Spatial Distribution of Rainfall in Indian Himalayas *–* A Case Study of Uttarakhand Region

Ashoke Basistha . D. S. Arya . N. K. Goel

Received: 7 September 2006 / Accepted: 12 November 2007 / Published online: 18 December 2007 \circledcirc Springer Science + Business Media B.V. 2007

Abstract Continuous rainfall data in grid format are required to run models for hydrological and agricultural research as well as water resources planning and management. The present work attempts to prepare a normal annual rainfall map in Himalayan region of India lying in Uttarakhand state at 1 km spatial resolution which currently is not available. In the region, India Meteorological Department maintains observatories/raingauge stations and data from 44 stations were used in this study. A comparative analysis of interpolation techniques like Inverse Distance Weighted, Polynomial, Splines, Ordinary Kriging and Universal Kriging shows that Universal Kriging with hole-effect model and natural logarithmic transformation with constant trend having Root Mean Square Error (RMSE) of 328.7 is the best choice. This is followed by Ordinary Kriging (RMSE 329.1), Splines (RMSE 392.4), Inverse Distance Weighted (RMSE 409.8) and Polynomial Interpolation (RMSE 418.5). Cross validation of the results shows the largest over prediction at Tehri rainfall station (62.5%) and largest under prediction at Nainital station (−36.5%). Physiographic zone wise, the least errors occur in the plains and the largest in the Great Himalayas. The spatial average rainfalls are 1,472 mm for Terai/Bhabar, 1,782 mm for the Shivalik ranges, 1,591 mm for the Lesser Himalayas and 1,635 mm for the Great Himalayan region. The mean areal rainfall in the region is 1,608 mm.

Keywords Interpolation . Kriging . Rainfall distribution . Himalaya . Uttarakhand

1 Introduction

Quantitative estimation of the spatial distribution of rainfall is required for various purposes like water resource management, hydrologic modelling, flood forecasting, climate change

A. Basistha : D. S. Arya (***) : N. K. Goel

Department of Hydrology, Indian Institute of Technology Roorkee, Roorkee, 247667 Uttarakhand, India e-mail: dsarya@rediffmail.com

A. Basistha e-mail: basistha_ashoke@yahoo.com

N. K. Goel e-mail: goelnfhy@iitr.ernet.in studies, water balance computations, soil moisture modelling for crop production, irrigation scheduling etc. Among all the hydrometeorological parameters, rainfall is the most difficult to predict due to its inherent variability in time and space (Guenni and Hutchinson [1998](#page-20-0)), especially for a complex mountainous terrain.

Rainfall studies in the Himalayan region have been limited due to the lack of information on rainfall at high altitudes (Singh et al. [1995](#page-21-0)). A review of literature reveals that studies related to spatial distribution were carried out in other parts of the Himalayan range (Higuchi et al. [1982](#page-20-0); Singh et al. [1995;](#page-21-0) Singh and Kumar [1997;](#page-21-0) Arora et al. [2006](#page-19-0)); however, the spatial variability in this region is still unexplored.

In the Himalayan ranges, rainfall decreases from south to north as monsoon moves across the mountain barrier. Also, there is a reduction from east to west along the path of travel of south west monsoon, as the number of days between its onset and withdrawal reduces (Kansakar et al. [2004\)](#page-20-0). According to Dhar and Bhattacharya ([1976\)](#page-20-0), the zone of maximum rainfall lies between 2,000 and 2,400 m above mean sea level (amsl); Barry ([1981\)](#page-19-0) describes it below 2,000 m while Upadhyay and Bahadur ([1982](#page-21-0)) indicates zone of maximum rainfall between 1,500 and 2,500 m amsl. However, researchers have a general consensus regarding high rainfall belt in the Shivaliks – foothills of the Himalayas.

2 Study Area and the Data Used

The Himalayan range confined in Uttarakhand state was chosen to prepare continuous normal rainfall data. It comprises of a geographical area of 53,484 km², lying in North India between coordinates 28°42′N to 31°28′N and 77°35′E to 81°05′E. The state shares its boundaries with the other Indian states Himachal Pradesh in northwest and Uttar Pradesh in South; and with Nepal and China on the southeast and northeast respectively. The state is divided in 13 districts grouped into two administrative divisions – Garhwal and Kumaon – as shown in Fig. [1.](#page-2-0)

In the study area, altitude varies from 175 to 7,409 m above mean sea level (GTOPO30 Digital Elevation Data). It is divided into four physiographic zones namely *Terai* (finer alluvium deposits) and *Bhabar* (a belt of pebbles) region (average height 175 m –600 m), Shivaliks (average height 600 m-1200 m), Lesser Himalayas (average height 1,200-3,000 m) and the Great Himalayas (snow-clad mountains having an average height of 3,000–7,000 m above Mean Sea Level) (Joshi [2004](#page-20-0)). The annual rainfall varies between 948 and 2,986 mm. Most of the rainfall in the region (60–85% of annual total) is received during *monsoon season* (June to September) (Climate of Uttar Pradesh [1989](#page-19-0)). The temporal distribution of rainfall over the year is shown in Fig. [2.](#page-2-0)

Normal annual rainfall data of 44 raingauge stations were used in this study (Climate of Uttar Pradesh [1989](#page-19-0)) based on the records during period 1901–1950. The raingauge stations were gradually reduced in the region and currently, India Meteorological Department is maintaining eight stations in the region. Locations of raingauge stations are shown in Fig. [1](#page-2-0).

3 Methodology

The commonly used interpolation methods from point data can be divided into three major groups: graphical, topographical and numerical (Daly et al. [1994\)](#page-20-0). Graphical methods include isohyet mapping and Thiessen polygon. Topographical methods involve the cor-

Fig. 1 Study area showing location of raingauge stations

Fig. 2 Temporal distribution of normal monthly rainfall in Uttarakhand

relation of point rainfall data with an array of topographic and synoptic parameters such as slope, exposure, elevation, location of barriers and wind speed and direction. The numerical methods are Inverse Distance Weighted (IDW), Kriging, Polynomial, Radial Basis Functions or Splines etc.

These methods have been used widely. Dirks et al. ([1998\)](#page-20-0), Nguyen et al. [\(1998](#page-20-0)), and Tomczak ([1998\)](#page-21-0) used simpler method like IDW to interpolate rainfall. The popularity of Spline technique is because it does not require the subjective approach and trials needed to fit a semivariogram model as in case of kriging (Hutchinson and Gessler [1994](#page-20-0)). None of the available methods (e.g. Thiessen Polygon, Inverse Distance Weighted, Polynomial, Splines, Ordinary Kriging and Universal Kriging) is able to fully account for climatic and spatial properties of rainfall (Creutin and Obled [1982\)](#page-19-0). However, kriging has been preferred by researchers all over the world particularly when applied on data obtained from low density and irregularly spaced network (Delfiner and Delhomme [1973;](#page-20-0) Tabios and Salas [1985](#page-21-0); Lebel et al. [1987;](#page-20-0) Saveliev et al. [1998;](#page-21-0) Goovaerts [1999](#page-20-0), [2000](#page-20-0); Campling et al. [2001;](#page-19-0) Lloyd [2005;](#page-20-0) Cheng et al. [2007](#page-19-0)).

Inclusion of elevation information in the form of a covariate (i.e., co-kriging) seems interesting to improve estimation efficiency in such a complex terrain. However, Thomas and Herzfeld [\(2004](#page-21-0)) describes that climatologically relevant processes, such as orographic lifting of air masses, are influenced more by morphological aspects and relative elevation differences of the local topography rather than by absolute altitude alone. As the relationship between rainfall and altitude varies over the entire region, co-kriging may not lead to improvement of results (Philips et al. [1992](#page-20-0)). Also, studies in Kosi basin in the Himalayas indicate the absence of any linear relationship between elevation, mean monsoon and annual rainfall (Dhar and Bhattacharya [1976](#page-20-0)). Goovaerts ([2000\)](#page-20-0) and Lloyd ([2005\)](#page-20-0) explains that incorporation of elevation data for rainfall estimation is beneficial when the value of correlation coefficient between rainfall and elevation is larger than 0.75. With a correlation coefficient of 0.16, co-kriging was not considered appropriate for this study.

Rather than depending solely on geostatistics for interpolation in such regions (Hevesi et al. [1992;](#page-20-0) Philips et al. [1992;](#page-20-0) Martinez-Cob [1995](#page-20-0), [1996;](#page-20-0) Papamichail and Metaxa [1996](#page-20-0); Tang et al. [1998](#page-21-0); Prudhomme and Reed [1999;](#page-21-0) Goovaerts [2000](#page-20-0); Naoum and Tsanis [2003](#page-20-0); Skirvin et al. [2003;](#page-21-0) Subyani [2004;](#page-21-0) Diodato [2005;](#page-20-0) Lloyd [2005\)](#page-20-0), researchers have employed regression models with multiple topographic parameters (Daly et al. [1994](#page-20-0), [2002](#page-20-0); Prudhomme and Reed [1998;](#page-21-0) Brunsdon et al. [2001;](#page-19-0) Johansson and Chen [2003;](#page-20-0) Vicente-Serrano et al. [2003](#page-21-0); Naoum and Tsanis [2004\)](#page-20-0) and even Principal Components of Digital Elevation Model (DEM) (Gyalistras [2003](#page-20-0); Thomas and Herzfeld [2004\)](#page-21-0).

These methodologies appear to be promising; however, the scope of present work is kept limited because the goal of this study was to produce a reasonably good continuous data set to bridge the knowledge gap. Thus, the objective of the study was to prepare a normal annual rainfall data set of the Himalayan region in the Uttarakhand state on 1 km grid resolution.

In order to find the most suitable method of interpolation for the region, various commonly used techniques were explored. These are described below and their relative merits discussed in Table [1.](#page-5-0)

3.1 Thiessen Polygon Method

Thiessen polygon method is a popular method to compute the mean areal rainfall. Though not suitable for hilly terrain, it has been employed for the sake of comparison. This method is primarily based on proximal mapping i.e, nearest distance neighbour (Tabios and Salas [1985\)](#page-21-0). The estimate of the process h_0 at any point of interest is equal to the observed value of the nearest sampling point in the area. Let

$$
d_{0j} = \sqrt{(x_0 - x_j)^2 + (y_0 - y_j)^2}, \quad j = 1, ..., n
$$
 (1)

where, $d_{0i} = \min (d_{01}, \ldots, d_{0n})$ That is, the subscript i is determined by searching for the minimum point station distance, so that $w_i=0$ for $j\neq i$ and $w_i=1$ for $j=i$.

3.2 Inverse Distance Weighted Method

The general equation for the inverse distance weighted method is

$$
z_0 = \frac{\sum_{i=1}^{s} z_i \frac{1}{d_i^k}}{\sum_{i=1}^{s} \frac{1}{d_i^k}}
$$
 (2)

where, z_0 is the estimated value at point 0, z_i is the z value at control point i, d_i is the distance between control point i and point 0, s is the number of control points used in estimations and k is the specified power (Chang [2002\)](#page-19-0). A low power (≤ 2) results in a greater contribution towards a grid point value of rainfall from distant gauges, indicating a low spatial variability and vice versa (Dirks et al. [1998](#page-20-0)).

3.3 Polynomial Method

In its simplest form this analysis uses the least square criterion to fit polynomials of successively higher order to spatially distributed data

(a) Linear Surface:

$$
X_n = b_0 + b_1 U + b_2 V \tag{3}
$$

(b) Quadratic Surface:

$$
X_n = c_0 + c_1 U + c_2 V + c_3 U^2 + c_4 UV + c_5 V^2 \tag{4}
$$

(c) Cubic Surface:

$$
X_n = d_0 + d_1 U + d_2 V + d_3 U^2 + d_4 U V + d_5 V^2 + d_6 U^3 + d_7 U^2 V + d_8 U V^2 + d_9 V^3
$$
\n⁽⁵⁾

where, X_n is the value of the spatially distributed variable and U, V are the geographic coordinates. The degree of fit given by the polynomial, and hence its effectiveness in describing the areal variability of the data, is given by the percentage reduction in sums of squares it achieves (Unwin [1969](#page-21-0)). A global Interpolation method uses every control point available to derive an equation. On the contrary, a Local Interpolation method uses a sample of control points within localized windows and thus fits local trends in estimating an unknown value (Chang [2002](#page-19-0)).

Kriging

External information can be combined to get the most External information can be combined to get the most out of expensive data. out of expensive data. Geostatistical methods are generally superior when there Geostatistical methods are generally superior when there part of the signal (Burrough and McDonnel 1998). part of the signal (Burrough and McDonnel [1998](#page-19-0)). unlike splines, such methods do not treat noise as is sufficient data to estimate a variogram because, is sufficient data to estimate a variogram because, unlike splines, such methods do not treat noise as

produced by Ordinary Kriging (Isaaks and Srivastava 1989) produced by Ordinary Kriging (Isaaks and Srivastava [1989](#page-20-0)). model, rather than Euclidian distance, it offers tremendous model, rather than Euclidian distance, it offers tremendous Kriging accounts both for the clustering of nearby samples continuity can be described and adequately captured in a possibilities for customizing the estimation method to the possibilities for customizing the estimation method to the Kriging accounts both for the clustering of nearby samples continuity can be described and adequately captured in a By considering statistical distance, through the variogram ariogram model, it is hard to improve on the estimates By considering statistical distance, through the variogram variogram model, it is hard to improve on the estimates and for their distance to the point being estimated. particular problem at hand. If the pattern of spatial and for their distance to the point being estimated. particular problem at hand. If the pattern of spatial

Another major problem with Thin Plate Splines is the Another major problem with Thin Plate Splines is the steep gradients in poor data areas (Chang 2002). steep gradients in poor data areas (Chang [2002\)](#page-19-0).

The quality of estimates produced by Ordinary Kriging depends like other interpolation algorithms, kriging tends to smooth out overestimation of small values and underestimation of large ones Like other interpolation algorithms, kriging tends to smooth out overestimation of small values and underestimation of large ones The quality of estimates produced by Ordinary Kriging depends local details of the spatial variability of the attribute, leading to local details of the spatial variability of the attribute, leading to spatial continuity. Ordinary Kriging with a poor model may spatial continuity. Ordinary Kriging with a poor model may on the time taken to choose an appropriate model of the on the time taken to choose an appropriate model of the (Goovaerts 1997). (Goovaerts [1997](#page-20-0)).

interpolation results is compared through cross validation, the yield the best predictions at unsampled locations as sample data, particularly when they are scarce and preferentially located, may particularly when they are scarce and preferentially located, may interpolation results is compared through cross validation, the yield the best predictions at unsampled locations as sample data, model that produces the best cross-validated results may not model that produces the best cross-validated results may not not be representative of the study area (Goovaerts 1997). not be representative of the study area (Goovaerts 1997). Though, normally the impact of different models on Though, normally the impact of different models on (Isaaks and Srivastava 1989). (Isaaks and Srivastava [1989\)](#page-20-0).

produce worse estimates than the other simpler methods

produce worse estimates than the other simpler methods

3.4 Splines Method

A spline is approximately a piecewise cubic polynomial that is continuous and has continuous first and second derivatives (Burrough and McDonnel [1998](#page-19-0)). The geostatistical analysis tool in ArcGIS software uses a set of n basis functions, one for each data location (Johnston et al. [2001](#page-20-0)). The predictor is a linear combination of the basis functions,

$$
\widehat{Z}(s_0) = \sum_{i=1}^{n} \omega_i \phi(|s_i - s_0|) + \omega_{n+1}
$$
\n(6)

where, $\phi(r)$ is a radial basis function, $r = |s_i - s_0|$ is Euclidean distance between the prediction location s_0 and each data location s_i , and $\{\omega_i : i = 1, 2, \ldots, n + 1\}$ are weights to be estimated.

The radial basis functions available with the software are (a) Completely regularised spline function, (b) Spline with tension function, (c) Multiquadric function, (d) Inverse Multiquadric function and (e) Thin-plate spline function. The optimal smoothing parameter can be found by minimising the root mean square prediction error using cross validation.

3.5 Kriging

The technique of kriging assumes that the spatial variation of an attribute is neither totally random nor deterministic (Chang [2002\)](#page-19-0). The theory of Regionalised Variables presumes that the spatial variation of any variable can be expressed as the sum of three major components:

- (a) A structural component, having a constant mean or trend.
- (b) A random, but spatially correlated component, known as the variation of the regionalized variable, and
- (c) A spatially uncorrelated random noise or error.

If x is a position in 1, 2 or 3 dimensions, the value of random variable Z at x is given by:

$$
Z(x) = m(x) + \epsilon'(x) + \epsilon''
$$
 (7)

where, $m(x)$ is a deterministic function describing the structural component of Z at $x, \infty(x)$ is the term denoting the stochastic locally varying but spatially dependent residuals from $m(x)$ and ϵ ["] is a residual, spatially independent Gaussian noise term having zero mean and variance γ^2 (Burrough and McDonnel [1998](#page-19-0)).

3.5.1 Ordinary Kriging

Ordinary Kriging (OK) is often associated with the acronym BLUE for "Best Linear Unbiased Estimator" (Isaaks and Srivastava [1989](#page-20-0)). The assumption of stationarity of difference and variance of differences, define the requirements for the intrinsic hypothesis of regionalized variable theory (Burrough and McDonnel [1998](#page-19-0)).

Some of the models available in Arc GIS Geostatistical Analyst, which have been found useful in spatial rainfall prediction are (Johnston et al. [2001](#page-20-0)) (a) Nugget effect, (b) Spherical, (c) Exponential, (d) Gaussian, (e) Hole effect, (f) K-Bessel and (g) J-Bessel.

Ordinary Kriging uses the fitted semivariogram directly in spatial interpolation. The general equation for estimating the z value at a point is

$$
z_0 = \sum_{i=1}^s z_x w_x \tag{8}
$$

where, z_0 is the estimated value, z_x are values at known points, w_x are weights associated with each known point and s is the number of known points used in estimation. Kriging also produces a variance measure for each estimated point to indicate the reliability of the estimation.

3.5.2 Universal Kriging

Universal Kriging (UK) assumes that spatial variation in z values has a drift or a structural component in addition to the spatial correlation between known points. Typically, Universal Kriging incorporates a trend surface equation in the kriging process. It can either be a first order polynomial or it can be a quadratic surface defined by a second order polynomial.

Fig. 3 Normal annual rainfall by Thiessen polygon method

4 Analysis of Rainfall Data

4.1 Criteria

Researchers have used different measures to choose amongst the interpolation schemes.

(a) Root Mean Square Error (RMSE = $\sqrt{\frac{n}{n}}\sum_{i=1}^{n} \mathcal{E}_i^{\tau}$) relates better to of the extremes (Gyalistras [2003](#page-21-0); Vicente-Serrano et al. 2003). $\frac{1}{n}\sum_{i=1}^{n} \varepsilon_i^2$ ¹) relates better to the estimation efficiency
prrano et al. 2003).

$$
\varepsilon_i = P_p - P_o \tag{9}
$$

 P_p and P_o representing the predicted and observed values respectively.

- (b) Mean Absolute Error (MAE = $\frac{1}{n} \sum_{i=1}^{n} |\varepsilon_i|$) is an indicator of overall performance of the internal tor (Daly 2006) $\frac{i=1}{1}$ interpolator (Daly [2006](#page-20-0))
- (c) The Standardised Mean Square Error (SMSE) is a measure of the consistency of the standard deviation σ_i and the corresponding error ε_i which should be in the range of $1 \pm 2\sqrt{2/n}$ (Delhomme [1978](#page-20-0)). If σ_i truly reflects the estimation error (from kriging) then it should be expected (Chua and Bras 1982) that then it should be expected (Chua and Bras [1982\)](#page-19-0) that

$$
\frac{1}{n}\sum_{i=1}^{n}\frac{\varepsilon_i^2}{\sigma_i^2} = 1\tag{10}
$$

(d) Percentage Error (PE = $\frac{P_p - P_o}{P_o} \times 100\%$) (Wei et al. [2005\)](#page-21-0) provides a better relative measure of the differences particularly when the spatial variation in rainfall is large (Price et al. [2000](#page-20-0)).

Fig. 4 Variation of RMSE with power and number of stations

This study considers the criteria mentioned above for selection of the best suitable method. Daly ([2006\)](#page-20-0) described that the physical explanation of the selected surface must be considered on top of all the above criteria. Hence, methodologies and parameters that generate surfaces lacking physical explanation were not considered for generation of normal rainfall data.

4.2 Thiessen Polygon

The weighted average normal annual rainfall in the region is 1,605 mm using Thiessen Polygon (or nearest neighbour) method, whereas the arithmetic average is 1,653 mm. The generated surface is shown in Fig. [3](#page-8-0). It has been prepared by distance allocation to the gauge station points in spatial analyst tool.

4.3 Inverse Distance Weighted Interpolation

The IDW method requires power and number of raingauge stations for interpolation. Figure [4](#page-9-0) shows the variation of RMSE with power varied from 1.0 to 6.0 in steps of 0.2 and maximum number of stations varied from 1 to 44 in steps of 1. It depicts a typical behaviour of errors

Fig. 5 Normal annual rainfall map using IDW interpolation method

dropping down suddenly when two points were included with power 1.0 (RMSE 395.2). However, the surface produced with two points in the neighbourhood shows serrated appearance which could not be explained physically. Another steep drop in RMSE is observed when power is increased from one to two. RMSE stabilised thereafter for higher powers. The error is also insensitive to the number of stations considered for interpolation in such cases.

Attempts were also made to include a minimum number of raingauge stations outside the search radius in case the stations were not available within the search radius. The minimum RMSE was 395.2 with power 1.0 and just 2 points in the neighbourhood. The surface with these parameters lacked the physical acceptability; hence was not considered for further analysis. The next best (RMSE=409.8, MAE=311.9) estimate was obtained by considering minimum 22 stations for interpolation at each point and is shown in Fig. [5](#page-10-0). The optimal power for this case is 1.8665. The computed weighted average areal rainfall in the region using IDW is 1,642 mm.

4.4 Polynomial Method

Interpolation of rainfall data with global and local polynomial methods was attempted. In case of global polynomial, the first order global polynomial trend surface was the best with

Fig. 6 Normal annual rainfall map using local polynomial interpolation method

RMSE value 515 having a bias of −0.3665. With higher orders of global polynomial, error kept on increasing and finally the surfaces became unexplainable. This first order surface also depicts a banded rainfall pattern parallel to the physiographic zones in the region.

The best local polynomial interpolation (linear) produced RMSE of 418.5, mean error of 12.3, MAE of 310.3 with forceful inclusion of 43 points in the neighbourhood. With increasing powers, RMSE increased continuously and bias fluctuated between positive and negative.

The normal annual rainfall map obtained through linear local polynomial interpolation was found to be superior and is presented in Fig. [6.](#page-11-0) The weighted average rainfall by this method was 1,595 mm.

4.5 Splines Method

The best Spline method was Inverse Multiquadric Spline with RMSE value 392.4, MAE value 282.3. The optimised parameter were 7,900.7 with 44 stations being forcefully included in the neighbourhood. This was followed closely by Spline with Tension with RMSE 397.8 and bias of 21.4. Further, RMSE of completely Regularised Spline was 397.9,

Fig. 7 Normal annual rainfall map using spline interpolation method

Multiquadric was 417.9 and Thin Plate Spline was 504.3. It is important that all the different types of Splines as mentioned above produced similar surfaces. Inclusion of anisotropy from secondary information produces zigzag toothy surfaces.

The normal annual rainfall map prepared using Inverse Multiquadric Spline is shown in Fig. [7.](#page-12-0) The weighted average rainfall obtained by interpolation using this method was 1595 mm.

4.6 Kriging

All the Kriging models (e.g. Nugget effect, Spherical, Exponential, Gaussian, Hole effect, K-Bessel and J-Bessel) were used for rainfall interpolation. From a few exploratory analyses, it was observed that log transformation makes the data near-normal. Thus, the log transformed data were used to improve the result. It has also been observed in the region that the rainfall varies from *Terai* and *Bhabar* region to *Shivaliks* and then to Lesser Himalayas and the Great Himalayas (Climate of Uttar Pradesh [1989](#page-19-0)). Therefore, anisotropy has been introduced in all the trials as suggested by Isaaks and Srivastava [\(1989](#page-20-0)).

Fig. 8 Normal annual rainfall map using Ordinary Kriging method

Fig. 9 Normal annual rainfall map using Universal Kriging method

4.6.1 Ordinary Kriging

The minimum RMSE (329.1) and MAE (243.9) was obtained using Hole Effect model with constant trend and log transformed data. The hole effect model relates to existence of two high valued rainfall fields (Goovaerts [2000](#page-20-0)): one at the Shivaliks and the other one at the foot of the Great Himalayas. The corresponding parameters used in Ordinary Kriging are as follows:

The weighted average rainfall by Ordinary Kriging was 1,608 mm. The corresponding map is shown in Fig. [8](#page-13-0).

4.6.2 Universal Kriging

The Hole Effect model under constant trend is found to be the best with RMSE of 328.7 and MAE of 243.1. For interpolation at each point, 16 neighbourhood stations were considered. The corresponding parameters of Universal Kriging are as follows:

The weighted average rainfall was computed as 1,608 mm. The interpolated normal annual rainfall map using these parameters is given in Fig. [9](#page-14-0).

5 Results and Discussion

The observed and predicted rainfall by various methods, error in prediction and percentage error for all stations are given in Table [2.](#page-15-0) It also shows the altitudes of raingauge stations. The statistics of prediction errors using various interpolation algorithms are given in Table 3. The error statistics clearly demonstrate that Kriging performs better in the region than other methods.

Universal Kriging with Hole Effect model and natural logarithmic transformation incorporating constant trend is found to be the best suitable method for interpolation of rainfall in the region (RMSE 328.7). This is followed by Ordinary Kriging (329.1), Splines (392.4), IDW (409.8) and Polynomial Interpolation (418.5). The range of correlation of rainfall is computed as 70 km – which means that rainfall at a location more than 70 km away from a raingauge station has all chances of being inaccurately predicted, even at annual level. The same has been observed by Upadhyay and Bahadur [\(1982](#page-21-0)), too.

In all cases, the *largest under prediction* occurs at Nainital station which records high annual rainfall and is surrounded by stations recording less rainfall. The *largest over* prediction occurs at Tehri station for all interpolation schemes except Polynomial, as shown in Table [2](#page-15-0). The largest over prediction for Polynomial interpolation occurs at Joshimath: it

Algorithm	Min	Max	Mean	Over prediction		Under prediction		Bias	RMSE
				Error	$%$ Error	Error	$%$ Error		
Thiessen Poly	948.2	2.986.0	1.604.6	$\overline{}$					
IDW	952.5	2.967.4	1.641.9	863.2	89.8	$-1.055.4$	40.6	60.7	409.8
Polynomial	1,196.8	2,282.0	1,594.8	1,120.2	115.2	$-1,041.9$	40.11	12.3	418.5
Spline	928.6	2.982.9	1.594.9	778.0	54.5	$-1,023.2$	39.4	20.2	392.4
OK	1,006.7	2,553.5	1.608.3	606.1	63.1	-947.3	36.5	7.3	329.1
UK	1,009.2	2.494.8	1,607.9	600.3	62.4	-948.3	36.5	5.9	328.7
Measured	948.2	2.986.0	1,652.5	$\overline{}$					

Table 3 Statistics of prediction errors using different interpolation schemes

Algorithm	Occurrence of absolute error more than $(\%)$							
	100 mm	200 mm	300 mm	400 mm	500 mm	600 mm		
IDW	72.7	56.8	38.6	31.8	25.0	15.9		
Polynomial	65.9	56.8	43.2	27.3	20.5	13.6		
Spline	65.9	43.2	36.4	31.8	20.5	13.6		
OK	72.7	47.7	29.5	15.9	13.6	9.1		
UK	70.5	45.5	27.3	18.2	11.4	9.1		

Table 4 Frequency of occurrence of absolute errors of different magnitudes

records low values compared to nearby stations, and being the last station towards north– east, is susceptible to edge effect (Table [1\)](#page-5-0). It is more appropriate to compare the estimates in terms of frequency of occurrence of errors of different magnitude from the point of view of water resources management planning. Table 4 shows that in 80% of the cases Kriging estimates are correct within a limit of ± 400 mm; and in 70% of the cases within ± 300 mm. Errors of higher magnitude occur only in around 10% of the cases. The analysis of spatial patterns of error distribution makes it apparent that the estimation errors are larger near the northern part for all the methods, where the network is sparse and extrapolation begins. The predictions are better in the south-eastern part, where the network is dense and area is relatively flat.

Zone	Min rain (mm)	Max rain (mm)	Rainfall range (mm)	Mean areal (mm)	Std. dev. (mm)
Physiographic regions					
The Great Himalayas	1,029.4	2,490.1	1,460.8	1,634.6	268.3
Lesser Himalayas	1,009.2	2,494.8	1,485.6	1,590.7	301.6
Shivaliks	1,279.9	2,352.2	1,072.3	1,781.9	213.7
Terai/Bhabar	1,064.6	2,092.4	1,027.8	1,472.4	251.9
Districts					
Almora	1,192.8	1,727.4	534.6	1,345.2	101.1
Bageshwar	1,301.4	2,494.4	1,193.0	1,973.7	283.6
Chamoli	1,029.4	2,177.2	1,147.9	1,497.5	229.6
Champawat	1,195.9	2,001.5	805.5	1,562.7	230.6
Dehradun	1,346.2	2,362.5	1,016.3	1,875.3	283.6
Garhwal	1,058.1	2,124.1	1,066.0	1,607.5	247.4
Haridwar	1,064.6	1,782.3	717.7	1,259.9	139.2
Nainital	1.308.8	2.105.5	796.7	1.738.0	203.7
Pithoragarh	1,205.6	2,494.8	1,289.3	1,715.2	285.9
Rudraprayag	1,114.2	2,179.3	1,065.1	1,667.9	272.7
Tehri Garhwal	1,009.2	2,325.8	1,316.6	1,576.6	391.9
Udham Singh Nagar	1,273.2	1,894.4	621.2	1,474.7	146.6
Uttarkashi	1,156.8	2,149.0	992.2	1,599.7	155.9
Administrative divisions					
Garhwal	1,009.2	2,362.5	1,353.3	1,580.2	280.9
Kumaun	1,192.8	2,494.8	1,302.0	1,650.6	292.7
Entire state					
Uttarakhand	1,009.2	2,494.8	1,485.6	1,607.9	287.7

Table 5 Predicted rainfall statistics for physiographic regions, districts and administrative divisions

Zone wise, the least errors occur over the plains and the largest over the Great Himalayas. Though all schemes depict the two high rainfall zones – one at Shivaliks and the other at the foot of the Great Himalayas – intervened by a zone of low rainfall, only geostatistical methods succeed to reveal its alignment parallel to the physiographic regions. An analysis of percentage error with elevation indicates minimum error in the elevation range of 200–400 m.

The mean areal rainfall computed by different methods ranges between 1,595 and 1,653 mm (Table [5\)](#page-18-0), of which the Universal Kriging estimate of 1,608 mm should be considered the mean areal rainfall for the region. The rainfall averages for the physiographic regions, districts, administrative divisions and entire region as obtained by Universal Kriging are provided in Table [5.](#page-18-0)

6 Conclusions

The major findings of the study are

- (a) The Universal Kriging method with Hole Effect model and natural logarithmic transformation incorporating constant trend is found to be the best suitable method for interpolation of rainfall in this region (RMSE 328.7).
- (b) At annual scale, there is a spatial correlation among neighbouring raingauges up to a distance of 70 km.
- (c) The mean areal rainfall of the state as per the best estimate is 1,608 mm.

A continuous data set of the average mean rainfall has been created which shall bridge the existing gap in knowledge of spatial distribution of rainfall over the Himalayan region lying in Uttarakhand region. The dataset may find its use in various hydrological studies and water resources management and planning studies.

Acknowledgements Authors thankfully acknowledge the Department of Science and Technology, Government of India, for providing financial support under U-PROBE project. The authors are also thankful to two anonymous reviewers for their constructive suggestions.

References

- Arora M, Singh P, Goel NK, Singh RD (2006) Spatial distribution and seasonal variability of rainfall in a mountainous basin in the Himalayan region. Water Resour Manage 20:459–508
- Barry RG (1981) Mountain weather and climate. Methuen & Co. Ltd, New York
- Brunsdon C, McClatchey J, Unwin DJ (2001) Spatial variations in the average rainfall–altitude relationship in Great Britain: an approach using geographically weighted regression. Int J Climatol 21:455–466
- Burrough PA, McDonnel RA (1998) Principles of geographical information systems. Oxford University Press, New York
- Campling P, Gobin A, Feyen J (2001) Temporal and spatial rainfall analysis across a humid tropical catchment. Hydrol Process 15:359–375
- Chang KT (2002) Introduction to geographic information systems. Tata McGraw Hill Publishing Company Limited, New Delhi
- Cheng SJ, Hsieh HH, Wang YM (2007) Geostatistical interpolation of space–time rainfall on Tamshui River basin, Taiwan. Hydrol Process 21:3136–3145
- Chua SH, Bras RL (1982) Optimal estimators of mean areal precipitation in regions of orographic influence. J Hydrol 57:23–48
- Climate of Uttar Pradesh (1989) India Meteorological Department, Govt. of India, Pune, p 380
- Creutin JD, Obled C (1982) Objective analyses and mapping techniques for rainfall fields: an objective comparison. Water Resour Res 18(2):413–431

Daly C (2006) Guidelines for assessing the suitability of spatial climate data sets. Int J Climatol 26:707–721

- Daly C, Neilson RP, Phillips DL (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J Appl Meteorol 33(2):140–158
- Daly C, Gibson WP, Taylor GH, Johnson GL, Pasteris P (2002) A knowledge-based approach to the statistical mapping of climate. Clim Res 22:99–113

Delfiner P, Delhomme JP (1973) Optimum interpolation by kriging. In: Davis JC, McCullagh MJ (eds) Display and analysis of spatial data. Nato Advanced Study Institute, John Wiley and Sons, London

Delhomme JP (1978) Kriging in the hydrosciences. Adv Water Resour 1(5):251–266

Dhar ON, Bhattacharya BK (1976) Variation of rainfall with elevation in the Himalayas – a pilot study. Indian J Power River Val Dev XXVI(6):179–185

Diodato N (2005) The influence of topographic co-variables on the spatial variability of precipitation over small regions of complex terrain. Int J Climatol 25:351–363

Dirks KN, Hay JE, Stow CD, Harris D (1998) High-resolution studies of rainfall on Norfolk Island, Part II: interpolation of rainfall data. J Hydrol 208(3–4):187–193

Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press, New York

- Goovaerts P (1999) Performance comparison of geostatistical algorithms for incorporating elevation into the mapping of precipitation. Geocomputation 99. [http://www.geovista.psu.edu/sites/geocomp99/Gc99/023/](http://www.geovista.psu.edu/sites/geocomp99/Gc99/023/gc_023.htm) [gc_023.htm](http://www.geovista.psu.edu/sites/geocomp99/Gc99/023/gc_023.htm). Accessed June 2004
- Goovaerts P (2000) Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J Hydrol 228:113–129
- Guenni L, Hutchinson MF (1998) Spatial interpolation of the parameters of a rainfall model from ground based data. J Hydrol 212–213:335–347
- Gyalistras D (2003) Development and validation of a high resolution monthly gridded temperature and precipitation data set for Switzerland (1951–2000). Clim Res 25:55–83
- Hevesi JA, Flint AL, Istok JD (1992) Precipitation estimation in mountainous terrain using multivariate geostatistics. Part II: isohyetal maps. J Appl Meteorol 31:677–688
- Higuchi K, Ageta Y, Yasunari T, Inoue J (1982) Characteristics of precipitation during the monsoon season in high-mountain areas of Nepal Himalaya. Hydrological Aspects of Alpine and High Mountain areas, IAHS Publication No. 138, 21–30
- Hutchinson MF, Gessler PE (1994) Splines more than just a smooth interpolator. Geoderma 62:45–67
- Hutchinson MF (1998) Interpolation of rainfall data with thin plate smoothing splines Part I: Two dimensional smoothing of data with short range correlation. J Geogr Inf Decis Anal 2(2):139–151
- Isaaks EH, Srivastava RM (1989) An introduction to applied geostatistics. Oxford University Press, New York Johansson B, Chen D (2003) The influence of wind and topography on precipitation distribution in Sweden: statistical analysis and modelling. Int J Climatol 23:1523–1535

Johnston K, Jay M, Hoef V, Krivoruchko K, Lucas N (2001) Using ArcGIS geostatistical analyst, ESRI

- Joshi SC (2004) Uttaranchal: environment and development a geo-ecological overview. Gyanodaya Prakashan, Nainital, Uttarakhand, p 426
- Kansakar SR, Hannah DM, Gerrard J, Rees G (2004) Spatial pattern in the precipitation regime of Nepal. Int J Climatol 24:1645–1659
- Lebel T, Bastin G, Obled C, Creutin JD (1987) On the accuracy of areal rainfall estimation: a case study. Water Resour Res 23(11):2123–2134
- Lloyd CD (2005) Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. J Hydrol 308:128–150
- Martinez-Cob A (1995) Estimation of mean annual precipitation as affected by elevation using multivariate geostatistics. Water Resour Manage 9:139–159
- Martinez-Cob A (1996) Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. J Hydrol 174:19–35
- Naoum, Tsanis (2003) Temporal and spatial variation of annual rainfall on the island of Crete, Greece. Hydrol Process 17:1899–1922
- Naoum S, Tsanis IK (2004) Orographic precipitation modelling with multiple linear regression. J Hydrol Eng 9(2):79–102
- Nguyen RT, Prentiss D, Shively JE (1998) Rainfall Interpolation for Santa Barbara County. [http://www.geog.](http://www.geog.ucsb.edu/dylan/rainfall/rainfall.html) [ucsb.edu/dylan/rainfall/rainfall.html](http://www.geog.ucsb.edu/dylan/rainfall/rainfall.html). Accessed September 2004
- Papamichail DM, Metaxa IG (1996) Geostatistical analysis of spatial variability of rainfall and optimal design of a rain gauge network. Water Resour Manage 10:107–127
- Philips DL, Dolph J, Marks D (1992) A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain. Agric For Meteorol 58:119–141
- Price DT, McKenney DW, Nalder IA, Hutchinson MF, Kesteven JL (2000) A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. Agric For Meteorol 101:81–94
- Prudhomme C, Reed DW (1998) Relationships between extreme daily precipitation and topography in a mountainous region: a case study in Scotland. Int J Climatol 18:1439–1453
- Prudhomme C, Reed DW (1999) Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in Scotland. Int J Climatol 19:1337–1356
- Rajagopalan B, Lall U (1998) Locally weighted polynomial estimation of spatial precipitation. J Geogr Inf Decis Anal 2(2):44–51
- Saveliev AA, Mucharamova SS, Piliugin GA (1998) Modeling of the daily rainfall values using surface under tension and kriging. J Geogr Inf Decis Anal 22:52–64
- Singh P, Kumar N (1997) Effect of orography on precipitation in the Western Himalayan Region. J Hydrol 199:183–206
- Singh P, Ramasastri KS, Kumar N (1995) Topographical influence on precipitation distribution in different ranges of Westerns Himalayas. Nord Hydrol 26:259–284
- Skirvin SM, Marsh SE, McClaran MP, Meko D (2003) Climate spatial variability and data resolution in a semi-arid watershed, South-eastern Arizona. J Arid Environ 54:667–686
- Subyani AM (2004) Geostatistical study of annual and seasonal mean rainfall patterns in the Southwest Saudi Arabia. Hydrol Sci J 49(5):803–817
- Tabios GQ, Salas JD (1985) A comparative analysis of techniques for spatial interpolation of precipitation. Water Resour Bull, Am Water Resour Assoc 21(3):365–380
- Tang C, Shindo S, Machida I (1998) Topographical effects on distributions of rainfall and ¹⁸O distributions: a case in Miyake Island, Japan. Hydrol Process 12:673–682
- Thomas A, Herzfeld UC (2004) REGEOTOP: new climatic data fields for East Asia based on localized relief information and geostatistical methods. Int J Climatol 24:1283–1306
- Tomczak M (1998) Spatial interpolation and its uncertainty using automated anisotropic Inverse Distance Weighting (IDW)-Cross-Validation/Jack knife Approach. J Geogr Inf Decis Anal 2(2):18–30
- Unwin DJ (1969) The areal extension of rainfall records: an alternative model. J Hydrol 7:404–414
- Upadhyay, Bahadur (1982) On some hydro meteorological aspects of precipitation in Himalayas. Proc. International Symposium on Hydrological Aspects of Mountainous Watersheds, School of Hydrology, University of Roorkee, Manglik Prakashan, Saharanpur, U.P., Vol.-I, I-58–I-65
- Vicente-Serrano SM, Saz-Sanchez MA, Cuadrat JM (2003) Comparative analysis of interpolation methods in the Middle Ebro Valley (Spain): application to annual precipitation and temperature. Clim Res 24:161–180
- Wei H, Li JL, Liang TG (2005) Study on the estimation of precipitation resources for rainwater harvesting agriculture in semi-arid land of China. Agric Water Manag 71:33–45