J. Inst. Eng. India Ser. A (November 2012–January 2013) 93(4):223–232
DOI 10.1007/s40030-013-0032-0

ORIGINAL CONTRIBUTION

Evaluation and Selection of Rain Gauge Network using Entropy

N. Vivekanandan · R. S. Jagtap

Received: 22 February 2012 / Accepted: 1 July 2013 / Published online: 17 July 2013

Abstract Establishment and maintenance of a rain gauge network (RGN) in any geographical region is required for planning, design and management of water resources projects so as to reduce chances of project failure and to minimize the economic risk. Setting up and maintaining a RGN is an evolutionary process, wherein a network is established early in the development of the geographical area; and the network is reviewed and upgraded periodically to arrive at the optimum network. The study shows the computation of transinformation index from marginal and conditional entropy values adopting normal distribution, which is used for evaluation of RGN with 25 rain gauge stations to derive the optimum network. Kolmogorov–Smirnov test is used for checking the adequacy of fitting of normal distribution to the recorded rainfall data. The paper presents that the derived network consisting of nineteen rain gauge stations with network density of 774 km^2 per gauging station is considered as the optimum RGN for upper Bhima basin, which satisfied the WMO recommended value of 600–900 km^2 per station.

Keywords Conditional entropy · Marginal entropy · Rain gauge · Transinformation index

Introduction

Monitoring networks established in different river basins collects data on hydrometeorological parameters such as rainfall, streamflow, water quality, sediment, etc.; which are required for water resources planning, design and management. Efficient information gathering system should assess the variability in hydrologic processes with spatial and temporal scales and reduce the costs of data collection. Evaluation of network involves determining optimal number of monitoring stations and their respective locations by which hydrological information can be maximized [1]. Setting up and maintaining a rain gauge network (RGN) is an evolutionary process, wherein a minimum network is established early in the development of the geographical area; and the network reviewed and upgraded periodically to arrive at an optimum network. For network optimisation, approaches commonly used include statistical approaches, user-survey technique, hybrid method and sampling strategies. Statistical approaches for RGN optimisation range from clustering techniques, spatial regression methods (in generalized least square framework) and entropy-based methods.

Entropy method quantifies the relative information content for the RGN, and has the advantage that it needs only rainfall data for evaluation. The method facilitates network design by quantifying the marginal contribution of each data collection node to the overall information provided by the network using an index termed as marginal entropy. In recent past, entropy method is widely applied by various researchers for evaluation of different hydrometeorological networks such as streamflow, rainfall, groundwater and water quality monitoring etc., to derive optimum networks. The researchers have applied entropy based approach to design stream gauge network for Pembina river basin in southern Manitoba, Canada [2]. They have adopted a directional informational transfer index to identify the most priority station amongst seven stations in the network. Previously, the entropy technique has been applied for the assessment of water quality monitoring networks for

N. Vivekanandan (✉) · R. S. Jagtap
Central Water and Power Research Station, Pune 411024, Maharashtra, India
e-mail: anandaan@rediffmail.com

Mississippi River in Louisiana [3]. They concluded that the concept of entropy is a promising method as it quantitatively measures the information produced by network, which will improve the efficiency and cost-effectiveness of current monitoring programs. The entropy and least square methods can be applied to evaluate the stream gauge network in the State of Illinois through an assessment of the transferring of information (transinformation) among gauging records for low, average, and high flow conditions [4]. They expressed that the hybrid combination of the entropy and GLS measures of regional value of stations would certainly provide more insight in assessment of stream gauging program and the station rankings based on the combined method can preserve correlation with rankings using both entropy and GLS.

The researchers have applied entropy theory to evaluate the stream flow network for Kizihrmak basin [5]. They compared the model uncertainty in stream flow network with transinformation index using normal, log-normal and gamma distributions. They also suggested that the ranking of log-normal distribution could be considered for selection of stream gauge network for Kizihrmak basin. The comparisons were made on the applications of mixed and continuous distributions in entropy theory to study about the rainfall variability and rainfall intermittency using transinformation index [6]. The study showed that the mixed distribution is suitable for evaluation of RGN for Choongju basin, Korea. Moreover, optimisation theory related to hydrological network studies iterated that the rain gauge station provided more than 50 % redundant information is considered to be insignificant and for possible discontinuation from the core network. The investigators have analyzed the application of a novel, process-oriented stream gauging method for generating rating curves to establish a stream gauging network in the Whitewater River basin in south central Kansas [7]. Entropy theory can be applied to the high density RGN of the urban area of Rome for evaluation of maximum information content achievable by a rainfall network for different sampling time intervals [8]. The scientists have applied entropy theory-based maximum information minimum redundancy criterion for evaluation of hydrometric network [9]. In the present study, normal distribution is used for computation of entropy values for the stations under consideration.

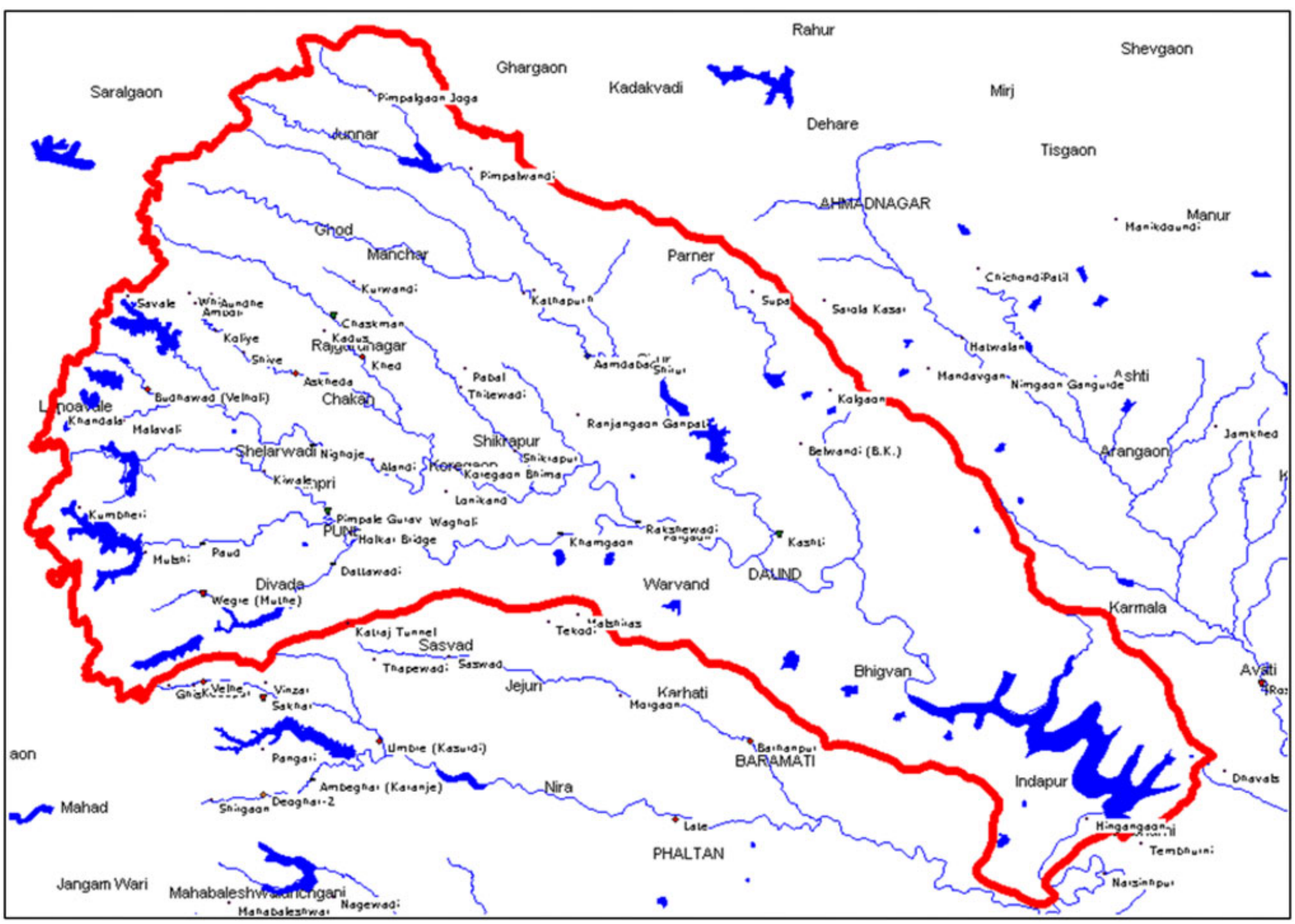


Fig. 1 Location map of the study area

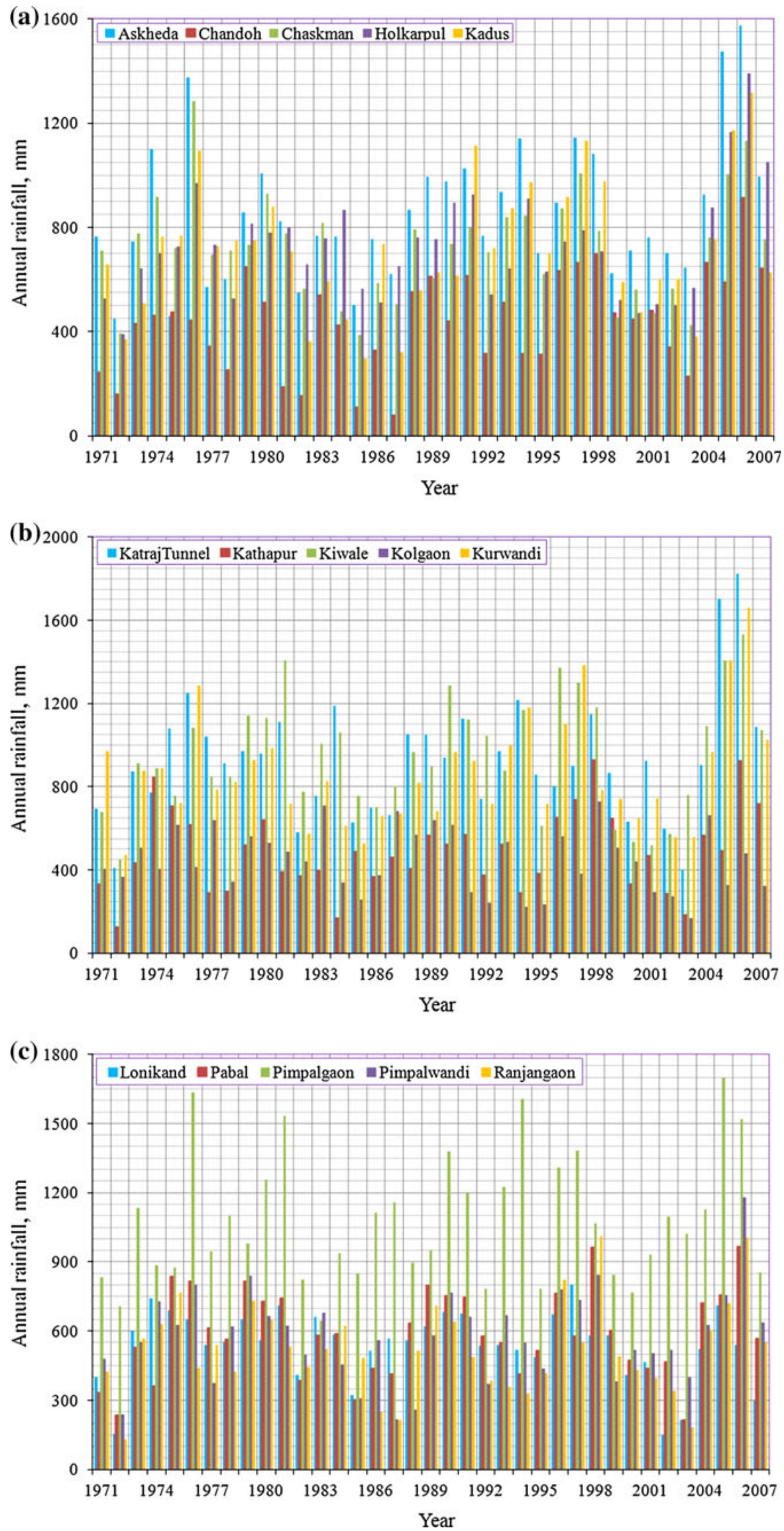
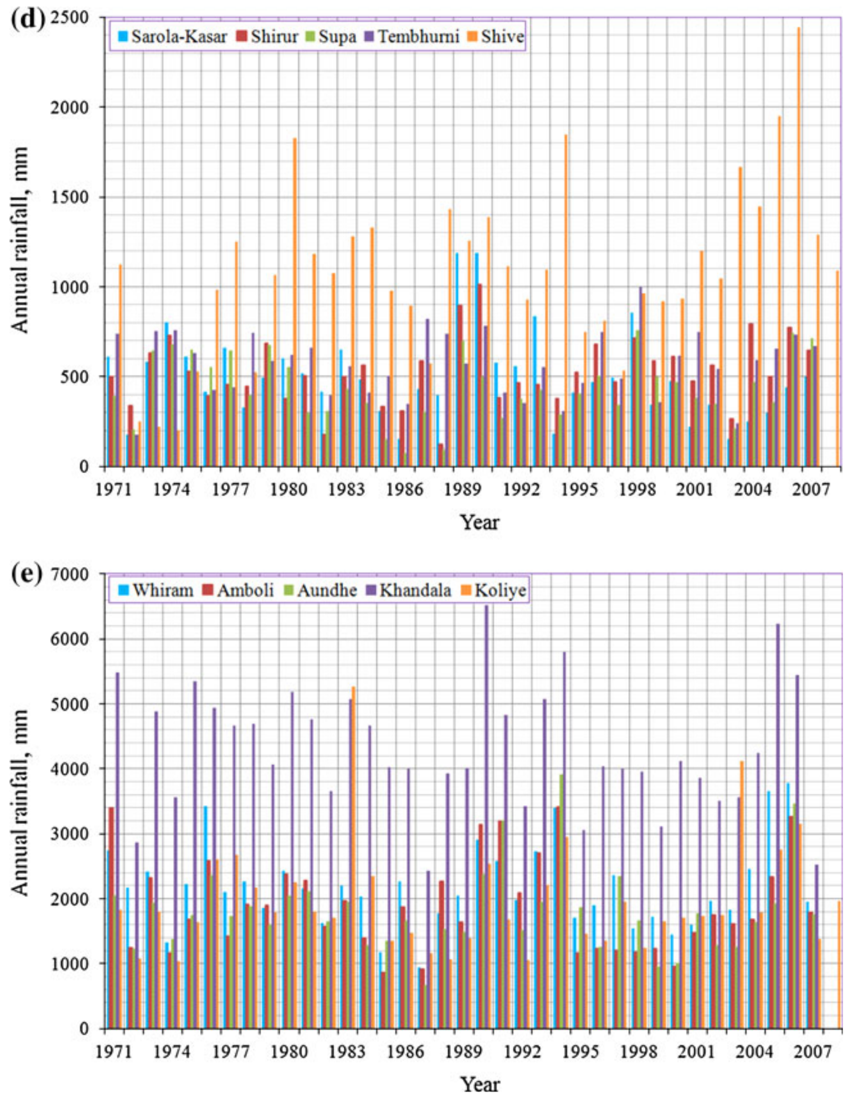


Fig. 2 a–e Plot of recorded annual rainfalls during 1971–2007 for 25 rain gauge stations of Upper Bhima basin

Fig. 2 continued



Transinformation index measures the redundant or mutual information between variables, and is computed from marginal and conditional entropy indices to derive an optimum network. Kolmogorov–Smirnov (KS) test is used for checking the adequacy of fitting of normal distribution to the recorded data. The methodology adopted in computation of KS statistics values and evaluation of RGN using entropy method for upper Bhima basin up to Ujjani reservoir is briefly described in the ensuing sections.

Methodology

Normal Distribution

The probability density function (PDF) and cumulative distribution function (CDF) of two-parameter normal distribution is given by:

$$f(X) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(X-\mu)^2}{\sigma^2}\right), \; -\infty < X < \infty, \; \sigma > 0 \quad (1)$$

$$F(X) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{(X-\mu)}{\sigma\sqrt{2}}\right)\right] \quad (2)$$

where μ and σ the parameters of the distribution [10]. The parameters are determined from the mean and standard deviation of the recorded data. Here, 'erf' indicates the error function of normal distribution.

Kolmogorov–Smirnov Test

The adequacy of fitting of normal distribution to the recorded data is judged by KS test and the theoretical description of KS statistics is expressed by:

Table 1 Statistical parameters of annual rainfall, computed values of KS statistics and marginal entropy values of 25 rain gauge stations given by normal distribution

Station ID	Rain gauge station	Statistical parameters, m^3/s			Computed values of KS statistics	Marginal entropy
		Mean (μ)	SD (σ)	CV, %		
A1	Askheda	855.8	265.6	31.0	0.147	7.001
A2	Aundhe	1,808.4	652.1	36.1	0.164	7.899
C1	Chaskman	722.3	202.9	28.1	0.096	6.732
H1	Holkarpul	729.5	206.2	28.3	0.099	6.748
K1	Katraj-tunnel	928.2	294.9	31.8	0.088	7.105
K2	Khandala	4,312.9	969.7	22.5	0.122	8.296
K3	Kiwale	950.5	277.4	29.2	0.068	7.044
K4	Kolgaon	448.4	153.7	34.3	0.081	6.454
K5	Koliye	1,971.5	868.0	44.0	0.217	8.185
K6	Kurwandi	862.8	265.1	30.7	0.124	6.999
L1	Lonikand	537.5	156.2	29.1	0.173	6.470
P1	Pimpalgaon	1,076.8	277.4	25.8	0.140	7.044
P2	Pimpalwandi	579.5	196.2	33.9	0.067	6.698
R1	Ranjangaon	522.2	199.3	38.2	0.091	6.714
S1	Shirur	523.1	186.6	35.7	0.087	6.648
S2	Shive	1,089.3	482.5	44.3	0.098	7.598
S3	Supa	437.3	183.0	41.9	0.117	6.629
T1	Tembhurni	571.4	181.4	31.7	0.117	6.620
W1	Whiram	2,181.7	651.4	29.9	0.121	7.898
A3	Amboli	1,862.9	783.0	42.0	0.112	8.082
C2	Chandoh	420.6	199.1	47.3	0.090	6.713
K7	Kadus	715.2	250.1	35.0	0.148	6.941
K8	Kathapur	487.0	197.6	40.6	0.097	6.705
P3	Pabal	588.5	191.5	32.5	0.109	6.674
S4	Sarola-Kasar	481.1	257.8	53.6	0.132	6.971

SD standard deviation, *CV* coefficient of variation [$(\sigma/\mu) \times 100$]; $KS_{37,0.05} = 0.218$

$$KS = \underset{i=1}{\overset{N}{\text{Max}}}(F_e(X_i) - F_D(X_i)) \quad (3)$$

Here $F_e(X_i) = (i - 0.44)/(N + 0.12)$ is the empirical CDF of X_i and $F_D(X_i)$ is the computed CDF of X_i. If the computed value of KS statistics given by normal distribution is less than the theoretical value at the desired significance level, then the distribution is considered to be suitable for fitting to the recorded rainfall data [11].

Concept of Entropy

A quantitative measure of the uncertainty associated with a probability distribution, or the information content of the distributions termed as, Shannon entropy, which can be mathematically expressed as:

$$H(X) = -k\sum P_i \ln(P_i) \quad (4)$$

where, H(X) is the entropy corresponding to the random variable X; k is a constant that has value equal to one, when natural logarithm is taken; and P_i represents the probability of i^{th} event of random variable X [12].

Marginal Entropy

Marginal entropy for the discrete random variable X is defined as:

$$H(X) = -k\sum_{i=1}^{N} P(X_i)\ln[P(X_i)] \quad (5)$$

where $P(X_i)$ is the probability of occurrence of X_i, computed by normal distribution, and N is the number of observations [13]. The marginal entropy H(X) indicates the amount of information or uncertainty that X has. If the variables X and Y are considered as independent, then the joint entropy [H(X, Y)] is equal to the sum of their marginal entropies defined by:

$$H(X, Y) = H(X) + H(Y) \quad (6)$$

Table 2 Transinformation index matrix based on marginal entropy given by normal distribution

Station ID	Marginal entropy	Transinformation index							
		S2	S3	S4	S5	S6	S7	S8	S9
A1	7.001	0.103	0.329	0.330	0.466	0.563	0.565	0.566	0.811
A2	7.899	0.256	0.282	0.297	0.328	0.432	0.436	0.450	0.476
C1	6.732	0.153	0.398	0.400	0.400	0.407	0.424	0.458	0.614
H1	6.748	0.123	0.282	0.287	0.505	0.507	0.509	0.526	0.602
K1	7.106	0.179	0.335	0.351	0.468	0.470	0.493	0.500	0.542
K2	8.296	–	–	–	–	–	–	–	–
K3	7.045	0.138	0.286	0.290	0.361	0.361	0.361	0.362	0.556
K4	6.454	0.011	0.087	0.156	0.163	–	–	–	–
K5	8.185	0.143	0.172	0.174	0.258	0.265	–	–	–
K6	6.999	0.154	0.382	0.410	0.453	0.483	0.495	0.497	0.642
L1	6.470	0.105	0.260	0.277	0.349	0.370	0.371	0.375	0.424
P1	7.044	0.181	0.219	0.241	0.260	0.276	0.281	0.283	–
P2	6.698	0.142	0.539	0.539	0.552	0.565	0.624	0.691	0.729
R1	6.714	0.072	0.488	0.502	0.519	0.563	0.581	0.632	0.674
S1	6.648	0.009	0.128	0.264	0.265	0.277	0.278	–	–
S2	7.598	0.116	0.117	0.121	–	–	–	–	–
S3	6.629	0.008	0.257	0.325	0.327	0.346	0.362	0.501	0.533
T1	6.620	0.016	0.176	0.212	0.222	0.270	0.301	0.323	0.323
W1	7.898	0.423	0.427	0.461	0.507	0.544	0.547	0.552	0.653
A3	8.082	0.304	0.309	0.311	0.373	0.392	0.411	0.416	0.416
C2	6.713	0.035	0.211	0.213	0.326	0.354	0.357	0.411	0.467
K7	6.941	0.134	0.380	0.413	0.426	0.482	0.482	0.487	0.609
K8	6.705	0.002	–	–	–	–	–	–	–
P3	6.674	0.089	0.356	0.363	0.385	0.476	0.501	0.517	0.540
S4	6.971	0.020	0.064	–	–	–	–	–	–

Station ID	Marginal entropy	Transinformation index							
		S10	S11	S12	S13	S14	S15	S16	S17
A1	7.001	0.811	0.819	0.850	0.855	0.958	0.969	0.970	1.212
A2	7.899	0.496	0.637	0.642	0.643	0.660	–	–	–
C1	6.732	0.614	0.654	0.682	0.769	0.790	0.799	0.847	1.142
H1	6.748	0.607	0.615	0.679	0.708	0.712	0.748	0.889	0.917
K1	7.106	0.542	0.542	0.647	0.791	0.836	0.883	0.883	0.889
K2	8.296	–	–	–	–	–	–	–	–
K3	7.045	0.557	0.560	0.676	0.676	0.677	0.693	–	–
K4	6.454	–	–	–	–	–	–	–	–
K5	8.185	–	–	–	–	–	–	–	–
K6	6.999	0.642	0.653	0.673	0.691	0.692	0.724	0.737	–
L1	6.470	0.430	0.455	–	–	–	–	–	–
P1	7.044	–	–	–	–	–	–	–	–
P2	6.698	0.730	0.739	0.739	0.768	0.859	0.883	0.936	0.936
R1	6.714	0.684	0.772	0.792	0.918	0.990	1.000	1.315	1.375
S1	6.648	–	–	–	–	–	–	–	–
S2	7.598	–	–	–	–	–	–	–	–
S3	6.629	0.533	0.543	0.552	–	–	–	–	–
T1	6.620	–	–	–	–	–	–	–	–
W1	7.898	0.754	0.938	0.954	0.975	0.990	1.008	1.011	1.415

Table 2 continued

Station ID	Marginal entropy	Transinformation index							
		S10	S11	S12	S13	S14	S15	S16	S17
A3	8.082	0.416	–	–	–	–	–	–	–
C2	6.713	0.467	0.546	0.559	0.605	–	–	–	–
K7	6.941	0.639	0.645	0.682	0.720	0.793	0.886	0.886	0.995
K8	6.705	–	–	–	–	–	–	–	–
P3	6.674	0.554	0.554	0.623	0.703	0.810	0.817	0.857	0.909
S4	6.971	–	–	–	–	–	–	–	–

Station ID	Marginal entropy	Transinformation index							
		S18	S19	S20	S21	S22	S23	S24	S25
A1	7.001	1.235	1.253	1.289	1.302	1.574	1.576	1.581	–
A2	7.899	–	–	–	–	–	–	–	–
C1	6.732	1.148	1.167	1.190	1.210	–	–	–	–
H1	6.748	1.138	1.139	1.177	1.291	1.376	–	–	–
K1	7.106	–	–	–	–	–	–	–	–
K2	8.296	–	–	–	–	–	–	–	–
K3	7.045	–	–	–	–	–	–	–	–
K4	6.454	–	–	–	–	–	–	–	–
K5	8.185	–	–	–	–	–	–	–	–
K6	6.999	–	–	–	–	–	–	–	–
L1	6.470	–	–	–	–	–	–	–	–
P1	7.044	–	–	–	–	–	–	–	–
P2	6.698	0.964	–	–	–	–	–	–	–
R1	6.714	1.404	1.445	1.461	1.464	1.486	1.487	–	–
S1	6.648	–	–	–	–	–	–	–	–
S2	7.598	–	–	–	–	–	–	–	–
S3	6.629	–	–	–	–	–	–	–	–
T1	6.620	–	–	–	–	–	–	–	–
W1	7.898	1.635	1.676	1.724	1.759	1.791	1.799	1.799	1.855
A3	8.082	–	–	–	–	–	–	–	–
C2	6.713	–	–	–	–	–	–	–	–
K7	6.941	0.995	1.049	1.150	–	–	–	–	–
K8	6.705	–	–	–	–	–	–	–	–
P3	6.674	0.964	1.014	–	–	–	–	–	–
S4	6.971	–	–	–	–	–	–	–	–

Si (i = 2, 3, 4, …, 25) indicate the steps involved in computation of transinformation index based on marginal entropy of first priority station, namely Khandala. From the table, the station having the least transinformation index at each step is considered as the next priority station to the first one

If the variables are stochastically dependent, then the joint entropy is less than its total entropy.

Conditional Entropy

Conditional entropy measures the entropy of a random variable Y, if one has already learned completely about the random variable X. The conditional entropy of Y on X is defined by:

$$H(Y/X) = H(X,\ Y) - H(X) \quad (7)$$

Conditional entropy value becomes zero, if the value of one variable is completely determined by the value of other variable. If the variables are independent, then $H(Y/X) = H(Y)$.

Transinformation Index

Transinformation is the form of entropy that measures the redundant or mutual information between variables [14]. Transinformation represents the amount of information, which is common to two stochastically dependent variables X and Y. The transinformation between X and Y is defined as:

$$T(X, Y) = H(X) + H(Y) - H(X, Y) = H(Y) - H\ (Y|X) \quad (8)$$

If the variables X and Y are dependent to each other, then the transinformation index $T(X, Y) = 0$.

Steps Involved in Computation of Entropy Measures

In practice, the existing sampling sites of a RGN can be arranged in the order of information content. In the ordered list thus obtained, the first station is the one where the highest uncertainty about the variable occurs, and the subsequent stations serve to reduce the uncertainty further. The steps involved in selecting the best combination of stations using entropy method are as follows:

(i) Let the data collection network under review, consists of M monitoring stations. The data series of the variable of interest at each station ($X_1, X_2,\ldots, X_M$) is represented by X_{ij}, where 'i' denotes the station identification number (i = 1, 2,…, M) and 'j' is for time period (j = 1, 2,…, N). The data length at all stations is assumed to be equal to N. The best fitted multivariate joint PDF for the subset ($X_1, X_2,\ldots, X_M$) of M monitoring stations is selected.

(ii) The marginal entropy of the variable $H(X_i)$ (i = 1, 2,…, M) for each station is calculated. The station with the highest marginal entropy is denoted as the first priority station $Pr(Xz_1)$. This is the location, where the highest uncertainty occurs about the variable and hence information-gain will be highest from the observations recorded at this site.

(iii) This station $Pr(Xz_1)$ is coupled with every other (M − 1) stations in the network to compute transinformation $T(X_i, Pr(Xz_1))$ with $X_i \neq Pr(Xz_1)$, i = 1, 2,… M; and to select that pair, which gives the least transinformation. The station that fulfils this condition is marked as the second priority location $Pr(Xz_2)$.

(iv) The pair $[Pr(Xz_1), Pr(Xz_2)]$ is coupled with every other (M − 2) station in the network to select a triplet with the least transinformation $T[X_i; Pr(Xz_1), Pr(Xz_2)]$. The same procedure is continued by successively considering combinations of three and more stations, and selecting the combination that produces the least transinformation. Finally, all M monitoring stations ($X_1, X_2,\ldots, X_M$) can be ranked in priority order to get $[Pr(Xz_1), Pr(Xz_2),\ldots, Pr(Xz_M)]$.

(v) It is possible to terminate the above process early, before carrying out for all M stations by selecting a particular threshold transinformation value as the amount of redundant information to be permitted in the network, such that sampling of the variable may be stopped at the stations that exceed the threshold to get optimum number of stations, which is less than M.

In the above procedure, the benefits for each combination of sampling sites are measured in terms of least transinformation or the highest conditional entropy produced by that combination. The above procedure helps to assess network configurations with respect to the existing stations. If new stations are to be added to the system, their locations may be selected again on the basis of the entropy method by ensuring maximum gain of information. The correlation coefficient of each monitoring station can be computed by using Eq. (9) based on transinformation index.

$$\text{Transinformation (T)} = -\frac{1}{2}\ln\left(1 - R^2\right) \quad (9)$$

where R represents the multiple correlation coefficient of X_i on $Pr(Xz_1), Pr(Xz_2),\ldots, Pr(Xz_M)$.

Application

The methodology detailed above has been applied to evaluate the RGN for upper Bhima basin up to Ujjani reservoir. The basin is located in the western part of Maharashtra between 18°03′N–19°24′N latitude and 70°20′E–75°18′E longitude. The geographical area of the basin is 14,712 km^2. Of the total geographical area under study, 25 % is hilly and/or highly dissected, 55 % plateau and 20 % is remaining plain area [15]. Figure 1 gives a location map of the study area. From scrutiny of historical rainfall data, it is noted that 25 rain gauge stations have concurrent data for the period 1971–2007; and considered in deriving the optimum RGN for the basin under study. The daily rainfall series was used to compute the annual rainfall series for each rain gauge station and delineated in Fig. 2a–e.

Results and Discussions

By applying the procedure detailed above, a computer program was developed and used to evaluate the RGN of upper Bhima basin. The program compute the KS statistics values, identify the first priority station based on marginal entropy, compute the conditional entropy with reference to first priority station and arrange the rain gauge stations in order of priority based on transinformation index. Table 1 gives the statistical parameters (μ, σ and CV), KS statistics values and marginal entropy values obtained from normal distribution for 25 rain gauge stations considered in the study.

From Table 1, it may be noted that the Khandala rain gauge station receives maximum average annual rainfall of about 4,313 mm when compared to the corresponding values of other rain gauge stations. Also, from Table 1, it may be noted that the CV values computed from average annual rainfall vary from about 23–54 %.

Based on GoF test results, it may be noticed that the computed values of KS statistics of all 25 rain gauge stations given by normal distribution are lesser than the theoretical value of 0.218 at 5 % significance level, and at this level, the normal distribution is considered to be acceptable for fitting annual rainfall data recorded at the rain gauge stations considered in the study. From Table 1, it transpires that the station is coupled with other 24 rain gauge stations individually to identify the next priority station in order, and to compute the transinformation index. Table 2 gives the transinformation index matrix obtained from normal distribution for the stations under study. Table 3 give the details on redundant information passed by each station based on transinformation index and its corresponding correlation coefficient values given by normal distribution.

From Table 3, it may be noted that Whiram station provides 100 % information, as indicated by the transinformation index value; and considered as the reference station for computation of redundant information for other stations. Also, from Table 3, it may be noted that the amount of redundant information from Kadus, Chaskman, Holkarpul, Ranjangaon, Askheda and Whiram stations are about 62, 65, 74, 80, 85 and 100 % respectively; and are proposed for possible discontinuation from the existing RGN. The study showed that the derived optimum network for the basin consists of 19 rain gauge stations with 774 km^2 per gauging station; which satisfies the WMO [16] recommended value of 600–900 km^2 per station for

Table 3 Details of redundant information passed by each station together with correlation coefficient based on normal distribution

Station ID	Rain gauge station	Transinformation index (T)	Optimum redundant information passed [(T/1.855 × 100]	Correlation coefficient $[1 - \exp(-2T)]^{0.5}$
K8	Kathapur	0.002	0.1	0.063
S4	Sarola-Kasar	0.064	3.5	0.347
S2	Shive	0.121	6.5	0.464
K4	Kolgaon	0.163	8.8	0.527
K5	Koliye	0.265	14.3	0.641
S1	Shirur	0.278	15.0	0.653
P1	Pimpalgaon	0.283	15.3	0.657
T1	Tembhurni	0.323	17.4	0.690
A3	Amboli	0.416	22.4	0.752
L1	Lonikand	0.455	24.5	0.773
S3	Supa	0.552	29.8	0.818
C2	Chandoh	0.605	32.6	0.838
A2	Aundhe	0.660	35.6	0.856
K3	Kiwale	0.693	37.4	0.866
K6	Kurwandi	0.737	39.7	0.878
K1	Katraj-tunnel	0.889	47.9	0.912
P2	Pimpalwandi	0.964	52.0	0.924
P3	Pabal	1.014	54.7	0.932
K7	Kadus	1.150	62.0	0.949
C1	Chaskman	1.210	65.2	0.955
H1	Holkarpul	1.376	74.2	0.968
R1	Ranjangaon	1.487	80.2	0.974
A1	Askheda	1.581	85.2	0.979
W1	Whiram	1.855	100.0	0.988

minimum density of RGN. The results are expected to be of assistance to stakeholders for decision making as regards RGN optimisation in the Upper Bhima basin.

Conclusions

This paper presented an entropy based procedure for evaluation of RGN of upper Bhima basin, consists of 25 rain gauge stations. The normal distribution was used to represent the data and then to compute the marginal and conditional entropy indices, and the transinformation index. The average annual rainfall recorded at the rain gauge stations varied from about 421–4,313 mm. The CV values based on average annual rainfall also varied from 23–54 %. The KS test results supported the use of normal distribution for fitting annual rainfall data recorded at the rain gauge stations considered in the study. The results showed that Khandala station is the first priority station with highest marginal entropy for upper Bhima basin. The results also showed that Whiram station provided 100 % redundant information and hence considered as the reference station for computation of redundant information passed by other stations. From the results of transinformation index given by normal distribution, it may be noted that the amount of redundant information given by Kadus, Chaskman, Holkarpul, Ranjangaon, Askheda and Whiram stations are about 62, 65, 74, 80, 85 and 100 % respectively; and are proposed for possible discontinuation from the core network while optimising RGN of upper Bhima basin up to Ujjani reservoir. Thus, the derived optimum network for Upper Bhima basin consists of 19 rain gauge stations, which satisfy WMO guidelines as regards minimum density for RGN. The results are expected to be of assistance to stakeholders for decision making as regards RGN optimisation in the river basin.

Acknowledgments The authors are grateful to the Central Water and Power Research Station, Pune, for providing the research facilities to carry out the study. The authors are also thankful to the Hydrometeorological Data Processing Division, Nashik, for making available the rainfall data of upper Bhima basin.

References

1. S.P. Richard, Review and Analysis of Stream Gauge Network for the Ontario Stream Gauge Rehabilitation Project. WSC report no. 01-2004, p. 96 (2004)
2. Y. Yang, D.H. Burn, An entropy approach to data collection network design. J. Hydrol. **157**(4), 307–324 (1994)
3. S. Ozkul, N.B. Harmanucioglu, V.P. Singh, Entropy based assessment of water quality monitoring networks. ASCE J. Hydrol. Eng. **5**(1), 90–99 (2000)
4. M. Markus, H.V. Knapp, G.D. Tasker, Entropy and generalized lease square methods in assessment of the regional value of stream gauges. J. Hydrol. **283**(1–4), 107–121 (2003)
5. N. Sarlak, U.A. Sorman, Evaluation and selection of stream flow network stations using entropy methods. Turk. J. Eng. Environ. Sci. **30**(2), 91–100 (2006)
6. C. Yoo, K. Jung, J. Lee, Evaluation of rain gauge network using entropy theory: comparison of mixed and continuous distribution function applications. ASCE J. Hydrol. Eng. **13**(4), 226–235 (2008)
7. A.C. Jordan, W.K. Jason, Establishing a multi-scale stream gauging network in the Whitewater river basin, Kansas, USA. Water Res. Manag. **24**(13), 3641–3664 (2010)
8. E. Ridolfi, V. Montesarchio, F. Russo, F. Napolitano, An entropy approach for evaluating the maximum information content achievable by an urban rainfall network. Nat. Hazards. Earth Syst. Sci. **11**(7), 2075–2083 (2011)
9. L. Chao, V.P. Singh, A.K. Mishra, Entropy theory-based criterion for hydrometric network evaluation and design: maximum information minimum redundancy. Water Res. Res. **48**(5), 1–15 (2012)
10. Y. Mogheir, V.P. Singh, Application of information theory to groundwater quality monitoring networks. Water Res. Manag. **16**(1), 37–49 (2002)
11. J. Zhang, Powerful goodness-of-fit tests based on the likelihood ratio. J. R. Stat. Soc. **64**(2), 281–294 (2002)
12. C.E. Shannon, A mathematical theory of communication. Bell Syst. Tech. J. **27**, 623–656 (1948)
13. J. Suhaila, A.A. Jemain, Fitting daily rainfall amount in peninsular Malaysia using several types of exponential distributions. J. Appl. Sci. Res. **3**(10), 1027–1036 (2007)
14. M. Zahrani-Al, T. Hussain, An algorithm for designing a precipitation network in the south-western region of Saudi Arabia. J. Hydrol. **205**(3–4), 205–216 (1998)
15. Government of Maharashtra (GOM), Report on RTDAS Network of Bhima and Krishna basins, vol. 1, p. 90 (1999)
16. World Meteorological Organization (WMO), Guide to Hydrological Practices, Report No. 168, Geneva (1994)