

Record Number: 8170
Author, Monographic: Mathier, L//Roy, R.//Perron, H.//Bobée, B.
Author Role:
Title, Monographic: Estimation régionale des débits de crue : détermination des régions homogènes par la méthode des régions d'influence
Translated Title:
Reprint Status:
Edition:
Author, Subsidiary:
Author Role:
Place of Publication: Québec
Publisher Name: INRS-Eau
Date of Publication: 1994
Original Publication Date: Juillet 1994
Volume Identification:
Extent of Work: 42
Packaging Method: pages incluant un annexe
Series Editor:
Series Editor Role:
Series Title: INRS-Eau, Rapport de recherche
Series Volume ID: 420
Location/URL:
ISBN: 2-89146-416-8
Notes: Rapport annuel 1994-1995
Abstract: Rapport rédigé pour le projet de recherche CRSNG stratégique méthodologie d'estimation régionale des débits de crue: application à la région Québec-Ontario
5.00\$
Call Number: R000420
Keywords: rapport/ ok/ dl

**ESTIMATION RÉGIONALE DES DÉBITS DE
CRUE: DÉTERMINATION DES RÉGIONS
HOMOGÈNES PAR LA MÉTHODE DES
RÉGIONS D'INFLUENCE.**

**ESTIMATION RÉGIONALE DES DÉBITS DE CRUE: DÉTERMINATION DES
RÉGIONS HOMOGENÈNES PAR LA MÉTHODE DES RÉGIONS D'INFLUENCE.**

Rapport rédigé pour

**Projet de recherche CRSNG stratégique
Méthodologie d'estimation régionale des débits de crue:
application à la région Québec-Ontario**

par

**Louis Mathier
René Roy
Hugues Perron
Bernard Bobée**

**Institut National de la Recherche Scientifique, INRS-Eau
2800, Einstein, CP 7500, Sainte-Foy, Québec, G1V 4C7**

**Projet stratégique
CRSNG - STR0118482**

Rapport de recherche No R-420

Juillet 1994

RÉSUMÉ

Ce projet stratégique a pour principal objectif de développer une méthodologie transposable d'estimation régionale des débits de crue et de l'appliquer aux données du Québec et de l'Ontario. Ce projet s'échelonne sur trois ans (du 1^{er} novembre 1991 au 31 octobre 1994) et est subventionné par le Conseil de recherches en sciences naturelles et en génie du Canada.

Une des étapes de ce projet consiste dans la détermination des régions homogènes par la méthode des régions d'influence (VOLET A: [DRH]₁). Le concept de régions géographiques homogènes a été fortement contesté car la proximité géographique des bassins versants n'est pas une garantie de leur similitude hydrologique. Nous considérons dans ce volet une méthode de détermination des régions homogènes au sens hydrologique.

Selon l'approche des régions d'influence, chaque site peut être considéré comme le centre de gravité d'une région homogène formée de sites dont les caractéristiques de crue sont similaires. Cette approche implique que l'on doive avoir un minimum d'information hydrologique au site pour lequel on souhaite définir la région d'influence.

Dans le cas où l'on ne dispose pas d'information hydrologique au site considéré, nous avons déterminé quelles pouvaient être les variables (climatologiques et physiographiques) susceptibles de pallier ce manque.

L'étape suivante consiste à calculer les distances entre les sites dans l'univers des caractéristiques sélectionnées. Nous définissons ensuite la région d'influence de chacune des stations de la banque de données.

REMERCIEMENTS

Nous tenons à remercier:

- le Conseil de recherche en sciences naturelles et en génie du Canada (CRSNG) pour son support financier;
- la Direction des ressources en eau et le Service de l'environnement atmosphérique d'Environnement Canada pour avoir mis à notre disposition les données de base pour ce projet;
- les professeurs, chercheurs et étudiants suivants qui ont participé à l'élaboration de ce projet: à l'INRS-Eau: Prof. P.F. Rasmussen et M. P. Trudel; à l'Université McGill: Prof. V.-T.-V. Nguyen et M. G. Pandey; à l'Université de Moncton: Prof. F. Ashkar et Dr T. Ouarda; à l'Université d'Ottawa: Prof. K. Adamowski, Dr D. Gingras et M. J.-L. Daviau; et à l'École Polytechnique de Montréal: Prof. J. Rousselle, Dr S. Birikundavy et Dr J. Ribeiro Corrêa.

Nous tenons également à remercier le Prof. J. Bernier, le Prof. G. Cavadias et le Dr P. Pilon pour leur aide et leurs judicieux conseils.

TABLE DES MATIÈRES

1	INTRODUCTION	1
1.1	Cadre du projet stratégique	2
1.2	Objectifs de l'étude	4
2	RÉGIONS D'INFLUENCE	6
2.1	Concept de voisinage	6
2.2	Méthode de Burn	7
2.2.1	Sélection des attributs	7
2.2.2	Standardisation	7
2.2.3	Calcul des distances euclidiennes	8
2.2.4	Détermination de la limite de la région d'influence	9
2.2.5	Fonction de poids	9
2.2.6	Estimation de Q_T	10
2.2.7	Exemples d'application	11
2.3	Limites de la méthode de Burn	12
3	APPLICATION	14
3.1	Les données	14
3.2	Application aux sites jaugés	16
3.2.1	Méthode de Burn ($[DRH]_{1,1}$)	16
3.2.2	Méthode de région d'influence ($[DRH]_{1,2}$)	19
3.3	Application aux sites non jaugés	23
3.3.1	Méthode des sites non jaugés ($[DRH]_{1,3}$)	23
4	CONCLUSION	27
5	RÉFÉRENCES	28
6	ANNEXE: LIMITE DES RÉGIONS D'INFLUENCE ET FONCTION DE POIDS ...	31

LISTE DES TABLEAUX

Tableau 3.1:	Stations et poids de la région d'influence des deux premières stations du fichier DRH_J.11	19
Tableau 3.2:	Stations et poids de la région d'influence des deux premières stations du fichier DRH_J.12	22
Tableau 3.3:	Stations et poids de la région d'influence des deux premières stations du fichier DRH_NJ.13	25
Tableau 6.1:	Paramètres utilisés pour déterminer la limite des régions d'influence et la fonction de poids des trois méthodes considérées	33

1 INTRODUCTION

On trouve au Canada une proportion importante des eaux douces du monde réparties à travers des régions montrant une grande diversité physiographique et climatique. La gestion et l'utilisation adéquate des eaux douces nécessitent une connaissance des écoulements aux sites où se posent les problèmes reliés tant à la planification des ouvrages de contrôle, qu'à l'opération des réservoirs, ou à la prévention des inondations. Cependant, en raison de la grande étendue du Canada, il serait trop coûteux d'installer un réseau de station permettant l'acquisition de mesures de débit en tous les points du territoire où il est nécessaire de disposer de données. C'est pourquoi, il arrive fréquemment que l'on ne dispose pas de suffisamment de données en des sites où l'on projette d'effectuer la gestion, le contrôle ou l'utilisation rationnelle de la ressource-eau.

Ce problème est particulièrement crucial dans le cas:

des débits de crue, en raison des risques associés à une mauvaise connaissance des écoulements;

- risque d'inondations et de pertes matérielles et de vies humaines en cas de sous-estimation des débits qui conduit à la construction de structures sous dimensionnées,
- coûts de construction trop élevés en cas de surestimation,

des débits d'étiage, en raison de la détérioration de la qualité de l'eau en période de bas débit et de l'impact qui en résulte sur l'alimentation en eau, la navigation et l'environnement.

Il est cependant possible, à partir de données à des stations appartenant à une même région hydrologique, d'effectuer une estimation régionale des débits à un site où l'on dispose de peu ou d'aucune données.

Le principal objectif de ce projet est de:

développer une méthodologie transposable d'estimation régionale et de l'appliquer aux données de crue du Québec et de l'Ontario. Cette région d'étude a été choisie en raison

de la densité et de la qualité des données disponibles et aussi en raison de la bonne connaissance qu'en ont les chercheurs de l'équipe.

Cette méthodologie pourra ensuite être appliquée à d'autres régions du Canada, dans d'autres pays et à d'autres données, les étiages ou les précipitations par exemple.

1.1 Cadre du projet stratégique

On est fréquemment amené à estimer un débit de crue Q_T correspondant à une période de retour T en des sites où l'on dispose de peu ou d'aucune information hydrologique. Dans ce cas, les approches statistiques traditionnelles (ajustement de lois statistiques ou représentation par un modèle de dépassement) ne peuvent être considérées. Une approche alternative consiste alors à utiliser l'information disponible à d'autres sites faisant partie de la même "région hydrologique homogène", c'est-à-dire ayant le même comportement hydrologique. De plus, plusieurs auteurs (Cunnane, 1989 par exemple) ont démontré que cette approche pouvait conduire à des estimations plus précises même dans le cas où l'on dispose à un site d'un nombre assez important d'observations.

Une approche d'estimation régionale des crues a été proposée initialement par Dalrymple dès 1960 (Index Flood Method); cette approche consiste à:

identifier des régions géographiques homogènes, ce qui a été critiqué car la proximité géographique n'est pas une garantie de l'homogénéité des caractéristiques hydrologiques (Whiltshire, 1986b; Cunnane, 1989; Cavadias, 1990),

déterminer dans la région une courbe standardisée régionale q_T . Le débit de période T au site j est alors $Q_T^j = \mu_j q_T$ où μ_j est un facteur d'échelle, qui peut être une fonction des caractéristiques physiographiques aux sites.

La méthode de Dalrymple (1960), et les applications qui en ont été faites (Chong et Moore,

1983, par exemple) ont été critiquées en ce qui concerne la détermination pratique des régions homogènes, le choix et l'estimation des paramètres de la loi régionale (Benson, 1962; Cunnane, 1989).

L'approche générale qui consiste en:

- A- la détermination des régions homogènes [DRH],
- B- l'utilisation d'une méthode d'estimation régionale [MER],

est cependant susceptible d'amélioration. Les problèmes qui se posent en pratique sont liés: à la détermination de régions homogènes au sens hydrologique, au choix de la méthode d'estimation régionale et à l'estimation du débit à des sites non jaugés.

C'est pourquoi, nous proposons, dans ce projet, de réexaminer globalement cette approche en considérant:

VOLET A: les différentes méthodes possibles de détermination des régions homogènes $[DRH]_i$ ($i = 1...p$),

VOLET B: les différentes méthodes possibles d'estimation régionale $[MER]_j$ ($j = 1...q$),

VOLET C: d'effectuer une intercomparaison globale des couples possibles $C_{ij} = [DRH]_i \times [MER]_j$ ($i = 1...p$ et $j = 1...q$).

À partir des données réelles d'une région du Canada (zone Québec - Ontario du bouclier canadien), l'objectif général du projet est de déterminer et de valider les méthodes C_{ij} les plus adéquates pour estimer Q_T à des sites jaugés où l'on a peu de données ainsi qu'à des sites où l'on dispose d'aucune information.

Une des étapes de ce projet (VOLET A) consiste donc dans la détermination des régions homogènes. Quatre méthodes ont été considérées:

$[DRH]_1$: Méthode des régions d'influence (Burn, 1988; 1990a; 1990b; Zrinji et Burn,

1994).

[DRH]₂ : Méthode d'analyse canonique des corrélations (Cavadias, 1989; 1990).

[DRH]₃ : Analyse factorielle des correspondances (Benzecri, 1973) et classification ascendante hiérarchique (Jambu, 1976).

[DRH]₄ : Méthode des L-moments (Adamowski *et al.* 1994).

La détermination des régions homogènes par la méthode des régions d'influence (Volet A: [DRH]₁) est le sujet principal de ce rapport. Les applications au projet stratégique des méthodes [DRH]₂, [DRH]₃ et [DRH]₄ sont respectivement présentées dans trois rapports de projet préparés par Ribeiro Corrêa et Rousselle (1994), Birikundavyi *et al.* (1994) et Adamowski *et al.* (1994).

1.2 Objectifs de l'étude

Le concept de régions géographiques homogènes a été fortement contesté car la proximité géographique des bassins versants n'est pas une garantie de leur similitude hydrologique (Whiltshire, 1986a; 1986b; Cunnane, 1989; Cavadias, 1990). Nous considérons ici une méthode de détermination des régions homogènes au sens hydrologique.

Selon l'approche des régions d'influence, chaque site peut être considéré comme le centre de gravité d'une région homogène formée de sites dont les caractéristiques de crue sont similaires. Cette approche implique que l'on doit avoir un minimum d'information hydrologique au site pour lequel on souhaite définir la région d'influence.

Dans le cas où l'on ne dispose pas d'information hydrologique au site considéré, nous avons déterminé quelles pouvaient être les variables (climatologiques et physiographiques) susceptibles de pallier ce manque. Aux sites où l'information hydrologique est disponible, les variables climatologiques peuvent également être utilisées comme supplément d'information.

L'étape suivante consiste à calculer les distances entre les sites dans l'univers des

variables sélectionnées. Pour cela, nous avons développé un programme de calcul des distances qui permet de définir la région d'influence d'un site donné selon trois cas:

- détermination de la région d'influence d'un site jaugé dans l'univers des caractéristiques hydrologiques et physiographiques telle que proposée par Burn (1990b) ([DRH_{1,1}]: Méthode de Burn);
- détermination de la région d'influence d'un site jaugé dans l'univers des caractéristiques hydrologiques, climatologiques et physiographiques ([DRH_{1,2}]: Méthode de région d'influence);
- détermination de la région d'influence d'un site non jaugé dans l'univers des caractéristiques climatologiques et physiographiques ([DRH_{1,3}]: Méthode des sites non jaugés).

La dernière étape consiste à identifier la région d'influence d'une station donnée selon la méthode utilisée.

2 RÉGIONS D'INFLUENCE

2.1 Concept de voisinage

La notion d'homogénéité régionale a été récemment remise en question (Wiltshire, 1986a; 1986b; Acreman et Wiltshire, 1989; Cunnane, 1989) au profit d'une méthode où l'on attribue à chaque station de jaugeage sa propre région homogène. Ainsi Cavadias (1989; 1990) et Burn (1990a; 1990b) ont énoncé les concepts de "*voisinage*" (CV) ou de "*région d'influence*" (ROI) qui représentent probablement les contributions les plus importantes à la régionalisation des caractéristiques de crue depuis les travaux de Dalrymple (1949; 1960).

Dans cette méthode, plutôt que de définir a priori une région homogène à l'intérieur de laquelle il sera possible de transférer de l'information, on détermine pour chaque site d'intérêt les stations à utiliser pour transférer l'information relative aux caractéristiques de crue. En régionalisation, les données reconstituées (à partir de l'information régionale) pour un site cible donné, appartiennent à une station virtuelle. Dans les approches classiques de définition des régions homogènes, la station virtuelle est positionnée (dans quelque espace que ce soit) à l'écart de la station cible. Selon la méthode de voisinage, la station cible coïncident avec la station virtuelle reconstituée en maintenant le centre de gravité de la région sur la station cible. Cette coïncidence est obtenue par le poids accordé aux stations voisines en fonction de leur distance avec la station cible (Roy, 1994).

Bien que la philosophie qui sous-tende les travaux de Cavadias et de Burn soit la même, les techniques statistiques employées pour déterminer les stations voisines et pour estimer Q_T à un site d'intérêt sont différentes. Cavadias (1989; 1990) propose une méthode de régionalisation basée sur l'analyse de corrélation canonique, où chaque station d'intérêt possède sa propre région hydrologique homogène. Les détails de l'application, dans le cadre de ce projet stratégique, de la méthode d'analyse canonique des corrélations (Volet A: [DRH]₂) sont donnés par Ribeiro Corrêa et Rousselle (1994).

2.2 Méthode de Burn

Burn (1990a; 1990b) propose une autre méthode de régionalisation qui a la particularité de considérer pour chaque station d'intérêt, sa propre région hydrologique homogène. On choisit les voisins selon la proximité (distance euclidienne) dans un espace de p attributs considérés comme étant pertinents en ce qui a trait à la signature hydrologique d'un bassin versant. Dans les sous-sections qui suivent, nous décrivons brièvement chacune des étapes de cette méthode qui fait l'objet du présent rapport.

2.2.1 Sélection des attributs

La sélection des attributs est effectuée selon la disponibilité des variables qui décrivent les caractéristiques physiques, climatologiques et hydrologiques du réseau de sites à l'étude. La sélection est basée sur le jugement et la connaissance qu'a l'utilisateur des facteurs qui caractérisent un comportement similaire des événements hydrologiques extrêmes. Cette sélection peut reposer sur une connaissance intuitive ou plus objectivement sur l'examen d'une matrice de corrélations incluant une mesure du débit extrême aux sites et certaines variables physiographiques, climatologiques et hydrologiques des bassins versants considérés.

2.2.2 Standardisation

La standardisation des attributs permet l'élimination des problèmes d'échelle et des unités de mesure entre les différentes caractéristiques sélectionnées. La standardisation est obtenue en divisant chaque valeur de l'attribut par l'écart type de ce même attribut calculé à partir de la série des valeurs observées à toutes les stations. On obtiendra ainsi une variable réduite.

2.2.3 Calcul des distances euclidiennes

La distance euclidienne D_{jk} entre les sites j et k est calculée de la façon suivante:

$$D_{jk} = \left[\sum_{i=1}^p (C_j^i - C_k^i)^2 \right]^{1/2} \quad (1)$$

où;

p nombre d'attributs considérés pour la mesure de la distance

C_j^i valeur standardisée de la mesure de l'attribut i au site j

C_k^i valeur standardisée de la mesure de l'attribut i au site k

Si on souhaite accorder davantage de poids à un attribut de façon à tenir compte de l'importance relative de certaines variables descriptives de la réponse hydrologique d'un bassin versant, on peut le faire à cette étape de la procédure. On obtient ainsi la distance pondérée D_{jk} :

$$D_{jk} = \left[\sum_{i=1}^p W_p (C_j^i - C_k^i)^2 \right]^{1/2} \quad (2)$$

où:

W_p poids appliqué à l'attribut p.

La distance obtenue à l'aide des équations 1 ou 2 permet de connaître la distance entre chacune des stations dans l'espace des p attributs considérés.

2.2.4 Détermination de la limite de la région d'influence

L'étape suivante consiste à sélectionner, pour chaque station j où l'on souhaite estimer Q_T , les stations les plus proches dans l'espace des p attributs choisis. On cherche alors une distance limite ou un seuil θ qui permet d'obtenir un compromis entre la quantité d'information (nombre de stations voisines) à considérer dans notre analyse régionale et l'homogénéité hydrologique du groupe de stations voisines sélectionnées. En d'autres termes, on souhaite tenir compte d'un maximum d'information régionale (c'est-à-dire retenir un maximum de voisins) tout en maximisant l'homogénéité du voisinage (ce qui revient à retenir un minimum de voisins). La valeur du seuil θ doit être choisie par l'utilisateur.

2.2.5 Fonction de poids

Comme les stations voisines ne sont pas toutes aussi près de la station j , Burn (1990a; 1990b) propose une fonction de poids qui prendra en compte l'importance relative de chacun des voisins (station k) dans l'estimation de Q_T au site j . Cette fonction de poids est définie de la façon suivante:

$$WF_{jk} = 1,0 - \left(\frac{D_{jk}}{TLH} \right)^n \quad (3)$$

où;

WF_{jk} poids de la station k dans le voisinage de la station j

TLH paramètre

n constante positive.

Le paramètre TLH donne la valeur de la fonction de poids à la limite supérieure de la région. La constante n précise le taux de décroissance du poids avec la distance entre les stations k et j . La méthode de Burn permet donc de sélectionner les stations voisines à inclure dans la région d'influence de la station j et d'affecter des poids relatifs à chacune

des stations k en fonction de leur distance à la station j dans la dimension des p attributs retenus. Ces poids sont utilisés lors de l'estimation de Q_T à la station j .

2.2.6 Estimation de Q_T

Contrairement à d'autres méthodes de régionalisation, où la détermination des régions homogènes et la méthode d'estimation régionale de Q_T représentent deux étapes distinctes de la méthodologie, la méthode de régionalisation proposée par Burn forme un tout. En effet, les poids WF_{jk} qui sont déterminés lors de l'élaboration de la région d'influence, peuvent être utilisés directement pour estimer les paramètres de la distribution régionale, qui est à son tour utilisée pour l'estimation de Q_T au site d'intérêt. Burn (1990a; 1990b) suggère l'utilisation de la loi GEV, avec estimation des paramètres par la méthode des moments pondérés, connue en anglais sous le nom de "probability weighted moment" (PWM), pour l'estimation de Q_T . La loi GEV / PWM a également été sélectionnée pour l'estimation régionale de Q_T dans le cadre du présent projet stratégique, "VOLET B: Méthode d'estimation régionale, [MER]₁: Distribution paramétrique régionale" (Rasmussen, 1994).

Dans cette méthode, les trois paramètres de la loi GEV sont estimés à partir des trois premiers moments pondérés des observations. Dans un premier temps, on calcule les PWM standardisés des échantillons de débits maximums annuels de chacune des stations de la région d'influence. Ensuite, les trois premiers PWM de la région d'influence (PWM régionaux) sont calculés à partir des PWM de toutes les stations de la région d'influence selon le poids WF_{jk} associé à chacune des stations du voisinage. Les PWM régionaux sont alors utilisés pour estimer les paramètres de la distribution régionale standardisée. Finalement, la distribution régionale standardisée est utilisée pour estimer la valeur de q_T^j pour une valeur de T donnée. Le débit au site j de période T est alors $Q_T^j = M_0^j q_T^j$ où M_0^j est la moyenne du débit maximum annuel.

Il est important de noter que les résultats de l'application de la méthode des régions d'influence présentés à la section 3 de ce rapport sont exclusivement liés à l'étape de

détermination des régions homogènes. Par contre, dans le cadre du projet stratégique, la méthode complète de régionalisation proposée par Burn sera représentée par le couple $C_{1.1,1} = [DRH]_{1.1} \times [MER]_1$. Il en est de même pour les deux modifications que nous proposons, qui seront représentées par les couples $C_{1.2,1}$ et $C_{1.3,1}$ lors de l'intercomparaison.

2.2.7 Exemples d'application

Burn (1990b) explore par simulation différentes options pour déterminer le seuil de distance à partir duquel on limite le voisinage. Le choix d'un seuil très élevé (avec une petite valeur de n) conduit à la sélection de toutes les stations disponibles comme voisines de la station j . À l'opposé, un seuil très restrictif (avec une grande valeur de n) limite au minimum le nombre de sites voisins retenus. Les résultats obtenus à partir de ces seuils de distance sont comparés à ceux obtenus à partir de régions homogènes déterminées à l'aide d'une analyse d'agrégation. Pour des débits de période de retour $T = 25, 50, 100$ et 200 ans, la racine carrée des erreurs moyennes au carré ainsi que le biais d'estimation de Q_T sont systématiquement moins élevés pour le scénario où on retient un maximum d'information. Le fait de retenir beaucoup d'information complémentaire, même si le poids des stations éloignées est relativement faible ($WF_{jk} \rightarrow 0$), rend la procédure plus robuste.

Burn (1990a) compare les résultats de l'application de cette méthode à ceux obtenus précédemment (Burn, 1988), par l'approche classique de régionalisation qui consiste à déterminer des régions homogènes à l'aide d'une analyse en composantes principales, pour un groupe de 91 stations du sud du Manitoba. Les deux attributs hydrologiques retenus sont:

- le coefficient de variation (CV) des séries de débit maximum annuel
- le débit spécifique (QSPECI) (rapport du débit annuel moyen et de la superficie de drainage du bassin versant)

et deux attributs physiographiques de localisation des stations;

- la longitude de la station de jaugeage
- la latitude de la station de jaugeage.

Les résultats de cette application montrent que la méthode de voisinage se compare avantageusement à celle proposée précédemment.

2.3 Limites de la méthode de Burn

La première limite de l'approche de voisinage se situe au niveau de la sélection de la quantité d'information à considérer lors de la reconstitution de Q_T (choix du seuil θ). On souhaite tenir compte d'un maximum d'information régionale (c'est-à-dire retenir un maximum de voisins) tout en maximisant l'homogénéité du voisinage (ce qui revient à retenir un minimum de voisins), deux objectifs pour lesquels les solutions sont contradictoires (Burn, 1988). Il serait donc intéressant de disposer d'une méthode objective pour déterminer la quantité de stations à inclure dans le voisinage.

Zrinji et Burn (1994) ont récemment développé une procédure permettant de déterminer le nombre optimal de stations voisines à inclure dans la région d'influence. Un test d'homogénéité est appliqué, de manière récursive, à chaque fois que l'on inclut une nouvelle station dans la région d'influence de la station d'intérêt. La région d'influence se limite au nombre de stations à partir duquel l'hypothèse d'homogénéité n'est pas acceptée. Le problème réside alors dans le choix du niveau de signification du test d'homogénéité. En effet la sélection préalable du niveau de sévérité aura une incidence directe sur la taille de la région d'influence. Le choix du niveau de signification du test d'homogénéité devient alors arbitraire et équivalent au choix du seuil θ .

L'autre limite propre à l'approche de Burn réside dans le fait qu'elle a été développée pour estimer Q_T à des sites où l'on dispose d'un minimum d'information hydrologique. Les exemples d'application présentés par l'auteur retiennent des attributs (CV et QSPECI) que l'on ne peut calculer qu'à des stations hydrométriques jaugeées. Selon Burn (1990b), cette limite pourrait être transgressée en remplaçant les attributs hydrologiques dans la méthode,

par des variables disponibles à des stations non jaugées.

En ce sens, Zrinji et Burn (1994) proposent une application de la méthode basée sur le calcul des distances à partir des caractéristiques physiographiques et climatiques des bassins versants qui sont disponibles pour des sites non jaugés. L'estimation des paramètres de la distribution statistique régionale (à partir des sites jaugés de la région d'influence) et l'estimation de q_T^j sont effectuées comme à la section 2.2.6. Par contre, l'estimation du débit au site j non jaugé ne peut pas être obtenue directement par $Q_T^j = M_0^j q_T^j$ puisque M_0^j , la moyenne du débit maximum annuel est inconnue au site non jaugé. Avant d'effectuer l'estimation de Q_T^j au site j non jaugé on doit estimer la valeur de M_0^j . Cette estimation de M_0^j repose sur une analyse de régression basée sur les caractéristiques physiographiques et climatologiques disponibles pour tous les sites de la région (le site non jaugé et les sites jaugés).

Nous présentons dans le chapitre suivant une application de la méthode des régions d'influence, pour déterminer des régions homogènes, à partir des données du Québec et de l'Ontario. Les régions d'influence ont été définies selon trois méthodes qui se distinguent par le choix des attributs utilisés pour l'élaboration du voisinage selon que le site d'intérêt est un site jaugé ou non jaugé.

3 APPLICATION

Les deux premières méthodes utilisées dans le cadre de ce projet stratégique pour la détermination des régions homogènes à partir de la méthode des régions d'influence sont appliquées aux sites jaugés. Dans la première méthode ($[DRH]_{1,1}$: Méthode de Burn), la région d'influence d'un site jaugé est définie dans l'univers des caractéristiques hydrologiques et physiographiques telle que proposée par Burn (1990b).

La deuxième méthode que l'on propose ($[DRH]_{1,2}$: Méthode de région d'influence) est issue d'une réflexion portant sur les variables climatologiques qui pourraient être utilisées comme information supplémentaire pour définir le voisinage dans la méthode de Burn. Les variables climatologiques ont été sélectionnées (Mathier *et al.*, 1993) selon leur influence sur les processus générateurs de crue. La région d'influence d'un site jaugé est alors définie dans l'univers des caractéristiques hydrologiques, climatologiques et physiographiques.

Finalement, la troisième méthode proposée ($[DRH]_{1,3}$: Méthode des sites non jaugés) porte sur une analyse des variables qui pourraient être substituées à CV et à QSPECI dans les applications de la méthode de Burn aux sites non jaugés pour lesquels ces informations ne sont pas disponibles. Dans ce cas, la région d'influence d'un site non jaugé est définie dans l'univers des caractéristiques climatologiques et physiographiques.

3.1 Les données

Les régions d'influence sont déterminées à partir des données réelles d'une région du Canada (zone Québec - Ontario du bouclier canadien). Nous disposons de données physiographiques, hydrologiques et météorologiques pour 179 stations du Québec et de l'Ontario. Mathier *et al.* (1993) présentent une description complète de l'ensemble des données du projet stratégique. Pour toutes ces stations, nous disposons d'une série d'au moins 20 ans de débit maximum printanier qui sert de base à l'élaboration des régions d'influence. Le débit maximum printanier est défini comme suit:

- pour chaque année, on considère la série chronologique des débits journaliers moyens pour la période printanière (du premier janvier au 30 juin);
- dans cette série, on choisit la plus grande valeur de débit moyen mesurée.

Cette valeur représente alors le débit maximum printanier (QMAXPRI) déterminé à partir des débits moyens journaliers pour la période de janvier à juin.

Dès le départ, il faut considérer deux cas d'élaboration des régions d'influence suivant que l'on considère la station cible comme étant un site jaugé (utilisation de l'information hydrologique au site) ou comme étant un site non jaugé (aucune information hydrologique au site). En réalité tous les sites de l'étude contiennent suffisamment de données ($N \geq 20$) pour obtenir une estimation adéquate de Q_T au site sans avoir recours à une estimation régionale. Cette estimation de Q_T au site est alors utilisée dans le projet stratégique comme base de comparaison (valeur théorique cible).

Pour les fins de la simulation (intercomparaison), dans le cas des stations jaugées, on considère une partie de l'information hydrologique disponible au site cible. Le choix du nombre d'années de données hydrologiques à considérer pour la simulation (niveau d'information), doit être suffisant pour être considéré dans le modèle régional mais insuffisant pour obtenir une estimation locale adéquate de Q_T . Le nombre d'années de données à considérer à la station cible a été fixé à $N_s = 10$ ans. En effet, ce nombre d'années nous apparaît suffisant pour être considéré dans le contexte d'un modèle régional, mais insuffisant pour obtenir une estimation locale adéquate de Q_T . Pour choisir les 10 ans de données de manière aléatoire, on tire une année de départ parmi les années de données disponibles ($N \geq 20$) à chaque station, en s'assurant qu'il y a au moins 9 autres années de données subséquentes. Si ce n'est pas le cas, on effectue un nouveau tirage aléatoire. Donc pour le cas des stations jaugées, on considère que la station utilisée comme cible, c'est à dire celle où l'on veut obtenir l'estimation régionale $Q_{T,r}$ qui sera comparée à l'estimation au site $Q_{T,s}$, est un site qui contient 10 ans de données (10 valeurs de QMAXPRI). Les 178 autres stations sont alors des sites entièrement jaugés (contiennent les valeurs de QMAXPRI pour toutes les années disponibles, $N \geq 20$).

Dans le cas des stations non jaugées, on fait l'hypothèse que la station cible ne contient aucune information hydrologique. On considère donc que la station utilisée comme cible est un site non jaugé (aucune valeur de QMAXPRI). Les 178 autres stations sont alors des sites jaugés (contiennent toutes les valeurs de QMAXPRI, $N \geq 20$).

3.2 Application aux sites jaugés

Les deux premières méthodes présentées dans les sections suivantes sont appliquées au cas des sites jaugés à l'étude dans le projet stratégique.

3.2.1 Méthode de Burn ([DRH]_{1,1})

La première méthode ([DRH]_{1,1} : Méthode de Burn) est utilisée pour déterminer la région d'influence des stations de l'étude (site jaugé) à partir des caractéristiques hydrologiques et physiographiques des bassins versants. La méthode utilisée est celle proposée par Burn (1990a; 1990b).

Étape 1: Sélection des attributs

Les attributs hydrologiques et physiographiques retenus sont les mêmes que ceux utilisés par Burn (1990a). Les deux attributs hydrologiques sont:

- le coefficient de variation (CV) des séries de débits maximums printaniers, qui caractérise la distribution des crues,
- le débit spécifique (QSPECI) (rapport du débit maximum printanier moyen et de la superficie de drainage du bassin versant), qui caractérise le volume des crues.

Les deux attributs physiographiques de localisation des stations tiennent compte de la proximité géographique entre les sites:

- la longitude de la station de jaugeage,
- la latitude de la station de jaugeage.

Étape 2: Standardisation

La standardisation est obtenue en divisant chaque valeur de l'attribut par l'écart type de ce même attribut calculé à partir de la série des valeurs observées à toutes les stations. Lorsqu'une station est utilisée comme site cible, elle contient 10 valeurs sélectionnées de QMAXPRI, sinon elle contient $N \geq 20$ valeurs de QMAXPRI. Donc, pour une même station, les valeurs des attributs QSPECI et CV calculées à partir de QMAXPRI, seront différentes selon que l'on considère cette station comme un site cible ($N_s = 10$ valeurs de QMAXPRI) ou non ($N \geq 20$ valeurs de QMAXPRI). Au site cible, le calcul de l'écart type des séries et des valeurs de QSPECI et CV standardisés est donc effectué à partir des valeurs de QSPECI et de CV estimées en considérant les 10 valeurs sélectionnées de QMAXPRI. Si la station n'est pas un site cible, les valeurs de QSPECI et de CV utilisées sont celles calculées à partir de toutes les valeurs de QMAXPRI disponibles.

Étape 3: Calcul des distances euclidiennes

Le calcul des distances euclidiennes introduit dans le programme informatique est basé sur l'équation 2. Toutefois, dans l'application pratique nous avons considéré $W_p = 1,0$ et ainsi les poids appliqués aux attributs (W_p) ne sont pas utilisés, ce qui revient à utiliser l'équation 1.

Étape 4: Détermination de la limite de la région d'influence

La détermination de la limite des régions d'influence est effectuée selon l'option 1 proposée par Burn (1990b). Dans cette option toutes les stations situées à l'intérieur d'une distance limite sont incluses dans le voisinage de la station cible. Si le nombre de stations dans la région d'influence est inférieur au nombre de stations désirées, la distance limite est augmentée afin d'obtenir plus de stations dans le voisinage.

Le choix de la distance seuil d'appartenance à une région est donné par:

$$\theta_j = \theta_L \quad NS_j \geq NST \quad (4)$$

$$\theta_j = \theta_L + (\theta_U - \theta_L) \left(\frac{NST - NS_j}{NST} \right) \quad NS_j < NST \quad (5)$$

où:

- θ_j distance seuil d'appartenance à une région pour le site j
- θ_L limite inférieure du seuil d'appartenance à une région
- θ_U limite supérieure du seuil d'appartenance à une région
- NS_j nombre de stations dans la région d'influence du site j pour la valeur de θ_j donnée
- NST nombre minimum de stations requises dans la région d'influence

Suite à l'examen de la matrice des distances entre les sites pour la méthode $[DRH]_{1,1}$ (voir Chapitre 6, Annexe), les limites suivantes ont été fixées:

- limite inférieure du seuil d'appartenance à une région, $\theta_L = 1,4$
- limite supérieure du seuil d'appartenance à une région, $\theta_U = 3,5$
- nombre minimum de stations requises dans la région d'influence, $NST = 20$ sites

Étape 5: Fonction de poids

Le calcul des poids WF_{ij} est effectué à partir de l'équation 3. Le paramètre TLH dicte la valeur de la fonction de poids à la limite supérieure de la région. On doit avoir $TLH \geq \theta_U$, si $TLH = \theta_U$, une station située à la limite supérieure de la région n'aura aucune contribution dans l'estimation régionale. Plus TLH est grand plus la contribution des dernières stations augmente. Pour la méthode $[DRH]_{1,1}$, la valeur de TLH est fixée à 3,6 (voir Chapitre 6, Annexe). La valeur du paramètre n, qui indique le taux de décroissance

des poids avec la distance du site cible est fixée à 2,5 (voir Chapitre 6, Annexe).

Étape 6: Résultat

La liste des stations (avec leur poids respectif) qui composent les régions d'influence des sites à l'étude pour la méthode de Burn ($[DRH]_{1,1}$) est disponible dans un fichier nommé DRH_J.11. Le tableau 3.1 donne les stations et les poids de la région d'influence des deux premières stations du fichier DRH_J.11.

Tableau 3.1: Stations et poids de la région d'influence des deux premières stations du fichier DRH_J.11

#01BD002							
02QA002	0.992835	02NE011	0.974264	01BH002	0.974141	02PG004	0.971224
01BF001	0.969050	02UC002	0.968360	02VC001	0.960366	02PB006	0.955293
02RD003	0.946287	02WA001	0.945660	02PJ007	0.944176	02PL005	0.944156
02PJ030	0.942602	02QC001	0.941502	02VA001	0.936052	02NE007	0.935572
02RF001	0.929854	02PL001	0.929323	02PA007	0.916493	02RF002	0.916319
02RD002	0.908369						
#01BF001							
02QC001	0.995191	02PG004	0.992490	02VA001	0.984494	02QB011	0.982691
01BH002	0.980152	02PJ007	0.975976	01BD002	0.969050	02PL005	0.967907
02QA002	0.966425	01BH005	0.966092	02PB006	0.964188	02PD012	0.964142
02PJ030	0.957432	02OE027	0.943603	02PL001	0.943178	01BH001	0.941309
02OG026	0.940124	02NE011	0.917552	02WA001	0.906950	02MC001	0.895369
02OE018	0.889852	02LB006	0.889413				

3.2.2 Méthode de région d'influence ($[DRH]_{1,2}$)

La deuxième méthode ($[DRH]_{1,2}$: Méthode de région d'influence) est utilisée pour déterminer la région d'influence des stations de l'étude (site jaugé) à partir des caractéristiques hydrologiques, physiographiques et météorologiques des bassins versants.

Étape 1: Sélection des attributs

Les attributs hydrologiques et physiographiques retenus sont les mêmes que ceux utilisés à la section 3.2.1 et proposés par Burn (1990a). Les deux attributs hydrologiques sont:

- le coefficient de variation (CV) des séries de débits maximums printaniers,
- le débit spécifique (QSPECI).

Les deux attributs physiographiques de localisation des stations sont:

- la longitude de la station de jaugeage,
- la latitude de la station de jaugeage.

Pour chaque station, nous avons ajouté à ces attributs, quatre variables météorologiques qui sont associées aux différents processus générateurs des crues:

- la médiane de la moyenne des températures moyennes journalières (deg. C), des 5 jours précédant la crue (DTY5),
- la moyenne de la précipitation totale (mm), des 5 jours précédant la crue (YPR5),
- la moyenne de la neige au sol (cm), 5 jours avant la crue (YNS5),
- la précipitation totale moyenne annuelle (mm) (PTMA).

Toutes les caractéristiques météorologiques sont évaluées à partir des données de la station climatologique la plus voisine de la station hydrométrique, parmi celles qui ont une période d'observations concomitantes suffisantes avec la station hydrométrique. Ces variables météorologiques sont décrites en détail dans le rapport de Mathier *et al.* (1993).

Étapes 2 et 3

Les étapes 2 et 3 sont réalisées de la même manière que celles présentées à la section 3.2.1.

Étape 4: Détermination de la limite de la région d'influence

Le choix de la distance seuil d'appartenance à une région est donné par les équations 4 et 5. Suite à l'examen de la matrice des distances entre les sites pour la méthode $[DRH]_{1,2}$ (voir Chapitre 6, Annexe) les limites suivantes ont été fixées:

- limite inférieure du seuil d'appartenance à une région, $\theta_L = 2,5$
- limite supérieure du seuil d'appartenance à une région, $\theta_U = 5,6$
- nombre minimum de stations requises dans la région d'influence, NST = 20 sites

Étape 5: Fonction de poids

Le calcul des poids WF_{ij} est effectué à partir de l'équation 3. Pour la méthode $[DRH]_{1,2}$, la valeur de TLH est fixée à 5,7 (voir Chapitre 6, Annexe). La valeur du paramètre n , qui indique le taux de décroissance des poids avec la distance du site cible est fixée à 2,5 (voir Chapitre 6, Annexe).

Étape 6: Résultat

La liste des stations (avec leur poids respectif) qui composent les régions d'influence des sites à l'étude pour la méthode de région d'influence ($[DRH]_{1,2}$) est disponible dans un fichier nommé DRH_J.12. Le tableau 3.2 donne les stations et les poids de la région d'influence des deux premières stations du fichier DRH_J.12.

Tableau 3.2: Stations et poids de la région d'influence des deux premières stations du fichier DRH_J.12

#01BD002							
02PJ007	0.962572	02PL005	0.949888	02PC009	0.922641	02JB003	0.920469
02PL001	0.915500	02JB013	0.911785	02OE027	0.906658	01BH002	0.893297
02EA005	0.888612	02EC002	0.879507	02OE018	0.877687	02LB006	0.869488
02WA001	0.863890	02PG004	0.862355	02PA007	0.861842	02LB007	0.861008
02LC043	0.856825	02LH004	0.849008	02QC001	0.848980	02LB008	0.848497
02OG026	0.840771	02EA010	0.839709	02UC002	0.839065	02FA001	0.835521
02HM004	0.835411	02VC001	0.828681	02MC001	0.826042	02HJ001	0.822182
02FB007	0.821816	01BF001	0.820779	02NE011	0.817913	01BH005	0.816037
02PJ030	0.810824	02QA002	0.808071	02LH002	0.804682	02HM005	0.797915
02VA001	0.792395	02RD003	0.789109	02RF002	0.788718	02JB004	0.785690
03AC002	0.783561	02HL004	0.783015	02RF001	0.775608	02GB007	0.775488
02ED003	0.771785	01BH001	0.771783	02RF006	0.769821	02CF007	0.767766
02LD001	0.761915	02PD012	0.759166	02DD008	0.758432	04NA001	0.755691
02RD002	0.754537	02LD002	0.753444	02HA006	0.750788	02FC002	0.742279
03AC001	0.733608	03DD002	0.732750	02LD005	0.730720	02GA010	0.717525
03AD001	0.713464	02OJ007	0.712702	02FC001	0.710060	02OA035	0.707211
02QB011	0.706748	03DD003	0.703298	02BF002	0.702825	02FF002	0.701910
02HB004	0.695078	02BF001	0.689246	03AC004	0.689228	02GG002	0.685754
02OJ001	0.681320	02CA002	0.679359	02GA018	0.674840	03EA001	0.674231
03DC002	0.670065	02CD001	0.669916	02GG005	0.664456	02HB012	0.660648
02PB006	0.660315	02PE009	0.658454	02GG004	0.654090	03AB002	0.643994
02GA017	0.643385	02HC033	0.642685	02GB009	0.639438	02GD020	0.636544
04LJ001	0.632374	02XC001	0.631025	02BB002	0.630147	02BB003	0.626440
04MF001	0.622662	02FC011	0.614478	02HD002	0.614312		
#01BF001							
01BH005	0.961208	02VA001	0.947246	02WA001	0.942756	02PG004	0.936657
02UC002	0.935481	01BH002	0.924805	02PA007	0.915100	02PE009	0.910693
01BH001	0.909735	02NE011	0.907053	02VC001	0.900016	02DD008	0.899187
02QA002	0.896186	02RD003	0.881334	02PJ030	0.877672	02EA010	0.874070
02MC001	0.856228	02BF002	0.853712	02HL004	0.850495	03AC002	0.844553
02LD001	0.844115	02RF002	0.840300	02LH004	0.834329	02OE018	0.833991
02RF001	0.833021	02QB011	0.823843	02RD002	0.823567	02PB006	0.821506
01BD002	0.820779	02BF001	0.816982	02HM005	0.816865	02CA002	0.812988
02RF006	0.806192	02CF007	0.801102	02LB007	0.797992	02LH002	0.797254
02JB004	0.794072	02CD001	0.785091				

Pour certaines stations du Québec et de l'Ontario, la détermination des régions d'influence n'est pas possible. Ce problème est dû à l'absence de données météorologiques pour certaines stations de la région de départ. Ces stations apparaissent à la fin du fichier DRH_J.12.

3.3 Application aux sites non jaugés

La dernière méthode présentée dans cette section est appliquée au cas des sites non jaugés à l'étude dans le projet stratégique.

3.3.1 Méthode des sites non jaugés ([DRH]_{1,3})

La troisième méthode que nous avons proposée ([DRH]_{1,3} : Méthode des sites non jaugés) est utilisée pour déterminer la région d'influence d'un site non jaugé à partir des caractéristiques météorologiques et physiographiques. Ici, les variables météorologiques sont utilisées comme substitut aux variables hydrologiques (CV, QSPECI) pour le calcul des distances et pour estimer la moyenne du débit maximum annuel M_0^j lors de l'estimation de Q_T (voir sections 2.2.6 et 2.3). La méthode que nous proposons est similaire à celle récemment proposée par Zrinji et Burn (1994).

Étape 1: Sélection des attributs

Les attributs physiographiques et météorologiques retenus sont les mêmes que ceux utilisés à la section 3.2.2. Les deux attributs physiographiques de localisation des stations sont:

- la longitude de la station de jaugeage,
- la latitude de la station de jaugeage.

Nous avons ajouté à ces attributs, quatre variables météorologiques qui sont associées aux différents processus générateurs des crues:

- la médiane pour la station de la moyenne des températures moyennes journalières (deg. C), des 5 jours précédant la crue (DTY5),
- la moyenne pour la station de la précipitation totale (mm), des 5 jours précédant la crue (YPR5),
- la moyenne pour la station de la neige au sol (cm), 5 jours avant la crue (YNS5),
- la précipitation totale moyenne annuelle (mm) (PTMA).

Ces variables météorologiques sont décrites en détail dans le rapport de Mathier *et al.* (1993).

Étape 2: Standardisation

La standardisation est obtenue en divisant chaque valeur de l'attribut par l'écart type de ce même attribut calculé à partir de la série des valeurs observées à toutes les stations.

Étape 3: Calcul des distances euclidiennes

L'étape 3 est réalisée de la même manière que celle présentée à la section 3.2.1.

Étape 4: Détermination de la limite de la région d'influence

Le choix de la distance seuil d'appartenance à une région est donné par les équations 4 et 5. Suite à l'examen de la matrice des distances entre les sites pour la méthode [DRH]_{1,3} (voir Chapitre 6, Annexe) les limites suivantes ont été fixées:

- limite inférieure du seuil d'appartenance à une région, $\theta_L = 2,0$
- limite supérieure du seuil d'appartenance à une région, $\theta_U = 5,0$
- nombre minimum de stations requises dans la région d'influence, NST = 20 sites

Étape 5: Fonction de poids

Le calcul des poids WF_{ij} est effectué à partir de l'équation 3. Pour la méthode $[DRH]_{1,3}$, la valeur de TLH est fixée à 5,1 (voir Chapitre 6, Annexe). La valeur du paramètre n, qui indique le taux de décroissance des poids avec la distance du site cible est fixée à 2,5 (voir Chapitre 6, Annexe).

Étape 6: Résultat

La liste des stations (avec leur poids respectif) qui composent les régions d'influence des sites à l'étude pour la méthode des sites non jaugés ($[DRH]_{1,3}$) est disponible dans un fichier nommé DRH_NJ.13. Le tableau 3.3 donne les stations et les poids de la région d'influence des deux premières stations du fichier DRH_NJ.13.

Tableau 3.3: Stations et poids de la région d'influence des deux premières stations du fichier DRH_NJ.13.

#01BD002							
02PJ007	0.969242	02LC043	0.965003	02PL005	0.951300	02PC009	0.944344
02OE027	0.930842	02OE018	0.921111	02JB003	0.919681	02PL001	0.917143
02LB008	0.904114	02JB013	0.901903	02EC002	0.878074	01BH002	0.875547
02LB006	0.858885	02EA005	0.855806	01BH005	0.855131	02PA007	0.853749
02QC001	0.839109	02OG026	0.838682	02LB007	0.838393	02PG004	0.837393
02WA001	0.836449	02LH004	0.833192	01BH001	0.828128	02HJ001	0.805308
02UC002	0.803594	02EA010	0.797091	01BF001	0.789675	02LH002	0.789294
02FA001	0.788995	02FB007	0.788555	02VC001	0.788028	02HM004	0.787685
02JB004	0.785145	02MC001	0.780764	02PE009	0.780292	02PJ030	0.778148
02ED003	0.772482	02VA001	0.761446	02NE011	0.760941	02RF006	0.759769
04NA001	0.759037	02PD012	0.758682	03AC002	0.757671	02QA002	0.754033
02RF002	0.752389	03AC001	0.748011	02RD003	0.745994	02HM005	0.737324
02RF001	0.736873	02LD002	0.735290	02OA035	0.733356	02HL004	0.732412
02LD001	0.731693	02PD013	0.730830	02HD002	0.728365	02PD014	0.727187
02RD002	0.726661	02OJ007	0.725003	03DD002	0.722925	02HD007	0.719770
02LD005	0.716287	02GB007	0.712281	02CF007	0.711939	02HA006	0.703787
03AD001	0.703263	02OJ001	0.699918	02QB011	0.699329	02DD008	0.696104
02FF002	0.679570	02FC002	0.679209	03DD003	0.677878	02GA010	0.672557

Tableau 3.3 (suite)

02GD019	0.666383	03AC004	0.664224	02GG005	0.661907	02GA018	0.649591
03EA001	0.649052	02HD006	0.647793	02HD009	0.645310	03DC002	0.631769
02FC001	0.627276	02GD020	0.623303	02GG002	0.618376	02HB004	0.618347
02GA017	0.617149	02BF002	0.615619	02GD010	0.612981	02GB009	0.612884
03AB002	0.612046	02HB012	0.608829	02CD001	0.606045	02BF001	0.604541
02CA002	0.603436	02GG004	0.598373	02FC011	0.579582	02GC010	0.577981
02GC018	0.577498	02PB006	0.571813				
#01BF001							
02PE009	0.993577	02UC002	0.972500	01BH005	0.966829	02PA007	0.966175
02WA001	0.959452	02VA001	0.933726	02VC001	0.930901	02RD003	0.919679
02LD001	0.918686	02PG004	0.916688	01BH001	0.915814	01BH002	0.915065
03AC002	0.907854	02NE011	0.903523	02RF006	0.897637	02JB004	0.894564
02LH004	0.893028	02RF002	0.891829	02RF001	0.885847	02QA002	0.885247
02RD002	0.885154	02HL004	0.883662	02LH002	0.871143	02DD008	0.870455
03AC001	0.857158	02PJ030	0.852188	02OA035	0.850720	02LD002	0.850646
02EA010	0.850165	02CD001	0.838290	02BF002	0.837885	03AD001	0.836766
02OJ007	0.825258	04NA001	0.818212	02OE018	0.816265	02LD005	0.814991
02MC001	0.814742	01BD002	0.789675	02OJ001	0.789160	02CF007	0.788822
02QB011	0.782248	02BF001	0.779606	02PB006	0.771355	02HM005	0.768555
02HJ001	0.763648	03EA001	0.760698	02CA002	0.753422	02LB007	0.753162
02PC009	0.739101	02BB002	0.736254	02GG004	0.734350	02EA005	0.730239

Pour certaines stations du Québec et de l'Ontario, la détermination des régions d'influence n'est pas possible. Ce problème est ici aussi dû à l'absence de données météorologiques pour certaines stations de la région de départ. Ces stations apparaissent à la fin du fichier DRH_NJ.13.

4 CONCLUSION

Les trois fichiers de régions d'influence seront utilisés dans le projet stratégique pour l'intercomparaison des différents couples C_{ij} possibles. L'utilisation des régions d'influence et des poids WF_{jk} associés à chaque station de ces régions, couplée à la procédure d'estimation de Q_T au site d'intérêt par la loi GEV / PWM ([MER]_i), représente une procédure complète de régionalisation. Pour la majorité des autres méthodes d'estimation régionale (MER) considérées dans ce projet stratégique, l'utilisation explicite des poids n'est pas possible lors de l'estimation régionale de Q_T . Pour ces MER, les régions d'influence seront quand même utilisées, sans toutefois considérer les poids régionaux.

5 RÉFÉRENCES

ACREMAN, M.C. and S.E. WILTSHIRE (1989). The regions are dead; long live the regions. Methods of identifying and dispensing with regions for flood frequency analysis. IAHS Publication, No 187: 175-188.

ADAMOWSKI, K., GINGRAS, D. and P.J. PILON (1994). Regional L-moment analysis, University of Ottawa, Department of Civil Engineering.

BENSON, M.A. (1962). Evolution of methods for evaluating the occurrence of floods. USGS Water supply paper 1550. A.

BENZECRI, J.P. (1973). Analyse des données. Dunod, Paris. Vol. 2, 619 p.

BIRIKUNDAVYI S., ROUSSELLE, J., et V.T.V. NGUYEN (1994). Détermination des régions homogènes pour le Québec et l'Ontario: une approche par l'analyse des correspondances et la classification ascendante hiérarchique. Département de génie civil, Section hydraulique, École Polytechnique, 50 p.

BURN, D.H. (1988). Delineation of groups for regional flood frequency analysis. Journal of Hydrology, 104: 345-361.

BURN, D.H. (1990a). An appraisal of the "region of influence" approach to flood frequency analysis. Hydrological Sciences Journal, 35 (2): 149-165.

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

CAVADIAS, G.S. (1989). Regional flood estimation by canonical correlation. Paper presented to the 1989 Annual Conference of the Canadian Society for Civil Engineering, St-John's, Terre-Neuve.

- CAVADIAS, G.S. (1990). The canonical correlation approach to regional flood estimation. Regionalisation in hydrology. Proc. of the Ljubljana, Symposium IAHS, Publ. No 191: 171-178.
- CHONG, S. and S.M. MOORE (1983). Flood frequency analysis for small watersheds in southern Illinois. Water Resources Bulletin, 19 (2): 277-282.
- CUNNANE, C. (1989). Statistical distributions for flood frequency analysis operational hydrology. WMO, Report No. 33.
- DALRYMPLE, T. (1949). Regional flood frequency. Presentation at the 29th annual meeting of the Highway Research Board, 22 p.
- DALRYMPLE, T. (1960). Flood frequency analyses. US Geological Survey, Water supply paper 1543.A, 77 p.
- GINGRAS, D., ADAMOWSKI, K. and P.J. PILON (1994). Regional Flood equations for the Provinces of Ontario and Quebec. Water Resources Bulletin, vol. 30(1): 55-67.
- JAMBU, M. (1976). Sur l'interprétation mutuelle d'une classification hiérarchique et d'une analyse des correspondances. Rev. Statist. Appl., 24 (2): 45-73.
- MATHIER, L., ROY, R., BOBÉE, B., PERRON, P. et V. FORTIN (1993). Estimation régionale des débits de crue: description des banques de données hydrométriques, météorologiques et physiographiques pour le Québec et l'Ontario. INRS-Eau, rapport interne no I-123, 79 p.
- RASMUSSEN, P. (1994). The Index Flood Method for regional estimation of T-year events. INRS-Eau, 12 p.

RIBEIRO CORRÉA, J. et J. ROUSSELLE (1994). Détermination des régions homogènes par la méthode d'analyse canonique des corrélations. Département de génie civil, Section hydraulique, École Polytechnique, 49 p.

ROY, R. (1994). Régionalisation des caractéristiques de crue - utilisation d'une méthode combinant les approches déterministes et stochastiques. Thèse de doctorat en sciences de l'eau, INRS-Eau, Université du Québec.

WILTSHIRE, S.E. (1986a). Regional flood frequency analysis I: Homogeneity statistics. *Hydrological Sciences Journal*, 31 (3): 321-333.

WILTSHIRE, S.E. (1986b). Regional flood frequency analysis II: Multivariate classification of drainage basins in Britain. *Hydrological Sciences Journal*, 31 (3): 335-346.

ZRINJI, Z., BURN, D.H. (1994). Flood frequency analysis for ungauged sites using a region of influence approach. *Journal of Hydrology*, 153: 1-21.

6 ANNEXE: LIMITE DES RÉGIONS D'INFLUENCE ET FONCTION DE POIDS

Pour les trois méthodes utilisées dans ce projet ([DRH_{1,1}]: Méthode de Burn; [DRH_{1,2}]: Méthode de région d'influence; [DRH_{1,3}]: Méthode des sites non jaugés), la détermination de la limite des régions d'influence est effectuée selon l'option 1 proposée par Burn (1990b). Dans cette option toutes les stations situées à l'intérieur d'une distance limite sont incluses dans le voisinage de la station cible. Si le nombre de stations dans la région d'influence est inférieur au nombre de stations désirées, la distance limite est augmentée afin d'obtenir plus de stations dans le voisinage. Comme les stations voisines ne sont pas toutes aussi près de la station cible, Burn (1990a; 1990b) propose une fonction de poids qui prendra en compte l'importance relative de chacun des sites voisins (station k) dans l'estimation de Q_T à la station cible.

Détermination de la limite de la région d'influence

Le choix de la distance seuil d'appartenance à une région est donné par:

$$\theta_j = \theta_L \quad NS_j \geq NST$$

$$\theta_j = \theta_L + (\theta_U - \theta_L) \left(\frac{NST - NS_j}{NST} \right) \quad NS_j < NST$$

où:

- θ_j distance seuil d'appartenance à une région pour le site j
- θ_L limite inférieure du seuil d'appartenance à une région
- θ_U limite supérieure du seuil d'appartenance à une région
- NS_j nombre de stations dans la région d'influence du site j pour la valeur de θ_j donnée
- NST nombre minimum de stations requises dans la région d'influence

Afin de déterminer la limite de la région d'influence, les paramètres θ_L , θ_U et NST doivent être choisis par l'utilisateur. Dans un premier temps, pour nous aider à fixer la valeur de NST, nous avons calculer le nombre moyen de stations par région pour toutes les autres méthodes de détermination des région homogènes (DRH_2 , DRH_3 , DRH_4) considérées dans ce projet stratégique. Ce nombre est égal à 17 sites par région en moyenne. Nous avons alors fixé la valeur de NST à 20 sites. Pour la région à l'étude, ce nombre représente un compromis entre la quantité d'information régionale souhaitée et l'homogénéité entre les station appartenant à une même région.

Ensuite, pour nous guider dans le choix de la valeur de θ_L , nous avons calculé à partir de la matrice des distances de chaque méthode, la distance moyenne du 20^e site (D_{20}) des régions d'influence associées à chaque station à l'étude. En fixant la valeur de θ_L légèrement supérieurs à D_{20} ($\theta_L = D_{20} + 0,1$), nous nous assurons d'avoir en moyenne 20 sites par région. Finalement, quelques essais ont été effectués pour trouver une valeur de θ_U adéquate afin de nous assurer de toujours obtenir au moins 20 sites par région d'influence. Le tableau 6.1 donne la valeur des paramètres utilisés pour déterminer la limite des régions d'influence des trois méthodes considérées.

Fonction de poids

Le poids qui prend en compte l'importance relative de chacun des voisins dans l'estimation de Q_T au site cible est définie de la façon suivante:

$$WF_{jk} = 1,0 - \left(\frac{D_{jk}}{TLH} \right)^n$$

où:

WF_{jk} poids de la station k dans le voisinage de la station j

TLH paramètre

n constante positive

Le paramètre TLH dicte la valeur de la fonction de poids à la limite supérieure de la région. La valeur du paramètre n , indique le taux de décroissance des poids avec la distance du site cible. La méthode permet donc d'affecter des poids relatifs à chacune des stations k en fonction de leur distance à la station j dans la dimension des p attributs retenus. Ces poids sont utilisés lors de l'estimation de Q_T à la station j .

On doit avoir $TLH \geq \theta_U$, si $TLH = \theta_U$, une station située à la limite supérieure de la région n'aura aucune contribution dans l'estimation régionale. Plus TLH est grand plus la contribution des dernières stations augmente. Pour les trois méthodes, la valeur de TLH est fixée à $\theta_U + 0,1$. Ainsi, la valeur du paramètre TLH est supérieure à la valeur de la distance du voisin le plus éloigné de façon à ce que toutes les stations voisines sélectionnées aient un poids positif et non nul. La constante n précise le taux de décroissance du poids avec la distance entre les stations k et le site cible j . Pour les trois méthodes, la valeur de n a été fixée à 2,5. Cette valeur permet l'attribution d'un poids élevé aux stations situées le plus près du site cible, et une décroissance rapide des poids avec la distance. Le tableau 6.1 donne la valeur des paramètres utilisés pour le calcul de la fonction de poids des trois méthodes considérées.

Tableau 6.1 Paramètres utilisés pour déterminer la limite des régions d'influence et la fonction de poids pour les trois méthodes considérées

Modèle	NST	D_20	θ_L	θ_U	TLH	n
[DRH _{1,1}]	20	1,3	1,4	3,5	3,6	2,5
[DRH _{1,2}]	20	2,4	2,5	5,6	5,7	2,5
[DRH _{1,3}]	20	1,9	2,0	5,0	5,1	2,5