Université du Québec
INRS - Eau, Terre et Environnement

Impacts de l'inclusion de paramètres de récession linéaire et non-linéaire dans un modèle régional d'estimation des débits d'étiage

Par
Christian Charron

Mémoire présenté pour l'obtention du grade de
Maître ès sciences (M. Sc.)

Jury d'évaluation

Examinateur externe	Faye F.E. Hicks Département de génie civil et environnemental Université d'Alberta
Examinateur interne	André St-Hilaire INRS-Eau, Terre et Environnement
Directeur de recherche	Taha B.M.J. Ouarda INRS-Eau, Terre et Environnement
Codirecteur de recherche	Ousmane Seidou INRS-Eau, Terre et Environnement

RÉSUMÉ

Plusieurs études sur l'estimation de caractéristiques d'étiage à l'aide d'un modèle régressif ont démontré que les performances étaient améliorées avec l'inclusion d'une variable dérivée de caractéristiques de récession des bassins versants. L'objectif de cette étude est de mesurer les impacts de l'inclusion de variables de récession dans des modèles régionaux d'estimation des étiages appliqués à 85 bassins versants du territoire de la province de Québec (Canada). Deux variables différentes ont été utilisées selon que la relation volume-débit considérée était linéaire ou non-linéaire. Les résultats obtenus montrent que le fait d'inclure une variable de récession améliore les estimations et que les performances obtenues avec la variable du modèle non-linéaire du réservoir sont supérieures à celles obtenues avec la variable du modèle linéaire du réservoir. Cependant, pour estimer un paramètre de récession à un site cible, des données hydrologiques doivent être disponibles. Cette étude cherche en plus à quantifier la valeur d'un paramètre de récession estimé à l'aide d'une courte série de données hydrologiques. Des simulations avec des modèles régionaux incluant des variables de récession ont été effectuées pour le cas où quelques années seulement de données hydrologiques sont disponibles au site cible. Les résultats montrent que les performances convergent rapidement vers celles obtenues en considérant les séries de données hydrologiques complètes à mesure que le nombre d'années utilisées pour estimer les paramètres augmente.

Étudiant

Directeur de recherche

REMERCIEMENTS

La réalisation de ce mémoire n'aurait pas pu être menée à terme sans l'aide de plusieurs personnes. J'aimerais donc remercier ici tous ceux qui m'ont aidé et encouragé durant mes travaux de recherche.

Je remercie d'abord mon directeur de recherche Taha Ouarda pour avoir cru en mes capacités et pour tous les précieux conseils qu'il m'a donnés tant sur le plan de la démarche scientifique que sur le plan professionnel.

Merci aussi à tous les membres de l'équipe de la chaire de recherche en hydrologie statistique pour toute l'aide qu'ils m'ont apportée durant mes études. Parmi eux, je remercie plus spécifiquement mon co-directeur de recherche Ousmane Seidou pour sa collaboration.

Je tiens à exprimer toute ma reconnaissance à ma mère pour m'avoir encouragé à poursuivre mes études. Je suis également reconnaissant envers les membres de ma famille et les amis qui m'ont supporté durant ma maîtrise.

Finalement, j'exprime ma gratitude envers tous les membres du jury pour avoir accepté d'évaluer ce mémoire.

TABLE DES MATIÈRES

Partie 1 : Synthèse

INTRODUCTION

Pour des besoins en gestion des ressources hydriques, la connaissance des caractéristiques d'étiage est d'une importance capitale. Les domaines d'applications, où de telles statistiques sont requises, sont vastes et comprennent l'approvisionnement en eau potable, la conservation d'un débit minimal pour assurer la survie des espèces aquatiques (débit réservé), la protection des organismes vivants lors de déversements de contaminants et la gestion de centrales hydro-électriques. Une statistique fréquemment employée est le quantile d'étiage $Q_{d,T}$: la valeur du débit d'étiage sur d jours consécutifs associée à une période de retour de T années. Le présent mémoire est une étude sur les impacts de l'inclusion de paramètres de récession dans un modèle régional d'estimation des quantiles d'étiage.

Lorsqu'une série de données hydrologiques disponible à un site d'intérêt s'étend sur plusieurs années, les quantiles d'étiage peuvent être estimés par une méthode d'analyse fréquentielle avec l'information locale. Cependant, lorsque peu ou aucune information hydrologique locale n'est disponible, l'utilisation d'une méthode d'estimation fréquentielle régionale est alors requise. Ces méthodes définissent d'abord habituellement des regroupements de bassins homogènes. Ensuite, un transfert au site cible de l'information régionale des sites jaugés appartenant à la même région hydrologique est effectué à l'aide d'une méthode d'estimation régionale. Pour ce faire, des variables indépendantes caractérisant la physiographie et la météorologie des bassins versants sont souvent utilisées pour le regroupement des stations et pour le transfert de l'information hydrologique. La qualité de l'estimation au site d'intérêt dépend alors beaucoup de la capacité de ces variables indépendantes à caractériser le régime hydrologique à faible débit.

Plusieurs auteurs mentionnent, parmi les facteurs explicatifs importants des débits d'étiages, la géologie et l'hydrogéologie du bassin versant. Ceci s'explique par le fait qu'en période de sécheresse, les rivières sont alimentées principalement par le déversement dans le lit de la rivière de l'eau souterraine emmagasinée dans l'aquifère du bassin versant. La géologie de l'aquifère devient alors un facteur important influençant la décharge. En pratique, il est cependant difficile d'exprimer des caractéristiques de ce type par une ou des variables explicatives. Des recherches ont démontré que certains indices caractérisant la récession du cours d'eau principal du bassin versant pouvaient être utilisés en remplacement des caractéristiques géologiques des bassins versants. Des indices de récession ont fréquemment été utilisés dans des modèles régressifs, entre autres par Bingham (1986), Vogel et Kroll

(1992, 1996) et Kroll et al. (2004). Ces études ont démontré que ces indices étaient d'importantes variables explicatives pour les quantiles d'étiage.

Dans la présente étude, des variables de récession ont été calculées et incluses dans des modèles régionaux d'estimation des étiages. Une méthode pour identifier les périodes de récession a été développée en se basant sur les travaux de Vogel et Kroll (1992). Deux paramètres de récession ont été utilisés: un provenant d'un modèle linéaire du réservoir et l'autre d'un modèle non-linéaire du réservoir. Des regroupements de stations en voisinages propres à chacune des stations avec la méthode d'analyse canonique des corrélations (ACC) ont aussi été effectués.

Cette présente section du mémoire est la synthèse de la contribution de l'étudiant à l'avancement des connaissances reliées au domaine de recherche. Le chapitre 1 de cette synthèse présente la contribution propre à l'étudiant lors du travail de recherche ayant mené à l'article scientifique présenté en deuxième partie. Le chapitre 2 présente la situation de ces contributions par rapport aux autres travaux scientifiques cités dans l'article. Cette synthèse se termine par une conclusion résumant les principales contributions du travail de recherche.

CHAPITRE 1

CONTRIBUTION DE L'ÉTUDIANT

Une revue de littérature a été effectuée par l'étudiant et les résultats sont présentés à la section 2 de l'article en deuxième partie. L'étudiant a identifié plusieurs études où des indices caractérisant la récession ont été utilisés dans des modèles régressifs visant l'estimation de quantiles d'étiage. Les indices de récession employés assument habituellement la relation linéaire suivante:

$$S = kQ \tag{1}$$

où S est le volume stocké, Q est le débit sortant et k est une constante en unité de temps. Le débit au temps t s'exprime alors par la simple relation exponentielle :

$$Q_t = Q_o e^{-t/k} \tag{2}$$

Cependant, de récentes études ont démontré que la relation non-linéaire :

$$S = aQ^b \tag{3}$$

où b est un exposant entre 0 et 1 et a est un coefficient, est plus réaliste (Moore, 1997, Wittenberg, 1999). L'équation du débit au temps t s'exprime alors par :

$$Q_t = Q_o \left[1 + \frac{(1-b)Q_o^{1-b}}{ab} t \right]^{1/(b-1)} \tag{4}$$

Dans les études consultées, aucun des indices de récession utilisés dans un cadre relatif à l'estimation régionale de caractéristique d'étiage n'assume que la relation (3) s'applique.

La revue de littérature a permis de constater que les variables de récession étaient pour plusieurs études estimées uniquement pour des bassins jaugés. Ces études montrent l'importance d'utiliser une variable de récession mais ne donnent pas le moyen d'en estimer la valeur pour une station non-jaugée. Dans d'autres études, l'estimation de ces variables à un site cible était obtenue à l'aide de cartes géographiques délimitant des surfaces d'indice de récession homogène (Bingham, 1986, Arihood et Glatfelter, 1991). Vogel et Kroll (1992) suggèrent qu'une variable de récession peut aussi être estimée avec un hydrogramme de courte durée. Cependant, aucune étude ne quantifie la validité d'une telle procédure.

Après la consultation des études au cours de la revue de littérature, l'étudiant a proposé d'utiliser le paramètre *a* du modèle non-linéaire comme un indice de récession et de l'inclure dans un modèle régional. Comme le modèle non-linéaire est jugé plus réaliste, il espérait que ce paramètre caractériserait mieux les bassins versants. Un objectif principal de cette étude était donc de comparer les performances obtenues avec les paramètres *a* et *k* inclus séparément dans des modèles régionaux. L'étudiant a aussi proposé de calculer les paramètres de récession en utilisant des séries hydrologiques de courtes périodes et d'évaluer les performances obtenues en utilisant ces valeurs au site cible dans les modèles régionaux. L'objectif était de quantifier les impacts sur les estimations des quantiles lorsqu'un paramètre est estimé sur une courte période.

La méthodologie employée dans l'étude est décrite à la section 3 de l'article. À partir des paramètres *k* et *a* estimés à tous les bassins versants jaugés du cas d'étude, les variables de récession *K* et *A* ont été créées respectivement. Six modèles régionaux différents ont été définit selon les diverses variables incluses dans le modèle et selon qu'un regroupement des stations en voisinages avec l'ACC a été effectué ou non. Ainsi, pour trois modèles, aucun regroupement n'a été effectué et la régression s'est effectuée au site cible à l'aide de toutes les autres stations. Ces modèles se distinguent entre eux selon qu'aucune variable de récession n'a été incluse, que la variable *K* a été incluse ou bien que la variable *A* a été incluse. Pour les trois autres modèles, les trois mêmes configurations de variables distinguent les modèles entre eux, mais en plus, des regroupements de stations homogènes ont été effectués avec la méthode de l'ACC.

Pour chaque modèle, le choix des variables explicatives incluses s'est fait par une analyse des corrélations entre les variables explicatives et les quantiles spécifiques. Ces derniers sont les quantiles dont les valeurs à chaque bassin versant ont été divisées par la superficie pour réduire l'influence du facteur d'échelle. La table des corrélations entre les variables explicatives a aussi servi à identifier les variables redondantes. Cette méthode de sélection des variables est similaire a celle employée par Ouarda et al. (2005).

La méthode de regroupement utilisée dans le modèle régional de cette étude est celle de l'analyse canonique des corrélations (ACC). Cette méthode est décrite en détails par Ouarda et al. (2001). Elle avait été appliquée à l'origine pour l'estimation des quantiles de crues mais a aussi été appliquée récemment aux quantiles d'étiages par Ouarda et al. (2005). Cette méthode définit des variables canoniques à partir des groupes de variables hydrologiques et explicatives. Cela permet d'identifier, dans l'espace des variables canoniques hydrologiques obtenu, des bassins au régime hydrologiquement similaire autour d'un site cible non-jaugé. La taille du voisinage dépend alors de la spécification d'un paramètre qui défini un seuil de confiance. La méthode d'estimation régionale utilisée est la régression multiple telle qu'utilisée dans les diverses études de régionalisation avec des variables de récession.

Les performances de chaque modèle ont été évaluées par une méthode de rééchantillonage consistant à considérer chaque station à tour de rôle comme non-jaugée et à estimer les variables hydrologiques à chaque fois. Ensuite, des indices de performance tels que le biais relatif et la racine de l'erreur quadratique relative moyenne ont été calculés avec les valeurs des estimations régionales et locales. Ces indices sont définis par les relations suivantes :

$$rBIAS = \frac{1}{N}\sum_{i=1}^{N}\left[(\hat{q}_i - q_i)/q_i\right]\times 100 \tag{5}$$

$$rRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[(\hat{q}_i - q_i)/q_i\right]^2}\times 100 \tag{6}$$

où $\hat{q}_i$ est l'estimation régionale du quantile au site i, q_i est l'estimation locale du quantile au site i et N est le nombre total de stations. Ces deux indices sont fréquemment utilisés en hydrologie, notamment par Ouarda et al. (2001).

Afin de mesurer l'impact de l'estimation d'un paramètre de récession sur une courte période, l'étudiant et son directeur ont défini une nouvelle méthode de rééchantillonage. Pour ce faire, les estimations des paramètres de récession ont été obtenues sur une fenêtre mobile d'un nombre d'années n allant de 1 an à 5 ans. L'étudiant a modifié les algorithmes des modèles régionaux pour remplacer la valeur d'un paramètre donné au site cible par chacune des estimations de ce paramètre pour une valeur n donnée. L'étudiant a adapté la formule de calcul des indices de performances des équations (5) et (6) puisque plusieurs estimations étaient obtenues à chaque site. Ces nouveaux indices sont les suivants:

$$rBIAS_n = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{P_{i,n}}\sum_{j=1}^{P_{i,n}}\left[(\hat{q}_{ij} - q_i)/q_i\right]\right)\times 100 \tag{7}$$

$$rRMSE_n = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{P_{i,n}}\sum_{j=1}^{P_{i,n}}\left[(\hat{q}_{ij} - q_i)/q_i\right]^2\right)}\times 100 \tag{8}$$

où $P_{i,n}$ est le nombre d'estimations du paramètre de récession considéré au site i pour n années et q_{ij} est l'estimation régionale j obtenue au site i.

L'estimation d'un paramètre de récession pour un bassin versant donné requière une analyse de récession. Cette analyse consiste à délimiter des segments de récession durant les périodes de sécheresse et à estimer les paramètres k et a pour chaque segment. Pour sélectionner les segments de récession, il existe plusieurs méthodes. Traditionnellement, des méthodes graphiques étaient employées. Ces méthodes ont le désavantage d'être lentes, subjectives et approximatives. À l'aide des ordinateurs, des méthodes de sélection automatique ont été développées. Généralement, des critères de départ et de longueur sont définis pour sélectionner des segments à débit uniquement décroissant. Ensuite, le calcul d'un paramètre de récession pour un bassin versant donné peut se faire en prenant la moyenne des estimations obtenues.

L'étudiant a développé un algorithme de sélection automatique des segments de récession. La méthode de sélection employée consiste à sélectionner des segments de débits moyens sur trois jours qui sont uniquement décroissants et de longueur d'au moins 10 valeurs. La méthode employée a été fortement inspirée par celle décrite dans Vogel et Kroll (1992). L'étudiant l'a cependant adaptée aux particularités du cas d'étude. En effet, une investigation du cycle hydrologique de plusieurs stations a montré une grande variabilité temporelle des périodes de sécheresses selon la latitude. La méthode développée par l'étudiant est basée sur l'identification des périodes de sécheresse à l'aide du débit maximum de la crue et du débit minimum d'étiage pour une saison donnée. À cela s'ajoute des critères reliés à des valeurs seuils de débit maximal et minimal à respecter pour chaque segment. Cette méthode serait plus appropriée que l'approche définissant les périodes de sécheresses par dates fixes comme dans Vogel et Kroll (1992) pour le grand territoire d'étude en raison de la variabilité du régime hydrologique rencontrée. L'étudiant a estimé les paramètres k et a pour les deux périodes de sécheresse associées aux débits d'étiage estival et hivernal.

Le cas d'étude considéré est présenté dans la section 4 de l'article. Les variables hydrologiques, physiographiques et météorologiques du cas d'étude proviennent de stations du Québec utilisées dans l'étude de Ouarda et al. (2005). Au cours de cette dernière étude, les séries de débits minimums annuels des stations ont été sujettes à des tests d'indépendance, de stationnarité et d'homogénéité pour obtenir des séries indépendantes et identiquement distribuées. L'analyse fréquentielle locale des stations a été effectuée lors de cette même étude. Les quantiles ont été calculés en ajustant la loi statistique sélectionnée aux données de chaque station. Le choix de la loi la plus appropriée pour une station donnée s'est fait à l'aide du critère bayésien d'information. Les mêmes séries journalières utilisées par Ouarda et al. (2005) pour le calcul des quantiles d'étiages ont été utilisées dans cette présente étude par l'étudiant pour l'analyse de récession ayant conduit à la création des variables de récession.

Les résultats de l'étude, obtenus par l'étudiant, sont présentés à la section 5 de l'article. L'analyse des résultats a été effectuée par l'étudiant en collaboration avec son directeur de recherche. Ce dernier a suggéré à l'étudiant de produire certains graphiques et résultats supplémentaires à ceux proposés initialement par l'étudiant afin d'offrir une analyse plus complète. La rédaction de l'article a été effectuée par l'étudiant. Son directeur de recherche a ensuite apporté quelques modifications au texte.

CHAPITRE 2

SITUATION DE LA CONTRIBUTION SCIENTIFIQUE PAR RAPPORT AUX TRAVAUX CITÉS

Dans cette étude, une variable provenant d'un paramètre d'un modèle non-linéaire du réservoir a été utilisée dans un modèle régressif pour estimer des quantiles d'étiage. Ceci n'a pas, à la connaissance de l'étudiant, été tenté précédemment. En effet, dans toutes les études consultées, les auteurs assumaient un modèle linéaire du réservoir pour définir des indices de récession. Cette contribution de l'étudiant est importante puisque les résultats obtenus dans cette étude démontrent que les performances sont significativement améliorées en utilisant le paramètre du modèle non-linéaire au lieu de celui du modèle linéaire.

Dans la majorité des études consultées sur la régionalisation avec une variable de récession, celle-ci était calculée à des stations jaugées en considérant toute la série des mesures hydrologiques disponibles. Qu'en est-il lorsque les débits à une station ont été mesurés sur une courte période ? Peuvent-ils permettre d'obtenir une estimation adéquate d'un paramètre de récession? L'impact de la longueur de la série de données hydrologiques sur les estimations restait inconnu jusqu'à présent. Cette étude tente de combler cette lacune en calculant les performances moyennes obtenues lorsque la longueur de la série de données au site cible est très courte. Les indices de performance des équations (7) et (8) y sont présentés en fonction du nombre d'années n considérés. Les résultats obtenus démontrent une convergence rapide des indices de performance avec le nombre d'années n.

Dans cette étude, les variables de récession ont été calculées pour les deux périodes d'étiage séparément (hivernale et estivale). Deux variables ont ainsi été obtenues pour chaque paramètre de récession. Dans plusieurs études, les variables de récession sont calculées durant une seule saison (Bingham, 1986, Tallaksen, 1989, Vogel et Kroll, 1992). L'impact du choix de la saison sur les performances du modèle n'avait pas été étudié précédemment. Les résultats de cette étude ont démontré que les variables calculées durant une saison donnée constituaient de meilleures variables explicatives pour les variables hydrologiques caractérisant des événements de faibles débits durant la même saison.

Dans les études consultées utilisant des variables de récession dans des modèles régressifs, aucun regroupement de stations homogènes n'est effectué. Dans cette étude, une méthode de regroupement de stations en voisinages a été appliquée. Les performances obtenues avec les regroupements ont été comparées avec les performances obtenues sans regroupement. Les résultats montrent en général de meilleurs résultats avec le regroupement en voisinages.

RÉFÉRENCES DE LA SYNTHÈSE

Arihood, L.D. and Glatfelter, D.R., 1991. Method for estimating low-flow characteristics of ungaged streams in Indiana. U.S. Geological Survey Water-Supply Paper No. 2372, 22 pp.

Bingham, R.H., 1986. Regionalization of low-flow characteristics of Tennessee streams. U.S. Geological Survey Water-Resources Investigation Report 85-4191, 63 pp.

Kroll, C., Luz, J., Allen, B. and Vogel, R.M., 2004. Developing a watershed characteristics database to improve low streamflow prediction. Journal of Hydrologic Engineering 9 (2), 116-125.

Moore, R.D., 1997. Storage-outflow modelling of streamflow recessions, with application to a shallow-soil forested catchment. Journal of Hydrology 198, 260-270.

Ouarda, T.B.M.J., Girard, C., Cavadias, G.S. and Bobée, B., 2001. Regional flood frequency estimation with canonical correlation analysis. Journal of Hydrology 254, 157-173.

Ouarda, T.B.M.J., Jourdain, V., Gignac, N., Gingras, H., Herrara, H. et Bobée, B., 2005. Développement d'un modèle hydrologique visant l'estimation des débits d'étiage pour le Québec habité. INRS-ETE, Rapport de recherche No. R-684-f1, 174 pp.

Tallaksen, L.M., 1989. Analysis of time variability in recessions. IAHS Publ. 187, 85-96.

Vogel, R.M. and Kroll, C.N., 1992. Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. Water Resources Research 28 (9), 2451-2458.

Vogel, R.M. and Kroll, C.N., 1996. Estimation of baseflow recession constants. Water Resources Management 10, 303-320.

Wittenberg, H., 1999. Baseflow recession and recharge as nonlinear storage processes. Hydrological Processes 13, 715-726.

Partie 2 : L'article scientifique

RÉSUMÉ EN FRANÇAIS

Voir le résumé à la page iii.

Impacts of the inclusion of linear and non-linear baseflow recession parameters in a regional low flow model

Christian Charron*
Taha B.M.J. Ouarda

Canada Research Chair on the Estimation of Hydrological Variables
Hydro-Québec/NSERC Chair in statistical Hydrology
INRS-ETE, University of Quebec
490, de la Couronne
Quebec (Quebec) G1K 9A9, CANADA

*Corresponding author: Email: christian_charron@ete.inrs.ca
Tel: +1 418 654 2530 x4481

Submitted to Journal of Hydrology

March 2007

Abstract: Several studies on the prediction of low flow statistics with regional regressive models have shown that improvements can be obtained with the inclusion of a variable derived from characteristics of catchment baseflow recessions. The purpose of this paper is to investigate the impacts of the inclusion of recession variables in low flow regional models applied to 85 gaged catchments over a large area of the province of Quebec (Canada). Two different recession variables were used depending on whether a linear or a non-linear storage-outflow relationship for the aquifer was considered. Results show that the inclusion of a recession variable improves the estimation of regional models and that including a variable from the parameter of a non-linear reservoir model gives better performances than one from the parameter of a linear reservoir model. However, a hydrological record is needed to estimate a recession parameter. This paper seeks also to evaluate the efficiency of recession parameter estimates obtained from short hydrographs. Regional models with recession variables were tested for the case where only few years of hydrological data are available to compute the recession characteristic at target sites. Results indicate that performances converge rapidly to those obtained with complete hydrographs as the number of years used to estimate the parameters increases.

Keywords: Regional estimation, Low flow, Baseflow, Recession analysis, Canonical correlation analysis

1. INTRODUCTION

It is of major importance to engineers and water managers to estimate the frequency of low flow events at a given site of interest. Applications that need such information are related to water quantity and quality management and include water supply, hydropower production, dilution of pollution discharge and aquatic wildlife protection. When a long at site hydrologic record is available at a site of interest, it is used to compute low flow statistics using local frequency analysis. However, when little or no data are available, a regional approach needs to be employed (Hamza et al., 2001, Kroll et al. 2004, Laaha and Blöschl, 2006). The most used low flow statistics in hydrology are the quantiles $Q_{d,y}$ of the minimum mean discharge over d-days corresponding to a return period of y-years.

For the purpose of regional estimation, a multiple regression is often used to estimate low flow quantiles with independent variables characterizing physiographic catchment properties and meteorologic conditions. Among the various explanatory variables generally used, those representing geological and hydrogeological characteristics are often stated to have a big influence on low flow regimes (Bingham, 1986, Tallaksen, 1989, Vogel and Kroll, 1992, Smakhtin, 2001). However, variables describing geological conditions in a catchment are hard to establish and difficult to quantify (Demuth and Hagemann, 1994). To compensate for the lack of information on catchment characteristics that influence low flows, many authors have used indirect variables derived from catchment recessions to improve the efficiency of regressive models and obtain better estimates of low flow statistics (Bingham, 1986, Tallaksen, 1989, Arihood and Glatfelter, 1991, Vogel and Kroll, 1992, 1996, Demuth and Hagemann, 1994, Kroll et al., 2004).

In this paper, regional models including variables characterising the recession were applied to catchments over a wide area in the province of Quebec (Canada). The low flow statistics estimated in this study are low flow quantiles from the summer and winter dry seasons. Two recession variables were defined depending on whether the parameter estimated is the solution of a linear reservoir model or a non-linear reservoir model. Each recession variable was included in a configuration of explanatory physiographical and meteorological variables of the regional model. Parameters from the linear model have already been used in low flow regressive models but a parameter from the non-linear model was never used in the present context. An objective of this paper is to compare the performances obtained with each of these parameters. In addition, this paper seeks also to investigate if valuable estimates of the recession variables can be obtained from modest hydrographs with record data that are only of few years long and if they can lead to reliable regional estimations. Attempts to delineate homogenous neighbourhoods in the regional model were also performed to see if this can improve the performances.

In the next section of this paper, we present a review of the theoretical background providing information regarding the various recession characteristics used as independent variables in regressive models, the delineation of recession segments, the modeling of recession curves and the delineation of homogenous neighbourhoods. In the third section, the methodology is exposed with details concerning the recession analysis method, the regional model and the performance criteria used. The forth section presents the application area. In the fifth section, the results are presented and discussed. Finally, the conclusions and recommendations are given.

2. REVIEW AND THEORETICAL BACKGROUND

2.1 Recession characteristics used as independent variables in regressive models

Streamflow discharge, during a period of no or little precipitation, is a complex mechanism involving the release of water from natural storage sources. During floods, caused by important precipitations or snow melt, water is stored in the aquifer of a catchment. When the recharge ceases, groundwater is gradually depleted into the stream where it intersects the water table. This is called a recession, corresponding to a downward curve on a hydrograph. During a recession, recharge from surface flow becomes null or negligible and most of the streamflow is composed of the baseflow that is defined by the portion of flow coming from groundwater storage and other delayed sources (Hall, 1968). Aside from groundwater, additional sources include lakes, marshes, snow and ice, and stream channel and bank storage (Hall, 1968). However, in many cases, the majority of natural gains to streamflow during low flow periods are derived from groundwater storage (Smakthin, 2001). Recession curves are then representative of the storage-outflow relationship of the aquifer. It is expected that the recessions are influenced by geological and hydrogeological characteristics of the catchment aquifer. With this assumption, recession descriptors have been used as an indirect representation of these characteristics.

Several variables representing recession characteristics have been used as independent variables in regressive models. In these studies, the assumption was made that the storage-outflow relationship is linear. The baseflow is then described by an exponential function approximated by a straight line on a semi-log scale. Using this model, different parameters have been derived from recession curves. Vogel and Kroll (1992) have, from a conceptual streamflow model, found that a recession constant, defined by the decrease rate of the recession curve, is an important parameter to explain low flow quantiles in addition to the area and the density of the basin. In Kroll et al. (2004), this constant has led to important improvements in performances compared with regressive models including other catchment characteristics. Arihood and Glatfelter (1991) included in a regressive model a flow-duration ratio defined by the 20% flow-duration divided by the 90% flow-duration computed from a flow-duration curve that is a graphical representation of the discharge versus the percentage of time a given value is equalled or exceeded. Bingham (1986) used a recession-index, which gives a measure in units of days of the time required for a recession curve to decrease by a factor of 10 on a log scale. The exponential model is implicitly considered in his study because curves that approximate straight lines on semi-log plots are selected.

To estimate a recession characteristic, local hydrological data are needed. Consequently, it cannot be estimated for ungaged basins. However, authors have tried to estimate it for ungaged basins by relating a given recession characteristic to information on geological characteristics. Demuth and Hagemann (1994) estimated the recession constant with a regressive model that includes hydrogeological indices as independent variables. Bingham (1986) and Arihood and Glatfelter (1991) used maps to retrieve recession characteristics for ungaged catchments where areas of similar recession indices have their boundaries delineated with maps of geological formations. The recession index for a given catchment is retrieved by a weight-average of values associated with recession areas within the boundaries of the catchment total area. In the present study, the adopted approach is to estimate recession characteristics from very short data series.

2.2 Delineation of recession segments

The computation of recession characteristics at gaged stations is usually done with a recession analysis. This involves the delineation of baseflow recession segments from hydrographs and the computation of recession characteristics from these segments. In practice, the interpretation of hydrographs is complicated by the fact that during a recession period, recharge events can often interrupt a recession and produce many recession segments of different lengths. Another interpretative complication comes from the fact that the different components of streamflow that are surface flow, interflow and baseflow are difficult to quantify at a given time. Given these considerations, various researchers have developed methods to delineate baseflow recession segments from hydrographs.

Traditionally, graphical techniques are used for recession analysis. They are however subjective and applicable only for a few analyses because they are time consuming. For a large database, automated methods are preferred. Several methods have been proposed in the literature. They usually take only decreasing portions of hydrographs in which starting and duration criteria are defined. The minimal length of individual recession segments can vary usually between 4 days to 10 days (Tallaksen, 1995). A portion at the beginning of recession segments can also be removed to avoid the presence of surface flow. In Vogel and Kroll (1992), a portion of 30% at the beginning of each segment is removed to ensure that the surface flow is negligible. This is an arbitrary value that has been chosen because it has given segments that were the most in concordance with the linear hypothesis for basins under their study. In Hammond and Han (2006), a window that considered only the past n hours was used to fit the recession models to the data. Results showed that using a window instead of all the data from the peak of the curve reduces the error related to the fit to the data.

Recession characteristics show high variability over the course of a year. This is principally caused by the climate seasonal variation (Tallaksen, 1995, Wittenberg, 2003). For this reason, recession segments are often taken within a given period of the year (Bingham 1986, Tallaksen, 1989, Vogel and Kroll, 1992, Brandes et al. 2005). Dry seasons are then often defined by a fixed starting date and ending date. In Vogel and Kroll (1992), recession constants are calculated for the summer period between July and October. Bingham (1986) and Brandes et al. (2005) had rather chosen to take the period between November and March in an attempt to minimise the effect of evapotranspiration which is less important during the winter season.

To characterise baseflow behaviour, selected segments should as much as possible be composed of baseflow. Several techniques have been developed to separate baseflow from total outflow. Nathan and McMahon, (1990a) stated two main type of baseflow separation techniques: those that assume that the baseflow responds to a storm event concurrently with surface runoff and those that account for the effects of bank storage and assume that the baseflow recession continues after the time when surface runoff begins. The assumption made within the techniques of the second type causes the theoretical baseflow to peak a certain time after the total hydrograph peaks. Chen (2006) demonstrated that the influence of the bank storage delaying effect can considerably postpone the baseflow peak. This influence is considered in the definition of the recession analysis method employed in the present study.

After having sorted out the recession segments, recession parameters are computed. To do this, two general approaches are frequently used: calculating the parameters for each segment and taking the mean of all individual estimations, or grouping all curves together to create a master recession curve and calculating the parameters for this master curve. The mean of the parameter values of all the segments gives a good approximation of the real value of a recession statistic as long as their number is large. The master recession approach is used in an attempt to overcome the problem of the high variability in the recession behaviour of individual segments (Tallaksen, 1995). However, some authors doubt the physical validity of the master recession technique. Chapman (2003) stated that recessions of such long duration are physically impossible, as they are always interrupted by recharge events. For Wittenberg (1999), this method tends to ignore the more curved shape of the individual recession curves.

2.3 Recession curve modeling

For recession analysis, the linear assumption of the storage-outflow relationship has often been made. The solution of the recession curve for the linear reservoir model is then a simple exponential equation. Non-linearity in recession curves has also been observed and several models that handle non-linearity have been proposed in the literacy. When considering a power-law in the storage-outflow relationship, non-linear equations of recession curves can be derived. Such equations have been used for modeling recession curves by several authors

(Wittenberg, 1994, Moore, 1997, Chapman, 1999). These equations have given, in general, a better fit to the data than the simple exponential equation.

Another approach is to consider that the recession outflow comes from multiple storage sources that add together (Hall, 1986). Moore (1997) proposed a model composed of two parallel linear reservoirs. Comparing this model with the power-law reservoir model and the linear reservoir model, he concluded that a model with two linear reservoirs gives a better fit to the data, followed by the power-law reservoir model. Wittenberg (1999) stated that the physical basis of considering the presence of two or more independent storage zones does not hold for most catchments and that a better fit is observed because there are more parameters in the equation.

The most used equation to describe the baseflow recession is the exponential function

$$Q_t = Q_0 e^{-t/k} \quad (1)$$

where Q_t is the outflow at a time t, Q_0 the initial outflow and k is a recession constant representing the residence time of water in the aquifer (k has the dimension of time) (Hall, 1986). Equation (1) in its exponential form is often used to describe recession curves because of its simplicity. Equation (1) is the solution of the continuity equation:

$$dS / dt = I - Q \quad (2)$$

where S is the water storage, Q the outflow and I the inflow. To obtain Equation (1), we make assumptions that inflow I is null for the recession period and that the relation between the storage and the outflow is explained by the linear relation

$$S = kQ \quad (3)$$

Equation (1) is logarithmically transformed to obtain a linear function

$$\ln(Q) = \ln(Q_0) - t / k \quad (4)$$

The parameter k can then be easily obtained by estimating the slope of the line in a semi-log scale.

It is also possible to consider the non-linear storage-outflow relation:

$$S = aQ^b \tag{5}$$

by adding an exponent b and where a is the recession coefficient. The following solution is then obtained (Wittenberg ,1999)

$$Q_t = Q_o \left[1 + \frac{(1-b)Q_o^{\,1-b}}{ab} t \right]^{1/(b-1)} \tag{6}$$

Chapman (1999, 2003) expressed Equation (6) as a function of residence time, which he expressed as $\tau_0 = S_0 / Q_0$, where S_0 is the initial storage. Unlike the recession constant k of the linear model, τ_0 is not a constant value since it depends on the initial storage S_0 and the outflow Q_0. However, the coefficient a and the exponent b are constant parameters for a given reservoir. The recession coefficient a characterizes the storage-outflow relation. According to Wittenberg (1999), this parameter is related to the porosity, hydraulic conductivity and morphometric properties of the catchment. If the exponent b in equation (5) is set to the unit value, a becomes equivalent to k. In practice, the recession parameter estimates obtained for a given catchment can present considerable variations through the year. According to Tallaksen (1995), these variations depend on physical factors but also on the particular recession model and calculation procedure chosen.

The linear reservoir depletion model of Equation (1) has been largely employed to model recession curves. Several recent studies indicated that non-linear models are more realistic (Moore, 1997, Wittenberg, 1999, Chapman, 2003). Some authors have fitted non-linear equations from the power-law model of Equation (5) to recession discharge data. The value of the exponent b ranged from 0.3 to 0.4 for 10 out of 11 catchments in Chapman (2003). In Wittenberg (1994), a mean value of 0.4 was obtained. Wittenberg (1999) stated that a value of 0.5 is typical for average cases and suggested to calibrate the factor a with b fixed to this value.

2.4 Delineation of homogenous regions

Regionalisation methods usually involve two steps: defining groups of homogeneous stations and applying an information transfer method over the proper groups. As in the case of flood regionalisation, grouping stations provides generally better estimates because stations in the same group are expected to have similar hydrological responses. Certain delineation methods allow one to define geographically contiguous regions. This kind of approach can involve the delineation on the basis of geographic considerations or on the basis of the similitude in residuals obtained by a regressive model (Smakhtin, 2001, Laaha and Blöschl, 2005). In reality, two basins can be hydrologically similar without being geographically close. Other methods allow one to define groups of catchments that are not necessarily contiguous. Delineation is then made on the basis of physiographic and climatic characteristics. Multivariate statistical analysis methods such as weighted cluster analysis are then used (Nathan and McMahon, 1990, Laaha and Blöschl, 2006).

Another promising multivariate approach is canonical correlation analysis (CCA). It has been applied in the field of flood regionalisation by Ouarda et al. (2000, 2001) and it has been proven to be applicable for low flow regionalisation in Ouarda et al. (2005). This method defines for each target station, a different set of homogenous stations (neighbourhoods). This had the advantage of maximising the similarity between the catchments in the neighbourhood and the target site. Neighbourhoods can also be defined using the region of influence approach (Burn, 1990). The neighbourhood approach was found in GREHYS (1996a, b) to be superior to approaches delineating fixed sets of stations for regional flood frequency analysis. Ouarda et al. (2005) obtained similar conclusions for regional low flow frequency analysis. In the present study, estimations with neighbourhoods defined with this method will be obtained. Estimations will also be obtained with all stations without applying a regrouping method, for comparison purposes.

The purpose of the CCA is the assignation of a neighbourhood of hydrologically similar stations to an ungaged catchment, based on physiographical and meteorological characteristics. This method is explained in detail in Ouarda et al. (2001). CCA is used to characterize the relation between two sets of random variables. In the present case, the set $X = (X_1, X_2, ..., X_n)$ contains explanatory variables and the set $Y = (Y_1, Y_2, ..., Y_r)$ contains low flow quantiles where $n \geq r$. Canonical variables W and V are linear combinations of each set: $W = \alpha' Y$ and $V = \beta' X$. In CCA, the values of the coefficients of linear combinations α and β are such that the correlation matrix between the canonical variables is diagonal and the correlations are maximized. Mathematical procedures to obtain vectors α and β are given in Ouarda et al. (2001). Prior to carrying out a CCA, original variables are normalized and standardized. The application of a CCA gives p solution-triplets (λ_i, V_i, W_i), where λ_i is the correlation between V_i and W_i. The number p is the rank of the correlation matrix of variables X and Y, and is generally equal to the number r of hydrological variables. For each gaged drainage basin B_k, $v_{i,k}$ and $w_{i,k}$ denote the computed corresponding values of the canonical variables V_i and W_i. For an ungaged basin B_0, the canonical score v_0 can be computed with values of the basin explanatory variables but the corresponding hydrological canonical score w_0 is unknown.

In order to obtain a simplified expression for the conditional density $w_0 = (W \mid V = v_0)$, the vector of a pair of canonical variables is considered to be jointly 2p-normally distributed, that is:

$$\begin{pmatrix} W \\ V \end{pmatrix} \sim N_{2p}(0, L) \tag{7}$$

where L is the covariance matrix of the canonical variables. L is given by:

$$L = \begin{bmatrix} I_p & \Lambda \\ \Lambda & I_p \end{bmatrix} \quad (8)$$

where Λ is the correlation matrix between W and V, and I_p is the identity matrix of dimensions $p \times p$. If Equation (7) is met, then the conditional density function of W is:

$$(W \mid V = v_0) \sim N_p(\Lambda v_0, I_p - \Lambda^2) \quad (9)$$

It can be shown that distances from the mean position to other stations in the canonical space is a Mahalanobis distance

$$D^2 = (w - \Lambda v_0)'(I_p - \Lambda\Lambda)^{-1}(w - \Lambda v_0) \quad (10)$$

and that this distance has a Chi-2 distribution with p degrees of freedom. This property is used to delineate a homogenous neighbourhood around the mean position where all the stations within the distance given by the Chi-2 quantile for a given probability $1-\phi$ are included in the neighbourhood. To identify the value of ϕ that gives the corresponding optimal neighbourhoods, Ouarda et al. (2001) applied a jackknife resampling procedure. Performance statistics of quantile estimations with the regional model was computed the parameter ϕ varying over a certain range of values. Performance indices were obtained by dividing the performances for each value ϕ by those obtained with all the stations included in the neighbourhood. The optimal value of ϕ was chosen for that these indices are the nearest possible to zero.

3. STUDY METHODOLOGY

3.1 Recession analysis method

The recession analysis method used in this study consists in selecting recession segments from the hydrographs and computing the estimates of the parameters k and a from each segment. With the mean value of these parameters at each station, two recession variables K and A are respectively obtained. Given the large number of stations and years of data available in the present study, recession segments are sorted out with an automated procedure. The one applied here is based on the approach described by Vogel and Kroll (1992) where segments of at least 10 values of decreasing 3-day moving average streamflow during a given period are selected. To minimise surface runoff components, 30 % of the beginning of each segment is subtracted.

Unlike Vogel and Kroll (1992), the dry periods are not defined with fixed dates in this study. An investigation of the hydrographs of the diverse stations of the present study shows a great variability in the starting and ending dates of the dry seasons caused by the high climate variability over the wide study area considered. Indeed, the spring flood season, for instance, starts considerably later for stations located further north. For the spring/summer season, flood peaks happen generally between April and June and lowest flows occur between July and October. For the fall/winter seasons, flood peaks happen generally between October and December and lowest flows occur between January and April. However, a general shift in dry seasons occurs as stations are located at larger latitudes. This spatial variability makes it difficult to define dry seasons with fixed dates. It is done here at a given station by considering the hydrograph portion between the dates where the principal flood peak occurs at $t_{q\max}$ and where the lowest flow occurs at $t_{q\min}$ during the spring/summer and the fall/winter season. The bank storage delaying effect discussed in section 2.2 is considered to be important after the main flood peak. This is handled here by defining a threshold value that should not be exceeded by the recession curve peak. It was decided to set this value to 60% of the flow difference between the maximal flow and the minimal flow plus the minimal flow (i.e. $0.6(q(t_{q\max})-q(t_{q\min}))+q(t_{q\min})$, where $q(t)$ is the outflow at time t).

Simulations of this procedure on the catchments of the present study area have shown that small segments at the very end of dry seasons often produce higher parameter estimates than the other values for the same catchment. These segments were recession curves with peak discharge values very near the minimal flows observed for the corresponding dry season. To minimise the impact of such values on recession variable estimates, recession curves having their peak below the threshold value corresponding to 2.5% of the flow difference between the maximal flow and the minimal flow plus the minimal flow are left out (i.e. $0.025(q(t_{q\max}) - q(t_{q\min})) + q(t_{q\min})$). Both of these threshold criteria restrict the recession segments selected to a narrower window than those obtained by considering all the ones included between $t_{q\min}$ and $t_{q\max}$. Because recessions present time variability, these restrictions decrease the parameter estimates variance. They thus provide more robustness to parameter estimates and will also helps obtain more reliable estimates when parameters are obtained from short length data series.

In this study, Equations (1) and (6) from respectively the linear and the non-linear storage outflow models are both fitted to the recession segments. For the linear model, Equation (4) is solved by a linear regression where the parameters are estimated with least squares. To solve Equation (6), the exponent b is set to 0.5 for all catchments as suggested in Wittenberg (1999). Flow in volume per second (m^3/s) is converted to height over a unit area per day (mm/d). The parameters Q_0 and a are estimated by a least-square method in which the iterative method of Gauss-Newton is used. A line search procedure is used in conjunction with this method to minimise the objective function along the direction of search obtained after each iteration of the Gauss-Newton method.

The hydrograph of the Rimouski River at the station 022003 between March and November 1999 is presented in Figure 1 for illustrative purposes. The hydrograph shows a main spring flood peak caused by snow melt and several smaller peaks caused by important precipitations. The procedure of recession segments delineation described previously was applied for the spring/summer period. Five segments were extracted by the method. The modeling of the first four recession segments is presented in Figure 2 with both recession reservoir models. It can be seen that for these segments, the linear model gives a good approximation although the non-linear model seems to give a slightly better fit to the data.

3.2 Regional models

In this section, the regional models applied to the database are presented. Two general approaches of regional estimation are employed. For the first one, neighbourhoods are delineated with the CCA method presented in the section 2.4. The hydrological variables included in the set of variables Y for the CCA are the low flow quantiles $Q_{7,2}$, $Q_{7,10}$ and $Q_{30,5}$. In the second approach, the regional information transfer is applied to a given target site with all the stations of the study area. The regional information transfer method used for low flow quantile estimation is multiple linear regression where the parameters of the regressive model are estimated with a least-squares procedure.

Overall, six regional models are defined depending on which explanatory variables are used and whether neighbourhoods are delineated or not. For the first three models (CCA, K_CCA, A_CCA), neighbourhoods are delineated with the canonical correlation analysis method. The CCA model includes solely physiographical and meteorological (physio/meteo) variables. The K_CCA and A_CCA models include respectively the variables K and A in addition to the physiographical and meteorological variables. Regional estimations are also obtained with all the stations of the database without delineation of optimal neighbourhoods (ALL, K_ALL, A_ALL). In this case, the model ALL includes solely the physiographical and meteorological variables and the K_ALL and A_ALL models include also variables K and A respectively. These six models are applied to all stations of the study area to estimate the quantiles corresponding to both low flow seasons.

To choose the explanatory variables that are most important to explain low flow statistics, correlations between normalized explanatory variables and normalized specific quantiles are computed. A specific quantile is the value of a quantile divided by the catchment area. Where high correlations are observed, this indicates that the variables involved are important explanatory variables. A table of the correlations between explanatory variables is also computed to investigate dependence among them.

Optimal neighbourhoods are then delineated for models where a CCA is involved. The optimisation procedure explained in section 2.3 has been modified in this study. The reason is that, when a neighbourhood is defined with a value of the parameter ϕ, it may happen that the number of stations in the neighbourhood is not large enough to be able to carry out an appropriate multiple regression. The modified procedure includes, in the neighbourhood of a target site, the *m* stations with the lowest Mahalanobis distance. This ensures that adequate estimations will be obtained at all stations of the study area. A jackknife resampling procedure, similar to the one presented in section 2.3, is applied to identify the optimal value of *m*.

3.3 Performance criteria

To access performances of the regional models, a jacknife resampling procedure is performed. Each gaged station is successively considered ungaged and is removed from the data base. A regional model is then applied to obtain an estimate of the quantiles at this target site with the remaining stations. This operation is repeated for all stations of the data base. Two performance measures are then calculated: the relative bias and the relative root mean square error, defined by:

$$rBIAS = \frac{1}{N}\sum_{i=1}^{N}\left[(\hat{q}_i - q_i)/q_i\right]\times 100 \qquad (11)$$

$$rRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[(\hat{q}_i - q_i)/q_i\right]^2}\times 100 \qquad (12)$$

where q_i and $\hat{q}_i$ are respectively the at-site and the regional estimates of the quantile at the catchment *i* and *N* is the number of gaged stations.

Performance criteria for models with a recession variable were adapted to represent the level of performance that can be achieved when recession variables are estimated with short data series. For this, the recession parameters are computed on a moving average window of *n* years for all the gaged stations with *n* ranging from 1 to 5 years. For a given station *i* with y_i years of recorded data, $P_{i,n} = y_i - n + 1$ estimations of each quantile are obtained. It is assumed that only years with a complete record during the dry season are considered. Usually, at least one segment is extracted by the selection procedure at each season. The regional models K_CCA, A_CCA, K_ALL and A_ALL are then applied to all gaged stations but with the value of the recession parameter at the target site *i* being successively equal to each one of the $P_{i,n}$ estimates corresponding to each value of *n*. Performance statistics corresponding to each model are then computed for each value of *n*.

Performance criteria are adapted to account for the fact that several estimations of quantiles are obtained at each target station. The new performance criteria are given, for each value of *n*, as follows:

$$rBIAS_n = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{P_{i,n}}\sum_{j=1}^{P_{i,n}}\left[(\hat{q}_{ij} - q_i)/q_i\right]\right)\times 100 \tag{13}$$

$$rRMSE_n = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{P_{i,n}}\sum_{j=1}^{P_{i,n}}\left[(\hat{q}_{ij} - q_i)/q_i\right]^2\right)}\times 100 \tag{14}$$

where $\hat{q}_{ij}$ is the regional estimation obtained with the recession variable estimate *j* at the catchment *i*. In the case where there is only one estimation of the quantile q_i, Equations (13) and (14) reduce to Equations (11) and (12).

4. STUDY AREA

The regional estimation models described in section 3.2 are applied to 85 gaged stations in the province of Québec (Canada). The locations of the gaged stations are presented in Figure 3. The diameters of the circles are proportional to the catchment area which varies from 572 to 96,600 km^2 with a median value of 3077 km^2. The stations cover a large area in the southern part of the province of Quebec (Canada). They are located between the 45°N and 55°N. The largest catchments are located in the northern part of the region of study. The average flow record size is 32 years of data. Winter mean temperatures for the study area range between -10°C in the south to -21°C in the north, and summer mean temperatures range between 20°C in the south to 12°C in the north. Annual hydrographs in the area are characterized by an important spring flood caused by snowmelt. The summer dry season that follows is subject to some streamflow peaks caused by spring and summer precipitations. Important rainstorms usually occur during the fall flood season. This is followed by a winter dry season caused by the lack of liquid precipitations as well as the soil being often frozen.

The hydrological, physiographical and meteorological variables used in the present case study came from a frequency analysis study by Ouarda et al. (2005). In this study, only stations with at least 10 years of data and with a natural regime have been selected. These stations were submitted to a stationarity test (Kendall test), an independence test (Wald-Wolfowitz test), a homogeneity test for the mean (Wilcoxon test) and a homogeneity test for the variance (Levene test). Stations that did not pass one of these tests were removed from the database. In the same study, hydrologic data were fitted with appropriate statistical distributions. The selection of the distribution that gives the best fit to the hydrological data for each station was based on the Bayesian information criterion. Distributions considered include the GEV, Gumbel, Weibull, the two and three parameter Lognormal, Gamma, Pearson, Log-Pearson III and Generalized Pareto distributions. The same hydrological record is used in the present study to compute the recession variables with the method presented in section 3.1 for both winter and summer dry seasons. The explanatory variables available for the stations of this case study are presented in Table 1. Recession variables A and K computed for each season are denoted A_S and K_S for the summer dry season, and A_W and K_W for the winter season.

Variables were normalized before use in the regional models. To access the normality of the variables, the Filliben's statistic was computed. It gives the degree of linearity of the variable on a normal probability plot. Its value ranges between zero and one with higher values indicating a better normality of the data. Observed variables that are not normal were normalized with the appropriate transformation (natural logarithm or a power transformation). Proper transformations applied to explanatory variables are shown in Table 1.

5. RESULTS

In the present section, the results of the investigation to identify the important explanatory variables to be included in the regional models are first presented. This is followed by a presentation and analysis of the performances of the regional models applied to the catchments of the study area.

5.1 Important explanatory variables

The correlations between the explanatory variables and the specific quantiles $SQ_{7,2}$, $SQ_{7,10}$ and $SQ_{30,5}$ from both seasons were computed and the results are presented in Table 2. These results are then used to select important explanatory variables. Given the number of stations in the study area, the degree of significance for a correlation is 0.28 at a confidence level of 1% using the t-test. In Table 2, correlations higher than this value are presented in bold characters. It is assumed that the drainage area variable (AREA) is always included in the models because of its prominent importance. The correlations between the explanatory variables are also computed and the results are presented in Table 3. The results are used to identify dependant variables. Correlations higher than or equal to 0.8 are identified in Table 3.

For the summer low flows, the physio/meteo variables that are most correlated with the specific quantiles are PLAC, PTMS, PTMA, NDH27, DDH13, DDBZ and LAT. Within this set, some strongly correlated variables can be discarded. The variables NDH27, DDH13, DDBZ, LAT are highly correlated together as are the precipitation variables PTMS and PTMA. Only one variable in each of these sets needs to be retained. The temperature variable selected is DDH13 since the other variables of the set are more correlated to other explanatory variables such as the precipitation variables and PLAC. From the set of precipitation variables, PTMA is selected because it is more correlated to the specific quantiles than PTMS. The models without recession variable (ALL and CCA) include then the variables AREA, PLAC, DDH13, PTMA and MCN.

The recession variables are among the most correlated with the summer specific quantiles. Those computed from the summer seasons have higher correlations than those from the winter seasons and the variable A_x has higher correlations than K_x for a given season x. The recession variables from summer recessions are then preferred because they are more correlated. The variable A_s is then included in the models A_ALL and A_CCA and the variable K_s in the models K_ALL and K_CCA. All recession variables are strongly correlated to PLAC. Consequently, the variable PLAC is discarded in models that include a recession variable. This can be explained by the fact that lakes are a component of the baseflow by continuing to feed the stream during a recession period.

For the winter low flows, the correlations between the physio/meteo explanatory variables and the specific quantiles are rather weak except for PLAC and AREA. This indicates that additional variables have only small impacts on estimations. Nevertheless, the temperature variable DDBZ and the precipitation variable PTMA have significant correlations and are included in the models ALL and CCA along with AREA and PLAC.

The correlations of the specific winter quantiles with the recession variables are superior to those with the physio/meteo explanatory variables. However, they are significantly lower than the correlations between the recession variables and the specific summer quantiles. The highest correlation values to the specific winter quantiles are obtained with the winter recession variable A_w and the lowest are obtained with the variable K_w. As with the summer quantiles, the recession variable A_w, extracted from the same season as those of the winter specific quantiles considered, is more correlated to these quantiles than the variable A_s extracted from a different season. However, this is not true for the variable K_x for which the variable K_s is slightly more correlated to the winter specific quantiles than K_w. This indicates that the linear model is probably weaker to model winter recessions. The variable A_w is then included in the models A_ALL and A_CCA and the variable K_s in the models K_ALL and K_CCA. Again, the variable PLAC is discarded in models that include a recession variable. A summary of the variable configurations for each regional model is presented in Table 4. In order to estimate the impact of using the variable K_s instead of the variable K_w for winter quantiles, it was also decided to produce winter quantile estimates based on the inclusion of the variable K_w. Two new models are then added and are denoted by Kw_ALL and Kw_CCA.

5.2 Performance evaluation

Table 5 and 6 compare the performances of the quantile estimations obtained by the application of the different regional models. Results show that adding the recession variable A or K to the model CCA or ALL always significantly improve the performances. Lower

$rRMSE$ values are always obtained with the variable A instead of the variable K while the $rBIAS$ values are often lower with A. These performance differences are more important for the winter quantile estimates than for the summer quantile estimates. Results show also that the delineation of neighbourhoods with the CCA usually improves the performances of the models for summer quantiles. This is not the case for winter quantiles where the performances are generally very similar. This seems to indicate that the overall level of homogeneity in the region of study is higher for winter low flows. This is probably also due to the lack of strongly correlated variables with the winter specific quantiles. The overall best performances are obtained with the model A_CCA for summer quantile estimates and with A_ALL and A_CCA for winter quantile estimates. In Table 6, the comparison of the performances obtained with the variable K_s and K_w shows generally similar values with a slight edge for the model using K_s values.

Figures 4 and 5 show the performances of the regional models when the recession variables at target sites are estimated for a given number n of years. The performances obtained for the ungaged case with the models ALL and CCA are also shown for comparison purposes. In that case, the figures show constant values independently of the number n. Results indicate that the $rRMSE_n$ values of the models with a recession variable (K_CCA, A_CCA, K_ALL, and A_ALL) decrease as n increases whereas the $rBIAS_n$ values gradually increases as n increases.

To access the improvement obtained by the use of a recession parameter, the performances of the models with a recession variable (K_CCA, A_CCA, K_ALL, and A_ALL) are compared with the performances obtained by their corresponding ungaged models CCA and ALL. The $rBIAS_n$ values of models with a recession variable are always better for any value of n for quantile estimations of both dry seasons. These models show also better performances for the $rRMSE_n$ values even with just one year of hydrological data for the summer quantile estimations. For the models that include the winter recession variable A_W (A_CCA and A_ALL), it takes 2 or 3 years to get better $rRMSE_n$ values for winter quantile estimations and it takes 2 years for the models with the summer recession variable K_S (K_CCA and K_ALL) to get better $rRMSE_n$. The performances of models with K_S are better than those with A_W for the first two years. The fact that it takes more years to gain valuable estimations with the recession variable A_W indicates a higher variability in the individual estimates of the parameter a than those of the parameter k for the winter period. The overall results indicate that a valuable estimate of a recession variable can be obtained with only a few years of data.

The fact that the $rBIAS_n$ increases as n increases seems to be related to the estimates of log-transformed parameters k and a that are obtained with increasing values of n. For illustrative purposes, we also present in Figure 6 the average means of log-transformed parameters k and a and the average variances of parameters k and a obtained at each station during the spring/summer season as a function of the number of years n. As the parameter estimates increase, the regression based quantile estimates increase also and the quantile estimation bias increases. This counterintuitive increase can be explained by the fact that for a given concave function $\varphi(X)$, which is the logarithm function in the case of this study, the distribution of the variable $Y = \varphi(X)$ is stretched-in by this function. This tends to produce lower values of the expected value of Y as the variance of X grows because the distortion caused by the function $\varphi(X)$ becomes more important. This is what happened with the log-transformed parameter estimates of this study: as the variance of parameter estimates decreases, the mean of log-transformed parameter estimates increases. What is mainly of interest to the hydrologist is the quantile estimation RMSE. Even though quantile estimates are more biased as n increases, their RMSE decreases. This is due to the decrease in estimation variance. For illustrative purposes, we present in Figure 7 the mean of mean square errors (MSE), biases^2 and variances of summer quantile $Q_{7,10}$ estimates obtained with the regional model K_CCA as a function of the number of years n.

6. CONCLUSIONS AND FUTURE WORK

In this paper, a regional model to estimate low flow statistics was developed. This model uses estimates of recession parameters as explanatory variables. Two different parameters were estimated: the recession constant k from a linear reservoir model and the recession coefficient a from a power-law reservoir model where the exponent b has been set to the typical value of 0.5. These parameters were estimated for both summer and winter dry seasons.

The investigation of the important variables that explain low flow statistics through correlations between explanatory variables and low flow specific quantiles leads to the conclusion that the recession variables are important explanatory variables and correlations are higher with variables from the recession coefficients a considering the power-law reservoir model. Generally, recession variables computed for a given dry season are more appropriate for the estimation of low flow statistics during that same season. The study results clearly demonstrated that the inclusion of a recession variable in a regional model improves the performance of the regional estimator. The performances obtained are in all cases better with models that include a recession coefficient from the non-linear reservoir model.

Recession parameters can be obtained only for gaged catchments. However, it is possible to estimate these parameters with short hydrographs. This paper aimed also to evaluate the performances obtained with recession parameters estimated with hydrographs of a few years only. Results of the application of regional models with hydrograph lengths ranging from one to five years have shown that performances converge rapidly to those obtained when the parameters are estimated from a long data record. These performances are better than those obtained with ungaged models even with only one year of data on average in the case of summer low flows. In the case of winter low flows, better performances than those of ungaged models are obtained within two or three years. This shows that it is possible to combine local hydrological information with regional information by estimating a recession parameter at a partially gaged station and using it in a regional model.

Considering these results, it is of interest to improve methods to estimate recession parameters. Improvement can result principally from the selection of a proper reservoir model and from the recession analysis method. Better reservoir models, in agreement with the real reservoir storage-outflow relationship should be developed. For instance, the exponent *b* of the power-law model was set arbitrary to 0.5 in the present study but no study was been done on the validity of this assumption for the catchments of the study area. Other reservoir models, such as multiple reservoir models that consider various loss and gain sources that affect the streamflow could be used. Other improvements can come from the development of recession analysis methods. For instance, recession segments should be representative of baseflow recession discharges, i.e. should represent portions of flow that are free of surface flow and interflow. Other methods for the delineation of homogenous regions should be considered. For instance, methods based on seasonality characteristics should be considered.

ACKNOWLEDGMENTS

The financial support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) is gratefully acknowledged

REFERENCES

Arihood, L.D. and Glatfelter, D.R., 1991. Method for estimating low-flow characteristics of ungaged streams in Indiana. U.S. Geological Survey Water-Supply Paper No. 2372, 22 pp.

Bingham, R.H., 1986. Regionalization of low-flow characteristics of Tennessee streams. U.S. Geological Survey Water-Resources Investigation Report 85-4191, 63 pp.

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

Brandes, D., Hoffmann, J.G. and Mangarillo J.T., 2005. Base flow recession rates, low flows, and hydrologic features of small watersheds in Pennsylvania, USA. Journal of the American Water Resources Association 41 (5), 1177-1186.

Chapman, T.G., 1999. A comparison of algorithms for stream flow recession and baseflow separation. Hydrological Processes 13, 701-714.

Chapman, T.G., 2003. Modelling stream recession flows. Environmental Modelling & Software 18, 683-692.

Chen, X.D., Chen, Q.Y. and Chen, X.H., 2006. Simulation of baseflow accounting for the effect of bank storage and its implication in baseflow separation. Journal of Hydrology 327, 539-549.

Demuth, S. and Hagemann, I., 1994. Estimation of flow parameters applying hydro-geological area information. IAHS Publ. 221, 151-157.

Groupe de recherche en hydrologie statistique (GREHYS), 1996a. Presentation and review of some methods for regional flood frequency analysis. Journal of Hydrology 186, 63-84.

Groupe de recherche en hydrologie statistique (GREHYS), 1996b. Inter-comparison of regional flood frequency procedures for Canadian rivers. Journal of Hydrology 186, 85-103.

Hall, F.R., 1968. Base-flow recessions - a review. Water Resources Research 4 (5), 973-983.

Hammond, M. and Han, D., 2006. Recession curve estimation for storm event separations. Journal of Hydrology 330, 573-585.

Hamza, A., Ouarda, T.B.M.J., Durrans, S.R. and Bobée, B., 2001. Development of tail and scale invariance models for the regional estimation of low flows. Canadian Journal of Civil Engineering 28, 291-304.

Kroll, C., Luz, J., Allen, B. and Vogel, R.M., 2004. Developing a watershed characteristics database to improve low streamflow prediction. Journal of Hydrologic Engineering 9 (2), 116-125.

Laaha, G. and Blöschl, G., 2006. A comparison of low flow regionalisation methods - catchment grouping. Journal of Hydrology 323, 193-214.

Moore, R.D., 1997. Storage-outflow modelling of streamflow recessions, with application to a shallow-soil forested catchment. Journal of Hydrology 198, 260-270.

Nathan, R.J. and McMahon, T.A., 1990a. Evaluation of automated techniques for base flow and recession analyses. Water Resources Research 26 (7), 1465-1473.

Nathan, R.J. and McMahon, T.A., 1990b. Identification of homogeneous regions for the purposes of regionalisation. Journal of Hydrology 121, 217-238.

Ouarda, T.B.M.J., Haché, M., Bruneau, P. et Bobée B., 2000. Regional flood peak and volume estimation in a northern Canadian basin. ASCE, Journal of Cold Regions Engineering 14 (4), 176-191.

Ouarda, T.B.M.J., Girard, C., Cavadias, G.S. and Bobée, B., 2001. Regional flood frequency estimation with canonical correlation analysis. Journal of Hydrology 254, 157-173.

Ouarda, T.B.M.J., Jourdain, V., Gignac, N., Gingras, H., Herrara, H. et Bobée, B., 2005. Développement d'un modèle hydrologique visant l'estimation des débits d'étiage pour le Québec habité. INRS-ETE, Rapport de recherche No. R-684-f1, 174 pp.

Smakhtin, V.U., 2001. Low flow hydrology: a review. Journal of Hydrology 240, 147-186.

Tallaksen, L.M., 1989. Analysis of time variability in recessions. IAHS Publ. 187, 85-96.

Tallaksen, L.M., 1995. A review of baseflow recession analysis. Journal of Hydrology 165, 349-370.

Vogel, R.M. and Kroll, C.N., 1992. Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. Water Resources Research 28 (9), 2451-2458.

Vogel, R.M. and Kroll, C.N., 1996. Estimation of baseflow recession constants. Water Resources Management 10, 303-320.

Wittenberg, H., 1994. Nonlinear analysis of flow recession curves. IAHS Publ. 221, 61-67.

Wittenberg, H., 1999. Baseflow recession and recharge as nonlinear storage processes. Hydrological Processes 13, 715-726.

Wittenberg, H., 2003. Effects of season and man-made changes on baseflow and flow recession: case studies. Hydrological Processes 17, 2113-2123.

LIST OF SYMBOLS

k	recession constant
a	recession coefficient in the power law reservoir model
b	exponent in the power law reservoir model
Q	catchment outflow
I	catchment inflow
S	catchment water storage
W	hydrological canonical variable
V	explanatory canonical variable
Y	set of hydrological variables
X	set of explanatory variables
Λ	correlation matrix between W and V
ϕ	probability that a value is exceed by a Chi-2 random variable with p degrees of freedom
D^2	Mahalanobis distance

Table 1
Explanatory variables available for all stations of the study area

Explanatory variable	**Units**	**Notation**	**Transformation**
Drainage area	km^2	AREA	LN
Mean slope of the drainage area	deg.	MSLP	LN
Percentage of forested area	%	PFOR	X^4
Percentage of lakes	%	PLAC	$X^{1/2}$
Mean annual precipitation	mm	PTMA	none
Mean summer-fall precipitation	mm	PTMS	none
Mean annual degree-days below zero	deg.d	DDBZ	X^{-1}
Mean annual degree-days higher than 13 °C	deg.d	DDH13	none
Number of days where temperatures are higher than 27 °C	d	NDH27	none
Mean curve number (soil characteristic)	-	MCN	none
Latitude of the gaging station	deg.	LAT	none
Recession constant (summer season)	d	K_s	LN
Recession coefficient (summer season)	$mm^{1/2}d^{1/2}$	A_s	LN
Recession constant (winter season)	d	K_w	LN
Recession coefficient (winter season)	$mm^{1/2}d^{1/2}$	A_w	LN

Table 2
Correlations between normalized specific quantiles and normalized explanatory variables

Explanatory variables	**summer**			**winter**		
	$SQ_{7,2}$	$SQ_{7,10}$	$SQ_{30,5}$	$SQ_{7,2}$	$SQ_{7,10}$	$SQ_{30,5}$
PLAC	**0.69**	**0.67**	**0.64**	**0.55**	**0.65**	**0.55**
MCN	**-0.58**	**-0.56**	**-0.54**	-0.15	-0.24	-0.12
PTMS	**-0.61**	**-0.61**	**-0.56**	-0.15	-0.28	-0.14
NDH27	**-0.68**	**-0.69**	**-0.72**	-0.07	-0.13	-0.05
MSLP	-0.02	-0.01	0.02	-0.25	-0.28	-0.25
PTMA	**-0.69**	**-0.67**	**-0.64**	-0.21	**-0.36**	-0.22
DDBZ	**-0.87**	**-0.86**	**-0.86**	**-0.29**	**-0.44**	**-0.29**
PFOR	0.03	0.02	0.03	-0.26	-0.24	**-0.29**
DDH13	**-0.82**	**-0.82**	**-0.83**	-0.15	-0.26	-0.13
LAT	**0.82**	**0.81**	**0.81**	0.11	0.25	0.11
K_s	**0.82**	**0.81**	**0.76**	**0.57**	**0.68**	**0.56**
K_w	**0.78**	**0.77**	**0.71**	**0.55**	**0.67**	**0.53**
A_s	**0.89**	**0.88**	**0.84**	**0.57**	**0.67**	**0.55**
A_w	**0.82**	**0.81**	**0.76**	**0.68**	**0.78**	**0.66**

Bold values correspond to those higher than 0.28.

Table 3
Correlations between normalized explanatory variables

	PLAC	MCN	PTMS	NDH27	MSLP	PTMA	DDBZ	PFOR	DDH13	LAT	K_S	K_W	A_S	A_W
AREA	0.71	-0.22	-0.52	-0.38	-0.29	-0.69	-0.72	-0.28	-0.51	0.59	0.75	0.65	0.75	0.66
PLAC		-0.23	-0.56	-0.27	-0.41	-0.69	-0.71	-0.40	-0.48	0.58	**0.84**	**0.80**	**0.83**	**0.82**
MCN			0.50	0.42	-0.25	0.49	0.49	-0.39	0.54	-0.50	-0.47	-0.49	-0.52	-0.48
PTMS				0.49	0.04	**0.88**	0.69	-0.07	0.67	-0.76	-0.72	-0.74	-0.72	-0.67
NDH27					-0.37	0.40	0.67	-0.22	**0.93**	-0.74	-0.42	-0.35	-0.53	-0.36
MSLP						0.21	0.11	0.34	-0.23	0.04	-0.26	-0.25	-0.19	-0.26
PTMA							0.79	0.08	0.61	-0.78	**-0.81**	-0.79	**-0.80**	-0.74
DDBZ								0.03	**0.82**	**-0.85**	-0.77	-0.76	**-0.82**	-0.75
PFOR									-0.16	-0.03	-0.18	-0.05	-0.11	-0.08
DDH13										**-0.87**	-0.60	-0.55	-0.69	-0.55
LAT											0.72	0.66	0.77	0.62
K_S												**0.92**	**0.99**	**0.93**
K_W													**0.91**	**0.98**
A_S														**0.92**

Bold values correspond to those higher than or equal to 0.8.

Table 4
Configuration of explanatory variables used in each model

Model	Explanatory variables	
	Summer	**Winter**
ALL, CCA	AREA, PLAC, MCN, PTMA, DDH13	AREA, PLAC, PTMA, DDBZ
K_ALL, K_CCA	AREA, MCN, PTMA, DDH13, K_s	AREA, PTMA, DDBZ, K_s
A_ALL, A_CCA	AREA, MCN, PTMA, DDH13, A_s	AREA, PTMA, DDBZ, A_w

Table 5
Performances (%) of the different regional models (summer quantiles)

	ALL	K_ALL	A_ALL	CCA	K_CCA	A_CCA
rBIAS (%)						
$Q_{7,2}$	5.6	4.8	4.0	4.6	2.4	**1.9**
$Q_{7,10}$	8.2	6.6	6.0	9.0	**3.4**	**3.4**
$Q_{30,5}$	5.6	5.2	4.7	5.9	2.8	**2.6**
rRMSE (%)						
$Q_{7,2}$	35.7	31.7	29.0	32.3	29.3	**27.1**
$Q_{7,10}$	44.3	37.9	36.5	43.6	34.8	**34.4**
$Q_{30,5}$	35.2	32.9	31.5	34.3	30.8	**29.7**

Bold values correspond to best performances.

Table 6
Performances (%) of the different regional models (winter quantiles)

	ALL	K_ALL	Kw_ALL	A_ALL	CCA	K_CCA	Kw_CCA	A_CCA
rBIAS (%)								
$Q_{7,2}$	2.9	2.2	2.4	1.4	2.7	1.4	1.9	**1.1**
$Q_{7,10}$	3.8	3.0	3.0	1.8	3.9	1.8	2.3	**1.3**
$Q_{30,5}$	3.1	2.5	2.7	1.7	3.1	1.7	2.2	**1.4**
rRMSE (%)								
$Q_{7,2}$	24.2	21.9	22.0	**17.3**	23.8	21.9	22.3	17.7
$Q_{7,10}$	29.4	26.6	25.7	**20.1**	29.1	25.9	25.9	20.3
$Q_{30,5}$	26.1	24.2	24.1	**19.6**	26.5	24.1	24.4	19.9

Bold values correspond to best performances.

FIGURE CAPTIONS

Figure 1. Hydrograph of the 022003 station on the Rimouski River (dotted lines are recession segments).

Figure 2. The first four recession segments in Figure 1 fitted with the linear model (dashed line) and the non-linear model (dotted line).

Figure 3. Location of catchments used in this study. Center of circle denotes catchment gauge location. The diameter of the circle is proportional to catchment area.

Figure 4. $rBIAS_n$ and $rRMSE_n$ for summer quantile estimates by the application of the regional models when the recession variables are computed on a given number of years. Quantiles estimated are: $Q_{7,2}$ (a, b); $Q_{7,10}$ (c, d) and $Q_{30,5}$ (e, f).

Figure 5. $rBIAS_n$ and $rRMSE_n$ for winter quantile estimates by the application of the regional models when the recession variables are computed on a given number of years. Quantiles estimated are: $Q_{7,2}$ (a, b); $Q_{7,10}$ (c, d) and $Q_{30,5}$ (e, f).

Figure 6. Average means of log-transformed parameter estimates k and a and average variances of parameter estimates k and a obtained at each station during the spring/summer season as a function of the number of years n. (a) Mean of means of $\ln(k)$, (b) mean of variances of k, (c) mean of means of $\ln(a)$ and (d) mean of variances of a.

Figure 7. Mean of mean square errors (MSE), biases^2 and variances of estimates of summer quantiles $Q_{7,10}$ obtained at each station with the regional model K_CCA as a function of the number of years n.

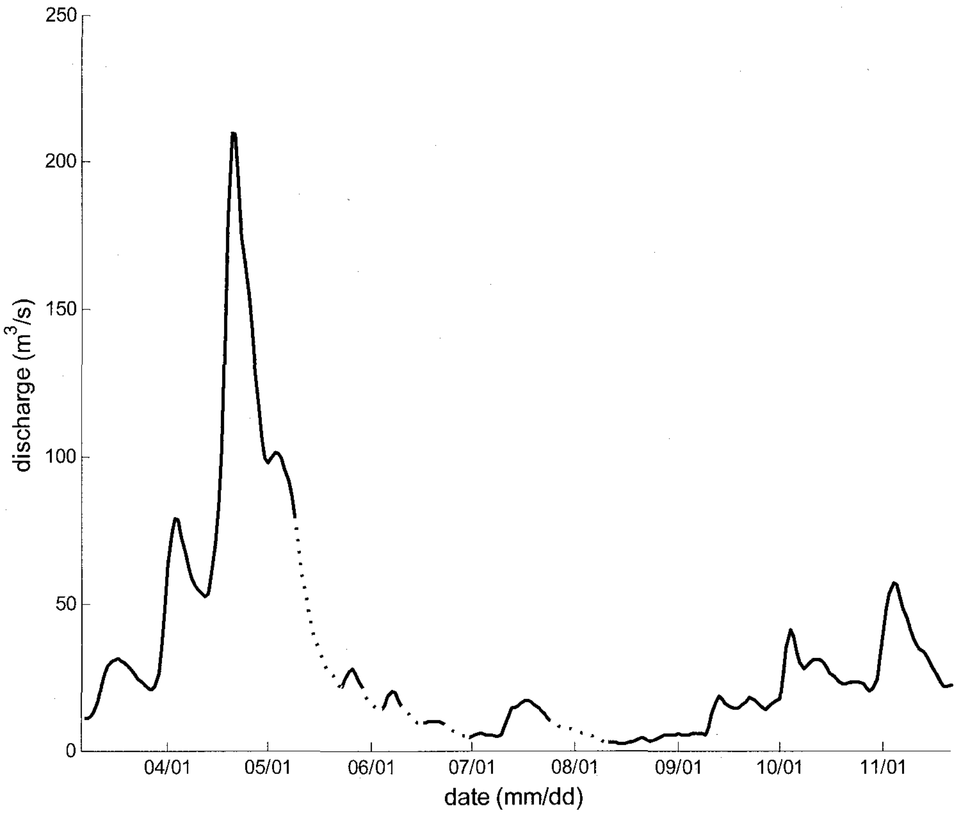

Figure 1. Hydrograph of the 022003 station on the Rimouski River (dotted lines are recession segments).

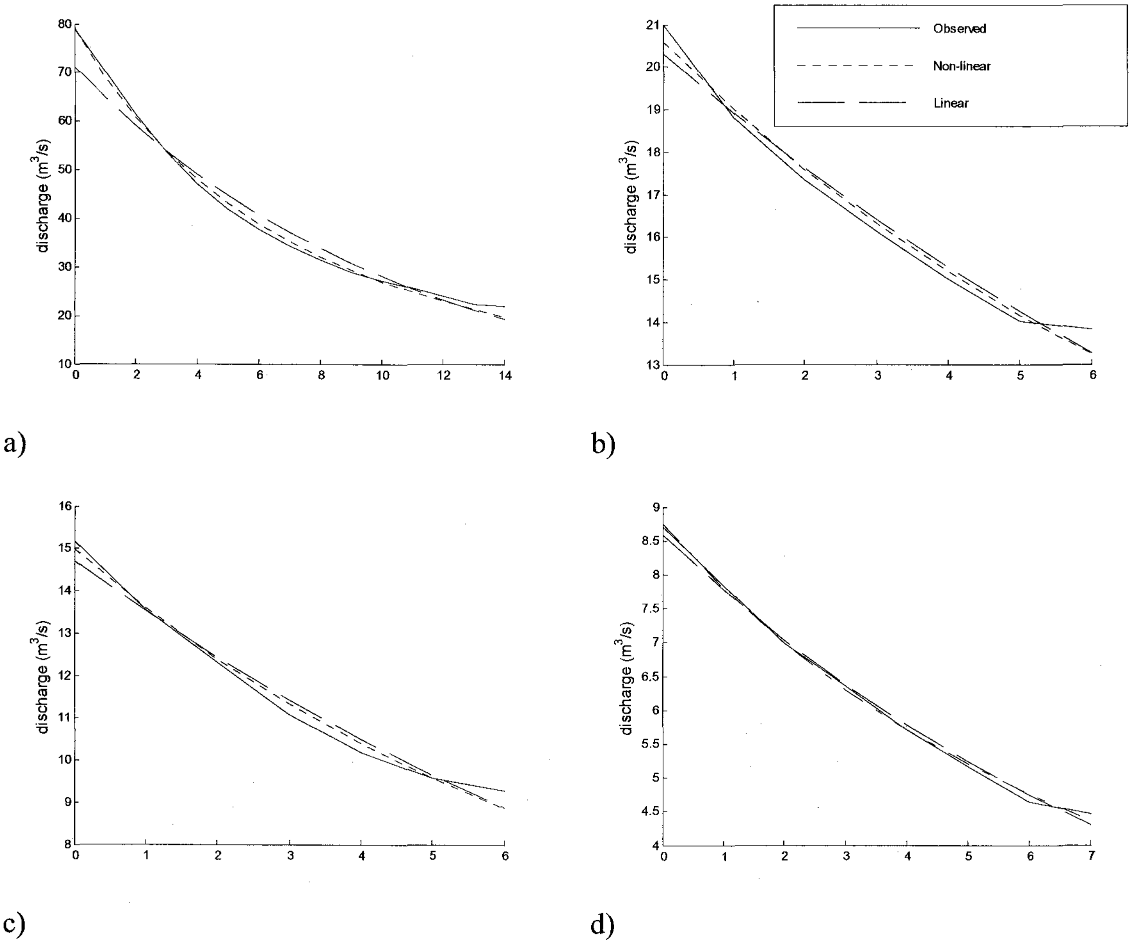

Figure 2. The first four recession segments in Figure 1 fitted with the linear model (dashed line) and the non-linear model (dotted line).

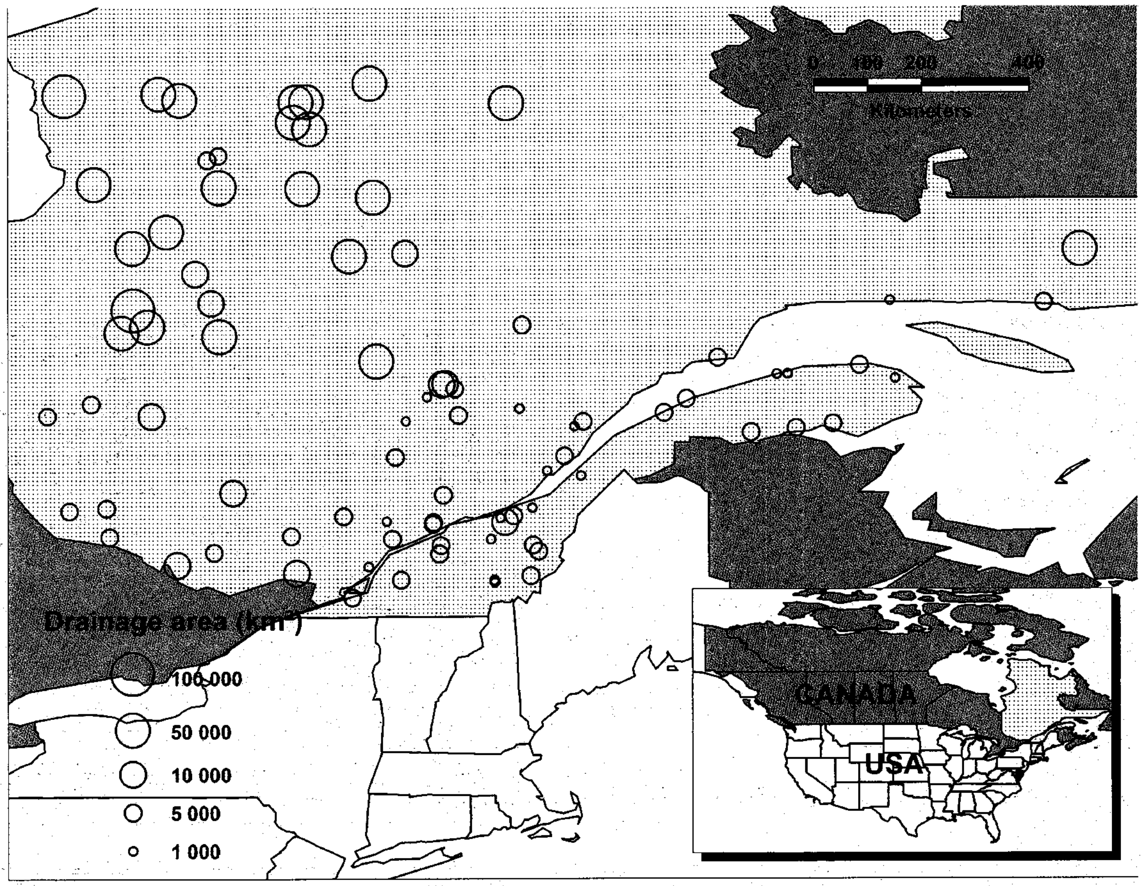

Figure 3. Location of catchments used in this study. Center of circle denotes catchment gauge location. The diameter of the circle is proportional to catchment area.

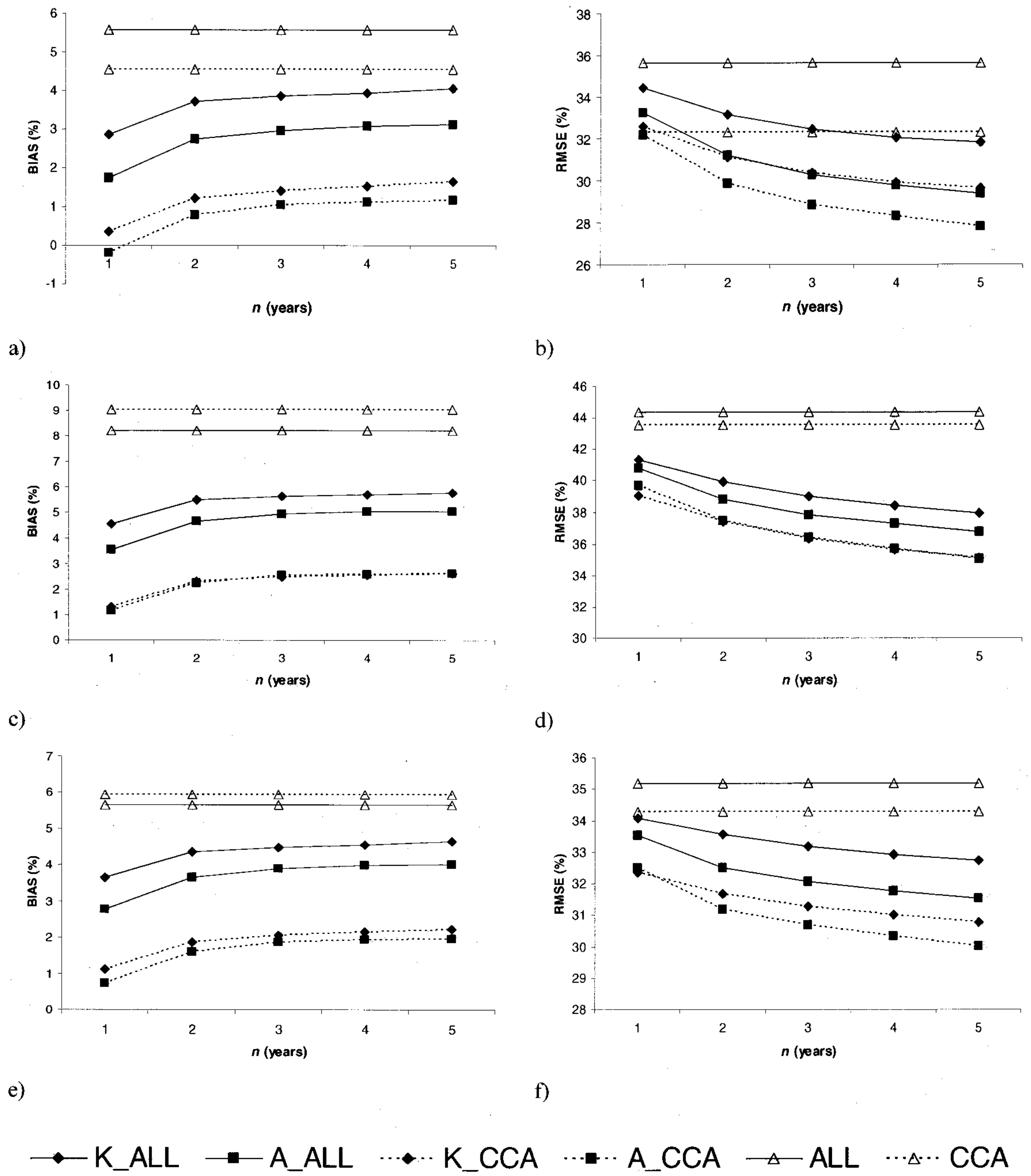

Figure 4. $rBIAS_n$ and $rRMSE_n$ for summer quantile estimates by the application of the regional models when the recession variables are computed on a given number of years. Quantiles estimated are: $Q_{7,2}$ (a, b); $Q_{7,10}$ (c, d) and $Q_{30,5}$ (e, f).

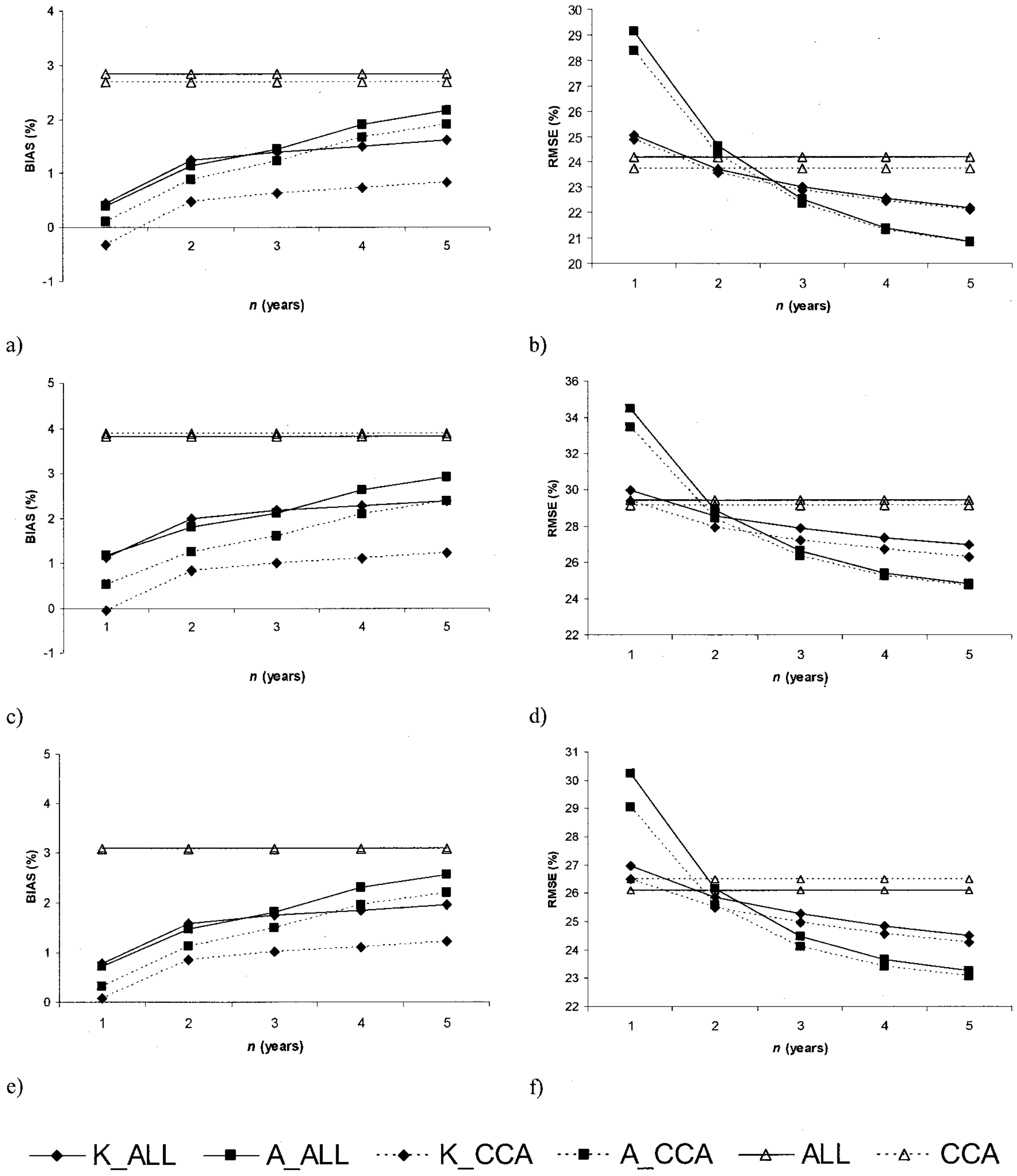

Figure 5. $rBIAS_n$ and $rRMSE_n$ for winter quantile estimates by the application of the regional models when the recession variables are computed on a given number of years. Quantiles estimated are: $Q_{7,2}$ (a, b); $Q_{7,10}$ (c, d) and $Q_{30,5}$ (e, f).

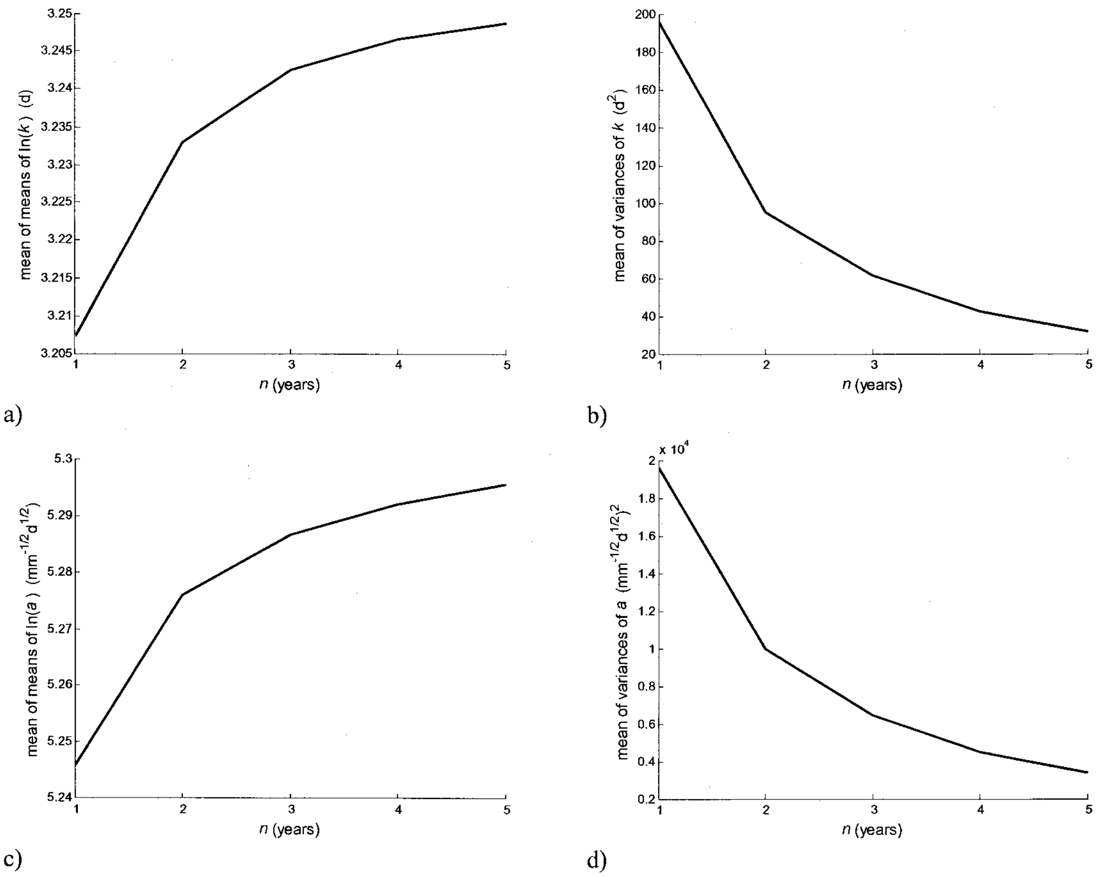

Figure 6. Average means of log-transformed parameter estimates *k* and *a* and average variances of parameter estimates *k* and *a* obtained at each station during the spring/summer season as a function of the number of years *n*. (a) Mean of means of ln(*k*), (b) mean of variances of *k*, (c) mean of means of ln(*a*) and (d) mean of variances of *a*.

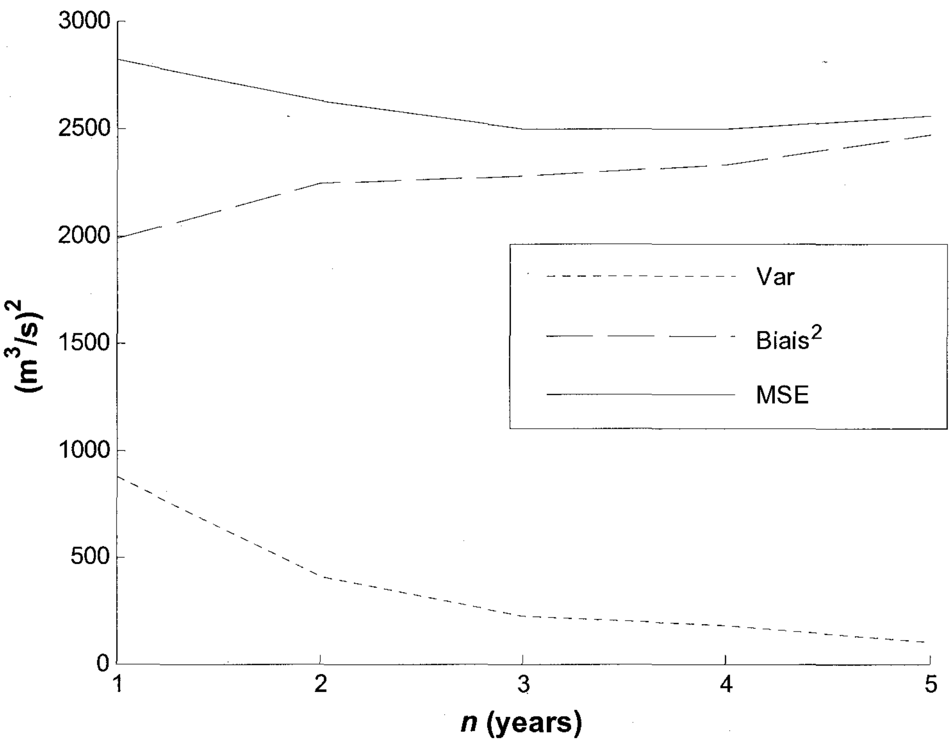

Figure 7. Mean of mean square errors (MSE), biases^2 and variances of estimates of summer quantiles $Q_{7,10}$ obtained at each station with the regional model K_CCA as a function of the number of years *n*.

CONCLUSION

L'article présenté dans ce mémoire apporte des contributions importantes au domaine de recherche relatif à l'utilisation de variables hydrologiques de récession dans des modèles régionaux d'estimation de variables d'étiage. Une première contribution importante de l'article est d'avoir inclus un paramètre d'un modèle non-linéaire du réservoir dans un modèle régional. Cette approche est originale, puisque dans les études consultées, seul le modèle linéaire était considéré. Les résultats obtenus ont démontré en général une amélioration des performances avec ce paramètre comparativement à celles avec un paramètre du modèle linéaire. Une deuxième contribution importante est d'avoir évalué les performances de modèles régionaux lorsque les variables de récession sont estimées sur de courtes périodes de temps. Les résultats obtenus ont permis de conclure que l'information locale contenue dans l'estimation d'un paramètre de récession pour un bassin partiellement jaugé peut être utilisée en combinaison avec de l'information régionale. Une autre contribution est aussi d'avoir observé d'après les résultats de cette étude que les variables de récession calculées durant une période de sécheresse donnée expliquent mieux les épisodes de faibles débits ayant lieu lors de cette même période de sécheresse. Ce résumé des contributions permet de conclure que les résultats présentés dans l'article scientifique sont importants pour l'avancement des connaissances dans le domaine scientifique étudié.