NEW BRUNSWICK HYDROMETRIC NETWORK ANALYSIS AND RATIONALIZATION

MARCH 2016

UNIVERSITÉ DE MONCTON
EDMUNDSTON MONCTON SHIPPAGAN

Pêches et Océans Canada
Fisheries and Oceans Canada

NEW BRUNSWICK HYDROMETRIC NETWORK

ANALYSIS AND RATIONALIZATION

New Nouveau Brunswick
ENVIRONMENTAL TRUST FUND

March 2016

This page intentionally left blank.

Research Team

Participated in this study:

UNIVERSITÉ DE MONCTON
EDMUNDSTON MONCTON SHIPPAGAN

FACULTY OF ENGINEERING
DEPARTMENT OF CIVIL ENGINEERING:

Jasmin Boisvert

Carole Compaore

Nassir El-Jabi

Noyan Turkkan

FACULTY OF SCIENCES
DEPARTMENT OF MATHEMATICS AND STATISTICS:

Salah-Eddine El Adlouni

Pêches et Océans Canada
Fisheries and Oceans Canada

Daniel Caissie

Université du Québec
Institut national de la recherche scientifique
Eau, Terre et Environnement

Alida Nadege Thiombiano

This page intentionally left blank.

Warning:

Neither the organizations named in this Report, nor any person acting on behalf of any of them, assume any liability for the misuse or misunderstanding of the information presented in this study. The user is expected to make the final evaluation of the appropriateness of the technique and the accuracy of the data and calculations in his or her own set of circumstances.

Avertissement:

Les organisations énumérées dans ce rapport ou toute personne agissant en leurs noms déclinent toute responsabilité pour le mauvais emploi ou la mauvaise interprétation des renseignements contenus dans cette étude. Il incombe aux utilisateurs d'évaluer la pertinence des techniques et l'exactitude des données et calculs dans les circonstances qui s'appliquent

Table of Contents

List of Tables

List of Figures

List of Acronyms and Symbols

$a(i)$	: Regression parameter between station i and all other stations
$b(i)$	: Regression parameter between station i and all other stations
CA	: Clustering Analysis
CNHN	: Canadian National Hydrometric Network
GEV	: Generalized Extreme Value distribution
GEVkapMax	: GEV shape parameter fitted to the maximum annual flow data
GEVkapMin	: GEV shape parameter fitted to the minimum annual flow data
$G(i)$	: Matrix of data from all stations other than i
$H(X)$	: Discrete form of entropy of the continuous, random variable X
$H(X,Y)$	: Joint entropy between X and Y (bivariate case)
$H(Y)$	: Discrete form of entropy of the continuous, random variable Y
κ (kappa)	: GEV shape parameter
k	: Discrete data interval for the variable X
K	: Finite number of class intervals for the corresponding variable X
l	: Discrete data interval for the variable Y
L	: Finite number of class intervals for the corresponding variable Y
Max	: Maximum annual flow
Min	: Minimum annual flow
NBHN	: New Brunswick Hydrometric Network
NG	: North Group (result of clustering)
PC	: Principal Component
PCA	: Principal Component Analysis
$p(x_k)$	: Probability of x_k, based on the empirical frequency of the variable X
$p(x_k, y_l)$	: Joint probability of an outcome corresponding to k for X and l for Y
S1	: Mean flow for January to March (Winter)
S2	: Mean flow for April to June (Spring)
S3	: Mean flow for July to September (Summer)
S4	: Mean flow for October to December (Fall)
SG	: South Group (result of clustering)

$T(X,Y)$	: Transinformation (mutual information) between X and Y
UR	: Indicates a station not ranked by entropy
WMO	: World Meteorological Organization
X	: Continuous, random variable
x_k	: An outcome corresponding to k
Y	: Continuous, random variable
y_l	: An outcome corresponding to l
$\hat{Z}$	: Quantity of information at station i, derived from linear regression
$Z(i)$	: Actual quantity of information contained at station i

Abstract

The availability and quality of hydrometric data is of great importance to the management of water resources, as well as the prediction of flood and drought events. The spatial distribution and density of hydrometric gauging stations are important for precision when estimating design flows, both for gauged and ungauged basins. The lengths of records are also important. Many examples can be found in scientific literature that show that an overly dense (redundant) network as well as an under developed (sparse) network can cause inaccurate simulations of hydrological phenomena. The objective of this study is to propose a methodology for the rationalization of the New Brunswick Hydrometric Network. A Hierarchical Clustering was first used to divide the province into two sections (North and South) based on latitude and high flow timing. After which a Principal Component Analysis was used in an attempt to identify important hydrological attributes that explain a significant amount of the variance found in flows, but was ultimately deemed inconclusive. Instead, the GEV shape parameter, fitted to the annual maximum flow series of each gauging station, was used to split each group into three homogenous subgroups, based on each station's value of the GEV shape parameter. Lastly, an Entropy method was used to rank the importance of each station in their group (North or South), by computing the amount of information that is shared between stations. A station with a lot of shared information is redundant, and therefore less important, whereas a station with very little to no shared information is unique, and thus very important. The ranking of stations by importance can be a useful decisional tool when deciding which stations can be discontinued or displaced, particularly in a budget reduction scenario.

Résumé

La disponibilité et la qualité des données hydrométriques est d'une grande importance pour la gestion des ressources en eau, ainsi que la prévision des crues et des étiages. La distribution spatiale et la densité des stations hydrométriques sont importantes pour la précision lors de l'estimation des débits de conception, tant pour les cours d'eau jaugées que non jaugées. Les longueurs des enregistrements sont également importantes. De nombreux exemples peuvent être trouvés dans la littérature scientifique, qui montre qu'un réseau dense (redondant) ou un réseau faible (peu de stations hydrométriques), peuvent causer des simulations inexactes des phénomènes hydrologiques. L'objectif de cette étude est de proposer une méthodologie pour la rationalisation du réseau hydrométrique du Nouveau-Brunswick. On a d'abord utilisé une approche hiérarchique afin diviser la province en deux secteurs dits homogènes (Nord et Sud) en fonction de la latitude et de l'occurrence des débits extrêmes maxima. Après quoi, une analyse en composantes principales a été utilisée dans une tentative d'identifier les attributs hydrologiques importants qui expliquent un part important de la variance trouvée dans les débits des cours d'eau. Cette dernière mais a été jugée non concluante. Au lieu de cela, le paramètre de forme de la fonction de répartition des valeurs extrêmes (GEV), des séries de débit maximal annuel de chaque station de jaugeage, a été utilisé pour diviser chaque groupe en trois sous-groupes homogènes, basées sur la valeur du paramètre de forme de la GEV de chaque station. Enfin, une méthode d'entropie a été utilisée pour classer l'importance de chaque station dans leur groupe (Nord ou Sud), en calculant la quantité d'information qui est partagée entre les stations. Une station qui comporte beaucoup d'information commune avec autres stations, est considérées redondante, et donc moins importante, tandis qu'une station avec très peu ou pas d'information partagée est considérée unique, et donc très importante. Le classement des stations par ordre d'importance peut être un outil décisionnel utile au moment de décider quelles stations peuvent être interrompues ou déplacées, en particulier dans un scénario de réduction du réseau.

1. Preamble

The importance of hydrometric gauging station networks for surface water monitoring is well established, given the usefulness of collected hydrometric data for decision making related to water resources management around the world (Hannah et al. 2011). However, the density of these networks is still being impacted by the shift of social and economic priorities of governments, like that observed in Canada (Burn 1997; Coulibaly et al. 2013; Mishra and Coulibaly 2009). In fact, Pilon et al. (1996) showed that, through the 1990s, data collection from Canadian National Hydrometric Network (CNHN) declined mainly due to financial pressure that impacted the budget of relevant agencies. More recently, Coulibaly et al. (2013) noticed that only 12% of the Canadian terrestrial area, the majority of which is in the southern portion of the country, is covered by hydrometric networks that meet the minimum standards according to the World Meteorological Organization (WMO) physiographic guidelines. Moreover, 49% of the Canadian terrestrial area is gauged by a sparse network and the remaining 39% is ungauged. Although the negative implications of this may not be immediately apparent, many water resource decisions, project designs and project management rely on information gained by hydrometric gauging stations. In other words, short-comings in a gauging network can lead to greater hydrological uncertainty, which can lead to inefficient project design and resource management, which in turn can have diverse consequences. For example, uncertainty could lead to over-designing, which adds unnecessary extra project costs. In addition, under-designing is also a possibility, which could lead to project failure. Poor resource management can also impact the population as well as the environment. Although reducing the amount of gauging stations available is not ideal according to WMO guidelines, financial and budget restraints may make it necessary. Therefore it seems an evaluation of the network must be undertaken in order to properly analyze options for station reduction or displacement to minimize information loss, thus optimizing the network, such as was done for the Ontario hydrometric network (Ouarda et al. 1996).

Mishra and Coulibaly (2009) provided a review of common methodologies developed to address hydrometric network design or redesign in response to this growing challenge for governments and data users. Using the entropy concept, Mishra and Coulibaly (2010) provided an evaluation of hydrometric network density and the worth of each station, in major watersheds across Ontario, Quebec, Alberta, New Brunswick and Northwest territories. Their study highlighted the generally deficient status of hydrometric networks, mainly over the northern part of Ontario and Alberta, and in the Northwest regions. The entropy concept, derived from Shannon information theory (Shannon 1948), assesses the information content of each gauging station of a given network in relation to all other stations of that network. It was adapted to suit hydrological concerns by Hussain (1987; 1989). Its applications showed its usefulness for optimal hydrometric network design (Alfonso et al. 2013; Li et al. 2012; Mishra and Coulibaly 2010; Singh 1997; Yeh et al. 2011). Nevertheless, multivariate analysis methods such as principal component analysis (PCA) and clustering analysis (CA) remain useful statistical tools in the hydrometric network rationalization process. These methods are commonly used to identify homogeny in a dataset, and potentially form groups of similar individuals (in this case hydrometric gauging stations), which is an important step for network rationalization and optimization (Daigle et al. 2011; Khalil and Ouarda 2009). For example, Morin et al. (1979) derived groups of homogenous precipitation stations from Eaton river sub-basin located in Quebec, using PCA. Their analysis allowed them to propose a better interpolation of spring and summer precipitation amounts from a less redundant network. For their part, Khalil et al. (2011) used PCA to select variables that better explain water quality in the Nile Delta watershed, and CA to extract different sub-hydrological units in order to better perform their assessment and redesign of the water quality monitoring network. Van Groenewood (1988) also used PCA to divide the New Brunswick into ten climatic regions. PCA and CA are used in most studies as statistical tools for preliminary datasets preparation (Burn and Goulter 1991; Ouarda et al. 1996).

Network optimization cannot be accomplished by solely using these purely statistical approaches mentioned above. There are other factors that must be taken into account. For example, a gauging station attached to a hydroelectric facility may not be statistically important in a network, but would most likely not be removed. Data user needs and perception must be integrated in any analysis of a network. It has been recommended and integrated in previous (Burn 1997; Coulibaly et al. 2013; Davar and Brimley 1990). Environment Canada and New Brunswick Dept. of Municipal Affairs and Envir. (1988) investigated accuracy requirements identified by users in order to define a minimum and target networks. They considered mean, low and high flows in this approach, which consisted of developing regional equations for each of the three categories. They initially identified 16 homogenous regions in the province, considering that there should be a small, medium, and large gauged basin in each homogenous area. This implied that 48 stations, plus an additional 6 for larger regions (total of 54 stations), was identified as a minimum network. They also identified a target network, this time considering that 10 stations were necessary per region in order to properly define regional regression equations. However, they also refined the initial 16 homogenous regions into 7 regions. This implied that 70 stations (plus an additional 7 for variations in size) were suggested as the target (total of 77 stations). They concluded that it was important to coordinate hydrographic gauging with meteorological gauging, that more gauging was necessary for smaller catchments, and that the central part of the province lacked gauging stations. Overall, their recommendation was to add 17 stations to reach what they considered to be a minimum network, with another 9 stations in addition to those 17 to reach what they considered to be a good target network. They also evaluated the hydrometric network using an audit approach, through which a ranked prioritization of stations was provided based on the hydrometric, socio-economic and environmental worth of each station according to data user perceptions. They also considered site characteristics, economic activity, federal and provincial commitments, special needs, as well as a station's regional and operational users in their audit approach. (Davar and Brimley 1990) used a similar approach to identifying a minimum and target network as Environment Canada and New Brunswick Deptarment of Municipal Affairs and Environment (1988), but their audit was slightly different. The existing stations and proposed new stations were evaluated using an audit approach, based on site characteristics, client needs (regional hydrology and operational), and regional water resource importance. They created different scenarios that had different impacts and values (based on audit points) in function of different costs (adding, removing, or maintaining the amount of gauging stations in the network). Overall, their recommendations included : reallocating resources to meet the minimum network; create a committee for ongoing planning and analysis, as well as communication with the user community; emphasize the importance of regional hydrology; coordinate with other related data gathering, such as water quality and atmospheric data;

2. Objectives and Aims

Hydrometric network rationalization and optimization is still a relevant challenge in Canada. The required assessment must define and integrate appropriate criteria for each region for the network to be properly updated. It is in this context that the present study aims to propose a rationalization of the hydrometric gauging network of New Brunswick (NB). This will be accomplished using the mentioned Clustering Analysis (CA) and Principal Component Analysis (PCA) as a preliminary evaluation of hydroclimatic behaviour and homogeny between gauging stations, as well as the entropy concept to quantify the importance of each station regarding information content. In order to have a more complete rationalization process, data managers and users should be consulted for their input on station importance.

3. New Brunswick Hydrometric Network

The hydrometric gauging station network being analyzed by this study is the New Brunswick Hydrometric Network (NBHN). There are also a few gauging stations located in Québec and in Maine (U.S.) that can be considered relevant to New Brunswick, since the watersheds of some rivers located in New Brunswick are partially located outside the province. The current network, as identified by Environment Canada, contains 67 stations. Of these 67 stations, 46 are active and 21 are discontinued. Table A1 located in Annexe A lists these stations, as well as some of their relevant properties.

The first measurements taken in the province were in 1918. The major expansion of the network occurred in the late 1960's, continuing in the early 1970's. This was caused by an increased demand for data for water supply, fisheries, and flood forecasting (Davar et al. 1990). Many stations were originally established to suit specific needs, often short-term. After their objectives were completed, these stations were kept in service. This method of network expansion was considered acceptable at the time (Davar et al. 1990). Although this method did in fact create an expanded network, it is not necessarily the most effective method. Since new stations were added in locations for a specific purpose (i.e. a single project), little consideration was given to the network as a whole. This implies that new stations may have been placed in similar locations to existing stations, causing redundancy in the information measured. Similarly, some areas of the network may have been lacking measurements, such as : a particular section of the province, a certain climatic region, or a certain range of catchment areas. If no specific need to add a stations in these areas arose, than they might remain under developed. Thus the need to analyze and optimize the network; one of the objectives of this study.

4. Numerical Analysis

It should be noted that for all the methods used in this study, the specific discharge (discharge per unit area; m^3/s per m^2) will be used as opposed to using the flow (m^3/s). This is done due to the fact that during some analyses, such as clustering analysis or principal component analysis, drainage area becomes an overwhelmingly dominant variable when it comes to explaining flow rates. Consequentially, all other variables (e.g. precipitation, latitude, temperature, etc.) become insignificant in comparison, defeating the purpose of these analyses.

4.1 Clustering Analysis

4.1.1 Objective

The objective of the clustering analysis is to divide, in a preliminary context, the province's gauging stations into groups which share similar traits. This has the intention of facilitating the analysis that will follow (PCA and entropy) by dividing the network into smaller, homogenous groups of stations. Rationalization and optimization assessment of the network has been shown to be better conducted with the division of a network into climatic regions (Burn and Goulter 1991; Khalil et al. 2011).

4.1.2 Methodology

The attributes from which similarities will be defined need to be specified for clustering analysis (Burn and Goulter 1991). Once this is done, clusters are formed by grouping similar observations together in such a way that variance is minimized within a cluster and maximized between clusters (Khalil and Ouarda 2009). The division of the complete network into clusters is done using hierarchical agglomerative clustering (based on Euclidean distance), accomplished using R software toolbox (R Core Team 2015). In this type of clustering, each individual station is initially considered as being its own cluster. Afterwards, an iterative process is used in which only the two most similar clusters (least Euclidean distance between two clusters of all possible combinations) are joined together to form one new cluster per iteration. This is repeated until a single cluster remains, containing all the individuals. In this study, two attributes were used for the clustering analysis : latitude of each station, and high flow timing. The latter is computed as the 30-day period with the highest mean flow (moving average). For example, if the highest mean was computed between April 4th and May 3rd, then the timing of the high flow would be considered as April 4th. The two attributes (latitude and timing) were chosen with the purpose of dividing the province based on climate. The high flow timing is typically dependent on temperature, due to snowmelt. The northern part of the province is typically cooler than the southern part. As such, using latitude and high flow timing, it is expected that the province will be divided into clusters in a north-south manor. All 67 stations identified by Environment Canada were used in this analysis

4.1.3 Results

Two clusters were formed in the hierarchical clustering analysis, based on high flow timing. A dendrogram is formed using the hierarchical clustering technique (Figure 1). The two major groups formed by the clustering analysis can be seen on this figure (identified in red).

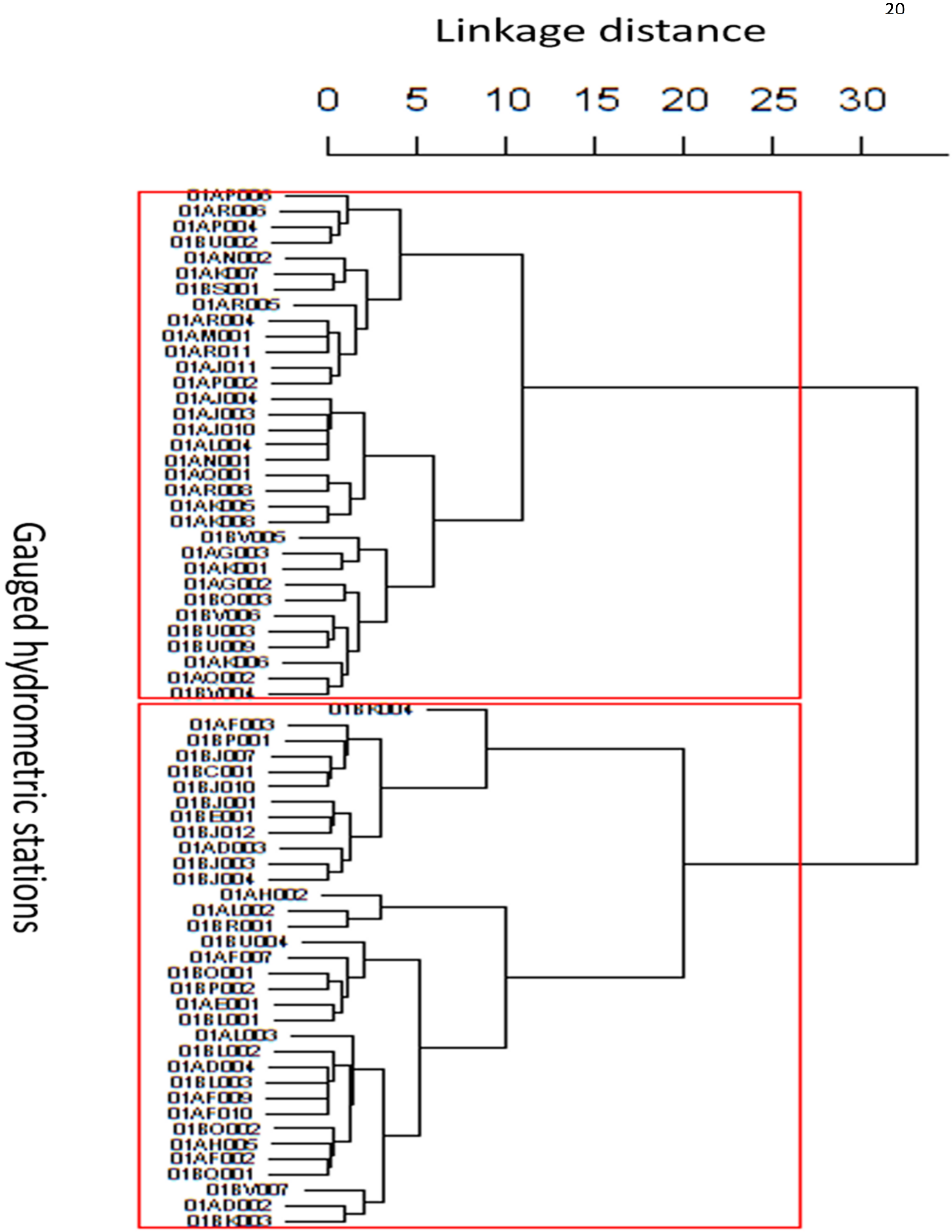

Figure 1. Hierarchical clustering of NB gauged hydrometric stations.

Each horizontal bar connecting two stations (or groups) corresponds to the maximum difference in timing of the stations within the two connected groups. For example, the stations AP4 and BU2 (3rd and 4th from the left), are connected by a horizontal line positioned at a value of close to 0, implying they have very similar high flow timing and latitude. Furthermore, station AR6 is connected to the previously mentioned group of two stations by a line positioned at a value of close to 1, indicating a difference in Euclidean distance (timing and latitude) between AR6 and the other two stations of close to 1, which is also a small distance. It should be noted that the method used for clustering was the complete linkage method. This simply means that the distance between clusters is calculated as the maximum possible Euclidean distance between a pair of stations, one from each cluster. This is important when selecting which two clusters to join together in an iteration, since other methods could use the minimum distance (single linkage), average distance (mean linkage), or other criterion, possibly yielding different results.

Since the groups are mostly positioned in a north-south fashion, the two groups are named North Group (NG) and South Group (SG). These two groups will be analyzed separately in the analyses that follow (principal component analysis, entropy). Figure 2 presents a map of these stations in the province. Looking at this map, it appears that there is a horizontal section of the province at around 46.5º of approximately 35 km in width that traverses the province where no gauging stations are present. This line also seems to divide the north from the south in terms of high flow timing. As such, it should be noted that the results of the clustering analysis were slightly modified for the final classification into the two groups (NG and SG). Stations BV7, BU4, AL3, and AL2 had flow timings similar to the North Group, despite being more southern stations. These stations were analysed part of the South Group, as they were a significant distance from the north, and typically surrounded by southern stations. Similar reasoning was applied to station BO3, which was clustered in the south, but located in the north. It was analysed part of the North Group. Similar reasoning could have been applied to stations AG2 and AG3 as well. However, these two stations are very close to the perceived divisional line mentioned above, and there are no other stations close to them. Taking that into account, they were not switched from their original cluster (SG) in favour of NG, as was done with BO3, but instead were allowed to remain in the South Group. Of the 67 stations used for the clustering, 31 were placed in NG and 36 in SG. It should be noted that the stations in Figure 2 that are in gray have been discontinued and are no longer active

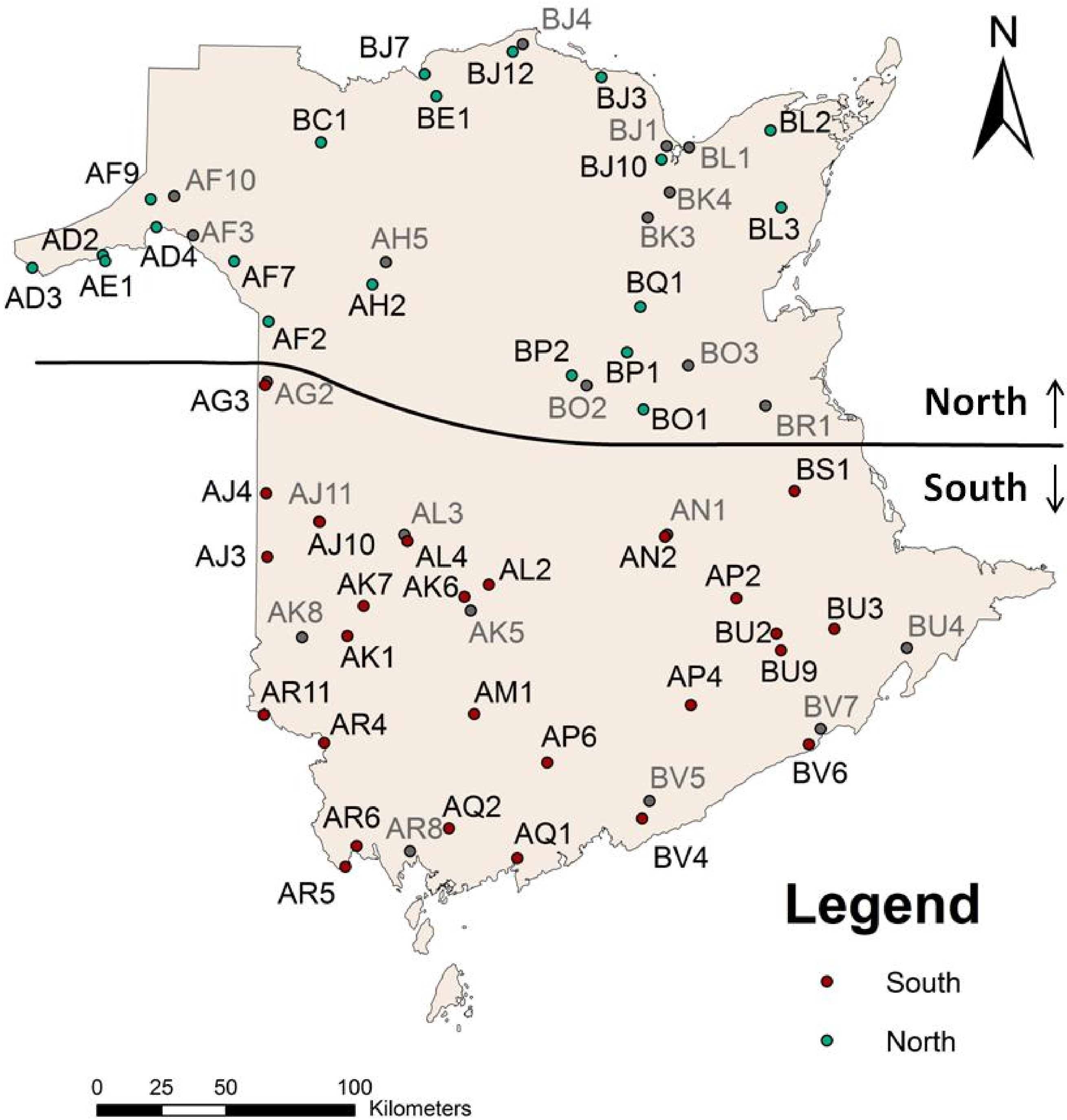

Figure 2. New Brunswick hydrometric gauging station network, with North-South division based on clustering results. Inactive stations are shown in gray.

4.2 Principal Component Analysis

4.2.1 Objective

Principal component analysis is commonly used to reduce the dimensionality of a dataset through the creation of new subsets of uncorrelated variates, called principal components (PCs) (Daigle et al. 2011; Westra et al. 2007). This is helpful in selecting the best attributes to explain variance in a dataset, from which homogenous subsets of interested variable observations can be derived (Khalil et al. 2011; Khalil and Ouarda 2009). In this study, PCA was used to identify which hydrological attributes, among several predefined possibilities, that better explained the variance in river flows. The purpose of the identification of these attributes is to further subdivide the clusters from the clustering analysis into smaller groups of homogenous data. Unlike the clustering analysis, which divided the province into the North and South groups, this division is not at all based on proximity. The North and South groups will be divided into sub groups not by geographical lines, but grouped together by common traits, regardless of position. It is important to note that each group has its importance. Therefore when analyzing which gauging stations are of little importance and can be removed, it is advisable to not remove the majority or entirety of a single group, even if they are considered to be the least statistically important. It would be preferable to remove a few of the least important stations per group, as opposed to several from the same group. It should be noted that of the 67 stations used in the clustering analysis, AD4 (NG) and BV7 (SG) were removed from the principal component analysis, given poor quality of their data (short record length and interpolated data). Therefore the principal component analysis was carried out with the remaining 65 stations; NG containing 30 stations, SG containing 35 stations.

4.2.2 Methodology

Eight attributes were chosen as characteristics to describe the data series : annual maximum (Max) and minimum (Min); mean flow for January to March (winter season; S1), April to June (spring season; S2), July to September (summer season; S3), and October to December (autumn season; S4); and the shape parameter (kappa; κ) of the Generalized Extreme Value (GEV) distribution fitted to the annual maxima time series (GEVkapMax) and the annual minima time series (GEVkapMin). These parameters are all describe flow and flow patterns. It was judged more reasonable to use these as opposed to other variables, such as precipitation and temperature, to group the gauging stations together, since the flow indirectly contains this information.

The first PC (PC1) is defined in such a way that it maximizes the explained variance of the dataset. The higher level PCs (PC2, PC3. etc.) are computed from the residuals of all previous PCs (Daigle et al. 2011). For example, PC2 is calculated from the residuals of PC1, PC3 is calculated from the residuals of PC1 and PC2. The first two PCs typically explain the majority of the variance, and thus only these two will be analyzed in this study. In addition, the contribution of each original attribute to each PC is quantifiable. This implies that the attributes with the highest impact on flow information can be defined. Once these attributes are defined, the stations can be analyzed in function of these attributes in order to further divide the clusters into smaller groups, either by finding patterns or similarities, or by using significant values or thresholds related to an attribute.

3.2.3 Results

Since PCs are orthogonal to one another, the first two PCs can be represented as a plane, with the horizontal axis being PC1 and the vertical axis being PC2. Each of the original metrics can be plotted in this space using their respective contributions to each PC. This representation, called a correlation circle, can give a visualization of the importance of each metric. Figures 3a and 3b show the correlation circles for the North Group and South Group respectively.

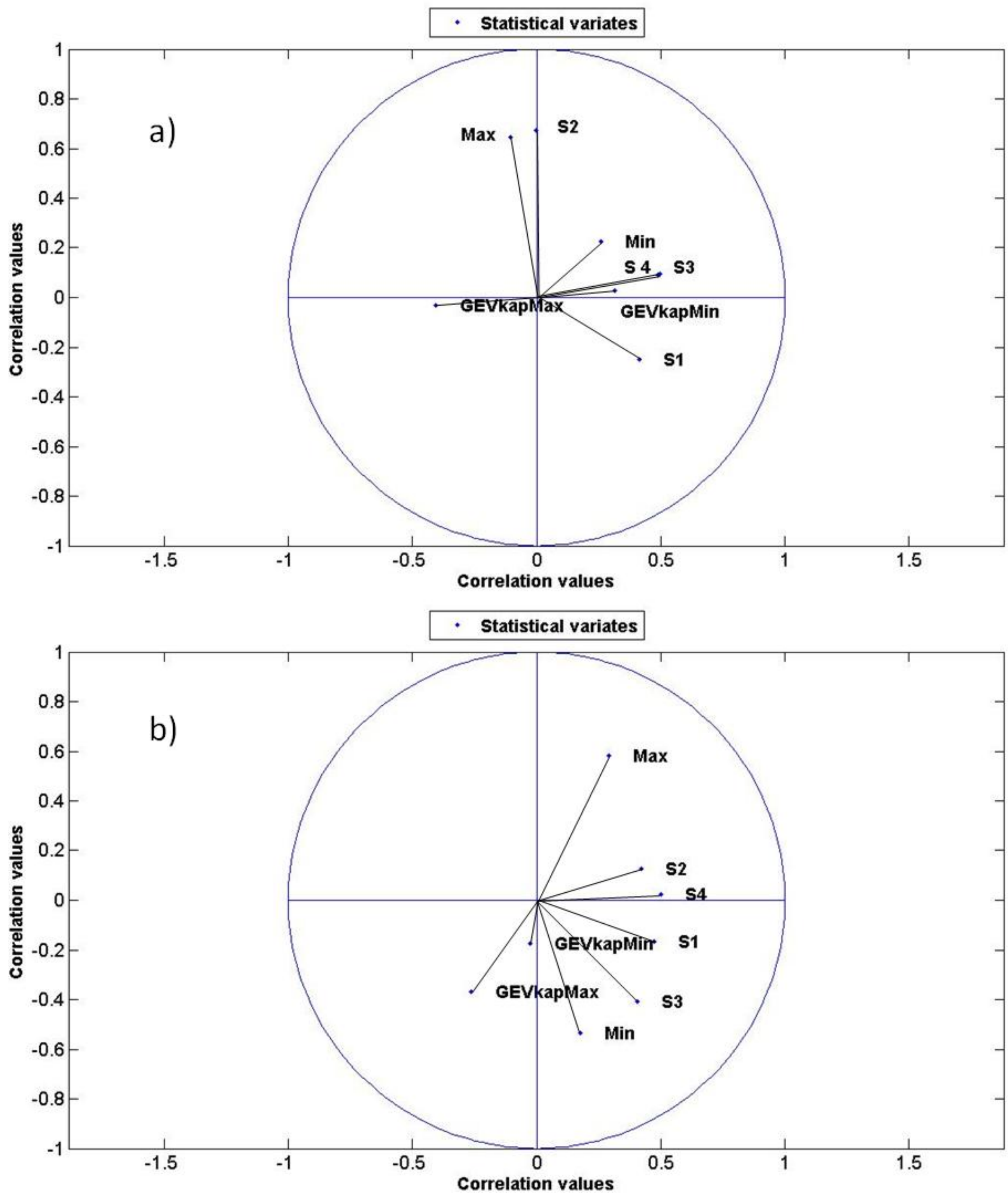

Figure 3. PCA metric Correlation Circles for a) North Group and b) South Group.

Many different observations can be made from these figures. The projection of a metric's vector onto each PC is the strength of the correlation between the metric and the PC. For example, S4 in the South group has a correlation to PC1 and PC2 of 0.50 and 0.03 respectively, whereas Max has a correlation to PC1 and PC2 of 0.29 and 0.58 respectively. In this example S4 has a high correlation to PC1, which is the most important of the PCs, but no correlation to PC2, which still explains a significant portion of the variance. In contrast, Max is highly correlated to PC2, and also has a significant correlation to PC1. As such, it could be argued that Max is the most important metric, even though it is not the most correlated to PC1, due to the length of its vector being the largest. Another observation can be made using the angle between metric vectors. If the angle is close to 0º, than they are positively correlated. If the angle is close to 180º, they are negatively correlated. If the angle is close to 90º or 270º than there are not at all correlated. In the South Group, Max and GEVkapMax would be negatively correlated, whereas Max and S3 would have almost no correlation. In order for any of these observation to be considered accurate, the correlation values (length of vectors) of the involved metrics has to be strong. Since the results of the Principal Component Analysis showed that none of the metrics in either the North or South groups are particularly strong or dominant, it would be difficult to define subgroups for the respective regions using one of these metrics. As such the results of the Principal Component Analysis are inconclusive, and another method will be used to sub divide the North and South Groups.

4.2.4 GEV Shape Parameter

In a study characterizing natural flow regimes and environmental flows in New Brunswick (El-Jabi et al. 2015), it was found that the GEV distribution was an appropriate distribution to model the annual maximum and minimum flows at most of the gauging stations in New Brunswick, using the Anderson-Darling test. Consequentially, it seems that GEVkapMax and GEVKapMin are good for characterising flows in the province. As such, differences and similarities in these values between stations would be interesting to investigate. Since the annual maximum flows were particularly well modeled by the GEV distribution, and the maximum flows are generally of more interest, the GEVkapMax was deemed as the metric to be used for dividing the North Group and South Group into smaller homogenous subgroups. The GEV probability density function is shown by Equation 1.

$$f(x)=\frac{1}{\alpha}[1-\frac{\kappa}{\alpha}(x-u)]^{\frac{1}{\kappa}-1}\exp\{-[1-\frac{\kappa}{\alpha}(x-u)]^{\frac{1}{\kappa}}\} \quad (1)$$

where x is a random variable in this case the specific discharge, κ is the shape parameter, α is the scale parameter, and u is a position parameter. In addition, the following restriction applies : $x<u+\alpha/\kappa$ if $\kappa>0$; $x>u+\alpha/\kappa$ if $\kappa<0$. The shape parameter, as suggested by its name, is responsible for the shape of the distribution. This means that depending on the parameter, the distribution can be symmetrical ($\kappa=0$), asymmetrical with a heavy left tail ($\kappa<0$), or asymmetrical with a heavy right tail ($\kappa>0$). The GEV shape parameter (kappa) has three statistically significant categories. These are used to subdivide the North Group and South Group each into three subgroups. The first category (NG1 and SG1), where kappa is between]-0.33; +0.33[, has a mean, variance and a skew that can be computed. The second category (NG2 and SG2), where kappa is between]-0.5; -0.33] or [+0.33; +0.5[, has a skew that is infinite. The third category (NG3 and SG3), where kappa is between]-∞; -0.5] or [0.5; ∞[, has an infinite variance as well as an infinite skew. It should be noted that a negative GEV shape parameter (kappa) value produces a positive skew (heavy left side of the distribution), which is most common in hydrology. Table 1 lists the six groups and the stations that belong to them. Figure 4 shows the position of the stations and to which group they belong.

Table 1. Division of the North and South Groups into subgroups based on the GEVkapMax parameter.

NG1 Kap ϵ]-0.33 ; +0.33[	NG2 Kap ϵ]-0.5 ; -0.33[	NG3 Kap < -0.5	SG1 Kap ϵ]-0.33 ; +0.33[	SG2 Kap ϵ]-0.5 ; -0.33[	SG3 Kap < -0.5
AF7	BO2	BL1	AK1	AR11	AN2
BQ1	AF3	BR1	AP2	AG2	AR8
BO1	BL3	AH5	AG3	BU2	
AD3	BL2	AF9	AL4	AK5	
AH2	BO3	BJ4	AR5	AJ4	
BJ3	BE1		AM1	AK8	
BC1	BJ1		AR4	AJ11	
BJ7	BJ10		AR6		
BP1	BK3		BV6		
AE1	BK4		BU3		
AD2			BS1		
AF2			AL2		
AF10			AP4		
BJ12			AK7		
BP2			AQ2		
			AP6		
			AN1		
			AJ3		
			AL3		
			AJ10		
			AQ1		
			AK6		
			BU4		
			BU9		
			BV4		
			BV5		

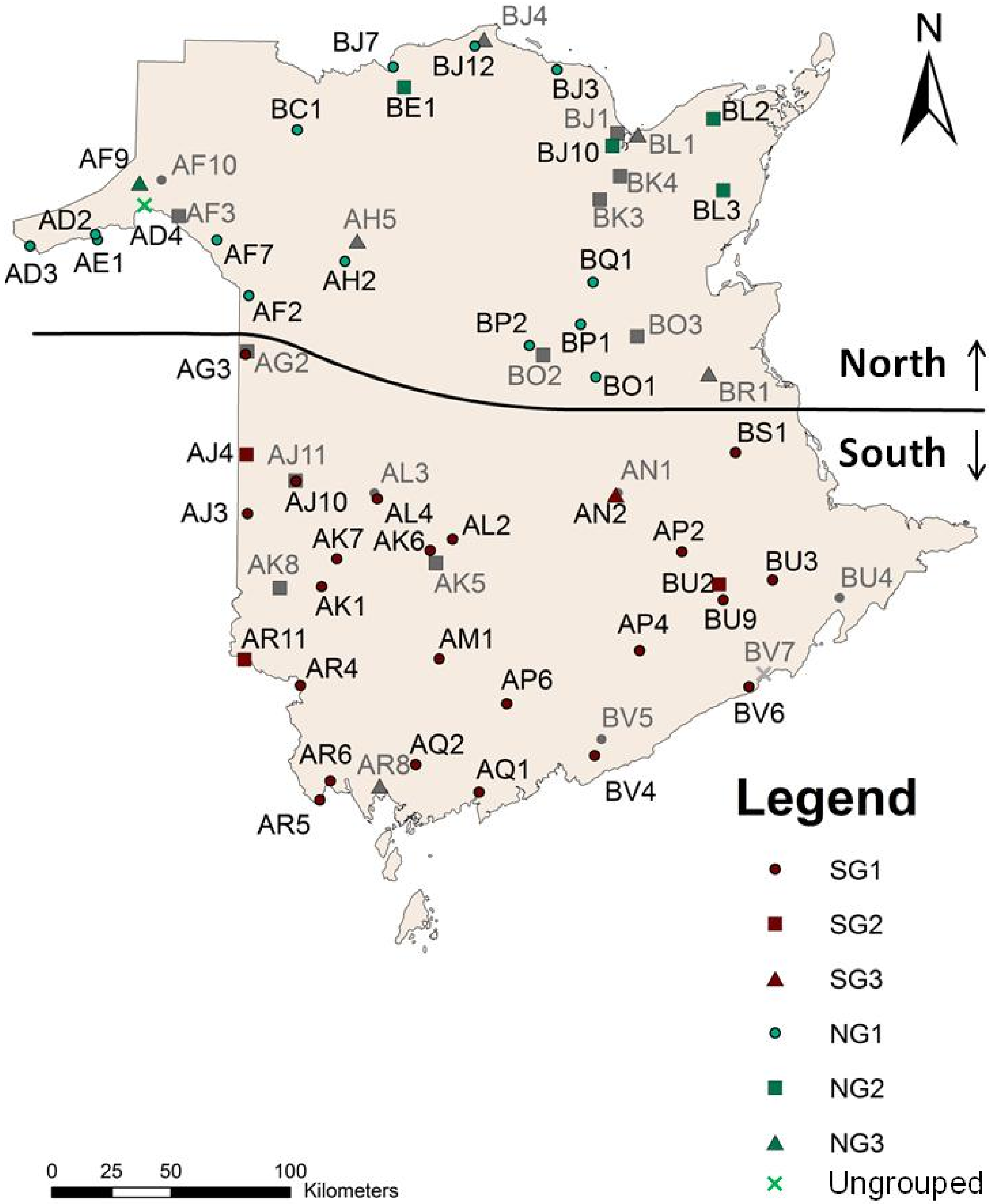

Figure 4. Division of the North and South Groups into subgroups based on the GEVkapMax parameter. Inactive stations are shown in gray.

4.3 Entropy Analysis

4.3.1 Objective

The objective of the entropy concept analysis is to quantify the information contained in the random variable (specific discharge) measured at the different gauging stations. This is important since it provides an objective criterion to describe each station. However, it is actually the measure of transinformation that is of particular interest in this study. The measure of transinformation, a function of marginal entropy and joint entropy, indicates if the same information is measured by multiple stations (redundancy), or if the information measured by a station is unique (optimal). This gives an idea of the relative importance of each station, given the principles of information maximization (Hussain 1987; 1989; Singh 1997; Mishra and Coulibaly 2010). This allows for better decision making when it comes to choosing if a station should be removed, displaced, or continued. For example, a station that only measures information that is also contained in other stations is highly redundant, adds no value to the network, and can be removed with minimal loss of information. In contrast, a station whose information is unique is highly valuable to the network, and should not be removed. It should be noted that a limitation of this method is the fact that the data from each stations has to be in the same time period (of at least 20 years), and the whole period must be covered. Therefore, a period of time was chosen where the greatest amount of stations had taken measurements for the full period. The window chosen was 1976-1995. Of the 65 stations used for the principal component analysis, 53 remain for the entropy analysis (23 in NG, 30 in SG). The annual maximum specific discharge is used for the entropy analysis.

4.3.2 Methodology

A station malfunctioning for a few days or even months is not uncommon. Therefore, it is important before proceeding to the entropy calculations to deal with missing data. To complete the data, a correlation matrix between stations with missing values and stations without them is constructed (Mishra and Coulibaly 2010), using a linear regression analysis (Ouarda et al. 1996). The individual station with complete data that showed the maximum correlation with a station having missing data was used to fill the data.

The transinformation (or mutual information) $T(X,Y)$ is described in Equation 2 as the information about a predicted variable transferred by the knowledge of a predictor (Mishra and Coulibaly 2010) as follows:

$$T(X,Y) = H(X) + H(Y) - H(X,Y) \tag{2}$$

In Equation 2, $T(X,Y)$ is the transinformation, whereas $H(X)$ and $H(Y)$ are the discrete form of entropy of the continuous random variables X and Y, as described by Equation 2, formulated by Shannon (1948) and updated by Hussain (1987; 1989) for use with hydrological time series data.

$$H(X) = -\sum_{k=1}^{K} p(x_k)\log\left[p(x_k)\right] \tag{3}$$

This information coefficient only gives a measure of information from the concerned random variable; hence the importance of joint entropy between the interested variables (flow time series), as described by Equation 3 as $H(X,Y)$ for the bivariate case. This allows the measurement of the overall information retained by random variables (Li and al. 2012). The logical extension can be made for the multivariate case.

$$H(X,Y) = -\sum_{k=1}^{K}\sum_{l=1}^{L} p(x_k, \mathrm{y}_l)\log\left[p(x_k, \mathrm{y}_l)\right] \tag{4}$$

In the above equations, k and l denote a discrete data interval for the variables X and Y, respectively; K and L are the finite number of class intervals for the corresponding variables with the general assumption that $K = L$; x_k is an outcome corresponding to k; $p(x_k)$ is the probability of x_k and is based on the empirical frequency of the variable X; $p(x_k, y_l)$ is the joint probability of an outcome corresponding to k for X and l for Y. In the case where the entropy concept is being applied to a hydrometric gauged network, the variable X becomes $Z(i)$; the actual quantity of information contained at station i. The variable Y becomes $\hat{Z}$ the quantity of information at station *i*, but this time derived from the linear regression demonstrated in Equation 5.

$$\hat{Z} = a(i) + b(i) * G(i) \tag{5}$$

In this equation, $G(i)$ is a matrix of data from all other stations, $a(i)$ and $b(i)$ are the parameters of the regression between station i and all other stations, assuming a linear relation between stations is deemed appropriate. The transinformation becomes $T\left(Z, \hat{Z}\right)$ (Burn 1997; Mishra and Coulibaly 2010). The data used for all these computations is the annual series of maximum monthly specific discharge. Since the entropy analysis is performed over a 20 year window, each station has a data series of 20 points, each one representing the average specific discharge for the month with the highest average specific discharge of that year.

Once the transinformation has been evaluated for each station, it can be used to rank station in order of importance (Li et al. 2012; Yeh et al. 2011). Stations with smaller transinformation values are the most important stations, since they contain little redundant information, and thus get ranked the highest (1 being the most important).

4.3.3 Results

The results of the entropy computation are presented in Tables 2a and 2b for the North Group and South Group respectively. The rank of the stations is also included in the table, and is simply the order of the value of $T\left(Z,\hat{Z}\right)$, from lowest (rank 1; most important station) to highest (rank 28; least important station). It is important to note that stations BL1, AK1, and AP2 are considered to be the most important stations, given that their values of $H\left(\hat{Z}\right)$ and $H\left(Z,\hat{Z}\right)$ are zero. This implies that the information measured by these stations is unique, and consequentially very important.

Table 2a. Entropy values and ranking of each station (North Group).

Station	$H(Z)$	$H(G)$	$H(Z,G)$	$T(Z,\hat{Z})$	R
AD2	2,2253	2,1050	2,7520	1,5784	20
AD3	1,7926	2,2071	2,7499	1,2499	8
AE1	2,1478	2,2071	2,7876	1,5673	19
AF2	2,1744	2,2681	2,8520	1,5905	21
AF3	2,0100	2,0673	2,9253	1,1520	5
AF7	2,2071	2,0100	3,1765	1,0406	2
AH2	2,1266	2,2253	3,0058	1,3462	10
AH5	2,1744	2,1499	2,9303	1,3939	14
BC1	1,9233	1,9416	2,3876	1,4773	16
BE1	1,8623	2,1233	2,6253	1,3602	12
BJ1	2,1266	2,2499	2,9926	1,3839	13
BJ3	1,9171	2,2071	2,7681	1,3561	11
BJ7	1,9623	2,0058	2,4855	1,4826	17
BJ10	2,1499	2,2253	2,9765	1,3987	15
BL1	1.5694	-	-	-	0*
BL2	2,2071	2,0681	3,0681	1,2071	7
BL3	2,0681	1,9416	2,8233	1,1865	6
BO1	1,8744	2,1644	2,9142	1,1245	4
BO2	1,8449	1,8744	2,7499	0,9694	1
BO3	2,0681	2,0428	2,7876	1,3233	9
BP1	2,1499	2,2071	2,8520	1,5050	18
BQ1	2,0100	2,0428	2,9876	1,0652	3
BR1	2,0428	2,0681	2,7876	1,3233	9

*A rank of 0 means the station's information is unique, thus very important.

Table 2b. Entropy values and ranking of each station (South Group).

Stations	$H(Z)$	$H(G)$	$H(Z,G)$	$T(Z,\widehat{Z})$	R
AG2	2,2253	2,1478	3,1681	1,2050	7
AG3	1,8253	1,9876	2,9876	0,8253	1
AJ3	1,9876	2,0794	2,4926	1,5744	21
AJ4	2,0303	2,0058	2,4694	1,5668	19
AJ10	2,2071	2,2071	2,5765	1,8377	25
AJ11	2,1644	2,2499	2,6499	1,7644	24
AK1	2.1449	-	-	-	0*
AK5	1,9303	1,9303	2,5071	1,3536	14
AK6	2,1233	2,0681	2,2855	1,9058	28
AK7	2,1926	2,1644	2,9303	1,4266	16
AK8	2,1499	2,1303	2,7058	1,5744	22
AL2	1,8253	2,0100	2,4926	1,3428	13
AL3	2,0694	2,1499	2,5897	1,6295	23
AL4	2,1121	2,0855	3,1142	1,0834	2
AM1	1,6989	2,0694	2,6549	1,1134	4
AN1	2,1926	2,1050	2,7253	1,5723	20
AN2	2,2071	2,2253	2,5338	1,8987	27
AP2	2.1449	-	-	-	0
AP4	1,8478	2,1926	2,6765	1,3639	15
AP6	2,2171	2,0549	2,8171	1,4549	18
AQ1	2,1171	2,1499	2,4142	1,8527	26
AQ2	2,0681	2,0694	2,6926	1,4449	17
AR4	1,9623	2,1744	3,0142	1,1224	6
AR5	1,9050	2,2071	3,0058	1,1063	3
AR6	2,0428	2,1499	2,9303	1,2623	9
AR11	2,1744	1,9623	3,0142	1,1224	5
BS1	2,1121	2,2499	3,0303	1,3316	12
BU2	2,2071	2,1499	3,1142	1,2427	8
BU3	2,0673	2,1449	2,8926	1,3196	11
BV6	2,0428	2,0855	2,8499	1,2784	10

*A rank of 0 means the station's information is unique, thus very important.

Tables 3a and 3b show the ranking of the stations divided into their respective groups. It is important to remember that removing the majority or entirety of a group is not advisable, since each group has its statistical importance. It would be preferable to remove a few of the least important stations per group, as opposed to several from the same group, even if the stations from a single group are ranked lower by the entropy analysis. Figure 5a and Figure 5b show the positions of these stations and their ranks for the North and South respectively. Figure 6 shows the ranks of the stations of the current network (only active stations).

Table 3a. Entropy values and ranking of each station per subgroup (Nouth Group).

NG1 (Rank)		**NG2 (Rank)**		**NG3 (Rank)**	
AF7	(2)	BO2	(1)	BL1	**(0)**
BQ1	(3)	AF3	(5)	BR1	(9)
BO1	(4)	BL3	(6)	AH5	(14)
AD3	(8)	BL2	(7)		
AH2	(10)	BO3	(9)		
BJ3	(11)	BE1	(12)		
BC1	(16)	BJ1	(13)		
BJ7	(17)	BJ10	(15)		
BP1	(18)				
AE1	(19)				
AD2	(20)				
AF2	(21)				
AF10	**(UR)***	BK3	**(UR)**	AF9	**(UR)**
BJ12	**(UR)**	BK4	**(UR)**	BJ4	**(UR)**
BP2	**(UR)**				

*__UR__ indicates that the station was excluded from the entropy analysis

Table 3b. Entropy values and ranking of each station per subgroup (South Group).

SG1 (Rank)		SG2 (Rank)		SG3 (Rank)	
AK1	**(0)**	AR11	5	AN2	27
AP2	**(0)**	AG2	7		
AG3	(1)	BU2	8		
AL4	(2)	AK5	14		
AR5	(3)	AJ4	19		
AM1	(4)	AK8	22		
AR4	(6)	AJ11	24		
AR6	(9)				
BV6	(10)				
BU3	(11)				
BS1	(12)				
AL2	(13)				
AP4	(15)				
AK7	(16)				
AQ2	(17)				
AP6	(18)				
AN1	(20)				
AJ3	(21)				
AL3	(23)				
AJ10	(25)				
AQ1	(26)				
AK6	(28)				
BU4	**(UR)**			AR8	**UR**
BU9	**(UR)**				
BV4	**(UR)**				
BV5	**(UR)**				

***UR** indicates that the station was excluded from the entropy analysis

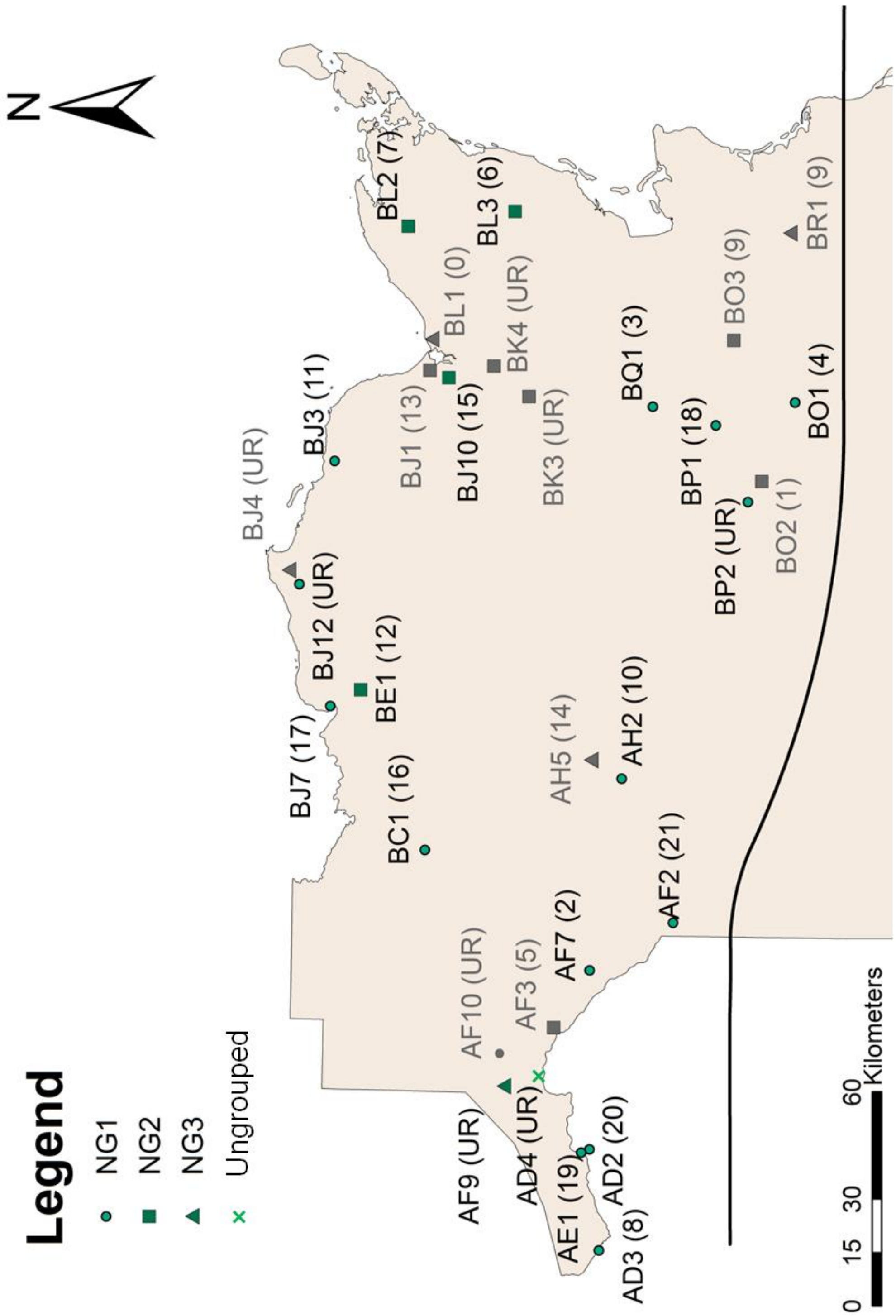

Figure 5a. Map of gauging stations (North), as well as their group and rank. Inactive stations are shown in gray.

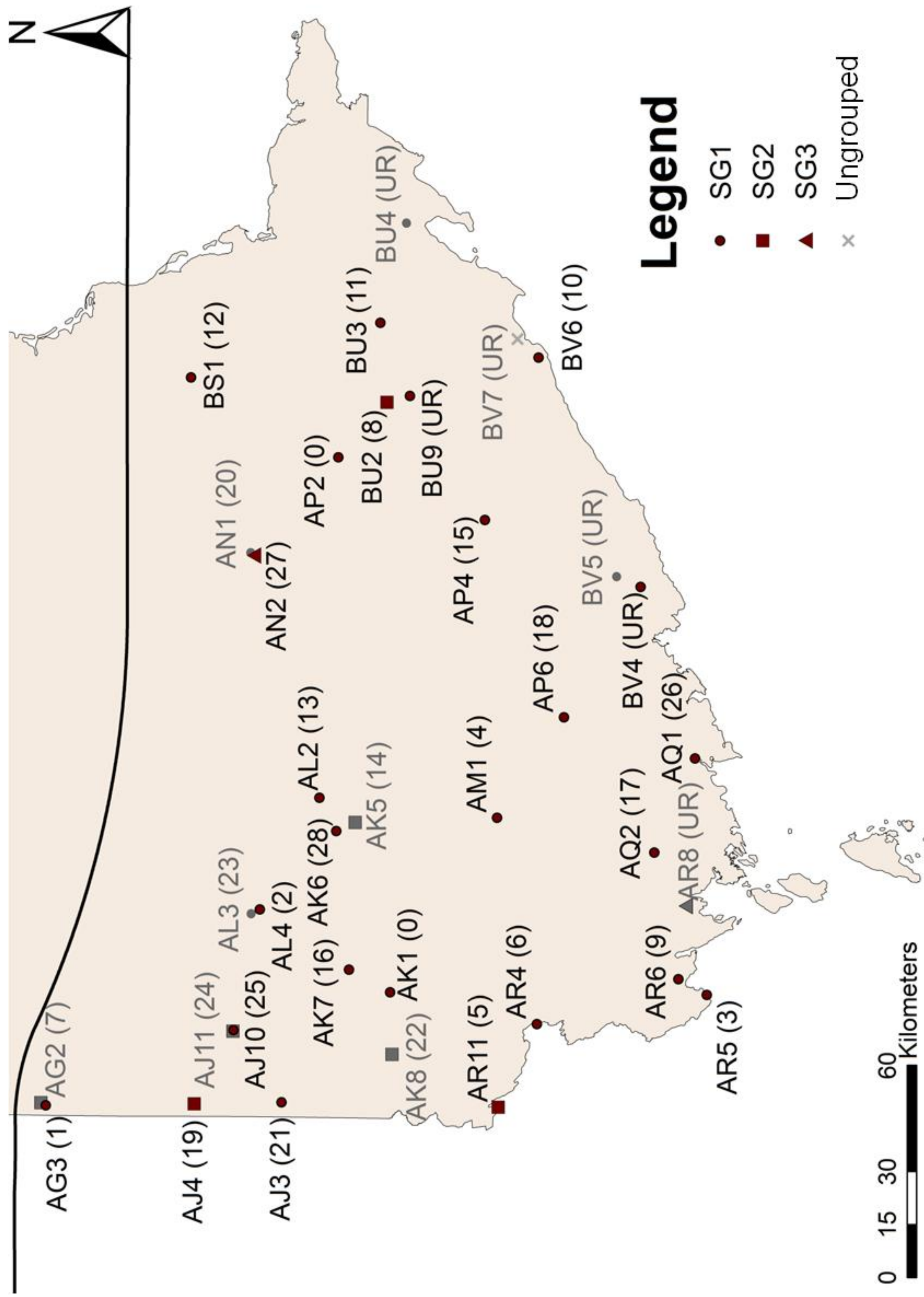

Figure 5b. Map of gauging stations (South), as well as their group and rank. Inactive stations are shown in gray.

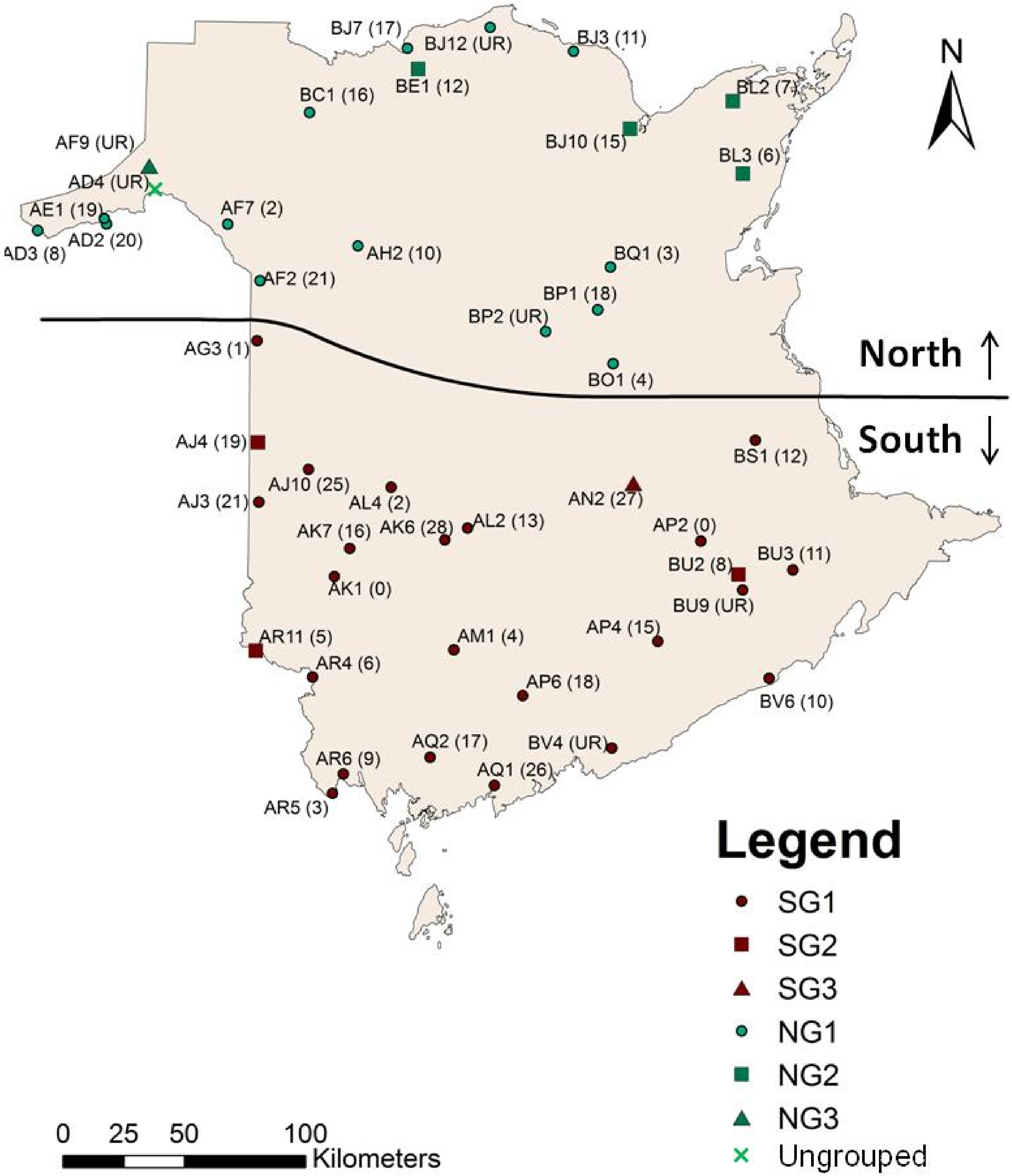

Figure 6. Ranks of the stations of the current network (only active stations).

4.3.4 Stations excluded from Entropy

Of the 67 stations initially identified as being part of the New Brunswick network of hydrometric gauging stations, only 53 were analyzed by the entropy method. The remaining 14 stations must also be dealt with. These stations are listed in Table 4.

Table 4. Stations excluded from the Entropy analysis.

Station N°	Station Name	Active	Record length (years)	Drainage Area (Km^2)	Mean Annual Flow (m^3/s)
AD4	SAINT JOHN RIVER AT EDMONSTON	Yes	46	15500	200.41
AF9	IROQUOIS RIVER AT MOULIN MORNEAULT	Yes	21	182	4.09
AF10	GREEN RIVER AT DEUXIEME SAULT	No	16	1030	28.65
AR8	BOCABEC RIVER ABOVE TIDE	No	14	43	1.10
BJ4	EEL RIVER NEAR EEL RIVER CROSSING	No	17	88.6	2.08
BJ12	EEL RIVER NEAR DUNDEE	Yes	29	43.2	0.94
BK3	NEPISIGUIT RIVER AT NEPISIGUIT FALLS	No	31	1840	33.92
BK4	NEPISIGUIT RIVER NEAR PABINEAU FALLS	No	18	2090	45.09
BP2	CATAMARAN BROOK AT REPAP ROAD BRIDGE	Yes	24	28.7	0.64
BU4	PALMERS CREEK NEAR DORCHESTER	No	20	34.2	0.92
BU9	HOLMES BROOK SITE NO.9 NEAR PETITCODIAC	Yes	17	6.2	0.12
BV4	BLACK RIVER AT GARNET SETTLEMENT	Yes	52	40.4	1.32
BV5	RATCLIFFE BROOK BELOW OTTER LAKE	No	12	29.3	0.99
BV7	UPPER SALMON RIVER AT ALMA	No	13	181	7.28

Many of the stations listed in Table 4 are already inactive. No reasoning or analysis will be applied to these stations, since it is assumed that they will not be reactivated. This leaves six stations that need to be dealt with. Stations AD4 and BV5 have long record lengths (46 and 52 years respectively) and therefore should be kept, since such a long record length is not common in the province. Station AF9 is part of NG3, which is a small group, and is the only member of this group in the northwest of New Brunswick. It may be wise to keep AF9, particularly if other stations of this group are already being removed. Station BJ12 is unremarkable and is located near BJ3, BJ4 and BJ7. Therefore it could be removed, if these stations are being kept. Station BP2 has a small drainage area (28.7 km^2) and a reasonably long record length (24 years). It is also near the center of the province, where there seems to be a lack of gauging stations (see Figure 2). It is otherwise unremarkable and there are other stations near it. It can be kept or removed, depending on what other nearby stations are being removed. Very similar reasoning and conclusions can be applied to station BU9.

5. Conclusion

Water management requires an optimal hydrometric network, as shown by the growing interest for hydrometric network evaluation and rationalization, in order to address challenges ahead in monitoring and data collection network stations. The present study provides a contribution to support decision makers, like data users and monitoring networks managers, in the process of selecting optimal representative stations for New Brunswick hydrometric network. Davar and Brimley (1990) proposed the first ranked prioritization of NB hydrometric network stations based on an audit approach, recommending the addition of stations to complete the hydrometric gauging network. More recently, Mishra and Coulibaly (2010) gave an overview using the transinformation index which included New Brunswick as a whole, giving an idea of the priority of each station. Coulibaly et al. (2013) also compared Canada to the WMO guidelines and found that most of the territory, including many parts of New Brunswick, are deficient when it comes to hydrometric gauging stations.

The present study proceeded by first dividing New Brunswick into two groups, using clustering analysis based on high flow timing. This had the effect of creating a mostly north-south division. However, this division is not a perfectly horizontal line dividing the north and the south, seeing as some northern stations had high flow timings similar to southern stations, and vice-versa. Principal component analysis was then used on both the North Group and South Group separately, but the results were inconclusive. It was then suggested to use the GEV shape parameter (maximum annual flow series) to split each group into three sub-groups. The purpose of these divisions was to avoid suggesting the complete or majority removal of stations from a single homogenous group, since removing a few stations of each group would be preferable. Finally, an entropy analysis was done to quantify the amount of information that was redundant at each station, thereby quantifying the importance of each station, based on its measurement of unique information. This allowed the ranking of each station in order of importance, which in turn allows the prioritization of stations, thus allowing the removal or displacement of the proper stations that would allow for a more optimal network. Some reasoning and analysis was done regarding the stations that did not meet the criteria for entropy analysis to better judge whether or not they are important.

With the selection of the more essential stations, with a good spatial repartition and a variety in the data they collect, the optimized reduced network can be more efficient for monitoring and data collection than if no optimization were done. Indeed, the optimal network suggested was designed taking into account regional climatic homogeneity (Burn 1997), similarity in hydrologic information between stations (Daigle et al. 2011; Morin et al. 1979), and availability of maximum information at each station with minimum dependency between them (Alfonso et al. 2013; Li et al. 2012). However, it is important to also take into account information about each station's worth using, for example, expert knowledge in order to make advised choices of an optimal network design (Hannah et al. 2011). For example, a statistically insignificant station according to the entropy analysis could in fact be very important because of its use in conjunction with a hydroelectric dam. Similar elements to this example can be helpful through consultations with data users and managers, in order to properly design a rationalized hydrometric network for NB.

6. Recommendation

As previously mentioned, it is not recommended to remove the majority or entirety of a subgroup. This is particularly the case for NG3 and SG3 as they are the subgroups with the least amount of stations, so removing even just a few can be the majority. It is instead preferable to remove some stations from each subgroup, as opposed to many from one subgroup.

Consideration should also be given to reactivating some of the more important station that have already been discontinued. This can be accomplished by removing a higher quantity of less important stations than what is necessary, allowing some of those removed stations to be displaced to better locations.

It is recommended when choosing which stations to remove or displace that a separate evaluation be done using existing regional regression equations. An analysis of these regressions should be done to see how they would be affected if a few selected stations were to be removed from the computation. This can give additional insight as to whether or not a station should be removed or kept.

Acknowledgements

This study was funded by the New Brunswick Environmental Trust Fund. The authors remain thankful to Mr. Darryl Pupek and Dr. Don Fox from the Department of Environment New Brunswick for their support.

References

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

Burn, D.H. 1997. Hydrological information for sustainable development. *Hydrological Sciences Journal* 42(4):481-492. DOI: 10.1080/02626669709492048

Burn, D.H., and I. C. Goulter. 1991. An approach to the rationalization of streamflow data collection networks. *Journal of Hydrology* 122:71-91.

Coulibaly, P., J. Samuel, A. Pietroniro, and D. Harvey. 2013. Evaluation of Canadian national hydrometric network density based on WMO 2008 standards. *Canadian Water Resources Journal* 38(2):159-167. DOI: 10.1080/07011784.2013.787181

Daigle, A., A. St-Hilaire, D. Beveridge, D. Caissie, and L. Benyahya. 2011. Multivariate analysis of the low-flow regimes in eastern Canadian rivers. *Hydrological Sciences Journal* 56(1):51-67. DOI: 10.1080/02626667.2010.535002

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

EL-Jabi, N., Turkkan, N., and D. Caissie. 2015. Characterisation of Natural Flow Regimes and Environmental Flows Evaluation in New Brunswick. *New Brunswick Environmental Trust Fund.*

Environment Canada and New Brunswick Dept. of Municipal Affairs and Envir. 1988. New Brunswick hydrometric network evaluation. Dartmouth, Nova Scotia, Canada

Hannah, D. M., Demuth, S., van Lanen, H. A. J., Looser, U., Prudhomme, C., Kerstin, S., and Tallaksen, L. M. 2011. *Hydrol. Process*. 25:1191-1200.

Husain, T. 1989. Hydrologic uncertainty measure and network design. *Water Resources Bulletin* 25(3):527-534.

Hussain, T. 1987. Hydrologic network design formulation. *Canadian Water Resources Journal* 12(1):44-63. DOI: 10.4296/cwrj1201044

Khalil, B., and T.B.M.J Ouarda. 2009. Statistical approaches used to assess and redesign surface water-quality-monitoring networks. *J.Environ.Monit* 11:1915-1929.

Khalil, B., Ouarda, T.B.M.J., and A. St-Hilaire. 2011. A statistical approach for the assessment and redesign of the Nile Delta drainage system water-quality-monitoring locations. *J. Environ. Monit.* 13:2190.

Li, C., V.P. Singh, and A.K. Mishra. 2012. Entropy theory-based criterion for hydrometric network evaluation and design: maximum information minimum redundancy. *Water Resources Research* 48, W05521. DOI: 10.1029/2011WR011251.

Mishra, A.K., and P. Coulibaly. 2009. Developments in hydrometric network design: a review. *Reviews of Geophysics* 47 (RG2001): 1-24. DOI: 10.1029/2007RG000243.

Mishra, A.K., and P. Coulibaly. 2010. Hydrometric network evaluation for Canadian watersheds. *Journal of hydrology* 380:420-437.

Morin, G., J.P. Fortin, W. Sochanska, J.P. Lardeau, and R. Charbonneau. 1979. Use of principal component analysis to identity homogenous precipitation stations for optimal interpolation. *Water Resources Research*. 15(6): 1841-1850. DOI: 10.1029/WR015i006p01841.

Ouarda et al. 1996. Rationalisation du réseau hydrométrique de la province de Québec pour le suivi des changements climatiques. *INRS-Eau*. 476:1-80.

Pilon, P.J., T.J. Day, T.R. Yuzyk, and R.A. Hale. 1996. Challenges facing surface water monitoring in Canada. *Canadian Water Resources Journal* 21: 157-164. DOI: 10.4296/cwrj210157

R Development Core Team. 2015. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. http://www.R-project.org.

Shannon, C.E. 1948. A mathematical theory of communication. *Bell Syst. Tech. J.* 27:379-423, 623-656.

Singh, V. P. 1997. The use of entropy in hydrology and water resources. *Hydrological Processes* 11:587-626.

Westra, S., C. Brown, U. Lall, and A. Sharma. 2007. Modeling multivariate hydrological series: Principal component analysis or independent component analysis? *Water Resources Research* 43, W06429. DOI: 10.1029/2006WR005617

Yeh, H-C., Y-C. Chen, C. Wei, R-H. Chen. 2011. Entropy and kriging approach to rainfall network design. *Paddy Water Environ 9:343-355*. DOI: 10.1007/sl0333-010-0247-x

Van Groenwoud, H.V. 1984. The climatic regions of New Brunswick : a multivariate analysis of meteorological data. *Can. J. For. Res.* 14: 389-394.
dfg

Appendix

Table A1. Hydrometric monitoring stations in New Brunswick

Station N°	Station Name	Latitude	Longitude	Start	End	Active	Drainage Area (Km^2)	Mean Annual Flow (m^3/s)
01AD002	SAINT JOHN RIVER AT FORT KENT	47 15 29	68 35 45	1926	2012	Yes	14700	278.412
01AD003	ST. FRANCIS RIVER AT OUTLET OF GLASIER LAKE	47 12 23	68 57 20	1951	2013	Yes	1350	24.353
01AD004	SAINT JOHN RIVER AT EDMONSTON	47 21 38	68 19 29	1968	2013	Yes	15500	200.412
01AE001	FISH KENT NEAR FORT KENT	47 14 15	68 34 58	1981	2013	Yes	2260	44.615
01AF002	SAINT JOHN RIVER AT GRAND FALLS	47 02 20	67 44 23	1930	2012	Yes	21900	412.517
01AF003	GREEN RIVER NEAR RIVIERE-VERTE	47 20 06	68 08 06	1962	1993	No	1150	26.296
01AF007	GRANDE RIVIERE AT VIOLETTE BRIDGE	47 14 49	67 55 16	1977	2012	Yes	339	7.023
01AF009	IROQUOIS RIVER AT MOULIN MORNEAULT	47 27 28	68 21 24	1991	2011	Yes	182	4.087
01AF010	GREEN RIVER AT DEUXIEME SAULT	47 28 14	68 14 08	1995	2010	No	1030	28.650
01AG002	LIMESTONE RIVER AT FOUR FALLS	46 49 42	67 44 35	1967	1993	No	199	3.646
01AG003	AROOSTOOK RIVER NEAR TINKER	46 48 58	67 45 07	1975	2012	Yes	6060	109.908
01AH002	TOBIQUE RIVER AT RILEY BROOK	47 10 22	67 12 38	1954	2011	Yes	2230	47.689
01AH005	MAMOZEKEL RIVER NEAR CAMPBELL RIVER	47 15 03	67 08 32	1972	1990	No	230	4.061
01AJ003	MEDUXNEKEAG RIVER NEAR	46 12 58	67 43 40	1967	2012	Yes	1210	24.567

Table A1. Hydrometric monitoring stations in New Brunswick

Station N°	Station Name	Latitude	Longitude	Start	End	Active	Drainage Area (Km^2)	Mean Annual Flow (m^3/s)
	BELLEVILLE							
01AJ004	BIG PRESQUE ISLE STREAM AT TRACEY MILLS	46 26 18	67 44 18	1967	2011	Yes	484	9.757
01AJ010	BECAGUIMEC STREAM AT COLDSTREAM	46 20 27	67 27 54	1973	2011	Yes	350	7.430
01AJ011	COLDSTREAM AT COLDSTREAM	46 20 32	67 28 09	1973	1993	No	156	3.182
01AK001	SHOGOMOC STREAM NEAR TRANS CANADA HIGHWAY	45 56 36	67 19 13	1918	2012	Yes	234	4.917
01AK005	MIDDLE BRANCH NASHWAAKSIS STREAM NEAR ROYAL ROAD	46 02 06	66 42 05	1965	1993	No	26.9	0.536
01AK006	MIDDLE BRANCH NASHWAAKSIS STREAM AT SANDWITH'S FARM	46 04 58	66 43 58	1966	2011	Yes	5.7	0.101
01AK007	NACKAWIC STREAM NEAR TEMPERANCE VALE	46 02 55	67 14 22	1967	2011	Yes	240	4.899
01AK008	EEL RIVER NEAR SCOTT SIDING	45 56 12	67 32 49	1974	1993	No	531	10.503
01AL002	NASHWAAK RIVER AT DURHAM BRIDGE	46 07 33	66 36 40	1962	2012	Yes	1450	35.572
01AL003	HAYDEN BROOK NEAR NARROWS MOUNTAIN	46 17 56	67 02 13	1970	1993	No	6.48	0.176
01AL004	NARROWS MOUNTAIN BROOK NEAR NARROWS MOUNTAIN	46 16 37	67 01 17	1972	2011	Yes	3.89	0.227
01AM001	NORTH BRANCH OROMOCTO RIVER AT TRACY	45 40 25	66 40 58	1962	2011	Yes	557	12.187
01AN001	CASTAWAY STREAM NEAR CASTAWAY	46 17 54	65 42 43	1972	1993	No	34.4	0.872
01AN002	SALMON RIVER AT CASTAWAY	46 17 26	65 43 21	1974	2012	Yes	1050	21.577
01AP002	CANAAN RIVER AT EAST	46 04 20	65 21 59	1925	2011	Yes	668	13.250

Table A1. Hydrometric monitoring stations in New Brunswick

Station N°	Station Name	Latitude	Longitude	Start	End	Active	Drainage Area (Km^2)	Mean Annual Flow (m^3/s)
	CANAAN							
01AP004	KENNEBECASIS RIVER AT APOHAQUI	45 42 05	65 36 06	1961	2011	Yes	1100	25.258
01AP006	NEREPIS RIVER NEAR FOWLERS CORNER	45 30 12	66 19 08	1976	2011	Yes	293	6.680
01AQ001	LEPREAU RIVER AT LEPREAU	45 10 11	66 28 05	1916	2013	Yes	239	7.209
01AQ002	MAGAGUADAVIC RIVER AT ELMCROFT	45 16 24	66 48 24	1917	2013	Yes	1420	32.935
01AR004	ST. CROIX RIVER AT VANCEBORO	45 34 08	67 25 47	1928	2013	Yes	1080	20.914
01AR005	ST. CROIX RIVER AT BARING	45 08 12	67 19 05	1975	2013	Yes	3550	74.965
01AR006	DENNIS STREAM NEAR ST. STEPHEN	45 12 35	67 15 45	1966	2012	Yes	115	2.740
01AR008	BOCABEC RIVER ABOVE TIDE	45 11 35	66 59 56	1966	1979	No	43	1.096
01AR011	FOREST CITY STREAM BELOW FOREST CITY DAM	45 39 51	67 44 00	1975	2013	Yes	357	15.479
01BC001	RESTIGOUCHE RIVER BELOW KEDGWICK RIVER	47 40 01	67 28 59	1962	2012	Yes	3160	66.282
01BE001	UPSALQUITCH RIVER AT UPSALQUITCH	47 49 56	66 53 13	1918	2012	Yes	2270	40.144
01BJ001	TETAGOUCHE RIVER NEAR WEST BATHURST	47 39 21	65 41 37	1922	1995	No	363	7.674
01BJ003	JACQUET RIVER NEAR DURHAM CENTRE	47 53 52	66 01 47	1964	2012	Yes	510	10.225
01BJ004	EEL RIVER NEAR EEL RIVER CROSSING	48 00 52	66 26 18	1967	1983	No	88.6	2.079
01BJ007	RESTIGOUCHE RIVER ABOVE RAFTING GROUND BROOK	47 54 31	66 56 53	1968	2012	Yes	7740	155.840
01BJ010	MIDDLE RIVER NEAR	47 36 30	65 43 18	1981	2012	Yes	217	4.266

Table A1. Hydrometric monitoring stations in New Brunswick

Station N°	Station Name	Latitude	Longitude	Start	End	Active	Drainage Area (Km^2)	Mean Annual Flow (m^3/s)
	BATHURST							
01BJ012	EEL RIVER NEAR DUNDEE	47 59 16	66 29 26	1984	2012	Yes	43.2	0.935
01BK003	NEPISIGUIT RIVER AT NEPISIGUIT FALLS	47 24 24	65 47 42	1921	2005	No	1840	33.921
01BK004	NEPISIGUIT RIVER NEAR PABINEAU FALLS	47 29 40	65 40 50	1957	1974	No	2090	45.089
01BL001	BASS RIVER AT BASS RIVER	47 39 00	65 34 40	1965	1991	No	175	3.156
01BL002	RIVIERE CARAQUET AT BURNSVILLE	47 42 20	65 09 19	1969	2012	Yes	173	3.510
01BL003	BIG TRACADIE RIVER AT MURCHY BRIDGE CROSSING	47 26 08	65 06 25	1970	2012	Yes	383	7.994
01BO001	SOUTHWEST MIRAMICHI RIVER AT BLACKVILLE	46 44 09	65 49 32	1918	2012	Yes	5050	114.668
01BO002	RENOUS RIVER AT McGRAW BROOK	46 49 17	66 06 53	1965	1995	No	611	14.649
01BO003	BARNABY RIVER BELOW SEMIWAGAN RIVER	46 53 19	65 35 44	1973	1995	No	484	9.681
01BP001	LITTLE SOUTHWEST MIRAMICHI RIVER AT LYTTLETON	46 56 09	65 54 26	1951	2012	Yes	1340	32.001
01BP002	CATAMARAN BROOK AT REPAP ROAD BRIDGE	46 51 23	66 11 24	1989	2012	Yes	28.7	0.641
01BQ001	NORTHWEST MIRAMICHI RIVER AT TROUT BROOK	47 05 41	65 50 11	1961	2012	Yes	948	20.857
01BR001	KOUCHIBOUGUAC RIVER NEAR VAUTOUR	46 44 36	65 12 17	1930	1995	No	177	3.746
01BS001	COAL BRANCH RIVER AT BEERSVILLE	46 26 38	65 03 53	1964	2011	Yes	166	3.678
01BU002	PETITCODIAC RIVER NEAR PETIT CODIAC	45 56 47	65 10 05	1961	2011	Yes	391	7.933

Table A1. Hydrometric monitoring stations in New Brunswick

Station N°	Station Name	Latitude	Longitude	Start	End	Active	Drainage Area (Km^2)	Mean Annual Flow (m^3/s)
01BU003	TURTLE CREEK AT TURTLE CREEK	45 57 34	64 52 40	1962	2010	Yes	129	3.622
01BU004	PALMERS CREEK NEAR DORCHESTER	45 53 14	64 30 59	1966	1985	No	34.2	0.921
01BU009	HOLMES BROOK SITE NO.9 NEAR PETITCODIAC	45 53 16	65 08 48	1995	2011	Yes	6.2	0.117
01BV004	BLACK RIVER AT GARNET SETTLEMENT	45 18 23	65 50 57	1960	2011	Yes	40.4	1.321
01BV005	RATCLIFFE BROOK BELOW OTTER LAKE	45 22 04	65 48 42	1960	1971	No	29.3	0.993
01BV006	POINT WOLFE RIVER AT FUNDY NATIONAL PARK	45 33 30	65 00 57	1964	2011	Yes	130	5.019
01BV007	UPPER SALMON RIVER AT ALMA	45 36 40	64 57 22	1967	1979	No	181	7.277

New Nouveau Brunswick
ENVIRONMENTAL TRUST FUND

UNIVERSITÉ DE MONCTON
EDMUNDSTON MONCTON SHIPPAGAN

Pêches et Océans Canada
Fisheries and Oceans Canada