ (mm/day).

result suggests that in terms of rmse and cc, the model is performing consistently better compared to 3B42 daily rainfall. It is to be mentioned that the mechanism for simulation of the rainfall by the model and the estimation of daily rainfall by the satellite sensors is completely different. The simulation method is based primarily on the outcome of the underlying physics of the convective systems, whereas the 3B42 rainfall is based on the interaction of infrared and microwave radiation with the hydrometeors.

4.3. Assessment of the Developed Model in Terms of Interannual Variation of Rainfall at Different Time Scales. To further assess the performance of the developed rainfall simulation model, rainfall accumulation at larger time scale, namely, monthly and intraseasonal (during active and break periods of the monsoon) is analysed and compared with the observed rainfall. The rainfall at larger time scale is calculated by applying the bias correction.

4.3.1. Interannual Variation of Monthly Rainfall. Interannual variation of monthly simulated rainfall and observed rainfall for 10 years during 1995–2004 is shown in Figure 9. The error statistics of the monthly simulated rainfall with respect to observed rainfall is provided in the figure panel. It is observed that the simulated rainfall is closely following the observed rainfall during the pre-monsoon, monsoon, and postmonsoon seasons with significant cc (0.88). Overall, the simulated rainfall is matching

reasonably well during the years 1995, 2001, 2002, 2003, and 2004 (particularly in the month of August). Whereas, during the years 1997, 1999, and 2000, the simulated rainfall is overestimated compared to the observed rainfall and in the years 1996, 1998, and 2004 (particularly in the month of June and July), it is underestimated. In order to understand the interannual variability of the error statistics of the monthly simulated rainfall with respect to the observed rainfall, the number of rainy days over seven different rainfall categories at the increment of 10 mm/day along with the total number of rainy days, annual rainfall, and monsoon rainfall during 1995–2004 are provided in Table 13. It is observed that there is significant overestimation in the year 2000. During this year, the observed rainfall was significantly low (712 mm) compared to other years. It is attributed to the limitation of the positive bias for the simulated rainfall during the presence of nonrainy days. The underestimation of rainfall during the years 1996, 1998, and 2004 (particularly in the month of June and July) is attributed to significant occurrence of higher daily rainfall events. During these three years, the number of rainfall days with >60 mm/days were 5, 6, and 5, respectively, whereas in the remaining years, it is varying in the range 0–4. Overall, the result suggests that the overestimation took place during the years when there was a shortfall of observed rainfall and underestimation took place during the years when the extreme high rainfall events (rainfall ≥ 60 mm/day) were prevalent. It is attributed to the limitation of the rainfall simulation model with respect to its positive bias, particularly in case of no rainfall situation and also its inability to simulate the extreme

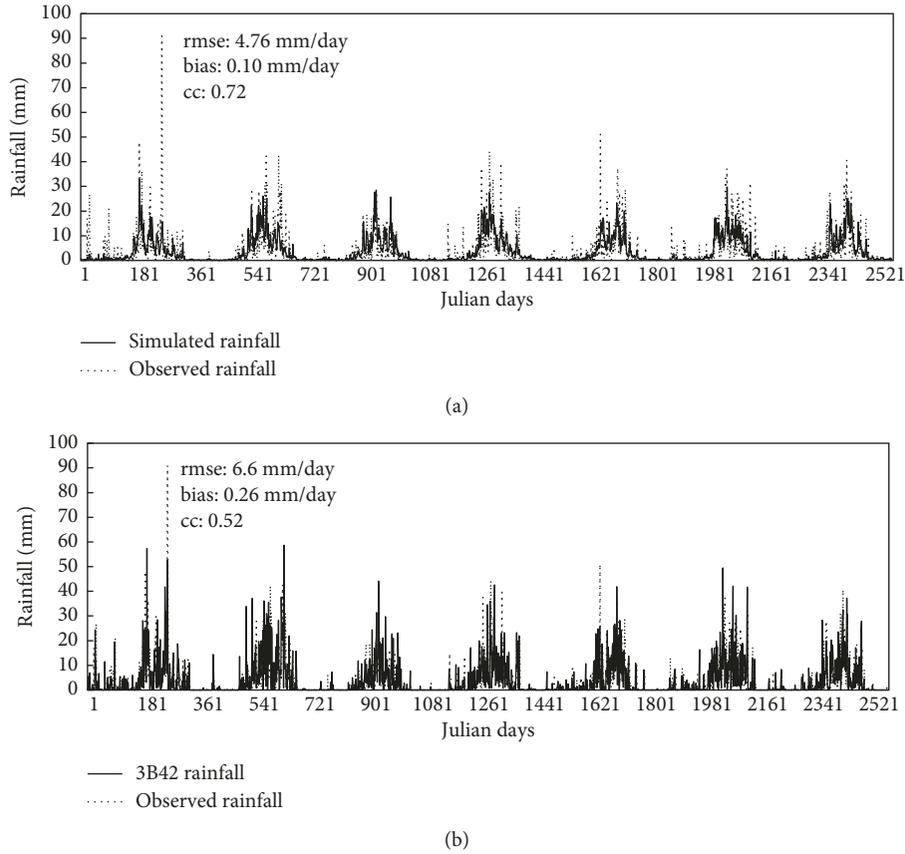


FIGURE 8: Variation of aerial average of (a) simulated and observed daily rainfall and (b) 3B42 and observed daily rainfall over the monsoon region during 1998–2004.

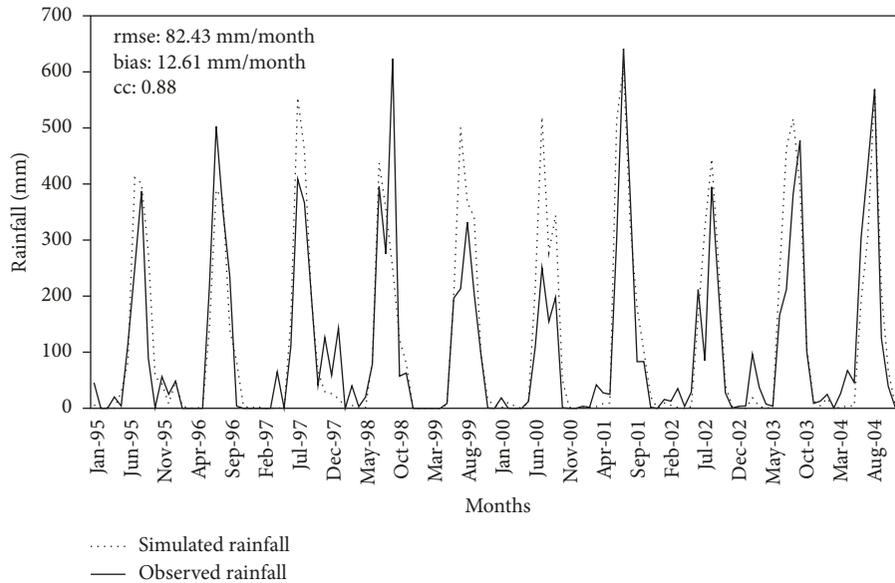


FIGURE 9: Variation of monthly simulated and observed rainfall during 1995–2004.

rainfall events. In terms of spatiotemporal averaging of the simulated rainfall, the performance of the MLP model improves considerably. The results also suggest that the error statistics is predominantly affected by the characteristics of rainfall.

4.3.2. *Intraseasonal Variation of Rainfall (Active-Break Days)*. The Indian summer monsoon is characterized by a strong intraseasonal variation of rainfall, indexed as an active and break periods [59]. During the active and break periods, the normalized anomaly of rainfall over the core

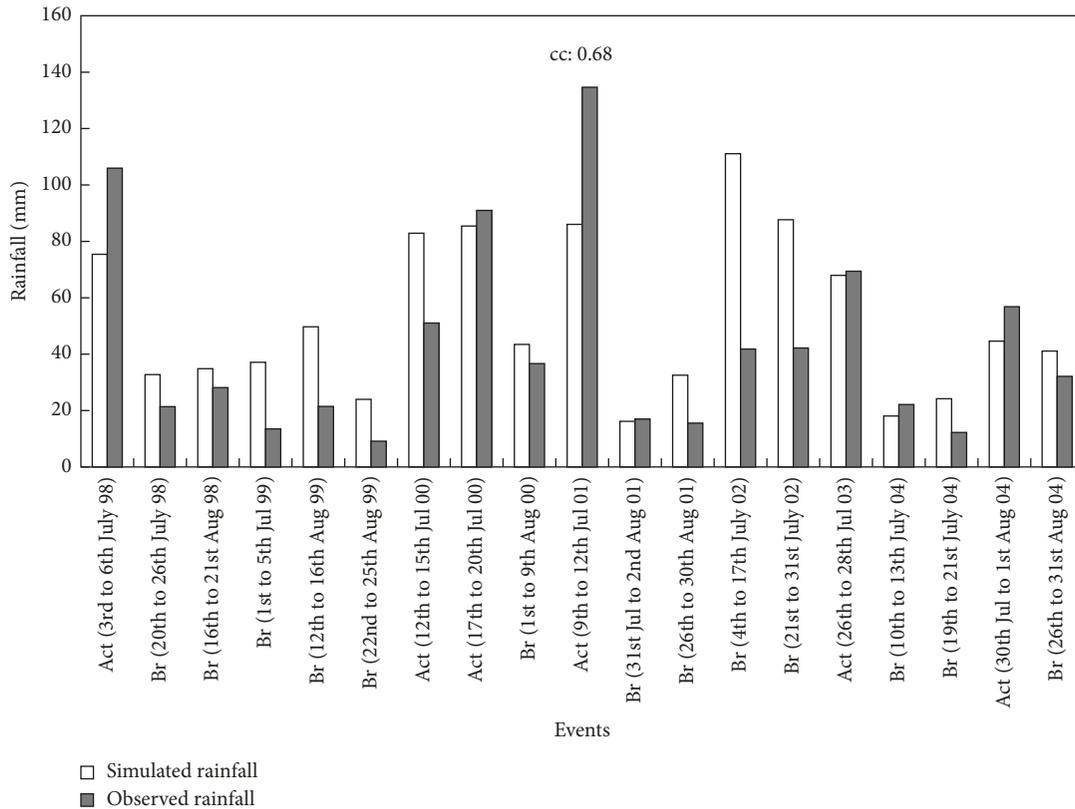


FIGURE 10: Interannual variation of simulated and observed rainfall during the active and break periods of southwest monsoon during 1998–2004.

monsoon region (18° – 28° N; 65° – 88° E) exceeds +1 or is less than -1 , respectively [60]. The active and break days are defined during the peak monsoon months of July and August. It is also to be mentioned that this criterion has to be satisfied at least for three consecutive days. During 1998 to 2004, total active and break days are considered from previous works [59, 60]. To further assess the performance of the model, the intraseasonal variation of the simulated rainfall is studied in terms of active and break periods. The performance of the model is compared with observed rainfall during these periods. During this period, the simulated rainfall and observed rainfall for each active and break events are shown in Figure 10. The cc (0.68) is quite significant. It suggests that the model is able to simulate the active and break periods with reasonably good accuracy.

5. Summary, Discussion, and Conclusion

In the present work, an attempt is made to explore a viable alternate approach to supplement existing rain parameterization schemes in the NWP and rainfall simulation models. The idea of the present work is inspired by the concept of hybrid environmental model, where a deterministic numerical model can be combined with an ANN-based parameterization scheme which takes into account the underlying physical processes [61]. In the present work, daily rainfall is simulated by using four concurrent meteorological parameters, namely, geopotential height, specific humidity,

zonal wind, and meridional wind, during 00, 06, 12, and 18 GMT at different pressure levels, that is, 1000, 925, 850, and 700 mb pressure levels by using the feedforward network. The architecture of the feedforward MLP model consists of 64 nodes, 10 nodes, and 1 node in the input, hidden, and output layers, respectively. It is a one-step simulation approach. In terms of rmse and cc, the model is performing better compared to 3B42, whereas in terms of bias, the performance of 3B42 is better compared to the model. In this regard, it is to be mentioned that 3B42 rainfall is a global rainfall product, whereas the developed model is region specific. The model is able to simulate daily rainfall at neighbourhood locations with reasonably good accuracy. The model simulation improves significantly for areal average over the core monsoon region. The model is able to capture monthly, its interannual and intraseasonal variation with reasonably good accuracy. The applicability of the model outside the core monsoon region is yet to be seen. Nevertheless, the model is not without its limitations. Two major limitations of this MLP model are (i) its inability to capture extreme high rainfall events and (ii) simulation of marginal rainfall, even during the no rain situation. Error statistics of simulated rainfall is significantly affected by rainfall characteristics. When there are more extreme rainfall events, there is an underestimation of rainfall, and during the predominantly low rainfall events, there is overestimation of rainfall. Some error may also arise due to the uncertainties in the input parameters. Underestimation of rainfall in case of more extreme rainfall events is due to the fact that the model

fails to capture the heavy rain events. During the training period, there are very less number of heavy rain situations and hence, it is not able to capture heavy rain. By repetition of heavy rain episodes in the training period, this limitation may be overcome. Overestimation of rainfall during the predominantly low rainfall events is due to the cumulative effect of simulation of marginal rainfall in no rain situation. By putting a threshold, the simulated marginal rainfall in no rain situation may be removed, and it may improve overestimation of rainfall during periods of scanty rainfall.

The present work is different from the majority of the work reported on the forecasting of rainfall at different time scale, where the time series data of either rainfall or meteorological parameters are utilized to forecast the time-lagged rainfall. In the present scenario of rainfall forecasting, researchers have utilized various sets of surface meteorological variables such as minimum and maximum temperature [35, 39, 41, 43–46]. Some other researchers have utilized extended meteorological parameters in addition to temperature variables such as relative humidity and previous day rainfall [62]. In some cases, rainfall forecasting is carried out by taking into account historical time series of rainfall only [48]. Majority of these works have utilized various form of feedback ANN, such as time delay network, recurrent network, and wavelet network to take into account temporal variation of the input variables.

The present effort is a novel approach in the sense that the simulation of daily rainfall is a hybrid approach, which blends the statistical and physical model by taking into account the underlying physical processes. For example, it takes into account the convection process, by virtue of the height profiles of the considered meteorological parameters, unlike the extensively used historical time series rainfall or surface-based meteorological data to simulate rainfall. Our nature of objective falls under the pattern mapping of a static input and output data. For the present objective, the feed-forward networks are reasonably suitable.

Overall, the developed model is performing reasonably well. This approach has a potential to be used as a rain parameterization scheme in the dynamical atmospheric and coupled models to simulate daily rainfall. Nevertheless, the present approach can also be used for multistep prediction of rainfall.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The financial support to Sanjay Sharma from Ministry of Human Resource Development (MHRD), Government of India, under *Rashtriya Uchchatar Shiksha Abhiyan* (National Higher Education Mission) is gratefully acknowledged. The

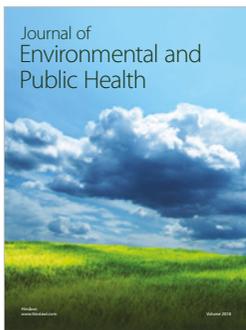
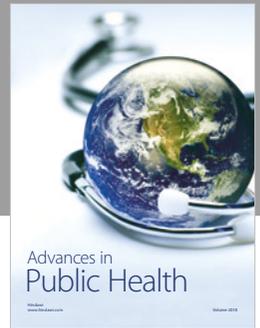
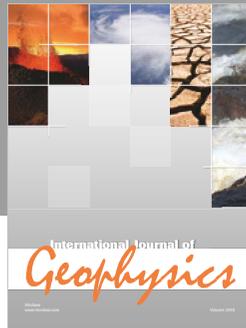
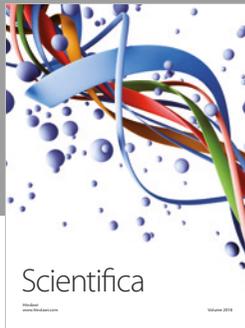
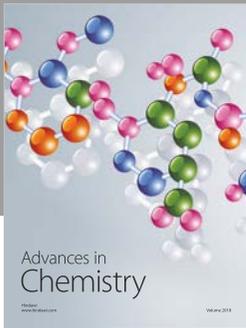
authors are thankful to the principal of Kohima Science College (autonomous) for providing necessary facilities to carry out the present work. The authors acknowledge NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, for providing the meteorological data. The authors are thankful to the TRMM project and the Indian Meteorological Department for providing the gridded daily rainfall products.

References

- [1] D. Dutta, S. Sharma, G. K. Sen et al., "An artificial neural network based approach for estimation of rain intensity from spectral moments of a Doppler Weather Radar," *Advances in Space Research*, vol. 47, no. 11, pp. 1949–1957, 2011.
- [2] C. Kummerow and L. Giglio, "A passive microwave technique for estimating rainfall and vertical structure information from space. Part I: algorithm description," *Journal of Applied Meteorology*, vol. 33, no. 1, pp. 3–18, 1994.
- [3] D. R. Johnson, T. H. Zapotocny, F. R. Reames, B. J. Wolf, and R. B. Pierce, "A comparison of simulated precipitation by hybrid isentropic-sigma and sigma models," *Monthly Weather Review*, vol. 121, no. 7, pp. 2088–2114, 1993.
- [4] H. H. Kwon, U. Lall, and J. Obeysekera, "Simulation of daily rainfall scenarios with interannual and multidecadal climate cycle for South Florida," *Stochastic Environmental Research and Risk Assessment*, vol. 23, no. 7, pp. 879–896, 2009.
- [5] R. I. Maidment, D. I. F. Grimes, R. P. Allan, H. Greatrex, O. Rojas, and O. Leo, "Evaluation of satellite and model reanalysis rainfall estimates for Uganda," *Meteorological Applications*, vol. 20, pp. 308–317, 2013.
- [6] K. A. Emanuel, *Atmospheric Convection*, Oxford University Press, Oxford, UK, 1994.
- [7] P. Berg, J. O. Haerter, P. Thejll, C. Piani, S. Hagemann, and J. H. Christensen, "Seasonal characteristics of the relationship between daily precipitation intensity and surface temperature," *Journal of Geophysical Research*, vol. 114, p. D18102, 2009.
- [8] I. A. Lund, "Correlations between areal precipitations and 850 mill bar geopotential height," *Monthly Weather Review*, vol. 99, no. 9, pp. 691–697, 1971.
- [9] B. R. Beckmann and T. A. Buishand, "Statistical downscaling relationship for precipitation in the Netherlands and North Germany," *International Journal of Climatology*, vol. 22, no. 1, pp. 15–32, 2002.
- [10] S. S. Dugam, "Use of meridional wind index at east coast of India for monitoring drought and flood monsoon conditions over Indian region," *Earth Science India*, vol. 3, pp. 69–72, 2010.
- [11] S. B. Kakade and S. S. Dugam, "Impact of cross-equatorial flow on intra-seasonal variability of Indian summer monsoon rainfall," *Geophysical Research Letters*, vol. 35, no. 12, p. L12805, 2008.
- [12] R. Hashima, C. Roy, S. Motamedi et al., "Selection of meteorological parameters affecting rainfall estimation using neuro-fuzzy computing methodology," *Atmospheric Research*, vol. 171, pp. 21–30, 2016.
- [13] J. Abbot and J. Marohasy, "Input selection and optimization for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks," *Atmospheric Research*, vol. 138, pp. 166–178, 2014.
- [14] S. Chattopadhyay, "Feed forward artificial neural network model to predict the average summer-monsoon rainfall in India," *Acta Geophysica*, vol. 55, no. 3, pp. 369–382, 2007.

- [15] M. P. Darji, V. K. Dabhi, and H. B. Prajapati, "Rainfall forecasting using neural network: a survey," in *Proceedings of 2015 International Conference on Advances in Computer Engineering and Applications*, Zhengzhou, China, July 2015.
- [16] M. N. French, W. F. Krajewski, and R. R. Cuykendall, "Rainfall forecasting in space and time using a neural network," *Journal of Hydrology*, vol. 137, no. 1–4, pp. 1–31, 2002.
- [17] K. K. Htike and O. O. Khalifa, "Rainfall forecasting models using Focused time-delay neural networks," in *Proceedings of International Conference Computer and Communication Engineering (ICCCCE) on IEEE*, Kuala Lumpur, Malaysia, 2010.
- [18] N. Q. Hung, M. S. Babel, S. Weesakul, and N. K. Tripathi, "An artificial neural network model for rainfall forecasting in Bangkok, Thailand," *Hydrology and Earth System Sciences*, vol. 13, no. 8, pp. 1413–1425, 2009.
- [19] A. Kumar, A. Kumar, R. Ranjan, and S. Kumar, "A rainfall prediction model using artificial neural network," in *Proceedings of Control and System Graduate Research Colloq. (ICSGRC)*, pp. 82–87, Shah Alam, Malaysia, July 2012.
- [20] F. Mekanik, M. A. Imteaz, S. Gato-Trinidad, and A. Elmahdi, "Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes," *Journal of Hydrology*, vol. 503, pp. 11–21, 2013.
- [21] R. D. Braddock, M. L. Kremmer, and L. Sanzogni, "Feed forward artificial neural network model for forecasting rainfall-runoff," *Environmetrics*, vol. 9, no. 4, pp. 419–432, 1998.
- [22] P. Hettiarachchi, M. J. Hall, and A. W. Minns, "The extrapolation of artificial neural networks for the modelling of rainfall-runoff relationships," *Journal of Hydroinformatics*, vol. 7, no. 4, pp. 291–296, 2005.
- [23] A. S. Tokar and P. A. Johnson, "Rainfall-runoff modeling using artificial neural networks," *Journal of Hydrologic Engineering*, vol. 4, no. 3, pp. 232–239, 1999.
- [24] Y. M. Chiang, F. J. Chang, B. J. D. Jou, and F. P. Lin, "Dynamic ANN for precipitation estimation and forecasting from radar observation," *Journal of Hydrology*, vol. 334, no. 1–2, pp. 250–261, 2007.
- [25] R. Xiao and V. Chandrasekhar, "Development of a neural network based algorithm for rainfall estimation from radar observations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 1, pp. 160–171, 1997.
- [26] T. Bellerby, M. Todd, D. Kniveton, and C. Kidd, "Rainfall estimation from a combination TRMM precipitation radar and GOES multispectral satellite imagery through an use of Artificial neural network," *Journal of Applied Meteorology*, vol. 39, no. 12, pp. 2115–2128, 2000.
- [27] D. Dutta, S. Sharma, J. Das, and R. M. Gairola, "Development of a regional rain retrieval algorithm for exclusive mesoscale convective systems over peninsular India," *Advances in Space Research*, vol. 49, no. 12, pp. 1750–1762, 2012.
- [28] D. I. F. Grimes, E. Coppola, M. Verdecchia, and G. Visconti, "A neural network approach to real time rainfall estimation for Africa using satellite data," *Journal of Hydrometeorology*, vol. 4, no. 6, pp. 1119–1133, 2003.
- [29] D. K. Sarma, M. Konwar, S. Sharma et al., "An artificial neural network based integrated regional model for rain retrieval over land and ocean," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, pp. 1689–1696, 2008.
- [30] D. K. Sarma, M. Konwar, J. Das, S. Pal, and S. Sharma, "A soft computing approach for rainfall retrieval from TRMM microwave imager," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 12, pp. 2879–2885, 2005.
- [31] F. J. Tapidor, C. Kidd, K. L. Hsu, and F. Marzano, "Neural networks in satellite rainfall estimation," *Meteorological Applications*, vol. 11, no. 1, pp. 83–91, 2004.
- [32] S. J. Burian, S. R. Durrans, S. J. Nix, and R. E. Pitt, "Training artificial neural network to perform rainfall disaggregation," *Journal of Hydrologic Engineering*, vol. 6, no. 1, pp. 43–51, 2001.
- [33] S. J. Burian and S. R. Durrans, "Evaluation of an artificial neural network rainfall disaggregation model," *Water Science and Technology*, vol. 45, no. 2, pp. 99–104, 2002.
- [34] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: the state of the art," *International Journal of Forecasting*, vol. 14, no. 1, pp. 35–62, 1998.
- [35] S. K. Nanda, D. P. Tripathy, S. K. Nayak, and S. Mohapatra, "Prediction of rainfall in India using Artificial Neural Network (ANN) models," *International Journal of Intelligent Systems and Applications*, vol. 5, no. 12, pp. 1–22, 2013.
- [36] G. Shrivastava, S. Karmakar, and M. K. Kowar, "BPN model for long range forecast of monsoon rainfall over a very small geographical region and its verification for 2012," *Geofizika*, vol. 30, no. 2, pp. 155–172, 2013.
- [37] V. K. Somvanshi, O. P. Pandey, P. K. Agrawal, N. V. Kalanker, M. R. Prakash, and R. Chand, "Modeling and prediction of rainfall using Artificial Neural Network and ARIMA techniques," *Journal of Indian Geophysical Union*, vol. 10, no. 2, pp. 141–151, 2006.
- [38] S. Chattopadhyay, "Anticipation of summer monsoon rainfall over India by Artificial Neural Network with Conjugate Gradient Descent Learning, Network with Conjugate Gradient Descent Learning," 2006, <http://arXiv-preprint-nlin/0611010:2-14>.
- [39] S. Chattopadhyay and M. Chattopadhyay, "A soft computing technique in rainfall forecasting," in *Proceedings of International Conference on IT*, pp. 19–21, HIT, Sapporo, Japan, 2007.
- [40] P. Goswami and Srividya, "A novel neural network design for long range prediction of rainfall pattern," *Current Science*, vol. 70, no. 6, pp. 447–457, 1996.
- [41] P. Guhathakurta, "Long lead monsoon rainfall prediction for meteorological subdivisions of India using deterministic artificial neural network model," *Meteorology and Atmospheric Physics*, vol. 101, no. 1–2, pp. 93–108, 2008.
- [42] H. D. Navone and H. A. Ceccatto, "Predicting Indian summer monsoon rainfall: a neural network approach," *Climate Dynamics*, vol. 10, no. 6–7, pp. 302–10, 1994.
- [43] A. K. Sahai, M. K. Soman, and V. Satyan, "All India summer monsoon rainfall prediction using an Artificial Neural Network," *Climate Dynamics*, vol. 16, no. 4, pp. 291–302, 2000.
- [44] C. Venkatesanet, S. D. Raskar, S. S. Tambe, B. D. Kulkarni, and R. N. Keshavamurty, "Prediction of all India summer monsoon rainfall using Error-Back-Propagation Neural Networks," *Meteorology and Atmospheric Physics*, vol. 62, no. 3–4, pp. 225–240, 1997.
- [45] P. Nastos, A. Paliatsos, K. Koukouletsos, I. Larissi, and K. Moustris, "Artificial neural networks modeling for forecasting the maximum daily total precipitation at Athens, Greece," *Atmospheric Research*, vol. 144, pp. 141–150, 2014.
- [46] C. L. Wu, K. W. Chau, and C. Fan, "Prediction of rainfall time series using Modular Artificial Neural Networks coupled with data preprocessing techniques," *Journal of Hydrology*, vol. 389, no. 1, pp. 146–167, 2010.
- [47] C. L. Wu and K. W. Chau, "Prediction of rainfall time series using modular soft computing methods," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 3, pp. 997–1007, 2013.
- [48] R. V. Ramana, B. Krishna, S. Kumar, and N. Pandey, "Monthly rainfall prediction using wavelet neural network

- analysis,” *Water Resources Management*, vol. 27, no. 10, pp. 3697–3711, 2013.
- [49] R. P. Shukla, K. C. Tripathi, A. C. Pandey, and I. M. L. Das, “Prediction of Indian summer monsoon rainfall using Nino indices: a neural network approach,” *Atmospheric Research*, vol. 102, no. 1-2, pp. 99–109, 2011.
- [50] M. Rajeevan, J. Bhate, J. D. Kale, and B. Lal, “High resolution daily gridded rainfall data for the Indian region: Analysis of break and active monsoon spells,” *Current Science*, vol. 91, pp. 296–306, 2006.
- [51] J. B. Wijngaard, K. M. Tank, and G. P. Konnen, “Homogeneity of 20th century European daily temperature and precipitation series,” *International Journal of Climatology*, vol. 23, no. 6, pp. 679–692, 2003.
- [52] A. N. Pettitt, “A non-parametric approach to the change-point detection,” *Applied Statistics*, vol. 28, no. 2, pp. 126–135, 1979.
- [53] H. Alexandersson, “A homogeneity test applied to precipitation data,” *Journal of Climatology*, vol. 6, no. 6, pp. 661–675, 1986.
- [54] T. A. Buishand, “Some methods for testing the homogeneity of rainfall records,” *Journal of Hydrology*, vol. 58, no. 1-2, pp. 11–27, 1982.
- [55] E. M. Kalnay, M. Kanamitsu, R. Kistler et al., “The NCEP/NCAR 40-year reanalysis project,” *Bulletin of the American Meteorological Society*, vol. 77, no. 3, pp. 437–471, 1996.
- [56] D. E. Rumelhart, G. E. Hintom, and R. J. Williams, “Learning representations of back propagation errors,” *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.
- [57] G. P. Compo, J. S. Whitaker, P. D. Sardeshmukh et al., “The twentieth century reanalysis project,” *Quarterly Journal of the Royal Meteorological Society*, vol. 137, no. 654, pp. 1–28, 2011.
- [58] G. J. Huffman, R. F. Adler, D. T. Bolvin et al., “The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales,” *Journal of Hydrometeorology*, vol. 8, no. 1, pp. 38–55, 2007.
- [59] S. B. Saha, S. Sen Roy, S. K. Roy Bhowmik, and P. K. Kundu, “Intra-seasonal variability of cloud amount over the Indian subcontinent during the monsoon season as observed by TRMM precipitation radar,” *Geofizika*, vol. 31, no. 1, pp. 29–53, 2014.
- [60] M. Rajeevan, S. Gadgil, and J. Bhate, “Active and break spells of the Indian summer monsoon,” *Journal of Earth System Science*, vol. 119, no. 3, pp. 229–247, 2010.
- [61] V. M. Krasnopolsky, “Neural network emulations for complex multidimensional geophysical mappings: Application of neural network techniques to atmospheric and oceanic satellite retrievals and numerical modeling,” *Reviews of Geophysics*, vol. 45, no. 3, pp. 1–34, 2007.
- [62] V. K. Dabhi and S. Chaudhary, “Hybrid wavelet-postfix-GP Model for rainfall prediction of Anand Region of India,” *Advances in Artificial Intelligence*, vol. 2014, Article ID 717803, 11 pages, 2014.



Hindawi

Submit your manuscripts at
www.hindawi.com

