B. K. Banik[1], A. Saiara[2], and S. F. Tabassum[3]

[1]Professor, Department of Civil and Environmental Engineering, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh. (ORCid: 0000-0002-3966-7430) Email: bijit-cee@sust.edu

[2]Department of Civil and Environmental Engineering, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh. Email: saiaraafiya@gmail.com

[3]Department of Civil and Environmental Engineering, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh. Email: sarah.sustcee@gmail.com

Corresponding author: Bijit Kumar Banik (bijit-cee@sust.edu)

Author's contribution:

B. K. Banik: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing

A. Saiara: Data curation, Formal Analysis, Investigation, Software, Validation, Visualization, Writing – original draft (Partly)

F. Tabassum: Data curation, Formal Analysis, Investigation, Visualization, Writing – original draft (Partly)

Key Points:

- The greedy algorithm is found very effective in designing the discharge monitoring network.
- Monitors have been placed where there is a connection of hydrological features like tributaries and wetlands with the River.
- Four monitoring stations are not enough for the Surma River.

Abstract

In developing countries like Bangladesh, river discharge monitoring networks are designed unseemly, operated poorly, and often fail to reach their purposes resulting in the unavailability of sufficient data to describe the behavior of such systems. In these cases, water-related decisions may create problems for the environment, the regional economy, and society. This paper has investigated the application of Shannon's Information Theory to design and evaluate an efficient discharge monitoring network for the Surma River. A 1-D model has been formulated to extract all discharge data at different points of Surma River using MIKE 11. The appropriate monitoring station locations were determined by optimizing two conflicting objective functions (joint entropy and total correlation) using the Non-dominated Sorting Genetic Algorithm-II and Greedy algorithm. The study demonstrates that an informative yet less redundant monitoring network configuration can be found through the greedy algorithm.

Plain Language Summary

In developing countries like Bangladesh, river discharge monitoring networks are designed unseemly, operated poorly, and often fail to reach their purposes, resulting in the unavailability of sufficient data to describe the behavior of such systems. In these cases, water-related decisions (such as distributing water resources among different stakeholders and flood forecasting) may negatively impact the environment, the regional economy, and society. This paper has investigated the application of Shannon's Information Theory to design and evaluate an efficient discharge monitoring network for the Surma River. The Surma River has only four monitoring stations in place. The study demonstrates that the number and positions of those monitoring stations are not optimal, and a more informative yet less redundant network is possible.

1 Introduction

The hydrologic cycle plays a pivotal role in managing water resources and controlling the climate (Fekete et al., 2012). Among different components of the hydrologic cycle, river discharge plays a significant role. In various parts of the world, river systems are the primary source of fresh water supply. River discharge is also the most accurately measured component of the water cycle where appropriate monitoring stations are in place (Grabs et al., 1996; Gutowski et al., 1997; Hagemann & Dümenil, 1997). A well-managed River discharge monitoring network is a prerequisite in allocating the water resources among different stakeholders, flood forecasting, hydraulic structure design, estimation of reservoir capacity, computing river water balance. On the other hand, a poorly managed monitoring system or insufficient information on river discharge and water-related decisions may negatively impact the environment, the regional economy, and society (Mason et al., 2003).

Bangladesh is a riverine country with 257 rivers, including 59 transboundary rivers (Islam, 2016). There are 344 water level monitoring stations in 127 rivers, while the rest are unmonitored (BWDB, 2020). Small rivers are very susceptible to an intensive precipitation event as they respond to such events much quicker than large rivers. Flash flood, a common phenomenon in Bangladesh, especially in the northeastern part, can cause severe flooding within some hours (BWDB, 2020). In Bangladesh, like other developing countries, monitoring networks are designed unseemly, operated poorly, and often fail to reach their purposes, resulting in the unavailability of sufficient data to describe the behavior of such systems (Alfonso et al., 2013). Besides, due to financial constraints and shifting monitoring priority, there is a continuous decline in water monitoring (Pilon et al., 1996; Mishra & Coulibaly, 2009). Therefore, finding the optimal number and their position is critical in designing a monitoring network.

There are some general recommendations for the placement of the monitoring network (Rodda, 2011), such as point of a regular and stable river bed, where parallel velocities remain uniform throughout a cross-section and avoid bent reaches, strong backwater effects, flow bifurcations, and aquatic growth (Alfonso et al., 2013). Besides, some methods in the literature exist that deal with the design and evaluation of monitoring river discharge data. Statistical method (Moss & Karlinger, 1974; Moss & Tasker, 1991), entropy-based method (Husain, 1989; Yang & Burn, 1994), and methods that include direct surveys to assess users' needs are the most common (Davar & Brimley, 2007). Furthermore, one paper regarding monitoring network design was derived from water quality data (Telci et al., 2009). A comprehensive review of available methods for evaluating monitoring networks is presented by Mishra and Coulibaly (2009). Mishra and Coulibaly (2009) conclude that among those three basic types of methodology, entropy-based methods have got the most attention from researchers and are among the most promising approaches for network design. Entropy-based methods are the byproducts of the Information theory, which was first introduced by Shannon (1948) to measure the information content in a dataset and applied to diverse research areas, including solving water resource problems. A second comprehensive review by Keum et al. (2017) mainly focuses on applying the entropy-based method in designing water monitoring network design. They categorize four areas where the application of entropy concept was used to design the water monitoring network. Those are (1) precipitation (Ridolfi et al., 2011; Yeh et al., 2011; Awadallah, 2012; Mahmoudi-Meimand et al., 2016); (2) streamflow and water level (Alfonso et al., 2010a; Alfonso et al., 2010b; Alfonso et al., 2013; Alfonso et al., 2014; Stosic et al., 2017; Werstuck & Coulibaly, 2017); (3) water quality (Lee et al., 2014; Banik et al., 2015; Banik et al., 2017a; Banik et al., 2017b; Boroumand & Rajaee, 2017); and (4) soil moisture and groundwater networks (Uddameri & Andruss, 2014; Kornelsen & Coulibaly, 2015; Leach et al., 2016; Hosseini & Kerachian, 2017;). Though few studies have been found in designing the discharge/water level monitoring network in the River (Mahjouri & Kerachian, 2011; Alfonso et al., 2013; Lee et al., 2014; Mishra & Coulibaly, 2014; Stosic et al., 2017; Mokin et al., 2018), there is no study found in the literature in the case of the Bangladeshi rivers.

This paper focuses on a case study of designing and evaluating the discharge monitoring station of the Surma river using the entropy-based method, which was first introduced by Amorocho & Espildora (1973) in the water resource field. In the first phage, a 1-D model has been developed for the Suram river to extract the time series of discharge data. Afterward, two entropy contents (Joint entropy and total correlation) have been used to design and evaluate the optimal placement of the monitoring stations in the Surma river. Nondominated sorting genetic algorithm II (NSGA-II by Dev et al., 2002) and Greedy algorithm (Alfonso et al., 2013; Banik et al., 2017a) have been used to optimize the monitoring network.

2 Materials and Methods

The River is a natural stream whose behavior can be significantly affected by the presence of tributaries. Especially in the case of discharge, their effect can be significant. The optimal monitoring network would be monitors that provide the maximum information content and capture independent information. There are two objectives. The first objective, providing maximum information content at each gauging site, can be achieved by maximizing the joint entropy of selected gauges. To fulfill the second objective of minimizing dependency or redundancy, the concept of Total Correlation has been used. The amount of information among two variables $X_1,\ X_2$ is known as the Joint Entropy and can be expressed as:

$H\left(X_1, X_2\right) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p\left(x_{1i}, x_{2j}\ \right)\ \log p\left(x_{1i}, x_{2j}\right)$ (1)

where $p\left(x_{1i}, x_{2j}\right)$ is the joint distribution between variables X_1 and X_2, n and m are the number of elementary events in X_1 and X_2 respectively.

Total Correlations among N variables can be expressed as:

$C\left(X_1, X_2, \ldots.., X_N\right) = \sum_{i=1}^{N} H\left(X_i\right) - H(X_1, X_2, \ldots.., X_N)$ (2)

In this case, the optimization problem can be expressed in following mathematical formulation.

min $\{C(X_1,\ X_2,\ \ldots\ldots\ldots.,\ X_N)\}$ (3)

max $\{H(X_1,\ X_2,\ \ldots\ldots\ldots.,\ X_N)\}$

Here the decision variables represent the geographical location of N gauges. Two approaches have been used to solve this optimization problem. One is the multi-objective optimization using NSGA-II, while the other is the greedy algorithm. More detailed information on this theory can be found elsewhere (Alfonso et al., 2013 and Banik et al., 2015).

The methodology consists of three parts.

Part 1: Generating time series of discharge data.

Part 2: Quantization and evaluation of entropy pattern.

Part 3: Optimization process.

2.1 Generating time series of discharge data

In order to prepare time-series data that can be used for information theory analysis, a 1D-hydrodynamic model was developed using MIKE11. The model includes 150 km of Surma River with points placed approximately every 500 m. The hydrological data from July 2016 to June 2017 was used with the complete data records at the tributaries and hydrologic stations. Discharge data obtained from the Sylhet station was used to calibrate the model. Once the model was calibrated, discharge data was extracted from the result file. The model's description and its calibration are described in the paper's third section.

2.2 Quantization and evaluation of entropy pattern

Although there are some nonparametric methods to estimate mutual information (see, e.g., Moon et al. 1995), a histogram-based frequency analysis is used in this paper. For this purpose, the discharge data matrix must be quantized first. Quantization is a procedure of constraining a continuous set of values to a discrete set. The quantization process rounds a value x to its nearest lowest integer multiple of a, namely x_q to give:

$x_q = a[\frac{2x+a}{2a}]$ (4)

The calculation of joint entropy and total correlation is related to probability calculation, which requires the information to be integer values. So, the quantized data matrix of this case will have the following properties:

The minimum value of the data matrix will be zero, and there will be no negative values. It will be quantized by another factor to round the values to their nearest integer value.

The entropy pattern is different for different sections of the River, and it is calculated from the quantized discharge data with the help of MATLAB. It is further used to develop an entropy map that shows high and low entropy zones. The possible location of the monitors will be at the places where a change in entropy takes place.

2.3 Optimization process

- Multi-objective optimization

The problem can be solved by posing it as a multi-objective optimization problem (MOOP). This problem provides a set of quasi-optimal, non-dominated solutions that draw a Pareto front. Generally, objective functions conflict with each other, and therefore a solution that satisfies all the objectives at a time may not exist. The MOOP searches for a set of decision variables that simultaneously satisfy constraints and optimize objective function values. MOOP has been successfully used to solve water-related optimization problems (e.g., Alfonso et al., 2010a; Barreto et al., 2009; Preis & Ostfeld, 2010). This paper uses NSGA-II, an elitist non-dominated sorting genetic algorithm for multi-objective optimization (Deb et al., 2002). It utilizes Simulated Binary Crossover (SBN) and Polynomial Mutation as genetic-related operations. Two objectives were chosen where joint entropy was maximized, and total correlation was minimized.

- Rank-based greedy algorithm

The greedy algorithm is a single objective optimization procedure. Alfonso et al. (2013) have used two single objectives (Joint entropy and total correlation) separately while optimizing the discharge monitoring network of Magdalena River. In this paper, we also considered both objective functions. Moreover, as the natural process is complex, a single objective optimization might not fulfill the expected result. To benefit from multi-objective optimization in a greedy framework, we introduced another objective function in terms of fitness score (shown in equation 5) in the range of 0-1 by combining two objectives, the joint entropy, and total correlation.

$\text{fitness} = \min\left\{\left[\left(1\text{-}\frac{\mathrm{C_{max}\text{-}C}}{\mathrm{C_{max}\text{-}C_{min}}}\right) + \left(1\text{-}\frac{\mathrm{H\text{-}H_{min}}}{\mathrm{H_{max}\text{-}H_{min}}}\right)\right] \div 2\right\}$ (5)

C_{max} and C_{min} are the system's maximum and minimum total correlation, respectively, while H_{max} and H_{min} are the system's maximum and minimum joint entropy, respectively.

The first two single objective algorithms are shown in figure 1a), while the other that combines both algorithms into a fitness score is shown in figure 1b). The most informative sensor will be chosen as the first one in all three algorithms, while the remaining sensors will be selected either with the specific objective function (equation 3) or with the fitness function (equation 5).

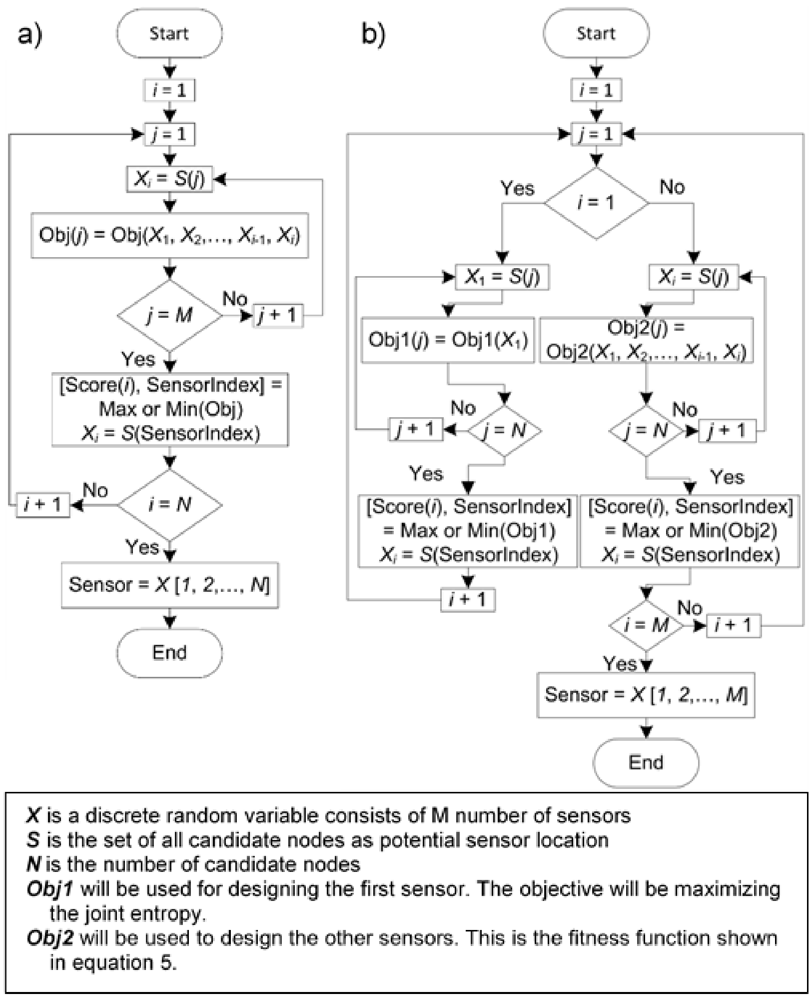
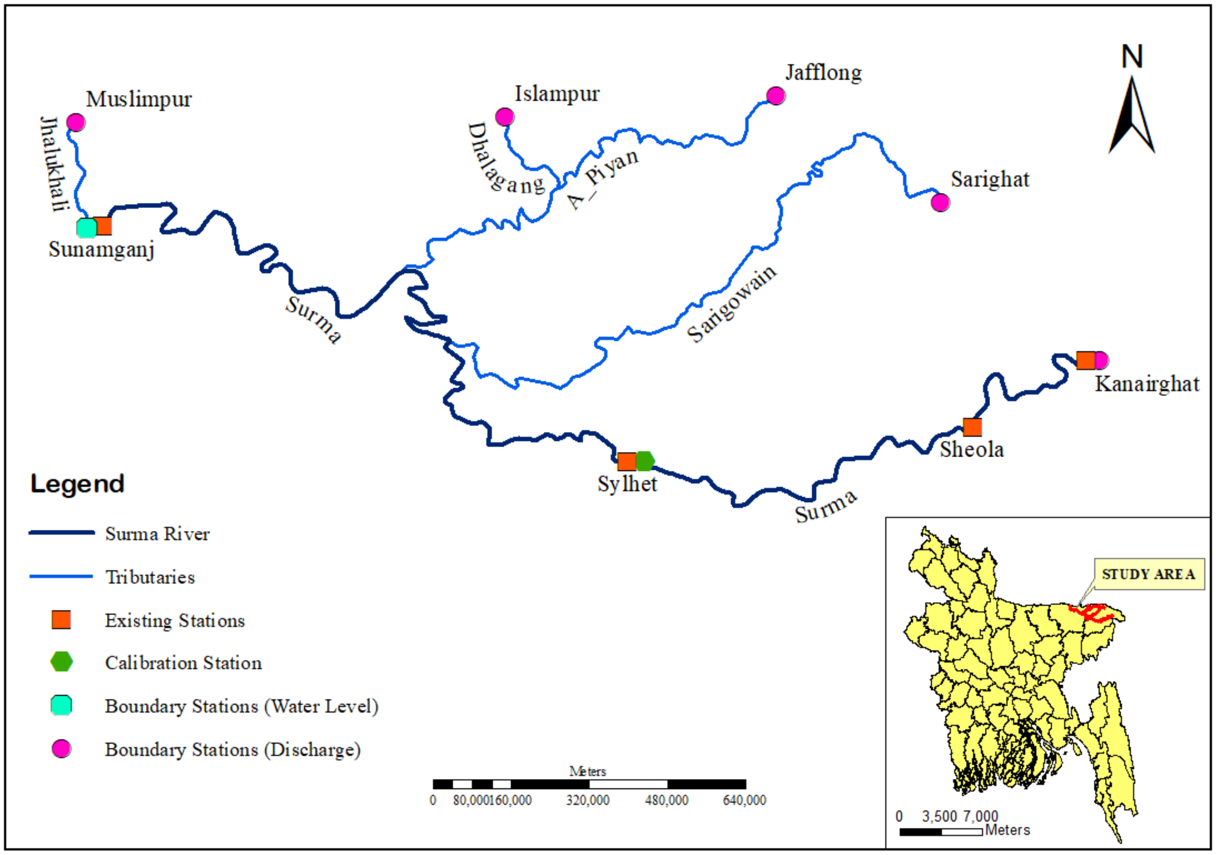


Fig.1: Flowchart of the greedy algorithm. (a) To solve single objective function (b) To solve two objectives as fitness function

3 Study area, model development, and calibration

3.1 Study area

Bangladesh is located in South Asia, slightly above sea level in the Bay of Bengal, which is formed by the confluence of three of Asia's greatest rivers, the Ganges, Brahmaputra, and Meghna Rivers, forming the world's largest water basin and delta (Hossain et al., 2021). This riverine country is divided into three major river systems- Ganges-Padma, the Brahmaputra-Jamuna, and Meghna-Surma. The Surma River is one of the most important rivers of the Meghna-Surma system, and the system originates in the Indian hills of Shillong and Meghalaya. The primary source is the Barak River which has a large catchment area in eastern Assam's ridge and valley topography bordering Myanmar. When it reaches the Bangladeshi border at Amalshid in Sylhet district, it splits into the Surma and Kushiyara rivers. The Surma River, which flows north of the Sylhet basin, receives Right Bank tributaries from Shillong's Khasia and Jaintia Hills. These steep, flashy rivers originate in one of the world's wettest regions. From the Tripura Hills, the Kushiyara receives left bank tributaries, the most important of which is the Manu. There are numerous internal drainage depressions (haors), meandering flood channels, and abandoned river courses between the Surma and Kushiyara, frequently inundated during the monsoon season. The two rivers merged at Markuli and became the Meghna after passing through Bhairab. The Padma and the Meghna meet in Chandpur, where they flow to the sea under tidal impact.



Fi

2: Study area showing different boundaries and calibration point

The study area (Figure 2) of this research has been chosen from the start of the Surma River at Kanaighat, Sylhet to Sunamganj, which is 150 km long. The model included tributaries which are the rivers Dhalagang, Jhalukhali, Piyan (active) and Sarigowain. Among these, Dhalagang does not discharge into the Surma River. Instead, it is a tributary of the Piyan (active) River, so the data from this tributary has only been used to build the model. These tributaries were included as point sources of discharge. The upstream boundary condition for the Surma river is the discharge series at Kanaighat, and the downstream boundary is the water level series at Sunamganj. Discharges were obtained through rating curves at Sunamganj, Muslimpur, Islampur, Jaflong, Sarighat, and Kanaighat. After providing five files necessary for the model; a river network file, a cross-section file, a boundary file, and a simulation file, the model was run with a time step of 1 minute. Figure 2 also shows four existing monitoring stations within the 150 km course of the Surma river.

3.2 Model development

A 1D-hydrodynamic model is necessary to design and evaluate discharge monitoring networks. Therefore, the upper part of the Surma River (Kanaighat to Sunamganj) has been modeled on the MIKE11. Rainfall events have not been included in this model. The analysis of the River under the response of rainfall events is beyond the objectives of this study. However, this developed model can be updated and complemented for other uses.

The information of water levels and discharges of river stations, river network, and bathymetry is obtained from the Institute of Water Modelling (IWM), Bangladesh, to develop the 1D- hydrodynamic model of the Surma River between Kanaighat and Sunamganj. The data used are described in detail as follows.

- Water level and discharge

Daily and multiannual water level and discharge data for the upper Surma River and its tributaries range from 1978 to 2017. The model was built from July 2016 to June 2017.

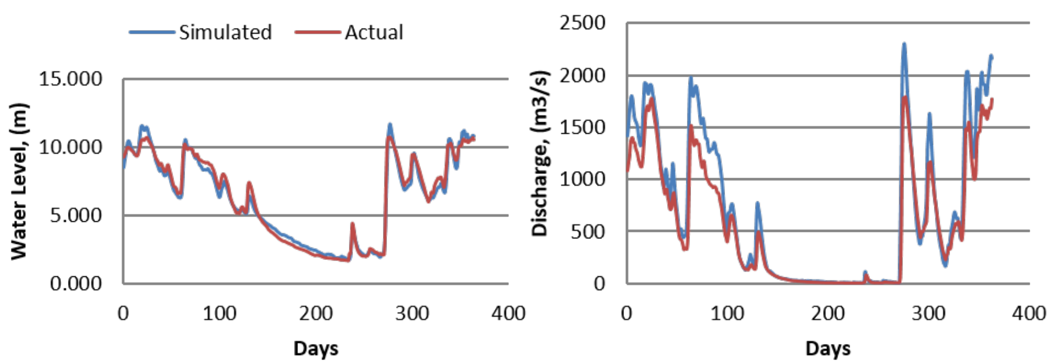
- Boundary condition

The boundaries of the model are the 2016-2017 discharge of Kanaighat (upstream) and the 2016-2017 water level of Sunamganj (downstream). The time series of the four tributaries (Muslimpur, Islampur, Jaflong, and Sarighat) were included as point sources.

- River network, bathymetry, and cross-sections

The model of the Surma river has been developed using a network point every 500 meters. Thus, the Surma river contains 301 points with chainage ranging from 0 km at Kanaighat (upstream) to 150 km at Sunamganj (downstream). The bathymetry and cross-section information of all those 301 points were collected from the Department of Surveying of IWM, Bangladesh.

3.3 Model calibration



In hydrologic modeling research, proper model calibration reduces uncertainty in model simulations (Engel et al., 2007). Calibration is traditionally done manually.

Fig. 3: Comparison of actual and simulated data of water level and discharge at Sylhet station

It entails modifying model input parameter values to generate simulated results within a given range of measured data. The discharge data at Sylhet station was used to calibrate the model. After calibrating the model, a comparison of actual and simulated data at Sylhet station was performed to ensure that the River's flows and water levels remained consistent (Figure 3). Moreover, the performance of the calibrated model was evaluated through three parameters, namely, RMSE-observations standard deviation ratio (RSR), Nash-Sutcliffe efficiency (NSE), and Percent bias (PBIAS). The standard values and the obtained values for the calibrated model are shown in table 1. From the table, it is clear that the performance of the calibrated model is excellent.

Table 1: Standard performance rating of three parameters with obtained value

Performance Rating	RSR	NSE	PBIAS (%)	Reference
Very Good	$0 \sim 0.5$	$0.75 \sim 1$	± 10	Moriasi et al. (2007)
Good	$0.5 \sim 0.6$	$0.65 \sim 0.75$	$\pm 10 \sim \pm 15$	
Satisfactory	$0.6 \sim 0.7$	$0.5 \sim 0.65$	$\pm 15 \sim \pm 25$	
Unsatisfactory	> 0.7	< 0.5	± 25	
Model	0.0036	0.9964	0.741	

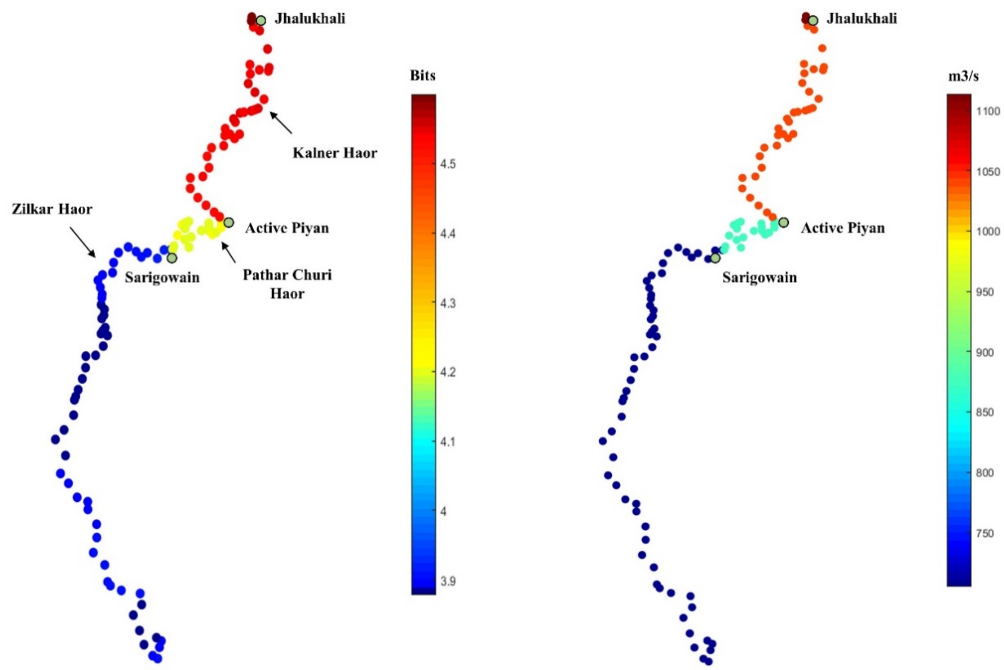
4 Results

In this section, the entropy map of the Surma river has been produced and analyzed using the data created after calibrating the 1D model. Afterward, the results obtained from two optimization processes have been discussed and compared. Finally, a sensitivity analysis on the quantization parameter has been performed to evaluate its impact on the optimization process.

4.1 Analysis of the entropy map of Surma River

The entropy map obtained from the discharge time series data and the map of mean discharge of Surma river are shown in figure 4, where the Jhalukhali tributary is staying downstream. The first thing to be noticed is that the entropy increases at points where tributaries discharge into the River. This finding aligns with Alfonso et al. (2013), where the authors found the same phenomena in the case of the Magdalena River in Columbia. The tributaries Sarigowain, Piyan (active), and Jhalukhali show significant increments in entropy due to their strong influence in terms of discharge. Most of the upstream of the Surma River shows the same entropy with very little change because no tributaries flow into it. Intriguingly, the lowest entropy value occurs in this part of the River. The lowest entropy is because the discharge range is minimal in this part, and upon applying the equation for quantization with a = 100 m^3/s the resulting quantized series has only a few unique values in the frequency analysis the variation in entropy is very little here.

Secondly, small decreases of entropy can also be seen as the effect of some haors, which are Zilkar Haor, Pathar Churi Haor, and Kalner Haor. Haors are the wetlands located in the northeastern part of Bangladesh, and they receive surface runoff water from rivers during the monsoon. However, the change of entropy that occurs here because of the contribution of



Fi

4: (a) Entropy map in bits for a = 100 m^3/s and (b) mean discharge map of 2016-17 in m^3/s for Surma River

discharge from the River to the haors. It is not that significant because it only happens during the rainy season.

Thirdly, some changes in entropy can be observed at the central part, beginning of the upstream, and in a tiny downstream region of the River. Sylhet City lies beside the center portion of the River, and many small natural channels discharge the City's runoff into the River, which is the cause of change in entropy in the central part of the River. The most likely reason for changes near upstream and downstream is the existence of small haors in these regions. Haor is a complex system of reservoirs that absorb the peak flows of the River, thus lowering the discharge range and diminishing the entropy.

The entropy map gives a complete overview of changes in the entropy of the

Surma River. Entropy is continuously increasing from upstream to downstream. A significant change in entropy occurred when the tributary Sarigowain met Surma River. Before this point, the entropy was almost constant because no significant flow was added or subtracted from the River. After this point, the entropy increases because of other tributaries' influence. Some decreases in the entropy happen due to the presence of wetlands. However, they are tiny compared to the entropy increases due to the inflows from the tributaries. Finally, at the point where the Jhalukhali River discharges into the Surma River, the entropy is the maximum.

4.2 Multi-objective optimization approach

The multi-objective optimization approach is solved using NSGA-II (Deb et al., 2002). For this approach, the evolutionary parameters must be defined, namely, the number of populations and generations. The number of decision variables (number of monitors to be placed along the River) must also be specified. A sensitivity analysis was performed on both evolutionary parameters, and it was found that 100 population and 100 generations were the optimal numbers. A series of simulations were run with the populations and generations (P, G): (100,100) and the number of monitors 2-13 to get the optimal number of monitors for the Surma River network. Joint entropy was maximized to find out the optimal number. The optimization result can be found in figure 5. It is evident from the figure that after eight monitors, the joint entropy does not significantly increase. Therefore, eight monitors will be the optimal number for the Surma River. Further experiments were carried out to find the optimal position of the monitors in which four different populations and generations were used with the following combinations (P, G): (100,50), (100,100), (200,50), (200,100) for monitors from 7 to 10.

[CHART]

Fig. 5: Optimal number of monitors for Surma River

Figure 6 shows the optimization result for 7-10 monitors. The Pareto front is shown with black filled circle in all four figures. From those four figures, it is evident that the increment of very little joint entropy will cost a substantial total correlation after eight monitors. In other words, the monitors are giving shared information, making those additional monitors redundant. As multi-objective optimization produces a set of near-optimal solutions, choosing a particular solution requires further criteria or judgment. For instance, the amount of information gathered by the set of monitors might be more critical than minimizing the redundant information and vice versa. At the same time, some might search for a compromise solution between two objectives. For this reason, three different solutions (A, B & C) have been chosen to show their position geographically. Solution A and C are extreme-end solutions (minimum total correlation and maximum joint entropy, respectively), while the other is a compromised solution considering the minimum distance from the origin.

The sensor placements of solutions A, B, and C, as mentioned in figure 7, are

NSGA-A, NSGA-B, and NSGA-C, respectively, for ten monitor cases. NSGA-A is the extreme solution where the minimum value of total correlation is considered, and NSGA-C is the extreme solution in terms

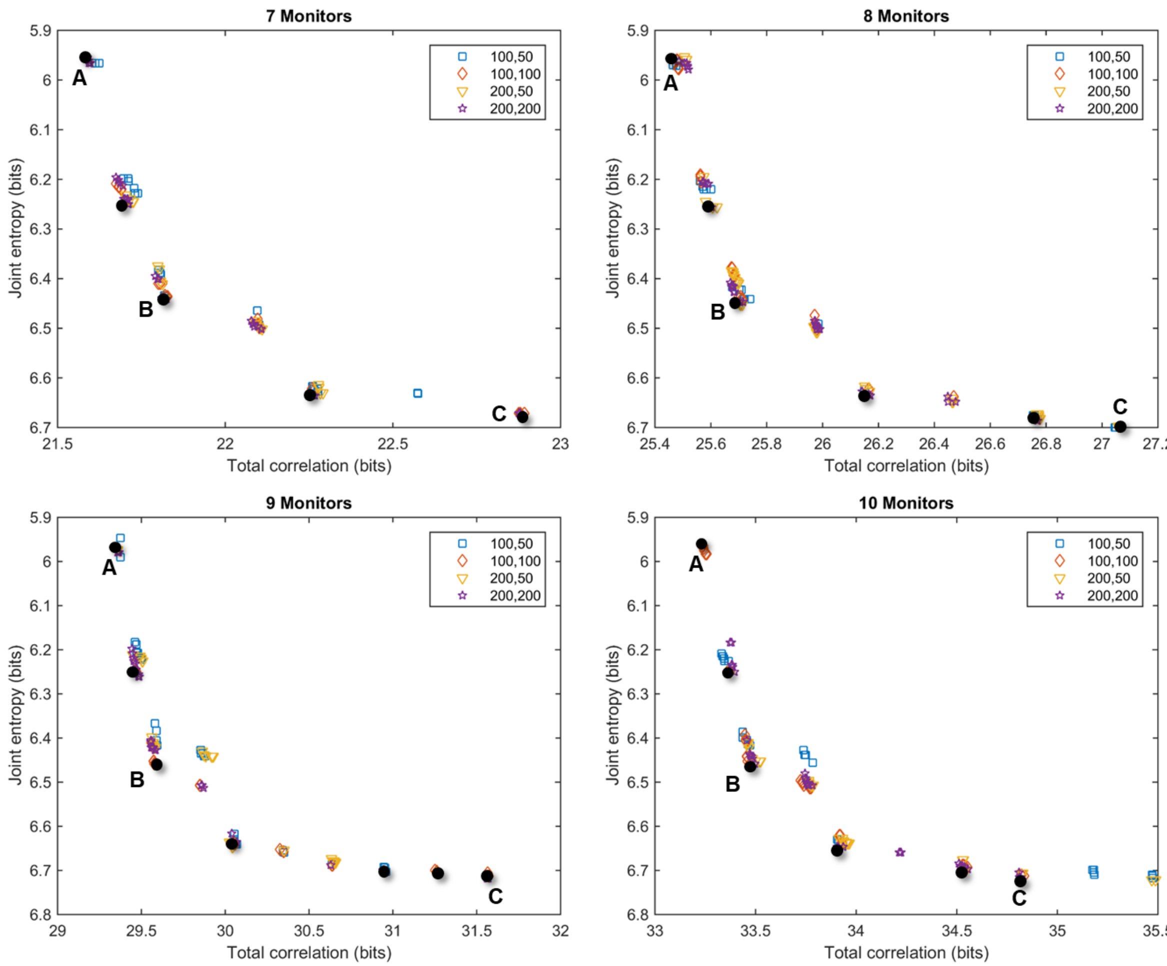
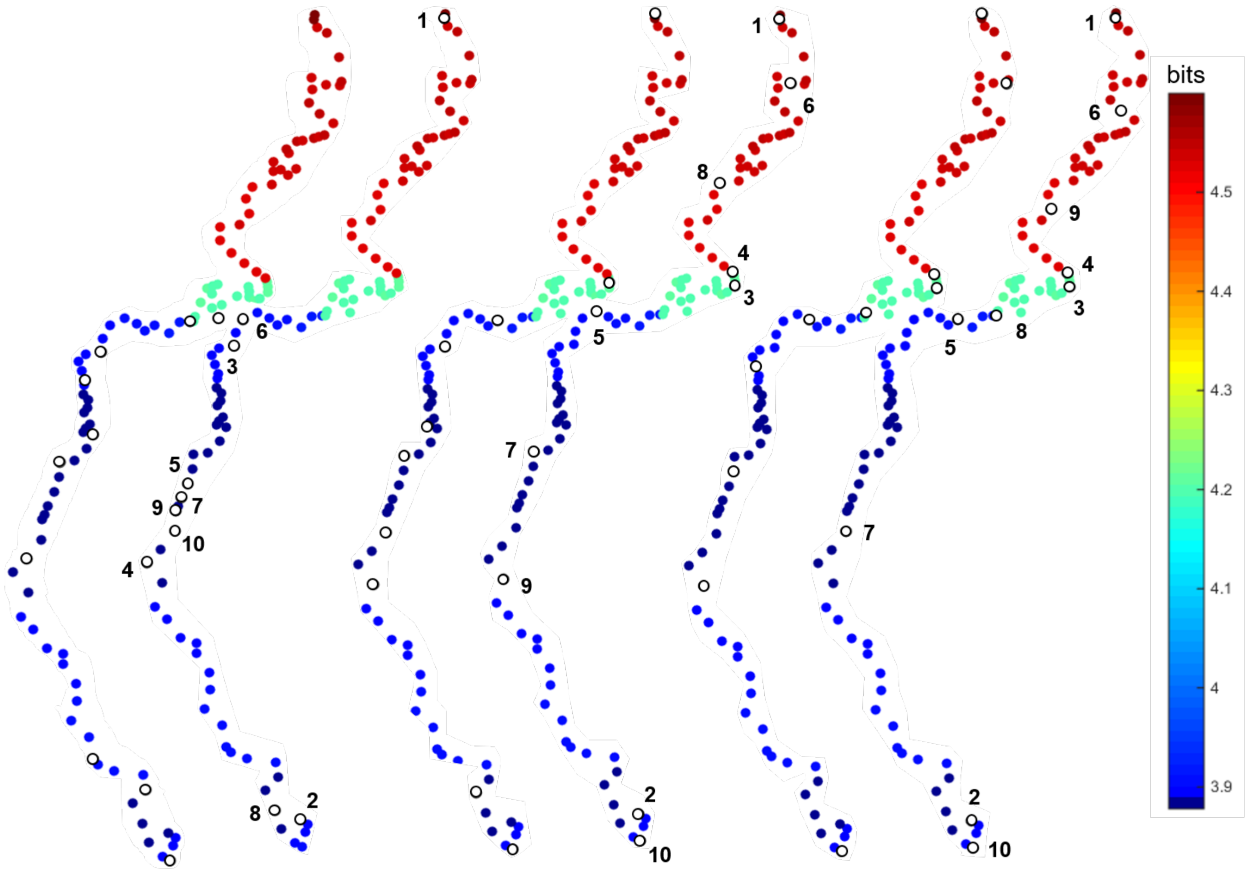


Fig. 6: Pareto front obtained from the NSGA-II optimization showing extreme and compromised solutions for 7-10 monitors. The legend shows the number of population and generation used during optimizations.

of maximum joint entropy. The other solution NSGA-B is the compromise between two objectives. One thing is clear from the figure that the more we go from solution NSGA-A to NSGA-C, the more the monitors are placed downstream. The solution NSGA-C maximizes the joint entropy, and more entropy can be found in the downstream end. Another observation is that the monitors are placed where there is a change in entropy for most cases. A similar observation was also reported in Alfonso et al. (2013).

4.3 Greedy optimization approach

Three greedy approaches have been studied in the paper. They are minimization of total correlation (Greedy-TC), maximization of joint entropy (Greedy-JE), and the minimization of the fitness score (equation 5), which is a combination of two objective functions (Greedy-JE_TC). The first two are a conventional single objective greedy algorithm, while the last can be viewed as a multi-objective optimization within a greedy framework. In all three cases, the first monitor was chosen to have the most entropy value. The chronological placement of ten



NSGA-A	Greedy-TC
5.983	6.18791
33.25	33.3856

NSGA-B	Greedy-JE_TC
6.458	6.710793
33.47	35.14624

NSGA-C	Greedy-JE
6.721	6.723856
35.47	35.464

Solution
Joint Entropy
Total Correlation

F

7: Sensor placement of six different solutions on the entropy map coming from NSG-II and Greedy optimizations

monitors for all three approaches are shown in figure 7. From a conceptual point of view, the three solutions NSGA-A, NSGA-B, and NSGA-C are analogous to the Greedy-TC, Greedy-JE_TC, and Greedy-JE, respectively, and are compared in figure 7. Few exciting observations can be made from figure 7. First, the changes in the distribution of monitors from left to right for both NSGA and Greedy are pretty similar. Second, monitors are placed at the upstream and downstream ends in five out of six solutions. Third, solution NSGA-A or Greedy-TC might not be a good solution since the monitors are concentrated in a specific location of the River. Fourth, looking at the joint entropy and total correlation of all six solutions, the greedy algorithm performs slightly better than the NSGA-II, especially Greedy-JE is a clear winner over NSGA-C. Banik et al. (2017b) also found similar results in the case of a wastewater network. They claimed greedy algorithm performs better over NSGA-II, especially at the extreme end of the Pareto front. Fifth, the Greedy-JE solution seems to cover all the vital River features in the forms of either tributary (Jhalukhali, Active Piyan, and Sarigowain), hoar (Kalner, Pathor Churi and Zilkar), or the City water discharges (Sylhet City Corporation). Therefore, if the discharge monitoring stations are placed in those positions, the water resources in that region can be effectively and efficiently managed.

There are four existing discharge monitoring stations in the Surma River. Those are currently placed in 1, 7, 10, and one close to 10, as shown in solution Greedy-JE of figure 7. The joint entropy covered by those four sensors is only 6.049 bits, significantly less than the system joint entropy of 6.803 bits. At the same time, we found from figure 5 that eight monitors will be appropriate for Surma River. So, skipping the last two monitors, which give a joint entropy of 6.69 bits, the Greedy_JE solution will be a good option for this River.

5 Conclusions

A discharge monitoring station is critical for the effective management of water resources. An insufficient number of monitors and the unplanned position may lead to inefficient flood management and inappropriate allocation of water resources among different stakeholders. A case study on designing the discharge monitoring network of Surma River is presented in this paper. The following key findings were obtained from this study.

- After analyzing the entropy map, it has been revealed that the entropy increases where the tributaries meet the River and decreases where the portion of water channels away from the River (e.g., haor). In the later part of the analysis, the monitors were chosen (especially the Greedy-JE solution) where any hydrological feature meets the River.
- Currently, there are four monitors for the Surma River, which has been found insufficient from the information theory point of view. Eight monitors could be sufficient, as further increases in number do not significantly increase the information content.
- NSGA-II will not give a definitive answer to the optimal network design as two conflicting objectives were fought against each other. To get a definite answer, one needs to apply an additional criterion on the near-optimal solutions (Pareto front). In that case, the solution with maximum

joint entropy could be a better choice as it will cover most of the system's information.

- Very interesting to observe that the Greedy-JE solution places the monitors where other hydrological features interact with the River. Under normal circumstances, the monitoring campaign also follows a similar fashion.

Acknowledgments

The authors are grateful to the Research Center, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh, for funding the research work.

Open Research

All the data used to prepare the figures can be found in google drive using the following link. https://drive.google.com/drive/folders/1diIZVdaqgSSnwiT4RRoEk3jei_-fvLFt?usp=sharing. In addition to that, the authors will willingly provide any further information without any restrictions.

References

Alfonso, L., Lobbrecht, A., & Price, R. (2010a). Information theory–based approach for location of monitoring water level gauges in polders. *Water Resources Research*, *46*(3).

Alfonso, L., Lobbrecht, A., & Price, R. (2010b). Optimization of water level monitoring network in polder systems using information theory. *Water Resources Research*, *46*(12).

Alfonso, L., He, L., Lobbrecht, A., & Price, R. (2013). Information theory applied to evaluate the discharge monitoring network of the Magdalena River. *Journal of Hydroinformatics*, *15*(1), 211-228.

Alfonso, L., Ridolfi, E., Gaytan-Aguilar, S., Napolitano, F., & Russo, F. (2014). Ensemble entropy for monitoring network design. *Entropy*, *16*(3), 1365-1375.

Amorocho, J., & Espildora, B. (1973). Entropy in the assessment of uncertainty in hydrologic systems and models. *Water Resources Research*, *9*(6), 1511-1522.

Awadallah, A. G. (2012). Selecting optimum locations of rainfall stations using kriging and entropy. *International Journal of Civil & Environmental Engineering IJCEE-IJENS*, *12*(1), 36-41.

Banik, B. K., Alfonso, L., Torres, A. S., Mynett, A., Di Cristo, C., & Leopardi, A. (2015). Optimal placement of water quality monitoring stations in sewer systems: An information theory approach. *Procedia Engineering*, *119*, 1308-1317.

Banik, B. K., Alfonso, L., Di Cristo, C., & Leopardi, A. (2017a). Greedy algorithms for sensor location in sewer systems. *Water*, *9*(11), 856.

Banik, B. K., Alfonso, L., Di Cristo, C., Leopardi, A., & Mynett, A. (2017b). Evaluation of different formulations to optimally locate sensors in sewer systems.

Journal of Water Resources Planning and Management, *143*(7), 04017026.

Boroumand, A., & Rajaee, T. (2017). Discrete entropy theory for optimal redesigning of salinity monitoring network in San Francisco bay. Water Science and Technology: *Water Supply*, *17*(2), 606-612.

BWDB (2020). *Analysis of water level along the Brahmaputra-Jamuna, Ganga-Padma & Surma-Meghna river syetem.* Surface Water Processing Branch BWDB, 72, Green Road, Dhaka.

Davar, Z. K., & Brimley, W. A. (1990). Hydrometric network evaluation: audit approach. *Journal of Water Resources Planning and Management*, *116*(1), 134-146.

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. M. T. (2002). A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation*, *6*(2), 182-197.

Engel, B., Storm, D., White, M., Arnold, J., & Arabi, M. (2007). A hydrologic/water quality model Applicati1 1. *Journal of the American Water Resources Association*, *43*(5), 1223-1236.

Fekete, B. M., Looser, U., Pietroniro, A., & Robarts, R. D. (2012). Rationale for monitoring discharge on the ground. *Journal of Hydrometeorology*, *13*(6), 1977-1986.

Grabs, W., de Couet, T., & Pauler, J. (1996). Freshwater fluxes from continents into the world oceans based on data from the global runoff database. *Fed. Inst. of Hydrology.*

Gutowski Jr, W. J., Chen, Y., & Ötles, Z. (1997). Atmospheric water vapor transport in NCEP–NCAR reanalyses: Comparison with river discharge in the central United States. *Bulletin of the American Meteorological Society*, *78*(9), 1957-1970.

Hagemann, S., & Dümenil, L. (1997). A parametrization of the lateral waterflow for the global scale. *Climate dynamics*, *14*(1), 17-31.

Hossain, M. S., Yasir, M., Wang, P., Ullah, S., Jahan, M., Hui, S., & Zhao, Z. (2021). Automatic shoreline extraction and change detection: A study on the southeast coast of Bangladesh. *Marine Geology*, *441*, 106628.

Hosseini, M., & Kerachian, R. (2017). A Bayesian maximum entropy-based methodology for optimal spatiotemporal design of groundwater monitoring networks. *Environmental monitoring and assessment*, *189*(9), 1-24.

Husain, T. (1989). Hydrologic uncertainty measure and network design 1. *JAWRA Journal of the American Water Resources Association*, *25*(3), 527-534.

Islam, S. N. (2016). Deltaic floodplains development and wetland ecosystems management in the Ganges–Brahmaputra–Meghna Rivers Delta in Bangladesh.

Sustain. Water Resour. Manag. 2, 237-256. https://doi.org/10.1007/s40899-016-0047-6

Keum, J., Kornelsen, K. C., Leach, J. M., & Coulibaly, P. (2017). Entropy applications to water monitoring network design: a review. *Entropy*, *19*(11), 613.

Kornelsen, K. C., & Coulibaly, P. (2015). Design of an optimal soil moisture monitoring network using SMOS retrieved soil moisture. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(7), 3950-3959.

Leach, J. M., Coulibaly, P., & Guo, Y. (2016). Entropy based groundwater monitoring network design considering spatial distribution of annual recharge. *Advances in water resources*, *96*, 108-119.

Lee, C., Paik, K., & Kim, J. H. (2014). Efficient method for optimal placing of water quality monitoring stations for an ungauged basin. *Journal of environmental management*, *132*, 24-31.

Mahjouri, N., & Kerachian, R. (2011). Revising river water quality monitoring networks using discrete entropy theory: the Jajrood River experience. *Environmental monitoring and assessment*, *175*(1), 291-302.

Mahmoudi-Meimand, H., Nazif, S., Ali Abbaspour, R., & Faraji Sabokbar, H. (2016). An algorithm for optimisation of a rain gauge network based on geostatistics and entropy concepts using GIS. *Journal of Spatial Science*, *61*(1), 233-252.

Mason, P. J., Manton, M., Harrison, D. E., Belward, A., Thomas, A. R., & Dawson, D. K. (2003). The second report on the adequacy of the global observing systems for climate in support of the UNFCCC. *GCOS Rep*, *82*, 74.

Mishra, A. K., & Coulibaly, P. (2009). Developments in hydrometric network design: A review. *Reviews of Geophysics*, *47*(2).

Mishra, A. K., & Coulibaly, P. (2014). Variability in Canadian seasonal streamflow information and its implication for hydrometric network design. *Journal of Hydrologic Engineering*, *19*(8), 05014003.

Mokin, V. B., Kryzhanovskyi, Y. M., Yascholt, A. R., Sobko, B. Y., Slobodianiuk, O. V., Ławicki, T., & Aizhanova, A. (2018, October). The method of multi-criteria ranking of monitoring stations for water discharge in rivers for determining priorities of their location. In *Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2018* (Vol. 10808, p. 108082Q). International Society for Optics and Photonics.

Moss, M. E., & Karlinger, M. R. (1974). Surface water network design by regression analysis simulation. *Water Resources Research*, *10*(3), 427-433.

Moss, M. E., & Tasker, G. D. (1991). An intercomparison of hydrological network-design technologies. *Hydrological Sciences Journal*, *36*(3), 209-221.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885-900.

Preis, A., & Ostfeld, A. (2010). Multi-objective contaminant sensor network design for water distribution systems. *J Water Resour. Plan. and Manage.*, *134*(4) 366-377.

Pilon, P. J., Yuzyk, T. R., Hale, R. A., & Day, T. J. (1996). Challenges facing surface water monitoring in Canada. *Canadian Water Resources Journal*, *21*(2), 157-164.

Ridolfi, E., Montesarchio, V., Russo, F., & Napolitano, F. (2011). An entropy approach for evaluating the maximum information content achievable by an urban rainfall network. *Natural Hazards and Earth System Sciences*, *11*(7), 2075-2083.

Rodda, J. C. (2011). Guide to Hydrological Practices: vol. I: Hydrology–From Measurement to Hydrological Information, and vol. II: Management of Water Resources and Application to Hydrological Practices, 2008 and 2009, WMO 168, World Meteorological Organziation, Geneva, Switzerland; Price CHF70. 00; ISBN 978-92-63-10168-6.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379-423.

Shen, H., & McBean, E. (2011). Pareto optimality for sensor placements in a water distribution system. *Journal of Water Resources Planning and Management*, *137*(3), 243-248.

Stosic, T., Stosic, B., & Singh, V. P. (2017). Optimizing streamflow monitoring networks using joint permutation entropy. *Journal of Hydrology*, *552*, 306-312.

Telci, I. T., Nam, K., Guan, J., & Aral, M. M. (2009). Optimal water quality monitoring network design for river systems. *Journal of Environmental Management* , *90*, 2987–2998

Uddameri, V., & Andruss, T. (2014). A GIS-based multi-criteria decision-making approach for establishing a regional-scale groundwater monitoring. *Environmental earth sciences*, *71*(6), 2617-2628.

Werstuck, C., & Coulibaly, P. (2017). Hydrometric network design using dual entropy multi-objective optimization in the Ottawa River Basin. *Hydrology Research*, *48*(6), 1639-1651.

Yeh, H. C., Chen, Y. C., Wei, C., & Chen, R. H. (2011). Entropy and kriging approach to rainfall network design. *Paddy and Water Environment*, *9*(3), 343-355.

Yang, Y., & Burn, D. H. (1994). An entropy approach to data collection network design. *Journal of hydrology*, *157*(1-4), 307-324.