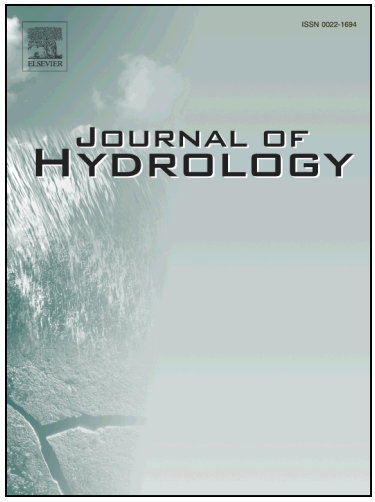


Research papers

Optimization of rainfall networks using information entropy and temporal variability analysis

Wenqi Wang, Dong Wang, Vijay P. Singh, Yuankun Wang, Jichun Wu, Lachun Wang, Xinqing Zou, Jiufu Liu, Ying Zou, Ruimin He

PII: S0022-1694(18)30086-6
DOI: https://doi.org/10.1016/j.jhydrol.2018.02.010
Reference: HYDROL 22559

To appear in: *Journal of Hydrology*

Received Date: 14 October 2017
Revised Date: 4 February 2018
Accepted Date: 6 February 2018

Please cite this article as: Wang, W., Wang, D., Singh, V.P., Wang, Y., Wu, J., Wang, L., Zou, X., Liu, J., Zou, Y., He, R., Optimization of rainfall networks using information entropy and temporal variability analysis, *Journal of Hydrology* (2018), doi: https://doi.org/10.1016/j.jhydrol.2018.02.010

Title: Optimization of rainfall networks using information entropy and temporal variability analysis

Authors: Wenqi Wang[1], Dong Wang[*1], Vijay P. Singh[2], Yuankun Wang[*1], Jichun Wu[1], Lachun Wang[3], Xinqing Zou[3], Jiufu Liu[4], Ying Zou[4], Ruimin He[5]

[1]Key Laboratory of Surficial Geochemistry, Ministry of Education, Department of Hydrosciences, School of Earth Sciences and Engineering, State Key Laboratory of Pollution Control and Resource Reuse, Nanjing University, Nanjing, P.R. China

[2]Department of Biological and Agricultural Engineering, Zachry Department of Civil Engineering, Texas A & M University, College Station, TX77843, USA

[3]School of Geographic and Oceanographic science, Nanjing University, Nanjing, P.R. China

[4]Nanjing Hydraulic Research Institute, Nanjing, P.R. China

[5]State Key Laboratory of Hydrology, Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing, P.R. China

(*Corresponding author: Dong Wang, wangdong@nju.edu.cn; Yuankun Wang, yuankunw@nju.edu.cn)

- **Key points:**
- Temporal variability analysis of optimal rainfall network from MIMR
- Optimal network varies significantly under sliding time series with fixed window
- The framework of dynamic network evaluation is proposed for decision support
- Optimal networks from dry season and wet season are compared

Abstract: Rainfall networks are the most direct sources of precipitation data and their optimization and evaluation are essential and important. Information entropy can not only represent the uncertainty of rainfall distribution but can also reflect the

correlation and information transmission between rainfall stations. Using entropy this study performs optimization of rainfall networks that are of similar size located in two big cities in China, Shanghai (in Yangtze River basin) and Xi'an (in Yellow River basin), with respect to temporal variability analysis. Through an easy-to-implement greedy ranking algorithm based on the criterion called, Maximum Information Minimum Redundancy (MIMR), stations of the networks in the two areas (each area is further divided into two subareas) are ranked during sliding inter-annual series and under different meteorological conditions. It is found that observation series with different starting days affect the ranking, alluding to the temporal variability during network evaluation. We propose a dynamic network evaluation framework for considering temporal variability, which ranks stations under different starting days with a fixed time window (1-year, 2-year, and 5-year). Therefore, we can identify rainfall stations which are temporarily of importance or redundancy and provide some useful suggestions for decision makers. The proposed framework can serve as a supplement for the primary MIMR optimization approach. In addition, during different periods (wet season or dry season) the optimal network from MIMR exhibits differences in entropy values and the optimal network from wet season tended to produce higher entropy values. Differences in spatial distribution of the optimal networks suggest that optimizing the rainfall network for changing meteorological conditions may be more recommended.

Key words: Entropy; rainfall network design; temporal variability; dynamic network

evaluation framework; maximum information minimum redundancy

1. Introduction

A hydrometric network is a data collection system, which is of fundamental significance for research and development. Mishra and Coulibaly (2009) reviewed several commonly used methods for hydrometric network design and evaluation, including: (1) statistically based, (2) information theory-based, (3) user survey, (4) hybrid, (5) physiographic components, and (6) sampling strategies. Chacon-Hurtado et al. (2017) made a review of rainfall and streamflow sensor network design with respect to different criteria, proposed a framework for classifying the design methods, and suggested a generalized procedure for optimal network design. In recent years, hydrometric network evaluation and optimization methods, particularly based on entropy, have received significant attention. Keum et al. (2017) reviewed studies that applied entropy theory in water monitoring network design and evaluation, especially for publications after 2009 which were not covered in the review by Mishra and Coulibaly (2009). Several groups of methods, using entropy theory, have been developed for network design and optimization. Fahle et al. (2015) divided most methods into three main groups: (1) derivation of minimum distance between stations or formation of regional information maps (e.g. Husain, 1989; Masoumi and Kerachian, 2010; Su et al., 2014); (2) provision of optimal sets or ranking of stations or iterative updating until reaching a certain threshold relative to specific objective functions for existing networks (e.g. Markus et al., 2003; Alfonso et al., 2010a; Li et al., 2012b; Mishra and Coulibaly, 2014; Stosic et al., 2017); and (3) multi-objective

optimization for simultaneously fulfilling several objectives of network design (e.g. Alfonso et al., 2010b; Samuel et al., 2013; Alfonso et al., 2014; Xu et al., 2015; Leach et al., 2015; Leach et al., 2016; Keum and Coulibaly, 2017a, 2017b). Sometimes hybrid methods are also applied and different groups of methods are compared. A brief literature review on hydrometric network design and optimization based on entropy is now given.

Langbein (1954) first proposed that monitoring network efficiency should be based on space-time measure of information. Caselton and Zidek (1984) chose optimal stations from maximum information transimission of the whole network. Husain (1987) used an objective function for network information based on bivariate normal, multivariate normal, and lognormal distributions but limited it to reducing stations from a dense network. Later he improved the method, generalized it for multivariate gamma distribution, and presented an approach based on transinformation-distance correlation for expansion of a sparse network (Husain, 1989). More improvements using information transmission and distance correlation have since been made for network design and optimization (e.g. Mogheir and Singh, 2003; Mogheir et al., 2006; Masoumi and Kerachian, 2008; Masoumi and Kerachian, 2010; Owlia et al., 2011; Su and You, 2014).

Many other measures originating in entropy theory have been applied to network evaluation, such as information transfer index (ITI), directional information transfer (DIT), coefficient of non-transferred and transferred information (t_1 and t_0), ensemble entropy, permutation entropy, and others (see Krstanovic and Singh, 1992a, 1992b;

Yang and Burn, 1994; Markus et al., 2003; Ridolfi et al., 2011; Alfonso et al., 2014; Stosic et al., 2017). Additionally, network evaluation and optimization have encompassed various types of hydrometric networks, including water quality and groundwater monitoring networks (Harmancioglu and Alpaslan, 1992; Ozkul et al., 2000; Mogheir and Singh, 2002; Mogheir et al., 2004a; Mogheir et al., 2004b; Mogheir et al., 2006; Guo et al., 2010; Leach et al., 2016), rainfall networks (Krstanovic and Singh, 1992a, 1992b; Yoo et al., 2008; Chen et al., 2008; Su and You, 2014; Xu et al., 2015; Yeh et al., 2017), streamflow gauge networks (Markus et al., 2003; Mishra and Coulibaly, 2010; Alfonso et al., 2012; Li et al., 2012b; Samuel et al., 2013; Mishra and Coulibaly, 2014; Leach et al., 2015; Stosic et al., 2017) and water level gauge networks (Alfonso et al., 2010a, 2010b; Mondal and Singh, 2012; Alfonso et al., 2014; Fahle et al., 2015). Among different types of hydrometric networks, regional rainfall networks provide a basis for analyzing rainfall characteristics, flood forecasting, water level monitoring, and design of hydraulic structures. Thus, evaluation and optimization of rainfall networks remain an important issue.

In addition, the effect of temporal variability and spatial differences on network design and optimization has received significant attention recently (e.g. Wei et al., 2014; Mishra and Coulibaly, 2014; Fahle et al., 2015). Climate change, human activities, and nonstationarity of hydrological processes cannot be overlooked during hydrometric network design and management, especially for rainfall events which are highly heterogeneous, localized, and strongly influenced by geographical, topographical, and climate factors (Zhang et al., 2014; Song et al., 2015; Gu et al.,

2016; Wang et al., 2016). Different temporal conditions can have appreciable impacts on hydrometric network design, meaning that an optimal network may only be optimal during the time when observations were made. Though temporal variability affects network optimization and its transferability to changing conditions, to the best of the authors' knowledge, studies on rainfall networks with respect to this temporal variability are limited. Accordingly, evaluation of rainfall networks under different times should be analyzed to complement the design procedure and to further strengthen the reliability and flexibility of network optimization.

The objectives of this study therefore are to (1) propose a framework for assessing the temporal variability of optimized rainfall networks by Maximum Information Minimum Redundancy (MIMR) criterion during changing periods with fixed window lengths; (2) examine how large the temporal variability is in terms of entropy values and rankings and quantify its effect on rainfall station ranking by ranking disorder index (RDI); and (3) investigate impacts of meteorological conditions on network optimization (especially for wet season and dry season) and compare the resulting optimal networks. The framework proposed in this study and related discussion and results will be useful for decision makers for obtaining optimal rainfall networks considering temporal variability, specifically, for designing and optimizing a rainfall network with a dynamic perspective. In addition, before conducting network optimization, we compare different quantization methods for entropy estimation and briefly discuss the impact of data length, which has been proposed and discussed in entropy-related hydrological studies (Mishra et al., 2011; Mishra and Coulibaly, 2014;

Keum and Coulibaly, 2017b).

The paper is organized as follows. First, in section 2, we introduce information entropy concepts and their estimation. In section 3, natural and geographical conditions of the study areas as well as current rainfall networks are described. In section 4, methodologies, including MIMR algorithm, RDI, and a dynamic network evaluation framework considering temporal variability, are presented. In section 5, we provide results and discussion on our rainfall networks. Conclusions and outlook are presented in the last section 6.

2. Entropy Concepts

Shannon (1948) laid the mathematical foundation of entropy as a measure of information or uncertainty. Entropy quantifies the uncertainty of a random variable via its probability distribution. In hydrometeorology, hydrology and water resources, frequently used entropy measures include marginal entropy, joint entropy, transinformation (mutual information), and total correlation.

2.1. Basic entropy measures

For entropy measures used in network design and optimization, there are basically three questions we need to address: (1) How much information would a station or several stations contain? (2) How much information can one station or several stations transmit to other stations? (3) How much information is shared by several stations? These three questions are answered by marginal entropy, joint entropy, transinformation, and total correlation.

Consider a discrete variable X with N possible results x_i ($i = 1,2,3, \dots, N$), where the probability of occurrence of x_i is p_i ($0 \leq p_i \leq 1$) and $\sum p_i = 1$. The marginal entropy is defined as:

$$H(X) = H(p_1, p_2, p_3, \dots, p_N) = -k \sum_{i=1}^{N} p_i \log p_i \quad (1)$$

where k is a constant and in general it can be taken as unity (Singh, 2013). The marginal entropy quantifies the average reduction in uncertainty of X or its probability distribution. Since variance is a measurement of dispersion or variability of a variable, it has been compared with entropy by Wei (1987) and Ebrahimi et al. (1999). While variance measures concentration only around the mean, entropy measures diffuseness of the probability irrespective of the location of concentration and may also be related to higher-order moments of a distribution, leading to a closer characterization of the probability density function (Mishra et al., 2009). There is therefore no unique relation between variance and entropy. For two discrete random variables X and Y with the joint probability of x_i and y_j denoted as $p(x_i, y_j) = p_{ij}, i = 1,2, \dots, N; j = 1,2, \dots, M$, the joint entropy between them can be defined as:

$$H(X, Y) = -\sum_{i=1}^{N} \sum_{j=1}^{M} p_{ij} \log p_{ij} \quad (2)$$

Transinformation describes transmission or sharing information between two random variables, which is defined as:

$$T(X, Y) = H(X) + H(Y) - H(X, Y) \quad (3)$$

For more than two variables ($X_1, \dots, X_n$), multivariate joint entropy with $p_{i_1, \dots, i_n}$ ($i_1 = 1,2, \dots, N_1, i_2 = 1,2, \dots, N_2, \dots, i_n = 1,2, \dots, N_n$) being an n-dimensional probability distribution can be defined as:

$$\mathrm{H}(X_1, \ldots, X_n) = -\sum_{i_1=1}^{N_1} \sum_{i_2=1}^{N_2} \ldots \sum_{i_n=1}^{N_n} p_{i_1 \ldots i_n} \log p_{i_1 \ldots i_n} \quad (4)$$

Total correlation describes information redundancy within multiple variables, which is an extension of transinformation and is defined as (McGill, 1954; Watanabe, 1960):

$$\mathrm{C}(X_1, \ldots, X_n) = \sum_{i=1}^{n} H(X_i) - H(X_1, \ldots, X_n) \quad (5)$$

To illustrate bivariate and multivariate joint entropy, transinformation, total correlation and relationships among them more directly, venn diagrams are provided in Figure 1 for clarity. Note that all the measures introduced above are with a logarithmic base of 2, so the units of entropy are 'bits'. Also, if p_i equals zero then we stipulate that 0log0=0.

Insert **Figure 1** here.

2.2. Estimation of entropy measures

As the probability distribution is contained in the definition of entropy, the key point is to estimate the probability distribution. The methods of probability distribution estimation can mainly be classified into two types: (1) fitting a probability distribution function (PDF) to data and its parameter estimation, and (2) nonparametric density estimation without assuming a distribution. In the first type, typical distributions, such as normal, lognormal, and gamma, have been mostly applied (e.g., Husain, 1989; Krstanovic and Singh, 1992b; Harmancioglu and Alpaslan, 1992; Yoo et al., 2008). When it comes to multivariate case, the calculations become quite difficult for other distributions. Further, many natural phenomena, like streamflow and precipitation, are heavy tail distributed (Li et al., 2012a) and the distribution assumptions may be debatable due to the unsuitable selection of distributions, while nonparametric density estimation avoids this step in the entropy estimation (Fahle et al., 2015). Among

different nonparametric estimators, the most predominant and frequently used is the binning estimator, since it is easy to understand and compute (e.g., Markus et al., 2003; Alfonso et al., 2010a; Li et al. 2012b). The main idea of this method is to discretize approximately a continuous random variable and there are different methods of discretization, for example, histogram method and mathematical floor function. For the histogram method, bins are often selected to group observation data into different classes and then use frequency as a substitute for probability. In MATLAB 2016a version the function "histcounts" which sets a binwidth parameter automatically or manually can be used for the histogram method. The mathematical floor function is another method of discretization (Alfonso et al., 2010a), which directly transfers continuous signals into discrete pulses and the conversion is performed by floor brackets ([.]) as:

$$X_q = a\left[\frac{2x+a}{2a}\right] \quad (6)$$

where x denotes the analog (e.g., continuous) value, and X_q denotes the quantized discrete value, while *a* denotes the bin width.

Other commonly used nonparametric methods, especially for estimating mutual information, include kernel density estimation (KDE) (e.g., Yang and Burn, 1994; Mishra and Coulibaly, 2014), and k nearest neighbor (KNN) (e.g., Kraskov et al., 2004; Khan et al., 2007; Fahle et al., 2015). However, due to their limited applications and sophisticated calculations in multivariate entropy estimation, we finally chose the binning estimator for our entropy calculations.

Generally, the bin width selection has always been an important issue, for it directly affects the approximate probability distribution and consequently affects the entropy value (e.g., Amorocho and Espildora, 1973; Scott, 1979; Shimazaki and Shinomoto, 2007; Ruddell and Kumar, 2009). The influence of different bin widths (or class

intervals) on the entropy value has been discussed by Singh (1997) and careful selection of class interval and sampling interval are suggested, because the entropy value decreases as the class interval or the sampling interval increases. Both fixed bin width and optimal bin width can be chosen and no standard rules for bin size (class interval) exist but empirical formulas can be referred to Papana and Kugiumtzis (2009). Scott (1979) suggested an optimal bin width as

$$\mathrm{a} = 3.49\sigma(\mathrm{x})N^{-\frac{1}{3}} \quad (7)$$

where σ denotes the standard deviation of an observation series of X (e.g. daily series of one station), and N denotes the sampling size (Scott, 1979).

Sturge (1926) presented the optimal bin width as

$$\mathrm{a} = \frac{R_x}{1+\log_2 N} \quad (8)$$

where R_x is the range of the observation series of X, and N is the sampling size of X. Bendat and Piersol (1966) developed another method for determining the optimal bin width as

$$\mathrm{a} = \frac{R_x}{1.87(N-1)^{0.4}} \quad (9)$$

where R_x is the range of the observation series of X and N is the sampling size of X. There have been some comparisons between different bin width selections (e.g., Fahle et al. 2015; Keum and Coulibaly, 2017a). Fahle et al. (2015) indicated that a fixed bin width was better suited, since a temporary rise in nearly constant water level increased the standard deviation and corresponding larger optimal bin width and lower marginal entropy (loss of information), which is against the initial intention of water level monitoring, i.e., detecting a change in the water level is more important. In the case of Keum and Coulibaly (2017a), they compared two fixed bin widths and two optimal bin widths (Sturges, 1926; Scott, 1979) for entropy values and found no strong

concordance among the four quantization cases (two fixed bin widths and two optimal bin widths) in entropy values. Though similar relative rankings were found for precipitation stations, most streamflow stations yielded very different relative rankings under different quantization cases. No matter which specific method is applied, the estimation method will influence the entropy value, and hence rankings (Fahle et al., 2015; Keum and Coulibaly, 2017a), and the influence can be specific for different cases so we will also compare different methods.

For multivariate entropy estimation, it is necessary to estimate the multivariate joint distribution to calculate the joint entropy. Alfonso et al. (2010b) and Li et al. (2012b) applied a grouping property (see Kraskov et al., 2005) for the estimation of multivariate joint entropy. The basic idea for this method lies in generating a new variable containing the information equal to that of the original variables. It can be verified that the information amount keeps invariant. Under this merging method, multiple variables can be converted to a nested variable and then multivariate joint entropy can be computed. For example, H(X, Y, Z) can be rewritten as H(A, Z)=H(<X,Y>, Z) with <·> denoting the merging operator. In a similar way, multivariate joint entropy with n variables can be transformed as:

$$\begin{aligned} H(X_1,X_2,\ldots, X_n) &= H(<X_1,X_2>, X_3,..., X_n) \\ &= H(<<X_1,X_2>,X_3>,\ldots, X_n) \\ &= H(<\ldots<<X_1,X_2>,X_3>,\ldots,X_{n-1}>, X_n) \\ &= \ldots \\ &= H(<\ldots<\ldots<<X_1,X_2>,X_3>,\ldots,X_{n-1}>, X_n>). \end{aligned} \quad (10)$$

Then, the total correlation can be computed by Equation (5) or also by a grouping property. The variable merging method is straightforward and for detailed information the reader can refer to Alfonso et al. (2010b) and Li et al. (2012b). Our study will

apply this merging method for multivariate entropy calculation.

3. Material

The Yangtze River and the Yellow River are the two most important rivers in China. Their watersheds exhibit greatly different geographical and hydrological characteristics. We selected two typical big cities, Shanghai in Taihu Lake basin (the Yangtze River) and Xi'an in Wei River basin (the Yellow River) as case studies, since they have similar rainfall network sizes. These two big cities have dense population and are both highly urbanized. Therefore, it is important to study the hydrological and meteorological characteristics of these areas and evaluate rainfall networks therein.

3.1. Study Area

Shanghai is located on the west coast of the Pacific Ocean, along the eastern Asian continent, in the front of the Yangtze River delta. It is east of East China Sea, south of Hangzhou Bay, west of Jiangsu Province and Zhejiang Province, north of the mouth of the Yangtze River, and also the Yangtze River and East China Sea connect here. By the end of 2003, Shanghai covered an area of 6340.5 square kilometers, accounting for 0.06% of the total area of the country. Most of the area is covered by flat plains, being part of the Yangtze River delta with an average altitude of 4 meters above the sea level. Shanghai has north subtropical maritime monsoon climate and is also the convergence zone of the north and south cold-warm air masses. Alternately affected by cold and warm air and marine moist air, Shanghai has humid climate and four distinct seasons with abundant sunshine and precipitation and its wet season extends from May to September. Characterized by three relatively low rainfall periods (midsummer, late autumn, and winter) and three relatively high rainfall periods (spring rainy, Meiyu, and autumn rainy), the mean annual rainfall is about

1123.7 mm and it is one of the areas where rainstorms occur frequently, with 85% of them occurring in the flood season and mostly concentrated during June to September. Huangpu River is the biggest river that forms the land water system of Shanghai, which originates from Taihu Lake with a total length of 113 kilometers and a width of 300~770 meters, flowing through the urban area and becoming an important water transportation route in Shanghai.

Xi'an is located in the central part of the Yellow River basin. It is the political, economic and cultural center of Shaanxi Province. The city covers an area of 10108 square kilometers, of which the urban area is about 1066 square kilometers. The landform of Xi'an includes four basic types: plain, loess tableland, hill and mountain, with the main geomorphic characteristics of the city being high in the south and low in the north. Another characteristic is its clear boundary between plain and mountain. Qinling Mountains and Wei River Plain constitute the main typical topography of this area. Wei River Plain is formed by the alluviation of Wei River (the largest tributary of the Yellow River) and its tributaries, which is also known as the Guanzhong Plain. Hills are mainly distributed in the south of Mount Li, Lantian County and Lintong County, covering an area of about 740 square kilometers, accounting for 7.41% of the city's total area. Mountains include the north slope of the Qinling Mountains and Mount Li, covering an area of 4875.6 square kilometers, accounting for 48.84% of the city's total area. Xi'an is located in the transitional zone between humid climate of the southeast coast and arid climate of the northwest inland and therefore has two types of climate characteristics. It has warm temperate semi-humid continental monsoon climate and the wet season extends approximately from May to September. The mean annual rainfall is about 537.5~1028.4 mm and the spatial distribution of rainfall shows significant differences, with more rainfall in southern Qinling

Mountains than in northern Wei River Plain. Interannual variability of rainfall is also significant, for wet year exceeding the average by 15%~33% and dry year being only 56.7%~77.4% of the average.

3.2. Dataset and Network

In total, 47 rainfall stations in Shanghai (except Chongming district) and 53 rainfall stations in Xi'an and their daily precipitation data of ten years (from 2006 to 2015) were chosen for a case study. Shanghai is further divided into two parts by the Huangpu River, which flows through the area and splits Shanghai into two parts: east of Huangpu River (SEH) with 23 rainfall stations and west of Huangpu River (SWH) with 24 rainfall stations. Xi'an is also divided into two parts according to topography. One part is plain with 24 rainfall stations (XPL) and another part is mountains and hills with 29 rainfall stations (XMH). Here we should note that two stations (No. 10, 11 in SWH) located in Kunshan, Jiangsu Province are included, since they are close to other stations in the network system. Similarly, three stations (No. 4, 5, 11 in XPL) located in Xianyang are included. Rainfall networks in two study areas with DEM values are briefly shown in Figure 2. Detailed geographical coordinate information of stations can be referred to in Appendix A.

Insert **Figure 2** here.

For prior knowledge of the two study areas, we first computed the network density and referred to hydrological guides for rainfall gauges provided by WMO [The Guide to Hydrological Practices, 2008] and Ministry of Water Resources of the People's Republic of China [Technical Regulations for Hydrologic Network Design, SL34-2013], as shown in Table 1.

Insert **Table 1** here.

According to WMO, the recommended minimum densities of precipitation stations for coastal, mountains, hilly area and interior plains are 900, 250, 575, and 575 km^2 per station, respectively. According to the guidelines of China, the control area of a single station should not be greater than 200 km^2 (150 km^2 especially in the plain river network region). Thus, the rainfall network densities of these two cities shown in Table 1 already meet both global and regional demands for recommended minimum network densities. Since it is not readily definable whether to add or reduce stations, we used the existing network for temporal analysis, especially for network reduction scenarios. However, the framework of analysis can be applied and generalized for network augmentation or relocation, if data in ungauged area is available by simulation or interpolation. Meanwhile, it should be realized that regulations are often strict not in terms of station density, but in the quality of data for providing information about the hydrological system (Chacon-Hurtado et al., 2017). Since one of the main objectives is to present a framework for dynamic rainfall network evaluation, we will mainly focus on rankings of importance for existing rainfall network and its corresponding network performance in terms of entropy values.

4. Methodology

In this section, an information theory-based method applied for rainfall network evaluation and design is introduced, named Maximum Information Minimum Redundancy (MIMR) greedy ranking algorithm, and a ranking disorder index (RDI) for rainfall stations based on normalized apportionment entropy is supplemented for

identifying stations with high temporal variability of the rankings. Besides, the framework of dynamic rainfall network evaluation mainly considering temporal variability analysis is developed for further discussion.

4.1. MIMR Algorithm

MIMR (Li et al., 2012b) greedy ranking criterion provides an easy-to-implement way to solve the network optimization problem, mainly suitable for selecting the best set of stations from a dense network or to obtain ranking of importance for stations in the network. The objective function contains three items: joint entropy (H), transinformation (T) and total correlation (C) and the optimization problem is formulated as:

$$max\ \mathrm{MIMR} = \lambda_1 \left(H\left(X_{s_1}, X_{s_2}, \dots, X_{s_k}\right) + \sum_{i=1}^{m} T\left(\langle X_{s_1}, X_{s_2}, \dots, X_{s_k}\rangle; X_{F_i}\right) \right) - \lambda_2 C\left(X_{s_1}, X_{s_2}, \dots, X_{s_k}\right)$$

subject to:

$$\mathrm{k} + \mathrm{m} = \mathrm{N}$$

$$\mathrm{H}\left(X_{S_1}, X_{S_2}, \dots, \mathrm{X}_{\mathrm{S_k}}\right) \geq \mathrm{Pct} * \mathrm{H}(X_{S_1}, X_{S_2}, \dots X_{S_k}, X_{F_1}, X_{F_2}, \dots, X_{F_m}) \qquad (11)$$

where k is the number of selected stations (denoted as a set of S), m is the number of unselected stations (denoted as a set of F), N is the total number of stations in the current network, and λ_1, λ_2 are trade-off weights between information and redundancy and $\lambda_1 + \lambda_2 = 1$. Pct is the percentage of joint entropy, representing the ratio between joint entropy for the selected stations and joint entropy for all stations in the existing network. The first item in MIMR denotes the total information contained by the selected stations in the network (Equation (4)), the second item denotes the information that can be obtained about unselected stations from the selected stations (Equation (3)) and the third item denotes the information redundancy among the

selected stations (Equation (5)). To solve the optimization problem and avoid computation burden from global search, Li et al. (2012b) generalized a greedy selection algorithm. First we identify the central station with the highest marginal entropy, and then rankings of stations can be obtained by implementing the MIMR criterion step by step. As a threshold (Pct) is set, e.g., obtaining 95% of the total joint entropy in the existing network, we can get an optimal subset of stations for the final decision. After selecting the last station, rankings of all stations can be determined and used for further variability analysis. For more details about MIMR readers can refer to Li et al. (2012b). Since the first goal of network design is generally accepted as maximizing information of the network system, λ_1 is usually larger than λ_2, suggested by Li et al. (2012b) as 0.8. We made sensitivity analysis on different information-redundancy trade-off weights and found that most stations kept stable with λ_1 varying from 0.5 to 1 (λ_2 varying from 0.5 to 0). To sum up, the default parameter set for our MIMR selection is given by: $\lambda_1 = 0.8$, $\lambda_2 = 0.2$ and Pct = 95%.

4.2. Ranking Disorder Index

To assess the dispersion of ranks for a specific rainfall station by MIMR, we adopted a ranking disorder index (RDI) for identifying rainfall stations whose importance is mostly affected by temporal change. The definition of the index originates from apportionment entropy (AE) (Maruyama and Kawachi, 1998; Kawachi et al., 2001) as:

$$\mathrm{AE} = -\sum_{i=1}^{n} \frac{r_i}{N} \log_2\left(\frac{r_i}{N}\right) \quad (12)$$

where n is the number of all possible ranks for a station (so it is equal to the number of stations in a network), N is the number of ranks under different times and r_i is the number of a certain *i*th rank. As the maximum values of AE is $\log_2(\mathrm{n})$, RDI is

calculated by normalized apportionment entropy (nAE) defined as:

$$\mathrm{RDI} = \mathrm{nAE} = \mathrm{AE}/\log_2(\mathrm{n}) \quad (13)$$

The RDI is an index for the uncertainty and fluctuation of the rank assigned to a rainfall station (Fahle et al., 2015), which mainly considers the impact of temporal change on rankings of rainfall stations. Thus, a higher RDI reflects a station whose importance is more affected by temporal variability. So more careful attention should be paid to such station as its importance can vary a lot during a relatively short period.

4.3. Dynamic network evaluation framework

One of the main objectives of this study is to present a dynamic network evaluation framework from a given length of time series. Figure 3 shows an overview of the proposed framework and a brief step-by-step description is given below:

Step 1. Obtain rainfall data at all stations in a network.

Step 2. Preprocess data for different time windows (e.g., 1-year, 2-year, 5-year).

Step 3. Do MIMR ranking with different data series from step 2.

Step 4. Obtain network performances (entropy values and rankings) for sliding series.

Step 5. Compute ranking disorder index (RDI) for all stations with rankings in step 4.

Step 6. Rank all stations in descending order of RDI.

Step 7. Identify ranking variations for individual stations.

Step 8. Obtain final suggestions for network design.

Insert **Figure 3** here.

For data preprocessing in step 2, we mainly focus on shifting inter-annual time window. To assess temporal variability from the inter-annual aspect, Figure 4 shows how sliding inter-annual observation series (fixed window length of 1-year, 2-year

and 5-year with a 10-day increment) are obtained from a continuous ten-year series (2006-2015). To assess the temporal variability from the seasonal aspect, we divide the entire data set into wet season (5 months, May to September) and dry season (7 months, October to April). To see the seasonal variability effect on average and avoid years with extreme precipitation, we use a 50-month series (5 months for one year, 10 years in total) for wet season and 70-month series (7 months for one year, 10 years in total) for dry season.

Insert **Figure 4** here.

5. Results and Discussion

5.1. Entropy estimation and data length impact

Entropy values are affected by different binning parameters when using the binning estimator. We compared fixed and optimal bin widths. Considering the fixed bin width, *a* in Equation (6) was set as 1 mm/d and 3 mm/d. For optimal bin width, three methods (Scott, Sturge, Bendat & Piersol) were used to calculate the optimal bin width for each station. Table 2 shows that entropy values diverged for different bin widths, especially for joint entropy and total correlation with fixed and optimal bin widths. The differences between two fixed widths (a=1 mm/d and a=3 mm/d) suggested that a larger bin width led to smaller entropy values, which agreed with the results from Singh (1997). For the optimal bin size suggested by Scott, Bendat & Piersol and fixed bin width as a=1 mm/d, the maximum marginal entropy and average marginal entropy were a little larger than the values calculated using Sturge's method

and the fixed bin width as a=3 mm/d. Though the maximum marginal entropy affected the first step in MIMR ranking, the differences between different methods fell in a narrow range. Meanwhile, both fixed bin widths with a=1 mm/d and a=3 mm/d had joint entropies lower than 4 bits, while the other three methods yielded average joint entropies higher than 5 bits. As the saturated value for the joint entropy was equal to $\log_2(n) = 11.83$ bits (n is the number of data points given by ten-year daily series, which is equal to 3652), the joint entropies estimated by fixed bin widths with a=1 mm/d and a=3 mm/d were much lower than the saturated joint entropy. The total correlation was also subject to different bin widths and even related to the number of variables (stations). Therefore, we provided not only the total correlation but also the values divided by the number of stations (TC/N). Though entropy values would change under different binning estimations, it was interesting to find that the relative differences between areas were maintained. For instance, all the entropy values of XPL were the smallest among the four subareas no matter which binning estimation was used. In addition, the entropy values of SEH and SWH were always similar regardless of different binning sizes. Going one step further, basic statistical properties of precipitation in the four subareas can serve as a reference (Figure 5). From Figure 5, XPL showed the lowest mean precipitation, lowest maximum precipitation and lowest standard deviation (std. dev.). Additionally, SEH and SWH had similar mean precipitation and standard deviation. Hence, it can be assumed that the average rainfall and its variability roughly determined the intrinsic characteristic of a network, which would be consistently quantified by entropy values. As a consequence,

compared with other three subareas, rainfall network in certain areas, such as interior plain (XPL) with both less rainfall and lower variability, would generate less information. In total, relative differences for entropy values among different areas were maintained with five estimation methods. To avoid multifarious analysis with different estimation methods, we finally chose Scott's method for entropy estimation in the following sections as it yielded the largest marginal entropy and joint entropy.

Insert **Table 2** here.

Insert **Figure 5** here.

In order to see the impact of data length on entropy values, we calculated average marginal entropy, joint entropy and total correlation with increasing length of data series. Figure 6 shows that variations of entropy values are similar for the four subareas. Generally, the values became larger with increasing data length and then tended to stabilize. It should also be noted that the increasing trends were not monotonous and fluctuations may emerge at certain data lengths (e.g., around 500 days, 1200 days and 1500 days). This agrees with the previous conclusion that both joint entropy and total correlation were sensitive to time series length (Keum and Coulibaly, 2017b). As they finally recommended the use of at least 10 years of daily time series data to avoid any significant information loss of the network, we also used a ten-year series for MIMR selection to get an optimal network for reference.

However, shorter lengths of series (1-year, 2-year and 5-year) were still used for assessing temporal variations of the optimized network with MIMR due to limited length of data series available. We also noticed that the entropy values had nearly reached 70%, 80% and 90% of the values estimated by 10-year series using 1-year, 2-year and 5-year series respectively. More importantly, network performance was subject to inter-annual variability and it was necessary to capture and analyze the variability with relatively shorter time series.

Insert **Figure 6** here.

5.2. Optimal networks under sliding series using MIMR

Before optimizing rainfall networks under sliding series, primary MIMR ranking and selection results based on the entire 10-year series are shown in Table 3. Generally, higher percentage of stations were selected for XPL and XMH than those for SEH and SWH, though MIMR values for XPL and XMH were relatively lower. In addition, the lower MIMR values were mainly due to lower transinformation in XPL and XMH.

Insert **Table 3** here.

In order to conduct temporal variability analysis on network optimization, it is necessary to identify impacts of sliding time series on optimal rainfall networks using MIMR. The graphs in Figure 7 show the dispersion pattern and temporal evolution

trend of the information (joint entropy) and redundancy (total correlation) content for optimal networks in four subareas with different time series, which are from moving time windows of 1 year, 2 years and 5 years. For the sake of comparison, the ranges of coordinate axes were deliberately set as the same under different time windows. From Figure 7 (left column), a longer time window brought less diffusive pattern of joint entropy and total correlation. Meanwhile, the quantity of joint entropy and total correlation also generally increased with longer time windows (shown as scatters in three subplots from top to bottom gradually shifting to upper right of the subplot), which again confirmed the data length impact discussed in section 5.1. Considering the temporal evolution of entropy values, the middle and right columns in Figure 7 revealed that the fluctuations could be more significant with 1-year series and 2-year series. In total, the results revealed that the performance of the optimal network using MIMR could fluctuate a lot, especially under shorter duration. However, it was found that the variation trend of joint entropy would be basically maintained for three time windows. From Figure 7 (middle column), variation curves with different time windows (especially for 1-year and 2-year, as the total number of series was greater than the number of 5-year series) showed similar patterns and periodicity for joint entropy. Yet the similarity was not much comparable for total correlation under three time windows. This demonstrated that changing series indeed affected the information content of the optimal network. In such case, a finer time window may capture this variability and periodicity more easily and optimization based on long time series may weaken such temporal variability. In addition, two subareas in the same city (i.e., SEH

and SWH in Shanghai, XPL and XMH in Xi'an) also showed similar variation patterns for joint entropy, so we only chose SWH and XPL for further discussion.

Insert **Figure 7** here.

For SWH (Figure 8(a)), six scenarios with high and low joint entropy were chosen for comparison. Scenario 1, 2 and 3 represented the optimal networks of the 80-th series (series started from around 2008-03-01) for three window lengths with relatively high joint entropy. Scenario 4, 5 and 6 represented the optimal networks of the 120-th series (series started from around 2009-04-05) for three window lengths with relatively low joint entropy. The selected stations and information content were provided in Table 4. Similarly, six scenarios from different series (Scenarios 1, 2, 3 are the 55-th series started from around 2007-06-25 and Scenarios 4, 5, 6 are the 110-th series started from around 2008-12-26) were also chosen for XPL (Figure 8(b), Table 5).

Insert **Figure 8** here.

Insert **Table 4** here.

Insert **Table 5** here.

Basically, the differences between high and low joint entropy gradually fell in a narrower range with increasing window length. However, the pattern was not the same for total correlation. Also, the optimal network was highly impacted by temporal variability, and changes in entropy values would also correspond to different selected stations. Much fewer stations (only 6 and 7 stations) were selected for 1-year series than in other scenarios in SWH (Table 4). However, the differences between the number of selected stations in different scenarios were not much significant in XPL (Table 5). Regarding specific stations selected in the optimal network, there was no strong pattern among all scenarios, but more stations tended to appear in both selected sets of scenarios with high and low joint entropy as the fixed window length increased (Table 4 and 5, marked as red numbers). Specifically, only one station (No. 5) appeared both in selected sets of Scenario 1 and 4, while six stations (No. 17, 4, 3, 5, 18, 14) appeared both in selected sets of Scenario 2 and 5 and the number for Scenario 3 and 6 was eight (stations No. 2, 20, 5, 18, 24, 16, 6, 4). In XPL, since the selected number was much larger, selected sets of stations were much more similar in both scenarios of the same length with high and low joint entropy. Specifically, the numbers (marked as red) were 12 stations for Scenario 1 and 4, 13 stations for Scenario 2 and 5, and 14 stations for Scenario 3 and 6. Naturally, larger selected number of stations would lead to less different selected set. The differences in the selected network were much more noticeable in SWH (Shanghai) since the number of selected stations would be more affected by the window length of the series. Generally, high or low joint entropy signified different specific rainfall distribution patterns

during a given time period. Therefore, the sliding series using MIMR may also help find the changing point in rainfall distributions for rainfall network optimization.

5.3. Ranking variation and dynamic network evaluation

Apart from variations of entropy values, rankings by MIMR of all stations also fluctuated with time lapse. From bubble plots (Figure 9 and Figure 10) with the bubble size representing the relative frequencies of different ranges of the assigned rank (in descending order, i.e., the top 20% rank meaning the most important), it can be concluded that a rainfall station can either be high-ranked or low-ranked during different times. Results showed that no clear fixed ranks were found for most stations in Shanghai as most bubbles on the same horizontal lines were of similar sizes in Figure 9(a) and 9(b), meaning that most stations could be ranked with any order especially for 1-year and 2-year sliding series. In other words, it can be concluded that the importance of stations fluctuated a lot and exhibited great variability for 1-year and 2-year sliding series. For example, bubbles of station No. 5, 15, 16 in SEH were almost of similar sizes for different ranks (Figure 9(a)), meaning that the rank of the station can be either high or low in different times. Similar results were found for many other stations in SEH, e.g., station No. 22 and station No. 23. Meanwhile, we also found that as the time window got longer, the differences between relative frequencies of different ranks of some stations became a little more clear. For instance, station No. 14 (SEH) was mostly ranked the last 20% and station No. 19 (SEH) was mostly ranked 20%~40% for 5-year time window (Figure 9(c)). It can be explained

that when a longer time window was used for temporal variability analysis, the variation in rainfall series would be more moderate with lower percentage of days changed and the optimization results would accordingly be less affected and fluctuating. In SWH (Figure 9(d), 9(e), 9(f)), similar results again warned that rankings of rainfall stations can vary a lot and getting an optimal result only on a fixed long observation period may bring risk at times. Compared with SEH and SWH, differences between relative frequencies of different assigned ranks for a specific station were a little more noticeable in XPL and XMH, and longer time window again brought more diverging bubble sizes. For instance, station No. 6 (XPL, Figure 10(c)) was always ranked top 20% while station No. 7 (XPL, Figure 10(c)) was mostly ranked last 20%. It can be assumed that the temporal variability of rankings might be a little lower for stations in Xi'an than stations in Shanghai.

Insert **Figure 9** here.

Insert **Figure 10** here.

Considering dynamic network evaluation, RDI of all stations under different time windows were calculated. For the convenience of discussion, only the top five stations with highest RDI were selected for assessment (see in Table 6). Stations with the highest RDI were not the same under different time windows but some stations would appear at the top in RDI order under different window lengths, e.g., station No. 20

(SEH) and station No. 22 (XPL). Generally, a shorter window length would yield a higher RDI for the same station. For example, station No. 20 in SEH yielded RDIs of 0.96, 0.93 and 0.82 respectively for 1-year, 2-year and 5-year windows. The results by RDI again confirmed the variation patterns under different window lengths shown in the bubble plots (Figure 9 and Figure 10).

Insert **Table 6** here.

Moreover, ranking variations for top five stations are shown in Figure 11 and Figure 12 under different sliding time windows. The ranks of these stations showed high temporal variability under sliding series, indicated by a broad range of assigned color referring to the specific rank during the observation period. On one hand, for instance, station No. 22 (SEH) was in the selected set based on the 10-year series (see in Table 3). However, during 1-year series starting from 2006 to around 2008 (Figure 11(1)), it was frequently ranked last (except several blue lines), while after around June in 2009 it may become more important, indicated by more blue lines. On the other hand, though station No. 20 (SEH) was not selected for the optimal network based on the 10-year series (see in Table 3), it was found to be relatively important during the sliding series starting from around April in 2009 regardless of the window length (Figure 11(1), 11(2), 11(3)). Another case is station No. 17 (SWH) which was selected in the optimal network (see in Table 3). However, its rank fluctuated a lot both in 1-year and 2-year windows of the sliding series (Figure 11(4), 11(5), especially from

2008/03/11 to 2010/05/20), as seen from either dark blue or dark red lines. Similar results can be found in XPL (station No. 13, 18, etc. Figure 12(1), 12(2), 12(3)) and XMH (station No. 1, 3, 4, etc. Figure 12(4), 12(5), 12(6)). Generally, the variation patterns were not very regular and it was also very difficult to predict when a station would be more important. However, we gained an insight from the results that the ranking from MIMR based on a fixed observation period may not be much reliable if temporal variability is considered. As a consequence, we should be more careful when using the MIMR-based ranking (maybe other ranking criteria) for network design due to the significant temporal variability impact. The dynamic network evaluation can be viewed as a supplement for MIMR selection, based on fixed long time series. Moreover, we tend to suggest avoiding the reduction of stations with great ranking variability in the final network, if economic conditions allow.

Insert **Figure 11** here.

Insert **Figure 12** here.

5.4. Comparison between optimal networks under wet and dry seasons

Temporal variability is partly due to forcing factors and changing patterns for rainfall distribution both in space and time. Therefore, we discuss rainfall optimization under different meteorological conditions. The optimal network results and relative network performances were found to be different in wet season and dry season (see Table 7).

Results for the optimized network with the entire series (ten-years) are also provided for reference. In general, wet season produced more information than did dry season, for the joint entropy was approximately 2 bits higher in wet season than in dry season in four subareas. Moreover, wet season even generated higher joint entropy than the entire series, which might be explained by relatively more concentrated rainfall and variability during this period. For example, the average daily rainfall for Shanghai was 4.56 mm during wet season and 2.31 mm during dry season. Meanwhile, the average variances were 160.58 for wet season and 52.12 for dry season. Though the same numbers of stations were selected in wet and dry seasons for SWH, a little more total correlation (redundant information) was also obtained in wet season. The range of differences of the total correlation was even larger for XPL (about 15 bits) in two seasons. Furthermore, more transinformation in wet season suggested that the information transition ability was higher during this period for the four subareas. In total, the MIMR values in different cases signified that the optimal network from wet season may yield higher information efficiency than that from dry season.

Insert **Table 7** here.

Despite different entropy values and rankings, the majority of selected stations were similar under three conditions (wet season, dry season and entire series), considering the relatively high overlapping ratio in Table 7. We noticed that the optimal network of the wet season was slightly more different from the optimal network of the entire

series in the central area of Shanghai (Figure 13(1)). For instance, stations No. 7, 8, 9 and 19 (in SWH) were included during wet season but were unselected in the optimal network of dry season and entire series. Fewer stations in the southwestern area of Xi'an (only station No. 2, 4, 8 in XMH, Figure 13(4)) were selected in the optimal network of the wet season than the other two optimal networks. Generally, the rainfall distribution pattern during wet season could be much different from dry season or entire series, which may explain the dissimilarities among different optimal networks under different meteorological conditions. Thus, some of the selected or unselected stations would be temporarily of importance or redundancy. So we again suggest that great caution should be exercised when dealing with the optimization of a rainfall network, since the elimination of some stations based on fixed entire series would cause significant information loss or redundancy under other conditions. In general, optimizing the rainfall network to make it more adapted to the changing meteorological conditions may be more recommended according to our results.

Insert **Figure 13** here.

6. Conclusions

This study focused on rainfall network optimization considering temporal variability. Results from the study suggest that temporal variability of rainfall network optimization may have a great impact on relative network performance (entropy values and station rankings), especially under sliding inter-annual series. For two study areas, optimal networks tended to be more different under shorter shifting time

windows. However, the information content of optimal networks under different lengths of sliding windows showed similar and even periodic variation patterns, meaning that the temporal variability can be captured by the sliding series. Application of the proposed framework can help get an insight into network optimization with respect to temporal variability and even identify the changing point for network optimization. Under such situation, the optimized rainfall network may not be the same as the optimal one based on the entire series, which especially weakens our confidence in the reduced network based on static boundary conditions. Additionally, the proposed framework of rainfall network evaluation aims to complement the primary MIMR optimization approach. This is because it does not change the procedure of MIMR selection but adds an independent dynamic evaluation process to make rankings for stations under different times. With the obtained rankings, we can identify stations whose importance is most likely to be impacted by temporal changes with high RDI. Consequently, design of an optimal network should be careful enough to avoid risks of losing information under different temporal cases. On one hand, for those stations whose ranks fluctuate a lot even under longer window length, we suggest greater caution should be exercised on the decision of reducing them just based on an optimal result from the entire series. On the other hand, for those stations which are always ranked last in the sliding series, more certainty can be gained on removing them. Still, we should point out that a longer window length can be adopted (e.g. 10-year) in the future if a longer series is available. Also, other optimization approaches may be applied in the dynamic network evaluation

framework for comparison and verification in future studies.

Differences between optimal networks from wet season and dry season revealed that meteorological conditions may impact the resulting optimal network, and rainfall network optimization based on fixed entire series would cause information loss or redundancy under different conditions. As meteorological conditions can partly be seen as one of the inner forcing factors for temporal variability of rainfall network, the study on network optimization under different meteorological conditions is critical for better understanding and recognition of dynamic network design. Since our division and classification was relatively simple and rough, more careful and multifarious categories of meteorological conditions can be defined and compared in the future. Generally, how to relate rainfall network optimization to its regional dynamic rainfall distribution characteristics, in order to derive a more integrated and practical optimization scheme under changing environment, is a future direction that deserves consideration.

Acknowledgments

This study was supported by National Natural Science Fund of China (No. 41571017, 51679118, 91647203), National Key Research and Development Program of China (2017YFC1502704, 2016YFC0401501).

Appendix A

SEH			SWH		
Station No.	latitude	longitude	Station No.	latitude	longitude
1	121°00′	30°54′	1	120°55′	31°07′
2	121°10′	30°54′	2	120°54′	31°01′
3	121°19′	30°52′	3	121°00′	31°05′
4	121°08′	30°47′	4	121°06′	31°01′
5	121°17′	30°48′	5	121°09′	30°58′
6	121°18′	30°58′	6	121°14′	30°58′
7	121°13′	30°56′	7	121°07′	31°07′
8	121°23′	30°59′	8	121°11′	31°06′
9	121°34′	30°57′	9	121°17′	31°07′
10	121°44′	30°59′	10	120°57′	31°19′
11	121°33′	31°15′	11	121°05′	31°18′
12	121°28′	31°10′	12	121°04′	31°16′
13	121°45′	31°07′	13	121°09′	31°20′
14	121°30′	31°01′	14	121°17′	31°19′
15	121°53′	31°00′	15	121°14′	31°16′
16	121°46′	31°01′	16	121°12′	31°28′
17	121°40′	31°19′	17	121°25′	31°03′
18	121°46′	31°13′	18	121°29′	31°18′
19	121°23′	30°45′	19	121°22′	31°08′
20	121°51′	30°52′	20	121°20′	31°25′
21	121°29′	31°00′	21	121°23′	31°20′
22	121°31′	30°50′	22	121°11′	31°17′
23	121°44′	30°52′	23	121°29′	31°22′
			24	121°30′	31°23′

XPL			XMH		
Station No.	latitude	longitude	Station No.	latitude	longitude
1	108°05′	34°11′	1	107°46′	33°51′
2	108°07′	34°07′	2	107°50′	33°51′
3	108°28′	34°10′	3	107°54′	33°55′
4	108°41′	34°26′	4	108°07′	33°53′
5	108°42′	34°19′	5	108°00′	33°49′
6	108°54′	34°04′	6	108°07′	33°50′
7	108°46′	34°06′	7	108°10′	33°51′
8	108°58′	34°05′	8	108°06′	33°58′
9	108°49′	34°06′	9	108°13′	34°03′
10	108°46′	34°14′	10	108°22′	34°02′
11	108°51′	34°24′	11	108°28′	33°51′

12	109°31′	34°11′	12	108°32′	33°49′
13	109°22′	34°09′	13	108°31′	33°55′
14	109°24′	34°03′	14	108°32′	34°01′
15	109°14′	34°13′	15	108°49′	33°51′
16	109°09′	34°14′	16	108°51′	33°54′
17	109°05′	34°18′	17	108°39′	33°56′
18	109°13′	34°02′	18	108°43′	34°00′
19	109°06′	34°08′	19	108°44′	33°53′
20	109°07′	34°32′	20	108°57′	33°56′
21	109°20′	34°22′	21	108°57′	33°59′
22	109°23′	34°17′	22	109°07′	33°58′
23	109°26′	34°25′	23	109°07′	34°00′
24	109°31′	34°16′	24	109°40′	34°09′
			25	109°30′	34°03′
			26	109°28′	34°04′
			27	109°30′	33°55′
			28	109°20′	33°54′
			29	109°22′	33°59′

References

Alfonso, L., Lobbrecht, A., Price, R., 2010a. Information theory–based approach for location of monitoring water level gauges in polders. Water Resour. Res. 45. W03528. http://dx.doi.org/10.1029/2009WR008101.

Alfonso, L., Lobbrecht, A., Price, R., 2010b. Optimization of water level monitoring network in polder systems using information theory. Water Resour. Res. 46. W12553. http://dx.doi.org/10.1029/2009WR008953.

Alfonso, L., Liyan, H., Lobbrecht, A., Price, R., 2012. Information theory applied to evaluate the discharge monitoring network of the Magdalena River. J. Hydroinform. 15 (1), 211-228. http://dx.doi.org/10.2166/hydro.2012.066.

Alfonso, L., Ridolfi, E., Gaytan-Aguilar, S., Napolitano, F., Russo, F., 2014. Ensemble entropy for monitoring network design. Entropy, 16(3), 1365-1375. http://dx.doi.org/10.3390/e16031365

Amorocho, J., Espildora, B., 1973. Entropy in the assessment of uncertainty in hydrologic systems and models. Water Resour. Res. 9, 1511-1522. http://dx.doi.org/10.1029/WR009i006p01511.

Bendat, S., Piersol, A., 1966. Measurements and Analysis of Random Data. John Wiley and Sons. New York.

Caselton, W. F., Zidek, J. V., 1984. Optimal monitoring network designs. Stat. Probab. Lett. 2(4): 223-227. https://doi.org/10.1016/0167-7152(84)90020-8.

Chacon-Hurtado, J. C., Alfonso, L., Solomatine, D. P., 2017. Rainfall and streamflow sensor network design: a review of applications, classification, and a proposed framework. Hydrol. Earth Syst. Sci., 21(6), 3071. https://doi.org/10.5194/hess-21-3071-2017.

Chen, Y. C., Wei, C., Yeh, H. C., 2008. Rainfall network design using kriging and entropy. Hydrol. Process. 22: 340-346. http://dx.doi.org/10.1002/hyp.6292.

Ebrahimi, N., Maasoumi, E., Soofi, E. S., 1999. Ordering univariate distributions by entropy and variance. J. Econometrics. 90(2), 317-336. http://dx.doi.org/10.1016/S0304-4076(98)00046-3.

Fahle, M., Hohenbrink, T. L., Dietrich, O., Lischeid, G., 2015. Temporal variability of the optimal monitoring setup assessed using information theory. Water Resour. Res. 51, 7723-7743. http://dx.doi.org/10.1002/2015WR017137.

Gu, X. H., Zhang, Q., Singh, V. P., Shi, P. J., 2017. Nonstationarity in timing of extreme precipitation across China and impact of tropical cyclones. Glob. Planet. Chang. 149, 153-165. https://doi.org/10.1016/j.gloplacha.2016.12.019.

Guo, Y. S., Wang, J. F., 2010. Spatial analysis on the layout of groundwater quality monitoring network. 18th International Conference on Geoinformatics, Geoinformatics, 2010. https://doi.org/10.1109/GEOINFORMATICS.2010.5567881.

Harmancioglu, N. B., Alpaslan, N., 1992. Water quality monitoring network design: A problem of multi-objective decision making. Water Resour. Bull., 28(1), 179–192. http://dx.doi.org/10.1111/j.1752-1688.1992.tb03163.x.

Husain, T., 1987. Hydrologic network design formulation. Can. Water Resour. J. 12 (1), 44–63. http://dx.doi.org/10.4296/cwrj1201044.

Husain, T., 1989. Hydrologic uncertainty measure and network design. Water Resour. Bull. 25 (3), 527–534. http://dx.doi.org/10.1111/j.1752-1688.1989.tb03088.x.

Kawachi, T., Maruyama, T., Singh, V. P., 2001. Rainfall entropy for delineation of water resources zones in Japan. J. Hydrol. 246(1), 36-44. https://doi.org/10.1016/S0022-1694(01)00355-9.

Keum, J., Coulibaly P., 2017a. Information theory-based decision support system for integrated design of multivariable hydrometric networks. Water Resour. Res., 53, https://doi.org/10.1002/2016WR019981.

Keum, J., Coulibaly, P., 2017b. Sensitivity of entropy method to time series length in hydrometric network design. J. Hydrol. Eng., 22(7), 04017009. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001508.

Keum, J., Kornelsen, K. C., Leach, J. M., Coulibaly, P., 2017. Entropy applications to water monitoring network design: a review. Entropy, 19(11), 613. https://doi.org/10.3390/e19110613.

Khan, S., Bandyopadhyay, S., Ganguly, A. R., Saigal, S., Erickson, D. J., Protopopescu, V., Ostrouchov, G., 2007. Relative performance of mutual information estimation methods for quantifying the dependence among short and noisy data, Phys. Rev. E, 76(2), 026209. http://dx.doi.org/10.1103/PhysRevE.76.026209.

Kraskov, A., Stöegbauer, H., Grassberger, P., 2004. Estimating mutual information. Phys. Rev. E, 69, 066138. http://dx.doi.org/10.1103/PhysRevE.69.066138.

Kraskov, A., Stöegbauer, H., Andrzejak, R. G., Grassberger, P., 2005. Hierarchical clustering based on mutual information. Europhys. Lett., 70, 278. http://dx.doi.org/10.1209/epl/i2004-10483-y.

Krstanovic, P. F., Singh, V. P., 1992a. Evaluation of rainfall networks using entropy: I. Theoretical development. Water Resour. Manage. 6, 279-293. http://dx.doi.org/10.1007/BF00872281.

Krstanovic, P. F., Singh, V. P., 1992b. Evaluation of rainfall networks using entropy: II. Application. Water Resour. Manage. 6, 295-314. http://dx.doi.org/10.1007/BF00872282.

Langbein, W. B., 1954. Stream gaging networks. Int. Assoc. of Hydrol. Gen. Assem., Rome., Publ. 38, pp. 293-303.

Leach, J. M., Kornelsen, K. C., Samuel, J., Coulibaly, P., 2015. Hydrometric network

design using streamflow signatures and indicators of hydrologic alteration, J. Hydrol., 529(3), 1350–1359, http://dx.doi.org/10.1016/j.jhydrol.2015.08.048.

Leach, J. M., Coulibaly, P., Guo, Y., 2016. Entropy based groundwater monitoring network design considering spatial distribution of annual recharge, Adv. Water Resour., 96, 108–119, http://dx.doi.org/10.1016/j.advwatres.2016.07.006.

Li, C., Singh, V. P., Mishra, A. K., 2012a. Simulation of the entire range of daily precipitation using a hybrid probability distribution. Water Resour. Res. 48. W03521. http://dx.doi.org/10.1029/2011WR011446.

Li, C., Singh, V. P., Mishra, A. K., 2012b. Entropy theory-based criterion for hydrometric network evaluation and design: Maximum information minimum redundancy. Water Resour. Res. 48. W05521. http://dx.doi.org/10.1029/2011WR011251.

Markus, M., Knapp, H. V., Tasker, G. D., 2003. Entropy and generalized least square methods in assessment of the regional value of streamgages. J. Hydrol. 283(1-4). http://dx.doi.org/10.1016/S0022-1694(03)00244-0.

Maruyama, T., Kawachi, T., 1998. Evaluation of rainfall characteristics using entropy. Journal of Rainfall Catchment Systems. 4(1), 7-10.

Maruyama, T., Kawachi, T., Singh, V. P., 2005. Entropy-based assessment and clustering of potential water resources availability. J. Hydrol. 309(1), 104-113. https://doi.org/10.1016/j.jhydrol.2004.11.020.

Masoumi, F., Kerachian, R., 2010. Optimal redesign of groundwater quality monitoring networks: A case study. Environ. Monit. Assess. 161(1-4), 247-257. http://dx.doi.org/10.1007/s10661-008-0742-3.

McGill, W., 1954. Multivariate information transmission. Psychometrika. 19, 97-116. http://dx.doi.org/10.1007/BF02289159.

Ministry of Water Resources, 2013. Technical Regulations for Hydrologic Network Design, SL34-2013. P. R. China.

Mishra, A. K., Özger, M., Singh, V. P., 2009. An entropy-based investigation into the variability of precipitation. J. Hydrol. 370(1), 139-154. https://doi.org/10.1016/j.jhydrol.2009.03.006.

Mishra, A. K., Özger, M., Singh, V. P., 2011. Association between uncertainties in meteorological variables and water-resources planning for the state of Texas. J. Hydrol. Eng. 16(12), 984-999. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000150.

Mishra, A.K., Coulibaly, P., 2009. Developments in hydrometric network design: A review. Rev. Geophys. 47 (2). http://dx.doi.org/10.1029/2007RG000243. RG2011.

Mishra, A.K., Coulibaly, P., 2010. Hydrometric network evaluation for Canadian watershed. J. Hydrol. 380 (3–4), 420–437. http://dx.doi.org/10.1016/j.jhydrol.2009.11.015.

Mishra, A.K., Coulibaly, P., 2014. Variability in Canadian seasonal streamflow information and its implication for hydrometric network design. J. Hydrol. Eng. 19 (8), 05014003. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000971.

Mogheir, Y., Singh, V. P., 2002. Application of information theory to groundwater quality monitoring networks. Water Resour. Manage. 16(1): 37-49. http://dx.doi.org/10.1023/A:1015511811686.

Mogheir, Y., de Lima, J. L. M. P., Singh, V. P., 2003. Assessment of spatial structure of groundwater quality variables based on the entropy theory. Hydrol. Earth Syst. Sci. 7(5): 707-721. http://dx.doi.org/10.5194/hess-7-707-2003.

Mogheir, Y., de Lima, J. L. M. P., Singh, V. P., 2004. Characterizing the spatial variability of groundwater quality using the entropy theory: I. Synthetic data.

Hydrol Process. 18(11): 2165-2179. http://dx.doi.org/10.1002/hyp.1465.

Mogheir, Y., de Lima, J. L. M. P., Singh, V. P., 2004. Characterizing the spatial variability of groundwater quality using the entropy theory: II. Case study from Gaza Strip. Hydrol Process. 18:2579-2590. http://dx.doi.org/10.1002/hyp.1466.

Mogheir, Y., Singh, V. P., de Lima, J. L. M. P., 2006. Spatial assessment and redesign of a groundwater quality monitoring network using entropy theory, Gaza Strip, Palestine. Hydrogeol. J. 14 (5), 700-712. http://dx.doi.org/10.1007/s10040-005-0464-3.

Mondal, N. C., Singh, V. P., 2012. Evaluation of groundwater monitoring network of Kodaganar River basin from Southern India using entropy. Environ. Earth Sci. 66(4), 1183-1193. http://dx.doi.org/10.1007/s12665-011-1326-z.

Moss, M. E., 1979. Some basic considerations in the design of hydrologic data networks. Water Resour. Res. 15(6), 1673-1676. http://dx.doi.org/10.1029/WR015i006p01673.

Owlia, R. R., Abrishamchi, A., Tajrishy, M., 2011. Spatial-temporal assessment and redesign of groundwater quality monitoring network: A case study. Environ. Monit. Assess. 172(1-4), 263-273. http://dx.doi.org/10.1007/s10661-010-1332-8.

Ozkul, S., Harmancioglu, N. B., Singh, V. P., 2000. Entropy-based assessment of water quality monitoring networks. J. Hydrol. Eng. 5(1), 90-100. http://dx.doi.org/10.1061/(ASCE)1084-0699(2000)5:1(90).

Papana, A., Kugiumtzis, D., 2009. Evaluation of mutual information estimators for time series. Int. J. Bifurcation Chaos. 19(12), 4197-4215. http://dx.doi.org/10.1142/S0218127409025298.

Ridolfi, E., Montesarchio, V., Russo, F., Napolitano, F., 2011. An entropy approach for evaluating the maximum information content achievable by an urban rainfall

network. Nat. Hazards Earth Syst. Sci. 11(7): 2075-2083. http://dx.doi.org/10.5194/nhess-11-2075-2011.

Ruddell, B. L., Kumar, P., 2009. Ecohydrologic process networks: 1. Identification. Water Resour. Res. 45. W03419. http://dx.doi.org/10.1029/2008WR007279.

Samuel, J., Coulibaly, P., Kollat, J., 2013. CRDEMO: combined regionalization dual-entropy-multiobjective optimization for hydrometric networks design. Water Resour. Res. 49 (12), 8070-8089. http://dx.doi.org/10.1002/2013WR014058.

Scott, D. W., 1979. On optimal and data-based histograms. Biometrika. 66(3), 605-610, http://dx.doi.org/10.1093/biomet/66.3.605.

Shannon, C. E., 1948. A mathematical theory of communication, Bell Syst. Tech. J., 27(3), 379–423, http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x.

Shimazaki, H., Shinomoto, S., 2007. A method for selecting the bin size of a time histogram. Neural Comput. 19(6), 1503-1527. http://dx.doi.org/10.1162/neco.2007.19.6.1503.

Singh, V. P., 1997. Effect of class-interval size on entropy. Stochastic Hydrol. Hydraul. 11(5), 423-431. http://dx.doi.org/10.1007/BF02427927.

Singh, V. P., 2013. Entropy theory and its application in environmental and water engineering. John Wiley & Sons.

Song, X. Y., Song, S. B., Sun, W. Y., Mu, X. M., Wang, S. Y., Li, J. Y., Li, Y., 2015. Recent changes in extreme precipitation and drought over the Songhua River Basin, China, during 1960-2013. Atmos. Res. 157, 137-152. https://doi.org/10.1016/j.atmosres.2015.01.022.

Stosic, T., Stosic, B., Singh, V. P., 2017. Optimizing streamflow monitoring networks using joint permutation entropy. J. Hydrol., 552, 306-312. https://doi.org/10.1016/j.jhydrol.2017.07.003

Sturges, H. A., 1926. The choice of a class interval, J. Am. Stat. Assoc., 21(153), 65 – 66.

Su, H.T., You, G.J.Y., 2014. Developing an entropy-based model of spatial information estimation and its application in the design of precipitation gauge networks. J. Hydrol. 519, 3316-3327. https://doi.org/10.1016/j.jhydrol.2014.10.022.

Wang, J., Liang, Z. M., Wang, D., Liu, T., Yang, J., 2016. Impact of climate change on hydrologic extremes in the upper basin of the Yellow River Basin of China. Adv. Meteorol. http://dx.doi.org/10.1155/2016/1404290.

Watanabe, S., 1960. Information theoretical analysis of multivariate correlation. IBM J. Res. Dev. 4(1), 66–82. https://doi.org/ 10.1147/rd.41.0066.

Wei, C., Yeh, H. C., Chen, Y. C., 2014. Spatiotemporal scaling effect on rainfall network design using entropy. Entropy. 16(8), 4626-4647. https://doi.org/10.3390/e16084626

Wei, Y., 1987. Variance, entropy and uncertainty measure. In Proceedings of the Survey Research Methods Section. pp. 609-610. Am. Stat. Assoc., Alexandria, Va.

World Meteorological Organization (WMO), 2008. Guide to Hydrological Practices. Volume I: Hydrology – From Measurement to Hydrological Information, sixth ed. WMO-No. 168, Geneva, Switzerland.

Yang, Y., Burn, D.H., 1994. An entropy approach to data collection network design. J. Hydrol. 157, 307-324. https://doi.org/10.1016/0022-1694(94)90111-2.

Yeh, H. C., Chen, Y. C., Chang, C. H., Ho, C. H., Wei, C., 2017. Rainfall network optimization using radar and entropy. Entropy, 19(10), 553. https://doi.org/10.3390/e19100553.

Yoo, C., Jung, K., Lee, J., 2008. Evaluation of rain gauge network using entropy theory: comparison of mixed and continuous distribution function applications. J.

Hydrol. Eng. 13 (4), 226-235. https://doi.org/10.1061/(ASCE)1084-0699(2008)13:4(226).

Zhang, J. G., Wang, H. M., Singh, V. P., 2011. Information Entropy of a Rainfall Network in China. In: Wu D., Zhou Y. (Eds.). Modeling Risk Management for Resources and Environment in China. Computational Risk Management. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-18387-4_2.

Zhang, Q., Peng, J. T., Singh, V. P., Li, J. F., Chen, Y. Q. D., 2014. Spatio-temporal variations of precipitation in arid and semiarid regions of China: The Yellow River basin as a case study. Glob. Planet. Chang. 114, 38-49. https://doi.org/10.1016/j.gloplacha.2014.01.005.

List of Tables

List of Figures

Figure 9. Ranking variations for stations in Shanghai (SEH and SWH) during sliding observation period with 1-year, 2-year and 5-year time windows

Figure 10. Ranking variations for stations in Shanghai (XPL and XMH) during sliding observation period with 1-year, 2-year and 5-year time windows

Figure 11. Ranking variation with sliding time window for 5 stations in SEH and SWH

Figure 12. Ranking variation with sliding time window for 5 stations in XPL and XMH

Figure 13. Optimal networks of wet season, dry season and entire series

Table 1. Recommended minimum densities of stations (by WMO and China) and current densities of stations for study areas

The Guide to Hydrological Practices, 2008, WMO		Study areas			
Physiographic unit	Recommended minimum densities of stations (area in km^2 per station)	SEH	SWH	XPL	XMH
		Current densities (area in km^2 per station)			
Coastal	900	108	113		
Mountains	250				
Hilly/undulating	575				194
Interior plains	575			182	
Small islands	25				
Urban areas	–				
Polar/arid	10000				
Technical regulations for hydrologic network design, SL34-2013, P.R. China		Study areas			
Area(km^2)	Number of stations	SEH	SWH	XPL	XMH
		Current number of stations			
<10	2				
50	3				
100	4				
200	5				
500	7				
1000	8				
2000	10	23	24		

3000	12			24	29
Geographical type of area	Control area per station(km^2)	Current densities (area in km^2 per station)			
plain river network	150	108	113		
other area	200			182	194

Note: SEH: east of Huangpu River in Shanghai; SWH: west of Huangpu River in Shanghai; XPL: plain area in Xi'an; XMH: mountain and hilly area in Xi'an.

Table 2. Entropy values with different binning estimation (unit: bits)

Binning estimation	Study area	Maximum marginal entropy	Average marginal entropy	Joint entropy	Total correlation	TC/N
a=1 mm/d	SEH	2.41	2.33	3.71	49.84	2.17
	SWH	2.37	2.31	3.81	51.59	2.15
	XPL	1.94	1.6	2.73	35.62	1.48
	XMH	2.21	1.83	2.9	50.14	1.73
a=3 mm/d	SEH	1.69	1.63	2.93	34.62	1.51
	SWH	1.66	1.61	2.94	35.67	1.49
	XPL	1.36	1.13	2.09	25.1	1.05
	XMH	1.56	1.31	2.38	35.52	1.22
Scott	SEH	**2.52**	**2.34**	**6.26**	47.63	2.07
	SWH	**2.63**	**2.37**	**6.16**	50.6	2.11
	XPL	**1.99**	**1.65**	**5.66**	33.83	1.41
	XMH	**2.24**	**1.87**	**6.99**	47.15	1.63
Sturge	SEH	1.58	1.53	5.62	29.58	1.29
	SWH	1.57	1.51	5.54	30.75	1.28
	XPL	1.27	1.01	5.21	19.1	0.8
	XMH	1.49	1.18	6.67	27.5	0.95
Bendat & Piersol	SEH	2.3	2.17	6.22	43.71	1.9
	SWH	2.29	2.14	6.13	45.18	1.88
	XPL	1.89	1.52	5.62	30.95	1.29
	XMH	2.18	1.75	6.98	43.86	1.51

Note:

1. SEH: east of Huangpu River in Shanghai; SWH: west of Huangpu River in Shanghai; XPL: plain area in Xi'an; XMH: mountain and hilly area in Xi'an.
2. The ten-year series (2005-2016) was used to calculate the entropy values.
3. TC: total correlation; N: number of stations.

Table 3. MIMR results for study areas with entire 10-year series

	SEH	SWH	XPL	XMH
Selected number of stations	12	13	17	21
Set of selected stations	{21,22,17,1,15,11,19,5,10, 13,18,2}	{12,4,18,17,16,3,24,13, 6,2,20,5,14}	{6,22,5,2,14,11,8,20,1, 21,17,10,15,9,18,12,4}	{15,6,23,28,24,17,26,19,3,25,18,29,10,21,22,11,7,20,8,27,5}
Percentage (selected/total)	52.17%	54.17%	70.83%	72.41%
Joint entropy	5.97	5.87	5.38	6.66
Transinformation	25.65	26.01	12.20	15.90
Total correlation	21.88	24.40	21.57	31.14
MIMR value	20.92	20.63	9.75	11.82

Note: SEH: east of Huangpu River in Shanghai; SWH: west of Huangpu River in Shanghai; XPL: plain area in Xi'an; XMH: mountain and hilly area in Xi'an.

Table 4. Six scenarios for SWH

Scenarios	Scenario 1	Scenario 2	Scenario 3
Starting date	2008/3/1	2008/3/1	2008/3/1
Ending date	2009/3/1	2010/3/1	2013/3/1
Joint entropy	**5.39**	**5.24**	**5.60**
Transinformation	32.35	22.15	26.20
Total correlation	**6.60**	**20.19**	**20.54**
MIMR value	28.87	17.88	21.40
Selected number	6	13	12
Selected set	{12,4,6,11,18,9,2,20,8,15,21,14,1,17}	{12,15,6,7,11,8,9,2,21,1,20,14,17,10,8,1,17,3}	{6,5,24,2,21,11,9,1,20,14,17,10,8,22,18,15,16,4}
Scenarios	Scenario 4	Scenario 5	Scenario 6
Starting date	2009/4/5	2009/4/5	2009/4/5
Ending date	2010/4/5	2011/4/5	2014/4/5
Joint entropy	**4.15**	**4.45**	**5.14**
Transinformation	29.03	25.91	22.77
Total correlation	**8.41**	**13.87**	**23.77**
MIMR value	24.86	21.51	17.58
Selected number	7	10	13
Selected set	{10,17,24,2,11,5,18}	{10,17,24,14,5,18,22,3,4,20}	{11,5,18,1,16,4,24,3,13,6,2,20,9}

Table 5. Six scenarios for XPL

Selected set	{6,16,2,12,11,21,20,9,8,10,1,22,17,24,14,15,5}	{6,23,2,13,11,10,9,21,1,8,20,17,14,5,22,15,4}	{6,23,2,12,11,9,21,10,1,20,14,22,17,8,13,15,4}

Table 6. RDI for top 5 stations in descending order in study areas

Window length			Top 5 stations with highest RDI				
SEH	1-year	Station No.	18	22	16	20	5
		RDI	0.97	0.96	0.96	0.96	0.95
	2-year	Station No.	23	5	20	9	4
		RDI	0.94	0.93	0.93	0.92	0.92
	5-year	Station No.	23	20	4	5	15
		RDI	0.87	0.82	0.81	0.81	0.79
SWH	1-year	Station No.	18	2	20	23	17
		RDI	0.98	0.96	0.96	0.95	0.94
	2-year	Station No.	2	17	1	14	3
		RDI	0.95	0.92	0.92	0.92	0.91
	5-year	Station No.	4	8	14	21	23
		RDI	0.93	0.86	0.83	0.83	0.83
XPL	1-year	Station No.	12	18	22	8	24
		RDI	0.96	0.93	0.92	0.89	0.88
	2-year	Station No.	8	18	22	10	13
		RDI	0.89	0.89	0.88	0.85	0.83
	5-year	Station No.	13	8	9	22	14
		RDI	0.84	0.79	0.79	0.72	0.69
XMH	1-year	Station No.	21	4	5	27	3
		RDI	0.92	0.91	0.90	0.90	0.89
	2-year	Station No.	2	1	3	5	6
		RDI	0.88	0.88	0.86	0.85	0.85
	5-year	Station No.	2	4	1	21	8
		RDI	0.88	0.83	0.79	0.78	0.77

Note: SEH: east of Huangpu River in Shanghai; SWH: west of Huangpu River in Shanghai; XPL: plain area in Xi'an; XMH: mountain and hilly area in Xi'an.

Table 7. Entropy values of optimal networks under different meteorological conditions

	Meteorological condition	Total station number	Selected number	Overlapping ratio	Joint entropy	Transinformation	Total correlation	MIMR value
SEH	wet season	23	12	75%	6.35	25.29	21.15	21.08
	dry season		13	69.23%	4.88	21.92	22.57	16.92
	entire series		12	—	5.97	25.65	21.88	20.92
SWH	wet season	24	14	57.14%	6.3	23.45	26.06	18.59
	dry season		14	78.57%	4.75	22.22	25.12	16.56
	entire series		13	—	5.87	26.01	24.40	20.63
XPL	wet season	24	19	73.68%	6.29	10.38	30.40	7.26
	dry season		19	89.47%	4.09	6.31	16.76	4.97
	entire series		17	—	5.38	12.20	21.57	9.75
XMH	wet season	29	19	78.95%	7.49	23.01	32.66	17.87
	dry season		23	78.26%	5.18	9.31	24.90	6.61
	entire series		21	—	6.66	15.90	31.14	11.82

Note:

1. The overlapping ratio (wet season or dry season) is calculated from selected number of stations that were the same as those in the network for entire series divided by total selected number in this season.
2. SEH: east of Huangpu River in Shanghai; SWH: west of Huangpu River in Shanghai; XPL: plain area in Xi'an; XMH: mountain and hilly area in Xi'an.

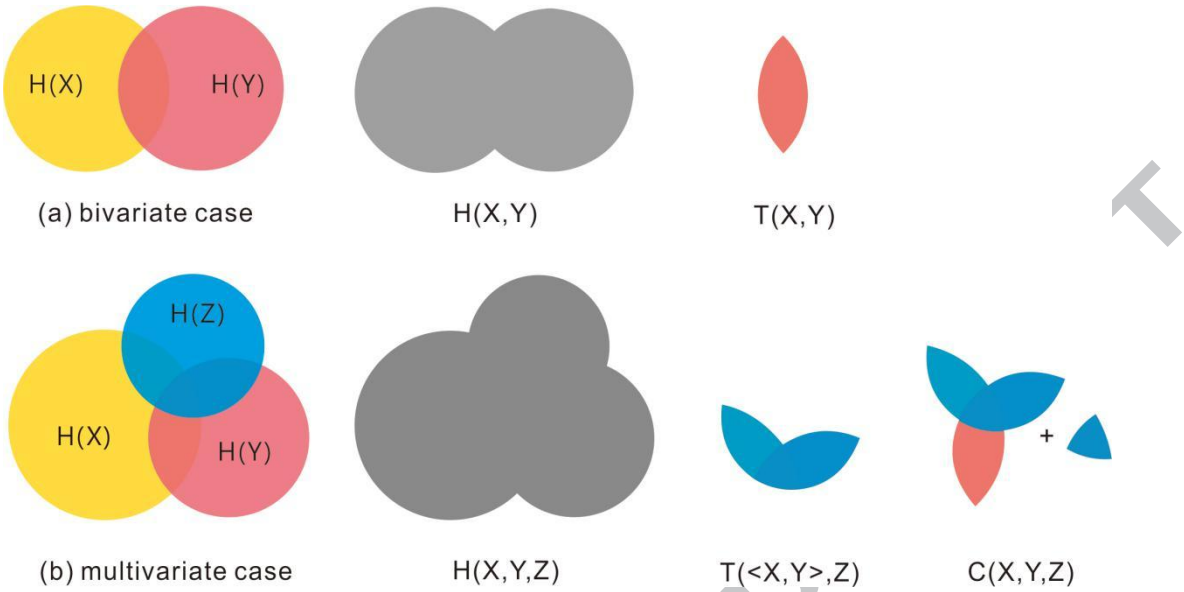


Figure 1. Venn diagrams for bivariate and multivariate entropy

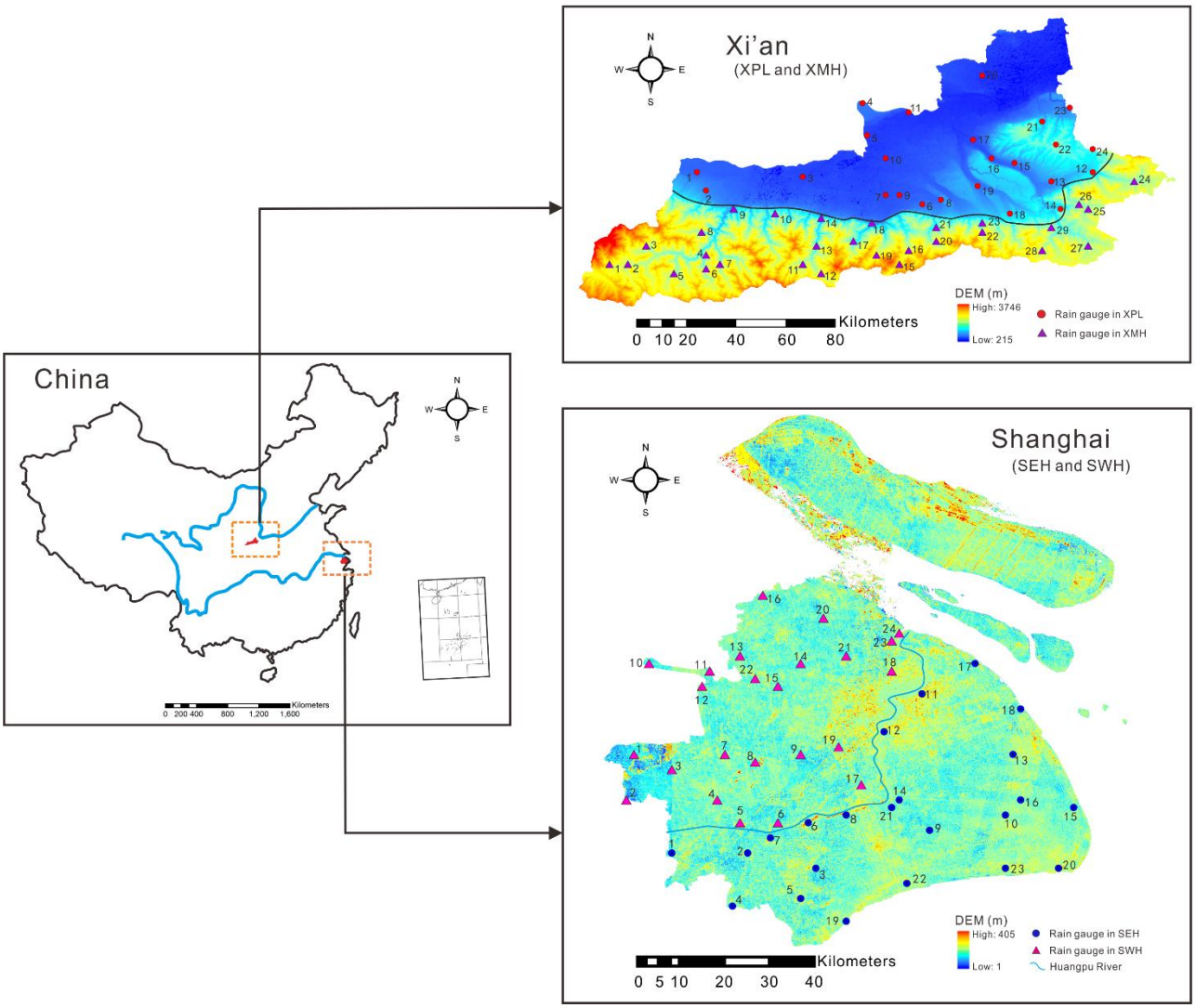


Figure 2. Study areas and rainfall networks in Shanghai and Xi'an

Note: XPL: plain area in Xi'an; XMH: mountain and hilly area in Xi'an; SWH: west

of Huangpu River in Shanghai; SEH: east of Huangpu River in Shanghai.

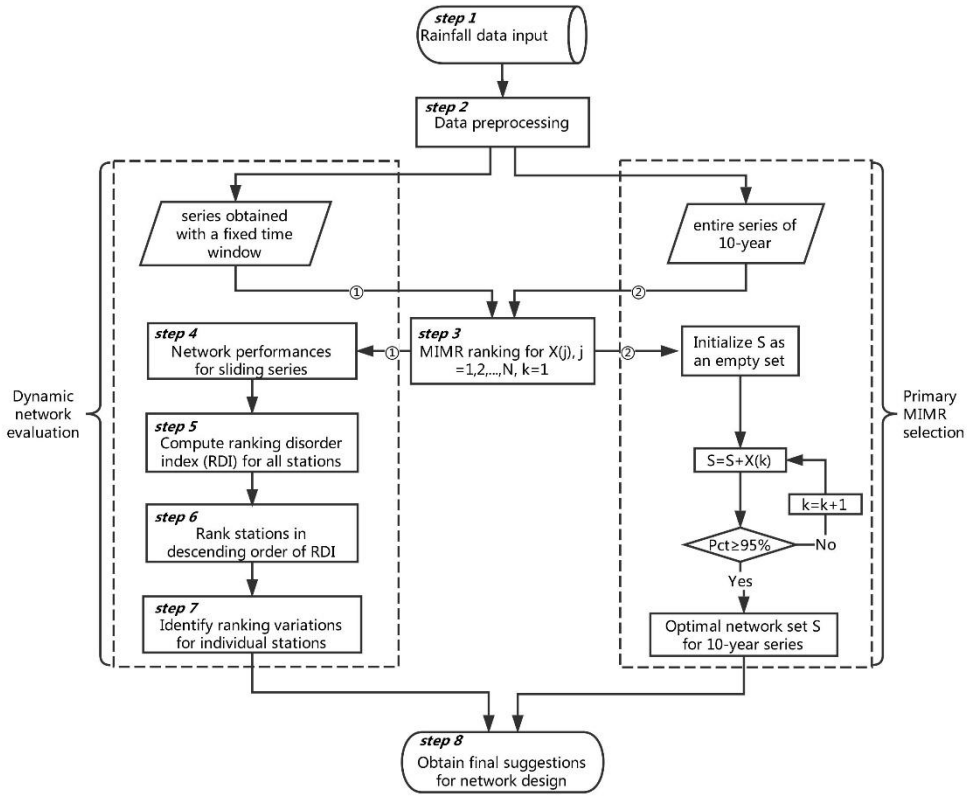


Figure 3. Flowchart of dynamic network evaluation framework

Note:

1. Candidate stations are denoted as X(i), i=1, 2, … , N (N is the total number of stations in the current network).
2. S is the set of selected stations in the optimal network (initialized as an empty set before MIMR selection.
3. Pct is a parameter of stopping criterion in MIMR selection, see in section 4.1.

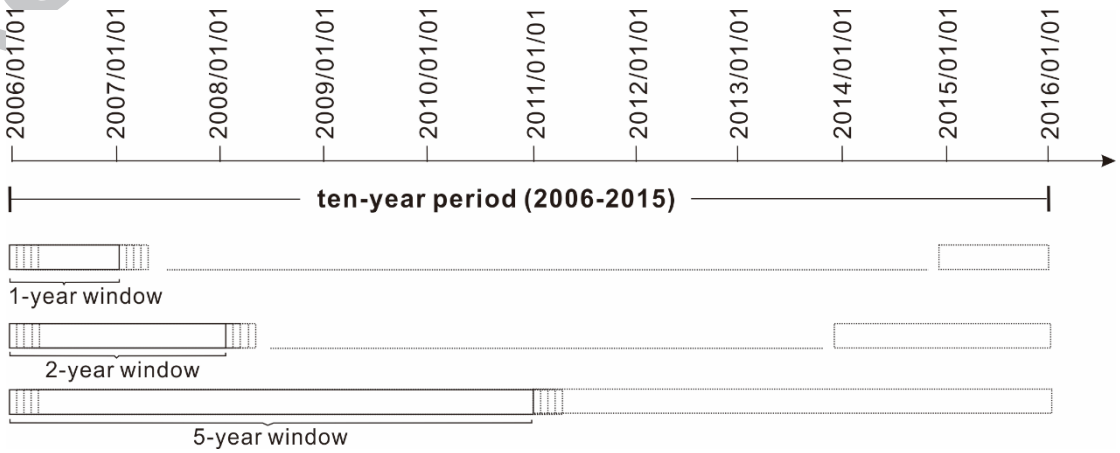
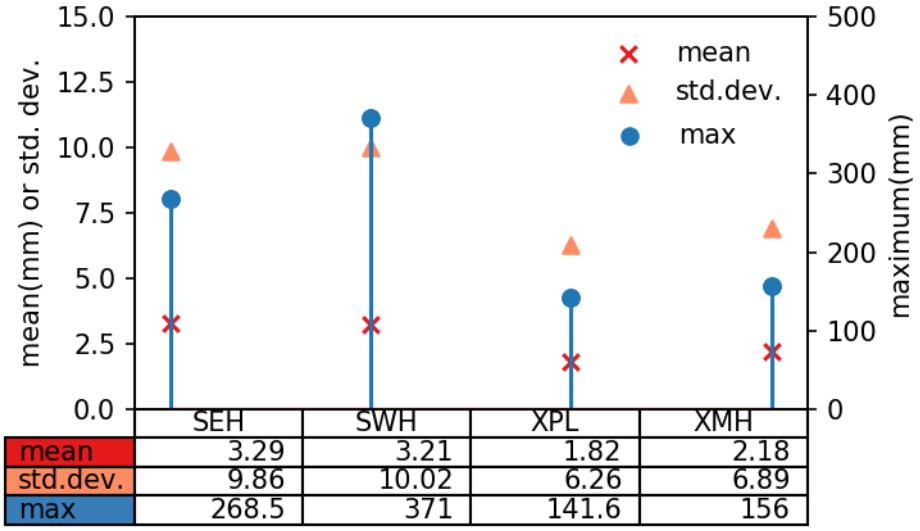


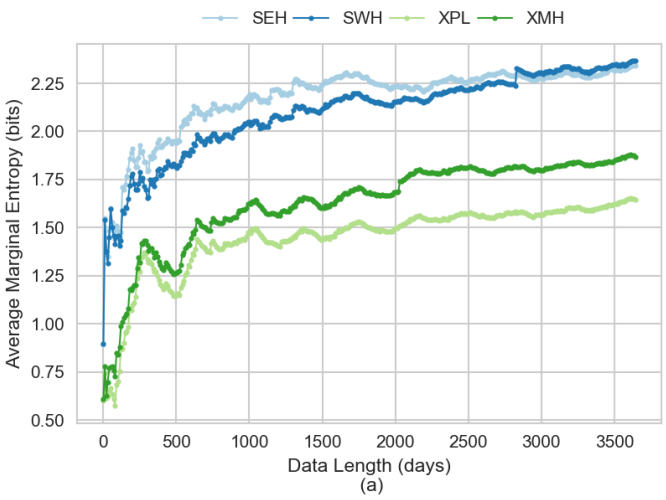
Figure 4. Diagram for fixing time window length as 1, 2, 5-year and shifting starting days



	SEH	SWH	XPL	XMH
mean	3.29	3.21	1.82	2.18
std.dev.	9.86	10.02	6.26	6.89
max	268.5	371	141.6	156

Figure 5. Basic statistical properties of precipitation

Note: SEH: east of Huangpu River in Shanghai; SWH: west of Huangpu River in Shanghai; XPL: plain area in Xi'an; XMH: mountain and hilly area in Xi'an.



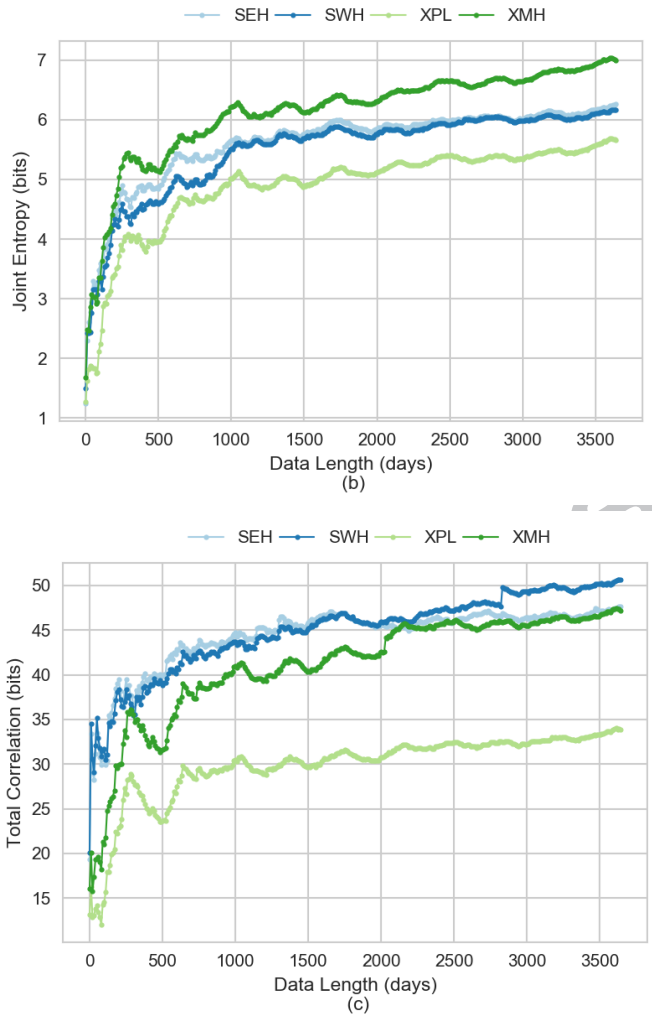


Figure 6. Variation of entropy values under different lengths of time series (a) average marginal entropy; (b) joint entropy; (c) total correlation

Note: SEH: east of Huangpu River in Shanghai; SWH: west of Huangpu River in Shanghai; XPL: plain area in Xi'an; XMH: mountain and hilly area in Xi'an.

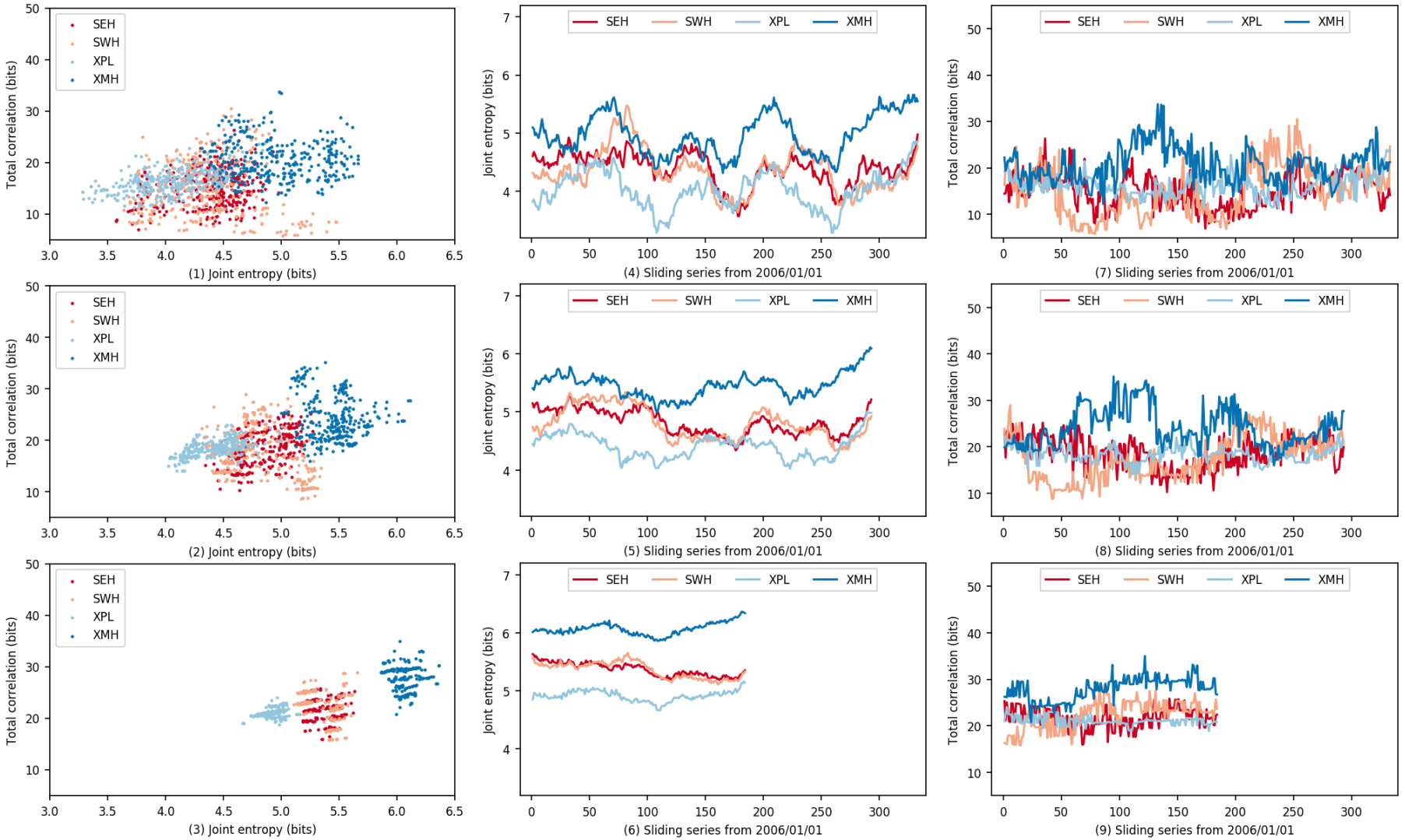


Figure 7. Entropy values for optimal networks by MIMR under sliding series with window lengths of 1-year, 2-year and 5-year

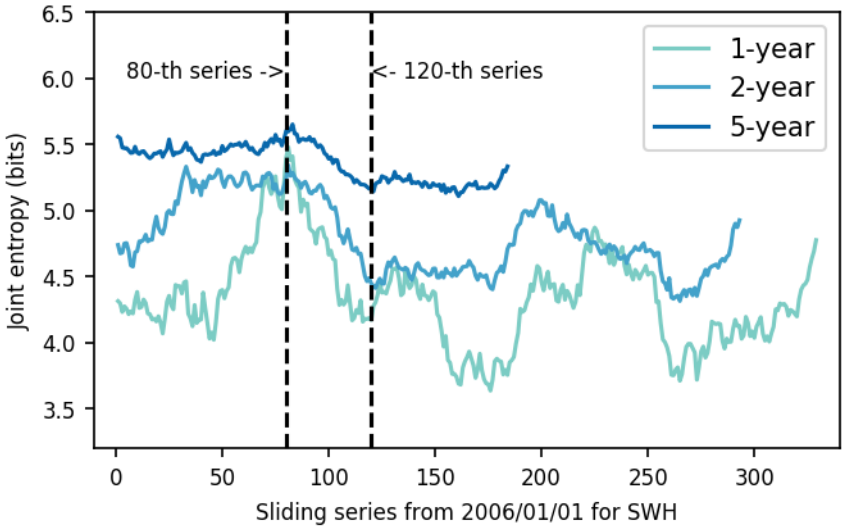
Note:

1. Description of subplots:

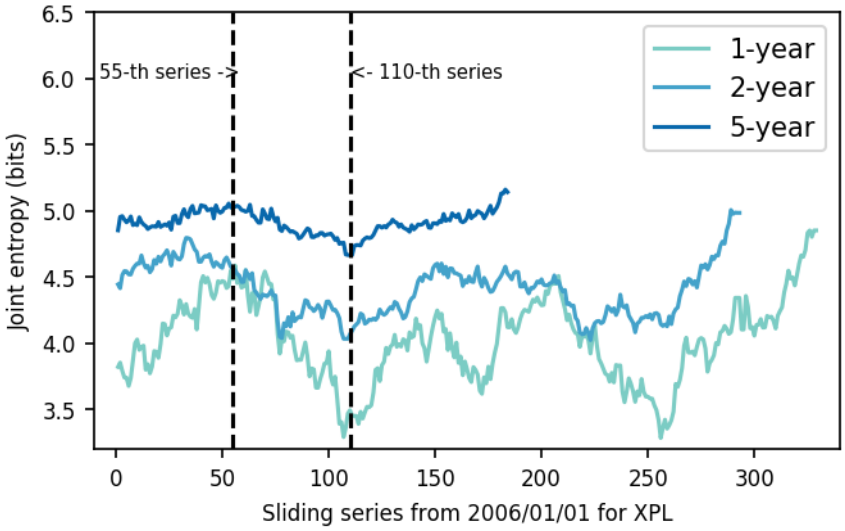
(1)	(4)	(7)
(2)	(5)	(8)
(3)	(6)	(9)

(1) Joint entropy and total correlation of optimal networks from 1-year sliding series
(2) Joint entropy and total correlation of optimal networks from 2-year sliding series
(3) Joint entropy and total correlation of optimal networks from 5-year sliding series
(4) Variations of joint entropy of optimal networks from 1-year sliding series
(5) Variations of joint entropy of optimal networks from 2-year sliding series
(6) Variations of joint entropy of optimal networks from 5-year sliding series
(7) Variations of total correlation of optimal networks from 1-year sliding series
(8) Variations of total correlation of optimal networks from 2-year sliding series
(9) Variations of total correlation of optimal networks from 5-year sliding series

2. SEH: east of Huangpu River in Shanghai;
SWH: west of Huangpu River in Shanghai;
XPL: plain area in Xi'an;
XMH: mountain and hilly area in Xi'an.



(a)



(b)

Figure 8. Scenarios selected from sliding series with high and low joint entropy for SWH and XPL

Note: SWH: west of Huangpu River in Shanghai; XPL: plain area in Xi'an.

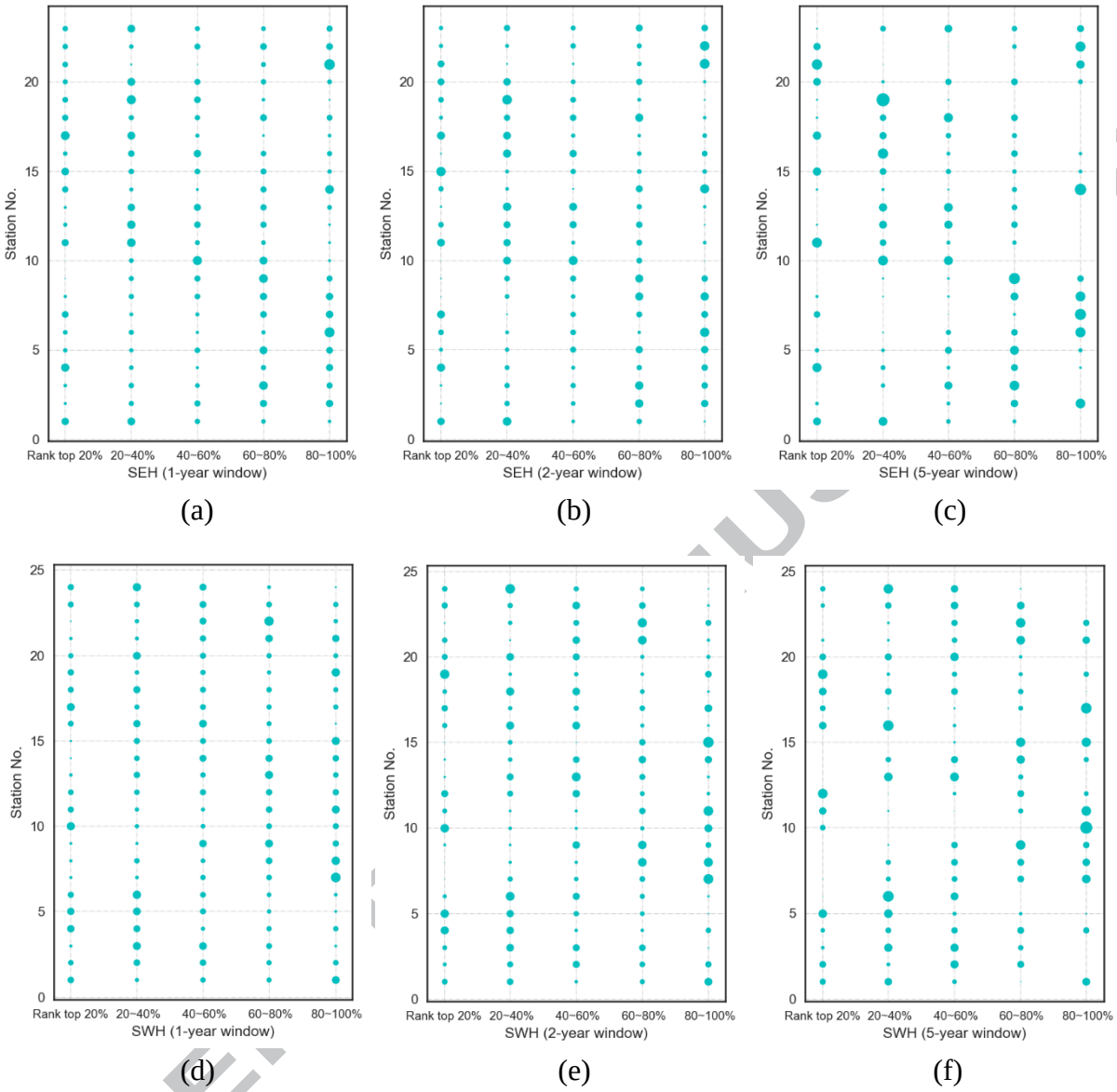


Figure 9. Ranking variations for stations in Shanghai (SEH and SWH) during sliding observation period with 1-year, 2-year and 5-year time windows

Note:

1. The size of bubble represented the frequency of the occurrence of the ranks, e.g., if the station was ranked top 20% for n times and the total number of ranking sets was N (equal to the number of sliding series), then the frequency was n/N.
2. SEH: east of Huangpu River in Shanghai; SWH: west of Huangpu River in Shanghai.

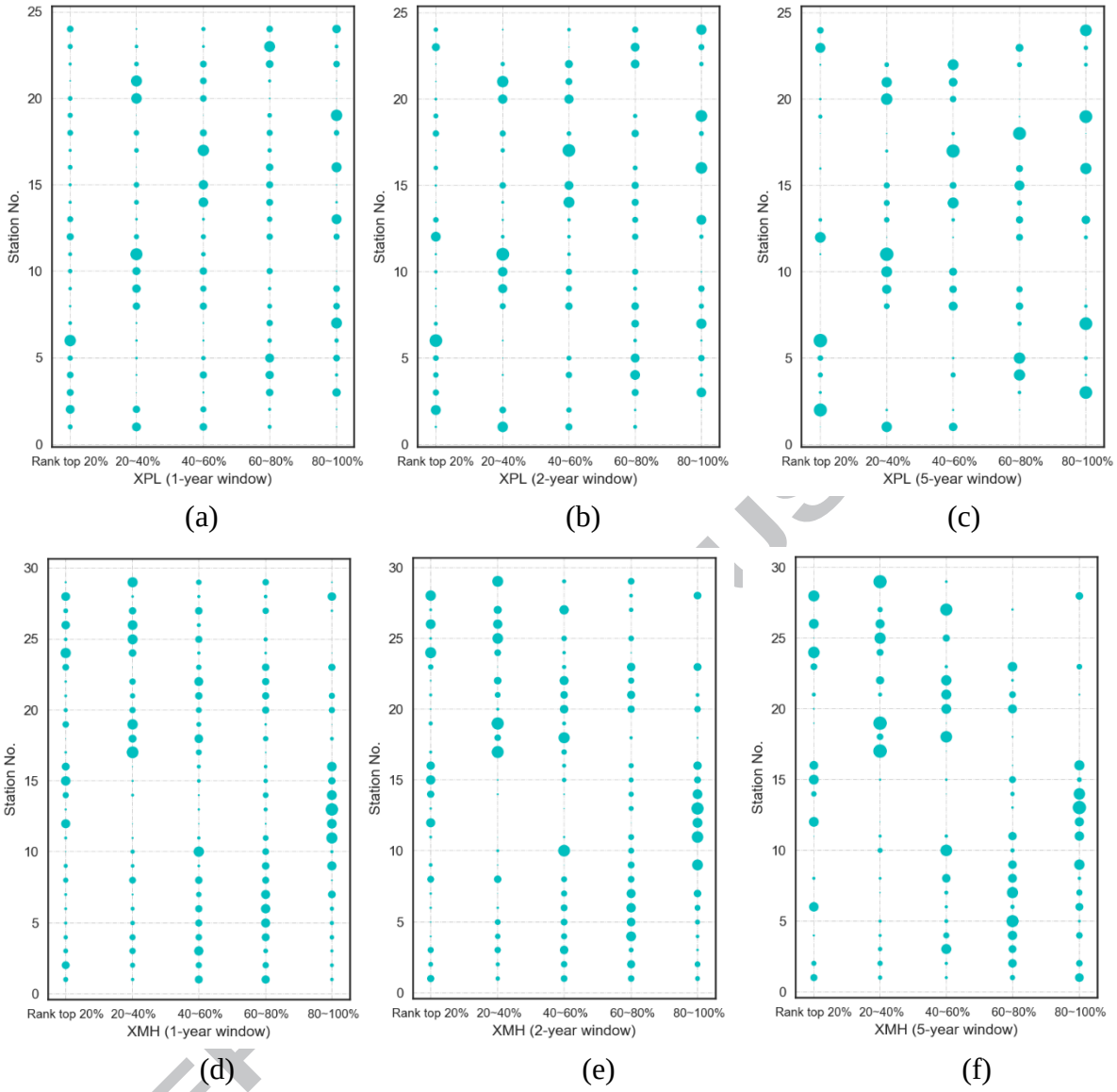
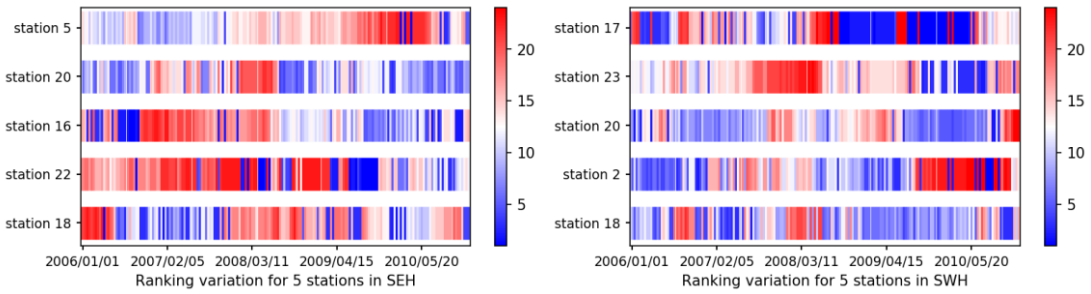


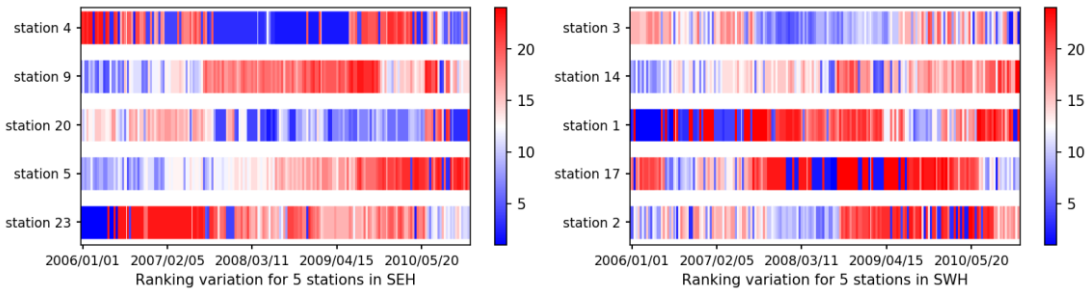
Figure 10. Ranking variations for stations in Xi'an (XPL and XMH) during sliding observation period with 1-year, 2-year and 5-year time windows

Note:

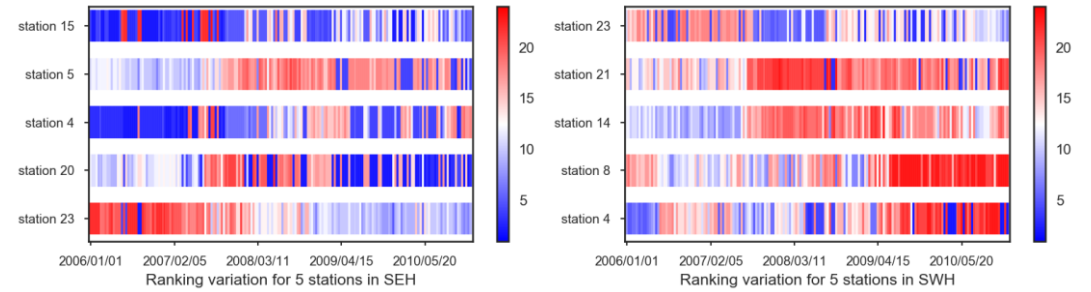
1. The size of bubble represented the frequency of the occurrence of the ranks, e.g., if the station was ranked top 20% for n times and the total number of ranking sets was N (equal to the number of sliding series), then the frequency was n/N.
2. XPL: plain area in Xi'an; XMH: mountain and hilly area in Xi'an.



(1) SEH, 1-year window (4) SWH, 1-year window



(2) SEH, 2-year window (5) SWH, 2-year window

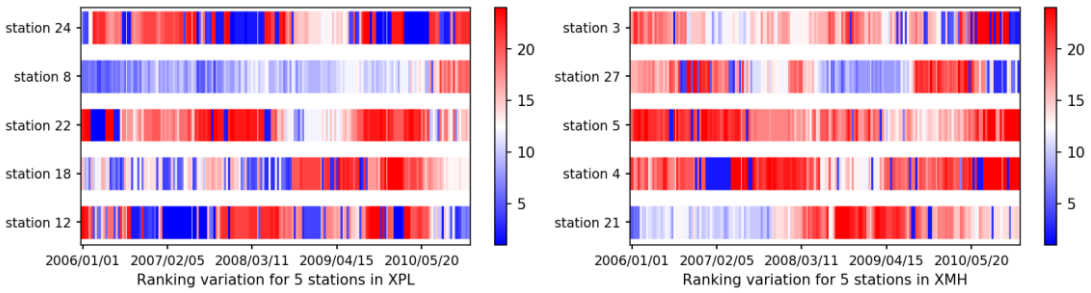


(3) SEH, 5-year window (6) SWH, 5-year window

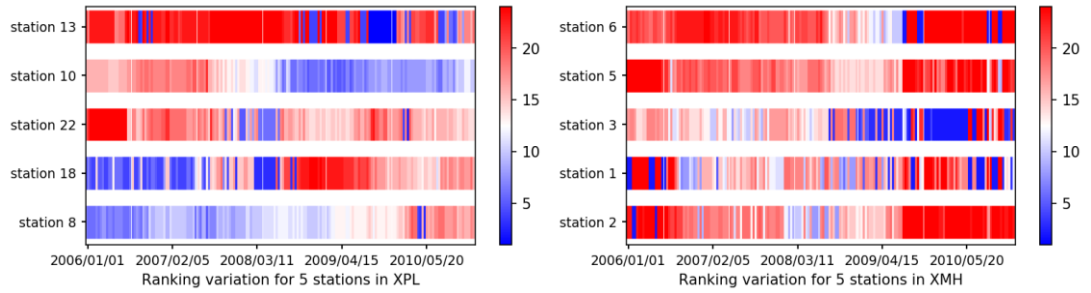
Figure 11. Ranking variation with sliding time window for 5 stations in SEH and SWH

Note:

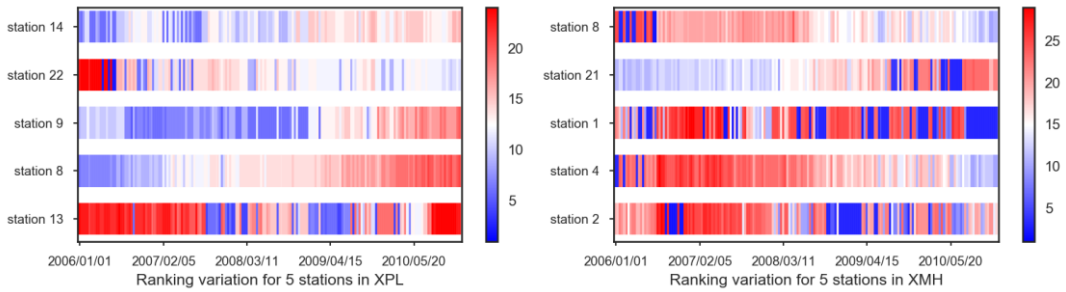
1. SEH: east of Huangpu River in Shanghai; SWH: west of Huangpu River in Shanghai.
2. A lower rank represents higher importance of a station, and vice versa.



(1) XPL, 1-year window (4) XMH,1-year window



(2) XPL, 2-year window (5) XMH, 2-year window



(3) XPL, 5-year window (6) XMH, 5-year window

Figure 12. Ranking variation with sliding time window for 5 stations in XPL and XMH

Note:

1. XPL: plain area in Xi'an; XMH: mountain and hilly area in Xi'an.
2. A lower rank represents higher importance of a station, and vice versa.

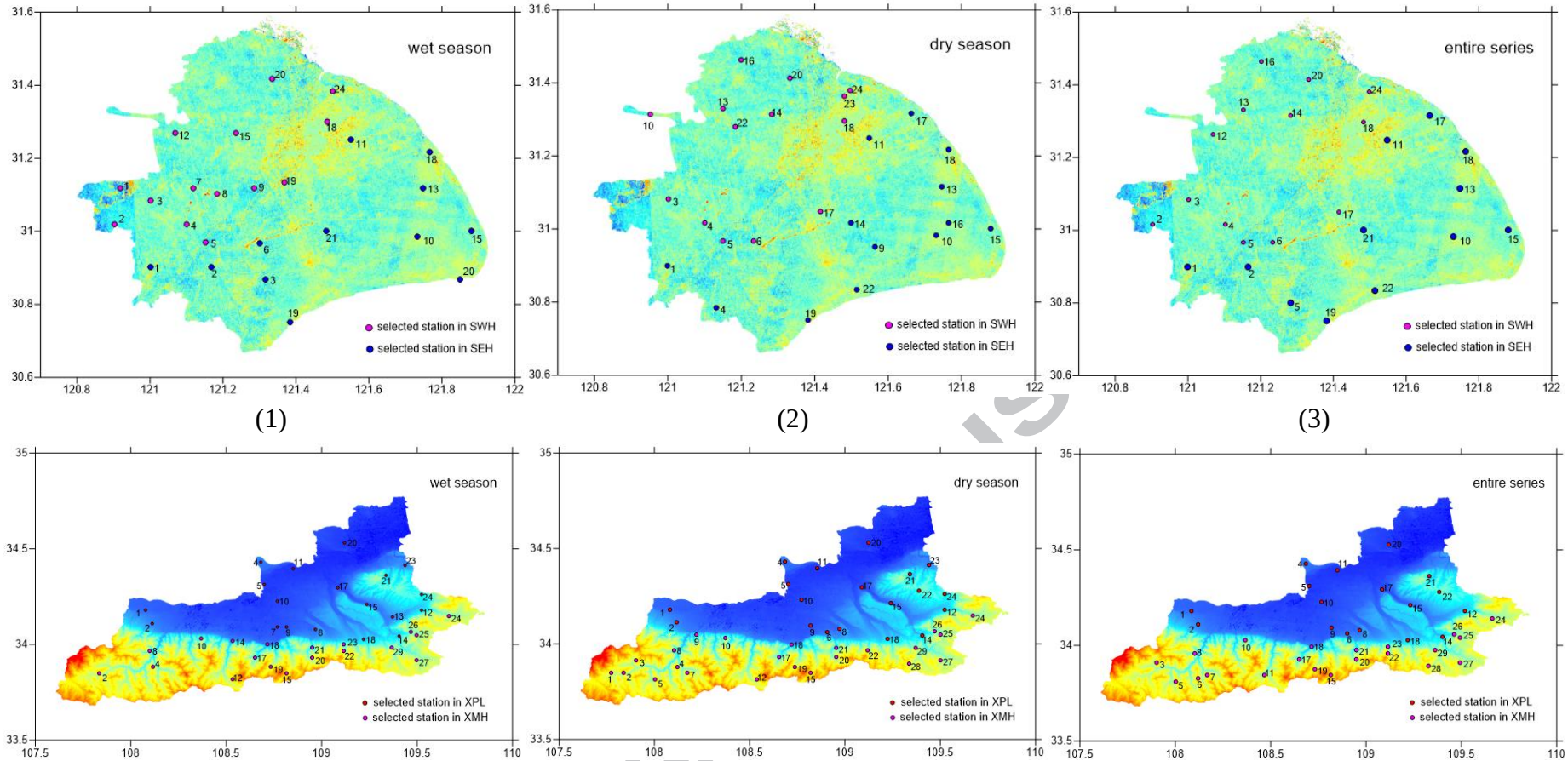


Figure 13. Optimal networks of wet season, dry season and entire series (Shanghai: (1), (2), (3); Xi'an: (4), (5), (6))