International Conference "Water, Environment, Energy and Society" (WEES–2009)
New Delhi, 12–16 January 2009

Geostatistical Analysis of Extreme Rainfalls of Andhra Pradesh, India

B.R. Krishna Murthy[1] and G. Abbaiah
Civil Engineering Department, JNT University
Kakinada, Andhra Pradesh, INDIA
E-mail: [1]bachu.krishna82@gmail.com

ABSTRACT: Rainfall is a hydrological phenomenon that varies in magnitude in space as well as in time and requires suitable tools to predict values in space and time. The identification of spatial rainfall pattern is an essential task for hydrologists, climatologists as well as regional and local planners and managers. This is due to the variability of both the temporal spatial distribution of rainfall. Multi-day rainfall events are an important cause of recent severe flooding in the India in general and in Andhra Pradesh in particular and are required for the design of structures such as dams, urban drainage systems and flood defences and cause failures to occur. Daily rainfall data from a network of 51 one degree gird stations for the period 1951–2003 has been in used for the study. On the basis of rainfall data for the heaviest storms that occurred in different part of Andhra Pradesh during the period 1951–2003, estimates of one-day, two-day, three-day, five-day and seven-day rainfalls were made. The main objectives of this work are: (1) to analyze and model the annual maximum rainfalls of various durations (2) to analyze and model the spatial variability of rainfall, (3) to interpolate kriging maps for different durations, and (4) to compare prediction errors and prediction variances with those of kriging methods for different durations. L-moment ratio diagrams have been used to identify candidate regional distribution of the data. Generalized Extreme Value distribution found to be the representative distribution. Parameters were estimated using maximum likelihood method. Return period quantiles were estimated using the fitted distribution for each station. Using these rainfall estimates, a geo-statistical analysis was performed. Rainfall surfaces have been predicted using ordinary kriging method. It was observed that the rainfall data is skewed and Box-cox transformation has been used for converting the skewed data to normal. It was found that one-day peak rainfall over the region varied from 25 mm to 360 mm. It is observed that the trend is present in all the cases, the first order polynomial fits well for all durations No-directional effects were observed in the region. The Spherical model fits well for higher order return periods of 50, 100 and 200 year return periods for one day and two day durations and for all return periods for five day and seven duration storms, where as the Gaussian model fits lower return periods up to 20 year for one-day, two-day duration storms. Fitted model resulted in a Mean Error (ME) varied in the range of –1.85 to –0.07, (which is very near to zero), Mean Square Error (MSE) altered in the range of 22.2 to 90.8 , (which is very low as compared to the variance of the data), Kriged Reduced Mean Square Error (KRMSE) of changed from 0.9547 to 1.09 , (which is very near to 1) and a Kriged Reduced Mean Error (KRME) varied in the range of –0.165 to –0.0153, (which is near to zero) for one-day duration events. The exploratory data analysis, variogram model fitting, and generation of prediction map through kriging were accomplished by using ESRI'S ArcGIS and geostatistical analyst extension.

Keywords: Geostatistical Analysis, Kriging, Annual Peak Rainfall, Multi-Day Durations.

INTRODUCTION

Rainfall information is an important input in the hydrological modeling, predicting extreme precipitation events such as draughts and floods, estimating quantity and quality of surface water and groundwater. However, in most cases, the network of the precipitation measuring stations is sparse and available data are insufficient to characterize the highly variable precipitation and its spatial distribution. This is especially true in the case of developing countries like India, where the complexity of the rainfall distribution is combined with the measurement difficulties. Therefore, it is necessary to develop methods to estimate rainfall in areas where rainfall has not been measured, using data from the surrounding weather stations. In recent years, there has been a growing interest in extreme events such as droughts, floods, etc., within the hydrological community. By definition, extreme events are rare but they do still occur and records are eventually broken. In this paper, we first carried out uni-variate analysis of extreme rainfalls of one to seven day durations and then rainfall is predicted using geo-statistical analyses over Andhra Pradesh. Regional distribution is identified with the L-moment ratio diagram. The method of L-moments introduced by Hosking (1990) is increasingly being used by hydrologists for flood frequency analysis. The L-moments are analogous to the conventional moments, but they have the theoretical

[1]Conference speaker

advantages of being able to characterize a wider range of distributions and, when estimated from a sample, of being more robust to the presence of outliers in the data. Hosking and Wallis (e.g., 1997) also point out that L-moments are less subject to bias in estimation than conventional moments.

Deterministic interpolation techniques create surfaces from measured points, based on either the extent of similarity (e.g., Inverse Distance Weighted) or the degree of smoothing (e.g., Radial Basis Functions). These techniques do not use a model of random spatial processes. Deterministic interpolation techniques can be divided into two groups, global and local. Global techniques calculate predictions using the entire dataset. Local techniques calculate predictions from the measured points within neighborhoods, which are smaller spatial areas within the larger study area.

Geostatistics assumes that at least some of the spatial variations of natural phenomena can be modeled by random processes with spatial autocorrelation. Geostatistical techniques produce not only prediction surfaces but also error or uncertainty surfaces, giving an indication of how good the predictions are. Many methods are associated with geostatistics, but they are all in the Kriging family. Ordinary, Simple, Universal, probability, Indicator, and Disjunctive kriging, along with their counterparts in cokriging, are available in the geostatistical analysis. Kriging is divided into two distinct tasks: quantifying the spatial structure of the data and producing a prediction. Quantifying the structure, known as variography, is where a spatial dependence model is fit to the data. To make a prediction for an unknown value for a specific location, kriging will use the fitted model from variography, the spatial data configuration, and the values of the measured sample points around the prediction location. The geostatistical analysis provides many tools to help determine which parameters to use, and also provides reliable defaults that can be used to make a surface quickly.

On the basis of rainfall data for the heaviest storms that occurred in different parts of Andhra Pradesh during the period 1951–2003, estimates of one-, two-, three-, five- and seven- day maximum rainfall for AP have been made. The main objectives of this study are: (1) to analyze and model the annual maximum rainfalls of various durations (2) to analyze and model the spatial variability of rainfall, (3) to interpolate kriging maps for different durations, and (4) to compare prediction errors and prediction variances with those of kriging methods for different durations.

STATE OF THE ART REVIEW

Fowler H.J and C.G. Kilsby (2003) used multi-day rainfall events in their study on regional frequency analysis of UK extreme rainfall from 1991 to 2000. They have used GEV distribution in the estimation of growth curves. Temporal changes in 1-, 2-, 5- and 10-day annual maxima are examined with L-moments using both a 10 year moving window and the fixed decades of 1961–70, 1971–80, 1981–90 and 1991–2000. Little change is observed at 1 and 2 days duration, but significant decadal-level changes are seen in 5- and 10-day events in many regions. Keshav P. Bhattarai, (2005) used the method of L-moments and the Generalized Extreme Value (GEV) distribution for flood frequency analysis for Irish river flow data. Jaiswal *et al.* (2003) used L-moment approach for flood frequency modeling.

Two graphical tools used in the earlier studies to assist in distribution selection are the sample average and a line of best-fit through the sample L-moment ratios. Hosking and Wallis (1995) have used the sample average, while the line of best-fit method was introduced by Vogel and Wilson (1996). These two graphical methods are subjective and are not a replacement for more objective and complex methods like those of Hosking and Wallis (1993) which take into account the sampling variability related to the sample size of the regional data. However, they do provide a quick visual assessment of which distribution may provide a good fit to the data.

Geostatistical methods have been shown to be superior to several other estimation methods, such as Thiessen polygon, polynomial interpolation, and inverse distance method by Creutin and Obled (1982), Tabios and Salas (1985). Basic concepts of the kriging technique and its application to natural phenomenon have been reviewed by the ASCE Task Committee (1990a, b). One of the advantages of the geostatistical methods is to use available additional information to improve precipitation estimations.

Kriging has been used in soil science (Vieria *et al.*, 1981; Berndtsson and Chen 1994; Bardossy and Lehmann 1998); Geostatistical methods were successfully used to study spatial distributions of precipitation by Dingman *et al.* (1988), Hevesi *et al.* (1992), extreme precipitation events by Chang (1991) and contaminant distribution with rainfall by Eynon (1998); Venkatram (1998). Krishna Murthy *et al.* (2007) applied geo-statistical interpolation techniques to estimate annual and seasonal rainfalls of Andhra Pradesh. Other studies in hydrology include, Creutin and Obled 1982; Storm *et al.,* 1988; Ahmed and de Marsily, 1989; Germann and Joss, 2001; Araghinejad and Burn, 2005; and atmosphere science (Bilonick, 1988; Casado *et al.,* 1994; Merino *et al.,* 2001). Kriging of groundwater levels was carried out by Delhomme (1978); Volpi and Gambolati (1978); Virdee and Kottegoda (1984); Kumar and Ahmed (2003), Vijay Kumar and Ramadevi (2006). Detailed information about geostatistical procedures can be found in literature Isaaks and Srivastava (1989), ESRI, (2001).

METHODS AND MATERIALS

Hydrologic variables exhibit substantial dependence over a wide range of temporal and spatial scales, and it is anticipated that their extremes do as well. Following sections describe in brief the methods adopted for the study.

L-moment Ratio Diagram

Generally the distribution selection process, using L-moment ratio diagrams, involves plotting the sample L-moment ratios as a scatter plot and comparing them with theoretical L-moment ratio curves of candidate distributions. L-moment ratio diagrams have been suggested as a useful tool for discriminating between candidate distributions to describe regional data (Hosking and Wallis, 1997). Numerous authors (Vogel and Wilson, 1996) have used L-moment ratio diagrams as part of their distribution selection process for regional data. Two graphical methods are often used in the distribution selection process, the sample average and a line of best-fit through the sample L-moment ratios.

Generalized Extreme Value Distribution

EVT is the branch of statistics which describes the behavior of the largest data observations. The historical cornerstone of EVT is the Generalized Extreme Value (GEV) distribution. The GEV distribution subsumes all three different extreme-value distributions (i.e., EV type I, II and III), to which the largest/smallest value from a set of independent and identically distributed random variables asymptotically tends. Consistently, several recent studies showed that the GEV distribution is a suitable statistical model for representing the frequency regime of rainfall extremes over the whole study area. The CDF (cumulative distribution function) of the GEV distribution is written as,

$$F_X(x) = \exp\left\{-\left[1-\frac{k(x-\xi}{\alpha}\right]^{1/k}\right\}, k \neq 0 \qquad \dots (1)$$

and

$$F_X(x) = \exp\left\{-\exp\left[-\frac{(x-\xi}{\alpha}\right]\right\}, k = 0 \qquad \dots (2)$$

while the quantile x (F) can be written as,

$$x(F) = \xi + \alpha\{1-(-\log F)^k\}/k, k \neq 0 \qquad \dots (3)$$

and

$$x(F) = \xi + \alpha \log(-\log F), k = 0 \qquad \dots (4)$$

where ξ is the location parameter, α the scale parameter, k the shape parameter of the distribution. When $k = 0$ the GEV distribution is equal to the Gumbel distribution.

Geostatistical Analysis

Although details on the kriging techniques are well documented (Isaaks and Srivastava 1989), a brief account of the relevant methods used is prescribed here. Kriging is a spatial interpolation method which is widely used in meteorology, geology, environmental sciences, agriculture etc. It incorporates models of spatial correlation, which can be formulated in terms of covariance or semivariogram functions. Parameters of the model viz., partial sill, nugget, range were estimated by minimizing the squared differences between empirical semivariogram values and theoretical model.

The first step in statistical data analysis is to verify three data features: dependency, stationarity and distribution. If the data are independent, it makes little sense to analyze them geostatistically. If data are not stationary, they need to be made so, usually by data detrending and data transformation. Geostatistics works best when input data are Gaussian. If not, data have to be made close to Gaussian distribution.

Preliminary analysis of the data will help in selecting the optimum geostatistical model. Preliminary analysis includes identifying the distribution of the data, looking for global and local outlier, global trends and examining the spatial correlation and covariance among the multiple datasets. Exploratory data analysis will help in accomplishing these tasks. With information on dependency, stationarity and distribution, one can proceed to the modeling step of geostatistical analysis, kriging.

Ordinary Kriging

The ordinary kriging model is,

$$Z(s) = \mu + \varepsilon(s) \qquad \dots (5)$$

Where $s = (X, Y)$ is a location and $Z(s)$ is the value at that location. The model is based on a constant mean μ for the data (no trend) and random errors $\varepsilon(s)$ with spatial dependence. It is assumed that the random process $\varepsilon(s)$ is intrinsically stationary. The predictor is formed as a weighted sum of the data,

$$Z(s_0) = \sum_{i=1}^{N} \lambda_i Z(s_i) \qquad \dots (6)$$

Where $Z(si)$ is the measured value at the ith location,
λi is an unknown weight for the measured value at the ith location
$S0$ is the prediction location.

In ordinary kriging, the weight λi depends on the semivariogram, the distance to the prediction location and the spatial relationships among the measured values around the prediction location.

Spatial Dependency

The goal of geostatistical analysis is to predict values where no data have been collected. The analysis will work on spatially dependent data. If the data are spatially independent, there is no possibility to predict values between them. Semi-variogram/covariance cloud is used to examine spatial correlation. Semi-variogram and covariance functions change not only with distance but also with direction. Anisotropy will help in studying the directional effects and identifying the optimal direction. Spatial dependency is given by,

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^{N} \left[Z(x_i + h) - Z(x_i)^2 \right] \qquad \dots (7)$$

where x, and $x + h$ are sampling locations separated by a distance h, $Z(x)$ and $Z(x + h)$ are measured values of the variable Z at the corresponding locations.

Data stationarity has been tested, data variance is constant in the area under investigation; and the correlation (covariance or semi-variogram) between any two locations depends only on the vector that separates them, not their exact locations.

Data has been tested for the presence of trend and directional effects. Global trend is represented by a mathematical polynomial which has been removed from the analysis of the measurements and added back before the predictions are made.

The shape of the semivariogram/covariance curve may also vary with direction (anisotropy) after the global trend is removed or if no trend exists. Anisotropy differs from the global trend because the global trend can be described by a physical process and modeled by a mathematical formula, The cause of anisotropy (directional influence) in the semivariogram is not usually known, so it is modeled as random error. Anisotropy is characteristic of a random process that shows higher autocorrelation in one direction than in another. For anisotropy, the shape of the semi-variogram may vary with direction. Isotropy exists when the semivariogram does not vary according to direction.

Transformations

Certain geostatistical interpolation assumes that the underlying data is normally distributed. Kriging relies on the assumption of stationarity. This assumption requires in part that all data values come from distributions that have the same variability. Transformations can be used to make the data normally distributed and satisfy the assumption of equal variability of data. Data brought to normal with help of suitable transformations. Some of the transformations adopted are Box-Cox, logarithmic, square-root transformation.

Box-Cox transformation,

$$Y(s) = \frac{\left(Z(s)^{\lambda} - 1\right)}{\lambda} \qquad \dots (8)$$

for $\lambda \neq 0$

Square root transformation occurs when $\lambda = ½$.

The log transformation is usually considered as a part of Box-Cox transformation when $\lambda = 0$,

$$Y(s) = \ln(Z(s)) \qquad \dots (9)$$

for $Z(s) > 0$ and ln is the natural logarithm. The log transformation is often used where the data has a positively skewed distribution and presence of very large values.

Cross-Validation

It gives an idea of how well the model predicts the unknown values. The objective of cross-validation is to help make an informed decision about which model provides the most accurate predictions.

The adequacy of the fitted models was checked on the basis of validation tests. In this method, known as jackknifing procedure, kriging is performed at all the data points, ignoring, in turn, each one of them one by one. Differences between estimated and observed values are summarized using the cross-validation statistics (de Marsily and Ahmed 1987): Mean Error (ME), Mean Squared Error (MSE), and Kriged Reduced Mean Error (KRME), and Kriged Reduced Mean

Square Error (KRMSE). If the semivariogram model and kriging procedure adequately reproduce the observed value, the error should satisfy the following criteria,

$$ME = \frac{1}{N}\sum_{i=1}^{N}\left[Z^{*}(x_i) - Z(x_i)\right] \cong 0 \quad \ldots (10)$$

$$ME = \frac{1}{N}\sum_{i=1}^{N}\left(Z^{*}(x_i) - Z(x_i)\right)^2 \text{ minimum} \quad \ldots (11)$$

$$KRME = \frac{1}{N}\sum_{i=1}^{N}\left[Z^{*}(x_i) - Z(x_i)/\sigma_{ki}\right] \cong 0 \quad \ldots (12)$$

$$KRMSE = \frac{1}{N}\sum_{i=1}^{N}\left[(Z^{*}(x_i) - Z(x_i))^2/\sigma_{ki}^{2}\right] \cong 1 \quad \ldots (13)$$

where, $z^*(xi)$, z(xi) and are the estimated value, observed value and estimation variance, respectively, at points xi. N is the sample size. As a practical rule, the MSE should be less than the variance of the sample values and KRMSE should be in the range $1 \pm 2\sqrt{2/N}$.

STUDY AREA

Andhra Pradesh the "Rice Bowl of India", is a state in southern India. It lies between 12°41′ and 22°N latitude and 77° and 84°40′E longitude, and is bordered with Maharashtra, Chhattisgarh and Orissa in the north, the Bay of Bengal in the East, Tamil Nadu in the south and Karnataka in the west. Andhra Pradesh is the fifth largest state in India by area and population. It is the largest and most populous state in South India. The state is crossed by two major rivers, the Godavari and the Krishna. The study has been limited, by necessity, to daily data, as sub-daily data are not generally available with sufficient coverage and length of record. However, daily data are adequate for the purposes of this study, since attention is focused on multi-day events.

Andhra Pradesh can be broadly divided into three unofficial geographic regions, namely Kosta (Coastal Andhra/Andhra), Telangana and Rayalaseema. Telangana lies west of the Ghats on the Deccan plateau. The Godavari River and Krishna River rise in the Western Ghats of Karnataka and Maharashtra and flow east across Telangana to empty into the Bay of Bengal in a combined river delta. Kosta occupies the coastal plain between Eastern Ghats ranges, which run through the length of the state, and the Bay of Bengal. Rayalaseema lies in the southeast of the state on the Deccan plateau, in the basin of the Penner River. It is separated from Telangana by the low Erramala hills, and from Kosta by the Eastern Ghats.

The rainfall of Andhra Pradesh is influenced by the South-West and North-West and North-East monsoons. The normal annual rainfall of the state is 925 mm. Major portion of the rainfall (68.5%) is contributed by South-West monsoon (June-Sept) followed by 22.3% byNorth-East monsoon (Oct.–Dec.). The rest of the rainfall (9.2%) is received during the winter and summer months. The rainfall distribution in the three regions of the state differs with the season and monsoon. The influence of south west monsoon is predominant in Telangana region (764.5 mm) followed by Coastal Andhra (602.26 mm) and Rayalaseema (378.5 mm). The North-East monsoon provides a high amount of rainfall (316.8 mm) in Coastal Andhra area followed by Rayalaseema (224.3 mm) and Telangana (97.1 mm). There are no significant differences in the distribution of rainfall during the winter and hot weather periods among the three regions.

RESULTS AND DISCUSSIONS

One degree daily data for 51 stations for 1951–2003 obtained from India Meteorological Department were used for the study. On the basis of rainfall data for the heaviest storms that occurred in different parts of Andhra Pradesh during the period 1951–2003, estimates of one-, two-, three-, five- and seven- day maximum rainfall for AP have been made. The spatial distribution of selected stations is presented in Figure 1.

The first step in statistical analysis is to investigate descriptive characteristics of the data. Descriptive analysis can help the investigators to have a preliminary judgment of the data and to decide suitable approaches for further analysis. The most important descriptive statistics are mean, standard deviation and Coefficient of Variation (CV), calculated as standard deviation divided by mean. In hydrology, however, there are two other important moments namely coefficient of skewness (Cs) as the measure of symmetry and coefficient of kurtosis (Ck) as the measure of shape of frequency function. Sample L-CV, L-skewness moments were determined for annual maximum rainfalls for one, two, three, five and seven day events. Firstly, the three L-moment ratios L-CV, L-skewness and L-kurtosis were determined for the AM series at each site using a routine from Hosking (1997). Figure 2 shows a plot of L-CV against L-skewness for the all grid stations. It can be seen that, generally, central coastal regions display a greater L-CV value than north, south regions, suggesting higher variability in these regions. The highest L-skewness values are found in central coastal region. These fall as a move is made northwards, to much lower values in Southern regions. In simple terms, this suggests that more intense rainfall events are experienced in central coastal and north coastal regions of AP. It was found that one-day maximum rainfall over the Andhra Pradesh region varied from 25 mm to 360 mm, with high values in the central coastal and north coastal regions and low values in the Rayalaseema (interior parts) region.

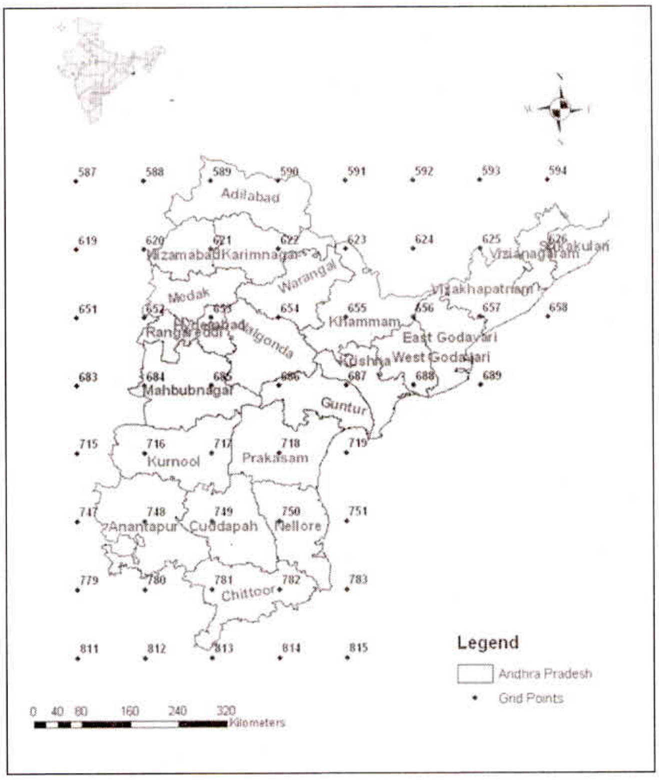


Fig. 1: Location map of study area

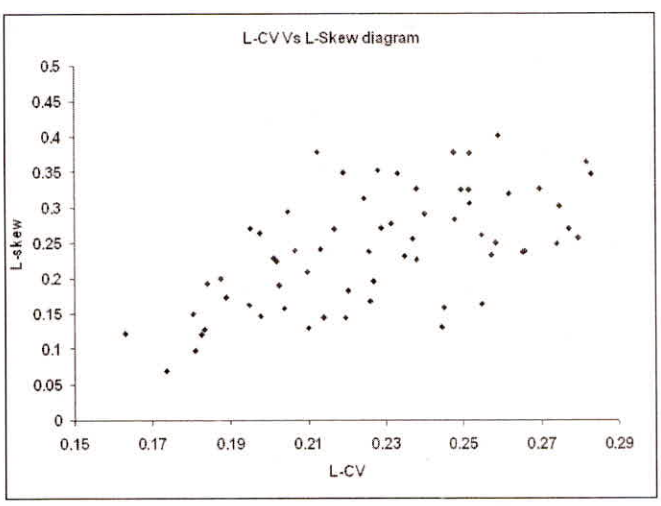
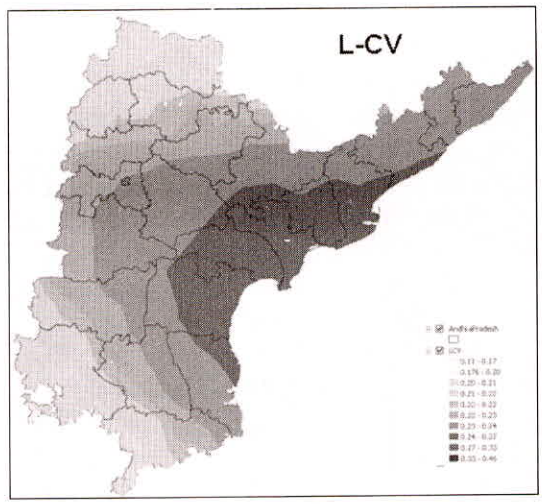
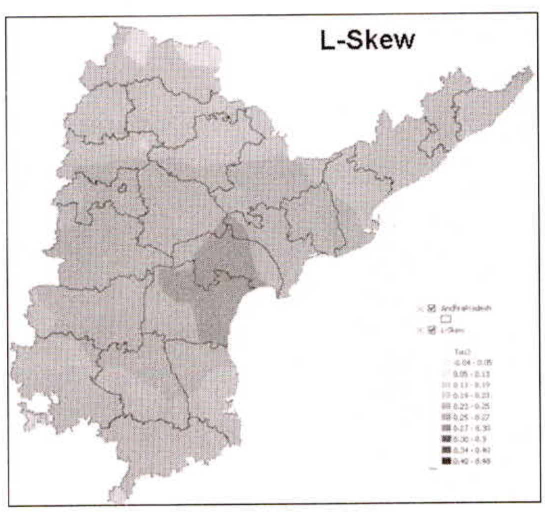


Fig. 2: L-CV and L-skew diagram for annual maximum daily rainfal

The proximity of the sample average (for regions with equal periods of record) or the record length weighted average (for regions with unequal periods of record) to a particular candidate distributions theoretical curve or point in L-skewness-L-kurtosis space has been interpreted as an indication of the appropriateness of that distribution to describe the regional data (Vogel *et al.*, 1993a; Hosking & Wallis, 1995). The diagram of L-moment ratios (see e.g., Hosking and Wallis, 1993) reported in Figure. 4 shows that the theoretical relationship between L-skewness (L-Cs) and L-kurtosis (L-Ck) for the GEV distribution is very close to the sample average L-Cs and L-Ck values for all storm duration of interest, therefore indicating that the GEV distribution is a suitable parent distribution.



(a) L-CV variation



(b) L-skew variation

Fig. 3: Spatial variation of L-CV and L-Skew for annual maximum daily rainfall, (a) L-CV variation (b) L-skew variation

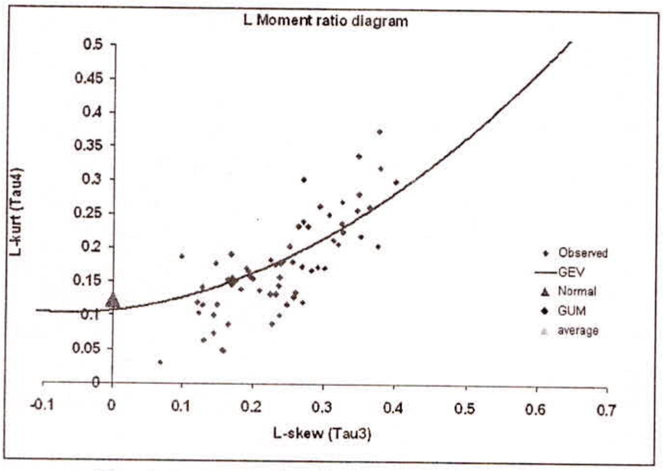


Fig. 4: L-moment ratio diagram for annual maximum daily rainfall

Figure 5 shows the variation of GEV shape parameter across Andhra Pradesh. According to results obtained, there are four main spatial groups of annual maximum rainfall over Andhra Pradesh, comprising central coastal Andhra Pradesh (Prakasam and Guntur districts), North Coastal Andhra (East, West Godavari, Vizaynagaram, Visakhapatnam, and Srikakulam), North AP (Telangana) and South Andhra Pradesh (Rayalaseema). Figure 6 shows time series plot of annual maximum one day rainfall for four grid locations 621 (Nizamabad), 657 (Visakhapatnam), 718 (Prakasam) and 748 (Anantapur) for the period 1951–2003, representing four regions having characteristics of high rainfall—high variability, high rainfall—medium variability, medium rainfall—medium variability and low- rainfall—low variability respectively. District names are given in the parentheses.

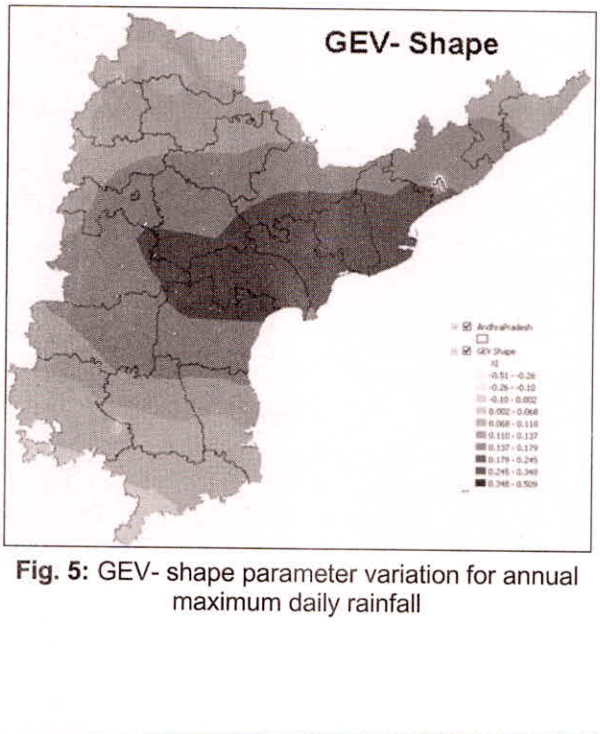


Fig. 5: GEV- shape parameter variation for annual maximum daily rainfall

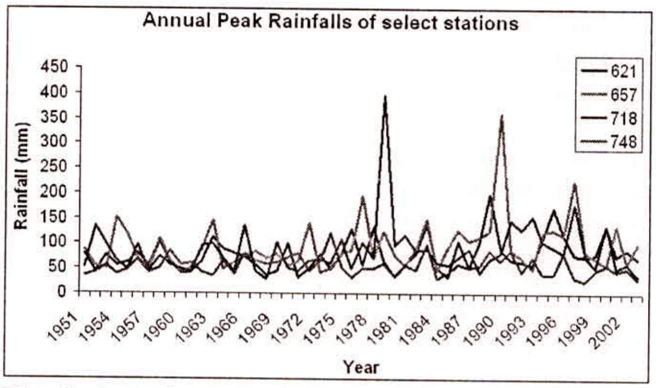


Fig. 6: Annual maximum rainfall for four grid locations 621 (Nizamabad), 657 (Visakhapatnam), 718 (Prakasam) and 748 (Anantapur) for the period 1951–2003

The quantile estimation was performed by using the Generalized Extreme Value (GEV) distribution. A quantile–quantile (Q–Q) plot for the GEV distribution (Figure. 7) indicates that the GEV fit is reasonably adequate. In all the reasons the GEV distributions appear to be acceptable. Frequency analysis of rainfall extremes using the maximum annual rainfall values for durations of one to seven day.

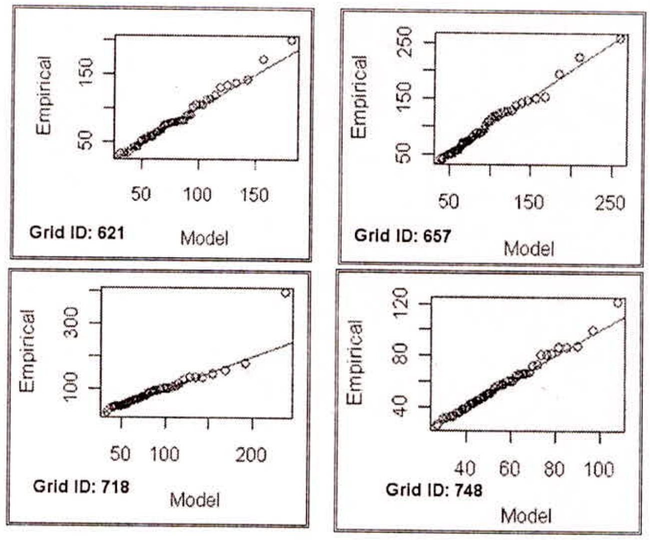


Fig. 7: Q–Q plot for fit of GEV distribution to annual daily maximum rainfalls for four grid locations 621 (Nizamabad), 657 (Visakhapatnam), 718 (Prakasam) and 748 (Anantapur). The solid blue line indicates the fit by a Generalized Extreme Value distribution and empirical estimates are given by circles

Basic statistics of the data annual maximum rainfalls for one-, two-, three-, five and seven day durations are shown in Table 1. It shows rainfall over AP is skewed and increases with durations of the storm durations. It was observed that skew coefficient is least from one duration storms and maximum seven day durations for a given return period and is positive for all durations. Skew coefficient is 1.046 for one day–10 year return period and 1.8638 for seven day–10 year return period. Box-Cox transformation has been used for converting the skewed distributed rainfall to normal and their coefficients are shown in Table 1. The transformation

Table 1: Basic Statistics of the Data Annual Maximum Rainfalls for One-, Two-, Three-, Five and Seven Day Durations

Item	*RPs*	*Average*	*Sd*	*Before Transformation*		*Box-Cox Parameter*	*After Transformation*	
				Skew	*Kurtosis*		*Skew*	*Kurtosis*
One day	2	84.5	29.90	1.2911	2.5607	–0.0589	–0.00100	3.380
	5	117.9	38.67	1.1864	2.0442	–0.1677	–0.00143	3.089
	10	142.7	45.31	1.0466	1.3663	–0.1454	0.00016	3.027
	20	169.2	53.48	0.9268	0.7703	–0.0858	0.00078	3.009
	50	208.1	69.19	0.9301	0.8275	–0.0373	0.00022	3.022
	100	241.8	87.15	1.1021	1.7848	–0.0556	–0.00084	3.078
	200	280.1	112.80	1.3862	3.3648	–0.1063	–0.00124	3.138
Two day	2	122.7	49.56	1.5938	3.4695	–0.0753	–0.00589	3.684
	5	170.5	63.64	1.5421	3.4493	–0.1596	–0.00770	3.461
	10	205.2	72.96	1.3804	2.8118	–0.1238	–0.00581	3.413
	20	241.7	83.37	1.1747	1.8671	–0.0491	–0.53971	4.036
	50	294.5	102.64	1.0261	1.0996	0.0301	–0.07080	3.392
	100	339.5	125.34	1.1881	2.0001	0.0264	0.00146	3.459
	200	390.2	159.47	1.6475	4.9412	–0.0249	–0.00224	3.697
Three Day	2	147.2	64.01	1.7689	3.9918	–0.0933	–0.01048	3.907
	5	203.1	81.80	1.7564	4.1567	–0.2080	–0.01572	3.633
	10	243.3	93.39	1.6205	3.6595	–0.1945	–0.01266	3.541
	20	285.0	105.72	1.4121	2.7478	–0.1278	–0.00686	3.463
	50	344.8	126.71	1.1469	1.5381	–0.0163	–0.00017	3.296
	100	395.0	149.56	1.0957	1.4123	0.0370	0.00092	3.208
	200	450.8	182.13	1.2600	2.4960	0.0408	0.00147	3.233
Five Day	2	182.0	87	1.9246	4.3442	–0.1303	–0.01778	4.037
	5	247.4	109	1.9188	4.4217	–0.2762	–0.02450	3.738
	10	293.4	122	1.8170	4.1107	–0.2841	–0.02008	3.608
	20	340.4	136	1.6367	3.4306	–0.2281	–0.01348	3.519
	50	406.6	158	1.3272	2.1370	–0.1053	–0.00376	3.338
	100	461.1	180	1.1443	1.3553	–0.0229	–0.00068	3.178
	200	520.9	212	1.1312	1.4096	0.0149	0.00028	3.102
Seven day	2	211.0	106	1.9848	4.5196	–0.1323	–0.01894	4.043
	5	283.4	131	1.9656	4.4284	–0.2870	–0.02780	3.778
	10	333.1	146	1.8638	4.0730	–0.3036	–0.02366	3.641
	20	382.8	160	1.6888	3.4048	–0.2577	–0.01611	3.527
	50	451.2	182	1.3782	2.1473	–0.1436	–0.21260	3.559
	100	506.6	204	1.1639	1.2543	–0.0576	–0.00139	3.128
	200	566.1	234	1.0781	0.9321	–0.0062	0.00021	3.001

greatly reduced the skew and the transformed series can be treated as nearly normal for further analysis and transformed values of skewness and kurtosis are also shown in the Table 1. Normality plot of three–100 year return period rainfall is shown in Figure 8.

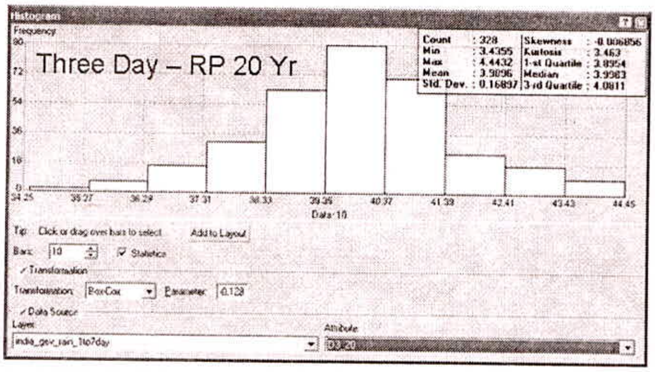


Fig. 8: Normality plot after Box-Cox transformation for three day—20 year return period rainfalls

The trend analysis enables to identify the presence/ absence of trends in the input dataset. If a trend exists in the data, it is the nonrandom (deterministic) component of a surface that can be represented by a mathematical formula and removed from the data. Once the trend is removed, the statistical analyses have been performed on the residuals. The trend will be added back before the final surface is created so that the predictions will produce meaningful results. By removing the trend, the analysis that is to follow will not be influenced by the trend, and once it is added back a more accurate surface will be produced. 3D perspective trend plot for two day 100 year return period rainfall is shown in Figure 9. It is observed that the trend is present in all the cases, and the first order polynomial fits well and all fits are tabulated in Table 2. It is further observed that trend varies from South-South-West to North-North-East direction.

The Semi-variogram has been used to examine the spatial autocorrelation among the measured sample points. In spatial autocorrelation, it is assumed that rainfalls that are close to one another are more alike. The Directional influences are also examined. No directional effects were found in this analysis. The Spherical model fits well for higher order return periods of 50, 100 and 200 year return periods for one day and two day durations and for all return periods for five day and seven duration storms, where as the Gaussian model fits lower return periods up to 50 year for one-day, two-day duration storms. For three day duration storms, Spherical model for lower return periods and Gaussian model for higher order return periods fits well. The best fit equations are tabulated in Table 2.

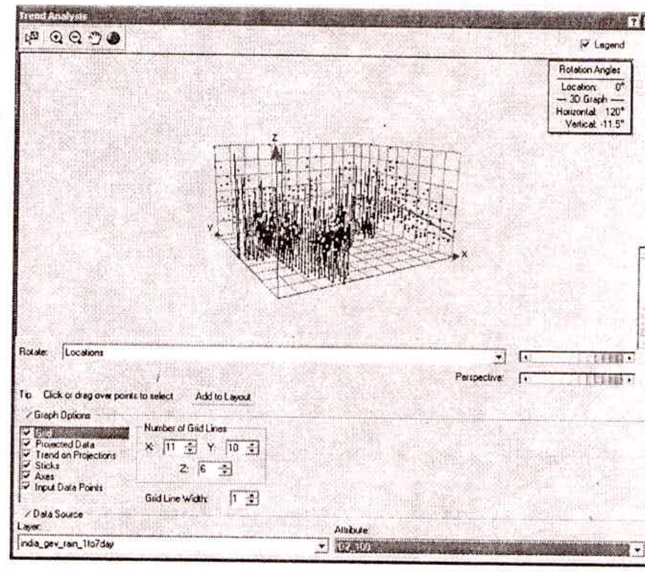


Fig. 9: Three dimensional trend plot for two day 100 year return period rainfall

After identifying the best fit variogram model, taking into account de-trending and directional influences in the data, prediction surfaces are generated for one-day to seven- rainfall events for return periods ranging from 2 to 100 year. The qualities of prediction map have been examined by creating the prediction standard error surface. The Prediction standard errors quantify the uncertainty for each location in the surface. A simple thumb rule is that 95% of the time, the true value of the surface will be within the interval formed by the predicted value +/– 2 times the predicted standard error if the data are normally distributed. It has been observed that the locations near the sample points have low error. For these analyses, ESRI's Geo-statistical analyst extension has been used.

The cross-validation gives an idea of how well the model predicts the unknown values. For all points, cross-validation sequentially omits a point, predicts its value using the rest of the data, and then compares the measured and predicted values. The calculated statistics serve as diagnostics that indicate whether the model is reasonable for map production. The criteria used for accurate prediction in the cross-validation are: the mean error should be close to zero, the root mean square error and average standard error should be as small as possible and the root mean square standardized error should be close to 1.

The cross-validation results are shown in Table 3. The Cross-validation statistics showed that the predicted values are reasonable for map production. Fitted model

Table 2: Kriging—Model Equations

Item	*Return Period*	*Model Equation*	*Nugget*	*Trend*
One day	2	0.044775*Gaussian(28.646)+0.044102*Nugget	0.0441	First
	5	0.0084719*Gaussian(28.646)+0.014357*Nugget	0.0144	First
	10	0.0097998*Gaussian(28.646)+0.016691*Nugget	0.0167	First
	20	0.017144*Gaussian(28.646)+0.030977*Nugget	0.0310	First
	50	0.029251*Spherical(28.646)+0.053391*Nugget	0.0534	First
	100	0.026164* Spherical (6.5822)+0.040905*Nugget	0.0409	First
	200	0.021804*Spherical(5.8902)+0.021771*Nugget	0.0218	First
Two day	2	0.031562*Gaussian(28.646)+0.04335*Nugget	0.0434	First
	5	0.011795*Gaussian(28.646)+0.016115*Nugget	0.0161	First
	10	0.01561*Gaussian(28.646)+0.021442*Nugget	0.0214	First
	20	0.0002354*Gaussian(28.646)+0.00042821*Nugget	0.0004	First
	50	0.0069599*Spherical(28.646)+0.001986*Nugget	0.0020	First
	100	0.074446*Spherical(28.646)+0.12872*Nugget	0.1287	First
	200	0.038997*Spherical(28.646)+0.086643*Nugget	0.0866	First
Three Day	2	0.028529* Spherical (28.646)+0.038245*Nugget	0.0382	First
	5	0.0075421*Spherical(28.646)+0.0084891*Nugget	0.0085	First
	10	0.0078485*Spherical(28.646)+0.0088607*Nugget	0.0089	First
	20	0.015277*Gaussian(28.646)+0.020605*Nugget	0.0206	First
	50	0.054698*Gaussian(28.646)+0.078368*Nugget	0.0784	First
	100	0.10295*Gaussian(28.646)+0.16483*Nugget	0.1648	First
	200	0.10826*Gaussian(28.646)+0.20464*Nugget	0.2046	First
Five Day	2	0.022225*Spherical(28.646)+0.024093*Nugget	0.0241	First
	5	0.003634*Spherical(28.646)+0.0040799*Nugget	0.0041	First
	10	0.0044177*Spherical(28.646)+0.0048725*Nugget	0.0049	First
	20	0.0050869*Spherical(28.646)+0.0057627*Nugget	0.0058	First
	50	0.020997*Spherical(28.646)+0.02453*Nugget	0.0245	First
	100	0.057118*Spherical(28.646)+0.071605*Nugget	0.0716	First
	200	0.092905*Spherical(28.646)+0.1317*Nugget	0.1317	First
Seven day	2	0.021979*Spherical(28.646)+0.024798*Nugget	0.0248	First
	5	0.0030601*Spherical(28.646)+0.0036329*Nugget	0.0036	First
	10	0.0021852*Spherical(28.646)+0.0026237*Nugget	0.0026	First
	20	0.0034349*Spherical(28.646)+0.004142*Nugget	0.0041	First
	50	0.012724*Spherical(28.646)+0.015871*Nugget	0.0159	First
	100	0.036785*Spherical(28.646)+0.049252*Nugget	0.0493	First
	200	0.070431*Spherical(28.646)+0.10653*Nugget	0.1065	First

Table 3: Kriging Model Cross-Validation Results

Item	*RPs*	*Mean*	*Root-Mean-Square*	*Average Standard Error*	*Mean Standardized*	*Root-Mean-Square-Standardized*
One day	2	–0.07	22.2	23.02	–0.0153	0.9754
	5	–1.14	29.8	31.10	–0.0182	0.9757
	10	–1.34	35.6	37.50	–0.0187	0.9660
	20	–1.51	42.9	45.57	–0.0177	0.9547
	50	–1.57	56.0	58.90	–0.1655	0.9559
	100	–1.04	69.1	67.90	–0.0132	1.0220
	200	–1.85	90.8	83.63	–0.0225	1.0900
Two day	2	1.38	36.0	36.61	–0.0098	0.9912
	5	1.97	47.3	48.61	–0.0150	0.9978
	10	–2.17	55.0	57.66	–0.0143	0.9788
	20	–1.23	64.8	71.32	0.0004	0.9910
	50	0.40	74.1	79.90	–0.0016	1.0000
	100	–2.30	102.1	105.20	–0.0106	0.9747
	200	–2.70	136.1	131.90	–0.0011	1.0280
Three Day	2	–1.77	46.0	45.84	–0.0066	1.0070
	5	–2.27	56.7	58.00	–0.0162	1.0050
	10	–2.58	66.0	68.38	–0.0170	0.9965
	20	–3.12	80.7	83.40	–0.0143	0.9993
	50	–3.30	100.9	104.90	–0.0139	0.9864
	100	–3.44	123.6	126.90	–0.0136	0.9938
	200	–3.63	156.2	155.70	–0.0135	1.0190
Five Day	2	–2.28	59.2	57.46	–0.0092	1.0200
	5	–3.39	76.1	74.04	–0.0192	1.0510
	10	–3.91	86.8	85.59	–0.0214	1.0400
	20	–4.14	98.8	99.30	–0.0214	1.0290
	50	–4.41	119.2	121.70	–0.0212	1.0130
	100	–4.73	142.3	144.20	–0.2210	1.0190
	200	–5.26	175.6	173.30	–0.0238	1.0460
Seven day	2	–2.83	73.3	69.36	–0.0082	1.0430
	5	–4.17	93.0	88.17	–0.0187	1.0770
	10	–4.72	105.3	101.30	–0.0208	1.0720
	20	–5.00	117.9	115.70	–0.0202	1.0520
	50	–5.17	138.8	139.50	–0.0186	1.0260
	100	–5.36	161.4	162.90	–0.0182	1.0210
	200	–5.70	193.3	192.70	–0.0186	1.0360

resulted in a Mean Error (ME) varied in the range of –1.85 to –0.07, (which is very near to zero), Mean Square Error (MSE) altered in the range of 22.2 to 90.8 , (which is very low as compared to the variance of the data), Kriged Reduced Mean Square Error (KRMSE) of changed from 0.9547 to 1.09, (which is very near to 1) and a Kriged Reduced Mean Error (KRME) varied in the range of –0.165 to –0.0153, (which is near to zero) for one-day duration events. Here the bracketed quantities refer to the requirements to consider a model as adequate. Similarly it was observed that for other storm durations, cross-validation statistics varied in a permissible range. The above cross validation results show that the chosen models are adequate.

Predicted rainfall surfaces and predicted error estimates are prepared using Geo-statistical analyst extension in ArcGIS environment. Figure 10 shows

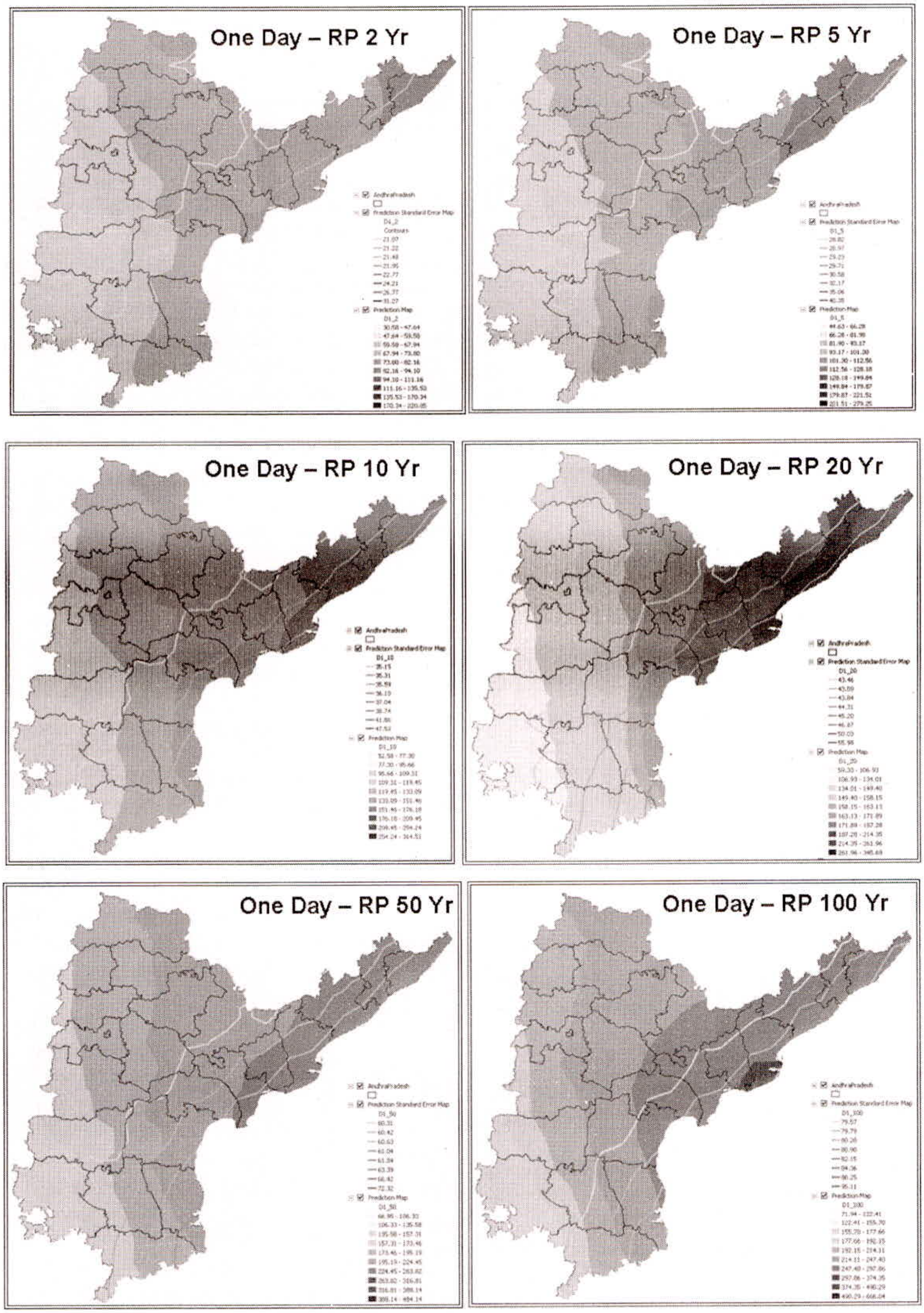


Fig. 10: Predicted rainfall surfaces and error maps for one duration maximum rainfall. Grid surface is the predicted rainfall surface and contours indicate the predicted error

predicted rainfall surfaces and error maps for one duration annual maximum rainfalls. Grid surface is the predicted rainfall surface and contours indicate the predicted error for 2, 5, 10, 20, 50 and 100 year return period. As seen from Figure 10, the estimation variance is low at in the central parts of study area and increase towards the boundaries. It indicates that the estimated rainfalls are highly reliable in the middle of the study area and at or near the boundary; these are not reliable to the same extent.

SUMMARY

In this study, annual maximum rainfalls for multi-day durations during the period 1951–2003 were used. It was found that the annual extremes are suitably described by a Generalized Extreme Value distribution. Return-period estimates presented here for different regions of the AP, using the most recent rainfall data, will allow the reassessment of the risk of failure of existing structures and facilitate the design of new structures incorporating better risk or uncertainty estimates. Kriging, a type of geostatistical techniques, is applied to multi-day rainfall estimates for different return periods. The rainfall surfaces were predicted using ordinary kriging method. It was observed that the rainfall data is skewed and Box-cox transformation has been used for converting the skewed data to normal. It is observed that the trend is present and first order polynomial fits well in all the cases. The cross-validation statistics showed that the predicted values are reasonable for map production. Finally, the realistic prediction surfaces and prediction error maps are generated.

REFERENCES

Ahmed, S. and de Marsily, G. (1989). "Cokriged estimates of transmissivities using jointly water level data", In: Armstrong, M. (ed.), *Geostatistic*, Kluwer Academic Pub., 2: 615–628.

ASCE Task Committee (1990a). "Review of geostatistics in geohydrology, I: Basic concepts." *J. of Hydraulic Engg.* (ASCE), Vol. 116, 612–632.

Araghinejad, S. and Burn, D.H. (2005). "Probabilistic fore-casting of hydrological events using geostatistical analysis." *Hydrological Sciences Journal*, Vol. 50, 837–856.

ASCE Task Committee (1990b). "Review of geostatistics in geohydrology, II: Applications." *J. of Hydraulic Engg.* (ASCE), Vol. 116,633–658

Bilonick, R.A. (1988). "Monthly hydrogen ion deposition maps for the North Eastern US from July 1982 to Sept. 1984." *Atmos. Environ.*, Vol. 22, 1909–1924.

Berndtsson, R. and Chen, H. (1994). "Variability of soil water content along a transect in a desert area." *J. of Arid Environ.*, Vol. 27, 127–139.

Bardossy, A. and Lehmann, W. (1998). "Spatial distribution of soil moisture in a small catchment. Part I: geostatistical analysis." *J. of Hydro.*, Vol. 206, 1–15.

Casado, L.S., Rohani, S., Cardelino, C.A. and Ferrier, A.J. (1994). "Geostatistical analysis and visualization of hourly ozone data." *Atmos. Envirn.*, Vol. 28, 2105–2118.

Chang, T. J. (1991). "Investigation of Precipitation Droughts by Use of Kriging Method." *J. of Irrig. and Drain. Engg.* 11 7(6), 935–943.

Creutin, J.D. and Obled, C. (1982). "Objective analysis and mapping techniques for rainfall fields: An objective comparison." *Water Res. Res.* 18(2), 4 13–43 1.

Delhomme, J.P. (1978). "Kriging in the hydroscience." *Adv. in Water Resour.*, Vol. 1,251–266.

De Marsily, G. and Ahmed, S. (1987). "Application of kriging techniques in groundwater hydrology." *J. Geol. Soc. of India*, Vol. 29, 57–82.

Dingman, S.L., Seely-Reynolds, D.M. and Reynolds 111, R.C. (1988). "Application of kriging to estimate mean annual precipitation in a region of orographic influence." Water Resour. Bull. 24(2), 329–339.

ESRI (2001). "Using ArcGIS Geostatistical Analyst".

Eynon, B.P. (1988). "Statistical analysis of precipitation chemistry measurements over the eastern United States. Part 11: Kriging analysis of regional patterns and trends." *J. Appl. Meteor.* 27, 1334–1343.

Fowler H.J. and Kilsby, C.G. (2003). "A Regional Frequency Analysis of United Kingdom Extreme Rainfall from 1961 to 2000", *Int. J. Climatol.* 23: 1313–1334 (2003).

Germann, U. and Joss, J. (2001). "Variograms of radar reflectivity to describe the spatial continuity of apline precipitation." *J. of Applied Meteorology*, Vol. 40,1042–1059.

Hevesi, J.A., Istok, J.D and Flint, A.L. (1992). "Precipitation estimation in mountainous terrain using multivariate geostatistics. Part I: Structural analysis." *J.Appl. Meteor.* 3 1, 661–676.

Hosking, J.R.M. and Wallis, J.R. (1988). "The Effect of Inter-Site Dependence on Regional Flood Frequency Analysis", *Water Resour. Res.*, 24, 588–600.

Hosking, J.R.M. and Wallis, J.R. (1993). "Some statistics useful in regional frequency analysis", *Water Resour. Res.*, 29(2), 271–281.

Hosking, J.R.M., Wallis, J.R. (1997). "Regional Frequency Analysis: An Approach Based on L-Moments." Cambridge University Press.

Jenkinson, A.F. (1995). "The frequency distribution of the annual maximum (or minimum) of meteorological elements", *Q.J. Royal Meteorol.* Soc., 81.

Isaaks, E.H. and Srivastava, R.M. (1989). "An Introduction to Applied Geostatistics": Oxford University Press, New York.

Jaiswal, R.K., Goel, N.K., Singh, P. and Thomas, T. (2003). "L-moment based Flood Frequency Modelling", *Institution of Engineers Journal, India*, Vol. 84. 6–10.

Keshav, P. Bhattarai (2005). "Flood Frequency Analysis of Irish River Flow Data using Variants of L-Moments", *National Hydrology seminar*, 2005.

Kumar, D. and Ahmed, S. (2003). "Seasonal behaviour of spatial variability of groundwater level in a granitic aquifer in monsoon climate." *Current Science*, Vol. 84, 188–196.

Krishna Murthy, B.R. and Abbaiah, G. (2007). "Geostatistical analysis for estimation of mean Rainfalls in Andhra Pradesh, India", *International Journal of Geology*, issue 3, Vol 1, pp. 35-51.

Merino, G.G., Jones, D., Stooksbury, D.E. and Hubbard, K.G. (2001). "Determination of semivariogram models to krige hourly and daily solar irradiance in Western Nebraska." *J. of Applied Meteorology*, Vol. 40, 1085–1094.

Storm, B., Jenson, K.H. and Refsgaard, R.C. (1988). "Estimation of catchment rainfall uncertainty and its influence on runoff prediction." *Nordic Hydrology*, Vol. 19, 79–88.

Tabios, G.Q. and Salas, J.D. (1985). "A comparative analysis of techniques for spatial interpolation of precipitation." *Water Resources Research* (Check).

Venkatram, A. (1988). "On the use of kriging in the spatial analysis of acid precipitation data." *Atmospheric environment* 22(9), 1963–1975.

Vijay Kumar and Remadevi (2006). "Kriging of Groundwater Levels"—*A Case Study Journal of Spatial Hydrology* Vol. 6, No. 1 (2006), pp. 81–94.

Volpi, G. and Gambolati, G. (1978). "On the use of main trend for the kriging technique in hydrology." *Adv. Water Resour.*, Vol. 1, 345–349.

Virdee, T.S. and Kottegoda, N.T. (1984). "A brief review of kriging and its application to optimal interpolation and observation well selection." *Hydro. Sci. J.*, Vol. 29, 367–387.