

Review

Entropy Applications to Water Monitoring Network Design: A Review

Jongho Keum ^{1,*}, Kurt C. Kornelsen ², James M. Leach ¹ and Paulin Coulibaly ^{1,2}

¹ Department of Civil Engineering, McMaster University, Hamilton, ON L8S4L8, Canada; leachjm@mcmaster.ca (J.M.L.); couliba@mcmaster.ca (P.C.)

² School of Geography and Earth Sciences, NSERC Canadian FloodNet, McMaster University, Hamilton, ON L8S4L8, Canada; kornelkc@mcmaster.ca

* Correspondence: jkeum@mcmaster.ca; Tel.: +1-905-525-9140 (ext. 27875)

Received: 16 October 2017; Accepted: 10 November 2017; Published: 15 November 2017

Abstract: Having reliable water monitoring networks is an essential component of water resources and environmental management. A standardized process for the design of water monitoring networks does not exist with the exception of the World Meteorological Organization (WMO) general guidelines about the minimum network density. While one of the major challenges in the design of optimal hydrometric networks has been establishing design objectives, information theory has been successfully adopted to network design problems by providing measures of the information content that can be deliverable from a station or a network. This review firstly summarizes the common entropy terms that have been used in water monitoring network designs. Then, this paper deals with the recent applications of the entropy concept for water monitoring network designs, which are categorized into (1) precipitation; (2) streamflow and water level; (3) water quality; and (4) soil moisture and groundwater networks. The integrated design method for multivariate monitoring networks is also covered. Despite several issues, entropy theory has been well suited to water monitoring network design. However, further work is still required to provide design standards and guidelines for operational use.

Keywords: entropy; water monitoring; network design; hydrometric network; information theory; entropy applications

1. Introduction

Water monitoring networks account for all aspects of the water-related measurement system including precipitation, streamflow, water quality, groundwater, soil moisture, etc. [1–3]. Adequate water monitoring networks and quality data from them comprise one of the first and primary steps towards efficient water resource management. The basic principles of water monitoring network design have simply been a number of monitoring stations, locations of the stations and data period or sampling frequency [4,5]. Recent technological advances have allowed gradual transitions from manual sampling to the automated observations, while some water quality parameters still require field and/or lab analyses of water or other environmental samples. One may expect that the more data we collect, the more water resource problems are solved efficiently. However, this is not always true because irrelevant, inadequate or inefficient data in the wrong location or at the wrong time can inhibit the quality of a dataset [1,6,7]. More seriously, the decline of water monitoring networks has been a general trend due to financial limitations and changes of monitoring priority [8–10]. Therefore, determining the adequate number of monitoring stations and their locations has become critical to network design. However, a standardized methodology for a proper water monitoring network design process has not been drawn yet due to the practical and socioeconomic complexity in diverse design cases [1,11].

The existing reviews have investigated the broad range of the water monitoring network design methodologies, such as statistical analysis, spatial interpolation, application of information theory, optimization techniques, physiographic analysis, user survey or expert recommendations and combinations of multiple methods [4,6,10–15]. A prior comprehensive review by Mishra and Coulibaly [10] reviewed evidence of declining hydrometric network density, highlighted the importance of quality data from well-designed networks and considered a range of approaches by which networks were designed. They also compared statistical, spatial interpolation, physiographic, sampling-driven and entropy-based approaches to hydrometric network design. Mishra and Coulibaly [10] were able to draw several conclusions about the importance of high quality hydrometric data for water resource management that remain valid. They also concluded that one of the most promising approaches for network design was the application of entropy methods highlighting early studies using the principle of maximum entropy and information transfer. Therefore, this review focuses on the recent studies that have applied information theory. Information theory was initially developed by Shannon in 1948 [16] to measure the information content in a dataset and has been applied to solve water resource problems. Recently, its applications extended to water monitoring network design by adopting the concept that the entropy would be able to explain the inherent information content in a monitoring station or a monitoring network. The basic objective has naturally been to have the maximum amount of information. In other entropy approaches, stations in a monitoring network would have the least sharable or common information, which is called transinformation. To achieve this, the stations should be as independent from each other as possible.

The scope of this paper includes water monitoring network design and evaluation studies that (1) applied entropy theory in the design process and (2) were published after the existing comprehensive review by Mishra and Coulibaly in 2009 [10]. To the best of our knowledge, no review exists that has focused on entropy applications to water monitoring network design previous to the 2009 study by Mishra and Coulibaly [10]. However, there has been considerable progress in the application of entropy theory to monitoring network design following the previous review, including new entropy-based measures, optimization techniques and approaches to estimating information content at ungauged stations. Therefore, a need was identified to consolidate knowledge and recent advances on the subject. For publications prior to 2010, the reader is referred to Mishra and Coulibaly [10]. This review firstly describes entropy concepts and various terms that are typically used in network design. The previous studies are then summarized by categorizing the type of networks; i.e., precipitation, streamflow and water level, soil moisture and groundwater and water quality monitoring networks. The integrated design method for multiple types of networks is also reviewed. We define some terminology hereafter to ensure a common understanding for the readers.

The term network evaluation is used when the network quality is assessed without changing any station, while network design is a general term that suggests some changes in stations. Specifically, network design includes network reduction, network expansion and network redesign. Network reduction is applied where some monitoring stations should be removed from the network. On the other hand, if financial flexibility meets the monitoring needs, further stations can be added to the existing network, called network expansion. Network redesign refers to rearranging stations without changing the number of stations. The term optimal network is to be used only if the network consists of optimal locations of stations that are identified by the actual use of an optimization technique.

2. Definitions of Entropy Terms as Applied to Water Monitoring Networks

2.1. Entropy Concept

In thermodynamics, entropy has been understood as a measure of disorder or randomness of a system. Shannon [16] extended the entropy concept to information theory by recognizing that uncertainty in a system will be decreased when information is added to the system. Therefore, the term entropy in information theory introduced by Shannon [16] in 1948 describes the amount of

information content in a random variable. The likelihood of an event is typically given by probability p . If a probability of an event is very high, such as 0.9999 or one, one will not be surprised, but can certainly anticipate the outcome. On the other hand, any low probability event is highly uncertain, so that a considerable amount of information can be given if this happens. Hence, the information from an event that occurred is inversely related to its probability, $1/p$ [17]. Suppose that there are two independent events A and B with their probabilities p_A and p_B , respectively. The probability of the joint occurrence of the events A and B can be $p_A p_B$, and the information gained by the joint event is then $1/(p_A p_B)$. However, the sum of information from each individual event is not equal to the information from the joint event, that is:

$$\frac{1}{p_A p_B} \neq \frac{1}{p_A} + \frac{1}{p_B} \quad (1)$$

The only transition that will make both sides of Equation (1) be equal is the logarithm [17–19], which can be written as:

$$\log \frac{1}{p_A p_B} = \log \frac{1}{p_A} + \log \frac{1}{p_B} = -\log p_A - \log p_B \quad (2)$$

Likewise, Tribus [20] showed that the uncertainty of an event with probability p is $-\log p$, which became a basis of the Shannon entropy, which is further described hereafter.

2.2. Marginal Entropy

When information is provided in a system, one can expect that the uncertainty of the system would be reduced; therefore, the amount of information that was given to a system by knowing a variable is called marginal entropy. If a random variable X is expected to have N outcomes with a probability distribution $P = \{p_1, p_2, \dots, p_N\}$, the (weighted) average information provided by the N joint events is given by:

$$H(X) = -p_1 \log p_1 - p_2 \log p_2 - \dots - p_N \log p_N = -\sum_{i=1}^N p_i \log p_i, \quad \sum_{i=1}^N p_i = 1, p_i \geq 0 \quad (3)$$

where $H(X)$ is the marginal entropy of a random variable X . Any base of the logarithm can be used in Equation (3), the choice depending on the problem given. In binary questions (i.e., yes or no questions), the base of two should be used, and the corresponding unit of entropy is bit. Similarly, unit trit for base 3, unit nat for base e and unit decibels or decit for base 10 are some example units of information. Recall that this review covers entropy applications for hydrometric network design, and the expected answer of the design process can be either “use/include/install the station” and “do not use/include/install the station” for the network to be optimal. Therefore, the logarithm in Shannon entropy calculation for hydrometric network design is most appropriate with a base of two. Then, the $H(X)$ value from Equation (3) will be understood as the information contents of a station X that can be delivered if installed.

If a variable K has a known value, the probability of an event will be one, while all the other alternative probabilities are zero. The information content in the variable K , $H(K)$, will be zero from Equation (3) representing that there is no uncertainty or a certain outcome. On the other hand, if a variable U has a uniform distribution (i.e., probability of each event is equal, $1/N$), the entropy of the variable U will be:

$$H(U) = \log N \quad (4)$$

The value of Equation (4) is often called as maximum entropy or saturated entropy. These two entropies, $H(K)$ and $H(U)$, define the minimum and maximum boundaries of entropy values, that is:

$$0 \leq H(X) \leq \log N \quad (5)$$

2.3. Multivariate Joint Entropy

While the marginal entropy described in Section 2.2 explains a univariate entropy, one can imagine how to calculate entropy values in a bivariate or a multivariate case. The total information contents from N variables can be calculated by using joint probability instead of univariate probability in Equation (3), given by:

$$H(X_1, X_2, \dots, X_N) = - \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \dots \sum_{i_N=1}^{n_N} p(x_{1,i_1}, x_{2,i_2}, \dots, x_{N,i_N}) \log_2 p(x_{1,i_1}, x_{2,i_2}, \dots, x_{N,i_N}) \quad (6)$$

where $H(X_1, X_2, \dots, X_N)$ is the joint entropy of N variables, $p(x_{1,i_1}, x_{2,i_2}, \dots, x_{N,i_N})$ is the joint probability of N variables and n_1, n_2, \dots, n_N are the numbers of class intervals of corresponding variable distributions [21]. If all variables are stochastically independent, the joint entropy from Equation (6) will be equal to the sum of marginal entropies, which becomes the maximum value of joint entropy. Therefore, the joint entropy is bounded by [21]:

$$0 \leq H(X_1, X_2, \dots, X_N) \leq \sum_{i=1}^N H(X_i) \leq N \log_2 N \quad (7)$$

2.4. Conditional Entropy

Conditional entropy explains a measure of information content of one variable that is not deliverable by other variables. If two random variables, A and B , are correlated, providing information from one variable may clear some uncertainty that the other variable has. In the case of no correlation between variables, the conditional entropy is equal to marginal entropy. That is:

$$H(A|B) = H(A, B) - H(B) \leq H(A) \quad (8)$$

where $H(A|B)$ is conditional entropy of the variable A when the information contents of the variable B is given. One can rewrite Equation (8) as:

$$H(A, B) = H(A|B) + H(B) = H(B|A) + H(A) \quad (9)$$

Furthermore, conditional entropy can be also presented mathematically using joint and conditional probabilities and Bayes theorem as:

$$H(A|B) = - \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} p(a_i, b_j) \log p(a_i|b_j) = - \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} p(a_i, b_j) \log \frac{p(a_i, b_j)}{p(b_j)} \quad (10)$$

2.5. Transinformation

The two variables, A and B , described in Section 2.4 will have some common or shared information, which is called transinformation or mutual information, because they are correlated.

$$T(A, B) = H(A) - H(A|B) = H(B) - H(B|A) = T(B, A) \quad (11)$$

where $T(A, B)$ is transinformation between the variables A and B . The larger the transinformation is, the higher those variables depend on each other. In other words, the transinformation indicates how much information content is transferrable from other variables. Similar to Equation (10), transinformation shall be written as [19]:

$$T(A, B) = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} p(a_i, b_j) \log \frac{p(a_i|b_j)}{p(a_i)} = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} p(a_i, b_j) \log \frac{p(a_i, b_j)}{p(a_i)p(b_j)} \quad (12)$$

Transinformation is typically used for measuring mutual information between two variables or two groups of variables as the generalized form for multivariate transinformation is given as:

$$\begin{aligned} T[(X_1, X_2, \dots, X_k); (X_{k+1}, X_{k+2}, \dots, X_N)] \\ = H(X_1, X_2, \dots, X_k) - H[(X_1, X_2, \dots, X_k)|(X_{k+1}, X_{k+2}, \dots, X_N)] \end{aligned} \quad (13)$$

2.6. Total Correlation

While transinformation and mutual information have the same definition, total correlation is not equivalent to them as the total correlation is a simple estimate that defines the amount of shared information typically of multiple variables. Simply, the total correlation is defined by the difference between the sum of marginal entropy of N variables and their joint entropy [22,23], which is given as:

$$C(X_1, X_2, \dots, X_N) = \sum_{i=1}^N H(X_i) - H(X_1, X_2, \dots, X_N) \quad (14)$$

If $N = 2$ in Equation (14), the total correlation will be equal to the transinformation or mutual information. However, the transinformation is only meaningful to two random variables as shown in Equations (11) to (13); therefore, the total correlation and the transinformation values would be different if $N > 2$.

2.7. Other Entropy Terms

The entropy terms described above (i.e., marginal entropy, joint entropy, conditional entropy, transinformation and total correlation) are the basic measures that have been typically used in entropy applications to water monitoring network design. While many studies developed specific approaches and applied for case studies using the basic entropy terms, some have extended the terms beyond them by deriving from or combining the basic measures. The detailed descriptions of the extended entropy terms are not included in this review, but briefly explained when needed in Section 3. Interested readers may refer to the original references.

3. Applications of Entropy to Water Monitoring Network Design

This section summarizes the recent applications of entropy theory to design water monitoring networks. The review was categorized by the types of networks, such as precipitation, streamflow or water level, soil moisture or groundwater and water quality networks. Then later, a hybrid design method for multivariate water monitoring networks was discussed. Table 1 presents brief summaries including network types, methods and key findings of the selected research articles that applied entropy theory for designing the water monitoring network and were published in 2010 or after to cover the most recent contributions since the existing review [10].

Table 1. Summary of significant contributions to water monitoring network design using entropy (author alphabetical order).

Authors/Year	Network Types	Study Areas	Methods/Entropy Measures	Key Findings
Alameddine et al., 2013 [24]	Water quality	Neuse River Estuary, NC, USA	-Total system entropy -Standard violation entropy -Multiple attribute decision making process -Analytical hierarchical process	-Networks designed using total system entropy and violation entropy of dissolved oxygen were similar -When measured water quality parameters have a low probability of violating water quality standards, their violation entropy is less informative
Alfonso et al., 2010 [25]	Water level	Pijnacker Region, The Netherlands	-Directional information transfer (DIT)	-Introduced total correlation for determining multivariate dependence in water monitoring network design -Information content and redundancy is dependent on the DIT between monitoring stations (DIT_{XY} or DIT_{YX})
Alfonso et al., 2010 [26]	Water level	Pijnacker Region, The Netherlands	-Max(Joint Entropy) min(Total Correlation) -Non-dominated Sorting Genetic Algorithm II (NSGA-II)	-Total correlation should be combined with joint entropy to get most information out of monitoring network
Alfonso et al., 2013 [27]	Streamflow	Magdalena River, Colombia	-Max(Joint Entropy) min(Total Correlation) -Rank-based iterative approach	-Rank method is useful in finding extremes on Pareto front -When iteratively selecting stations, the information content of the network is not guaranteed to be maximum if the network contains the station with the most information
Alfonso et al., 2014 [28]	Water level	North Sea, The Netherlands	-Max(Joint Entropy) min(Total Correlation) -Ensemble entropy -NSGA-II	-By creating an ensemble of solutions through varying the bin size of the initial Pareto optimal solution set, the authors highlight the uncertainty related to choosing bin size
Boroumand and Rajaei, 2017 [29]	Water quality	San Francisco Bay, CA, USA	-Transinformation-distance (T-D) curve	-Using T-D curve they were able to reduce the network from 37 to 21 monitoring stations. -New network covered entire study area without having redundant data
Brunsell, 2010 [30]	Precipitation	Continental United States	-Relative entropy -Wavelet multi-resolution analysis	-The temporal scaling regions identified (1) synoptic, (2) monthly to annual, (3) interannual patterns -Little correlation between relative entropy and annual precipitation except for breakpoint at 95° W Lat
Fahle et al., 2015 [31]	Water level/Groundwater level	Spreewald region, Germany	-MIMR, max(Joint Entropy + Transinformation – Total Correlation) -Subsets of time series data	-Found using subsets of the available time series data could better identify important stations -Showed water levels across network react similarly during high precipitation and are more unique during dry periods -Consequently method can allow for design of network which focuses on floods or droughts

Table 1. Cont.

Authors/Year	Network Types	Study Areas	Methods/Entropy Measures	Key Findings
Hosseini and Kerachian, 2017 [32]	Groundwater level	Dehgolan plain, Iran	-Marginal entropy -Data fusion of spatiotemporal kriging and ANN model -Value of information (VOI)	-Network reduction from 52 to 42 (35 high priority and 7 low priority) stations while standard deviation of average estimation error variance stayed the same -Found sampling frequency of high priority stations should be every 20 days and low priority should be every 32, based on analysis of stations selected using VOI
Hosseini and Kerachian, 2017 [33]	Groundwater level	Dehgolan plain, Iran	-Bayesian maximum entropy (BME) -Multi-criteria decision making based on ordered weighted averaging	-Network reduction from 52 to 33 stations while standard deviation of average estimation error variance stayed the same -Sampling frequency increased from 4 weeks to 5 weeks
Keum and Coulibaly, 2017 [34]	Precipitation/Streamflow	Columbia River basin, BC, Canada. Southern Ontario, Canada	-Dual Entropy and Multiobjective Optimization (DEMO) to max(Joint Entropy) and min(Total Correlation)	-Found that networks obtain significant amount of information from 5 to 10 years of data periods, and total correlation tends to be stabilized within 5 years by applying daily time series -Recommended minimum 10 years data periods for designing precipitation or streamflow networks using daily time series
Keum and Coulibaly, 2017 [35]	Integrated	Southern Ontario, Canada	-DEMO to max(Joint Entropy), min(Total Correlation), and max(Conditional Entropy) -Sturge, Scott and rounding binning methods	-Precipitation and streamflow networks were designed simultaneously. -Binning methods were compared and concluded that the optimal networks can be altered due to the binning methods
Kornelsen and Coulibaly, 2015 [36]	Soil Moisture	Great Lakes Basin, Canada-USA	-DEMO to Max(Joint Entropy) min(Total Correlation) -SMOS satellite data	-Optimum networks were different for ascending and descending overpasses -Combining overpass data resulted in complimentary spatial distribution of stations
Leach et al., 2015 [37]	Streamflow	Columbia River basin, BC, Canada. Southern Ontario, Canada	-DEMO to Max(Joint Entropy) min(Total Correlation) -Streamflow signatures -Indicators of hydrologic alteration (IHA)	-Found that including streamflow signatures as design objective increases network coverage in headwater areas. -Found including IHAs increases network coverage in downstream and urban areas.
Leach et al., 2016 [38]	Groundwater level	Southern Ontario, Canada	-DEMO to Max(Joint Entropy) min(Total Correlation) -Annual recharge	-Found that considering spatial distribution of annual recharge can improve network coverage

Table 1. Cont.

Authors/Year	Network Types	Study Areas	Methods/Entropy Measures	Key Findings
Lee, 2013 [39]	Water quality	Hagye Basin, South Korea	-Marginal entropy analogous cost function -Genetic algorithm	-Developed computationally efficient way to design a monitoring network in an ungauged basin
Lee et al., 2014 [40]	Water quality	Sanganmi Basin, South Korea	-Multivariate transinformation -Genetic algorithm	-Developed method based on maximizing information content to design a water quality monitoring network in a sewer system
Li et al., 2012 [41]	Streamflow/Water level	Brazos River basin, USA. Pijnacker, The Netherlands	-MIMR, max(Joint Entropy + Transinformation – Total Correlation)	-Developed maximum information minimum redundancy method (MIMR) -Found it to be better at locating high information content stations for a monitoring network
Mahjouri and Kerachian, 2011 [42]	Water quality	Jajrood River, Iran	-Information transfer index (ITI) distance and time curves -Micro genetic algorithm (MGA)	-The MGA was used to find the optimal combination of monitoring stations which minimize the temporal and spatial ITI -Found that the sampling frequency and number of stations could be increased in the monitoring network
Mahmoudi-Meimand et al., 2016 [43]	Precipitation	Karkheh, Iran	-Transinformation entropy -Kriging error variance -Weighted cost function to select from Monte Carlo generated networks	-Consideration of spatial analysis error and transinformation entropy improved network design
Masoumi and Kerachian, 2010 [44]	Groundwater quality	Tehran, Iran	-Transinformation-distance (T-D) curve -Transinformation-time (T-T) curve -C-mean clustering -Hybrid genetic algorithm (HGA)	-Developed different T-D curves based on homogeneous clusters of existing monitoring stations -Used HGA to find optimal network with maximum spatial coverage and minimum transinformation -Showed that sampling frequency could be optimized in the same way
Memarzadeh et al., 2013 [45]	Water quality	Karoon River, Iran	-Information transfer index (ITI) distance curve -Homogenous zone clustering -Dynamic factor analysis (DFA)	-Increased monitoring network without increasing redundant information

Table 1. Cont.

Authors/Year	Network Types	Study Areas	Methods/Entropy Measures	Key Findings
Mishra and Coulibaly, 2010 [46]	Streamflow	Selected basins across Canada	-Transinformation index -Marginal, joint, and transinformation	-Used information theory to highlight critical areas across Canada in need of monitoring -Found that several watersheds are information deficient and would benefit from increased monitoring
Mishra and Coulibaly, 2014 [47]	Streamflow	Selected basins across Canada	-Transinformation index -Seasonal streamflow information (SSI)	-Evaluated and highlighted the effects of seasonal climate on streamflow network design
Mondal and Singh, 2012 [48]	Groundwater level	Kodaganar River basin, India	-Marginal entropy, joint entropy, transinformation -Information transfer index (ITI)	-Identified high priority monitoring stations using marginal entropy -ITI was used to evaluate monitoring network, showed that it could be reduced
Samuel et al., 2013 [49]	Streamflow	St. John and St. Lawrence River basins, Canada	-Combined Regionalization-DEMO -Max(Joint Entropy) min(Total Correlation)	-Proposed combined regionalization dual entropy multi-objective optimization approach to design of minimum optimal network that meets World Meteorological Organization (WMO) guidelines -Found that the location of new monitoring stations added to a network depends on the current network density
Santos et al., 2013 [50]	Precipitation	Portugal	-ANN sensitivity analysis -Mutual Information criteria -K-means with Euclidean distance	-Compared three clustering methods to reduce station density -Best method was case dependent -All subset networks reproduced spatial precipitation pattern
Stosic et al., 2017 [51]	Streamflow	Brazos River, TX, USA	-Joint permutation entropy	-Used joint permutation entropy to account for ordering of time series data to better account for station information -Found that the most efficient measurement window was seven days when compared to daily and monthly
Su and You, 2014 [52]	Precipitation	Shihmen Reservoir Taiwan	-Developed 2D transinformation-distance (T-D) model -T-D model used to interpolate network information	-Network designed by maximizing additional information provided by station given regionalized transinformation -Temporal scale has significant influence on information delivery

Table 1. Cont.

Authors/Year	Network Types	Study Areas	Methods/Entropy Measures	Key Findings
Uddameri and Andruss, 2014 [53]	Groundwater level	Victoria County Groundwater Conservation District, TX, USA	-Marginal entropy -Monitoring priority index (MPI)	-Compared MPI found using kriging to MPI found using marginal entropy -Showed entropy derived MPI to be more conservative measure
Wei et al., 2014 [54]	Precipitation	Taiwan University Experimental Forest, Taiwan	-Joint Entropy of hourly, monthly, dry/wet months and annual rainfall at 1, 3, 5 km grids	-Station priority changes at different spatiotemporal scales -Temporal scales have more significant changes on joint entropy values than spatial scales -Long time and short spatial scales require fewer stations for stable joint entropy
Werstuck and Coulibaly, 2016 [55]	Streamflow	Ottawa River Basin, Canada	-Transinformation index -DEMO to Max(Joint Entropy) min(Total Correlation)-Streamflow signatures -Indicators of hydrologic alteration (IHA)	-Compared regionalized data from McMaster University-Hydrologiska Byråns Vattenbalansavdelnin (MAC-HBV) and Inverse Distance Weighting—Drainage Area Ratio (IDW-DAR) and found IDW-DAR to be more adequate for generating synthetic time series for potential monitoring stations -Critical areas highlighted by TI index method were the same areas where additional stations were added using DEMO method
Werstuck and Coulibaly, 2017 [56]	Streamflow	Ottawa River Basin, Canada	-Transinformation index -DEMO to max(Joint Entropy) min(Total Correlation)	-Transinformation index analysis is not significantly affected by scaling -Scaling effects are noticeable when DEMO method was applied
Xu et al., 2015 [57]	Precipitation	Xiangjiang River Basin, China	-Mutual Information (MI) of rain gauges -Designed network by $\min(\Sigma[MI]), \min(\text{bias}), \max(\text{NSE})$ -Resampled rainfall used in Xinanjiang and SWAT models	-Lumped model performance was stable with different Pareto optimal networks -Distributed model performance improves with number of stations
Yakirevich et al., 2013 [58]	Groundwater quality	OPE3 research site, Maryland, USA	-Principle of minimum cross entropy (POMCE) -Hydrus-3D	-Using POMCE with two variants of Hydrus-3D, additional monitoring stations were added where the difference between the models was greatest

3.1. Precipitation Networks

The design of a representative precipitation monitoring network is an important and still challenging task for which an entropy approach is well suited. High quality precipitation information is necessary for streamflow and flood forecasting, surface water management, agricultural management, climate process understanding and many other applications. However, precipitation is well known to be highly variable in both space and time [59] and often statistically represented by highly skewed distributions [60] making the application of parametric analysis methods difficult. These challenges also extend to entropy-based approaches for precipitation monitoring. For example, the marginal entropy has been found to be well correlated with total precipitation in northern Brazil because the probability distribution in regions with higher rainfall tended to be more uniform and less skewed [61]. In contrast, Mishra et al. [59] found that the marginal disorder index (MDI), which is the ratio of observed entropy to the maximum possible entropy at a given site, was inversely related to mean annual rainfall in the U.S. state of Texas, where MDI was found to vary seasonally. Brunsell [30] studied the entropy from monitoring stations across the United States where little correlation was found between precipitation and marginal entropy with the exception of a breakpoint in entropy at -95° longitude corresponding to high temporal variability precipitation patterns. It has also been noted by several studies that the temporal sampling of precipitation is an important consideration for calculating entropy and for designing precipitation networks [54,59]. At finer timescales (hourly to daily), precipitation is highly variable resulting in higher overall entropy, whereas longer time periods (monthly to annual) have less variability resulting in lower marginal entropy [30,52,59]. The dependence on spatial and temporal scales has also been identified in a network design application. Wei et al. [54] prioritized potential stations in Central Taiwan to maximize the joint entropy of the network at hourly, monthly and annual temporal scales, as well as 1-, 3- and 5-km spatial scales. They found that priority stations changed with both spatial and temporal scales, where changes in temporal scales resulted in more significant changes in station priority than spatial-rescaling. The decrease in entropy at longer timescales also had an impact on station density where fewer stations were required to reach a stable joint entropy value for longer time scales [54]. These findings demonstrate the important first consideration of network objectives when determining the spatial and temporal sampling used to calculate entropy. However, the research on this topic is still limited, and more work is needed to provide robust guidance on sampling strategies.

Several approaches have been proposed to design or redesign a precipitation monitoring network using one or more entropic measures. Many of these approaches are initialized by building a network around a central station usually selected as the station with the highest marginal entropy [43,62–64]. In urban Rome, Ridolfi et al. [62] selected stations for the precipitation network by sequentially finding the next station that minimized the conditional entropy of the network and adding that station to the network. A similar approach was taken by Yeh et al. [63] to expand a precipitation network in Taiwan. Hourly rainfall data were normalized with a Box-Cox transform and kriging used to interpolate rainfall to candidate grid cells. The joint entropy of the network was calculated using an analytic equation for joint entropy valid for normal data [65], and stations were added sequentially that had the lowest conditional entropy with the rest of the network. The final number of stations needed by the network was accepted when 95% of the network information was captured [63]. Awadallah [64] applied multiple entropy measures sequentially to add stations to a precipitation network. The first new stations were selected as those with the highest entropy. The second station was chosen to minimize the mutual information and the third as the station that maximized conditional entropy.

The aforementioned approaches all sequentially add single stations to a monitoring network based on a single criterion. Mahmoudi-Meimand et al. [43] presented a methodology to add stations to a network based on a multi-variate cost function. Precipitation data were spatially interpolated from existing stations using the kriging approach where the kriging error associated with the rainfall estimation is calculated as the kriging error variance. Their method selected the station that maximized transinformation entropy and minimized error variance using a weighted average of both measures as

an objective [43]. This approach balanced the information content in the network with the errors in the interpolation method. Xu et al. [57] used a multi-objective approach to simultaneously select a subset of stations that minimized the sum of pairwise mutual information, minimized bias and maximized Nash–Sutcliffe efficiency. Solutions were generated via Monte Carlo sampling, and network solutions falling along the Pareto front were found as compromise solutions. Coulibaly and Keum [66] and Samuel and Coulibaly [67] also used a multi-objective approach to add stations to snow monitoring networks in Canada. Their approach used a genetic algorithm to find networks that maximized the joint entropy and minimized the total correlation of the network to form a Pareto front of optimal network designs, some of which also included network cost in the optimization [35,67].

A challenge to an entropy-based approach to adding stations to a precipitation network is the requirement to have data available for candidate points. For precipitation, this can be challenging because data at shorter time scales in particular are well known to be non-normal. Most studies use the kriging approach for interpolation [43,63,64] and address the need for normally distributed data using a Box-Cox transform. Samuel and Coulibaly [67] addressed the interpolation problem by using the external data from the Snow Data Assimilation System (SNODAS) for candidate stations. Su and You [52] presented a unique approach to adding stations that maximized the information content of the network. In most literature cases, entropic measures at ungauged sites are determined by interpolating observations of precipitation across a watershed. Su and You [52] calculated the transinformation between neighbouring stations to develop a 2D transinformation-distance relationship. In contrast to transferring data to ungauged stations, this approach transferred transinformation to ungauged stations and selected a site with the maximum transinformation. This approach should be further tested and contrasted with the data transfer approach.

As previously stated, precipitation data are of critical importance for a variety of applications. Despite this, few studies have explored the impact of precipitation networks designed with an entropy approach for actual water resource applications. Applications found in the literature have taken the reasonable approach of using entropy to reduce network density for comparison to a network that included all stations. In Portugal, Santos et al. [50] compared artificial neural networks, K-means clustering and mutual information (MI) criteria for reducing the density of a precipitation network for drought monitoring at different time scales. They found the best performing reduction method was case dependent depending on the region and time scale applied, but noted that all methods performed well. They also found that all subset networks could reliably reproduce the spatial precipitation pattern. Xu et al. [57] used the multi-objective approach previously described to select a subset of precipitation stations from a dense network in the Xiangjiang River Basin in China. Rainfall from the subset networks was used to force the lumped Xinanjiang hydrological model [68] and the distributed SWAT hydrological model [69]. The author's found that lumped model performance became stable with a subset of 20 to 25 stations, whereas the distributed model's performance continued to increase as more stations were added to the network [57]. These analyses are important to demonstrate the utility of precipitation networks and the advantages of entropy-based approaches in designing precipitation networks.

3.2. Streamflow and Water Level Networks

Water quantity monitoring, such as streamflow rates and water level, is one of the essential tasks for water management to prevent damage to nature and human beings from flooding. A successful floodplain management or flood forecasting and warning system can be feasible through expert forecasters who implement well-calibrated models and reliable tools using quality data [70]. The design of water quantity monitoring network has been well implemented because of not only the good performance of entropy-based methods, but also the unaffectedness by the zero effect, which is caused by discontinuity of probability density function due to zero values in data, except for the ephemeral or intermittent streams. To deal with the zero effect in entropy calculations, Chapman [71] and Gong et al. [60] separated the marginal entropy Equation (3) to nonzero terms and zero values,

which are certain. While Gong et al. [60] summarized the possible issues in entropy calculations from hydrologic data as effects due to zero values, histogram binning including skewness consideration and measurement errors, some studies noticed that the length and the location of time window also affect entropy calculations and the corresponding network design. Fahle et al. [31] observed the temporal variability of station rankings by shifting the time window for the design of water level network of a ditch system in Germany. Mishra and Coulibaly [47] also found the dependency of the seasonality on the efficiency of hydrometric networks. Stosic et al. [51] found an inverse relationship between the network density and sampling time interval as the larger number of monitoring stations is required if the time interval is shorter and vice versa. Keum and Coulibaly [34] analyzed the temporal changes of entropy measures and optimal networks by applying daily time series for streamflow network design. They found that the information gain of a monitoring network is not significant when the length of time series is longer than 10 years, and the total correlation tends to stabilize within five years of data. The optimal networks using the data lengths of 5, 10, 15 and 20 years also show that there are no significant differences in the results from 10 years or longer while the optimal network using five years of data was evidently different from others. Werstuck and Coulibaly [56] analyzed scaling effects by considering two study areas. Specifically, one study area is a small watershed, which is a part of another study area. After applying the transinformation analysis and the multi-objective optimization, they concluded that the optimal networks tend to be affected by scaling while transinformation index does not.

Mishra and Coulibaly [46] evaluated the effects of the class intervals and the infilling missing data by applying the linear regression method to daily time series and concluded that the station rankings based on the transinformation values were not significantly changed. Li et al. [41] also investigated the changes of station rankings based on the maximum information minimum redundancy (MIMR) approach and obtained the similar conclusion. However, Fahle et al. [31] and Keum and Coulibaly [35] drew the opposite opinion that station rankings can be affected by the binning method that defines the class intervals. The conflict comes from the selection of the binning methods compared. The former group applied different parameters to a single binning method, the mathematical floor function. However, the latter group compared other binning methods with the floor function. Considering that Alfonso et al. [25] found that the design solutions were not common in some cases from the sensitivity analysis of the parameter of the mathematical floor function, it is not recommended to use a specific binning method without any consideration.

As discussed in the review of precipitation networks in Section 3.1, network redesign and network expansion require data at candidate locations, which are ungauged. Alfonso et al. [27] applied a one-dimensional hydrodynamic model to generate the discharge time series. The model estimated discharge at each segment, which divides rivers with approximately 200 m increments longitudinally. The use of hydrodynamic model enabled to determine the critical monitoring locations in the main stream and its tributaries. On the other hand, Samuel et al. [49] combined regionalization techniques with entropy calculation in order to estimate the discharge at candidate locations. They compared the performance of various regionalization methods including not only a conceptual hydrologic model, but also spatial proximity, physical similarity and their combinations with drainage area ratio. Based on the performance statistics by applying multiple basins, inverse distance weighting coupled with drainage area ratio performed the best, and this conclusion has been adopted in several studies [34,35,37,55].

Some studies have extended the entropy applications for the streamflow monitoring network design. Stosic et al. [51] proposed the concept of permutation entropy, which is able to differentiate based on the order of sequential observations, as well as the histogram frequency in basic Shannon entropy measures. Even though histograms from two different observations are the same, the permutation entropy value tends to be higher if there are more variations between time steps. However, the network design studies using the permutation entropy are still limited. On the other hand, Leach et al. [37] applied additional features to the network design. While the common objectives in water monitoring network design using an optimization technique are to maximize the information

and to minimize the redundancy in the network, they additionally considered the physical properties of watersheds, such as the streamflow signatures [72,73] and the indicators of hydrologic alterations (IHAs) [74,75]. After the comparison of the optimal streamflow monitoring networks with and without considerations of the streamflow signatures and IHAs, it was concluded that inclusion of basin physical characteristics yielded a better coverage of the selected locations of the optimal networks.

3.3. Soil Moisture and Groundwater Networks

Soil moisture is a critical water variable as the interface between the atmosphere and subsurface. Unfortunately, the monitoring of soil moisture is very sparse compared to its spatial variability. To design an optimum network for monitoring soil moisture in the Great Lakes Basin, Kornelsen and Coulibaly [36] proposed using data from the Soil Moisture and Ocean Salinity (SMOS) satellite [76] to design a soil moisture monitoring network using the DEMO algorithm of Samuel et al. [49]. Grid cells were selected to add monitoring stations that optimally maximized joint entropy while minimizing total correlation using only the satellite data. The ascending and descending overpasses were found to contain different information, and the spatial distribution of a network designed with both overpasses was found to contain complimentary features from both datasets [36].

Groundwater monitoring allows for a better understanding of the hydrogeology in an area. This is achieved through groundwater quality and quantity monitoring. Groundwater quality monitoring is used to detect contaminant plumes or for long-term monitoring (LTM) of post remediation effects, and groundwater quantity monitoring is used to determine available water for drinking, irrigation and industry. However, monitoring groundwater is inherently difficult due to physical barriers between observers and the water. Through the understanding of subsurface flow physics and with flow and contaminant transport models such as MODFLOW, MODPATH and MT3D [77–79], we can simulate the behaviour of groundwater. Unfortunately, our simulations are not always accurate, and the models require real-world observations to be calibrated and validated. Due to constraints such as accessibility and financial cost, it is not feasible to monitor at every possible location in an area of interest. It is instead ideal for an optimal monitoring network to be designed to allow for the best placement of monitoring stations and to determine the ideal measurement frequencies. The merit of using information theory entropy has been shown in several cases of groundwater network design [31–33,38,44,48,53,58].

Various methods that utilize information theory entropy have been developed for use in designing optimal groundwater monitoring networks. These include the use of entropy measures in both single and multi-objective optimization problems and are used in network reduction [32,33], expansion [38,58] and redesign [44], as well as have been used to highlight vulnerable areas in an area that should be monitored [53]. In identifying vulnerable areas in the Victoria County Groundwater Conservation District (VCGCD) in Texas, USA, Uddameri and Andruss [53] developed a monitoring priority index (MPI) based on a weighted stakeholder preference to highlight the areas of interest. They compared kriging standard deviation and marginal entropy as metrics to characterize groundwater variability and found entropy to be the more conservative metric.

In areas where there is excessive monitoring, Mondal and Singh [48] showed the information transfer index (ITI), the quotient of joint entropy and transinformation, could be used to evaluate the existing monitoring network. Through this evaluation, redundant monitoring stations (wells) could be identified and removed from the groundwater monitoring network. It may also be the case that the existing groundwater monitoring network is not adequate and additional monitoring stations are needed. Yakirevich et al. [58] developed a method that utilizes minimum cross entropy (MCE) to sequentially add monitoring stations to a network. MCE was used as a metric to quantify the difference between two variants of a Hydrus-3D model [80], and the monitoring stations were added to the network where the difference between models was largest. A multi-objective approach for adding monitoring stations to a groundwater monitoring network was applied by Leach et al. [38], which utilized two entropy measures, total correlation and joint entropy and a metric used to quantify the

spatial distribution of annual recharge; the results of which were used to develop maps that highlight areas in which additional monitoring stations should be added. The majority of network design experiments look at the entire available time series when calculating entropy measures; however, Fahle et al. [31] showed that using a combination of MIMR and subsets of the data series could be more ideal. The subsets were used to represent the intra-annual variability of groundwater levels. This method identified locations which were consistently important through each subset and found that monitoring stations showed similarities during wet periods and uniqueness during dry periods. Fahle et al. [31] also suggest that a consequence of using subsets of data allows for the design of a network, which can be focused on floods or droughts.

One issue that can arise with entropy-based methods is the need for lengthy data series to produce accurate measures of entropy. Unfortunately, the area of interest for new monitoring stations will not have available data for all possible locations. To work around this limitation transinformation-distance (T-D) curves have been applied in the design of optimal groundwater monitoring networks [44,81]. In these studies, T-D curves were developed for sub areas within the desired study area based in different clustering methods. Additionally, Masoumi and Kerachian [44] showed that this method could be applied temporally as transinformation-time curves which could then be used to optimize the temporal sampling frequency of the stations. It should be noted that both previously mentioned studies were applied in the same study area using slightly different methods for clustering monitoring stations, and both produced different groundwater monitoring networks that could be considered optimal. This highlights an issue with optimal monitoring network design in that it can be subjective and does not have a singular solution. A comparison of Hosseini and Kerachian [32,33] also illustrates this issue, where through the use of different entropy measures, marginal entropy and Bayesian maximum entropy and optimization techniques, one experiment found the optimal monitoring network included 42 monitoring stations while the other only included 33 stations.

3.4. Water Quality Networks

The importance of water quality monitoring networks is their ability to assist in identifying those parameters that exceed water quality standards. Several water quality monitoring strategies, including two methods that utilized entropy measures [42,45], were recently reviewed by Behmel et al. [15]. This review found that identifying a single approach to water quality monitoring network design would be virtually impossible. Despite this, various applications of the transinformation-distance curve methods have shown promise in the optimal redesign and reduction of water quality monitoring networks [29,42,45]. Lee [39] found that by maximizing the multivariate transinformation between chosen and unchosen stations, using the storm water management model to simulate the total suspended solids and a GA for optimization, an optimal water quality network could be designed for a sewer system. Banik et al. [82] compared information theory, detection time and reliability measures for the design of a sewer system monitoring network through both single and multi-objective optimization approaches. It was shown that for a small monitoring network, the methods had similar performances, while the single objective detection time-based method had slightly better performance when the number of monitoring station is larger. Alameddine et al. [24] used exceedance probabilities to determine violation entropy of dissolved oxygen and chlorophyll-a in the Neuse River estuary. Along with violation entropy, the total system entropy was used as a measure to identify areas of importance of monitoring. A multi-objective optimization scheme based on expert assigned weights was used to develop a compromise solution from the three entropy measures. Ultimately, the method allowed for the identification of high uncertainty areas, which would benefit from future water quality monitoring. Data availability is an issue when using entropy methods, particularly when attempting to use them in the design of a monitoring network in an ungauged basin. To address this, Lee et al. [40] developed a method that uses a measure analogous to marginal entropy. This method uses characteristics of the basin such as the length and number of reaches in the river network as part of the cost function, which is then optimized using a combined GA and filtering algorithm. This was

shown to be a computationally-efficient method for use in optimal network design of an ungauged river basin.

3.5. Integrated Network Design

To the best of our knowledge, almost all of the previous studies about water monitoring network design have focused on a specific network type (i.e., considering a single hydrologic variable in each study) as reviewed in the previous sections. However, considering that hydrologic processes are interconnected in a water cycle, there are causes and effects between hydrologic variables. For instance, if a noticeable amount of precipitation occurs, streamflow or groundwater level is likely increased; hence, the information content of a variable may affect that of other variables. Keum and Coulibaly [35] developed a multivariate network design method by taking conditional entropy as the measure of information that is independent to a given variable. In their study, the method designed precipitation and streamflow monitoring networks simultaneously. Specifically, the method followed the traditional multi-objective approach that maximizes joint entropy and minimizes total correlation, but added another objective that maximizes conditional entropy of streamflow network given precipitation network to mimic the direction of the water cycle as streamflow may fluctuate due to precipitation. After comparing the integrated design with the single-variable design, their results showed that the effectiveness of network integration mostly came from reducing the number of additional precipitation stations. It was also found that the integrated network design approach allows adding a precipitation station at a location that will benefit the stream gauge network.

4. Conclusions and Recommendations

It is evident that successful water management cannot be achieved without proper water monitoring networks. Although there has been much progress in network design methods and applications, a standardized design methodology has not yet emerged. After the pioneering invention of information theory in the 1940s, entropy concepts have been applied in various applications with recent efforts on network design problems. The unique benefit of this approach is that a water monitoring network can be evaluated or designed based on the information the network monitors, which is in contrast to the set station densities proposed by WMO guidelines; the advantage of the former being that a network could be better tailored to specific applications or optimized to provide the most gain at densities lower than those suggested in WMO guidelines. In addition, when combined with multi-objective optimization techniques, users' specific criteria can be included in the optimal network design process.

This manuscript provides a comprehensive review of the recent research attainments and their applications in entropy-based water monitoring network design. The literature has demonstrated the use of various information theory measures and adaptations thereof for use in network design with an emerging consensus that the goal of these network design methods is to select the stations that provide the most information to the monitoring network while simultaneously being independent of each other. Through rigorous testing, information theory has proven to be a robust tool to use when evaluating and designing an optimal water monitoring network. However, when it comes to evaluating the optimal design, there are still issues that need to be addressed.

The first is that an optimal monitoring network design can be found based on specified design criteria; however, the practical application of the new optimal monitoring network is rarely evaluated in a hydrologic or other model [11,57]. This type of numerical experiment is an important requirement to evaluate the utility of a network rather than just identifying its optimality or information content. Further, it is an important exercise to identify the benefits of entropy-based network designs in order to convince decision makers of the importance of adopting entropy approaches.

Another issue with the optimal network is that it can be subjective, based on choices made in the calculation of entropy and the design method chosen, especially when additional objective functions are considered in the design. This extends to the method selected for finding the optimal

monitoring network, whether it is found using an iterative method where one station is added at a time or a collection of stations is added all at once. Research has also shown that data length, catchment scale and ordering can influence the design of an optimal network [31,34,51]. Finally, when using discrete entropy, the binning method has been shown to influence the final network design [35]. The influence of binning on entropy calculation has received greater attention in other geophysical network design applications [83–85], and similar consideration should be given in the field of water resources, particularly owing to the unique and difficult nature of water variables (e.g., streamflow, precipitation) spatial and temporal distribution [30,60]. Thus, explicit consideration is needed when choosing the bins based on the intended application of the monitoring network and further research to provide guidance specific to water monitoring networks. Therefore, despite the possibility of finding an optimal network design in a formal sense, the subjectivity induced by the designer's choices, and the lack of standardized design methods, must be recognized. Future research should focus on comparative studies among multiple entropy design methods, discretization approaches and data characteristics. The current literature provides many novel entropy design approaches and the evolution of concepts, but rigorous comparisons are critical to provide generic guidelines for network design. Despite the potential sources of subjectivity identified, entropy methods remain one of the most objective approaches for network design.

In particular, more work is needed on spatial and temporal scaling of data for entropy calculation to provide robust guidance to decision makers. Many new methods and optimization techniques have been reviewed herein, but few examples were found in the literature that explored the data characteristics used in those techniques. Further research is required to provide guidance on the proper length of data in water monitoring network design [34], the sampling frequency of the data [54] and the spatial scale at which information should be measured for various monitoring network applications.

The aforementioned issues are considered crucial gaps that need to be filled to enable practical recommendations or guidelines for a widespread adoption of entropy approaches for designing optimal water monitoring networks. In addition, the comparative studies of entropy-based methods reviewed herein should be robustly compared to network design methods from other disciplines, such as geostatistics, to identify areas of equivalence and disparity [10]. Considerable advances have occurred over the past decade as reviewed herein, and measures derived from Shannon's base equation [16] have reached a high level of maturity for the task of network design. We challenge the research community to put a similar creativity into the joint consideration of the nexus of data characteristics, network design and applications, all of which are intricately linked.

Acknowledgments: This research was supported by the Natural Science and Engineering Research Council (NSERC) of Canada, Grant NSERC Canadian FloodNet (NETGP-451456).

Author Contributions: The structure of this review paper was built and the contents were written by all authors. Specifically, Jongho Keum contributed Sections 1, 2, 3.2 and 3.5. Kurt C. Kornelsen contributed Sections 3.1 and 3.3. James M. Leach contributed Sections 3.3, 3.4 and 4. Paulin Coulibaly defined the scope of the review and supervised and reviewed the entire manuscript. Although the primary contributing sections are as indicated above, all authors cross-reviewed other co-authors' sections. All authors have read and approved the final manuscript.

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

References

1. Langbein, W.B. Overview of Conference on Hydrologic Data Networks. *Water Resour. Res.* **1979**, *15*, 1867–1871. [[CrossRef](#)]
2. Herschy, R.W. *Hydrometry: Principles and Practice*, 2nd ed.; John Wiley and Sons Ltd.: Chichester, UK, 1999.
3. Boiten, W. *Hydrometry*; A.A. Balkema Publishers: Lisse, The Netherlands, 2003.
4. Nemeč, J.; Askew, A.J. Mean and variance in network-design philosophies. In *Integrated Design of Hydrological Networks (Proceedings of the Budapest Symposium)*; Moss, M.E., Ed.; International Association of Hydrological Sciences Publication: Washington, DC, USA, 1986; pp. 123–131.

5. Rodda, J.C.; Langbein, W.B. *Hydrological Network Design—Needs, Problems and Approaches*; World Meteorological Organization: Geneva, Switzerland, 1969.
6. World Meteorological Organization. *Casebook on Hydrological Network Design Practice*; Langbein, W.B., Ed.; World Meteorological Organization: Geneva, Switzerland, 1972.
7. Davis, D.R.; Duckstein, L.; Krzysztofowicz, R. The Worth of Hydrologic Data for Nonoptimal Decision Making. *Water Resour. Res.* **1979**, *15*, 1733–1742. [[CrossRef](#)]
8. Pilon, P.J.; Yuzyk, T.R.; Hale, R.A.; Day, T.J. Challenges Facing Surface Water Monitoring in Canada. *Can. Water Resour. J.* **1996**, *21*, 157–164. [[CrossRef](#)]
9. U.S. Geological Survey. *Streamflow Information for the Next Century—A Plan for the National Streamflow Information Program of the U.S. Geological Survey*; U.S. Geological Survey: Denver, CO, USA, 1999.
10. Mishra, A.K.; Coulibaly, P. Developments in Hydrometric Network Design: A Review. *Rev. Geophys.* **2009**, *47*. [[CrossRef](#)]
11. Chacon-hurtado, J.C.; Alfonso, L.; Solomatine, D.P. Rainfall and streamflow sensor network design: A review of applications, classification, and a proposed framework. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 3071–3091. [[CrossRef](#)]
12. Moss, M.E. *Concepts and Techniques in Hydrological Network Design*; World Meteorological Organization: Geneva, Switzerland, 1982.
13. Van der Made, J.W.; Schilperoort, T.; van der Schaaf, S.; Buishand, T.A.; Brouwer, G.K.; van Duyvenbooden, W.; Becinsky, P. *Design Aspects of Hydrological Networks*; Commissie voor Hydrologisch Onderzoek TNO: The Hague, The Netherlands, 1986.
14. Pyrcce, R.S. *Review and Analysis of Stream Gauge Networks for the Ontario Stream Gauge Rehabilitation Project*, 2nd ed.; Watershed Science Centre: Peterborough, ON, Canada, 2004.
15. Behmel, S.; Damour, M.; Ludwig, R.; Rodriguez, M.J. Water quality monitoring strategies—A review and future perspectives. *Sci. Total Environ.* **2016**, *571*, 1312–1329. [[CrossRef](#)] [[PubMed](#)]
16. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
17. Batty, M. Space, scale, and scaling in entropy maximizing. *Geogr. Anal.* **2010**, *42*, 395–421. [[CrossRef](#)]
18. Singh, V.P. *Entropy Theory in Hydrologic Science and Engineering*; McGraw-Hill Education: New York, NY, USA, 2015.
19. Lathi, B.P. *An Introduction to Random Signals and Communication Theory*; International Textbook Company: Scranton, PA, USA, 1968.
20. Tribus, M. *Rational Descriptions, Decisions and Designs*; Irvine, T.F., Hartnett, J.P., Eds.; Pergamon Press: Oxford, UK, 1969.
21. Krstanovic, P.F.; Singh, V.P. Evaluation of rainfall networks using entropy: I. Theoretical development. *Water Resour. Manag.* **1992**, *6*, 279–293. [[CrossRef](#)]
22. McGill, W.J. Multivariate information transmission. *Psychometrika* **1954**, *19*, 97–116. [[CrossRef](#)]
23. Watanabe, S. Information Theoretical Analysis of Multivariate Correlation. *IBM J. Res. Dev.* **1960**, *4*, 66–82. [[CrossRef](#)]
24. Alameddine, I.; Karmakar, S.; Qian, S.S.; Paerl, H.W.; Reckhow, K.H. Optimizing an estuarine water quality monitoring program through an entropy-based hierarchical spatiotemporal Bayesian framework. *Water Resour. Res.* **2013**, *49*, 6933–6945. [[CrossRef](#)]
25. Alfonso, L.; Lobbrecht, A.; Price, R. Information theory-based approach for location of monitoring water level gauges in polders. *Water Resour. Res.* **2010**, *46*. [[CrossRef](#)]
26. Alfonso, L.; Lobbrecht, A.; Price, R. Optimization of water level monitoring network in polder systems using information theory. *Water Resour. Res.* **2010**, *46*. [[CrossRef](#)]
27. Alfonso, L.; He, L.; Lobbrecht, A.; Price, R. Information theory applied to evaluate the discharge monitoring network of the Magdalena River. *J. Hydroinform.* **2013**, *15*, 211–228. [[CrossRef](#)]
28. Alfonso, L.; Ridolfi, E.; Gaytan-Aguilar, S.; Napolitano, F.; Russo, F. Ensemble Entropy for Monitoring Network Design. *Entropy* **2014**, *16*, 1365–1375. [[CrossRef](#)]
29. Boroumand, A.; Rajaei, T. Discrete entropy theory for optimal redesigning of salinity monitoring network in San Francisco bay. *Water Sci. Technol. Water Supply* **2017**, *17*, 606–612. [[CrossRef](#)]
30. Brunzell, N.A. A multiscale information theory approach to assess spatial-temporal variability of daily precipitation. *J. Hydrol.* **2010**, *385*, 165–172. [[CrossRef](#)]

31. Fahle, M.; Hohenbrink, T.L.; Dietrich, O.; Lischeid, G. Temporal variability of the optimal monitoring setup assessed using information theory. *Water Resour. Res.* **2015**, *51*, 7723–7743. [[CrossRef](#)]
32. Hosseini, M.; Kerachian, R. A data fusion-based methodology for optimal redesign of groundwater monitoring networks. *J. Hydrol.* **2017**, *552*, 267–282. [[CrossRef](#)]
33. Hosseini, M.; Kerachian, R. A Bayesian maximum entropy-based methodology for optimal spatiotemporal design of groundwater monitoring networks. *Environ. Monit. Assess.* **2017**, *189*, 433. [[CrossRef](#)] [[PubMed](#)]
34. Keum, J.; Coulibaly, P. Sensitivity of Entropy Method to Time Series Length in Hydrometric Network Design. *J. Hydrol. Eng.* **2017**, *22*. [[CrossRef](#)]
35. Keum, J.; Coulibaly, P. Information theory-based decision support system for integrated design of multi-variable hydrometric networks. *Water Resour. Res.* **2017**, *53*, 6239–6259. [[CrossRef](#)]
36. Kornelsen, K.C.; Coulibaly, P. Design of an Optimal Soil Moisture Monitoring Network Using SMOS Retrieved Soil Moisture. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 3950–3959. [[CrossRef](#)]
37. Leach, J.M.; Kornelsen, K.C.; Samuel, J.; Coulibaly, P. Hydrometric network design using streamflow signatures and indicators of hydrologic alteration. *J. Hydrol.* **2015**, *529*, 1350–1359. [[CrossRef](#)]
38. Leach, J.M.; Coulibaly, P.; Guo, Y. Entropy based groundwater monitoring network design considering spatial distribution of annual recharge. *Adv. Water Resour.* **2016**, *96*, 108–119. [[CrossRef](#)]
39. Lee, J.H. Determination of optimal water quality monitoring points in sewer systems using entropy theory. *Entropy* **2013**, *15*, 3419–3434. [[CrossRef](#)]
40. Lee, C.; Paik, K.; Yoo, D.G.; Kim, J.H. Efficient method for optimal placing of water quality monitoring stations for an ungauged basin. *J. Environ. Manag.* **2014**, *132*, 24–31. [[CrossRef](#)] [[PubMed](#)]
41. Li, C.; Singh, V.P.; Mishra, A.K. Entropy theory-based criterion for hydrometric network evaluation and design: Maximum information minimum redundancy. *Water Resour. Res.* **2012**, *48*. [[CrossRef](#)]
42. Mahjouri, N.; Kerachian, R. Revising river water quality monitoring networks using discrete entropy theory: The Jajrood River experience. *Environ. Monit. Assess.* **2011**, *175*, 291–302. [[CrossRef](#)] [[PubMed](#)]
43. Mahmoudi-Meimand, H.; Nazif, S.; Abbaspour, R.A.; Sabokbar, H.F. An algorithm for optimisation of a rain gauge network based on geostatistics and entropy concepts using GIS. *J. Spat. Sci.* **2016**, *61*, 233–252. [[CrossRef](#)]
44. Masoumi, F.; Kerachian, R. Optimal redesign of groundwater quality monitoring networks: A case study. *Environ. Monit. Assess.* **2010**, *161*, 247–257. [[CrossRef](#)] [[PubMed](#)]
45. Memarzadeh, M.; Mahjouri, N.; Kerachian, R. Evaluating sampling locations in river water quality monitoring networks: Application of dynamic factor analysis and discrete entropy theory. *Environ. Earth Sci.* **2013**, *70*, 2577–2585. [[CrossRef](#)]
46. Mishra, A.K.; Coulibaly, P. Hydrometric Network Evaluation for Canadian Watersheds. *J. Hydrol.* **2010**, *380*, 420–437. [[CrossRef](#)]
47. Mishra, A.K.; Coulibaly, P. Variability in Canadian Seasonal Streamflow Information and Its Implication for Hydrometric Network Design. *J. Hydrol. Eng.* **2014**, *19*. [[CrossRef](#)]
48. Mondal, N.C.; Singh, V.P. Evaluation of groundwater monitoring network of Kodaganar River basin from Southern India using entropy. *Environ. Earth Sci.* **2011**, *66*, 1183–1193. [[CrossRef](#)]
49. Samuel, J.; Coulibaly, P.; Kollat, J.B. CRDEMO: Combined Regionalization and Dual Entropy-Multiobjective Optimization for Hydrometric Network Design. *Water Resour. Res.* **2013**, *49*, 8070–8089. [[CrossRef](#)]
50. Santos, J.F.; Portela, M.M.; Pulido-Calvo, I. Dimensionality reduction in drought modelling. *Hydrol. Process.* **2013**, *27*, 1399–1410. [[CrossRef](#)]
51. Stosic, T.; Stosic, B.; Singh, V.P. Optimizing streamflow monitoring networks using joint permutation entropy. *J. Hydrol.* **2017**, *552*, 306–312. [[CrossRef](#)]
52. Su, H.T.; You, G.J.Y. Developing an entropy-based model of spatial information estimation and its application in the design of precipitation gauge networks. *J. Hydrol.* **2014**, *519*, 3316–3327. [[CrossRef](#)]
53. Uddameri, V.; Andruss, T. A GIS-based multi-criteria decision-making approach for establishing a regional-scale groundwater monitoring. *Environ. Earth Sci.* **2014**, *71*, 2617–2628. [[CrossRef](#)]
54. Wei, C.; Yeh, H.C.; Chen, Y.C. Spatiotemporal scaling effect on rainfall network design using entropy. *Entropy* **2014**, *16*, 4626–4647. [[CrossRef](#)]
55. Werstuck, C.; Coulibaly, P. Hydrometric network design using dual entropy multi-objective optimization in the Ottawa River Basin. *Hydrol. Res.* **2016**, *48*, 1–13. [[CrossRef](#)]

56. Werstuck, C.; Coulibaly, P. Assessing Spatial Scale Effects on Hydrometric Network Design Using Entropy and Multi-Objective Methods. *JAWRA J. Am. Water Resour. Assoc.* **2017**, in press.
57. Xu, H.; Xu, C.-Y.; Sæthun, N.R.; Xu, Y.; Zhou, B.; Chen, H. Entropy theory based multi-criteria resampling of rain gauge networks for hydrological modelling—A case study of humid area in southern China. *J. Hydrol.* **2015**, *525*, 138–151. [[CrossRef](#)]
58. Yakirevich, A.; Pachepsky, Y.A.; Gish, T.J.; Guber, A.K.; Kuznetsov, M.Y.; Cady, R.E.; Nicholson, T.J. Augmentation of groundwater monitoring networks using information theory and ensemble modeling with pedotransfer functions. *J. Hydrol.* **2013**, *501*, 13–24. [[CrossRef](#)]
59. Mishra, A.K.; Özger, M.; Singh, V.P. An entropy-based investigation into the variability of precipitation. *J. Hydrol.* **2009**, *370*, 139–154. [[CrossRef](#)]
60. Gong, W.; Yang, D.; Gupta, H.V.; Nearing, G. Estimating information entropy for hydrological data: One dimensional case. *Water Resour. Res.* **2014**, *50*, 5003–5018. [[CrossRef](#)]
61. Silva, V.; da Silva, V.d.P.R.; Belo Filho, A.F.; Singh, V.P.; Almeida, R.S.R.; da Silva, B.B.; de Sousa, I.F.; de Holanda, R.M. Entropy theory for analysing water resources in northeastern region of Brazil. *Hydrol. Sci. J.* **2017**, *62*, 1029–1038.
62. Ridolfi, E.; Montesarchio, V.; Russo, F.; Napolitano, F. An entropy approach for evaluating the maximum information content achievable by an urban rainfall network. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 2075–2083. [[CrossRef](#)]
63. Yeh, H.C.; Chen, Y.C.; Wei, C.; Chen, R.H. Entropy and kriging approach to rainfall network design. *Paddy Water Environ.* **2011**, *9*, 343–355. [[CrossRef](#)]
64. Awadallah, A.G. Selecting optimum locations of rainfall stations using kriging and entropy. *Int. J. Civ. Environ. Eng* **2012**, *12*, 36–41.
65. Ahmed, N.A.; Gokhale, D.V. Entropy expressions and their estimators for multivariate distributions. *IEEE Trans. Inf. Theory* **1989**, *35*, 688–692. [[CrossRef](#)]
66. Coulibaly, P.; Keum, J. *Snow Network Design and Evaluation for La Grande River Basin*; Hydro-Quebec: Hamilton, ON, Canada, 2016.
67. Samuel, J.; Coulibaly, P. *Design of Optimum Snow Monitoring Networks Using Dual Entropy Multiobjective Optimization (DEMO) with Remote Sensing Data: Case Study for Columbia River Basin*; BC-Hydro: Hamilton, ON, Canada, 2013.
68. Zhao, R.J. The Xinanjiang model applied in China. *J. Hydrol.* **1992**, *135*, 371–381.
69. Arnold, J.G.; Srinivasan, R.; Muttiah, R.S.; Williams, J.R. Large Area Hydrologic Modeling and Assessment Part I: Model Development. *J. Am. Water Resour. Assoc.* **1998**, *34*, 73–89. [[CrossRef](#)]
70. United Nations. *Guidelines for Reducing Flood Losses*; Pilon, P.J., Ed.; United Nations: Geneva, Switzerland, 2004.
71. Chapman, T.G. Entropy as a measure of hydrologic data uncertainty and model performance. *J. Hydrol.* **1986**, *85*, 111–126. [[CrossRef](#)]
72. Yadav, M.; Wagener, T.; Gupta, H. Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Adv. Water Resour.* **2007**, *30*, 1756–1774. [[CrossRef](#)]
73. Sawicz, K.; Wagener, T.; Sivapalan, M.; Troch, P.A.; Carrillo, G. Catchment classification: Empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 2895–2911. [[CrossRef](#)]
74. Richter, B.D.; Baumgartner, J.V.; Powell, J.; Braun, D.P. A Method for Assessing Hydrologic Alteration within Ecosystems. *Conserv. Biol.* **1996**, *10*, 1163–1174. [[CrossRef](#)]
75. Monk, W.A.; Peters, D.L.; Allen Curry, R.; Baird, D.J. Quantifying trends in indicator hydroecological variables for regime-based groups of Canadian rivers. *Hydrol. Process.* **2011**, *25*, 3086–3100. [[CrossRef](#)]
76. Kerr, Y.H.; Waldteufel, P.; Wigneron, J.P.; Delwart, S.; Cabot, F.; Boutin, J.; Escorihuela, M.J.; Font, J.; Reul, N.; Gruhier, C.; et al. The SMOS Mission: New Tool for Monitoring Key Elements of the Global Water Cycle. *Proc. IEEE* **2010**, *98*, 666–687. [[CrossRef](#)]
77. Harbaugh, A.W. *MODFLOW-2005, The U.S. Geological Survey Modular Ground-Water Model—The Ground-Water Flow Process*; U.S. Geological Survey: Reston, VA, USA, 2005.
78. Pollock, D.W. *User Guide for MODPATH Version 7—A Particle-Tracking Model for MODFLOW*; U.S. Geological Survey: Reston, VA, USA, 2016.

79. Bedekar, V.; Morway, E.D.; Langevin, C.D.; Tonkin, M.J. *MT3D-USGS Version 1: A U.S. Geological Survey Release of MT3DMS Updated with New and Expanded Transport Capabilities for Use with MODFLOW*; U.S. Geological Survey: Reston, VA, USA, 2016.
80. Šimůnek, J.; van Genuchten, M.T.; Šejna, M. *The HYDRUS Software Package for Simulating Two- and Three-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Porous Media*; PC Progress: Prague, Czech Republic, 2012.
81. Owlia, R.R.; Abrishamchi, A.; Tajrishy, M. Spatial-temporal assessment and redesign of groundwater quality monitoring network: A case study. *Environ. Monit. Assess.* **2011**, *172*, 263–273. [[CrossRef](#)] [[PubMed](#)]
82. Banik, B.K.; Alfonso, L.; di Cristo, C.; Leopardi, A.; Mynett, A. Evaluation of Different Formulations to Optimally Locate Sensors in Sewer Systems. *J. Water Resour. Plan. Manag.* **2017**, *143*. [[CrossRef](#)]
83. Ruddell, B.L.; Kumar, P. Ecohydrologic process networks: 1. Identification. *Water Resour. Res.* **2009**, *45*. [[CrossRef](#)]
84. Ruddell, B.L.; Kumar, P. Ecohydrologic process networks: 2. Analysis and characterization. *Water Resour. Res.* **2009**, *45*, 1–14. [[CrossRef](#)]
85. Kang, M.; Ruddell, B.L.; Cho, C.; Chun, J.; Kim, J. Agricultural and Forest Meteorology Identifying CO₂ advection on a hill slope using information flow. *Agric. For. Meteorol.* **2017**, *232*, 265–278. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).