Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

RESEARCH ARTICLE

Engineering Reports

Comparison of Geospatial Interpolation Techniques for Assessing Spatio-Temporal Variation of Rainfall

Srinivas Bankuru[1]| Sandeep Samantaray[2]| Abinash Sahoo[2]| Omkesh Tripathy[1]| Dillip K Ghose*[3]|

[1]M.Tech Scholar, Department of Civil engineering, National Institute of Technology Silchar, Assam, India.
[2]Ph. D Scholar, Department of Civil engineering, National Institute of Technology Silchar, Assam, India.
[3]Assistant Professor, Department of Civil Engineering, Department of Civil engineering, National Institute of Technology Silchar, Assam, India.

Correspondence
*Assistant Professor, Department of Civil Engineering, National Institute of Technology Silchar, Assam, India, Email: dillipghose_2002@yahoo.co.in

Funding Information
Project Name: No funding information available

Abstract
Precipitation is an important phenomenon which contributes in the constant supply of water over entire earth. Atmospheric water accounts for less than 0.001% of total water yet it is responsible for the constant supply throughout the globe. It is important to know the distribution of precipitation along with space to know the pattern of precipitation spatially. In order to know this spatial pattern five different geospatial interpolation techniques totaling to 20 different models are applied for 30 years (1988 - 2018) of monthly average precipitation. These models are compared to know which one of these gives the best resemblance of the phenomena. Six performance measures, MAE, MBE, MSE, RMSE, ME and R^2 are used to compare the different models. The model for which error is minimum (close to zero) and efficiency is maximum (close to unity) are preferable. After application of various models, it was found that IDW technique with weight parameter of 3 gives the best result with MBE of -0.1397, MAE of 2.9372, MSE of 13.0708, RMSE of 3.6154, ME of 0.7842 and R^2 of 0.7744. Other models that performed well were Universal kriging and RBF. After evaluating the best model, error in the estimation of data by that model was also carried out to know the locations where error is intense. It is seen that where the precipitation is intense the errors associated increases. Temporal variation of rainfall is equally important to know have a clearer picture about the pattern of precipitation spatially as well as with seasonally. Therefore, after figuring out the best model, temporal variation of precipitation was also determined showing monthly variation of rainfall. So, after plotting spatial and temporal variation of precipitation it becomes easier for us to determine the precipitation at places which are not gauged.

KEYWORDS
Geospatial, Precipitation, Temporal variation.

1 | INTRODUCTION

Geographical studies are necessary for planning and forecasting. Hence, it is necessary that a thorough study of topographical features in the study area should be done. This will help to know the basic parameters to be kept in mind for the purpose of planning and forecasting

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

events such as floods and extreme rainfall events. Precipitation is one of basic and most important parameters if it comes to hydrological modelling. Therefore, precipitation characteristics of an area have to be examined properly for modelling of runoff and other hydrological events. Characteristics of precipitation are known only by measuring it accurately. Traditional way of measuring precipitation involves usage of rain gauges. This method requires placing of rain gauges at different locations and recording the readings manually or automatically every twenty-four hours. This is a bit tedious work and placement of rain gauges is not easy if the topography of the region is uneven. Problems such as maintenance of rain gauge are required from time to time. On the other hand, if precipitation is accompanied by strong winds then the amount of rainfall recorded is not accurate. Another precipitation measuring technique is Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) which gives an alignment based successive plan for joining rainfall gauges from various satellites, and in addition check investigations where possible, at satisfactory scales (0.25° X 0.25°). As per Huffman et al. (2006) initial approval outcomes are as per the following: the TMPA gives sensible execution over month to month scales. The TMPA, has bring down expertise in effectively indicating moderate and light events on brief time interims, just the same as other fine scale observations. Illustrations are given of a surge occasion and diurnal cycle assurance. Joyce et al. (2004) talks about a system using which half-hourly worldwide precipitation estimates achieved from remote sensing satellite are circulated by movement vectors got from geostationary satellite information. The (CMORPH) utilizes movement vectors got from half-hourly interval geostationary satellite IR symbolism to circulate generally higher amount of precipitation measurement got from passive microwave information. Moreover, the shape and intensity of the rainfall highlights are altered amid the interval in which microwave sensor examines by using time weighted direct linear interpolation. This procedure yields both spatial and temporal total microwave-inferred rainfall investigations, autonomous of the infrared temperature field. It was seen that CMORPH made improvement in averaging estimations done by microwave. CMORPH also improved those techniques which uses microwave and infrared data to estimate precipitation usually when passive microwave information is not available. According to Ebert et al. (2007) satellite measurements of precipitation event are most exact amid summer and at lower latitudes, while the NWP models indicate most noteworthy expertise a midwinter and at higher latitudes. As a rule, the more the precipitation regime inclines toward profound convection, the more (less) precise the satellite (model) measurements are. The approval over the Joined States additionally recommends that in general the IR-PMW blended satellite measurements performed and also radar as far as every day precipitation. We accentuate that these outcomes apply to precipitation gauges made (for the most part) over land, at day by day time scales and - 25 km spatial scales. The exactness would surely be diverse for shorter eras and could enhance or break down depending upon the regime. The satellite precipitation evaluations might be more exact over the sea than over land in light of the fact that the PMW calculations can have the advantage of the microwave emission channels. Consequently, the decisions with respect to relative exactness of models versus satellite assessments ought to be rethought for oceanic rainfall, maybe utilizing TRMM precipitation radar data measurements as approval for month to month models and satellite precipitation aggregations.

In hilly regions it is not easy to maintain a proper rain gauging network because of topographical hindrance. Therefore, precise measurement of rainfall and runoff cannot be done in such regions. In such regions precipitation over the whole region can be determined using interpolation. It requires selection of a proper interpolation technique and then utilising that technique to determine the precipitation over the entire region. Thus, if we have a

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

precipitation data measured in a region using rain gauges, then we can utilise that data set for interpolation and also determine the extent of precipitation over that region. The approach utilised for interpolation is geo-statistical technique, this differs from the classical statistics. The difference between classical statistics and geo-statistical is that classical approach assumes that every single data from a group of data is independent and it does not tell about the other data while geo-statistical approach assumes the spatial data which is being utilised has a correlation between them this correlation is a function of distance between the gauging stations. Geo-statistical techniques have become very popular for interpolation of data where there is a scarcity of data. Geographical Information System (ArcGIS 10.2.2) provides its users with a variety of interpolation techniques. Each of these techniques has their own advantages and disadvantages. One cannot decide prior to use any techniques. After a proper examination of the data and the topographical features of a region, any one of the techniques is utilised to perform the interpolation. The superiority of geo-statistical methods over classical one can be inferred from previous study done for optimizing monitoring networks.

2. STUDY AREA

Assam is a state located in India and has very undulated topography. The area of the state is 78,438 km^2 and is divided into 33 numbers of districts. There are 28 number of rain gauges installed and maintained by meteorological department of Assam. These are the only means using which the precipitation is measured in the state. The undulated topography causes a difficulty in measuring the precipitation uniformly over the entire region. Assam receives a large amount of precipitation during monsoon period and it reaches to a maximum of 400mm. Hence, the measurement of precipitation becomes vital for calculation of runoff, and flood predictions.

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

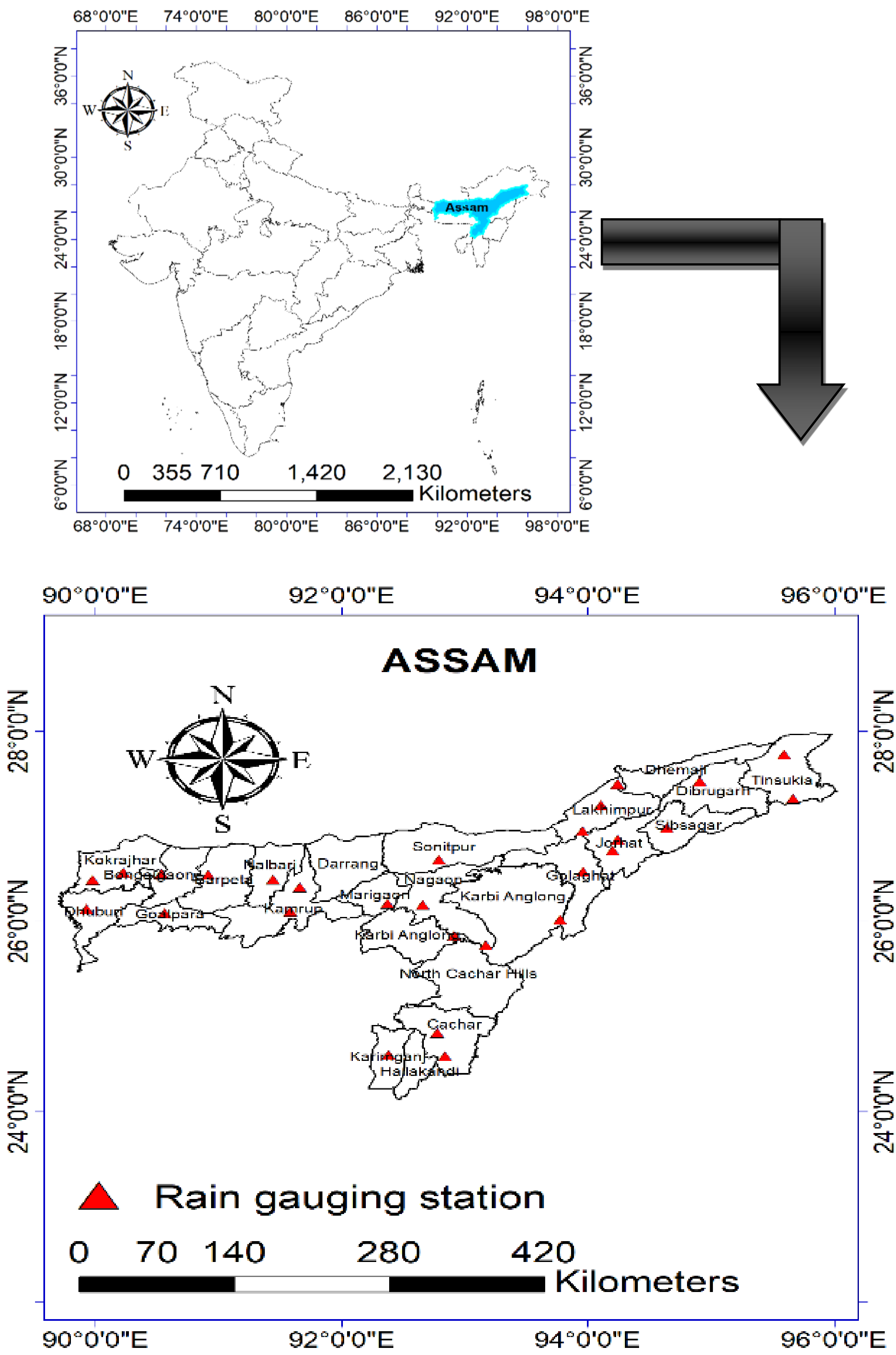


Figure 1 Showing the geographical location of the study area and the rain gauges.

Precipitation data of 28 gauging stations ranging for 30 years (1988-2018) was taken from Regional Meteorological Centre, Guwahati. This data was utilised for making the prediction maps for the region.

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

3. MATHEDOLOGY

3.1 Geospatial Interpolation

The geo-statistical techniques are different from the classical statistical techniques. Geo-statistical methods are more reliable here because these methods assume that there is an auto correlation among the data sets. When precipitation occurs, it is obvious that the area closer to them will also be getting ample amount of rainfall, thus it indicates that there is always some relation for the precipitation data in the region surrounded by an area experiencing precipitation.

ArcGIS uses geo statistical extension to incorporate geospatial interpolation. There is several interpolation techniques that this extension provides all of which are based on the same principle of Tobler law. It states that all things are inter-related while things that are nearer are more similar among themselves as compared to things which are distant from each other. Therefore, geo-statistical interpolation is based on notion of spatial autocorrelation. When the interpolation is done using this technique first it determines the spatial autocorrelation among the spatial data. The correlation determines

- Resemblance of data (objects) in the search neighborhood.
- Degree of spatial correlation of data with itself in space.
- How the variables are dependent and their degree.

All the interpolation methods give different result. No two methods will give same result.

- Deterministic Interpolation: Creates a surface on the basis of observed data sets in the search radius. It uses mathematical equations to determine the smoothness of the observed interpolation surface grid.
- Geo-statistical Interpolation: Creates a surface on the basis of observed data sets in the search radius. It uses mathematical equations to determine the smoothness of the observed interpolation surface grid.

3.2 Interpolation Methods

3.2.1 Global polynomial

In this method a smooth surface is fitted using a polynomial function on the basis of input points. There is a gradual change in the surface. This method can be assumed as fitting a piece of paper between the plotted points; these points are elevated points, elevation being equal to their respective values. It's obvious that the paper will not pass through all the points. There will be a few sets of points which shall not lie on the paper. There are some points which are above and some which are below the paper. However, when we measure the height by which the points are raised above the paper and take the summation of all such points, it is found that this summation is equal to the summation of height of all points falling below the surface of the paper.

The interpolating polynomial is as follows:

$$P(x_k) = \sum_{i=0}^{n-1} a_i x_k^i \quad (1)$$

1

Where $P(x_k)$ is the interpolating polynomial for n distinct points (x_k, f_k)

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

3.2.2 Local polynomial

While global polynomial interpolation fits a polynomial to the whole surface, local polynomi al interpolation fits many polynomials, each within specific neighbourhoods. We can define the search neighbourhood using the size and shape, number of neighbours and configuration of the sector. A single plane is fitted through the points for the first order global polynomial interpolation, for the second order the surface is fitted along with a bend in it. Similarly, third order polynomial has two bends in it. But when there is a surface with sloping terrain or an irregular slope, a single polynomial surface does not fit well. More than one polynomial surface is required to represent the surface more precisely. LPI fits a polynomial of specified order with the help of input data for a defined neighbourhood. The neighbourhoods overlap and the value obtained at the centre of the neighbourhood is the predicted value.

3.2.3 Inverse Distance Weighted

IDW was first proposed by Shephard (1961), he stated this method to be a deterministic. This method assumes that observation points farther away from the estimation points have less impact over the estimated value while the observation points nearer to estimated points have more impact on the estimated value. Hence it can be seen that the estimated value is inversely proportional to the distance of observed value. IDW function uses a parameter defined as weight (W_j), which assigns the impact of an observed station to any other point of measurement. The parameter (W_j) uses a parameter p, which act as a decay function. As the value of p increases the weight of the stations which are farther decreases. For p=0, the predicted value will be the arithmetic mean of the stations in the search neighbourhood because the weight of all stations become equal for p=0. For very high value of p, the values are estimated taking into effect only the immediate surrounding stations. In this paper the values are estimated using the p=2, 3 and 4.

If r (x, y) is any arbitrary point within the region of interpolation then interpolated value is given by:

$$P(r) = \sum_{j=1}^{N} W(r_j)(f(r_j) \tag{2}$$

where P(r) = interpolated value at

$$W(r_j) = \frac{d_j r^{-p}}{\sum_{j=1}^{N} d_j r^{-p}} \tag{3}$$

Where $d_j(r)$ = $\sqrt{(x-x_j)^2+(y-y_j)^2}$ distance between r and r_j; p = decay determining parameter.

A comprehensive discussion has been made by Kravchenko and Bullock (1999) for the choice of parameter p, the effect of weight parameter for IDW interpolation is discussed by Cecilio and Pruski (2003), a similar discussion has been made by Vicento-Serrano (2003) over the importance of weight parameter in IDW function for prediction model.

3.2.4 Radial Basis Function

RBF is a deterministic interpolation technique; the interpolation surface is formed in such a way that it passes through all the observation points. Wong et al (2002) and Fornberg (2006) has defined the function as:

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

$$P(r) = \sum_{j=1}^{N} \lambda_j s\,(||r - r_j||) \tag{4}$$

Where s(.) definite positive RBF; ||.|| = Euclidian norm; λ_j = set of unknown weights

$\lambda_j$is calculated using $P(r_j) = f\,(r_j) \quad j = 1,2\ldots\ldots., N$ (5)

(4) and (5) together form a new system of equation which is the form

$$S\Lambda = \Theta \tag{6}$$

Where ***S*** is a NxN matrix of RBF values it is also known as interpolation matrix and $\Lambda = [\lambda_j]$ whose weights are unknown $[f_j]$ = column matrix of observed values.

The interpolation widely depends upon the basis function chosen. The available choices of basic functions are thin plate spline, multi-log, inverse multi-quadric, natural cubic spline. The basis function further depends upon Euclidean distance between r and r_j, and a smoothing parameter "R". Hardy (1971) has discussed how to evaluate R. Later, Folly (1987) and Franke (1982) also discussed the range of values that can be taken for R. In this paper R is taken as $R^2= d^2/(25N)$. Where d is the diagonal distance of the grid in which the interpolation is taking place.

3.2.5 Kriging

Kriging is an interpolation technique and woks in a similar manner as IDW but the difference is the weight parameter is calculated using statistical relations obtained from semi-variogram model. The semi-variogram uses an empirical equation given by

$$\lambda(\mathrm{H}) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [f(r_j) - f(r_j + h)]^2 \tag{7}$$

Where $N(h)$ is the number of sample pair $(f(r_j)$

$f(r_j+h))$separated by distance h.

This empirical function is fitted into a particular semi-variogram model; there are a number of models. Sironvalle (1980) has discussed about quadratic model. Kitanidis (1997) has discussed about linear model and is most widely used model. Cressie (1991) discussed about an exponential, rational quadratic and wave model. Paninatier (1996) has discussed about Gaussian, power and spherical models. Olea (2000) has discussed about pent spherical and cubic model. The function is fitted for the input points within the search radius and then the value is predicted at a specified point. Unlike IDW, the predicted values can exceed the numerical value of sample points.

3.3 Evaluation of Interpolation Methods

We are estimating rainfall at ungagged locations with different models as discussed above. There are as many as twenty models which are utilised in this work to assess the rainfall at ungagged locations. Therefore, it is necessary to know the ability of these models to predict rainfall accurately. But there is a challenge to find out which one of these models is giving most reliable result. For this we have done cross validation of the estimated value with the observed value. Cross validation is carried out following Vicento-Serrano et al (2003) and Muller et al (2004). Borges et al (2016) and Xie et al (2011) stated that cross validation should be performed by removing one of the stations and then applying one of the interpolation techniques to measure the predicted value at the removed station taking into

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

account the other observing stations in the search radius. This method is repeated for the remaining stations and the predicted value is noted down after measuring the estimated rainfall at all the stations comparisons is carried out to see how efficiently the model is predicting. The comparison was done following the performance measures discussed by Luo et al (2008), Tabio and Salas (1985) and Li and Heap (2011). Cross validation is performed using seven performance measures for all the twenty methods. The seven measures are as follows

Mean bias error

$$\text{MBE} = \frac{1}{N}\sum_{j=1}^{N}[F(r_j) - f(r_j)] \tag{8}$$

Mean absolute error

$$\text{MAE} = \frac{1}{N}\sum_{j=1}^{N}|F(r_j) - f(r_j)| \tag{9}$$

Mean squared error

$$\text{MSE} = \frac{1}{N}\sum_{j=1}^{N}[F(r_j) - f(r_j)]^2 \tag{10}$$

Root mean square error

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}[F(r_j) - f(r_j)]^2} \tag{11}$$

Model efficiency

$$\text{ME} = 1 - \frac{\sum_{j=1}^{N}[F(r_j) - f(r_j)]^2}{\sum_{j=1}^{N}[f(r_j) - \overline{f(r)}]^2} \tag{12}$$

Coefficient of determination

$$R^2 = \left\{\frac{\left[\sum_{j=1}^{N} F(r_j) - \overline{F(r)}\right]\left[\sum_{j=1}^{N} f(r_j) - \overline{f(r)}\right]}{\sqrt{\sum_{j=1}^{N}[F(r_j) - \overline{F(r)}]^2}\sqrt{\sum_{j=1}^{N}[f(r_j) - \overline{f(r)}]^2}}\right\}^2 \tag{13}$$

Where N is the number of observed data points

$f(r_j)$ is the observed value at station

$F(r_j)$ is the value estimated at station r_j

$\overline{f(r)}$ is mean of all observed values.

Different interpretations are drawn from the above formulated performance measures. Coefficient of determination is adopted very commonly as a performance measuring criterion. Li and Heap (2008) has discussed that R^2 solely cannot represent the performance of any model; hence other measures are also used to compare the performances of models used in this study. As per Hallak and Pereira (2011), MBE does not provide a clear picture about the individual errors; this is because the positive and negative error gets cancelled out. To overcome this MAE is calculated, Fox (1981) introduced MSE is another performance measure which is used to check the accuracy of models to predict, but according to Ponce-Hernandez (2006) MAE shows the error on large scale because the errors are squared and

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

errors get amplified. RMSE is another parameter which is commonly used for measuring the performance of models. Willmot (1982) found that RMSE is one of the best tools to measure the error of a model as it gives us a summary about the average difference of values between observed ones and predicted ones. Model efficiency (ME) is also taken into consideration for evaluation of the models. According to Nash and Sutcliffe (1970) ME can take any value in the range of (-∞, 1] and best model is the one which is closer to unity. For the model having ME as zero, it can be inferred that the predicted values are just the mean of the observed values in search radius. For ME less than zero it should be inferred that mean of observed values are better estimates than the predicted one (Krause et al., 2005).

4. RESULTS AND DISCUSSION

The ranking of the models was done as per the values obtained by the above-mentioned performance measures (Table 1). For the best model the errors (MBE, MAE, MARE, and RMSE) must approach a value closer zero while ME and R^2 should approach to a value closer to unity.

Table 1 Measures used to evaluate performances of interpolation methods of 30 years of monthly precipitation data in Assam, India

Interpolation method	MBE	MAE	MSE	RMSE	ME	R^2	RANK
Inverse Distance Weighted							
P = 2	-0.2298	3.0101	15.1298	3.8897	0.7410	0.7521	7
P = 3	-0.1397	2.9372	13.0708[a,b]	3.6154[a, b]	0.7842[a,b]	0.7844	1
P = 4	-0.0644[a]	2.9148[a]	24.4566	4.9453	0.7832	0.7833[a,b]	2
Radial Basis Function							
Thin Plate Spline	0.1328	4.1512	30.2661	5.5015	0.4819	0.6600	18
Completely regularised spline	-0.0743	2.8734	13.7976[a]	3.7145[a]	0.7638[a]	0.7600	5
Inverse multiquadric	-0.1073	2.9028	14.1010	3.7563	0.7584	0.7622	6
Multi quadratic	0.0394[a,b]	3.0719	16.0417	4.0052	0.7253	0.7315	8
Spline with tension	0.0742	2.8710[a,b]	13.8226	3.7179	0.7634	0.7655[a]	4
GPI							
1st order	-0.1264	3.8162	20.1777[a]	4.1920[a]	0.6546	0.6673[a]	12
2nd order	-0.0537[a]	3.4556[a]	23.2702	4.8239	0.6016	0.6244	17
3rd order	0.0584	5.5779	72.3190	8.5041	0.2380	0.2739	20
4th order	0.2093	4.9294	44.6667	6.6833	0.2553[a]	0.4630	19
Constant	-0.0554[a]	3.7822	19.9012	4.4611	0.6593	0.6623	15
LPI							
Exponential	0.2138	3.4242[a]	17.8100[a]	4.2202[a]	0.6957[a]	0.6969[a]	10
Gaussian	0.1314	3.6406	19.0747	4.3674	0.6734	0.6759	11
Polynomial	0.1151	3.7187	19.8030	4.4501	0.6600	0.6602	16
Quartic	0.0878	3.7142	19.5123	4.4173	0.6660	0.6670	14
Epanechnikov	0.1273	3.7913	19.6326	4.4309	0.6639	0.6668	13
Kriging							
Simple	-0.1594[a]	2.9950	16.9510	4.1172	0.7098	0.7178	9
Universal	-0.2620	2.9743[a]	13.6904[a]	3.700[a]	0.7653[a]	0.7696[a]	3

To find out the variation of precipitation in the region (Assam) monthly average rainfall data of 28 stations were taken for a period of 30 years ranging from January 1988 to December 2018. Spatial variation of rainfall of each month is evaluated using Geospatial analyst tool of

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

ArcGIS (version 10.2.2). Five interpolation techniques IDW, GPI, RBF, LPI, Kriging were used to find out the spatial variation of rainfall over the entire region of Assam. The results obtained by all the techniques were cross validated by finding out the errors and efficiency of all the models using the performance measures discussed above. Table 1 shows the performance of all the models. The rank column in the table tells the overall performance of all the interpolation techniques on the basis of performance measures. From the table we can see IDW at p=3 is the most efficient model while 3rd order GPI is least efficient out of the 20 models utilised. Second best model is again from IDW for weight parameter p = 4. The other models which give reliable results are Simple Kriging and RBF with basis function as spline with tension. However, in some cases user is constrained to use a particular technique only, for such cases also table 1 should be used as a guide to choose the best parameter of a particular model. From table we find out GPI gives its best result when the polynomial order is taken as one. Performance of LPI is good when the kernel function is chosen as Gaussian; similarly for Kriging the results are more favourable for simple as compared to that of universal. Fig 2- Fig 6 shows how the rainfall varies spatially in Assam during month of January using 5 models with the best parameters. Obviously as discussed above IDW will represent the phenomenon more accurately compared to remaining four models.

Fig 7 – Fig 9 shows how the errors are distributed spatially while predicting the rainfall. It is observed that the regions where the precipitation is more intense shows relatively higher error compared to the region where precipitation is less intense. It is clearly visible that the districts like Dhemaji, Tinsukia, Lakhimpur and Dibrugarh are receiving higher amount of rainfall and the errors associated is also higher in these regions.

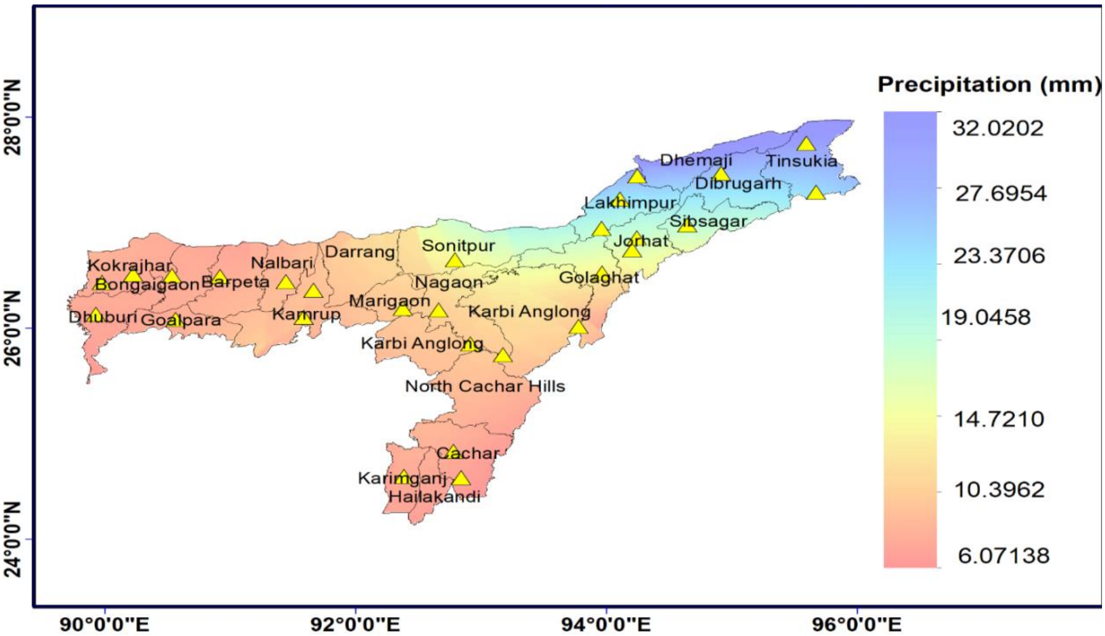


Figure 2 showing spatial variation of rainfall during month of January by LPI

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

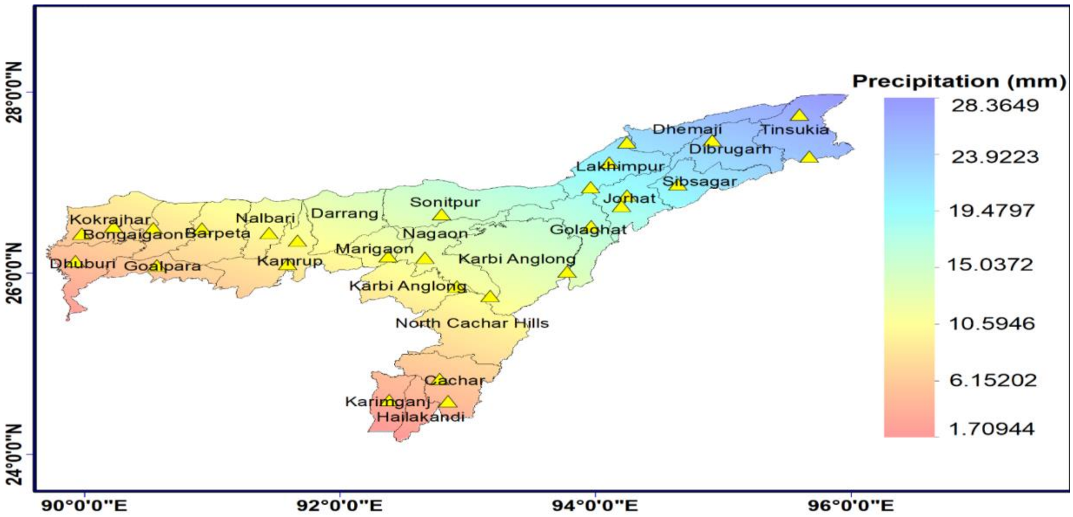


Figure 3 showing the spatial variation of rainfall during month of January by GPI

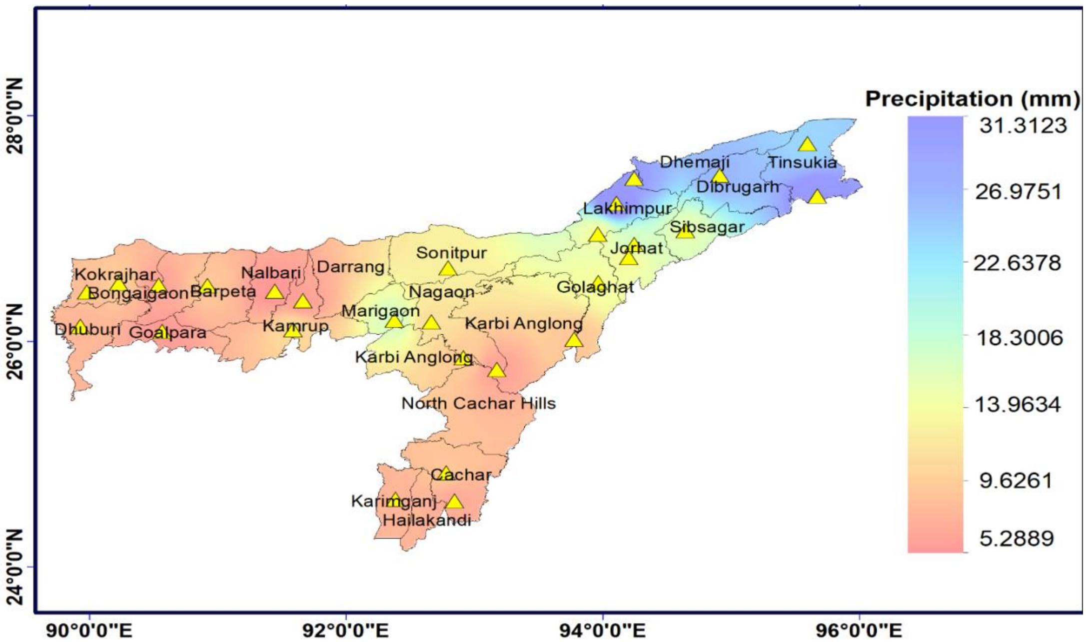


Figure 4 showing spatial variation of rainfall during month of January by IDW

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

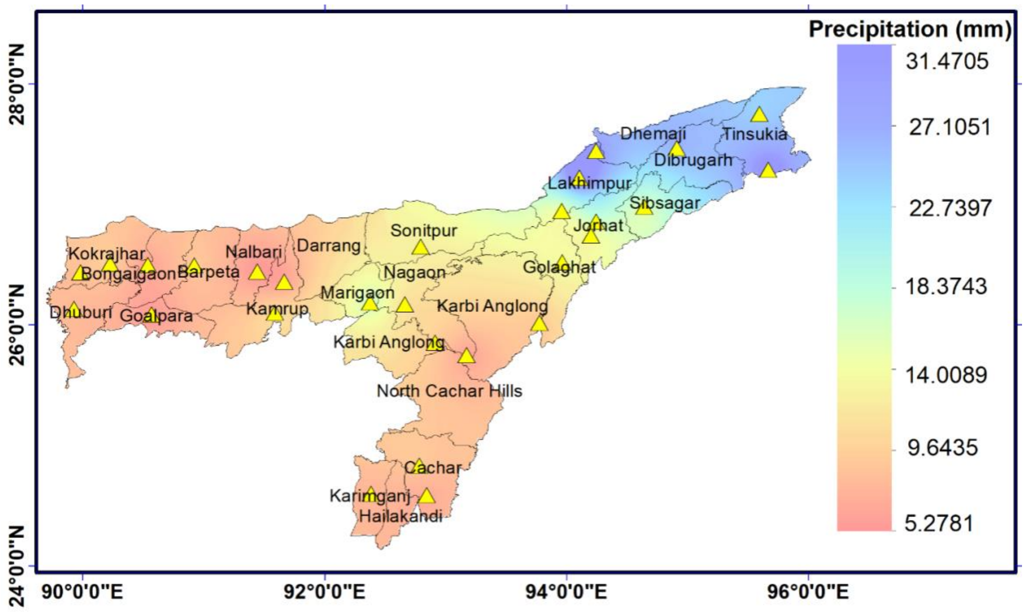


Figure 5 showing spatial variation of rainfall during month of January by RBF

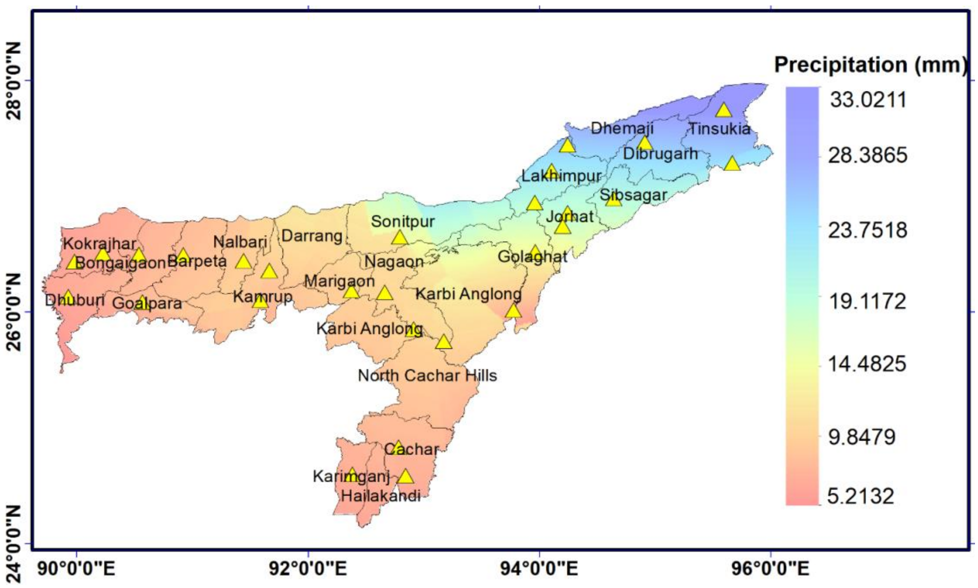


Figure 6 showing spatial variation of rainfall by Kriging

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

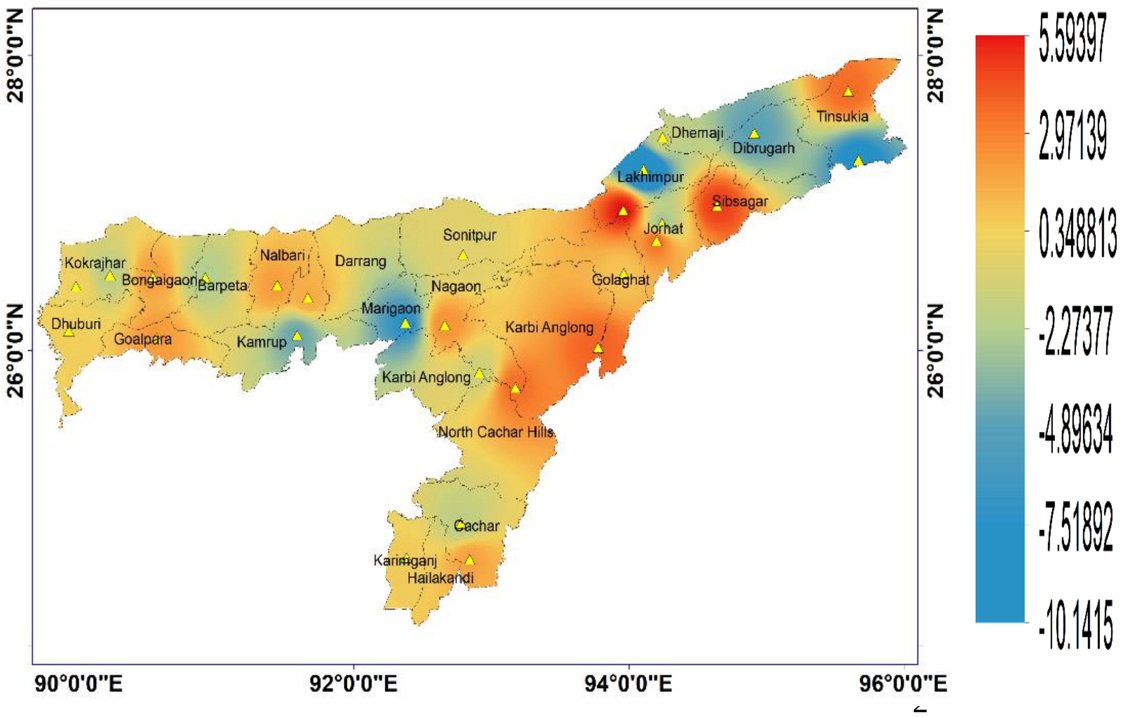


Figure 7 Demonstrating about the MBE in prediction of rainfall at various locations in Assam

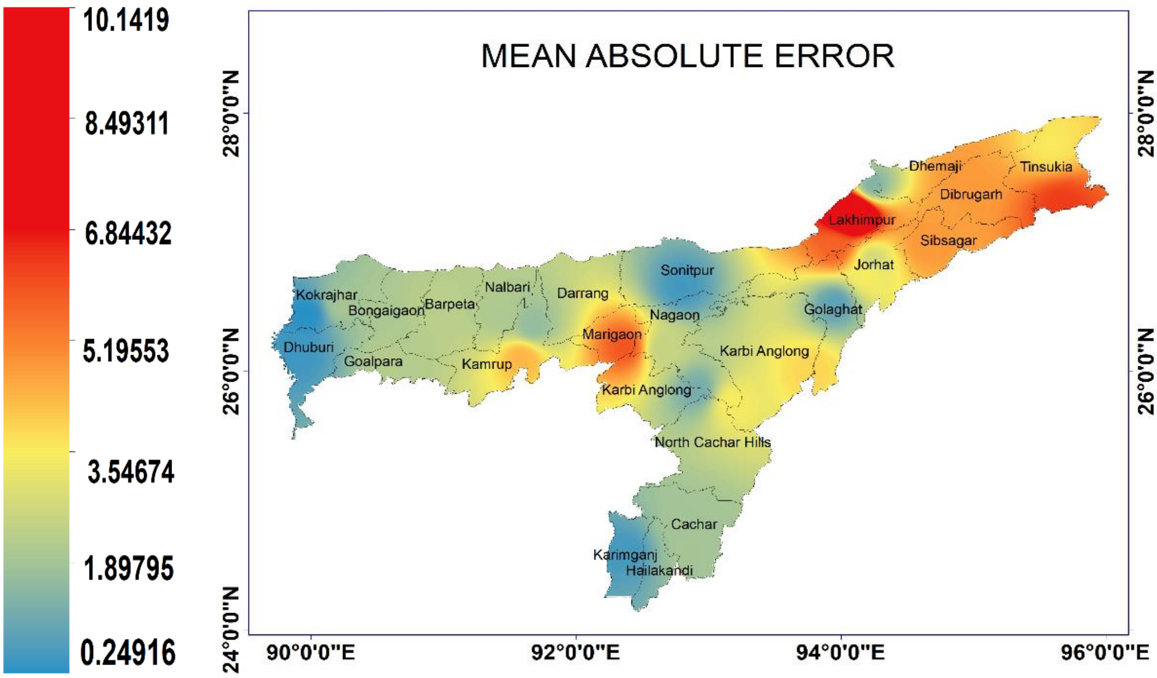


Figure 8 Demonstrating about the MAE in prediction of rainfall at various locations in Assam

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

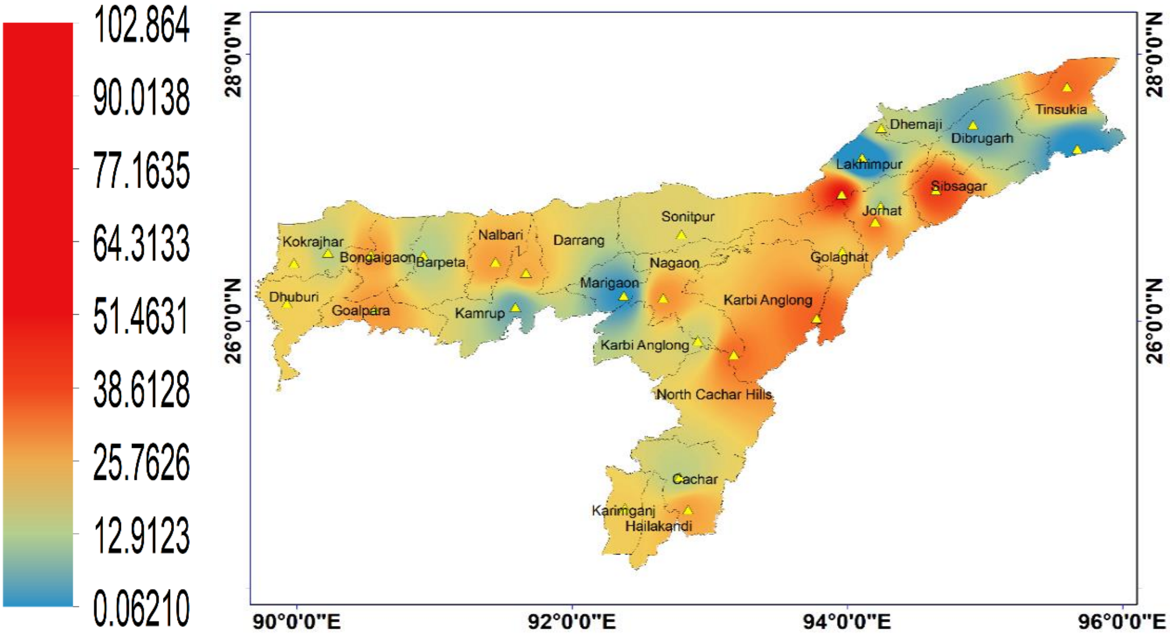
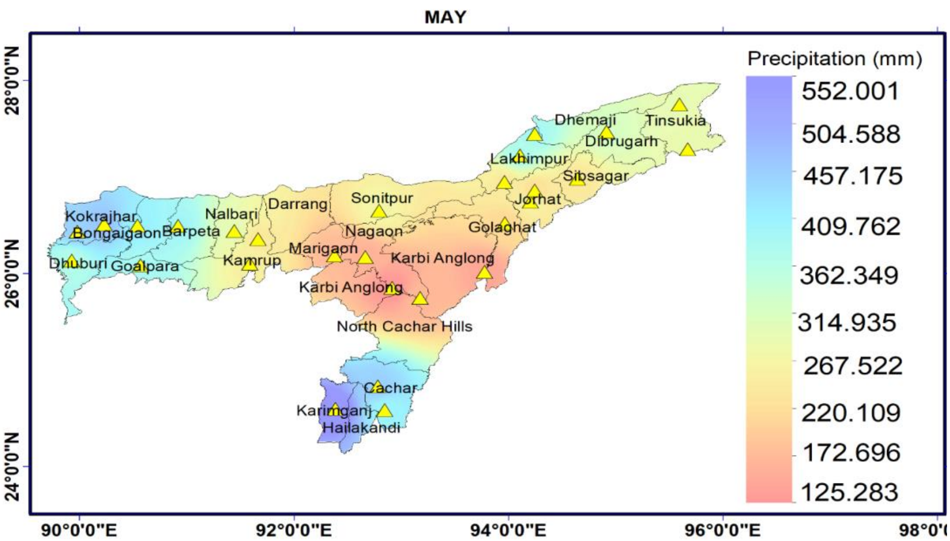


Figure 9 Demonstrating the MSE in prediction of rainfall at various locations in Assam

4.1 Temporal Variation of Rainfall during Monsoon and Non-Monsoon Season

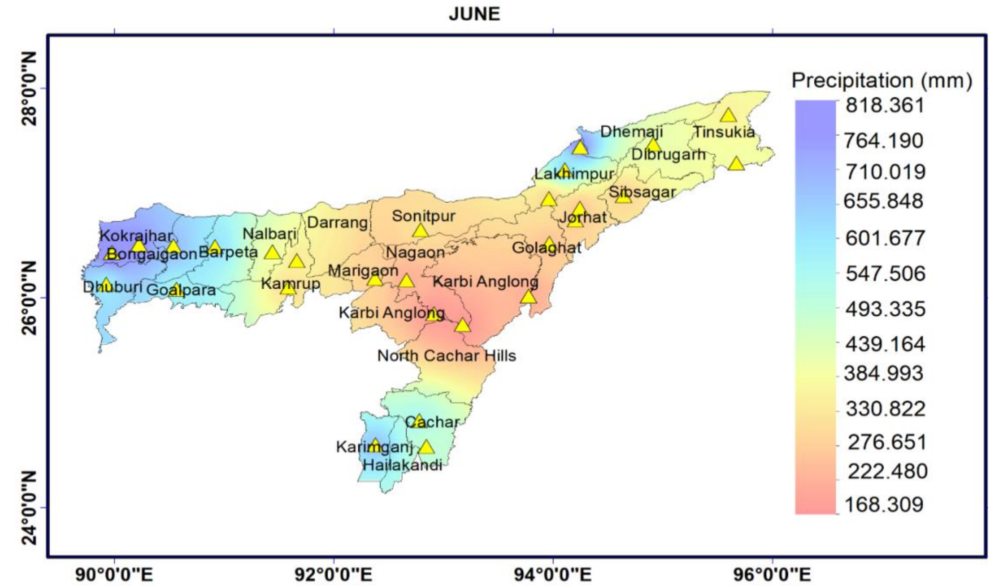
It is necessary to assess the seasonal variation of rainfall to know the months in which maximum and minimum precipitation is received. Spatial and temporal variation together will give us a clearer picture about the precipitation trend for the study area. It is obvious that all the districts will not receive the same maximum precipitation at the same time everywhere. To analyse it, it is necessary that we should plot the temporal variation of the rainfall in the study area. The figures below show how the monthly average rainfall varies in Assam from January till December. As we have discussed that IDW gives a more reliable result in comparison to other methods, is used to plot the average monthly spatial variation of precipitation in Assam.

The monsoon season in Assam starts with May and lasts till October. Therefore, knowledge of spatial variation and amount of rainfall during these seasons is more crucial. The figures below show the spatio-temporal variation of rainfall during monsoon season.

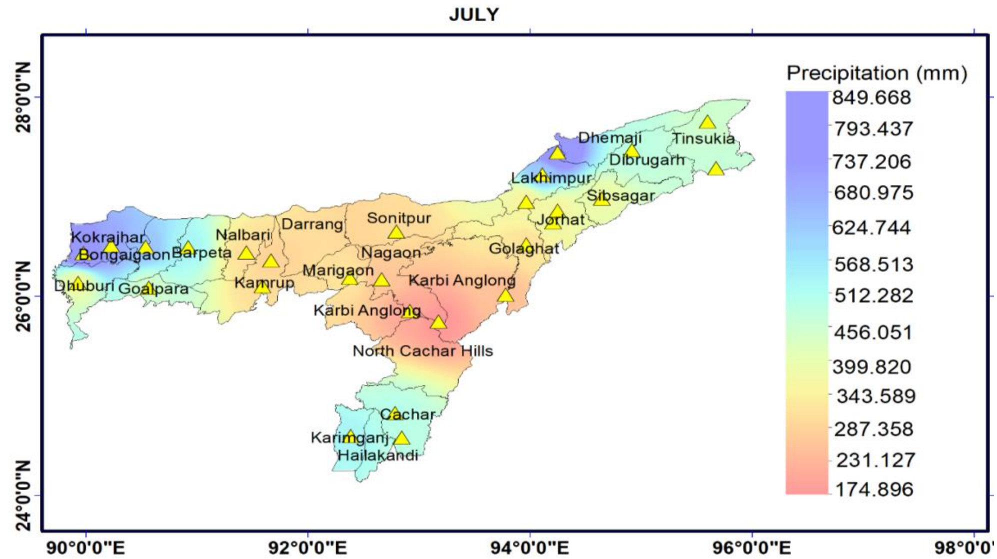


(a)

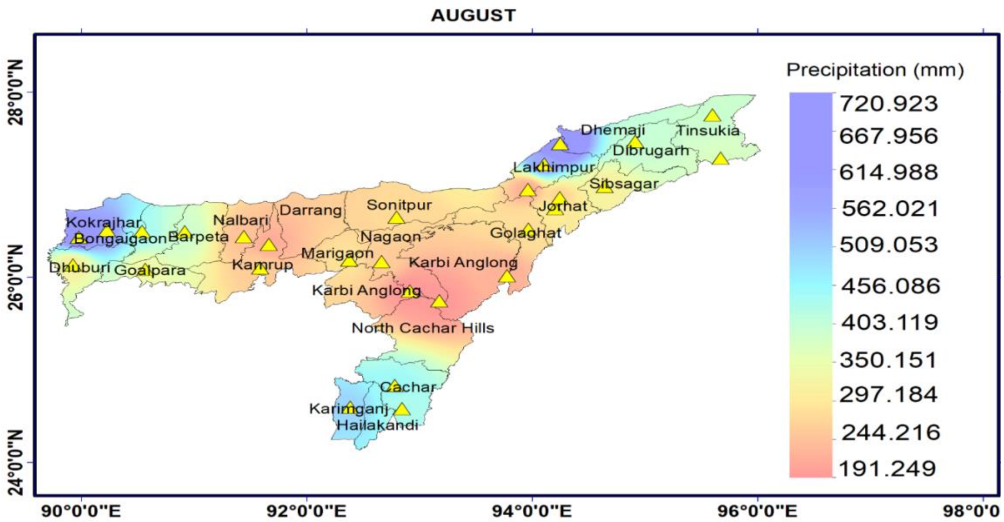
Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx



(b)

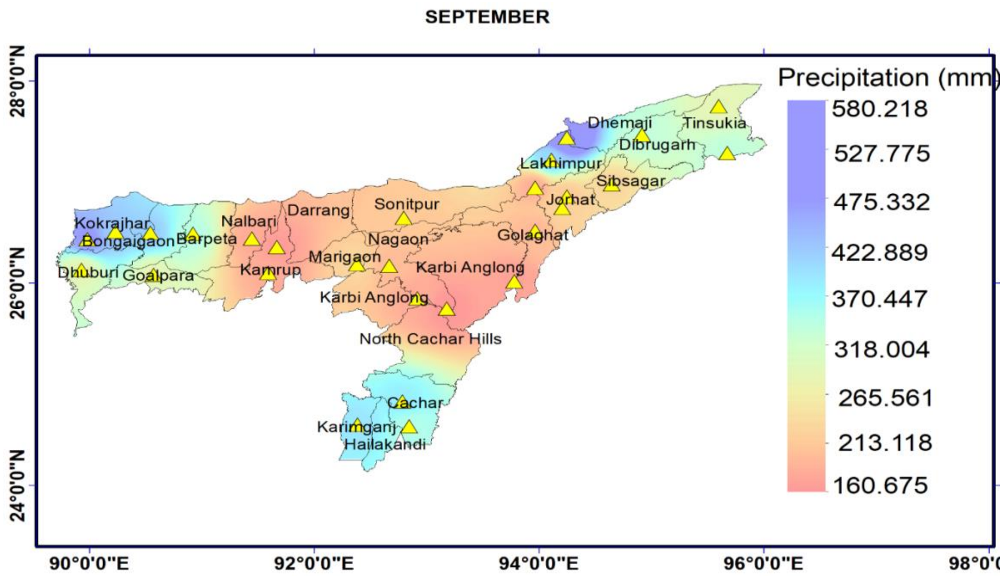


(c)

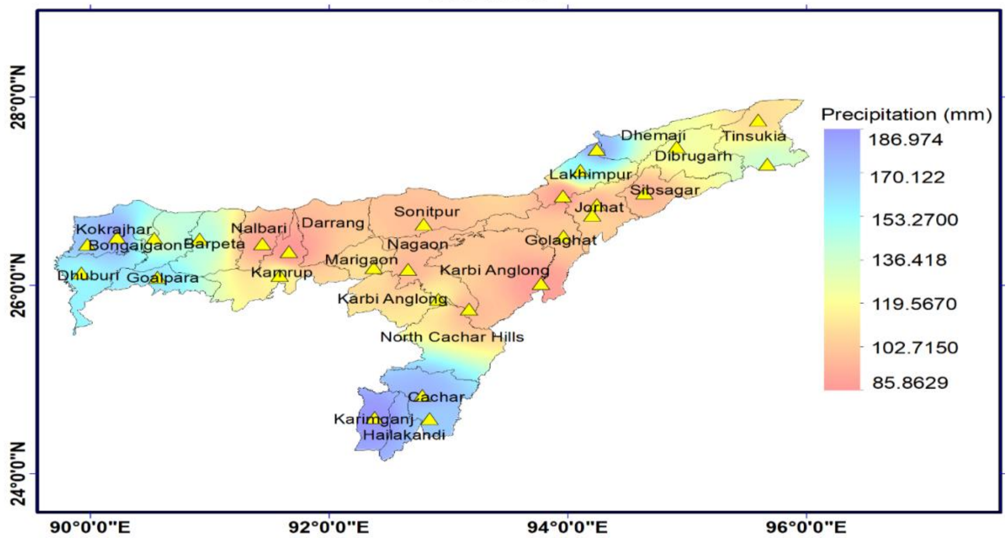


(d)

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx



(e)

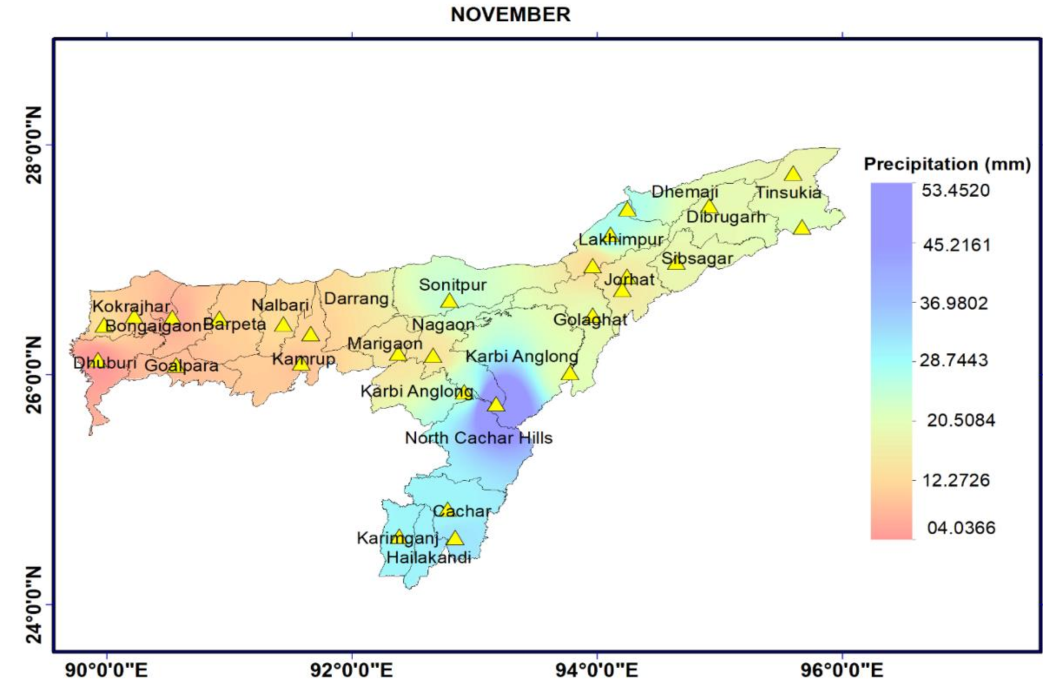


(f)

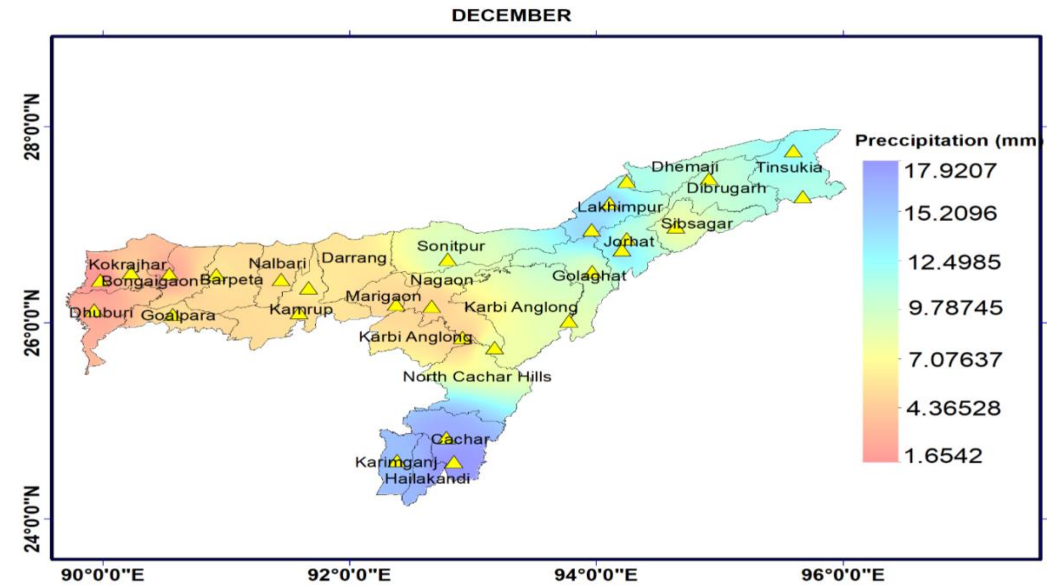
Figure 10 Spatial variation of rainfall during month of (a) May, (b) June, (c) July, (d) August, (e) September, (f) October

It can be depicted from the figure that in May, (Figure 10.a) there is a lot of variation of rainfall with space in the whole region; the amount of rainfall is as high as 552mm in Karimganj to relatively low precipitation amounting to only 125 mm in regions of Karbi Anglong. Middle Assam receives 125 mm-172 mm of rainfall while upper Assam receives 267 mm to 315 mm of rainfall. The region receives the maximum rainfall during the months of June-August and shows a large variation in amount of precipitation occurring in the region along with the space (Figure 10. b, c, d). The monthly average rainfall received during June and July is almost same and varies from 170 mm to 850 mm. regions close to Lakhimpur and Kokrajhar receives the maximum rainfall which amounts to 800 – 850 mm in each of the months of June and July, while in August the same district receives the rainfall of about 700mm. Districts of middle Assam receives relatively a lesser amount of rainfall (170 mm – 190 mm) when compared to districts like Lakhimpur and Kokrajhar. There is a significant decrease in the amount of rainfall received in October, as this month is considered as offset for monsoon in Assam. The Figures below show the spatio-temporal variation in the region during non-monsoon season i.e. November-April.

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

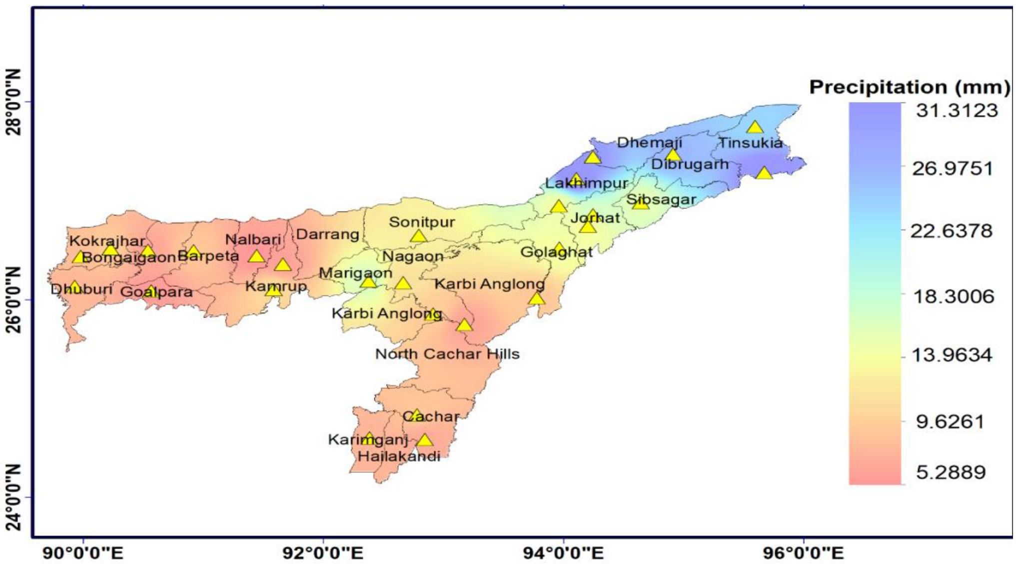


(a)

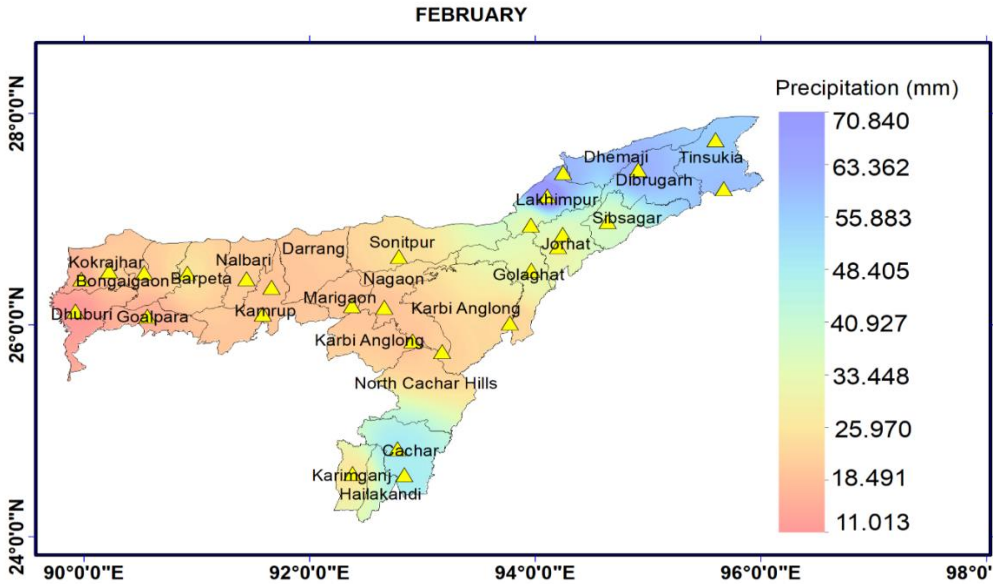


(b)

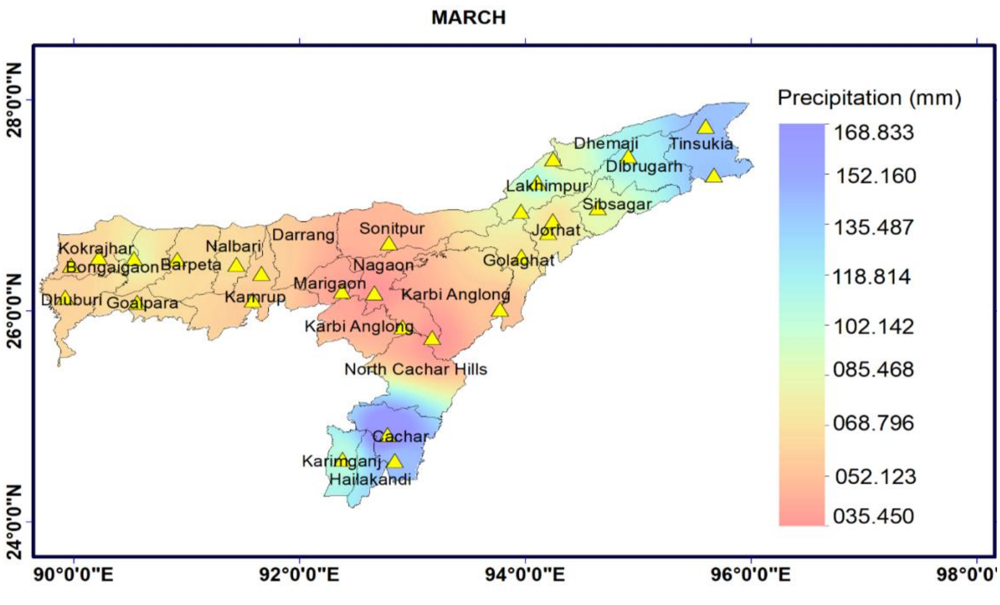
Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx



(c)

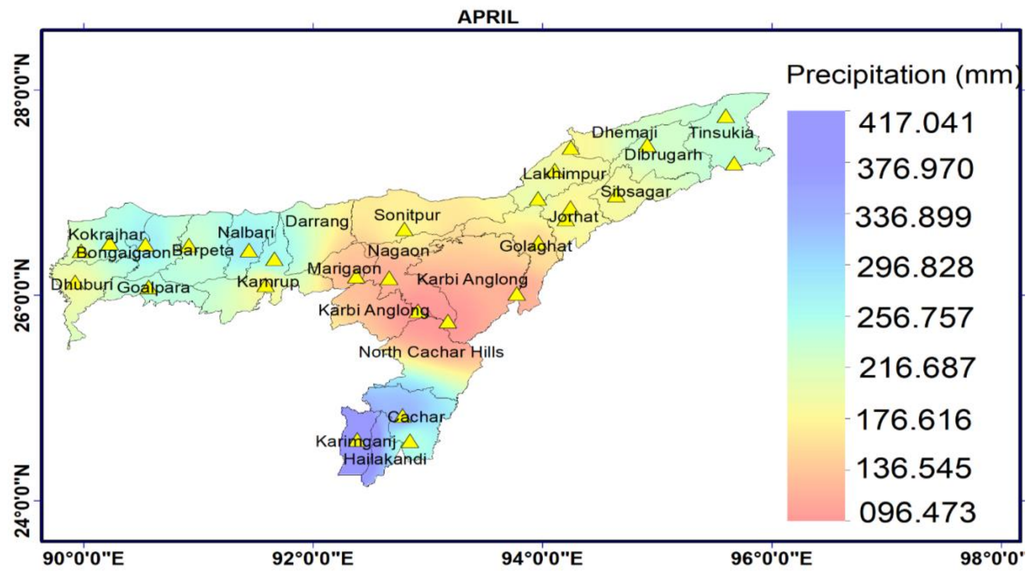


(d)



Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

(e)



(f)

Figure 11 Spatial variation of rainfall during month of (a) November, (b) December, (c) January, (d) February, (e) March, (f) April

There is a decrease in the amount of rainfall received during November- December. November receives a maximum rainfall of 53 mm in region of North Cachar hills, while the other districts receive precipitations as low as 4mm (Dhubri region). Assam receives the minimum amount of rainfall in December (Figure 11. b) with highest (17 mm) of rainfall in Cachar region to minimum (1-2 mm) in Kokrajhar region. In January, (Figure 11. c) we can see that the precipitation varies from 5 mm to 32 mm, with districts of upper Assam (Dhemaji, Tinsukia and Dibrugarh) receiving the maximum rainfall while Cachar, Nalbari and Kokrajhar receives minimum. While, in month of February (Figure 11. d) there is a subtle increase in the amount of precipitation received for upper Assam and the value reaches to 60 mm to 70 mm and minimum reaches to 10 mm to 15 mm in Kokrajhar and Dhubri districts of lower Assam. In March, (Figure 11. e) the region starts to receive more precipitation and the maximum rainfall received varies between 150 mm to 170 mm, in districts of upper Assam. In April (Figure 11. f) whole of the region receives rainfall of substantial amount with regions surrounded by state boundary receiving maximum rainfall amounting to 375 mm to 415 mm in Cachar region, while the middle Assam is still receiving relatively low precipitation which is 100 mm – 110 mm.

CONCLUSION

The following conclusions are come from the above research;

- Geospatial interpolation is used in order to represent the spatial variation of rainfall in the Assam region.
- Comparison of five geospatial interpolation techniques (amounting to a total of 20) applied for monthly average precipitation data for state of Assam; India for a period of 30 years ranging from 1988 to 2018 proved that choice of interpolation technique highly affects the results.
- From the comparison it was found that IDW for weight parameter p = 3 gave the best performance as compared to all other methods. There are few other methods which performed well they were universal Kriging and RBF.

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

- The cross validation of the results shows that the predicted values are close to the observed values and these results can be efficiently used for research works that requires precipitation data for any place which is un-gauged.
- This work will also act as guidance to opt for the best technique among various geospatial interpolations available for interpolation.
- Using the best interpolation method obtained (IDW at p=3) temporal variation of the precipitation is determined.
- Figure 10-11 shows spatial and temporal variation of rainfall. It is found that precipitation is intense in Lakhimpur district of upper Assam and Kokrajhar district of lower Assam which reaches to 820mm - 850mm during monsoon season.
- Figures 10 illustrate that during Monsoon, middle Assam (Nagaon, Karbi Anglong, Sonitpur) experienced less precipitation as compared to other regions. Middle Assam which is 250mm – 350mm
- Temporal variation of rainfall shows that minimum rainfall is observed during December (Figure 11. b) which amounts to a maximum of 15mm – 17mm in Cachar region and 1mm - 4mm in lower Assam region.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] Borges, P. A., Franke, J., Anunciação, Y. M. T., Weiss Bernhofer, H. "Comparison of spatial interpolation methods for the estimation of precipitation distribution in Distrito Federal, Brazil," Theor. Appl. Climatol. 123 (1–2), pp. 335–348, 2016.

[2] Cecílio, R. A., Pruski, F. F. "Interpolação dos parâmetros da equaçãoo de chuvasintensas com uso do inverso de potências da distância," Rev. Bras. EngenhariaAgrícola e Ambiental 7 (3), pp. 501–504, 2003.

[3] Cressie, N. Statistics for spatial data. New York: Wiley, 1991.

[4] Ebert, E.E., Janowiak, J.E., Kidd, C. "Comparison of near-real-time precipitation estimates from satellite observations and numerical models," American meteorological society, pp. 47-64, 2007.

[5] Foley, T. A. "Interpolation and approximation of 3-D and 4-D scattered data," *Comput. Math. Appl.* 13 (8), pp. 711–740, 1987.

[6] Fornberg, B., Larsson, E., Wright, G. "A new class of oscillatory radial basis functions," Comput. Math. Appl. 51 (8), pp. 1209–1222, 2006.

[7] Fox, D. G. "Judging air quality model performance," *Bull. Am. Meteorol. Soc.* 62 (5), pp. 599–609, 1981.

[8] Franke, R. "Scattered data interpolation: Tests of some methods,"*Math. Comput.* 38 (157), pp. 181–200, 1982.

[9] Hallak, R., Pereira Filho, A. J. "Metodologia para análise de desempenho de simulações de sistemasconvectivosnaregiãometropolitana de São Paulo com o modelo ARPS: sensibilidade a variações com osesquemas de adveção e assimilação de dados." Rev. Bras. Meteorologia, 26 (4), 591–608, 2011.

[10] Hardy, R. L. "Multiquadric equations of topography and other irregu-lar surfaces," J. Geophys. Res. 76 (8), pp. 1905–1915, 1971.

[11] Hernandez-Stefanoni, J. L., Ponce-Hernandez, R. "Mapping the spatial variability of plant diversity in a tropical forest: Comparison of spatial interpolation methods," Environ. Monit. Assess, 117 (1–3), pp. 307–334, 2006.

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

[12] Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F., Wolff, D.B. "The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales," Journal of hydrometeorology, 8, pp. 38- 55, 2007.
[13] Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P. "CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution," Journal of hydrometeorology, 5, pp. 487-503, 2004.
[14] Kitanidis, P. K. "*Introduction to geostatistics: Applications in hydro- geology,"* London: Cambridge University Press, 1997.
[15] Krause, P., Boyle, D., Bäse, F. "Comparison of different effi-ciency criteria for hydrological model assessment," Adv. Geosci. 5, pp. 85–97, 2005.
[16] Kravchenko, A., Bullock, D. G. "A comparative study of interpolation methods for mapping soil properties," *Agron. J.* 91 (3), pp. 393– 400, 1999.
[17] Li, J., Heap, A. D. "A review of spatial inter-polation methods for environmental scientists," Symonston, Australia: Geoscience Australia, 137, 2008.
[18] Li, J., Heap, A. D. "A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and im-pact factors," Ecol. Inf. 6 (3), pp. 228–241, 2011.
[19] Luo, W., Taylor, M., Parker, S. "A comparison of spatial inter-polation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales," Int. J. Climatol. 28 (7), pp. 947–959, 2008.
[20] Mueller, T. G., Pusuluri, N. B., Mathias, K. K., Cornelius, P. L., Barnhisel, R. I., Shearer, S. A. "Map quality for ordinary kriging and inverse distance weighted interpolation," Soil Sci. Soc. Am. J. 68 (6), pp. 2042–2047, 2004.
[21] Nash, J., Sutcliffe, J. V. "River flow forecasting through conceptual models. Part i—A discussion of principles," J. Hydrol. 10 (3), pp. 282–290, 1970.
[22] Olea, R. A. "Geostatistics for engineers and earth scientists," Tech-nometrics, 42 (4), pp. 444–445, 2000.
[23] Paninatier, Y. "Variowin: Software for spatial data analysis in 2-D," New York, Springer, 1996
[24] Shepard, D. "A two-dimensional interpolation function for regularly spaced data." In Proc., 1968 23rd ACM National Conf., 517-524. New York: ACM, 1968.
[25] Sironvalle, M. A. "The random coin method: Solution of the problem of the simulation of a random function in the plane," Math. Geol. 12 (1), pp. 25–32, 1980.
[26] Tabios, G. Q., Salas, J. D. "A comparative analysis of techniques for spatial interpolation of precipitation," J. Am. Water Resour. Assoc. 21 (3), pp. 365–380, 1985.
[27] Vicente-Serrano, S. M., Saz-Sanchez, S., Cuadrat, J. M. "Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): Application to annual precipitation and temperature," Clim. Res. 24 (2), pp. 161–180, 2003.
[28] Wang, S., Huang, G. H., Lin, Q. G., Li, Z., Zhang, H., Fan. Y. R. "Comparison of interpolation methods for estimating spatial distribu-tion of precipitation in Ontario, Canada," Int. J. Climatol. 34 (14), pp. 3745–3751, 2014.
[29] Willmott, C. J. "Some comments on the evaluation of model perfor-mance," Bull. Am. Meteorol. Soc. 63 (11), pp. 1309–1313, 1982.
[30] Wong, S., Hon, Y., Golberg. M. "Compactly supported radial basis functions for shallow water equations," *Appl. Math. Comput.* 127 (1), pp. 79–101, 2002.