Université du Québec
Institut National de la Recherche Scientifique
centre Eau, Terre Environnement

Étude des variables hydrologiques dans un cadre multivarié et dans un contexte de changement

Par

Mohamed Aymen Ben Aissia

Thèse présentée pour l'obtention du grade de
Philosophiae doctor (Ph.D.) en sciences de l'eau

Jury d'évaluation

Examinateur externe	Ali A. Assani Université du Québec à Trois-Rivières
Examinateur externe	Musandji Fuamba École polytechnique de Montréal
Examinateur interne	André St-Hilaire INRS-ETE
Codirecteur de recherche	Taha B.M.J. Ouarda INRS- ETE & Masdar Institute of Science and Technology
Directeur de recherche	Fateh Chebana INRS-ETE

REMERCIEMENTS

Cette thèse est le fruit de plusieurs années de travail durant lesquelles il m'a été donné de recevoir un grand support de plusieurs personnes. Je voudrais tout d'abord remercier mon directeur Fateh Chebana et mon codirecteur Taha Ouarda sans qui cette thèse n'aurait pu être accomplie. J'ai été extrêmement privilégié puisqu'ils m'ont tous deux supervisé de manière très complémentaire et attentive. Je désire donc remercier Fateh, mon directeur et mon mentor, d'avoir dirigé mon travail de thèse et pour m'avoir aidé à me préparer pour ma carrière. Tu m'as poussé à dépasser mes limites. Je suis conscient que si je suis plus fort et plus confiant aujourd'hui, tu es en grande partie responsable. Je remercie également mon codirecteur Taha pour sa confiance, la plupart du temps plus grande que ma propre confiance, et pour m'avoir donné des conditions de travail et des opportunités exceptionnelles. Je tiens également à te remercier pour ta patience, ta compréhension et tes encouragements. Mon expérience au doctorat a été tellement enrichie par ta présence. J'espère sincèrement, Taha et Fateh, pouvoir un jour vous rendre la pareille.

J'aimerais remercier également Hydro-Québec et le Conseil de Recherches en Sciences Naturelles et en Génie (CRSNG) pour leur financement partiel de ma thèse.

Je suis également redevable à plusieurs collègues de travail : Jean-Xavier Giroux, Christian Charron et Martin Durocher pour leur collaboration à mes travaux. Merci à ma femme, Rimeh, pour sa patience et sa compréhension, ma fille Acil, à toute ma famille, et à mes amis de l'INRS.

Enfin, j'aimerais remercier spécialement ma mère, pour son amour inconditionnel, sa confiance en moi, et le support moral qu'elle m'a accordés tout au long de ma vie. C'est à ma mère que je dédie cette thèse.

AVANT-PROPOS

Cette thèse présente les travaux de recherche menés au cours de mes études doctorales. La structure de la présente thèse suit la structure standard des thèses par articles de l'INRS-ETE. La première partie de la thèse comporte une synthèse générale des travaux effectués. Cette synthèse a pour objectif de survoler la méthodologie adoptée et les principaux résultats obtenus au cours de la thèse. La deuxième partie contient trois articles acceptés, soumis ou sur le point d'être soumis à des revues internationales avec comité de lecture ainsi qu'un rapport qui sera déposé au service de documentation de l'INRS-ETE.

RÉSUMÉ

Les études portant sur les évènements hydrologiques sont cruciales vu leurs nombreux impacts sur l'économie et l'environnement. Ces études nécessitent l'analyse et la modélisation des variables hydrologiques décrivant l'évènement en question. Généralement, les événements hydrologiques sont caractérisés par différentes variables corrélées, par exemple volume, pointe et durée de crue, intensité, durée et sévérité des tempêtes et durée et ampleur des sécheresses. Par conséquent, l'analyse de chacune de ces variables séparément (univariée) ne peut pas fournir une évaluation complète et adéquate des risques pour ces événements. D'autre part, plusieurs études ont rapporté des changements temporels des caractéristiques du régime hydrologique des cours d'eau. Les changements climatiques peuvent être à l'origine de tels changements.

Dans cette thèse, l'objectif général visé est de déterminer, adapter et appliquer les méthodes les plus prometteuses pour l'analyse et la modélisation, dans des cas pratiques, des variables hydrologiques à la fois dans un cadre multivarié et dans un contexte de changement climatique. Plus précisément, la présente thèse vise à combler un vide pour certaines des étapes (imputation des données manquantes et détection des points de rupture) et éléments manquants (étude régionale et évolution spatiale) pour une analyse fréquentielle multivariée complète des variables hydrologiques. L'analyse fréquentielle est un des plus importants outils statistiques qui englobe l'analyse et la modélisation des variables étudiées. Par conséquent, l'analyse fréquentielle multivariée a été retenue comme le cadre général de cette thèse. Par ailleurs, la dépendance entre les variables hydrologiques représente un aspect très important quand on considère des séries multivariées. L'étude de l'évolution temporelle et spatiale de la dépendance sur des séries hydrologiques simulées et observées est, alors, d'une utilité majeure.

Dans la première étude de la thèse, une revue bibliographique sur les techniques d'imputation des données manquantes dans des séries multivariées a été réalisée. À partir de cette revue, on a déterminé les techniques les plus prometteuses qui ont été comparées et appliquées sur des cas réels. La comparaison entre les techniques a été réalisée par l'évaluation de la performance de chaque technique par Jackknife. Les résultats montrent que la dépendance entre les composantes des séries multivariées joue un rôle important dans le choix et la performance des techniques d'imputation.

Ensuite, une étude de détermination, adaptation, comparaison et application des tests de détection des points de rupture dans un cadre multivarié a été réalisée. Les tests considérés sont basés sur les fonctions de profondeur. Ils ont été comparés par une étude de simulation et puis appliqués sur deux cas d'étude. Les résultats montrent que les performances des tests considérés sont, en général, influencés par divers facteurs, comme la taille de l'échantillon et l'amplitude du changement. Certains tests sont problématiques en cas d'échantillon de petite taille et d'autres sont conseillés pour des fonctions de profondeur spécifiques.

Par la suite, l'évolution temporelle de la dépendance entre la pointe, le volume et la durée de crue a été étudiée dans un contexte de changement climatique. Cette étude a été réalisée sur des séries observées, résultats de réanalyse et simulées de deux bassins versants dans la province de Québec, Canada : soit le réservoir Baskatong et la rivière Romaine et ce, en considérant HYDROTEL et HSAMI comme modèles hydrologiques. Les résultats montrent que le comportement des mesures de dépendance varie en fonction du modèle hydrologique, du modèle général du climat et du scénario d'augmentation des gaz à effet de serre. De plus, le comportement des séries issues de tau de Kendall et de rho de Spearman est similaire.

Enfin, une approche développée récemment en analyse fréquentielle régionale dans un cadre multivarié a été appliquée sur la région de côte Nord dans la province de Québec, Canada. Cette approche, comme extension du modèle de l'indice de crue, est basée sur les copules et les quantiles multivariés. Dans ce travail, l'accent est mis sur les aspects pratiques de cette approche en l'appliquant sur la pointe et le volume de crue. Les résultats montrent que l'approche proposée performe bien sur les cas d'études qu'on a considérés et que l'erreur du modèle dépend aussi bien du risque, de la discordance bivariée que de la discordance univariée.

ARTICLES, RAPPORT ET CONTRIBUTION DE CHAQUE AUTEUR

Ben Aissia, M.A., F. Chebana, T. B. M. J. Ouarda (2014b) : Missing data imputation in multivariate hydrological frequency analysis. À soumettre.

Ben Aissia, M.A., F. Chebana, T. B. M. J. Ouarda (2014d) : Détection des points de rupture multivariés en hydrologie. Rapport de recherche à déposer.

Ben Aissia, M.A., F. Chebana, T. B. M. J. Ouarda, L. Roy, P. Bruneau and M. Barbet (2014a) : Dependence evolution of the main flood characteristics in a context of climate change. *Journal of Hydrology*. Accepté

Ben Aissia, M.A., F. Chebana, T. B. M. J. Ouarda, P. Bruneau and M. Barbet (2014c) : Bivariate index-flood model for a northern case study in Quebec, Canada. *Hydrological Sciences Journal*. Accepté

Dans le premier article, M.A. Ben Aissia a effectué la revue de littérature sur les techniques d'imputation des données manquantes et le choix des méthodes à inclure dans la comparaison. L'étude de comparaison, l'application au cas réel et la rédaction de l'article ont été réalisées par M.A. Ben Aissia. F. Chebana et T.B.M.J. Ouarda ont révisé la version finale du manuscrit. L'idée originale provient de T.B.M.J. Ouarda, F. Chebana et M.A. Ben Aissia.

Dans le rapport, le choix des tests considérés ainsi que l'étude de simulation ont été faits par F. Chebana. M.A. Ben Aissia a vérifié et complété les simulations, a effectué l'étude de comparaison et l'application et a rédigé le rapport. F. Chebana et T.B.M.J. Ouarda ont révisé le rapport. L'idée originale provient de F. Chebana.

Dans le deuxième article, M.A. Ben Aissia a réalisé l'étude d'évolution de la dépendance et a rédigé l'article. F. Chebana, T.B.M.J. Ouarda, L. Roy, P. Bruneau et M. Barbet ont fourni leurs commentaires durant l'exécution du travail. F. Chebana et T.B.M.J. Ouarda ont révisé le manuscrit. L'idée originale provient de F. Chebana et T.B.M.J. Ouarda.

Dans le troisième article, M.A. Ben Aissia a réalisé l'étude d'analyse fréquentielle régionale multivariée et a rédigé l'article. F. Chebana, T.B.M.J. Ouarda, P. Bruneau et M. Barbet ont fourni leurs commentaires durant l'exécution du travail. F. Chebana et T.B.M.J. Ouarda ont révisé le manuscrit. L'idée originale provient de F. Chebana et T.B.M.J. Ouarda.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

PARTIE 1 : SYNTHÈSE

1 Introduction

1.1 Contexte

Les événements extrêmes, tels que les crues, les ouragans et les sécheresses ont des conséquences économiques, environnementales et sociales graves. L'estimation adéquate de l'occurrence de ces événements est primordiale en raison des risques associés. L'analyse fréquentielle (AF) est un des outils fondamentaux pour étudier les événements extrêmes. L'AF est un ensemble de méthodes statistiques qui ont comme objectif principal d'estimer les probabilités d'occurrence en utilisant les mesures d'événements passés.

En hydrologie, les événements extrêmes sont, généralement, caractérisés par plusieurs variables corrélées. Par exemple, les crues sont décrites par leur volume, pointe et durée (e.g. Ashklar 1980; Yue et al. 1999; Ouarda et al. 2000; Yue 2001; De Michele et al. 2005; Chebana et Ouarda (2009a, 2009b); Ben Aissia et al. 2011; Chebana et Ouarda 2011a). Ces études ont souligné l'importance de considérer ces variables conjointement. En effet, une AF univariée n'est pas en mesure de fournir une évaluation complète de la probabilité d'occurrence de l'événement considéré et peut engendrer des pertes de vies humaines ou de biens associés à une sous-estimation, ou une augmentation de coût de la construction associée à une surestimation.

Une AF est composée de quatre étapes principales : a) analyse descriptive et explicative des données ainsi que la détection des valeurs aberrantes, b) vérification des

hypothèses de base, y compris la stationnarité, l'homogénéité et l'indépendance, c) modélisation et estimation et d) analyse et évaluation des risques. Dans un cadre univarié, ces étapes ont été largement considérées (e.g. Rao et Hamed 2000). Dans un contexte multivarié, les deux étapes préliminaires (a et b) ont attiré beaucoup moins d'attention que les deux dernières (Chebana et Ouarda 2011b). Les deux étapes préliminaires ont un impact significatif sur le choix du modèle approprié dans une AF multivariée. Par exemple, une non-stationnarité dans les données peut engendrer des tendances possibles dans certaines ou toutes les parties de la distribution multivariée (copule et marginales). Par conséquent, ignorer certaines étapes de l'AF peut conduire à un modèle inadéquat et donc à des résultats erronés.

La dépendance entre les variables est l'élément clé dans une étude multivariée. Bien que la dépendance entre les principales caractéristiques de crue soit largement étudiée (pour une station donnée et sur une période historique fixe), l'évolution de cette dépendance dans le temps ou dans l'espace n'a pas été étudiée, en particulier dans le contexte de changement climatique pour des variables hydrologiques.

Par ailleurs, l'estimation des récurrences des crues à des sites d'intérêt où l'on dispose de peu ou même d'aucune information hydrologique (sites non jaugés) s'avère souvent nécessaire. Pour remédier à ce problème, on peut avoir recours à des méthodes de régionalisation des débits extrêmes qui permettent d'utiliser les données disponibles à d'autres sites de la même région que le site non jaugé. L'AF régionale dans un cadre univarié a été largement étudiée (e.g. Stedinger et Tasker 1986; Hosking et Wallis 1993; Ouarda et al. 2001). Récemment Chebana et Ouarda (2009b) ont développé une

méthodologie d'AF bivariée régionale basée sur le modèle d'indice de crue. Cependant, cette méthodologie n'a pas été testée sur un cas réel.

1.2 Problématique

L'étude des variables hydrologiques a fait l'objet de nombreux travaux de recherches existantes en littérature. Cependant, ces variables sont généralement considérées dans un cadre univarié. Récemment, avec l'utilisation des techniques multivariées, plusieurs travaux ont considéré l'étude conjointe de ces variables hydrologiques (e.g. Little et Rubin 2002). L'AF est une des méthodes les plus utilisées lorsqu'il s'agit d'étudier les évènements extrêmes en hydrologie. Elle permet d'estimer l'occurrence d'un évènement en modélisant les variables qui le décrivent. L'AF de crue a été retenue comme le cadre général de cette thèse.

Dans la littérature, plusieurs méthodes d'AF ont été développées dans le cadre univarié (voir e.g. Cunnane 1987; Rao et Hamed 2000). Cependant, les crues sont caractérisées par plusieurs variables (principalement la pointe Q, le volume V et la durée D) et une estimation adéquate des récurrences des crues nécessite la connaissance des informations complètes sur cet événement. D'où la nécessité des techniques de modélisation multivariée qui permettent de considérer les caractéristiques de crue conjointement.

Les études des événements hydrologiques extrêmes dans un contexte multivarié sont de plus en plus nombreuses (e.g. Ashklar 1980; Ouarda et al. 2000; De Michele et al. 2005; Salvadori et De Michele 2010; Ben Aissia et al. 2011; Chebana et Ouarda 2011a; Chebana 2013; Volpi et Fiori 2014). Cependant, la littérature grandissante en AF multivariée hydrologique est orientée vers l'étape de la modélisation en négligeant les autres étapes préliminaires tel que souligné dans Chebana et al. (2013) et Chebana

(2013). La contribution de la présente thèse est de combler ce vide par la détermination, l'adaptation et l'application des techniques multivariées pour certaines de ces étapes et éléments manquants pour une AF multivariée complète.

Le traitement des données manquantes et les tests de rupture dans les séries hydrologiques multivariées n'ont jamais été considérés dans la littérature hydrologique malgré leurs disponibilités en statistique et leurs importances dans la modélisation hydrologique. Ces étapes pourraient avoir un impact significatif dans les choix du modèle d'AF approprié. L'analyse des données disponibles comporte, généralement, une analyse descriptive, la détection de points aberrants et l'imputation des données manquantes. Récemment Chebana et Ouarda (2011b) ont mis l'accent sur l'analyse descriptive des données et la détection des points aberrants dans un cadre multivarié. Cependant, à notre connaissance, l'imputation des données manquantes n'a pas encore été étudiée dans le cadre multivarié et dans le contexte d'AF hydrologique.

D'autre part, les études d'AF multivariées réalisées se sont concentrées sur les études locales (un seul site) et en considérant une période de temps fixe (historique). Du point de vue spatial, récemment, une nouvelle méthodologie d'AF régionale bivariée (plusieurs sites) a été développée par Chebana et Ouarda (2009b). Cette méthodologie représente la version multivariée du modèle d'indice de crue basé sur les copules et les quantiles multivariés. Cependant, cette méthodologie n'a pas été testée sur un cas réel. Dans la présente thèse, dans le but de traiter les aspects pratiques de la méthode, on l'applique sur un cas réel et on évalue ça performance en utilisant la méthode de jacknife.

En ce qui concerne l'aspect temporel, plusieurs études ont rapporté des changements temporels critiques des caractéristiques du régime hydrologique des cours d'eau (e.g. Lins et Slack 1999; Douglas et al. 2000; IPCC 2001; Zhang et al. 2001). En ignorant la présence des changements ou en considérant les variables séparément, les méthodes d'AF peuvent conduire à des sous-estimations ou surestimations. Ainsi, il devient de plus en plus important de considérer des outils statistiques multivariés permettant de tester l'existence d'un changement.

Parmi les changements possibles, on trouve les changements climatiques à long termes qui peuvent avoir des effets sur les événements hydrologiques des bassins versants au Québec. Les changements climatiques peuvent affecter les crues extrêmes en termes d'amplitude, de fréquence ou de durée. Ainsi la dépendance entre les variables caractérisant une crue peut également être affectée. Ainsi, une étude de l'évolution de la dépendance entre les principales caractéristiques de crues dans un contexte de changement climatique est très utile et n'a jamais été considérée.

Les quatre sujets traités dans cette thèse sont tous nouveaux, à différents degrés, en AF multivariée des variables hydrologiques. En effet, malgré leurs disponibilités en statistique et leurs importances dans la modélisation hydrologique, le traitement des données manquantes et les tests de rupture dans les séries hydrologiques multivariées n'ont jamais été considérés dans la littérature hydrologique. D'autre part, l'étude de l'évolution de la dépendance est également nouvelle étant donné que, dans la littérature, on considère la dépendance ponctuelle (une valeur d'un coefficient de corrélation ou un paramètre de copule pour toute la durée de l'étude) ou rarement la comparaison entre deux périodes fixes où on compare deux valeurs (e.g. Ben Aissia et al. 2011).

1.3 Objectifs

Les objectifs de la thèse de doctorat sont :

- Déterminer, adapter et appliquer les méthodes les plus prometteuses pour l'analyse et la modélisation des variables hydrologiques à la fois dans un contexte de changement et dans un cadre multivarié, qui permettront notamment :
  - L'étude de l'évolution de la dépendance en fonction du temps et de l'espace;
  - La détection des points de rupture en utilisant des tests basés sur les fonctions de profondeur;
  - L'étude d'imputation des données manquantes dans un cadre d'AF multivarié;
- Adapter une méthodologie d'AF régionale basée sur l'indice de crue sur un cas réel.
- Développer les aspects pratiques des méthodes utilisées dans toute la thèse et évaluer leurs performances.

1.4 Méthodologie

Pour répondre aux objectifs de ma thèse, une étude bibliographique sur les données manquantes a été réalisée. Dans cette étude, l'importance de traiter les données manquantes en AF multivariée est mise en évidence suivie par la comparaison entre les méthodes d'imputation univariées et multivariées les plus prometteuses (Ben Aissia et al. 2014b). Deux types d'applications multivariées sont effectués dans l'article pour différentes caractéristiques de crue et multi-site pour différentes stations.

Ensuite, afin de comprendre l'importance de la dépendance entre caractéristiques de crue, une étude sur l'évolution de la dépendance entre les principales caractéristiques de crue (Q, V et D) dans un contexte de changement climatique a été réalisée et a constitué l'article de Ben Aissia et al. (2014a). Dans cette partie, on s'intéresse à l'évolution temporelle continue de la dépendance entre les trois principales caractéristiques de crues (Q, V et D) en utilisant une fenêtre mobile de 30 ans. Trois mesures de dépendances sont utilisées soit la corrélation de Pearson, le tau de Kendall et le rho de Spearman. Deux bases de données provenant des deux bassins versant, i.e. Baskatong et Romaine, sont considérées. De plus, deux modèles hydrologiques (HSAMI et HYDROTEL), trois modèles climatiques (CGCM3, HADCM3 et ECHAM5) ainsi que trois scénarios d'augmentation de gaz à effet de serre (A2, B1 et B2) sont considérés dans les bases de données. L'étude de l'évolution de la dépendance comporte la détermination des bandes de confiance des séries de dépendance, l'étude de stationnarité et la détermination des points de rupture.

Par la suite, une étude de détection des points de rupture dans les données hydrologiques multivariés est réalisée (Ben Aissia et al. 2014d). Les tests considérés sont principalement basés sur la notion de fonction de la profondeur à l'exception d'un test qui est utilisé pour comparaison. Le but est d'étudier la détection des points de rupture dans un contexte hydrologique et dans un cadre multivarié (en considérant deux séries conjointement). Une étude de simulation qui tient compte des contraintes hydrologiques est effectuée pour comparer la puissance de tests considérés. De plus, une application est présentée pour montrer l'aspect pratique des tests considérés.

Finalement, une approche développée récemment (Chebana et al. 2009) en analyse fréquentielle régionale dans un cadre multivarié a été appliquée sur la région de la Côte Nord dans la province de Québec, Canada. Cette approche, une extension du modèle de l'indice de crue, est basée sur les copules et les quantiles multivariés. Les aspects pratiques de ce modèle sont considérés dans cette étude. Ces aspects comportent : les tests d'ajustement pour les copules et lois marginales, l'estimation des paramètres associés, l'estimation de l'indice de crue et l'interprétation des quantiles multivariés. De plus, une comparaison entre les différentes analyses fréquentielles (locales/régionales et univariées/multivariées) a été réalisée (Ben Aissia et al. 2014c).

Le choix de la Côte Nord dans la dernière étude est justifié principalement par le régime d'écoulement des stations existantes. En effet, dans la région de Côte Nord les stations ont un régime d'écoulement naturel au lieu d'un régime artificiel dans plusieurs stations au sud du Québec. Il y a aussi la disponibilité des stations et la diversité d'une station à l'autre de point de vue longueur des séries et caractéristiques des bassins

versants. L'intérêt à cette région peut être expliqué également par l'importance de son potentiel de production hydroélectrique (Romaine) et la biodiversité de l'habitat naturel.

1.5 Organisation de la synthèse

Le présent document est organisé comme suit : le chapitre 2 présente une revue de littérature générale sur l'étude des variables hydrologiques dans un cadre multivarié de l'AF hydrologique. Le chapitre 3 présente la méthodologie adoptée pour réaliser les objectifs de la thèse. Les principaux résultats obtenus sont présentés dans le chapitre 4. Enfin, la conclusion et les perspectives de recherches sont présentées dans le chapitre 5. Les articles et le rapport produits au cours de cette thèse sont présentés dans la deuxième partie.

2 Revue de littérature

Ce chapitre présente une revue de littérature générale sur l'étude des variables hydrologiques dans un cadre multivarié dans le contexte de l'AF. Il est constitué d'une brève revue de l'historique de l'AF hydrologique multivariée et une description des différentes étapes de l'AF en général.

2.1 Analyse fréquentielle hydrologique multivariée

L'AF régionale et locale est un ensemble d'outils statistiques couramment utilisés pour l'analyse des phénomènes hydrologiques extrêmes. Souvent, l'AF réalisée s'est concentrée sur une seule variable, comme la pointe ou le volume de crue, afin d'évaluer le risque associé. Cunnane (1987) et Rao et Hamed (2000) peuvent être consultés pour des études bibliographiques approfondies sur l'AF des crues. Cependant, un évènement extrême est décrit par plusieurs variables corrélées tel que la pointe, le volume et la durée de crue (see e.g. Chebana 2013). Par conséquent, l'AF nécessite la considération du risque conjoint de ces variables. Par exemple, le risque conjoint de Q et V représente un plus grand intérêt pour la construction d'un barrage que le risque de ces deux variables séparées.

Récemment, l'approche de l'AF multivariée hydrologique attire de plus en plus d'attention où on traite les variables conjointement (e.g. Ashklar 1980; Kelly et Krzysztofowicz 1997; Yue et al. 1999; Yue 2001; Chebana et al. 2009; Chebana et Ouarda 2009a; Ben Aissia et al. 2011). L'approche multivariée permet de tenir compte de la dépendance entre les différentes variables décrivant l'évènement telles que Q et V pour

les crues. Initialement, le modèle normal multivarié a été employé pour sa simplicité (distribution marginale normale), mais dans le cas des extrêmes, la loi normale n'est pas appropriée (e.g. Goel et al. 1998). Ensuite, d'autres distributions multivariées ont été proposées. Leur désavantage est que les distributions marginales doivent appartenir à la même famille (e.g Escalante-Sandoval et Raynal-Villaseñor 1998; Yue et al. 1999; Yue 2001). La notion de copule a été introduite dans le but de résoudre ce problème. En effet, la copule permet la modélisation de la structure de dépendance entre des variables indépendamment des lois marginales de ces variables (Sklar 1959).

L'AF peut être divisée en quatre groupes: locale-univariée, régionale-univariée, locale-multivariée et régionale-multivariée. Les deux premiers groupes ont été largement étudiés (Dalrymple 1960; Stedinger et Tasker 1986; Kite 1988; e.g. Burn 1990; Hosking et Wallis 1993; Durrans et Tomic 1996; Nguyen et Pandey 1996; Alila (1999, 2000); Ouarda et al. 2001; Chebana et Ouarda 2008; Ouarda 2013). Récemment, l'AF locale-multivariée a pris de plus en plus d'importance dans la littérature hydrologique (Shiau 2003; De Michele et al. 2005; Grimaldi et Serinaldi 2006; Zhang et Singh 2006; Wang et al. 2009; Chebana et Ouarda 2009a; Chebana 2013). Enfin, rares sont les études d'AF régionale-multivariée mais elles commencent à prendre de l'ampleur (Chebana et Ouarda 2007; Chebana et al. 2009; Sadri et Burn 2011; Chebana et Ouarda 2011a).

2.2 Différentes étapes de l'analyse fréquentielle

L'AF locale (univariée ou multivariée) est composée de quatre étapes principales : (a) l'analyse exploratoire des données qui comprend la détection des valeurs aberrantes, l'estimation (ou imputation) des données manquantes et l'analyse descriptive des

données, (b) la vérification des hypothèses de base telles que la stationnarité, l'indépendance et l'homogénéité, (c) la modélisation de l'événement extrême et l'estimation des paramètres correspondants et (d) l'estimation et l'analyse du risque associé. Dans le cas univarié, ces étapes ont été généralement considérées et incluses dans les études (e.g. Kite 1988; Bobée et Ashkar 1991; Rao et Hamed 2000; Yue et al. 2002; Khaliq et al. 2006). Cependant, dans le cas multivarié, les étapes (c) et (d) ont reçu plus d'attention que les deux premières étapes. En effet, à notre connaissance, Chebana et Ouarda (2011b) sont les seuls qui ont étudié les statistiques descriptives des séries multivariées telles que la moyenne, la variance, l'asymétrie et l'aplatissement ainsi que la détection des valeurs aberrantes dans un cadre d'AF multivarié hydrologique. Pour l'étape (b), les mêmes auteurs ont réalisé une revue de littérature et une application des tests de détection de tendance monotones multivariées (Chebana et al. 2013). Cette étude est considérée comme le premier travail dans le cadre de l'étape b et dans un contexte d'AF multivariée. D'autre part, l'étape (c) a été largement étudiée dans la littérature (e.g. Shiau 2003; Zhang et Singh 2006; Karmakar et Simonovic 2009; Chebana et Ouarda 2009b; Salarpour et al. 2013). Concernant l'étape d, plusieurs travaux ont étudié la notion de période de retour dans un contexte multivarié (e.g. Shiau 2003) et Chebana et Ouarda (2011a) ont examiné l'estimation des quantiles multivariés.

Par ailleurs, une AF régionale est composée de deux principales parties soit la délimitation de la région hydrologique homogène et l'estimation régionale (e.g. Ouarda 2013). Dans un contexte multivarié, la délimitation d'une région hydrologiquement homogène a été étudiée par Chebana et Ouarda (2007). Ils ont proposé des tests de discordance et d'homogénéité multivariés basés sur les L-moments multivariés et les

copules. Chebana et al. (2009) ont considéré l'aspect pratique de ces tests en les appliquant sur une région de la Côte Nord de Québec, Canada. Chebana et Ouarda (2009) ont proposé une procédure d'AF régionale-multivariée en se concentrant sur la partie de l'estimation régionale. La procédure proposée représente une version multivariée du modèle de l'indice de crue où il a été évalué sur la base d'une étude de simulation. Notons que les quatre étapes de l'AF locale sont présentes dans la partie de l'estimation régionale de l'AF régionale.

3 Détermination, adaptation, comparaison et application des techniques d'analyse et de modélisation des variables hydrologiques

Ce chapitre présente une discussion sur les approches et les méthodologies des articles et rapport présentés dans la deuxième partie de cette thèse. Ainsi, ce chapitre est organisé comme suit : les travaux réalisés dans le cadre de l'étude de l'évolution de la dépendance entre les principales caractéristiques de crue sont présentées. Ensuite, une étude sur l'imputation des données manquantes en AF multivariée est discutée. Par la suite, une étude de simulation et d'application des tests multivariés de détection des points de rupture est présentée. Les tests considérés sont basés sur les fonctions de profondeur. Enfin, les aspects pratiques de l'AF régionale basée sur l'indice de crue dans un cadre multivarié sont discutés.

3.1 Imputation des données manquantes en AF univariée et bivariée

L'objectif de la présente section est de souligner l'importance de traiter les données manquantes en AF hydrologique multivariée par une revue de littérature, l'application de méthodes d'imputation multivariée et une comparaison des méthodes d'imputation univariée et multivariée. Deux types d'applications multivariées sont présentées, multi-variables pour les différentes caractéristiques de crue et multi-sites pour plusieurs stations.

3.1.1 Donnée manquante - Revue de littérature

Le traitement des données manquantes a été largement étudié dans la littérature statistique (e.g. Allison 2001; Little et Rubin 2002; Molenberghs et Kenward 2007; Enders 2012; Graham 2012). Plusieurs méthodes d'imputation des données manquantes ont été développées et peuvent être séparées selon deux domaines soit : le domaine des séries chronologiques, i.e. analyse des données sur une période de temps et le domaine de l'AF, i.e. analyse des données qui sont périodiques. Aussi bien en univarié qu'en multivarié, la reconstruction des données manquantes a été largement étudiée dans le domaine des séries chronologiques (e.g. Gleason et Staelin 1975; Gyau-Boakye et Schultz 1994; Hughes et Smakhtin 1996; Abebe et al. 2000; Jeffrey et al. 2001; Ng et al. 2009; Han et Li 2010; Honaker et King 2010; Marlinda et al. 2010). Cependant, la reconstruction des données manquantes dans le domaine de l'AF a reçu moins d'attention (e.g. Kelly et al. 2004; Erol 2011; Peterson et al. 2011).

Comme en statistique, les travaux de reconstitution des données manquantes en hydrologie se sont concentrés sur celles en séries chronologiques telles que les séries de débit et de précipitation. Le

Tableau 1 présente la répartition des études de reconstitution des données manquantes en hydrologie et ailleurs en statistiques avec quelques références. L'étude de reconstruction des données manquantes dans un cadre de l'AF multivariée en hydrologie n'a pas été complétée dans le passé et donc, sera un des objectifs de la présente étude.

Tableau 1 : Résumé des domaines d'étude des données manquantes avec quelques références

Contexte		Application/ développement	
		Statistique	**Hydrologie**
Univarié	Séries chronologiques	Largement étudié : Chow and Lin (1976) Azen et al. (1989) Gelason and Staelin (1975)	Largement étudié : Lettenmayer (1980) Jefferey et al. (2001) Teegavarapu and Chandramouli (2005)
	Analyse fréquentielle	Largement étudié : Kodituwakku et al. (2011) Erol (2011)	Rarement étudié : Fleig et al. (2010) Peterson et al. (2011)
Multivarié	Séries chronologiques	Largement étudié : Hopke et al. (2001) Honaker et al. (2010) Frane (1976)	Assez étudié : Ng et al. (2009) Kalteh and Hjorth (2009)
	Analyse fréquentielle	Rarement étudié : Kelly et al. (2004)	**N'a pas été étudié**

Les techniques univariées de reconstruction des données manquantes ont été largement étudiées en hydrologie, comme l'utilisation de la moyenne de la série ou la moyenne d'une partie de la série (e.g. Linacre 1992), les modèles de séries temporelles (e.g. Lettenmaier 1980), l'interpolation spatiale et/ou temporelle (e.g. Filippini et al. 1994), la régression (e.g. Kuligowski et Barros 1998), l'imputation hot deck (e.g. Srebotnjak et al. 2012) et la pondération inverse à la distance (e.g. ASCE 1996). Dans le cadre multivarié, les applications en hydrologie des techniques de reconstruction des données manquantes peuvent être divisées en 3 groupes : (1) des versions multivariées des techniques univariées telles que : le modèle de régression multivarié (e.g. Simonovic 1995) et les modèles de séries temporelles multivariées (e.g. Bennis et al. 1997); (2) les méthodes basées sur les données (« data driven methods ») incluant les réseaux de neurones artificiels (ANN) (e.g. Raman et Sunilkumar 1995), la méthode des k-voisins

les plus proches (K-NN) (e.g. Kalteh et Hjorth 2009); (3) les méthodes basées sur des modèles statistiques tels que l'algorithme espérance-maximisation (EM) (e.g. Ng et al. 2009) et l'algorithme de l'imputation multiple (MI) (e.g. Ng et al. 2009). La comparaison entre les différentes méthodes de reconstitution des données manquantes dans les séries chronologiques a été l'objectif des plusieurs études telle que Kalteh et Hjorth (2009) et Coulibaly et Evora (2007). Les résultats des quelques études de comparaison sont présentés dans l'article 1 en deuxième partie.

Dans la présente étude, six méthodes d'imputation ont été considérées, dont trois méthodes univariées: la substitution par la moyenne, l'interpolation linéaire et l'arbre de régression pas-à-pas; ainsi que trois méthodes multivariées: la carte auto-organisatrice (« self-organizing map ») qui forme une classe des ANN fondée sur des méthodes d'apprentissage non-supervisées, l'algorithme EM régularisé (REGEM) et l'algorithme MI.

3.1.2 Méthodes d'imputation considérées

Dans la présente section, on décrit brièvement les différentes méthodes d'imputation considérées. L'article en question dans la partie 2 contient les descriptions détaillées de ces méthodes.

Substitution par la moyenne (MS) : C'est la méthode d'imputation la plus simple. Elle consiste à remplacer les données manquantes par la moyenne de la série. Cette méthode a été utilisée dans plusieurs études d'AF multivariées hydrologiques telle que Wang et al. (2009) et Kao et Chang (2012).

Interpolation linéaire (LI) : Cette méthode consiste à tracer une ligne droite entre les valeurs observées avant et après les données manquantes puis les estimer par interpolation. Elle a été considérée par Fleig et al. (2011) en AF univariée régionale. Cependant elle n'a pas été considérée dans une AF multivariée.

Arbre de régression pas-à-pas (SRT) : L'algorithme de la SRT, développé par Huang et Townshend (2003), est une amélioration de l'arbre de régression en utilisant des modèles de régression linéaires avec une sélection des variables explicatives pas-à-pas. Initialement, toutes les observations se trouvent dans le nœud de départ de l'arbre. Les observations sont partagées jusqu'à ce que les nœuds soient considérés terminaux. Un nœud est partagé lorsqu'il s'ensuit une amélioration de la somme des carrés des erreurs des estimations.

Carte auto-organisatrice (SOM) : La méthode SOM (aussi appelée carte de Kohonen) fait partie de la famille des algorithmes ANN. Elle est basée sur des méthodes d'apprentissage non-supervisées (Kohonen et al. 1996). Plusieurs travaux récents ont montré la performance de cette méthode dans l'estimation des données manquantes dans les séries chronologiques en hydrologie telle que Adebayo et Rustum (2012) et Mwale et al. (2012). La méthode SOM permet de transformer, de façon non linéaire, l'espace de données, généralement de grande dimension, à l'espace de représentation qui est une carte discrète de deux dimensions. Les nœuds de la carte sont disposés géométriquement selon une topologie fixée a priori. L'apprentissage de la carte se fait de façon itérative similaire à un algorithme d'apprentissage séquentiel. Au début, les vecteurs poids entre les deux espaces doivent être initialisés à chaque neurone. Puis les vecteurs d'entrée sont comparés avec les neurones de la carte SOM pour trouver les correspondances les plus

proches qui sont appelées les meilleures unités de reconnaissance (BMU). À cette fin, la distance euclidienne est le critère le plus communément utilisé. Cette procédure doit être réitérée plusieurs fois jusqu'à ce que le nombre optimal d'itérations soit atteint ou que les critères d'erreur spécifiés soient atteints. La valeur de la donnée manquante est le BMU dans le nœud correspondant.

Algorithme Espérance-maximisation régularisé (REGEM) : L'algorithme EM est une méthode itérative d'estimation des données manquantes se basant sur le maximum de vraisemblance (Dempster et al. 1977). L'algorithme REGEM est une forme particulière de l'algorithme EM basée sur des modèles de régression entre les données observées et celles manquantes (Schneider 2001). C'est un algorithme itératif en deux principales étapes : l'espérance (E) et la maximisation (M). Plus précisément, il consiste à : (1) remplacer les données manquantes par les valeurs estimées par la dernière itération; (2) estimer, compte tenu des données observées et les paramètres estimés de la régression actuelle, le vecteur des moyennes et la matrice de covariance des données et (3) réestimer les données manquantes en assumant que les nouveaux paramètres sont corrects. L'algorithme consiste à l'itération de ces trois étapes jusqu'à la convergence c'est-à-dire lorsque les variations du vecteur des moyennes et la matrice de covariance sont inférieures à un seuil prédéfini. Les estimations initiales des paramètres du modèle sont obtenues à partir de la base de données complète après la substitution des données manquantes par la moyenne. L'algorithme REGEM a été largement utilisé pour l'imputation des données manquantes dans les séries multivariées normalement distribuées (Little et Rubin 2002). Cependant, les travaux traitant les données manquantes en hydrologie par l'algorithme REGEM sont rares (e.g. Kalteh et Hjorth 2009). A notre

connaissance, l'algorithme REGEM n'a pas été considéré en AF multivariée en hydrologie.

Imputation multiple (MI) : La méthode MI est une procédure assez simple pour imputer les données manquantes multivariées (Rubin 1987). L'idée de base de cette méthode est d'abord de générer plusieurs séries complétées par la génération de plusieurs valeurs possibles pour chaque donnée manquante, puis d'analyser chaque série séparément. Le nombre de bases de données complètes à générer dépend du pourcentage des données manquantes. Toutefois, selon Schafer (1997), cinq estimations des données manquantes fournissent généralement des estimations non biaisées. Comme l'algorithme REGEM, la méthode MI a été largement utilisée pour la reconstruction des données manquantes dans les séries multivariées normalement distribuées (Little et Rubin 2002). Cependant, son application en hydrologie est rare (e.g. Kalteh et Hjorth 2009) en particulier dans l'AF hydrologique multivariée.

3.1.3 Performances des méthodes d'imputation

Pour évaluer la performance des méthodes d'imputation, on utilise la procédure de rééchantillonnage jackknife. Elle consiste à considérer chaque valeur comme manquante en la retirant temporairement de la série et l'estimer à partir du reste des données. Les critères utilisés pour évaluer les performances sont la racine de l'erreur quadratique moyenne relative (*RRMSE*) et le biais relatif moyen (*MRB*) défini respectivement par :

$$RRMSE = \frac{100}{n}\sqrt{\sum_{i=1}^{n}\left(\frac{\hat{x}_i - x_i}{x_i}\right)^2}, \; x_i \neq 0 \qquad (1)$$

$$MRB = \frac{100}{n}\sum_{i=1}^{n}\left(\frac{\hat{x}_i - x_i}{x_i}\right), x_i \neq 0 \qquad (2)$$

avec $\hat{x}_i$ est la valeur imputée et x_i est la valeur observée.

3.1.4 Application

Dans cette section, les méthodes d'imputation décrites précédemment sont appliquées sur deux situations : multi-variables pour les différentes caractéristiques de crue à la même station et multi-sites pour plusieurs stations avec la même caractéristique de crue.

Multi-variables : Dans cette application, les principales caractéristiques de crue sont considérées, soit Q, V et D extraites à partir des séries de débit journalier de trois stations hydrométriques avec régimes naturels. Ces stations sont *Moisie*, *Magpie* et *Romaine* situées dans la région de la Côte Nord. La Figure 1 montre la localisation géographique des trois stations tandis que le Tableau 2 présente les informations générales sur ces stations.

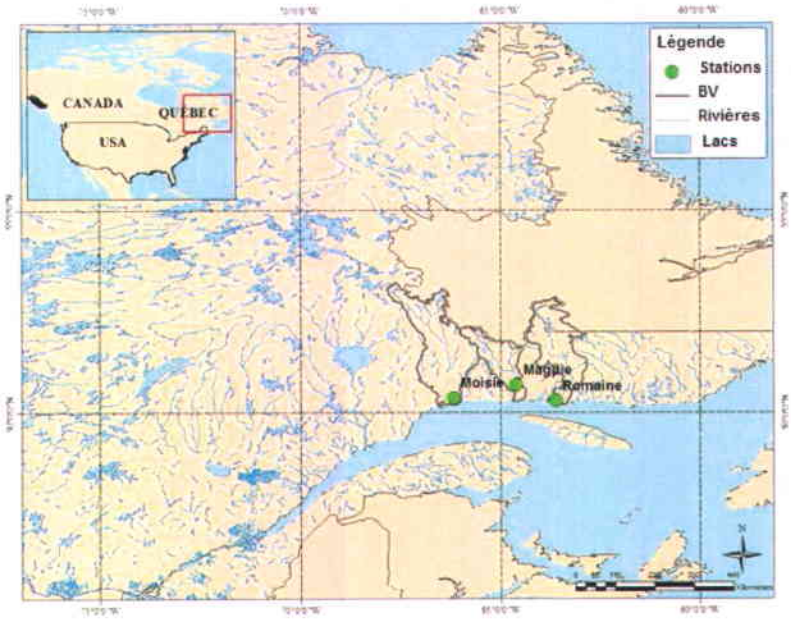


Figure 1 : Localisation géographique des stations Moisie, Magpie et de la Romaine

Tableau 2 : Informations générales des stations Moisie, Magpie et de la Romaine

Nom de la station	Moisie	Magpie	Romaine
Numéro de la station	72301	73503	73801
Latitude	50 21 09	50 41 08	50 18 28
Longitude	-66 11 12	-64 34 43	-63 37 07
Période d'enregistrement (#années)	1979-2004 (26)	1979-2004 (26)	1979-2004 (26)
Données manquantes	1999, 2000	-	-
Superficie (Km^2)	19 012	7 201	12 922
Débit moyen (m3s-1)	391.62	163.56	282.89

Pour chaque station, les méthodes d'imputation univariée, à savoir MS, LI et SRT sont appliquées à chaque série de Q, V et D alors que pour les méthodes multivariées, les séries considérées sont les couples (Q,V), (V,D) et (Q,D). Les performances des méthodes d'imputation sont évaluées en utilisant les deux stations qui ne présentent pas

de données manquantes, soit *Magpie* et *Romaine*, et les méthodes d'imputation sont appliquées pour estimer les données manquantes dans la station *Moisie*.

Dans le présent travail, on assume qu'une seule valeur peut être manquante dans chacune des séries. Par conséquent, trois situations peuvent se présenter (Figure 2): (i) une seule série présente des données manquantes, (ii) chaque composante de la série possède une donnée manquante dans le même évènement, et (iii) chaque composante de la série présente une donnée manquante, mais pas au même événement. Noter que les performances de MS et LI restent les mêmes dans les trois situations, car ils traitent chaque série séparément. Les méthodes SRT, SOM et REGEM ne peuvent pas être utilisées dans (ii) parce qu'elles exigent au moins une observation pour chaque événement. Par conséquent, seule la méthode MI peut être appliquée pour (ii). Les situations (i) et (ii) sont étudiées en détail dans la présente étude, tandis que (iii) a fait l'objet d'un exemple, car il peut être considéré comme une combinaison des deux cas de (i).

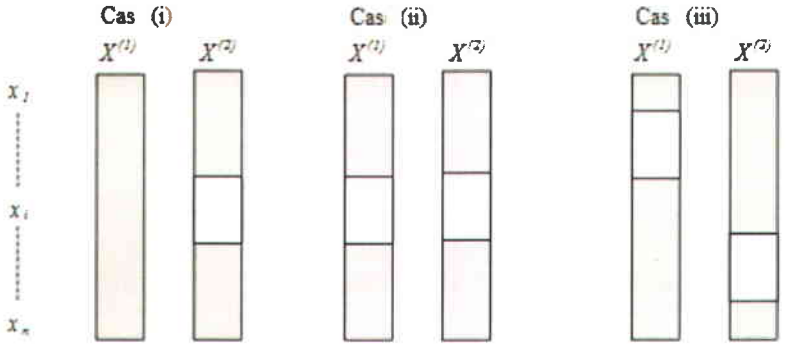


Figure 2 : Exemples de situation de données manquantes.

La couleur grise correspond aux données observées

Multi-sites : Dans cette section, nous examinons l'application des méthodes d'imputation considérées à des situations multi-sites où chaque site peut être considéré comme une variable associée à une série. Les trois sites utilisés dans l'application précédente sont utilisés ici également. Le Tableau 3 présente les corrélations entre les différentes stations de Q, V et D. Pour chaque série individuelle, les méthodes d'imputation univariée et multivariée sont appliquées. Dans la présente application, on se concentre sur la première et la deuxième situation décrites dans la section précédente, pour évaluer les performances des méthodes d'imputation des données manquantes. Le but ici est de comparer les performances des méthodes d'imputation par les différentes valeurs de dépendance. Comme dans l'application multi-variables, nous utilisons les deux stations sans données manquantes, à savoir *Magpie* et *Romaine* et les trois séries multivariées: $(Q_{Magpie}, Q_{Romaine})$, $(V_{Magpie}, V_{Romaine})$ et $(D_{Magpie}, D_{Romaine})$.

Tableau 3 : Corrélation entre Q, V et D

Stations		Variables		
		D	V	Q
Moisie				
	D	1	0.66	-0.07
	V		1	0.59
	Q			1
Magpie				
	D	1	0.44	-0.20
	V		1	0.70
	Q			1
Romaine				
	D	1	0.18	-0.36
	V		1	0.77
	Q			1

3.2 Détection des points de rupture multivariés en hydrologique

Dans la présente section, on s'intéresse aux tests de détection des points de rupture. Ces tests sont basés sur les fonctions de profondeur. Une étude de simulation dans le contexte hydrologique est effectuée pour comparer la puissance des tests considérés. En outre, une application est présentée afin de montrer l'aspect pratique des tests considérés selon des contraintes hydrologiques. Notons que la nonstationnarité est une notion générale qui peut se manifester, entre autres, en présence de points de rupture.

3.2.1 Concept

Un point de rupture peut être défini comme la position dans laquelle au moins une des caractéristiques d'un modèle statistique (par exemple, la moyenne, la variance ou la tendance) subit un changement brusque (Seidou et al. 2007). Un grand nombre de techniques peuvent être trouvées dans la littérature pour identifier la date d'un changement potentiel et pour vérifier si le changement est significatif ou non. La plupart des méthodes utilisent des tests statistiques pour détecter les changements dans les pentes des modèles de régression linéaire (Easterling et Peterson 1995; Vincent 1998; Lund et Reeves 2002). Cependant, les approches de détection des points de rupture ne sont pas toutes basées sur des tests statistiques. Par exemple Wong et al. (2006) ont utilisé la méthode relationnelle (Moore 1979; Deng 1989) pour la détection d'un seul point de rupture dans les séries de débits. Pour une revue de littérature approfondie sur les techniques de détection des points de rupture en hydrologie et climatologie, le lecteur peut se référer à Peterson et al. (1998).

Pour définir un point de rupture, soit $(x_i)_{i=1,\ldots,n}$ un échantillon de données de d-dimension et $1 < s < n$ un point de rupture possible. Si s existe, alors l'échantillon peut être divisé en deux sous-échantillons de tailles s et $m = n\text{-}s$ tel que:

$$\begin{aligned}(y_1,\ldots,y_s) &= (x_1,\ldots,x_s)\\ (z_1,\ldots,z_m) &= (x_{s+1},\ldots,x_n)\end{aligned} \tag{3}$$

Soit G_1 et G_2 les fonctions de distribution des deux sous-échantillons respectivement. Les deux fonctions de distribution G_1 et G_2 ont la même forme sauf pour le paramètre de localisation, c'est-à-dire : $G_1(x) = G_2(x+\boldsymbol{\delta})$ pour tout $x \in R^d$ avec $\boldsymbol{\delta} \in R^d$ un vecteur constant. Par conséquent, lorsqu'on teste la présence d'un point de rupture au point s dans la série $(x_i)_{i=1,\ldots,n}$, l'hypothèse nulle et l'hypothèse alternative sont respectivement:

$H_0 : \boldsymbol{\delta} = 0$; il n'y a pas un point de rupture (4)

$H_1 : \boldsymbol{\delta} \neq 0$; il existe un $j \in \{1,\cdots,d\}$ tel que $\boldsymbol{\delta}_j \neq 0$. (5)

3.2.2 Description des tests considérés

Dans le présent document, plusieurs tests de détection des points de rupture sont considérés. Tous ces tests sont basés sur la fonction de profondeur, sauf un (le test de Cramér), qui est considéré à titre de comparaison. La fonction de profondeur est une notion statistique utile dans l'inférence non paramétrique des données multivariées. Trois différentes fonctions de profondeur sont considérées à savoir: Mahalanobis, Simplicial et Half-space notées respectivement par MD, SD et TD. Une brève description des tests

considérés est présentée dans ce qui suit. Pour plus de détail sur ces tests ainsi qu'une revue bibliographique sur les études de comparaisons de performance entre ces tests est présentée dans le rapport en question dans la section 2.

Le test de Cramér (C-test) : Le test de Cramér utilisé dans la présente étude a été développé par Baringhaus et Franz (2004) et représente une généralisation du test univarié proposé par Cramér (1928). Ce test est basé sur la différence entre la distance euclidienne entre les deux sous-échantillons et la moitié de la somme de toutes les distances euclidiennes d'un même sous-échantillon.

Le test M : Selon Li et Liu (2004), le point le plus profond d'une distribution est un paramètre de localisation (comme la médiane). Par conséquent, si G_1 et G_2 sont des distributions identiques, alors leurs points les plus profonds θ_{G_1} et θ_{G_2} sont identiques également. Soit la fonction de profondeur D, si G_1 et G_2 sont identiques alors $D_{G_2}(\theta_{G_1}) = D_{G_1}(\theta_{G_2})$ par contre, si la différence entre les valeurs de $D_{G_2}(\theta_{G_1})$ et $D_{G_1}(\theta_{G_2})$ dépasse un seuil limite alors G_1 et G_2 ne sont pas identiques et donc l'existence d'un point de rupture est confirmé.

Le test T : Li et Liu (2004) ont décrit une approche graphique appelée DD-plot pour comparer les paramètres de localisation de deux sous-échantillons. Dans cette approche graphique, les deux axes représentent les valeurs des fonctions de profondeur de deux sous-échantillons. Lorsque les deux sous-échantillons suivent exactement la même distribution, le DD-plot correspond à une ligne diagonale qui passe par l'origine. Dans le cas contraire, le DD-plot présente une forme de feuille avec l'extrémité vers l'origine. Le

test T est basé sur l'approximation de la distance entre l'extrémité de la feuille et l'origine de la figure-DD.

Le test de Wilcox (Test W) : Le test W a été développé par Wilcox (2005). Il est basé sur l'idée que, sous l'hypothèse nulle, les médianes des deux sous-échantillons doivent être similaires. La statistique de test permet non seulement de savoir si les deux médianes sont différentes, mais aussi de combien.

Les tests d'indice de qualité (tests QIA et QIB) : Liu et Singh (1993) ont proposé un test de détection des points de rupture de type Wilcox des rangs signés. Il permet de tester l'existence d'un changement simultanément dans les paramètres de localisation et d'échelle. On peut déterminer la valeurs-p de façon asymptotique (QIA) ou en utilisant le bootstrap (QIB).

Le test de Zhang (test Z) : Récemment, Zhang et al. (2009) ont développé un nouveau test basé sur la statistique des tests d'indice de qualité (*QIA* et *QIB*) qui permet de détecter les points de rupture dans une série multivariée.

3.2.3 Étude de simulation

L'objectif de cette étude de simulation est d'évaluer et de comparer la performance des tests présentés précédemment dans un contexte hydrologique, par exemple pour les séries des caractéristiques de crue Q et V. En général, ces tests sont utilisés pour des séries de distribution Normal, Cauchy ou t qui ne sont pas appropriée dans l'AF hydrologique multivariée. Ces tests peuvent être affectés par plusieurs facteurs. Dans cette étude de simulation, nous étudions l'impact de la longueur des séries n (taille de

l'échantillon), ainsi que le degré de changement (amplitude de la rupture) dans chaque composante de la série multivariée.

Certaines études antérieures ont montré que les séries de Q et V peuvent être marginalement représentées par une distribution de Gumbel. Alors que la dépendance entre Q et V peut être modélisée par une copule de Gumbel (e.g. Yue et al. 1999; Shiau 2003; Chebana et Ouarda 2009a; Ben Aissia et al. 2011). Ces distributions ont été considérées pour générer les séries de simulations.

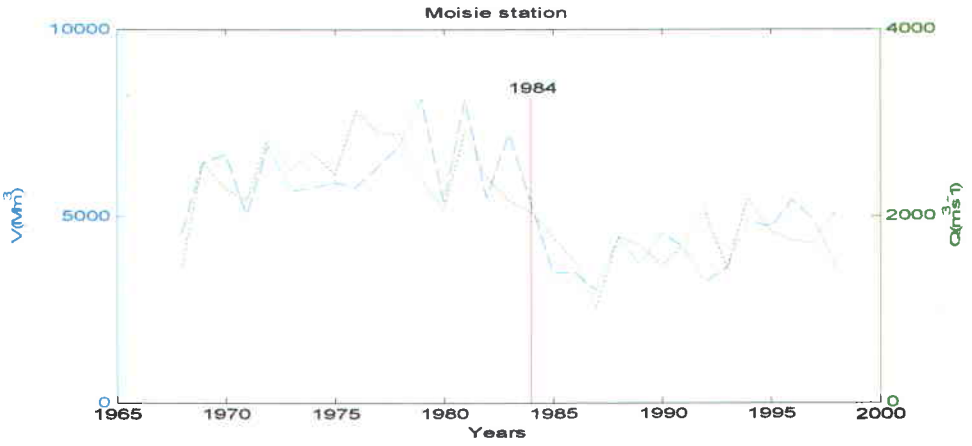
Pour les paramètres de ces distributions, nous avons sélectionné ceux des études de simulation dans Chebana et Ouarda (2007, 2009b). L'étude de simulation effectuée est constituée de deux étapes. Lors de la simulation, nous avons généré un grand nombre N d'échantillons pour évaluer les effets de différents facteurs sur la performance des tests. Trois tailles d'échantillons sont considérées soit n = 30, 50 et 80 correspondant aux points de rupture s = 10, 20 et 30 respectivement. Pour chaque taille de l'échantillon, plusieurs amplitudes de changement dans les paramètres de localisation sont considérées soit δ = 10, 20, -20, 40 et 70%.

3.2.4 Application

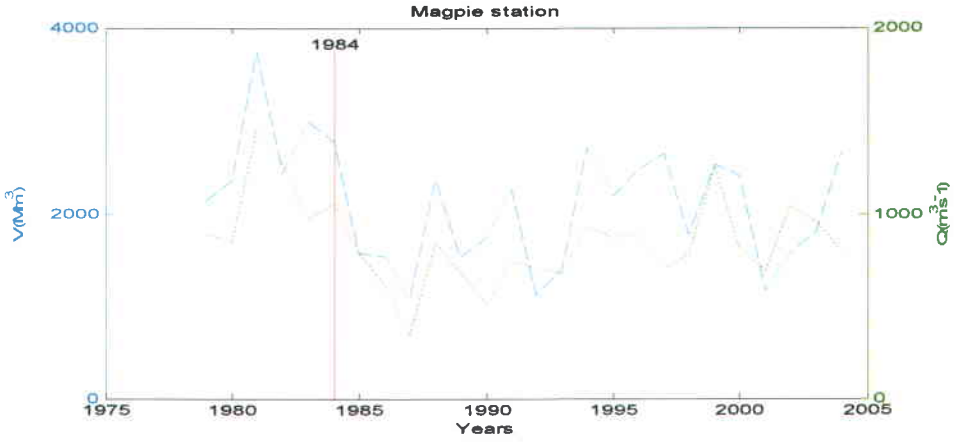
Dans cette section, les tests présentés et évalués sont appliqués sur trois stations de la côte Nord de Québec (*Moisie*, *Magpie* et *Romaine*). La localisation géographique des trois stations ainsi que leurs informations générales sont présentées dans la Figure 1 et le Tableau 2.

En utilisant le débit journalier des trois stations, les caractéristiques des crues de printemps Q et V sont calculées pour chaque année. La Figure 3 montre les séries

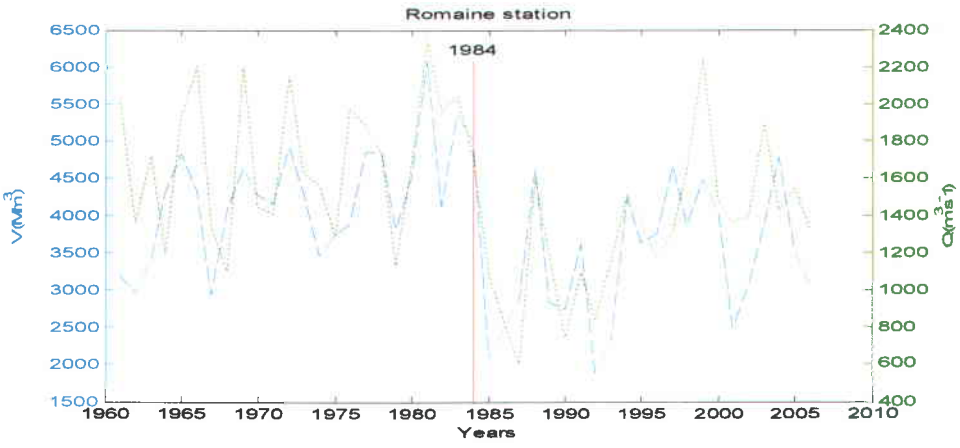
disponibles de Q et V des trois stations. Étant donné que ces stations sont géographiquement proches (Figure 1), il est possible qu'un changement (point de rupture) soit observé dans les trois stations autour de la même année. D'après la Figure 3, il semble que l'année 1984 pourrait être un point de rupture.



a)



b)



c)

Figure 3 : Les séries de Q et V des stations a) Moisie, b) Magpie et c) Romaine

3.3 Évolution de la dépendance entre les principales caractéristiques de crue dans un contexte de changement climatique

L'objectif de cette section de la thèse est d'évaluer et d'analyser la dépendance entre Q & V, entre V & D et entre Q & D en utilisant une fenêtre mobile de 30 ans sur une période allant de 1961 à 2070 pour les séries simulées et de 1969 à 2000 ou 2009 (selon le bassin versant) pour les séries historiques. Notons que la dépendance est une notion forte qui concerne toute la distribution jointe des variables concernées. Cependant, cette dépendance est généralement mesurée avec des coefficients qui la résument comme le coefficient de corrélation de Pearson, le tau de Kendall et le rho de Spearman. Ces trois mesures de dépendance sont utilisées dans la présente étude. Ainsi, on obtient pour chaque couple de variables et mesure de dépendance une série différente. L'étude des séries de dépendance obtenues comporte la détermination des bandes de confiance, l'étude de stationnarité et la détermination des points de rupture. Pour effectuer les comparaisons des résultats obtenus, les bassins versants de Baskatong et de la Romaine ont été considérés. De plus, pour générer les séries de débits futurs, deux modèles hydrologiques : HSAMI (e.g. Bisson et Roberge 1983)et HYDROTEL (Fortin et al. 1995) ont été considérés. Ces modèles hydrologiques sont alimentés par des séries météorologiques issues de trois modèles climatiques (CGCM3, HADCM3 et ECHAM5) suivant trois scénarios d'augmentation de gaz à effet de serre (A2, B1 et B2) (Bates et al. 2008). Le choix des bassins versant, des modèles hydrologiques, des modèles climatiques et des scénarios climatiques a été faite suivant la disponibilité des données.

3.3.1 Zone d'étude et bases de données

Les deux bassins versants considérés (Baskatong et Romaine) sont situés au Québec (voir Figures 1 et 2 de l'article 2 dans la partie 2). La superficie des bassins versants est 13 040 Km^2 pour Baskatong et 14 500 Km^2 pour Romaine. Le régime hydrologique dans les deux bassins versants est caractérisé principalement par une importante fonte de neige au printemps. De nombreux ouvrages hydrauliques sont construits ou sur le point d'être construits sur les deux réservoirs. On cite par exemple le barrage Mercier sur le réservoir Baskatong et 4 grands barrages (Romaine-1, Romaine-2, Romaine-3 et Romaine-4) sur la rivière Romaine.

La base de données de chacun des sites contient des séries de débit observé et simulé. Pour Baskatong, on a deux modèles hydrologiques soit HSAMI et HYDROTEL alors que pour la Romaine, seul HSAMI est considéré, mais avec trois scénarios de changements climatiques.

Plus précisément, la base de données de Baskatong est divisée en trois catégories soit les données observées, les données de réanalyse ERA-40 et les données de simulation CGCM3. Pour les deux dernières catégories, deux modèles hydrologiques sont considérés soit HSAMI et HYDROTEL. Pour la Romaine, sept séries de débit ont été étudiées, la première contient les débits observés, les six autres séries sont simulées par HSAMI en utilisant les données météorologiques des trois modèles climatiques CGCM3, HADCM3 et ECHAM5 suivant les scénarios d'augmentation des gaz à effet de serre A2 & B1 pour CGCM3 et ECHAM5 et A2 & B2 pour HADCM3. Pour plus de détails, veuillez voir l'article en question dans la partie 2.

3.3.2 Méthodologie

La méthodologie adoptée dans la présente étude comporte l'extraction des caractéristiques des crues soit Q, V et D pour chacune des séries considérées (débit journalier). À partir de ces séries, on détermine des séries de dépendance en utilisant une fenêtre mobile de trente ans et les trois mesures de dépendance. Le choix de trente ans est adopté pour intégrer l'étendue moyenne des variabilités interannuelles, tel que l'ENSO (El Nino Southern Oscillation). Enfin, une étude des séries de dépendance est réalisée.

Pour l'extraction des caractéristiques de crue, on utilise l'algorithme de Pacher qui se base sur l'analyse des hydrogrammes annuels cumulatifs en ajustant les pentes d'une approximation linéaire. Cet algorithme permet d'identifier les dates de début et de fin de crue et, par conséquent, les principales caractéristiques de crues Q, V et D.

Afin de mesurer la dépendance entre ces caractéristiques, on a considéré la corrélation de Pearson, le tau de Kendall et le rho de Spearman. Les expressions et définitions de ces mesures sont détaillées dans l'article 2 présenté dans la partie 2. Les bandes de confiance des séries de dépendance sont estimées par la méthode *BCa* (biais corrigé et accéléré e.g. DiCiccio et Efron 1996). L'avantage de cette méthode est qu'elle est non biaisée et assez rapide en temps de calcul.

Pour l'étude de la stationnarité des séries de dépendance, on a considéré le test de Mann-Kendall (e.g. Mann 1945; Kendall 1975) ainsi que cinq de ces variantes soit : Mann-Kendall avec Block Bootstrap (BB) (e.g. Hipel et McLeod 2005; Khaliq et al. 2009), pre-whitening (PW) (e.g. Douglas et al. 2000; Zhang et al. 2001), trend-free pre-whitening (TFPW) (e.g. Yue et al. 2002), variance correction 1 et variance correction 2

(VC1 et VC2) (e.g. Bayley et Hammersley 1946; Hamed et Ramachandra Rao 1998; Yue et Wang 2004). Ces variantes du même test sont considérées pour tenir compte de l'autocorrélation dans les séries.

Enfin, le test bayésien de détection des points de rupture développé par Seidou et Ouarda (2007) a été considéré. Il est basé sur un modèle de régression linéaire multiple.

3.4 Analyse fréquentielle régionale avec l'indice de crue multivarié

Récemment Chebana et Ouarda (2009) ont proposé une procédure d'AF régionale-multivariée en se concentrant sur la partie de l'estimation régionale. La procédure proposée représente une version multivariée du modèle de l'indice de crue qui a été évalué sur la base d'une étude de simulation. Les aspects pratiques de la procédure proposée par Chebana et Ouarda (2009) font l'objet de la présente étude à travers une application à un cas d'étude, le premier du genre. Pour y parvenir, les données observées de débits et de diverses variables physio-météorologiques dans une région de N=26 sites de la Côte Nord de Québec, sont utilisées (Chebana et al. 2009). La pointe Q et le volume V de crues sont les deux variables à étudier conjointement.

3.4.1 Méthodologie

La procédure de l'AF régionale-multivariée est composée de sept étapes:

1- Délimitation de la région homogène

Cette étape consiste à identifier et à exclure les sites discordants en appliquant le test de discordance multivarié, puis, à vérifier l'homogénéité de la région avec le reste des sites en utilisant le test d'homogénéité. Ces tests sont proposés par Chebana et Ouarda (2007). En pratique, il est très difficile de trouver une région parfaitement homogène. D'après Hosking et Wallis (1997), une homogénéité approximative de la région est suffisante pour appliquer l'AF régionale, en particulier le modèle de l'indice de crue. Cette première étape a été réalisée dans Chebana et al. (2009) sur la même région et ses résultats seront utilisés dans cette étude.

2- Standardisation des séries

Soit N' le nombre des sites dans la région homogène (ou possiblement homogène). Pour chaque site i de la région, on calcule les paramètres locaux de position μ_{iX} et μ_{iY} i=1,..., N' puis on standardise les séries (x_{ij},y_{ij}) j=1,...,n_i par les équations suivantes :

$$x'_{ij} = \frac{x_{ij}}{\mu_{iX}}, y'_{ij} = \frac{y_{ij}}{\mu_{iY}} \qquad (6)$$

où (x_{ij},y_{ij}) est un couple de valeurs observées pour la station i et l'année j.

3- Choix de la distribution bivariée régionale

Cette étape consiste à déterminer la distribution bivariée adéquate pour la région homogène. La distribution bivariée est composée de la copule et de deux distributions marginales. Cette étape consiste à :

a) Rassembler les données de tous les sites (x_{ij},y_{ij}) avec j=1,...,n_i, i = 1, ..., N'et n_i le nombre d'observations du site i. pour construire une seule série de la région (x_k'', y_k'')

$k = 1,...,n;\ n = \sum_{i=1}^{N_1'} n_i$ qui sera utilisée pour la détermination de la copule et les distributions marginales;

b) Identifier les distributions marginales adéquates (pour X et Y) en utilisant des tests d'adéquation graphiques et les deux critères numériques AIC et BIC.

c) Déterminer la copule adéquate de la région à partir du test graphique proposé par Genest et Rivest (1993). Lorsque plusieurs copules sont adéquates, une des façons de choisir la meilleure copule est d'appliquer une version appropriée du critère AIC sur les copules adéquates.

4- Estimer les paramètres de la distribution bivariée ainsi que les paramètres régionaux

En utilisant les données d'un site i, on estime les paramètres des distributions marginales ainsi que ceux de la copule choisie dans l'étape 3. Pour les distributions marginales, on utilise la méthode des L-moments pour estimer les paramètres. Tandis que pour la copule on utilise la méthode de pseudo-maximum de vraisemblance qui est la meilleure selon la littérature (Besag 1975; Genest et al. 1995; Shih et Louis 1995; Kim et al. 2007). Ensuite, pour chaque paramètre de la distribution bivariée, on estime le paramètre régionale $\hat{\theta}_k^{(R)}$ par :

$$\hat{\theta}_k^{(R)} = \frac{\sum_{i=1}^{N_1'} n_i \hat{\theta}_k^{(i)}}{\sum_{i=1}^{N_1'} n_i}, \quad k = 1,...,s \tag{7}$$

avec s est le nombre des paramètres à estimés.

5- Estimation des courbes des différentes combinaisons du quantile bivarié $\hat{q}_{v,q}(p)$ définie par :

$$q_{XY}(p)=\begin{Bmatrix}(x,y)\in R^2 \quad \text{such that} \quad x=F_X^{-1}(u), \\ y=F_Y^{-1}(v);\ u,v\in[0,1]:C(u,v)=p\end{Bmatrix}$$

où F_X et F_Y sont les fonctions de répartitions de X et Y respectivement et p est le risque situé entre 0 et 1.

6- Estimation de l'indice de crue

L'indice de crue, noté μ, est estimé en utilisant un modèle de régression multivariée multiple :

$$\log(\mu)=E\times\log(A)+\varepsilon \tag{8}$$

où $\mu=(\mu_X,\mu_Y)$, A est la matrice des caractéristiques physiographiques des bassins versants, E est la matrice des paramètres à estimer et ε est l'erreur du modèle.

7- Estimation du quantile bivarié

Pour un site d'intérêt l (non Jaugé), on estime le quantile bivarié régional en multipliant le vecteur d'indice de crue par le quantile bivarié régional :

$$\left(\hat{Q}_{xy}(p)\right)_l=\begin{pmatrix}\mu_{lX}\\ \mu_{lY}\end{pmatrix}\hat{q}_{xy}(p),\ \ 0<\mathrm{p}<1 \tag{9}$$

Pour vérifier la performance du modèle adopté dans cette étude, on utilise la méthode de validation croisée pour calculer les trois critères de performance suivant :

Biais régional, Biais régional absolu et Erreur régionale quadratique utilisés dans Chebana et Ouarda (2009).

3.4.2 Application

L'application de la méthodologie adoptée a été réalisée sur une région de la Côte Nord de Québec. Le nombre de sites dans cette région est de 26 stations. Les longueurs des séries sont entre 14 et 48 et les superficies des bassins versants varient entre 489 km^2 et 15 600 km^2. La Figure 4 montre les localisations géographiques des stations ainsi que la variation spatiale de la corrélation entre Q et V. Tandis que le Tableau 4 présente des informations générales sur les stations et les valeurs de discordance Q, V et (Q,V)

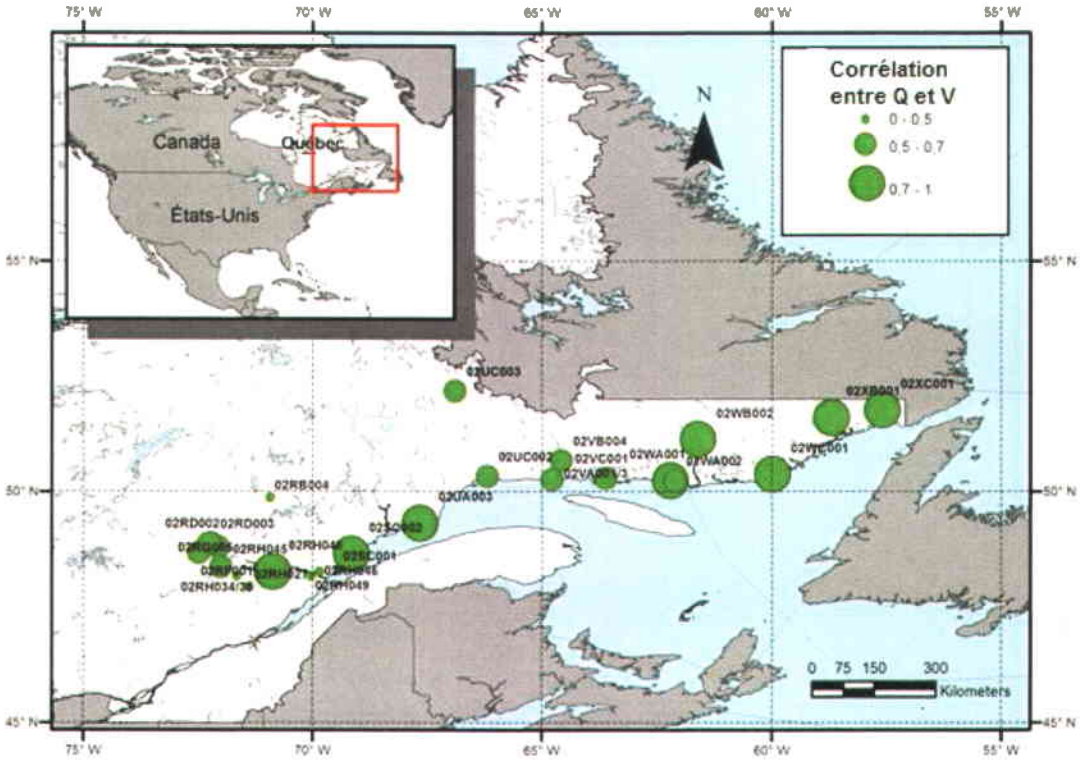


Figure 4 : Localisation géographique des stations

Tableau 4 : Informations générales sur les stations de la région d'étude

#	Nom de la station	*superficie* (Km^2)	n_i	Statistique de discordance		
				V	Q	(V,Q)
1	Petit Saguenay	729	24	0.80	0.40	1.09
2	Des Ha Ha	564	19	3.60	4.44	3.88
3	Aux Écorces	1120	34	0.16	2.42	0.69
4	Pikauba	489	34	0.89	1.16	1.22
5	Métabetchouane	2270	30	1.22	1.23	1.59
6	Petite Péribonka	1090	31	0.26	0.45	0.98
7	Chamouchouane (Ashuapmushuan)	15 300	43	0.13	0.14	0.26
8	Mistassibi	8690	39	0.32	0.78	0.88
9	Mistassini	9620	43	0.62	0.19	0.53
10	Manouane	3720	23	0.55	0.47	2.38
11	Valin	740	31	0.40	0.46	2.37
12	Ste-Marguerite	1100	21	1.50	0.55	1.30
13	DesEscoumins	779	19	1.14	1.81	1.27
14	Portneuf	2580	20	0.99	0.32	1.06
15	Godbout	1570	30	0.91	0.89	1.29
16	Aux-Pékans	3390	16	3.19	0.38	2.25
17	Tonerre	674	48	0.51	1.65	2.25
18	Magpie	7200	27	0.12	1.23	1.11
19	Romaine	13 000	48	1.62	0.48	0.57
20	Nabisipi	2060	25	1.12	0.64	0.54
21	Aguanus	5590	19	1.53	0.84	3.07
22	Natashquan	15 600	39	0.28	0.39	1.02
23	Etamamiou	2950	19	1.06	1.33	1.32
24	St Augustin	5750	14	0.62	0.67	0.92
25	St Paul	6630	25	0.31	1.35	1.11
26	Moisie	19000	39	1.16	0.32	0.54

La méthodologie adoptée ainsi que les différents tests et méthodes utilisés dans cette étude sont détaillés dans l'article en question présenté dans la partie 2.

4 Résultats et discussions

Dans ce chapitre, les résultats généraux des quatre travaux réalisés durant ma thèse sont présentés brièvement. Les résultats sont détaillés dans les articles et rapport respectifs présentés dans la partie 2.

4.1 Données manquantes

Les résultats de la revue de littérature montrent que la problématique des données manquantes a été largement étudiée dans différents domaines. De nombreuses méthodes d'imputation des données manquantes ont été développées dans les deux contextes univarié et multivarié. En hydrologie, le traitement des données manquantes dans les séries chronologiques a été largement étudié. Cependant, dans le domaine d'AF rare sont les travaux qui ont étudié l'imputation des données manquantes dans un contexte univarié et aucune dans le contexte multivarié.

Les résultats montrent qu'en général, les méthodes d'imputation ont tendance à surestimer les données manquantes. La performance de la méthode MI est meilleure que celle de SRT ou REGEM. La méthode de SOM conduit à de bonnes performances en particulier pour V et D, bien que son rendement est généralement inférieur à celui de SRT, REGEM et MI. Les performances des méthodes MS et LI sont plus faibles que celles des autres méthodes.

Dans le cas d'une forte dépendance entre les composantes de la série, les méthodes MI, SRT et REGEM produisent des performances élevées. La méthode MI semble donner les meilleures performances, mais les deux autres méthodes conduisent à une

performance presque similaire. La performance de la méthode SOM est inférieure à celles des MI, SRT et REGEM et supérieure à celles de LI et MS. Cette situation peut changer en cas de faible dépendance entre les variables, où la méthode SOM performe mieux que MI, SRT et REGEM. En cas d'indépendance, les méthodes univariées, telles que MS et LI, sont plus appropriées que celles multivariées.

4.2 Détection des points de rupture

Dans la présente section, on présente en premier lieu les résultats des simulations et par la suite ceux de l'application en deux cas d'étude. Dans le cas d'aucun changement dans tous les paramètres, de bonnes performances sont obtenues pour les tests M, T, W et C quelle que soit la fonction de profondeur. Par contre, les tests QIB, QIA et Z sont problématiques surtout lorsque n=30. On remarque une faible puissance des tests pour des faibles amplitudes de changement dans le paramètre de position. En effet, 10% de changement dans un ou les deux paramètres de position n'est pas détecté par les tests considérés. La performance des tests croit, généralement, avec la taille de l'échantillon et l'amplitude du changement. Il a été remarqué que le test C est sensible à l'ordre de grandeur de la composante qui présente un changement dans la série. Les meilleures performances sont obtenues pour les tests M, T, W, tandis que les tests QIB, QIA et Z peuvent être problématiques, surtout pour de faible amplitude de changement.

Au niveau de l'application de ces tests à un cas d'étude, l'année 1984 (Figure 3), pourrait être un point de rupture pour (V,Q). Pour vérifier l'existence du point de rupture à cette date, on a appliqué les tests considérés aux dates 1983, 1984 et 1985. À partir des résultats détaillés dans le rapport, on peut conclure que, pour la station *Moisie*, les trois

années peuvent être considérées comme des points de rupture. Considérant la station *Magpie*, seule l'année 1984 peut être considérée comme un point de rupture. De la Figure 3-b, on peut voir que le changement autour de 1984 n'est pas très clair. Notons que le court segment avant le changement peut avoir un effet sur la puissance des tests considérés à détecter le changement. Enfin, pour la *Romaine*, les années 1983 et 1985 peuvent correspondent à des points de rupture tandis que 1984 est probablement un point de rupture.

Les tests considérés permettent de tester, pour une date donnée, l'existence d'un point de rupture dans une série multivariée. En pratique, les tests peuvent confirmer à l'existence de point de rupture pour des dates différentes comme, par exemple, la station *Moisie* où les tests montrent que les années 1983, 1984 et 1985 peuvent être considérées comme des points de rupture. Une manière de déterminer la date probable du point de rupture est de déterminer la date où les valeurs-p des tests M, T, W, QIB, QIA et Z sont les plus faibles ou bien la valeur-p du test C la plus élevée. D'après les résultats détaillés dans l'article en question, on conclut que l'année 1984 présente, probablement, un point de rupture.

4.3 Étude de l'évolution de la dépendance

Pour le réservoir Baskatong, les résultats montrent que, généralement, les deux modèles hydrologiques HSAMI et HYDROTEL surestiment la dépendance entre Q & V et sous-estiment la dépendance entre V & D et entre Q & D (Figure 3 de l'article). À partir de la Figure 4 de l'article on note que les mesures de dépendance entre les principales caractéristiques de crues varient en fonction du temps et par conséquent, la

forme de l'hydrogramme varie également. De faibles valeurs de dépendance entre Q & V à la fin de la période de simulation (1961-2070) sont remarquées, ce qui montre que les hydrogrammes des crues futures peuvent être différents de l'hydrogramme typique où on observe une forte dépendance entre Q & V. Cette différence peut être due à des hydrogrammes aplatis ou pointus. On remarque également que lorsque les mesures de dépendance entre Q & V diminuent, celles entre V & D augmentent et vice versa (Figure 4 de l'article).

Pour la rivière Romaine, la Figure 5 dans l'article montre que l'allure des séries de dépendance entre Q & V et celles entre Q & D se ressemblent. Généralement, les modèles climatiques sous-estiment la dépendance entre Q & V et entre Q & D et surestiment celle entre V & D et les résultats de modèle HADCM3 sont les plus proches des séries observées. Pour les mesures de dépendance entre Q & V entre 2020 et 2050, on remarque deux groupes de séries. Le premier contient les séries du scénario B1 qui présentent des faibles valeurs de dépendance entre Q & V tandis que le deuxième contient les séries des scénarios A2 et B2 caractérisées par des fortes valeurs de dépendance entre Q & V. Par conséquent, la dépendance entre Q & V est très liée au scénario d'augmentation des gaz à effet de serre.

L'étude des bandes de confiance (à 95%) montre que ces deux bandes suivent la forme générale de la courbe de dépendance associée. Par contre, la largeur de ces bandes dépend de la variation de la série ainsi que du scénario d'augmentation des gaz à effet de serre.

Concernant la stationnarité, les résultats montrent qu'en utilisant le test de Mann-Kendall, la plupart des séries de dépendance sont non stationnaires. Par ailleurs, avec les

tests PW et TFPW, la majorité des séries sont stationnaires. Pour les tests BB, VC1 et VC2, toutes les séries sont stationnaires. La raison de la non-concordance de ces résultats est le fait que le test de Mann-Kendall ne tient pas compte de l'autocorrélation dans la série et que les deux tests PW et TFPW se basent sur l'hypothèse que les observations sous-jacentes génèrent un mécanisme conforme au modèle autorégressif d'ordre 1 ce qui n'est pas toujours vrai (Khaliq et al. 2009).

Différents points de rupture (entre 2 et 5 points) ont été détectés dans les séries de dépendance provenant des séries simulées. Pour les séries de dépendances obtenues à partir des données observées, au maximum un point de rupture a été observé. Ceci est dû au fait qu'elles sont moins longues que celles provenant des données simulées et qu'on utilise la même taille de la fenêtre mobile pour les divers types de données.

4.4 Analyse fréquentielle régionale multivariée

Les résultats de l'étude de discordance et d'homogénéité (Chebana et al. 2009) montrent que si on enlève les deux sites « De Ha Ha » et « Aguanus » de la région d'étude, la région est homogène pour V, hétérogène pour Q et possiblement homogène pour (V,Q).

Pour le choix des distributions marginales, les résultats montrent que les distributions adéquates sont Gumbel pour Q et GEV pour V. La t-copule est la copule la plus adéquate d'après la procédure de sélection adoptée. Une très bonne performance (R^2> 0,96) a été obtenue pour le modèle de régression multivarié multiple qui sert à estimer l'indice de crue.

En analysant les résultats obtenus, on remarque une concordance entre les résultas de l'AF multivariée et ceux de l'AF univariée. En effet, cela est dû au fait que les quantiles univariés sont les extrémités de la courbe des quantiles multivariés. On remarque également que la performance du modèle d'AF multivariée régionale est bonne où les biais moyens sont inférieurs à 10% pour un risque de 0,99. Plus précisément, la performance du modèle univarié de Q est plus faible à la fois que celle du modèle univarié de V et celle du modèle bivarié de (V,Q). Cette constatation est due à l'hétérogénéité de la région si on considère Q uniquement. Le biais du modèle d'indice de crue multivarié dépend aussi bien du risque p, de la discordance bivariée et de la discordance univarée.

5 Conclusions et perspectives de recherche

5.1 Conclusions

L'analyse et la modélisation adéquate des variables hydrologiques sont cruciales pour l'étude des événements extrêmes. Un des outils de base de l'analyse des événements extrêmes est l'AF qui a été largement étudiée dans le cadre univarié. Par ailleurs, généralement les évènements extrêmes sont caractérisés par plusieurs variables corrélées qui peuvent subir des changements de différentes sources (changement climatique par exemple). D'où la nécessité d'étudier l'AF et ces différentes étapes dans un cadre multivarié et dans un contexte de changement. Ceci correspond à l'objectif général de ma thèse. Pour atteindre cet objectif, on a réalisé quatre études qui constitueront trois articles et un rapport dans ma thèse.

La première étape de l'AF est l'analyse descriptive et explicative des données. Dans ce cadre on a réalisé deux études : la première consiste à déterminer, adapter, comparer et appliquer les méthodes d'imputation des données manquantes dans des séries multivariées tandis que la deuxième consiste à étudier l'évolution temporelle de la dépendance entre les principales caractéristiques de crue.

L'étude d'imputation des données manquantes montre que l'utilisation des techniques d'imputation multivariées, telles que MI, REGEM et SOM, donne des résultats meilleurs que celles des techniques univariées. La dépendance entre les composantes des séries multivariées joue un rôle important dans la performance des techniques d'imputation. Dans le cas d'indépendance entre les composantes, les

méthodes d'imputation univariées donnent des performances meilleures que celles multivariées.

La deuxième étape de l'AF est la vérification des hypothèses de base. Dans le cadre de cette étape, on a réalisé une étude de détermination, adaptation, comparaison et application des tests de rupture multivariés basés sur les fonctions de profondeur. Les résultats montrent un bon comportement des tests en général. En effet, la puissance des tests augmente avec l'amplitude du changement et avec la taille de l'échantillon. Les performances des tests considérés sont, en général, influencées par divers facteurs, comme la taille de l'échantillon et l'amplitude du changement. De plus, le test C est spécialement influencé par l'ordre de grandeur des variables. Les tests QIA, QIB et Z peuvent être problématiques pour les échantillons de petite taille et ils surestiment l'erreur de première espèce. Pour des faibles amplitudes de changement, les tests considérés ne détectent pas le point de rupture quelle que soit la taille de l'échantillon. En général, les tests M, T et W donnent les meilleurs résultats et peuvent être recommandés en pratique. Pour les échantillons de petite taille, la fonction de profondeur MD est suggérée pour les tests M et T tandis que la fonction de la profondeur TD est préférée pour le test W.

Les résultats de l'étude de l'évolution de la dépendance entre les principales caractéristiques de crues montrent que le comportement des mesures de dépendance varie en fonction du modèle hydrologique, du modèle climatique global et du scénario d'augmentation des gaz à effet de serre. De plus, le comportement des séries résultantes de tau de Kendall et de rho de Spearman est similaire. Les séries de dépendance sont stationnaires et présentent des tendances partielles.

Enfin, on a réalisé une étude d'adaptation de l'analyse fréquentielle régionale basée sur le modèle de l'indice de crue multivarié à un cas réel. L'objectif est de développer les aspects pratiques d'une étude de simulation réalisée par Chebana et Ouarda (2009). Une approche d'estimation pratique des quantiles régionaux multivariés se basant sur l'indice de crue a été proposée. Pour vérifier la performance de cette approche, une validation croisée a été réalisée. Les résultats montrent que le biais du modèle proposé dépend aussi bien du risque p, de la discordance bivariée et de la discordance univarée. Le modèle performe bien sur le cas d'études qu'on a considérés.

Enfin, bien que le modèle d'indice de crue multivarié est développé dans Chebana et Ouarda (2009b) dans le cadre de l'AF régionale multivariée, il n'est appliqué sur un cas réel que dans la présente thèse. Le fait que ces études soient nouvelles en hydrologie montre qu'il peut y avoir des limitations et aussi des perspectives de recherche dans ces directions.

5.2 Perspectives de recherche

Pour un risque p, le quantile bivarié estimé par l'AF multivariée est représenté par un nombre infini des couples (par exemple les couples (Q,V)). Le choix du couple optimal qui peut être considéré dans un cas pratique n'a pas été sujet à des travaux approfondis. Ce choix peut être relié au coût de réalisation, à l'objectif du projet (production hydro-électrique, protection contre les inondations, dimensionnement d'un réseau de drainage,...) ou aux caractéristiques hydrologiques de la station. Une étude approfondie qui permet de déterminer le ou les couples optimaux à considérer lors d'une étude pratique est envisageable.

Les techniques d'imputation des données manquantes considérées dans ma thèse représentent les méthodes les plus prometteuses qui existent dans la littérature. Le développement de nouvelles méthodes mieux adaptées au contexte d'AF et au cadre multivarié est souhaitable. On peut, par exemple, développer une méthode d'imputation basée sur la modélisation de la dépendance par les copules.

Dans le cadre de l'étape de vérification des hypothèses de base de l'AF multivariée, on a étudié les tests de détection des points de rupture basés sur les fonctions de profondeur. L'étude d'autres types de tests comment, par exemple, les tests basés sur l'approche bayésienne est une alternative intéressante.

Malgré que la crue est généralement caractérisée par trois variables soit Q, V et D, dans ma thèse, on s'est limité au cas bivarié puisque, à nos connaissances, il n'existe pas encore des outils fiables qui permettent de modéliser la dépendance trivariée convenablement. Le développement d'un tel outil nous permet d'avoir une information plus complète sur l'évènement extrême.

6 Notation

Symbole	Définition
A	Matrice des caractéristiques
A1	Scénario climatique
AF	Analyse fréquentielle
AIC et BIC	Critères d'adéquatement numériques
ANN	Réseaux de neurones artificiels
BB	Block bootstrap
BCa	Méthode du biais corrigé et accéléré
B1 et B2	Scénario climatique
BMU	meilleures unités de reconnaissance
C	Test de Cramér
CGCM	Modèle climatique globale couplé (Canada)
E	*E* Matrice des paramètres
ECHAM	Modèle climatique globale(Allemagne)
EM	méthode d'Espérance-Maximisation
ENSO	Oscillation australe d'El Nino
F_X et F_Y	Fonctions de répartitions de X et Y
D	Durée de la crue
G_1 et G_2	Fonctions de profondeur
HADCM	Modèle couplé du centre Hadley (Royaume-Uni)
K-NN	Méthode des K-voisins les plus proches
LI	Interpolation linéaire
M	Test M
M	$m = n\text{-}s$
MD	Mahalanobis
MI	Méthode d'imputation multiple
MS	Substitution par la moyenne
MRB	Biais relatif moyen
N	Longueur de la série
P	Risque
PW	Pre-whitening
Q	Pointe de la crue
$Q_{V,Q}$	Quantile bivarié de Q et V
$q_{V,Q}$	Courbe de quantile de Q et V
QIA et QIB	Tests d'indice de qualité (tests

REGEM	Méthode EM régularisée
RRMSE	Erreur quadratique moyenne
S	Position du point de rupture
SD	Simplicial
SOM	Carte auto-organisatrice
SRT	Arbre de régression pas-à-pas
T	Test T
TD	Half-space
TFPW	Trend-free pre-whitening
u et v	Vecteur entre 0 et 1
V	Volume de la crue
VC1 et VC 2	Correction de la variance 1 et 2
W	Test de Wilcox
Z	Test de Zhang
X et Y	Séries multivariées
x_i, y_i et z_i	Valeurs observées
$\hat{x}_t$	Valeur estimée
δ	Paramètre de localisation dans les fonctions de profondeurs
μ_{iX} et μ_{iY}	Paramètres locaux de position
ε	Erreur

7 Références

Abebe, A. J., D. P. Solomatine et R. G. W. Venneker (2000). "Application of adaptive fuzzy rule-based models for reconstruction of missing precipitation events." *Hydrological Sciences Journal* 45(3): 425-436.

Adebayo, A. J. et R. Rustum (2012). "Self-organising map rainfall-runoff multivariate modelling for runoff reconstruction in inadequately gauged basins." *Hydrology Research* 43(5): 603-617.

Alila, Y. (1999). "A hierarchical approach for the regionalization of precipitation annual maxima in Canada." *Journal of Geophysical Research D: Atmospheres* 104(D24): 31645-31655.

Alila, Y. (2000). "Regional rainfall depth-duration-frequency equations for Canada." *Water Resources Research* 36(7): 1767-1778.

Allison, P. D. (2001). *Missing Data*, SAGE Publications. 93 Pages

ASCE (1996). *Hydrology Handbook*. New York, American Society of Civil Engineers. 784 Pages

Ashklar, F. (1980). Partial duration series models for flood analysis. Montreal, Qc, Canda, Ecole Polytech of Montreal.

Azen, S. P., M. van Guilder et M. A. Hill (1989). "Estimation of parameters and missing values under a regression model with non-normally distributed and non-randomly incomplete data." *Statistics in Medicine* 8(2): 217-228.

Baringhaus, L. et C. Franz (2004). "On a new multivariate two-sample test." *Journal of Multivariate Analysis* 88(1): 190-206.

Bates, B., Z. W. Kundzewicz, S. Wu et J. Palutikof (2008). Climate Change and Waterr. Technical Paper of the Intergovenmental Panel on Climate Change, Intergovernmental Panel on Climate Change, IPCC Secretariat. Geneva: 210 pp.

Bayley, G. V. et J. M. Hammersley (1946). "The "Effective" Number of Independent Observations in an Autocorrelated Time Series." *Supplement to the Journal of the Royal Statistical Society* 8(2): 184-197.

Ben Aissia, M. A., F. Chebana et T. B. M. J. Ouarda (2014b). "Missing data imputation in univariate and multivariate hydrological frequency analysis – Review and application." *À soumettre*.

Ben Aissia, M. A., F. Chebana et T. B. M. J. Ouarda (2014d). Multivariate shift testing for hydrological applications. Québec, INRS-ETE (Rapport à déposer).

Ben Aissia, M. A., F. Chebana, T. B. M. J. Ouarda, P. Bruneau et M. Barbet (2014c). "Bivariate index-flood model for a northern case study." *Hydrological Sciences Journal* Accepté.

Ben Aissia, M. A., F. Chebana, T. B. M. J. Ouarda, L. Roy, P. Bruneau et M. Barbet (2014a). "Dependence evolution of hydrological characteristics, applied to floods in a climate change context in Quebec." *Journal of hydrology* En révision.

Ben Aissia, M. A., F. Chebana, T. B. M. J. Ouarda, L. Roy, G. Desrochers, I. Chartier et É. Robichaud (2011). "Multivariate analysis of flood characteristics in a climate change context of the watershed of the Baskatong reservoir, Province of Québec, Canada." *Hydrological Processes* 26(1): 130-142.

Bennis, S., F. Berrada et N. Kang (1997). "Improving single-variable and multivariable techniques for estimating missing hydrological data." *Journal of Hydrology* 191(1–4): 87-105.

Besag, J. (1975). "Statistical Analysis of Non-Lattice Data." *Journal of the Royal Statistical Society. Series D (The Statistician)* 24(3): 179-195.

Bisson, J. L. et F. Roberge (1983). Prévision des apports naturels: Expérience d'Hydro-Québec. Paper presented at the Workshop on flow predictions, . Toronto.

Bobée, B. et F. Ashkar (1991). *The gamma family and derived distributions applied in hydrology*, Water Resources Publications. 203 Pages

Burn, D. H. (1990). "Evaluation of regional flood frequency analysis with a region of influence approach." *Water Resources Research* 26(10): 2257-2265.

Chebana, F. (2013). Multivariate Analysis of Hydrological Variables. *Encyclopedia of Environmetrics*, John Wiley & Sons, Ltd. DOI: 10.1002/9780470057339.vnn044

Chebana, F. et T. B. M. J. Ouarda (2007). "Multivariate L-moment homogeneity test." *Water Resour. Res.* 43.

Chebana, F. et T. B. M. J. Ouarda (2008). "Depth and homogeneity in regional flood frequency analysis." *Water Resources Research* 44(11).

Chebana, F. et T. B. M. J. Ouarda (2009a). "Multivariate quantiles in hydrological frequency analysis." *Environmetrics* 9999(9999): n/a.

Chebana, F. et T. B. M. J. Ouarda (2009b). "Index flood–based multivariate regional frequency analysis." *Water Resour. Res.* 45.

Chebana, F. et T. B. M. J. Ouarda (2011a). "Multivariate quantiles in hydrological frequency analysis." *Environmetrics* 22(1): 63-78.

Chebana, F. et T. B. M. J. Ouarda (2011b). "Depth-based multivariate descriptive statistics with hydrological applications." *Journal of Geophysical Research: Atmospheres* 116(D10): D10120.

Chebana, F., T. B. M. J. Ouarda, P. Bruneau, M. Barbet, S. El-Adlouni et M. Latraverse (2009). "Multivariate homogeneity testing in a northern case study in the province of Quebec, Canada." *Hydrological Processes* 23(12): 1690-1700.

Chebana, F., T. B. M. J. Ouarda et T. C. Duong (2013). "Testing for multivariate trends in hydrologic frequency analysis." *Journal of Hydrology* 486(0): 519-530.

Chow, G. C. et A.-L. Lin (1976). "Best Linear Unbiased Estimation of Missing Observations in an Economic Time Series." *Journal of the American Statistical Association* 71(355): 719-721.

Coulibaly, P. et N. D. Evora (2007). "Comparison of neural network methods for infilling missing daily weather records." *Journal of Hydrology* 341(1–2): 27-41.

Cramér, H. (1928). "On the composition of elementary errorsé: II, Statistical applications." *Skandinavisk Aktuarietidskrift* 11: 141-180.

Cunnane, C. (1987). *Review of statistical models for flood frequency estimation.*

Dalrymple, T. (1960). "Flood frequency analysis." *U.S. Geol. Surv. Water Supply Pap.* 1543 A: 80p.

De Michele, C., G. Salvadori, M. Canossi, A. Petaccia et R. Rosso (2005). "Bivariate statistical approach to check adequacy of dam spillway." *Journal of Hydrologic Engineering* 10(1): 50-57.

Dempster, A. P., N. M. Laird et D. B. Rubin (1977). "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society. Series B (Methodological)* 39(1): 1-38.

Deng, J. L. (1989). "Introduction to Grey System Theory." *J. Grey Syst.* 1(1): 1-24.

DiCiccio, T. J. et B. Efron (1996). "Bootstrap confidence intervals " *Statistical Science* 11: 189-228.

Douglas, E. M., R. M. Vogel et C. N. Kroll (2000). "Trends in floods and low flows in the United States: impact of spatial correlation." *Journal of Hydrology* 240(1-2): 90-105.

Durrans, S. R. et S. Tomic (1996). "Regionalization of low-flow frequency estimates: An Alabama case study." *Journal of the American Water Resources Association* 32(1): 23-37.

Easterling, D. R. et T. C. Peterson (1995). "A new method for detecting undocumented discontinuities in climatological time series." *International Journal of Climatology* 15(4): 369-377.

Enders, C. K. (2012). *Applied Missing Data Analysis*, Guilford Publications. 377 Pages

Erol, S. (2011). "Time-Frequency Analyses of Tide-Gauge Sensor Data." *Sensors* 11(4): 3939-3961.

Escalante-Sandoval, C. A. et J. A. Raynal-Villaseñor (1998). "Multivariate estimation of floods: The trivariate Gumbel distiribution." *Journal of Statistical Computation and Simulation* 61(4): 313-340.

Filippini, F., G. Galliani et L. Pomi (1994). "The estimation of missing meteorological data in a network of automatic stations." *Transactions on Ecology and the Environmental Modelling & Software* 4: 283-291.

Fleig, A. K., L. M. Tallaksen, H. Hisdal et D. M. Hannah (2011). "Regional hydrological drought in north-western Europe: linking a new Regional Drought Area Index with weather types." *Hydrological Processes* 25(7): 1163-1179.

Fortin, J. P., R. Moussa, C. Bocquillon et J. P. Villeneuve (1995). "Hydrotel, a distributed hydrological model compatible with remote sensing and geographical information systems." *Revue des Sciences de l'Eau* 8(1): 97-124.

Frane, J. (1976). "Some simple procedures for handling missing data in multivariate analysis." *Psychometrika* 41(3): 409-415.

Genest, C., K. Ghoudi et L.-P. Rivest (1995). "A semiparametric estimation procedure of dependence parameters in multivariate families of distributions." *Biometrika* 82(3): 543-552.

Genest, C. et L.-P. Rivest (1993). "Statistical Inference Procedures for Bivariate Archimedean Copulas." *Journal of the American Statistical Association* 88(423): 1034-1043.

Gleason, T. et R. Staelin (1975). "A proposal for handling missing data." *Psychometrika* 40(2): 229-252.

Goel, N. K., S. M. Seth et S. Chandra (1998). "Multivariate modeling of flood flows." *Journal of Hydraulic Engineering* 124(2): 146-155.

Graham, J. W. (2012). *Missing Data: Analysis and Design*, Springer. 342 Pages

Grimaldi, S. et F. Serinaldi (2006). "Asymmetric copula in multivariate flood frequency analysis." *Advances in Water Resources* 29(8): 1155-1167.

Gyau-Boakye, P. et G. A. Schultz (1994). "Filling gaps in runoff time series in West Africa." *Hydrological Sciences Journal* 39(6): 621-636.

Hamed, K. H. et A. Ramachandra Rao (1998). "A modified Mann-Kendall trend test for autocorrelated data." *Journal of Hydrology* 204(1-4): 182-196.

Han, Y. et N. Li (2010). Interpolation of missing hydrological data based on BP-Neural Networks. *Information Science and Engineering (ICISE)*.

Hipel, K. W. et A. I. McLeod (2005). *Time Series Modelling of Water Resources and Environmental Systems*, Elsevier Science. 1013 Pages

Honaker, J. et G. King (2010). "What to do about missing values in time-series cross-section data." *American Journal of Political Science* 54(2): 561-581.

Hopke, P. K., C. Liu et D. B. Rubin (2001). "Multiple Imputation for Multivariate Data with Missing and Below-Threshold Measurements: Time-Series Concentrations of Pollutants in the Arctic." *Biometrics* 57(1): 22-33.

Hosking, J. R. M. et J. R. Wallis (1993). "Some statistics useful in regional frequency analysis." *Water Resources Research* 29(2): 271-281.

Hosking, J. R. M. et J. R. Wallis (1997). *Regional frequency analysis : an approach based on L-moments*. Cambridge, Cambridge University Press. xiii, 224 p.

Huang, C. et J. R. G. Townshend (2003). "A stepwise regression tree for nonlinear approximation: Applications to estimating subpixel land cover." *International Journal of Remote Sensing* 24(1): 75-90.

Hughes, D. A. et V. Smakhtin (1996). "Daily flow time series patching or extension: a spatial interpolation approach based on flow duration curves." *Hydrological Sciences Journal* 41(6): 851-871.

IPCC (2001). Climate Change 2001: Synthesis Report. A Contribution of Working Groups I, II, and III to the Third Assessment Report of the Integovernmental Panel on Climate Change [Watson, R.T. and the Core Writing Team (eds.)]. Cambridge, United Kingdom, and New York, NY, USA, Cambridge University: 398 p.

Jeffrey, S. J., J. O. Carter, K. B. Moodie et A. R. Beswick (2001). "Using spatial interpolation to construct a comprehensive archive of Australian climate data." *Environmental Modelling & Software* 16(4): 309-330.

Kalteh, A. M. et P. Hjorth (2009). "Imputation of missing values in a precipitation–runoff process database." *Hydrology Research* 40(4): 420–432.

Kao, S. C. et N. B. Chang (2012). "Copula-based flood frequency analysis at ungauged basin confluences: Nashville, tennessee." *Journal of Hydrologic Engineering* 17(7): 790-799.

Karmakar, S. et S. P. Simonovic (2009). "Bivariate flood frequency analysis. Part 2: A copula-based approach with mixed marginal distributions." *Journal of Flood Risk Management* 2(1): 32-44.

Kelly, E., F. Sievers et R. McManus (2004). "Haplotype frequency estimation error analysis in the presence of missing genotype data." *BMC Bioinformatics* 5(1): 188.

Kelly, K. et R. Krzysztofowicz (1997). "A bivariate meta-Gaussian density for use in hydrology." *Ashkar, F.* 11(1): 17-31.

Kendall, M. G. (1975). *Rank Correlation Methods*, Griffin, London. 202 Pages

Khaliq, M. N., T. B. M. J. Ouarda, P. Gachon, L. Sushama et A. St-Hilaire (2009). "Identification of hydrological trends in the presence of serial and cross correlations: A review of selected methods and their application to annual flow regimes of Canadian rivers." *Journal of Hydrology* 368(1-4): 117-130.

Khaliq, M. N., T. B. M. J. Ouarda, J. C. Ondo, P. Gachon et B. Bobée (2006). "Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review." *Journal of Hydrology* 329(3–4): 534-552.

Kim, G., M. J. Silvapulle et P. Silvapulle (2007). "Comparison of semiparametric and parametric methods for estimating copulas." *Computational Statistics and Data Analysis* 51(6): 2836-2850.

Kite, G. (1988). *Frequency and Risk Analyses in Hydrology*. Colo., USA, Water Resources Publications. 257 Pages

Kodituwakku, S., R. A. Kennedy et T. D. Abhayapala (2011). Time-frequency analysis compensating missing data for Atrial Fibrillation ECG assessment. *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*.

Kohonen, T., E. Oja, O. Simula, A. Visa et J. Kangas (1996). "Engineering applications of the self-organizing map." *Proceedings of the IEEE* 84(10): 1358-1384.

Kuligowski, R. J. et A. P. Barros (1998). "Using artificial neural networks to estimate missing rainfall data 1." *JAWRA Journal of the American Water Resources Association* 34(6): 1437-1447.

Lettenmaier, D. P. (1980). "Intervention analysis with missing data." *Water Resour. Res.* 16(1): 159-171.

Li, J. et R. Y. Liu (2004). "New Nonparametric Tests of Multivariate Locations and Scales Using Data Depth." *Statistical Science* 19(4): 686-696.

Linacre, E. (1992). *Climate Data and Resources - A Reference and Guide.* Routledge, London and New York. 384 Pages

Lins, H. F. et J. R. Slack (1999). "Streamflow trends in the United States." *Geophysical Research Letters* 26(2): 227-230.

Little, R. J. A. et D. B. Rubin (2002). *Statistical Analysis With Missing Data*. New Jersey, Wiley Interscience Publication. 381 Pages

Liu, R. Y. et K. Singh (1993). "A Quality Index Based on Data Depth and Multivariate Rank Tests." *Journal of the American Statistical Association* 88(421): 252-260.

Lund, R. et J. Reeves (2002). "Detection of Undocumented Changepoints: A Revision of the Two-Phase Regression Model." *Journal of Climate* 15(17): 2547-2554.

Mann, H. B. (1945). "Nonparametric tests against trend." *Econometric* 13: 245-259.

Marlinda, A. M., M. S. Siti et H. Sobri (2010). "Restoration of Hydrological Data in the Presence of Missing Data via Kohonen Self Organizing Maps." *InTech*: 223-242.

Molenberghs, G. et M. Kenward (2007). *Missing Data in Clinical Studies*, Wiley. 526 Pages

Moore, R. E. (1979). *Methods and Applications of Interval Analysis*. Philadelphia, USA, SIAM. 190 Pages

Mwale, F. D., A. J. Adeloye et R. Rustum (2012). "Infilling of missing rainfall and streamflow data in the Shire River basin, Malawi – A self organizing map approach." *Physics and Chemistry of the Earth, Parts A/B/C* 50–52(0): 34-43.

Ng, W., U. Panu et W. Lennox (2009). "Comparative Studies in Problems of Missing Extreme Daily Streamflow Records." *Journal of Hydrologic Engineering* 14(1): 91-100.

Nguyen, V. T. V. et G. Pandey (1996). "A new approach to regional estimation of floods in Quebec." *Proceedings of the 49th Annual Conference of the CWRA*: 587-596.

Ouarda, T. B. M. J. (2013). Hydrological Frequency Analysis, Regional. *Encyclopedia of Environmetrics*, John Wiley & Sons, Ltd. DOI: 10.1002/9780470057339.vnn043

Ouarda, T. B. M. J., C. Girard, G. S. Cavadias et B. Bobée (2001). "Regional flood frequency estimation with canonical correlation analysis." *Journal of Hydrology* 254(1-4): 157-173.

Ouarda, T. B. M. J., M. Hache, P. Bruneau et B. Bobee (2000). "Regional flood peak and volume estimation in northern Canadian basin." *Journal of Cold Regions Engineering* 14(4): 176-191.

Peterson, H. M., J. L. Nieber et R. Kanivetsky (2011). "Hydrologic regionalization to assess anthropogenic changes." *Journal of Hydrology* 408(3–4): 212-225.

Peterson, T. C., D. R. Easterling, T. R. Karl, P. Groisman, N. Nicholls, N. Plummer, S. Torok, I. Auer, R. Boehm, D. Gullet, L. Vincent, R. Heino, H. Tuomenvirta, O. Mestre, T. Szentimrey, J. Salinger, E. J. Førland, I. Hanssen-Bauer, H. Alexandersson, P. Jones et D. Parker (1998). "Homogeneity adjustments of in situ atmospheric climate data: A review." *International Journal of Climatology* 18(13): 1493-1517.

Raman, H. et N. Sunilkumar (1995). "Multivariate modelling of water resources time series using artificial neural networks." *Hydrological Sciences Journal* 40(2): 145-163.

Rao, A. R. et K. Hamed (2000). *Flood Frequency Analysis*. Boca Raton, CRC Press. 376 Pages

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*, John Wiley & Sons, Inc. 358 Pages

Sadri, S. et D. H. Burn (2011). "A Fuzzy C-Means approach for regionalization using a bivariate homogeneity and discordancy approach." *Journal of Hydrology* 401(3-4): 231-239.

Salarpour, M., Z. Yusop, F. Yusof, S. Shahid et M. Jajarmizadeh (2013). "Flood frequency analysis based on t-copula for Johor river, Malaysia." *Journal of Applied Sciences* 13(7): 1021-1028.

Salvadori, G. et C. De Michele (2010). "Multivariate multiparameter extreme value models and return periods: A copula approach." *Water Resources Research* 46(10): W10501.

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. London, Chapman & Hall. 448 Pages

Schneider, T. (2001). "Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values." *Journal of Climate* 14(5): 853-871.

Seidou, O., J. J. Asselin et T. B. M. J. Ouarda (2007). "Bayesian multivariate linear regression with application to change point models in hydrometeorological variables." *Water Resources Research* 43(8): W08401.

Seidou, O. et T. B. M. J. Ouarda (2007). "Recursion-based multiple changepoint detection in multiple linear regression and application to river streamflows." *Water Resour. Res.* 43(7): W07404.

Shiau, J. T. (2003). "Return period of bivariate distributed extreme hydrological events." *Stochastic Environmental Research and Risk Assessment* 17(1-2): 42-57.

Shih, J. H. et T. A. Louis (1995). "Inferences on the association parameter in copula models for bivariate survival data." *Biometrics* 51(4): 1384-1399.

Simonovic, S. P. (1995). "Synthesizing missing streamflow records on several Manitoba streams using multiple nonlinear standardized correlation analysis." *Hydrological Sciences Journal* 40(2): 183-203.

Sklar, A. (1959). "Fonctions de répartition à n dimensions et leurs marges." *Publ. Inst. Statist. Univ. Paris* 8: 229-231.

Srebotnjak, T., G. Carr, A. de Sherbinin et C. Rickwood (2012). "A global Water Quality Index and hot-deck imputation of missing data." *Ecological Indicators* 17(0): 108-119.

Stedinger, J. R. et G. D. Tasker (1986). "Regional hydrologic analysis, 2 model-error estimators, estimation of sigma and log-pearson type 3 distributions." *Water Resources Research* 22(10): 1487-1499.

Teegavarapu, R. S. V. et V. Chandramouli (2005). "Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records." *Journal of Hydrology* 312(1–4): 191-206.

Vincent, L. A. (1998). "A Technique for the Identification of Inhomogeneities in Canadian Temperature Series." *Journal of Climate* 11(5): 1094-1104.

Volpi, E. et A. Fiori (2014). "Hydraulic structures subject to bivariate hydrological loads: Return period, design, and risk assessment." *Water Resources Research*: n/a-n/a.

Wang, C., N.-B. Chang et G.-T. Yeh (2009). "Copula-based flood frequency (COFF) analysis at the confluences of river systems." *Hydrological Processes* 23(10): 1471-1486.

Wilcox, R. R. (2005). "Depth and a multivariate generalization of the Wilcoxon-Mann-Whitney test." *American Journal of Mathematical and Management Sciences* 25(3-4): 343-363.

Wong, H., B. Q. Hu, W. C. Ip et J. Xia (2006). "Change-point analysis of hydrological time series using grey relational method." *Journal of Hydrology* 324(1–4): 323-338.

Yue, S. (2001). "A bivariate extreme value distribution applied to flood frequency analysis." *Nordic Hydrology* 32(1): 49-64.

Yue, S., T. B. M. J. Ouarda, B. Bobée, P. Legendre et P. Bruneau (1999). "The Gumbel mixed model for flood frequency analysis." *Journal of Hydrology* 226(1-2): 88-100.

Yue, S., P. Pilon et G. Cavadias (2002). "Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series." *Journal of Hydrology* 259(1-4): 254-271.

Yue, S. et C. Wang (2004). "The Mann-Kendall Test Modified by Effective Sample Size to Detect Trend in Serially Correlated Hydrological Series." *Water Resources Management* 18(3): 201-218.

Zhang, C., Z. Lin et J. Wu (2009). "Nonparametric tests for the general multivariate multi-sample problem." *Journal of Nonparametric Statistics* 21(7): 877-888.

Zhang, L. et V. Singh (2006). "Bivariate Flood Frequency Analysis Using the Copula Method." *Journal of Hydrologic Engineering* 11(2): 150-164.

Zhang, X., K. D. Harvey, W. D. Hogg et T. R. Yuzyk (2001). "Trends in Canadian streamflow." *Water Resour. Res.* 37(4): 987-998.

PARTIE 2: ARTICLES ET RAPPORT

8 Article 1 : Missing data imputation in univariate and multivariate hydrological frequency analysis – Review and application

Missing data imputation in univariate and multivariate hydrological frequency analysis – Review and application

M.-A. Ben Aissia[1]*

F. Chebana[1]

T. B. M. J. Ouarda [1,2]

[1] Statistical Hydroclimatology Research Group, 490, Rue de la Couronne, Quebec, Qc, Canada, G1K 9A9.

[2] Institute Center for Water and Environment (iWATER), Masdar Institute of Science and Technology. PO Box 54224, Abu Dhabi, UAE.

* Corresponding author: mohamed.ben.aissia@ete.inrs.ca

March 2014

Abstract

Water resources planning and management require complete data sets of a number of hydrological variables, such as flood peaks and volumes. However, hydrologists are often faced with the problem of missing data (MD) in hydrological databases. Several methods are used to deal with the imputation of MD. During the last decade, multivariate approaches have gained popularity in the field of hydrology, especially in hydrological frequency analysis. The MD problem arises also in multivariate time series. However, treating the MD remains neglected in the multivariate HFA literature wherein the focus has been mainly on the modeling component. For a complete analysis and in order to optimize the use of data, MD should also be treated in the multivariate setting prior to modeling and inference. Imputation of MD in the multivariate hydrological framework is useful to improve the quality of the estimation. Indeed, the dependence between the series is additional information to be included in the imputation process. The objective of the present paper is to highlight the importance of treating MD in multivariate hydrological frequency analysis by reviewing and applying multivariate imputation methods and by comparing univariate and multivariate imputation methods. Two types of multivariate applications are performed in the present work, multi-variable for multiple flood attributes and multi-site for multiple locations. The results indicate that, in both cases, the performance of imputation methods can be improved in the multivariate setting, compared to mean substitution and interpolation methods, especially in the case of highly correlated variables.

1 Introduction

The availability of hydrological data of adequate quality and length is vital for optimal water resources planning and management. In practice, hydrological studies suffer from missing data (MD) caused for instance by budget cuts, equipment failures, errors in measurements and natural hazards (Kalteh et Hjorth 2009). This is generally the case for hydrological variables such as rainfall and streamflow, particularly for extreme conditions such as northern watersheds where equipment failures in remote locations are often identified and fixed with a delay.

The MD in hydrological variables, such as flood peak (Q) and volume (V), are, generally caused by missing streamflow observations during floods. Since the length and the duration of MD are random, three situations may occur: (i) only the V set contains MD and the Q set is complete and vice-versa; (ii) series of V and Q contain MD but not at the same event; (iii) V and Q contain MD at the same event. Indeed, when a data gap exists at a gauged station, associated data are generally observed in one or more neighboring stations such as in other tributaries of the same river.

Generally, hydrological data are characterized by several correlated variables, such as Q and V (e.g. Zhang et Singh 2006; Chebana et Ouarda 2011a). These correlated variables are considered simultaneously in a multivariate framework, see e.g. Chebana (2013) for an explanation of the importance and the justification of jointly considering all variables associated to an event such as in hydrological frequency analysis (HFA).

In HFA, as an essential and commonly used approach for the analysis and prediction of hydrological extreme events, we are frequently faced with the MD problem which can

66 affect the reliability of the results if it is not correctly handled. Generally, before
67 proceeding to any hydrological analysis it is relevant to ensure that the quality of the data
68 is adequate through an exploratory analysis, outlier detection and MD estimation. The
69 presence of MD was highlighted for several hydrometeorological variables such as
70 streamflow (Ng et al. 2009) and precipitation (Makhnin et McAllister 2009).

71 Multivariate HFA is composed of four main steps: (a) Carry out the exploratory analysis
72 including outlier detection, MD estimation and descriptive analysis, (b) Verify HFA
73 assumptions, (c) model the extreme events and estimate the corresponding parameters,
74 and (d) estimate and analyze the risk (see Table 1 for an overview). Recently, Chebana et
75 al. (2013) described these steps and focused on testing multivariate trends in HFA. Under
76 the framework of step (a), the statistical features and the shape of the data are
77 investigated in a multivariate setting by Chebana et Ouarda (2011b). However, MD
78 estimation is generally ignored in multivariate HFA. Consequently, MD estimation in the
79 multivariate setting is a missing step in the HFA.

80 Ignoring the MD estimation in multivariate HFA may lead to a loss of information which
81 may result in inappropriate decisions regarding, for instance, the design of hydraulic
82 infrastructures. Consequently, it is necessary to estimate MD in the multivariate HFA to
83 avoid or reduce the unnecessary construction costs associated with overestimation and
84 the loss of human lives associated with underestimation.

85 The paper is organized as follows. Literature review and general considerations in
86 missing data are presented in Section 2. Section 3 deals with an experimental study

87 describing imputation methods used in the present work and treating two applications on
88 real data. Conclusions are reported in Section 4.

89 2 Literature review

90 The MD estimation is also called infilling (e.g. Abudu et al. 2010), reconstruction (e.g.
91 Kim et Pachepsky 2010), completion (e.g. Ramos-Calzado et al. 2008), patching (e.g.
92 Hughes et Smakhtin 1996) or imputation (e.g. Schneider 2001). It is largely studied in the
93 time domain analysis, i.e. analyzing the data over a time period, (see e.g. Gyau-Boakye et
94 Schultz 1994; Hughes et Smakhtin 1996; Abebe et al. 2000; Han et Li 2010; Marlinda et
95 al. 2010). However, in the frequency domain analysis, i.e. analyzing the data that are
96 periodic over time, the MD handling problem has received less attention (e.g. Peterson et
97 al. 2011).

98 Several imputation methods have been developed to treat MD in both time domain
99 analysis and frequency domain analysis. In time domain analysis, the use of imputation
100 methods has received considerable attention in hydrology and elsewhere in statistics (see
101 e.g. Gleason et Staelin 1975; Jeffrey et al. 2001; Ng et al. 2009; Honaker et King 2010).
102 However, the imputation of MD in frequency domain analysis has received less attention
103 (see e.g. Kelly et al. 2004; Erol 2011; Peterson et al. 2011). Table 2 gives a summary of
104 MD frameworks with some references. In frequency domain analysis studies, MD
105 estimation is largely treated in the univariate setting (e.g. Kodituwakku et al. 2011)
106 whereas in the multivariate setting, studies are relatively rare (e.g. Kelly et al. 2004).

107 Handling MD in HFA is generally ignored or treated separately for each series. The most
108 common practices in HFA are to discard missing observations (see e.g. Overeem et al.

109 2009; Westra et al. 2012) or to impute by the mean of the variable in each missing value
110 (see e.g. Özçelik et Benzeden (2010) and Peterson et al. (2011)). Fleig et al. (2011) used
111 more sophisticated univariate methods, such as, interpolation and regression to estimation
112 MD in HFA. Consequently, MD estimation in multivariate HFA has not been adequately
113 studied yet. Multivariate imputation methods are useful, in particular in hydrology, to
114 improve the quality of the estimation and to provide more accurate imputed values by
115 including variable dependence.

116 In hydrology, imputation techniques of time domain analysis are extensively treated in
117 the univariate and multivariate setting. Table 3 shows an overview of the main imputation
118 techniques in missing hydrological data with a number of references, as well as the
119 advantages and disadvantages of each method. Univariate methods are largely treated in
120 hydrology and include mean or subgroup mean imputation (e.g. Linacre 1992), time
121 series analysis (e.g. Lettenmaier 1980), spatial or temporal interpolation (e.g. Filippini et
122 al. 1994), regression (e.g. Kuligowski et Barros 1998), hot-deck imputation (Srebotnjak
123 et al. 2012) and inverse distance heightening method (ASCE 1996).

124 Multivariate techniques of MD estimation are largely considered in hydrology in time
125 domain analysis as well and can be gathered in three groups: (1) multivariate versions of
126 univariate methods including, for instance, the multivariate version of the regression
127 model (e.g. Simonovic 1995) or the time series analysis (e.g. Bennis et al. 1997); (2) data
128 driven methods including the Artificial Neural Networks (ANNs), e.g. Raman et
129 Sunilkumar (1995) and the k-nearest neighborhood (K-NN) approach, e.g. Kalteh et
130 Hjorth (2009); and (3) model-based approaches including the Expectation-Maximization
131 (EM) (see e.g. Ng et al. 2009) algorithm and the Multiple Imputation (MI) approach (see

132 e.g. Ng et al. 2009). In HFA, handling MD is generally treated in the univariate setting
133 using, for instance, mean substitution (MS) or linear interpolation (LI). The different
134 multivariate imputation methods are applied in hydrological time domain analysis but
135 they have not been used in the HFA multivariate setting.

136 Several studies focused on comparing MD imputation methods in multivariate time series
137 analysis, such as Kalteh et Hjorth (2009) and Coulibaly et Evora (2007). Kalteh et Hjorth
138 (2009) compared five multivariate methods to impute missing values in precipitation-
139 runoff databases. The considered methods are self-organizing maps (SOM) which is an
140 unsupervised ANNs method, multilayered ANN, multivariate K-NN, regularized EM
141 algorithm (REGEM) and MI method. They found that SOM and the multivariate K-NN
142 methods provide the most robust and reliable results. The ability of the SOM to produce
143 reliable estimates of missing hydrological data is also demonstrated in Adebayo et
144 Rustum (2012) and Mwale et al. (2012). On the other hand, Coulibaly et Evora (2007)
145 compared six ANN methods to impute missing daily weather records. These methods are
146 the multilayer perceptron (MLP) network, the time-lagged feed forward network (TLFN),
147 the generalized radial basis function (RBF) network, the recurrent neural network (RNN)
148 and its variant the time delay recurrent neural network (TDRNN), and the counter
149 propagation fuzzy-neural network (CFNN). They found that MLP, TLFN and CFNN
150 methods can provide the most accurate estimates of the missing precipitation values.

151 In the present study several univariate and multivariate methods are used to investigate
152 the performance of multivariate methods against univariate ones in the case of several
153 types of MD patterns and different dependence levels. In the univariate context several
154 methods can be used (Table 3). The MS and LI methods have been used in HFA studies

such as Fleig et al. (2011) and Peterson et al. (2011). Another method that is used is the stepwise regression tree method (SRT) which is a regression model in several nodes. This method was shown to be an efficient technique to impute univariate MD (see e.g. Kim et Pachepsky 2010).

Low accuracy rate in multivariate data

According to Table 3, four multivariate imputation methods are generally used in hydrology in time domain analysis. The first method is the K-NN which not recommended in the multivariate HFA context since, it has low accuracy rate in multivariate data. The second one is the ANN method and its several variants (see e.g. Coulibaly et Evora 2007). According to Kalteh et Hjorth (2009), Adebayo et Rustum (2012) and Mwale et al. (2012), among the ANNs methods, the SOM method leads to good performances. As a third method, we have the EM algorithm. It is originally developed by Dempster et al. (1977) and received some modifications such as the Expectation Conditional Maximization (e.g. Meng et Rubin 1993), the Expectation Conditional Maximization Either (e.g. Liu et Rubin 1994), Alternating Expectation Conditional Maximization (e.g. Meng et Van Dyk 1997), Parameter-Expanded expectation-maximization (e.g. Liu et Rubin 1998) and the REGEM algorithm (e.g. Schneider 2001). The latter is the most commonly used in hydrology (e.g. Kalteh et Hjorth 2009). Finally, the MI method which consists of the imputation of several values (usually 3-5 times) for each MD using an appropriate imputation model (e.g. Patrician 2002). The MI method is rarely used in hydrology (e.g. Kalteh et Hjorth 2009).

176 Several softwares handling MD imputation in the multivariate context are available. In
177 particular, a number of R-packages can be used depending on the imputation method, for
178 instance *AMELIA*, *CLASS*, *MICE*, *NORM*, *VIM*, or *MI*. A number of other packages have
179 also been developed for other environments for example: *S+MissingData* for S-PLUS,
180 *ice* for Stata, *PROCMI* for SAS and *SOM toolbox* for Matlab.

181 3 Experimental study

182 Based on the previous literature review, in this paper, six imputation methods are used for
183 hydrological variables in HFA framework. These techniques are MS, LI, SRT, SOM,
184 REGEM and MI.

185 *3.1 General considerations*

186 Let $(X_i)_{i=1,\cdots,n} = (X_i^{(1)}, X_i^{(2)}, \cdots, X_i^{(d)})'_{i=1,\cdots,n}$ be a continuous d-dimensional sample from a
187 stochastic process $(d \geq 1, n \geq d)$, where "'" denotes the matrix transpose. Let
188 $x_i = (x_i^{(1)}, x_i^{(2)}, \cdots, x_i^{(d)})'$ be an observation from X_i, such as flood peak Q and volume V,
189 at time i. Each series $X^{(k)}, k = 1, \cdots, d,$ can be written as $X^{(k)} = (X_{obs}^{(k)}, X_{mis}^{(k)})$, where $X_{obs}^{(k)}$
190 represents the observed part and $X_{mis}^{(k)}$ denotes the missing part. Before imputing MD it is
191 important to know how and where the MD occurred, in the series. For this we refer
192 respectively to MD mechanisms and MD patterns.

3.1.1 Missing data mechanisms

The MD mechanism determines how the MD is produced. It is a potential factor that could affect the imputation results (Zhu et al. 2012). There are three types of MD mechanisms (Little et Rubin 2002):

- *Missing completely at random (MCAR)*

In this case, MD is unrelated to both the observed or unobserved values in the series. Let $x_i^{(k)}$ be the value of $X^{(k)}$, $k=1,\cdots,d$, at time i and $p\left(x_i^{(k)}\right)$ the probability that $x_i^{(k)}$ is missing. Under MCAR assumption, $p\left(x_i^{(k)}\right)$ can be expressed as

$$p\left(x_i^{(k)} \mid X_{obs}^{(k)}, X_{mis}^{(k)}\right) = p\left(x_i^{(k)}\right) \quad (1)$$

meaning that $p\left(x_i^{(k)}\right)$ is independent of both the observed ($X_{obs}^{(k)}$) and unobserved ($X_{obs}^{(k)}$) parts of $X^{(k)}$.

- *Missing at random (MAR)*

It refers to the case where the incomplete data depend on observed values but not on unobserved values. The probability $p\left(x_i^{(k)}\right)$ can be expressed as

$$p\left(x_i^{(k)} \mid X_{obs}^{(k)}, X_{mis}^{(k)}\right) = p\left(x_i^{(k)} \mid X_{obs}^{(k)}\right) \quad (2)$$

The MAR mechanism occurs when the probability of an observation having a missing value for a component may depend on the available values, but not on the MD themselves.

- *Not missing at random (NMAR)*

In this mechanism, the probability of an observation having a missing value could depend on the observed values as well as unobserved values.

Most of MD in hydrological modeling may be attributed to MCAR or MAR cases (Gill et al. 2007; Kalteh et Hjorth 2009). They are also called ignorable response mechanisms because the reasons for MD can be ignored during the analysis. Model-based methods require the MCAR or MAR assumption (Kalteh et Hjorth 2009).

3.1.2 Missing data pattern

MD imputation methods depend also on the MD pattern which describes where data are observed or missed in the series. Some of these methods apply to any pattern of MD, whereas others are limited to special ones. Several MD patterns exist in the literature such as *multivariate nonresponse* where a set of series are all observed or missing on the same set of cases, *monotone pattern* where the series can be arranged so that all $X^{(j+1)},\cdots,X^{(k)}$ are missing for cases where $X^{(j)}$ is missing, for all $j=1,\cdots,k-1$ and *general pattern* where the MD typically have a random pattern (see e.g. Little et Rubin 2002 for more details).

The methods for handling MD in the case of multivariate nonresponse, or monotone patterns can be easier than the methods for general pattern. In the present study, we consider multivariate hydrological datasets where MD are inside the series. Figure 1 illustrates three possibilities of MD patterns in the bivariate case. The three possibilities are: (i) only one missing values are present in one of the two series, (ii) two missing value is present and located in the same event, and (iii) two missing vales are present but not at the same event. The method to treat these three possibilities is not the same.

3.2 *Considered imputation methods*

3.2.1 *Mean substitution (MS)*

The MS method is the simplest imputation technique. It consists in replacing each missing value in the series $X^{(k)}$, $k=1,\cdots,d$ by the corresponding mean of each component. This imputation method is used in multivariate HFA studies such as Wang et al. (2009) and Kao et Chang (2012)

3.2.2 *Linear Interpolation (LI)*

One of the simplest methods to impute MD is the LI method. It consists of drawing a straight line between observed values before and after the gap and then estimating MD values by interpolation. In univariate regional HFA, this method was used by Fleig et al. (2011). However, to the best knowledge of the authors, it was not used to estimate MD in multivariate HFA.

3.2.3 *Stepwise Regression Trees (SRT)*

The SRT algorithm developed in Huang et Townshend (2003) consists in fitting, in each node of a regression tree, a stepwise regression model (e.g. Miller 2002). Initially, all the data are in the first node of the tree and the partition of the samples into subsets is made recursively until no remaining nodes can be further split. The split of a node into two subsets is made when splitting reduces the residual sum of squares (RSS), such that:

$$RSS = \sum_{i=1}^{n} (\hat{x}_i - x_i), x \in \Re^d \quad (3)$$

where n is the number of observations in the subset; and x_i and $\hat{x}_i$ represent the observed and predicted series from fitting a stepwise regression model. The *RSS* in a given node,

255 before splitting, is noted *RSSN*. The *RSS* of the left and right node after splitting are
256 computed and denoted *RSSL* and *RSSR*, respectively. The sum of *RSSL* and *RSSR* is the
257 total residual sum of squares denoted by *RSST*. The RSST is computed for all possible
258 splits, and the one leading to the smallest *RSST* is conserved and noted *RSSM*. If the split
259 improves the predictions, it will be conserved. Therefore, a measure of the improvement
260 (*I*) from splitting is calculated as

261 $$I = \frac{RSSN - RSSM}{RSSN} \cdot 100\% \quad (4)$$

262 The split is conserved if *I* is larger than the fixed minimum improvement values (I_{min}) and
263 if there are as many observations in the nodes resulting from splitting as the predefined
264 minimum node size (n_{min}). This procedure of splitting continues recursively until all
265 nodes are considered terminal, i.e. the number of observations in that node (n) is smaller
266 than n_{min} or *I* is smaller than I_{min}. To split a node, I_{min} is fixed to 1%. This value was also
267 used in Huang et Townshend (2003) and Beaulieu et al. (2012). We use n_{min} of: 3, 4, 5, 6,
268 7, 10 or 15 observations. The value of n_{min} leading to the model with the best
269 performance is chosen.

270 When SRT is made, the MDs are estimated using a regression model into the
271 corresponding node.

272 ### *3.2.4 Self-Organizing map (SOM)*

273 The SOM method, also called feature map or Kohnen map, is the most widely used of the
274 ANN algorithms designed for unsupervised pattern recognition applications (Kohonen et
275 al. 1996). The ability of the SOM technique in the estimation of missing univariate and

276 multivariate hydrological data was also demonstrated in several studies, see e.g. Adebayo
277 et Rustum (2012) and Mwale et al. (2012). However, these applications are made in time
278 domain analysis. In the present study, this method is applied in the HFA context. The
279 principal goal of the SOM is to transform, in a nonlinear way, a high dimensional input
280 layer to a two dimensional discrete map. A typical structure of a two-dimensional SOM
281 consists of a multi-dimensional input layer and the competitive or output layer. Both of
282 these layers are fully interconnected. The neurons in the input layer are connected to all
283 output layers via weight vectors. Therefore, similar input patterns are represented by the
284 same output neurons, or by one of its neighbors (Back et al. 1998). The SOM can be
285 viewed as a tool for reducing the amount of data by clustering nonlinear statistical
286 relationships between high dimensional data into a simple relationship on a two
287 dimensional display (Kohonen et al. 1996). This method preserves the most important
288 relationship of the original data elements. This implies that, during the mapping, not
289 much information is lost which makes the SOM method a very good tool for prediction.
290 Note that, for prediction values outside the range used for the extrapolation, the SOM
291 method cannot be used because, as with most data-driven methods, SOM is a very poor
292 extrapolator (Adeloye et al. 2011).

293 The training of the SOM is iterative and is hence similar to a sequential training
294 algorithm. In the training algorithm, the whole database is presented to the map before
295 any updates are made while in the sequential training, the weights are updated vector by
296 vector. The SOM procedure can be summarized as follows: at the beginning of the
297 training, weight vectors must be initialized to each neuron and the input vectors are
298 compared with the SOM neurons to find the closest matches which are called the best

299 matching units (BMUs). The Euclidean distance is the most commonly used criterion.
300 This procedure must be iterated several times until the optimal number of iterations is
301 reached or the specified error criteria are attained. The MD are obtained as their
302 corresponding values in the BMU.

303 ### *3.2.5 Regularized Expectation-Maximization algorithm (REGEM)*

304 The Expectation Maximization (EM) algorithm is a very general iterative method for
305 Maximum Likelihood (ML) estimation in MD problems (Dempster et al. 1977). The EM
306 algorithm is proposed for several contexts. The REGEM method (Schneider 2001), as a
307 particular form of the EM algorithm, is based on estimated regression models between
308 missing and available data. The REGEM method is an iterative algorithm based on E step
309 (Expectation) and M step (Maximization). This method consists of: (1) replacing MD by
310 estimated values; (2) estimating, given the observed data and current estimated regression
311 parameters, the mean vector and the covariance matrix of the data; (3) Re-estimating the
312 MD assuming the new parameters are correct. The algorithm consists of iterating these 3
313 steps until convergence i.e. when the variations of the mean vector and the covariance
314 matrix are lower than a predefined threshold. The initial estimates of the model
315 parameters are obtained from the complete database after substituting the missing values
316 with the mean.

317 During the past few decades, the REGEM algorithm was intensively used for MD
318 imputation on multivariate normally distributed series (e.g. Little et Rubin 2002).
319 However, the literature dealing with the application of the REGEM algorithm in
320 hydrology is very sparse (e.g. Kalteh et Hjorth 2009). The REGEM algorithm has not yet
321 been applied in HFA.

3.2.6 *Multiple imputation (MI)*

MI is a fairly straightforward procedure for imputing multivariate MD (Rubin 1987). It provides a useful strategy for dealing with datasets that have MD (Klebanoff et Cole 2008; Sterne et al. 2009). It has been and continues to be developed theoretically and adapted and implemented in numerous statistical problems such as measurement error (e.g. Yucel et Zaslavsky 2005; Reiter et Raghunathan 2007). The basic idea of this method is to first generate several completed data sets by generating several possible values for each MD, and then to analyze each dataset separately. The number of completed datasets to be generated depends on the extent of the missing data. However, according to Schafer (1997), five completed datasets typically provide unbiased estimates. The Schafer's (1999) NORM software, which was used in the present study, uses the data augmentation algorithm to generate five possible values for each MD. The multivariate normal distribution is used to generate imputations. The data augmentation algorithm treats parameters and MD as random variables and simulates random values of parameters and MD from their conditional distribution.

like the REGEM method, the MI technique has been used intensively for MD imputation on multivariate normally distributed variables (e.g. Little et Rubin 2002). However, its application in hydrology is sparse (e.g. Kalteh et Hjorth 2009) especially in multivariate HFA.

3.3 *Employed Softwares*

In the present paper, Matlab codes are developed for MS, Li and SRT methods. The SOM imputation method was carried out by the SOM toolbox which can be downloaded from http://research.ics.aalto.fi/software/somtoolbox/. The Matlab code used for REGEM

345 method can be downloaded from http://www.clidyn.ethz.ch/imputation/ index.html while
346 for MI method, the R-package *NORM* is used.

347 ### *3.4 Performances of imputation methods*

348 To evaluate the accuracy of the imputation methods, their performances are evaluated by
349 a jackknife resampling procedure. It consists in considering each value as a missing one
350 by removing it temporarily from the series. The criteria employed to evaluate the
351 performances are the Relative Root-Mean Squared Error $(RRMSE)$ (see e.g. Chebana et
352 Ouarda 2008) and the mean relative bias (MRB) (see e.g. Beaulieu et al. 2012) defined
353 by:

354 $$RRMSE = \frac{100}{n}\sqrt{\sum_{i=1}^{n}\left(\frac{\hat{x}_i - x_i}{x_i}\right)^2},\ x_i \neq 0 \qquad (5)$$

355 $$MRB = \frac{100}{n}\sum_{i=1}^{n}\left(\frac{\hat{x}_i - x_i}{x_i}\right), x_i \neq 0 \qquad (6)$$

356 where $\hat{x}_i$ is the imputed value and x_i is the observed one.

357 These performance measures were chosen to provide a measure for the deviation of the
358 estimated values from the observations (RRMSE) and to indicate whether the imputation
359 method tends to overestimate or underestimate the observations (MRB).

360 ### *3.5 Applications*

361 In this section, the previously developed imputation methods are applied on two
362 situations. The first one deals with different hydrological variables for the same site,
363 whereas the second one deals with different sites for the same variable. These two

364 examples are given for illustrative purposes in order to emphasize the MD aspects.
365 However, in a real-world study, the analysis should be more extensive and consider the
366 complete HFA procedures.

367 ### *3.5.1. Multi-variable application*

368 In this application, the main flood characteristics are considered, i.e. Q, V and D
369 (duration) on three stations characterized by their natural regime. These stations are
370 located in the Côte Nord region of the province of Quebec, Canada. The first station,
371 namely *Moisie* station (reference number 072301), is located on the Moisie River at 1.5
372 km upstream of the QNSLR bridge with a drainage area of 19 012 km^2. Data series of Q,
373 V and D are available from 1979 to 2004 with missing values in 1999 and 2000. The
374 *Magpie* station (reference number 073503) is the second station, which is located at the
375 outlet of Magpie Lake. The drainage basin of the *Magpie* station has an area of 7 201 km^2
376 and complete data are available from 1979 to 2004. The third station is the *Romaine*
377 station (reference number 073801) located at 16.4 km from the Chemin-de-fer bridge on
378 Romaine River, with a drainage area of 12 922 km^2. Available Q, V and D series are from
379 1979 to 2004 with no MD. Figure 2 and Table 4 present respectively the location and
380 general information about the considered stations. The correlations between Q, V and D
381 for each station are presented in Table 5.

382 For each station, the univariate imputation methods, i.e. MS, LI and SRT are applied to
383 each series of Q, V and D. For the multivariate imputation methods, the considered series
384 are (Q,V), (V,D) and (Q,D). The performance of the imputation methods is evaluated
385 using the two stations with no MD, i.e. *Magpie* and *Romaine* station, and the imputation
386 methods are applied to estimate MD in *Moisie* station.

387 To evaluate the performance of an imputation method, it is assumed that only one value
388 can be missing in each series. Consequently, three situations can occur: (i) only one series
389 has MD (Case (i) in Figure 1); (ii) each series has MD in the same event (Case (ii) in
390 Figure 1); and (iii) each series has MD but not for the same event (Case (iii) in Figure 1).
391 Note that MS and LI performances remain the same in the three situations, since they
392 deal with each series separately. The SRT, SOM and REGEM methods cannot be used in
393 (ii) because they require at least one observation for each event. Consequently, only the
394 MI method is applied for (ii). The situations (i) and (ii) are studied in detail in the present
395 study while (iii) is the subject of an example hereafter since it can be considered as a
396 combination of two cases of (i). The obtained RRMSE and MRB values of imputation
397 methods are given in Table 6.

398 Table 6 indicates that the RRMSE of SRT, REGEM and MI methods are close and range
399 between 16.14% and 17.77% for Q, 14.69% and 17.33% for V and 13.38% and 17.75%
400 for D. Generally, the performance of the MI method is better than that of SRT or
401 REGEM. The SOM method leads to good performances especially for V and D although
402 its performance is generally lower than that of SRT, REGEM or MI. The performances of
403 MS and LI methods are lower compared to the rest of the methods. When comparing
404 *Magpie* and *Romaine* results, we see that the behavior of the imputation methods is
405 similar in both stations. Generally, the MRB is positive or negative but close to zero.
406 Therefore, the imputation methods tend to overestimate the MD.

407 For the situation (ii) the MI method is the only one that can be applied and it leads,
408 generally, to poor performances with the RRMSE ranging between 25.14% and 61.49%.

409 In this case, MD imputation by MI is not recommended and the use of univariate methods
410 will lead to better results.

411 For the third situation (iii), we have (n-1)*(n-1) possibilities of MD location. Applying
412 imputation methods for each possibility and computing the mean of RRMSE and MRB
413 for each method may lead to a global result which is close to that of (i), since the
414 situation (iii) is a combination of two cases of (i). However the situation (iii) raises the
415 question of the sensitivity of imputation methods to MD. To evaluate that sensitivity we
416 eliminate one value in one series (in the present study we chose the value in the middle of
417 the series) and then we use the jackknife resampling procedure with the other series.
418 Table 7 shows the mean and variance of the relative error of 1991's Q, V and D
419 estimation in *Magpie* by eliminating, one by one, the values of Q at *Romaine*. Table 7
420 indicates that the relative error of imputation methods when estimating the missing value
421 in *Magpie* series depends on the existence of MD in the *Romaine* series. This dependence
422 is not very important since the relative error does not exceed ± 0.2.

423 ***3.5.2. Multi-site Application***

424 In the present section we consider the application of the above imputation methods to
425 multi-site situations where each site can be seen as a variable. The three sites used in the
426 previous section are used here. Table 8 presents the dependence between different
427 stations for Q, V and D. We note a high dependence between the three stations. For each
428 individual series, the univariate and multivariate imputation methods are applied. In the
429 present application, we study the same three situations detailed in the previous section.
430 Here we focus on the first and the second situation to evaluate the performance of MD
431 imputation methods. The goal here is to compare the performance of imputation methods

for different dependence values. As in the Multi-variable application, we use the two stations with no MD, namely *Magpie* and *Romaine*. Three multivariate series are used: (Q,V), (V,D) and (Q,D). The jackknife resampling procedure is used in order to evaluate the performance of imputation methods. Results are detailed in Table 9.

Table 9 indicates that, for (Q,V) series, the MI method has the best performance followed by REGEM, SRT and SOM. This result confirms the findings of the previous application. Indeed, in the multi-variable application the considered variables are highly correlated which is the case for Q and V. For (V,D) series, high performance is observed for MI for *Magpie* station followed by REGEM, SRT and SOM for *Romaine* station followed by MI. Table 8 indicates that the dependence between V and D is 0.44 for *Magpie* and 0.18 for *Romaine,* which explains the raison why SRT, MI and REGEM do not perform well in the *Romaine* station. The low dependence between Q and D leads to low performances for all imputation methods and, in this case, the simple methods such as MS or LI could be preferred to the other methods. This application highlights the importance of the dependence between series for MD imputation methods in the multivariate framework. Indeed, in the case of high dependence between variables, multivariate MD imputation methods lead to high performances. This performance decreases with the dependence between variables and in case of no dependence, univariate methods are more appropriate than multivariate ones. In the second situation (ii) results of the MI method lead to low performances whatever the dependence between variables.

In *Moisie* station we have two MDs which are 1999 and 2000. To estimate these MDs we use the three stations in the region. Imputation results are presented in Figure 3 which presents also the Q of the three considered stations. Figure 3 shows that results of

455 REGEM, SRT and MI methods are very close while SOM results are a little bit different.
456 The results of MS and LI are also close. While the LI result is the interpolation in straight
457 line between the observed values just before and after MD, SRT, MI, REGEM and SOM
458 methods seem to reproduce the behavior of Q in *Magpie* and *Romaine* stations for the
459 same period. Indeed, we see an increase of Q in *Magpie* and *Romaine* stations for the first
460 MD and then a decrease for the second. This form is reproduced by the SRT, MI,
461 REGEM and SOM methods. Since Q is highly correlated in the three stations (Table 8)
462 we can expect that the behavior of the three series is similar.

463 Based on the results presented in the two applications, we can conclude that in case of
464 high dependence between variables the MI, SRT and REGEM methods give adequate
465 estimates of MD. The MI method seems to give the best estimation but the two other
466 methods lead to an almost similar performance. The performance of the SOM method is
467 lower than those of MI, SRT and REGEM and better than LI and MS methods. This
468 situation can change in case of low dependence between variables, where the SOM
469 method can lead to better performances than those of MI, SRT and REGEM. In case of
470 low dependence between variables, the highest performances can be obtained by the LI or
471 MS methods. SOM is the most sensitive method to MD in the multivariate series.

472 4 Conclusions

473 The main objective of this study is to show the importance of MD imputation in
474 multivariate (multi-variable and multi-site) hydrological series, to compare univariate and
475 multivariate imputation methods and to present imputation methods that can be
476 considered in multivariate HFA. Imputation methods reduce the loss of information

477 which may lead to suboptimal results and hence to inappropriate decisions regarding, for
478 instance, risk estimation of extreme event.

479 A number of univariate and multivariate imputation methods are presented and applied to
480 the multivariate HFA context. These methods are generally used in time domain analysis.
481 The application of these methods on a number of multi-variable and multi-site series in
482 the Côte Nord of Quebec, Canada, indicates that using MI, SRT and REGEM imputation
483 methods can improve the performance relative to the imputation when variables are
484 highly correlated. On the basis of the above comparison it can be recommended to
485 consider the MI, REGEM or SRT methods to impute multivariate correlated data and the
486 SOM method in case of low dependence. The dependence between variables has a major
487 impact on the imputation quality.

488 In the present study we focused on the bivariate case. Since imputation methods used
489 additional information from available data to estimate MD, the use of high dimensional
490 data can increase the performance of imputation methods especially in multivariate
491 regional HFA were available data of a homogenous region can be useful for the
492 imputation purpose.

493 **Acknowledgments**

494 This study was sponsored mainly by the Natural Sciences and Engineering Research
495 Council (NSERC) of Canada and the Canada Research Chair Program.

496

497

References

Abebe, A. J., D. P. Solomatine and R. G. W. Venneker (2000). "Application of adaptive fuzzy rule-based models for reconstruction of missing precipitation events." *Hydrological Sciences Journal* 45(3): 425-436.

Abudu, S., A. S. Bawazir and J. P. King (2010). "Infilling missing daily evapotranspiration data using neural networks." *Journal of Irrigation and Drainage Engineering* 136(5): 317-325.

Adebayo, A. J. and R. Rustum (2012). "Self-organising map rainfall-runoff multivariate modelling for runoff reconstruction in inadequately gauged basins." *Hydrology Research* 43(5): 603-617.

Adeloye, A. J., R. Rustum and I. D. Kariyama (2011). "Kohonen self-organizing map estimator for the reference crop evapotranspiration." *Water Resources Research* 47(8): W08523.

ASCE (1996). *Hydrology Handbook*. New York, American Society of Civil Engineers. 784 Pages

Azen, S. P., M. van Guilder and M. A. Hill (1989). "Estimation of parameters and missing values under a regression model with non-normally distributed and non-randomly incomplete data." *Statistics in Medicine* 8(2): 217-228.

Back, B., K. Sere and H. Vanharanta (1998). "Managing complexity in large data bases using self-organizing maps." *Accounting, Management and Information Technologies* 8(4): 191-210.

Beaulieu, C., S. Gharbi, T. B. M. J. Ouarda, C. Charron and M. Aissia (2012). "Improved Model of Deep-Draft Ship Squat in Shallow Waterways Using Stepwise Regression Trees." *Journal of Waterway, Port, Coastal, and Ocean Engineering* 138(2): 115-121.

Bennis, S., F. Berrada and N. Kang (1997). "Improving single-variable and multivariable techniques for estimating missing hydrological data." *Journal of Hydrology* 191(1–4): 87-105.

Bobée, B. and F. Ashkar (1991). *The gamma family and derived distributions applied in hydrology*, Water Resources Publications. 203 Pages

Chebana, F. (2013). Multivariate Analysis of Hydrological Variables. *Encyclopedia of Environmetrics*, John Wiley & Sons, Ltd. DOI: 10.1002/9780470057339.vnn044

Chebana, F. and T. B. M. J. Ouarda (2008). "Depth and homogeneity in regional flood frequency analysis." *Water Resources Research* 44(11): n/a-n/a.

Chebana, F. and T. B. M. J. Ouarda (2011a). "Multivariate quantiles in hydrological frequency analysis." *Environmetrics* 22(1): 63-78.

Chebana, F. and T. B. M. J. Ouarda (2011b). "Depth-based multivariate descriptive statistics with hydrological applications." *Journal of Geophysical Research: Atmospheres* 116(D10): D10120.

Chebana, F., T. B. M. J. Ouarda and T. C. Duong (2013). "Testing for multivariate trends in hydrologic frequency analysis." *Journal of Hydrology* 486(0): 519-530.

Chow, G. C. and A.-L. Lin (1976). "Best Linear Unbiased Estimation of Missing Observations in an Economic Time Series." *Journal of the American Statistical Association* 71(355): 719-721.

Coulibaly, P. and N. D. Evora (2007). "Comparison of neural network methods for infilling missing daily weather records." *Journal of Hydrology* 341(1–2): 27-41.

Cunnane, C. and V. Singh (1987). Review of statistical models for flood frequency estimation. *Hydrologic Frequency Analysis*. V. P. Singh, Reidel: 49-95.

Dempster, A. P., N. M. Laird and D. B. Rubin (1977). "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society. Series B (Methodological)* 39(1): 1-38.

546 Erol, S. (2011). "Time-Frequency Analyses of Tide-Gauge Sensor Data." *Sensors* 11(4): 3939-
547 3961.
548 Filippini, F., G. Galliani and L. Pomi (1994). "The estimation of missing meteorological data in a
549 network of automatic stations." *Transactions on Ecology and the Environmental*
550 *Modelling & Software* 4: 283-291.
551 Fleig, A. K., L. M. Tallaksen, H. Hisdal and D. M. Hannah (2011). "Regional hydrological
552 drought in north-western Europe: linking a new Regional Drought Area Index with
553 weather types." *Hydrological Processes* 25(7): 1163-1179.
554 Frane, J. (1976). "Some simple procedures for handling missing data in multivariate analysis."
555 *Psychometrika* 41(3): 409-415.
556 Gill, M. K., T. Asefa, Y. Kaheil and M. McKee (2007). "Effect of missing data on performance of
557 learning algorithms for hydrologic predictions: Implications to an imputation technique."
558 *Water Resour. Res.* 43(7): W07416.
559 Gleason, T. and R. Staelin (1975). "A proposal for handling missing data." *Psychometrika* 40(2):
560 229-252.
561 Gyau-Boakye, P. and G. A. Schultz (1994). "Filling gaps in runoff time series in West Africa."
562 *Hydrological Sciences Journal* 39(6): 621-636.
563 Han, Y. and N. Li (2010). Interpolation of missing hydrological data based on BP-Neural
564 Networks. *Information Science and Engineering (ICISE)*.
565 Honaker, J. and G. King (2010). "What to do about missing values in time-series cross-section
566 data." *American Journal of Political Science* 54(2): 561-581.
567 Hopke, P. K., C. Liu and D. B. Rubin (2001). "Multiple Imputation for Multivariate Data with
568 Missing and Below-Threshold Measurements: Time-Series Concentrations of Pollutants
569 in the Arctic." *Biometrics* 57(1): 22-33.
570 Huang, C. and J. R. G. Townshend (2003). "A stepwise regression tree for nonlinear
571 approximation: Applications to estimating subpixel land cover." *International Journal of*
572 *Remote Sensing* 24(1): 75-90.
573 Hughes, D. A. and V. Smakhtin (1996). "Daily flow time series patching or extension: a spatial
574 interpolation approach based on flow duration curves." *Hydrological Sciences Journal*
575 41(6): 851-871.
576 Jeffrey, S. J., J. O. Carter, K. B. Moodie and A. R. Beswick (2001). "Using spatial interpolation
577 to construct a comprehensive archive of Australian climate data." *Environmental*
578 *Modelling & Software* 16(4): 309-330.
579 Kalteh, A. M. and P. Hjorth (2009). "Imputation of missing values in a precipitation–runoff
580 process database." *Hydrology Research* 40(4): 420–432.
581 Kao, S. C. and N. B. Chang (2012). "Copula-based flood frequency analysis at ungauged basin
582 confluences: Nashville, tennessee." *Journal of Hydrologic Engineering* 17(7): 790-799.
583 Kelly, E., F. Sievers and R. McManus (2004). "Haplotype frequency estimation error analysis in
584 the presence of missing genotype data." *BMC Bioinformatics* 5(1): 188.
585 Khaliq, M. N., T. B. M. J. Ouarda, J. C. Ondo, P. Gachon and B. Bobée (2006). "Frequency
586 analysis of a sequence of dependent and/or non-stationary hydro-meteorological
587 observations: A review." *Journal of Hydrology* 329(3–4): 534-552.
588 Kim, J.-W. and Y. A. Pachepsky (2010). "Reconstructing missing daily precipitation data using
589 regression trees and artificial neural networks for SWAT streamflow simulation." *Journal*
590 *of Hydrology* 394(3–4): 305-314.
591 Kite, G. (1988). *Frequency and Risk Analyses in Hydrology*. Colo., USA, Water Resources
592 Publications. 257 Pages
593 Klebanoff, M. A. and S. R. Cole (2008). "Use of multiple imputation in the epidemiologic
594 literature." *American Journal of Epidemiology* 168(4): 355-357.

595 Kodituwakku, S., R. A. Kennedy and T. D. Abhayapala (2011). Time-frequency analysis
596 compensating missing data for Atrial Fibrillation ECG assessment. *Acoustics, Speech and*
597 *Signal Processing (ICASSP), 2011 IEEE International Conference on.*
598 Kohonen, T., E. Oja, O. Simula, A. Visa and J. Kangas (1996). "Engineering applications of the
599 self-organizing map." *Proceedings of the IEEE* 84(10): 1358-1384.
600 Kuligowski, R. J. and A. P. Barros (1998). "Using artificial neural networks to estimate missing
601 rainfall data 1." *JAWRA Journal of the American Water Resources Association* 34(6):
602 1437-1447.
603 Lettenmaier, D. P. (1980). "Intervention analysis with missing data." *Water Resour. Res.* 16(1):
604 159-171.
605 Linacre, E. (1992). *Climate Data and Resources - A Reference and Guide.* Routledge, London
606 and New York. 384 Pages
607 Little, R. J. A. and D. B. Rubin (2002). *Statistical Analysis With Missing Data.* New Jersey,
608 Wiley Interscience Publication. 381 Pages
609 Liu, C. and D. B. Rubin (1994). "The ECME Algorithm: A Simple Extension of EM and ECM
610 with Faster Monotone Convergence." *Biometrika* 81(4): 633-648.
611 Liu, C. and D. B. Rubin (1998). "Ellipsoidally symmetric extensions of the general location
612 model for mixed categorical and continuous data." *Biometrika* 85(3): 673-688.
613 Makhnin, O. V. and D. L. McAllister (2009). "Stochastic precipitation generation based on a
614 multivariate autoregression model." *Journal of Hydrometeorology* 10(6): 1397-1413.
615 Marlinda, A. M., M. S. Siti and H. Sobri (2010). "Restoration of Hydrological Data in the
616 Presence of Missing Data via Kohonen Self Organizing Maps." *InTech*: 223-242.
617 Meng, X.-L. and D. Van Dyk (1997). "The EM Algorithm—an Old Folk-song Sung to a Fast
618 New Tune." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*
619 59(3): 511-567.
620 Meng, X. L. and D. B. Rubin (1993). "Maximum likelihood estimation via the ECM algorithm: A
621 general framework." *Biometrika* 80(2): 267-278.
622 Miller, A. (2002). *Subset Selection in Regression*, Taylor & Francis. 256 Pages
623 Mwale, F. D., A. J. Adeloye and R. Rustum (2012). "Infilling of missing rainfall and streamflow
624 data in the Shire River basin, Malawi – A self organizing map approach." *Physics and*
625 *Chemistry of the Earth, Parts A/B/C* 50–52(0): 34-43.
626 Ng, W., U. Panu and W. Lennox (2009). "Comparative Studies in Problems of Missing Extreme
627 Daily Streamflow Records." *Journal of Hydrologic Engineering* 14(1): 91-100.
628 Overeem, A., T. A. Buishand and I. Holleman (2009). "Extreme rainfall analysis and estimation
629 of depth-duration-frequency curves using weather radar." *Water Resources Research*
630 45(10): W10424.
631 Özçelik, C. and E. Benzeden (2010). "Regionalization approaches for the periodic parameters of
632 monthly flows: a case study of Ceyhan and Seyhan River basins." *Hydrological*
633 *Processes* 24(22): 3251-3269.
634 Patrician, P. A. (2002). "Multiple imputation for missing data†‡." *Research in Nursing & Health*
635 25(1): 76-84.
636 Peterson, H. M., J. L. Nieber and R. Kanivetsky (2011). "Hydrologic regionalization to assess
637 anthropogenic changes." *Journal of Hydrology* 408(3–4): 212-225.
638 Raman, H. and N. Sunilkumar (1995). "Multivariate modelling of water resources time series
639 using artificial neural networks." *Hydrological Sciences Journal* 40(2): 145-163.
640 Ramos-Calzado, P., J. Gómez-Camacho, F. Pérez-Bernal and M. F. Pita-López (2008). "A novel
641 approach to precipitation series completion in climatological datasets: application to
642 Andalusia." *International Journal of Climatology* 28(11): 1525-1534.
643 Rao, A. R. and K. H. Hamed (2000). *Flood Frequency Analysis.* Boca Raton, CRC Press. 376
644 Pages

Reiter, J. P. and T. E. Raghunathan (2007). "The Multiple Adaptations of Multiple Imputation." *Journal of the American Statistical Association* 102(480): 1462-1471.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*, John Wiley & Sons, Inc. 358 Pages

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. London, Chapman & Hall. 448 Pages

Schafer, J. L. (1999). "NORM: Multiple Imputation of Incomplete Multivariate Data under a Normal Model, version 2. Software for Windows 95/98/NT, available at: http://www.stat.psu.edu/jls/misoftwa.html."

Schneider, T. (2001). "Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values." *Journal of Climate* 14(5): 853-871.

Simonovic, S. P. (1995). "Synthesizing missing streamflow records on several Manitoba streams using multiple nonlinear standardized correlation analysis." *Hydrological Sciences Journal* 40(2): 183-203.

Srebotnjak, T., G. Carr, A. de Sherbinin and C. Rickwood (2012). "A global Water Quality Index and hot-deck imputation of missing data." *Ecological Indicators* 17(0): 108-119.

Sterne, J. A., I. R. White, J. B. Carlin, M. Spratt, P. Royston, M. G. Kenward, A. M. Wood and J. R. Carpenter (2009). "Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls." *BMJ (Clinical research ed.)* 338.

Teegavarapu, R. S. V. and V. Chandramouli (2005). "Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records." *Journal of Hydrology* 312(1–4): 191-206.

Wang, C., N.-B. Chang and G.-T. Yeh (2009). "Copula-based flood frequency (COFF) analysis at the confluences of river systems." *Hydrological Processes* 23(10): 1471-1486.

Westra, S., R. Mehrotra, A. Sharma and R. Srikanthan (2012). "Continuous rainfall simulation: 1. A regionalized subdaily disaggregation approach." *Water Resources Research* 48(1): W01535.

Yucel, R. M. and A. M. Zaslavsky (2005). "Imputation of Binary Treatment Variables With Measurement Error in Administrative Data." *Journal of the American Statistical Association* 100(472): 1123-1132.

Yue, S., P. Pilon and G. Cavadias (2002). "Power of the Mann–Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series." *Journal of Hydrology* 259(1–4): 254-271.

Zhang, L. and V. Singh (2006). "Bivariate Flood Frequency Analysis Using the Copula Method." *Journal of Hydrologic Engineering* 11(2): 150-164.

Zhu, B., C. He and P. Liatsis (2012). "A robust missing value imputation method for noisy data." *Applied Intelligence* 36(1): 61-74.

Tables

Table 1: Main HFA steps in the univariate and multivariate frameworks with some references

Main HFA steps	Framework	
	Univariate	Multivariate
(i) Exploratory analysis:	For instance:	For instance:
- Outlier detection	Cuanne and Singh (1987)	Chebana and Ouarda (2011) for outlier detection in descriptive analysis
- Missing data imputation	Rao and Hamed (2000)	**The specific aim of the present paper: missing data imputation in the multivariate setting**
- Descriptive analysis	Kite (1988)	
(ii) Checking the HFA assumptions	For instance:	For instance:
	Yue et al. (2002)	Chebana et al. (2013) for testing multivariate trends in HFA
	Khaliq et al .(2006)	
(iii) Modeling and estimation	For instance:	For instance:
	Cuanne and Singh (1987)	Shiau (2006)
	Bobée and Ashkar (1991)	Zhang and Singh (2006)
(iv) Risk evaluation and analysis	For instance:	For instance:
	Rao and Hamed (2000)	Shiau (2003)
		Chebana and Ouarda (2011)

692

Table 2: Summary of missing data frameworks with some references

Framework		Fields	
		Statistic	Hydrology
Univariate	Time series analysis	Large body of literature: Gelason and Staelin (1975) Chow and Lin (1976) Azen et al. (1989)	Large body of literature: Lettenmayer (1980) Jefferey et al. (2001) Teegavarapu and Chandramouli (2005)
	Frequency analysis	Large body of literature: Erol (2011) Kodituwakku et al. (2011)	Sparce body of literature: Fleig et al. (2011) Peterson et al. (2011)
Multivariate	Time series analysis	Large body of literature: Frane (1976) Hopke et al. (2001) Honaker et al. (2010)	Large body of literature: Ng et al. (2009) Kalteh and Hjorth (2009)
	Frequency analysis	Sparce body of literature: Kelly et al. (2004)	**The specific aim of the present paper**

693

694

695

696

697

Table 3: Overview of imputation methods in MD problems

Technique	Description	When to be used	Advantages	Disadvantages	Studies
Univariate setting					
Mean substitution	Missing data are replaced by the mean	Less than 10% of data are missing	Easy to use	Underestimates the variance and the degree of freedom	Linacre (1992);
Subgroup mean substitution	Missing data are replaced by the mean of a subgroup	When it is easy to define subgroups	Gives better estimates when compared to mean substitution	Underestimates the variance, subgroup are defined arbitrarly	Linacre (1992);
Time series analysis	Determines the model and the corresponding parameters and then estimates missing data	High autocorrelation	Takes into consideration the the temporal variablity in the data	The necessity to define, a priori, the functional form of the relationships	Lettenmaier (1980);
Interpolation	Interpolate two points of data one immediately before the gap and the other soon after the gap and interpolating the missing data	Only suitable in stable periods and short lengths of the gap	Gives better estimates of statistical inference when well used	Limited to special case that rarely occures	Filippini et al. (1994);
Regression	Estimate parameters of the regression and use them to estimate missing data	Data sets exhibiting significant temporal patterns	Estimated data preserves deviation from the mean and the shape of the available	Could distort the number of degrees of freedom. Difficult to use in noisy data sets	Kuligowski and Barros (1998);
Hot-deck imputation	Replace missing data with value from a similar case	Data are missing in certain patterns.	Missing values are replaced with real values	Problematic if no other case is closely related to the missing value	Srebotnjak et al. (2012);
Inverse distance weighting	Define the neighborhood and the weighting parameters. Then estimate missing data by spatial interpolation using weighting	Stations are highly correlated	Gives better estimates of statistical inference when well used	Problematic with the existence of negative autocorrelation	ASCE (1996);
Multivariate setting					
Artifical neural networks	Determine the architecture of the ANN, estimate parameters and estimate missing data.	When assumptions about the missing data mechanism can not be made and in case of non-linear relationships between variables.	Ability to model complex patterns without a prior knowledge of the underlying process.	Nemerous parameters to estimate and gives unrealistic results when such noise is available in the data.	Raman and Sunilkumar (1995)
k-nearest neighbor	Estimate the missing data based on the closest training examples in the feature space.	When the feature space does not require the selection of a predeterminated model.	Flexible and missing values are replaced with real values.	Low accuracy rate in multidimensional data and computation cost is quite high.	Kalteh & Hjorth (2009);
Expectation Maximisation	Estimate model parameters by iterative process that continues until there is convergence	When distribution assumption are realistic.	Increased accurancy if model is correct.	Strict assumptions, comlex algorithm and takes time to converge.	Ng et al. (2009);
Multiple imputation	Specify an appropriate imputation model, estimate more than one imputed value for each of the missing data.	When assumptions are realistic.	The variablility of the imputed values can be considered.	Strict assumptions and takes time to converge.	Ng et al. (2009);

700 **Table 4: General information of *Moisie*, *Magpie* and *Romaine* stations**

701

Station name	Station number	Latitude	Longitude	Period of records (#years)	Missing data	Area (Km^2)	Mean streamflow (m^3s^{-1})
Moisie	072301	50 21 09	-66 11 12	1979-2004 (26)	1999, 2000	19 012	391.62
Magpie	073503	50 41 08	-64 34 43	1979-2004 (26)	-	7 201	163.56
Romaine	073801	50 18 28	-63 37 07	1979-2004 (26)	-	12 922	282.89

702

703

Table 5: Correlations between *Q*,*V* ans *D*

Stations		Variables		
		D	V	Q
Moisie				
	D	1	0.66	-0.07
	V		1	0.59
	Q			1
Magpie				
	D	1	0.44	-0.20
	V		1	0.70
	Q			1
Romaine				
	D	1	0.18	-0.36
	V		1	0.77
	Q			1

704

705

706

Table 6: RRMSE and BRM of univariate and multivariate imputation methods for the multi-variable case

Method		RRMSE		MRB	
		Magpie	Romaine	Magpie	Romaine
			Q		
Univariate	MS	41.09	46.26	9.73	11.93
	LI	32.33	30.73	5.65	4.63
	SRT	16.65	17.68	2.94	2.12
Multivariate	SOM	28.24	27.40	5.13	4.18
	REGEM	16.85	17.77	3.05	2.27
	MI(i)	16.14	17.31	0.83	-4.08
	MI (ii)	56.54	55.55	31.02	25.23
			V		
Univariate	MS	42.54	39.25	12.81	11.49
	LI	37.64	31.08	8.94	6.52
	SRT	17.31	15.26	1.63	2.10
Multivariate	SOM	19.37	17.88	-1.07	2.69
	REGEM	17.33	15.31	1.71	2.16
	MI(i)	16.98	14.69	-0.17	-2.20
	MI (ii)	61.49	49.99	36.91	26.43
			D		
Univariate	MS	28.21	22.59	5.70	6.41
	LI	28.67	22.58	5.48	5.01
	SRT	17.54	13.81	1.94	3.31
Multivariate	SOM	15.04	17.52	-1.58	5.05
	REGEM	17.75	13.77	2.07	3.45
	MI(i)	17.21	13.38	-0.12	-2.50
	MI (ii)	38.42	25.14	21.70	11.04

MI (i) and MI (ii) correspond to the MI method for situations (i) and (ii) respectively (see section 4.1). Gray color indicates the methods with smallest RRMSE or BRM for each variable.

Table 7: Mean and variance of the relative error of imputation methods applied in situation (iii)

	Q		V		D	
	Mean	**Variance**	**Mean**	**Variance**	**Mean**	**Variance**
SRT	-0.055	0.00007	-0.096	0.00002	-0.085	0.0001
SOM	-0.044	0.00547	-0.108	0.00409	-0.087	0.0004
REGEM	-0.052	0.00004	-0.096	0.00002	-0.091	0.0002

Table 8: Correlations between the considered stations

Variables		**Stations**		
		Moisie	**Magpie**	**Romaine**
Q				
	Moisie	1	0.85	0.69
	Magpie		1	0.87
	Romaine			1
V				
	Moisie	1	0.81	0.76
	Magpie		1	0.92
	Romaine			1
D				
	Moisie	1	0.65	0.74
	Magpie		1	0.73
	Romaine			1

721 **Table 9: RRMSE and MRB of univariate and multivariate imputation methods for the multi-site case**

Methods		(Q,V)				(V,D)				(Q,D)			
		RRMSE		MRB		RRMSE		MRB		RRMSE		MRB	
		V	Q	V	Q	V	D	V	D	Q	D	Q	D
a) Magapie													
Univarite	**MS**	42.54	41.09	12.81	9.73	28.21	42.54	5.70	12.81	41.09	28.21	9.73	5.70
	LI	37.64	32.33	8.94	5.65	28.67	37.64	5.48	8.94	32.33	28.67	5.65	5.48
	SRT	29.22	26.82	7.28	4.19	27.28	46.67	5.58	13.68	41.09	28.21	9.73	5.70
Multivariate	**SOM**	34.24	32.31	9.18	5.52	25.56	39.93	2.49	7.35	47.40	26.21	8.67	4.17
	REGEM	29.50	27.34	7.57	4.44	25.91	44.24	5.11	12.63	41.59	28.79	10.30	5.99
	MI(i)	27.77	24.37	3.89	-3.51	23.83	39.91	1.83	1.08	39.31	26.76	6.04	-1.94
	MI (ii)	61.49	43.55	36.91	14.45	38.42	40.82	21.70	9.18	56.54	25.48	31.02	-7.99
b) Romaine													
Univarite	**MS**	39.25	46.26	11.49	11.93	22.59	39.25	6.41	11.49	46.26	22.59	11.93	6.41
	LI	31.08	30.73	6.52	4.63	22.58	31.08	5.01	6.52	30.73	22.58	4.63	5.01
	SRT	24.43	28.49	6.20	3.25	22.88	43.93	6.28	14.33	50.25	25.86	15.56	8.17
Multivariate	**SOM**	27.27	31.71	5.24	3.21	19.26	29.72	1.07	4.04	41.16	31.28	6.16	7.78
	REGEM	24.60	28.70	6.35	3.49	22.89	42.35	6.37	13.60	48.00	24.32	14.35	6.07
	MI(i)	23.35	26.63	3.38	-4.75	20.24	38.58	3.10	2.67	42.25	22.31	8.52	-1.71
	MI (ii)	56.20	52.06	33.72	20.44	34.59	33.71	23.32	1.07	65.42	20.76	36.55	-10.22

722

723 MI (i) and MI (ii) correspond to the MI method for situations (i) and (ii) respectively (see section 4.1). Gray color indicates the
724 methods with smallest RRMSE or BRM for each variable.

Figures

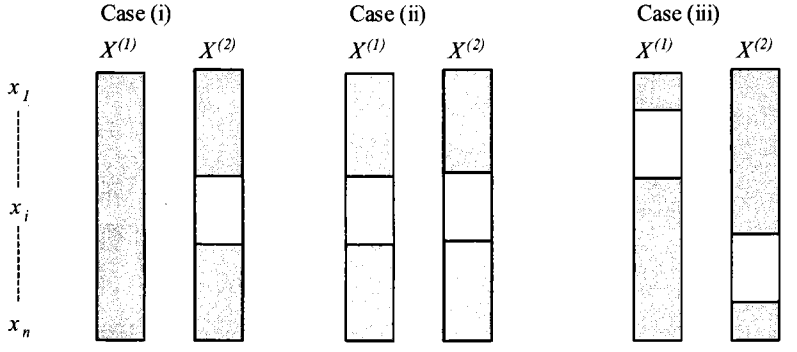
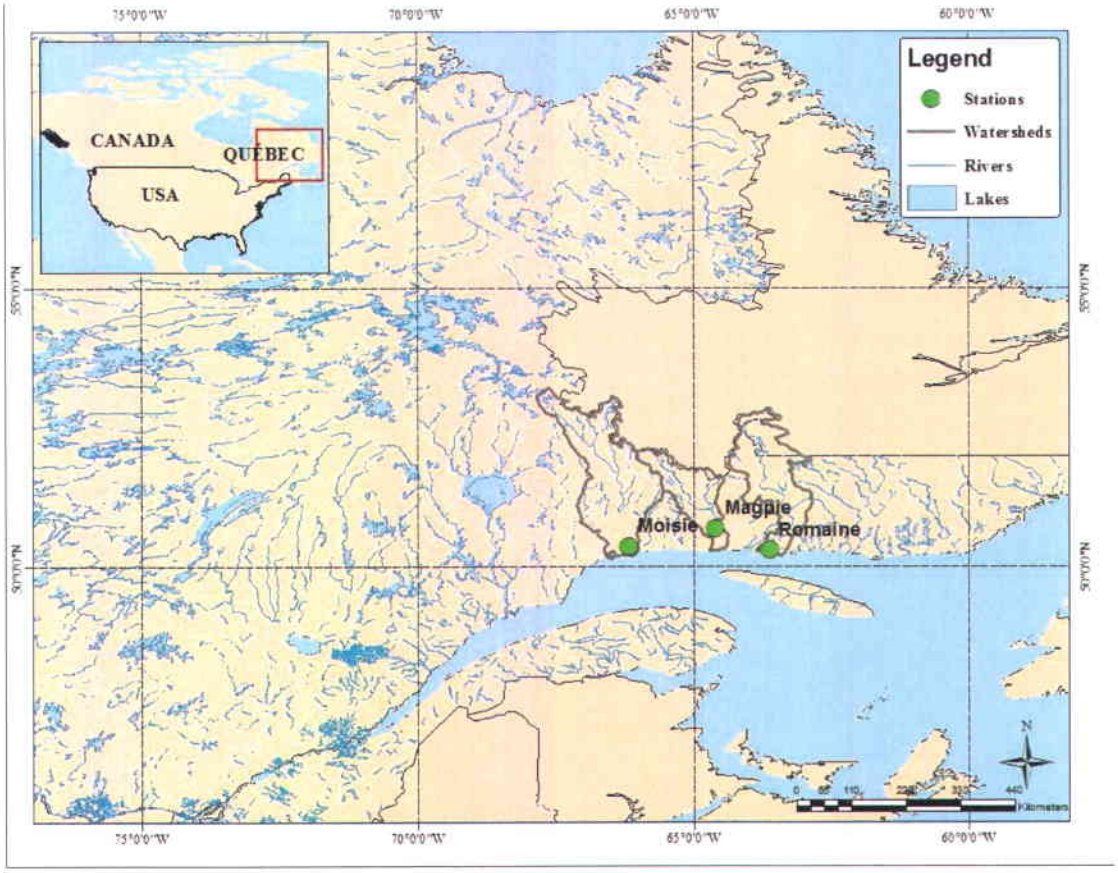


Figure 1: Examples of missing data patterns.

Gray color corresponds to observed data.

731

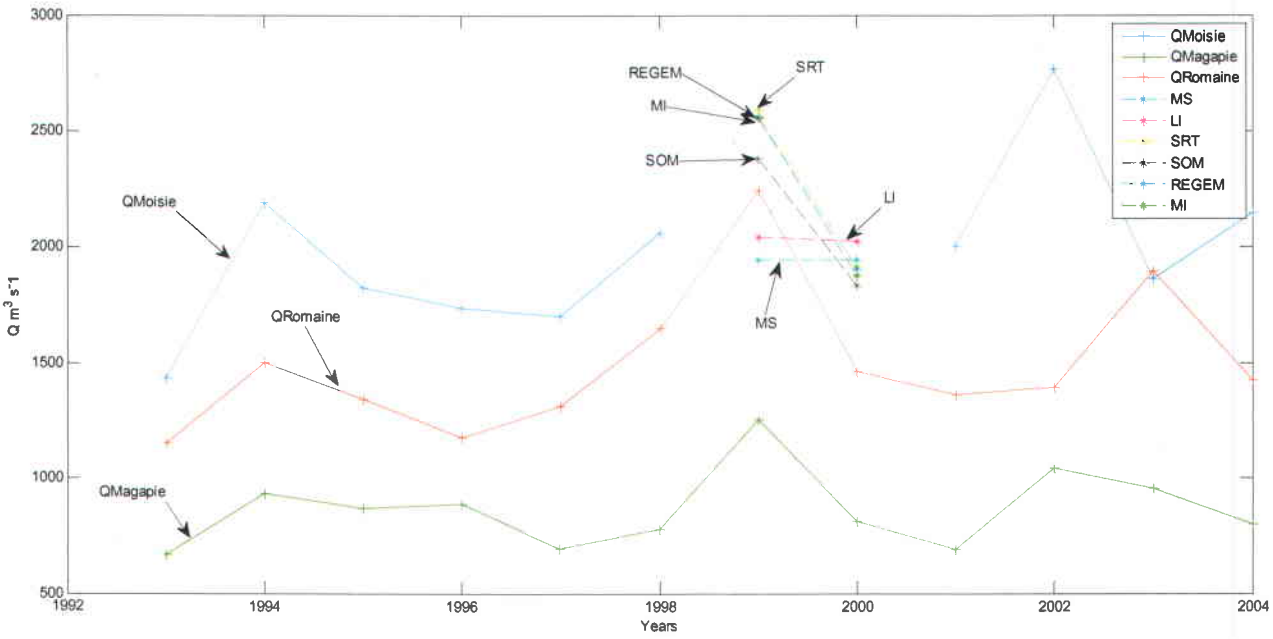


732

733 **Figure 2: Geographical locations of *Moisie*, *Magpie* and *Romaine* stations**

734

735



736

737 **Figure 3: The Q series of the three studied stations and the estimation of MD in**
738 **Moisie**

9 Rapport : Détection des points de rupture multivariés en hydrologie

Détection des points de rupture multivariés en hydrologie

Rapport de recherche

Par

Mohamed Aymen Ben Aissia[1]

Fateh Chebana[1]

Taha B. M. J. Ouarda [1,2]

[1] Groupe de recherche en hydroclimatologie statistique

490, Rue de la Couronne, Québec, Qc, Canada, G1K 9A9.

[2] Institute Center for Water and Environment (iWATER)

Masdar Institute of Science and Technology,

PO Box 54224, Abu Dhabi, UAE.

Mars 2013

Abstract

In hydrology, several events can be described with a number of dependent characteristics, such as floods through their peak, volume and duration. Frequency analysis (FA) can be used to model these events. FA is mainly based on homogeneity, independence and stationarity assumptions. In the multivariate case, the checking of these assumptions is often neglected. In the present paper, we focus on tests to detect shift in multivariate hydrological data. The considered shift tests are mainly based on the notion of depth function except for one which is considered for comparisons purposes. The goal here is to study shift tests in the hydrologic context. A simulation study in hydrological context is performed to evaluate and compare the power of these tests. In addition, a hydrological application is provided to show practical aspect of the considered tests. The power of the considered tests is influenced by various factors, such as the sample size, the shift amplitude, magnitude of the series and the location of the shift in the series.

1. Introduction

To perform most statistical analysis in several fields, such as in hydrology and climatology, a number of fundamental assumptions are required. More precisely, preliminary testing for stationarity, homogeneity and independence is a necessary step in any hydrologic frequency analysis (HFA) (e.g. Rao et Hamed 2000). The presence of shifts in data series is highlighted in several hydrometeorological studies such as floods (Seidou et al. 2007) and precipitations (Beaulieu et al. 2008). Because of the growing evidence of climate change, the common assumption of stationarity of hydrologic phenomena no longer holds. Several recently published works pointed out shifts or trend changes in hydrologic time series (e.g. Burn et Hag Elnur 2002; Woo et Thorne 2003; Salinger 2005). Other reasons could also contribute to jump changing hydrological series, for instance, flood shifts could be due to an abrupt change in a watershed or a river system caused by natural or anthropogenic actions on the physical environment such as deforestation and construction of hydraulic structures (e.g. Bobée et Ashkar 1991; Seidou et al. 2007).

The analysis of multivariate events is of particular interest in many applied fields, such as hydrology. Indeed, complex hydrological events, such as floods, droughts and storms are multivariate events characterized by a number of correlated variables. For instance, volume (V), peak (Q) and duration (D) describe floods (Yue et al. 1999; Shiau 2003; Chebana et Ouarda 2011). The use of univariate HFA can lead to inaccurate estimation of the risk associated to a given event (Yue et al. 1999; Chebana et Ouarda 2011). Recently, several studies adopted the multivariate framework to treat hydrological extreme events, see e.g. Chebana et Ouarda (2011) for a summary and recent references.

The HFA is composed of four main steps: i) descriptive and explanatory analysis, ii) checking the basic assumptions including stationarity, homogeneity and independence, iii) modeling and estimation, and iv) risk evaluation and analysis. In the univariate setting, these steps are extensively treated (e.g. Rao et Hamed 2000; Chebana 2013). In the multivariate context, the first two preliminary steps (i and ii) attracted considerably less attention than the last two. For an overview of step i) in multivariate framework the reader can be referred to Chebana (2013). The hypothesis testing (step ii) is generally ignored in the hydrological literature in multivariate setting, e.g. (Kao et Govindaraju 2007; Song et Singh 2009; Vandenberghe et al. 2010).

The hypothesis testing step has a significant impact on the selection of the appropriate model in multivariate HFA. Indeed, a non-stationarity in the data should include possible trends in some or all parts of the multivariate distribution (copula and margins) (e.g. Chebana et al. 2013). Therefore, ignoring the testing step may lead to inaccurate model and hence to wrong results which lead to inappropriate decisions regarding, for instance, the design of hydrological engineering works. In order to avoid the loss of human lives and property associated to underestimation, or the surplus cost of construction associated to overestimation, it is necessary to include hypothesis testing step in the multivariate HFA (see e.g. Chebana et al. 2012; Chebana 2013).

The notion of non-stationarity is very wide and includes in particular the presence of one or several shifts in the data. Recently, Chebana et al. (2013) provided a review and application of multivariate nonparametric tests for monotonic trends and presented approaches that can be considered as a preliminary step in a complete multivariate HFA. Chebana et al. (2013) mentioned that, for multivariate hydrological data, various type of non-stationarity can be found for which appropriate tests should by reviewed, compared and applied.

The available literature on shift detection is either for hydrological variables but only in the univariate setting or in multivariate setting in a general statistical framework without hydrological focus. In the latter situation, the comparisons and evaluations of the proposed tests are based on scenarios not adapted to the hydrological context (e.g. sample size, distributions) where the comparisons are performed on simple situations and only a limited number of such tests. Therefore, evaluation and comparisons representative of the hydrological reality and constraints by including all the available tests are required. Hence, the aim of this paper is to focus on shift concept of hydrological variables within the multivariate HFA context.

Several shift tests are based on depth function. The latter is a statistical notion to measure the *depth* or *outlyingness* of a given point with respect to a multivariate data cloud or its underlying distribution. Depth functions were developed in the seventies and received increasing interest (Tukey 1975; Liu 1990; e.g. Zuo et Serfling 2000; Mizera et Müller 2004; Zuo et Cui 2005; Lin et Chen 2006; Liu et Singh 2006; Chebana and Ouarda 2008; 2011; Singh et Bárdossy 2012). The depth function provides a scale-standardized measure of the position of any data point relative to the center of the distribution due to its affine-invariant property (Li et Liu 2004). For the location shift, this property allows us to view the depth-based test statistics as scale-standardized measures. Therefore, the depth-based tests can be performed without the difficulty of estimating the variance of the null sampling distributions. Instead, the decision rule is derived by obtaining p-values using the idea of permutation.

The report is organized as follows. Section 2 introduces definitions and notations related to the shift concept. The considered tests are described in section 3. The simulation study to evaluate the performance of these tests is given in Sections 4. Section 5 presents an application of the reviewed tests in the paper on hydrological data. The conclusions of the study and a number of perspectives are reported in Section 6.

2. Shift concept

A shift can be defined as the position where at least one feature of a statistical model (e.g., location, scale, intercept and trend) undergo an abrupt change (Seidou et al. 2007). A large number of techniques can be found in the literature to find the date of a potential shift and to check if the shift is significant or not. Most of the methodologies use statistical hypothesis testing to detect shift in slopes or intercept of linear regression models (Easterling et Peterson 1995; Vincent 1998; Lund et Reeves 2002). For instance, Solow (1987), Easterling et Peterson (1995), Vincent (1998), Lund et Reeves (2002) and Wang (2003) all used the Fisher test to compare a model with and without a shift. The Student and Wilcoxon tests can be applied sequentially to detect shifts in data series (Beaulieu et al. 2007).

Note that not all shift approaches are based on hypothesis testing. For instance, Wong et al. (2006) used the grey relational method (Moore 1979; Deng 1989) for single shift detection in stream flow data series. In some rare cases, other curve fitting methods are used (e.g. Sagarin et Micheli 2001; Bowman et al. 2006). Extensive reviews of shift detection and correction methodologies in hydrology and climate sciences can be found in work by Peterson et al. (1998) and Beaulieu et al. (2007).

To define a shift, let $(x_i)_{i=1,\ldots,n}$ be a given d-variate dataset and $1 < s < n$ be the location of a possible shift. If such s exists, the series is divided in two subsamples with sizes s and $m = n\text{-}s$ such that:

$$\begin{aligned}(y_1,\ldots,y_s) &= (x_1,\ldots,x_s)\\ (z_1,\ldots,z_m) &= (x_{s+1},\ldots,x_n)\end{aligned} \qquad (1)$$

Denote G_1 and G_2 respectively the distribution functions of these two subsamples. The two distributions G_1 and G_2 have the same form, except for location, i.e. $G_1(x) = G_2(x+\delta)$ for all $x \in R^d$ where $\delta \in R^d$ is a constant vector. Consequently, when testing the presence of a shift at a point s of the series $(x_i)_{i=1,\ldots,n}$, the null and alternative hypotheses are respectively:

H_0 : $\delta = 0$ i.e. there is no location shift (2)

H_1 : $\delta \neq 0$ i.e. there are two different subsamples at least in one coponent of δ (3)

3. The selected tests

In the present paper, several tests to detect shift in location are considered. All these tests are based on depth function, except one (the C-test), which is considered for comparison purposes. Three different depth functions are considered namely: Mahalanobis, simplicial and half-space respectively denoted by MD, SD and TD. In principle, each depth-based test can be defined using any available depth function. However, some of these tests are originally defined and their properties are studied on the basis of a specific depth function. Table 1 presents a summary of the presented tests.

3.1. Description of tests

In the present section, the considered multivariate tests to detect shift are described as well as the method to evaluate their performance in the literature.

The Cramér test (The C-test)

The Cramér test is a two-sample test proposed by Baringhaus et Franz (2004). It is a generalisation of the univariate test proposed by Cramér (1928). However it is more appropriate to detect shifts in location. This test is based on difference of Euclidian distances between the

observations of the two different subsamples and the half sum of all Euclidian distances of observations of a same subsample. The corresponding statistic of test is given by

$$C = \frac{sm}{s+m}\left[\frac{1}{sm}\sum_{i=1}^{s}\sum_{j=1}^{m}\|y_i - z_i\| - \frac{1}{2s^2}\sum_{i,j=1}^{s}\|y_i - y_j\| - \frac{1}{2m^2}\sum_{i,j=1}^{m}\|z_i - z_j\|\right] \quad (4)$$

where $\|y_i - z_j\|$ is the Euclidian distance between i^{th} observation of first subsample and j^{th} observation of second subsample. The null hypothesis H_0 is rejected for large value of C. A large value of C means that distance between observations of two subsamples is large and consequently, the two subsamples are different. To calculate a p-value, the bootstrapping method is used.

The M-test

According to Li et Liu (2004), the deepest point of a distribution is a location parameter. Consequently, if G_1 and G_2 are identical distributions, they would have the same deepest point, that is, respectively deepest point θ_{G_1} and θ_{G_2} should be the same. In addition, for a given depth function D, we have $D_{G_2}(\theta_{G_1}) = D_{G_1}(\theta_{G_2})$. However, if there is an important change in location, θ_{G_1} and θ_{G_2} would be different and θ_{G_2} is located far away from the subsample with the distribution G_1 for which the depth value $D_{G_1}(\theta_{G_2})$ with respect to G_1, is smaller, and conversely. Based on this idea, Li et Liu (2004) proposed the statistic:

$$M = \min\left\{D_{G_2}(\theta_{G_1}), D_{G_1}(\theta_{G_2})\right\} \quad (5)$$

Li et Liu (2004) used the Simplicial depth function, but other depth functions can be used. Indeed, Li et Liu (2004) suggest Mahalanobis depth for elliptical distribution. They specified that simplicial and half-space depth can be used with any distribution.

The null hypothesis H_0 is rejected for small value of M. To approximate the corresponding p-value, Li et Liu (2004) proposed Fisher's permutation test (Snedecor et Cochran 1967).

The T-test

Li et Liu (2004) described a graphical approach called DD-plot (for depth-depth) to compare the location of two subsamples. In DD-plot, the two axes represent the depth values of two subsamples. When the two subsamples follow exactly the same distribution, the DD-plot is a diagonal line that passes by the origin as illustrated in Figure 1-a. However, if there is a location change, the graph has a form of leaf with its tip pointing toward the origin (Figure 1-b). the more the location change is important; the closer tip to the origin (Figure 1-c). The T-test is based on an approximation of the distance between the tip and the origin of the DD-plot. Define the set of points:

$$\Omega = \left\{ x_i \mid i \in \{1,...,n\}, \text{there is no } x_j : D_{G_1}(x_j) \geq D_{G_1}(x_i) \text{ and } D_{G_2}(x_j) \geq D_{G_2}(x_i) \right\} \quad (6)$$

Then we find the point $x_{\min}$ of Ω such that:

$$\left| D_{G_1}(x_{\min}) - D_{G_2}(x_{\min}) \right| = \min_{x \in \Omega} \left| D_{G_1}(x) - D_{G_2}(x) \right| \quad (7)$$

If there are several points $x_{\min}$, we take the mean of the corresponding coordinates.

The point found in (7) is an approximation of leaf-tip point of the DD-plot. The test statistic is then given by

$$T = \frac{D_{G_1}(x_{\min}) + D_{G_2}(x_{\min})}{2} \quad (8)$$

Even though, the distance of the leaf-tip to the origin is approximately $\sqrt{2}T$, the use of the statistic T is equivalent. Similar to the M-test, Li et Liu (2004) used the Simplicial depth function

for the T-test and also the Mahalanobis and half-space depths can be used. The p-value is obtained using the Fisher's permutation test.

The Wilcox test (The W-test)

The W-test was developed by Wilcox (2005). As the M-test, the W-test is based on the idea that under null hypothesis, the medians of the two subsamples must be similar. To define the W-test statistic, first the difference of each component is calculated $d_{ij}^{(u)} = z_i^{(u)} - y_j^{(u)}, u = 1,...,d;\ i = 1,...,s;\ j = 1,...,m$ to constitute the vector $d_{ij} = \left(d_{ij}^{(1)},\cdots,d_{ij}^{(d)}\right)$. Wilcox (2005) defined the test statistic by:

$$W = \frac{D_F(\mathbf{0})}{\max\limits_{i=1,...,s;j=1,...,m} D_F(d_{ij})} \tag{9}$$

where F is the distribution of the set of vectors d_{ij} and D is the half-space depth function. Under the null hypothesis, we have $W = 1$, whereas under the alternative hypothesis, $W < 1$. The asymptotic distribution of W is unknown, but Wilcox (2005) proposed some critical values C_μ for significance levels $\mu = 0.01; 0.025; 0.05; 0.10$. The values of C_μ are derived empirically from simulations using a least squares regression method, and under the assumption of normality. The null hypothesis is rejected when W is below than C_μ.

The quality index test (The QIA- and QIB-tests)

Liu et Singh (1993) developed a Wilcoxon-type rank test based on data depth. This test can detect location shift and / or positive scale shift. The statistic of this test is given by:

$$Q_a = \frac{1}{n}\sum_{i=1}^{m} \#\left\{y \in \{y_1,\cdots,y_s\} : D_G(y) \le D_G(z_i)\right\} \tag{10}$$

Under the null hypothesis, $Q_a = 0.5$. If there is a shift in location $Q_a < 0.5$. Liu et Singh (1993) used Mahalanobis depth. Zuo et He (2006) found that under some regularity conditions, the asymptotic distribution of Q_a calculated with Mahalanobis, half-space or projection depth is normal $N(\mu,\sigma^2)$ with mean $\mu = 0.5$ and variance $\sigma^2 = (s^{-1} + m^{-1})/12$. In the present study, the asymptotic (QIA test) and bootstrap (QIB test) methods are used to evaluate the p-value.

The Zhang test (The Z-test)

Zhang et al. (2009) developed a new test based on statistic Q_a defined in (10) of quality index test that detects shifts. The statistic of this test is given by

$$Z = \frac{6}{n}\sum s \times m(Q_a - 0.5)^2 \qquad (11)$$

where Q_a is the quality index statistic defined in (10). Zhang et al. (2009) used the Mahalanobis depth to define Z. To find the asymptotic distribution of Z, we define the matrix A:

$$A = \begin{bmatrix} 1-p_1 & \sqrt{p_1 p_2} \\ \sqrt{p_1 p_2} & 1-p_2 \end{bmatrix} \qquad (12)$$

where $p_i = \frac{n_i}{n}$, i = 1 or 2 and n_i is the number of observations in the i^{th} subsample. Let r be the rank of A. The nonzero eigenvalues of A are denoted by $\lambda_1,..,\lambda_r$. Under H_0, Z follows asymptotically a sum of independent chi-square distributions:

$$Z \approx \lambda_1\chi^2(1) + \lambda_2\chi^2(1) + ... + \lambda_r\chi^2(1) \qquad (13)$$

This relation is also valid for half-space and projection depth functions. The asymptotic method is used to evaluate the corresponding p-value.

3.2. The p-value computation

The p-value of a given test is a simple criterion commonly used by practitioners to decide for the acceptance or rejection of a target null hypothesis. The p-value is based on the distribution of the statistics of the test, generically denoted S. For some of the considered tests in the present study, the asymptotic or the exact distribution of test statistic is unknown or difficult to obtain. Consequently, approximations of the distribution of test statistics, under the null hypothesis are required. To this end, resampling methods are used. In the present paper, permutation method (Snedecor et Cochran 1967) and bootstrap method are used. They are briefly described below where more details can be found in Good (2005) and Hallin et Ley (2006).

To apply the permutation method, the observations should be exchangeable i.e. the observations are independent and identically distributed (see e.g. Efron et Tibshirani 1994). This method consists in permuting a large number n_p times the sample $(x_i)_{i=1,\ldots,n}$ *without replacement*. For each permuted sample, the s first elements constitute first subsample and the remaining ones constitute the second subsample. The test statistic S is calculated for each permutation $\left(S^*_{i,i=1,\ldots,n_p}\right)$.

If the null hypothesis should be rejected for small value of statistic test, the p-value is the proportion of $\left(S^*_{i,i=1,\ldots,n_p}\right)$ smaller or equal to the value S_{obs} obtained from the original sample.

The bootstrap method is similar to the permutation method, except that the sample $(x_i)_{i=1,\ldots,n}$ is resampled *with replacement* and the independence assumption is necessary (see e.g. Efron et Tibshirani 1994).

3.3. Literature based comparisons

Some performance comparisons of the above tests are presented in the literature. The M- and T-tests, given respectively in (5) and (8), have been compared to Hotelling (1947) T^2 test by Li et Liu (2004). The Hotelling's T^2 test is a largely used parametric test to detect shift location (e.g. Ye et al. 2002). For normal samples, the powers of these three tests are found to be comparable, whereas for Cauchy samples, the M- and T- tests are more powerful than the Hotelling's test. Moreover, in this case, the M-test outperformed the T-test.

Liu et Singh (2006) compared also the quality index test (10) to Hotelling's test. For normal samples, the performances of the two tests are similar, while for Cauchy and Exponential samples the quality index test outperformed the Hotelling's test. Baringhaus et Franz (2004) found that the C-test (4) performs almost as well as Hotelling's test for normal and non-normal samples.

These comparisons and evaluations are not appropriate for the hydrological applications, since the adopted data are not representative the hydrological reality where sample sizes are generally short, the distributions are mainly of extremes such as Gumbel and GEV.

4. Simulation study

The objective of this simulation study is to evaluate and compare the performance of the previously presented tests to the hydrological context, such as flood series based on Q and V, small samples encountered in hydrology. In addition, these tests are used on Normal, Cauchy and t distributions which are not commonly used in multivariate HFA.

4.1. Adaptation to floods

The previously presented tests can be applied to hydrological events such as floods, rain storms and droughts. In this paper, we focus on floods. Floods can be described by their peak Q, volume

V and duration D, which can be correlated. Indeed, according for instance to Yue (2001) there is generally a closed correlation between Q and V, between Q and D and a little correlation between Q and D. In the present paper, the above considered tests are used to detect location shifts in Q and V.

According to Sklar's (1959) result, a bivariate distribution can be composed by marginal distributions and a copula. Some previous studies showed that Q and V series can be marginally fitted by a Gumbel distribution (e.g. Yue et al. 1999; Yue 2001; Yue et Rasmussen 2002; Shiau 2003; Chebana et Ouarda 2007). The cumulative Gumbel distribution is given by:

$$F(x) = \exp\left\{-\exp\left(-\frac{x-\beta}{\sigma}\right)\right\},\ x \text{ and } \beta \text{ real},\ \sigma > 0 \qquad (14)$$

Moreover, the dependence between Q and V can be represented by Gumbel logistic model (e.g. Yue et al. 1999; Shiau 2003; Chebana et al. 2009; Aissia et al. 2012), expressed according to the fallowing copula:

$$C_b(x,y) = \exp\left\{-\left[\left(-\log(x)\right)^b + \left(-\log y\right)^b\right]^{1/b}\right\},\ b \geq 1 \text{ and } 0 \leq x,\ y \leq 1 \qquad (15)$$

Note that $b = 1/\sqrt{1-\rho}$ where ρ is the usual correlation coefficient (see e.g. Gumbel et Mustafi 1967; Genest et Rivest 1993).

The presented tests may be affected by several factors. In the simulation study, we study the impact of the record length n (sample size) as well as the degree of change (shift amplitude) in each component of multivariate series.

For the simulation study, we generate samples (Q, V) according to models (14) and (15). By considering the Gumbel distribution as marginal for both Q and V, the corresponding parameters are denoted by:

- σ_{Q1} and β_{Q1} for respectively scale and location parameters for Q of s first observations (before the shift)

- σ_{Q2} and β_{Q2} for respectively scale and location parameters for Q after the shift

Similarly, we defined the parameters of V (σ_V, β_V) and the parameter b of the logistic Gumbel copula.

For the G distribution before shift, we selected the parameters of the Skootamatta basin in Ontario, (Canada) identified by Yue et Rasmussen (2002) and employed for simulation studies by Chebana and Ouarda (2007; 2009). Consequently, σ_{Q1}= 15.85, β_{Q1}= 51.85, σ_{V1}= 300.22 β_{V1} = 1239.8 and m_1= 1.414.

We study the effect of different parameters on the performance of the tests record length (n: sample size) and the amplitude of shifts in location parameters β, since the tests are mainly designed to detect shifts in the location. Usually, the dependence parameter appears in the copula whereas the location and scale parameters are present in the marginal distributions (Hobæk Haff et al. 2010). For location shift, let $G_1(x) = G_2(x+\delta)$ where δ = (δ_Q, δ_V) respectively the shifts in location of Q and in location of V.

4.2. Simulation design

The conducted simulation study is constituted of two steps. As a first step, we generated a large number N of samples to evaluate the effects of different factors on the performance of the tests. Three samples sizes are considered which are n = 30, 50 and 80 corresponding to s=10, 20 and 30 respectively. For each sample size, several amplitudes of location shift are considered which are δ= 10, 20,-20, 40 and 70%. We generated samples as follows:

I. *No change in all parameters*: All the parameters of the distribution are the same before and after the shift. This allows to obtain samples under the null hypothesis (no shift) and therefore, for each record length n, we calculated the probability of first kind error (α);

II. *Change in location parameters:* The distribution before shift (G_1) is the same that after shift (G_2), except for location parameters β in the marginal. We considered three cases:

   a. Change only in location of Q: δ_Q= 10,20,40 and 70%;

   b. Change only in location of V: δ_V = 10,20 and 40%;

   c. Change in the location of Q and V simultaneously: (δ_Q, δ_V)= (10,10), (20,20), (20,-20), (40,40), and (70,70)%.

For the evaluation of p-values, based on the permutation and the bootstrap methods, we used n_p = 500 permutations or bootstrap samples. This value of n_p is proposed by Li et Liu (2004) for the M- and T-tests and it is superior to the value 200 proposed by Baringhaus et Franz (2004) for the C-test.

As a second step of simulation study, we evaluated the performance of each test on the basis of the estimation $\hat{\alpha}$ of α and the power of considered tests. In the present study, we fixed α = 5%. Consequently, we reject H_0 if p-value is less than 5%. As a number of replications, we considered N=3000 which is higher than those used by Li et Liu (2004), Wilcox (2005) and Zhang et al. (2009).

Note that, when simulating, a problem related to the set Ω occurred with the T-test. Indeed, the set Ω given in (6) can be empty. In fact, Ω is rarely empty in general with Simplicial and Mahalanobis depths, but it is often empty with Half-space depth. This issue has not been mentioned or considered in Li et Liu (2004). These cases are excluded from the present computations.

4.3. Simulation results

In order to avoid repetition and for notation simplicity, the depth function will be written in the test index when it is needed. For example, M_{TD}-test is the M-test with TD depth function.

I. First kind error estimation

The estimates $\hat{\alpha}$ of α for the considered tests are presented in Table 2. Since we fixed critical level at $\alpha = 5\%$, a performing test should have $\hat{\alpha}$ as close as possible to 5%. From Table 2, we see that $\hat{\alpha}$ tends generally to 5% when n increases. Values of $\hat{\alpha}$ for M-test are close to 5% except M_{TD} and M_{SD} with the case $(n,s)=(30,10)$. The T- and C-tests have $\hat{\alpha}$ around 5% whatever the sample size. The W-test underestimates α while QIB-, QIA- and Z-tests overestimate it. However, the QIB_{SD}-, QIA_{TD}- and Z_{TD}- tests have $\hat{\alpha}$ higher than 20% when $(n,s)=(30,10)$ which means that they tend to reject H_0 when it is true.

II. Power evaluation

Table 3 summarises the simulation results for shift detection tests for several shift amplitude in Q, V and (Q,V). In general, these results show good behaviour of the tests in terms of power where the power increases with the shift amplitude δ and with the sample size n. In the present paper, a test power is considered high when it exceeds 95%.

For n=30, Table 3 (part a) shows that high powers are, generally, recorded for large shift is amplitudes i.e. $(\delta_Q, \delta_V) = (70,0)$ or $(\delta_Q, \delta_V) = (70,70)$. For M- and T-tests, best powers are recorded with MD depth function. The TD depth function gives best powers for W-, QIA- and Z-tests while for QIB-test, it is with SD depth function. However, as seen before, QIB_{SD}-, QIA_{TD}- and Z_{TD}-tests are problematic with the estimation of α. Note that the depth function that provides best test power is not necessary the one with which the test is originally defined, e.g M- and T-tests. For the C-test, the power depends on the variable on which the shift is occurred. Indeed, a

shift only on Q leads to low power of the C-test which is opposite to when the shift is either on V or in (Q,V). This is due to the difference in the first term in (4), which can be affected by the scale of the series In the case of floods, Q and V series have very different scales where a change in Q does not have a great effect on the test statistic and the opposite for V (and hence for (Q,V)). We can conclude that C-test is more sensitive to a change on V than to a change on Q. This situation has not been revealed in previous studies since the simulations are based on variables of the same nature and scale.

For n=50, from Table 3 (part b), we can see that high powers are obtained starting form (δ_Q, δ_V)=(0,40). For each test, the depth function that leads to best power when n=30 are generally the same when n=50. The powers when n=50 are generally higher than when n=30 with some exceptions: for QIB-, QIA-, and Z-tests with (δ_Q, δ_V)=(0,10), (10,0), (10,10). (0,20), (20,0) or (20,20).

Table 3 (part c) gathers the simulation results of the presented tests when n=80. Results show that high powers are observed starting from (δ_Q, δ_V)=(20,-20) for M-, T- and W_{TD}-tests. For the M-test, results are similar of the three considered depth functions for each shift amplitude whereas for the other tests, depth functions leading to the highest powers for n=80 are the same for n=30 or 50. Generally, performances of the tests increase when the shifts of V and Q have different signs. For instance, powers for (δ_Q, δ_V)= (20,-20) are higher than those of (δ_Q, δ_V)= (20,20) for all tests. Note that C-test power increases with n except when the shift is located only in Q.

From these results one can conclude that, generally best results are obtained by M-, T- and W-tests (with power high or equal to that of the rest of tests). For low sample size, high powers are observed for large shift amplitude (70%) while for large sample size, high powers are observed

starting from (δ_Q, δ_V)=(20,-20)%. For low shift amplitude (10%), low powers are recorded for all the considered tests. Figure 2 illustrates the applicability of considered tests for the combinations of the studied sample sizes and shift amplitudes.

From the present simulation study, the following general observations can be made:

- The C-test is more sensitive to a change on V than on a change on Q;
- For small sample size (n=30), high power is observed only for high shift amplitude;
- For large sample size (n=80), best powers are observed for M-, T- and W-tests;
- The QIB-, QIA- and Z-tests can be problematic especially for low shift amplitudes.
- For first kind error estimation, QIB_{SD}-, QIA- and Z_{TD}-tests are problematic, especially when n=30. Good performances are observed for M-, T-, W- and C-tests whatever the depth function;
- For low shift amplitudes (δ_Q, δ_V)=(0, 10), (10, 0) or (10, 10), powers are low. This means that 10% of one or both location parameters change is not detected by the considered tests;

5. Application

In this section, the previously considered tests are applied on three stations with natural flow data series. These stations are located in the Côte Nord of the province of Québec, Canada. The first one is *Moisie* station (reference number 072301) and is located on Moisie River at 1.5 km upstream of the Québec North Shore Labrador Railway (Q.N.S.L.R.) bridge with a drainage basin area of 19 012 km^2. Data series are available from 1968 to 1998. The *Magpie* station (reference number 073503) is the second station, which is located at the outlet of Magpie Lake. Its drainage basin has an area of 7 201 km^2 and available data are from 1979 to 2004. The third

station is the *Romaine* station (reference number 073801) located at 16.4 km from the Chemin-de-fer bridge on Romaine River, with a drainage basin area of 12 922 km^2 and data series available from 1961 to 2006. Figure 3 and Table 4 present respectively the geographical position and general information about the considered stations.

Spring flood characteristics Q and V are extracted from daily streamflow series for each station. Figure 4 shows time series of Q and V of the three stations. Since these stations are geographically close to each other (Figure 3), it is expected that a shift can be observed in the three stations. From Figure 4 we can see that a shift can be located in Q and V around 1984 for the three stations. Therefore, the previously presented tests are applied for each station in 1984.

Statistics and p-values of the considered tests are summarized in Table 5. Note that for the W-test the conclusion is presented as (1: there is a shift, 0: no shift) instead of the p-value since it is based on critical thresholds (Wilcox 2005).

Results show that all considered tests are in agreement with the existence of shift in *Moisie* station, for instance the p-values of T-, QIB-, QIA-, Z- and C-tests are less than 1%. For *Magpie* station, the M-test is the only test which does not detect the presence of a shift for all depth functions whereas the T-test indicates a shift with all depth functions. This can be explained by the fact that for small sample size (Table 3-a) the power of the M-test is lower than that of the T-test. Considering *Romaine* station, only T_{SD}-, QIB_{TD}-, QIB_{MD}- and Z_{TD}-tests cannot confirm the existence of a shift in the year 1984.

From the results of the three stations one can conclude that, the year 1984 can be considered as a shift for *Moisie* and *Romaine* stations and can be a shift for *Magpie* station. From Figure 4-b one can see that a shift in 1984 is not very clear in Magpie station and the short sample data before the shift can has an affect the power of considered tests. Since these stations are geographically proximate (Figure 3) one can say that 1984 is probably a shift for all these stations.

6. Conclusions

The aim of this paper is to study shift detection in the multivariate setting of hydrological variables by comparing the power of several tests in hydrologic context and by adapting these tests for hydrological practice. Shift detection is required to ensure the validity of HFA assumptions (homogeneity and stationarity) and hence leads to the selection of the appropriate multivariate distribution. All the considered tests are based on data depth, except the C-test, which is considered for comparison purposes. A simulation study, taking into account the hydrological context, is performed to evaluate and compare the power of the considered tests to detect shift in location parameter of Q, V and (Q,V). In addition, these tests are applied to three stations.

In general, the powers of these tests increase with the shift amplitude and with the sample size. However, the QIA-, QIB- and Z-tests may be problematic for small sample sizes and they overestimate the first kind error α. For low shift amplitudes, the considered tests do not perform whatever the sample size. On the basis of the above comparison, and considering the nature of hydrological data, it can be recommended to use the M-, T- and W-tests. More precisely, for small sample sizes, the MD depth function is preferred for M- and T-tests while TD depth function is preferred for W-test.

Application of the considered tests in a real data shows their ability to detect multivariate shift in three stations. From this application we deduce that tests performance can be affected by the length of sub-series before or after the shift.

ACKNOWLEDGEMENTS

The authors thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for the financial support and Marjolaine Dubé for her assistance.

Reference

Aissia, M. A. B., F. Chebana, T. B. M. J. Ouarda, L. Roy, G. Desrochers, I. Chartier and É. Robichaud (2012). "Multivariate analysis of flood characteristics in a climate change context of the watershed of the Baskatong reservoir, Province of Québec, Canada." *Hydrological Processes* 26(1): 130-142.

Baringhaus, L. and C. Franz (2004). "On a new multivariate two-sample test." *Journal of Multivariate Analysis* 88(1): 190-206.

Beaulieu, C., T. B. M. J. Ouarda and O. Seidou (2007). "A review of homogenization techniques for climate data and their applicability to precipitation series." *Hydrological Sciences Journal* 52(1): 18-37.

Beaulieu, C., O. Seidou, T. B. M. J. Ouarda, X. Zhang, G. Boulet and A. Yagouti (2008). "Intercomparison of homogenization techniques for precipitation data." *Water Resources Research* 44(2): W02425.

Bobée, B. and F. Ashkar (1991). "The gamma family and derived distributions applied in hydrology." *Water Resources Publication. Littleton, Colorado, USA.*

Bowman, A. W., A. Pope and B. Ismail (2006). "Detecting discontinuities in nonparametric regression curves and surfaces." *Statistics and Computing* 16(4): 377-390.

Burn, D. H. and M. A. Hag Elnur (2002). "Detection of hydrologic trends and variability." *Journal of Hydrology* 255(1–4): 107-122.

Chebana, F. (2013). Multivariate Analysis of Hydrological Variables. *Encyclopedia of Environmetrics*, John Wiley & Sons, Ltd. DOI: 10.1002/9780470057339.vnn044

Chebana, F., S. Dabo-Niang and T. B. M. J. Ouarda (2012). "Exploratory functional flood frequency analysis and outlier detection." *Water Resources Research* 48(4): W04514.

Chebana, F. and T. B. M. J. Ouarda (2007). "Multivariate L-moment homogeneity test." *Water Resources Research* 43(8).

Chebana, F. and T. B. M. J. Ouarda (2008). "Depth and homogeneity in regional flood frequency analysis." *Water Resources Research* 44(11).

Chebana, F. and T. B. M. J. Ouarda (2009). "Index flood–based multivariate regional frequency analysis." *Water Resources Research* 45(10): W10435.

Chebana, F. and T. B. M. J. Ouarda (2011). "Multivariate quantiles in hydrological frequency analysis." *Environmetrics* 22(1): 63-78.

Chebana, F., T. B. M. J. Ouarda, P. Bruneau, M. Barbet, S. El Adlouni and M. Latraverse (2009). "Multivariate homogeneity testing in a northern case study in the province of Quebec, Canada." *Hydrological Processes* 23(12): 1690-1700.

Chebana, F., T. B. M. J. Ouarda and T. C. Duong (2013). "Testing for multivariate trends in hydrologic frequency analysis." *Journal of Hydrology* 486(0): 519-530.

Cramér, H. (1928). "On the composition of elementary errorsé: II, Statistical applications." *Skandinavisk Aktuarietidskrift* 11: 141-180.

Deng, J. L. (1989). "Introduction to Grey System Theory." *J. Grey Syst.* 1(1): 1-24.

Easterling, D. R. and T. C. Peterson (1995). "A new method for detecting undocumented discontinuities in climatological time series." *International Journal of Climatology* 15(4): 369-377.

Efron, B. and R. J. Tibshirani (1994). *An Introduction to the Bootstrap*, Taylor & Francis.

Genest, C. and L.-P. Rivest (1993). "Statistical Inference Procedures for Bivariate Archimedean Copulas." *Journal of the American Statistical Association* 88(423): 1034-1043.

Good, P. (2005). *Permutation, Parametric and Bootstrap Tests of Hypotheses*, Springer New York. 315

Gumbel, E. J. and C. K. Mustafi (1967). "Some Analytical Properties of Bivariate Extremal Distributions." *Journal of the American Statistical Association* 62(318): 569-588.

Hallin, M. and C. Ley (2006). Permutation Tests. *Encyclopedia of Environmetrics*, John Wiley & Sons, Ltd. DOI: 10.1002/9780470057339.vap009.pub2

Hobæk Haff, I., K. Aas and A. Frigessi (2010). "On the simplified pair-copula construction — Simply useful or too simplistic?" *Journal of Multivariate Analysis* 101(5): 1296-1310.

Hotelling, H. (1947). Multivariate quality control: Illustrated by the air testing of sample bomb sight. *In Selected Techniques of Statistical Analysis for Scientific and Industrial Research and Production and Management Engineering*. McGraw-Hil. New York: 111-184.

Kao, S.-C. and R. S. Govindaraju (2007). "A bivariate frequency analysis of extreme rainfall with implications for design." *J. Geophys. Res.* 112(D13): D13119.

Li, J. and R. Y. Liu (2004). "New Nonparametric Tests of Multivariate Locations and Scales Using Data Depth." *Statistical Science* 19(4): 686-696.

Lin, L. and M. Chen (2006). "Robust estimating equation based on statistical depth." *Statistical Papers* 47(2): 263-278.

Liu, R. Y. (1990). "On a Notion of Data Depth Based on Random Simplices." *The Annals of Statistics* 18(1): 405-414.

Liu, R. Y. and K. Singh (1993). "A Quality Index Based on Data Depth and Multivariate Rank Tests." *Journal of the American Statistical Association* 88(421): 252-260.

Liu, R. Y. and K. Singh (2006). Rank tests for multivariate scale difference based on data depth. *Data Depth: robust multivariate analysis, computational geometry, and applications*. R. Y. S. Liu, K. and Souvaine, D.L., Àmerican Mathematical Society: 17-35.

Lund, R. and J. Reeves (2002). "Detection of Undocumented Changepoints: A Revision of the Two-Phase Regression Model." *Journal of Climate* 15(17): 2547-2554.

Mizera, I. and C. H. Müller (2004). "Location-scale depth." *Journal of the American Statistical Association* 99(468): 949-966.

Moore, R. E. (1979). *Methods and Applications of Interval Analysis*. Philadelphia, USA, SIAM.

Peterson, T. C., D. R. Easterling, T. R. Karl, P. Groisman, N. Nicholls, N. Plummer, S. Torok, I. Auer, R. Boehm, D. Gullet, L. Vincent, R. Heino, H. Tuomenvirta, O. Mestre, T. Szentimrey, J. Salinger, E. J. Førland, I. Hanssen-Bauer, H. Alexandersson, P. Jones and D. Parker (1998). "Homogeneity adjustments of in situ atmospheric climate data: A review." *International Journal of Climatology* 18(13): 1493-1517.

Rao, A. R. and K. H. Hamed (2000). *Flood Frequency Analysis*. Boca Raton, CRC Press. 376

Sagarin, R. and F. Micheli (2001). "Climate change in nontraditional data sets." *Science* 294(5543): 811.

Salinger, M. J. (2005). "Climate Variability and Change: Past, Present and Future – An Overview." *Climatic Change* 70(1-2): 9-29.

Seidou, O., J. J. Asselin and T. B. M. J. Ouarda (2007). "Bayesian multivariate linear regression with application to change point models in hydrometeorological variables." *Water Resources Research* 43(8): n/a-n/a.

Shiau, J. T. (2003). "Return period of bivariate distributed extreme hydrological events." *Stochastic Environmental Research and Risk Assessment* 17(1-2): 42-57.

Singh, S. K. and A. Bárdossy (2012). "Calibration of hydrological models on hydrologically unusual events." *Advances in Water Resources* 38(0): 81-91.

Sklar, A. (1959). *Fonctions de répartition à n dimensions et leurs marges*.

Snedecor, G. W. and W. G. Cochran (1967). *Statistical Methods*, Iowa State University Press.
Solow, A. R. (1987). "Testing for Climate Change: An Application of the Two-Phase Regression Model." *Journal of Climate and Applied Meteorology* 26(10): 1401-1405.
Song, S. and V. P. Singh (2009). "Meta-elliptical copulas for drought frequency analysis of periodic hydrologic data." *Stochastic Environmental Research and Risk Assessment*: 1-20.
Tukey, J. W. (1975). Mathematics and the picturing of data. *International Congress of Mathematicians*, Vancouver, B. C., 1974, Canad. Math. Congress,.
Vandenberghe, S., N. E. C. Verhoest and B. De Baets (2010). "Fitting bivariate copulas to the dependence structure between storm characteristics: A detailed analysis based on 105 year 10 min rainfall." *Water Resour. Res.* 46(1): W01512.
Vincent, L. A. (1998). "A Technique for the Identification of Inhomogeneities in Canadian Temperature Series." *Journal of Climate* 11(5): 1094-1104.
Wang, X. L. (2003). "Comments on "Detection of Undocumented Changepoints: A Revision of the Two-Phase Regression Model"." *Journal of Climate* 16(20): 3383-3385.
Wilcox, R. R. (2005). "Depth and a multivariate generalization of the Wilcoxon-Mann-Whitney test." *American Journal of Mathematical and Management Sciences* 25(3-4): 343-363.
Wong, H., B. Q. Hu, W. C. Ip and J. Xia (2006). "Change-point analysis of hydrological time series using grey relational method." *Journal of Hydrology* 324(1–4): 323-338.
Woo, M.-k. and R. Thorne (2003). "Comment on 'Detection of hydrologic trends and variability' by Burn, D.H. and Hag Elnur, M.A., 2002. Journal of Hydrology 255, 107–122." *Journal of Hydrology* 277(1–2): 150-160.
Ye, N., S. M. Emran, Q. Chen and S. Vilbert (2002). "Multivariate statistical analysis of audit trails for host-based intrusion detection." *IEEE Transactions on Computers* 51(7): 810-820.
Yue, S. (2001). "A bivariate gamma distribution for use in multivariate flood frequency analysis." *Hydrological Processes* 15(6): 1033-1045.
Yue, S., T. B. M. J. Ouarda, B. Bobée, P. Legendre and P. Bruneau (1999). "The Gumbel mixed model for flood frequency analysis." *Journal of Hydrology* 226(1-2): 88-100.
Yue, S. and P. Rasmussen (2002). "Bivariate frequency analysis: Discussion of some useful concepts in hydrological application." *Hydrological Processes* 16(14): 2881-2898.
Zhang, C., Z. Lin and J. Wu (2009). "Nonparametric tests for the general multivariate multi-sample problem." *Journal of Nonparametric Statistics* 21(7): 877-888.
Zuo, Y. and H. Cui (2005). "Depth weighted scatter estimators." *Annals of Statistics* 33(1): 381-413.
Zuo, Y. and X. He (2006). "On the limiting distributions of multivariate depth-based rank sum statistics and related tests." *Annals of Statistics* 34(6): 2879-2896.
Zuo, Y. and R. Serfling (2000). "General notions of statistical depth function." *Annals of Statistics* 28(2): 461-482.

Tables

Table 1: Summary of the presented tests

	Reference	Designed to detect	p-value evaluation	Used depth functions	Comparison from the literature	
					For normal samples	For non-normal samples
C-test Eq. (4)	Baringhaus and Franz (2004)	Location and/or scale shift	Bootstrap	NA	The C-test performs almost as well as Hotelling test	
M-test Eq. (5)	Li and Liu (2004)	Location shift	Permutation	- Half-space - Mahalanobis - Simplicial*	The powers of M-test, T-test and Hotteling tests are comparable	The M-test outperformed the T-test and both are more powerful than the Hotelling test
T- test Eq. (8)	Li and Liu (2004)	Location shift	Permutation	- Half-space - Mahalanobis - Simplicial*		
W-test Eq. (9)	Wilcox (2005)	Location shift	Critical thresholds given in Wilcox (2005)	- Half-space* - Mahalanobis - Simplicial	NA	
Q-test Eq. (10)	Liu and Singh (1993)	Location and/or positive scale shift	Bootstrap or asymptotic	- If asymptotic p-value: Half-space or Mahalanobis* - If bootstrap p-value: Half-space, Mahalanobis* or Simplicial	The performances of the Q- and Hotelling tests are similar	The Q-test outperformed theHotelling one
Z-test Eq. (11)	Zhang et al. (2009)	Multiple location and/or scale shift	Asymptotic	- Half-space - Mahalanobis*	NA	

*with which the test is originally defined

Table 2 : The $\hat{\alpha}$ estimated for the considered tests and for each sample size.

n	*s*	M			T			W			QIB			QIA		Z		C
		TD	MD	SD*	TD	MD	SD*	TD*	MD	SD	TD	MD*	SD	TD	MD*	TD	MD*	
30	10	1.3	5.2	0.1	3.9	5.5	5.0	2.7	0.2	1.5	10.1	6.7	86.5	46.2	19.3	22.0	7.6	5.1
50	20	3.8	5.8	5.4	4.6	5.5	5.3	3.9	0.1	0.4	8.4	6.6	48.0	29.9	12.4	12.1	5.8	5.4
80	30	4.1	5.1	5.0	5.0	5.2	5.1	2.6	0.0	0.2	6.8	6.0	27.1	22.8	9.9	8.2	4.6	6.0

with *n*: sample size, *s*: shift, *: the depth function with which the test is originally defined. Gray color indicates that $\hat{\alpha}$ is close to 5% (between 3% and 7%).

Table 3 : Power comparison for the considered tests to detect shifts in Q, V or (Q,V).

δ_Q	δ_V	M			T			W			QIB			QIA		Z		C
		TD	MD	SD*	TD	MD	SD*	TD*	MD	SD	TD	MD*	SD	TD	MD*	TD	MD*	
a)								$(n,s) = (30,10)$										
0	10	2.5	**10.5**	0.1	7.8	**11.6**	7.8	**7.0**	0.6	4.0	10.0	8.0	**84.5**	**42.4**	19.9	**27.1**	9.6	**13.9**
10	0	0.8	**8.4**	0.1	5.6	**9.0**	7.7	**5.1**	0.4	2.7	9.5	6.6	**85.4**	**43.1**	19.8	**24.1**	7.7	4.0
10	10	2.0	**10.6**	0.2	7.4	**11.8**	8.7	**7.2**	0.5	4.1	8.4	7.0	**83.3**	**39.3**	19.6	**26.8**	9.3	**14.2**
0	20	3.0	**27.1**	0.3	23.6	**32.6**	23.7	**27.1**	5.3	18.6	16.2	14.9	**89.6**	**53.0**	32.4	**48.7**	20.2	**40.5**
20	0	2.1	**17.8**	0.2	15.5	**22.2**	14.8	**16.6**	2.1	10.4	13.2	10.1	**87.7**	**48.9**	25.6	**37.5**	13.6	5.1
20	20	5.7	**26.3**	0.3	21.6	**29.7**	20.1	**25.1**	4.9	17.0	11.4	11.8	**83.7**	**41.6**	24.3	**47.9**	19.5	**38.9**
20	-20	13.1	**54.6**	0.4	49.4	**65.0**	41.5	**60.1**	21.6	47.9	45.6	38.2	**95.9**	**84.1**	61.6	**75.7**	46.6	**40.6**
0	40	17.5	**77.3**	0.9	71.1	**86.5**	65.1	**84.6**	51.3	76.7	54.1	53.6	**97.0**	**82.8**	72.7	**91.9**	74.9	**91.7**
40	0	14.1	**60.5**	0.8	53.1	**67.5**	46.2	**65.0**	26.9	54.0	36.8	35.4	**94.0**	**72.1**	55.2	**80.3**	51.6	5.5
40	40	17.1	**74.3**	0.7	65.7	**80.6**	63.8	**80.6**	43.2	71.1	39.8	43.3	**94.7**	**69.8**	60.2	**91.8**	72.0	**92.3**
70	0	22.6	**96.6**	1.0	87.4	**98.4**	84.2	**98.6**	86.2	96.6	81.0	83.4	**99.7**	**95.4**	92.7	**99.4**	96.1	6.3
70	70	23.5	**98.8**	1.4	86.9	**99.2**	90.8	**99.2**	93.5	97.5	83.7	88.6	**99.9**	94.7	**94.8**	**99.9**	99.4	**99.9**
b)								$(n,s) = (50,20)$										
0	10	11.2	**17.4**	15.8	15.5	**19.0**	15.3	**16.2**	1.6	4.5	9.2	8.9	**46.3**	**29.0**	15.1	**20.4**	7.5	**21.3**
10	0	7.7	**12.7**	11.3	10.8	**13.4**	10.4	**10.4**	0.7	2.5	7.4	7.1	**44.2**	**25.7**	12.7	**17.5**	7.2	5.4
10	10	11.1	15.1	**15.2**	13.6	**17.5**	13.7	**14.9**	1.0	3.9	5.5	6.4	**39.0**	**21.3**	11.6	**20.8**	8.5	**22.3**
0	20	38.6	**48.5**	48.4	46.9	**55.5**	42.7	**55.8**	14.1	27.9	15.2	17.2	**57.9**	**40.3**	26.9	**52.0**	23.6	**63.5**
20	0	24.7	33.0	**33.5**	31.0	**39.2**	28.9	**37.4**	6.3	14.4	11.4	13.3	**51.6**	**33.6**	21.1	**37.9**	14.4	5.1
20	20	37.9	45.4	**47.2**	42.8	**49.7**	39.7	**52.8**	11.1	25.6	8.4	11.6	**43.5**	**26.0**	18.7	**55.4**	24.3	**65.1**
20	-20	79.9	**86.7**	84.4	85.8	**91.2**	81.2	**92.9**	56.7	76.9	63.4	58.8	**89.5**	**87.1**	70.8	**88.1**	65.5	**64.5**
0	40	95.9	**97.5**	97.4	97.0	**98.4**	94.8	**99.2**	88.1	95.8	65.1	73.9	**93.4**	**84.8**	81.3	**98.8**	92.2	**99.6**
40	0	82.3	**87.5**	86.0	86.3	**91.5**	81.2	**93.0**	59.8	78.8	41.3	48.0	**80.8**	**68.4**	59.7	**90.5**	69.8	6.1
40	40	**96.8**	96.3	**96.8**	94.5	**97.5**	92.6	**98.8**	79.3	93.4	45.9	57.0	**84.0**	**67.7**	66.6	**98.9**	90.2	**99.7**
70	0	99.8	**99.9**	**99.9**	99.3	**100.0**	99.5	**100.0**	99.5	99.9	92.5	96.5	**99.5**	97.8	**98.2**	**100.0**	99.9	6.8
70	70	**100.0**	**100.0**	**100.0**	98.3	**100.0**	**100.0**	**100.0**	99.8	99.9	93.9	98.0	**99.8**	98.0	**98.7**	**100.0**	**100.0**	**100.0**
c)								$(n,s) = (80,30)$										
0	10	21.3	22.8	**23.5**	22.3	**26.3**	20.9	**22.2**	1.4	3.3	7.4	8.3	**25.0**	**22.2**	12.7	**17.2**	7.9	**30.0**
10	0	12.4	15.3	**16.8**	15.5	**17.9**	13.4	**13.0**	0.6	1.5	5.3	5.7	**23.5**	**19.2**	10.0	**11.7**	5.9	4.7
10	10	20.0	22.5	**24.3**	20.6	**23.1**	19.3	**21.2**	1.1	2.9	3.6	4.9	**17.3**	**12.9**	8.0	**20.2**	8.2	**31.3**
0	20	69.0	**70.9**	70.4	69.4	**75.5**	64.0	**76.4**	21.9	38.1	17.0	21.0	**40.7**	**37.3**	28.8	**59.7**	33.0	**82.5**
20	0	48.0	**51.6**	50.5	49.0	**56.2**	43.1	**54.1**	8.8	17.5	11.9	14.3	**33.1**	**29.7**	20.7	**36.6**	17.1	5.3
20	20	66.7	64.9	**67.4**	65.1	**69.0**	59.1	**73.4**	16.6	33.0	7.6	12.9	**23.1**	**19.6**	18.0	**66.2**	31.5	**84.0**
20	-20	97.3	**97.8**	97.0	97.7	**98.6**	95.6	**99.3**	78.9	90.8	78.3	74.9	**89.2**	**93.0**	82.7	**94.9**	82.0	**85.2**
0	40	**99.9**	**99.9**	99.8	99.9	**100.0**	99.6	**100.0**	97.4	99.5	79.0	87.3	**94.7**	**91.1**	90.8	**99.8**	98.8	**100.0**
40	0	**98.5**	98.2	98.2	98.3	**99.2**	96.2	**99.5**	81.2	92.8	51.7	60.5	**78.5**	**73.6**	69.6	79.0	**86.4**	5.9
40	40	**99.9**	99.7	99.8	**99.7**	**99.7**	99.1	**100.0**	92.6	98.9	50.1	66.5	**81.1**	69.5	**73.2**	**100.0**	98.8	**100.0**
70	0	**100.0**	**100.0**	**100.0**	99.9	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	98.3	99.5	**99.9**	99.6	**99.7**	**100.0**	**100.0**	6.7
70	70	**100.0**	**100.0**	**100.0**	99.7	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	97.8	99.8	**99.9**	99.5	**99.9**	**100.0**	**100.0**	**100.0**

with n: sample size, s: shift location, δ_Q: shift amplitude in Q, δ_V: shift amplitude in V and *: the depth function with which the test is originally defined. Gray color indicates a test power higher than 95%. Numbers written in bold and underlined indicate the best power of each test for the corresponding (δ_Q, δ_V).

Table 4 : General information about *Moisie*, *Magpie* and *Romaine* stations

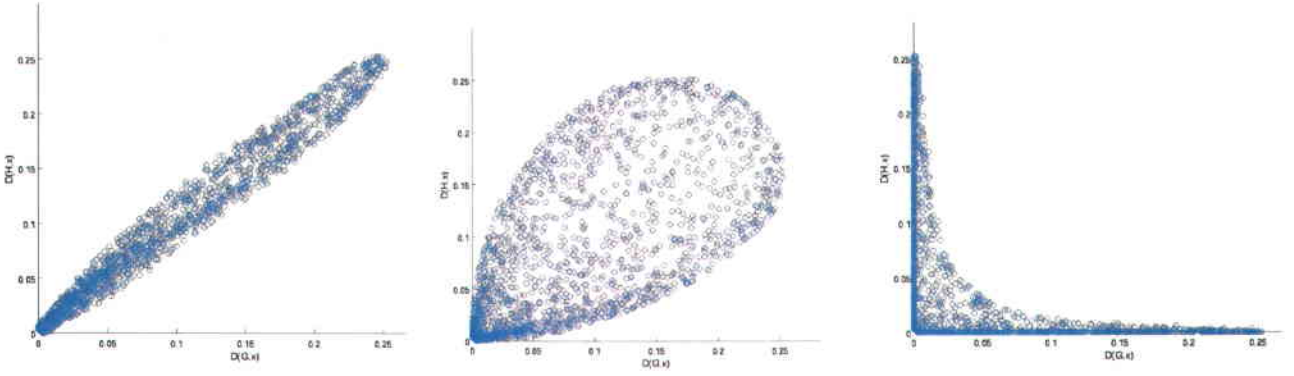
Station name	Station number	Latitude	Longitude	Period of records (#years)	Area (Km^2)
Moisie	072301	50 21 09	-66 11 12	1968-1998 (31)	19 012
Magpie	073503	50 41 08	-64 34 43	1979-2004 (26)	7 201
Romaine	073801	50 18 28	-63 37 07	1961-2006 (46)	12 922

Table 5 : Tests statistic and p-value of M-, T-, QIB-, QIA-, Z- and C-test and decision (1: shift, 0: no shift) of W-test

Tests		M			T			W			QIB			QIA		Z		Cramer
		TD	MD	SD*	TD	MD	SD*	TD*	MD	SD	TD	MD*	SD	TD	MD*	TD	MD*	
Moisie	Stat	0.08	0.00	0.00	0.32	0.10	0.19	0.22	0.06	0.06	0.10	0.06	0.00	0.10	0.06	15.99	20.50	9353.21
	p-val	0.00	0.01	0.05	0.00	0.00	0.00	1	1	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Magpie	Stat	0.3	0.00	0.00	0.6	0.2	0.3	0.5	0.3	0.3	0.3	0.19	0.00	0.3	0.19	0.99	2.92	1132.5
	p-val	0.16	0.54	0.53	0.08	0.09	0.07	0	0	0	0.47	0.32	0.02	0.11	0.02	0.32	0.09	0.03
Romaine	Stat	0.5	0.1	0.1	0.7	0.3	0.3	0.63	0.5	0.6	0.4	0.31	0.1	0.4	0.31	2.5	4.08	2478.4
	p-val	0.03	0.07	0.07	0.03	0.09	0.19	0	1	1	0.10	0.11	0.04	0.05	0.01	0.11	0.04	0.01

Gray color indicates that the corresponding test detect a shift in the corresponding year.

Figures



a) b) c)

Figure 1: DD-plot for a) two identical subsamples, b) two different subsamples and c) two very different subsamples.

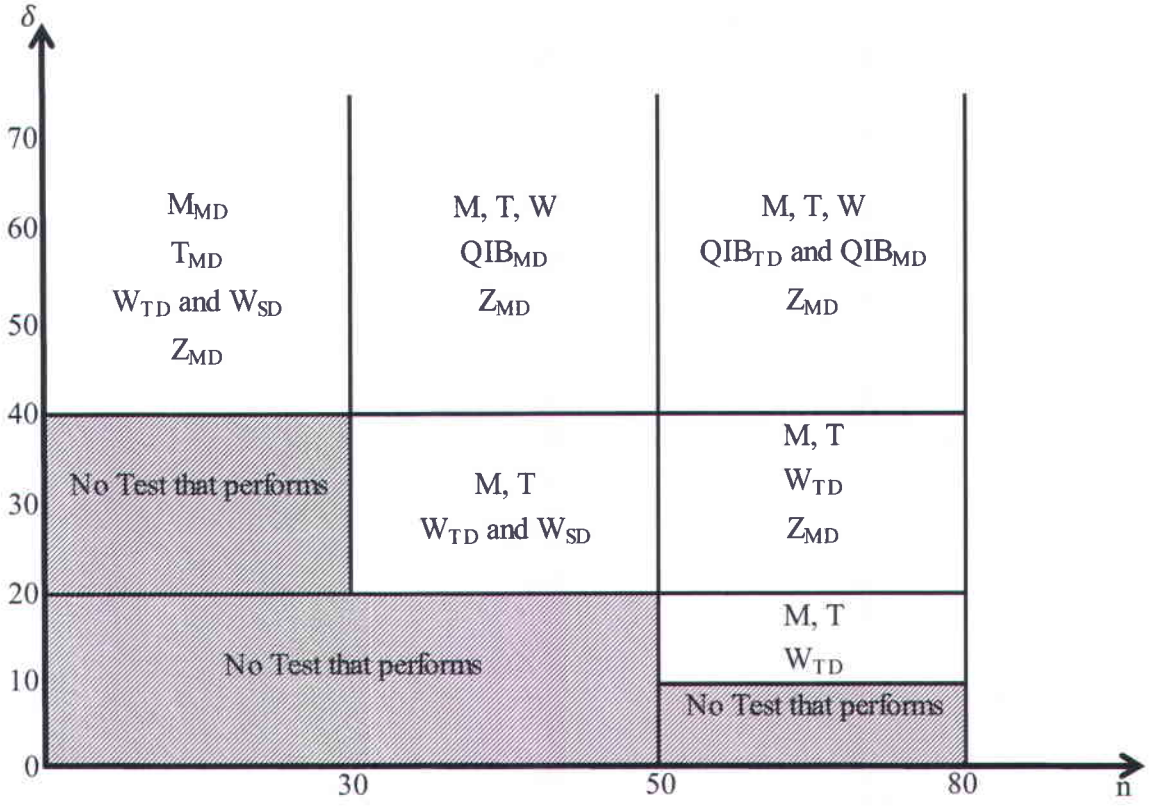


Figure 2 : Diagram of the applicability of considered tests for studied sample lengths and shift amplitudes.

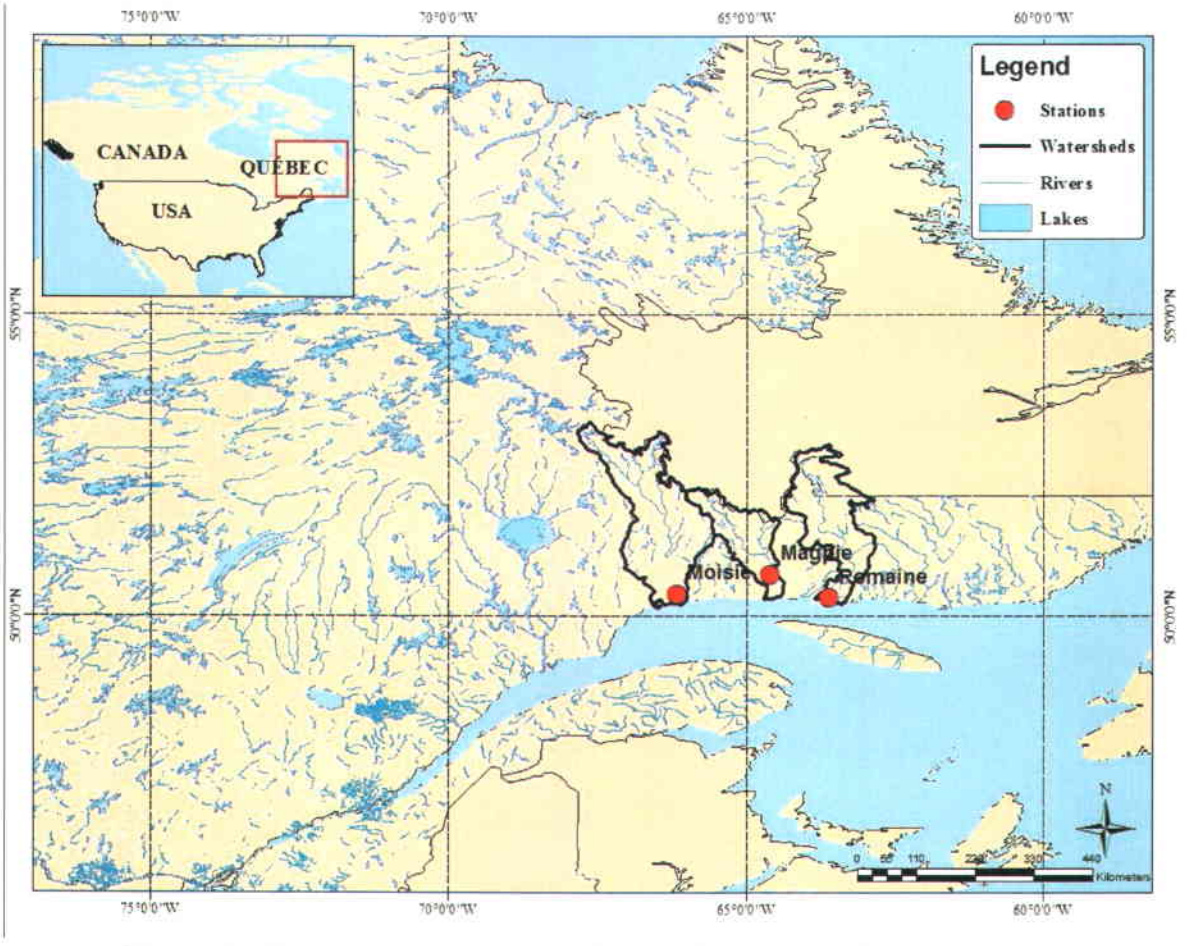
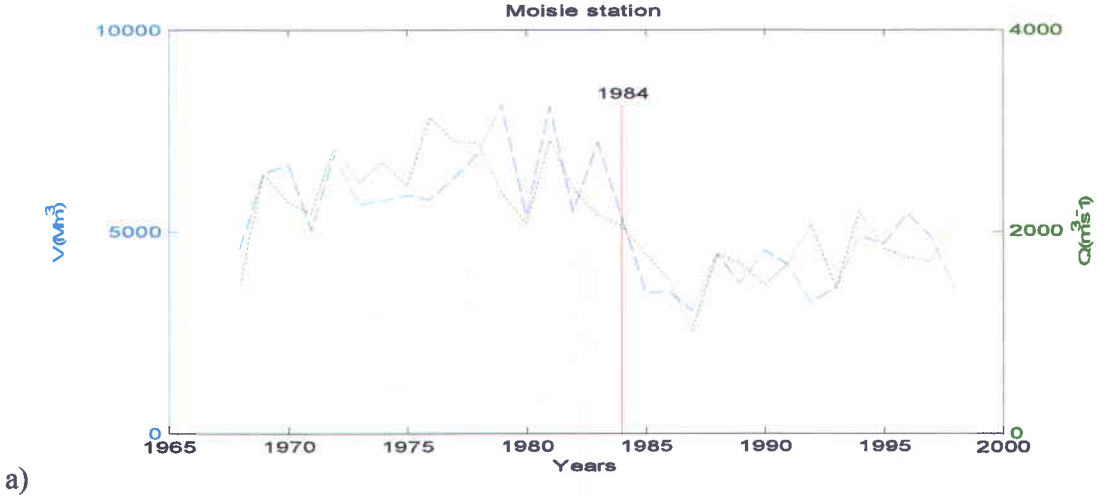
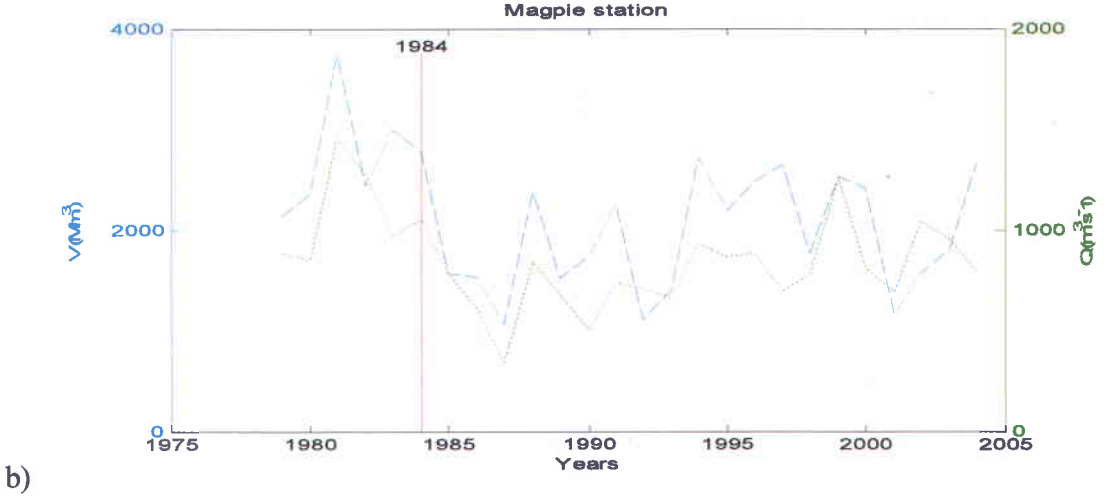


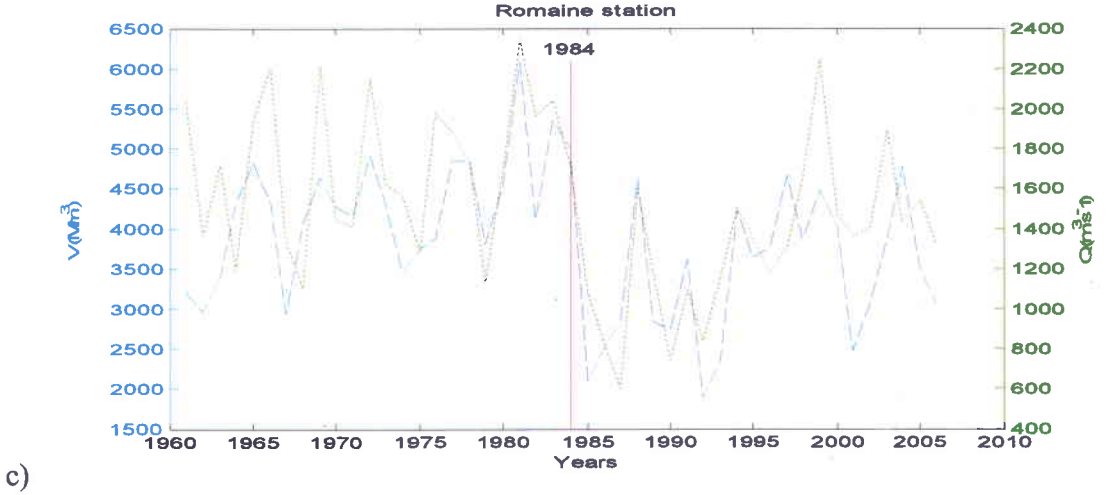
Figure 3 : Geographical location of *Moisie*, *Magpie* and *Romaine* stations.



a)



b)



c)

Figure 4 : The *V* and *Q* time series of a) *Moisie*, b) *Magpie* and c) *Romaine* stations.

10 Article 2: Dependence evolution of hydrological characteristics, applied to floods in a climate change context in Quebec

Dependence evolution of hydrological characteristics, applied to floods in a climate change context in Quebec

M.-A. Ben Aissia* [1]

F. Chebana [1]

T. B. M. J. Ouarda [1,2]

L. Roy[3]

P. Bruneau[3]

M. Barbet [3]

[1] Canada Research Chair on the Estimation of Hydro-meteorological Variables, INRS-ETE,
490, de la Couronne, Québec, (Québec)
Canada, G1K 9A9

[2] Masdar Institute of Science and Technology,
Abu Dhabi, United Arab Emirates.

[3] Hydro-Québec Équipement
855, Ste-Catherine East, 19th Floor, Montreal (Quebec)
Canada, H2L 4P5

* Corresponding author: Mohamed Aymen Ben Aissia
Ph.D student at INRS-ETE
490 de la Couronne Québec(Québec) G1K 9A9
Tel : (418)654-2530 # 4469
email : mohamed.ben.aissia@ete.inrs.ca

August 2013

Abstract

Generally, a hydrological event such as floods, storms and droughts can be described as a multivariate event with mutually dependent characteristics. In the literature, two types of studies are performed focusing either on the evolution of one variable or more but separately, or on the joint distribution of two or more variables on a fixed window period. The main aspect of multivariate analysis is the dependence between the studied variables. It is important to study the evolution of this dependence over a long period, especially in studies dealing with climate change (CC). The aim of the present study is to evaluate and analyze the dependence evolution between hydrological variables with an emphasis on the following flood characteristics, peak (Q), volume (V) and duration (D). This analysis includes confidence interval determination, stationarity analysis and change-point detection over a moving window series of three dependence measures. Two watersheds are considered along with observed and simulated flow data, obtained from two hydrological models. Results show that the dependence between the main flood characteristics over time is not constant and not monotonic. The corresponding behavior is sensitive to the choice of hydrological model, to climate scenarios and to the global climate model being used. The dependence of (Q, V) decreases when that of (V, D) increases. Moreover, the two considered hydrological models generally overestimate the dependence of (Q, V) and underestimate the dependence of (V, D) and (Q, D). All simulated dependence series are stationary over the whole period and present several break-points corresponding to short trends. This study allows also to check the ability of hydrological models, and if necessary, to recalibrate them to correctly simulate the dependence historically and in the future.

Keywords: Dependence measures, Dependence evolution, Flood characteristics, Climate Change, Stationarity, Break-point detection, Confidence interval and Québec.

46 # I. Introduction

47 A number of engineering design planning, design, and management activities require a detailed
48 knowledge of hydrological variables through their characteristics, which are mutually correlated.
49 Various studies showed the importance of considering the dependence structure between these
50 characteristics. For instance, Cordova & Rodriguez-Iturbe (1985) showed that the correlation
51 between rainfall duration and average intensity had a non-negligible effect on the storm generated
52 runoff. Kao & Govindaraju (2007) quantified the effect of dependence between rainfall duration
53 and average intensity on runoff. Therefore, dependence between hydrological variables can
54 influence flood flow quantiles (Goel et al. 2000).

55 In earlier studies (e.g. Wood (1976) and Chan & Bras (1979)), the use of joint distribution was
56 often accompanied with the assumption of independence between different variables. This
57 assumption was fairly inconvenient, and is frequently not supported by the data (Kao &
58 Govindaraju 2007). Recently, increasing attention has been given to multivariate analysis in
59 which, the main component is the dependence between variables.

60 A number of hydrological studies considered either one or more variables and studied their future
61 evolution (e.g. Reynard et al. 2001; Bronstert 2003). Recently, Ben Aissia et al. (2011)
62 compared, in a climate change (CC) context, eight spring flood characteristics on two different
63 periods of 30 years: observed (1971-2000) and future simulated (2041-2070). In previous studies,
64 attention was often given to modeling the joint distribution of two or more variables on a fixed
65 window period, usually historical period, by evaluating one value of the corresponding
66 correlation coefficient or the copula parameter (e.g. Shiau 2003; Zhang & Singh 2006; Chebana
67 & Ouarda 2011). However, the evolution of dependence between hydrological characteristics,
68 over a long period has not been considered in the literature.

69 In the present study, we are interested in the temporal evolution of the dependence of the
70 characteristics of hydrological variables in a multivariate framework. Three dependence measures
71 are considered, namely: the Pearson's correlation (r), Kendall's tau (τ) and Spearman's rho (ρ).
72 Based on these measures, moving window series are analyzed. The methodology includes
73 descriptive statistical analysis, confidence interval determination, stationarity analysis and
74 change-point detection. An application of the proposed procedure, to a case-study from the
75 province of Quebec (Canada), is performed. In this study we focus on the main flood
76 characteristics: peak Q, volume V, and duration D.
77 The paper is organized as follows. The methodology is presented in the second section. The third
78 section contains a description of the study area and the available data. Results and discussions are
79 reported in the fourth section and the conclusions are presented in the last section.

80 II. Methodology

81 In the present section we describe the proposed methodology in its general form applied on
82 hydrological variables. This section contains the evaluation of variable dependence
83 characteristics, confidence interval determination, stationarity testing and break-point detection.

84 *II.1. Determination of the dependence*

85 From a physical point of view, and supported by the hydrological literature, the dependence is
86 generally significant between Q & V and less significant between V & D, but not significant
87 between Q & D (e.g. Yue et al. 1999). This dependence varies from one site to another. The most
88 commonly used dependence measure coefficients are the Pearson's correlation (Hollander &
89 Wolfe 1973), Kendall's tau (Kendall 1975) and Spearman's rho (Best & Roberts 1975). These
90 dependence measures are considered in the present study. Let $x=(x_1,\ldots,x_N)$ and

$y=(y_1,\ldots,y_N)$ be the two datasets of interest (e.g. V and Q) where N is the length of the dataset. Let $(x_1^*,x_2^*,\cdots,x_n^*)$ and $(y_1^*,y_2^*,\cdots,y_n^*)$ be the n series for each variable derived by using a moving windows of length q (e.g. q=30 years in the application below) where $n=N-q+1$ and $x_1^*=(x_1,x_2,\cdots,x_q),x_2^*=(x_2,x_3,\cdots,x_{q+1}),\cdots,x_n^*=(x_{N-q+1},x_{N-q+2},\cdots,x_N)$. The three dependence measures are then computed for the n series.

The Pearson's correlation coefficient (r) is defined by:

$$r_k=\frac{v_{x_k^*y_k^*}}{\sigma_{x_k^*},\sigma_{y_k^*}};k=1,\ldots,n \quad (1)$$

where $v_{x_k^*y_k^*}$ is the covariance between x_k^* and y_k^* and $\sigma_{x_k^*},\sigma_{y_k^*}$ are the standard deviations of x_k^* and y_k^* respectively.

The Kendall' tau (τ) coefficient is given by:

$$\tau_k=\frac{(\text{number of concordant pairs})-(\text{number of discordant pairs})}{n(n-1)/2},\ k=1,..,n \quad (2)$$

For i,j=1,...,q, $(x_k^*(i),y_k^*(i))$ and $(x_k^*(j),y_k^*(j))$ are said to be concordant if $x_k^*(i)>x_k^*(j)$ and $y_k^*(i)>y_k^*(j)$ or if $x_k^*(i)<x_k^*(j)$ and $y_k^*(i)<y_k^*(j)$. They are said to be discordant, if $x_k^*(i)>x_k^*(j)$ and $y_k^*(i)<y_k^*(j)$ or if $x_k^*(i)<x_k^*(j)$ and $y_k^*(i)>y_k^*(j)$. If $x_k^*(i)=x_k^*(j)$ or $y_k^*(i)=y_k^*(j)$, the pair is neither concordant, nor discordant.

Spearman's rho (ρ) coefficient is defined as:

$$\rho_k=1-\frac{6\sum_{i=1}^{n}d_k^2(i)}{n(n^2-1)};\quad d_k(i)=x_k^*(i)-y_k^*(i); \quad (3)$$

II.2. Confidence interval determination

A number of methods may be used to determine the statistical CI of a series in hydrology, for

107 instance: asymptotic approximation, bias-corrected and accelerated (BCa), bootstrap-t,
108 approximate bootstrap confidence (ABC) and calibration (DiCiccio & Efron 1996). According to
109 Efron (1987), the BCa is the best method since it is unbiased and it is not time consuming. This
110 method consists in assessing the lower and upper bounds of the CIs at the confidence level
111 $\alpha \in]0,1[$ through the expressions :

$$ql = \Phi\left(z_0 + \frac{z_0 + z^{1-\alpha/2}}{1 - a\left(z_0 + az^{1-\alpha/2}\right)} \right) \tag{4}$$

$$qu = \Phi\left(z_0 + \frac{z_0 + z^{\alpha/2}}{1 - a\left(z_0 + az^{\alpha/2}\right)} \right) \tag{5}$$

112 Here Φ is the standard normal c.d.f, z_0 and a represent bias-correction and acceleration
113 coefficients respectively and are estimated by (DiCiccio & Efron 1996):

$$z_0 = \phi^{-1}\left\{ \frac{\#\left\{\hat{\theta}^* < \hat{\theta}\right\}}{B} \right\} \tag{6}$$

$$a = \frac{1}{6} \frac{\sum_{i=1}^{q} U_i^3}{\left(\sum_{i=1}^{q} U_i^2\right)^{3/2}} \tag{7}$$

114 where $\theta = (\theta_1, \theta_2, \cdots, \theta_n)$ is the dependence measure of interest (θ can be r, τ or ρ), $\hat{\theta}$ is an
115 estimate of θ based on the observed data, $\hat{\theta}^*$ is a bootstrap replication of $\hat{\theta}$ obtained by
116 resampling, B is the number of bootstrap samples, and U is the vector defined by:

$$U_i = (n-1)\left(\hat{\theta} - \hat{\theta}_i\right), i = 1, ..., q \tag{8}$$

117 $\hat{\theta}_{(i)}$ is the estimate of $\hat{\theta}$ based on the reduced data set

$(x_i, y_i) = \left(\left(x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_q \right), \left(y_1, y_2, \cdots, y_{i-1}, y_{i+1}, \cdots, y_q \right) \right)$.

The procedure of the BCa method is the following:

1- Estimate θ by $\hat{\theta} = d(x)$, d can be r, τ or ρ

2- Resample B times to get $(x_1^*, x_2^*, ..., x_B^*)$

3- Evaluate the statistics $\theta_b^* = d(x_b^*)$, $b = 1, ..., B$

4- Sort θ_b^* as $\theta_{(1)}^* \le \theta_{(2)}^* \le \cdots \le \theta_{(B)}^*$

5- Estimate z_0 and a

6- Evaluate ql and qu (equations 4 and 5).

II.3. Stationarity testing

In order to study the stationarity of dependence series, we apply the Mann-Kendall test (Mann 1945; Kendall 1975). The Mann-Kendall test (*MK*) detects the existence of a trend in a series. The test statistic is given by:

$$Z = \begin{cases} \dfrac{S-1}{\left[\mathrm{var}(S)\right]^{1/2}} & if \quad S > 0 \\ 0 & if \quad S = 0 \\ \dfrac{S+1}{\left[\mathrm{var}(S)\right]^{1/2}} & if \quad S < 0 \end{cases} \qquad (9)$$

where

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \mathrm{sgn}\left(x_j - y_k \right) \qquad (10)$$

and

$$\mathrm{sgn}(x) = \begin{cases} +1 & if\ x > 0 \\ 0 & if\ x = 0 \\ -1 & if\ x < 0 \end{cases} \tag{11}$$

One of the assumptions of the *MK* test is that the data are independent, so there is no trend or serial correlation measure among the observations. However, since we use a moving window in the evaluation of dependence series, these series should present an autocorrelation. Therefore, the use of the original *MK* test may not be appropriate. A number of MK test versions have been suggested to overcome the problem of autocorrelation, see e.g. Hamed and Ramachandra (1998), Yue et al. (2002) and Khaliq et al. (2009) and references therein. Among these tests, we select five to be applied in the present study: Block–Bootstrap (*BB*), Pre-whitening (*PW*), Trend-free pre-whitening (*TFPW*) and two Variance correction tests (*VC*1 and *VC*2). A brief description of these tests is provided below.

Block-Bootstrap (BB)

The test based on Block–Bootstrap is suggested in the case of autocorrelation (e.g. Hipel & McLeod 2005; Khaliq et al. 2009). This test is based on block resampling which was introduced by Carlstein (1986). In this approach, the original data are resampled in predefined blocks m of length l where we suppose $n = ml$. If $\theta = (\theta_1, \theta_2, ..., \theta_n)$ is the series, we set $z_1 = (\theta_1, ..., \theta_l)$, $z_2 = (\theta_{l+1}, ..., \theta_{2l})$, ..., $z_m = (\theta_{(m-1)l+1}, ..., \theta_{ml})$ giving blocks $z_1, z_2, \cdots, z_m$. The idea that underlies the Block-Bootstrap is that if the blocks are long enough, the *MK* test statistics $Z^* = (Z_1, Z_2, ..., Z_m)$ of resampled series have approximately the same distribution as the values Z calculated from the original series (Davison & Hinkley 2009). The optimal block length can be derived automatically using the algorithm of Politis & White (2004). A correction of this algorithm was made by Patton et al. (2009). The corrected algorithm is used in the present paper

154 to determine the optimal block length. The p-value of the MK Block–Bootstrap test is defined by
155 Davidson & Hinkley (2009):

$$p = \Pr\left(Z^* \geq Z \mid H_0\right) \tag{12}$$

156 *Pre-whitening (PW)*

157 To overcome the problem of autocorrelation in the considered series, the pre-whitening approach
158 was suggested (e.g. Douglas et al. 2000; Zhang et al. 2001). The main steps of this approach are:
159 1) compute the lag-1 serial correlation coefficient r_1, 2) if r_1 is found to be non-significant for a
160 significance level α, then the trend identification test (MK in this study) is applied to the original
161 time series, otherwise 3) the trend identification test is applied to the pre-whitened time series
162 $\left(\theta_2 - r_1\theta_1, \theta_3 - r_1\theta_2, \ldots, \theta_n - r_1\theta_{n-1}\right)$.

163 *Trend-free pre-whitening (TFPW)*

164 Fleming and Clarke (2002) and Yue et al. (2002) found that the magnitude of the trend can be
165 affected by the pre-whitening approach when the data are not transformed, i.e. the removal of the
166 positive (or negative) AR(1) process deflates (or inflates) the existing trend. To remedy this
167 problem, Yue et al. (2002) suggested the *TFPW* approach for taking into account the effect of
168 serial correlation. The required steps to implementing this approach are: 1) for a given series,
169 estimate the slope of the trend, 2) de-trend the dependence series and estimate the first serial
170 correlation coefficient r'_1 from the de-trended series, 3) if r'_1 is non-significant, then the *MK*
171 trend test is applied to the original time series, otherwise 4) the *MK* trend test is applied to the de-
172 trended pre-whitened series recombined with the estimated slope of trend from step 1.

173 *Variance corrections (VC1 and VC2)*

174 Yue et al. (2002) demonstrated that the presence of serial correlation in a time series does not
175 alter the mean of the *MK* test statistic *Z* as well as its asymptotic normality, but it may change the

176 dispersion of the distribution of *Z*. The existence of positive (or negative) autocorrelation
177 increases (or decreases) the variance of *Z*. Bayley & Hammersley (1946), Hamed &
178 Ramachandra (1998) and Yue & Wang (2004) proposed corrections of the variance of the *MK*
179 test statistic *Z* using an effective sample size that reflects the effect of serial correlation on the
180 variance of *Z*. The modified variance of the *MK* test statistic is given by:

$$Var^*(Z) = CF \times Var(Z) \tag{13}$$

181 with *CF* is a correction factor and *Var(Z)* is the variance of the *MK* test statistic *Z* for the original
182 data series. The correction factors proposed by Hamed & Ramachandra (1998) (*CF1*) and Yue &
183 Wang (2004) (*CF2*) are:

$$CF_1 = 1 + \frac{2}{n(n-1)} \sum_{k=1}^{n-1} (n-k)(n-k-1)(n-k-2) r_k^R \tag{14}$$

184 and

$$CF_2 = 1 + 2 \sum_{k=1}^{n-1} (1 - k/n) r_k \tag{15}$$

185 where r_k and r_k^R are respectively the lag-k serial correlation coefficient and the ranks of the data.

186 ***II.4. Break-point detection:***

187 To examine the existence of break-points in the dependence series, we use the Bayesian method
188 of multiple break-point detection in multiple linear regressions (Seidou & Ouarda 2007). This
189 method determines the number of break-points and their locations as well as the trends in each
190 segment of the series. The Bayesian procedure assumes an improper non informative prior as a
191 distribution for the parameters representing the break-points. In the present study, the regression
192 model is formulated as:

$$\begin{cases} \theta = \alpha_1 t + \alpha_2 + \varepsilon \text{ if } t \le \gamma \\ \theta = \alpha_1' t + \alpha_2' + \varepsilon' \text{ if } t > \gamma \end{cases} \tag{16}$$

193 where θ represents the dependence series, t is the time in years, α_1, α_2, α_1' and α_2' are the

194 parameters to be estimated and γ is the break-point time.

195 III. Region and datasets

196 The presented methodology is applied to a flood case-study in the province of Quebec, Canada.
197 Two stations representing two watersheds are considered, that is Baskatong reservoir and the
198 Romaine River (figure 1) with respectively watershed areas of 13 040 km^2 and 14 500 km^2. The
199 hydrological regime is dominated by snowmelt runoff, which occurs generally in the spring
200 season from April to June. At the outlet of the Baskatong reservoir, Mercier dam is the largest
201 structure in the watershed. Four dams will be constructed on the Romaine River by the year 2020
202 with an installed capacity of 1550 megawatts (MW). Note that in Baskatong reservoir, the
203 gauging station is located upstream of the reservoir. Consequently, the observed flow is not
204 controlled by the dam.
205 For the watershed of the Baskatong reservoir, Q, V and D are evaluated using three categories of
206 available flow datasets. The observed flow series from January 1st 1969 to December 31st 2000
207 represent the first category. The second category, is represented by two simulated flow data sets
208 obtained by the hydrological models HSAMI (e.g. Bisson & Roberge 1983) and HYDROTEL
209 (e.g. Fortin et al. 1995) with meteorological data as inputs of these two models. Meteorological
210 data are driven using the atmospheric fields from ERA-40 reanalysis as boundary conditions,
211 Simulated by the Canadian Regional Climate Model (CRCM) (e.g. Brochu & Laprise 2007;
212 Music & Caya 2007) for the period from January 1st 1969 to December 31st 2000. Two simulated
213 flow data sets obtained by HSAMI and HYDROTEL for the period from January 1st 1961 to
214 December 31st 2070 constitute the third category. The two hydrological data sets have been
215 validated by comparing the observed and simulated flows. Meteorological data used as inputs to
216 the hydrological models are simulated by the CRCM driven by the atmospheric fields from the

217 Coupled Global Climate Model (CGCM3), according to the assumptions of the climate change
218 scenario A2 (Bates et al. 2008).
219 For the Romaine River watershed, two available dataset categories are used to evaluate Q, V and
220 D. The observed flow series from January 1st 1956 to December 31st 2009 represent the first
221 category. The second category is represented by six simulated flow data sets obtained by HSAMI
222 derived by the atmospheric fields from three Coupled Global Climate Models, which are
223 CGCM3, HADCM3 and ECHAM5. For each model, two series are simulated following the
224 assumption of two climate change scenarios. These scenarios are A2 & B1 for CGCM3, A2 & B2
225 for HADCM3 and A2 & B2 for ECHAM5 (Bates et al. 2008). Datasets of the second category
226 are simulated on the period from January 1st 1961 to December 31st 2070. All the above datasets
227 are summarized in table 1.
228 The dependence is calculated on the basis of the above three measures over a moving average of
229 thirty years (q = 30 years). The choice of this value of q in the present context is explained by the
230 fact that, over thirty years, the extended average of long-term climatic cycles such as the El Niño-
231 Southern Oscillation (ENSO) can be taken into account.
232 The main flood characteristics are Q, V and Q which can be obtained from the corresponding
233 hydrograph (figure 2). To extract these characteristics from the associated daily flow series, we
234 use the algorithm of Pacher (2006). This algorithm is based on the analysis of cumulative annual
235 hydrographs to find different periods within a year; dates are determined to maximize the fit
236 between a linear approximation and the cumulative hydrograph for each period of the year,
237 usually four or five for typical hydrographs in Quebec. It then determines the start, peak and end
238 of spring flood dates. Then, the values of Q, V and D are determined using the same procedure as
239 in Ben Aissia et al. (2011).

IV. Results and discussion

The considered dependence series, described in the previous section, are obtained for the different categories of datasets. Then, these series are analyzed as indicated in the methodology section, namely: confidence intervals, stationarity testing and break-point detection.

IV.1. Study of the dependence

For each type of dependence measure between Q & V, V & D and Q & D, different figures of dependence series are presented. For Baskatong, the historical period is short (only 32 years) and if we use a moving window of 30 years, we only obtain 3 values of the dependence. In order to further examine the dependence in the historical period, we consider a shorter moving window with 15 years length presented in figure 3. The dependence series CGCM3S (A2) and CGCM3Y (A2) (table 1) during the simulation period (1961-2070) are presented in figure 4. The DH series is presented in figure 4 in order to facilitate the comparison between different dependence series. Figure 5 presents all dependence series obtained for the Romaine watershed. This figure includes nine sub-figures: three for each couple of dependence (Q & V, V & D and Q & D). Each one of the nine sub-figures shows the dependence series for the simulated datasets and the observed series (table 1).

From the figures of dependence series (figures 4 and 5) we remark that the behavior of Kendall's tau and Spearman's rho series is similar. Then, we present the corresponding results of these two dependence measures at the same time.

a) Dependence between Q & V

The Pearson's correlation (r): Figures 4a and 5a, indicate that, in the historical period, the correlation series of historical data is lower than that simulated by HSAMI and higher than that simulated by HYDROTEL for the Baskatong reservoir. To explain this result, figure 6 presents

263 the scatter plots of (Q, V) for DH, CGCM3S (A2) and CGCM3Y (A2) during a 30 year period:
264 1970-1999. We observe that the general shape of the graph follows a straight line. We note that
265 the two years 1981 and 1982 of CGCM3Y (A2) differ from the rest of the series and can cause
266 the low correlation value of CGCM3Y (A2) during 1970-1999. To check this, we calculate the
267 Pearson's correlation of CGCM3Y (A2) during 1970-1999 without 1981 and 1982 and found that
268 it increases from 0.54 to 0.65 and becomes larger than the Pearson's correlation value of DH
269 during 1970-1999. This confirms that 1981 and 1982 cause the low correlation of CGCM3Y (A2)
270 during the reference period. Figure 6 indicates that the flood of 1981 is characterized by a low V
271 and high Q (peaked hydrograph) for CGCM3Y (A2). This may be due to a rapid snow melt for
272 HYDROTEL, a larger storage in the catchment, or less snow accumulation. For 1982, figure 6
273 shows also that CGCM3Y (A2)'flood is characterized by a low Q and high V (flat hydrograph).
274 The study of the flood hydrograph of 1982 (not presented here in order to alleviate the paper)
275 shows that for CGCM3S (A2) and CGCM3Y (A2) high liquid precipitations are recorded shortly
276 after the flood of 1982. These precipitations lead to a renewed increase of flow to a level that is
277 higher for CGCM3Y (A2) than for CGCM3S (A2). On the other hand, in 1982 the flood peak of
278 CGCM3Y (A2) occurs later than that of CGCM3S (A2). Therefore, the renewed increase of flow
279 was considered as part of the spring flood for CGCM3Y (A2) and not for CGCM3S (A2). This
280 explains the high V for CGCM3Y (A2) in 1982. For DH, we do not record high liquid
281 precipitation during the flood of 1982. Hence we can conclude that the low correlation value for
282 CGCM3Y (A2) compared to CGCM3S (A2) and DH over the reference period is due to the
283 unusual shape of the flood hydrographs of 1981 and 1982 simulated by HYDROTEL. This
284 finding shows the importance of the hydrograph shape on the dependence between Q and V.
285 For the Romaine River, the dependence series of historical data are generally higher than those
286 simulated (figure 5a). The dependence series of HADCM3 (B2) are the closest to those of

287 historical data. The difference between correlation series derived from the same climate model,
288 e.g. CGCM3S (A2) and CGCM3Y (A2), may to be unexpected since they share the same inputs
289 fot hydrological models. This can be explained by the fact that here we compare the correlation
290 series rather than the flow series.
291 For the Baskatong reservoir, we observe that the correlation series of CGCM3S (A2) and
292 CGCM3Y (A2) over the simulation period (1961-2070) increase during the historical period,
293 remain stable for some years and then decrease. To explain these results, one considers two
294 moving windows of 30 years with high dependence measures of CGCM3Y (A2) for the first
295 (1995-2024) and low ones for the second (2033-2062). Figure 7 shows the scatter plots of (Q, V)
296 for CGCM3Y (A2) over the two periods. We note that over 1995-2024 the scatter plot of (Q, V)
297 follows approximately a straight line whereas over 2033-2062 it is more dispersed. To further
298 explain these results, we present in figure 8 the minimum and maximum temperatures, the
299 precipitations and the daily flows for two specific years: 2039 and 2041 which cause (among
300 others) the dispersion over 2033-2062. Indeed, 2039 is characterized by high Q and low V
301 whereas 2041 is characterized by low Q and high V. Figure 9 shows that the high Q and low V
302 observed in 2039 are mainly caused by a strong increase in temperature which led to a high
303 snowmelt. For 2041 we observe during the flood a short period of decrease in temperature
304 leading to a reduction of snowmelt hence the low Q and the high V. Thus, we can conclude from
305 this example that low dependence values can be explained by peaked hydrographs which were
306 caused in this case by a high increase in temperature during a short period.
307 For the Romaine River, the correlation series are stable in the historical period and from 2020 we
308 see a separation: CGCM3 (B1) and ECHAM5 (B1) decrease, HADCM3 (A2), HADCM3 (B2),
309 ECHAM5 (A2) and CGCM3 (B1) increase. To explain this finding, we consider one moving
310 window of 30 years (2025-2054) with high dependence measures of CGCM3 (A2), HADCM3

311 (A2), HADCM3 (B2) and ECHAM5 (A2) and low dependence measures of CGCM3 (B1) and
312 ECHAM5 (B1). Figure 10 shows that, for CGCM3 (B1) and ECHAM5 (B1), there are few floods
313 with low Q and low V and a considerable number of years with low Q and high V which
314 correspond to low dependence.
315 The Kendall's tau (τ) and Spearman's rho (ρ): Figures 4b, 4c, 5b and 5c show tau and rho in the
316 Baskatong and the Romaine, respectively. From figures 4b and 4c, we note that over the
317 historical period (1969-2000), generally, CGCM3S (A2) and CGCM3Y (A2) are higher than DH.
318 Since only the Pearson's correlation of CGCM3S (A2) is higher than that of DH, then we can
319 conclude that, in general, the two hydrological models are not able to represent the observed
320 dependence. Over the simulation period (1961-2070), the tau series of CGCM3Y (A2) increases
321 during the period 2000 to 2010, followed by a decrease to values lower than those of the
322 historical period. The tau series of CGCM3S (A2) increases in the early 2000, remains stable
323 until 2020 and then decreases gradually. For the Romaine, we note that both tau and rho of the
324 simulated data are lower than those of the historical data. Furthermore, we note that tau and rho
325 of the simulated data HADCM3 (A2) and HADCM3 (B2) are higher than the rest of the
326 simulated series and remain stable during the simulation period (1961-2070). On the other hand,
327 the tau and rho series of simulated data CGCM3 (A2) and CGCM3 (B1) have a maximum and
328 minimum around 2040 respectively while those of ECHAM5 (A2) remain stable during the
329 simulation period and ECHAM5 (B1) reach a maximum between 2000 and 2010 and a minimum
330 between 2020 and 2030. The separation into two groups of series for the correlation is also
331 observed for the Kendall's tau and Spearman's rho. It can be explained in the same way.
332 From figures 4a-c, we note that the simulation period can be divided into two periods: during the
333 first one (1961-2015), CGCM3Y (A2) and CGCM3S (A2) are close (especially for the Pearson's
334 correlation) and during the second one (2015-2070) CGCM3S (A2) is clearly higher than

335 CGCM3Y (A2). To interpret this behavior, figure 12 presents the scatter plots of (Q, V) for
336 CGCM3S (A2) and CGCM3Y (A2) during these two periods. We notice that during the period
337 1961-2015, the scatter plot of CGCM3Y (A2) and CGCM3S (A2) follows approximately a
338 straight line. For the second period 2015-2070 we note that the scatter plot of CGCM3S (A2)
339 follows a more linear pattern than CGCM3Y (A2) which is more dispersed. Indeed, the years
340 2025, 2039, 2041 and 2055 are remarkably far from the rest of the points which means that the
341 hydrographs of these years are flat or peaked. To verify this finding, we calculate the Pearson's
342 correlation of CGCM3Y (A2) during 2015-2070 excluding 2025, 2039, 2041 and 2055 and we
343 find that it increases from 0.51 to 0.68 and becomes close to the Pearson' correlation value of
344 CGCM3S (A2) during the same period. Thus, the low dependence values for CGCM3Y (A2) are
345 mainly due to the particularity of hydrograph shapes over 2025, 2039, 2041 and 2055.
346 Consequently, the difference between the two dependence series CGCM3S (A2) and CGCM3Y
347 (A2) over the two periods 1961-2015 and 2015-2070 is due to the corresponding difference in
348 dispersion of the scatter plot of (Q, V) which can be caused by the particularity of the hydrograph
349 shape of some years.

350 *b) Dependence between V & D:*

351 The Pearson's correlation (r): For the historical period, the correlation series of historical data are
352 higher of those simulated in the Baskatong reservoir (figure 4d) while the opposite occurred in
353 the Romaine River (figure 5d). For both watersheds, we observe a slight increase in the
354 correlation series during the simulation period (1961-2070). Correlation series of ERA40Y are
355 generally higher than the one associated to ERA40S in the Baskatong reservoir (figure 3). The
356 situation is different for the Romaine (figure 5d) where we note, before 2030, an increase in the
357 correlation for CGCM3 (A2), HADCM3 (A2) and HADCM3 (B2), and a decrease followed by

358 an increase for CGCM3 (B1), ECHAM5 (A2) and ECHAM5 (B1). After 2030 all the series
359 decrease except CGCM3 (B1) that continues to increase.
360 The Kendall's tau (τ) and Spearman's rho (ρ): tau and rho of the Baskatong and the Romaine are
361 presented in figures 4e, 4f, 5e and 5f respectively. Figures 4e and 4f show a large difference
362 between the tau and rho series of the historical data and those of corresponding simulated data.
363 We observe also a slightly upward trend of tau and rho for the two series CGCM3S (A2) and
364 CGCM3Y (A2). On the other hand, from figures 5e and 5f, we note that the dependence series of
365 the simulated data follow generally a cyclic trend which is more apparent for ECHAM5 (B1).
366 We notice that, when the dependence between Q & V decreases, the dependence between V & D
367 increases. To explain this behavior, we consider the following reasoning: Consider a 30 year
368 period where the dependence measures between Q & V are low. This is reasonable as we have
369 already shown that there are floods with high Q and low V and flood with low Q and high V. To
370 obtain the first case with high Q and low V, we should have a short D so that the combination of
371 low V and short D will help to increase the dependence measures between D & V over the period.
372 The same results can be formulated for the second situation i.e. low Q and high V.

373 *c) Dependence between Q & D*

374 The Pearson's correlation (r): Figures 4g and 5g present the correlation between Q & D for the
375 Baskatong and the Romaine respectively. For the historical period, figure 4g shows that the
376 correlation in the historical data is higher than the correlation in the simulated data. For the
377 simulation period (1961-2070), we see that the correlation series of CGCM3S (A2) and
378 CGCM3Y (A2) have two maxima with dates and magnitudes that vary from one series to
379 another. For the Romaine, figure 5g shows that the correlation series of the simulated data are
380 generally lower than those of historical data. On the other hand, the correlation series are

generally stable in the simulation period except for CGCM3 (A2), HADCM3 (A2) and HADCM3 (A2) where an associated maximum is present respectively between 2020-2030, 2035-2040 and 2000-2010. We can also see that the correlation series related to HADCM3 (A2) and HADCM3 (B2) are generally higher than the rest of the correlation series.

The Kendall's tau (τ) and Spearman's rho (ρ): From figures 4h, 4i, 5h and 5i we note a high variation and a slight upward trend in tau and rho series for Baskatong and Romaine. In the historical period, both dependence measures are higher for historical data than for simulated data. Dependence series of simulated data HADCM3 (A2) and HADCM3 (B2) are the highest of the dependence series.

From these results we conclude that: 1) the behavior of the Kendall's tau and the Sperman's rho are similar, 2) the dependence between flood characteristics evolves with time, 3) the dependence series is sensitive to years with flat or peaked hydrographs, 4) the choice of hydrological model, climate model or climate change scenario affects the behavior of the dependence series, 5) in general, HSAMI and HYDROTEL overestimate the dependence between Q and V and underestimate that between V and D or between Q and D, and 6) when the dependence between Q and V increases the dependence between V and D decreases and vice versa.

IV.2. Confidence intervals

For each dependence series, we determine the corresponding confidence interval (CI). For space limitations and to avoid repetitions, a selected number of dependence series with their CIs are presented. For the Baskatong reservoir, we present in figure 11 the CI of the three dependence series between Q and V corresponding to the series of CGCM3S (A2). For the Romaine River, the CIs for Pearson's correlation series between Q & V are presented for CGCM3 (A2), CGCM3

404 (B1), HADCM3 (A2), HADCM3 (B2), ECHAM5 (A2) and ECHAM5 (B1) series (figure 12).
405 For the Baskatong reservoir we note that the width of the CI depends on the series variability.
406 Indeed, for HSAMI, we remark that between the years 1990 and 2000 the dependence series
407 between Q & V are more variable than in the rest of the series and have a wider confidence band.
408 The same remark is observed in dependence series between Q & D but during 2000-2030.
409 Generally, when the dependence measure decreases or increases suddenly the associated CI
410 becomes wider.
411 For the Romaine River we note that the CIs of dependence series between Q & V depend on the
412 series variability and climate change scenario (A2, B1 and B2). Indeed, we remark that CIs of
413 dependence series CGCM3 (B1) and ECHAM5 (B1) are narrower than those of the remaining
414 simulated series. CIs for dependence series of simulated data issued from A2 are narrow in the
415 reference period and wider afterwards. However, the opposite occurred for CIs of the simulated
416 data from B1. For dependence series between V & D, CIs seem to be steady during the simulation
417 period except for HADCM3 (A2) which has a CI that is reduced over the years. CIs of
418 dependence series between Q & D of simulated data HADCM3 (A2) and HADCM3 (B2) are
419 wider than those of the rest of the data.
420 Generally, the width of the CI increases when a sudden change in dependence series is observed.
421 This can be explained by an increase of the series variation during the sudden change.

422 ***IV.3. Stationarity***

423 In this section the stationarity of the dependence series over all simulated period is tested using
424 the six trend tests described above. These tests are applied on the two dependence series
425 CGCM3S (A2) and CGCM3Y (A2) for the Baskatong reservoir and on all studied dependence
426 series for the Romaine. Note that, for the Baskatong reservoir, the dependence series over the

427 historical period are very short and it is not appropriate to apply trend tests to them. Since the
428 number of series is different in the two watersheds, we present the stationary results in four
429 tables: one for the Baskatong reservoir and three for the Romaine. Table 2 shows the p-values
430 and the corresponding decisions regarding the stationarity hypothesis for the six tests applied to
431 the dependence series CGCM3Y (A2) and CGCM3S (A2) in the Baskatong reservoir. According
432 to the *MK* test, only three series are stationary, i.e. for Q & V with simulated data by HSAMI, for
433 Q & D with simulated data by HYDROTEL and for the rho of Spearman series Q & D with
434 simulated data by HYDROTEL. However, results of the five other tests show that all studied
435 series are stationary. Therefore, the non stationary behavior detected by the *MK* test is due to the
436 autocorrelation of studied series.
437 For the Romaine River, tables 3a-c present the p-values and the stationarity decision for
438 dependence series of Q & V, V & D and Q & D respectively. From table 3a, we see that,
439 according to the *MK* test, eight dependence series are stationary: DH for the correlation of
440 Pearson; DH, CGCM3 (A2), HADCM3 (A2) and ECHAM5 (B1) for the tau of Kendall and DH,
441 CGCM3 (A2) and ECHAM5 (B1) for the rho of Spearman. From the same table, we see that,
442 according to the *PW* method, DH for correlation of Pearson and rho of Spearman and HADCM3
443 (B2) for tau of Kendall are non-stationary. According to the *TFPW* method, two Spearman's rho
444 series are non-stationary: DH and ECHAM5 (A2). For BB, VC1 and VC2 methods, all
445 dependence series between Q & V are stationary. Note that the *PW* and *TFPW* approaches are
446 based on the assumption that the underlying observation generating mechanism conforms to an
447 autoregressive process of order one, which is not always true. Consequently, based on the results
448 of *BB*, *VC1* and *VC2*, we can conclude that dependence series between Q & V are stationary.
449 The results of tables 3b and 3c are similar to those of table 3a. Finally, we can conclude that all
450 dependence series studied in this paper are stationary over the whole period. However, they can

451 present trends or changes over short periods. The detection of these short-term trends is the object
452 of the next section.

453 ***IV.4. Bayesian analysis for break-point detection***

454 In this section, we apply the Bayesian analysis of break-point detection. Because of the high
455 number of dependence series and the similarity between figures, selected results of the break-
456 point detection are shown in figure 13 for the Baskatong reservoir and figure 14 for the Romaine.
457 We note that all dependence series obtained from simulated data present a number of break-
458 points (between 2 and 5 points). For Romaine, dependence series from DH present at most one
459 break-point. This can be explained by the fact that the DH series are shorter than the simulated
460 ones. Although all dependence series are stationary, they present a number of break-points.
461 Therefore, dependence series present some local trends. These trends are not long enough to
462 cause a non stationarity but represent the different phases of dependence behavior.
463 We note that the presence of a number of short periods where a break point is observed for most
464 series For instance, during the few years around 2005 for CGCM3S (A2) (also 2013 for
465 CGCM3Y (A2)) a break point is recorded for most dependence series. To investigate this result,
466 we consider for example the years around 2005 for CGCM3S (A2) for which we note a high
467 increase in dependence between Q & V. The period of dependence increase is 2001-2008. Since
468 we use a 30 years moving window, for each transition from window to the next we remove one
469 year and we add another. For the period of moving window of 2001 to 2008, the removed years
470 are 1987 to 1994 and the added years are 2016 to 2023. We find that break-points recorded
471 around 2005 for CGCM3S (A2) are the result of the succession of a few years that increase the
472 dependence between Q and V.

V. Conclusions

The aim of this paper is to study the evolution of the dependence between hydrological variables, such as floods and droughts, where more than one dependent feature characterize each variable in a multivariate framework. A particular interest of this study is in a climate change context over a long period. The proposed study considers the construction of the dependence series using a moving window on the basis of three measures, the evaluation of the corresponding confidence intervals, trend testing and break-point detection.

The above procedure is applied to flood features in two locations in Quebec, i.e. the Baskatong reservoir and the Romaine River. The analysis is performed on the dependence between Q & V, V & D and Q & D using historical as well as simulated flow data series from two hydrological models (HSAMI and HYDROTEL). Results show that dependence between main flood characteristics evolve over time and are not constant. In general, HSAMI and HYDROTEL overestimate the dependence between Q and V and underestimate that between V and D or between Q and D. Furthermore, despite the fact that HSAMI and HYDROTEL are driven by the same data, the behavior of their dependence series is different. All simulated dependence series are stationary and present several break-points. In general, this study allows pushing the limits of hydrological models by checking their ability to correctly simulate the whole hydrological events including their features as well as their dependence, and if necessary to revisit these models.

Acknowledgments

The authors thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and Hydro-Québec for the financial support. The CRCM data have been generated and supplied by the Ouranos consortium.

REFERENCES

Bates, B., Z. W. Kundzewicz, S. Wu and J. Palutikof (2008). Climate Change and Waterr. Technical Paper of the Intergovenmental Panel on Climate Change, Intergovernmental Panel on Climate Change, IPCC Secretariat, . Geneva 210 pp.

Bayley, G. V. and J. M. Hammersley (1946). "The "Effective" Number of Independent Observations in an Autocorrelated Time Series." *Supplement to the Journal of the Royal Statistical Society* **8**(2): 184-197.

Ben Aissia, M. A., F. Chebana, T. B. M. J. Ouarda, L. Roy, G. Desrochers, I. Chartier and É. Robichaud (2011). "Multivariate analysis of flood characteristics in a climate change context of the watershed of the Baskatong reservoir, Province of Québec, Canada." *Hydrological Processes*: n/a-n/a.

Best, D. J. and D. E. Roberts (1975). "Algorithm AS 89: The Upper Tail Probabilities of Spearman's rho." *Applied Statistics* **24**: 377-379.

Bisson, J. L. and F. Roberge (1983). "Prévision des apports naturels: Expérience d'Hydro-Québec." *Paper presented at the Workshop on flow predictions, Toronto*

Brochu, R. and R. Laprise (2007). "Surface water and energy budgets over the Mississippi and Columbia River basins as simulated by two generations of the Canadian regional climate model." *Atmosphere-Ocean* **45**(1): 19 - 35.

Bronstert, A. (2003). "Floods and Climate Change: Interactions and Impacts." *Risk Analysis* **23**(3): 545-557.

Carlstein, E. (1986). "The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence." *The Annals of Statistics* **14**(3): 1171-1179.

Chan, S.-O. and R. L. Bras (1979). "Urban storm water management: Distribution of flood volumes." *Water Resour. Res.* **15**(2): 371-382.

Chebana, F. and T. B. M. J. Ouarda (2011). "Multivariate quantiles in hydrological frequency analysis." *Environmetrics* **22**(1): 63-78.

Córdova, J. R. and I. Rodríguez-Iturbe (1985). "On the probabilistic structure of storm surface runoff." *Water Resour. Res.* **21**(5): 755-763.

Davison, A. C. and D. V. Hinkley (2009). Bootstrap methods and their application. New York, *Cambridge university press*.

DiCiccio, T. J. and B. Efron (1996). "Bootstrap confidence intervals " *Statistical Science* **11**: 189-228.

Douglas, E. M., R. M. Vogel and C. N. Kroll (2000). "Trends in floods and low flows in the United States: impact of spatial correlation." *Journal of Hydrology* **240**(1-2): 90-105.

Efron, E. (1987). "Better bootstrap confidence intervals." *Journal of the American Statistical Association* **82**: 171–185.

Fleming, S. W. and G. K. C. Clarke (2002). "Autoregressive Noise, Deserialization, and Trend Detection and Quantification in Annual River Discharge Time Series." *Canadian Water Resources Journal* **27**(3): 335-354.

Fortin, J. P., R. Moussa, C. Bocquillon and J. P. Villeneuve (1995). "Hydrotel, un modèle hydrologique distribué pouvant bénéficier des données fournies par la télédétection et les systèmes d'information géographique." *Revue des sciences de l'eau* **8**(1): 97-124.

Goel, N. K., R. S. Kurothe, B. S. Mathur and R. M. Vogel (2000). "A derived flood frequency distribution for correlated rainfall intensity and duration." *Journal of Hydrology* **228**(1–2):

541 56-67.
542 Hamed, K. H. and A. Ramachandra Rao (1998). "A modified Mann-Kendall trend test for
543 autocorrelated data." *Journal of Hydrology* **204**(1-4): 182-196.
544 Hipel, K. W. and A. I. McLeod (2005). Time Series Modelling of Water Resources and
545 Environmental Systems.
546 Hollander, M. and D. A. Wolfe (1973). Nonparametric Statistical Metho, *Wiley*.
547 Kao, S.-C. and R. S. Govindaraju (2007). "A bivariate frequency analysis of extreme rainfall with
548 implications for design." *De Michele, C.* **112**(D13): D13119.
549 Kao, S.-C. and R. S. Govindaraju (2007). "Probabilistic structure of storm surface runoff
550 considering the dependence between average intensity and storm duration of rainfall
551 events." *Water Resour. Res.* **43**(6): W06410.
552 Kendall, M. G. (1975). Rank Correlation Methods, *Griffin, London*.
553 Khaliq, M. N., T. B. M. J. Ouarda, P. Gachon, L. Sushama and A. St-Hilaire (2009).
554 "Identification of hydrological trends in the presence of serial and cross correlations: A
555 review of selected methods and their application to annual flow regimes of Canadian
556 rivers." *Journal of Hydrology* **368**(1-4): 117-130.
557 Kulkarni, A. and H. Von-Storch (1995). "Monte Carlo experiments on the effect of serial
558 correlation on the Mann-Kendall test of trend." *Meteorologische Zeitschrift* **4**(2).
559 Mann, H. B. (1945). "Nonparametric tests against trend." *Econometric* **13**: 245-259.
560 Music, B. and D. Caya (2007). "Evaluation of the Hydrological Cycle over the Mississippi River
561 Basin as Simulated by the Canadian Regional Climate Model (CRCM)." *Journal of*
562 *Hydrometeorology* **8**(5): 969-988.
563 Pacher, G. (2006). Détermination objective des paramètres des hydrogrammes. Montréal,
564 Ouranos: 8.
565 Patton, A., D. N. Politis and H. White (2009). "Correction to "Automatic Block-Length Selection
566 for the Dependent Bootstrap" by D. Politis and H. White." *Econometric Reviews* **28**(4):
567 372 - 375.
568 Politis, D. N. and H. White (2004). "Automatic Block-Length Selection for the Dependent
569 Bootstrap." *Patton, Andrew* **23**(1): 53 - 70.
570 Reynard, N. S., C. Prudhomme and S. M. Crooks (2001). "The Flood Characteristics of Large
571 U.K. Rivers: Potential Effects of Changing Climate and Land Use." *Climatic Change*
572 **48**(2): 343-359.
573 Seidou, O. and T. B. M. J. Ouarda (2007). "Recursion-based multiple changepoint detection in
574 multiple linear regression and application to river streamflows." *Water Resour. Res.*
575 **43**(7): W07404.
576 Shiau, J. T. (2003). "Return period of bivariate distributed extreme hydrological events."
577 *Stochastic Environmental Research and Risk Assessment* **17**(1-2): 42-57.
578 Wood, E. F. (1976). "An analysis of the effects of parameter uncertainty in deterministic
579 hydrologic models." *Water Resour. Res.* **12**(5): 925-932.
580 Yue, S., T. B. M. J. Ouarda, B. Bobée, P. Legendre and P. Bruneau (1999). "The Gumbel mixed
581 model for flood frequency analysis." *Journal of Hydrology* **226**(1-2): 88-100.
582 Yue, S., P. Pilon, B. Phinney and G. Cavadias (2002). "The influence of autocorrelation on the
583 ability to detect trend in hydrological series." *Hydrological Processes* **16**(9): 1807-1829.
584 Yue, S. and C. Wang (2004). "The Mann-Kendall Test Modified by Effective Sample Size to
585 Detect Trend in Serially Correlated Hydrological Series." *Water Resources Management*
586 **18**(3): 201-218.
587 Zhang, L. and V. P. Singh (2006). "Bivariate flood frequency analysis using the copula method."

588 *Journal of Hydrologic Engineering* **11**(2): 150-164.
589 Zhang, X., K. D. Harvey, W. D. Hogg and T. R. Yuzyk (2001). "Trends in Canadian
590 streamflow." *Water Resour. Res.* **37**(4): 987-998.
591
592
593

Tables and Figures

Table 1 : Data series notations

Watershed	Type	Hydrological model	CRCM	Scenario	Period	Notation
Baskatong	Observed	-	-	-	1969-2000	DH
	Simulated	HSAMI	ERA-40	-	1969-2000	ERA40S
			CGCM3	A2	1961-2070	CGCM3S (A2)
		HYDROTEL	ERA-40	-	1969-2000	ERA40Y
			CGCM3	A2	1961-2070	CGCM3Y (A2)
Romaine	Observed	-	-	-	1956-2009	DH
	Simulated	HSAMI	CGCM3	A2	1961-2070	CGCM3 (A2)
				B1	1961-2071	CGCM3 (B1)
			HADCM3	A2	1961-2072	HADCM3 (A2)
				B2	1961-2073	HADCM3 (B2)
			ECHAM5	A2	1961-2074	ECHAM5 (A2)
				AB1	1961-2075	ECHAM5 (B1)

599 **Table 2 : Stationarity tests of the dependence series in Baskatong reservoir.**

	Series	p-value					
		MK	BB	PW	TFPW	VC1	VC2
Pearson's correlation	CGCM3S (A2)_QV	0.06	0.54	0.64	0.43	0.77	0.99
	CGCM3Y (A2)_QV	0.00	0.45	0.37	0.69	0.41	0.96
	CGCM3S (A2)_VD	0.00	0.56	0.06	0.97	0.33	0.95
	CGCM3Y (A2)_VD	0.00	0.50	0.09	1.00	0.23	0.94
	CGCM3S (A2)_QD	0.00	0.54	0.24	0.74	0.38	0.96
	CGCM3Y (A2)_QD	0.34	0.52	0.64	0.87	0.88	0.99
Kendall's tau	CGCM3S (A2)_QV	0.00	0.53	0.59	0.78	0.64	0.98
	CGCM3Y (A2)_QV	0.00	0.46	0.54	0.96	0.46	0.97
	CGCM3S (A2)_VD	0.00	0.52	0.13	0.94	0.33	0.96
	CGCM3Y (A2)_VD	0.00	0.49	0.14	0.99	0.21	0.94
	CGCM3S (A2)_QD	0.00	0.51	0.63	0.58	0.36	0.96
	CGCM3Y (A2)_QD	0.02	0.59	0.24	0.79	0.71	0.98
Spearman's rho	CGCM3S (A2)_QV	0.01	0.51	0.77	0.60	0.70	0.98
	CGCM3Y (A2)_QV	0.00	0.46	0.38	0.88	0.42	0.96
	CGCM3S (A2)_VD	0.00	0.53	0.09	0.71	0.43	0.96
	CGCM3Y (A2)_VD	0.00	0.49	0.13	0.94	0.21	0.94
	CGCM3S (A2)_QD	0.00	0.53	0.47	0.72	0.41	0.96
	CGCM3Y (A2)_QD	0.07	0.59	0.35	0.77	0.78	0.99

600 The gray color indicates that the dependence series is non-stationary according to the
601 corresponding test
602

603 **Table 3a : Stationarity tests of the dependence series between Q & V in Romaine**

	Series	p-value					
		MK	BB	PW	TFPW	VC1	VC2
Pearson's correlation	DH	0.50	0.62	0.02	0.09	0.83	0.98
	CGCM3 (A2)	0.00	0.50	0.78	0.91	0.62	0.98
	CGCM3 (B1)	0.00	0.49	0.83	0.98	0.42	0.96
	HADCM3 (A2)	0.00	0.52	0.09	0.90	0.38	0.96
	HADCM3 (B2)	0.00	0.57	0.10	0.83	0.51	0.97
	ECHAM5 (A2)	0.00	0.55	0.32	0.31	0.40	0.96
	ECHAM5 (B1)	0.00	0.46	0.78	0.65	0.56	0.97
Kendall's tau	DH	1.00	0.96	0.05	0.37	1.00	1.00
	CGCM3 (A2)	0.41	0.56	0.52	0.26	0.90	0.99
	CGCM3 (B1)	0.00	0.52	0.49	0.13	0.26	0.95
	HADCM3 (A2)	0.05	0.57	0.43	0.44	0.76	0.99
	HADCM3 (B2)	0.00	0.66	0.05	0.63	0.42	0.96
	ECHAM5 (A2)	0.00	0.58	0.16	0.16	0.27	0.95
	ECHAM5 (B1)	0.58	0.50	0.44	0.76	0.93	1.00
Spearman's rho	DH	0.80	0.85	0.05	0.05	0.94	0.99
	CGCM3 (A2)	0.53	0.55	0.46	0.30	0.92	1.00
	CGCM3 (B1)	0.00	0.49	0.50	0.22	0.23	0.94
	HADCM3 (A2)	0.00	0.57	0.31	0.99	0.61	0.98
	HADCM3 (B2)	0.00	0.66	0.08	1.00	0.50	0.97
	ECHAM5 (A2)	0.00	0.58	0.20	0.05	0.28	0.95
	ECHAM5 (B1)	0.67	0.50	0.47	0.67	0.95	1.00

604 The gray color indicates that the dependence series is non-stationary according to the
605 corresponding test
606

607 **Table 3b : Stationarity tests of the dependence series between *V* & *D* in Romaine**

	Series	p-value					
		MK	BB	PW	TFPW	VC1	VC2
Pearson's correlation	DH	0.00	0.71	0.06	0.53	0.12	0.85
	CGCM3 (A2)	0.24	0.52	0.42	0.98	0.85	0.99
	CGCM3 (B1)	0.00	0.52	0.03	0.44	0.53	0.97
	HADCM3 (A2)	0.07	0.50	0.26	0.12	0.78	0.99
	HADCM3 (B2)	0.03	0.53	0.81	0.60	0.74	0.98
	ECHAM5 (A2)	0.02	0.57	0.04	0.17	0.70	0.98
	ECHAM5 (B1)	0.00	0.45	0.34	0.85	0.51	0.97
Kendall's tau	DH	0.00	0.80	0.02	0.41	0.25	0.89
	CGCM3 (A2)	0.68	0.54	0.88	0.88	0.95	1.00
	CGCM3 (B1)	0.00	0.52	0.05	0.95	0.24	0.94
	HADCM3 (A2)	0.00	0.45	0.97	0.20	0.41	0.96
	HADCM3 (B2)	0.00	0.46	0.08	0.19	0.65	0.98
	ECHAM5 (A2)	0.34	0.62	0.17	0.17	0.88	0.99
	ECHAM5 (B1)	0.02	0.45	0.81	0.83	0.72	0.98
Spearman's rho	DH	0.00	0.79	0.03	0.32	0.24	0.89
	CGCM3 (A2)	0.90	0.52	0.74	1.00	0.98	1.00
	CGCM3 (B1)	0.00	0.50	0.04	0.94	0.24	0.95
	HADCM3 (A2)	0.00	0.44	0.83	0.17	0.42	0.96
	HADCM3 (B2)	0.02	0.47	0.05	0.11	0.71	0.98
	ECHAM5 (A2)	0.18	0.55	0.11	0.19	0.83	0.99
	ECHAM5 (B1)	0.01	0.46	0.83	0.78	0.69	0.98

608 The gray color indicates that the dependence series is non-stationary according to the
609 corresponding test
610

611 **Table 3c : Stationarity tests of the dependence series between *Q* & *D* in Romaine**

	Series	p-value					
		MK	BB	PW	TFPW	VC1	VC2
Pearson's correlation	DH	0.00	0.66	0.49	0.11	0.07	0.83
	CGCM3 (A2)	0.66	0.47	0.30	0.20	0.94	1.00
	CGCM3 (B1)	0.00	0.52	0.76	0.47	0.59	0.98
	HADCM3 (A2)	0.00	0.53	0.87	0.24	0.60	0.98
	HADCM3 (B2)	0.03	0.49	0.35	0.41	0.73	0.98
	ECHAM5 (A2)	0.00	0.47	0.00	0.14	0.39	0.96
	ECHAM5 (B1)	0.01	0.53	0.08	0.56	0.68	0.98
Kendall's tau	DH	0.00	0.77	0.04	0.53	0.13	0.85
	CGCM3 (A2)	0.99	0.47	0.05	0.07	1.00	1.00
	CGCM3 (B1)	0.00	0.57	0.78	0.71	0.58	0.97
	HADCM3 (A2)	0.00	0.49	0.73	0.15	0.35	0.96
	HADCM3 (B2)	0.00	0.47	0.15	0.43	0.54	0.97
	ECHAM5 (A2)	0.00	0.51	0.01	0.20	0.59	0.97
	ECHAM5 (B1)	0.00	0.45	0.03	0.36	0.41	0.96
Spearman's rho	DH	0.00	0.73	0.12	0.79	0.09	0.84
	CGCM3 (A2)	0.73	0.48	0.09	0.09	0.96	1.00
	CGCM3 (B1)	0.00	0.57	0.99	0.69	0.65	0.98
	HADCM3 (A2)	0.00	0.49	0.66	0.13	0.41	0.96
	HADCM3 (B2)	0.00	0.50	0.09	0.52	0.52	0.97
	ECHAM5 (A2)	0.00	0.53	0.04	0.37	0.55	0.97
	ECHAM5 (B1)	0.00	0.46	0.02	0.51	0.29	0.95

612 The gray color indicates that the dependence series is non-stationary according to the
613 corresponding test
614

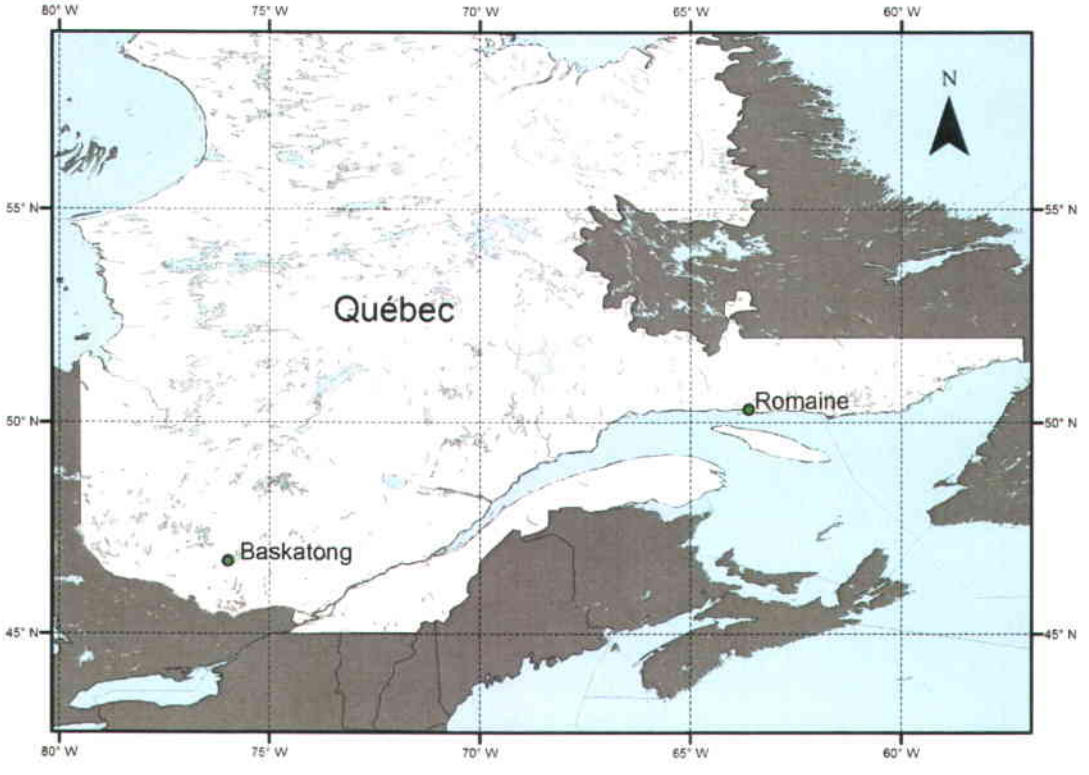
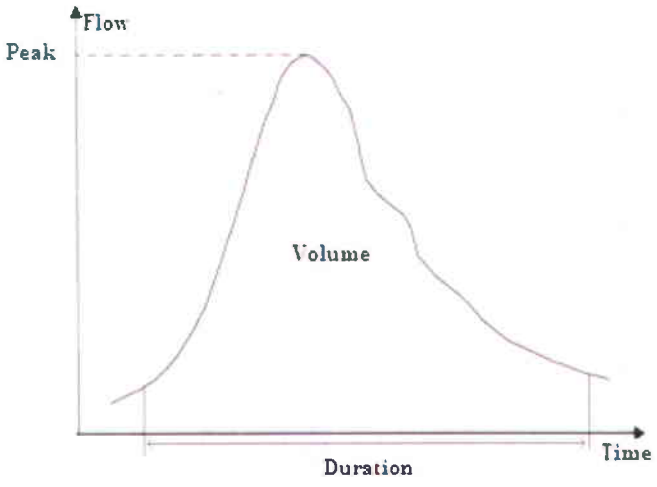


Figure 1: Location of the two stations representing the Baskatong reservoir and Romaine River

620



621

622 **Figure 2: Main flood characteristics**

623

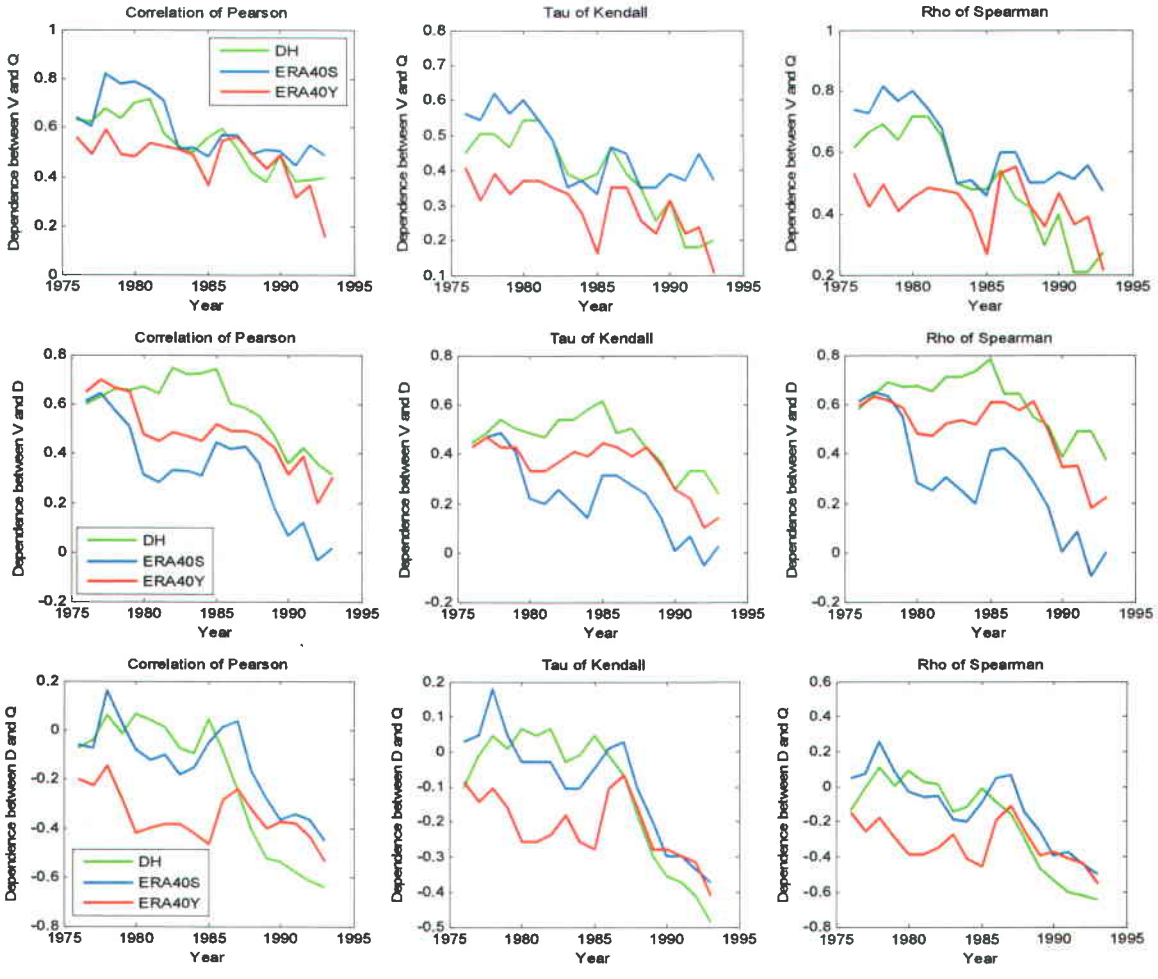


Figure 3: Dependence series for Baskatong reservoir in the historical period (moving window of 15 years) with a) dependence between *Q* & *V*, b) dependence between *V* & *D* and c) dependence between *Q* & *D*.

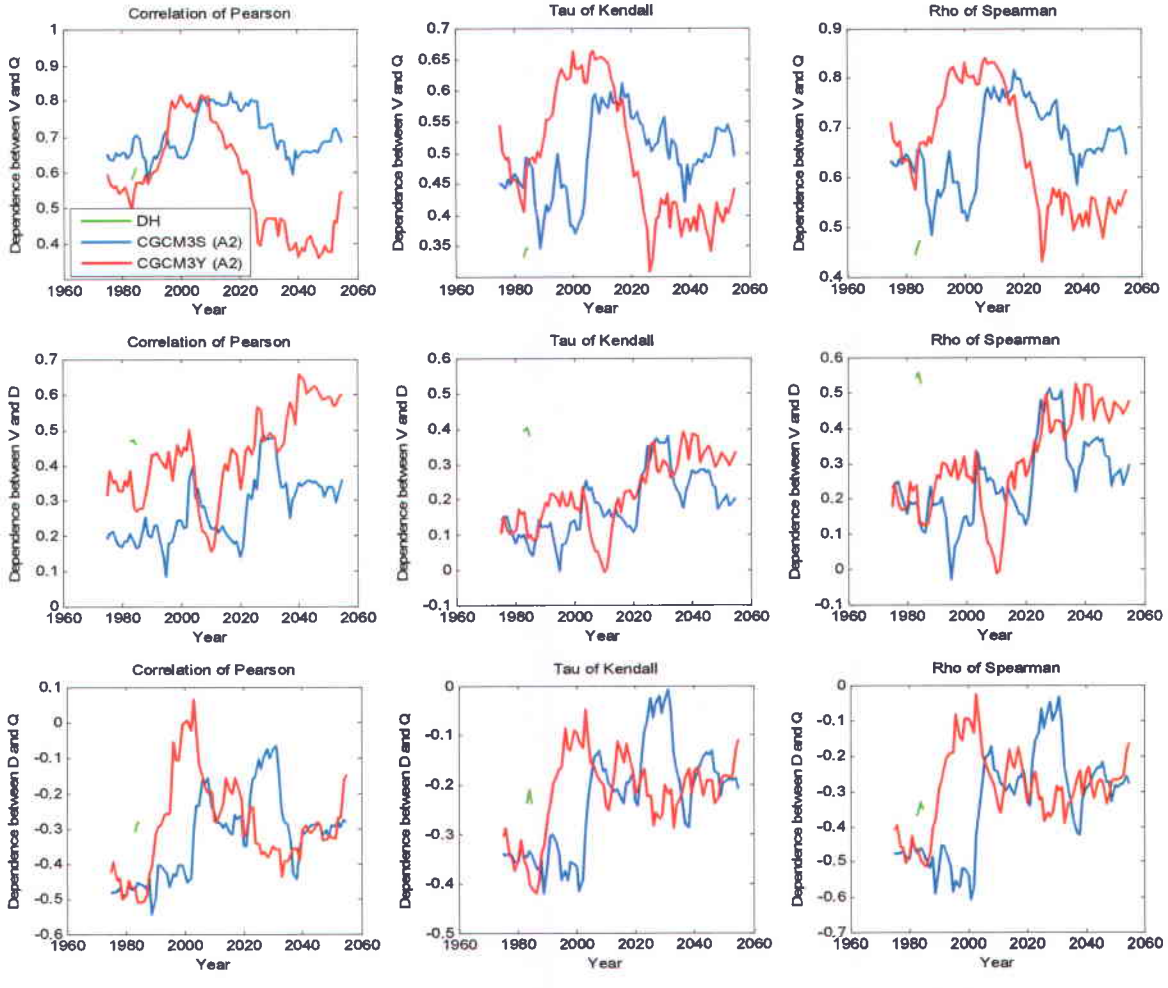


Figure 4 : Dependence series for Baskatong reservoir (moving window of 30 years) with a) dependence between *Q* & *V*, b) dependence between *V* & *D* and c) dependence between *Q* & *D*.

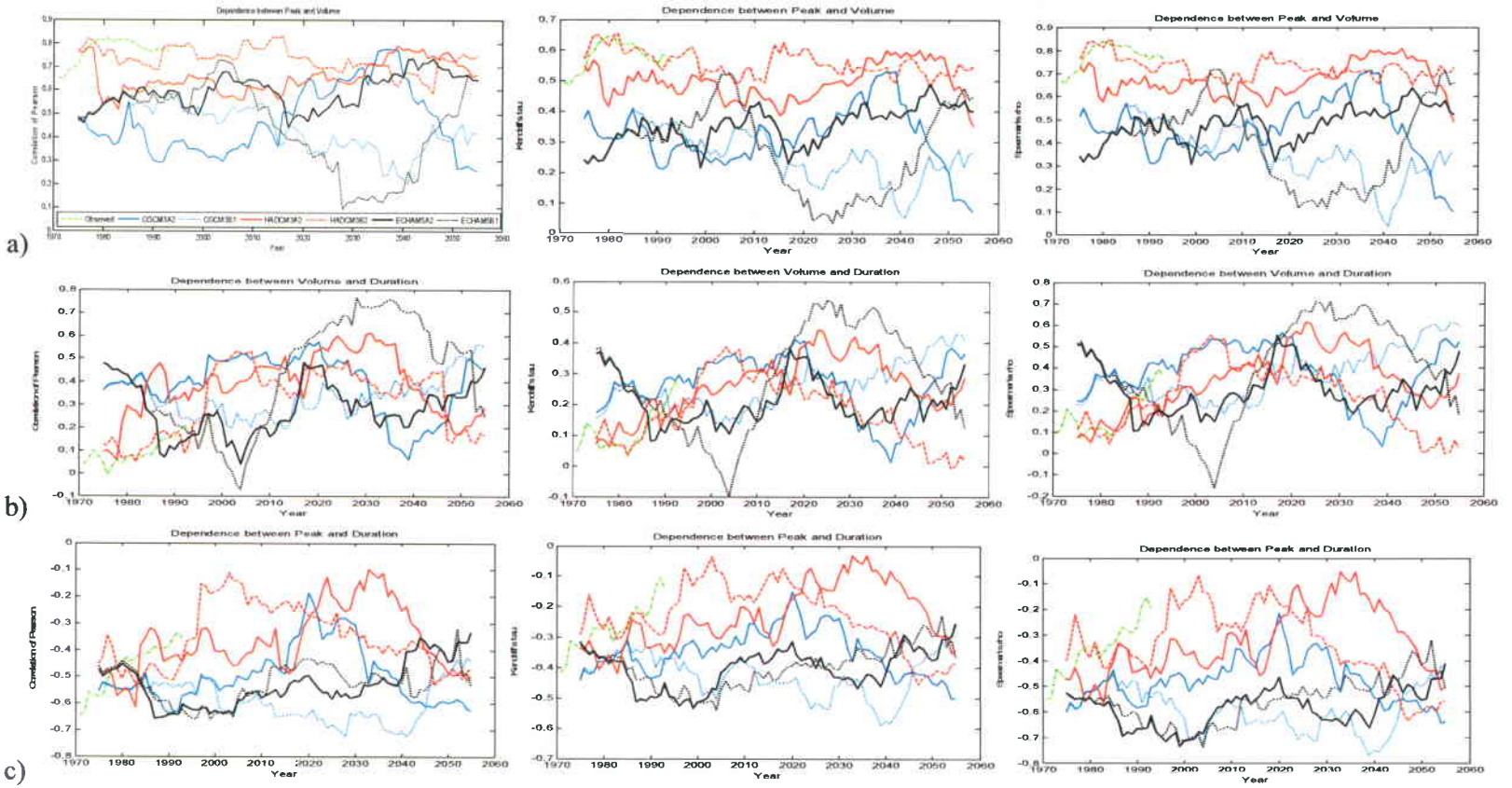


Figure 5 : Dependence series for Romaine watershed (moving window of 30 years) with a) dependence between *V*& *Q*, b) dependence between *V* & *D* and c) dependence between *Q* & *D*.

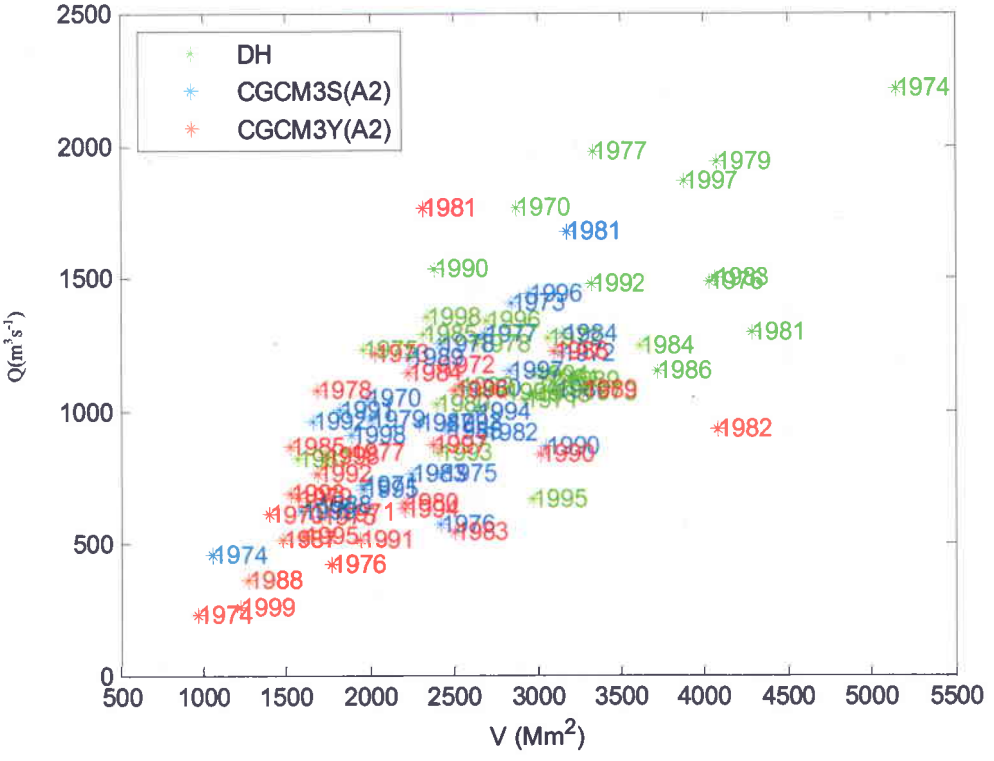
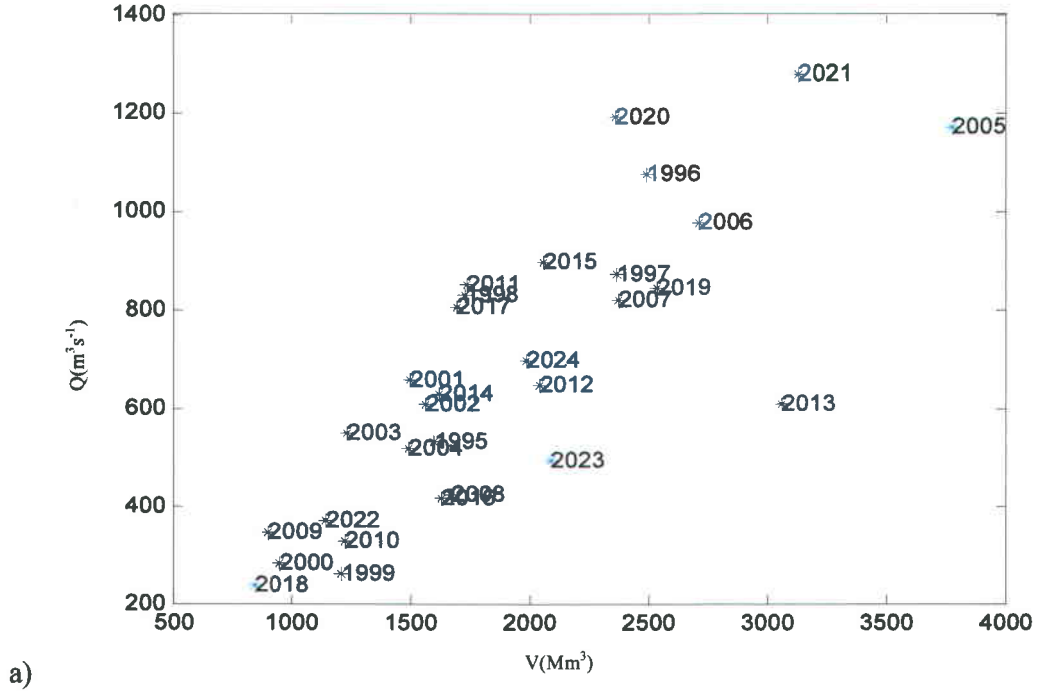
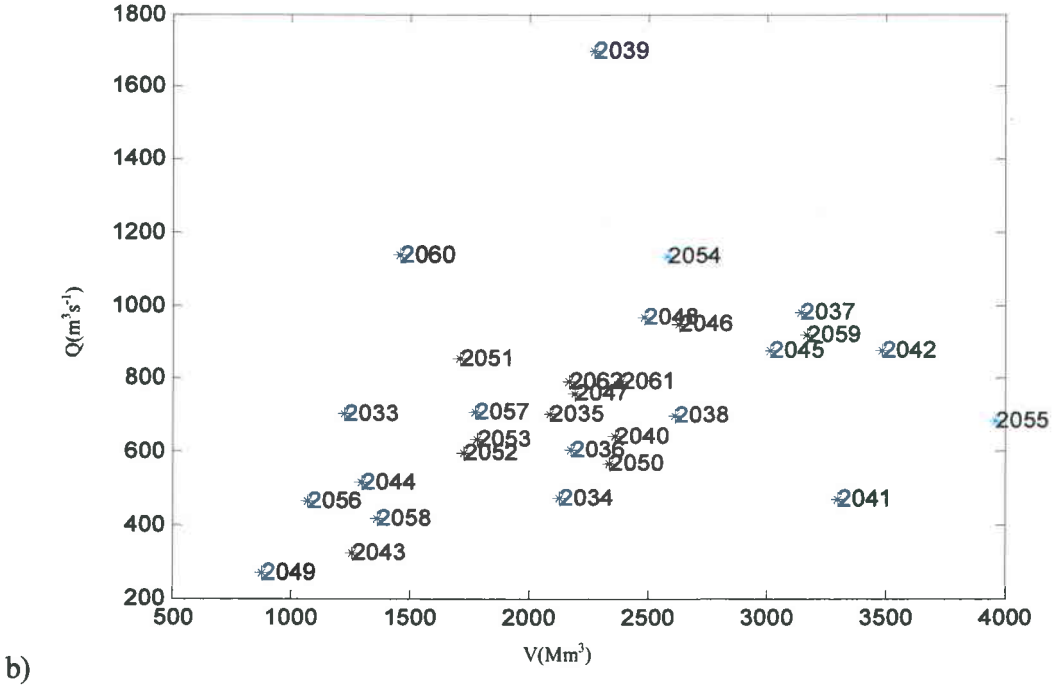


Figure 6 : Scatter plots of (*Q*, *V*) for DH, CGCM3S (A2) and CGCM3Y (A2) over 1970-1999 in Baskatong.

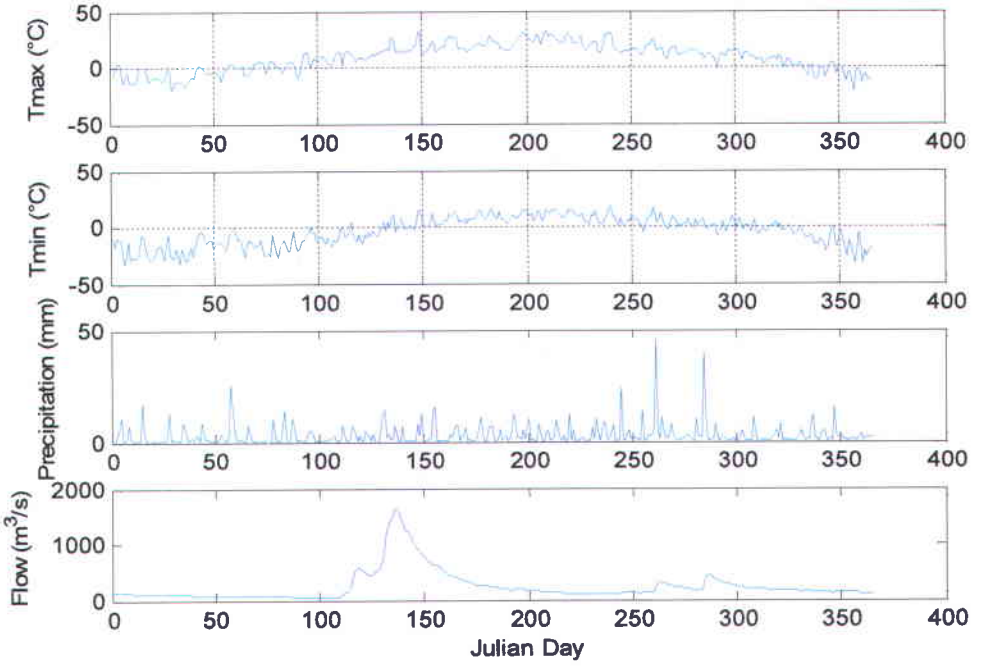


a)

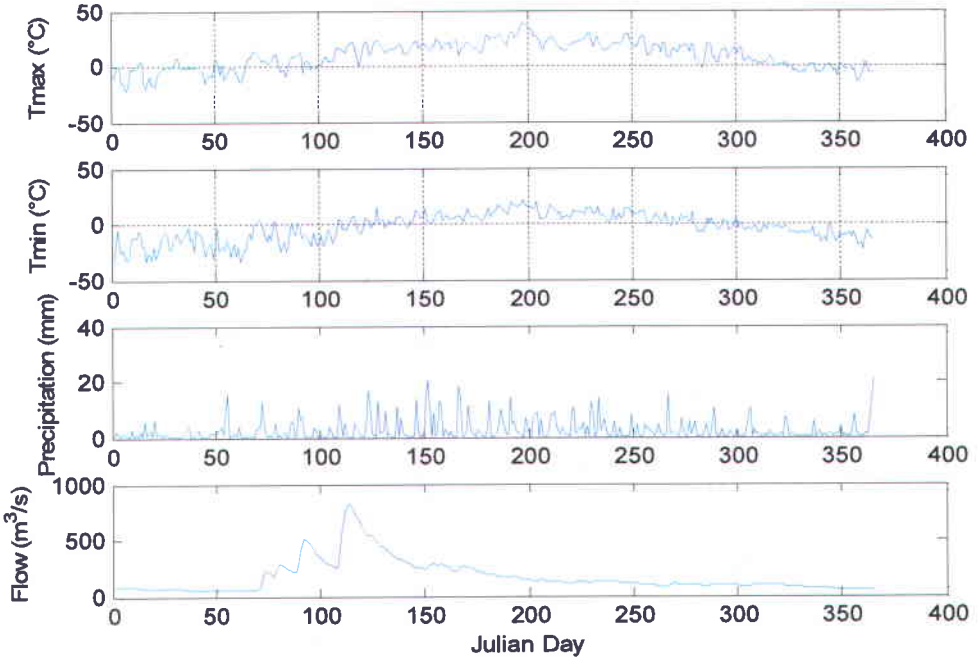


b)

Figure 7 : Scatter plots of (Q, V) for CGCM3Y (A2) over a) 1995-2024 and b) 2033-2062 in Baskatong



a)



b)

Figure 8 : Minimum and maximum temperature, precipitation and daily flow of CGCM3Y (A2) in Baskatong for a) 2039 and b) 2041

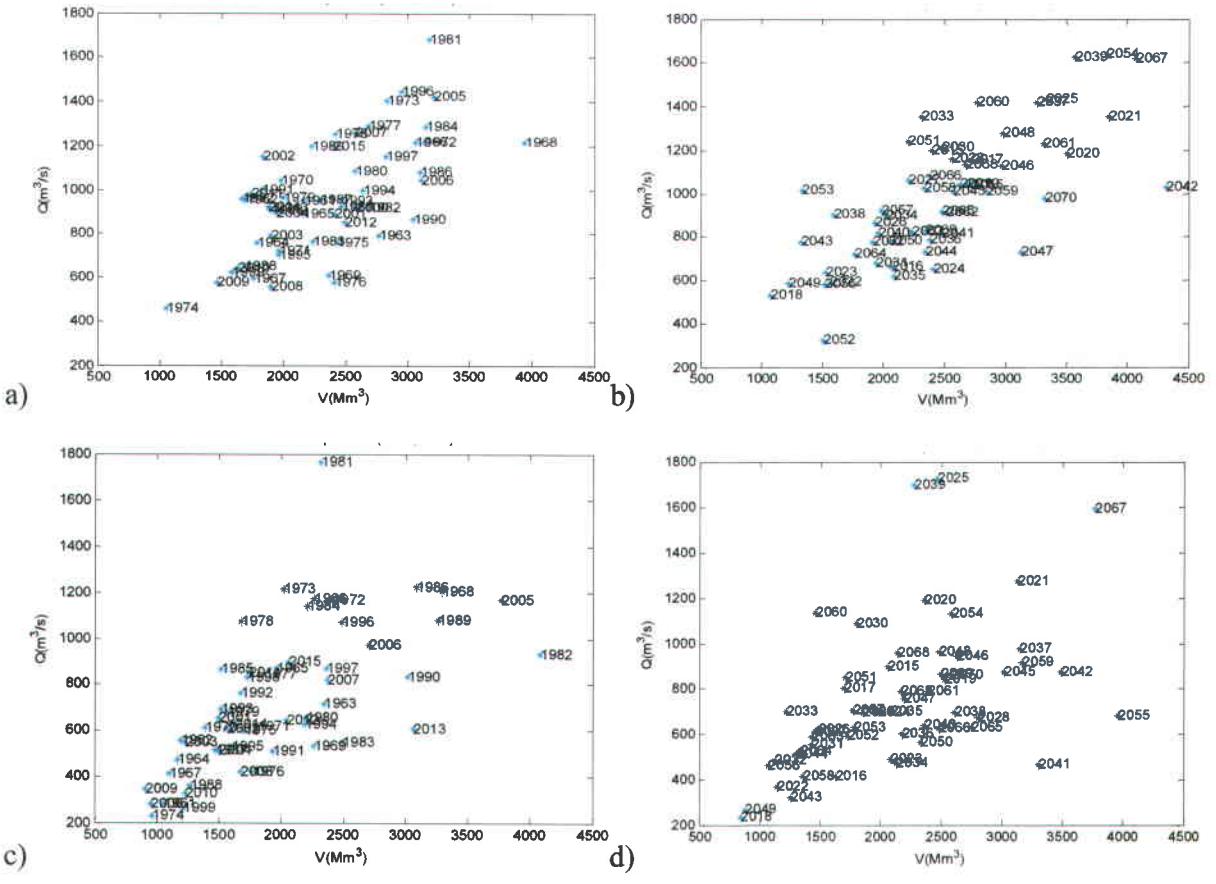


Figure 9 : Scatter plots of (Q, V) in Baskatong for a) CGCM3S (A2) over 1961-2015, b) CGCM3S (A2) over 2015-2070, c) CGCM3Y (A2) over 1961-2015 and d) CGCM3Y (A2) over 2015-2070

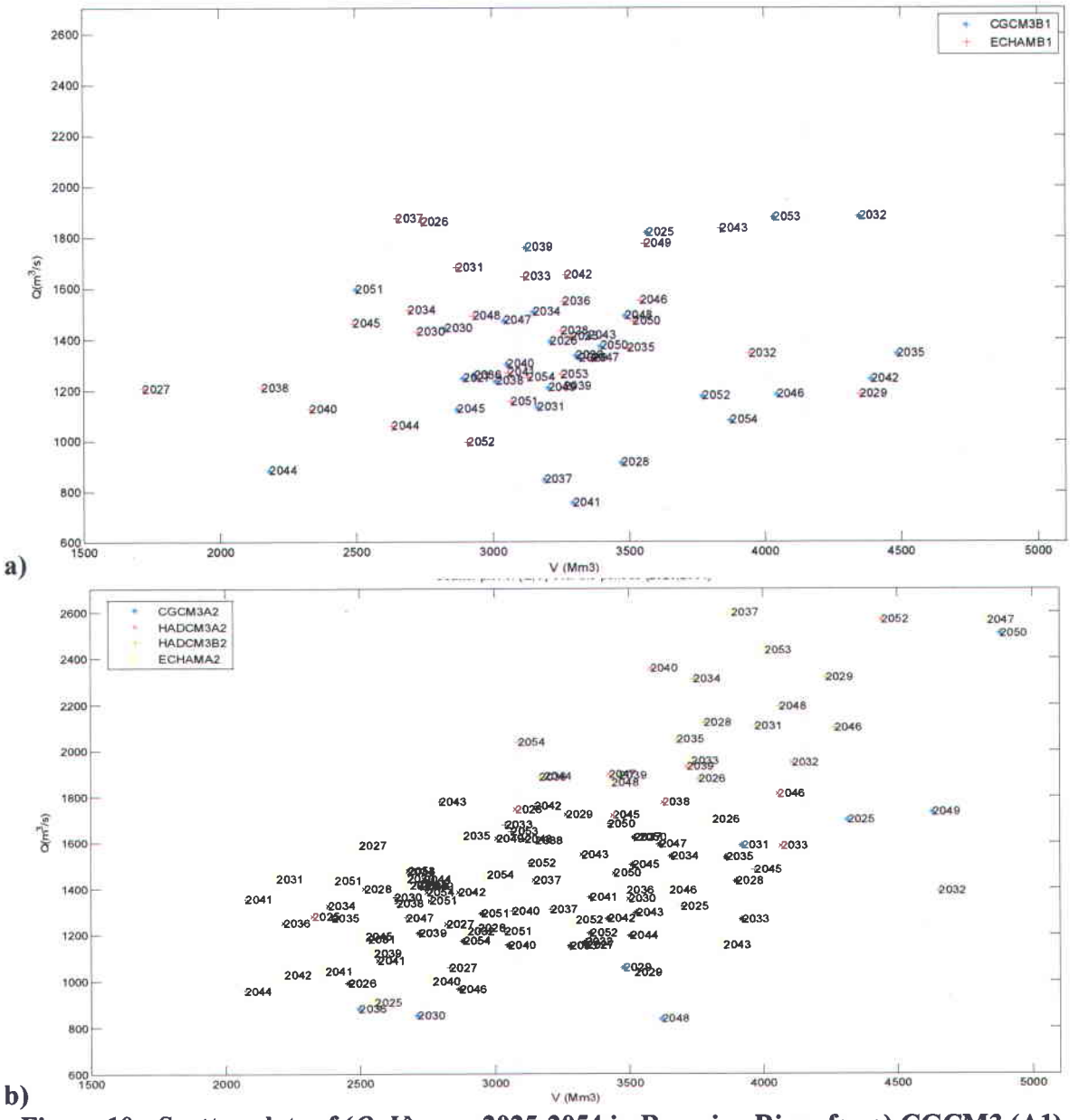


Figure 10 : Scatter plots of (Q, V) over 2025-2054 in Romaine River for a) CGCM3 (A1) and ECHAM5 (B1) and b) CGCM3 (A2), HADCM3 (A2), HADCM3 (B2), and ECHAM5 (A2)

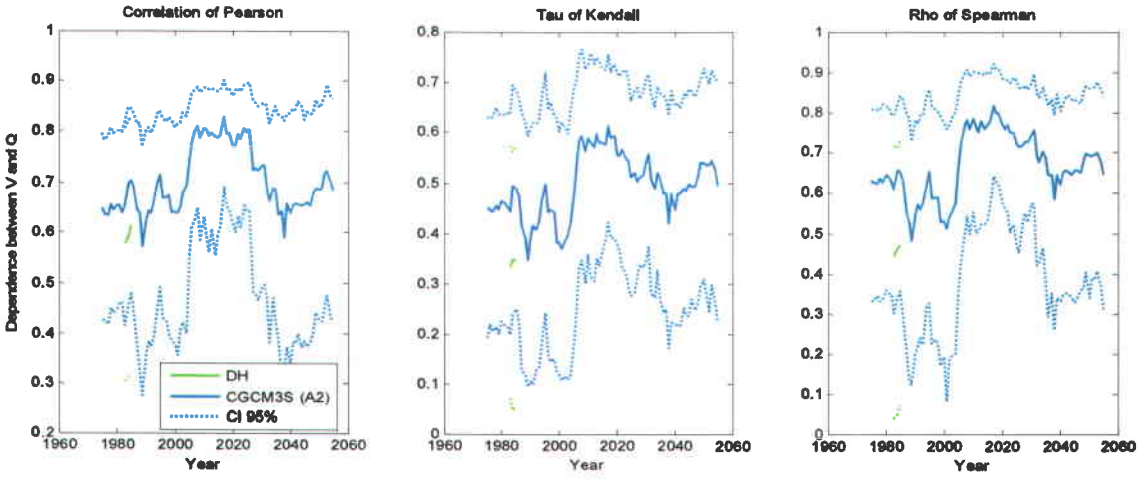


Figure 11 : CI for dependence series in Baskatong reservoir between *Q* & *V* for CGCM3S (A2)

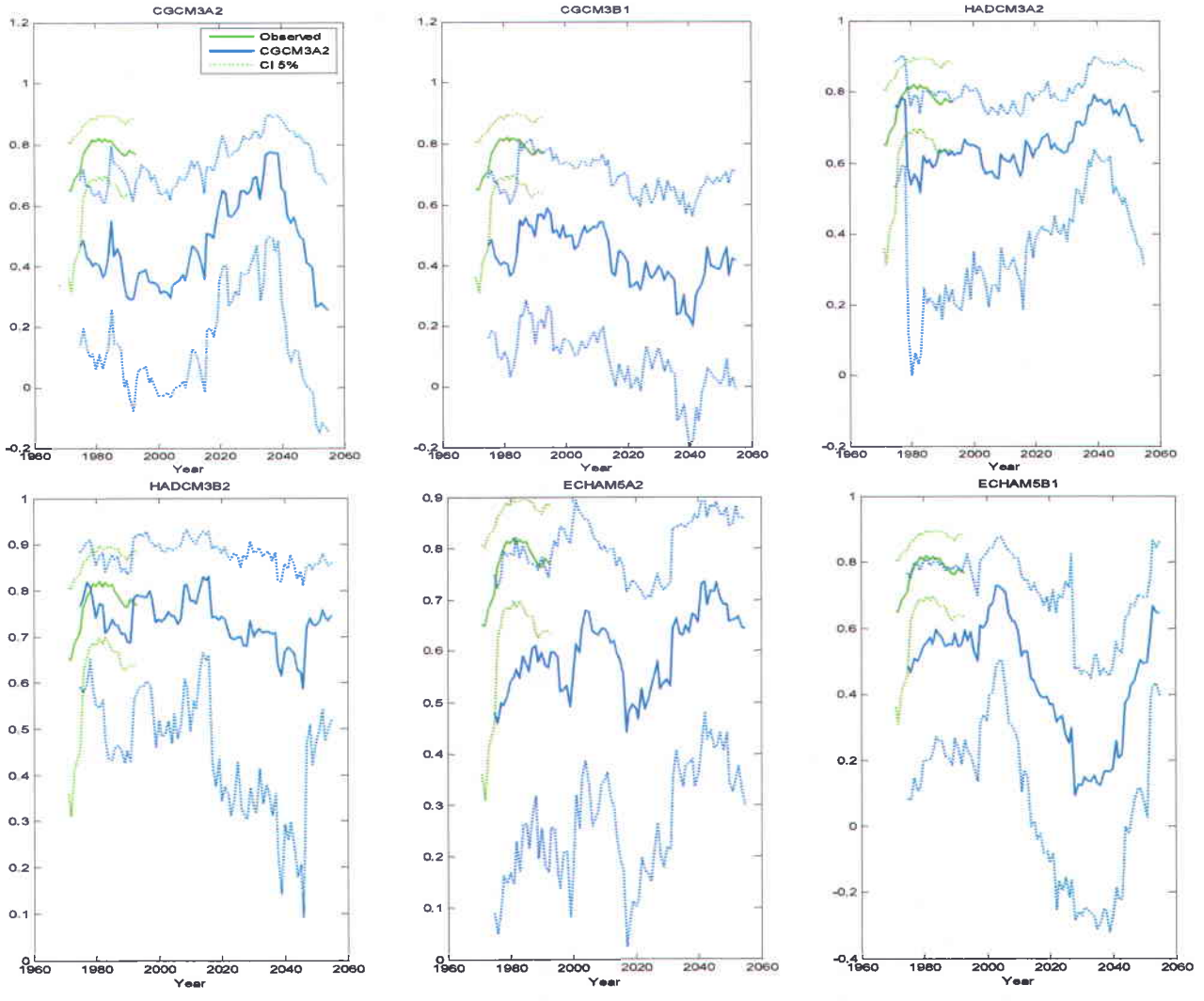


Figure 12 : CI for Pearson's correlation series between *Q* & *V* for Romaine River.

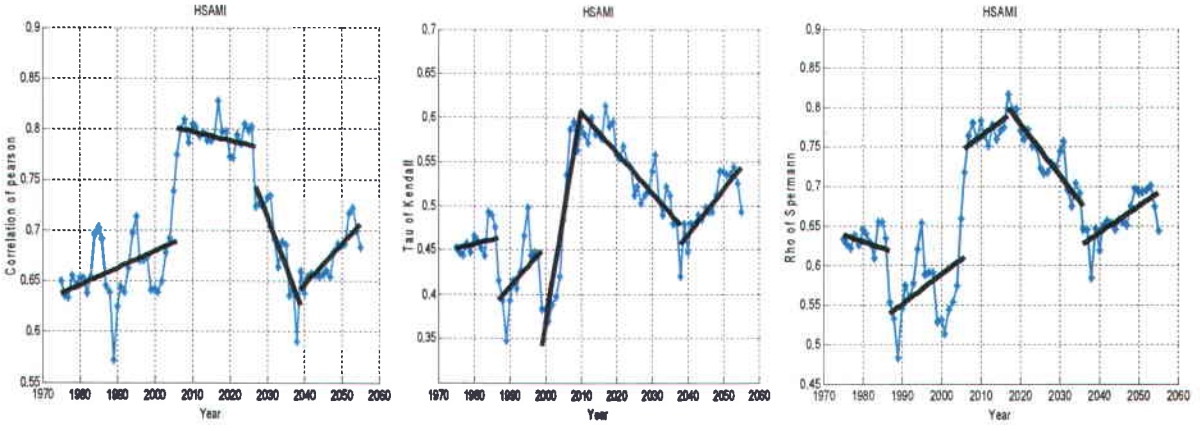


Figure 13: Break-points detection in considered dependence series between Q & V in Baskatong reservoir

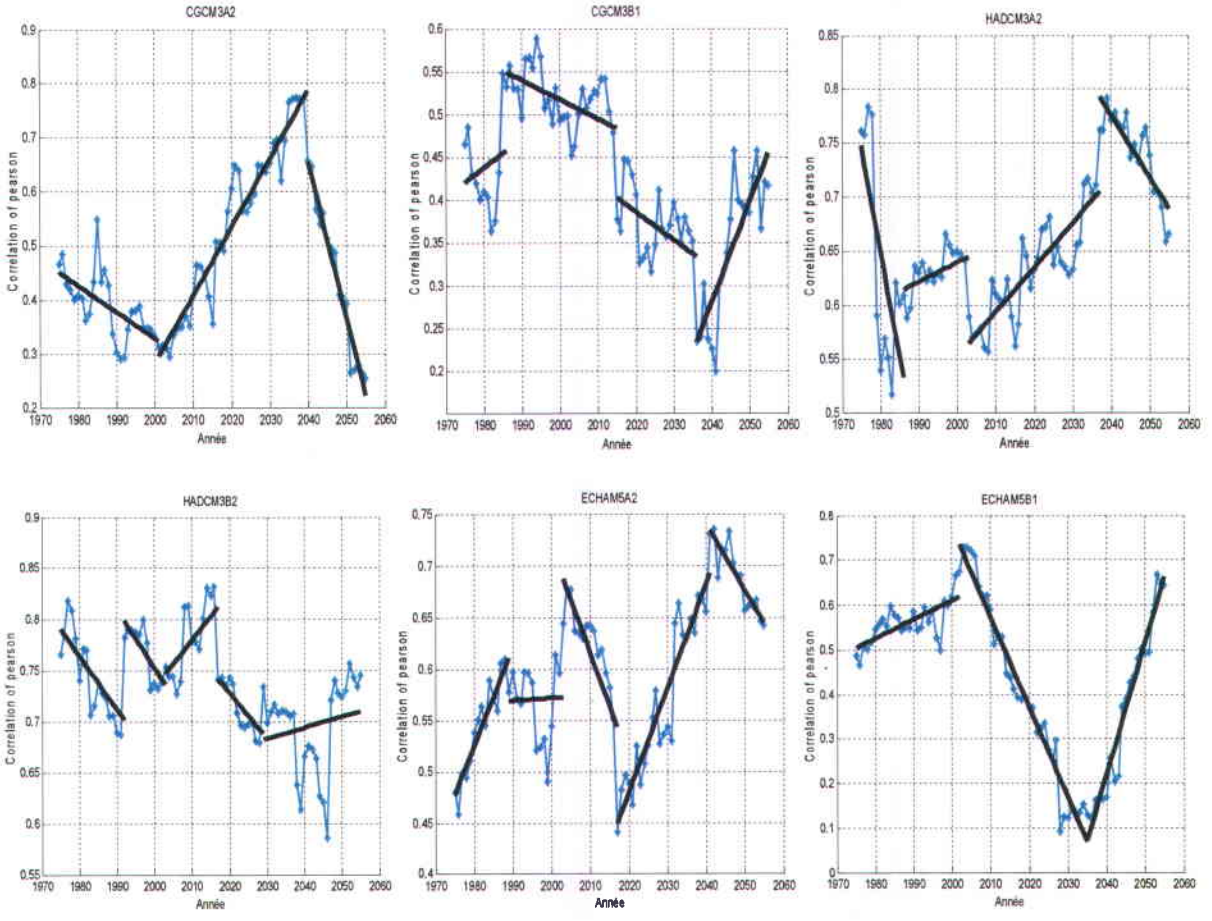


Figure 14 : Break-point detection in the Pearson's correlation series between *Q* & *V* for Romaine River.

11 Article 3: Bivariate index-flood model for a northern case study

Bivariate index-flood model for a northern case study

M.-A. Ben Aissia* [1]

F. Chebana [1]

T. B. M. J. Ouarda [1,2]

P. Bruneau [3]

M. Barbet [3]

[1] Statistics in Hydrology Working Group. Canada Research Chair on the Estimation of Hydrological Variables, INRS-ETE,

490, de la Couronne, Quebec, (Quebec) Canada, G1K 9A9.

[2] Institute Center for Water and Environment (iWATER)

Masdar Institute of Science and Technology,

PO Box 54224, Abu Dhabi, UAE.

[3] Hydro-Québec Équipement

855, Ste-Catherine East, 19th Floor, Montreal (Quebec) Canada, H2L 4P5.

* Corresponding author: mohamed.ben.aissia@ete.inrs.ca

October 2013

Abstract

Floods, as extreme hydrological phenomena, can be described by more than one correlated characteristic such as peak, volume and duration. These characteristics should be jointly considered since they are generally not independent. For an ungauged site, univariate regional flood frequency analysis (FA) provides a limited assessment of flood events. A recent study proposed a procedure for regional FA in a multivariate framework. This procedure represents a multivariate version of the index-flood model and is based on copulas and multivariate quantiles. The performance of the proposed procedure was evaluated by simulation. However, the model was not tested on a real-world case study data. In the present paper, practical aspects are investigated jointly for flood peak (Q) and volume (V) of a data set from the Côte-Nord region in the province of Quebec, Canada. The application of the proposed procedure requires the identification of the appropriate marginal distribution, the estimation of the index flood and the selection of an appropriate copula. The results of the case study show a good performance of the regional bivariate FA procedure. This performance depends strongly on the performance of the two univariate models and more specifically the univarite model of Q. Results show also the impact of the homogeneity of the region on the performance of the univariate and bivariate models.

41 # 1. Introduction and literature review

42 A flood can be described as a multivariate event whose main characteristics are peak,
43 volume and duration. Thus, the severity of a flood depends on these characteristics,
44 which are mutually correlated (Ashkar 1980, Yue et al. 1999, Ouarda et al. 2000, Yue
45 2001, Shiau 2003, De Michele et al. 2005, Zhang and Singh 2006, Chebana and Ouarda
46 2009, Chebana and Ouarda 2011). These studies show that these variables have to be
47 jointly considered.

48 The use of joint probabilistic behaviour of correlated variables is necessary to understand
49 the probabilistic characteristic of such events. Yue et al. (1999) used the bivariate
50 Gumbel mixed model with standard Gumbel marginal distributions to represent the joint
51 probability distribution of flood peak and volume, and flood volume and duration.
52 Ouarda et al (2000) were first to study the joint regional behaviour of flood peaks and
53 volume. To model flood peak and volume, Yue (2001) and Shiau (2003) used the
54 Gumbel logistic model with standard Gumbel marginal distributions. Recently, copulas
55 have been shown to represent a useful statistical tool to model the dependence between
56 variables. To model flood peak and volume with Gumbel and Gamma marginal
57 distribution respectively Zhang and Singh (2006) used the copula method, bivariate
58 distributions of flood peak and volume, and flood volume and duration in frequency
59 analysis (FA). Using the Gumbel–Hougaard copula, Zhang and Singh (2007) derived
60 trivariate distributions of flood peak, volume and duration in FA.

61 Generally, the record length of the available streamflow data at sites is much shorter than
62 the return period of interest and in some cases, there may not be any streamflow record at

these sites. Consequently, local frequency estimation is difficult and/or not reliable. Regional FA is hence commonly used to overcome this lack of data. It is based on the transfer of available information data from other stations within the same hydrologic region into a site where little or no data are available. The regional FA procedure was investigated with different approaches by several authors including Stedinger and Tasker (1986), Durrans and Tomic (1996), Nguyen and Pandey (1996), Hosking and Wallis (1997), Alila (1999, 2000) and Ouarda et al. (2001). GREHYS (1996a, 1996b) presented an intercomparison of various regional FA procedures.

In the literature, flood FA can be classified into four classes according to the univariate/multivariate and local/regional aspects. The local-univariate and regional-univariate classes were widely studied in the literature (Singh 1987, Wiltshire 1987, Burn 1990, Hosking and Wallis 1993, Hosking and Wallis 1997, Alila 1999, Ouarda et al. 2006, Nezhad et al. 2010). Recently, researchers have been increasingly interested in the multivariate case and many studies treated the problem of local-multivariate flood FA (Yue et al. 1999, Yue 2001, Shiau 2003, De Michele et al. 2005, Grimaldi and Serinaldi 2006, Zhang and Singh 2006, Chebana and Ouarda 2011). However, multivariate regional FA has received much less attention (Ouarda et al. 2000, Chebana and Ouarda 2007, Chebana and Ouarda 2009, Chebana et al. 2009).

The two main steps of the regional FA are the delineation of hydrological homogeneous regions and regional estimation (GREHYS 1996a). In the multivariate case, the delineation of hydrological homogeneous regions was treated by Chebana and Ouarda (2007). They proposed discordancy and homogeneity tests that are based on multivariate L-moments and copulas. Chebana et al. (2009) studied the practical aspects of these tests.

86 In univariate-regional FA, different methods were proposed to estimate extreme quantiles
87 such as regressive models and index-flood models (e.g. GREHYS 1996a, 1996b).
88 Chebana and Ouarda (2009) proposed a procedure for regional FA in a multivariate
89 framework. The proposed procedure represents a multivariate version of the index-flood
90 model. In this method, it is assumed that the distribution of flood characteristics (flood,
91 peak or volume) at different sites within a given flood region is the same except for a
92 scale parameter. Chebana and Ouarda (2009) adopted the multivariate quantile as the
93 curve formed by the combination of variables corresponding to the same risk (Chebana
94 and Ouarda 2011). In order to model the dependence between variables describing the
95 event they employed the copula. In the present paper, practical aspects of the proposed
96 procedure by Chebana and Ouarda (2009) are studied. Real data sets from sites in the
97 Côte Nord region in the northern part of the province of Quebec, Canada are used. Flood
98 peak and volume are the two variables studied jointly in the present study.

99 The next section presents the theoretical background, including the bivariate modelling,
100 univariate index-flood model and multivariate quantiles. The "Multivariate Index-flood
101 Model" section details the methodology of the adopted procedure with an emphasis on
102 practical aspects. The case study section presents the study procedure as well as the
103 obtained results. Concluding remarks are presented in the last section.

104 ## 2. Background

105 In this section, the background elements to apply the index-flood model in the
106 multivariate regional FA procedure are presented. Bivariate modelling including copulas

and marginal distributions, univariate index-flood model and multivariate quantiles are briefly described.

2.1. *Bivariate flood modelling and copulas*

In bivariate modelling, a joint bivariate distribution for the underlying variables has to be obtained. According to Sklar's theorem (1959), the bivariate distribution is composed of a copula and two marginal which are not necessarily similar.

In the remainder of the paper, we denote F_X and F_Y respectively the marginal distribution functions of given random variables X and Y, and F_{XY} the joint distribution function of the vector (X,Y).

a) Copula

Due to their ability to overcome the limitation of classical joint distributions, copulas have received increasing attention in various fields of science (see e.g. Nelsen 2006). Copulas are used to describe and model the dependence structure between the two random variables. A copula is an independent function of marginal distributions. For more details on copula functions, see for instance Nelsen (2006), Chebana and Ouarda (2007) and Salvadori et al. (2007). According to Sklar's (1959) theorem, we can construct the bivariate distribution F_{XY} with margins F_X and F_Y by:

$$F_{XY}(x,y) = C\left[F_X(x), F_Y(y)\right] \text{ for all real x and y} \qquad (1)$$

When F_X and F_Y are continuous, the copula C is unique.

Different classes of copulas are studied in the literature such as the Archimedean, Elliptical, Extreme Value (EV), Plackette and Farlie-Gumbel-Morgenstern (FGM)

127 copulas (see e.g. Nelsen 2006, Salvadori et al. 2007). The use of a copula requires the
128 estimation of its parameters as well as goodness-of-fit procedures. In addition, since in
129 hydrology we are particularly interested by the risk, the tail dependence of copulas is also
130 a factor to take into account.

131 *Copula parameter estimation:* Assuming the unknown copula C belongs to a parametric
132 family $C_0 = \{C_\theta : \theta \in R^q\}; q \geq 2$. The estimation of the parameter vector $\boldsymbol{\theta}$ is the first step to
133 deal with. In the case of one-parameter bivariate copula, a popular approach consists of
134 using the method of moment-type based on the inversion of Spearman's ρ and Kendall's
135 τ. Demarta and McNeil (2005) have shown that such approach may lead to
136 inconsistencies. The maximum pseudo-likelihood (MPL) approach is shown to be
137 superior to the other ones (Besag 1975, Genest et al. 1995, Shih and Louis 1995, Kim et
138 al. 2007) in which the observed data are transformed via the empirical marginal
139 distributions to obtain pseudo-observations on which the maximum-likelihood approach
140 is based to estimate the associated copula parameters (Genest et al. 1995). The advantage
141 of this approach is that it can provide greater flexibility than the likelihood approach in
142 the representation of real data. It consists in maximizing the log pseudo-likelihood:

$$\log L(\theta) = \sum_{i=1}^{n} \log c_\theta \left(\hat{U}_i\right) \tag{2}$$

143 where c_θ denotes the density of a copula $C_\theta \in C_0$, and $\hat{U}_k = \left(\hat{U}_{kX}, \hat{U}_{kY}\right)^T$ are the pseudo-
144 observation obtained from $(X_k, Y_k)^T$ given by:

$$\hat{U}_{kl} = R_{kl} / (n+1), \quad k = 1, \ldots, n; \quad l = X \text{ or } Y \tag{3}$$

145 with R_{kX} being the rank of X_k among $X_1, \ldots, X_n$ and R_{lY} being the rank of Y_l among $Y_1, \ldots, Y_n$.

146 *Goodness-of-fit test*: The most important step in copula modelling is the copula selection
147 by the goodness-of-fit test. Formally, one wants to test the hypotheses:

$$H_0 : C \in C_0 \quad \text{against} \quad H_1 : C \notin C_0 \tag{4}$$

148 Due to the novelty of copula modelling in flood FA, there is no common goodness-of-fit
149 test for copulas. One of the most commonly used goodness-of-fit tests and valid only for
150 Archimedean copulas is the graphic test proposed by Genest and Rivest (1993) based on
151 the K function given by

$$K_\phi(u) = u - \frac{\phi(u)}{\phi'(u)} \quad 0 < u < 1 \tag{5}$$

152 where ϕ is the generator function of the Archimedean copula. The K function can be
153 estimated by

$$\hat{K}(u) = \frac{1}{N}\sum_{i=1}^{N} 1_{[w_i \le u]} \quad \text{where}$$
$$w_i = \frac{1}{N-1}\sum 1_{[u_1^i < u_1^i,\, u_2^i < u_2^i]}, \ \mathrm{i} = 1,..,\mathrm{N} \tag{6}$$

154 for a given bivariate sample $(u_1^1, u_2^1), (u_1^2, u_2^2), \ldots, (u_1^N, u_2^N)$. Genest and Revest (1993) have
155 shown that $\hat{K}$ is a consistent estimator of K under weak regularity conditions. Note that
156 Archimedean copulas are widely employed in hydrology and particularly to model flood
157 dependence.
158 Recently, a relatively large number of goodness-of-fit tests were proposed (see e.g.
159 Charpentier 2007, Genest et al. 2009, for extensive reviews). Genest et al. (2009) carried
160 out a power study to evaluate the effectiveness of various goodness-of-fit tests and
161 recommended a test based on a parametric bootstrapping procedure which makes use of
162 the Cramer-von Mises statistic S_n (S_n goodness-of-fit test) :

$$S_n = \int n\{C_n(u,v) - C_{\theta_n}(u,v)\}^2 dC_n(u,v) \tag{7}$$

where C_n is the empirical copula calculated using n observation data. and $C_{\theta n}$ is an estimation of C obtained assuming $C \in C_0$. The estimation $C_{\theta n}$ is based on the estimator θ_n of θ such as the maximum pseudo-likelihood estimator given in (2).

b) AIC for copula

In some cases, results of the goodness-of-fit testing show that more than one copula provide a good fit to the data set. To select the most adequate copula, we use the AIC (Akaike's information criterion) proposed by Kim et al. (2008) in the context of copulas:

$$\begin{aligned} AIC &= -2\log(L(\hat{\theta}; X, Y)) + 2r; \\ L(\hat{\theta}; X, Y) &= \sum \log\left\{ c\left(F_X(X), F_Y(Y), \hat{\theta}\right)\right\} \end{aligned} \tag{8}$$

where $\hat{\theta}$ is the estimation of the copula parameter vector $\boldsymbol{\theta}$, r is the dimension of $\boldsymbol{\theta}$ and c is the copula density.

The copula which has the lowest AIC value is the most adequate copula for the data set.

c) Marginal distributions

To selection of the most appropriate marginal distribution (for X and for Y). The choice of the appropriate distribution is based on the Chi-square goodness-of-fit test, graphics and selection criteria (AIC see e.g. Akaike (1973) and BIC see e.g. Schwarz (1978)). For parameter estimation, a number of methods are available in the literature to estimate marginal distribution parameters; such as, the method of moments, the maximum likelihood method and the L-moments method.

180 *2.2. Univariate Index-flood model*

181 Introduced by Dalrymple (1960), the index-flood model was used initially for regional
182 flood prediction. It is also used to model other hydrological variables including storms
183 and droughts (e.g. Pilon 1990, Hosking and Wallis 1997, Hamza et al. 2001, Grimaldi
184 and Serinaldi 2006). This model is based on the assumption of the homogeneity of the
185 considered region and all the sites in the region have the same frequency distribution
186 function apart from a scale parameter specific to each site. Let N_s be the number of sites
187 in the region. The model gives the quantile $Q_i(p)$ corresponding to the non-exceedance
188 probability p at site i as:

$$Q_i(p) = \mu_i q(p), \quad i = 1,....,N_s \quad \text{and} \quad 0 < p < 1 \tag{9}$$

189 where μ_i corresponds to the index flood and q is the regional growth curve.

190 The index flood parameter μ_i can be estimated using a number of approaches (Hosking
191 and Wallis 1997). For instance, Brath et al. (2001) used three models of estimating the
192 index flood parameter. These models are multi-regression model, rational model and
193 geomorphoclimatic model. They show that best results are given by considering the
194 multi-regression model of the form:

$$\hat{\mu}_i = a_0 A_1^{a_1} A_2^{a_2} A_3^{a_3} ... A_{np}^{a_n} \tag{10}$$

195 in which a_i are coefficients to be estimated, and $A_i, ..., A_{np}$ represent an appropriate set of
196 morphological and climatic characteristics of the basin such as watershed area and slope
197 of the main channel.

198 *2.3. Multivariate quantiles*

199 Unlike to the well-known univariate quantile, the multivariate quantile has received less
200 attention in hydrology. Despite that, a few studies proposed multivariate quantile
201 versions. For details, the reader is referred to Chebana and Ouarda (2011). The p^{th}
202 bivariate quantile curve for the direction ε is defined as:

$$q_{XY}(p,\varepsilon)=\left\{(x,y)\in R^2 : F(x,y)=p\right\} \tag{11}$$

203 with $p\in I$ is the risk and *F(x,y)* is the bivariate cumulative distribution function given
204 by:

$$F(x,y)=\Pr\left\{X\le x, Y\le y\right\} \tag{12}$$

205 which represents the probability of the simultaneous non-exceedance event. Other events
206 can also be considered (see Chebana and Ouarda 2011 for more details).
207 The bivariate quantile in (10) is a curve corresponding to an infinity of combinations (*x*,*y*)
208 that satisfies $F(x,y)=p$. For the event $\left\{X\le x, Y\le y\right\}$, using (2) and (10), the quantile
209 curve can be expressed as follows:

$$q_{XY}(p)=\left\{\begin{array}{l}(x,y)\in R^2 \quad \text{such that} \quad x=F_X^{-1}(u),\\ y=F_Y^{-1}(v);\ u,v\in[0,1]: C(u,v)=p\end{array}\right\} \tag{13}$$

210 The index-flood model used in this paper is based on (12). The resolution of (12), using
211 copula and margin distribution, gives an infinity of combinations (x,y). These
212 combinations constitute the corresponding quantile curve. The main properties of the
213 index-flood model are (see Chebana and Ouarda 2011 for more details):

214 1. The marginal quantiles are special cases of the bivariate quantile curve. Indeed, they
215 correspond to the extreme scenarios of the proper part related to the event;

2. The bivariate quantile curve is composed of two parts: naïve part and proper part. The proper part is the central part whereas the naïve part is composed of two segments starting at the end of each extremity of the proper part;
3. When the risk p increases, the proper part of the bivariate quantile becomes shorter.

3. Multivariate index-flood model in practice

The following procedure is proposed by Chebana and Ouarda (2009) and represents a complete multivariate version of regional FA. Since Chebana and Ouarda (2009) represent a theoretical study, we propose in the present paper a methodology of application of this procedure on a real world case study. The multivariate index-flood model regional estimation requires the delineation of a homogeneous region.

The step of delineation of a homogeneous region is treated by Chebana and Ouarda (2007) in the multivariate case. Based on multivariate L-moments, they proposed statistical tests of multivariate discordancy D and homogeneity H. The practical aspects of these tests are studied in Chebana et al. (2009).

The estimation procedure of the extreme event by the multivariate index-flood model is developed by Chebana and Ouarda (2009). It consists in extending the index-flood model to a multivariate framework using copula and multivariate quantiles. In this step, the homogeneity of the region is assumed. Indeed, non-homogeneous sites must be removed in the first step.

Let N' be the number of sites in the homogeneous region with record length n_i at site i, $i=1,...,N'$. The goal is to estimate, at the target site l, the bivariate and marginal quantiles corresponding to a risk p.

Let (x_{ij}, y_{ij}) for $i=1,...,N'; j=1,...,n_i$, be the data where x and y represent the observations of the considered variables. Let q_p be the regional growth curve which represents a quantile curve common to the whole region.

The complete procedure of determination of the bivariate quantile curve for an ungauged site is described as follows:

1. Identify the homogeneous region to be used in the estimation as follows: to identify and remove discordant sites, apply the multivariate discordancy test D and check the homogeneity of the remaining sites by the homogeneous test H. In practice, it's very difficult to find an exactly homogeneous region. According to Hosking and Wallis (1997), approximate homogeneity is sufficient to apply a regional FA, in the multivariate framework, this procedure was developed by Chebana and Ouarda (2007) and results will be used in this paper.
2. Assess the location parameters μ_{iX} and μ_{iY} $i=1,...,N'$ and standardize the sample (x_{ij},y_{ij}), $j=1,...,n_i$ to be:

$$x'_{ij} = \frac{x_{ij}}{\mu_{iX}}, y'_{ij} = \frac{y_{ij}}{\mu_{iY}} \tag{14}$$

3. Select the bivariate distribution which is composed of a copula and two margins. In this step, our goal is to identify adequate marginal distributions and copula for the whole region to fit the standardized data (x_{ij}', y_{ij}'). This step is described as follows:

a) Collect the data from the homogeneous region to get a sample $(x_k", y_k")$ $k = 1,...,n;\ n = \sum_{i=1}^{N'} n_i$. This sample will be used to select the marginal distributions and copula.

b) Identify the adequate marginal distributions (for X and for Y) using the AIC, BIC and graphical criteria.

c) Select the adequate copula using the graphic test proposed by Genest and Rivest (1993) and the AIC criterion.

4. For each site i, $i=1,\ldots,N'$, estimate the parameters of marginal distributions and copula family selected in step 3. For the copula family, the MPL method is used to estimate the copula parameter. However, for marginal distributions, the estimation method depends on the marginal distribution. Let $\hat{\theta}_k^{(i)}$ be the estimator of the k^{th} parameter from the standardized data of the ith site $k=1,\ldots, s$; s is the number of parameters to be estimated, $i = 1,...,N'$. Obtain the weighted regional parameter estimators:

$$\hat{\theta}_k^r = \frac{\sum_{i=1}^{N'} n_i \hat{\theta}_k^{(i)}}{\sum_{i=1}^{N'} n_i}, \quad k = 1,...,s \tag{15}$$

5. For a given value of risk p, estimate different combinations of the estimated growth curve $\hat{q}_{x,y}(p)$ from (12) using the fitted copula with the corresponding weighted regional parameter $\hat{\theta}_k^{(R)}$with $k=1,\ldots, s$.

6. Estimate the index flood parameter by a multivariate multiple regression model

$$\log(\mu) = E \times \log(A) + \varepsilon \tag{16}$$

where μ is the index flood vector, A is the matrix of watershed physiographic characteristics, E is the matrix of coefficients to estimate and ε is the error. The estimation of index flood can be separated into two steps:

a) Choice of physiographic characteristics: the aim of this step is to select, the optimal set of physiographic characteristics to be considered in the model. Here, the order of characteristics in the selected set is important. The method of multivariate stepwise regression based on the Wilks statistics was used (see e.g. Rencher 2003).

b) Estimation of the coefficients E: the method of maximum likelihood is used (Meng and Rubin 1993).

7. Multiply each growth curve combination with the vector of index flood of the target l: μ_{lX} and μ_{lY}

$$\left(\hat{Q}^{r}_{xy}(p)\right)_{l} = \begin{pmatrix} \mu_{lX} \\ \mu_{lY} \end{pmatrix} \hat{q}_{xy}(p), \quad 0 < p < 1 \tag{17}$$

Hence, the obtained result in (17) is an estimation of the bivariate regional quantile associated to the risk p.

To evaluate the performance of the regional FA models, Hosking and Wallis (1997) suggested an assessment procedure that involves generation of regional average L-moments through a Monte Carlo simulation. This procedure is based on the Jackknife resampling procedure (e.g. Chernick 2012). It consists in considering each site as an ungauged one by removing it temporarily from the region and estimating the bivariate and univariate regional quantiles for various nonexceedance probabilities p in the simulations. This is similar, for instance, to Ouarda et al (2001) in the regional frequency

295 analysis context. At the *mth* repetition, the regional growth curves and the site *i* quantiles
296 are computed.
297 As indicated in Chebana and Ouarda (2009), the performance of the corresponding
298 bivariate regional FA model cannot be evaluated on the basis of the usual performance
299 evaluation criteria. The evaluation is based on the deviation between the regional and
300 local quantile estimated curves. The quantile curve is denoted by (x, $G_p(x)$). The relative
301 error between the regional and local quantile curves is given by:

$$R_p(x) = \frac{G_p^r(x) - G_p^l(x)}{G_p^l(x)} \tag{18}$$

302 where exponents *r* and *l* referring respectively to regional and local quantile curves.
303 This relative difference represents vertical point-wise distances between the two quantile
304 curves. In order to evaluate the estimation error for a site *I,* Chebana and Ouarda (2009)
305 proposed the bias and root-mean-square error respectively given by

$$B_i(p) = \frac{100}{M}\sum_{m=1}^{M} REI_m{}^*(p) \text{ and } R_m(p) = 100\sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(REI_m(p)\right)^2} \tag{19}$$

306 where *M* is the number of simulations, REI* and REI are the two relative integrated error
307 of the simulation *m* defined respectively by

$$REI^*(p) = \frac{1}{L_p}\int_{QC_p} R_p(x)\,dx,\ 0 < p < 1 \tag{20}$$

$$REI(p) = \frac{1}{L_p}\int_{QC_p} \left|R_p(x)\right|dx,\ 0 < p < 1 \tag{21}$$

308 with L_p is the length of the proper part of the true quantile curve QC_p for the risk *p*.
309 To summarize these criteria over the sites of the region, it is possible to average them to
310 obtain the regional bias, the absolute regional bias and the regional quadratic error given
311 respectively by

$$RB(p)=\frac{1}{N'}\sum_{i=1}^{N'}B_i$$
$$ARB(p)=\frac{1}{N'}\sum_{i=1}^{N'}|B_i| \quad (22)$$
$$RRMSE(p)=\frac{1}{N'}\sum_{i=1}^{N'}R_i$$

312 ## 4. Case study

313 The application of the index-flood model in a multivariate regional FA framework
314 concerns a regional data set of interest for the Hydro-Québec Company. The two main
315 flood characteristics, that is, volume V and peak Q are jointly considered. These flood
316 features are random by definition since they are based on the flood starting and ending
317 dates. The latter are obtained using an automatic method which consists in the analysis of
318 cumulative annual hydrographs by adjusting the slopes with a linear approximation (e.g.
319 Ben Aissia et al. 2012). The employed data is used in Chebana et al. (2009). They are
320 from sites in the Côte Nord region in the northern part of the province of Quebec,
321 Canada. The number of sites in the region is N=26 stations with record lengths n_i between
322 14 and 48 years. More information about the data is given in Table 1. Figure 1 presents
323 the geographical location and the correlation coefficient between Q and V for the
324 underlying sites.

325 ### *4.1. Study procedure:*

326 The procedure of the study is composed of the following eight steps:

327 1. Delineate the homogeneous region;
328 2. Assess the location parameters μ_{iV} and μ_{iQ} for $i = 1,\ldots, N'$ given by (13);

3. Select a family of regional multivariate distributions to fit the standardized data of the whole region;
4. For each site in the homogeneous region, estimate the parameters of the marginal distributions and copula family. Estimate the regional parameter estimator $\hat{\theta}_k^{(R)}$ by (14);
5. Estimate different combinations of the estimated growth curve $\hat{q}_{v,q}(p)$ from (12);
6. Estimate the index flood by a multiregression model (15);
7. Using (16), estimate the bivariate regional quantiles associated to the risk p;
8. For each flood characteristic, estimate the univariate regional growth curve and using (8) estimate the univariate regional quantile;
9. Evaluate the performance of the regional models (univariate and bivariate) by Monte Carlo simulation.

4.2. Result and discussion

In this section, results of the application of the adopted procedure are presented. First, results of the multivariate homogeneity study are briefly presented followed by the results of the index-flood regional estimation.

Discordancy and homogeneity

The employed data are the same used in Chebana et al. (2009) and the discordancy and homogeneity results are presented in that reference and in Table 1. Results show that:

- Sites 2 and 16 are discordant for V;
- Site 2 or sites 2 and 3 are discordant for Q;

- Sites 2 and 21 are discordant for (V,Q).

The two sites 2 and 21 are eliminated to allow application of the respective homogeneity test. Table 2 presents the homogeneity test values for the region for V, Q and (V,Q) after removing the two discordant sites (2 and 21). From Table 2, according to the statistic H, we conclude that the region is homogeneous for V, heterogeneous for Q and could be homogeneous for (V,Q).

Identification of marginal distributions

In regional FA, a single frequency distribution is fitted from the whole standardized data. In general, it will be difficult to get a homogeneous region, consequently there will be no single "true" marginal distribution that applies to each site (Hosking and Wallis 1997). Therefore, the aim is to find a marginal distribution that will yield accurate quantile estimates for each site. The scale factor of this marginal distribution changes from one site to another.

Figure 2 shows that the adequate marginal distributions are Gumbel for Q and GEV for V. Results for the appropriate marginal distributions are in agreement with those of similar studies e.g. Cunnane and Nash (1971) and De Michele and Salvadori (2002).

Identification of copula

Table 1 indicates that the dependence between V and Q varies from 0.34 to 0.82 while Figure 1 shows that the dependence variability is scattered over the entire study area. The graphic test based on the K function (5) with the estimate (6) is applied for the three Archimedean copulas: Gumbel, Frank and Clayton. This test leads to fitting the Frank

371 copula to the bivariate data for the studied region. The illustration of this fitting is
372 presented in Figure 3.

373 The AIC and p-value of the S_n goodness-of-fit test described earlier and proposed by
374 Kojadinovic and Yan (2009) are also calculated for the commonly considered copulas in
375 hydrology. However, direct results show that none of the commonly used copulas in
376 hydrology can be accepted. Even though, the graphic test based on the K function
377 indicates excellent fitting with Frank copula, the S_n goodness-of-fit test rejects this
378 copula, as well as the other ones being considered. First, the reason may be that
379 numerical tests tend to be narrowly focused on a particular aspect of the relationship
380 between the empirical copula and the theoretical copula and often try to compress that
381 information into a single descriptive number or test result (see e.g. NIST 2013). Second,
382 the test is widely and successfully applied to at-site hydrological studies which is not the
383 case for regional studies where the total sample size is very large (here n=714). The
384 performance of S_n goodness-of-fit test could be affected when the sample size is large as
385 indicated in Genest et al. (2009). In addition, in terms of application, Vandenberghe et al.
386 (2010) indicated limitation of this test for long sample size like in rainfall. Therefore, to
387 overcome this situation, this test is applied to the data series of each site separately. This
388 is justified since regional FA assumes the same distribution in each site apart from a scale
389 factor (see e.g. Hosking and Wallis 1997, Ouarda et al. 2008). However, according to
390 Hosking and Wallis (1997), it is difficult in practice to have a single distribution which
391 provides a good fit for each site. The goal is hence to find a distribution that will yield
392 accurate quantile estimates for all sites. For the present case-study, results (Table 3) show
393 that Frank is the most accepted copula in the study sites (accepted by the S_n goodness-of-

fit test for 20 sites and sorted best by AIC for 17 sites among 24 sites). Frank copula has already been shown to be adequate to model the dependence between flood V and Q in a number of hydrological studies (see e.g. Grimaldi and Serinaldi 2006). Finally, based on the above arguments (at-site Goodness-of-fit selection, regional graphic test based on the K function, regional and at-site AIC, hydrological literature), the Frank copula is selected for the present case-study. Therefore, the appropriate copula is *Frank* defined by:

$$C_\gamma(u,v) = \frac{1}{\ln\gamma}\ln\left[1+\frac{\left(\gamma^u-1\right)\left(\gamma^v-1\right)}{\gamma-1}\right]; \quad 0\le\gamma; \quad 0<u,v<1 \qquad (23)$$

where γ is the parameter to be estimated. The choice of the adequate copula is in agreement with those of similar studies e.g Lee et al. (2012).

Estimation of parameters associated to margins and copula

Parameters of marginal distributions and copula for each site and their corresponding confidence intervals are presented in Figure 4 while Table 4 showing the regional parameters of the marginal distributions and copula determined by (14). The MPL is employed for the copula parameter. For the Gumbel distribution, μ and σ represent, respectively, the location and scale parameters whereas for the GEV distribution, μ, σ and k represent respectively the location, scale and shape parameters. The ML method is used to estimate the Gumbel parameters while the generalized ML (Martins and Stedinger 2000) is used to estimate the GEV parameters.

414 *Index flood estimation*

415 To estimate the index flood $\hat{\mu}_Q$ of the peak and $\hat{\mu}_V$ of the volume, we use the
416 multiregression model described by (9). The available morphologic and climatic
417 characteristics, used as explicative or input variables in the model are: watershed area in
418 km^2 (*BV*), mean slope of the watershed in % (*BMBV*), percentage of forest in % (*PFOR*),
419 percentage of area covered by lakes in % (*PLAC*), annual mean of total precipitation in
420 mm (*PTMA*), summer mean of liquid precipitation in mm (*PLME*), degree days above
421 zero in degree Celsius (*DJBZ*), absolute value of mean of minimum temperatures in
422 January ($Tmin_{jan}$), February ($Tmin_{feb}$), March ($Tmin_{mar}$) and April ($Tmin_{apr}$), absolute
423 value of mean of maximum temperatures in January ($Tmax_{jan}$), February ($Tmax_{feb}$), March
424 ($Tmax_{mar}$) and April ($Tmax_{apr}$), and mean of cumulative precipitation in January
425 ($PRCP_{jan}$), February ($PRCP_{feb}$), March ($PRCP_{mar}$) and April ($PRCP_{apr}$).

426 The selection of the significant variables to be included in model (9) is based on the
427 stepwise method. Which led to the selection of *BV*, $Tmin_{jan}$, $Tmax_{fev}$ and $PRCP_{fev}$. The
428 model coefficients are estimated by the ML method. Then, the model built is given by:

$$\begin{aligned} \hat{\mu}_Q &= -4.05 \cdot BV^{0.09} \cdot T\min_{jan}{}^{-1.33} \cdot T\max_{feb}{}^{1.04} \cdot PRCP_{feb}{}^{0.79} \\ \hat{\mu}_V &= 6.68 \cdot BV^{1.00} \cdot T\min_{jan}{}^{-3.31} \cdot T\max_{feb}{}^{1.55} \cdot PRCP_{feb}{}^{0.14} \end{aligned} \qquad (1)$$

429 Note that *BV* is already selected in similar studies (e.g. Brath et al. 2001) which is not the
430 case for $Tmin_{jan}$, $Tmax_{feb}$ and $PRCP_{feb}$.

431 Model performance is evaluated by the following criteria: coefficient of determination
432 (R^{2*}), relative root-mean-square error ($RRMSE^*$) and mean relative bias (MRB^*) defined
433 by:

$$R^{2*} = 1 - \frac{\sum_{i=1}^{N'} (\hat{\chi}_i - \chi_i)^2}{\sum_{i=1}^{N'} (\chi_i - \bar{\chi})^2} \quad (2)$$

$$RRMSE^* = 100 \sqrt{\frac{1}{N'-1} \sum_{i=1}^{N'} \left(\frac{\hat{\chi}_i - \chi_i}{\chi_i} \right)^2} \quad (3)$$

$$MRB^* = 100 \frac{1}{N'} \sum_{i=1}^{N'} \left(\frac{\hat{\chi}_i - \chi_i}{\chi_i} \right) \quad (4)$$

with $\hat{\chi}_i$ and χ_i represent the estimated and calculated (mean of observed data in underling site) index flood respectively, and N' is the number of sites.

The criteria R^{2*}, RRMSE* and MRB* are evaluated on the basis of a cross-validation of the model with Jackknife. Results are presented in Table 5. The obtained values of R^2 are higher than 0.95 which shows the high performance of the built model in (9). This performance is confirmed by the low values of R*RMSE* and *MRB* in Table 5.

Bivariate and univariate growth curve estimation

The bivariate regional growth curve is estimated for each risk value p by (12) and by using the regional parameters of the bivariate distribution. On the other hand, univariate regional growth curves of V and Q are estimated directly using regional parameters of marginal distributions. Figure 5 shows the univariate and bivariate estimated growth curves corresponding to nonexceedance probabilities p =0.9, 0.95, 0.99, 0.995 and 0.999 as well as the quantile curve in the unit square and the marginal distributions for Q and V. Univariate regional growth curves of V and Q are also presented in Table 6. Univariate and bivariate quantiles can be assessed by multiplying growth curves by the corresponding index flood (16).

451 *Model performances*

452 As described above, the accuracy of the quantile estimates of the three regional models:
453 univariate of V (V-model), univariate of Q (Q-model) and bivariate of (V,Q) (VQ-model)
454 is assessed using a Monte Carlo simulation procedure. The record lengths of the
455 simulated sites are assumed to be the same as those of observed data and the number of
456 simulations is set to be M=500. To illustrate these results, we present in Figure 6 the
457 univariate and bivariate quantiles of three sites derived from one simulation (M=1) and
458 from the sample data, as well as quantile curves in the unit square and the local and
459 regional marginal distributions of Q and V. Figure 6 shows that, generally, the
460 performance of the two univariate models and the bivariate model decrease with the risk
461 level and depends on the discordancy values. Indeed, for Mistassibi (Figure 6 a) the
462 performance of the V-model is higher than that of the Q-model which is in harmony with
463 the two discordance values of V and Q and with the difference between marginal
464 distributions (local and regional) of Q and V in the side panels. The performance of the
465 bivariate model depends mainly on marginal distributions. Indeed, a small difference in
466 the marginal distribution leads to possible wide shifts in the quantile curve. However, the
467 unit square curves indicate very less effect. Figure 7 illustrates the bivariate quantiles
468 (Regional and the 500 simulations) corresponding to a nonexceedance probability of
469 p=0.9 for the Petit Saguenay station. Figure 7 shows that, in the Petit Saguenay station,
470 the simulated bivariate quantile curves form a surface which includes (but not in the
471 middle) the regional bivariate quantile curve. Table 7 presents the univariate and bivariate
472 model performances of the corresponding nonexceedance probability p = 0.90, 0.95,

473 0.99, 0.995 and 0.999. The univariate and bivariate model performances in each site are
474 presented in Figure 8.

475 Table 7 shows that the V-model performs well, since all performance criteria are less than
476 16% for all values of p. However, the performance of the Q-model is lower compared to
477 that of the V-model where for instance, for p = 0.999, the RRMSE is larger than 21%.
478 This conclusion can also be drawn from Figure 8 where the performance criteria of the Q-
479 model are clearly higher than those of the V-model for all values of p. This conclusion
480 can be explained by the fact that the region is heterogeneous for Q. On the other hand, the
481 performance of the VQ-model is, generally, somewhat lower than the Q-model. This
482 conclusion is confirmed by Figure 8 where we see a close performance criteria for the
483 VQ-model and Q-model. One can explain this by the fact that the univariate quantiles are
484 special cases of bivariate quantiles, since they correspond to the extreme scenario of the
485 proper part related to the event. Then the performance of the univariate models has an
486 effect on the performance of the bivariate model. Since the performance criteria of the Q-
487 model are higher than those of the V-model then effects of the Q-model performance on
488 the QV-model is more important than the effects of the V-model performance. On the
489 other hand, from Figure 8 we observe that the performance behaviour criteria of the VQ-
490 model and Q-model are similar to those of Gumbel parameters (Figure 4 a), especially for
491 the scale parameter (σ). Consequently, a variation of the Gumbel parameters has an effect
492 on the Q-model performance and therefore an effect on the VQ-model performance.

493 Performance criteria corresponding to the VQ-model are less than 19% for the highest
494 considered risk level p = 0.999 (Table 7). Values of these performance criteria are larger
495 than those obtained by Chebana and Ouarda (2009). Indeed, unlike their simulation study,

496 the performance of the bivariate model is affected by the error of the index flood
497 estimation as well as parameter estimations. Generally the performance criteria increase
498 with the value of the risk p (Table 7 and Figure 8). An exception is recorded between
499 p=0.995 and p=0.999 where performance criteria of the VQ-model are higher for
500 p=0.995. This finding can be explained by the curse of dimensionality in the multivariate
501 context, where the central part of a distribution contains little probability mass compared
502 to the univariate framework (for more details see Scott 1992, Chebana and Ouarda 2009).

503 In order to further explain the results, we plot in Figure 9 the RRMSE of each model (for
504 p=0.99) with respect to the corresponding discordancy values. Ideally we should find an
505 increasing relation between the RRMSE of each model and the corresponding
506 discordance. This relation is observed only for the V-model (Figure 9a) since the studied
507 region is homogeneous for V, heterogeneous for Q and could be homogeneous for (V,Q).
508 To find out other factors that have an impact on the model performance, we present in
509 Figure 10 the RRMSE of the VQ-model (for p=0.99) with respect to watershed area and
510 the correlation between V and Q. Figure 10a shows that high RRMSE values are seen for
511 small watersheds whereas Figure 10b shows that sites with $\rho(V,Q)> 0.6$ have a good
512 performance (RRMSE of the order of 10%) with the exception of Godbout (site number
513 15) which has ρ=0.75 and high RRMSE. Godbout is one of the four sites that have a high
514 value of the Gumbel scale parameter and a high RRMSE of the Q-model and the VQ-
515 model.

516 The quantile curve, for a given risk p, leads to infinite combinations of (Q,V) associated
517 to the same return period. However, they could be not equal in practice or in practical
518 point of view (Chebana and Ouarda 2011). Recently, Volpi and Fiori (2012) proposed a

519 methodology to identify a subset of the quantile curve according to a fixed probability
520 percentage of the events, on the basis of their probability of occurrence; see Volpi and
521 Fiori (2012) for more details. As an illustrative example, the Chamouchouane station is
522 considered. Figure 11 presents the curves and the limits with probability (1-α)=0.95.

523 ## 5. Conclusions and perspectives

524 The procedure for regional FA in a multivariate framework is applied to a set of sites
525 from the Côte-Nord region in the northern part of the province of Quebec, Canada. This
526 procedure is proposed by Chebana and Ouarda (2009) and represents a multivariate
527 version of the index-flood model. It is based on copulas and multivariate quantiles.
528 Chebana and Ouarda (2009) evaluated the proposed model based on a simulation study.
529 In the present paper, practical aspects of this model are presented and investigated jointly
530 for the flood peak and volume of the considered data set.

531 Results show that the appropriate fitted marginal distributions are Gumbel for Q and
532 GEV for V as well as the Frank copula for their dependence structure. The multi-
533 regressive proposed method to estimate the index flood is shown to lead to a high
534 performance. The performance of the two univariate models is in accordance with the
535 quality of the region (homogeneity test). Indeed, the studied region is homogenous for V
536 and heterogeneous for Q where the performance of the V-model is higher than that of the
537 Q-model. The high performance of the V-model is confirmed by a relation between their
538 performance criteria and the discordance values of V in each site whereas the low
539 performance of the Q-model is mainly caused by the variation of the marginal
540 distribution parameters. This is a logical consequence of the heterogeneity of the region

for Q. The performance of the two univariate models increases with the risk level p. For the bivariate model, the performance criteria are less than 19% which indicates the high performance of the proposed procedure to estimate bivariate quantiles at ungauged sites. This performance increases, generally, with the risk level p and is affected by the performance of the Q-model. Results show also that high values of the performance criteria of the bivariate regional model are seen for small watershed and for sites with low correlation between V and Q. From this study it is concluded that a good performance of the bivariate model requires good performance of the two univariate models. This means that we should have a homogeneous region for both univariate variables.

The considered method estimates the bivariate quantile as combinations that constitute the quantile curve for a given risk level p. A method to select the appropriate combination(s) for a specific application is of interest and should be developed in future efforts. Furthermore, the adaptation of the model to the estimation of other hydrological phenomena such as drought and the consideration of others homogenous regions can be conducted by considering the appropriate distributions, copulas and events to be studied.

ACKNOWLEDGEMENTS

The authors thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and Hydro-Québec for the financial support.

563

564 # REFERENCES

565

566 Akaike, H., 1973. "Information measures and model selection." *Information measures*
567 *and model selection*, 50: 277-290.
568 Alila, Y., 1999. "A hierarchical approach for the regionalization of precipitation annual
569 maxima in Canada." *Journal of Geophysical Research D: Atmospheres*,
570 104(D24): 31645-31655.
571 Alila, Y., 2000. "Regional rainfall depth-duration-frequency equations for Canada."
572 *Water Resources Research*, 36(7): 1767-1778.
573 Ashkar, F., 1980. Partial duration series models for flood analysis. Montreal, Qc, Canda,
574 Ecole Polytech of Montreal.
575 Ben Aissia, M. A., F. Chebana, T. B. M. J. Ouarda, L. Roy, G. Desrochers, I. Chartier
576 and É. Robichaud, 2012. "Multivariate analysis of flood characteristics in a
577 climate change context of the watershed of the Baskatong reservoir, Province of
578 Québec, Canada." *Hydrological Processes*, 26(1): 130-142.
579 Besag, J., 1975. "Statistical Analysis of Non-Lattice Data." *Journal of the Royal*
580 *Statistical Society. Series D (The Statistician)*, 24(3): 179-195.
581 Brath, A., A. Castellarin, M. Franchini and G. Galeati, 2001. "Estimating the index flood
582 using indirect methods." *Estimation de l'indice de crue par des méthods*
583 *indirectes*, 46(3): 399-418.
584 Burn, D. H., 1990. "Evaluation of regional flood frequency analysis with a region of
585 influence approach." *Water Resources Research*, 26(10): 2257-2265.
586 Charpentier, A., Fermanian, J.-D., Scaillet, O., 2007. The estimation of copulas: Theory
587 and practice. New York, *Risk Books*.
588 Chebana, F. and T. B. M. J. Ouarda, 2007. "Multivariate L-moment homogeneity test."
589 *Water Resources Research*, 43.
590 Chebana, F. and T. B. M. J. Ouarda, 2009. "Index flood–based multivariate regional
591 frequency analysis." *Water Resources Research*, 45.
592 Chebana, F. and T. B. M. J. Ouarda, 2011. "Multivariate quantiles in hydrological
593 frequency analysis." *Environmetrics*, 22(1): 63-78.
594 Chebana, F., T. B. M. J. Ouarda, P. Bruneau, M. Barbet, S. E. Adlouni and M.
595 Latraverse, 2009. "Multivariate homogeneity testing in a northern case study in
596 the province of Quebec, Canada." *Hydrological Processes*, 23(12): 1690-1700.
597 Chernick, M. R., 2012. "The jackknife: A resampling method with connections to the
598 bootstrap." *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(2): 224-
599 226.
600 Cunnane, C. and J. E. Nash, 1971. "Bayesian estimation of frequency of hydrological
601 events." *Internationa Association of Hydrological Sciences Publications*, 100: 47-
602 55.
603 Dalrymple, T., 1960. Flood-frequency analyses. Washington, D.C., *U.S. G.P.O.*

De Michele, C. and G. Salvadori, 2002. "On the derived flood frequency distribution: analytical formulation and the influence of antecedent soil moisture condition." *Journal of Hydrology*, 262(1-4): 245-258.

De Michele, C., G. Salvadori, M. Canossi, A. Petaccia and R. Rosso, 2005. "Bivariate statistical approach to check adequacy of dam spillway." *Journal of Hydrologic Engineering*, 10(1): 50-57.

Demarta, S. and A. J. McNeil, 2005. "The t copula and related copulas." *International Statistical Review*, 73(1): 111-129.

Durrans, S. R. and S. Tomic, 1996. "Regionalization of low-flow frequency estimates: an Alabama case study." *Journal of the American Water Resources Association*, 32(1): 23-37.

Genest, C., K. Ghoudi and L. Rivest, 1995. "A semiparametric estimation procedure of dependence parameters in multivariate families of distributions." *Biometrika*, 82(3): 543-552.

Genest, C., B. Rémillard and D. Beaudoin, 2009. "Goodness-of-fit tests for copulas: A review and a power study." *Insurance: Mathematics and Economics*, 44(2): 199-213.

Genest, C. and L.-P. Rivest, 1993. "Statistical Inference Procedures for Bivariate Archimedean Copulas." *Journal of the American Statistical Association*, 88(423): 1034-1043.

GREHYS, 1996a. "Presentation and review of some methods for regional flood frequency analysis." *Journal of Hydrology*, 186(1-4): 63-84.

GREHYS, 1996b. "Inter-comparison of regional flood frequency procedures for Canadian rivers." *Journal of Hydrology*, 186(1-4): 85-103.

Grimaldi, S. and F. Serinaldi, 2006. "Asymmetric copula in multivariate flood frequency analysis." *Advances in Water Resources*, 29(8): 1155-1167.

Hamza, A., T. B. M. J. Ouarda, S. R. Durrans and B. Bobée, 2001. "Development of scale invariance and tail models for the regional estimation of low-flows." *Canadian Journal of Civil Engineering*, 28(2): 291-304.

Hosking, J. R. M. and J. R. Wallis, 1993. "Some statistics useful in regional frequency analysis." *Water Resources Research*, 29(2): 271-281.

Hosking, J. R. M. and J. R. Wallis, 1997. Regional frequency analysis : an approach based on L-moments. Cambridge, *Cambridge University Press*.

Kim, G., M. J. Silvapulle and P. Silvapulle, 2007. "Comparison of semiparametric and parametric methods for estimating copulas." *Computational Statistics and Data Analysis*, 51(6): 2836-2850.

Kim, J.-M., Y.-S. Jung, E. Sungur, K.-H. Han, C. Park and I. Sohn, 2008. "A copula method for modeling directional dependence of genes." *BMC Bioinformatics*, 9(1): 225.

Kojadinovic, I. and J. Yan, 2009. "A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems." *Statistics and Computing*: 1-14.

Lee, T., R. Modarres and T. B. M. J. Ouarda, 2012. "Data based analysis of bivariate copula tail dependence for drought duration and severity." *Hydrological Processes*: in press.

649 Martins, E. S. and J. R. Stedinger, 2000. "Generalized maximum-likelihood generalized
650 extreme-value quantile estimators for hydrologic data." *Water Resources*
651 *Research*, 36(3): 737-744.
652 Meng, X. L. and D. B. Rubin, 1993. "Maximum likelihood estimation via the ecm
653 algorithm: A general framework." *Biometrika*, 80(2): 267-278.
654 Nelsen, R. B., 2006. An introduction to copulas. New York, *Springer*.
655 Nezhad, M. K., K. Chokmani, T. B. M. J. Ouarda, M. Barbet and P. Bruneau, 2010.
656 "Regional flood frequency analysis using residual kriging in physiographical
657 space." *Hydrological Processes*, 24(15): 2045-2055.
658 Nguyen, V.-T.-V. and G. Pandey, 1996. A new approach to regional estimation of floods
659 in Quebec. Proceedings of the 49th Annual Conference of the CWRA, Quebec
660 City, *Collection Environnement de l'Université de Montréal*.
661 NIST, 2013. e-Handbook of Statitical Methods.
662 http://www.itl.nist.gov/div898/handbook/pmd/section4/pmd44.htm, [accessed
663 september 2013].
664 Ouarda, T. B. M. J., K. M. Bâ, C. Diaz-Delgado, A. Cârsteanu, K. Chokmani, H. Gingras,
665 E. Quentin, E. Trujillo and B. Bobée, 2008. "Intercomparison of regional flood
666 frequency estimation methods at ungauged sites for a Mexican case study."
667 *Journal of Hydrology*, 348(1–2): 40-58.
668 Ouarda, T. B. M. J., J. M. Cunderlik, A. St-Hilaire, M. Barbet, P. Bruneau and B. Bobée,
669 2006. "Data-based comparison of seasonality-based regional flood frequency
670 methods." *Journal of Hydrology*, 330(1-2): 329-339.
671 Ouarda, T. B. M. J., C. Girard, G. S. Cavadias and B. Bobée, 2001. "Regional flood
672 frequency estimation with canonical correlation analysis." *Journal of Hydrology*,
673 254(1-4): 157-173.
674 Ouarda, T. B. M. J., M. Hache, P. Bruneau and B. Bobee, 2000. "Regional flood peak and
675 volume estimation in northern Canadian basin." *Journal of Cold Regions*
676 *Engineering*, 14(4): 176-191.
677 Pilon, P. J., 1990. The Weibull distrubution applied to regional low-flow frequency
678 analysis. Regionalization in Hydrology, Ljubljana, *IAHS Publ.*
679 Rencher, A. C., 2003. Methods of Multivariate Analysis, *John Wiley & Sons, Inc.*
680 Salvadori, G., N. Carlo De Michele, T. Kottegoda and R. Rosso, 2007. Extremes in
681 Nature: An Approach Using Copula. Dordrecht, *Springer, p292.*
682 Schwarz, G., 1978. "Estimating the Dimension of a Model." *Annals of Statistics*, 6(2):
683 461-464.
684 Scott, D. W., 1992. Multivariate density estimation : theory, practice, and visualization.
685 New York, *Wiley, p265.*
686 Shiau, J. T., 2003. "Return period of bivariate distributed extreme hydrological events."
687 *Stochastic Environmental Research and Risk Assessment*, 17(1-2): 42-57.
688 Shih, J. H. and T. A. Louis, 1995. "Inferences on the association parameter in copula
689 models for bivariate survival data." *Biometrics*, 51(4): 1384-1399.
690 Singh, V. P., 1987. Regional flood frequency analysis. Dordrecht, *D. Reidel Publishing,*
691 *p419.*
692 Sklar, A., 1959. "Fonction de répartition à n dimensions et leurs margins." *Institut*
693 *Statistique de l'Université de Paris*, 8: 229-231.

694 Stedinger, J. R. and G. D. Tasker, 1986. "Regional Hydrologic Analysis, 2, Model-Error
695 Estimators, Estimation of Sigma and Log-Pearson Type 3 Distributions." *Water*
696 *Resources Research*, 22(10): 1487-1499.
697 Vandenberghe, S., N. E. C. Verhoest and B. De Baets, 2010. "Fitting bivariate copulas to
698 the dependence structure between storm characteristics: A detailed analysis based
699 on 105 year 10 min rainfall." *Water Resources Research*, 46(1): W01512.
700 Volpi, E. and A. Fiori, 2012. "Design event selectionin bivariate hydrological frequency
701 analysis." *Hydrological science journal*, 57(8): 1-10.
702 Wiltshire, S. E. (1987). "Statistical techniques for regional flood-frequency analysis.
703 [electronic resource]."
704 Yue, S., 2001. "A bivariate extreme value distribution applied to flood frequency
705 analysis." *Nordic Hydrology*, 32(1): 49-64.
706 Yue, S., T. B. M. J. Ouarda, B. Bobée, P. Legendre and P. Bruneau, 1999. "The Gumbel
707 mixed model for flood frequency analysis." *Journal of Hydrology*, 226(1-2): 88-
708 100.
709 Zhang, L. and V. P. Singh, 2006. "Bivariate flood frequency analysis using the copula
710 method." *Journal of Hydrologic Engineering*, 11(2): 150-164.
711 Zhang, L. and V. P. Singh, 2007. "Trivariate Flood Frequency Analysis Using the
712 Gumbel-Hougaard Copula." *Journal of Hydrologic Engineering*, 12(4): 431-439.
713
714

715 **Tables and Figures**

716 # Tables

717 **Table 1: Discordancy statistic for each site (Chebana et al. 2009).**

#	Site name	BV (Km^2)	n_i	(V,Q) correlation coefficient	Discordancy statistic		
					V	Q	(V,Q)
1	Petit Saguenay	729	24	0.50	0.80	0.40	1.09
2	Des Ha Ha	564	19	0.73	3.60	4.44	3.88
3	Aux Écorces	1120	34	0.5	0.16	2.42	0.69
4	Pikauba	489	34	0.34	0.89	1.16	1.22
5	Métabetchouane	2270	30	0.54	1.22	1.23	1.59
6	Petite Péribonka	1090	31	0.62	0.26	0.45	0.98
7	Chamouchouane (Ashuapmushuan)	15 300	43	0.70	0.13	0.14	0.26
8	Mistassibi	8690	39	0.52	0.32	0.78	0.88
9	Mistassini	9620	43	0.52	0.62	0.19	0.53
10	Manouane	3720	23	0.39	0.55	0.47	2.38
11	Valin	740	31	0.42	0.40	0.46	2.37
12	Ste-Marguerite	1100	21	0.48	1.50	0.55	1.30
13	DesEscoumins	779	19	0.49	1.14	1.81	1.27
14	Portneuf	2580	20	0.80	0.99	0.32	1.06
15	Godbout	1570	30	0.75	0.91	0.89	1.29
16	Aux-Pékans	3390	16	0.54	3.19	0.38	2.25
17	Tonerre	674	48	0.64	0.51	1.65	2.25
18	Magpie	7200	27	0.66	0.12	1.23	1.11
19	Romaine	13 000	48	0.68	1.62	0.48	0.57
20	Nabisipi	2060	25	0.78	1.12	0.64	0.54
21	Aguanus	5590	19	0.60	1.53	0.84	3.07
22	Natashquan	15 600	39	0.75	0.28	0.39	1.02
23	Etamamiou	2950	19	0.82	1.06	1.33	1.32
24	St Augustin	5750	14	0.73	0.62	0.67	0.92
25	St Paul	6630	25	0.73	0.31	1.35	1.11
26	Moisie	19000	39	0.65	1.16	0.32	0.54

718 **Table 2 : Homogeneity after exclusion of the discordant sites**

	V	*Q*	*(V,Q)*
H	0.7052	2.4081	1.5234

719

720 **Table 3 : Results of *Sn* Goodness-of-fit test and AIC criterion for considered**
721 **copulas. Gray color indicates that Frank copula is accepted by *Sn* goodness-of-fit**
722 **test (p-value column) and has the smallest AIC (AIC column) for the corresponding**
723 **site.**

Site	Gumbel		Frank		Clayton		Galambos		Husler-Reiss		Placket	
	P-value	AIC	P-value	AIC	P-value	AIC	P-value	AIC	P-value	AIC	P-value	AIC
1	0.190	-75.0	0.078	-97.5	0.012	-55.2	0.237	-74.5	0.318	-74.1	0.034	-66.6
3	0.086	-106.2	0.081	-151.7	0.175	-72.2	0.095	-105.4	0.117	-104.8	0.086	-86.5
4	0.173	-98.6	0.154	-200.9	0.460	-70.8	0.183	-97.4	0.194	-96.6	0.191	-84.5
5	0.150	-114.0	0.137	-177.7	0.083	-82.3	0.169	-113.2	0.187	-112.4	0.068	-102.6
6	0.128	-146.8	0.133	-103.9	0.064	-102.3	0.030	-146.5	0.028	-145.9	0.044	-131.7
7	0.152	-210.9	0.054	-202.4	0.003	-140.0	0.159	-210.2	0.202	-209.1	0.020	-182.3
8	0.135	-142.9	0.207	-181.7	0.120	-95.0	0.148	-142.1	0.175	-141.3	0.135	-116.7
9	0.148	-175.2	0.404	-231.1	0.041	-115.7	0.155	-174.2	0.197	-173.2	0.214	-145.1
10	0.459	-48.0	0.231	-96.7	0.104	-36.6	0.480	-47.4	0.522	-47.1	0.218	-42.0
11	0.002	-101.2	0.017	-172.8	0.242	-71.4	0.002	-100.3	0.002	-99.5	0.017	-86.5
12	0.120	-66.7	0.113	-59.9	0.200	-48.7	0.114	-66.4	0.112	-66.2	0.180	-58.6
13	0.016	-71.4	0.037	-101.0	0.079	-56.2	0.016	-70.9	0.016	-70.4	0.036	-69.6
14	0.027	-101.0	0.058	24.1	0.019	-78.1	0.026	-101.1	0.029	-101.0	0.020	-98.7
15	0.120	-160.8	0.041	-164.2	0.048	-118.1	0.118	-160.3	0.130	-159.2	0.066	-153.5
16	0.243	-43.0	0.124	-56.0	0.227	-33.1	0.253	-42.7	0.249	-42.5	0.199	-39.1
17	0.048	-164.2	0.003	-230.3	0.000	-113.1	0.059	-163.2	0.069	-162.1	0.003	-143.3
18	0.092	-123.2	0.112	-105.3	0.255	-88.9	0.081	-122.7	0.069	-122.1	0.135	-113.4
19	0.214	-236.0	0.177	-180.7	0.150	-149.6	0.208	-235.3	0.222	-234.4	0.192	-193.7
20	0.352	-122.2	0.326	28.1	0.059	-90.4	0.366	-122.1	0.345	-121.9	0.321	-116.6
22	0.241	-193.4	0.215	-199.3	0.006	-131.4	0.254	-192.7	0.302	-191.6	0.022	-171.5
23	0.002	-100.3	0.011	-184.9	0.040	-85.4	0.002	-100.5	0.007	-101.2	0.002	-104.2
24	0.173	-64.5	0.091	29.3	0.002	-55.0	0.179	-64.5	0.206	-64.6	0.052	-64.8
25	0.138	-107.0	0.310	-111.4	0.031	-78.5	0.140	-106.6	0.140	-106.0	0.035	-99.4
26	0.168	-182.9	0.280	-192.1	0.041	-123.6	0.158	-182.2	0.193	-181.2	0.127	-159.7
Pooled data	0.0005	-329.35	0.045	-365.65	0.0005	-318.94	0.0005	-327.34	0.0005	-300.11	0.04785	-364.86

724

725

726 **Table 4: Regional parameters of marginal distributions and copula**

Marginal distribution					Copule
Peak (Gumbel)		Volume (GEV)			Frank
μ_r	σ_r	k_r	σ_r	μ_r	γ_r
1.16	0.33	0.16	0.28	0.88	2.06

727

728 **Table 5 : Performance criteria of multiregression index flood model**

	R^{2*}	$MRB^{*}(\%)$	$RRMSE^{*}(\%)$
Q	0.94	1.24	16.75
V	0.97	0.70	11.68

729

730 **Table 6 : Univariate regional growth curve values**

Marginal distribution	p				
	0.9	0.95	0.99	0.995	0.999
Volume (GEV)	1.40	1.53	1.77	1.85	2.02
Peak (Gumbel)	1.90	2.13	2.67	2.90	3.43

731

732

733 **Table 7 : Performance of the univariate and bivariate quantiles corresponding to**
734 **the nonexceedance probabilities 0.9, 0.95, 0.99, 0.995 and 0.999.**

735

Risk	Criterion	$(V.Q)$	V	Q
p=0.9	RB	1.99	-1.25	-0.94
	ARB	9.60	3.44	11.87
	RRMSE	15.25	7.68	13.88
p=0.95	RB	3.27	-1.34	-1.05
	ARB	11.25	4.32	13.56
	RRMSE	17.34	9.37	15.83
p=0.99	RB	2.49	-0.78	-1.26
	ARB	12.03	5.95	15.99
	RRMSE	17.93	13.56	18.79
p=0.995	RB	3.41	-0.20	-1.73
	ARB	12.87	7.25	16.93
	RRMSE	19.06	14.90	19.77
p=0.999	RB	3.23	0.56	-1.51
	ARB	12.21	7.89	18.95
	RRMSE	18.21	15.79	21,78

736

737

Figures

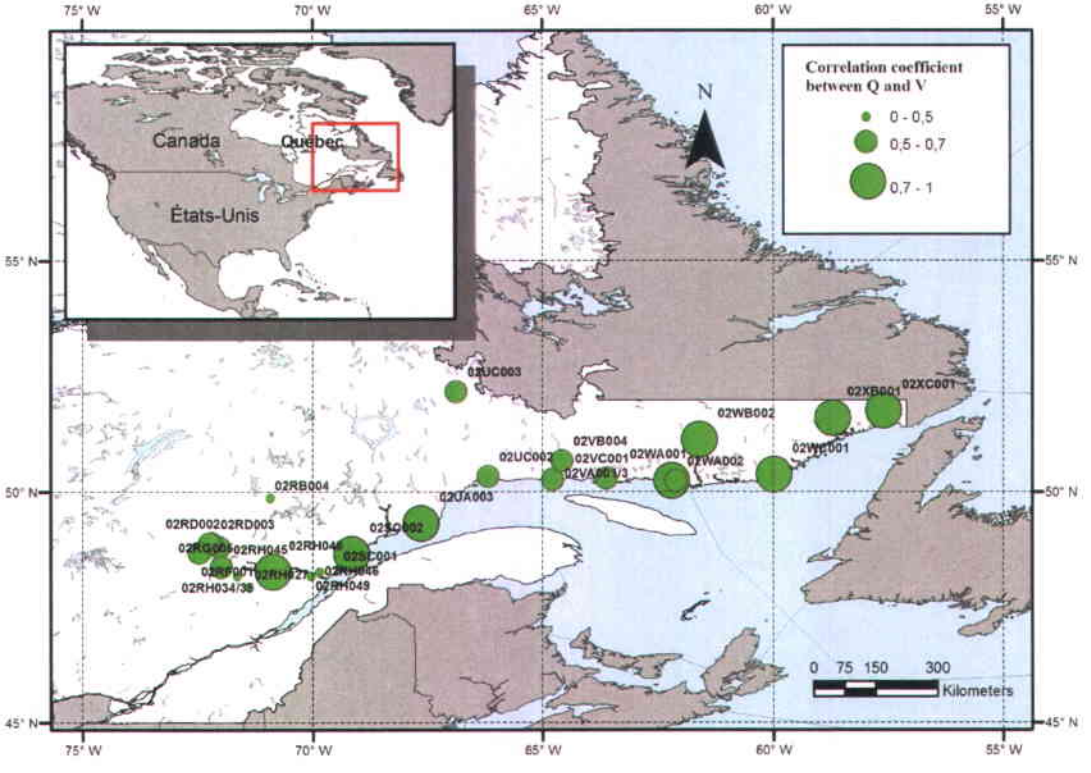
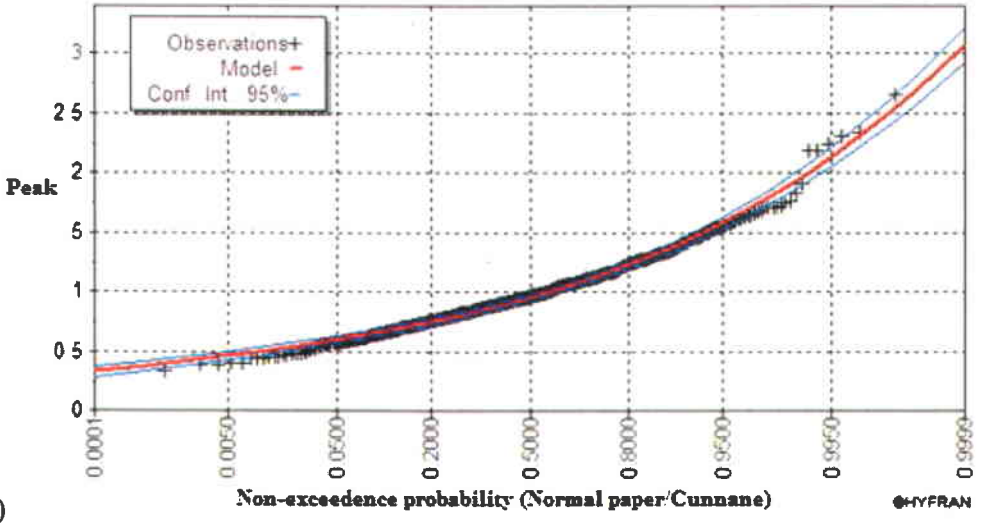
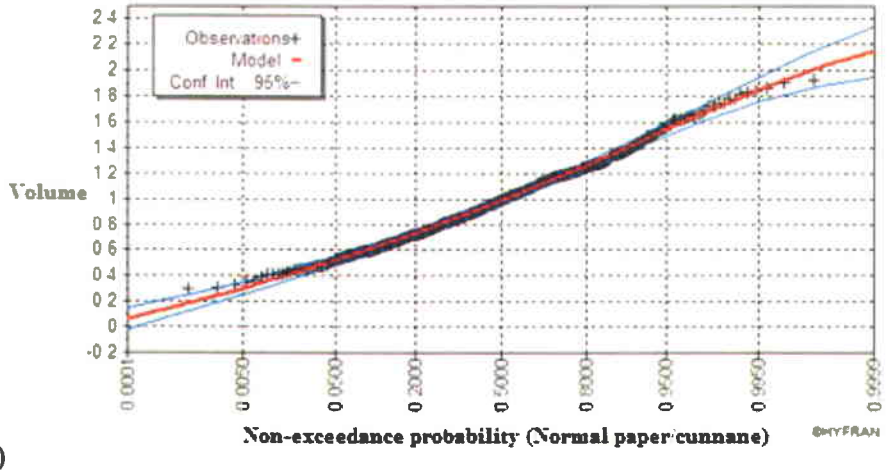


Figure 1 : Geographical chart of the location of the sites

744



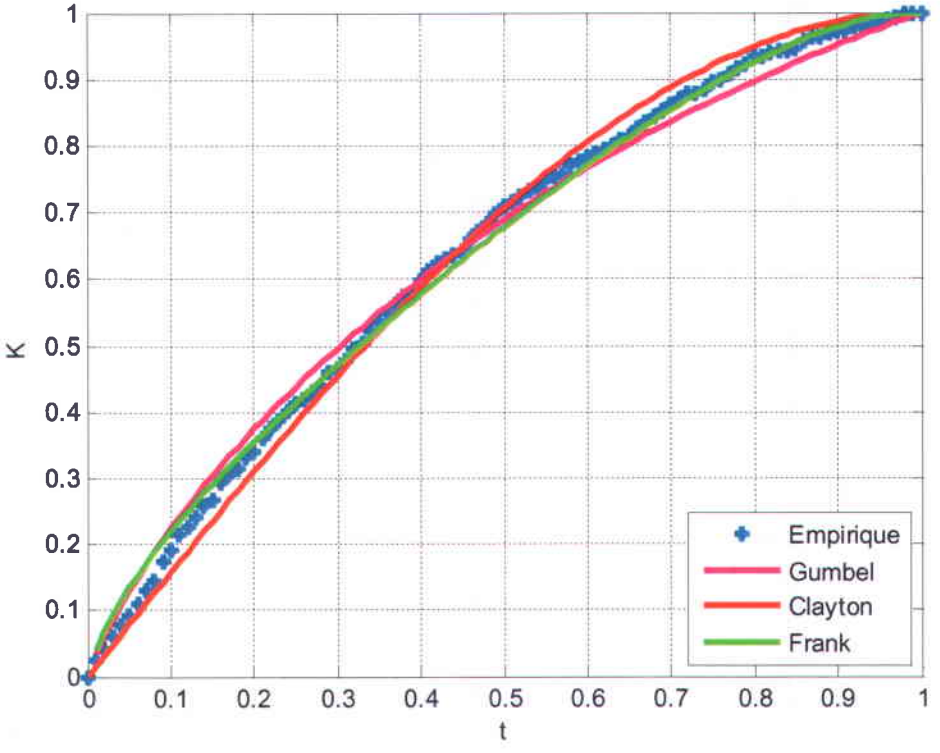
745 a)



746 b)

747 **Figure 2 : Fitting of the marginal distribution of a) *Q* and b) *V*.**

748

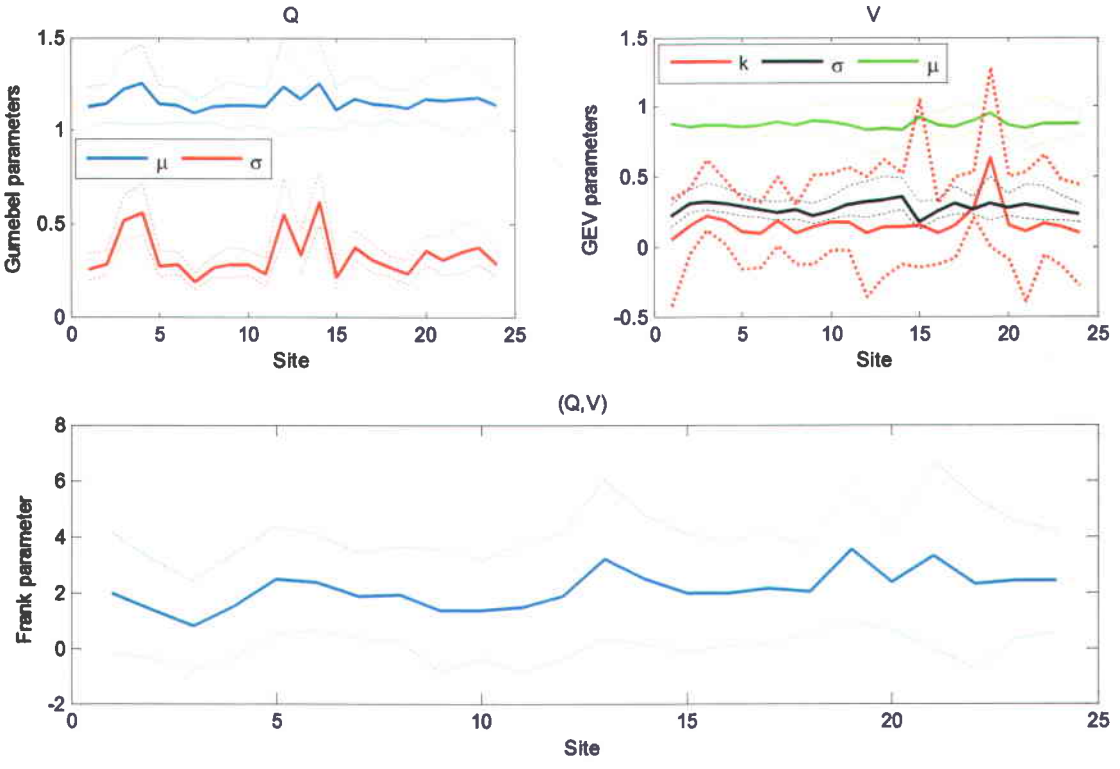


749

750 **Figure 3 : Copula fitting using K-function**

751

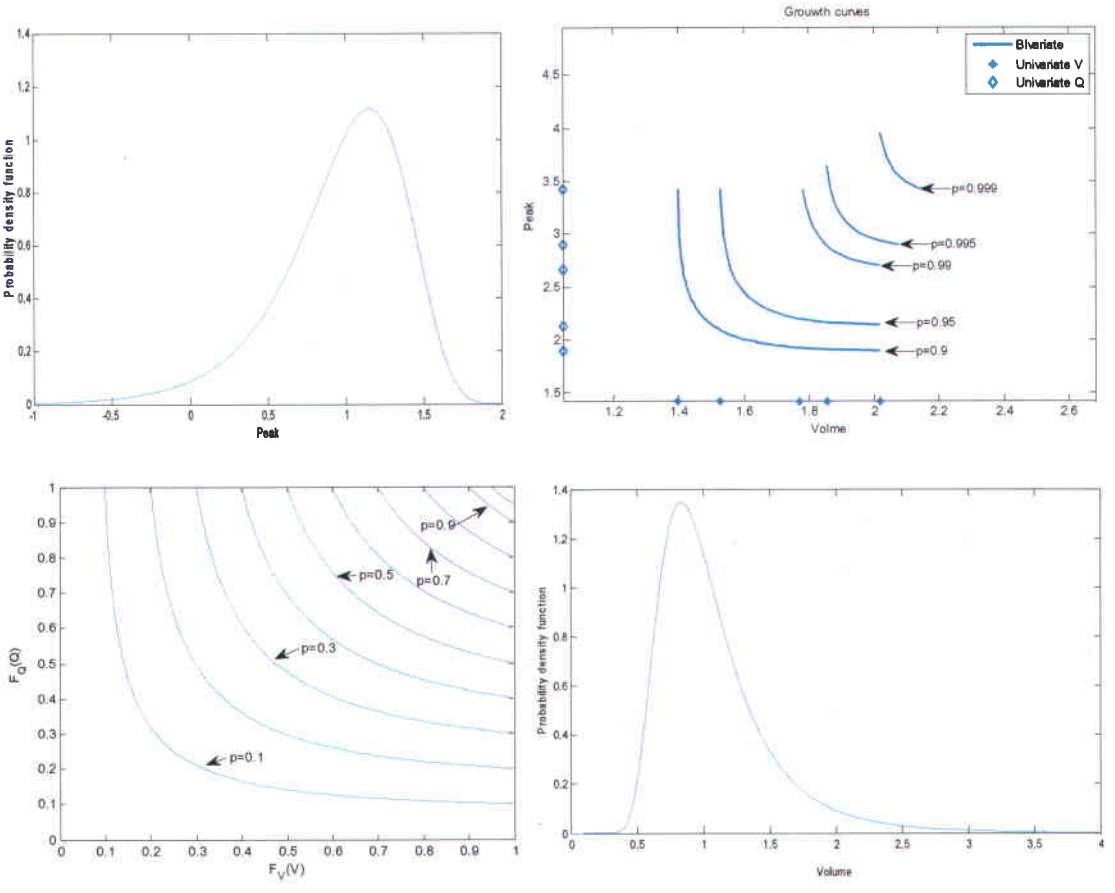
752



753

754 **Figure 4: Parameters of marginal distributions and copula. Dashed lines indicate**
755 **the confidence interval corresponding to each parameter**

756

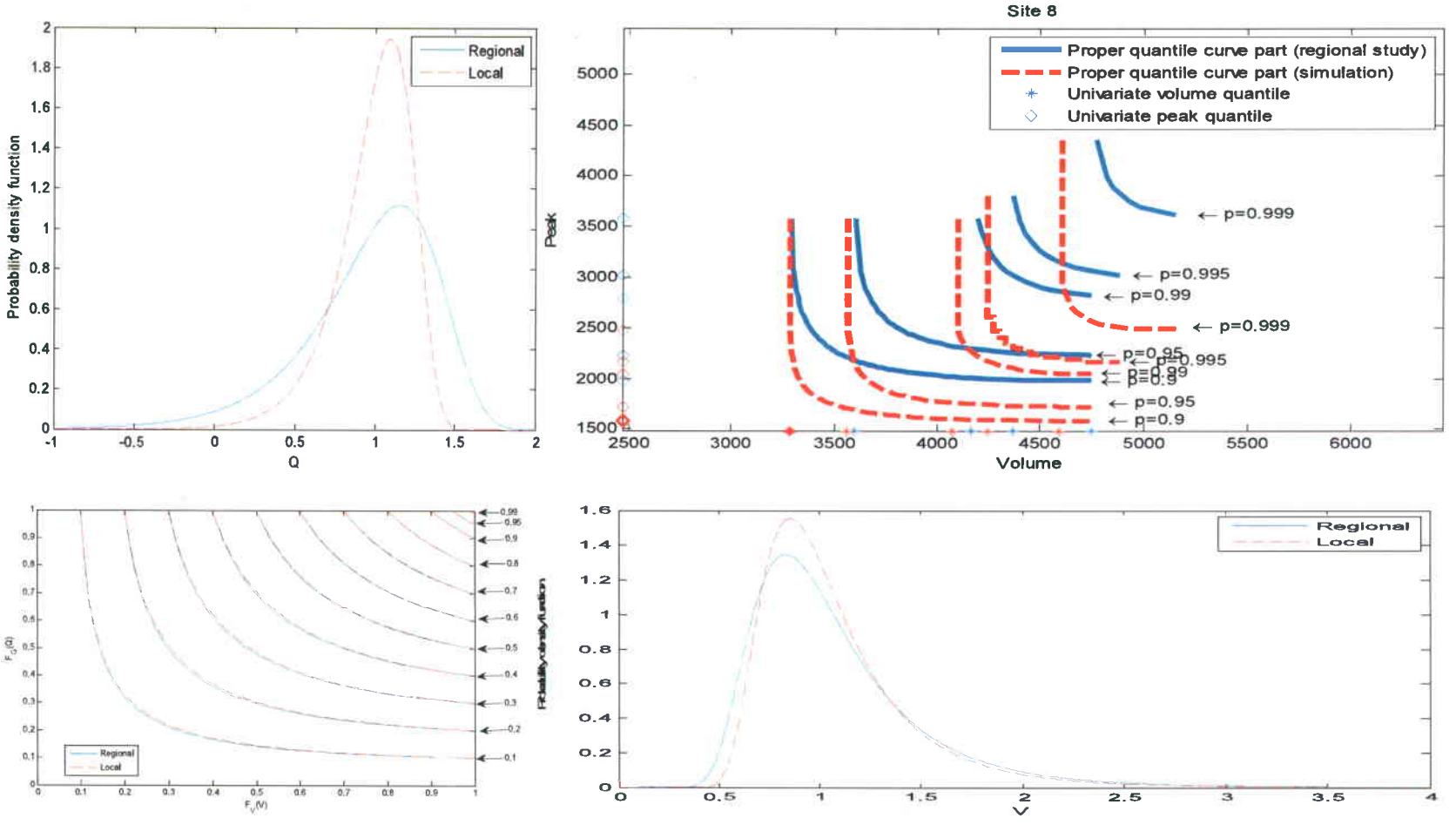


757

758

759 **Figure 5 : Estimated regional bivariate and univariate growth curves, quantile**
760 **curve in the unit square and the marginal distributions for Q and V**

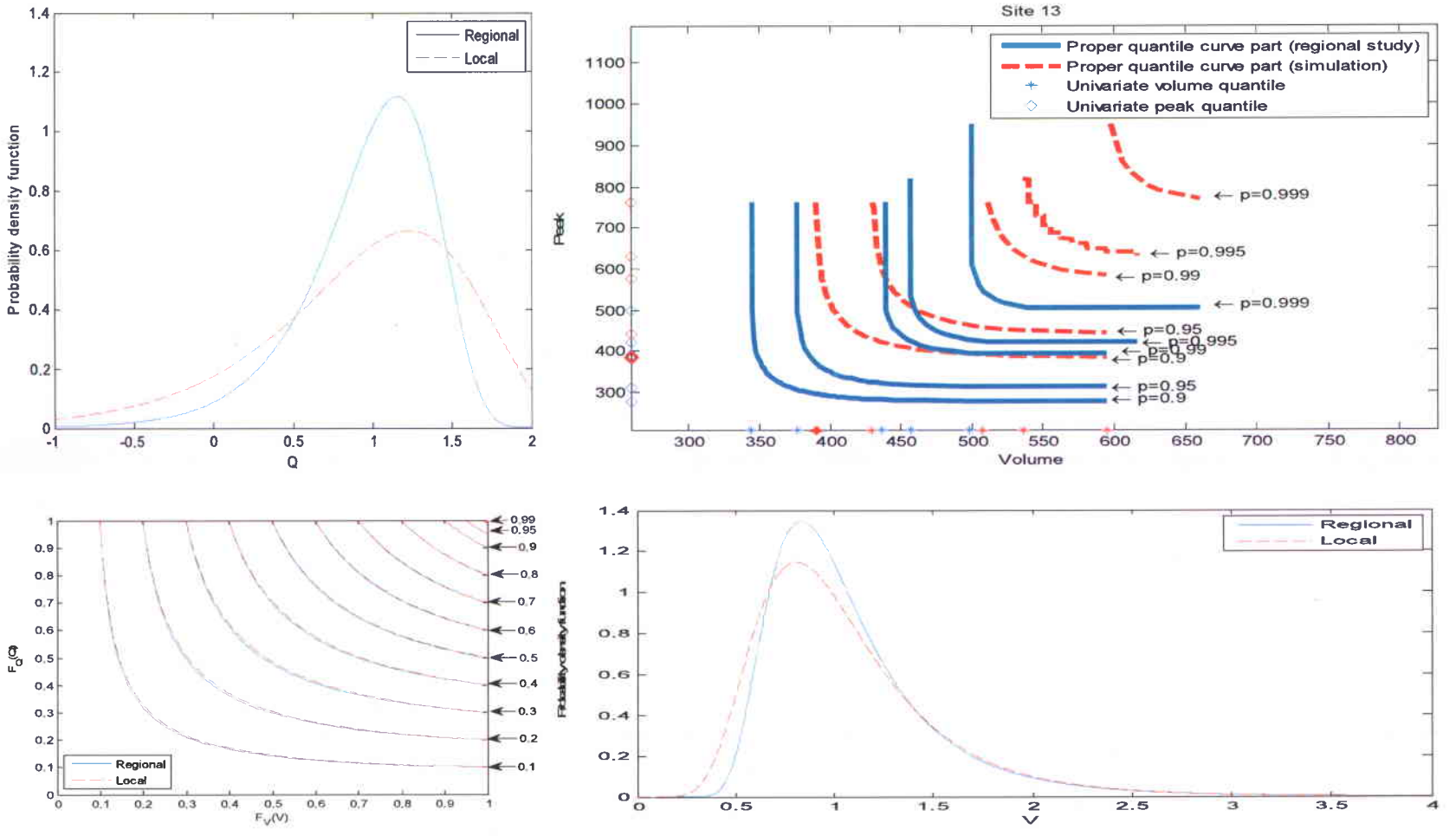
761



762

763 **Figure 6a : Univariate and bivariate quantiles corresponding to a nonexceedance probability p=0.9, 0.95, 0.99, 0.995 and 0.999**
764 **in Mistassibi, quantile curve in the unit square and side panels showing the marginal distributions (local and regional) of Q**
765 **and V**

766

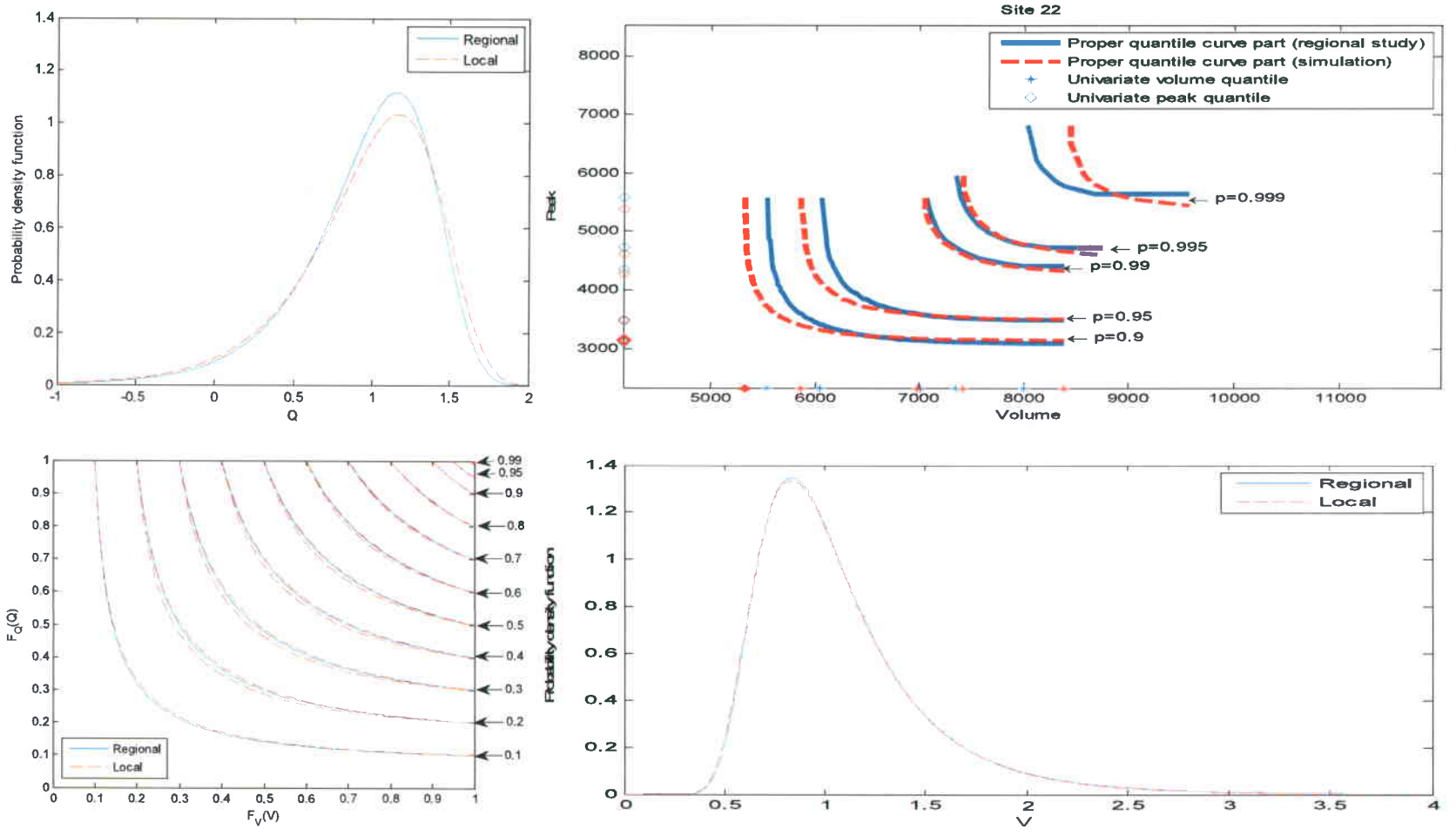


767

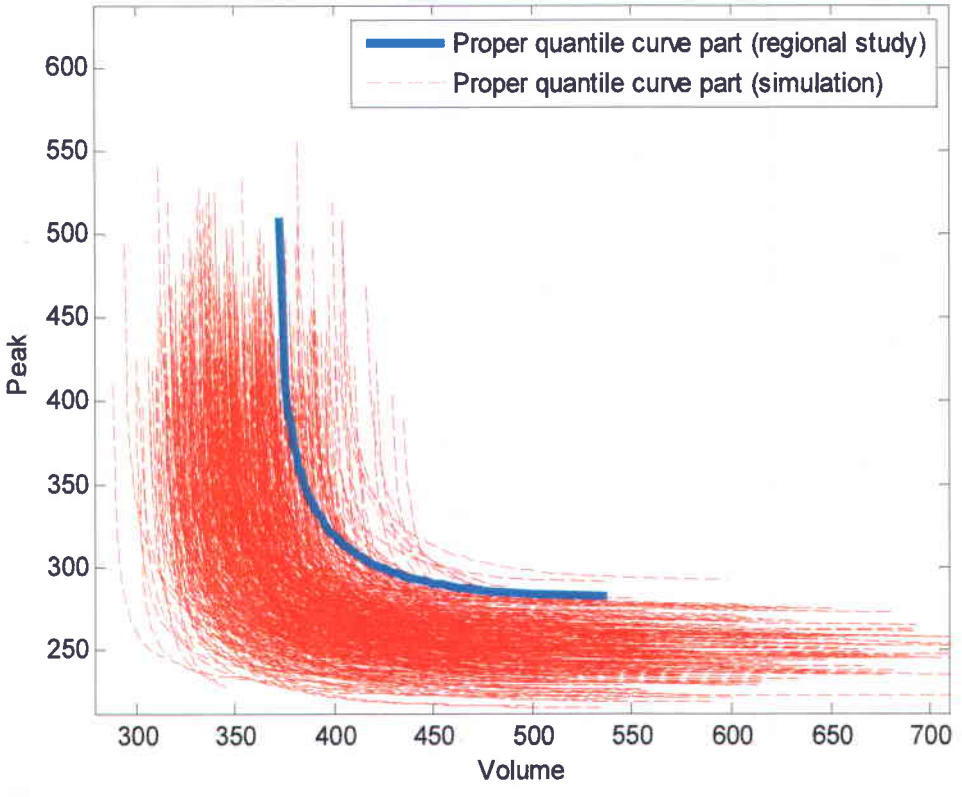
768 **Figure 6b: Univariate and bivariate quantiles corresponding to a nonexceedance probability *p*=0.9, 0.95, 0.99, 0.995 and 0.999**
769 **in Des Escoumins , quantile curve in the unit square and side panels showing the marginal distributions (local and regional) of**
770 ***Q* and *V*.**

771

772



773 **Figure 6c: Univariate and bivariate quantiles corresponding to a nonexceedance probability *p*=0.9, 0.95, 0.99, 0.995 and 0.999**
774 **in Natashquan , quantile curve in the unit square and side panels showing the marginal distributions (local and regional) of *Q***
775 **and *V*.**



776

777 **Figure 7 : Bivariate quantiles (Regional and the 500 simulation) corresponding to a**

778 **nonexceedance probability p=0.9 in the Petit Saguenay station.**

779

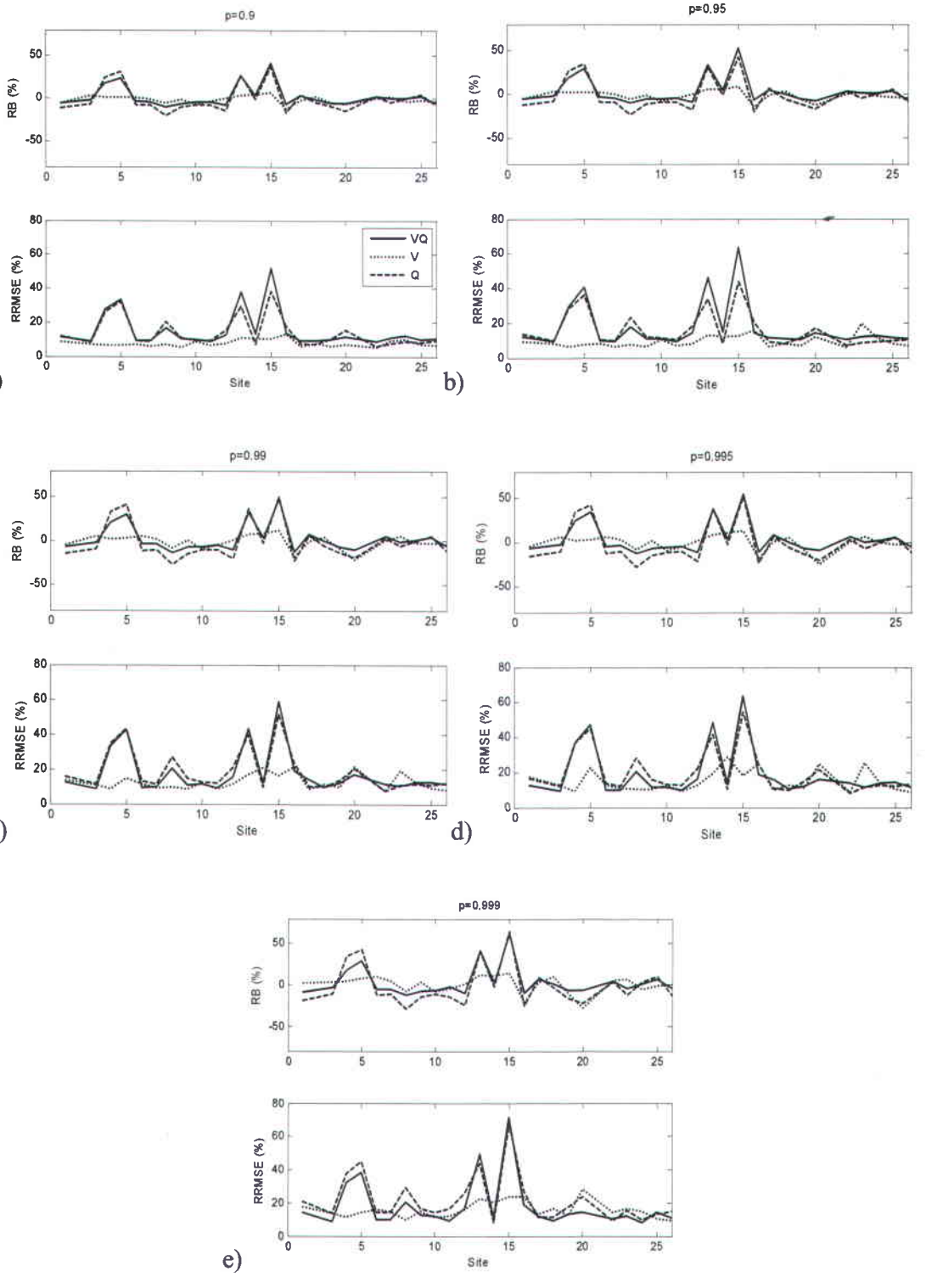
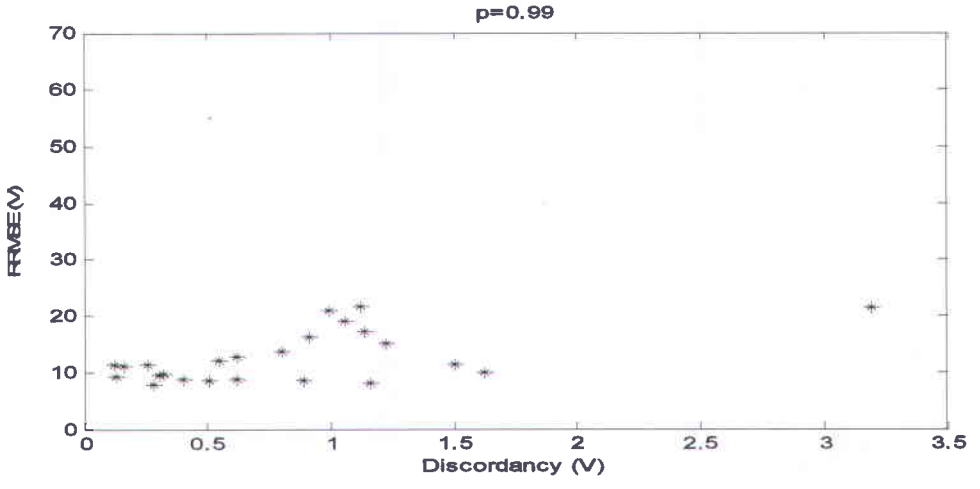
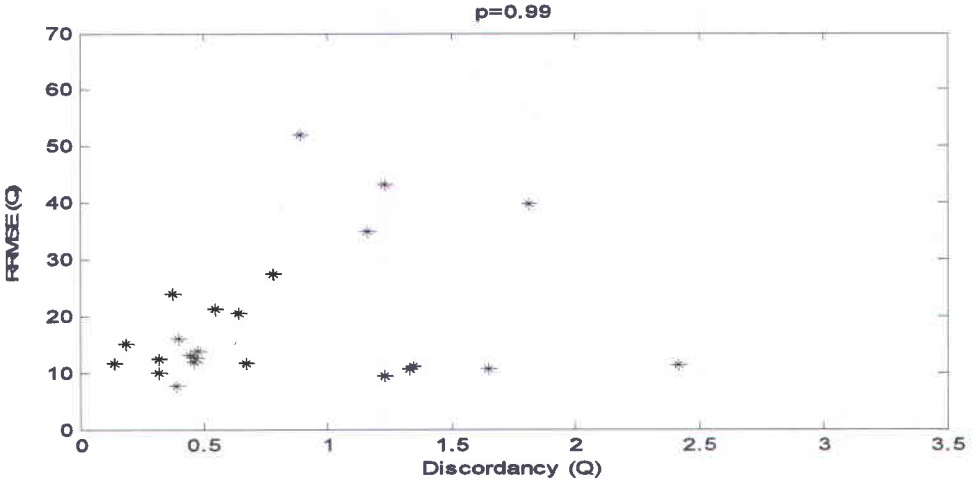


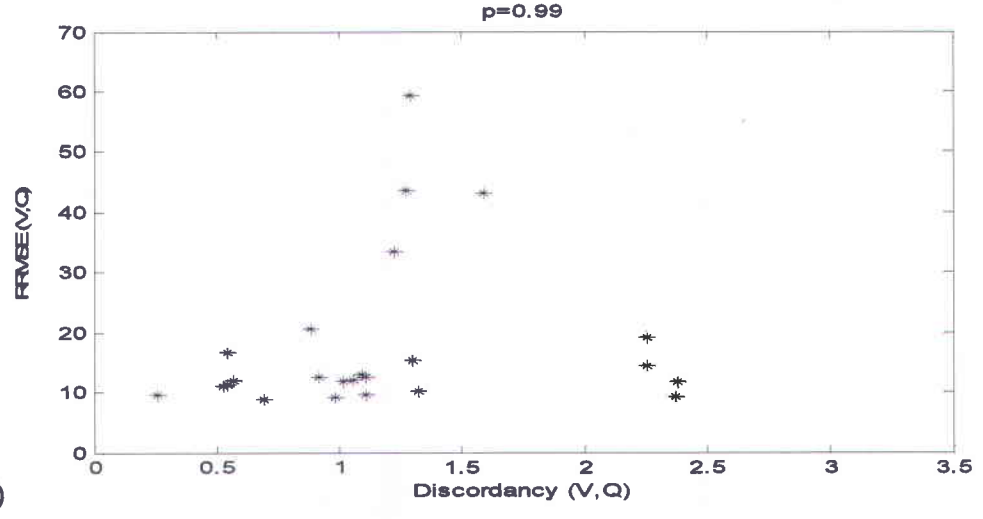
Figure 8 : Performance of the univariate and bivariate quantiles for each site with a) p=0.9, b) p=0.95, c) p=0.99, d) p=0.995 and d) p=0.999. Continuous line: VQ; dotted line: V and dashed line: Q.



786 a)



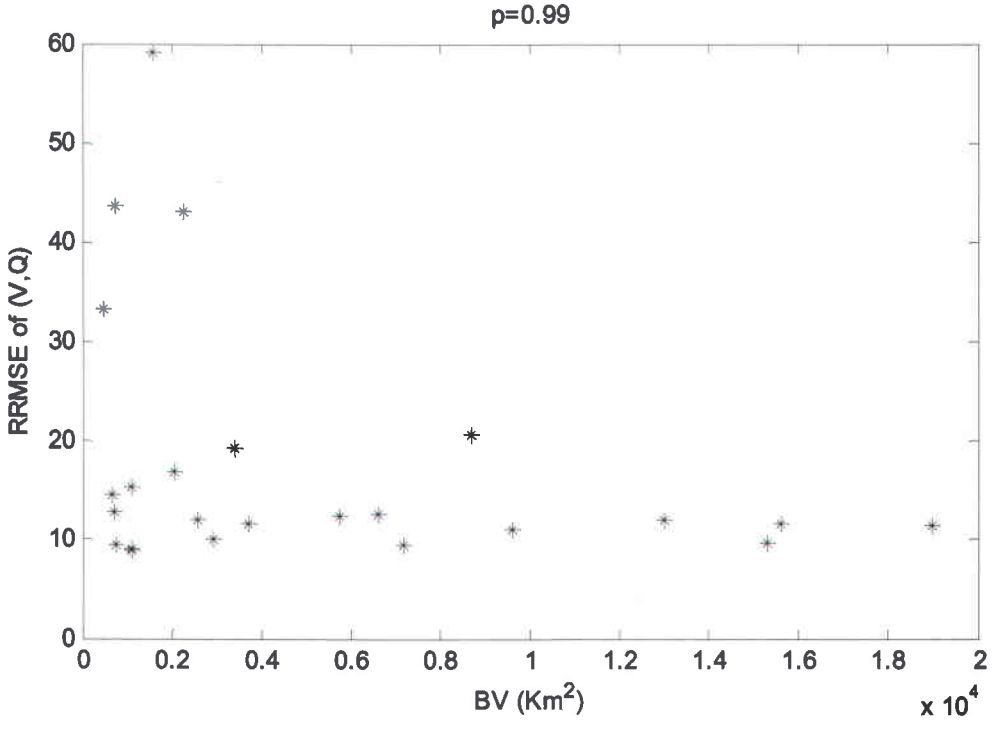
787 b)



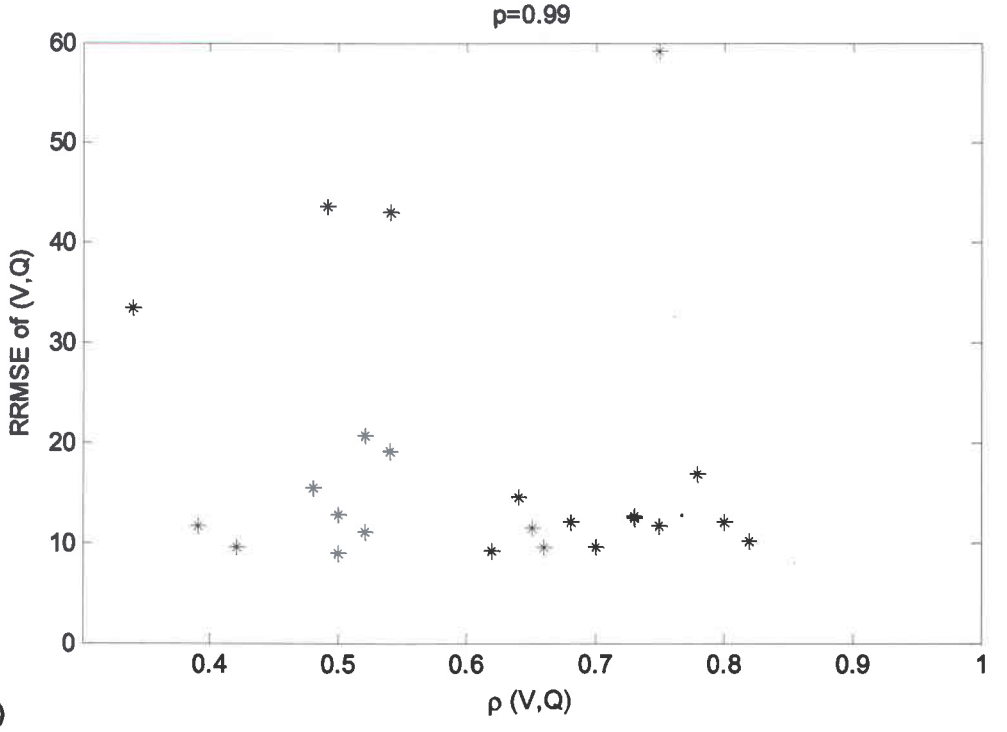
788 c)

789 **Figure 9 : RRMSE (%) of the three models with respect to the corresponding**
790 **discordance values for *p*=0.99: a) margin for V, b) margin for Q, and c) bivariate**

791

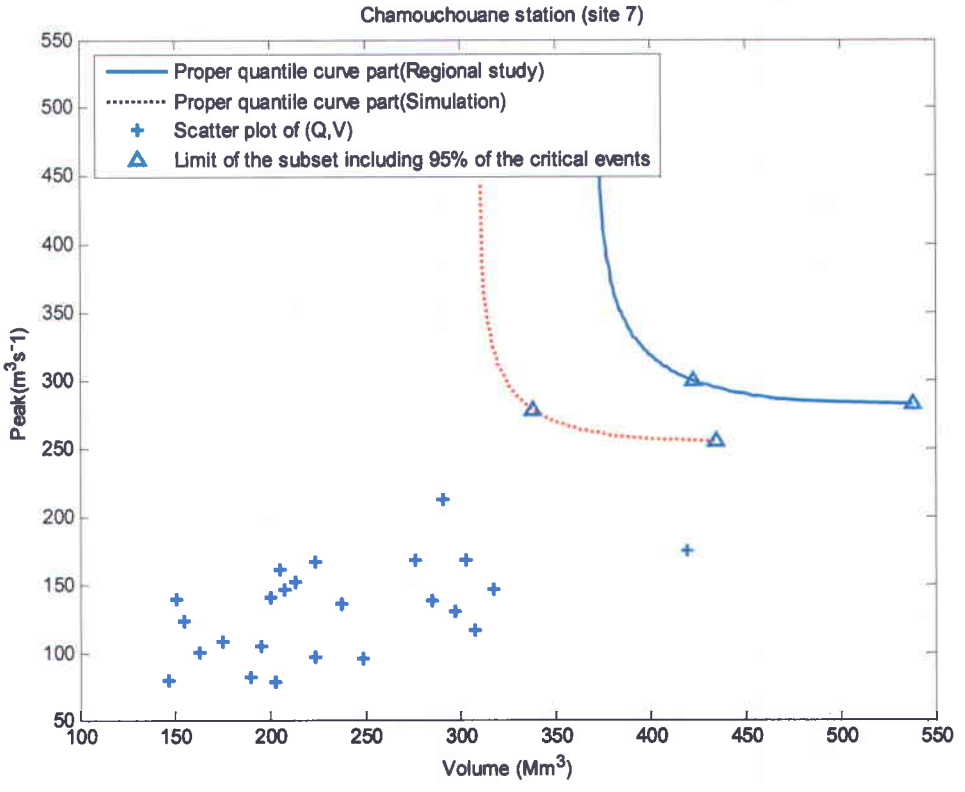


a)



b)

Figure 10 : RRMSE of bivariate quantile for p=0.99 with respect to a) watershed area (BV) and b) correlation between V and Q.



797

798 **Figure 11 : Bivariate quantiles of Chamouchouane station corresponding to a**
799 **nonexceedance probability p=0.9 with scatter plot of (Q,V) and the limit of subset**
800 **that includes the critical events with probability (1-α)=0.95. Simulation in dotted**
801 **line and sample data in solid line**