AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE

10.1002/2016WR019981

Information theory-based decision support system for integrated design of multivariable hydrometric networks

Jongho Keum[1] and **Paulin Coulibaly**[1,2]

[1]Department of Civil Engineering, McMaster University, Hamilton, Ontario, Canada, [2]School of Geography and Earth Sciences, McMaster University, Hamilton, Ontario, Canada

Key Points:

- A decision support system for integrated hydrometric networks design and evaluation was developed
- Entropy measures and multiobjective optimization were combined to find optimal networks simultaneously for multiple hydrologic variables
- Four quantization cases for calculating discrete entropy and their effects on the network design were also evaluated

Correspondence to:
J. Keum,
jkeum@mcmaster.ca

Citation:
Keum, J., and P. Coulibaly (2017), Information theory-based decision support system for integrated design of multivariable hydrometric networks, *Water Resour. Res.*, *53*, doi:10.1002/2016WR019981.

Received 21 OCT 2016
Accepted 24 JUN 2017
Accepted article online 5 JUL 2017

Abstract Adequate and accurate hydrologic information from optimal hydrometric networks is an essential part of effective water resources management. Although the key hydrologic processes in the water cycle are interconnected, hydrometric networks (e.g., streamflow, precipitation, groundwater level) have been routinely designed individually. A decision support framework is proposed for integrated design of multivariable hydrometric networks. The proposed method is applied to design optimal precipitation and streamflow networks simultaneously. The epsilon-dominance hierarchical Bayesian optimization algorithm was combined with Shannon entropy of information theory to design and evaluate hydrometric networks. Specifically, the joint entropy from the combined networks was maximized to provide the most information, and the total correlation was minimized to reduce redundant information. To further optimize the efficiency between the networks, they were designed by maximizing the conditional entropy of the streamflow network given the information of the precipitation network. Compared to the traditional individual variable design approach, the integrated multivariable design method was able to determine more efficient optimal networks by avoiding the redundant stations. Additionally, four quantization cases were compared to evaluate their effects on the entropy calculations and the determination of the optimal networks. The evaluation results indicate that the quantization methods should be selected after careful consideration for each design problem since the station rankings and the optimal networks can change accordingly.

1. Introduction

Gathering quality hydrologic data sets from reliable hydrometric networks is the first step toward sustainable and effective water resources planning and management. Although emerging satellite-based remote sensing techniques can provide valuable hydrologic information, adequate ground-based monitoring networks have always been essential to collect accurate in-situ data and to validate remote measurements. While the importance of establishing efficient monitoring networks is well documented [*Mishra and Coulibaly*, 2009; *Chacon-Hurtado et al.*, 2016], the shrinking hydrometric networks primarily due to financial budget cuts have been a global issue [*Pilon et al.*, 1996; *U.S. Geological Survey*, 1999; *Burn*, 2009; *Mishra and Coulibaly*, 2009]. Thus, it is essential to design efficient hydrometric networks, which can provide quality information with minimal redundancies by selecting the optimal number of monitoring stations as well as their locations [*Li et al.*, 2012]. Although the World Meteorological Organization (WMO) has suggested design guidelines for the minimum monitoring network densities [*World Meteorological Organization*, 2008], these guidelines are not about the optimal hydrometric network design which requires resorting to other methods.

Various methodologies have been investigated for the design of hydrometric networks [*Mishra and Coulibaly*, 2009; *Behmel et al.*, 2016; *Chacon-Hurtado et al.*, 2016]. In recent reviews on hydrometric network design, network evaluation methods can be broadly summarized and categorized in statistical analysis, spatial interpolation, information theory applications, basin physiographic characteristics, hybrid methods, and methods based on expert recommendations [*Mishra and Coulibaly*, 2009; *Chacon-Hurtado et al.*, 2016]. Also, *Keum and Kaluarachchi* [2015] combined a spatially distributed water quality model and the network density concept to determine adequate numbers of the salinity monitoring at a watershed scale in a large basin. *Halverson and Fleming* [2015] adapted complex network theory to hydrometric network design and conducted betweenness analyses to evaluate key stations. *Dai et al.* [2017] suggested a design method for rain gauge network using satellite rainfall. The authors applied principal component analysis to evaluate network redundancy and conducted cluster analysis to determine the potential locations of rain gauges.

Among the network design methodologies, the entropy-based design approach has been widely applied after the development of Shannon's information theory (see section 3.1 entropy terms) [*Shannon*, 1948; *Singh*, 1997]. Entropy concept was initially adopted to water resources for characterizing uncertainty in hydrologic time series [e.g., *Amorocho and Espildora*, 1973]. The entropy applications have been extended to hydrometric network design as *Caselton and Husain* [1980] utilized transinformation, which is the amount of information that can be transferred from or to others, to design a precipitation network. The networks interpreted by the transinformation tend to have smaller transinformation so that the stations can be more independent to each other and include less duplicated information. Since the initial studies, the entropy theory has been used to evaluate and design hydrometric networks [e.g., *Husain*, 1987, 1989; *Krstanovic and Singh*, 1992a,b; *Yang and Burn*, 1994]. In specific, *Alfonso et al*. [2010a, 2010b, 2013] evaluated water level and discharge monitoring networks by maximizing the joint entropy and minimizing the total correlation. *Mishra and Coulibaly* [2010] used net entropy to evaluate Canadian hydrometric networks. *Yoo et al*. [2008] evaluated rain gauge networks by comparing the applicability of rainfall variability and intermittency using daily data. *Samuel et al*. [2013] combined regionalization methods with entropy measures and multiobjective optimization to find optimal locations for streamflow network extension out of a set of potential stations. *Kornelsen and Coulibaly* [2015] expanded the entropy-based network design to the soil moisture monitoring using the Soil Moisture and Ocean Salinity (SMOS) data set. *Leach et al*. [2015, 2016] investigated the performance of entropy theory in the design of streamflow and groundwater network with additional objectives related to physical hydrologic characteristics. *Fahle et al*. [2015] evaluated monitoring networks using data subsets, which were given by different time windows and concluded that the optimal networks from the partial subsets show better performance than ones from the entire data sets. *Alfonso et al*. [2014] introduced the ensemble entropy method to replace a single entropy value from given time series to ensemble entropies from a probability distribution. *Alfonso and Price* [2012] brought the concept of the value of information from economics area to assess the worth of additional information. Although there have been many studies on hydrometric network designs using information theory including the works mentioned above, to the best of our knowledge, previous studies have only focused on a single variable, such as precipitation, water level, streamflow, or water quality, individually.

Hydrologic variables are interconnected in the water cycle. The correlation leads to overlapping information; hence, one variable can significantly change the quantity and quality of information about another variable. For example, an extreme precipitation event may cause the increase of streamflow or groundwater level. However, to date, the studies that include these interactions in the design of hydrometric networks are very few due to the lack of an appropriate methodology [*World Meteorological Organization*, 2008; *Mishra and Coulibaly*, 2009]. Therefore, the main goal of this study is to develop a new framework that can be used to design multivariable hydrometric networks simultaneously. This research will expand on the application of entropy theory for network design by developing a decision-making framework. The proposed methodology will provide comprehensive hydrologic information from the networks while accounting for the interconnection of the hydrologic variables and their monitoring networks in an integrated manner. As a demonstration of the proposed integrated network design methodology, the optimal hydrometric networks for streamflow and precipitation were designed in the combined Hamilton, Halton and Credit Valley watersheds in southern Ontario, Canada.

In addition, although many studies about entropy-based network design have used discrete entropy because of the difficulties in calculating joint entropy of non-Gaussian distributions [*Alfonso et al*., 2010a], only a few have discussed the sensitivity of the quantization of the continuous observations, which can alter the shape of histogram and entropy calculations. Some of the previous studies [e.g., *Markus et al*., 2003; *Alfonso et al*., 2010a; *Mishra and Coulibaly*, 2010; *Li et al*., 2012] calculated station rankings by changing class intervals and concluded that the rankings are less likely to be altered due to the different intervals. However, this conclusion was not from the different quantization methods, but from the different discretization parameters of a single method. On the other hand, *Fahle et al*. [2015] compared the Rounding method, which was adopted from the theory of communication systems to the hydrologic practices by *Alfonso et al*. [2010a], and Scott's method, which uses Gaussian density [*Scott*, 1979]. Because the latter tended to attenuate the differences between marginal entropies, they concluded the former more adequate. However, the validity of this conclusion to the multivariable network design was not verified given that the statistics of each variable are significantly different. Therefore, this study includes an additional objective which is to compare different quantization methods and their effects on the design of hydrometric networks.

2. Study Area and Monitoring Stations

The combined watersheds of the Hamilton, Halton and Credit Valley (HHCV) region in southern Ontario, Canada (Figure 1) was selected as study area. This region was initially selected as a pilot study area of entropy-based hydrometric network design through a joint project with Water Survey Canada, and was also used in previous studies [*Leach et al*., 2015; *Keum and Coulibaly*, 2017]. The HHCV watersheds are home of the most fast growing cities in south western Ontario. There is currently a plan for adding new precipitation stations in these watersheds. The results of this study could help this purpose. HHCV is located in the western part of the Lake Ontario basin, and its total drainage area is 2294 km^2. Approximately 80% of HHCV is rural agricultural or forested lands, and the remaining is urban area, which is mostly congregated along the lake shore. The landscape is mostly flat except the Niagara Escarpment. The average winter and summer temperatures are -6°C and 21°C, respectively, so that all four seasons are clearly defined in HHCV. The average annual precipitation is 910 mm, and the seasonal variation is not significant [*Kornelsen and Coulibaly*, 2013].

As shown in Figure 1, there are two precipitation stations inside HHCV that have sufficient length of continuous data (10 years in this study). Since the precipitation stations outside HHCV can also cover parts of the region, it is reasonable to include stations that are located outside but near the region. In this study, the seven precipitation stations, which have Thiessen polygons overlapping HHCV, were used. For streamflow monitoring, Environment and Climate Change Canada is operating 23 stations in HHCV, and they are distributed fairly well throughout the study area. In order to determine the optimal locations of new hydrometric stations, candidate stations need to be predefined. The potential locations of the candidate precipitation stations are chosen from Gridded Binary Code Edition 2 (GRIB2), which is an international standard of meteorological grid points for data transmission [*World Meteorological Organization*, 2003]. Specifically, the Regional Deterministic Prediction System in GRIB2 format was made on the 935 $\times$ 824 polar stereographic grids, which are equivalent to 10 km resolution at 60°N, in North America and its nearby water bodies [*Government of Canada*, 2015]. The number of potential precipitation stations (i.e., the number of the grid points in HHCV) then becomes 33. Since streamflow stations have to be representative of their drainage area, locating monitoring stations at the outlet of each catchment is reasonable. A previous hydrometric network design study by *Leach et al*. [2015] delineated catchments and defined their outlets as potential locations of streamflow stations in the same study area, HHCV. The basic criteria to define the potential locations were: the minimum catchment area should be larger than 0.1% of the total watershed area, and if the distance between two outlets is less than 2 km, only one outlet will be used as a potential location [*Samuel et al*., 2013; *Leach et al*., 2015]. In this research, the locations of the potential streamflow stations are the same as those of *Leach et al*. [2015], and the number of the stations is 137 (see Figure 1).

3. Network Design Framework

Figure 2 shows a schematic overview of the proposed network design framework. A brief description of each step and data preparation is given below, followed by the detailed methodology:

1. Obtain observed precipitation at existing stations (precipitation data at existing stations: EP).
2. Select the closest grid point to the potential precipitation stations and collect the data (precipitation data at potential stations: PP).
3. Obtain observed streamflow data at Environment and Climate Change Canada stations (streamflow data at existing stations: EQ).
4. Apply a regionalization method using observed streamflow data from (3).
5. Get streamflow estimates at potential streamflow stations from (4) (streamflow data at potential stations: PQ).
6. Quantize time series data (EP, EQ, PP, and PQ) to calculate discrete entropy values using Rounding, Scott's, and Sturges' methods.
7. Apply entropy-based optimization for the integrated design of precipitation and streamflow monitoring networks (7–1). Single-variable network designs are also conducted individually for comparison (7–2 and 7–3).
8. Find the optimal networks from the optimization results from (7).
9. Draw monitoring hot spot maps to determine the primary locations for installing new stations.

10.1002/2016WR019981

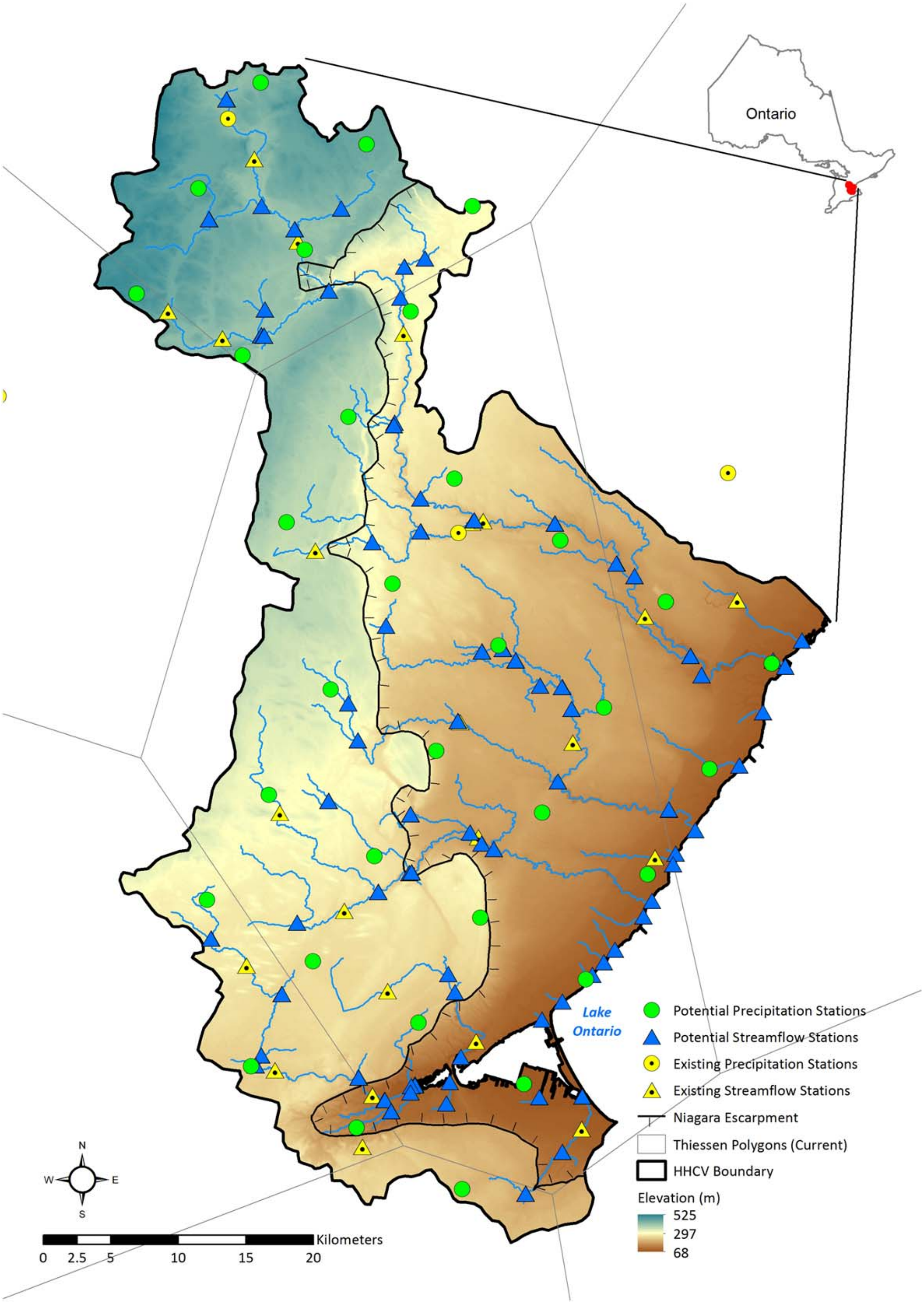

Figure 1. Study area: Hamilton, Halton, and Credit Valley region (HHCV).

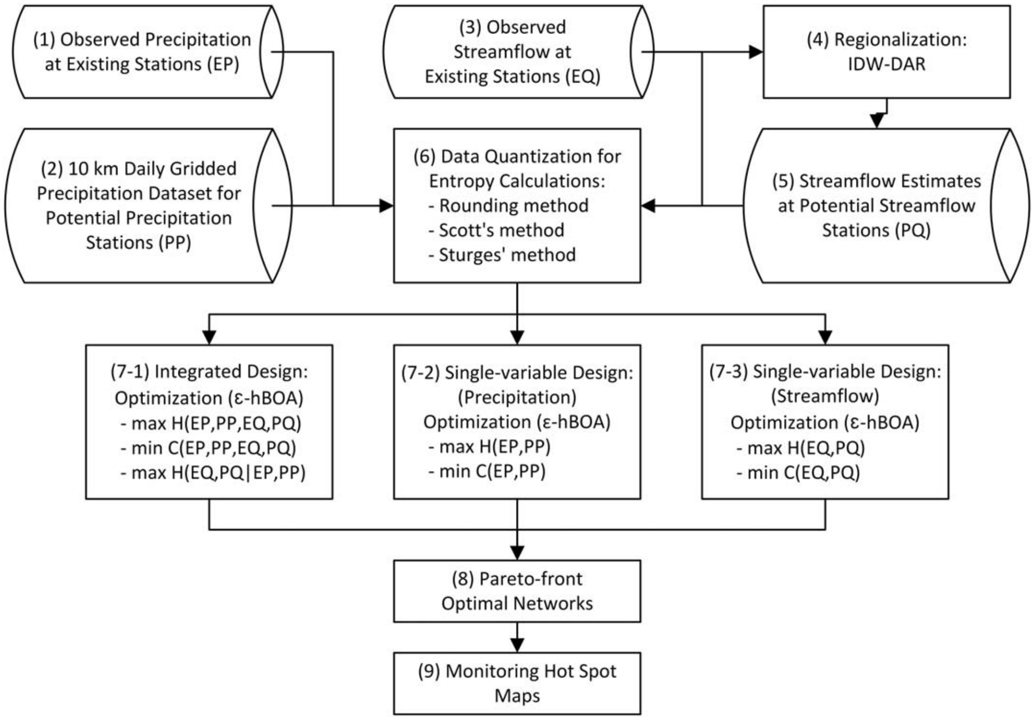

Figure 2. Flowchart illustrating the proposed framework for multivariable integrated network design.

3.1. Data Preparation

In the entropy-based hydrometric network design using daily time series, specifically for precipitation and streamflow, 10 years or more hydrologic data period is recommended to avoid potential loss of information due to short data period [*Keum and Coulibaly*, 2017]. Accordingly, daily precipitation and streamflow data from 2003 to 2012 (3653 days) were used in this study. The daily precipitation data at the existing stations (Box (1) in Figure 2) were obtained from the weather station data of Environment and Climate Change Canada [*Environment Canada*, 2014a]. Precipitation data for the potential stations in HHCV (Box (2) in Figure 2) were obtained from the 10 km daily gridded precipitation, which was created by the National Land and Water Information Service, Agriculture and Agri-Food, Canada. The gridded precipitation data were interpolated from daily observations of the Environment Canada climate stations using a thin-plate smoothing spline surface fitting method [*Natural Resources Canada*, 2015]. Again, while technological advances offer remotely sensed hydrologic data, there is still a need for calibration and validation using the ground-based measurements by hydrometric networks. This asserts the needs of efficient and optimal hydrometric network, even though there is a gridded precipitation data. Moreover, it should be noted that the gridded precipitation data set used in this study is not a satellite data but a ground-based interpolated data. Therefore, the gridded precipitation is only used for the time series of potential stations in the entropy calculations.

The streamflow data at 23 existing stations (Box (3) in Figure 2) were also retrieved from the Environment and Climate Change Canada database [*Environment Canada*, 2014b]. To apply entropy method for hydrometric design by adding new stations, a technique, such as regionalization approach, to transfer hydrologic information from the existing stations (or their drainage areas) to potential stations (or their drainage areas) is required. *Samuel et al*. [2013] examined several regionalization methods based on spatial proximity, physical similarity, drainage area ratio, and their combinations while designing hydrometric networks. The most recommended regionalization method is coupling spatial proximity and drainage area ratio (inverse distance weighting and drainage area ratio (IDW-DAR)) [*Samuel et al*., 2013]. While their conclusion was made after evaluating many regionalization methods, the IDW-DAR normally performs better in flat or hilly areas than mountainous regions. Considering that the HHCV is relatively flat except the Niagara Escarpment, it is reasonable to use the IDW-DAR method for generating streamflow data at the potential stations (Boxes (4) and (5) in Figure 2). The equation of IDW-DAR is given as

$$Q_p=\sum_{i=1}^{n}\omega_{i,p}\left(\frac{A_p}{A_i}\right)^{\alpha}Q_i,\quad \omega_{i,p}=\frac{h_{i,p}^{-2}}{\sum_{i=1}^{n}\left(h_{i,p}^{-2}\right)} \tag{1}$$

where Q_p and Q_i are the interpolated streamflow at the potential station p and the observed streamflow at the existing station i, n is the number of stations used in the interpolation, $\omega_{i,p}$ is the weighting factor, A_p and A_i are the drainage area of the potential station p and the existing station i, α is the exponential coefficient (1.0 used in this case), and $h_{i,p}$ is the distance between the existing station i and the potential station p, respectively.

3.2. Entropy Calculations and Data Quantization

In information theory, entropy is a measure of uncertainty. If there is uncertainty in a data set, a further observation will deliver additional information which can reduce the uncertainty. Therefore, the uncertainty corresponds to the amount of information as a counterpart. Shannon entropy, $H(X)$, provides a mathematical formula for explaining the entropy from a variable X (i.e., station X), which has a set of discrete probabilities, $p_1, \ldots, p_n$ [*Shannon*, 1948; *Singh*, 1997]. The marginal entropy of a variable X is given as

$$H(X)=-\sum_{i=1}^{n}p(x_i)\,\log_2 p(x_i) \tag{2}$$

where $H(X)$ is the marginal entropy of a discrete random variable X (bits), n is the total number of class intervals (or bins) in a histogram, and $p(x_i)$ is the occurrence probability of x_i at variable X in the ith class interval. The information content in a set of variables can be calculated in a similar manner. The multivariate joint entropy is given as:

$$H(X_1, X_2, \ldots, X_N)=-\sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2}\cdots\sum_{i_N=1}^{n_N}p\left(x_{1,i_1},x_{2,i_2},\ldots,x_{N,i_N}\right)\log_2 p\left(x_{1,i_1},x_{2,i_2},\ldots,x_{N,i_N}\right) \tag{3}$$

where $H(X_1, X_2, \ldots, X_N)$ is the joint entropy of N variables, $p\left(x_{1,i_1},x_{2,i_2},\ldots,x_{N,i_N}\right)$ is the joint probability of N variables, and $n_1, n_2, \ldots, n_N$ are the number of class intervals (or bins) of distributions [*Krstanovic and Singh*, 1992a]. The marginal entropies of a number of stations often contain duplicated information, so that the joint entropy should be less than the sum of marginal entropy unless every variable is independent of each other. In this case, although a station is added to a network, only parts of the station's marginal entropy can contribute the increase of the joint entropy of the network. Therefore, the amount of duplicated or sharable information in a network explains the redundancy or ineffectiveness of the network. Total correlation, C, is a simple estimate of the duplicated information in a system [*McGill*, 1954; *Watanabe*, 1960], and is defined as the difference between the sum of the marginal entropy and the joint entropy of N variables, given by:

$$C(X_1, X_2, \ldots, X_N)=\sum_{i=1}^{N}H(X_i)-H(X_1, X_2, \ldots, X_N) \tag{4}$$

where $C(X_1, X_2, \ldots, X_N)$ is the total correlation of N variables.

If variables X and Y are correlated, parts of information from X can be explained by the information of Y. The information of X that is independent from Y is defined by the conditional entropy of X given Y. The mathematic formulation of the conditional entropy of X given Y is

$$H(X|Y)=-\sum_{i=1}^{n}\sum_{j=1}^{m}p\left(x_i,y_j\right)\log_2 p\left(x_i|y_j\right)=H(X,Y)-H(Y) \tag{5}$$

where X and Y are the variables, and n and m are the numbers of class intervals of X and Y, respectively.

In the application of the entropy theory to hydrometric network design, a variable X and its entropy $H(X)$ describe the monitoring station X and its information content, respectively. Likewise, the joint entropy $H(X_1, X_2, \ldots, X_N)$ and the total correlation $C(X_1, X_2, \ldots, X_N)$ represent the total information and the redundant information, respectively, from a hydrometric network consisting of N stations. Various types of entropy are conceptually understandable using Venn diagram. Figure 3 describes the entropy terms of a

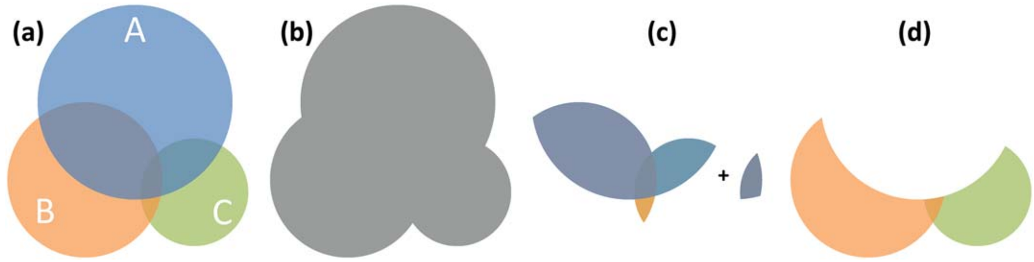

Figure 3. Illustrations describing (a) marginal entropy, (b) joint entropy, (c) total correlation, and (d) conditional entropy of B and C given A.

three-variable data set (three monitoring stations), A, B, and C. The amount of marginal entropy is shown as the size of each circle in Figure 3a. The total area drawn by all stations in Figure 3b represents the joint entropy, $H(A,B,C)$. This example has the first-order and second-order duplications as shown in Figure 3c, and the sum of them is the total correlation, $C(A,B,C)$. The conditional entropy of stations B and C given information from station A, $H(B,C|A)$, is shown in Figure 3d.

Recently, entropy calculations using discrete distributions have been more dominant than those of continuous distributions due to the unavoidable assumptions of distribution functions and difficulties of formularizing the joint entropy in many distributions [*Fahle et al*., 2015]. However, the discrete entropy calculations also require an assumption of data quantization while there is no correct answer which method should be used. This study compares different quantization methods based on their effectiveness in network design. First, the Rounding method, which has been commonly used in communication systems, converts continuous signals to a discrete pulse [*Alfonso et al*., 2010a].

$$x_r = i \cdot floor\left\lfloor \frac{2x+i}{2i} \right\rfloor \tag{6}$$

where x_r is the nearest multiples of i that are corresponding to x, $floor\lfloor\cdot\rfloor$ is the floor function which eliminates all digits after the decimal point. To verify the previous conclusions that changing i does not affect the network design significantly [*Markus et al*., 2003; *Mishra and Coulibaly*, 2010; *Li et al*., 2012; *Alfonso et al*., 2013], two multiplication factors, i.e., $i=1$ mm/d and $i=5$ mm/d, were used in this study for both precipitation and streamflow time series, respectively. Next, *Sturges* [1926] suggested an optimal class interval for statistical analyses based on the range and the number of data in a series

$$h_{st} = \frac{R_x}{1+log_2 N} \tag{7}$$

where h_{st} is the bin width from Sturges' method, R_x is the range of a series of x, and N is the total number of data points. *Scott* [1979] developed a calculation method of an optimal class interval based on the Gaussian density assumption and concluded that the method is still applicable to the other density functions

$$h_{sc} = 3.49\ s \cdot N^{-1/3} \tag{8}$$

where h_{sc} is the bin width from Scott's method and s is the standard deviation of a series of x.

Considering that this study covers multiple monitoring stations of two hydrologic variables, each hydrologic variable is individually quantized and combined into one set of time series for calculating entropy values. In specific, the ranges of precipitation and streamflow were separately used to calculate the bin widths from Sturges' method. Then, the bin widths are estimated at 7.8 mm/d for precipitation and 3.9 mm/d for streamflow time series, respectively. For Scott's method, the standard deviation of each station was firstly applied to calculate the bin width, which ranges from 1.4 to 1.8 mm/d for precipitation and 0.1 to 0.6 mm/d for streamflow, respectively. In turn, Sturges' method and the Rounding method with 5 mm/d of multiplication factor have coarse bin widths compared to Scott's method and the Rounding method with 1 mm/d. Since the discrete entropy terms are used, there is a discontinuity at each edge of the histogram bins, and the discontinuity is generally more significant when the bin widths are coarser [*Scott*, 2015]. However, the edge effects due to the different sizes of bin widths are not considered in this study.

3.3. Design Objectives and Multiobjective Optimization Approach

While the marginal entropy represents the information content by an individual station, the joint entropy represents the information from an entire network. In order to be an efficient network, the network should be able to provide a maximum amount of information with a minimum amount of redundancy. Therefore, the entropy applications to the network design have had two common principles: maximizing joint entropy and minimizing total correlation. There are two approaches that have frequently been used in the entropy-based network design; multiobjective optimization, and single objective maximum information minimum redundancy (MIMR). Multiobjective optimization typically has two objective functions, such as maximizing joint entropy and minimizing total correlation [e.g., *Alfonso et al*., 2010b; *Coulibaly et al*., 2013; *Ridolfi et al*., 2013; *Kornelsen and Coulibaly*, 2015; *Leach et al*., 2015]. The multiobjective problem typically produces a set of Pareto-front optimal solutions. On the other hand, the MIMR reformulate the multiobjective problem to a single objective optimization by merging the different criteria into one objective using weighting factors [e.g., *Li et al*., 2012; *Fahle et al*., 2015]. However, the weight for each objective should be assumed in advance. In this study, the former approach is applied not to make any prior assumptions but to compare various optimal networks in decision processes.

In addition to the entropy terms used in a single network design, conditional entropy is implemented to integrate multiple networks, which can monitor different but correlated hydrologic variables. For the example of precipitation and streamflow networks, the conditional entropy of streamflow given precipitation can represent the information from a streamflow network that cannot be obtained from the precipitation network. This is defined as

$$H(Q_1,\ Q_2,\ \ldots Q_N|P_1,\ P_2,\ \ldots P_M)=H(Q_1,\ Q_2,\ldots Q_N,P_1,\ P_2,\ \ldots P_M)-H(P_1,\ P_2,\ldots P_M) \tag{9}$$

where $Q_1,\ Q_2,\ \ldots Q_N$ are N streamflow monitoring stations and $P_1,\ P_2,\ \ldots P_M$ are M precipitation monitoring stations. By maximizing $H(Q_1,\ Q_2,\ \ldots Q_N|P_1,\ P_2,\ldots P_M)$, the streamflow network is able to avoid redundant information that can be given by the precipitation network. Typically, precipitation and streamflow time series are correlated with a time-lag. Therefore, shifting one of those time series by the time-lag could be delivering more meaningful information for the multivariable (precipitation and streamflow in this study) entropy calculations. However, time-lag correction between precipitation and streamflow is not considered in this study because the time of concentration at each station in HHCV is less than 1 day while daily precipitation and daily streamflow were used. For cases where the time of concentration is longer than the time step of time series, the time-lag should be considered in the data preparation process. Accordingly, the multiobjective optimization problem for the integrated network design objectives (Box (7–1) in Figure 2) is given by:

$$\max H(P,Q)=H\left[EP_1,\ldots,EP_{N_{EP}},PP_1,\ldots,PP_{N_{PP}},EQ_1,\ \ldots,EQ_{N_{EQ}},PQ_1,\ldots,PQ_{N_{PQ}}\right]$$

$$\min C(P,Q)=C\left[EP_1,\ \ldots,EP_{N_{EP}},PP_1,\ldots,PP_{N_{PP}},EQ_1,\ldots,EQ_{N_{EQ}},PQ_1,\ldots,PQ_{N_{PQ}}\right]$$

$$\max H(Q|P)=H\left[\left(EQ_1,\ \ldots,EQ_{N_{EQ}},PQ_1,\ldots,PQ_{N_{PQ}}\right)\ |\ \left(EP_1,\ \ldots,EP_{N_{EP}},PP_1,\ldots,PP_{N_{PP}}\right)\right]$$

subject to:

N_{EP} and N_{EQ} are the numbers of the existing precipitation and streamflow stations, respectively

$$N_{PP}\in\left\{1,\ 2,\ldots,N_{PP,\max}\right\}$$

$$N_{PQ}\in\left\{1,\ 2,\ldots,N_{PQ,\max}\right\} \tag{10}$$

where *EP*, *PP*, *EQ*, and *PQ* denote the existing precipitation stations, the potential precipitation stations, the existing streamflow stations, and the potential streamflow stations, respectively. N_{XX} is the number of stations in the station set of *XX* (i.e., *EP*, *PP*, *EQ*, or *PQ*) and $N_{PX,\max}$ is the total number of potential stations, respectively. In this study, the optimal networks are to be determined by adding new stations among the potential stations while all existing stations remain active to keep the value of the existing long-term monitoring data. The proposed method in this study is versatile so that it can be also used to resolve various network design problems: for example, moving existing stations to new locations without changing the total number of stations, shrinking network size by removing the most redundant stations. Simply, N_{EP} and N_{EQ} in equation (10) would not be equal to the numbers of the existing stations but vary based on the

Table 1. ε-hBOA Parameters and Decision Variables

Model Parameters and Decision Variables	Integrated Design	Precipitation Network	Streamflow Network
Initial population size	10,000	10,000	10,000
Minimum population size	10,000	10,000	10,000
Maximum population size	100,000	100,000	100,000
Archive injection rate	0.25	0.25	0.25
Number of decision variables (n = a + b)	200	40	160
a. Number of existing stations	30	7	23
b. Number of potential stations	170	33	137
Maximum generation (= 2n)	400	80	320
ε for joint entropy	0.0001	0.0001	0.0001
ε for total correlation	0.0001	0.0001	0.0001
ε for conditional entropy	0.0001		
Number of random seeds	50	50	50

designer's needs in these cases. For comparison purposes, the optimal integrated networks are compared to the traditional single-variable network design method (i.e., individual network design), which maximizes the joint entropy and minimizes the total correlation for the precipitation network and the streamflow network, separately (Boxes (7–2) and (7–3) in Figure 2).

The epsilon-dominance hierarchical Bayesian optimization algorithm (ε-hBOA) is used as the optimization technique for finding the optimal networks. The ε-hBOA is a multiobjective evolutionary algorithm that utilizes Bayesian networks, which replace the simulated binary crossover and the polynomial mutation of the epsilon-nondominated sorted genetic algorithm II (ε-NSGA-II) [*Deb*, 2000; *Kollat and Reed*, 2007; *Kollat et al*., 2008]. The initial application of the ε-hBOA was for the design of groundwater monitoring network, and this multiobjective optimization technique has been successfully used in many hydrometric network design problems [e.g., *Kollat et al*., 2008, 2011; *Samuel et al*., 2013; *Kornelsen and Coulibaly*, 2015; *Leach et al*., 2015, 2016; *Keum and Coulibaly*, 2017]. The mathematical and structural improvements of the ε-hBOA beyond the original nondominated sorted genetic algorithm (NSGA) [*Deb*, 1999, 2000; *Deb et al*., 2002] also include the epsilon dominance of the Pareto sets and simplifying the problem by hierarchical decompositions to subproblems [*Pelikan*, 2002; *Osman et al*., 2006; *Kollat et al*., 2008]. The epsilon-dominance technique alleviates the problem complexity by applying a threshold value (i.e., epsilon) to reduce the number of solutions in sorting process [*Laumanns et al*., 2002; *Kollat and Reed*, 2007]. The brief descriptions of the solution searching process of the ε-hBOA are as follows. First, the initial population is generated and sorted to find the fitness to each objective based on the Pareto-dominance. Next, a child population is generated by sampling the joint probability distribution of the Bayesian network, which is iteratively created by the parent population. Then, the Pareto ranking and the crowded binary tournament selection are applied to determine an elite child population. These elite children become the parents of a next generation, and this process continues until the termination criteria are met. The readers who are interested in ε-hBOA may refer *Kollat et al*. [2008] for the detailed information. The ε-hBOA parameters and the numbers of decision variables in this study are shown in Table 1.

4. Results and Discussion

4.1. Entropy and Relative Rankings

Discrete entropy varies with the quantization method. The entropy values can range from zero to the saturated entropy even when the same distribution is used. In specific, the widest bin width will have all data in one bin, which is a certain outcome, so the entropy value goes to zero from equation (2). On the other hand, the finest bin widths let every data be in the different bin if all items are unique. In this case, the entropy will be saturated, of which the value is given by $log_2(nd)$ where nd refers the number of data points. Therefore, the entropy ranges from zero bit to $log_2(3653)=11.83$ bits in this study case. From the entropy values shown in Table 2, Sturges' method using precipitation and the Rounding method with 5 mm/d as the multiplication factor using streamflow time series have the smallest entropies, respectively. The highest entropies of precipitation were given by the Rounding method with 1 mm/d as the multiplication factor while those of streamflow were from Scott's method. The total correlations were very different compared to the average marginal entropy or the joint entropy because the total correlation and the

Table 2. Changes of Entropy Values by the Choice of Quantization

Quantization Cases	Conditions	Variables	ns[a]	Average Marginal Entropy (bits)	Joint Entropy (bits)	Total Correlation (TC) (bits)	TC/ns (bits)
Rounding (i = 1 mm/d)	Existing	Precipitation	7	2.50	7.54	9.98	1.43
	Potential	Precipitation	33	2.69	6.90	82.00	2.48
	Existing	Streamflow	23	1.38	8.13	23.64	1.03
	Potential	Streamflow	137	1.40	8.12	183.14	1.34
Rounding (i = 5 mm/d)	Existing	Precipitation	7	1.66	6.65	4.95	0.71
	Potential	Precipitation	33	1.68	4.99	50.60	1.53
	Existing	Streamflow	23	0.13	1.15	1.81	0.08
	Potential	Streamflow	137	0.11	1.07	13.75	0.10
Sturges	Existing	Precipitation	7	0.66	2.61	2.00	0.29
	Potential	Precipitation	33	0.61	2.00	18.11	0.55
	Existing	Streamflow	23	0.20	1.60	3.02	0.13
	Potential	Streamflow	137	0.18	1.51	23.09	0.17
Scott	Existing	Precipitation	7	2.39	7.50	9.27	1.32
	Potential	Precipitation	33	2.61	6.83	79.35	2.40
	Existing	Streamflow	23	3.14	11.62	60.53	2.63
	Potential	Streamflow	137	3.28	11.82	437.63	3.19

[a]ns: the number of stations.

number of variables (i.e., stations) are correlated. To reduce the effects of this correlation, the average total correlation per station (TC/ns in Table 2 representing the expected amount of redundancy by adding one station) is also shown. Even for the TC/ns, there is no strong concordance over the quantization cases. For example, streamflow stations have higher TC/ns than precipitation stations using Scott's method while the opposite outcomes appear from the other three cases. It can also be assumed that adding one potential station to the streamflow network results on average in 3.19 bits of increased redundancy using Scott's method while in 0.10 bits using the Rounding method with 5 mm/d.

Considering that the entropy values from different quantization are different from each other despite using the same data, the station rankings and the optimal networks are more interesting than the entropy values in the network design processes. Figure 4 shows the relative rankings of potential precipitation and streamflow stations, respectively. The relative rankings were estimated based on the rankings of each quantization case, which was used to calculate the marginal entropy of each potential station. Regardless of the actual rankings, the relative ranking only compares whether the rankings from the four quantization cases are similar or not. Hence, if the quantization does not alter the intrinsic informational characteristics of each station, it is expected that the rankings from different quantization cases will remain similar. The relative rankings of the potential precipitation stations in Figure 4a are similar except few stations located in the southern HHCV resulting in relatively consistent outcomes expected regardless of the quantization cases. However, in Figure 4b, most stations yield different relative rankings. Specifically, the rankings from Scott's method are much higher than the other three cases in the northern area (e.g., Station IDs 1 to 30) while the opposite results are shown near the southeastern lakeshore. Only a few stations in the center of the HHCV near the Station ID 50 show similar rankings from four cases.

4.2. Multiobjective Optimization Results

For the precipitation network design from the individual design in Figure 5, the majority of the number of selected stations is relatively constant, which ranges from 5 to 15 approximately. This agrees with the previous conclusion that the optimal networks remained less affected by the quantization methods [*Markus et al.*, 2003; *Alfonso et al.*, 2010a; *Mishra and Coulibaly*, 2010; *Li et al.*, 2012]. However, the numbers of selected streamflow stations of the individual designs range from 1 to 26 using the Rounding method with 5 mm/d as the multiplication factor and Sturges' method, while they vary from 1 to 54 from the other two quantization cases. Additionally, the highest counts of the additional streamflow stations appear between 10 and 20 from the Rounding method with 1 mm/d as the multiplication factor while the other three methods have their peaks near 10 or fewer additional stations (see Figure 5).

The adequate selection of quantization is more important in the integrated design than the individual design since there can be a bias in entropy calculations from different variables. Specifically, daily precipitation is mostly zero but has several spikes during rain or snow days; however, streamflow time series may

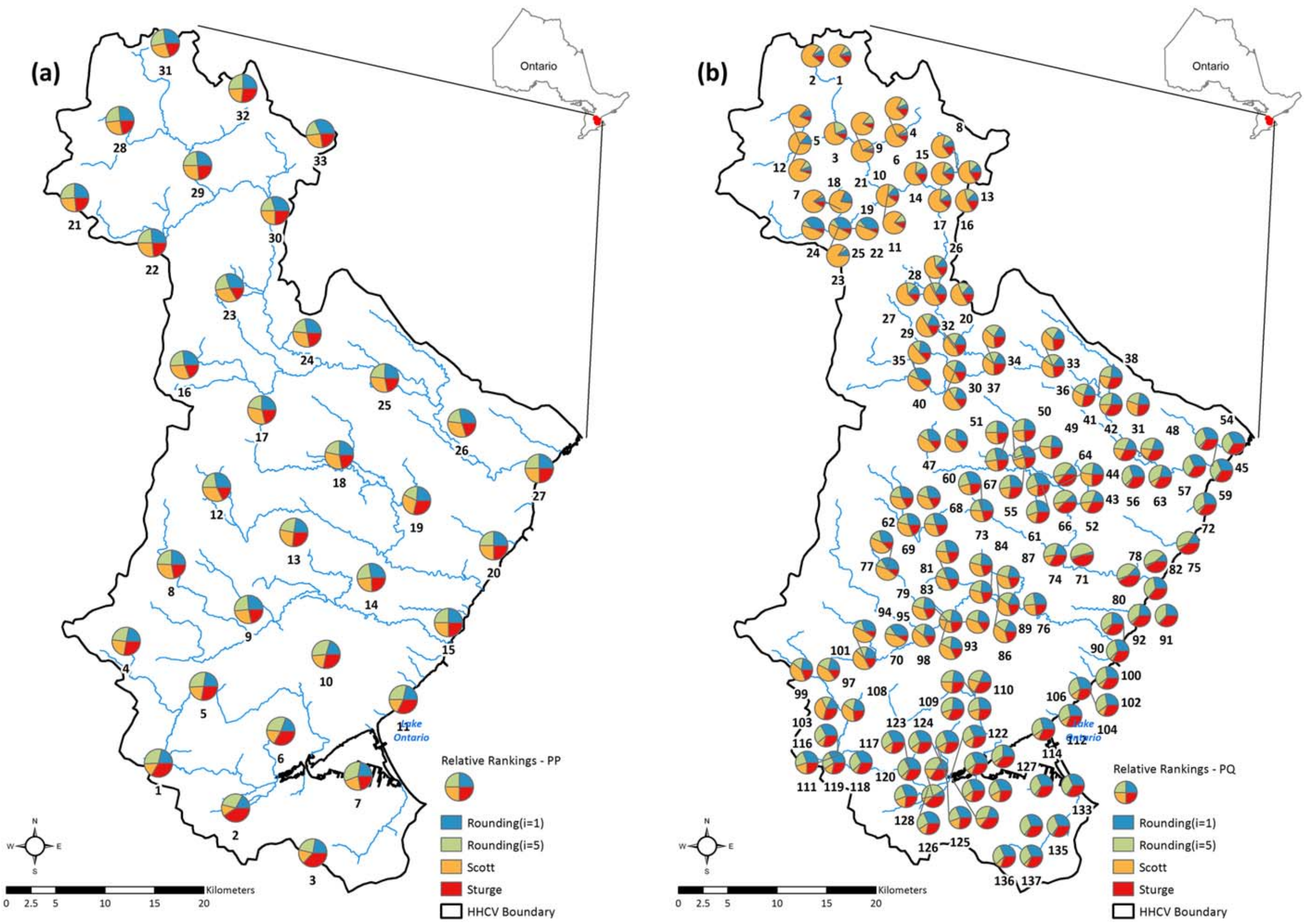

Figure 4. Relative rankings of the potential (a) precipitation and (b) streamflow stations in HHCV based on their marginal entropies. The small numbers near stations are their station IDs, and each pie chart shows the station's relative rankings. Equally divided pies indicate the rankings from the four quantization cases are identical, but a larger pie piece of a quantization case means higher ranking is given by that quantization.

have more fluctuations around baseflows and the smoothed and delayed runoffs during flood events. Therefore, the entropy values of two different hydrologic variables can be biased by how to discretize the continuous data. Recalling the marginal entropy, joint entropy and total correlation values shown in Table 2, the precipitation entropy and the streamflow entropy hold differing relationships. From Figure 5, the numbers of selected stations in the integrated network designs are mostly smaller than those of the individual network designs; hence, more efficient networks are expected from the integrated design. Specifically in the integrated design, most of the numbers of additional precipitation stations from the Rounding method with 1 mm/d as the multiplication factor were one to three while those are greater from the individual design. This bias is more noticeable when Scott's method was used. The number of additional precipitation stations is one for the complete Pareto-optimal solutions. Considering the individual design yields a range of 1 to 21 additional stations, the entropy characteristic of precipitation data set has been altered by Scott's method. In addition, the joint entropy of the optimal combined networks (jointly precipitation and streamflow networks) approximately range from 10.65 to 10.88 bits from the Rounding method ($i = 1$ mm/d), 7.45 to 8.10 bits from the Rounding method ($i = 5$ mm/d), 3.98 to 4.24 bits using Sturges' method, and 11.79 to 11.83 bits using Scott's method. Considering that the saturated entropy is 11.83 bits in this study, it can be assumed that the outcomes of Scott's method, which the highest joint entropy among the optimal networks met the saturated entropy, is less reliable than other quantization cases. The joint entropy values using Sturges' method are the lowest because the bin width of this quantization is the coarsest. Too coarse binning may not capture the characteristics of a time series. However, Sturges' method was able to provide similar histogram shapes between the integrated and the individual network designs in this study area; the

10.1002/2016WR019981

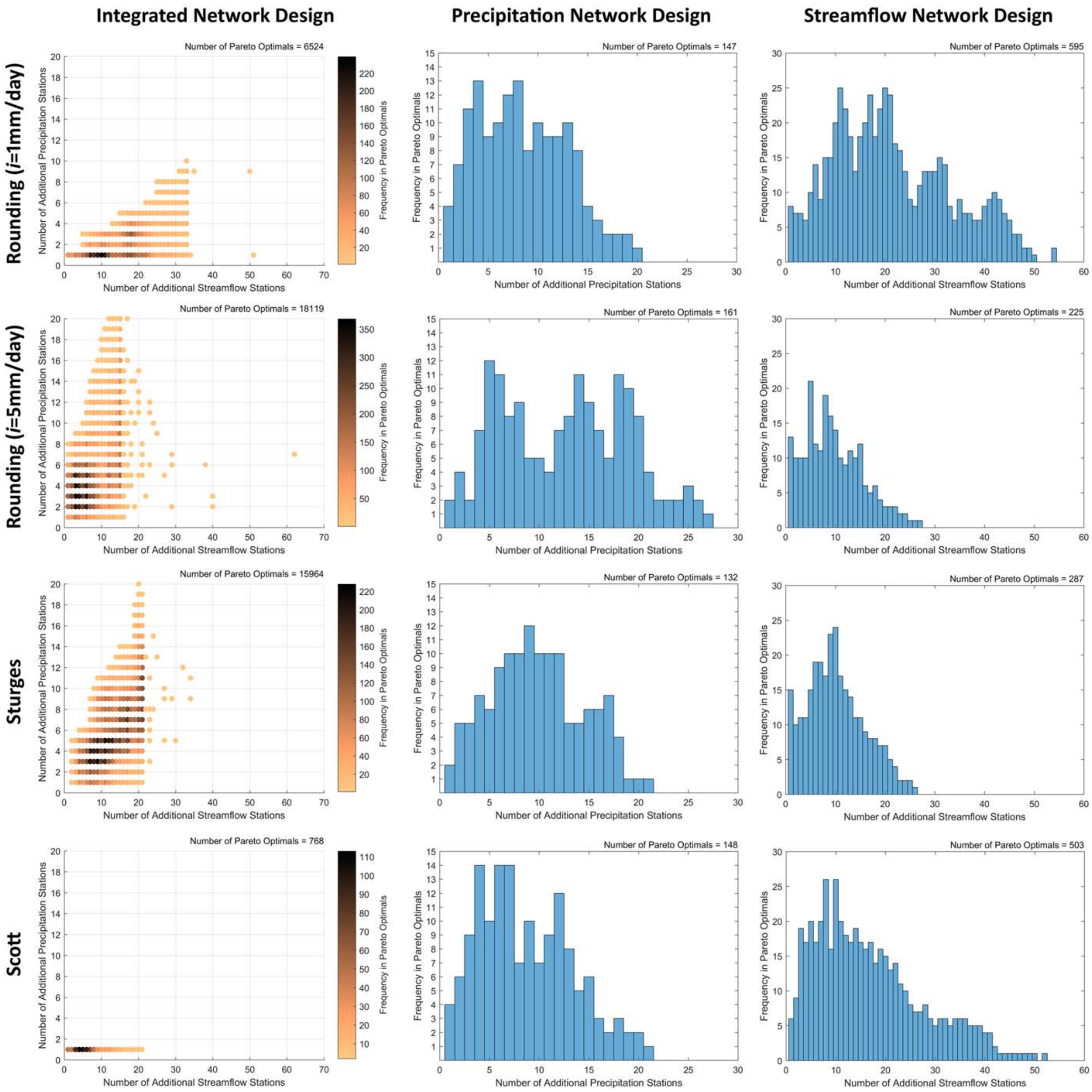

Figure 5. Histograms of the numbers of additional stations from the integrated and individual network designs.

majority is shown at 5–10 additional precipitation stations and 5–15 additional streamflow stations in HHCV.

Figures 6 and 7 show the Pareto-optimal solutions using Sturges' method and Rounding method with 5 mm/d as the multiplication factor, respectively. It should be noted that Figures 6a and 7a are three-dimensional plots, and each optimal solution (i.e., each dot) includes both the precipitation and streamflow networks in the HHCV altogether. As shown in Figure 6, the number of Pareto-optimal solutions from the integrated design using Sturges' method is 15,964. The joint entropy from 30 existing stations and 170 potential stations from both networks are estimated approximately between 3.95 and 4.25 bits. Total correlation of the entire networks and the conditional entropy of the streamflow network given the precipitation

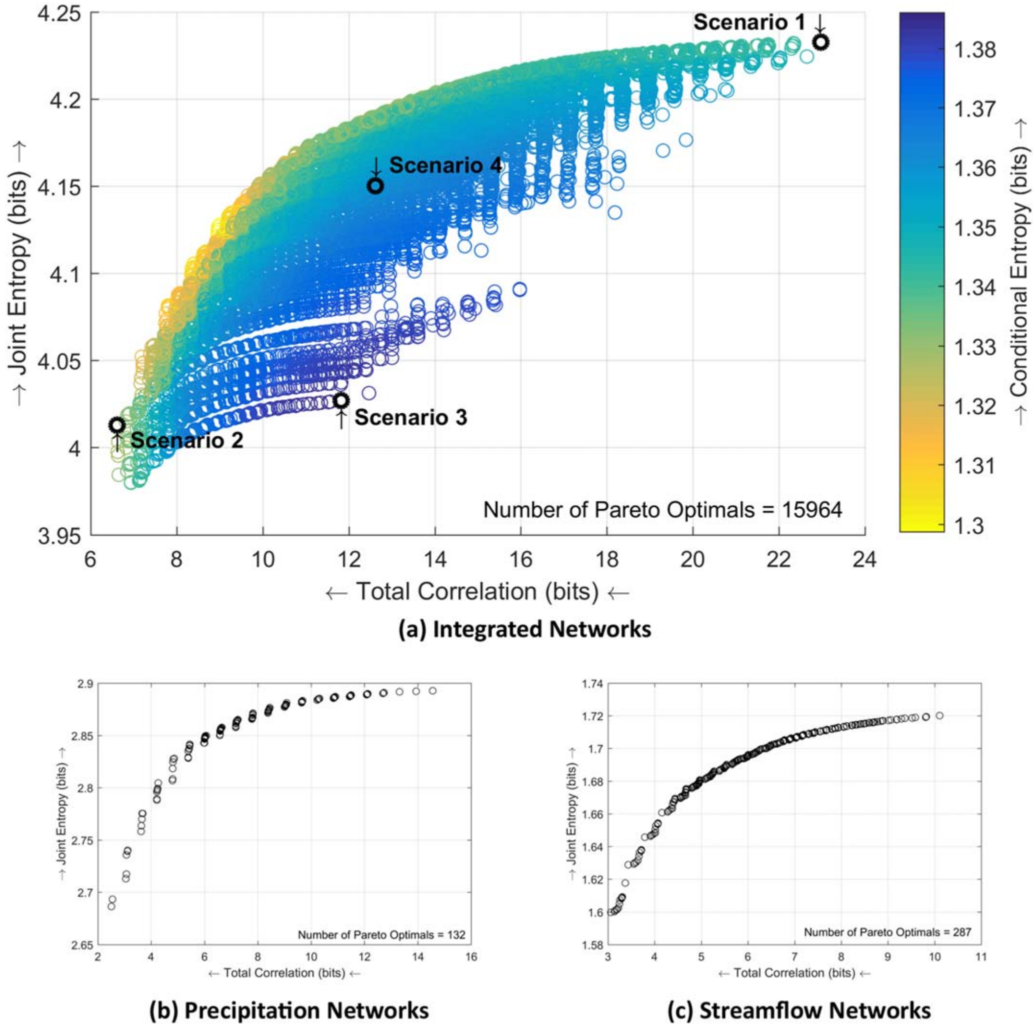

Figure 6. Pareto-front optimal solution sets from the integrated network design and the individual network designs using Sturges' method. The arrows in the axis titles explain the directions of the objectives.

network range from 7 to 23 bits and 1.30 to 1.38 bits, respectively. Even though the entropy values are different, the shape of the edge, where low conditional entropy appears, is similar to the Pareto fronts from the individual designs. As the conditional entropy increases, more optimal solutions were given by accepting trade-offs from joint entropy and total correlation.

Four example scenarios (1–4, see Figure 6a) were selected from the Pareto-front optimal solutions, and their spatial distributions of the selected monitoring stations were compared. Among the optimal solutions, Scenario 1 has the maximum joint entropy but the largest total correlation and medium conditional entropy. Scenario 2 has the minimum total correlation but lowest joint entropy and conditional entropy. Scenario 3 has the maximum conditional entropy of streamflow network given the precipitation network but the joint entropy and the total correlation are not the best. In other words, Scenarios 1–3 represent the extreme point of each objective, respectively. On the other hand, the three objective values are near their mediums in the Scenario 4. In the same manner, Scenarios 5–8 were chosen from the Pareto-optimal solutions using the Rounding method with 5 mm/d of multiplication factor (Figure 7). The detailed descriptions of each scenario including the objective function values are shown in Table 3.

The shapes of the Pareto-fronts and the number of the optimal solutions from different quantization cases are relatively similar. Scenarios 1 and 5, which are the optimal solutions having maximum joint entropy, has 40 and 41 additional stations while there are only 3 and 2 additional stations in Scenarios 2 and 6, respectively. The small numbers of stations in Scenarios 2 and 6, which have a feature of the minimum total

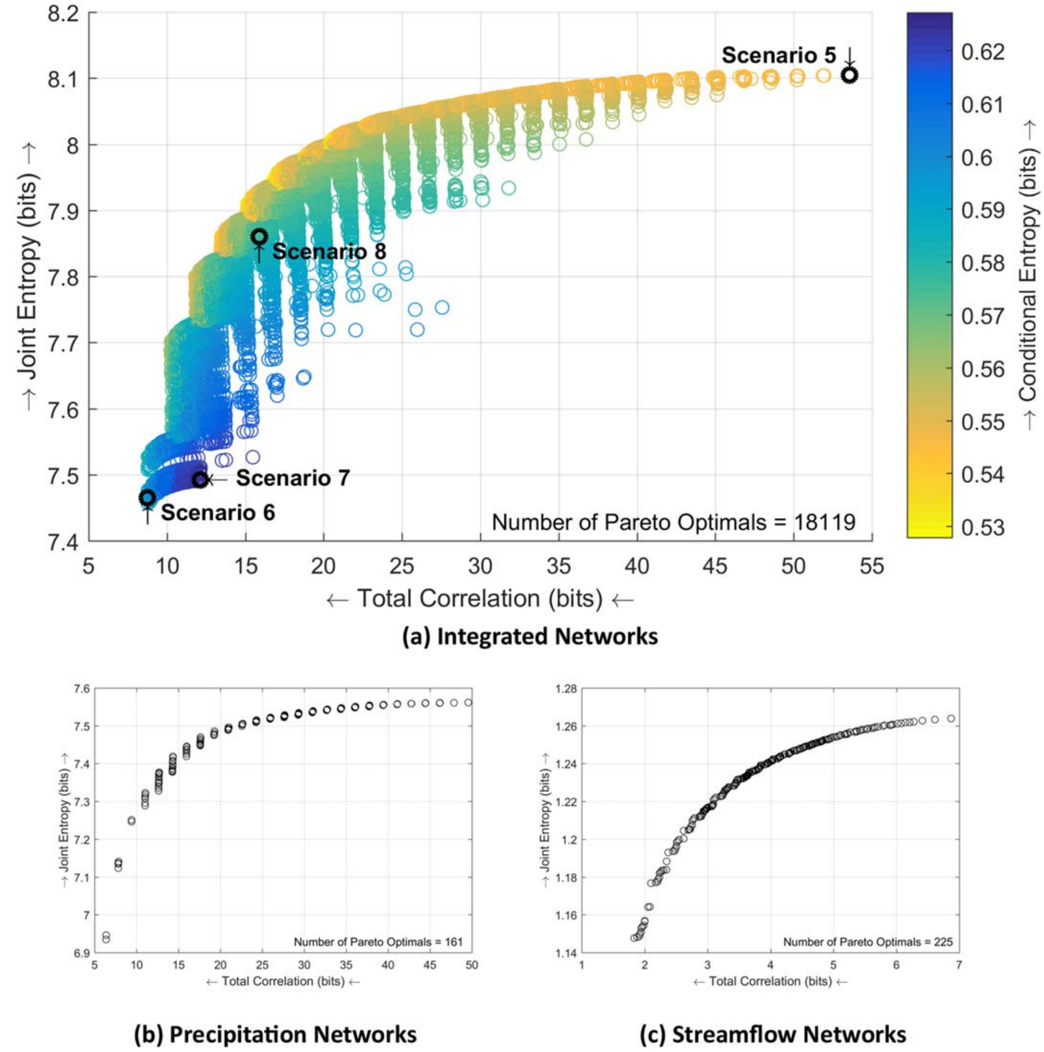

Figure 7. Pareto-front optimal solution sets from the integrated network design and the individual network designs using Rounding method with 5 mm/d as the multiplication factor. The arrows in the axis titles explain the directions of the objectives.

correlation, are reasonably acceptable based on the mathematical formulation of total correlation (equation (3)) because the total correlation increases or remains the same when one or a number of stations is added. Thus, the objective of minimizing total correlation also tends to minimize the number of stations. Scenarios 3 and 7 have relatively small number of precipitation stations compared to streamflow stations as the conditional entropy of streamflow stations given precipitation stations was maximized (see equation (6)).

Table 3. Scenario Descriptions

Scenarios	Quantization Cases	Features	Joint Entropy (bits)	Total Correlation (bits)	Conditional Entropy (bits)	Numbers of Selected Precipitation Stations	Numbers of Selected Streamflow Stations
1	Sturges	Max. joint entropy	4.2326	22.9699	1.3412	20	20
2	Sturges	Min. total correlation	4.0126	6.6144	1.3261	1	2
3	Sturges	Max. conditional entropy	4.0268	11.8257	1.3860	1	21
4	Sturges	Medium objective values	4.1501	12.6255	1.3598	5	15
5	Rounding (i = 5 mm/d)	Max. joint entropy	8.1047	53.5747	0.5442	26	15
6	Rounding (i = 5 mm/d)	Min. total correlation	7.4651	8.7608	0.5903	1	1
7	Rounding (i = 5 mm/d)	Max. conditional entropy	7.4918	12.1247	0.6272	1	16
8	Rounding (i = 5 mm/d)	Medium objective values	7.8604	15.9070	0.5793	4	12

10.1002/2016WR019981

From the results shown in Figures (6 and 7), and Table 3, there is no obvious difference in the shape of the Pareto-fronts and the number of optimal solutions due to the quantization cases even though the actual entropy values changed.

Alfonso et al. [2010a] analyzed the sensitivity of the multiplication factor of Rounding method and concluded that 25–50% of monitoring stations were commonly selected in the optimal networks regardless of the multiplication factors. On the other hand, *Fahle et al*. [2015] also calculated the station rankings by changing the multiplication factors of Rounding method and using Scott's method. They concluded that the effects of multiplication factor were evident. However, the changes of the station rankings were even more noticeable when different quantization methods were used. From the relative rankings and the optimization results using three quantization method and different multiplication factors for Rounding method, it can be concluded that the quantization methods affects both the station rankings and the optimal networks.

4.3. Optimal Networks

The locations of the selected stations from each example scenario using Sturges' method and Rounding method with 5 mm/d as the multiplication factor are shown in Figures 8 and 9, respectively. While there is no strong pattern of the selected precipitation stations across all scenarios, the selected streamflow stations are mostly found in the south to east HHCV along the western shore of Lake Ontario except Scenario 6. Since the joint entropy tends to increase when the number of the station is increased, the optimal networks that have higher joint entropy naturally consist of a larger number of stations. Therefore, the numbers of selected stations in Scenarios 1 and 5 are greater than other scenarios. When comparing the two quantization cases shown in Figures 8 and 9, more than half of the selected precipitation and streamflow stations appeared in both. This pattern is more noticeable to the selected streamflow stations in Scenarios 3 versus 7 and 4 versus 8, of which the number of selected stations are relatively large. However, if the number of selected stations is small such as Scenarios 2 and 6, the optimal networks become more different due to the quantization. Although every optimal solution was not compared completely, it can be assumed that the networks with a smaller number of additional stations are highly affected by the quantization while the effects become diminishing when the numbers of the additional stations are larger. While every point in the Pareto-fronts represents the optimal integrated network, larger networks, such as Scenarios 1, 4, 5, and 8, can be preferred for the purpose of avoiding the quantization effects.

4.4. Monitoring Hot Spot Maps

Since the number of Pareto-optimal solutions is large (e.g., more than 10,000 solutions from each integrated network design), monitoring hot spot maps are a simple and a direct way to understand the results in general. The occurrence frequency of each potential station was estimated by counting the number of solutions that the station was included in among the total number of Pareto-optimal solutions. In other words, the frequency represents the likelihood of a potential station to be selected in the optimal networks. For example, the potential streamflow station ID # 57 appeared 10,452 times among 15,964 Pareto-optimal solutions (or networks), so that the occurrence ratio of the station becomes 65%. The optimal solutions from the integrated designs include both precipitation and streamflow networks; however, the monitoring hot spot maps were generated for precipitation and streamflow stations separately. Figure 10 shows the monitoring hot spot maps from the results using Sturges' method. The areas with high-occurrence ratio values are referred to as "hot spots" where the additional station(s) should be primarily considered whatever the optimal network selected.

The hot spots are located very differently in the precipitation networks. Specifically, there is one large hot spot in the north of HHCV and the ratio gradually decreases to east and south bounds when using the integrated design, while hot spots of individual design are located sparsely along the HHCV boundary. The integrated design results appear interesting as the addition of precipitation stations in northern HHCV would provide a significant amount of new information to the streamflow network that is also independent from the information gained by the precipitation network itself. More interestingly, the spatial distribution of the hot spots for streamflow is very similar in the two design cases (integrated design versus individual design)—indicating that areas in critical need of additional streamflow stations are well identified by the integrated design approach with the additional benefit of knowing where to add precipitation stations that will enhance information content of the streamflow network. Thus the integrated design offers a way of

10.1002/2016WR019981

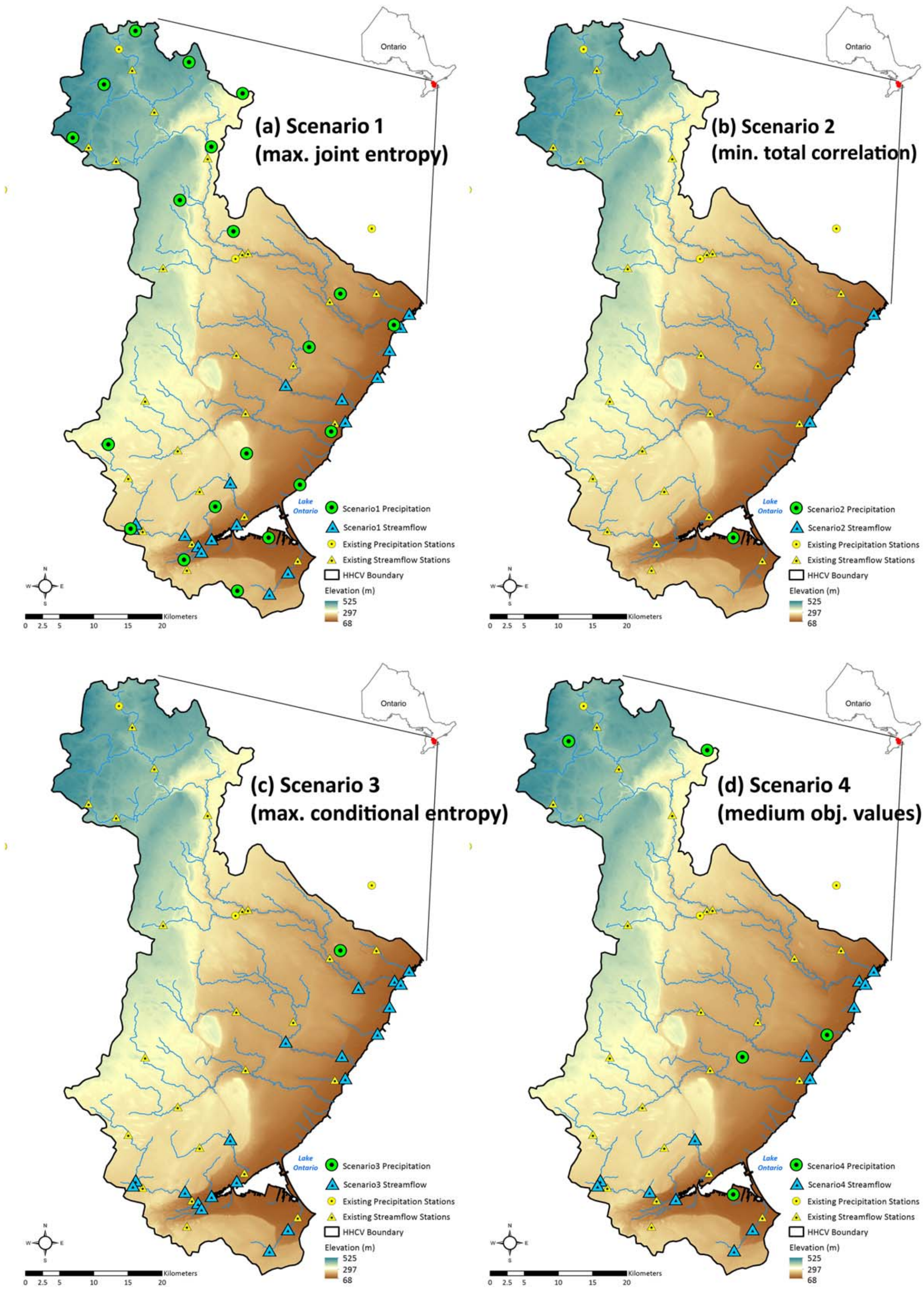

Figure 8. Locations of the selected precipitation and streamflow stations for each scenario using Sturges' method.

10.1002/2016WR019981

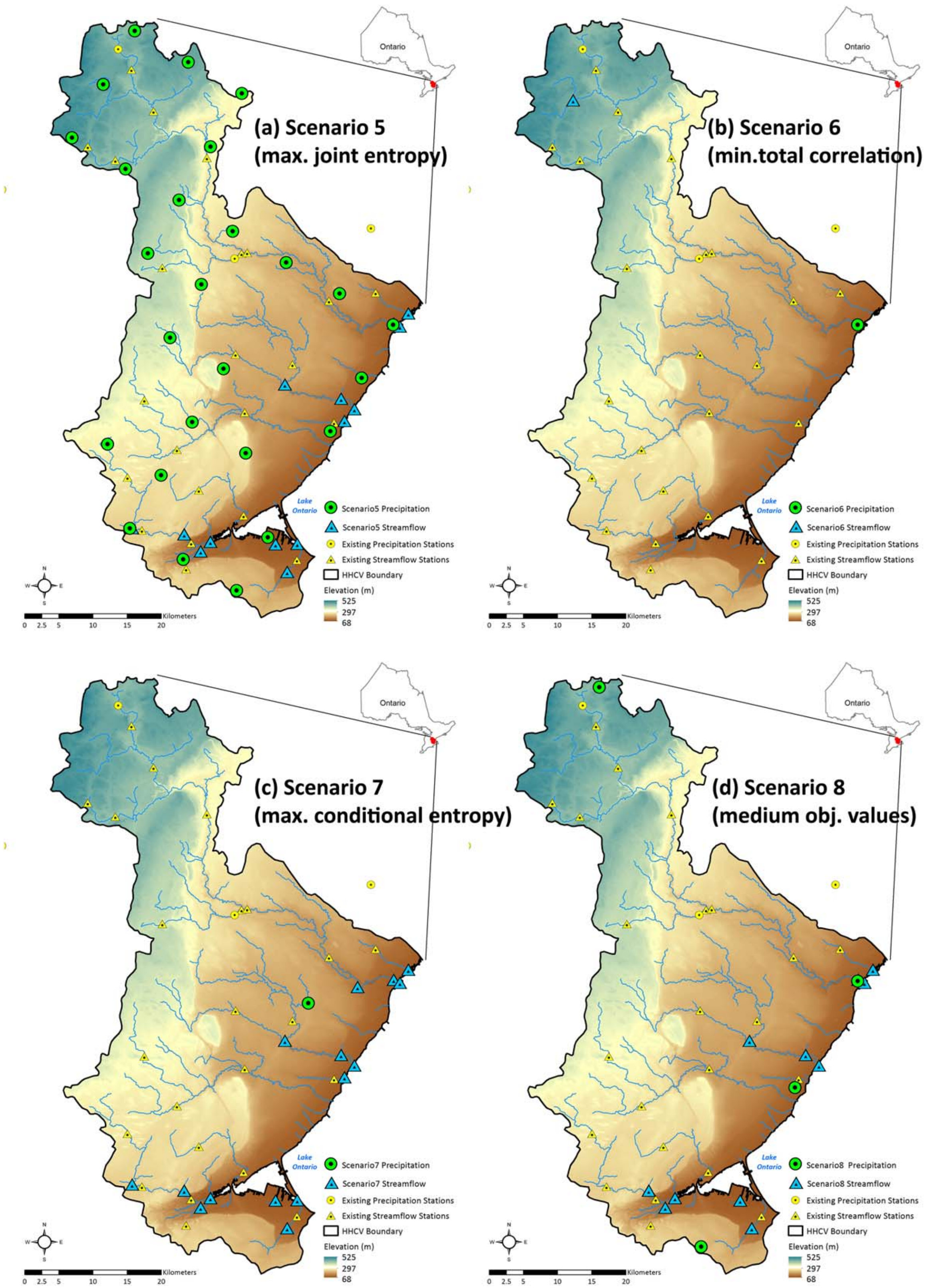

Figure 9. Locations of the selected precipitation and streamflow stations for each scenario using Rounding method with 5 mm/d as the multiplication factor.

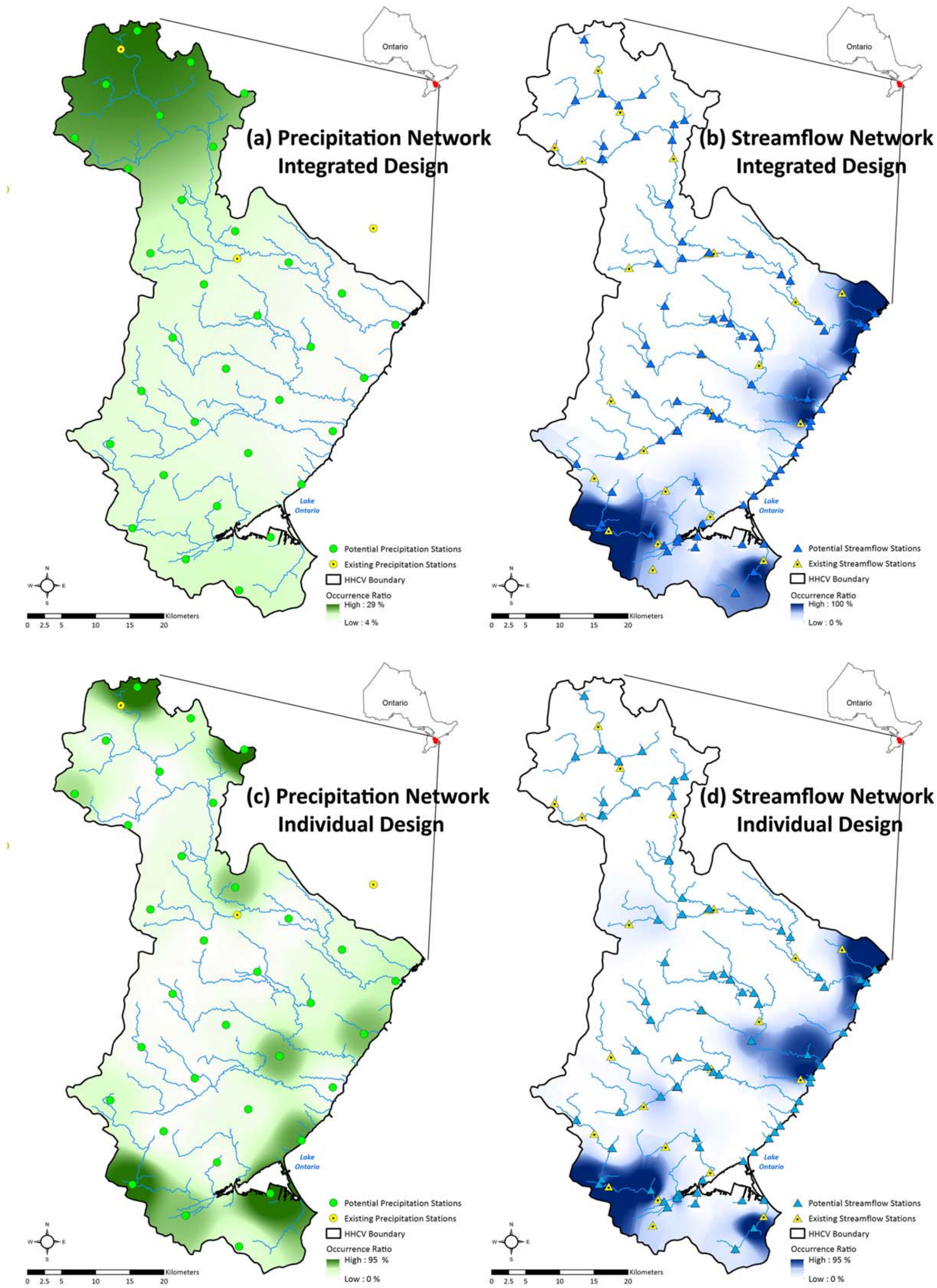

Figure 10. Monitoring hot spot maps of precipitation and streamflow networks from the integrated and the individual single-variable network design methods using Sturges' method.

identifying locations of additional precipitation stations that will benefit to the streamflow network. This knowledge could be important for decision making on new station location.

Considering the integrated design includes an additional objective, i.e., conditional entropy of the streamflow network given the precipitation network, the efficiency of the integrated network design was achieved mostly by reducing the number of additional precipitation stations, which may include redundant or duplicated information to the optimal streamflow networks. This explains the fewer hot spots compared to the individual network design results for precipitation. The integrated network design method provides an obvious contrast between the high-occurrence ratio areas and the low-occurrence ratio areas which may be good for decision-makers as they are provided with a more distinct choice in the selection of the additional station locations.

5. Conclusions

An integrated multivariable hydrometric network design method was proposed using entropy theory with a multiobjective optimization approach. The proposed decision support system was applied to the design of both precipitation and streamflow monitoring networks simultaneously by maximizing the joint entropy, minimizing the total correlation, and maximizing the conditional entropy of the streamflow network given the precipitation network in a watershed in southern Ontario. Application of the conditional entropy enabled the design of an integrated network of two variables (i.e., precipitation and streamflow) by maximizing the amount of streamflow network information which is distinct from the precipitation network. As the multiobjective optimization does not provide a single optimal solution but yields Pareto-optimal solutions, the ultimate choice of the adequate networks lies in the hands of the water resources managers. However, monitoring hot spot maps highlight the most needed potential stations. When comparing the integrated multivariable design with the individual variable design, the spatial distribution of hot spots is different for the precipitation network but similar for the streamflow network. This is because the integrated design permits to identify locations of additional streamflow stations that have distinct information contents from precipitation stations. Areas in critical need of additional streamflow stations are well identified by the integrated design approach with the additional benefit of knowing where to add precipitation stations that will benefit to the streamflow network. Such information is useful for decision makers in the selection of new station locations. The proposed methodology can be extended to two or more variables, such as precipitation, streamflow, groundwater level, snow depth, and water quality.

Additionally, the effects of quantization in calculating entropy were evaluated. From the analysis of the relative rankings based on the marginal entropy of each potential station, the quantization altered the rankings of potential streamflow stations significantly. When comparing the number of solutions and their histograms from the multiobjective optimization, Sturges' method and the Rounding method with 5 mm/d as the multiplication factor show similar results while Scott's method appears poor with the integrated design. The quantization also affects the relationship between the entropy measures as the low conditional entropy was found on low total correlation using Sturges' method while the opposite outcome occurred from the Rounding method with 5 mm/d as the multiplication factor. The differences were also found from the spatial distributions of the selected stations, specifically when the selected number of stations are small. While there is no specific guideline for the selection of quantization in the entropy calculations, a careful consideration based on the characteristics and statistics of the hydrologic variables is required.

Acknowledgments
This research was supported jointly by the Natural Science and Engineering Research Council (NSERC) of Canada, Water Survey Canada, Environment and Climate Change Canada, BC-Hydro, and Hydro-Québec. The data were obtained from Environment and Climate Change Canada, Water Survey Canada, and Credit Valley Conservation. Computational resources were provided by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET) and Compute/Calcul Canada. We would like to thank Dr. Joshua Kollat at Pennsylvania State University for providing the source code of the ε-hBOA. Data used in this study are available from the corresponding author upon request. We are also grateful to Dr. Marcus Fahle and three anonymous reviewers for their comments that helped to improve the initial manuscript.

References

Alfonso, L., and R. Price (2012), Coupling hydrodynamic models and value of information for designing stage monitoring networks, *Water Resour. Res.*, *48*, W08530, doi:10.1029/2012WR012040.

Alfonso, L., A. Lobbrecht, and R. Price (2010a), Information theory-based approach for location of monitoring water level gauges in polders, *Water Resour. Res.*, *46*, W03528, doi:10.1029/2009WR008101.

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

Alfonso, L., L. He, A. Lobbrecht, and R. Price (2013), Information theory applied to evaluate the discharge monitoring network of the Magdalena River, *J. Hydroinf.*, *15*(1), 211–228, doi:10.2166/hydro.2012.066.

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

Amorocho, J., and B. Espildora (1973), Entropy in the assessment of uncertainty in hydrologic systems and models, *Water Resour. Res.*, *9*(6), 1511–1522, doi:10.1029/WR009i006p01511.

Behmel, S., M. Damour, R. Ludwig, and M. J. Rodriguez (2016), Water quality monitoring strategies—A review and future perspectives, *Sci. Total Environ.*, *571*, 1312–1329, doi:10.1016/j.scitotenv.2016.06.235.

Burn, D. H. (2009), Hydrological information for sustainable development, *Hydrol. Sci. J.*, *42*(4), 481–492, doi:10.1080/02626669709492048.

Caselton, W. F., and T. Husain (1980), Hydrologic networks: Information transmission, *J. Water Resour. Plann. Manage. Div. Am. Soc. Civ. Eng.*, *106*(WR2), 503–520.

Chacon-Hurtado, J. C., L. Alfonso, and D. Solomatine (2016), Rainfall and streamflow sensor network design: A review of applications, classification, and a proposed framework, *Hydrol. Earth Syst. Sci. Discuss.*, *21*, 3071–3091, doi:10.5194/hess-2016-368.

Coulibaly, P., K. C. Kornelsen, J. M. Leach, and J. Samuel (2013), Hydrometric network design using indicators of hydrologic alteration: Pilot case study: Environment Canada, *Water Surv. of Can., Rep. K3D35–13-1310R*, Water Survey of Canada of Environment Canada, Hamilton, Ont., Canada.

Dai, Q., M. Bray, L. Zhuo, T. Islam, and D. Han (2017), A scheme for rain gauge network design based on remotely sensed rainfall measurements, *J. Hydrometeorol.*, *18*(2), 363–379, doi:10.1175/JHM-D-16-0136.1.

Deb, K. (1999), Multi-objective genetic algorithms: Problem difficulties and construction of test problems, *Evol. Comput.*, *7*(3), 205–230, doi:10.1162/evco.1999.7.3.205.

Deb, K. (2000), An efficient constraint handling method for genetic algorithms, *Comput. Methods Appl. Mech. Eng.*, *186*(2–4), 311–338, doi:10.1016/S0045-7825(99)00389-8.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan (2002), A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Trans. Evol. Comput.*, *6*(2), 182–197, doi:10.1109/4235.996017.

Environment Canada (2014a), *Climate Data*. [Available at http://climate.weather.gc.ca.]

Environment Canada (2014b), *Historical Hydrometric Data—WaterOffice*. [Available at https://wateroffice.ec.gc.ca/mainmenu/historical_data_index_e.html.]

Fahle, M., T. L. Hohenbrink, O. Dietrich, and G. Lischeid (2015), Temporal variability of the optimal monitoring setup assessed using information theory, *Water Resour. Res.*, *51*, 7723–7743, doi:10.1002/2015WR017137.

Government of Canada (2015), *10 km Resolution Numerical Data of the Regional Deterministic Prediction System (RDPS)—GRIB2 Format*. [Available at http://weather.gc.ca/grib/grib2_reg_10km_e.html.]

Halverson, M. J., and S. W. Fleming (2015), Complex network theory, streamflow, and hydrometric monitoring system design, *Hydrol. Earth Syst. Sci.*, *19*(7), 3301–3318, doi:10.5194/hess-19-3301-2015.

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

Husain, T. (1989), Hydrologic uncertainty measure and network design, *J. Am. Water Resour. Assoc.*, *25*(3), 527–534, doi:10.1111/j.1752-1688.1989.tb03088.x.

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

Keum, J., and J. J. Kaluarachchi (2015), Development of a decision-making methodology to design a water quality monitoring network, *Environ. Monit. Assess.*, *187*(7), 466, doi:10.1007/s10661-015-4687-z.

Kollat, J. B., and P. M. Reed (2007), A computational scaling analysis of multiobjective evolutionary algorithms in long-term groundwater monitoring applications, *Adv. Water Resour.*, *30*(3), 408–419, doi:10.1016/j.advwatres.2006.05.009.

Kollat, J. B., P. M. Reed, and J. R. Kasprzyk (2008), A new epsilon-dominance hierarchical Bayesian optimization algorithm for large multiobjective monitoring network design problems, *Adv. Water Resour.*, *31*(5), 828–845, doi:10.1016/j.advwatres.2008.01.017.

Kollat, J. B., P. M. Reed, and R. M. Maxwell (2011), Many-objective groundwater monitoring network design using bias-aware ensemble Kalman filtering, evolutionary optimization, and visual analytics, *Water Resour. Res.*, *47*, W02529, doi:10.1029/2010WR009194.

Kornelsen, K. C., and P. Coulibaly (2013), McMaster Mesonet soil moisture dataset: Description and spatio-temporal variability analysis, *Hydrol. Earth Syst. Sci.*, *17*, 1589–1606, doi:10.5194/hess-17-1589-2013.

Kornelsen, K. C., and P. Coulibaly (2015), Design of an optimal soil moisture monitoring network using SMOS retrieved soil moisture, *IEEE Trans. Geosci. Remote Sens.*, *53*(7), 3950–3959, doi:10.1109/TGRS.2014.2388451.

Krstanovic, P. F., and V. P. Singh (1992a), Evaluation of rainfall networks using entropy: I. Theoretical development, *Water Resour. Manage.*, *6*(4), 279–293, doi:10.1007/BF00872281.

Krstanovic, P. F., and V. P. Singh (1992b), Evaluation of rainfall networks using entropy: II. Application, *Water Resour. Manage.*, *6*(4), 295–314, doi:10.1007/BF00872282.

Laumanns, M., L. Thiele, K. Deb, and E. Zitzler (2002), Combining convergence and diversity in evolutionary multiobjective optimization, *Evol. Comput.*, *10*(3), 263–282, doi:10.1162/106365602760234108.

Leach, J. M., K. C. Kornelsen, J. Samuel, and P. Coulibaly (2015), Hydrometric network design using streamflow signatures and indicators of hydrologic alteration, *J. Hydrol.*, *529*(3), 1350–1359, doi:10.1016/j.jhydrol.2015.08.048.

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

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

Markus, M., H. Vernon Knapp, and G. D. Tasker (2003), Entropy and generalized least square methods in assessment of the regional value of streamgages, *J. Hydrol.*, *283*(1–4), 107–121, doi:10.1016/S0022-1694(03)00244-0.

McGill, W. J. (1954), Multivariate information transmission, *Psychometrika*, *19*(2), 97–116, doi:10.1007/BF02289159.

Mishra, A. K., and P. Coulibaly (2009), Developments in hydrometric network design: A review, *Rev. Geophys.*, *47*, RG2001, doi:10.1029/2007RG000243.

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

Natural Resources Canada (2015), *Regional, National and International Climate Modeling*. [Available at http://cfs.nrcan.gc.ca/projects/3.]

Osman, M. S., M. A. Abo-Sinna, and A. A. Mousa (2006), Epsilon-dominance based multiobjective genetic algorithm for economic emission load dispatch optimization problem, in *Power Systems Conference, 2006, MEPCON 2006, Eleventh International Middle East, Elev. Int. Middle East*, vol. 2, pp. 576–581.

Pelikan, M. (2002), Bayesian optimization algorithm: From single level to hierarchy, PhD thesis, Univ. of Illinois at Urbana-Champaign, Urbana, Ill.

Pilon, P. J., T. R. Yuzyk, R. A. Hale, and T. J. Day (1996), Challenges facing surface water monitoring in Canada, *Can. Water Resour. J.*, *21*(2), 157–164, doi:10.4296/cwrj2102157.

Ridolfi, E., L. Alfonso, G. Di Baldassarre, F. Dottori, F. Russo, and F. Napolitano (2013), An entropy approach for the optimization of cross-section spacing for river modelling, *Hydrol. Sci. J.*, *59*(1), 126–137, doi:10.1080/02626667.2013.822640.

Samuel, J., P. Coulibaly, and J. B. Kollat (2013), CRDEMO: Combined regionalization and dual entropy-multiobjective optimization for hydrometric network design, *Water Resour. Res.*, *49*, 8070–8089, doi:10.1002/2013WR014058.

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

Scott, D. W. (2015), *Multivariate Density Estimation: Theory, Practice, and Visualization*, 2nd ed., John Wiley, Hoboken, N. J.

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

Singh, V. P. (1997), The use of entropy in hydrology and water resources, *Hydrol. Processes*, *11*(6), 587–626.

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

U.S. Geological Survey (1999), Streamflow information for the next century, a plan for the national streamflow information program of the U.S. Geological Survey, *U.S. Geol. Surv. Open-File Rep. 99–456*, U.S. Geol. Surv., Denver, Colo.

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

World Meteorological Organization (2003), *Introduction to GRIB Edition 1 and GRIB Edition 2*.

World Meteorological Organization (2008), *Guide to Hydrological Practices*, Volume I *Hydrology—From Measurement to Hydrological Information, WMO-No. 168*, Sixth.

Yang, Y., and D. H. Burn (1994), An entropy approach to data collection network design, *J. Hydrol.*, *157*(1–4), 307–324, doi:10.1016/0022-1694(94)90111-2.

Yoo, C., K. Jung, and J. Lee (2008), Evaluation of rain gauge network using entropy theory: Comparison of mixed and continuous distribution function applications, *J. Hydrol. Eng.*, *13*(4), 226–235, doi:10.1061/(ASCE)1084-0699(2008)13:4(226).