

Article

Comparison of Entropy Methods for an Optimal Rain Gauge Network: A Case Study of Daegu and Gyeongbuk Area in South Korea [†]

Taeyong Kwon ¹, Junghyun Lim ¹, Seongsim Yoon ²  and Sanghoo Yoon ^{3,*}¹ Department of Statistics, Daegu University, Gyeongbuk 38453, Korea; td1022@daegu.ac.kr (T.K.); ljh90@theimc.co.kr (J.L.)² Korea Institute of Civil Engineering and Building Technology, Goyang 10223, Korea; ssyoon@kict.re.kr³ Division of Mathematics and Big Data Science, Daegu University, Gyeongbuk 38453, Korea

* Correspondence: statstar@daegu.ac.kr

[†] This work was supported by Korea Hydro and Nuclear Power Co., Ltd. (NO. 2018-Tech-20).

Received: 20 July 2020; Accepted: 11 August 2020; Published: 13 August 2020



Abstract: To reduce hydrological disasters, it is necessary to operate rain gauge stations at locations where the spatio-temporal characteristics of rainfall can be reflected. Entropy has been widely used to evaluate the designs and uncertainties associated with rain gauge networks. In this study, the optimal rain gauge network in the Daegu and Gyeongbuk area, which requires the efficient use of water resources due to low annual precipitation and severe drought damage, was determined using conditional and joint entropy, and the selected network was quantitatively evaluated using the root mean square error (RMSE). To consider spatial distribution, prediction errors were generated using kriging. Four estimators used in entropy calculations were compared, and weighted entropy was calculated by weighting the precipitation. The optimal number of rain gauge stations was determined by calculating the RMSE reduction and the reduction ratio according to the number of selected rain gauge stations. Our findings show that the results of conditional entropy were better than those of joint entropy. The optimal rain gauge stations showed a tendency wherein peripheral rain gauge stations were selected first, with central stations being added afterward.

Keywords: entropy; optimal rain gauge network; root mean square error; kriging

1. Introduction

Recently, there has been a significant increase in cases of damage caused by natural disasters globally, due to an increase in extreme weather events caused by climate change. Some of the major causes of socioeconomic damage on the Korean Peninsula are natural disasters related to drought, heavy rain, flooding, and flooding of dams [1]. Droughts, in particular, are a type of natural disaster that can cause long-term damage. Due to limited annual precipitation, the Daegu and Gyeongbuk area is susceptible to severe drought damage [2]. To respond to the risk of such damage, it is important to efficiently utilize and manage water resources through hydrological analysis. The efficient management of water resources can counteract rain-related natural disasters. Water resources can be efficiently managed by identifying information about the rainfall characteristics by basin and the spatial distribution of rainfall. In other words, water resources can be managed using rain information. No wonder information about rainfall with high observational accuracy is important. We need to construct a rain gauge network with high accuracy. The more rain gauge stations there are, the higher the accuracy of the observations can be. However, the higher the number of rain gauge stations, the higher the cost. A rain gauge station should be properly positioned so that it can obtain a lot of

information while minimizing costs to a small number of rain gauge stations. This could also allow a new rain gauge station to be operated at a point where observations are needed though the cost saved. Therefore, the optimal rain gauge network must be constructed based on the spatial distribution of rainfall, and the appropriateness of the constructed network must be evaluated.

Information entropy approaches have been widely used to evaluate the design and uncertainties of networks for collecting hydrological data [3–8]. The basis for designing a rain gauge network using entropy is that the network must have the least amount of overlapping information possible, which means that all rain gauge stations must be independent of each other, as far as possible. One study [9] used the maximum entropy value of rainfall information to develop a model that predicts long-term. One study [10] used the maximum value of entropy to determine the threshold of rainfall. One study [11] objectively evaluated the groundwater quality monitoring network of the Gaza strip, Palestine, using the concept of entropy. The study [12] investigated network designs for different Canadian basins using marginal entropy, joint entropy, and transmitted information. The study [13] showed that the entropy theory can help design the optimal water-level station in the Netherlands. The study [14] developed the maximum information minimum redundancy criterion for the evaluation and planning of water level monitoring networks. Additionally, the study [14] used information theory to quantify the common information among all the monitoring stations of the system. The study [15] evaluated the optimal stations for rainfall data using the information theory of Shannon. The optimal gauge station combination was proposed based on the maximum information transmission using joint entropy. The study [4] calculated the uncertainty of hydrological data based on entropy and evaluated the degree of uncertainty reduction. The studies [16,17] evaluated the gauge station network in Louisiana, USA, and defined the optimal number of stations. The study [6] approached the optimal gauge network design by evaluating the information flow measurement between stations based on entropy. The study [18] used entropy to find locations that could maximize optimal watercourse and hydrological information. The study [19] investigated temporal variations in the quantity and uncertainty of rainfall events using entropy. They showed that entropy is a means of probabilistically comparing and evaluating regional precipitation characteristics and that it can be used as a measure of evaluating water resources.

Entropy cannot reflect the spatial distribution of rain gauge networks because it is calculated based on the probability distribution of observation data. As such, studies have been conducted to reflect the spatial distribution of rain gauge stations along with entropy. The study [20] used the minimum variance estimation method of kriging to select the optimal rain gauge station locations in the basin. The study [21] proposed a method for optimal data collection at random locations through the variance reduction analysis based on kriging. Moreover, studies have been conducted using the entropy theory, which is a concept of information quantity that quantitatively represents a reduction in uncertainty. The study [22] used kriging and entropy to evaluate the optimal rain gauge stations considering the spatial distribution of stations. Kriging was used in the reconstruction of the precipitation of a basin to determine the spatial distribution of rain gauge stations. The optimal gauge station locations were then selected from the stations constructed in the basin through joint entropy. The study [23] secured spatiality through rainfall prediction using kriging and determined the minimum number of rain gauge stations required in a Taiwanese basin using joint entropy. The study [24] used kriging and joint entropy for the optimal network design of Shanghai. They created rainfall data using kriging and proposed an optimal network by calculating entropy.

Rain gauge stations have been evaluated using various methods, such as kriging, regression analysis, correlation analysis, and entropy. Entropy-based studies used conditional and joint entropies to select the optimal stations. In addition, several studies focused on the spatiality of stations because the evaluation of stations using entropy is performed through the empirical probabilities of observation data and does not reflect the spatial distribution of stations. However, no quantitative comparison was conducted to assess how much performance the optimal rain gauge network actually has based on entropy. In statistics and meteorology, prediction models are developed and evaluated through

differences between the observed and predicted values. In the case of continuous data, the root mean square error (RMSE) is mainly used as a model selection criterion. In this study, several estimators of entropy and weighted entropy were compared. A quantitative comparison of the rain gauge networks selected by various entropy methods allows the design of appropriate rain gauge networks. An entropy-based rain gauge network was quantitatively evaluated by predicting the precipitation values of unselected rain gauge stations based on the stations selected as the optimal rain gauge network using entropy. In addition, the performance of a rain gauge network selected as entropy was evaluated using the RMSEs of all rain gauge network combinations. Additionally, a method of determining the number of rain gauge stations using RMSE was proposed. The number of rain gauge stations that can yield the maximum efficiency was determined by calculating the RMSE reduction and the reduction ratio according to the number of rain gauge stations.

The structure of this paper is as follows. Section 2 presents the kriging theory for considering entropy and spatial distribution and discusses the methods of selecting the optimal rain gauge networks and determining the optimal number of rain gauge stations using RMSE. Entropy theory presents several estimators and weighted entropy. In Section 3, the optimal rain gauge network for 11 rain gauge stations in the Daegu and Gyeongbuk area is evaluated based on entropy, and the determination of the number of rain gauge stations for efficient operation is described. Conclusions are presented in Section 4.

2. Methodology

This section discusses the method of selecting the optimal rain gauge network using entropy. The optimal rain gauge network was selected through the conditional and joint entropies of the long-term precipitation observed from rain gauge stations. Entropy cannot consider the spatial distribution of rain gauge stations because uncertainty is calculated from the probability distribution of precipitation. To consider the spatial distribution, prediction errors were generated using ordinary kriging (OK) and leave-one-out cross validation (LOOCV). Rain gauge stations with small prediction errors were the ones with low entropy. The optimal rain gauge network was selected through the conditional and joint entropies of the prediction errors generated by OK. In addition, the precipitations of unselected stations were predicted using OK, and the optimal number of stations was determined using the RMSE. The operating system of the computer used in this study was Windows 10. The central processing unit (CPU) was i5-7500 and the RAM was 32 GB. The software used for data processing was R version 3.5.2. (Daegu university, Korea).

2.1. Research Flowchart

Figure 1 shows the methods used to select the optimal rain gauge stations considering conditional and joint entropies. Methods A and B were the calculation of conditional entropy from the empirical probability distribution of rain gauge data, and method C was the calculation of joint entropy. In addition, the prediction error of each rain gauge station was generated using kriging and LOOCV. As entropy decreases with an increase in the kriging interpolation performance, kriging and entropy were used for considering both the spatiality of rain gauge stations and the uncertainty of rain gauge data. Methods D and E were the calculation of conditional entropy from the probability distribution of prediction error, and method F was the calculation of joint entropy. Methods A and D were used for obtaining the entropy of unselected rain gauge stations using the stations selected through entropy. Methods B and E, on the other hand, were used for obtaining the entropy of selected rain gauge stations using unselected stations. These methods were calculated and compared for each of the entropy estimators. In addition, the entropy with weight was calculated and compared. We applied the monthly average precipitation, regional precipitation, and regional monthly precipitation as weights. Further, in order to focus on places with little rainfall in each region, 1-regional precipitation was applied as a weight, as opposed to regional precipitation. In addition, conditional and joint entropies based on the number of stations were calculated for each method, and the number that maximized the entropies

was determined. The mean absolute error (MAE), RMSE, bias, and correlation coefficient (Corr) were calculated using the predicted precipitations of the rain gauge stations that were not selected in the combination of stations with high entropies. Finally, the optimal number of rain gauge stations was determined using the quality of the RMSE decrease with an increase in the number of stations.

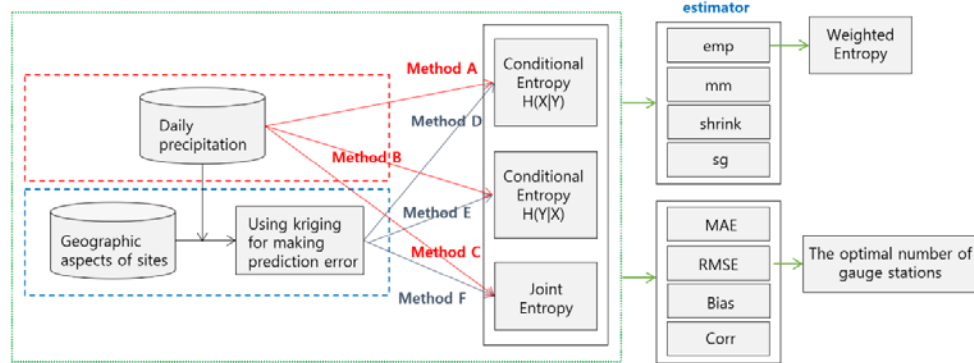


Figure 1. Research flowchart.

2.2. Entropy

Entropy is a method of quantitatively evaluating the degree of randomness using the probability distribution of data [25]. Entropy includes marginal entropy, conditional entropy, and joint entropy. Marginal entropy proposed by [25] is the amount of uncertainty of discrete random variable X , as shown in the following equation.

$$H(X) = - \sum_{n=1}^N p(x_n) \ln p(x_n), \quad (1)$$

where $p(x_n)$ is the empirical probability of x_n and $H(X)$ is the uncertainty of marginal entropy X . n is the number of the class intervals of X .

If there are two random variables X and Y , conditional and joint entropies are calculated through conditional and joint probabilities. When variable Y is given, the entropy of variable X is conditional entropy $H(X|Y)$, as shown below.

$$H(X|Y) = - \sum_{n=1}^N \sum_{m=1}^M p(x_n, y_m) \ln p(x_n | y_m), \quad (2)$$

where n and m of X and Y are the number of class intervals. $P(x_n, y_m)$ is the joint probability and $p(x_n | y_m)$ is the conditional probability. Joint entropy $H(X, Y)$ is calculated from the joint probability of random variables X and Y .

$$H(X, Y) = - \sum_{n=1}^N \sum_{m=1}^M p(x_n, y_m) \ln p(x_n, y_m). \quad (3)$$

Weights can be applied according to the purpose and study characteristics determined by the experimenter. The weight may be independent of or dependent on $p(x_n)$. Weight entropy is given by the following equation [26].

$$WH(X) = - \sum_{n=1}^N w_n p(x_n) \ln p(x_n), \quad (4)$$

where w_n is weight of event. Weight entropy is applicable to conditional and joint entropy. Rain gauge stations with high entropy are important as they have high uncertainties in rainfall observation.

The most common entropy calculation method estimates uncertainty from the empirical probability distribution of data. Several approaches have been developed and used to compensate for the shortcomings of these methods. In this study, three more methods were considered, other than the “emp” method using empirical probability.

“mm” is the bias-corrected empirical entropy estimator proposed by [27]. “shrink” is a shrinkage estimate of the entropy of a Dirichlet probability distribution [28]. “sg” is the entropy of a Dirichlet probability distribution proposed by [29]. In this study, entropy was calculated using the entropy, condentropy, and multiinformation functions of the infotheo package [30].

2.3. Ordinary Kriging (OK)

Kriging is a technique for predicting the values of unobserved points using the weighted combination of known surrounding values. Kriging was developed by [31] and mathematically established by [32]. Kriging defines the spatial interrelationship as a variogram calculated from the given data. The variogram $2\gamma(h)$ is a measure of spatial autocorrelation that represents the similarity of data present at a constant distance h .

$$2\gamma(h) = E[(z(x) - z(x+h))^2], \quad (5)$$

OK predicts the unknown value of point z_0 using z_i , which are n data values already known. OK resolves the bias problem of simple kriging by proposing the sum of weights λ_i as 1.

$$E(z_0) - E(\hat{z}_0) = E(z_0) - E\left(\sum_{i=1}^n \lambda_i z_i\right) = 0, \quad (6)$$

$$1 - \sum_{i=1}^n \lambda_i = 0, \quad (7)$$

Under this condition, weights are obtained so that the error variance can be minimized. The problem of obtaining the minimum and maximum values under constraints is solved by the Lagrange parameter method. The error variance equation of OK is as follows.

$$L(\lambda_1, \lambda_2, \dots, \lambda_n : n) = \sigma^2 - 2 \sum_{i=1}^n \lambda_i \sigma_{0i}^2 + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \sigma_{ij}^2 + 2w(1 - \sum_{i=1}^n \lambda_i), \quad (8)$$

where $\sigma^2 = \text{Var}(z_0)$, $\sigma_{0i}^2 = \text{Cov}(z_0, z_i)$, $\sigma_{ij}^2 = \text{Cov}(z_i, z_j)$.

where w is the Lagrangian factor and $L(\lambda_1, \lambda_2, \dots, \lambda_n : n)$ is the Lagrangian objective function. This study used a variogram based on the exponential model. In this study, OK was performed using the krige function of the gstat package [33].

2.4. Method for Selecting the Optimal Rain Gauge Network

The combination of two or more rain gauge stations that maximizes conditional entropy or joint entropy is considered an optimal combination. In this study, the observation data of unselected rain gauge stations were predicted using the selected rain gauge network and OK to evaluate the suitability of the selected combination of rain gauge stations. This is because the precipitation of unselected stations is quantitatively calculated when the selected rain gauge network is operated. The MAE, RMSE, bias, and Corr calculated from the differences between the observed and predicted values of unselected rain gauge stations were considered as the measures of evaluation.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n (|\hat{z}(x_i) - z(x_i)|), \quad (9)$$

$$\text{RMSE} = \frac{1}{n} \sum_{i=1}^n \sqrt{[\hat{z}(x_i) - z(x_i)]^2}, \quad (10)$$

$$\text{bias} = \frac{1}{n} \sum_{i=1}^n (\hat{z}(x_i) - z(x_i)), \quad (11)$$

$$\text{Corr} = \frac{\sum_{i=1}^n (z(x_i) - \bar{z}(x))(\hat{z}(x_i) - \bar{\hat{z}}(x))}{\sqrt{\sum_{i=1}^n (z(x_i) - \bar{z}(x))^2 \sum_{i=1}^n (\hat{z}(x_i) - \bar{\hat{z}}(x))^2}}, \quad (12)$$

where n is the number of data, $\hat{z}(x_i)$ is the predicted value at an arbitrary position x_i , and $z(x_i)$ is the observed value. In addition, $\bar{\hat{z}}(x_i)$ is the average of the predicted values and $\bar{z}(x_i)$ is the average of the observed values.

The optimal combination according to the determined number of stations can be obtained using the maximum value of entropy. A previous study presented a method of obtaining the optimal number of stations using the maximum informational content [34]. The study selected the combination with the maximum informational content among the various combinations of rain gauge stations based on joint entropy. The prediction performance of the selected combination, however, could not be reflected. This study proposes a method of determining the optimal number of stations through the RMSE calculated using OK. As the number of rain gauge stations increases, the RMSE decreases due to the improvement in the spatial interpolation performance. In other words, the appropriate number of rain gauge stations can be determined by evaluating the degree of the RMSE reduction due to the increase in the number of stations.

First, we considered the difference in the RMSE between the optimal combination of i rain gauge stations and that of $i - 1$ stations. The minimum value of the RMSE ($\text{RMSE}_{\min}$) was obtained when there were $n - 1$ rain gauge stations, and the maximum value of RMSE ($\text{RMSE}_{\max}$) was obtained when there was one rain gauge station. When the number of rain gauge stations increased by one, the average RMSE reduction was obtained from the range of RMSE ($\text{RMSE}_{\max} - \text{RMSE}_{\min}$), as follows.

$$\overline{\text{RMSE}} = \frac{\text{RMSE}_{\max} - \text{RMSE}_{\min}}{n - 1}. \quad (13)$$

Assuming that S_i is the reduction ratio of the RMSE of i rain gauge stations, S_i is calculated using the following equation.

$$S_i = \frac{\text{RMSE}_{i-1} - \text{RMSE}_i}{\overline{\text{RMSE}}}, \text{ where } i = 1, 2, \dots, n - 1. \quad (14)$$

In other words, when S_i is larger than 1, the number of rain gauge stations can be considered an efficient number because the RMSE reduction for i rain gauge stations is relatively larger than the average RMSE reduction. In addition, we calculated the RMSE reduction ratio R_i for i rain gauge stations. R_i represents the RMSE reduction when i rain gauge stations are operated, compared to the total RMSE reduction.

$$R_i = \frac{\text{RMSE}_{\max} - \text{RMSE}_i}{\text{RMSE}_{\max} - \text{RMSE}_{\min}} \times 100. \quad (15)$$

Assuming that x -axis is the number of stations and the y -axis is R_i , $y = 0$ holds for $x = 1$ and $y = 1$ for $x = n - 1$; the farthest point from the straight line of the two points is the optimal number of stations.

2.5. Study Area and Data

The Daegu and Gyeongbuk area is a large basin that occupies the southeast of the Korean Peninsula. Topographically, the Daegu and Gyeongbuk area appears to be affected by the mountains

of the nearby Taebaek mountain range. The limited annual precipitation results in a scarcity of water in the area. Furthermore, climate change has led to an increase in the number of days with heavy rainfall, leading to an increase in disaster risk awareness in the area [35]. Hence, identification of the spatial distribution of rain gauge stations installed in the area and the evaluation of their efficiency are important.

The study used 11 rain gauge stations installed by the Korea Meteorological Administration (KMA) in the Automated Surface Observing System (ASOS) of the Daegu and Gyeongbuk area, as shown in Figure 2. The daily precipitation data from 1988 to 2016 obtained from the meteorological data open portal site (<https://data.kma.go.kr>) were used as research data. The daily maximum precipitation was 516.4 mm, and the average annual precipitation ranged from 1018.2 to 1328.7 mm. The northern part of the study area had higher precipitation than the southern part (Table 1).

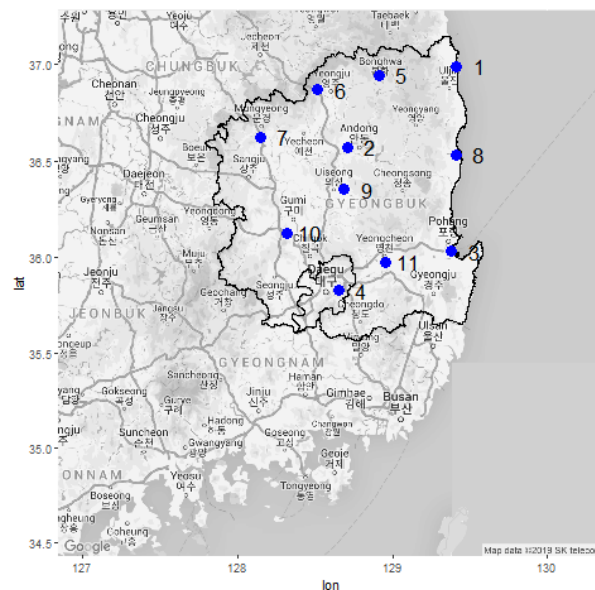


Figure 2. Location of rain gauge station in Daegu and Gyeongbuk (ASOS).

Table 1. Summary of ASOS.

No.	Locations	Period	Longitude	Latitude	Daily Maximum Precipitation	Average Annual Precipitation
1	Uljin	1988~2016	36.992	129.413	279.0	1171.1
2	Andong		36.573	128.707	170.7	1047.1
3	Pohang		36.032	129.380	516.4	1194.5
4	Daegu		35.878	128.653	225.8	1077.6
5	Bonghwa		36.944	128.914	250.0	1190.2
6	Yeongju		36.872	128.517	259.5	1328.7
7	Mungyeong		36.627	128.148	161.0	1281.5
8	Yeongdeok		36.533	129.409	296.0	1085.2
9	Uiseong		36.356	128.688	215.5	1018.2
10	Gumi		36.130	128.320	221.0	1091.9
11	Yeongcheon		35.977	128.951	159.0	1081.7

3. Result and Discussion

3.1. Rain Gauge Network Using Entropy

The precipitation data were converted to discrete spaces to calculate entropy. Figure 3 compares the entropies for class intervals of 100, 200, 300, and 400. Figure 3a shows the entropy of the observed precipitation, and Figure 3b shows the entropy of the prediction error. For both cases, the amount of entropy information increased as the class interval increased; however, the pattern of the maximum information amount according to the number of stations was the same. This study was conducted for a fixed class interval of 200.

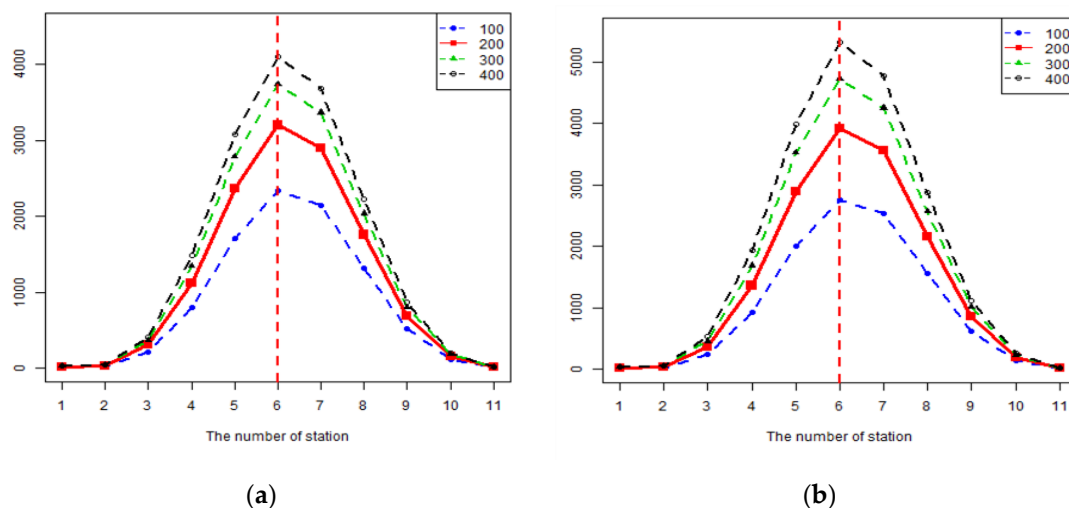


Figure 3. Maximum information of empirical probability distribution: (a) observation data; (b) prediction error.

To find the optimal rain gauge stations of the study area using entropy, the first optimal rain gauge station was selected using marginal entropy. First, entropy was calculated using the basic “emp” method. The rain gauge station installed in Mungyeong was selected as the first station, as its marginal entropy (2.441) was higher than those of other regions. Daegu (2.766) was selected based on the marginal entropy that used the prediction error.

Two or more optimal rain gauge stations can be determined using conditional and joint entropies. As conditional entropy (methods A, B, D, and E) sequentially determines the optimal rain gauge stations, it takes a relatively short time compared with the joint entropy methods. Joint entropy (methods C and F) calculates the entropies of all combinations and selects those with the maximum values. Therefore, they require relatively more computing time than conditional entropy. In addition, the combinations of rain gauge stations were selected irregularly. In contrast, the joint entropy of the prediction error that considered the spatiality of stations sequentially selected the optimal rain gauge stations in a manner similar to conditional entropy (Table 2).

The “mm” and “shrink” methods showed the same results in all cases as the “emp” method (Tables A1 and A2). The calculated entropy values were different, but the selected rainfall network was the same. However, there was a slight difference in the case of the “sg” method (Table A3). In the process of selecting two rain gauge stations with conditional entropy using prediction error (method E), the rain gauge network was selected differently. As a result, the estimator differed in terms of the estimation method, but there was no significant difference noted in this study.

3.2. Rain Gauge Network Using Weighted Entropy

We also considered weighted entropy, as precipitation varies seasonally, and entropy performance varies depending on the amount of precipitation. On heavy precipitation days, entropy was better to select the optimal rain gauge network. We applied monthly average precipitation, regional precipitation, 1-regional precipitation, and regional monthly precipitation as weights.

Table A4 shows the result of calculating the entropy using the monthly average precipitation as a weight. Methods A and F showed one + (better rank), method C showed two + and one – (worse rank). Method D showed one + and one –, and method E showed three – only. Method B did not show any change. In the calculation of entropy using observation data, we can see that a better rain gauge network was selected. However, it can be seen that the entropy calculation using the prediction error was slightly worse.

Table 2. Results of a select optimal rain gauge network of methods (emp).

Num.	Conditional Entropy (1)	Cases	Method A	Num.	Conditional Entropy (1)	Cases	Method D
1	2.441	11	7	1	2.766	11	4
2	1.590	10	7,(11)	2	1.953	10	4,(7)
3	0.873	9	7,11,(1)	3	1.029	9	4,7,(1)
4	0.471	8	7,11,1,(5)	4	0.479	8	4,7,1,(10)
5	0.236	7	7,11,1,5,(4)	5	0.334	7	4,7,1,10,(8)
6	0.187	6	7,11,1,5,4,(8)	6	0.253	6	4,7,1,10,8,(5)
7	0.166	5	7,11,1,5,4,8,(6)	7	0.175	5	4,7,1,10,8,5,(11)
8	0.116	4	7,11,1,5,4,8,6,(10)	8	0.113	4	4,7,1,10,8,5,11,(9)
9	0.095	3	7,11,1,5,4,8,6,10,(3)	9	0.074	3	4,7,1,10,8,5,11,9,(6)
10	0.079	2	7,11,1,5,4,8,6,10,3,(2)	10	0.054	2	4,7,1,10,8,5,11,9,6,(2)
11	0.047	1	7,11,1,5,4,8,6,10,3,2,(9)	11	0.026	1	4,7,1,10,8,5,11,9,6,2,(3)
Computing time: 1.00 s				Computing time: 1.03 s			
Num.	Conditional Entropy (2)	Cases	Method B	Num.	Conditional Entropy (2)	Cases	Method E
1	2.441	11	7	1	2.766	11	4
2	2.025	10	7,(3)	2	2.087	10	4,(5)
3	2.576	9	7,3,(1)	3	3.127	9	4,5,(3)
4	3.078	8	7,3,1,(4)	4	3.462	8	4,5,3,(8)
5	3.448	7	7,3,1,4,(8)	5	3.885	7	4,5,3,8,(9)
6	3.598	6	7,3,1,4,8,(5)	6	4.083	6	4,5,3,8,9,(6)
7	3.851	5	7,3,1,4,8,5,(9)	7	4.176	5	4,5,3,8,9,6,(2)
8	3.933	4	7,3,1,4,8,5,9,(10)	8	4.126	4	4,5,3,8,9,6,2,(7)
9	3.991	3	7,3,1,4,8,5,9,10,(6)	9	4.335	3	4,5,3,8,9,6,2,7,(10)
10	3.976	2	7,3,1,4,8,5,9,10,6,(11)	10	4.435	2	4,5,3,8,9,6,2,7,10,(11)
11	4.069	1	7,3,1,4,8,5,9,10,6,11,(2)	11	4.530	1	4,5,3,8,9,6,2,7,10,11,(1)
Computing time: 0.52 s				Computing time: 0.52 s			
Num.	Joint Entropy	Cases	Method C	Num.	Joint Entropy	Cases	Method F
1	2.441	11	7	1	2.766	11	4
2	0.918	55	4,11	2	0.996	55	4,11
3	2.387	165	2,6,7	3	2.683	165	4,10,11
4	4.098	330	2,5,6,7	4	4.789	330	4,7,10,11
5	5.895	462	2,7,9,10,11	5	7.101	462	1,4,7,10,11
6	7.774	462	2,6,7,9,10,11	6	9.376	462	1,2,4,7,10,11
7	9.664	330	2,5,6,7,9,10,11	7	11.596	330	1,2,4,6,7,10,11
8	11.558	165	2,4,5,6,7,9,10,11	8	13.806	165	1,2,4,6,7,9,10,11
9	13.193	55	2,4,5,6,7,8,9,10,11	9	16.010	55	1,2,4,6,7,8,9,10,11
10	14.913	11	1,2,4,5,6,7,8,9,10,11	10	18.170	11	1,2,4,5,6,7,8,9,10,11
11	16.443	1	1,2,3,4,5,6,7,8,9,10,11	11	20.265	1	1,2,3,4,5,6,7,8,9,10,11
Computing time: 24.79 s				Computing time: 23.67 s			

Table A5 shows the result of calculating the entropy using regional precipitation as a weight. Method A had one + and four – changes, method B had only two –, and method C had only four –. Methods D and E had two + and seven –, and method F had seven + and two –. Most methods showed worse results; however, method F showed improved results. Nevertheless, this was confirmed as a worse result than that of method B.

Table A6 shows the results of calculating the entropy using 1-regional precipitation as a weight. Method B had one –; method C had one +. Method D had four –, method E had two – and three +, and method F had five –. Method A was unchanged. There were some improvements in methods C and E, but most methods showed poor results.

Table A7 shows the results of calculating the entropy using regional monthly precipitation as a weight. Method A had one +, and Method B had two –. Method C had one + and two –, method D had one + and one –, and method E had two + and two –. Method F remained unchanged. There was a slight improvement in the results, except for those by method F, but there was also a deterioration. Although it looks similar, entropy had a bad appearance in method B, which was considered to be the best rank. The schematic of these results is shown in Table 3 below.

3.3. Discussion

Table 4 shows the top five station combinations by calculating the RMSEs for all combinations. In addition, the RMSEs of the rain gauge network combinations by the number of rain gauge stations are shown in Figure 4. When selecting one rain gauge station, the stations located in the centers of Daegu

and Gyeongbuk regions showed excellent prediction performance. However, if one selected more than two rain gauge stations, you can see that stations located outside Daegu and Gyeongbuk were mainly selected. In other words, the station when selecting one rain gauge station was not often selected in more than one combination. Rain gauge stations located in Pohang (3) were frequently selected from a combination of two stations, and rain gauge stations in Bonghwa (5) were also frequently selected. From the combination of three rain gauge stations, Yeongju (6) was added and selected, and from the combination of five rain gauge stations, Uljin (1) can be confirmed to be included in all combinations. Pohang, Bonghwa, Yeongju, and Uljin are rain gauge stations located along the north and the coast in Daegu and Gyeongbuk region. In several combinations of rain gauge stations, Uiseong (9), Gumi (10), and Yeongcheon (11) were added, and the prediction performance of the rain gauge network was improved.

Table 3. Results of increase and decrease of rank after application of weighted entropy.

Weight: Monthly Average Precipitation			Weight: Precipitation by 1-Station		
Method	+	−	Method	+	−
A	1	0	A	0	0
B	0	0	B	0	1
C	2	1	C	1	0
D	1	1	D	0	3
E	0	3	E	3	2
F	1	0	F	0	5

Weight: precipitation by station			Weight: monthly average precipitation by station		
Method	+	−	Method	+	−
A	1	4	A	1	0
B	0	2	B	0	2
C	0	4	C	1	2
D	2	7	D	1	1
E	2	7	E	2	2
F	7	2	F	0	0

Table 4. The best of all combinations of RMSE.

Num.	Rank	Combination	RMSE	Num.	Rank	Combination	RMSE
1	1	2	11.62	6	1	1,3,4,5,6,9	6.94
	2	9	11.77		2	1,3,4,5,7,9	7.00
	3	10	12.23		3	1,3,4,6,7,9	7.06
	4	11	12.50		4	1,3,4,5,7,10	7.10
	5	8	12.76		5	1,3,5,6,9,11	7.13
2	1	6,11	9.57	7	1	1,3,4,5,6,8,9	6.86
	2	2,3	9.75		2	1,3,4,5,6,7,9	6.89
	3	7,11	9.77		3	1,2,3,4,5,6,10	6.99
	4	3,5	9.82		4	1,3,4,5,7,8,9	7.02
	5	5,11	9.84		5	1,3,4,5,6,9,10	7.03
3	1	3,6,9	8.20	8	1	1,3,4,5,6,7,8,9	6.66
	2	4,6,8	8.38		2	1,3,4,5,6,7,9,10	6.75
	3	6,8,11	8.53		3	1,2,3,4,5,6,8,10	6.88
	4	4,5,8	8.54		4	1,3,5,6,7,9,10,11	6.88
	5	3,6,10	8.56		5	1,2,3,4,5,6,7,10	6.90
4	1	1,3,6,9	7.72	9	1	1,3,4,5,6,7,8,9,10	6.32
	2	3,4,6,9	7.93		2	1,2,3,4,5,6,7,8,10	6.52
	3	1,3,4,6	7.98		3	1,3,5,6,7,8,9,10,11	6.60
	4	2,3,4,6	7.98		4	1,3,4,5,6,7,8,9,11	6.70
	5	1,3,6,10	8.03		5	1,2,3,5,6,7,8,10,11	6.81
5	1	1,3,4,6,9	7.15	10	1	1,3,4,5,6,7,8,9,10,11	6.19
	2	1,2,3,4,6	7.29		2	1,2,3,4,5,6,7,8,10,11	6.45
	3	1,3,6,9,11	7.30		3	1,2,3,4,5,6,7,8,9,10	6.58
	4	1,3,4,5,7	7.46		4	1,2,3,5,6,7,8,9,10,11	7.16
	5	1,2,3,6,11	7.53		5	1,2,3,4,5,6,7,8,9,11	7.29

In the case of a rain gauge network using conditional entropy, Mungyeong (7) and Daegu (4) appear to have been selected first, and the performance seems to have deteriorated. In the case of method B, which was judged to be the best method for selecting the rain gauge network using entropy, the performance was improved with the addition of Bonghwa and Yeongju in the north. In other

words, it is considered that the rain gauge stations in the north are important locations for rain gauge networks in Daegu and Gyeongbuk.

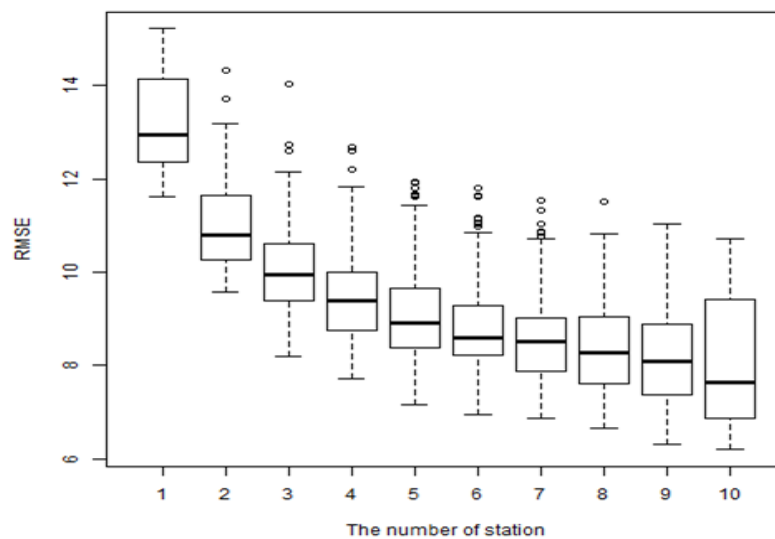


Figure 4. The performances of all combination gauge networks.

3.4. Optimal Number of Rain Gauge Stations

For comparing the six methods, the errors generated by predicting the average daily precipitation of the unselected stations based on the selected stations were evaluated through MAE, RMSE, bias, and Corr (Figure 5).

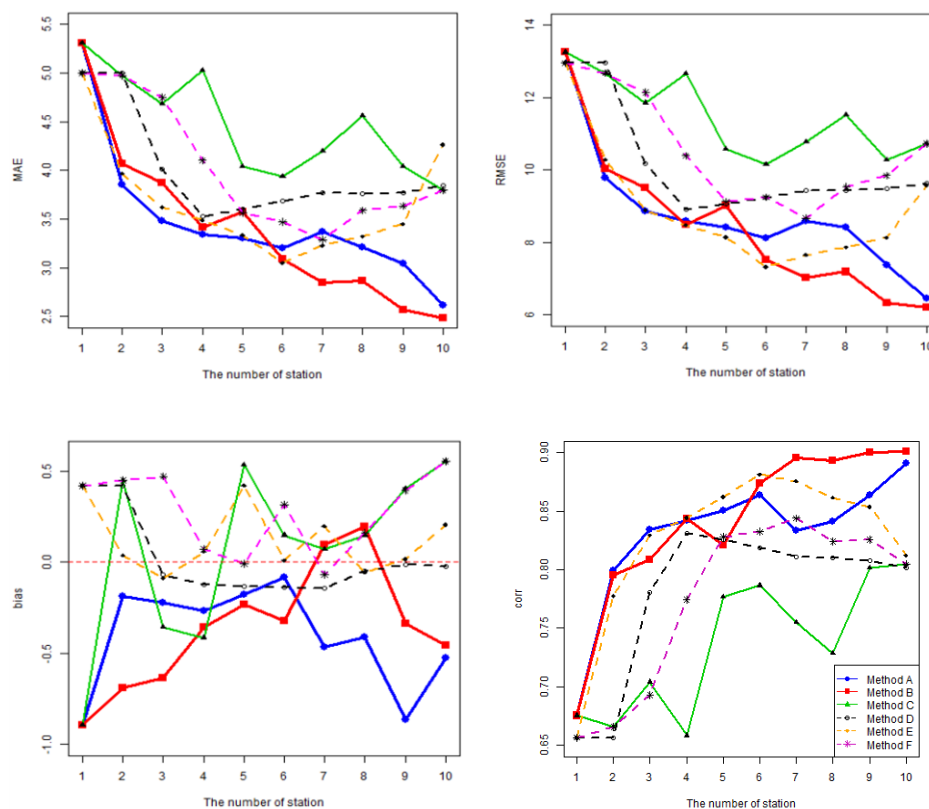


Figure 5. Result of MAE, RMSE, bias, and correlation coefficient (Corr) variation according to the numbers of stations for the rain gauge network calculated by different methods.

Conditional entropy methods (methods A, B, D, and E) generally exhibited similar results. For joint entropy methods (methods C and F), the results were similar until the number of stations was three; different results were observed for a higher number of stations. The conditional entropy methods exhibited better results than joint entropy because MAE and RMSE for these were lower, and Corr was higher. With regard to the bias, methods A and B showed a tendency to predict lower values than the observed values, while methods C, D, E, and F predicted higher values. For the RMSE, methods A and D exhibited similar results; however, they started to show different results when the number of stations reached nine. The RMSE of method A became lower and exhibited excellent performance, with Pohang being the location of the added station. For method D, Pohang was the last station added. This was because it exhibited the highest daily maximum precipitation of 516.4 mm based on the research data mentioned above. In method B, the rain gauge stations installed in coastal areas were first selected. As for the RMSE values of methods A and B, method A exhibited better results when the number of stations was less than six, and method B when the number of stations was six or more. In other words, methods A and B exhibited relatively good performance.

Table 5 and Figure 6 show the RMSE results for determining the optimal number of stations. S shows the actually decreasing amount under the assumption that RMSE decreases at equal intervals as the number of rain gauge stations increases. In other words, it is meaningful to find the values of S higher than one and higher than the average value. For method A, the numbers of stations where S was higher than one were 2, 3, 9, and 10. The number of stations that exhibited the highest value was two (4.436), followed by nine (1.334), three (1.173), and 10 (1.162). For method B, the numbers of stations where S was higher than one were 2, 4, 6, and 9. The number of stations that showed the highest value was two (4.119), followed by six (1.912), four (1.300), and nine (1.106).

Table 5. RMSE result to determine the optimal number of stations for the rain gauge network.

Method A			Method B		
No.	S	R	No.	S	R
2	4.436	49.28%	2	4.119	45.77%
3	1.173	62.32%	3	0.652	53.02%
4	0.337	66.05%	4	1.300	67.46%
5	0.220	68.50%	5	−0.670	60.01%
6	0.386	72.79%	6	1.912	81.26%
7	−0.613	65.98%	7	0.627	88.23%
8	0.238	68.62%	8	−0.206	85.94%
9	1.334	83.45%	9	1.106	98.23%
10	1.162	96.37%	10	0.159	100.00%

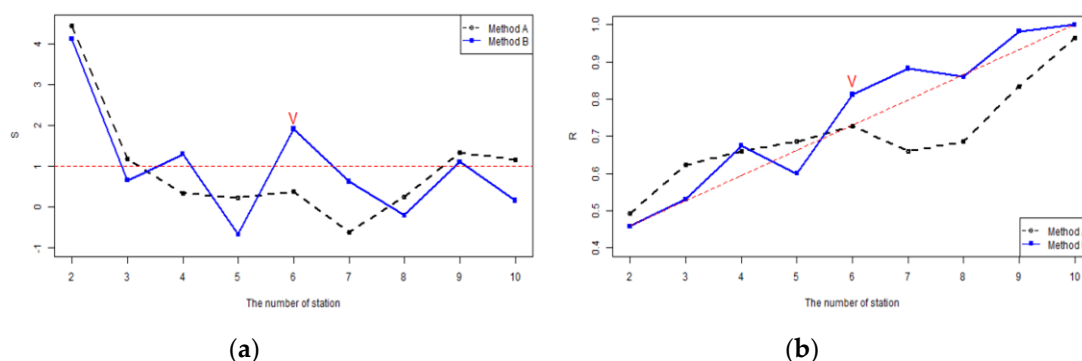


Figure 6. Optimal number of rain gauge networks using RMSE: (a) reduction of RMSE (S); (b) ratio of RMSE (R).

R is the predictive power of i rain gauge stations by comparing the total RMSE reduction with the RMSE reduction when i stations are used. For method A, R was 83.45% and represented more than 80% when the number of stations was nine or higher. For method B, R was 81.26% and represented more than 80% when the number of stations was six or higher. Subsequently, R did not significantly increase

even with an increase in the number of stations. Hence, six stations were judged to be appropriate. In addition, when the farthest R values from the average curve were compared, six and seven stations had the best predictive power.

The optimal number of stations was determined to increase the accuracy of observation while minimizing costs. If the minimum number of stations is operated considering the accuracy of rainfall observation, then the optimal number of stations is six. This number was selected based on the stations of method B, because method B exhibited excellent performance for up to six stations. The optimal rain gauge stations installed in the Daegu and Gyeongbuk area were Mungyeong, Pohang, Uljin, Daegu, Yeongdeok, and Bonghwa. Figure 7 shows the optimal rain gauge stations in the order of selection. The selected stations were located on the outskirts of the Daegu and Gyeongbuk area.

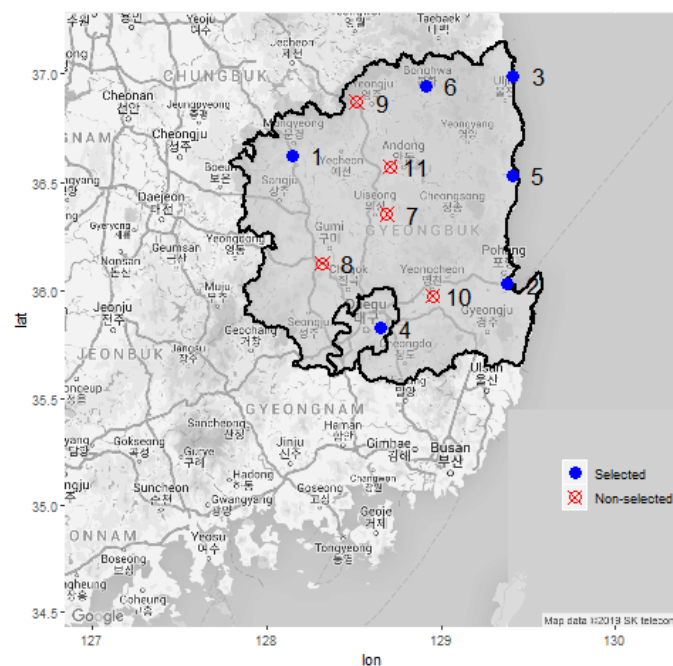


Figure 7. Results of the optimal rain gauge network selection.

4. Conclusions

In this study, the uncertainties were calculated and compared according to the various estimators of entropy. There was no significant difference between the entropy estimators “emp,” “mm,” “shrink,” and “sg.” In addition, weighted entropy was calculated to determine the uncertainty by applying monthly average precipitation, regional precipitation, and regional monthly precipitation as weights. In order to focus on regions with little rainfall, 1-regional precipitation was applied as a weight as opposed to regional precipitation. In this case, there was a tendency to improve, but also a tendency to get worse. There was no significant difference noted when using method B, which showed the best results.

The RMSEs of all combinations of rain gauge stations were calculated to determine the patterns of the top five combinations. The main locations for rain gauge stations were Pohang, Bonghwa, Yeongju, and Uljin, which are located along the north and the coast of Daegu and Gyeongbuk regions. Uiseong, Gumi, and Yeongcheon in the central region can also be considered as locations for rain gauge stations.

The study also proposed a method for determining the optimal number of stations by calculating the RMSE reduction and (S) reduction ratio (R). The optimal rain gauge network was selected using entropy, and the optimal number of rain gauge stations was determined through RMSE for the efficient operation of rain gauge stations. RMSE was obtained by predicting the unselected rain gauge stations using the observation data of the stations selected according to the optimal number of stations. When the value of S was higher than one according to the number of rain gauge stations, the case with R higher than 80% was

selected to determine the optimal number of rain gauge stations. When the optimal number of rain gauge stations was determined for the 11 stations in the Daegu and Gyeongbuk area, 80% performance could be obtained with six stations, and performance similar to that of all stations could be obtained with only six stations. The peripheral stations were selected first, followed by the central stations. The uncertainties of the peripheral stations were found to be higher because of large fluctuations.

It was confirmed that the conditional entropy of observation data was excellent for selecting the optimal rain gauge network and the proportion of RMSE was used for the selection of the optimal number of rain gauge stations. They can contribute to the improvement of the accuracy of hydrological research and the efficient operation of rain gauge stations. Future research will allow efficient rain gauge networks to be designed if radar data are used.

Author Contributions: Conceptualization, J.L. and S.Y. (Sanghoo Yoon); methodology, T.K., J.L. and S.Y. (Sanghoo Yoon); software, S.Y. (Seongsim Yoon); validation, T.K., S.Y. (Seongsim Yoon) and S.Y. (Sanghoo Yoon); formal analysis, T.K. and J.L.; investigation, T.K. and J.L.; resources, S.Y. (Seongsim Yoon) and S.Y. (Sanghoo Yoon); data curation, T.K. and S.Y. (Sanghoo Yoon); writing—original draft preparation, T.K. and J.L.; writing—review and editing, T.K., S.Y. (Seongsim Yoon) and S.Y. (Sanghoo Yoon); visualization, T.K.; supervision, S.Y. (Seongsim Yoon) and S.Y. (Sanghoo Yoon); project administration, T.K., S.Y. (Seongsim Yoon) and S.Y. (Sanghoo Yoon); funding acquisition, S.Y. (Seongsim Yoon) and S.Y. (Sanghoo Yoon). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Korea Hydro and Nuclear Power Co., Ltd., grant number 2018-Tech-20.

Acknowledgments: I would like to express my gratitude to all those who contributed to this study.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Results of a select optimal rain gauge network of methods (mm).

Num.	Conditional Entropy (1)	Cases	Method A	Num.	Conditional Entropy (1)	Cases	Method D
1	2.455	11	7	1	2.779	11	4
2	1.686	10	7,(11)	2	2.045	10	4,(7)
3	0.937	9	7,11,(1)	3	1.125	9	4,7,(1)
4	0.511	8	7,11,1,(5)	4	0.529	8	4,7,1,(10)
5	0.260	7	7,11,1,5,(4)	5	0.361	7	4,7,1,10,(8)
6	0.204	6	7,11,1,5,4,(8)	6	0.272	6	4,7,1,10,8,(5)
7	0.180	5	7,11,1,5,4,8,(6)	7	0.191	5	4,7,1,10,8,5,(11)
8	0.127	4	7,11,1,5,4,8,6,(10)	8	0.124	4	4,7,1,10,8,5,11,(9)
9	0.104	3	7,11,1,5,4,8,6,10,(3)	9	0.081	3	4,7,1,10,8,5,11,9,(6)
10	0.088	2	7,11,1,5,4,8,6,10,3,(2)	10	0.064	2	4,7,1,10,8,5,11,9,6,(2)
11	0.051	1	7,11,1,5,4,8,6,10,3,2,(9)	11	0.031	1	4,7,1,10,8,5,11,9,6,2,(3)
Computing time: 0.88 s				Computing time: 0.97 s			
Num.	Conditional Entropy (2)	Cases	Method B	Num.	Conditional Entropy (2)	Cases	Method E
1	2.455	11	7	1	2.779	11	4
2	2.097	10	7,(3)	2	2.164	10	4,(5)
3	2.722	9	7,3,(1)	3	3.284	9	4,5,(3)
4	3.265	8	7,3,1,(4)	4	3.671	8	4,5,3,(8)
5	3.661	7	7,3,1,4,(8)	5	4.127	7	4,5,3,8,(9)
6	3.835	6	7,3,1,4,8,(5)	6	4.347	6	4,5,3,8,9,(6)
7	4.103	5	7,3,1,4,8,5,(9)	7	4.459	5	4,5,3,8,9,6,(2)
8	4.196	4	7,3,1,4,8,5,9,(10)	8	4.428	4	4,5,3,8,9,6,2,(7)
9	4.266	3	7,3,1,4,8,5,9,10,(6)	9	4.652	3	4,5,3,8,9,6,2,7,(10)
10	4.258	2	7,3,1,4,8,5,9,10,6,(11)	10	4.759	2	4,5,3,8,9,6,2,7,10,(11)
11	4.359	1	7,3,1,4,8,5,9,10,6,11,(2)	11	4.865	1	4,5,3,8,9,6,2,7,10,11,(1)
Computing time: 0.56 s				Computing time: 0.63 s			
Num.	Joint Entropy	Cases	Method C	Num.	Joint Entropy	Cases	Method F
1	2.455	11	7	1	2.779	11	4
2	0.856	55	4,11	2	0.927	55	4,11
3	2.253	165	2,6,7	3	2.531	165	4,10,11
4	3.940	330	2,5,6,7	4	4.593	330	4,7,10,11
5	5.724	462	2,7,9,10,11	5	6.881	462	1,4,7,10,11
6	7.597	462	2,6,7,9,10,11	6	9.147	462	1,2,4,7,10,11
7	9.481	330	2,5,6,7,9,10,11	7	11.365	330	1,2,4,6,7,10,11
8	11.374	165	2,4,5,6,7,9,10,11	8	13.575	165	1,2,4,6,7,9,10,11
9	13.008	55	2,4,5,6,7,8,9,10,11	9	15.775	55	1,2,4,6,7,8,9,10,11
10	14.722	11	1,2,4,5,6,7,8,9,10,11	10	17.934	11	1,2,4,5,6,7,8,9,10,11
11	16.250	1	1,2,3,4,5,6,7,8,9,10,11	11	20.031	1	1,2,3,4,5,6,7,8,9,10,11
Computing time: 23.17 s				Computing time: 29.60 s			

Table A2. Results of a select optimal rain gauge network of methods (shrink).

Num.	Conditional Entropy (1)	Cases	Method A	Num.	Conditional Entropy (1)	Cases	Method D
1	2.431	11	7	1	2.760	11	4
2	1.508	10	7,(11)	2	1.878	10	4,(7)
3	0.821	9	7,11,(1)	3	0.951	9	4,7,(1)
4	0.438	8	7,11,1,(5)	4	0.439	8	4,7,1,(10)
5	0.216	7	7,11,1,5,(4)	5	0.317	7	4,7,1,10,(8)
6	0.173	6	7,11,1,5,4,(8)	6	0.242	6	4,7,1,10,8,(5)
7	0.154	5	7,11,1,5,4,8,(6)	7	0.165	5	4,7,1,10,8,5,(11)
8	0.106	4	7,11,1,5,4,8,6,(10)	8	0.106	4	4,7,1,10,8,5,11,(9)
9	0.088	3	7,11,1,5,4,8,6,10,(3)	9	0.067	3	4,7,1,10,8,5,11,9,(6)
10	0.072	2	7,11,1,5,4,8,6,10,3,(2)	10	0.045	2	4,7,1,10,8,5,11,9,6,(2)
11	0.044	1	7,11,1,5,4,8,6,10,3,2,(9)	11	0.022	1	4,7,1,10,8,5,11,9,6,2,(3)
Computing time: 1.06 s				Computing time: 1.31 s			
Num.	Conditional Entropy (2)	Cases	Method B	Num.	Conditional Entropy (2)	Cases	Method E
1	2.431	11	7	1	2.760	11	4
2	1.963	10	7,(3)	2	2.026	10	4,(5)
3	2.455	9	7,3,(1)	3	3.001	9	4,5,(3)
4	2.922	8	7,3,1,(4)	4	3.296	8	4,5,3,(8)
5	3.271	7	7,3,1,4,(8)	5	3.694	7	4,5,3,8,(9)
6	3.400	6	7,3,1,4,8,(5)	6	3.874	6	4,5,3,8,9,(6)
7	3.642	5	7,3,1,4,8,5,(9)	7	3.950	5	4,5,3,8,9,6,(2)
8	3.715	4	7,3,1,4,8,5,9,(10)	8	3.888	4	4,5,3,8,9,6,2,(7)
9	3.763	3	7,3,1,4,8,5,9,10,(6)	9	4.085	3	4,5,3,8,9,6,2,7,(10)
10	3.742	2	7,3,1,4,8,5,9,10,6,(11)	10	4.181	2	4,5,3,8,9,6,2,7,10,(11)
11	3.829	1	7,3,1,4,8,5,9,10,6,11,(2)	11	4.271	1	4,5,3,8,9,6,2,7,10,11,(1)
Computing time: 0.59 s				Computing time: 0.59 s			
Num.	Joint Entropy	Cases	Method C	Num.	Joint Entropy	Cases	Method F
1	2.431	11	7	1	2.760	11	4
2	0.972	55	4,11	2	1.056	55	4,11
3	2.501	165	2,6,7	3	2.815	165	4,10,11
4	4.234	330	2,5,6,7	4	4.956	330	4,7,10,11
5	6.040	462	2,7,9,10,11	5	7.287	462	1,4,7,10,11
6	7.926	462	2,6,7,9,10,11	6	9.573	462	1,2,4,7,10,11
7	9.820	330	2,5,6,7,9,10,11	7	11.797	330	1,2,4,6,7,10,11
8	11.716	165	2,4,5,6,7,9,10,11	8	14.007	165	1,2,4,6,7,9,10,11
9	13.353	55	2,4,5,6,7,8,9,10,11	9	16.213	55	1,2,4,6,7,8,9,10,11
10	15.077	11	1,2,4,5,6,7,8,9,10,11	10	18.371	11	1,2,4,5,6,7,8,9,10,11
11	16.609	1	1,2,3,4,5,6,7,8,9,10,11	11	20.466	1	1,2,3,4,5,6,7,8,9,10,11
Computing time: 27.93 s				Computing time: 21.57 s			

Table A3. Results of a select optimal rain gauge network of methods (sg).

Num.	Conditional Entropy (1)	Cases	Method A	Num.	Conditional Entropy (1)	Cases	Method D
1	2.429	11	7	1	2.754	11	4
2	1.506	10	7,(11)	2	1.873	10	4,(7)
3	0.819	9	7,11,(1)	3	0.948	9	4,7,(1)
4	0.438	8	7,11,1,(5)	4	0.438	8	4,7,1,(10)
5	0.216	7	7,11,1,5,(4)	5	0.312	7	4,7,1,10,(8)
6	0.173	6	7,11,1,5,4,(8)	6	0.237	6	4,7,1,10,8,(5)
7	0.153	5	7,11,1,5,4,8,(6)	7	0.616	5	4,7,1,10,8,5,(11)
8	0.106	4	7,11,1,5,4,8,6,(10)	8	0.104	4	4,7,1,10,8,5,11,(9)
9	0.088	3	7,11,1,5,4,8,6,10,(3)	9	0.067	3	4,7,1,10,8,5,11,9,(6)
10	0.072	2	7,11,1,5,4,8,6,10,3,(2)	10	0.046	2	4,7,1,10,8,5,11,9,6,(2)
11	0.044	1	7,11,1,5,4,8,6,10,3,2,(9)	11	0.022	1	4,7,1,10,8,5,11,9,6,2,(3)
Computing time: 0.89 s				Computing time: 1.02 s			
Num.	Conditional Entropy (2)	Cases	Method B	Num.	Conditional Entropy (2)	Cases	Method E
1	2.429	11	7	1	2.754	11	4
2	1.961	10	7,(3)	2	2.021	10	4,(3)
3	2.452	9	7,3,(1)	3	2.804	9	4,3,(5)
4	2.918	8	7,3,1,(4)	4	3.283	8	4,3,5,(8)
5	3.267	7	7,3,1,4,(8)	5	3.679	7	4,3,5,8,(9)
6	3.396	6	7,3,1,4,8,(5)	6	3.859	6	4,3,5,8,9,(6)
7	3.638	5	7,3,1,4,8,5,(9)	7	3.936	5	4,3,5,8,9,6,(2)
8	3.710	4	7,3,1,4,8,5,9,(10)	8	3.870	4	4,3,5,8,9,6,2,(7)
9	3.758	3	7,3,1,4,8,5,9,10,(6)	9	4.067	3	4,3,5,8,9,6,2,7,(10)
10	3.736	2	7,3,1,4,8,5,9,10,6,(11)	10	4.159	2	4,3,5,8,9,6,2,7,10,(11)
11	3.823	1	7,3,1,4,8,5,9,10,6,11,(2)	11	4.245	1	4,3,5,8,9,6,2,7,10,11,(1)
Computing time: 0.61 s				Computing time: 0.54 s			
Num.	Joint Entropy	Cases	Method C	Num.	Joint Entropy	Cases	Method F
1	2.429	11	7	1	2.754	11	4
2	0.971	55	4,11	2	1.056	55	4,11
3	2.498	165	2,6,7	3	2.812	165	4,10,11
4	4.229	330	2,5,6,7	4	4.953	330	4,7,10,11
5	6.035	462	2,7,9,10,11	5	7.283	462	1,4,7,10,11
6	7.919	462	2,6,7,9,10,11	6	9.565	462	1,2,4,7,10,11
7	9.812	330	2,5,6,7,9,10,11	7	11.786	330	1,2,4,6,7,10,11
8	11.706	165	2,4,5,6,7,9,10,11	8	13.995	165	1,2,4,6,7,9,10,11
9	13.343	55	2,4,5,6,7,8,9,10,11	9	16.201	55	1,2,4,6,7,8,9,10,11
10	15.067	11	1,2,4,5,6,7,8,9,10,11	10	18.362	11	1,2,4,5,6,7,8,9,10,11
11	16.598	1	1,2,3,4,5,6,7,8,9,10,11	11	20.454	1	1,2,3,4,5,6,7,8,9,10,11
Computing time: 24.27 s				Computing time: 25.14 s			

Table A4. Results of a select optimal rain gauge network of methods (weight: monthly average precipitation).

Num.	Conditional Entropy (1)	Rank	Method A	Num.	Conditional Entropy (1)	Rank	Method D
1	4.739	7	7	1	5.071	6	4
2	1.072	3	7,(11)	2	1.435	7	4,(7)
3	0.590	19	7,11,(1)	3	0.554	17	4,7,(1)
4	0.281	64	7,11,1,(5)	4	0.291	100(−)	4,7,1,(8)
5	0.130	84(+)	7,11,1,5,(8)	5	0.212	45(+)	4,7,1,8,(5)
6	0.111	94	7,11,1,5,8,(4)	6	0.149	54	4,7,1,8,5,(10)
7	0.098	176	7,11,1,5,8,4,(6)	7	0.101	106	4,7,1,8,5,10,(11)
8	0.067	91	7,11,1,5,8,4,6,(10)	8	0.061	98	4,7,1,8,5,10,11,(9)
9	0.055	14	7,11,1,5,8,4,6,10,(3)	9	0.033	39	4,7,1,8,5,10,11,9,(6)
10	0.041	2	7,11,1,5,8,4,6,10,3,(2)	10	0.016	11	4,7,1,8,5,10,11,9,6,(2)
11	0.031		7,11,1,5,8,4,6,10,3,2,(9)	11	0.008		4,7,1,8,5,10,11,9,6,2,(3)
Computing time: 0.86 s				Computing time: 0.92 s			
Num.	Conditional Entropy (2)	Rank	Method B	Num.	Conditional Entropy (2)	Rank	Method E
1	4.739	7	7	1	5.071	6	4
2	1.589	9	7,(3)	2	1.675	51(−)	4,(3)
3	1.810	53	7,3,(1)	3	2.111	21	4,3,(5)
4	2.137	47	7,3,1,(4)	4	2.413	42	4,3,5,(8)
5	2.377	255	7,3,1,4,(8)	5	2.684	59	4,3,5,8,(9)
6	2.427	27	7,3,1,4,8,(5)	6	2.782	12	4,3,5,8,9,(6)
7	2.613	4	7,3,1,4,8,5,(9)	7	2.807	54	4,3,5,8,9,6,(2)
8	2.656	15	7,3,1,4,8,5,9,(10)	8	2.656	65(−)	4,3,5,8,9,6,2,(10)
9	2.662	1	7,3,1,4,8,5,9,10,(6)	9	2.729	41(−)	4,3,5,8,9,6,2,10,(11)
10	2.614	1	7,3,1,4,8,5,9,10,6,(11)	10	2.801	10	4,3,5,8,9,6,2,10,11,(7)
11	2.667		7,3,1,4,8,5,9,10,6,11,(2)	11	2.904		4,3,5,8,9,6,2,10,11,7,(1)
Computing time: 0.63 s				Computing time: 0.57 s			
Num.	Joint Entropy	Rank	Method C	Num.	Joint Entropy	Rank	Method F
1	4.739	7	7	1	5.071	6	4
2	3.672	40(+)	2,7	2	3.756	50	4,11
3	7.784	156	2,6,7	3	8.191	161	4,10,11
4	11.981	330	2,5,6,7	4	12.855	275	4,7,10,11
5	16.225	413(+)	2,5,6,7,11	5	17.635	272	1,4,7,10,11
6	20.499	437(−)	2,5,6,7,9,11	6	22.302	340	1,2,4,7,10,11
7	24.784	324	2,5,6,7,9,10,11	7	26.905	189	1,2,4,6,7,10,11
8	29.071	165	2,4,5,6,7,9,10,11	8	31.487	105(+)	1,2,4,6,7,8,10,11
9	33.132	53	2,4,5,6,7,8,9,10,11	9	36.077	49	1,2,4,6,7,8,9,10,11
10	37.287	11	1,2,4,5,6,7,8,9,10,11	10	40.631	11	1,2,4,5,6,7,8,9,10,11
11	41.228		1,2,3,4,5,6,7,8,9,10,11	11	45.082		1,2,3,4,5,6,7,8,9,10,11
Computing time: 24.94 s				Computing time: 23.36 s			

Table A5. Results of a select optimal rain gauge network of methods (weight: precipitation by station).

Num.	Conditional Entropy (1)	Rank	Method A	Num.	Conditional Entropy (1)	Rank	Method D
1	1.229	7	7	1	1.143	7(−)	7
2	0.806	3	7,(11)	2	0.897	35(−)	7,(1)
3	0.548	19	7,11,(1)	3	0.533	17	7,1,(4)
4	0.379	64	7,11,1,(5)	4	0.432	114(−)	7,1,4,(6)
5	0.269	53(+)	7,11,1,5,(10)	5	0.351	172(+)	7,1,4,6,(11)
6	0.218	261(−)	7,11,1,5,10,(6)	6	0.270	216(−)	7,1,4,6,11,(2)
7	0.177	231(−F)	7,11,1,5,10,6,(4)	7	0.200	189(−)	7,1,4,6,11,2,(10)
8	0.138	91	7,11,1,5,10,6,4,(8)	8	0.170	132(−)	7,1,4,6,11,2,10,(5)
9	0.121	40(−)	7,11,1,5,10,6,4,8,(2)	9	0.099	40(−)	7,1,4,6,11,2,10,5,(8)
10	0.075	11(−)	7,11,1,5,10,6,4,8,2,(9)	10	0.082	2(+)	7,1,4,6,11,2,10,5,8,(3)
11	0.061		7,11,1,5,10,6,4,8,2,9,(3)	11	0.037		7,1,4,6,11,2,10,5,8,3,(9)
Computing time: 0.53 s				Computing time: 0.70 s			
Num.	Conditional Entropy (2)	Rank	Method B	Num.	Conditional Entropy (2)	Rank	Method E
1	1.229	7	7	1	1.143	7(−)	7
2	1.102	9	7,(3)	2	1.025	32(−)	7,(9)
3	1.390	53	7,3,(1)	3	1.393	138(−)	7,9,(5)
4	1.806	47	7,3,1,(4)	4	1.784	124(−)	7,9,5,(8)
5	2.143	255	7,3,1,4,(8)	5	1.959	40(+)	7,9,5,8,(3)
6	2.321	27	7,3,1,4,8,(5)	6	2.171	248(−)	7,9,5,8,3,(2)
7	2.586	4	7,3,1,4,8,5,(9)	7	2.440	188(−)	7,9,5,8,3,2,(10)
8	2.677	15	7,3,1,4,8,5,9,(10)	8	2.646	93(−)	7,9,5,8,3,2,10,(4)
9	2.747	24(−)	7,3,1,4,8,5,9,10,(2)	9	2.801	30	7,9,5,8,3,2,10,4,(6)
10	2.838	3(−)	7,3,1,4,8,5,9,10,2,(6)	10	2.992	3(+)	7,9,5,8,3,2,10,4,6,(1)
11	2.969		7,3,1,4,8,5,9,10,2,6,(11)	11	3.317		7,9,5,8,3,2,10,4,6,1,(11)
Computing time: 0.47 s				Computing time: 0.40 s			
Num.	Joint Entropy	Rank	Method C	Num.	Joint Entropy	Rank	Method F
1	1.229	7	7	1	1.143	7(−)	7
2	0.469	54(−)	6,7	2	0.309	54(−)	6,7
3	0.962	165(−)	5,6,7	3	0.618	76(+)	6,7,11
4	1.531	330	2,5,6,7	4	1.071	96(+)	1,6,7,11
5	2.101	462(−)	2,5,6,7,9	5	1.568	172(+)	1,4,6,7,11
6	2.744	462(−)	2,5,6,7,9,10	6	2.073	204(+)	1,4,6,7,10,11
7	3.438	324	2,5,6,7,9,10,11	7	2.623	189	1,2,4,6,7,10,11
8	4.151	165	2,4,5,6,7,9,10,11	8	3.141	30(+)	1,2,3,4,6,7,10,11
9	4.712	53	2,4,5,6,7,8,9,10,11	9	3.646	29(+)	1,2,3,4,6,7,9,10,11
10	5.336	11	1,2,4,5,6,7,8,9,10,11	10	4.154	6(+)	1,2,3,4,5,6,7,9,10,11
11	5.850		1,2,3,4,5,6,7,8,9,10,11	11	4.624		1,2,3,4,5,6,7,8,9,10,11
Computing time: 19.20 s				Computing time: 18.75 s			

Table A6. Results of a select optimal rain gauge network of methods (weight: precipitation by 1-station).

Num.	Conditional Entropy (1)	Rank	Method A	Num.	Conditional Entropy (1)	Rank	Method D
1	2.363	7	7	1	2.700	6	4
2	1.580	3	7,(11)	2	1.810	7	4,(7)
3	0.893	19	7,11,(1)	3	0.982	89(−)	4,7,(10)
4	0.490	64	7,11,1,(5)	4	0.554	98	4,7,10,(1)
5	0.246	125	7,11,1,5,(4)	5	0.374	241	4,7,10,1,(8)
6	0.194	94	7,11,1,5,4,(8)	6	0.283	363(−)	4,7,10,1,8,(11)
7	0.172	176	7,11,1,5,4,8,(6)	7	0.233	106	4,7,10,1,8,11,(5)
8	0.123	91	7,11,1,5,4,8,6,(10)	8	0.131	98	4,7,10,1,8,11,5,(9)
9	0.099	14	7,11,1,5,4,8,6,10,(3)	9	0.083	50(−)	4,7,10,1,8,11,5,9,(2)
10	0.083	2	7,11,1,5,4,8,6,10,3,(2)	10	0.064	11	4,7,10,1,8,11,5,9,2,(6)
11	0.048		7,11,1,5,4,8,6,10,3,2,(9)	11	0.032		4,7,10,1,8,11,5,9,2,6,(3)
Computing time: 1.05 s				Computing time: 1.81 s			
Num.	Conditional Entropy (2)	Rank	Method B	Num.	Conditional Entropy (2)	Rank	Method E
1	2.363	7	7	1	2.700	6	4
2	1.971	9	7,(3)	2	2.081	51(−)	4,(3)
3	2.552	53	7,3,(1)	3	2.877	21	4,3,(5)
4	3.074	47	7,3,1,(4)	4	3.533	87(−)	4,3,5,(6)
5	3.450	255	7,3,1,4,(8)	5	3.722	43(+)	4,3,5,6,(8)
6	3.620	27	7,3,1,4,8,(5)	6	4.074	12	4,3,5,6,8,(9)
7	3.871	14(−)	7,3,1,4,8,5,(10)	7	4.234	54	4,3,5,6,8,9,(2)
8	3.991	15	7,3,1,4,8,5,10,(9)	8	4.334	57	4,3,5,6,8,9,2,(7)
9	4.032	1	7,3,1,4,8,5,10,9,(6)	9	4.560	7(+)	4,3,5,6,8,9,2,7,(1)
10	4.009	1	7,3,1,4,8,5,10,9,6,(11)	10	4.525	5(+)	4,3,5,6,8,9,2,7,1,(11)
11	4.098		7,3,1,4,8,5,10,9,6,11,(2)	11	4.621		4,3,5,6,8,9,2,7,1,11,(10)
Computing time: 0.90 s				Computing time: 1.08 s			
Num.	Joint Entropy	Rank	Method C	Num.	Joint Entropy	Rank	Method F
1	2.363	7	7	1	2.700	6	4
2	0.894	50	4,11	2	0.960	50	4,11
3	2.273	156	2,6,7	3	2.563	161	4,10,11
4	3.901	330	2,5,6,7	4	4.448	295(−)	2,4,10,11
5	5.658	436	2,7,9,10,11	5	6.525	419(−)	2,4,7,10,11
6	7.472	423	2,6,7,9,10,11	6	8.657	340	1,2,4,7,10,11
7	9.305	323(+)	2,4,6,7,9,10,11	7	10.796	308(−)	1,2,4,7,9,10,11
8	11.151	165	2,4,5,6,7,9,10,11	8	12.924	162(−)	1,2,4,7,8,9,10,11
9	12.746	53	2,4,5,6,7,8,9,10,11	9	14.995	50(−)	1,2,4,5,7,8,9,10,11
10	14.420	11	1,2,4,5,6,7,8,9,10,11	10	16.981	11	1,2,4,5,6,7,8,9,10,11
11	15.907		1,2,3,4,5,6,7,8,9,10,11	11	18.924		1,2,3,4,5,6,7,8,9,10,11
Computing time: 26.61 s				Computing time: 32.86 s			

Table A7. Results of a select optimal rain gauge network of methods (weight: monthly average precipitation by station).

Num.	Conditional Entropy (1)	Rank	Method A	Num.	Conditional Entropy (1)	Rank	Method D
1	4.653	7	7	1	5.072	6	4
2	1.160	3	7,(11)	2	1.435	7	4,(7)
3	0.588	19	7,11,(1)	3	0.554	17	4,7,(1)
4	0.281	64	7,11,1,(5)	4	0.291	100(−)	4,7,1,(8)
5	0.130	84(+)	7,11,1,5,(8)	5	0.212	45(+)	4,7,1,8,(5)
6	0.111	94	7,11,1,5,8,(4)	6	0.149	54	4,7,1,8,5,(10)
7	0.098	176	7,11,1,5,8,4,(6)	7	0.101	106	4,7,1,8,5,10,(11)
8	0.067	91	7,11,1,5,8,4,6,(10)	8	0.061	98	4,7,1,8,5,10,11,(9)
9	0.055	14	7,11,1,5,8,4,6,10,(3)	9	0.033	39	4,7,1,8,5,10,11,9,(6)
10	0.041	2	7,11,1,5,8,4,6,10,3,(2)	10	0.016	11	4,7,1,8,5,10,11,9,6,(2)
11	0.031		7,11,1,5,8,4,6,10,3,2,(9)	11	0.008		4,7,1,8,5,10,11,9,6,2,(3)
Computing time: 0.87 s				Computing time: 0.90 s			
Num.	Conditional Entropy (2)	Rank	Method B	Num.	Conditional Entropy (2)	Rank	Method E
1	4.653	7	7	1	5.072	6	4
2	1.633	9	7,(3)	2	1.684	51(−)	4,(3)
3	1.867	53	7,3,(1)	3	2.113	21	4,3,(5)
4	2.240	47	7,3,1,(4)	4	2.419	42	4,3,5,(8)
5	2.398	255	7,3,1,4,(8)	5	2.686	59	4,3,5,8,(9)
6	2.448	27	7,3,1,4,8,(5)	6	2.784	12	4,3,5,8,9,(6)
7	2.643	14(−)	7,3,1,4,8,5,(10)	7	2.807	54	4,3,5,8,9,6,(2)
8	2.698	15	7,3,1,4,8,5,10,(9)	8	2.665	65(−)	4,3,5,8,9,6,2,(10)
9	2.666	1	7,3,1,4,8,5,10,9,(6)	9	2.732	11(+)	4,3,5,8,9,6,2,10,(1)
10	2.683	3(−)	7,3,1,4,8,5,10,9,6,(2)	10	2.830	7(+)	4,3,5,8,9,6,2,10,1,(11)
11	2.675		7,3,1,4,8,5,10,9,6,2,(11)	11	2.881		4,3,5,8,9,6,2,10,1,11,(7)
Computing time: 0.95 s				Computing time: 0.59 s			
Num.	Joint Entropy	Rank	Method C	Num.	Joint Entropy	Rank	Method F
1	4.653	7	7	1	5.072	6	4
2	3.548	54(−)	6,7	2	3.754	50	4,11
3	7.610	156	2,6,7	3	8.180	161	4,10,11
4	11.785	330	2,5,6,7	4	12.842	275	4,7,10,11
5	16.003	413(+)	2,5,6,7,11	5	17.613	272	1,4,7,10,11
6	20.265	437(−)	2,5,6,7,9,11	6	22.281	340	1,2,4,7,10,11
7	24.512	324	2,5,6,7,9,10,11	7	26.882	189	1,2,4,6,7,10,11
8	28.695	165	2,4,5,6,7,9,10,11	8	31.459	144	1,2,4,6,7,9,10,11
9	32.736	53	2,4,5,6,7,8,9,10,11	9	36.046	49	1,2,4,6,7,8,9,10,11
10	36.834	11	1,2,4,5,6,7,8,9,10,11	10	40.598	11	1,2,4,5,6,7,8,9,10,11
11	40.731		1,2,3,4,5,6,7,8,9,10,11	11	45.041		1,2,3,4,5,6,7,8,9,10,11
Computing time: 25.30 s				Computing time: 25.14 s			

References

1. Lee, M.; Hong, J.H.; Kim, K.Y. Estimating damage costs from natural disasters in Korea. *Nat. Hazards Rev.* **2017**, *18*, 04017016. [CrossRef]
2. Lee, J.H.; Kim, C.J. Derivation of drought severity-duration-frequency curves using drought frequency analysis. *J. Korea Water Resour. Assoc.* **2011**, *44*, 889–902. [CrossRef]
3. Amorocho, J.; Espildora, B. Entropy in the assessment of uncertainty in hydrologic systems and models. *Water Resour. Res.* **1973**, *9*, 1511–1522. [CrossRef]
4. Chapman, T.G. Entropy as a measure of hydrologic data uncertainty and model performance. *J. Hydrol.* **1986**, *85*, 111–126. [CrossRef]
5. Harmancioglu, N.; Yevjevich, V. Transfer of hydrologic information among river points. *J. Hydrol.* **1987**, *91*, 103–118. [CrossRef]
6. Yang, Y.; Burn, D.H. An entropy approach to data collection network design. *J. Hydrol.* **1994**, *157*, 307–324. [CrossRef]
7. Singh, V.P. The use of entropy in hydrology and water resources. *Hydrol. Process.* **1997**, *11*, 587–626. [CrossRef]
8. Kawachi, T.; Maruyama, T.; Singh, V.P. Rainfall entropy for delineation of water resources zones in Japan. *J. Hydrol.* **2001**, *246*, 36–44. [CrossRef]
9. Krstanovic, P.F.; Singh, V.P. A univariate model for long-term streamflow forecasting. *Stoch. Hydrol. Hydraul.* **1991**, *5*, 189–205. [CrossRef]
10. Montesarchio, V.; Ridolfi, E.; Russo, F.; Napolitano, F. Rainfall threshold definition using an entropy decision approach and radar data. *Nat. Hazards Earth Syst. Sci.* **2011**, *7*, 2061–2074. [CrossRef]
11. Mogheir, Y.; Singh, V.P.; De Lima, J.L.M.P. Spatial assessment and redesign of a groundwater quality monitoring network using entropy theory, Gaza strip, Palestine. *Hydrogeol. J.* **2006**, *14*, 700–712. [CrossRef]
12. Mishra, A.K.; Coulibaly, P. Hydrometric network evaluation for Canadian watersheds. *J. Hydrol.* **2010**, *380*, 420–437. [CrossRef]
13. Alfonso, L.; Lobbrecht, A.; Price, R. Information theory-based approach for location of monitoring water level gauges in polders. *Water Resour. Res.* **2010**, *46*, 374–381. [CrossRef]
14. Li, C.; Singh, V.P.; Mishra, A.K. Entropy theory-based criterion for hydrometric network evaluation and design: Maximum information minimum redundancy. *Water Resour. Res.* **2012**, *48*. [CrossRef]
15. Caselton, W.F.; Husain, T. Hydrologic networks: Information transmission. *J. Water Resour. Plan. Manag. Div.* **1980**, *106*, 503–520.
16. Krstanovic, P.F.; Singh, V.P. Evaluation of rainfall networks using entropy: I. Theoretical development. *Water Resour. Manag.* **1992**, *6*, 279–293. [CrossRef]
17. Krstanovic, P.F.; Singh, V.P. Evaluation of rainfall networks using entropy: II. Application. *Water Resour. Manag.* **1992**, *6*, 295–314. [CrossRef]
18. Al-Zahrani, M.; Husain, T. An algorithm for designing a precipitation network in the south-western region of Saudi Arabia. *J. Hydrol.* **1998**, *205*, 205–216. [CrossRef]
19. Maruyama, T.; Kawachi, T. Evaluation of rainfall characteristics using entropy. *J. Rainwater Catchment Syst.* **1998**, *4*, 7–10. [CrossRef]
20. Bastin, G.; Lorent, B.; Duque, C.; Gevers, M. Optimal estimation of the average rainfall and optimal selection of rain gauge locations. *Water Resour. Res.* **1984**, *20*, 463–470. [CrossRef]
21. Rouhani, S. Variance reduction analysis. *Water Resour. Res.* **1985**, *21*, 837–846. [CrossRef]
22. Awadallah, A.G. Selecting optimum locations of rainfall stations using kriging and entropy. *Int. J. Civ. Environ. Eng.* **2012**, *12*, 36–41.
23. Chen, Y.C.; Wei, C.; Yeh, H.C. Rainfall network design using kriging and entropy. *Hydrol. Process. Int. J.* **2008**, *22*, 340–346. [CrossRef]
24. Xu, P.; Wang, D.; Singh, V.P.; Wang, Y.; Wu, J.; Wang, L.; Zou, X.; Liu, J.; Zou, Y.; He, R. A kriging and entropy-based approach to rain gauge network design. *Environ. Res.* **2018**, *161*, 61–75. [CrossRef]
25. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
26. Guiaşu, S. Weighted entropy. *Rep. Math. Phys.* **1971**, *2*, 165–179. [CrossRef]
27. Miller, G. Note on the bias of information estimates. Available online: <https://www.scienceopen.com/document?vid=357d299f-62fa-4bda-8dd2-e4d5b5abde5d> (accessed on 20 July 2020).

28. Hausser, J.; Strimmer, K. Entropy inference and the James-Stein estimator, with application to nonlinear gene association networks. *J. Mach. Learn. Res.* **2009**, *10*, 1469–1484.
29. Schürmann, T.; Grassberger, P. Entropy estimation of symbol sequences. *Chaos Interdiscip. J. Nonlinear Sci.* **1996**, *6*, 414–427. [[CrossRef](#)]
30. Meyer, P.E. Information-theoretic variable selection and network inference from microarray data. *Univ. Libre Brux.* **2008**, *2*, 261–274.
31. Krige, D.G. A statistical approach to some basic mine valuation problems on the Witwatersrand. *J. S. Afr. Inst. Min. Metall.* **1951**, *52*, 119–139.
32. Matheron, G. Principles of geostatistics. *Econ. Geol.* **1963**, *58*, 1246–1266. [[CrossRef](#)]
33. Pebesma, E.J. Multivariable geostatistics in S: The gstat package. *Comput. Geosci.* **2004**, *30*, 683–691. [[CrossRef](#)]
34. Lee, J.H.; Byun, H.; Kim, H.S.; Jun, H.D. Evaluation of a raingauge network considering the spatial distribution characteristics and entropy: A case study of Imha dam basin. *J. Korean Soc. Hazard. Mitig.* **2013**, *13*, 217–226. [[CrossRef](#)]
35. Son, C.; Lee, S. Changes in precipitation characteristics over inland of Yeongnam region. *J. Clim. Res.* **2011**, *6*, 121–142.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).