Mausam, (1980), 31,2,261-266

551.577 : 551.501.

On normalisation of rainfall series

D.S. UPADHYAY, S. N. KATHURIA and S. PRASAD

Meteorological Office, New Delhi

(*Received* 9 *April* 1979)

ABSTRACT. An attempt has been made to normalise the annual rainfall series at eleven observatories in Bihar by applying λ transformation given as $y=(x^{\lambda}-1)/\lambda$. Variation of skewness and kurtosis with λ has been studied and isopleths for skewness and kurtosis have benn drawn. The areas of normal distribution have been separated out which reveal that in most of the cases, the rainfall series of the sample under study may be normalised by taking λ equal to 1.5 and 1.6

1. Introduction

The utility of normal distribution for bringing out various features of meteorological series has usually been over-emphasised. As experienced by the authors, the distribution of short and long duration rainfall series is very rarely normal by itself. Some transformations such as $y = \log(x+c)$ (Bakthavathsalu and Balasubramaniyan 1953), and $y=\sqrt{x}$ have been attempted for normalising rainfall series. In certain circumstances, these transformations might have given desired results but there is no sound base or theory which has been provided in their support and as such these remain more or less as random trials. Some authors also feel that the point rainfall data recorded over a long period of time are expected to follow gamma-distribution, with probability density function

$$f(x) = \frac{a^s}{T_s} x^{s-1} e^{-ax}, \; x \geqslant 0 \text{ and } s > 0 \qquad (1)$$

or a Pearson type X distribution (Upadhyay and Mishra 1977).

It is needless to say that transformation capable to normalise meteorological parameters in majority of the situations will be of immense help for analysis and computation of return periods. In the present paper, an attempt has been made for selecting an appropriate transformation for normalising annual rainfall series in eleven cases of large samples.

2. Data

The point annual rainfall recorded at eleven observatories situated in plains and plateau of Bihar State (Fig. 1) for sixty years (1901-1960) have been considered for the present transformation analysis. Since these stations have been recording continuous observations at the same point, each rainfall series may be assumed to be homogeneous. A few scanty missing observations, however, were linearly interpolated. The computations show that each series is significantly skewed and platykurtic with kurtosis ranging from 3.5 to 5.2.

3. Lamba Transformation

Let x_t represent a point rainfall time series for $t = 1,2,\ldots\ldots n$. Consider a transformed time series y_t given by

$$y_t = \frac{x_t^{\lambda}-1}{\lambda} \qquad (2)$$

The parameter λ may be so chosen that the skewness of y_t tends to zero. If we do not consider the kurtosis aspect of the distribution, y_t follows the normal distribution for all practical purposes. But, in cases where it is not appropriate to avoid kurtosis, the distribution of y_t will be a symmetrical one for which the tables showing the probabilities under different ordinates have already been computed by earlier authors (Box and Tiao 1973). It can be seen that the square root transformation appears as a special case of λ—transformation, where λ is taken to be equal to 0.5. Also from Eqn. (2) we have

$$\lambda \log x_t = \log(1+\lambda y_t) \approx \lambda y_t, \text{ where } \lambda y_t << 1 \qquad (3)$$

$$\text{Hence } y_t = \log x_t, \qquad (4)$$

Thus under the conditions given by Eqn. (3), this transformation reduces to the log-transformation.

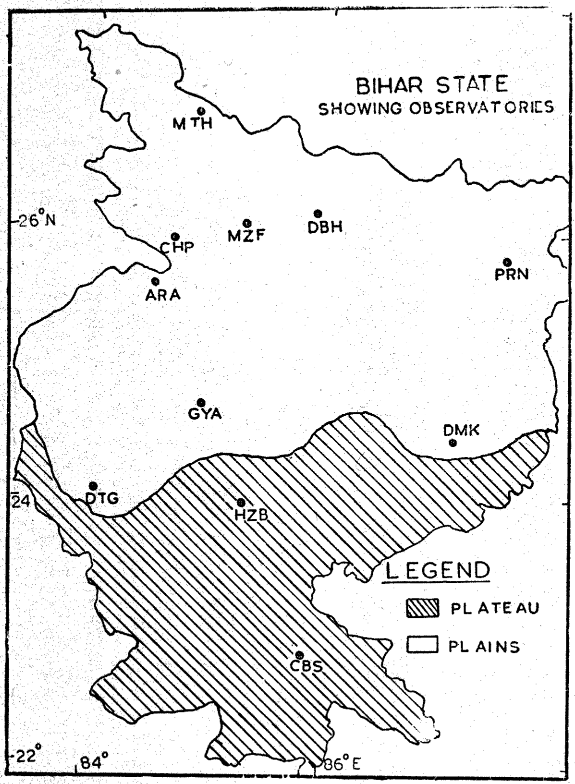

Fig. 1. Map Showing Observatiories under consideration in Bihar State

4. Computations and Results

For each of the eleven transformed series, skewness and kurtosis were computed for values of λ varying from 0.1 to 3.0 at an interval of 0.1. The results have been presented in Tables 1 and 2. The isopleths of skewness and kurtosis have been drawn on the plane showing observatories and the transformation parameter λ. The confidence bands which would normalise the distribution of y . were separated out and are marked as N (Figs. 2 and 3) on the analysed charts. These confidence bands were computed by considering the standard error of skewness, where skewness is given by :

$$r_1 = \sqrt{\mu^2_3/\mu^3_2} \qquad (5)$$

This standard error on the presumption of normality of population is given by $6/n$, where n is the number of observations. The standard error of excess of kurtosis (β_2—3.0) is $24/n$, where,

$$\beta_2 = \mu_4/\mu_2^2 \qquad (6)$$

One of the striking features of this analysis is that in the range $\lambda = 1.5$ to 1.7, seven observatories fall under the area of normal distribution in respect of both skewness and kurtosis whereas for the remaining four observatories this is true for the range from $\lambda = 1.8$ to 2.0. The remarkable point to note is that these samples automatically become mesokurtic in the process of minimising skewness. The isopleth analysis also suggests that the area of normal kurtosis is much larger than that of normal skewness. But the variation of skewness and kurtosis appear to follow the similar trend of variation as depicted in the Figs. 4, 5 and 6. It is evident from the figures that both the parameters fall rather sharply as λ increases upto a optimum value λ_o. As λ increases above λ_o the skewness and kurtosis increases steadily though the gradient is much less than their rate of fall in the range $\lambda < \lambda_o$. However, for the observatory at Muzzafarpur, the increase in kurtosis for $\lambda > \lambda_o$ is found to be almost insignificant.

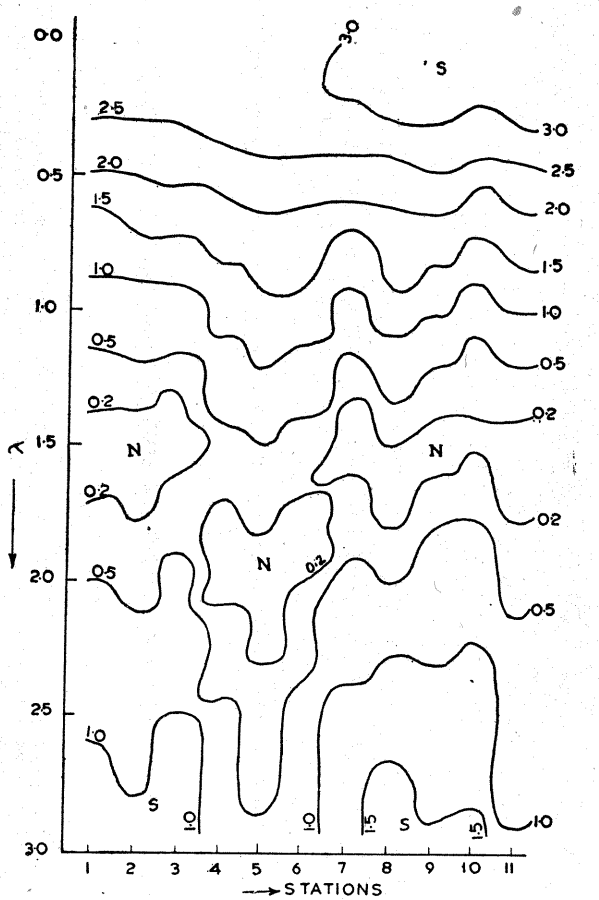

Fig. 2. Isopleths of skewness on the plane showing observatories and transformation parameter λ (N-Normal, S-Strongly skewed). Stations are : (1) Gaya, (2) Arrah, (3) Chapra, (4) Motihari, (5) Muzzafarpur, (6) Darbhanga, (7) Purnea, (8) Dumka, (9) Hazaribagh, (10) Daltonganj, (11) Chaibasa

TABLE 1

Variation of skewness with λ for various stations in Bihar

λ	Gaya	Arrah	Chapra	Moti-hari	Muzza-farpur	Darbh-anga	Purnea	Dumka	Hazari-bagh	Dalton-ganj	Chaibasa
0.1	2.96	2.97	2.97	2.98	2.98	2.98	3.41	3.42	3.43	3.41	3.42
0.2	2.80	2.85	2.84	2,87	2.89	2.89	3.24	3.28	3.31	3.24	3.27
0.3	2.57	2.67	2.65	2.72	2.75	2.75	2.98	3.08	3.11	2.97	3.05
0.4	2.30	2.47	2.41	2.53	2.58	2.58	2.67	2.83	2.86	2.66	2.77
0.5	2.02	2.19	2.15	2.32	2.39	2.38	2.34	2.56	2.58	2.31	2.47
0.6	1.75	1.93	1.88	2.10	2.18	2.17	2.01	2.29	2.27	1.96	2.16
0.7	1.49	1.67	1.61	1.80	1.97	1.95	1.70	1.03	1.96	1.63	1.86
0.8	1.25	1.42	1.35	1.67	1.76	1.73	1.41	1.77	1.66	1.31	1.57
0.9	1.04	1.19	1.11	1.46	1.56	1.52	1.14	1.53	1.27	1.04	1.31
1.0	0.84	0.98	0.89	1.27	1.37	1.32	0.91	1.30	1.10	0.79	1.06
1.1	0.67	0.79	0.69	1.09	1.20	1.13	0.69	1.08	0.85	0.56	0.85
1.2	0.50	0.61	0.50	0.92	1.03	0.94	0.50	0.88	0.62	0.36	0.65
1.3	0.36	0.45	0.33	0.77	0.88	0.78	0.33	0.69	0.42	0.18	0.48
1.4	0.22	0.30	0.18	0.62	0.73	0.62	0.17	0.50	0.22	0.02	0.32
1.5	0.10	0.16	0.04	0.49	0.60	0.47	0.03	0.33	0.05	0.13	0.18
1.6	0.02	0.04	0.10	0.36	0.48	0.33	0.10	0.16	0.11	0.26	0.05
1.7	0.13	0.07	0.21	0.24	0.37	0.20	0.22	0.00	0.25	0.39	0.07
1.8	0.23	0.18	0.33	0.13	0.27	0.08	0.34	0.16	0.39	0.50	0.17
1.9	0.33	0.28	0.43	0.03	0.17	0.03	0.45	0.32	0.51	0.61	0.27
2.0	0.43	0.37	0.53	0.07	0.08	0.13	0.55	0.47	0.63	0.71	0.36
2.1	0.52	0.46	0.62	0.16	0.00	0.23	0.65	0.62	0.74	0.81	0.45
2.2	0.61	0.54	0.71	0.25	0.07	0.32	0.75	0.76	0.84	0.91	0.52
2.3	0.70	0.62	0.79	0.34	0.15	0·41	0.84	0.91	0.94	01.00	0.60
2.4	0.78	0.69	0.87	0·42	0.21	0.49	0.93	1.05	1.03	1·08	0.67
2.5	0.87	0.77	0.95	0.50	0.27	0.57	1.02	1.20	1.12	1.17	0.73
2.6	0.95	0.83	1.02	0.58	0.33	0.65	1.10	1.34	1.21	1.26	0.80
2.7	1.03	0.90	1.09	0.66	0.39	0.72	1.19	1.48	1.29	1.34	0.86
2.8	1.11	0.96	1.16	0.73	0.45	0.79	1.28	1.62	1.37	1.42	0.91
2.9	1.19	1.02	1.23	0.80	0.50	0.86	1.36	1.76	1.45	1.50	0.97
3.0	1.27	1.08	1.29	0.87	0.54	0.92	1.44	1.90	1.53	1.59	1.02

It can be seen from Tables 1 and 2 that the optimum value of λ for Muzzaffarpur is significantly different from those of Gaya, Chapra and Arrah although these places are more or less homogeneous in respect of physical and climatoloical features. This difference may be attributed to the large variations in skewness and kurtosis in the actual precipitation observation recorded at these stations.

A frequency distribution of λ which reduces skewness of y_t between the normal range $\pm 11.76/n$ has also been studied and a histogram for this is given in Fig. 7. The distribution is bimodal. The value of λ equal to 1.5 and 1.6 can be accepted as most appropriate normal transformations for the samples under study. It may, however, be mentioned that no claim is made for unique determination of λ which could normalise all rainfall series for all the points in a given area. The aerial variation of λ is required to be studied in more detail with large sample size before an accepted trend could be established.

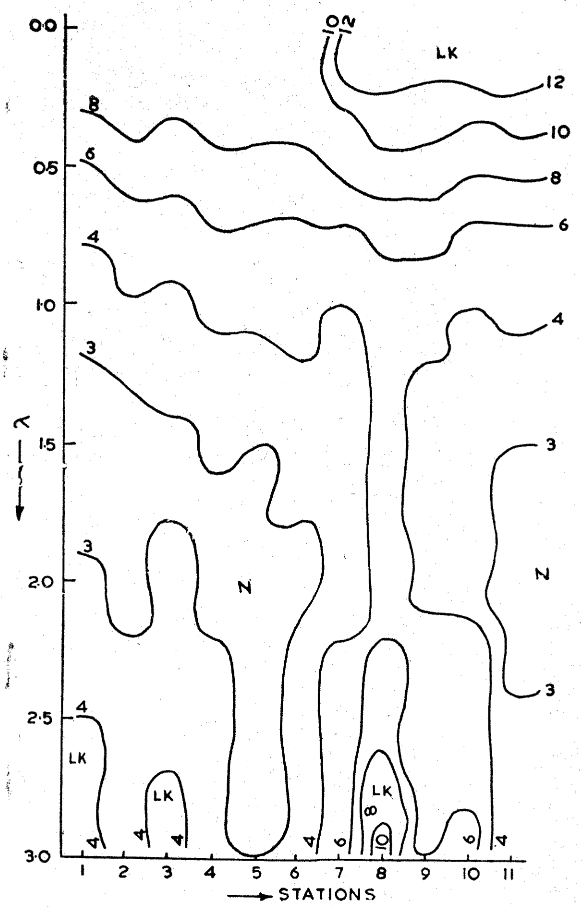

Fig. 3. Isopleths of kurtosis on the plane showing observatories and transformation parameter λ (N-Normal, LK-Lepto Kurtic). Stations are : (1) Gaya, (2) Arrah, (3) Chapra, (4) Motihari, (5) Muzzafarpur, (6) Darbhanga, (7) Purnea, (8) Dumka, (9) Hazaribagh, (10) Daltonganj, (11) Chaibasa

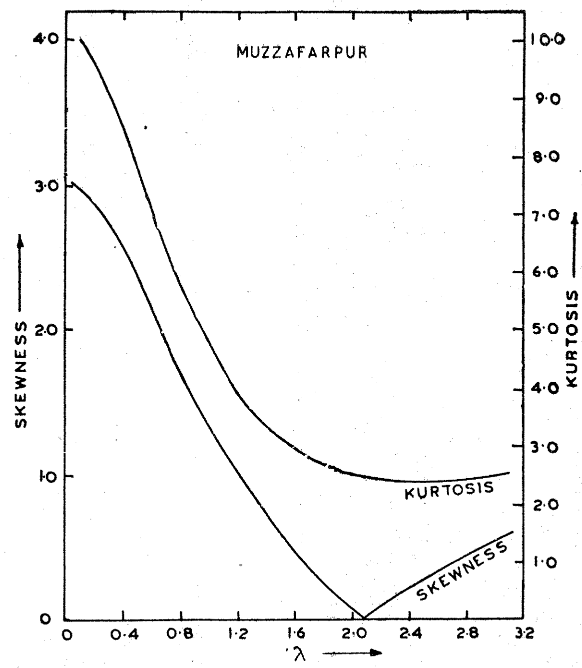

Fig 5. Variation of skewness and kurtosis with trans formation parameter λ for Muzzafarpur Observatory

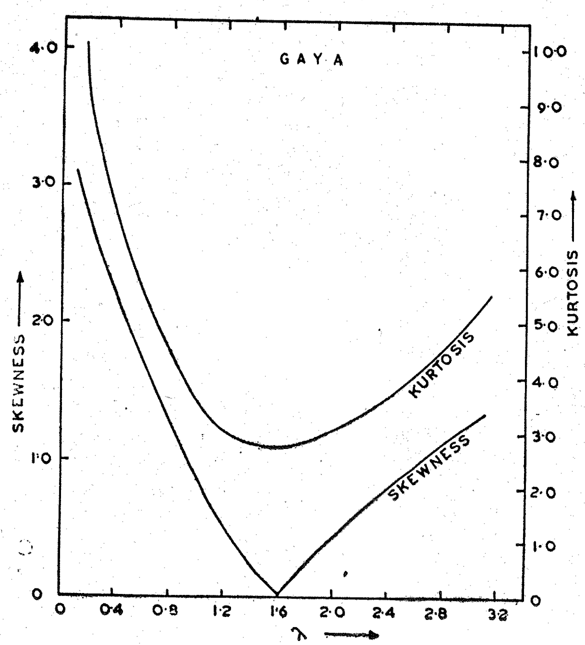

Fig. 4. Variation of skewness and kurtosis with transformation parameter λ for Gaya Observatory

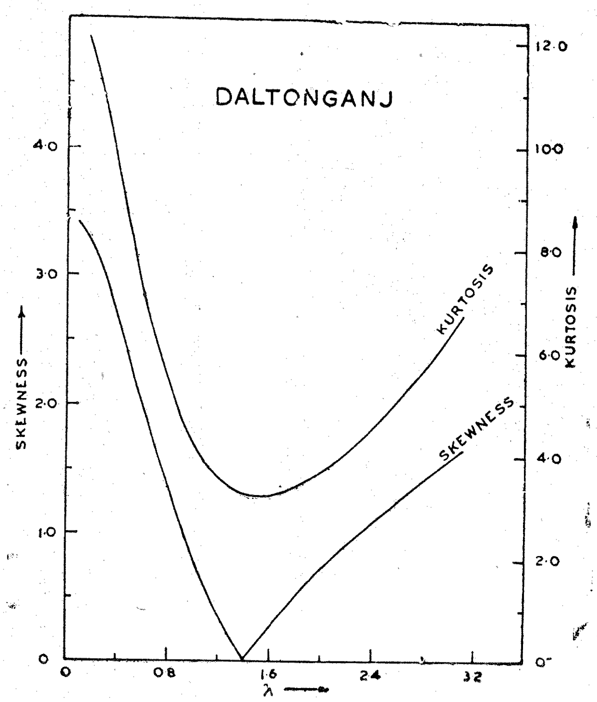

Fig. 6. Variation of skewness and kurtosis with transformation parameter λ for Daltonganj observatory

TABLE 2

Variation of kurtosis with λ for various stations in Bihar

λ	Gaya	Arrah	Chapra	Moti-hari	Muzza-farpur	Darbh-anga	Purnea	Dumka	Hazari-bagh	Daltan-ganj	Chaibasa
0.1	9.87	9.93	9.93	9.95	9.97	9.97	12.8	12.84	12.89	12.81	12.85
0.2	9.28	9.50	9.46	9.57	9.64	9.64	12.05	12.22	12.37	12.07	12.22
0.3	8.44	8.84	8.78	9.00	9.13	9.13	10.97	11.33	11.57	11.00	11.28
0.4	7.50	8.06	7.96	8.32	8.51	8.51	9.73	10.28	10.59	9.75	10.17
0.5	6.57	7.23	7.11	.58	7.83	7.83	8.47	9.20	9.52	8.47	9.00
0.6	5.72	6.42	6.29	6.84	7.13	7.14	7.30	8.18	8.46	7.28	7.87
0.7	4.99	5.67	5.54	6.15	6.46	6.47	6.29	7.25	7.46	6.25	6.85
0.8	4.38	5.01	4.89	5.52	5.83	5.85	5.44	6.45	6.56	5.39	5.95
0.9	3.89	4.45	4.36	4.97	5.26	5.29	4.76	5.78	5.80	4.70	5.20
1.0	3.51	3.99	3.92	4.49	4.75	4.80	4.23	5.24	5.17	4.18	4.59
1.1	3.23	3.63	3.59	4.09	4.31	4.38	3.83	4.83	4.66	3.79	4.10
1.2	3.02	3.34	3.34	3.76	3.93	4.02	3.54	4.52	4.28	3.52	3.71
1.3	2.88	3.12	3.16	3.50	3.61	3.73	3.34	4.32	3.99	3.34	3.42
1.4	2.80	2.96	3.04	3.30	3.35	3.50	3.22	4.21	3.80	3.25	3.20
1.5	2.76	2.86	2.97	3.14	3.12	3.32	3.16	4.17	3.68	3.21	3.04
1.6	2.76	2.79	2.94	3.02	2.94	3.16	3.15	4.21	3.62	3.23	2.92
1.7	2.79	2.76	2.94	2.95	2.79	3.08	3.18	4.31	3.62	3.29	2.85
1.8	2.86	2.76	2.97	2.90	2.68	3.01	3.26	4.48	3.66	3.39	2.81
1.9	2.95	2.78	3.03	2.88	2.58	2.97	3.36	4.70	3.73	3.52	2.80
2.0	3.06	2.82	3.11	2.89	2.51	2.96	3.49	4.97	3.84	3.68	2.81
2.1	3.19	2.88	3.20	2.92	2.46	2.97	3.65	5.30	3.97	3.86	2.84
2.2	3.34	2.95	3.30	2.96	2.42	3.00	3.83	5.67	4.12	4.07	2.88
2.3	3.51	3.03	3.42	3.02	2.40	3.04	4.04	6.08	4.29	4.29	2.93
2.4	3.69	3.13	3.55	3.10	2.39	3.10	4.26	6.54	4.47	4.53	2.99
2.5	3.89	3.23	3.68	3.19	2.38	3.17	4.50	7.04	4.67	4.80	3.06
2.6	4.11	3.34	3.83	3.28	2.39	3.25	4.75	7.59	4.88	5.07	3.13
2.7	4.34	3.46	3.97	3.39	2.41	3.34	5.03	8.17	5.11	5.37	3.22
2.8	4.58	3.59	4.13	3.51	2.43	3.44	5.32	8.78	5.34	5.68	3.30
2.9	4.84	3.72	4.29	3.64	2.45	3.55	5.62	9.43	5.53	6.01	3.40
3.0	5.10	3.85	4.45	3.78	2.49	3.67	5.94	10.12	5.83	6.35	3.49

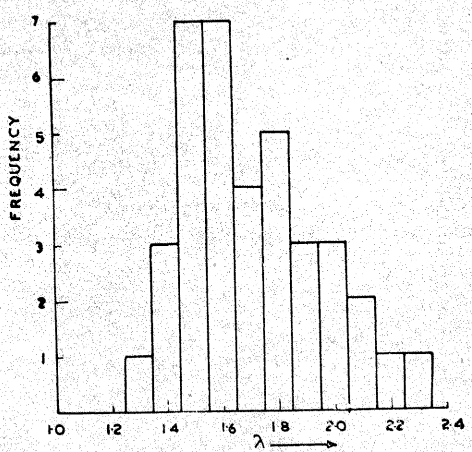

Fig. 7. Frequency distribution of optimum values of transformation parameter λ.

The differences between skewness and kurtosis of original and transformed series were tested for significance. For optimum value of λ such as 1.5, the differences are highly significant. For ten degrees of freedom, the computed values of t are 31.12 and 20.13 in respect of skewness and kurtosis, respectively. The large magnitude of computed t also suggests qualitatively the power of λ- transformation at λ_o.

5. Test of normality

To confirm the inferences drawn in the preceding paragraphs regarding the normality, the transformed series were subjected to rigorous test devised by Fisher (Rao 1952). The test statistics given by :

$$w_1 = r_1 \sqrt{\frac{(n+1)\,(n+2)}{6\,(n-2)}} \tag{7}$$

and

$$w_2 = \left(r_2 + \frac{6}{n+1}\right) \sqrt{\frac{(n+1)^2\,(n+3)\,(n+5)}{24\,n\,(n-2)\,(n-3)}} \tag{8}$$

TABLE 3

Values of test statistics w_1 and w_2 for different λ

λ		Gaya	Arrah	Chapra	Motihari	Muzzafarpur	Darbhanga	Purnea	Dumka	Hazaribagh	Daltonganj	Chaibasa
0.5	w_1	6.76	7.32	7.18	7.75	7.97	7.95	7.82	8.56	8.61	7.72	8.25
	w_2	6.62	7.80	7.59	8.43	8.88	8.88	10.02	11.35	11.92	10.03	10.92
1.0	w_1	2.81	3.27	2.97	4.23	4.58	4.39	3.02	4.32	3.67	2.62	3.55
	w_2	1.10	1.86	1.83	2.85	3.30	3.39	2.27	4.19	4.05	2.28	3.02
1.5	w_1	0.33	0.54	0.12	1.62	2.00	1.56	0.10	1.09	0.18	0.43	0.59
	w_2	0.26	0.11	0.11	0.42	0.40	0.74	0.45	2.27	1.39	0.55	0.28
2.0	w_1	1.42	1.23	1.76	0.23	0.28	0.44	1.83	1.56	2.09	2.37	1.20
	w_2	0.27	0.15	0.36	0.03	0.70	0.11	1.06	3.90	1.67	1.39	0.17

where, $r_1 = \sqrt{\mu_3^2/\mu_2^3}$

and

$$\text{excess of kurtosis } r_2 = \frac{\mu_4}{\mu_2^2} - 3 \quad (9)$$

follow standard normal distribution. For various λ, the values of w_1 and w_2 are summarised in Table 3.

It can be seen that (*i*) the original data (corresponding to $\lambda = 1.0$ are non-normal for all the 11 samples at 5% level, (*ii*) the root transformation ($\lambda < 1.0$) tends to drift the original series further away from normality, (*iii*) for $\lambda = 1.5$ most of the transformed sample follow normal distribution in respect of skewness and kurtosis both, (*iv*) for higher values of λ, the departure of the transformed series from normality exhibit a slow increasing pattern. But for practical purposes all the values of λ, from 1.5 to 2.0 seem to be more or less suitable for normalisation of the most of the samples under consideration.

However, the authors do not wish to claim on generalisation of these values of λ for annual precipitation series recorded for other areas. It only indicates the possibility of transforming precipitation time series of various durations including extreme values to follow on standard Gussain laws for certain values or intervals of the transformations parameter λ.

6. Conclusions

From the foregoing discussions, the following inferences may be drawn :

(*a*) For $\lambda \leqslant 1.0$, the skewness increases as λ decreases. In such cases, the application of $y = \sqrt{x}$ transformation will shift the original series further away from normality.

(*b*) Considering the regularities in values of λ, where most of the samples converge to normality, the transformation seems to be useful for general application. At $\lambda = 1.5$, where the transformed y_t series tends to normality in maximum number of cases, the transformation is equivalent to semicubical parabolic relationship given by :

$$x^3 = (1 + 3/2\, y)^2 \quad (10)$$

(*c*) In the distribution of extreme values, this transformation may also be successfully applied and and the return periods

$$T(y) = \frac{1}{1 - F(y)} \quad (11)$$

can be computed where $F(y)$ will be a normal distribution function of variate y.

Acknowledgements

The authors are highly obliged to Dr. A. S. Ramanathan, Director for his keen interest in the work. Thanks are due to Shri S.D.S. Abbi, Director, Publication for encouragement and to Shri V. K. Bhalla for assistance in computer programming

References

Bakthavathsalu, C. M. and Balasubramaniyan, C., 1953, *Indian J. Met. Goephys.*, 4, p. 249.

Box, G. E. P. and Tiao, G. C., 1973, *Bayesian Inference in Statistical Analysis*, Addison Wesley, Reading, Massachusetts.

Rao, C. R., 1952, *Advanced Statistical Methods in Biometry*, pp. 218-219.

Upadhyay, D. S. and Mishra, D. K., 1977, *Elementry Statistics*.