

Article

Spatiotemporal Scaling Effect on Rainfall Network Design Using Entropy

Chiang Wei ^{1,*}, Hui-Chung Yeh ² and Yen-Chang Chen ³

¹ The Experimental Forest, College of Bio-Resources and Agriculture, National Taiwan University, No.12 Chien-Shan Rd. Sec.1 Jushan Township, Nantou 55750, Taiwan

² Department of Natural Resources, Chinese Culture University, Taipei 11114, Taiwan;
E-Mail: hcyeh@faculty.pccu.edu.tw

³ Department of Civil Engineering, National Taipei University of Technology, Taipei 10608, Taiwan;
E-Mail: yenchen@ntut.edu.tw

* Author to whom correspondence should be addressed; E-Mail: d87622005@ntu.edu.tw;
Tel.: +886-49-2658412; Fax: +886-49-2631943.

Received: 28 January 2014; in revised form: 26 May 2014 / Accepted: 7 August 2014/

Published: 18 August 2014

Abstract: Because of high variation in mountainous areas, rainfall data at different spatiotemporal scales may yield potential uncertainty for network design. However, few studies focus on the scaling effect on both the spatial and the temporal scale. By calculating the maximum joint entropy of hourly typhoon events, monthly, six dry and wet months and annual rainfall between 1992 and 2012 for 1-, 3-, and 5-km grids, the relocated candidate rain gauges in the National Taiwan University Experimental Forest of Central Taiwan are prioritized. The results show: (1) the network exhibits different locations for first prioritized candidate rain gauges for different spatiotemporal scales; (2) the effect of spatial scales is insignificant compared to temporal scales; and (3) a smaller number and a lower percentage of required stations (PRS) reach stable joint entropy for a long duration at finer spatial scale. Prioritized candidate rain gauges provide key reference points for adjusting the network to capture more accurate information and minimize redundancy.

Keywords: scaling effect; rainfall network design; entropy

1. Introduction

The most crucial information required for planning, constructing, and operating hydraulic structures is rainfall data. The objective of a rainfall network is to design hydraulic structures efficiently and economically, according to the researched rainfall data [1]. However, because of topography, rain patterns, and effects of time, the spatial and temporal distributions of precipitation are uneven. Consequently, planning a suitable and optimal rain gauge station network is a challenging task. Research shows that even when two rain gauge stations are in close proximity to each other (5 km), the correlation coefficient of their precipitation time sequences may be lower than 0.5 [2]. This low correlation complicates the design and modulation of a rainfall network.

A reliable rainfall network can provide the immediate and precise precipitation data that is crucial to designing and placing hydraulic structures such as flood-prevention drainages. However, the number of rain gauge stations required may differ significantly from one watershed to another. In the past, stations were usually built without any standardization, according to the population density, assigned budget of each area, or the traffic feasibility to the stations. Hence, a problem of representative rainfall distribution and possible data redundancy may result. The World Meteorological Organization (WMO) [3] prescribes the minimum density of rainfall networks under various topographical and meteorological conditions, in particular, 25 km² per station on mountainous islands with irregular precipitation, one of which is Taiwan, the selected case in this study.

The optimization of a rainfall network has to satisfy two crucial criteria—spatial and temporal. Previous studies have mainly focused on the estimation of the location, distribution, and priority of rain gauge stations. However, determining the minimum number and optimal location of stations requires further investigation. Furthermore, the effect of spatiotemporal scaling on network design is yet to be analyzed. Applications to groundwater quality monitoring networks, stream gauge networks, and water distribution networks have increased in recent years. The methods used in network research related to entropy include least square methods and entropy [4], kriging [5], information entropy [6–22], and combined kriging and information entropy [23–25]. In particular, the information entropy approach has been widely adopted since the 1970s for hydrologic data collection network design and uncertainty evaluation [26–33].

The effects of heterogeneity and of scale and scaling continue to be significant issues in hydrologic research. Despite the progress that has been achieved in the past 20 years, considerable interest and much research remain centered on these issues [34]. Disaggregation-aggregation approaches were illustrated to link catchment-scale and point-scale state variables, which permit an empirical large-scale model that still retains some essence of small-scale physics [35]. The analysis of rainfall data from buoys on the ocean, the microwave imager (TMI) and precipitation radar (PR) on board the Tropical Rainfall Measuring Mission (TRMM) satellite shows that TMI and PR satellite data from $2.5^\circ \times 2.5^\circ$ areas centered on the buoys consistently agree well with the buoys on both annual and seasonal timescales, but the monthly point measurement of buoys exhibited large differences with TRMM data [36]. Singh [37] addressed the complete entropy theory and its applications in environmental and water engineering, and most importantly, several entropy measures, including directional information transfer index, total correlation, and maximum information minimum redundancy, were introduced. Nevertheless, few studies have been conducted to evaluate the possible spatiotemporal

scaling issue for network design. The proposed scheme for analyzing an optimal rainfall network involves kriging to generate the rainfall data of the candidate rain gauge stations, and entropy to evaluate the uncertainty in different combinations of spatial and temporal scales. In this study, we evaluate the spatiotemporal scaling effect on rain gauge network design and suggest optimal configuration. In particular, because records of several typhoon rainfall events, together with monthly, six dry and wet months, and annual rainfall, were analyzed, an optimal rain gauge network can be optimized by the hydrologic or climatic consideration, either short- or long-duration based. Once the optimal network design with maximum information and minimum redundancy is established, rainfall characteristics can be obtained to provide the key reference for the hydrologic planning for the watershed.

2. Methodology

This study involved the following steps: (1) determining the different combinations of spatial and temporal scales and delineating the study area; (2) applying kriging to existing rainfall data to generate the rainfall data of the candidate rain gauge stations for a certain combination; (3) determining the priority sequence of the candidate rain gauge stations and evaluating the minimum number required; and (4) summarizing the spatiotemporal scaling effect.

2.1. Spatiotemporal Scale

Stewart *et al.* [38] pointed out two distinct problems involved in scaling: (1) the requirement for a set of concepts that will allow the correct partitioning of the water balance at any given scale; and (2) the concepts that will allow information gathered at one scale to be used in making predictions at other scales. To solve these two problems, the possible scaling effect should be addressed first. In this study, spatial scales for 1×1 , 3×3 , and 5×5 km grids were partitioned to delineate a study area comprising 327.86 km². A total of 346, 45, and 20 grids were created, respectively. The center of each grid was assigned the location of a candidate rain gauge station. At each of the three different spatial scales, hourly records for typhoon events, monthly, six dry and wet months and annual rainfall were individually analyzed. Hourly data are used to investigate fluctuations of short duration for extreme events while the monthly, six dry and wet months and annual data depict the possible seasonal and annual trends or variations. Therefore, a total of fifteen combinations for different conditions are evaluated.

2.2. Kriging

Kriging is a geostatistical method for interpolating random spatial variations in rainfall data to estimate linear grid points [39–43]. In this study, we use the exponential model to fit the semi-variogram from the measurements of rainfall [24]:

$$\gamma(h) = b \left[1 - \exp\left(-\frac{h}{a}\right) \right] \quad (1)$$

where a denotes the range parameter, and b denotes the sill; that is, the critical variance in the spatial independence is as high as $3a$.

In particular, three independent short-term rainfall measurements from the 50 rain gauges were used to verify the kriging results.

2.3. Entropy

The information entropy introduced by Shannon [44,45] is based on probabilities. The entropy value is used to estimate uncertainties:

$$H(x) = -\sum_{i=1}^n p_i \log(p_i) \quad (2)$$

where $H(x)$ is the entropy value and p_i is probability.

Shannon's entropy [45] is a measure of information content, which depends on the current level of knowledge or uncertainty. Mathematically, the amount of information is inversely related to the probability of occurrence. The basic assumptions of the entropy are the amount of information, $I(p)$, being a real nonnegative measure, additive, and a continuous function of probability p . For the rational numbers, the function of $I(p)$ obeys the same formula as the log function. For any discrete probability distribution, Shannon's entropy is expressed as:

$$H(x_1) = -\sum_i p_i \ln p_i \quad (3)$$

where p_i is the probability of event x_i . Equation (3) refers only to the information state before receiving data. Thus, $H(x_1)$ measures the average amount of information. $H(x_1) = 0$ when the event is certain ($p_i = 0$ or 1) and there is no surprise. Because of the uniformity resulting in an inability to believe any outcome being more likely than any other, uniform distribution results correspond to maximum ignorance. Maximum entropy can be seen as a generalization of the classical principle of indifference and can be used to obtain unbiased probability assessments.

The rainfall information of two rain gauge stations may be overlapped. Therefore the rainfall information of the two rain gauge stations can become two variables x_1 and x_2 . Corresponding to Equation (3), the joint entropy of two variables is [24]:

$$H(x_1, x_2) = -\sum_i \sum_j p_{ij} \ln p_{ij} \quad (4)$$

For three variables of x_1 , x_2 , and x_3 , the joint entropy is [24]:

$$H(x_1, x_2, x_3) = -\sum_i \sum_j \sum_k p_{ijk} \ln p_{ijk} \quad (5)$$

where p_{ijk} is the joint probability of x_1 , x_2 , and x_3 .

When the x_1 rain gauge station is examined to record rainfall data, the remaining uncertainty of the x_2 rain gauge station will be exhibited by the conditional entropy. The probability of x_2 under the influence of x_1 's condition can be shown as below [24]:

$$p(x_2 | x_1) = p_{ji} = \frac{p_{ij}}{p_i} \quad (6)$$

Thus:

$$H(x_1, x_2) = H(x_2 | x_1) + H(x_1) \quad (7)$$

The joint entropy that can measure the amount of information of the joint events is derived by Equation (7) with conditional probability [24]:

$$H(x_2 | x_1) = H(x_1, x_2) - H(x_1) \quad (8)$$

where $H(x_2|x_1)$ is the conditional entropy of x_2 given x_1 . To find out the amount of mutual or overlapped information of the two stations, a transferable information calculation can be utilized to do so, as if using the x_1 rain gauge station to forecast information from the x_2 rain gauge station.

2.4. Optimization of Network Design

The significance of each rain gauge station in its network can be determined from its entropy value. The greater the value, the higher the uncertainty; thus, each station is prioritized according to the descending order of entropy values. After confirming the first station, which has the greatest entropy value, the rest are selected and added one at a time, according to the inferiority of the system and overlapped information. To minimize the system's uncertainty, the standardization of determining the second most important station in the sequence is set as [24]:

$$\text{Min}\{H(x_2) - H(x_2 | x_1)\} \quad (9)$$

Then, the n -th to be added is [24]:

$$\text{Min}\{H(x_1, x_2, \dots, x_{n-1}) - H[(x_n | x_1, x_2, \dots, x_{n-1})]\} \quad (10)$$

In the calculation, therefore, the greatest $H[(x_n | x_1, x_2, \dots, x_{n-1})]$ can be selected to arrange the order of data overlap of all the stations, in which the station with minimum overlap is the first to be added to the network, and the station with maximum overlap is the last to be added.

The sequence of prioritized stations determined by the entropy values can also be used as the sequence of station elimination. In each selection stage, the objective is to ascertain the maximum entropy value of each selected station. Stations are subsequently added according to the gradual increase in the joint entropies. However, the joint entropy does not increase sufficiently to show any differences when it reaches a certain number and attains a definite value. In other words, adding more stations has a very limited effect on the network system. The exponential model is applied to the correlation graph of the station numbers and $H[(x_n | x_1, x_2, \dots, x_{n-1})]$ to detect the critical data volume and the supposed number of stations. The coefficient k_m denotes the specific value of the m -th entropy value, as compared with all entropy values of the study area; it is assumed to be used as the reveal data volume of the m -th station. Assuming n stations in the study area and a number of basic stations have been selected, and the addition of new candidate rain gauge stations is prioritized on the basis of the entropy value, the definition of k_m in this study can be expressed as [24]:

$$k_m = \frac{H(x_1, x_2, \dots, x_m)}{H(x_1, x_2, \dots, x_m, \dots, x_{n-1}, x_n)}, \quad m < n \quad (11)$$

and $k_1, k_2, \dots, k_m, \dots, k_{n-1}, k_n < 1$. Hence, in determining the number of stations in a catchment area, a threshold value k_m^* must be set, and by setting a limit such as $k_m > k_m^*$, the number can be secured.

On the other hand, the threshold value is determined by the increasing efficiency revealed by the increment of k_m . In this study, k_m is set to 0.95, which is 95% of the information. Hence, if the number of rain gauge stations in the existing network is greater than that in the candidate network, those existing stations sequenced behind the candidate station are to be eliminated; otherwise, more stations are added.

In this study, there are fifteen combinations of different spatiotemporal scales. To evaluate the efficiency to the number k_m , we define the percentage of required stations reaching 95% information PRS (%) as:

$$PRS(\%) = \frac{n(k_m)}{n} \times 100\%, m < n \quad (12)$$

PRS is used to evaluate the efficiency of the sequence of prioritized stations. The number of prioritized stations may be the same at different combinations of spatiotemporal scales. However, lower PRS can be regarded as more efficient to reach the total 95% threshold value of measured uncertainty with fewer stations.

2.5. Study Area and Data

To illustrate and evaluate the proposed model, rainfall data within and near the National Taiwan University Experimental Forest (NTUEF) is used to demonstrate an optimal rainfall network. Located upstream on the Zhuoshui River, the geographical site lies between 23°48'49" and 23°28'10"N and 121°45'16" and 121°59'15"E, accounting for an approximate area of 327.86 km² (Figure 1). The catchment area extends from the Mt. Jade (elevation 3952 m, the highest peak in Taiwan) to Gueitsuto (220 m). Coursing through major forestland within the Chenyulan catchment, the river flows over 41.4 km, with an average slope of 2.7%. Most of the catchment is covered by mature forests, and the geological features of this region include a complex suite of rocks, such as granite, gneiss, schist, sandstone, conglomerate, and marl. Because of differences in altitude, the climate is divided into subtropical, warm temperate, cold temperate, subfrigid, and frigid zones. The mean annual temperature ranges between a low of 4 °C (Mt. Jade) and a high of 23 °C (Jushan). The NTUEF area usually experiences enhanced rainfall, and it receives an average (1992–2012) rainfall of 2408 mm; however, the rainfall is unevenly distributed, with more than 70% of it occurring between May and September. The annual rainfall increases from north to south and east to west, and consists of several centers (Figure 2a). During the period 1992 to 2012, the most severe rainfall event, Typhoon Morakot, which occurred 5 August to 9 August 2009, poured almost the entire average annual rainfall within a few days (Figure 2b). The largest rainfall recorded by the Alishan weather station (bottom left near the boundary in Figure 1) is 2884 mm, which caused a large-scale landslide and debris flow.

There are 50 rain gauge stations within or near the NTUEF territory; they are listed in Table 1. These stations are operated by the Central Weather Bureau and the NTUEF, and include the microclimate station and rain gauges set by the authors. Figure 1 also shows the locations of the 50 existing and candidate rain gauge station from 1 × 1, 3 × 3, and 5 × 5 km. For integrity, the rainfall data from 1992 to 2012 is used in this study. In addition, only hourly records of seven severe typhoon events during this period are

selected to evaluate the short-duration rainfall (Table 2). Three criteria were considered to select these seven typhoons (1) the most of typhoon period cover the study area; (2) its 24-h rainfall is approaching or over 600 mm; (3) recorded massive disaster such as landslide and debris flow occurred in the study area. The sample size for hourly, monthly, and annual rainfall (dry and six wet months) is 385, 252, and 21, respectively.

Figure 1. Candidate and existing rain gauges in the study area NTUEF.

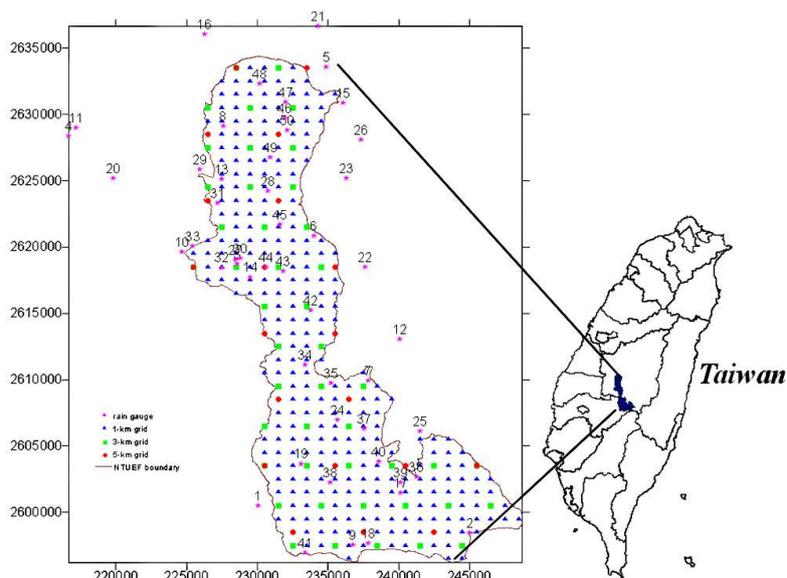


Figure 2. Contour maps of (a) average annual rainfall between 1992 and 2012; and (b) Typhoon Morakot rainfall of 5–9 August 2009 at the NTUEF area.

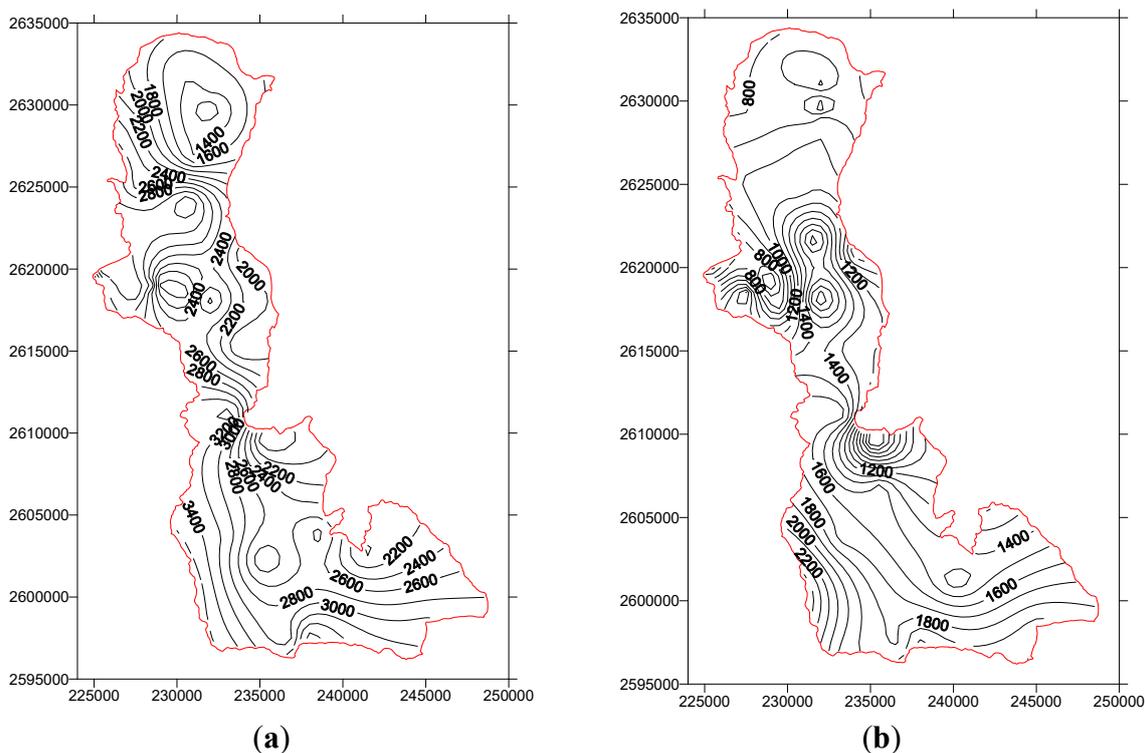


Table 1. Summary description of rain gauge stations in Study Area.

No.	Rain Gauge Station	Elevation (m)	TM2 (m)		Hourly Rainfall for Typhoon Events (mm)				Monthly Rainfall (mm)				Annual Rainfall (mm)			
			Easting	Northing	Maximum	Minimum	Mean	Standard Deviation	Maximum	Minimum	Mean	Standard Deviation	Maximum	Minimum	Mean	Standard Deviation
1	AliShan	2413	230,043	2,600,476	123.0	0.0	25.8	25.6	3346.0	0.0	336.6	438.1	5886.7	2196.5	4039.2	1140.9
2	Mt. Jade	3845	245,030	2,598,435	64.0	0.0	14.7	13.4	2189.9	0.0	254.8	299.0	4705.2	1702.7	3058.2	830.3
3	Xitou Nursery	1169	228,583	2,618,722	110.0	0.0	17.0	20.6	1770.0	0.0	202.3	246.1	4053.0	1291.0	2455.3	673.5
4	Jushan-NTU	156	216,693	2,628,383	145.0	0.0	10.2	18.1	1173.0	0.0	181.0	214.8	2821.6	1355.1	2221.0	449.4
5	Shueli-NTU	295	234,893	2,633,571	123.5	0.0	10.5	17.2	1512.5	0.0	150.7	200.4	2816.0	212.5	1835.9	724.6
6	Nemoupu-NTU	509	233,987	2,620,868	125.5	0.0	11.2	15.8	1008.0	0.0	153.3	175.4	2805.0	946.0	1820.9	491.0
7	Heshe-NTU	780	237,830	2,609,920	74.0	0.0	11.2	14.3	1258.0	0.0	154.2	185.0	2688.5	1062.0	1855.9	498.8
8	Chinshueigao-NTU	520	227,576	2,629,098	100.0	0.0	9.2	14.5	1271.6	0.0	187.6	222.1	4275.0	680.5	2234.6	852.4
9	Hsingouko	2540	236,749	2,597,543	112.5	0.0	16.8	15.9	2203.0	0.0	241.7	294.0	4524.5	787.0	2828.7	1068.2
10	Dann	1528	224,672	2,619,646	75.5	0.0	8.6	11.2	945.0	0.0	180.7	197.9	3154.0	773.0	2088.8	627.1
11	Jushan	151	217,157	2,629,012	170.0	0.0	8.9	16.9	1133.5	0.0	177.6	207.3	3205.0	613.0	2047.8	611.7
12	Wanshian	2403	240,080	2,613,075	85.0	0.0	12.8	15.0	1633.5	0.0	208.3	247.9	3642.0	924.0	2421.6	832.5
13	Phoenix Garden	878	227,485	2,625,117	141.0	0.0	11.9	18.4	1292.0	0.0	218.2	235.6	3671.0	948.0	2522.5	741.5
14	Xitou Observation	1771	229,514	2,617,731	61.0	0.0	9.2	9.6	1053.5	0.0	192.8	203.2	3139.0	909.5	2219.5	629.0
15	Long-Shen Bridge	339	236,100	2,630,858	130.5	0.0	9.0	15.0	900.0	0.0	164.5	179.4	2812.5	1133.5	1921.2	653.3
16	Ji-Ji	235	226,257	2,636,039	103.5	0.0	8.3	13.2	975.0	0.0	188.7	210.0	3100.5	1504.5	2256.7	923.4
17	GuanShan	1780	240,135	2,601,472	81.5	0.0	14.8	15.4	1171.5	0.0	227.7	225.2	3695.9	1296.5	2444.9	820.9
18	Pasture	2677	237,860	2,597,660	136.0	0.0	18.3	17.4	2383.5	0.0	304.1	387.7	5218.8	1653.5	3719.5	1025.3
19	Shenmu Village	1595	233,125	2,603,668	91.5	0.0	16.4	17.0	2141.5	0.0	260.1	330.5	4649.5	1653.5	3114.4	1664.2
20	Chungshinlun	661	219,839	2,625,192	63.5	0.0	9.4	12.6	1075.0	0.0	231.0	264.0	3682.0	1554.5	2731.8	1468.3
21	Shueli	593	234,295	2,636,644	110.0	0.0	10.2	17.5	911.0	0.0	193.9	218.3	3094.0	1451.0	2341.8	1276.9
22	Fongchiou	1151	237,647	2,618,491	84.5	0.0	11.1	14.3	1211.0	0.0	166.9	210.1	2938.0	1088.0	2021.5	1114.5
23	ShangAn	781	236,321	2,625,167	66.0	0.0	8.6	12.2	804.5	0.0	162.0	190.1	2914.0	1193.0	1973.3	1074.3
24	Hsin-shin Bridge	897	235,680	2,606,957	96.5	0.0	14.2	17.2	1751.5	0.0	193.9	266.1	3277.5	1291.0	2425.1	1297.9
25	Dongpu	887	241,493	2,606,091	67.0	0.0	10.2	11.8	1307.0	0.0	169.6	219.2	2917.0	1107.0	2092.9	1138.7
26	Siluang	1001	237,315	2,628,058	78.5	0.0	10.7	15.5	963.5	0.0	186.7	217.3	3061.0	1313.5	2193.8	1218.5

Table 1. Cont.

No.	Rain Gauge Station	Elevation (m)	TM2 (m)		Hourly Rainfall for Typhoon Events (mm)				Monthly Rainfall (mm)				Annual Rainfall (mm)			
			Easting	Northing	Maximum	Minimum	Mean	Standard Deviation	Maximum	Minimum	Mean	Standard Deviation	Maximum	Minimum	Mean	Standard Deviation
27	Xitou office	1156	228,453	2,619,028	56.0	0.0	13.5	12.9	1218.5	0.0	223.2	277.4	4005.5	1125.5	3053.3	1369.3
28	TienDi	787	230,728	2,624,199	53.5	0.0	11.3	12.6	1360.5	0.0	140.3	266.4	3122.5	137.0	2326.5	973.3
29	GuangHsin	645	225,917	2,625,831	49.5	0.0	9.8	12.2	1190.5	0.0	238.9	286.9	3628.5	329.0	3124.5	1299.3
30	No.3 Gully	1185	228,811	2,619,174	25.0	0.0	4.2	5.7	776.0	0.0	167.2	180.2	3339.5	712.5	1937.0	907.2
31	Neihu elementary school	772	227,181	2,623,316	52.0	0.0	9.8	11.7	1214.5	0.0	201.6	258.7	3560.5	901.0	3090.3	1157.6
32	Lower University Gully	1197	227,492	2,618,456	106.0	0.0	15.1	16.8	1663.0	0.5	255.6	402.7	3956.0	222.5	3334.2	1207.7
33	Wushio	1495	225,450	2,620,064	32.0	0.0	7.3	7.5	2296.5	0.0	212.4	394.3	3941.0	610.0	3400.8	1221.4
34	Yashanpin	1390	233,383	2,611,144	86.0	0.0	17.5	19.4	1872.5	0.0	269.3	367.1	4221.0	2264.5	3534.0	1478.1
35	Alibudon	1208	235,227	2,609,712	31.5	0.0	4.8	7.9	823.5	0.0	135.8	162.1	2876.0	530.5	1549.0	780.0
36	Salishian	1216	241,259	2,602,664	61.5	0.0	10.3	12.3	1211.5	0.0	208.0	274.3	3284.5	824.5	1941.2	799.8
37	Neuchangpin	1306	237,549	2,606,292	83.0	0.0	13.8	16.5	1500.0	0.0	185.3	247.9	2672.5	1768.5	2302.8	1008.6
38	Shenmu	1315	235,142	2,602,259	75.0	0.0	12.9	15.5	1490.5	0.5	297.5	413.5	3885.0	425.0	2185.3	955.8
39	32-compartment	1823	240,123	2,602,231	59.0	0.0	15.6	13.5	1714.5	0.0	223.5	298.9	3073.0	2050.5	2318.0	1157.2
40	30-compartment	2097	238,588	2,603,814	66.0	0.0	14.8	14.4	1725.0	0.0	227.2	341.3	4294.0	1134.5	2903.7	1244.7
41	29-compartment	2298	233,408	2,596,924	80.5	0.0	20.2	20.7	2307.5	0.0	347.0	466.6	5450.5	974.5	3453.3	1774.8
42	20-compartment	967	233,765	2,615,241	72.0	0.0	12.8	15.1	1372.5	15.0	174.8	282.1	2010.0	603.0	1456.5	573.6
43	21-compartment	1280	231,832	2,618,174	99.5	0.0	22.3	24.4	2243.5	7.0	296.3	433.4	2946.0	2048.5	2518.5	1023.4
44	22-compartment	892	230,636	2,618,475	79.0	0.0	13.7	17.1	1403.0	3.0	198.1	259.8	2859.5	1859.5	2129.6	877.4
45	24-compartment	1278	231,635	2,621,701	107.0	0.0	18.9	21.2	2042.0	0.0	171.1	339.8	2820.5	442.0	1582.8	797.0
46	13-compartment	454	231,953	2,629,686	51.5	0.0	7.9	11.0	708.5	8.5	178.2	193.9	2728.5	1003.5	1870.9	804.6
47	16-compartment	1002	232,038	2,630,932	71.0	0.0	10.6	14.9	303.0	2.5	114.4	203.5	1473.5	508.5	1058.4	458.3
48	17-compartment	454	230,194	2,632,283	51.0	0.0	8.9	10.4	343.0	0.0	166.6	259.3	1403.0	409.5	1110.8	300.3
49	11-compartment	1228	230,931	2,626,757	60.0	0.0	12.9	14.5	915.0	21.0	216.6	207.2	2998.0	1402.0	2219.9	929.7
50	9-compartment	1213	232,127	2,628,823	57.0	0.0	10.4	12.3	755.0	21.5	216.4	189.2	2391.5	1301.5	1839.6	762.0

Table 2. Typhoon events in this study between 1996 and 2012.

Typhoon	Date	Maximum Wind (m/s)	Rainfall Duration (h)	Damage (Billion, NT)
Herb	29 July–1 August 1996	53	44	39.3
Toraji	28–31 July 2001	38	24	14.7
Mindulle	28 June–3 July 2004	45	72	6.5
Kalmeigi	16–18 July 2008	33	32	3.4
Silaku	11–16 September 2008	51	75	5.6
Marakot	5–10 August 2009	40	96	47.7
Saola	30 July–3 August 2012	38	42	16.2

3. Result and Discussion

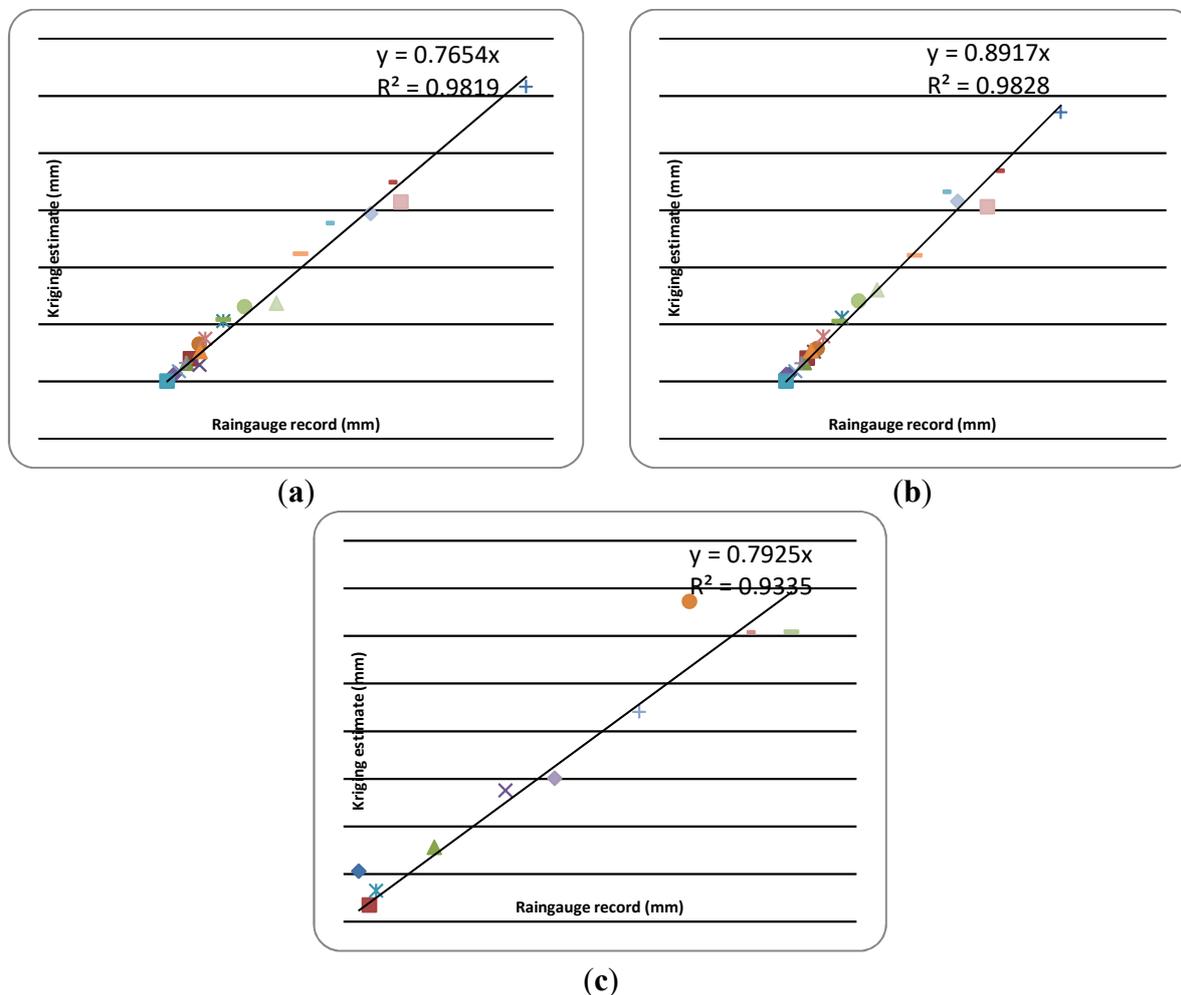
3.1. Validation of Kriging Estimates

Kriging was used in this study to estimate the rainfall at ungauged sites. However, how to validate the estimates using existing observation data becomes another important issue. The fitted parameters and kriging variance are listed in Table 3. From the basic statistical data in Table 1 and spatial distribution in Figure 2, the range of hourly, monthly, annual rainfall is quite large, the b sill parameter shows increase trend as the temporal scale enlarges; the a range parameter shows no significant variation except for six dry months. During six dry months, the rainfall regime dominated mainly by frontal rain lead to low rainfall intensity in winter and spring which yields wide influence range. Kriging variance in wet six months is far larger than that of dry six months, showing the significant rainfall amount distribution for dry (30%) and wet period (70%). Besides the 50 rain gauges, rainfall measurements of another three temporary rain gauge stations located in Xitou were used to validate the kriging estimates. By using the semi-variogram constructed from the 50 rain gauges, estimates of these three sites were then obtained and compared with actual point measurements. Monthly rainfall validations are shown in Figure 3. Although kriging results were underestimated by 11% to 24%, considering the high rainfall variability in mountainous areas, the correlation between estimates and measurements were high enough to be adjusted using simple regression.

Table 3. Details of kriging estimates.

Temporal Scale	b (Sill, mm ²)	a (Range Parameter, m)	Kriging Variance (mm ²)
Hour	165 ± 292	40,243 ± 25,538	21 ± 62
Month	23,529 ± 67,316	39,481 ± 67,316	3154 ± 7760
Dry six months	43,209 ± 52,813	50,926 ± 22,708	3198 ± 4184
Wet six months	583,324 ± 560,410	39,499 ± 27,171	64,250 ± 58,808
Annual	645,623 ± 654,175	31,337 ± 29,685	104,080 ± 107,613

Figure 3. Validation of monthly rainfall at three stations in Xitou Tract by Ordinary Kriging; (a) Phoenix 3.8 K Station (January 2004 to September 2005); (b) Liu Long Gully Station (January 2004 to September 2005); and (c) Upper Station of University Gully (December 2004 to September 2005).



3.2. Uncertainty Distributed in Space

The entropy was used in this study to analyze the uncertainty of rainfall for individual gauges. The result shows the spatial distribution differs from typhoon hourly, monthly, six dry and wet monthly and annual rainfall, as illustrated in Figure 4. For hourly data, the contour line is comparably smooth, and the entropy values increase from north (1.05) to south (1.50) and east (1.35) to west (1.95). For monthly rainfall, the contour line changes locally, and the values are higher than those of typhoon hourly values, especially in northern and eastern areas, which indicates higher uncertainty exists in temporal scales. As aforementioned, the rainy season in the study area ranges from May to September, resulting in uneven monthly distribution. In addition, the entropy contour line of annual rainfall is irregularly distributed. However, the pattern of the annual contour map is distinct from the other two, revealing the larger local variation even with small entropy values. Compared with Figure 2a, around the rainfall center in Figure 4c, the contour line is also comparably smooth. This indicates that the entropy value in surrounding areas with large rainfall is smaller. If the network design is based on the uncertainty, the priority in this area is not so important, and vice versa.

Figure 4. Entropy contour maps at (a) hourly; (b) monthly; (c) six dry monthly; (d) six wet monthly; and (e) annual scale.

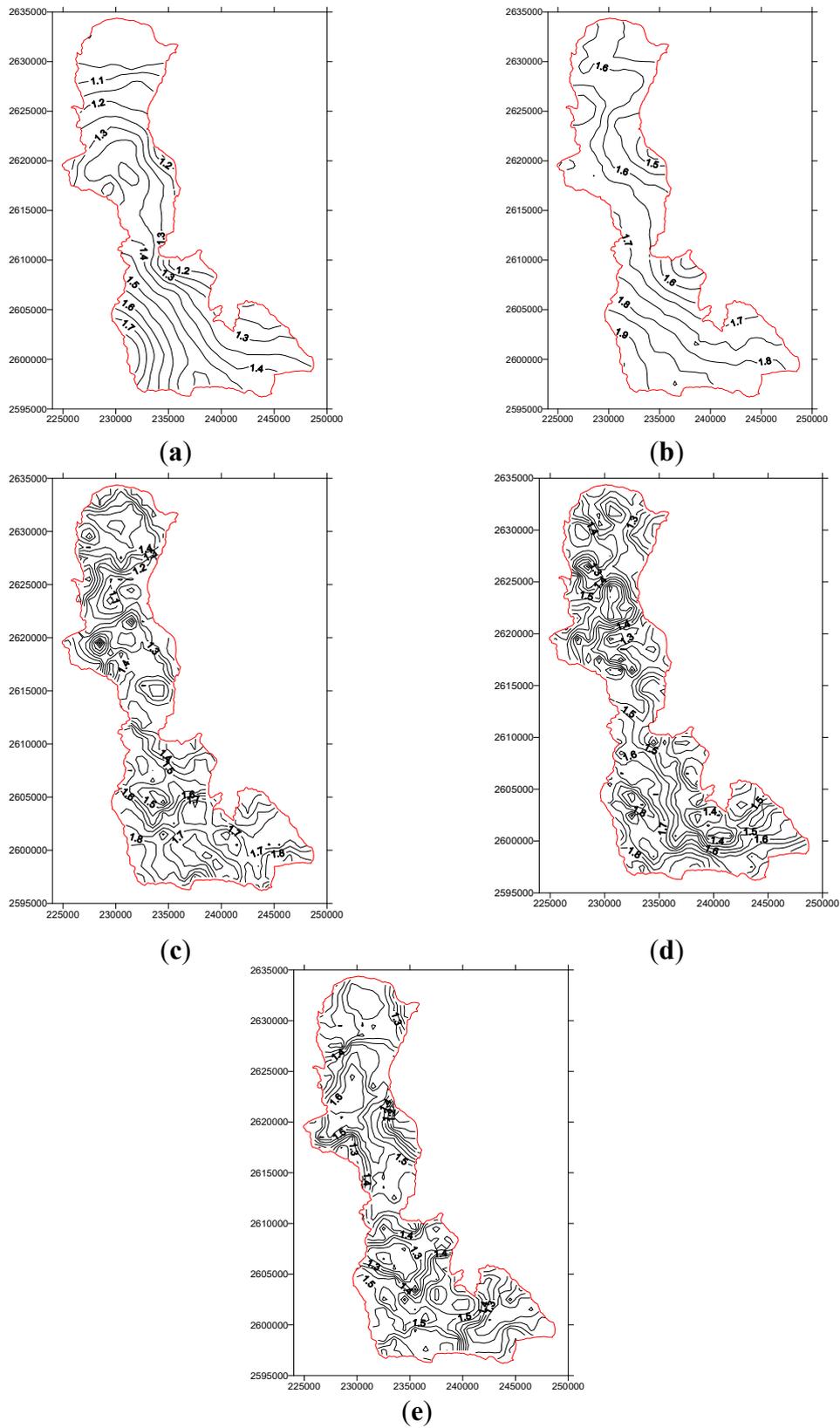
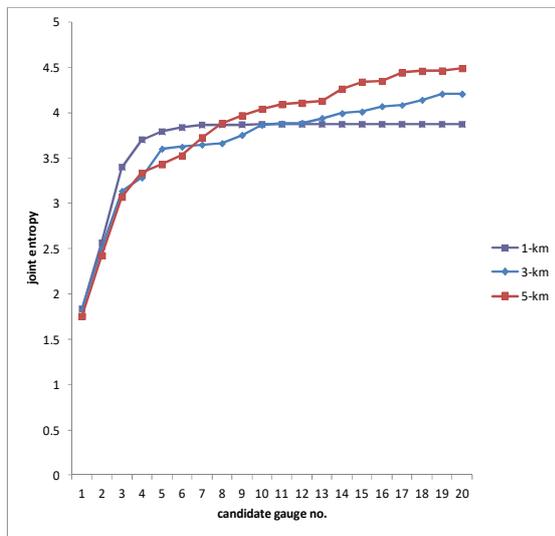
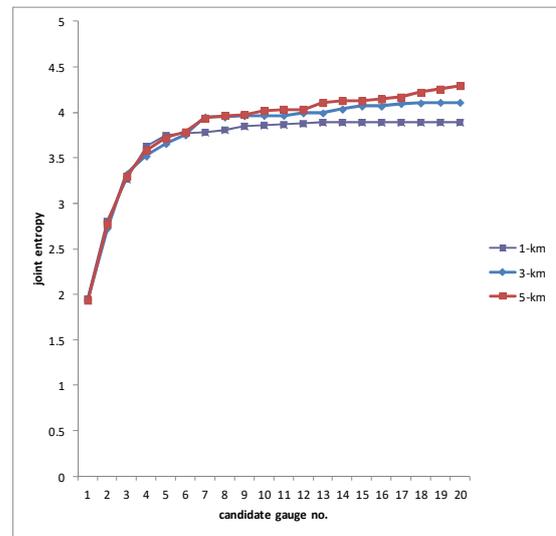


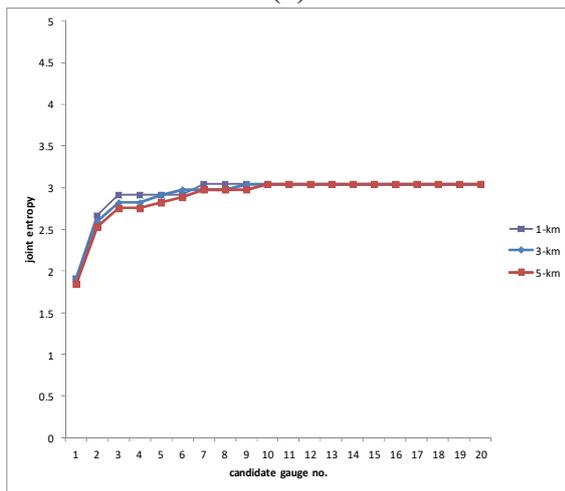
Figure 5. Variation of joint entropy for first 20 prioritized candidate gauges at (a) hourly, (b) monthly; (c) six dry monthly; (d) six wet monthly; and (e) annual scale.



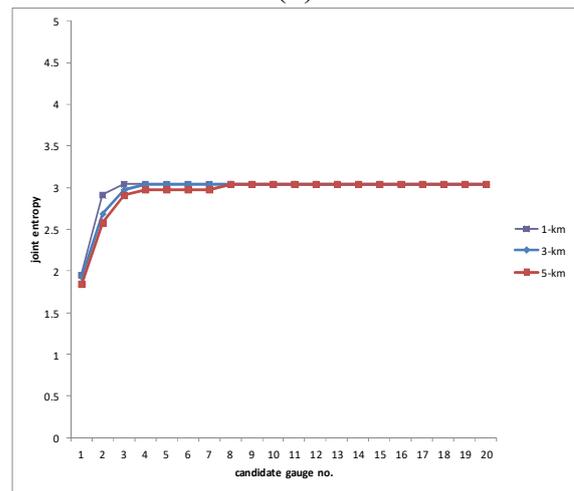
(a)



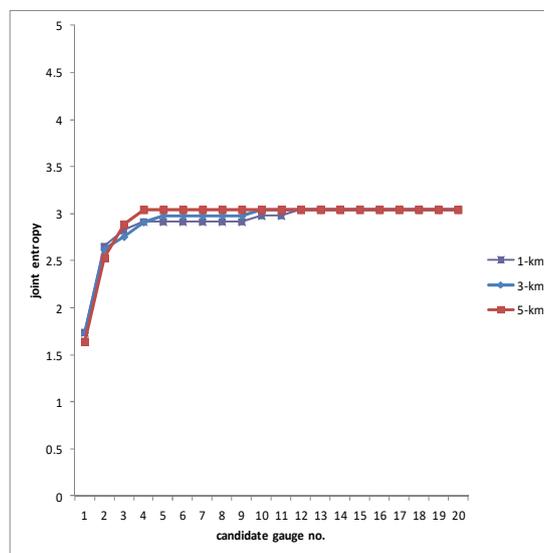
(b)



(c)



(d)



(e)

3.3 Spatial Scale Effect

If the same temporal scale was considered, the variation at the spatial scale is not so significant. The effect decreases as the temporal scale increases. Because of different candidate gauge stations for 1, 3 and 5-km, only the first 20 prioritized candidate rain gauge stations were compared for the spatial scale effect shown in Figure 5. For hourly data, the variation in the 5-km grid is larger than in the 3-km grid and in the 1-km grid. Before the eighth selected candidate rain gauge, the joint entropy value of 1-, 3-, and 5-km grid oscillates. However, the joint entropy of the 5-km grid is higher than that of 3-km grid and the 1-km grid after the eighth selected candidate rain gauges. The joint entropy of the 3-km grid is always higher than that of the 1-km grid. For the annual scale, no significant difference was found among the 1-, 3-, and 5-km grids. In general, for the first 20 prioritized candidate rain gauge stations, the maximum entropy of hourly, monthly, dry and wet six months and annual rainfall is around 4.5, 4.3, and 3.0, respectively. It implies the uncertainty of these two temporal scales are higher than the long-term scale and also implies that the spatial scale effect is less for long temporal scales and fewer rain gauges are required for long-term monitoring. In particular, the hourly data were collected during typhoon events, regarded as rainy and comparably large, and exhibiting diverse variability in high terrain relief. Within the 5-km grid, the joint entropy is larger because of more uncertainty existing in short-period rainfall. The rain gauge network is suggested to increase the number of stations to obtain more detailed variation for hydrologic design and rainfall forecasting.

3.4 Temporal Scale Effect

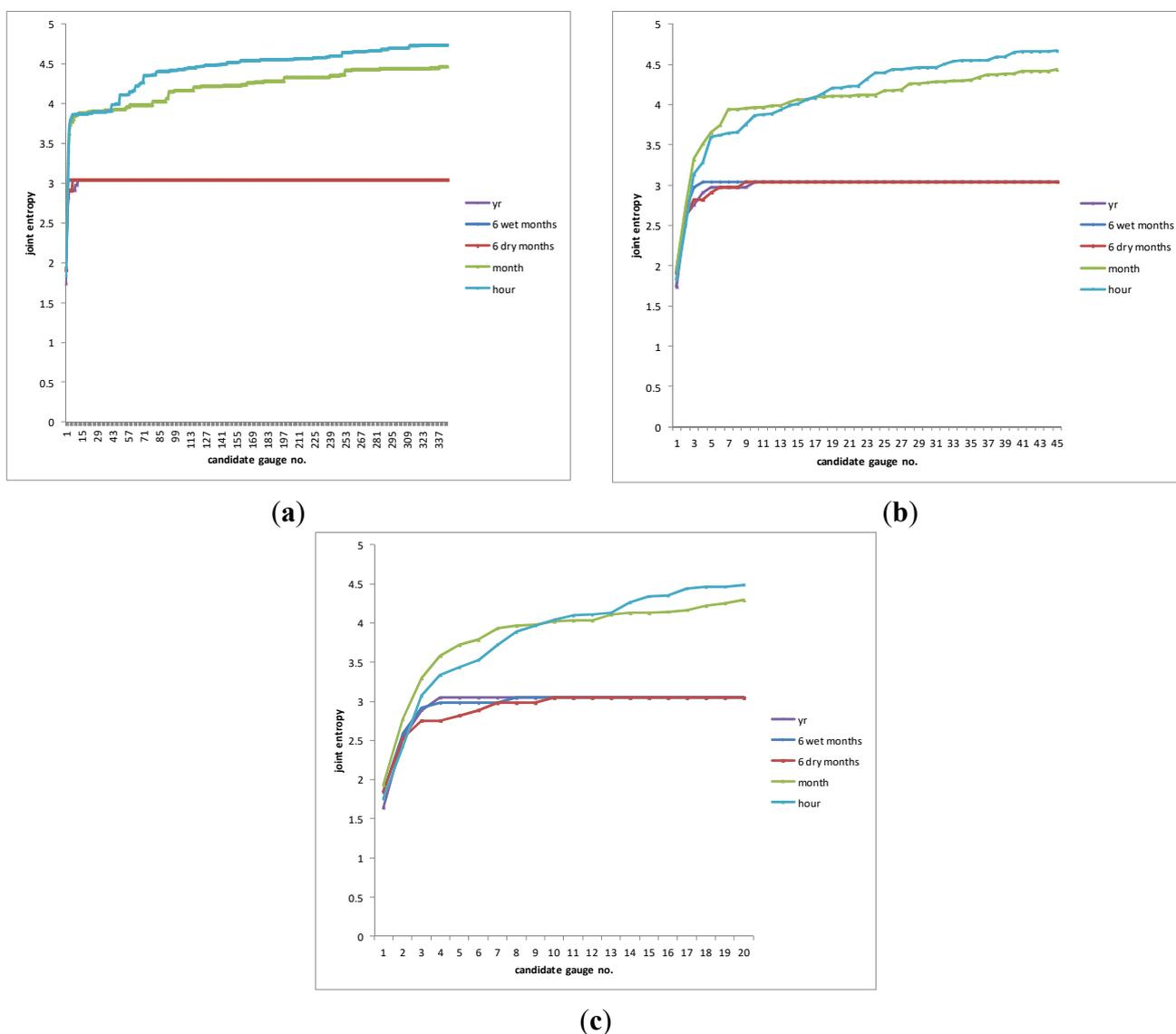
Based on the same spatial scale, more candidate rain gauge stations are needed to reach the stable value of joint entropy for a short temporal scale. In Figure 6, fewer gauge stations are required to reach the stable value of joint entropy. The trends coincide with each other between different spatial scales. In the 5-km grid, only 4 stations are needed to reach maximum joint entropy, but more than 300 gauges are needed to reach maximum joint entropy in the 1-km grid. The maximum joint entropy value did not change much between the 1-, 3-, and 5-km grids, but that of the annual rainfall, dry and wet six months (around 3) is separated from hourly and monthly rainfall. The spatial scaling effect is also proved again not to be as significant at the temporal scale.

3.5 Optimal Rain Gauge Station Network of the NTUEF Area

The priorities of the rain gauge stations are determined by calculating the joint entropies of the study area. The priority of the rain gauge stations obtained on the basis of the entropy can also become the sequence of removal of the stations. As aforementioned, different candidate rain gauge stations exist for different spatial scales. The number and PRS are calculated to express the percentage of required gauge stations at the three spatial scales listed in Table 4. Figure 7 illustrates the first ten prioritized gauge stations in 1-, 3-, and 5-km grids. Figure 7a demonstrates the first prioritized candidate station almost located at the southwestern corner with all five temporal scales at the 1-km grid. Hourly, monthly and six dry months temporal scale are located around rain gauge No. 1. The annual rainfall around this region is over 3600 mm and also with the largest rainfall in the study area. However, the first prioritized candidate station for six wet months and annual temporal scale is located at rain

gauge northern No.19 and eastern No.18, respectively. Three groups along the eastern boundary were clearly identified, implying that the existing rain gauge stations were crucial across temporal scales; however, no distinct groups were found in Figure 7b,c. More gauge stations are needed for these three concentrated groups. In addition, the second large prioritized candidate gauge stations were quite different across 1, 3, and 5 km. It can be inferred that the existing rain gauge stations not in the neighboring area of prioritized candidates should be addressed for the issue of stopping observation or abandonment. The decision for optimal rain gauge network can be made according to the prioritized and overlapped gauge stations across five temporal scales.

Figure 6. Variation of joint entropy vs. candidate gauge number for (a) 1-km scale; (b) 3-km scale; (c) 5-km scale.



If the number of stations in the network is greater than the minimum number of candidate stations, then the stations exceeding the minimum candidate number can be processed for elimination. In Table 4, PRS exceeding the threshold value of 95% across different temporal scales is smaller at finer spatial scales. Although the entropy value changes in Figure 4e, only three to four gauges are enough to

represent the variability of annual rainfall. The PRS of hourly, monthly, six dry and wet monthly and annual rainfall increases as the spatial scale enlarges. For the same spatial scale, the PRS for shorter temporal scale as hour and month is far larger than longer temporal scale; for the same temporal scale, the PRS increase as the spatial enlarges. It is noted the six rainy months need less candidate stations than six dry months which implies more variation and uncertainty of rainfall existed during dry seasons. According to the third category for the WMO standard aforementioned, the study area of 327.86 km² is equivalent to 13.1 gauges, very close to the number analyzed for hourly and monthly rainfall in the 5-km grid in Table 4. More rainfall information can be obtained as the required rain gauges increased for 3- and 1-km scale. However, for efficiency, 13 and 14 candidate stations at monthly and hourly at the 5-km grid, equivalent to one-fourth of existing rain gauge stations, are enough for general use, respectively; for hydrologic design and using the prioritized network at 3- or 1-km grid, the number will double or even more. Compromising the accuracy and network density, 13 candidate stations were identified as the optimal network according to the prioritized and overlapped gauge stations across all spatiotemporal scales in Figure 8.

Table 4. Number and Percentage of Required Stations (PRS) at different spatiotemporal scales.

Scale	Candidate Station Number	Hour	Month	Six Dry Months	Six Wet Months	Year
1-km	346	126 (36.4%)	143 (41.3%)	3(0.9%)	2(0.6%)	4 (1.1%)
3-km	45	26 (57.8%)	28 (62.2%)	5(11.1%)	3(6.7%)	4 (8.9%)
5-km	20	14 (70.0%)	13 (65.0%)	6(30%)	3(15%)	3 (15.0%)

Compared with Figure 7, these 13 candidate stations locate very closely with three concentrated groups found at 1-km scale. Kay and Kutiel [46] suggested a new approach in mapping climate maps of precipitation and found the actual rainfall field is more closely represented if more rainfall events and dense grid. In this study, the 1-km grid can capture more rainfall uncertainty than 3- and 5-km grid but with low PRS, indicating more candidate rain stations need to yield same accuracy. Kutiel and Kay [47] found no consistent recommendation of network design is best for all purposes. From Table 4 we demonstrate the PRS for fifteen combination of spatiotemporal scale, for best efficiency and low cost of rain gauge configuration, we should choose lowest PRS both at spatial and temporal scale, which means 1-km at six months or annual scale is best choice. However, it may only satisfy the evaluation or research for long-term climate and fail to capture need information for short-term such as hydrologic forecast.

Figure 7. First 10 prioritized candidate gauges in the study area at (a) 1-km scale; (b) 3-km scale; and (c) 5-km scale at different temporal scales.

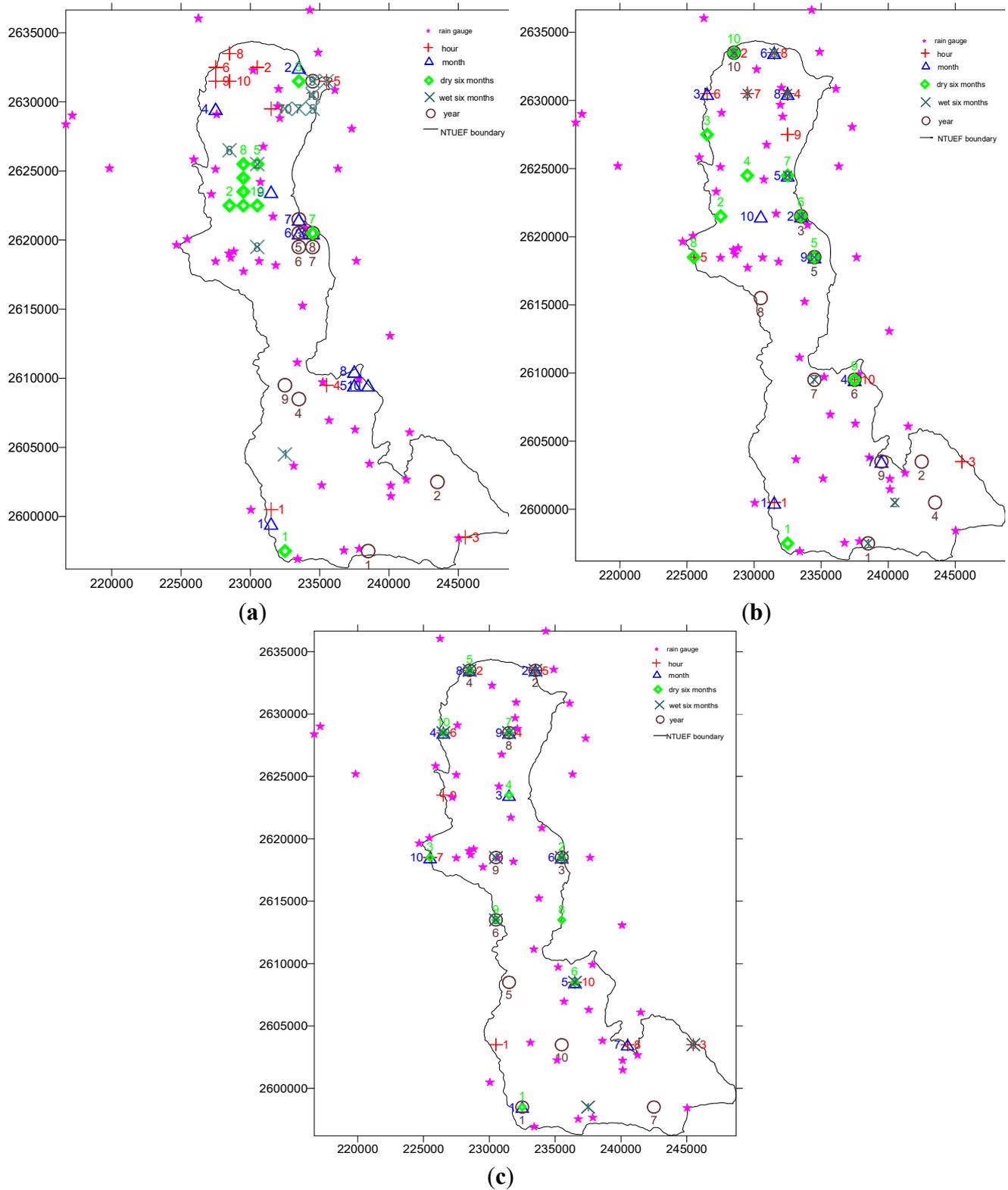
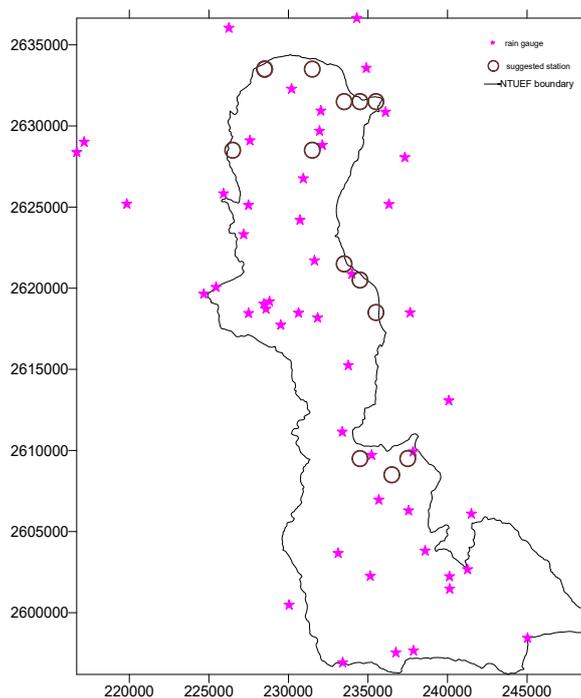


Figure 8. Locations of 13 suggested rain gauge candidates.

The authors did not analyze all the hourly rainfall records for three reasons. First, for hydrologic design and disaster warning and prevention, the records with “rain” are far more important than “no-rain.” Second, if all the hourly data are considered, the sample size will be larger than 183,000, which contains too many zeros or tiny rain records (e.g., 0.5 mm). The discrete distribution of data will cause the bias in calculating the entropy with respect to monthly and annual data in Equation (2). Third, the rainfall of typhoon events covered most of the study area for constructing semi-variogram in Equation (1), preventing inadequate semi-variogram resulting from rainfall only in some local areas. Despite above three reasons, hourly data analyzed in this study is still part of the whole dataset and just represent the network design for rainy hours. For No. 1 rain gauge (Alishan), there were only about half rainy days between 1992 and 2012 (total 7671 days). We do not include daily rainfall data for the same reasons even though it may be a suitable temporal scale between hour and month. Compared with the work by Cheng *et al.* [5] at hourly and annual scale, this study only includes hourly data for typhoons with the selecting criterion that over two thirds records are non-zero data, we neglected the other three major rainfall types, *i.e.*, convective, Mei-Yu and frontal rain. The major reason is the rain intensity of seven typhoon events in study area is large enough to represent the extreme rainfall condition in study area, in particular, the Typhoon Morakot in Aug 2009. From Table 1, 30% rain gauge stations (15 out of 50) its maximum hourly rainfall had reached 100 mm. The hourly spatial variability with annual rainfall is not significant different as the aforementioned study may lead to different area size 2200 km², rain gauge number 27, with respect to 328 km² and 50 in this study.

4. Conclusions

The proposed model can be used to design an optimal rainfall network that provides the majority of rainfall information as the existing rainfall network. Derived from the data of the existing rainfall network, it is used to analyze the spatiotemporal scaling effect and suggest an optimal candidate station

network by use of information entropy and kriging. In this study, we demonstrated that the candidate rainfall network is able to reduce the number of rain gauge stations, while accurately reflecting the location of precipitation. The results indicate that the optimal rainfall network at different combinations of spatiotemporal scales consists of candidate stations with capacities to provide 95% rainfall information of existed rain gauge stations. The relevant conclusions and propositions are as follows:

- (1) It exhibits different locations for first prioritized candidate rain gauges between spatiotemporal scales.
- (2) The effect of spatial scales is insignificant in comparison to temporal scales for network design. From the joint entropy value, the difference between hourly and monthly scales is more significant than the six dry, wet months and annual rainfall. However, the difference is significant across the spatial scale.
- (3) A smaller number and a lower percentage of required stations (PRS) are needed to reach stable joint entropy of long duration (six months or year) at finer spatial scale. Compromising the accuracy and network density, we suggest the optimal network design comprising of 13 candidate stations be suitable across all spatiotemporal scales.

Acknowledgments

This paper is based on the follow-on work supported by research project of the National Taiwan University Experimental Forest, Taiwan (NTUEF 2006-B06 and 2009-A05) and partial-financially support of Project no. MOST 103-2119-M-002-018- grant by Ministry of Science and Technology of Taiwan. In addition, special thanks go to the lab led by Hsin-Hsiung Chen for providing the rainfall data of three additional rain gauges.

Author Contributions

Chiang Wei was responsible for the research framework and managed the manuscript. Hui-Chung Yeh took charge of the code programming and data analysis. Yen-Chang Chen assisted with building the framework and conducting data analysis. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Hackett, O.M. National water data program. *J. Am. Water Work Assoc.* **1966**, *58*, 786–792.
2. Campbell, S.A. *Sampling and Analysis of Rain*; American Society of Testing and Materials: West Conshohocken, PA, USA, 1983.
3. World Meteorological Organisation (WMO). *Guide to Hydrometeorological Practices*; WMO Technical Paper 82; WMO: Geneva, Switzerland, 1970.

4. Markus, M.H.; Knapp, V.; Tasker, G.D. Entropy and generalized least square methods in assessment of the regional value of streamgauges. *J. Hydrol.* **2003**, *283*, 107–121.
5. Cheng, K.S.; Lin, Y.C.; Liou, J.J. Rain-gauge network evaluation and augmentation using geostatistics. *Hydrol. Process.* **2008**, *22*, 2554–2564.
6. Harmancioglu, N. Measuring the information content of hydrological processes by the entropy concept. *J. Civ. Eng. Facul. Ege Univ.* **1981**, 13–40.
7. Awumah, K.; Goulter, I. Assessment of reliability in water distribution networks using entropy based measures. *Stoch. Hydrol. Hydraul.* **1990**, *4*, 309–320.
8. Awumah, K.; Goulter, I.; Bhatt, S.K. Entropy-based redundancy measures in water-distribution networks. *J. Hydraul. Eng.* **1991**, *117*, 595–614.
9. Krstanovic, P.F.; Singh, V.P. Evaluation of rainfall network using entropy II: Application. *Water Resour. Manag.* **1992**, *6*, 295–314.
10. 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.
11. Ozkul, D.S.; Harmancioglu, N.B.; Singh, V.P. Entropy-based assessment of water quality monitoring networks. *J. Hydraul. Eng.* **2000**, *5*, 90–100.
12. Mogheir, Y.; Singh, V.P. Application of information theory to groundwater quality monitoring networks. *Water Resour. Manag.* **2002**, *16*, 37–49.
13. Mogheir, Y.; de Lima, J.L.M.P.; Singh, V.P. Characterizing the spatial variability of groundwater quality using the entropy theory: I. Synthetic data. *Hydrol. Process.* **2004**, *18*, 2165–2179.
14. Nunes, L.M.; Cunha, M.C.; Ribeiro, L. Groundwater Monitoring Network Optimization with Redundancy Reduction. *J. Water Resour. Plann. Manag.* **2004**, *130*, 33–43.
15. Masoumi, F.; Kerachian, R. Assessment of the groundwater salinity monitoring network of the Tehran region: Application of the discrete entropy theory. *Water Sci. Technol.* **2008**, *58*, 765–771.
16. Mogheir, Y.; Singh, V.P.; de Lima, J.L.M.P. Spatial assessment and redesign of a groundwater quality monitoring network quality using entropy theory, Gaza Strip, Palestine. *Hydrogeol. J.* **2006**, *14*, 700–712.
17. Yoo, C.; Jung, K.; Lee, J. Evaluation of rain gauge network using entropy theory: Comparison of mixed and continuous distribution function applications. *J. Hydrol. Eng.* **2008**, *13*, 226–235
18. Mogheir, Y.; de Lima, J.L.M.P.; Singh, V.P. Entropy and multi-objective based approach for groundwater quality monitoring network assessment and redesign. *Water Resour. Manag.* **2009**, *23*, 1603–1620.
19. Alfonso, L.; Lobbreche, A.; Price, R. Information theory-based approach for location of monitoring water level gauges in polder, *Water Resour. Res.* **2010**, *46*, W03528.
20. Alfonso, L.; Lobbrecht, A.; Price, R. Optimization of water level monitoring network in polder systems using information theory. *Water Resour. Res.* **2010**, *46*, W12553.
21. 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*, W5521.
22. Gong, W.; Gupta, H.V.; Yang, D.W.; Sricharan, K.; Hero, A.O. Estimating epistemic and aleatory uncertainty during hydrologic modeling: An information theoretic approach. *Water Resour. Res.* **2013**, *49*, 2253–2273.

23. Chen, Y.C.; Wei, C.; Yeh, H.C. Rainfall network design using kriging and entropy. *Hydrol. Process.* **2008**, *22*, 340–346.
24. Yeh, H.C.; Chen, Y.C.; Wei, C.; Chen, R.H. Entropy and Kriging Approach to Rainfall Network Design. *Paddy Water Environ.* **2011**, *9*, 343–355.
25. Awadallah, A.G. Selecting optimum locations of rainfall stations using kriging and entropy. *Int. J. Civil Environ. Eng.* **2012**, *12*, 36–41.
26. Amorocho, J.; Espildora, B. Entropy in the assessment of uncertainty in hydrologic systems and models. *Water Resour. Res.* **1973**, *9*, 1511–1522.
27. Chapman, T.G. Entropy as a measure of hydrologic data uncertainty and model performance. *J. Hydrol.* **1986**, *85*, 111–126.
28. Harmancioglu, N.; Yevjevich, V. Transfer of hydrologic information among river points. *J. Hydrol.* **1987**, *91*, 103–118.
29. Yang, Y.; Burn, D.H. An entropy approach to data collection network design. *J. Hydrol.* **1994**, *157*, 307–324.
30. Singh, V.P. The Use of Entropy in Hydrology and Water Resources. *Hydrol. Process.* **1997**, *11*, 587–626.
31. Kawachi, T.; Maruyama, T.; Singh, V.P. Rainfall entropy for delineation of water resources zones in Japan. *J. Hydrol.* **2001**, *246*, 36–44.
32. Mishra, A.K.; Coulibaly, P. Hydrometric network evaluation for Canadian watersheds. *J. Hydrol.* **2010**, *380*, 420–437.
33. Memarzadeh, M.; Mahjouri, N.; Kerachian, R. Evaluating sampling locations in river water quality monitoring networks: Application of dynamic factor analysis and discrete entropy theory. *Environ. Earth Sci.* **2013**, *70*, 2577–2585.
34. Sivapalan, M.; Grayson, R.; Woods, R. Preface Scale and scaling in hydrology. *Hydrol. Process.* **2004**, *18*, 1369–1371.
35. Viney, N.R.; Sivapala, M. A framework for scaling of hydrologic conceptualizations based on a disaggregation–aggregation approach. *Hydrol. Process.* **2004**, *18*, 1395–1408.
36. Serra, Y.L.; McPhaden, M.J. Multiple Time- and Space-Scale Comparisons of ATLAS Buoy Rain Gauge Measurements with TRMM Satellite Precipitation Measurements. *J. Appl. Meteorol.* **2003**, *42*, 1045–1059.
37. Singh, V.P. Entropy Theory and Its Application in Environmental and Water Engineering; John Wiley & Sons: Hoboken, NJ, USA, 2013.
38. Stewart, J.B.; Engman, E.T.; Feddes, R.A.; Kerr, Y. *Scaling up in Hydrology Using Remote Sensing*; John Wiley & Sons: Hoboken, NJ, USA, 1996.
39. Matheron, G. *Traite De Geostatistique Appliqué*, Tome 1; Editions Technip: Paris, France, 1962. (In French)
40. Rodriguez-Iturbe, I.; Sanabria, M.G.; Bras, R.L. A geomorphoclimatic theory of the instantaneous unit hydrograph. *Water Resour. Res.* **1982**, *18*, 877–886.
41. Chiles, J.-P.D.P. *Geostatistics-Modeling Spatial Uncertainty*; Willey: New York, NY, USA, 1999.
42. Wackernagel, H. *Multivariate Geostatistics*; Springer-Verlag: Berlin, Germany, 2003.
43. Kebaili, B.Z.; Chebbi, A. Comparison of two kriging interpolation methods applied to spatiotemporal rainfall. *J. Hydrol.* **2009**, *365*, 56–73.

44. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 623–656.
45. Shannon, C.E.; Weaver, W. *Mathematical Theory of Communication*; University of Illinois Press: Champaign, IL, USA, 1949.
46. Kay, P.A.; Kutiel, H. Some remarks on climatic maps of precipitation. *Clim. Res.* **1994**, *4*, 233–241.
47. Kutiel, H.; Kay, P.A. Effects of network design on climatic maps of precipitation. *Clim. Res.* **1996**, *7*, 1–10.

© 2014 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 license (<http://creativecommons.org/licenses/by/3.0/>).