Research Article

MODELLING OF SEASONAL AND ANNUAL RAINFALL OF WEST COAST DRAINAGE SYSTEM USING DATA TRANSFORMATION METHODS

*N. Vivekanandan
Central Water and Power Research Station, Pune 411024, Maharashtra, India
*Author for Correspondence

ABSTRACT
Estimation of seasonal and annual rainfall for a desired return period is of prime importance for planning of various water resources management projects. This paper details a study involving the use of five transformations such as square root, cube root (CRT), logarithmic, box-cox (BCT) and modified SMEMAX for modelling seasonal and annual rainfall of West Coast Drainage System (WCDS) of India. Kolmogorov-Smirnov test is employed for checking the applicability of the transformations to the recorded rainfall data. Model performance indicators such as root mean square error and correlation coefficient are used for selection of suitable transformation for modelling seasonal and annual rainfall of WCDS. The study shows that BCT is the best suited amongst five transformations for modelling monsoon, post-monsoon, winter and annual rainfall patterns while CRT for summer rainfall pattern of WCDS.

Key Words: Correlation, Kolmogorov-Smirnov, Rainfall, Root Mean Square Error, Transformation

INTRODUCTION
Rainfall analysis for any region is essential in planning and design of irrigation and drainage systems, and overall programme of command area development. As the distribution of rainfall varies over space and time, it is required to analyse the data covering long periods and recorded at various locations to obtain reliable information. Further, these data needs to be analysed in different ways depending on the problem under considerations. For example, analysis of consecutive days of rainfall is more relevant for drainage design of agricultural lands whereas analysis of weekly rainfall data is for planning of cropping pattern. Likewise, analysis of monthly, seasonal and annual data is more useful for water management practices. Deterministic and probabilistic approaches are commonly used in rainfall data analysis for estimation of expected rainfall of a given frequency (Hameed et al, 1982 and Jam and Singh, 1986).
Deterministic approach to rainfall modelling generally adopts a procedure wherein the river system is modelled by a set of variables, often representing the catchment characteristics, wherein each combination of variable values represents a unique state of condition of the system, with the manipulation of the variable simulating the movement of the system from state to state (Vivekanandan and Mathew, 2009). Probabilistic approach includes fitting of standard probability distributions to recorded rainfall data, and transformation methods wherein the techniques attempt to find a unique distribution in the transformed range regardless of the distribution of the original sample (May, 2004; Sharda and Das, 2005; and Bhaskar et al, 2006).
Transformation methods attempt to transform the original data, which has an arbitrary distribution into a group of data with a known distribution. The advantages of transformations methods are as follows: (i) These methods do not fit a distribution to the sample itself, no assumptions about the underlying distribution are required; (ii) There is no need to attempt to find sometime tenuous theoretical justification upon which many of the distributions are based; and (iii) These methods can be realized if a relatively simple distribution is used in the transformed range (Mohammad, 2011).
Number of transformations such as Square Root (SRT), Cube Root (CRT), Logarithmic (LGT), Box-Cox (BCT) and Modified SMEMAX (MSX) are widely used for modelling seasonal (monsoon, pre-monsoon, winter and summer) and annual rainfall. Research reports indicate that there is no general agreement for adopting a standard transformation for a region or a country for rainfall estimation though different hydrologists have recommended different transformation methods for modelling rainfall data (Panigrahi et al, 2002). This paper details a study for identifying the most suitable transformation for modelling seasonal and annual rainfall for West Coast Drainage System (WCDS) of India. Goodness-of-Fit (GoF) test involving

Research Article

Kolmogorov-Smirnov (KS) statistic is employed for checking the applicability of the transformation to the recorded rainfall data. Model performance indicators (MPIs) such as root mean square error (RMSE) and correlation coefficient (CC) are used for selection of suitable transformation for modelling seasonal and annual rainfall of WCDS.

MATERIALS AND METHODS

Transformation methods are used to transform the sample data to a near normal distribution, thus underlying it tractable by bringing into its purview a whole gamut of developments, irrespective of the underlying distribution of the sample data. The methodology adopted in estimating the rainfall by five transformations, GoF tests and MPIs are detailed below:

Square Root, Cube Root and Logarithmic Transformations
The parameters are determined by maximum likelihood method and used to estimate the rainfall (R_T) using back transformations of SRT, CRT and LGT, which can be expressed by:

$$R_T = (A + K_T B)^2 \text{ (for SRT)}; \quad R_T = (A + K_T B)^3 \text{ (for CRT) and } R_T = 10^{A+K_T B} \text{ (for LGT)} \qquad \dots \quad (1)$$

Here, A and B are parameters of the transformations (Islam and Kumar, 2003). K_T is the frequency factor corresponding to coefficient of skewness ($C_s=0$) and probability of exceedance for a given return period (T).

Box-Cox Transformation
Box-Cox transformation (BCT) is the generalized form of power transformation. According to this method, transformed variate (Y_i) is given as:

$$Y_i = \begin{cases} \dfrac{R_i^{\lambda} - 1}{\lambda}, \lambda \neq 0 \\ Log(R_i), \lambda = 0 \end{cases} \qquad \dots \quad (2)$$

where λ is a constant of transformation (Kumar and Devi, 1982). The value of λ is determined by trial and error method so that C_s of the transformed series becomes zero. The rainfall estimate (R_T) for a given return period can be obtained using back transformation as follows:

$$R_T = \begin{cases} (\lambda Y_T + 1)^{1/\lambda}, \lambda \neq 0 \\ Exp(Y_T), \lambda = 0 \end{cases} \quad \text{with } Y_T = \overline{Y} + K_T \sigma_Y \qquad \dots \quad (3)$$

Here, Y_T is the reduced variate of transformed data, $\overline{Y}$ and σ_Y are mean and standard deviation of the transformed data.

Modified SMEMAX Transformation
MSX is the modified version of SMEMAX (Small, MEdian, MAXimum) transformation (Prakash, 1981 and Rasheed et al, 1982). According to this method, the transformed variate (Y_i) is given as:

$$Y_i = \begin{cases} Y_M \left(1 - \dfrac{(R_M - R_i)}{(R_M - R_S)} \right), R_i \langle R_M \\ Y_M \left(1 - \dfrac{(R_M - R_i)}{(R_L - R_M)} \right), R_i \rangle R_M \end{cases} \qquad \dots \quad (4)$$

where $Y_M = ((R_M - R_S)/2CosA)$ and transformed angle $(A) = arctan((R_L - R_M)/(R_M - R_S))$. R_S, R_M and R_L are the smallest, median and maximum values of the recorded series of rainfall (seasonal or

Research Article

annual) respectively. The rainfall estimate (R_T) for a given return period can be obtained using back transformation $R_T = R_M + (R_M - R_S)[(Y_T/Y_M) - 1]$.

Goodness-of-Fit Test
The KS test statistics is defined by:

$$K = \underset{i=1}{\overset{N}{Max}}(F_e(R_i) - F_D(R_i)) \qquad \dots (5)$$

where $F_e(R_i) = (i - 0.35/N)$ is the empirical cumulative distribution function (CDF) of R_i, $F_D(R_i)$ is the computed CDF of R_i and N is the number of observations. The rejection region of KS test at the desired significance level 'η' is $K_C > K_{N,1-\eta}$. If the computed value (K_C) of KS test of the transformation is less than that of critical value at the desired significance level 'η' then the selected transformation is accepted to be adequate than any other transformations (Zhang, 2002).

Model Performance Indicators
RMSE and CC are used for selection of suitable transformation for modelling seasonal and annual rainfall of WCDS (Singh, 1987). Theoretical descriptions of RMSE and CC are given by:

$$RMSE = SQRT\left(\frac{1}{N}\sum_{i=1}^{N}\left(R_i - R_i^*\right)^2\right) \qquad \dots (6)$$

$$CC = \left(\frac{\sum_{i=1}^{N}\left(R_i - \overline{R}\right)\left(R_i^* - \overline{R^*}\right)}{\sqrt{\left(\sum_{i=1}^{N}\left(R_i - \overline{R}\right)^2\right)\left(\sum_{i=1}^{N}\left(R_i^* - \overline{R^*}\right)^2\right)}}\right) \qquad \dots (7)$$

where, R_i is the recorded rainfall (seasonal or annual), R_i^* is the estimated rainfall (seasonal or annual) by transformation, $\overline{R}$ is the mean value of R_i and $\overline{R^*}$ is the mean value of R_i^*.

RESULTS AND DISCUSSION
Study Area and Data Used
In the present study, an attempt has been made to estimate the seasonal and annual rainfall for WCDS using five transformations. WCDS comprises the river Narmada and Tapi, both of which flow in western direction of the country and drain the narrow belt of peninsular India. Further, west of the Western Ghats in the north of WCD system forms Narmada and Tapi river basins and the drainage of Gujarat. WCDS consists of twenty-four small rivers that are originating in the Sahayadri range and flows into the Arabian sea. The total drainage area of WCDS is 1,17,962 km^2. Figure 1 shows the location map of the study area. Seasonal and annual data in respect of WCDS related to the period 1901-2005 are used (IITM, 2007). Table 1 gives the summary statistics of recorded seasonal and annual rainfall of WCDS.

Estimation of Seasonal and Annual Rainfall
By applying the above-described procedures, the parameters of the transformed data of monsoon, post-monsoon, winter, summer and annual rainfall series were computed by five transformations and further used to estimate rainfall for different return periods from 2-year (yr) to 1000-yr. For BCT, transformation constant (λ) of 0.554 was used for transform the seasonal and annual rainfall data. For MSX, the transformation angle (A) of 0.662 was used to transform the data under study. Table 2 gives the parameters of the transformed data using SRT, CRT, LGT, BCT and MSX for seasonal and annual rainfall data series of WCDS. Tables 3-7 give the seasonal and annual rainfall estimates for different return periods given by five transformations for WCDS.

Research Article
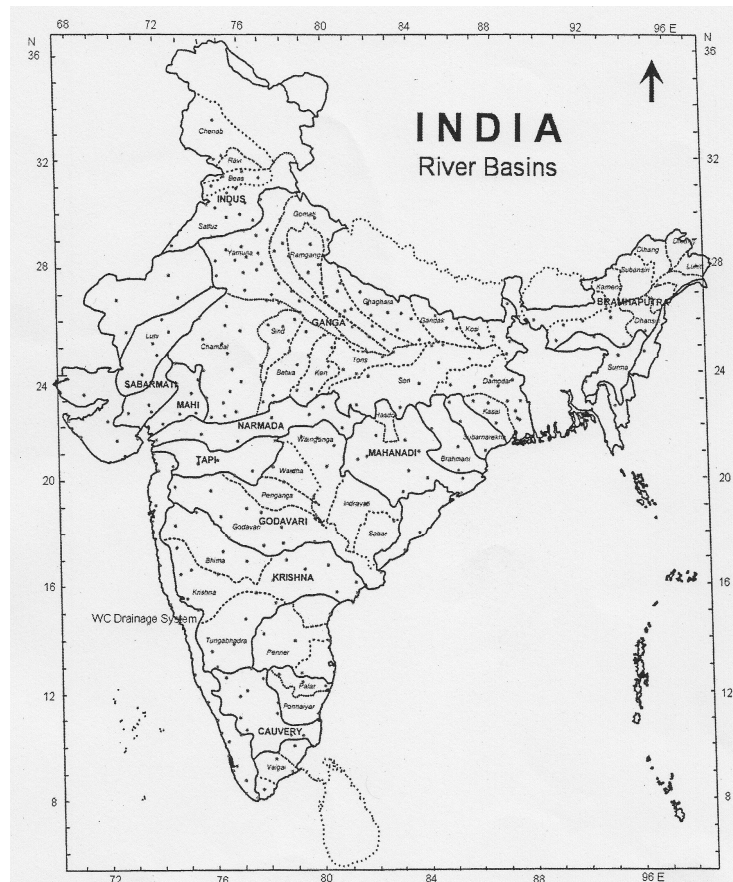
Figure 1: Location map of the study area

Table 1: Summary statistics of seasonal and annual rainfall of WCDS

Season/ Annual	Period	Statistical parameters (units in mm)	
		Mean	Std. Deviation
Monsoon	June-September	1,815.7	224.7
Post-monsoon	October-December	302.6	37.4
Winter	January-February	16.4	2.0
Summer	March-May	211.2	26.2
Annual	January-December	2,345.9	290.3

83

Research Article

Table 2: Parameters of the transformed data using SRT, CRT, LGT, BCT and MSX

Season/ Annual	SRT		CRT		LGT		BCT		MSX		
	A	B	A	B	A	B	$\overline{Y}$	σ_Y	R_S	R_M	R_L
Monsoon	42.5	12.2	3.256	3.256	3.256	0.054	113.3	7.9	1,281.5	1,805.2	2,398.4
Post-monsoon	17.4	6.7	2.478	2.478	2.478	0.054	40.9	2.9	213.6	300.9	399.7
Winter	4.0	2.5	1.212	1.212	1.212	0.054	6.7	0.6	11.6	16.3	21.7
Summer	14.4	5.9	2.321	2.321	2.321	0.054	33.1	2.4	149.0	209.9	278.9
Annual	48.3	13.3	3.367	3.367	3.367	0.054	130.9	9.1	1,655.7	2,332.3	3,098.7

Table 3: Rainfall estimates using five transformations for monsoon season

Return Period (yr)	Estimated rainfall (mm) using				
	SRT	CRT	LGT	BCT	MSX
2	1,806	1,816	1,802	1,810	1,805
5	1,997	2,010	2,002	2,003	1,982
10	2,101	2,117	2,115	2,108	2,075
25	2,214	2,236	2,242	2,222	2,173
50	2,289	2,314	2,329	2,297	2,237
100	2,357	2,386	2,409	2,365	2,294
200	2,420	2,454	2,486	2,429	2,347
500	2,498	2,537	2,581	2,507	2,410
1000	2,554	2,597	2,650	2,562	2,455

Table 4: Rainfall estimates using five transformations for post-monsoon season

Return Period (yr)	Estimated rainfall (mm) using				
	SRT	CRT	LGT	BCT	MSX
2	303	301	300	302	301
5	336	336	334	334	330
10	354	356	352	351	346
25	373	377	374	370	362
50	386	392	388	383	373
100	398	405	402	394	382
200	409	417	414	405	391
500	423	433	430	418	402
1000	433	444	442	427	409

Research Article

Table 5: Rainfall estimates using five transformations for winter season

Return Period (yr)	Estimated rainfall (mm) using				
	SRT	CRT	LGT	BCT	MSX
2	16	16	16	16	16
5	17	17	18	18	18
10	18	18	19	19	19
25	19	19	20	20	20
50	19	20	21	21	20
100	20	20	22	21	21
200	20	21	22	22	21
500	21	22	23	23	22
1000	21	22	24	23	22

Table 6: Rainfall estimates using five transformations for summer season

Return Period (yr)	Estimated rainfall (mm) using				
	SRT	CRT	LGT	BCT	MSX
2	207	205	210	210	210
5	240	266	233	233	230
10	258	302	246	245	241
25	278	343	261	258	253
50	291	372	271	267	260
100	304	400	280	275	267
200	315	426	289	282	273
500	329	459	300	291	280
1000	339	483	308	298	285

Table 7: Annual rainfall estimates using five transformations

Return Period (yr)	Estimated rainfall (mm) using				
	SRT	CRT	LGT	BCT	MSX
2	2,333	2,353	2,328	2,338	2,332
5	2,583	2,583	2,586	2,588	2,561
10	2,719	2,709	2,732	2,723	2,680
25	2,868	2,848	2,897	2,870	2,808
50	2,966	2,941	3,009	2,967	2,890
100	3,056	3,025	3,113	3,056	2,964
200	3,139	3,104	3,211	3,138	3,032
500	3,241	3,202	3,335	3,239	3,114
1000	3,314	3,271	3,424	3,310	3,172

From Tables 3-7, it may be noted that: (i) The rainfall estimates obtained using LGT for return periods above 10-yr are comparatively higher than the values given by other four transformations for monsoon and annual rainfall patterns of WCDS; (ii) The variation between the rainfall estimates given by all five transformations for winter season are very meager; (iii) The rainfall estimates obtained using CRT for return periods above 5-yr are comparatively higher than the corresponding values given by other four transformations for post-monsoon and summer rainfall patterns of WCDS; and (iv) The estimated rainfall by MSX is consistently less than the corresponding rainfall values given by other four transformations for seasonal and rainfall patterns of WCDS.

Research Article

Analysis Based on GoF Test

KS test statistics of five transformations for the data under study was computed by using Eq. (5) and given in Table 8.

Table 8: KS test statistic values for five transformations

Season/ Annual	K_C values for				
	SRT	CRT	LGT	BCT	MSX
Monsoon	0.032	0.034	0.030	0.031	0.051
Post-monsoon	0.031	0.033	0.032	0.032	0.054
Winter	0.038	0.037	0.038	0.039	0.036
Summer	0.126	0.128	0.109	0.085	0.098
Annual	0.032	0.034	0.029	0.031	0.017

From Table 8, it is noted that the computed values of KS test statistic for five transformations are less than of its critical value ($K_{105,0.05} = 0.133$) at five percent level. At this level, all five transformations considered in the study are acceptable for modelling seasonal and annual rainfall of WCDS.

Analysis Based on MPIs

RMSE and CC was used for selection of most suitable transformation for modelling seasonal and annual rainfall though KS test gave sufficient information on fitting of transformation to the data under study. Indices of RMSE and CC for five transformations were computed from Eqs. (6) and (7), and given in Table 9.

Table 9: Indices of RMSE and CC for seasonal and annual rainfall using five transformations

Season/ Annual	RMSE (mm)					CC				
	SRT	CRT	LGT	BCT	MSX	SRT	CRT	LGT	BCT	MSX
Monsoon	8.8	8.6	8.8	8.3	18.2	0.998	0.985	0.998	0.996	0.913
Post-monsoon	1.4	1.6	1.8	1.2	3.0	0.998	0.986	0.996	0.995	0.904
Winter	3.6	3.5	3.3	3.0	5.1	0.994	0.989	0.986	0.989	0.912
Summer	0.4	0.1	0.6	0.8	2.2	0.998	0.995	0.996	0.998	0.906
Annual	14.7	14.5	13.5	11.8	26.4	0.997	0.988	0.998	0.994	0.909

From CC values, it is noticed that there is generally a good correlation between the recorded and estimated rainfall using five transformations for seasonal and annual rainfall of WCDS; with the values varying from 0.904 to 0.998. From Table 9, it is noted that RMSE of BCT are minimum for monsoon, post-monsoon, winter and annual rainfall patterns of WCDS when compared with the corresponding indices of other four transformations. Also, from Table 9, it is noted that the value of RMSE for CRT is minimum when compared with other transformations for summer rainfall pattern of WCDS. Based on GoF test and MPIs, it is suggested that BCT could be employed for modelling rainfall for monsoon, post-monsoon, winter and annual rainfall patterns while CRT for summer rainfall pattern of WCDS.

Conclusions

The paper described a study on modelling seasonal and annual rainfall using five transformation methods and its applicability to WCDS of India. The results showed that the BCT is the best suited among five transformations for modelling monsoon, post-monsoon, winter and annual rainfall patterns while CRT for summer rainfall pattern of WCDS. The study compared the superiority of BCT and CRT over other transformations for WCDS through GoF test and MPIs. The results indicated that the variation on estimated rainfall given by the selected transformation method with reference to recorded rainfall is about 4 percent for seasonal rainfall and about10 percent for annual rainfall. The paper gives the results of the study reported herein and the methodology is expected to be of assistance to stakeholders for planning of water resources management projects at WCDS.

Research Article

REFERENCES
Bhakar S R, Bansal A K, Chhajed N and Purohit R C (2006). Frequency analysis of consecutive days maximum rainfall at Banswara, Rajasthan, Journal of Engineering and Applied Sciences, **1** (3), 64-67.

Hameed R Rasheed, Marur V, Ramamoorthy A S and Al-Dabbagh (1982). Modified SMEMAX transformation for frequency analysis, American Water Resources Association, **18** (3), 509 – 512.

IITM Research Report (2007). Characteristics of hydrological wet season over different river basins of India. Report No. 119

Islam A and Kumar A (2003). HYDRO: A program for frequency analysis of rainfall data, IIE Journal of Agricultural Engineering, **84**, 1-5.

Jam D and Singh V P (1986). A comparison of transformation methods for rainfall frequency analysis, American Water Resources Association, **22** (6), 903 – 910.

Kumar A and Devi R (1982). Statistical determination of design low flows – A comment, Journal of Hydrology, **58**, 175-177.

May W (2004). Variability and extremes of daily flood during the Indian summer monsoon in the period 1901-1989, Global & Planetary Change, **44**, 83-105.

Mohammad M Z (2011). The use of box-cox transformation technique in economic and statistical analyses, Journal of Emerging Trends in Economics and Management Sciences. **2** (1), 32-39.

Panigrahi B, Panda N S and Mull R (2002). Prediction of hydrological events for planning rainfed rice, Journal of Hydrological Sciences, **47** (3), 435-448.

Prakash A (1981). Statistical determination of design low flows, Journal of Hydrology, **51**, 109-118.

Rasheed H R, Marur V. Ramamoorthy and Al-Dabbagh S (1982). Modified SMEMAX transformation for frequency analysis, Water Resources Bulletin, **18** (3), 509-511.

Sharda V N and Das P K (2005). Modelling weekly flood data for crop planning in a sub-humid climate of India, Journal of Agricultural and Water Management, **76**, 120-138.

Singh V P (1987). On statistical intercomparison of EVI estimators by monte carlo simulation. Advanced Water Resources Publications, **10**, 87-105.

Vivekanandan N and Mathew F T (2009). Effect of data length on water resources assessment, Special issue of ISH Journal of Hydraulic Engineering, **15** (3), 28-40.

Zhang J (2002). Powerful goodness-of-fit tests based on the likelihood ratio. Journal of Royal Statistical Society Series B, **64** (2), 281-294.