

Université du Québec  
INRS-EAU

**ESTIMATION RÉGIONALE DES DÉBITS D'ÉTIAGE DE LA PROVINCE DE QUÉBEC  
(Développement des modèles régionaux de queues et d'invariance d'échelle)**

Par Abdelaziz Hamza

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

Jury d'évaluation

Examineur externe

S. Rocky Durrans  
Department of Civil and  
Environmental Engineering  
The University of Alabama  
Tuscaloosa, Alabama

Examineur interne

Monique Bernier  
INRS-Eau

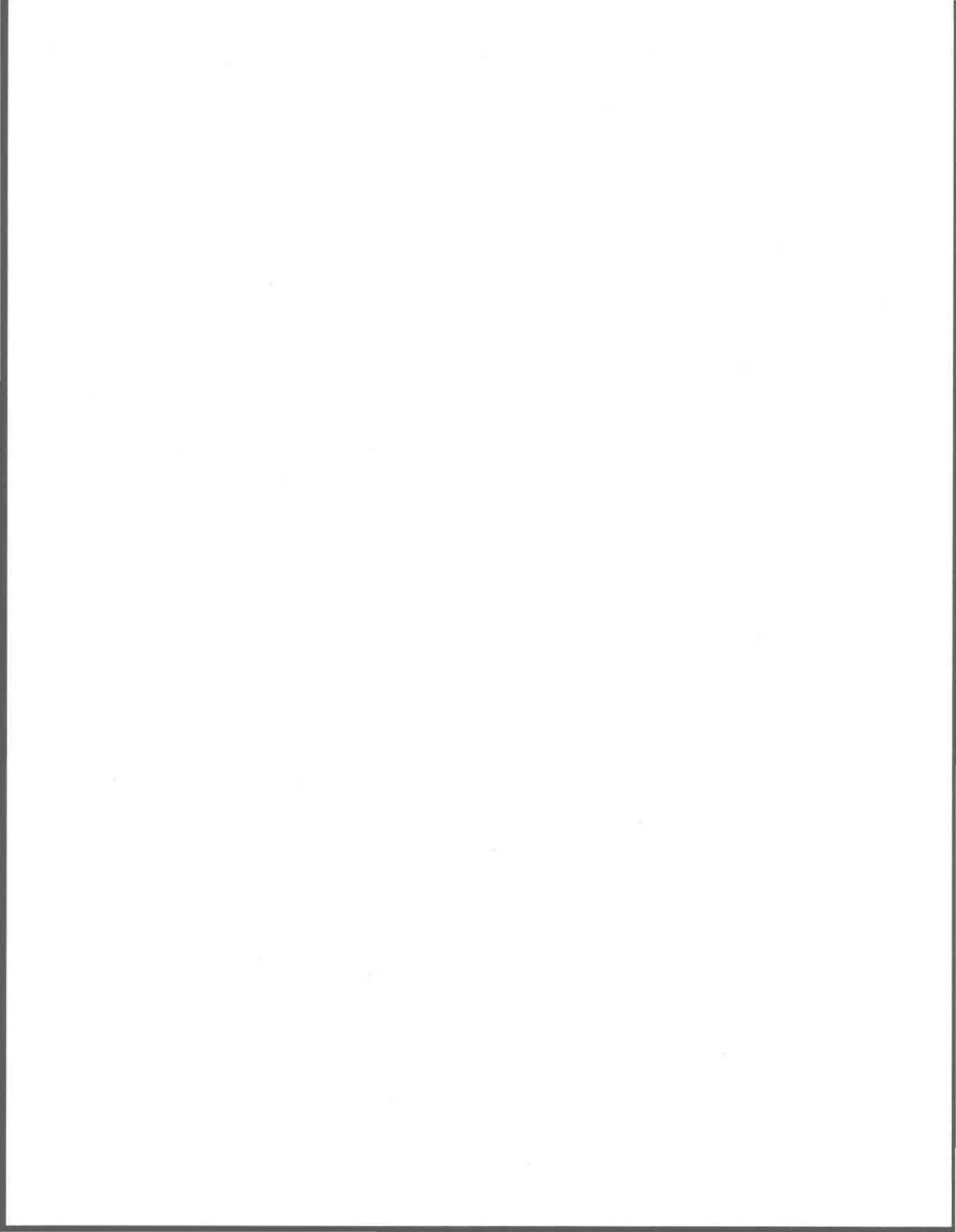
Directeur de recherche

Bernard Bobée  
INRS-Eau

Codirecteur de recherche

Taha, B.M.J Ouarda  
INRS-Eau

Février 1999



## SOMMAIRE

---

La connaissance des caractéristiques d'étiage est nécessaire pour la conception des politiques de gestion et de contrôle des ressources hydriques, pour la planification et la modélisation des réserves en eau, pour l'étude de la qualité, pour la planification des débits de régulation ainsi que pour l'analyse des impacts environnementaux et économiques. L'analyse de fréquence permet d'estimer  $X_T$ , débit de période de retour T, aux sites où l'on dispose de suffisamment d'information hydrométrique. Toutefois, la détermination d'un débit d'étiage de fréquence donnée (quantile), peut-être très souvent nécessaire, en des sites non jaugés ou partiellement jaugés.

La méthode que l'on propose de développer et d'appliquer dans le cadre du présent travail s'inscrit dans l'axe des contributions récentes à l'analyse régionale d'événements hydrologiques extrêmes. Cette méthodologie repose sur la combinaison de deux procédures :

- La méthode d'invariance d'échelle simple des séries régionales des débits d'étiage, basée sur l'analyse de la relation moments-surfaces et/ou de la relation quantiles-surfaces, en d'autres termes, cette méthode explique la variabilité spatiale des processus d'étiage par leur indexation sur une série de paramètres d'échelle, tel que la taille des bassins de drainage ; et

- L'analyse de fréquence des débits minimums avec l'approche par modèles de queue conditionnels, en fait, cette dernière accorde plus de poids à la partie inférieure de la distribution en définissant *a priori* un niveau de censure des données, communément appelé seuil  $u$ , « ceiling value ».

Cette procédure a été appliquée à un échantillon de 187 sites repartis dans la province de Québec, afin que l'analyse soit la plus fiable possible, c'est à dire, que plus le nombre de stations est grand, plus précises seront les analyses.

La première étape a consisté à vérifier les hypothèses de base d'une similitude d'échelle simple que doivent respecter les échantillons de débits minimums. L'analyse des caractéristiques des débits d'étiage de 187 sites répartis dans la province de Québec a montré que les moments non centraux d'ordre un à six suivent une invariance d'échelle simple par rapport à la superficie du bassin versant. Suite à ce résultat nous avons déterminé les bassins hydrologiquement semblables à partir d'une analyse des résidus du sixième moment. Le comportement des débits minimums des différentes régions a montré qu'eux aussi suivaient une invariance d'échelle simple.

La seconde étape de notre travail a porté sur l'application des modèles de queue des distributions Weibull et Lognormale pour l'estimation des débits d'étiage en des sites non jaugés. L'avantage de ces modèles réside en leur facilité d'utilisation, car on réduit le nombre de paramètres à estimer et ainsi on diminue la variance des estimateurs des quantiles. Cette méthode suppose la détermination *a priori* d'un seuil de censure  $u$ . L'estimation de ce seuil telle que présentée dans la littérature est assez subjective, car basée essentiellement sur un examen visuel des graphiques des probabilités empiriques. Une procédure originale basée sur la comparaison des rapports de quantiles calculés pour différents percentiles a été développée. Cette approche a mené au choix d'un seuil  $u$  standard pour tous les échantillons. Ce seuil est représenté par le percentile 60%. Nous avons ensuite testé la performance des modèles de queue en comparant leurs équations de régression régionale respectives à celles résultant du modèle régional de la distribution Weibull et Lognormale basée sur l'échantillon complet.

La variable d'étiage  $Q_{7,10}$ , a été choisie *a priori*, car c'est la variable la plus couramment utilisée en Amérique du Nord. Dans le but de tester davantage la performance de la méthodologie développée, nous avons choisi de l'appliquer à une autre variable qui est  $Q_{30,5}$ . Cette variable est utilisée fréquemment par le Ministère de l'Environnement du Québec ; les résultats de cette application se trouvent dans l'annexe B. Tout comme la variable  $Q_{7,10}$ , l'analyse des caractéristiques des débits d'étiage des 187 sites pour la variable d'étiage  $Q_{30,5}$  a montré que les moments non centraux d'ordre un à six suivaient une invariance d'échelle simple par rapport à la surface du bassin versant. L'analyse des

---

résidus du sixième moment a permis de distinguer deux régions homogènes : la première contenant 117 sites et la seconde région qui en contient 70. L'étude des débits minimums des différentes régions a montré qu'elles aussi suivaient une similitude d'échelle simple. Nous avons ensuite appliqué la théorie des modèles de queue des distributions Weibull et Lognormale pour l'estimation régionale des débits d'étiage. L'application de la procédure présentée au chapitre 4 pour la détermination d'un seuil de censure  $u$  pour la variable  $Q_{30,5}$ , a donné des résultats différents, relativement aux distributions Lognormale et Weibull. En effet, l'analyse des résultats montre que pour la distribution Lognormale (figure B.6), la considération de l'échantillon au complet (seuil = 100%) offre la meilleure estimation. Le modèle régional Lognormal non censuré est plus performant que les modèles censurés. Cette observation est corroborée par la figure B.6(b) où l'erreur quadratique moyenne est minimum pour le percentile 100%. Par contre, l'analyse de la figure B.5, représentant la détermination du seuil de censure  $u$  avec la distribution Weibull, montre que le choix du percentile 50% est plus approprié, le modèle de queue régional de la distribution de Weibull censurée à 50% est plus performant que les autres modèles. Dans ce cas, l'estimation régionale avec la distribution de Weibull non censurée ne performe pas de manière satisfaisante, car l'erreur quadratique moyenne qui lui est associée figure B.5(b) est la plus grande.



## DÉDICACES

---

*À ma mère dont les sacrifices consentis à ma formation et à mon éducation n'ont d'égal que le témoignage profond de ma reconnaissance.*

*À la mémoire de mon défunt père, que dieu lui accorde sa profonde miséricorde.*

*À ma famille*

*À mes amis(es)*

**A. Hamza**



## **REMERCIEMENTS**

---

Je tiens tout d'abord à exprimer une vive reconnaissance au professeur Bernard Bobée de l'INRS-Eau, titulaire de la chaire en Hydrologie Statistique - Hydro Québec, d'avoir accepté de diriger ce travail de recherche et d'avoir rendu possible la réalisation de ce mémoire.

Je remercie également mon codirecteur Taha B.M.J. Ouarda, d'avoir accepté de co-diriger ce travail ; ses conseils et son intérêt m'ont permis de persévérer et de mener à terme ce travail de recherche.

Je tiens à exprimer ma profonde gratitude à l'organisme INRS-Eau et aux professeurs Bernard Bobée et Taha B.M.J. Ouarda pour le soutien financier dont j'ai bénéficié.



## LISTE DES FIGURES

---

- Figure 2.1 Probabilités au non dépassement d'une loi normale centrée réduite.
- Figure 2.2 Illustration de la contribution des sites voisins à l'estimation régionale de  $X_T$ .
- Figure 3.1 Tracé des débits annuels  $Q_7$  de la rivière Coulonge (station 02kg001).
- Figure 4.1 Répartition géographique des 187 stations hydrométriques.
- Figure 4.2 Tracé des moments statistiques des débits d'étiage en fonction de la surface de drainage (187 sites).
- Figure 4.3 Ligne de régression des pentes (187 sites).
- Figure 4.4 Répartition des résidus sur une base géographique.
- Figure 4.5 Localisation géographique de 187 sites jaugés de la province de Québec.
- Figure 4.6a Tracé des moments statistiques des débits d'étiage en fonction de la surface de drainage pour la région homogène1 (108 sites).
- Figure 4.6b Ligne de régression des pentes correspondant à la région1 (108 sites).
- Figure 4.7a Tracé des moments statistiques des débits d'étiage en fonction de la surface de drainage pour la région homogène2 (79 sites).
- Figure 4.7b Ligne de régression des pentes correspondant à la région2 (79 sites).
- Figure 4.8 Relations entre les moments et les surfaces de drainage pour les deux régions homogènes.
- Figure 4.9a Relation entre le sixième moment de la région1 et la surface de drainage.
- Figure 4.9b Relation entre le sixième moment de la région2 et la surface de drainage.
- Figure 4.10a Comparaison du rapport  $Q_{10cens} / Q_{10emp}$  pour les différents percentiles.
- Figure 4.10b Erreur quadratique moyenne, Variance et Biais<sup>2</sup> pour les différents percentiles.
- Figure 4.11a Comparaison du rapport  $Q_{10cens} / Q_{10emp}$  pour les différents percentiles.
- Figure 4.11b Erreur quadratique moyenne, Variance et Biais<sup>2</sup> pour les différents percentiles.

- Figure 4.12 Comparaison entre  $Q_{10}$  estimé par le modèle de queue de Weibull et  $Q_{10}$  estimé par le modèle de Weibull complet.
- Figure 4.13 comparaison des droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$ .
- Figure 4.14 Équations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  (187 sites) avec Weibull censurée à 60%.
- Figure 4.15 Équations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 1 (108 sites) avec Weibull censurée à 60%.
- Figure 4.16 Équations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 2 (79 sites) avec Weibull censurée à 60%.
- Figure 4.17 Équations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  (187 sites) avec Lognormale censurée à 60%.
- Figure 4.18 Équations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 1 (108 sites) avec Lognormale censurée à 60%.
- Figure 4.19 Équations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 2 (79 sites) avec Lognormale censurée à 60%.
- Figure 4.20 Équations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  (187 sites) avec Weibull complète.
- Figure 4.21 Équations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 1 (108 sites) avec Weibull complète.
- Figure 4.22 Équations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 2 (79 sites) avec Weibull complète.
- Figure A1 Débits annuels  $Q_7$  de quelques sites de la Province de Québec.
- Figure A2 Débits annuels  $Q_7$  de quelques sites de la Province de Québec (suite).
- Figure A3 Répartition des résidus des moments 1, 2, 3, 4, 5 sur un plan géographique.
- Figure A4 Répartition du coefficient de variation des différentes régions considérées pour les 187 sites.
- Figure A5 Droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$  pour les différentes régions considérées (Weibull censurée à 60%).

- 
- Figure A6 Droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$  pour les différentes régions considérées (Lognormale censurée à 60%).
- Figure A7 Droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$  pour les différentes régions considérées (Weibull complète).
- Figure B1 Caractéristiques de l'invariance d'échelle simple en considérant les 187 sites comme échantillon homogène.
- Figure B2 Répartition des résidus des moments 1, 2, 3, 4, 5 et 6 sur un plan géographique.
- Figure B3 Répartition du coefficient de variation des différentes régions considérées pour les 187 sites.
- Figure B4 Distinction entre les différentes régions sur une base régressive.
- Figure B5 Estimation du seuil de censure  $u$  avec la distribution de Weibull.
- Figure B6 Estimation du seuil de censure  $u$  avec la distribution de Lognormale.
- Figure B7 Comparaison entre  $Q_{10}$  estimé par le modèle de queue de Weibull et  $Q_{10}$  estimé par le modèle de Weibull complet.
- Figure B8 Droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$  pour les différentes régions considérées (Weibull censurée à 50%).
- Figure B9 Droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$  pour les différentes régions considérées (Lognormale censurée à 50%).
- Figure B10 Droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$  pour les différentes régions considérées (Weibull complète).
- Figure B11 Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  (187 sites) avec Weibull censurée à 50%.
- Figure B12 Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 1 (117 sites) avec Weibull censurée à 50%.
- Figure B13 Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 2 (70 sites) avec Weibull censurée à 50%.

- Figure B14 Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  (187 sites) avec Lognormale censurée à 50%.
- Figure B15 Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 1 (117 sites) avec Lognormale censurée à 50%.
- Figure B16 Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 2 (70 sites) avec Lognormale censurée à 50%.
- Figure B17 Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  (187 sites) avec Weibull complète.
- Figure B18 Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 1 (117 sites) avec Weibull complète.
- Figure B19 Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 2 (79 sites) avec Weibull complète.
- Figure C1 histogramme et fonction de densité de probabilité.
- Figure C2 Fonction de distribution cumulée.

## **LISTE DES TABLEAUX**

---

Tableau 4.1 Estimation des paramètres représentant la relation entre les moments non centraux (MNC) des étiages et la surface du bassin versant.

Tableau 4.2 Sommaire des équations de régressions pour toute la province.

Tableau 4.3 Sommaire des équations de régressions pour région 1.

Tableau 4.4 Sommaire des équations de régressions pour la région 2.

Tableau A1 Quelques caractéristiques physiographiques et hydrologiques des 187 sites, avec la correspondance code provinciale vs code fédérale.

Tableau B1 Sommaire des équations de régressions pour toute la province.

Tableau B2 Sommaire des équations de régressions pour région 1.

Tableau B3 Sommaire des équations de régressions pour la région 2.



# TABLES DES MATIÈRES

---

<b>1. INTRODUCTION</b>	<b>1</b>
<b>1.1 INTRODUCTION</b>	<b>1</b>
<b>1.2 PROBLÉMATIQUE</b>	<b>4</b>
1.2.1 PROBLÉMATIQUE GÉNÉRALE (STATISTIQUE HYDROLOGIQUE)	4
1.2.2 PROBLÉMATIQUE STATISTIQUE	5
<b>1.3 OBJECTIFS</b>	<b>8</b>
<b>2. REVUE DE LITTÉRATURE</b>	<b>9</b>
<b>2.1 CONTRIBUTION À L'ESTIMATION LOCALE DES CARACTÉRISTIQUES D'ÉTIAGE</b>	<b>9</b>
2.1.1 DISTRIBUTIONS STATISTIQUES	10
2.1.1.1 Loi de Weibull à deux paramètres (W2)	10
2.1.1.2 loi Lognormale à deux paramètres (LN2)	11
2.1.2 ESTIMATION DE $X_T$ AUX SITES JAUGÉS	13
2.1.2.1 Méthode empirique	13
2.1.2.2 Méthodes statistiques	14
2.1.2.3 Analyse fréquentielle des débits d'étiage	17

<b>2.2 PREMIÈRES CONTRIBUTIONS À L'ESTIMATION RÉGIONALE DES CARACTÉRISTIQUES</b>	
<b>D'ÉTIAGE</b>	<b>21</b>
2.2.1 NOTION D'HOMOGENÉITÉ	22
2.2.2 MÉTHODE DE L'INDICE DE CRUE	24
2.2.3 MÉTHODES RÉGRESSIVES	26
<b>2.3 MÉTHODES ALTERNATIVES</b>	<b>28</b>
2.3.1 MODÈLES DÉTERMINISTES	28
2.3.2 HYDROGRAMME GÉOMORPHOLOGIQUE UNITAIRE	
(RODRIGUEZ-ITURBE ET VALDÉS, 1979)	29
2.3.3 MODÈLES MORPHOMÉTRIQUES	30
<b>3. DESCRIPTION DE LA MÉTHODOLOGIE UTILISÉE, COMBINAISON ENTRE INVARIANCE D'ÉCHELLE SIMPLE ET MODÈLES DE QUEUE (I.E.S/M.Q)</b>	<b>31</b>
<b>3.1 MOTIVATIONS À L'ORIGINE DE LA PROCÉDURE PROPOSÉE</b>	<b>31</b>
<b>3.2 THÉORIE DE L'INVARIANCE D'ÉCHELLE</b>	<b>34</b>
3.2.1 INVARIANCE D'ÉCHELLE SIMPLE	34
3.2.2 MODÈLE CASCADE LOGNORMAL	39
<b>3.3 THÉORIE DES MODÈLES DE QUEUE</b>	<b>45</b>
3.3.1 MODÈLE DE QUEUE CONDITIONNEL DE LA DISTRIBUTION WEIBULL	47
3.3.2 MODÈLE DE QUEUE CONDITIONNEL DE LA DISTRIBUTION LOGNORMALE	53

---

<b>4. APPLICATION DE LA MÉTHODE (I.E.S/M.Q)</b>	<b>59</b>
<b>4.1 DESCRIPTION DES DONNÉES ET PÉRIODES DE FONCTIONNEMENT</b>	<b>59</b>
<b>4.2 APPLICATION DE LA SIMILITUDE D'ÉCHELLE SIMPLE</b>	<b>62</b>
<b>4.3 DÉLIMITATION DES RÉGIONS HOMOGENES</b>	<b>64</b>
<b>4.4 APPLICATION DU MODÈLE DE QUEUE CONDITIONNEL DES DISTRIBUTIONS WEIBULL ET LOGNORMALE</b>	<b>75</b>
4.4.1 DÉTERMINATION DU SEUIL U	75
4.4.1.1 Estimation de u (avec la distribution Weibull censurée)	76
4.4.1.2 Estimation de u (avec la distribution Lognormale censurée)	78
4.4.2 ESTIMATION DE $X_T$ À UN SITE NON JAUGÉ	82
<b>5. CONCLUSIONS</b>	<b>97</b>
<b>6. RÉFÉRENCES</b>	<b>101</b>
<b>7. APPENDICES</b>	<b>107</b>



# 1. INTRODUCTION

---

## 1.1 Introduction

Les caractères hydroclimatologiques (tels que les précipitations, les crues et les étiages) d'une région montrent des variations *a priori* purement aléatoires bien qu'en fait elles soient centrées sur des moyennes dépendantes des conditions physiographiques, climatologiques et géographiques. Il est donc indispensable d'introduire en hydrologie, la statistique comme méthode procédant par analyse et synthèse, permettant de lier les paramètres statistiques des distributions aux caractéristiques naturelles étudiées, en leur donnant une signification physique.

L'estimation des événements extrêmes représente l'une des problématiques les plus importantes des sciences hydrologiques. Notons toutefois que l'événement extrême qui a fait l'objet du plus d'intérêt et qui a suscité de nombreuses études dans plusieurs perspectives, est le débit de crue. Ce phénomène naturel, est particulièrement crucial en raison des catastrophes qu'il peut occasionner :

- Risques d'inondations, pertes matérielles et surtout humaines en cas de sous estimation des débits de conception. En 1998, la crue du Yangtsé en Chine a causé des dégâts historiques engendrant plus de deux mille morts, des millions de sans abris et des dégâts matériels se chiffrant à plusieurs milliards de dollars. Au Québec, la crue du Saguenay, en 1996, a causé des dégâts matériels considérables et la perte de quelques vies humaines (deux). En 1993 au Etats-Unis, les crues exceptionnelles du Mississippi ont causé les pires dégâts de l'histoire de ce fleuve : Il y a eu une montée des eaux de près de 14 m, causant ainsi des pertes matérielles très importantes.

- Des coûts de construction élevés dans le cas d'une surestimation de ces débits.

Il n'en demeure pas moins que les phénomènes d'étiage suscitent de plus en plus d'intérêt. Que ce soient les agriculteurs, les ingénieurs hydrauliciens, les entreprises de production

d'énergie hydroélectrique ou les compagnies de navigation et même les consommateurs de la ressource-eau, tous s'entendent pour affirmer l'importance de mieux connaître le régime des rivières et des cours d'eau afin de prévoir et de combattre toute pénurie qui serait dommageable pour l'économie régionale et la santé publique. Bernier (1964) indique que *"Les besoins en eau du monde moderne qui croissent au rythme du développement économique imposent aux hommes d'avoir une connaissance aussi précise que possible du comportement des rivières en basses eaux. Il s'agit essentiellement d'une connaissance 'opérationnelle' propre à dégager des méthodes de prévisions des bas débits"*. La gestion et l'utilisation rationnelle des ressources en eau nécessitent une connaissance des écoulements aux sites où se posent les problèmes reliés à la gestion de la qualité de l'eau, à la navigation, à la production d'énergie hydroélectrique, ou à l'estimation de l'impact des conditions d'écoulement en basses eaux sur la vie aquatique. Par conséquent, les stations de jaugeages s'avèrent très importantes pour quantifier l'apport en eau afin d'évaluer l'intérêt d'un site du point de vue de son potentiel hydroélectrique, mais aussi pour des considérations sécuritaires et environnementales. Toute surestimation du débit d'étiage entraîne un déficit d'eau pour la centrale, et peut placer une entreprise dans une situation de défaillance énergétique, causer une augmentation des risques liés à la navigation, et engendrer des erreurs dans l'évaluation de la vulnérabilité de la vie aquatique ainsi que dans l'analyse de la qualité de l'eau. En pratique, on a recours à différentes techniques d'estimation des débits d'étiage, telle que l'ajustement de distributions statistiques à des débits minimums annuels. Cet outil statistique permet d'estimer les débits rares d'étiages et par conséquent évaluer le risque de rencontrer de tels événements hydrologiques extrêmes dans le futur. Il arrive fréquemment que les hydrologues soient amenés à estimer des caractéristiques d'étiage et particulièrement les valeurs de  $X_T$  débit d'une période de retour  $T$ , en des sites où l'on dispose de peu ou d'aucune information. Dans ce cas, on utilise une approche communément appelée estimation régionale, qui consiste à tirer profit de l'information hydrométéorologique mesurée à des sites voisins, où encore de la connaissance de caractéristiques physiographiques, que l'on peut facilement obtenir à un site non jaugé.

On définit au **chapitre 1**, après une introduction générale, la problématique reliée à l'estimation régionale des débits d'étiage, à savoir, les coûts engendrés par l'installation d'un réseau de stations de jaugeage de grande densité, l'estimation des débits de rivières dans des sites non jaugés et enfin la complexité générale reliée au phénomène d'étiage qui rend l'estimation des caractéristiques des débits minimums difficile. Une revue bibliographique permet de présenter au **chapitre 2**, dans un premier temps les méthodes d'estimation locale des caractéristiques des débits extrêmes, c'est à dire l'estimation de quantiles  $X_T$  en des sites jaugés. Par la suite, on présente une revue des méthodes de régionalisation utilisées depuis la proposition de l'approche indice de crue par Dalrymple (1949, 1960) pour estimer les caractéristiques des événements extrêmes et particulièrement les valeurs de  $X_T$ , débit de période de retour  $T$ , à un site non jaugé à partir d'informations obtenues à des sites géographiquement voisins. Benson (1962a) a ensuite proposé une alternative intéressante à la méthode de l'indice de crue en mettant en relation les caractéristiques physiographiques et climatiques avec  $X_T$ . L'utilisation de ces modèles régressifs devait devenir en Amérique du Nord la méthode de régionalisation la plus largement répandue. Bien que les approches citées, ci-dessus, aient été largement utilisées avec succès à travers le monde ces dernières décennies, il existe dans la littérature une classe de méthodes d'estimation régionale des caractéristiques des débits extrêmes basée sur une approche déterministe qui sont tout a fait différentes des approches proposées par Dalrymple et Benson. C'est notamment le cas pour les méthodes basées sur l'utilisation des modèles déterministes, et sur l'hydrogramme géomorphologique unitaire, qui sont brièvement décrites. Le **chapitre 3**, reprend dans le détail la description de la méthodologie utilisée nommée I.E.S/M.Q (acronyme pour méthode d'invariance d'échelle simple combinée à la méthode des modèles de queue), que l'on applique au **chapitre 4** aux données de débit d'étiage de 187 stations réparties dans la province de Québec, afin d'estimer  $X_T$  en des sites contenant peu de données ou aucune donnée. Cette approche consiste d'abord à s'assurer que les données de débits d'étiage du Québec suivent une invariance d'échelle simple, par la suite on applique la méthode des modèles de queue conditionnels, qui consiste à ajuster une distribution de probabilité aux données qui sont

inférieures ou égales à un certain seuil  $u$  (ceiling value), celui-ci est déterminé *a priori*. Deux modèles de queue conditionnels sont utilisés dans le présent travail (Durrans 1996) :

- Le modèle de queue conditionnel de la distribution Weibull, et
- Le modèle de queue conditionnel de la distribution Lognormale

On définit, finalement, pour chaque région une équation régionale  $Q_T$  en fonction d'une variable physiographique qui sera dans le présent contexte la surface de drainage du bassin versant. Notons qu'à notre connaissance, la surface de drainage est la seule variable physiographique qui a été utilisée dans ce type d'étude.

## 1.2 Problématique

### 1.2.1 Problématique générale (statistique hydrologique)

Les coûts importants associés à l'installation et au maintien de stations de mesures ont conduit en période de coupures budgétaires à la réduction de la densité du réseau de stations de jaugeage. Souvent les hydrologues doivent estimer le débit d'une rivière en un site où l'on ne dispose d'aucune information (site non jaugé) ou de peu d'information (site jaugé depuis peu de temps ou qui ne l'est plus), cela peut se produire en des sites où l'on projette d'ériger des structures (station hydroélectrique), d'effectuer des travaux (navigation) ou d'entreprendre des études pour la gestion de la qualité de l'eau. Il s'avère donc essentiel de disposer d'outils adéquats permettant d'estimer les événements extrêmes à un site particulier. Pour pallier au problème du manque d'information hydrologique, les hydrologues utilisent l'approche d'estimation régionale.

Cette approche consiste à utiliser l'information disponible à des stations appartenant à une même région hydrologique pour faire une estimation régionale des débits extrêmes (dans notre cas débits d'étiage) à un site où l'on dispose de peu ou d'aucune information. La première approche d'estimation régionale des crues (indice de crue) a été proposée dès 1949 par Dalrymple, qu'on distingue par trois grandes étapes:

- 1- Définition d'une région homogène sur une base géographique
- 2- Ajustement de distributions aux données de sites appartenant à cette région ; et
- 3- Développement à l'intérieur de la région homogène d'une relation régionale pour l'estimation de  $X_T$  aux sites non jaugés ;

Toutefois, depuis quelques années, plusieurs auteurs ont mis en évidence les faiblesses de cette approche et ont proposé des méthodologies alternatives intervenant à différents niveaux de ces procédures classiques, notamment en ce qui concerne la définition des régions homogènes et l'estimation de  $X_T$ . En effet, le choix de régions géographiques homogènes est discutable car la proximité géographique des bassins versants n'est pas à elle seule une garantie de l'homogénéité des caractéristiques hydrologiques. Aussi le concept de relation régionale conduit à accepter l'hypothèse que la même distribution statistique est adéquate pour tous les sites d'une même région homogène. Dalrymple (1960) a considéré la loi Gumbel (à 2 paramètres) pour représenter le débit de crue ce qui revient à admettre l'hypothèse qu'en tous les sites, on a une même valeur du coefficient de variation et d'asymétrie. Dans le cadre d'un projet commun du CRSNG, l'équipe de recherche de la chaire en hydrologie statistique en collaboration avec d'autres chercheurs ont recensé des méthodologies les plus performantes concernant l'estimation régionale des crues (Bobée, 1994).

### 1.2.2 Problématique statistique

Parmi les difficultés que peut rencontrer l'utilisateur dans la pratique de l'étude des étiages, on retrouve souvent, le choix de la variable d'étiage à retenir. Contrairement aux débits de crue qui sont un phénomène qui peut être représenté par un débit maximal instantané (ou sur une courte période, journalière par exemple). La préhension du phénomène d'étiage est beaucoup plus délicate, l'effet des multiples prélèvements, dérivations et rejets dont est l'objet un bief fluvial est considérable, mais surtout l'effet de la durée des périodes d'étiage influe considérablement sur le choix de la variable. Ce sont des interventions humaines qui sont généralement responsables des perturbations du débit d'étiage. Dans de telles

conditions le débit instantané n'a aucun sens, et l'on préfère définir l'étiage par : Un débit minimum annuel ou saisonnier de 1, 7, 14, 30, etc. jours consécutifs. L'estimateur des débits d'étiage le plus largement utilisé en Amérique du Nord est  $Q_{7,10}$  « l'étiage de sept jours consécutifs et de période de retour de dix ans », qui se définit comme le minimum annuel lissé sur sept jours consécutifs des débits journaliers, qui devrait être dépassé en moyenne à chaque dix ans (Riggs, 1980). En pratique, le peu de données d'étiage disponible ainsi que la complexité générale reliée au phénomène des basses eaux, rendent la détermination de l'estimateur du quantile d'étiage difficile. Cette complexité a historiquement forcé les hydrologues à utiliser des distributions de probabilité théorique à deux ou trois paramètres. De tels modèles peuvent produire des estimateurs de quantiles biaisés, et le nombre relativement grand des paramètres cause une large variance des estimateurs. Ceci arrive en général quand on traite un échantillon de petite taille. L'approche par régionalisation a été introduite pour surmonter le problème de manque de données, mais surtout pour estimer des événements extrêmes en des sites où aucune information hydrométrique n'est disponible (sites non jaugés).

Une des problématiques rencontrées dans la procédure classique de régionalisation, concerne le choix et l'ajustement d'une distribution statistique à l'échantillon. En effet, après avoir choisi et traité un échantillon de débits minimums annuels en respectant certaines hypothèses de bases (débit naturel, homogénéité, stationnarité, etc.), on doit procéder au choix de la distribution statistique qui s'ajuste le plus adéquatement à l'échantillon de débits minimums annuels. Pour ce faire plusieurs approches sont proposées :

- 1) Utilisation des tests statistiques d'adéquation (test du chi carré...etc.)
- 2) Utilisation des diagrammes de rapport de moments (Bobée et al, 1993)
- 3) Ce choix peut aussi reposer sur l'examen visuel d'un graphique où l'on trace les observations en fonctions de leur probabilités empirique et sur lequel on superpose les courbes théoriques de lois ajustées en estimant les paramètres à partir de l'échantillon (Riggs, 1972).

Alors que les deux premiers outils permettent à l'utilisateur d'éliminer les distributions nettement inadéquates, sans toutefois indiquer parmi les lois applicables celle qui convient le mieux. L'examen visuel, qui est plus subjectif, s'avère un outil indispensable lors du choix de la distribution statistique. Le *United States Geological Survey (U.S.G.S)*, admet que la distribution empirique des débits d'étiages est souvent complexe (Matalas, 1963). Plusieurs distributions de probabilités (Lognormale, Weibull, Gamma, etc.) ont été proposées comme étant appropriées pour l'analyse fréquentielle des débits minimums. En général l'ajustement est fait sur la base de toutes les données d'enregistrement recueillies à un site. L'utilisation de toutes les données disponibles est logique dans le cas où la distribution de la population apparaît « robuste » (c'est à dire est une distribution unimodale et raisonnablement bien décrite par les distributions classiques). Mais dans plusieurs cas, les populations des débits d'étiage apparaissent être composées de plus d'une distribution de probabilité (distributions mixtes). Des observations de cette nature ont été rapportées par plusieurs auteurs dont Riggs (1972). Comme pratique standard, le *U.S.G.S.* emploie l'approche par estimation graphique dans le cas où la distribution Log-pearson type3 n'ajuste pas de manière satisfaisante la série de données Riggs (1972). Toutefois, l'utilisation de l'estimation graphique n'est pas entièrement satisfaisante, essentiellement à cause de la subjectivité associée à l'examen visuel.

Devant une telle problématique, plusieurs modèles doivent être combinés pour traiter le problème du mélange de populations. Toutefois, il est difficile de sélectionner des distributions susceptibles d'être employées dans des modèles mixtes. Cela nécessite de considérer des distributions statistiques ayant un nombre élevé de paramètres, ce qui a pour effet d'augmenter les difficultés liées à l'estimation numérique des paramètres, et aussi d'induire des estimateurs de quantiles ayant une grande variance. Le développement et l'application d'approches alternatives pour l'estimation régionale des débits d'étiage apparaît donc utile et nécessaire.

## 1.3 Objectifs

L'objectif de cette recherche est de faire une estimation régionale des données de débits d'étiage de la province de Québec, en appliquant une procédure qui combine les méthodes d'invariance d'échelle simple et d'analyse des débits d'étiage avec les modèles de queue conditionnels. Les équations de régression régionale obtenues par cette méthode seront comparées à celles obtenues par des méthodes pour lesquelles l'ajustement des distributions de probabilités sera effectué sur l'ensemble des données. Durrans(1996) suggère que la méthode de l'estimation des étiages par les modèles de queue peut venir à bout de certaines difficultés rencontrées dans l'analyse fréquentielle des débits d'étiages.

Miller (1996) indique que « Si l'hypothèse d'invariance d'échelle simple est vérifiée pour les débits d'étiage alors la méthode de l'indice d'étiage peut-être utilisée pour la régionalisation des étiages en considérant la valeur du seuil  $u$  comme indice d'étiage». L'idée principale de la méthode est de s'intéresser uniquement à la partie inférieure de la distribution (d'où le nom queue de distribution), qui est la section où les débits sont les plus critiques et où l'on désire connaître les quantiles ayant des périodes de retour élevées.

Cette approche implique la sélection *a priori* d'un « niveau plafond », essentiellement un seuil qui définirait la limite supérieure de l'intervalle de données qui seraient employées, la deuxième étape consiste à ajuster une distribution conditionnelle sur les données situées en dessous de ce seuil. La description complète de la méthode sera présentée ultérieurement (cf. 3.3).

## 2. REVUE DE LITTÉRATURE

---

Dans un premier temps on expose les méthodes d'estimation locale des caractéristiques des débits extrêmes, en d'autres termes l'estimation de  $X_T$ , débit de période de retour  $T$ , à des sites jaugés.

Compte tenu du fait que l'information hydrométrique est souvent peu abondante sinon inexistantes à des sites où l'on souhaite connaître  $X_T$ , on présente ensuite les méthodes classiques d'estimation régionale des crues proposées par Dalrymple (1949,1960) et Benson (1962) ; ces méthodes ont été adaptées par certains auteurs tel que Pilon (1990) et Vogel et Kroll (1990) pour l'étude des caractéristiques d'étiage.

### 2.1 Contribution à l'estimation locale des caractéristiques d'étiage

L'estimation de  $X_T$  à des sites jaugés a nécessité plusieurs contributions significatives, tant au niveau du développement de distributions statistiques et de leurs méthodes d'ajustement, qu'au niveau des distributions non paramétriques où des modèles de dépassement. Les résultats de ces travaux permettent d'estimer avec davantage de précision les événements  $X_T$  de grandes périodes de retour  $T$  à des sites jaugés. Une définition mathématique d'une distribution statistique, ainsi que des méthodes d'ajustement des paramètres sont reportées en annexe C.

## 2.1.1 Distributions statistiques

### 2.1.1.1 Loi de Weibull à deux paramètres (W2)

La f.d.p d'une variable aléatoire X distribuée selon la loi de Weibull est :

$$f(x) = \frac{c}{\alpha} \left[ \frac{x}{\alpha} \right]^{c-1} \exp \left[ - \left( \frac{x}{\alpha} \right)^c \right], \quad x > 0 \quad (2.12)$$

avec les paramètres  $c > 0$  et  $\alpha > 0$ , lorsque  $c = 1$ , la distribution de Weibull devient une distribution exponentielle.

La fonction de distribution d'une variable aléatoire X distribuée selon la loi de Weibull est :

$$F(x) = 1 - \exp \left[ - \left( \frac{x}{\alpha} \right)^c \right], \quad x > 0 \quad (2.12a)$$

Le moment non centré d'ordre r de la variable X est : (Johnson et Kotz, 1970)

$$\mu'_r(x) = E[X^r] = \alpha^r \Gamma \left[ 1 + \frac{r}{c} \right] \quad (2.13)$$

On peut déduire le moment centré  $\mu_r(x)$  d'ordre r à partir des moments non centrés  $\mu'_r(x)$ . Il suffit d'appliquer la relation:

$$\mu_r(x) = E \left[ (x - \mu_1)^r \right] = \int_{-\infty}^{\infty} (x - \mu_1)^r f(x) dx \quad (2.14)$$

La moyenne et la variance de X sont telles que :

$$\mu_r = \alpha \Gamma\left(1 + \frac{1}{c}\right) \quad (2.15)$$

$$\mu_2 = \alpha^2 \left[ \Gamma\left(1 + \frac{2}{c}\right) - \Gamma\left(1 + \frac{1}{c}\right)^2 \right] \quad (2.16)$$

Les estimateurs du maximum de vraisemblance sont obtenus en maximisant la fonction :

$$\ln L(\alpha, c) = n \ln(c) - n \ln(\alpha) - \sum_{i=1}^n \left(\frac{x_i}{c}\right)^c + (c-1) \sum_{i=1}^n \left(\frac{x_i}{\alpha}\right) \quad (2.17)$$

Après quelques calculs, nous obtenons le système suivant :

$$\alpha = \left( \frac{1}{n} \sum_{i=1}^n x_i^c \right)^{1/c} \quad (2.17a)$$

$$c = \left[ \left( \sum_{i=1}^n x_i^c \ln x_i \right) \left( \sum_{i=1}^n x_i^c \right)^{-1} - \frac{1}{n} \sum_{i=1}^n \ln x_i \right]^{-1} \quad (2.17b)$$

### 2.1.1.2 loi Lognormale à deux paramètres (LN2)

Si la variable aléatoire  $Y = \ln(x)$  suit une distribution normale  $N(\mu_y, \sigma_y^2)$  alors  $X$  est distribué selon une loi Lognormale  $LN(\mu_y, \sigma_y^2)$  et sa f.d.p est donnée par :

$$f(x) = \frac{1}{x \sigma_y \sqrt{2\pi}} \exp \left[ - \frac{[\ln(x) - \mu_y]^2}{2\sigma_y^2} \right], \quad x > 0 \quad (2.18)$$

Le moment non-centré d'ordre  $r$  est donné par :

$$\mu'_r = \exp\left(r\mu_y + r^2 \frac{\sigma_y^2}{2}\right) \quad (2.19)$$

La moyenne ainsi que la variance sont déduites de l'équation précédente et prennent les expressions suivantes :

$$\mu'_1 = \exp\left(\mu_y + \frac{\sigma_y^2}{2}\right) \quad (2.20)$$

$$\mu_2 = \exp(2\mu_y + \sigma_y^2) \left(\exp(\sigma_y^2) - 1\right) \quad (2.21)$$

L'estimation des paramètres par la méthode du maximum de vraisemblance est obtenue en maximisant la fonction :

$$\ln L(\mu_y, \sigma_y^2) = -\frac{n}{2} \ln(2\pi\sigma_y^2) - \sum_{i=1}^n \ln(x_i) - \frac{1}{2\sigma_y^2} \sum_{i=1}^n [\ln(x_i) - \mu_y]^2 \quad (2.22)$$

Le maximum de  $\ln L(\mu_y, \sigma_y^2)$  est obtenu en résolvant le système d'équations :

$$\frac{\partial \ln L}{\partial \mu_y} = \frac{1}{\sigma_y^2} \sum_{i=1}^n [\ln(x_i) - \mu_y] = 0 \quad (2.23a)$$

$$\frac{\partial \ln L}{\partial \sigma_y} = -\frac{n}{\sigma_y} + \frac{1}{\sigma_y^3} \sum_{i=1}^n [\ln(x_i) - \mu_y]^2 = 0 \quad (2.23b)$$

La solution du système est :

$$\hat{\mu}_y = \frac{1}{n} \sum_{i=1}^n \ln(x_i) \quad (2.24)$$

$$\hat{\sigma}_y = \frac{1}{n-1} \sum_{i=1}^n [\ln(x_i) - \mu_y]^2 \quad (2.25)$$

L'estimateur de  $\sigma_y^2$  est multiplié par la quantité  $\frac{n}{n-1}$  pour tenir compte du biais.

Notons que la méthode du maximum de vraisemblance pour la loi Lognormale est équivalente à la méthode du maximum de vraisemblance appliquée à la loi normale en considérant un échantillon dont les données sont transformées en logarithme népérien ( $y_i = \ln x_i$ ).

## 2.1.2 Estimation de $X_T$ aux sites jaugés

### 2.1.2.1 Méthode empirique

On peut déterminer empiriquement la probabilité au non dépassement de certains événements hydrologiques mesurés, ou encore la magnitude d'un débit associé à une probabilité au non dépassement donnée, en supposant que l'on dispose d'une série de débits journaliers  $Q_j$  pour  $j$  jours à un site donné. En terme d'étiage annuel, si pour chacune des années, on extrait des séries de débits journaliers la valeur du débit minimum annuel  $Q_x$ , et qu'on ordonne chacune des valeurs minimales annuelles on peut estimer la probabilité d'un événement minimum annuel  $Q_T$  de période de retour  $T$  nécessairement inférieur au nombre d'années échantillonnées  $N$ , car c'est une approche basée sur le calcul des probabilités

empiriques. Cette approche empirique ne permet donc pas l'inférence à des valeurs de  $Q_T$  inférieures ou plus rares que la plus petite valeur de  $Q_x$  sur la période d'échantillonnage.

### 2.1.2.2 Méthodes statistiques

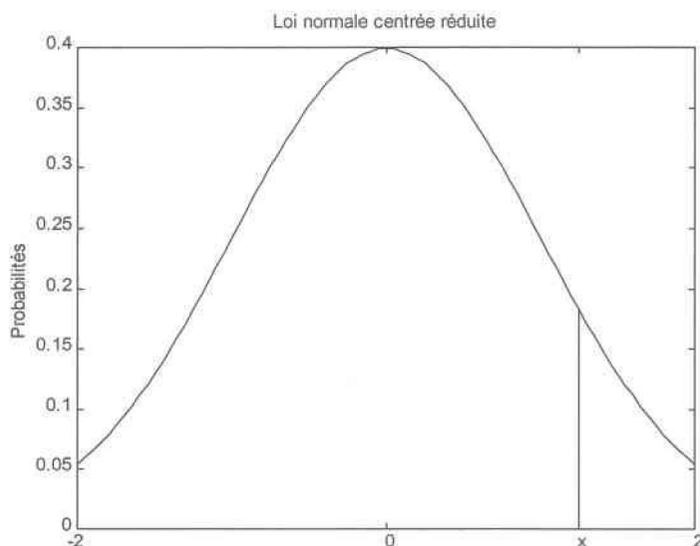
#### *Analyse de fréquence*

L'analyse fréquentielle des débits extrêmes annuels journaliers, représente un des outils statistiques les plus employés en hydrologie. Cette méthode consiste à ajuster une loi statistique à un échantillon d'observations caractérisant l'étiage, afin d'évaluer la relation entre l'estimation  $Q_T$  du débit de période de retour  $T$  et la probabilité au non dépassement qui lui est associée (voir figure 2.1) définie par le rapport  $p = 1/T$ . Il serait utile de donner une définition de la probabilité au non dépassement, d'un quantile, et de la période de retour :

$Q_T$  = Débit qui revient en moyenne toutes les  $T$  années.

$T$  = représente la période moyenne (en années) qui sépare l'occurrence des événements  $X$  supérieurs au seuil  $x$ .

Probabilité au non dépassement de  $x$  : c'est la probabilité qu'un événement ne dépasse pas la valeur  $x$ , qui représente dans notre cas un débit.



**Figure 2.1 Probabilités au non dépassement d'une loi normale centrée réduite**

Les événements rares sont ceux dont la période de retour est la plus grande, c'est à dire la probabilité au non dépassement la plus faible comme le montre la figure 2.1.

**i) Choix et traitement préliminaire d'un échantillon de débits minimums annuels qui respectent certaines hypothèses de bases.**

Après avoir extrait d'une série de débits journaliers mesurés les débits minimums pour chaque année, on doit s'assurer que la série des minimums annuels ainsi formée respecte certaines hypothèses de base afin que l'analyse de fréquence soit théoriquement acceptable. Ces hypothèses sont :

a/ *Débit naturel*, c'est à dire que les fluctuations sont naturelles et non pas induites artificiellement (rivière régularisée, débit de sortie d'un réservoir contrôlé, etc.). notons que dans notre cas on a toléré les débits qui sont influencés de manière journalière car on considère qu'un débit contrôlé de manière journalière n'influence pas la justesse des résultats.

b/ *Indépendance des observations de l'échantillon* (ordonné de manière chronologique), c'est à dire qu'aucun débit de la série n'est conditionné par les valeurs précédentes ou n'influence les débits suivants. Généralement les valeurs minimums annuelles de débit ne sont pas autocorrélées, ce qui n'est pas le cas pour un échantillon de valeurs journalières ou hebdomadaires où chaque observation dépend des précédentes et est reliée aux observations suivantes.

c/ *Homogénéité de l'échantillon*, c'est à dire que toutes les observations proviennent de la même population statistique.

d/ *Stationnarité de la série*, une série de données est dite stationnaire, si en dépit des fluctuations aléatoires dans les valeurs classées chronologiquement, on ne retrouve pas de variations significatives dans le temps. Dans les séries hydrologiques, ces variations temporelles peuvent être associées à la construction d'un ouvrage hydraulique ou au déplacement de la station hydrométrique, ce qui peut entraîner une discontinuité marquée dans la série (saut de la moyenne, ...etc.). la non stationnarité d'une série peut-être aussi associée à un changement dans l'utilisation du sol d'un bassin versant (déforestation, urbanisation, ...etc.), ce qui se traduit par une augmentation ou une diminution graduelle des valeurs de débits minimums annuels d'une série.

e/ *Valeurs singulières*, il arrive parfois que des erreurs grossières de mesure ou de transcription se produisent. Lorsqu'on détecte une valeur singulière à un niveau de signification donné, on doit s'assurer, avant de la corriger ou de la retrancher de l'échantillon, que cette donnée ne soit pas une valeur extrême engendré par une combinaison de causes météorologiques réelles enregistrées.

## ii) **Choix et ajustement d'une distribution statistique à l'échantillon.**

Une fois que les tests qui permettent de vérifier les cinq hypothèses ci-dessus, ont été effectués, on doit procéder au choix de la distribution statistique qui s'ajuste le plus fidèlement possible à l'échantillon de débits minimums annuels, et de la méthode d'estimation appropriée, ce qui permettra éventuellement d'estimer les débits  $X_T$ , plus rares que la plus petite valeur de la série.

### iii) Utilisation de la distribution ajustée pour estimer les débits $X_T$ de récurrence $T$ .

Après avoir retenu la distribution statistique et choisi la méthode d'ajustement, certaines caractéristiques statistiques de l'échantillon sont prises en compte de façon à estimer la distribution des débits minimums annuels. Cette relation permet ensuite d'extrapoler  $X_T$  pour des périodes de retour plus grandes. Des détails sur l'ajustement d'une distribution statistique sont présentés à l'annexe C.

### 2.1.2.3 Analyse fréquentielle des débits d'étiage

L'objectif principal des études d'analyse de fréquence, est d'assigner à des valeurs de débit une probabilité au non dépassement  $p$  ou période de retour  $T = 1/p$ . Et pour ce faire, deux approches peuvent être considérées, la première est une approche directe qui utilise les fonctions de répartition empiriques, et la seconde consiste en l'ajustement de distributions théoriques. Et ce, afin de calculer les débits d'étiage critiques  $X_T$  correspondant à une certaine période de retour  $T$ , définie comme étant le temps moyen entre deux occurrences d'étiages. Les études effectuées ont portées sur le minimum annuel du débit journalier ou les données hydrologiques sont supposées indépendantes et identiquement distribuées.

Parmi les difficultés que peut rencontrer l'utilisateur dans la pratique de l'étude des étiages, on trouve souvent le choix de la variable d'étiage à retenir. L'estimateur des débits d'étiage le plus largement utilisé en Amérique du Nord est le  $Q_{7,10}$ , qui correspond au minimum annuel lissé sur 7 jours consécutifs ayant une période de retour de 10 ans. Il est aussi d'usage d'analyser le minimum annuel du débit moyen sur une période supérieure à 7 jours consécutifs selon l'intensité de l'étiage enregistré. Ces périodes peuvent être 10, 30, 50, 60, 90 ou 120 jours. Dans le cadre de cette étude, la variable d'étiage  $Q_{7,10}$  a été choisie comme première variable d'étude *a priori*. Et pour des fins de comparaison des résultats, nous avons décidé d'appliquer la méthodologie du présent travail à une deuxième variable, dont le choix a été fait après consultation avec le Ministère de l'Environnement et de la Faune Québec, nous avons convenu de choisir la variable  $Q_{30,5}$  qui est la variable d'étiage la plus

souvent utilisée après la variable  $Q_{7,10}$  par cet organisme, et qui définit le minimum annuel de 30 jours consécutifs pour une période de retour de 5 ans.

Riggs (1972) décrit la procédure à suivre afin d'obtenir les caractéristiques fréquentielles des débits d'étiage. Il montre comment les caractéristiques de certains bassins influencent la moyenne et la variabilité annuelle des débits d'étiage.

Loganathan et al.(1986) classent les méthodes d'estimation des caractéristiques d'étiage en quatre catégories :

1/ La méthode de transformation, où les données originales sont normalisées et la loi normale est ensuite appliquée sur les données transformées.

2/ La distribution théorique, où une loi est ajustée directement aux données.

3/ La distribution empirique, où les données sont ordonnées et tracées selon une formule de probabilité empirique (plotting position) qui est utilisée pour estimer les probabilités de dépassement. On retrouvera au chapitre 4 une définition plus détaillée d'une probabilité empirique (cf. 4.41).

4/ Les méthodes de probabilités à base physique qui utilisent les relations hydrologiques et des distributions théoriques des variables aléatoires hydrologiques dépendantes.

Gumbel (1954) a été le premier à appliquer la théorie des valeurs extrêmes aux événements d'étiage et en particulier la loi de Weibull à trois paramètres (un paramètre de forme, un paramètre d'échelle et un paramètre de position), l'estimation doit être telle que l'estimateur du paramètre de position soit inférieur au plus faible débit minimum observé.

Matalas (1963) a effectué une étude de comparaison de quatre distributions théoriques asymétriques à trois paramètres, ayant une limite inférieure : la loi Gumbel, la loi Lognormale à trois paramètres (LN3), la loi Pearson type 3, et la loi Pearson type 5. Il en a conclu que la loi Gumbel et la loi Pearson type 3 s'appliquent mieux aux débits minimums annuels de 34 jours d'eau, avec comme méthode d'estimation des paramètres, la méthode

des moments, et ce, selon deux critères : la limite inférieure estimée ne doit pas être inférieure à la limite observée, et la relation entre le coefficient d'asymétrie et le coefficient d'aplatissement de la distribution théorique doit ressembler à cette même relation présente dans les données observées.

Joseph (1970) a effectué une étude de comparaison de quatre distributions théoriques : la loi Gamma, la loi Lognormale à deux paramètres (LN2), la loi Normale, et la loi de Weibull, appliquées au débit minimum annuel de la moyenne mobile sur 14 jours des débits journaliers et ce, pour 37 stations de mesure dont les données sont supposées indépendantes. La méthode du maximum du vraisemblance (MV) est appliquée pour estimer les paramètres de ces quatre lois. Il en a conclu que la loi Gamma suivie de la loi de Weibull et de la loi LN2 sont les plus adéquates. Les critères de comparaison sont le test d'ajustement du khi-deux, et la comparaison des relations des coefficients d'asymétrie et d'aplatissement.

Condie et Nix (1975) ont appliqué la loi Pearson type3, la loi Pearson type5, la loi Weibull et la loi LN3 à l'analyse des faibles débits de 38 rivières Canadiennes. Les débits journaliers minimaux sont utilisés, et afin d'assurer l'indépendance, l'année hydrologique débute le premier juin. Les paramètres sont estimés par successivement la méthode du maximum de vraisemblance et la méthode des moments. D'après leurs résultats, ils ont conclu que la loi Weibull est la plus adéquate (estimation par le maximum de vraisemblance), suivie par la loi Weibull (estimation par la méthode du plus faible débit observé). La loi est rejetée si la limite inférieure estimée est supérieure au plus faible débit minimum observé. Un avantage de la loi Weibull est qu'elle peut représenter les données qui ont un coefficient d'asymétrie négatif.

Au Québec, Hoang (1978) a trouvé que la loi log-Pearson type3 (LP3) s'applique le mieux aux minima annuels des débits journaliers pour l'ensemble de toutes les stations de mesure de la province de Québec.

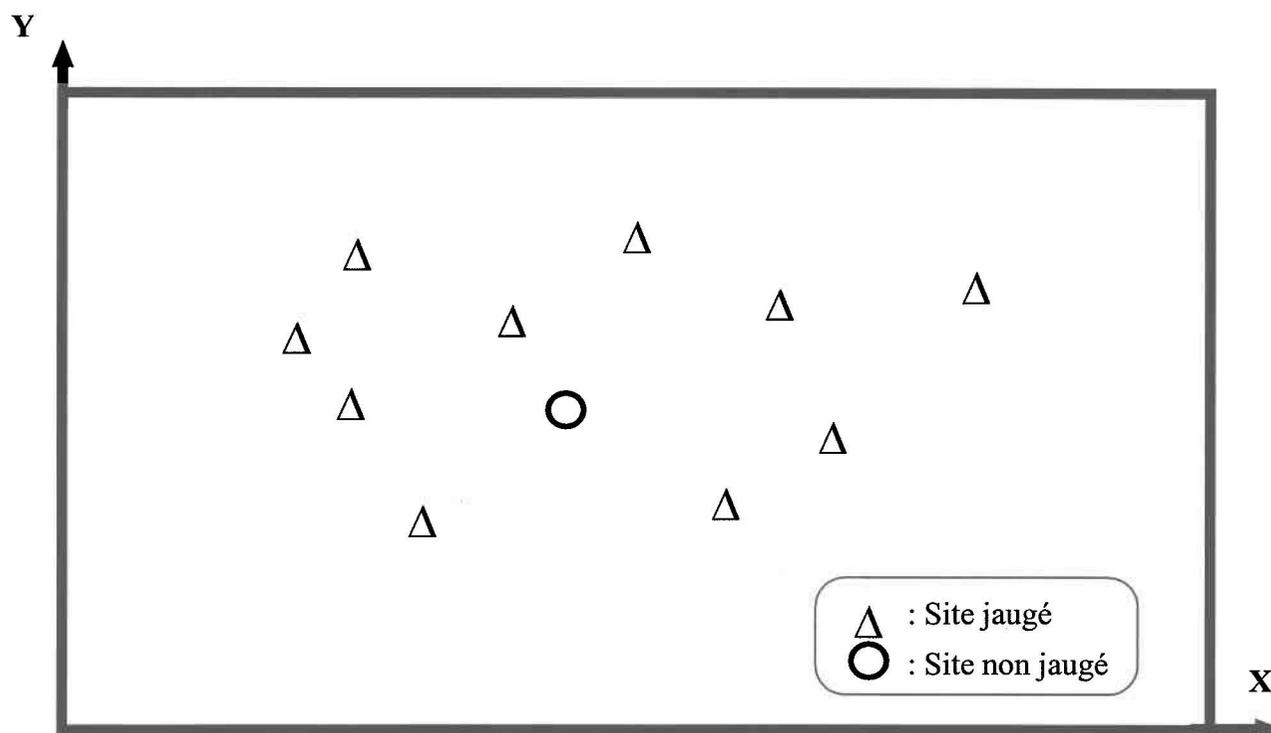
Tasker (1987) a utilisé la méthode du bootstrap (rééchantillonnage dans les distributions empiriques observées) afin d'estimer  $Q_7$  pour les périodes de retour de 10ans et de 20ans. Les méthodes basées sur une distribution théorique telle que la loi Weibull ou la loi LP3 donnent une erreur quadratique moyenne plus faible que les méthodes basées sur une distribution empirique.

Nathan et Mc Mahon (1990) ont étudié les méthodes d'estimation des paramètres de la loi Weibull pour le débit minimum annuel ainsi que pour les débits de  $m$  jours consécutifs  $Q_m$ ,  $m = 1, 7, 15, 30, 60, 120, 183, 284$ . Les trois méthodes sont, moments, maximum de vraisemblance et moments de probabilités pondérés. Ils ont trouvé que la méthode des moments de probabilités pondérés donnait des estimateurs peu satisfaisants pour les sécheresses de courtes durées. Quand à la méthode du maximum de vraisemblance produit des estimateurs de volumes de sécheresse qui sont souvent plus grand que ceux obtenus avec les deux autres méthodes. Finalement, ils recommandent l'utilisation de plus d'une méthode surtout pour les petits échantillons.

Vogel et Kroll (1989) ont utilisé la méthode intitulée « Probability-plot correlation coefficient (The PPCC hypothesis test) » qui est un test permettant d'évaluer les hypothèses de distribution de probabilité régionale des débits d'étiage en considérant que les sites sont indépendants entre eux, et retenir parmi ces distributions, celles qui s'ajustent le plus fidèlement aux données de débits d'étiage considérés. Aussi, pour distinguer parmi les procédures d'estimation des paramètres, celles qui performant le mieux. Originellement le PPCC a été développé pour tester les distributions à deux paramètres, telles que, la loi Normale, Lognormale (LN2), Weibull (W2), et la loi de Gumbel à deux paramètres. Ils ont comparé le PPCC pour les distributions Lognormale à deux et trois paramètres, Weibull à deux et trois paramètres, et la distribution log-Pearson type3, pour les séquences de débit minimum annuel  $Q_7$ , appliquée à 23 sites au Massachusetts. Ils ont montré en considérant un PPCC régional, que la distribution la plus adéquate pour représenter les minimums annuels  $Q_7$  au Massachusetts est la loi Lognormale à deux paramètres.

## 2.2 Premières contributions à l'estimation régionale des caractéristiques d'étiage

Les approches décrites précédemment impliquent que l'on dispose d'informations hydrologiques aux sites où l'on souhaite connaître la probabilité de rencontrer des événements hydrologiques extrêmes. Or il arrive que l'on doive estimer des événements extrêmes à des sites où aucune information hydrométrique n'est disponible, dans ce cas, il est nécessaire de tirer profit de l'information hydrométéorologique mesurée à des sites voisins, ou encore de la connaissance de caractéristiques physiographiques, que l'on peut facilement obtenir à un site non jaugé. La figure 2.3, illustre la contribution des sites voisins à l'estimation de  $X_T$  à un site non jaugé. Les stations sont positionnées dans un espace géographique ( $x$  = longitude,  $y$  = latitude), où l'information hydrométrique est disponible. La contribution de l'information des sites voisins à l'estimation de  $X_T$  permet d'établir une relation régionale moyenne qui correspond aux caractéristiques d'une station virtuelle qui serait située au centre de gravité de la région homogène. Pour cela, la détermination d'une région homogène doit être préalablement bien définie. Nous présentons dans ce qui suit une brève revue sur les méthodes récentes traitant de la notion d'homogénéité, ainsi que les méthodes d'estimation du quantile  $X_T$  à un site où l'on ne dispose d'aucune information hydrologique, les plus largement utilisées en hydrologie.



**Figure 2.2** Illustration de la contribution des sites voisins à l'estimation régionale de  $X_T$

### 2.2.1 Notion d'homogénéité

Dans cette section, nous présentons une revue des méthodes de délimitation des régions homogènes, établies originellement pour les crues, et que l'on envisage par la suite leur application aux étiages. La détermination des régions homogènes consiste à regrouper, pour une variable hydrologique donnée, les stations ayant le même comportement hydrologique. L'abondance de travaux récents portant sur la notion d'homogénéité régionale témoigne du caractère important de cette composante des procédures classiques de régionalisation. La définition *a priori* des sites faisant partie d'une région homogène selon des critères géographiques uniquement est souvent utilisée, mais cette approche n'est pas acceptable. En effet, la proximité géographique des bassins versants n'est pas à elle seule une garantie de l'homogénéité des caractéristiques hydrologiques, bien que la distribution spatiale de la

précipitation ait elle-même un caractère géographique. En effet, pour un bassin versant, la taille, la forme, le réseau de drainage, les caractéristiques du sol (lithologie) et de sa couverture végétale ainsi que le régime climatologique, s'avèrent appropriés comme critères de regroupement de bassins hydrologiquement homogènes. Parmi les nouvelles méthodes permettant de délimiter, pour le débit maximum annuel, les régions homogènes à l'intérieur desquelles les sites seraient considérés comme hydrologiquement homogènes, on trouve celles où l'on propose de considérer les caractéristiques physiographiques (Wiltshire, 1985) ou hydrologiques (Burn, 1989) des bassins versants. La définition de régions homogènes à partir de ces caractéristiques est basée sur le positionnement des sites dans des dimensions physiographiques, climatologiques ou hydrologiques. La région devient ainsi plus homogène du point de vue hydrologique (Bobée et al, 1994). Plus récemment on remettait en question la notion d'homogénéité régionale pour proposer le concept de « région d'influence » (Burn, 1989). Selon cette approche, plutôt que de définir *a priori* une région homogène à l'intérieur de laquelle il sera possible de transférer l'information souhaitée, on détermine pour chaque site non jaugé, un ensemble de stations pouvant être utilisées pour transférer l'information relative aux caractéristiques de crue au site d'intérêt. Nguyen et Pandey (1996), ont défini une nouvelle approche pour la délimitation des régions homogènes lors d'une étude pour l'estimation régionale des débits de crue au Québec. Cette procédure est basée sur le principe de la similarité des propriétés statistiques d'échelle des débits maximums annuels de chaque région. Concrètement, les auteurs ont trouvé que les moments non centraux d'ordre un à six suivaient une invariance d'échelle simple par rapport à la superficie du bassin versant. Ce résultat empirique a été utilisé pour déterminer les bassins hydrologiquement semblables. Ces conclusions sont appuyées par des évidences physiques telles que la proximité géographique des sites et la similarité des caractéristiques climatiques. Notons que cette méthode utilise la technique de résolution développée par Sparth (1982) appelée « Clusterwise Linear Régression ».

Une fois la région homogène déterminée, le transfert d'information des sites jaugés vers un site non jaugé à l'intérieur d'une même région peut alors être effectué. Plusieurs approches d'estimation régionale peuvent alors être considérées. Les méthodes du type stations-années

(Stedinger et Tasker, 1985) consistent à regrouper les données standardisées des différents sites de la région pour obtenir un échantillon représentatif de la distribution régionale. Ces méthodes sont très sensibles à la corrélation spatiale entre les données des différents sites d'une région. Les méthodes d'estimation du type indice de crue (index-flood) et par régression directe, identifiées comme les plus performantes par GREHYS (1996b) sont représentées brièvement dans les deux sections suivantes.

### 2.2.2 Méthode de l'indice de crue

La méthode de l'indice de crue pour l'estimation régionale a initialement été proposée par Dalrymple (1960). L'hypothèse de base de la méthode est que les données aux différents sites d'une région sont indépendantes et suivent la même distribution statistique à un facteur d'échelle près. Certaines caractéristiques statistiques (tels que le coefficient de variation ou le coefficient d'asymétrie) sont alors constantes à l'intérieur d'une même région homogène.

Dans ce qui suit, nous donnons le détail de la méthode telle que proposée initialement par Dalrymple (1960). Cette procédure peut-être divisée en deux grandes étapes :

#### Étape1 : Développement d'une courbe de fréquence régionale adimensionnelle.

- i) Pour un groupe de N stations voisines, on calcule  $X_T$  à chaque station (analyse de fréquence avec la loi de Gumbel)
- ii) On reporte graphiquement pour toutes les stations de la région le rapport,

$$x_T = \frac{X_T}{\bar{Q}_X} \quad (\text{ou } \bar{Q}_X \text{ est le débit maximum annuel moyen à la station considérée})$$

qui représente le débit relatif de période de retour T.

On considère que la région est homogène si les pentes des droites ne sont pas significativement différentes.

- iii) La courbe de fréquence régionale adimensionnelle est obtenue en prenant pour chaque valeur de T, la médiane des estimations  $x_T$ , obtenues aux stations de la région homogène.

La distribution « moyenne » standardisée de la région,  $F(x)$  est alors déterminée en regroupant les données standardisées des différents sites:

$$X_T = F_X^{-1}\left(1 - \frac{1}{T}; \alpha\right) \quad (2.32)$$

où T est la période de retour, et  $\alpha$  le vecteur des paramètres du modèle.

Etape2 : Etablissement d'une relation entre le débit maximum annuel moyen  $\bar{Q}_x$  et les caractéristiques physiographiques du bassin versant.

Généralement on considère la surface de drainage (A) du bassin versant :

$$\bar{Q}_x = \beta_0 + \beta_1 A \quad (2.33)$$

La courbe de fréquence adimensionnelle combinée au modèle indice de crue produit une courbe de fréquence applicable à n'importe quel site :

$$X_T = x_T(\beta_0 + \beta_1 A) \quad (2.34)$$

Différentes variantes de la méthode peuvent être présentées. Si un nombre limité d'années de données est disponible dans le site cible, la moyenne  $\bar{Q}_x$ , peut-être estimée directement à partir des données, ou bien comme une moyenne pondérée de l'estimation par régression sur les caractéristiques des bassins et de l'estimation à partir des données. La méthode de l'indice de crue a été délaissée pendant longtemps en faveur de méthodes régressives. Ceci était principalement dû aux limites de l'algorithme présenté par Dalrymple (1960), qui

propose la détermination des régions homogènes sur une base géographique et l'utilisation systématique de la distribution de Gumbel comme distribution régionale. Cependant, le principe général de cette approche fait l'objet d'un regain d'intérêt et de nouveaux développements durant les dernières années. Elle est maintenant considérée comme l'une des approches d'estimation régionale les plus efficaces (Wallis, 1980 ; Cunnane, 1989 ; Potter et Lettenmaier, 1990 ; Rasmussen, 1994 ; GREHYS, 1996b).

### **2.2.3 Méthodes régressives**

Benson (1962a) a proposé une alternative intéressante à la méthode de l'indice de crue, où plutôt que d'estimer la courbe de fréquence régionale et la relation entre  $X_T$  et une caractéristique physique quelconque, on mettrait directement en relation les variables hydrologiques,  $X_T$  et des variables explicatives telles que les caractéristiques physiographiques (surface de drainage, pente du cours d'eau principal, surface contrôlée par les lacs et les marées ...etc.) et climatiques (précipitation moyenne annuelle ...etc.) des bassins versants, cette méthode a l'avantage d'être simple, rapide et permet d'utiliser des distributions différentes pour représenter les débits extrêmes dans différents sites de la même région. De plus, elle n'est pas sensible à l'hétérogénéité qui peut exister dans la région considérée. De la même manière, on peut diviser cette procédure régressive en deux grandes étapes :

Étape1 : Pour un groupe de N stations voisines, on calcule  $X_T$  à chaque station par l'analyse de fréquence.

L'homogénéité régionale des stations est évaluée selon leur proximité géographique selon la méthode classique mais des améliorations ont été apporté depuis (GREHYS, 1994).

Étape2 : Etablissement d'une relation entre  $X_T$  et m caractéristiques physiographiques et climatiques  $Y_i$  du bassin versant.

Elle prend la forme d'une fonction puissance :

$$X_T = \beta_0 Y_1^{\beta_1} Y_2^{\beta_2} \dots Y_m^{\beta_m} \varepsilon \quad (2.35)$$

$Y_i$  : représente les caractéristiques climatiques et physiographiques

$\beta_i$  : étant les paramètres obtenus à partir de régression multiple

$\varepsilon$  : terme d'erreur

On peut effectuer une transformation logarithmique des variables afin de linéariser la relation :

$$\text{Log}X_T = \text{Log}\beta_0 + \beta_1 \text{Log}Y_1 + \beta_2 \text{Log}Y_2 + \dots + \beta_m \text{Log}Y_m + \text{Log}\varepsilon \quad (2.36)$$

posons  $e = \text{Log}\varepsilon = [\varepsilon_i]$  ; c'est la matrice des logarithmes des erreurs

$$\text{Log}X_T = \text{Log}\beta_0 + \beta_1 \text{Log}Y_1 + \beta_2 \text{Log}Y_2 + \dots + \beta_m \text{Log}Y_m + e \quad (2.36a)$$

Les paramètres  $\beta_m$  peuvent alors être estimés par la méthode classique des moindres carrés, ou par la méthode des moindres carrés généralisées (Stedinger et Tasker, 1985) pour remédier à la corrélation entre les valeurs des quantiles  $X_T$ . Il faut noter que la transformation logarithmique introduit généralement un biais car :

$$E[X_T] = E[\exp(\log X_T)] \neq \exp[E(\log X_T)] \quad (2.37)$$

Aussi, la linéarisation par transformation logarithmique ne peut pas être utilisée si l'équation de régression prend la forme d'une fonction puissance avec un terme d'erreur additif. Il est alors nécessaire d'utiliser un algorithme d'optimisation non linéaire pour estimer les paramètres du modèle (Ouarda et al, 1997a). Les deux procédures de

régionalisation précédentes sont largement utilisées dans les études de crue. Dans la pratique des études de l'estimation régionale des débits d'étiage, la méthode régressive a été usuellement appliquée aux débits relatifs. A l'annexe C, on retrouve les résultats de deux études d'estimation régionale des débits minimums, la première utilisant les modèles régressifs (Vogel et Kroll, 1990) et la seconde utilise la méthode de l'indice d'étiage qui est une transformation de la méthode de l'indice de crue (Pilon, 1990).

## **2.3 Méthodes alternatives**

Bien que les approches décrites précédemment aient été largement utilisées avec succès à travers le monde ces dernières décennies, certaines approches pour la régionalisation des débits exceptionnels sont tout à fait étrangères aux méthodes proposées par Dalrymple (1960) et Benson (1962a). C'est notamment le cas des méthodes basées sur l'utilisation des modèles déterministes, d'hydrogrammes géomorphologiques unitaires, où les modèles morphométriques. Dans ce qui suit, nous présentons une description succincte des principales contributions alternatives aux méthodes classiques d'estimation des événements extrêmes.

### **2.3.1 Modèles déterministes**

Une des méthodes d'estimation de débits de période de retour  $T$  consiste à utiliser un modèle déterministe du type pluie-débit croûton sec constitué d'un ensemble d'équations qui traduisent mathématiquement les processus impliqués au niveau du sol soumis à des contraintes météorologiques données. L'objectif de ce type de modèle est de reproduire les composantes de l'écoulement de l'eau lorsque celles-ci sont soumises à des conditions hydrométéorologiques données. Les modèles journaliers lorsqu'ils sont bien calibrés (adéquation entre le débit simulé par le modèle et le débit mesuré au site), permettent notamment de reproduire fidèlement l'hydrogramme annuel d'une rivière. Si l'on connaît le

débit journalier  $Q_j$  mesuré à l'exutoire d'un bassin versant ainsi que la séquence de données hydrométéorologiques journalières de ce bassin pour ces mêmes années, on peut alors ajuster les paramètres de la fonction de transfert du modèle déterministe de façon à ce que la réponse hydrologique modélisée reproduise ce qui a été mesuré. On sait en général que la nature même des phénomènes météorologiques permet de régionaliser plus adéquatement les variables météorologiques de bassins versants non jaugés que les débits. Il est alors possible d'estimer, pour un bassin versant donné, de longues séquences de variables météorologiques aux sites où l'on ne dispose que de très peu d'information hydrométrique et de simuler à l'aide du modèle déterministe pluie-débit les hydrogrammes annuels correspondant aux années de météorologie disponibles.

L'utilisation de séries météorologiques extrêmes (réelles ou simulées) permet d'estimer, à l'aide de modèles déterministes, l'ampleur d'événements hydrologiques beaucoup plus rares. Il reste cependant impossible d'associer une probabilité au dépassement à de tels événements, ce qui constitue l'une des principales faiblesses de cette approche.

### **2.3.2 Hydrogramme géomorphologique unitaire (Rodriguez-Iturbe et Valdés, 1979)**

Les caractéristiques géomorphologiques, et topologiques (géométrie en plan des réseaux de drainage) sont reconnues comme des facteurs conditionnant le régime hydrologique des bassins versants. Plusieurs auteurs ont proposé une méthode d'estimation de l'hydrogramme unitaire basée sur la géométrie des réseaux de drainage afin de caractériser la réponse hydrologique d'un bassin versant.

Une méthodologie d'estimation de l'hydrogramme géomorphologique unitaire de station non jaugés, consiste à générer à partir des variables hydrométéorologiques mesurées, ainsi que des caractéristiques des réseaux de drainage, des hydrogrammes géomorphologiques unitaires qui caractérisent la réponse hydrologique des bassins versants.

Dans ce cas aussi, l'inconvénient majeur de la méthode réside dans le fait qu'on peut difficilement estimer la magnitude d'événements  $X_T$  de grande période de retour.

### 2.3.3 Modèles morphométriques

Une manière intéressante de procéder à une régionalisation consiste à mettre à contribution les caractéristiques de forme de sections transversales de cours d'eau où l'on souhaite estimer  $X_T$  sans toutefois disposer d'information hydrologique à ce site. Cette méthode de type régressif, comme celle de Benson (1962a) consiste à établir des relations entre les quantiles d'étiage ou de crue et la morphométrie des cours d'eau plutôt qu'avec les caractéristiques physiques du bassin versant. La contribution des variables morphométriques à l'estimation de  $X_T$  à des sites non jaugés, bien qu'intéressante, s'avère très contraignante, puisqu'elle n'est applicable qu'à de petits bassins versants dont les cours d'eau, non régularisés, sont encaissés dans des matériaux meubles, et ainsi susceptibles d'être modelés par des événements hydrologiques importants.

Malgré les aspects intéressants de ces méthodes alternatives, différentes des procédures classiques de régionalisation, on retient de cette revue de littérature que les approches proposées par Dalrymple (1960) et de Benson (1962a) conduisent à des résultats satisfaisants, mais que l'on doit s'attarder davantage au choix de la distribution de probabilité à retenir et à celui de la variable d'étiage adéquate, surtout dans le cas des écoulements à débit minimum, et aussi insister sur la notion d'homogénéité régionale.

### **3. DESCRIPTION DE LA MÉTHODOLOGIE UTILISÉE, COMBINAISON ENTRE INVARIANCE D'ÉCHELLE SIMPLE ET MODÈLES DE QUEUE (I.E.S/M.Q)**

---

On présente dans ce chapitre une méthode originale d'estimation régionale des caractéristiques d'étiage, composée :

- D'une première partie traitant de la théorie d'invariance d'échelle, qui peut-être accomplie soit par analyse des moments statistiques de la distribution, soit par analyse des quantiles.

- D'une deuxième partie concernant le choix et l'ajustement d'une distribution statistique, approche par modélisation des queues de distribution, qui consiste à sélectionner un seuil  $u_i$ . Celui-ci définit la limite supérieure de l'ensemble des données qui seront employés pour l'estimation des paramètres de la distribution, à la station  $i$ .

#### **3.1 Motivations à l'origine de la procédure proposée**

Bien qu'il existe une littérature abondante à propos des modèles statistiques ou probabilistes d'estimation régionale sur les divers phénomènes naturels et, plus spécifiquement, sur les crues, très peu d'études portent sur le phénomène d'étiage. Le *U.S.G.S.*, met à jour et publie, périodiquement, un ensemble d'équations de régression régionale pour la prédiction des quantiles aux sites non jaugés. Ces équations de prédiction sont généralement développées sur une base provinciale (c'est à dire, état par état selon la région d'étude) pour la régionalisation des crues, et sont par la suite appliquées pour l'estimation régionale des débits d'étiage. Cependant, plusieurs études ont montré que la

prédiction des débits d'étiage est plus difficile que celle des débits de crue, à cause principalement, de la complexité générale attribuée au phénomène d'étiage. Cette difficulté est caractérisée par des erreurs standards relatives des quantiles d'étiage qui sont largement supérieures à celles des crues. Le concept d'invariance d'échelle qui a émergé récemment, a été introduit par plusieurs auteurs dans les champs hydrologique et climatologique, et a fait l'objet de plusieurs développements théoriques et d'applications afin de caractériser les propriétés des distributions des crues et des précipitations (Cadavid 1988 ; Gupta et Waymire 1990 ; Smith 1992 ; Gupta et al. 1994 ; Dawdy et Gupta 1995 ; Nguyen et Pandey 1996). Notons toutefois, que plusieurs applications de ces théories ont été effectuées en dehors du champ de l'hydrologie. Malheureusement, très peu d'intérêt a été accordé au phénomène d'étiage, et les seuls auteurs qui se sont penchés sur la question sont à notre connaissance Furey et Gupta (1995) et Miller (1996).

Le modèle de la théorie de l'invariance d'échelle est basé sur la paramétrisation simple impliquant la surface de drainage et peut représenter aussi bien une invariance d'échelle simple qu'une invariance d'échelle multiple. Dans ce qui suit, nous présentons une brève revue de littérature sur les études qui ont été faites jusqu'à présent dans ce domaine. Gupta et al. (1994) ont appliqué la méthode de la similitude d'échelle à la région du centre des Appalaches aux Etats-Unis ; cette étude permet de mettre en évidence les différences fondamentales entre la similitude d'échelle simple et la similitude d'échelle multiple. Les résultats obtenus sont plus efficaces que ceux obtenus par les méthodes classiques. Nguyen et Pandey (1996) ont trouvé que le regroupement des bassins similaires par la méthode de l'invariance d'échelle simple de la province de Québec a formé des régions géographiques bien définies ayant des caractéristiques climatiques distinctes. De plus, la meilleure estimation des quantiles de crue par la méthode d'invariance d'échelle a démontré la supériorité de cette technique par rapport à la méthode classique GEV/indice de crue (Hosking et Wallis, 1985). Ribeiro et Rousselle (1996) ont cependant souligné un inconvénient à l'approche de la similitude d'échelle simple. Celle-ci suppose en effet, que les coefficients de variation des distributions de débits de crue sont identiques à l'intérieur d'une région. Bien que des données empiriques montrent des propriétés statistiques qui

correspondent à des hypothèses simples de similitude d'échelle, les coefficients de variations des débits de crue dépendent de la zone de drainage. Ribeiro et Rousselle ont donc suggéré l'utilisation d'une méthode de Bayes empirique pour tenir compte de cette variation régionale du coefficient de variation. Enfin, ils ont utilisé une série de 109 débits de crue de la province de l'Ontario afin d'illustrer l'avantage d'utiliser une analyse par Bayes empirique subséquente à une modélisation par similitude d'échelle simple.

En pratique, les distributions empiriques des débits d'étiage sont souvent complexes, les populations apparaissent composées de plus d'une distribution, par conséquent, les lois de probabilités généralement utilisées en hydrologie ne suffisent pas à bien les modéliser. Des études récentes ont démontré que l'ajustement de différents modèles (Weibull, Lognormal, ...) aux queues des distributions des débits minimums peut générer des améliorations dans l'estimation des quantiles. Durrans (1996) a développé une procédure d'estimation des débits d'étiage avec les modèles de queue. Cette méthode, consiste à ajuster, pour un échantillon données, une distribution de probabilité aux données qui sont inférieures ou égales à un seuil  $u$  préalablement défini. Le choix de ce seuil est fait de façon à modéliser la queue inférieure de la distribution, ce qui revient à ignorer la partie centrale ainsi que la partie supérieure, représentant le reste des données. Une méthode originale pour l'estimation du seuil  $u$  a été utilisée dans cette étude. D'un point de vue analytique, cette procédure est moins subjective et offre des résultats satisfaisants. Durrans (1996) propose une estimation du seuil  $u$  en assignant une distribution empirique aux données originales. Dans le cadre du présent travail, nous proposons d'utiliser cette approche de modélisation des queues pour l'estimation des quantiles de débits d'étiage de la province de Québec, en la combinant avec le concept de similitude d'échelle. Une description détaillée de ces méthodes est effectuée dans les sections suivantes ; une application à un échantillon de 187 sites hydrométriques distribués à travers la province de Québec est également présentée.

## 3.2 Théorie de l'invariance d'échelle

Dans cette section, on présente la théorie d'invariance d'échelle pour l'estimation régionale des débits d'étiage, notons qu'à l'origine cette procédure a été développée pour l'estimation des débits de crue et qu'on se propose de l'appliquer pour les étiages. La méthode de la similitude d'échelle permet d'expliquer la variabilité spatiale des processus d'étiage en les indiquant par un paramètre d'échelle, qui est la taille des bassins de drainage. Sur cette base, une analyse régionale peut-être effectuée pour construire une courbe régionale unique (growth curve).

On peut considérer deux types d'invariance d'échelle :

- 1- invariance d'échelle simple (simple scaling)
- 2- invariance d'échelle multiple (multiscaling)

### 3.2.1 Invariance d'échelle simple

Les débits d'étiage suivent une invariance d'échelle simple s'ils satisfont deux conditions :

Condition d'homogénéité : Ce qui implique que tous les débits d'étiage dans une région proviennent de la même distribution statistique.

Condition d'échelle : Cette condition stipule que les débits d'étiage d'une station peuvent être transformés, moyennant une fonction d'échelle  $g(\lambda)$ , basée seulement sur la surface de drainage, pour obtenir la même distribution statistique que celle d'une autre station de la région.

Si ces deux hypothèses sont vérifiées, alors on sera en mesure de faire des prédictions de débits d'étiage à des sites non jaugés.

Soit  $Q(A)$  la variable aléatoire représentant les débits minimums d'une même région indexées par la surface de drainage  $A$ , et soit  $\lambda > 0$  un scalaire arbitraire.  $Q(A)$  est défini comme suivant une invariance d'échelle simple si pour une fonction d'échelle positive  $g(.)$  l'équation suivante est vérifiée :

$$Q(\lambda A) \stackrel{d}{=} g(\lambda)Q(A) \quad (3.1)$$

$d$  : représente l'égalité des distributions entre deux variables aléatoires.

$\lambda A$  : est l'ensemble déduit de  $A$  en multipliant ce dernier par le facteur  $\lambda > 0$

Pour évaluer la fonction d'échelle  $g(\cdot)$ , on prend  $|A| = 1$  représentant le bassin de référence unitaire, et considérons deux scalaires  $\lambda$  et  $\mu$  positifs. Appliquons l'équation (3.1) itérativement :

$$Q(\lambda\mu) \stackrel{d}{=} g(\lambda)Q(\mu) \stackrel{d}{=} g(\lambda)g(\mu)Q(1) \stackrel{d}{=} g(\lambda\mu)Q(1) \quad (3.2)$$

À partir de l'équation (3.1) on a

$$Q(\lambda\mu) \stackrel{d}{=} g(\lambda)Q(\mu) \quad (3.2a)$$

$$Q(\mu) \stackrel{d}{=} Q(\mu.1) \stackrel{d}{=} g(\mu).Q(1) \quad (3.2b)$$

À partir des équations (3.2a) et (3.2b) on peut déduire :

$$Q(\lambda\mu) \stackrel{d}{=} g(\lambda)g(\mu)Q(1) \quad (3.2c)$$

$$Q(\lambda\mu) \stackrel{d}{=} Q(\lambda\mu.1) \stackrel{d}{=} g(\lambda\mu).Q(1) \quad (3.2d)$$

On en déduit l'équation fonctionnelle :

$$g(\lambda\mu) = g(\lambda)g(\mu) \quad (3.3)$$

Cette fonction est appelée "fonction homogène", sa solution peut-être obtenue de la manière suivante :

Définissons,  $z(\log \lambda) = \log g(\lambda)$ , et prenons le logarithme des deux membres de l'équation (3.3). Nous avons alors:

$$\begin{aligned}\log[g(\lambda\mu)] &= \log[g(\lambda) \cdot g(\mu)] \\ \log[g(\lambda\mu)] &= \log[g(\lambda)] + \log[g(\mu)] \\ z[\log(\lambda\mu)] &= z[\log(\lambda)] + z[\log(\mu)] \\ z[\log(\lambda) + \log(\mu)] &= z[\log(\lambda)] + z[\log(\mu)]\end{aligned}\tag{3.4}$$

Une solution de cette équation fonctionnelle est,  $z(\log \lambda) = \theta \log \lambda$ , pour  $\theta$  constant, sachant que  $\log g(\lambda)$  est linéaire comme la fonction de  $\log \lambda$ . La solution de l'équation (3.3) est telle que ;

$$\log g(\lambda) = \theta \log(\lambda) = \log(\lambda^\theta)$$

et est donnée par *la loi de puissance* :

$$g(\lambda) = \lambda^\theta\tag{3.5}$$

$\theta$  : est un paramètre fondamental appelé exposant d'échelle. Il peut-être soit positif soit négatif.

Prenons  $\lambda = A$ , où A représente la surface de drainage. Il résulte des équations (3.1), (3.5) et d'après la définition de quantile ce qui s'en suit :

De l'équation (3.1), pour  $|A| = 1$ , nous avons :

$$Q(\lambda) = g(\lambda) \cdot Q(1)$$

De l'équation (3.5) pour  $\lambda = A$  :

$$g(\lambda) = A^\theta$$

Alors :

$$Q(A) = A^\theta Q(1)$$

$$q_p(A) = q_p(1) A^\theta \tag{3.6}$$

L'équation (3.6) montre l'étroite relation qui existe entre l'hypothèse de l'indice d'étiage et le concept d'invariance d'échelle simple. En effet, si on considère que la moyenne des débits d'étiage ( $\mu$ ) est une fonction déterministe de la surface de drainage A telle que  $\mu = A^\theta$ , et que le processus  $q_p(1)$  corresponde à la courbe de fréquence d'étiage adimensionnelle, alors l'hypothèse d'indice d'étiage est équivalente au concept d'invariance d'échelle simple.

Notons que si deux variables aléatoires X et Y suivent la même distribution statistique alors leurs moments sont les mêmes. D'après (3.1) et (3.5) on a en terme de moments statistiques ce qui suit :

$$Q(A) = A^\theta Q(1)$$

$$Q^K(A) = A^{K\theta} [Q(1)]$$

$$E[Q^K(A)] = A^{K\theta} E[Q^K(1)] \quad (3.7)$$

$$\text{Log}E[Q^K(A)] = K\theta \text{Log}A + C_K \quad (3.8)$$

où

$$C_K = \log E[Q^K(1)]$$

L'équation (3.8) peut se réécrire :

$$\beta(K) = \alpha(K) \log A + C_K \quad (3.8a)$$

où

$$\alpha(K) = K\theta ; \text{ et}$$

$$\beta(K) = \text{Log}E[Q^K(A)], \quad \text{quand } A = 1$$

les équations (3.7) et (3.8) montrent deux très importantes caractéristiques qui sont :

- i) log-log linéarité entre les moments statistiques et le paramètre d'échelle qui est ici la surface de drainage.
- ii) croissance linéaire de la pente  $\alpha(K) = \theta K$ , la pente est proportionnelle à l'ordre des moments statistiques considérés.

Ces deux caractéristiques jouent un rôle important dans l'analyse des données. Deux hypothèses peuvent être dérivées à partir de ces deux caractéristiques :

- Si les données exhibent des caractéristiques qui sont compatibles avec la croissance linéaire de la pente  $\alpha(K) = \theta K$ , et si l'équation (3.8a) est vérifiée, alors la distribution suit une invariance d'échelle simple.

- Par contre si la log-log relation de l'équation (3.8a) est vérifiée, mais que la pente ne soit pas proportionnelle à l'ordre des moments, on dit alors que la distribution suit le processus d'invariance d'échelle multiple (Multiscaling).

Les objectifs d'une théorie physique de la méthode d'invariance d'échelle simple est de ;

- a) donner une interprétation physique de l'exposant  $\theta$ , et
- b) évaluer la structure statistique de  $Q(A)$  par des considérations physiques impliquant divers mécanismes d'écoulement.

### 3.2.2 Modèle cascade Lognormal

Gupta et Waymire (1990), développèrent le modèle cascade Lognormal dans le but d'étudier les propriétés d'échelle des processus hydrologiques. Le développement de ce modèle suppose que les logarithmes des étiages sont distribués normalement, et selon l'équation (3.8a), cela suppose aussi que les moments de distribution sont fonction du logarithme de la surface de drainage

$$\ln(Q_i) \approx N(\mu_i, \sigma_i)$$

$$\ln(Q_i) \approx N(m_1 + m_2 \ln(A_i), r_1 - r_2 \ln(A_i)) \quad (3.9)$$

avec :  $\mu_i = m_1 + m_2 \ln(A_i)$  et  $\sigma_i = r_1 - r_2 \ln(A_i)$

$m_1$  et  $m_2$  sont des réels

$r_1$  et  $r_2$  sont des réels positifs

$$A_i < \exp\left(\frac{r_1}{r_2}\right)$$

Dans ce qui suit nous présentons un aperçu de la méthode du modèle cascade Lognormal, pour plus de détails se référer à l'article de Gupta et Waymire (1990), où Miller (1996).

La condition sur  $r_2$  est importante, car elle distingue entre le modèle d'invariance d'échelle simple ( $r_2 = 0$ ) et multiple ( $r_2 > 0$ ). Les moments du modèle cascade peuvent être déterminés de la manière suivante :

$$\ln\left(E[Q_i]^K\right) = \left[m_1 K + \frac{1}{2} r_1 K^2\right] + \left[m_2 K - \frac{1}{2} r_2 K^2\right] \ln(A_i) \quad (3.10)$$

Si  $r_2 = 0$  le modèle est donc d'invariance d'échelle simple car la pente  $m_2 K$  est proportionnelle à  $K$  avec :

$$g(s) = s^{m_2} \quad (3.11)$$

Si  $r_2 > 0$  le modèle est d'invariance d'échelle multiple avec une pente

$$b(K) = m_2 K - \frac{1}{2} r_2 K^2 \quad (3.12)$$

Les paramètres du modèle cascade Lognormale peuvent être calculés à partir des débits d'étiage annuels dans des sites jaugés.

$\phi = (m_1, m_2, r_1, r_2)$  représente le vecteur des paramètres du modèle

$m$  : le nombre des surfaces de drainage

$n_i$  : le nombre de débits d'étiage à la station  $i$  avec une surface de drainage  $A_i$

$Q_{ij}$  : représente le  $j^{\text{eme}}$  débit d'étiage à la station  $i$ .

L'estimation des paramètres  $(m_1, m_2, r_1, r_2)$  du modèle peut être faite à l'aide de la méthode du maximum de vraisemblance.

$$f(Q_i) = \frac{1}{Q_i \sigma_i \sqrt{2\pi}} \exp \left[ -\frac{(\ln Q_i - \mu_i)^2}{2\sigma_i^2} \right] \quad (3.13)$$

avec

$$\mu_i = m_1 + m_2 \ln(A_i) \quad \text{et} \quad \sigma_i = r_1 - r_2 \ln(A_i)$$

$$f(Q_i) = \frac{1}{\ln Q_i \sigma_i \sqrt{2\pi}} \exp \left[ -\frac{(\ln Q_i - m_1 - m_2 \ln(A_i))^2}{2\sigma_i^2} \right] \quad (3.13a)$$

La fonction de vraisemblance est obtenue de la manière suivante :

$$l(\phi) = \prod_{i=1}^m \prod_{j=1}^{n_i} f(Q_{ij}) \quad (3.13b)$$

$$l(\phi) = \prod_{i=1}^m \prod_{j=1}^{n_i} \left( \frac{1}{\ln Q_{ij} \sigma_i \sqrt{2\pi}} \exp \left[ -\frac{(\ln Q_{ij} - m_1 - m_2 \ln(A_i))^2}{2\sigma_i^2} \right] \right) \quad (3.13c)$$

En prenant le logarithme de la fonction vraisemblance on obtient l'équation suivante :

$$L(\phi) = -\sum_{i=1}^m \sum_{j=1}^{n_i} \ln(Q_{ij}) - \frac{1}{2} \sum_{i=1}^m n_i \ln(2\pi) - \frac{1}{2} \sum_{i=1}^m n_i \ln \sigma_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^{n_i} \frac{[\ln(Q_{ij}) - (m_1 + m_2 \ln(A_i))]^2}{\sigma_i} \quad (3.14)$$

Pour plus de détails se référer à Gupta et Waymire (1990) où Miller (1996).

Les estimateurs du maximum de vraisemblance des paramètres sont donnés par le système d'équations suivant, qui doit être résolu simultanément afin de maximiser la fonction log-vraisemblance (Gupta et Waymire, 1990):

$$\sum_{i=1}^m \sigma_i^{-1} \sum_{j=1}^{n_i} \ln(Q_{ij}) = m_1 \left[ \sum_{i=1}^m \frac{n_i}{\sigma_i} \right] + m_2 \left[ \sum_{i=1}^m \frac{n_i \ln(A_i)}{\sigma_i} \right] \quad (3.15)$$

$$\sum_{i=1}^m \frac{\ln(A_i)}{\sigma_i} \sum_{j=1}^{n_i} \ln(Q_{ij}) = m_1 \left[ \sum_{i=1}^m \frac{n_i \ln(A_i)}{\sigma_i} \right] + m_2 \left[ \sum_{i=1}^m \frac{n_i (\ln(A_i))^2}{\sigma_i} \right] \quad (3.16)$$

$$\sum_{i=1}^m \frac{n_i}{\sigma_i} = \sum_{i=1}^m \frac{\sum_{j=1}^{n_i} (\ln(Q_{ij}) - m_1 - m_2 \ln(A_i))^2}{\sigma_i^2} \quad (3.17)$$

$$\sum_{i=1}^m \frac{n_i \ln(A_i)}{\sigma_i} = \sum_{i=1}^m n_i \ln(A_i) \left( \frac{\sum_{j=1}^{n_i} (\ln(Q_{ij}) - m_1 - m_2 \ln(A_i))^2}{\sigma_i^2} \right) \quad (3.18)$$

Dans le cas d'une invariance d'échelle simple, c'est à dire  $r_2 = 0$ , les solutions du système d'équations peuvent être décrites par :

$$r_1 = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} [\ln(Q_{ij}) - m_1 - m_2 \ln(A_i)]^2}{\sum_{i=1}^m n_i} \quad (3.19)$$

$$m_1 = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} \ln(Q_{ij})}{\sum_{i=1}^m n_i} - m_2 \frac{\sum_{i=1}^m n_i \ln(A_i)}{\sum_{i=1}^m n_i} \quad (3.20)$$

$$m_2 = \frac{c_1}{c_2} \quad (3.21)$$

où

$$c_1 = \sum_{i=1}^m n_i \sum_{i=1}^m \sum_{j=1}^{n_i} \ln(Q_{ij}) \ln(A_i) - \sum_{i=1}^m n_i \ln(A_i) \sum_{i=1}^m \sum_{j=1}^{n_i} \ln(Q_{ij}) \quad (3.22)$$

$$c_2 = \sum_{i=1}^m n_i \sum_{i=1}^m n_i [\ln(A_i)]^2 - \left( \sum_{i=1}^m n_i \ln(A_i) \right)^2 \quad (3.23)$$

Notons que la procédure du maximum de vraisemblance décrite ci-dessus, suppose que les observations de débits d'étiage pour différents sites, sont indépendantes les unes des autres. En réalité les débits d'étiage sont fortement corrélés dans l'espace que ne le sont les débits de crue. Ceci est dû, principalement, au fait que la sécheresse est un phénomène régional, tandis que les crues sont plutôt des phénomènes locaux.

Comme on l'indique précédemment, la distinction entre une invariance d'échelle simple ( $r_2 = 0$ ) et une invariance d'échelle multiple ( $r_2 > 0$ ) réside dans le paramètre  $r_2$ . Pour déterminer si ce dernier est nulle ou non, on peut utiliser le test unilatéral suivant :

Hypothèse nulle :  $H_0 : r_2 = 0$

Hypothèse alternative :  $H_a : r_2 > 0$

Le test du rapport des vraisemblances (likelihood ratio test) peut-être utilisé pour tester l'hypothèse de nulle. La statistique du test  $X^2$  est donnée par :

$$X^2 = -2[\ln L(\phi_0) - \ln L(\phi)] \quad (3.24)$$

$\phi_0$  : représente le modèle réduit avec  $r_2 = 0$

$\phi$  : représente le modèle au complet

Pour les échantillons de grande taille, la statistique du test  $X^2$  est distribuée approximativement selon une loi chi-carré (Neter, et. Al., 1989).

Donc, la règle de décision pour le test de l'invariance d'échelle simple est :

Si  $X^2 \leq \chi^2(1 - \alpha, 1)$ , alors  $H_0$  est accepté

Si  $X^2 > \chi^2(1 - \alpha, 1)$ , alors  $H_0$  est rejeté

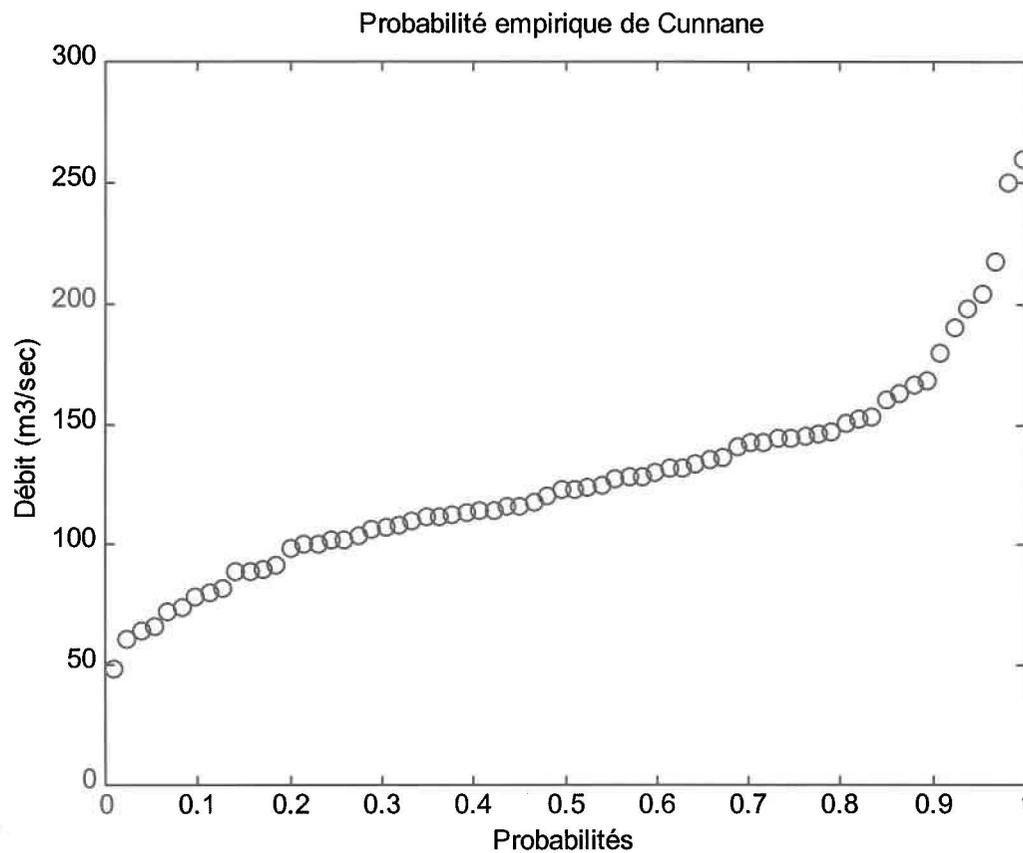
où  $\chi^2(1 - \alpha, 1)$  est la variable chi-carré de probabilité au non dépassement  $(1 - \alpha)$  et à un degré de liberté.

On présentera dans le chapitre suivant, une application de la méthodologie permettant de modéliser les débits minimums avec le modèle cascade Lognormal. Ce modèle peut représenter aussi bien une invariance d'échelle simple, qu'une invariance d'échelle multiple, et le facteur qui discrimine entre les deux concepts est le paramètre  $r_2$ .

Miller (1996), a appliqué le modèle cascade Lognormale couplé avec la théorie des modèles de queue aux données de débits d'étiage de 83 stations hydrométriques de l'état de l'Alabama. Cette première approche, a mené à des résultats indiquant que la combinaison des deux théories peut mener à des résultats intéressants.

### 3.3 Théorie des modèles de queue

Les distributions empiriques des débits d'étiage annuels sont souvent complexes. Cette considération a conduit les organismes chargés de la gestion et du contrôle des ressources hydriques à adopter l'approche d'estimation graphique dans la plupart des cas. Les méthodes graphiques ne sont cependant pas très satisfaisantes en raison de la subjectivité qui leur est associée. D'où la nécessité de développer des méthodes alternatives d'analyse.



**Figure 3.1** Tracé des débits annuels  $Q_7$  de la rivière Coulonge (station 02kg001)

La figure 3.1 représente une distribution de probabilité typique des débits d'étiage de la province de Québec. cette figure montre que les données de débits minimums peuvent

provenir de plus d'une distribution. Des observations de même nature ont été faites par Riggs (1972), Waylen et Woo (1987), Nathan et McMahon (1990), Tomic (1994) et Durrans et Tomic (1996). Cette figure suggère qu'il y a au moins deux distributions qui peuvent être combinées pour représenter les données de cette station, une qui correspond à la queue inférieure et l'autre à la queue supérieure de la distribution empirique. Dans les études de débits minimums la queue inférieure, est la région la plus critique pour les données de débits d'étiage. Les travaux de Durrans (1996) suggèrent que les modèles de queue peuvent venir à bout de certaines difficultés rencontrées dans l'analyse fréquentielle des débits minimums. L'avantage de cette approche réside en la facilité d'utilisation des modèles, car en ce faisant, on réduit le nombre de paramètres à estimer. L'auteur a considéré trois distributions pour la modélisation des débits d'étiage de l'état de l'Alabama : la Gamma, la Lognormale et la Weibull ; il a constaté que les deux premières distributions étaient difficiles et peu commodes à manipuler. L'utilisation de ces deux modèles ne conduit pas à « l'élimination » d'un paramètre du modèle, mais plutôt à l'établissement d'une relation (contrainte) entre les paramètres du modèle, c'est à dire l'expression d'un paramètre en fonction d'un autre. La distribution Weibull a été choisie en raison de sa commodité et de sa facilité d'utilisation. Dans le cadre de cette étude, et afin de comparer les résultats des différentes méthodes des modèles de queue, nous avons choisi d'en considérer deux :

***Modèle de queue conditionnel de la distribution Weibull***

***Modèle de queue conditionnel de la distribution Lognormale***

### 3.3.1 Modèle de queue conditionnel de la distribution Weibull (Durrans, 1996)

Considérons la variable aléatoire positive non nulle ( $Q$ ) représentant les débits minimums annuels de sept jours consécutifs, tirée de la distribution Weibull, dont la fonction densité de probabilité peut s'écrire sous la forme :

$$f_Q(q) = \alpha \lambda q^{\alpha-1} \exp(-\lambda q^\alpha) \quad (3.25)$$

où  $\alpha > 0$  est un paramètre de forme et  $\lambda > 0$  un paramètre d'échelle.

Cette distribution est ajustée en considérant dans l'échantillon de données les valeurs inférieures ou égales à un seuil  $u$ . Le choix de ce seuil est effectué de façon à modéliser uniquement la queue inférieure de la distribution des débits d'étiages et exclure de l'analyse les valeurs qui sont supérieures à ce seuil. En ce faisant, le modèle résultant donne une représentation inadéquate de la vraie population pour l'intervalle de données qui est plus grand que le seuil  $u$  ; cette approche justifie notre intérêt pour la queue inférieure de la distribution, et notre indifférence pour la région centrale et la queue supérieure de la distribution. La nouvelle variable aléatoire est définie comme :

$$X = \frac{Q}{u} \quad (3.26)$$

La fonction de densité de  $X$  sera :

$$f_X(x) = f_Q[q(x)] \left| \frac{dq}{dx} \right| = \alpha \xi x^{\alpha-1} \exp(-\xi x^\alpha) \quad (3.27)$$

$x > 0$  et  $\alpha > 0$

$\xi = \lambda u^\alpha$  ( $\xi > 0$ )

La fonction de distribution de  $x$  est :

$$F_X(x) = 1 - \exp(-\xi x^\alpha) \quad (3.28)$$

$$F_X(1) = 1 - \exp(-\xi) \quad (3.28a)$$

Posons que  $F_X(1) = p_u$ , alors le paramètre  $\xi$  est défini uniquement en terme de  $p_u$  comme :

$$\xi = -\ln(1 - p_u) \quad (3.29)$$

Si  $u$  est défini comme le 60<sup>ème</sup> percentile de la variable  $Q$  c'est à dire la valeur de la probabilité au non dépassement 0.6, on a  $u = q_{60}$ , alors

$$F_Q(u) = F_X(1) = p_u = 0.60 \quad (3.30)$$

donc à partir de (3.28a) on a :

$$F_X(1) = p_u = 0.60 = 1 - \exp(-\xi) \quad (3.31)$$

On en déduit

$$\xi = -\ln(1 - 0.60) \quad (3.31a)$$

donc

$$\xi = -\ln 0.40 = 0.91629 \quad (3.32)$$

La distribution conditionnelle de  $X$ , pour  $X \leq 1$ , notée  $G_X(x)$  peut-être obtenue par :

$$G_X(x) = \frac{F_X(x)}{F_X(1)} = \frac{F_X(x)}{0.60} = 1.667 F_X(x) \quad (3.33)$$

où l'intervalle de variation de  $x$  est tel que  $0 < x \leq 1$

L'expression générale (pour  $u$  quelconque) de la fonction de densité cumulée de la distribution conditionnelle est :

$$G_x(x) = \left\{ 1 - \exp\left[x^\alpha \ln(1 - p_u)\right] \right\} / p_u \quad (3.33a)$$

A partir des équations (3.28) et (3.32) l'équation (3.33) peut aussi s'écrire :

$$G_x(x) = 1.667 \left[ 1 - \exp(-0.91629x^\alpha) \right] \quad (3.34)$$

L'expression générale (pour  $u$  quelconque) de la fonction de densité de la distribution conditionnelle est :

$$g_x(x) = \frac{-\alpha \ln(1 - p_u) x^{\alpha-1} \exp\left[x^\alpha \ln(1 - p_u)\right]}{p_u} \quad (3.35)$$

Pour un percentile de 60% l'équation (3.35) devient :

$$g_x(x) = 1.52715 \alpha x^{\alpha-1} \exp(-0.91629x^\alpha) \quad (3.36)$$

Cette expression peut aussi être obtenue directement en dérivant l'équation (3.34).

Notons qu'en définissant le seuil  $u$ , et en effectuant la transformation pour la nouvelle variable  $X$ , ainsi que le développement de la distribution conditionnelle  $G_x(x)$ , nous avons par conséquent réduit l'intervalle d'étude à  $(0,1]$ , et éliminé le paramètre  $\lambda$  réduisant ainsi le modèle à un seul paramètre  $\alpha$ . Effectivement le paramètre  $\xi$  dépend uniquement du niveau de probabilité correspondant au seuil choisi (0.60 dans le présent cas), le paramètre  $\alpha$  sera donc estimé, sur la base des données observées qui sont inférieures ou égales à  $u$ .

Cependant, le seuil peut prendre d'autres valeurs selon la taille des échantillons de données de débits minimums. Dans le chapitre suivant consacré à l'application de la méthodologie aux données de la province de Québec, nous présentons une procédure de détermination du seuil  $u$  différente de celle proposée par Durrans (1996).

### **Estimation du paramètre $\alpha$**

Dans tout ce qui suit, notons que  $x_i$ ,  $i = 1, 2, 3, \dots, m$ , est employé pour signifier que les données (transformées) utilisées pour estimer  $\alpha$  sont celles inférieures ou égales au seuil  $u$ . En pratique, une estimation de  $u$  peut-être obtenue en assignant des probabilités empiriques aux données originales  $q_i$ , et en interpolant éventuellement. Les données  $x_i$  peuvent donc être calculées comme  $x_i = q_i/u$ , pour les  $m$  données originales qui sont inférieures ou égales au seuil  $u$ . La probabilité empirique correspond à une probabilité au non dépassement  $P_i$  et est obtenue à partir de la formule suggérée par Cunnane (1978) :

$$P_i = \frac{i - 0.4}{n - 0.2} \quad (3.37)$$

où  $P_i$  est la probabilité affectée à la valeur  $x_i$  dans l'échantillon classé par ordre croissant ( $x_1 < x_2 < \dots < x_m$ ),

$P_i$  ; probabilité au non dépassement associée à la  $i^{\text{ème}}$  plus petite donnée ( $i^{\text{ème}}$  ordre statistique),

$n$  ; taille de l'échantillon original de données.

Parmi les méthodes d'estimation des paramètres on distingue :

- la méthode du maximum de vraisemblance
- la méthode des moments
- la méthode des moments de probabilité pondérés

Notons que pour le modèle de queue de la distribution Weibull, la méthode des moments est identique à la méthode des moments pondérés, car on réduit l'estimation à un seul paramètre du modèle, dans ce cas les deux méthodes utilisent seulement l'estimation de la moyenne de l'échantillon.

Dans ce qui suit, nous présentons uniquement la méthode du maximum de vraisemblance pour estimer le paramètre  $\alpha$ , car c'est la méthode qui a été appliquée dans le chapitre suivant consacré à l'application de la méthodologie I.E.S/M.Q.

La fonction vraisemblance de la distribution conditionnelle peut s'écrire à partir de l'équation (3.35) de la manière suivante :

$$L = \prod_{i=1}^m (g_X(x_i)) \quad (3.37a)$$

$$L = \prod_{i=1}^m \left( \frac{-\alpha \ln(1-p_u) x_i^{\alpha-1} \exp[x_i^\alpha \ln(1-p_u)]}{p_u} \right) \quad (3.37b)$$

Le logarithme de la fonction vraisemblance de la distribution conditionnelle peut s'écrire de la manière suivante (Durrans, 1996) :

$$\ln L = m \ln \alpha - m \ln p_u + m \ln[-\ln(1-p_u)] + (\alpha-1) \sum_{i=1}^m \ln x_i + \ln(1-p_u) \sum_{i=1}^m x_i^\alpha \quad (3.38)$$

$m$  : nombre de données inférieures au seuil  $u$

L'estimation du paramètre  $\alpha$  satisfait  $\frac{\partial \ln L}{\partial \alpha} = 0$

$$\frac{m}{\alpha} + \sum_{i=1}^m \ln x_i + \ln(1-p_u) \sum_{i=1}^m x_i^\alpha \ln x_i = 0 \quad (3.39)$$

La solution de cette équation peut-être obtenue relativement facilement par la méthode des approximations successives.

Prenons, comme exemple,  $p_u = 0.60$ . On trouve :

$$\ln L = m \ln 1.52715 + m \ln \alpha + (\alpha - 1) \sum_{i=1}^m \ln x_i - 0.91629 \sum_{i=1}^m x_i^\alpha \quad (3.40)$$

La dérivée de l'expression précédente par rapport à  $\alpha$  donne :

$$\frac{m}{\alpha} + \sum_{i=1}^m \ln x_i - 0.91629 \sum_{i=1}^m x_i^\alpha \ln x_i = 0 \quad (3.41)$$

### Estimation de quantiles

L'estimation du quantile  $q_p$  de la distribution de Q pour n'importe quel percentile p qui est plus petit ou égal à  $p_u$ , peut-être évaluée avec des modèles de queue comme :

$$q_p = ux_p \quad (3.42)$$

ou  $x_p$  est le  $p^{\text{ième}}$  percentile de la distribution  $F_X(x)$  ou bien le  $1.667p^{\text{ième}}$  percentile de la distribution conditionnelle  $G_X(x)$ , pour  $u = q_{60}$ .

Un estimateur de  $x_p$  est obtenu en réarrangeant la distribution cumulative  $F_X(x)$  donnée par l'équation (3.28):

$$F_X(x) = 1 - \exp(-\xi x_p^\alpha) = p$$

$$\xi = -\ln(1 - p_u)$$

$$p = 1 - \exp(x_p^\alpha \ln(1 - p_u))$$

$$\ln(1 - p) = x_p^\alpha \ln(1 - p_u)$$

$$x_p^\alpha = \frac{\ln(1 - p)}{\ln(1 - p_u)}$$

$$x_p = \left[ \frac{\ln(1 - p)}{\ln(1 - p_u)} \right]^{\frac{1}{\alpha}} \quad (3.43)$$

Notons que l'équation  $q_p = ux_p$  rappelle l'expression de la méthode d'estimation régionale des débits maximums appelée, indice de crue (Darlymple, 1960),  $u$  représentant l'indice de crue (indice d'étiage dans le présent contexte). À partir de cette remarque, si l'hypothèse de l'invariance d'échelle simple est vérifiée, nous pouvons utiliser la procédure de l'indice d'étiage avec  $u$  représentant l'indice d'étiage. Smith (1989) lors d'une étude sur la régionalisation des crues dans la région centrale des Appalaches, a employé une approche de modélisation similaire à celle décrite précédemment, dans laquelle il a choisi de prendre le 90<sup>ème</sup> percentile comme indice de crue de la région.

### 3.3.2 Modèle de queue conditionnel de la distribution Lognormale (Durrans, 1996)

La densité de probabilité de la distribution Lognormale s'écrit :

$$f_Q(q) = \frac{1}{q\sigma_y\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln q - \mu_y}{\sigma_y}\right)^2\right] \quad (3.44)$$

où  $\mu_y$  et  $\sigma_y > 0$  sont respectivement la moyenne et l'écart type de la variable aléatoire

$$Y = \ln Q$$

Définissons une nouvelle variable :

$$X = Q / u; \quad \text{où } u \text{ est le seuil}$$

La fonction de distribution de probabilité (f.d.p) de  $X$  s'écrit :

$$f_X(x) = \frac{1}{(xu)\sigma_y\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln(xu) - \mu_y}{\sigma_y}\right)^2\right] (u) \quad (3.44a)$$

La fonction de distribution cumulative de  $X$  s'écrit :

$$F_X(x) = \frac{1}{\sigma_y\sqrt{2\pi}} \int_0^x \frac{1}{t} \exp\left[-\frac{1}{2}\left(\frac{\ln t - \mu_*}{\sigma_y}\right)^2\right] dt \quad (3.45)$$

$$\mu_* = \mu_y - \ln u$$

Notons que  $F_X(1) = p_u$ , Ceci implique qu'il y a une relation entre  $\mu_*$ ,  $\sigma_y$  et  $p_u$

Cette contrainte est de la forme (Durrans, 1996):

$$\sigma_y = \frac{\ln u - \mu_y}{\phi^{-1}(p_u)} = -\frac{\mu_*}{\phi^{-1}} \quad (3.46)$$

$\phi^{-1}(\cdot)$ , c'est l'inverse de la fonction de distribution normale standardisée

La f.d.p du modèle conditionnel de queue est :

$$g_X(x) = \frac{1}{xp_u \sigma_y \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{\ln x}{\sigma_y} + \phi^{-1}(p_u) \right)^2 \right] \quad (3.47)$$

Notons que seul le paramètre  $\sigma_y$  a besoin d'être estimé sur la base des données qui sont en dessous du seuil  $u$ .

### Calcul des probabilités :

$$F_X(x) = \frac{1}{\sigma_y \sqrt{2\pi}} \int_0^x \frac{1}{t} \exp \left[ -\frac{1}{2} \left( \frac{\ln t - \mu_*}{\sigma_y} \right)^2 \right] dt \quad (3.48)$$

posons :

$$z = \ln t; \quad dz = \frac{1}{t} dt \quad \Rightarrow \quad t dz = dt$$

$$F_X(x) = \frac{1}{\sigma_y \sqrt{2\pi}} \int_{-\infty}^{\ln(x)} \frac{1}{t} \exp \left[ -\frac{1}{2} \left( \frac{z - \mu_*}{\sigma_y} \right)^2 \right] t dz \quad (3.49)$$

posons :

$$w = \frac{z - \mu_*}{\sigma_y}; \quad dw = \frac{1}{\sigma_y} dz \quad \Rightarrow \quad \sigma_y dw = dz$$

$$F_X(x) = \frac{\sigma_y}{\sigma_y \sqrt{2\pi}} \int_{-\infty}^{\frac{\ln(x) - \mu_*}{\sigma_y}} \exp \left[ -\frac{w^2}{2} \right] dw \quad (3.50)$$

$$F_X(x) = \int_{-\infty}^{\frac{\ln(x) - \mu_*}{\sigma_y}} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{w^2}{2}\right] dw \quad (3.51)$$

On reconnaît la fonction de densité cumulée de la loi normale centrée réduite

$$F_X(x) = \Phi\left(\frac{\ln(x) - \mu_*}{\sigma_y}\right) - \Phi\left(\frac{-\mu_*}{\sigma_y}\right) \quad (3.52)$$

### Estimation du paramètre $\sigma_y$

La fonction vraisemblance du modèle de queue Lognormal s'écrit de la manière suivante :

$$L = \prod_{i=1}^m (g_X(x_i)) \quad (3.52a)$$

$$L = \prod_{i=1}^m \left( \frac{1}{x p_u \sigma_y \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left( \frac{\ln x}{\sigma_y} + \phi^{-1}(p_u) \right)^2 \right] \right) \quad (3.52b)$$

Le logarithme de la fonction vraisemblance du modèle de queue Lognormal s'écrit pour

$x_1 < x_2 < \dots < x_m$  (Durrans 1996):

$$\begin{aligned} \ln L = & -m \ln p_u - m \ln \sigma_y - \frac{m}{2} \ln(2\pi) - \frac{m}{2} [\phi^{-1}(p_u)]^2 \\ & - \sum_{i=1}^m \ln x_i - \frac{1}{2\sigma_y^2} \sum_{i=1}^m (\ln x_i)^2 - \frac{\phi^{-1}(p_u)}{\sigma_y} \sum_{i=1}^m \ln x_i \end{aligned} \quad (3.53)$$

$m$  : Nombre de données inférieures ou égales au seuil  $u$

En dérivant  $\ln L$  par rapport à  $\sigma_y$  on obtient :

$$\frac{\partial \ln L}{\partial \sigma_y} = -\frac{m}{\sigma_y} + \frac{1}{\sigma_y^3} \sum_{i=1}^m (\ln x_i)^2 + \frac{\phi^{-1}(p_u)}{\sigma_y^2} \sum_{i=1}^m (\ln x_i) = 0 \quad (3.53a)$$

$$-m\sigma_y^2 + \phi^{-1}(p_u) \sum_{i=1}^m (\ln x_i) \sigma_y + \sum_{i=1}^m (\ln x_i)^2 = 0 \quad (3.53b)$$

ce qui peut s'écrire :

$$A\sigma_y^2 + B\sigma_y + C = 0 \quad (3.54)$$

où

$$A = -m \quad (3.55)$$

$$B = \phi^{-1}(p_u) \sum_{i=1}^m \ln x_i \quad (3.56)$$

et

$$C = \sum_{i=1}^m (\ln x_i)^2 \quad (3.57)$$

$B^2$  étant toujours positif, et comme  $A$  et  $C$  sont de signe contraire, les racines de cette équation du second degré seront toujours réelles et de signe contraire (une racine positive et une négative). Comme  $\sigma_y$  est positif. On retient donc la racine positive.

**Estimation de quantiles**

$$q_p = ux_p \quad (3.58)$$

$$x_p = \exp[\mu_* + \sigma_y \phi^{-1}(p)] \quad (3.59)$$

$$x_p = \exp\{\sigma_y [\phi^{-1}(p) - \phi^{-1}(p_u)]\} \quad (3.60)$$

## **4. APPLICATION DE LA MÉTHODE (I.E.S/M.Q)**

---

Dans ce chapitre, on présente les résultats de l'application de la méthodologie décrite au chapitre précédent, concernant la méthode de l'invariance d'échelle simple (3.2.1) et des modèles de queue conditionnel des distributions Weibull (3.3.1) et Lognormale (3.3.2).

### **4.1 Description des données et périodes de fonctionnement**

Les données utilisées pour l'application de la méthode (I.E.S/M.Q) proviennent de la banque de données « HYDAT » d'environnement Canada. On a considéré initialement 287 stations hydrométriques réparties dans la province de Québec, pour lesquelles on dispose de mesures de débits journaliers et de la superficie du bassin versant correspondant. 187 stations de jaugeages ont été finalement retenues. Deux raisons ont motivé la sélection de ces stations :

- 1) La période de fonctionnement des stations et,
- 2) Le régime de l'écoulement

Nous avons choisi une période d'enregistrement supérieure ou égale à 10 ans et un régime d'écoulement naturel ou influencé seulement de manière journalière.

Se situant entre les latitudes 45°12' et 60°19' et les longitudes 57°6' et 79°1', les bassins versants de la région d'étude ont des superficies allant de 0.69 km<sup>2</sup> à 146 000 km<sup>2</sup>. D'un point de vue géographique, ces stations sont réparties à travers la province du Québec (Figure 4.1). Les stations ont des séries d'observations variant de 10 ans à 78 ans. Elles seront considérées hydrologiquement homogènes au sens proposé par Gupta et al (1994).

Les distributions de débits sont supposées dépendre de la surface des bassins versants (effet d'échelle). L'information sur la position géographique n'est pas prise en compte lors de l'étude des distributions des débits d'étiage. Cependant, les résultats obtenus seront analysés pour identifier toute influence de la localisation géographique. Ceci sera éventuellement utile si un regroupement géographique des différentes régions est considéré approprié.

Remarque :

Nous avons repéré les stations par un numéro d'ordre pour alléger les graphiques et les cartes. On retrouvera la correspondance entre le code provincial et fédéral à l'annexe A (Tableau A1).

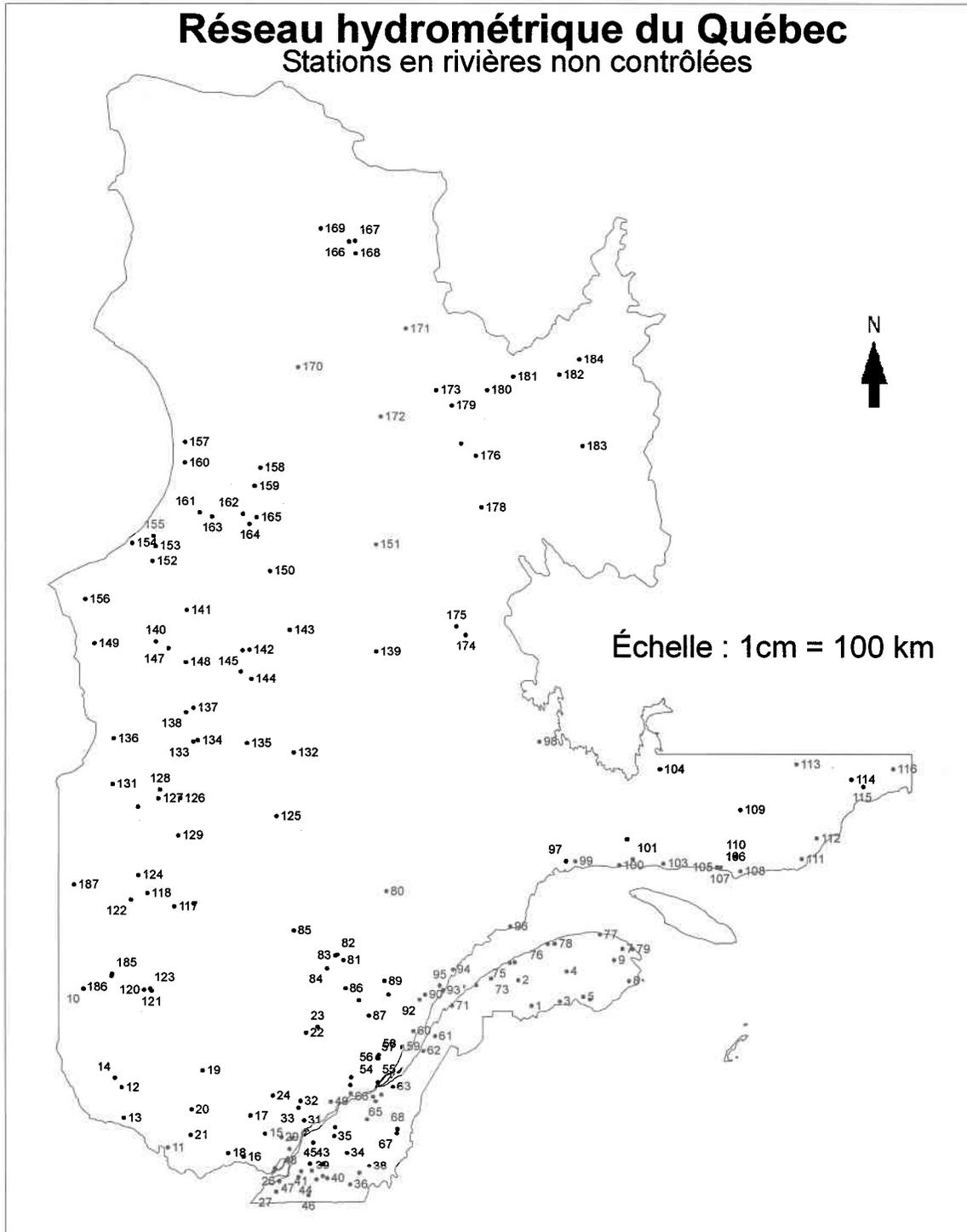


Figure 4.1 Répartition géographique des 187 stations hydrométriques.

## 4.2 Application de la similitude d'échelle simple

Dans ce qui suit, pour illustrer l'application de l'approche proposée, nous avons dans un premier temps calculé les six premiers moments non centraux (MNC) des séries de débits d'étiage pour chaque station contenue dans la région d'étude.

La figure 4.2 représente le tracé du logarithme des MNC en fonction du logarithme des surfaces de drainage.

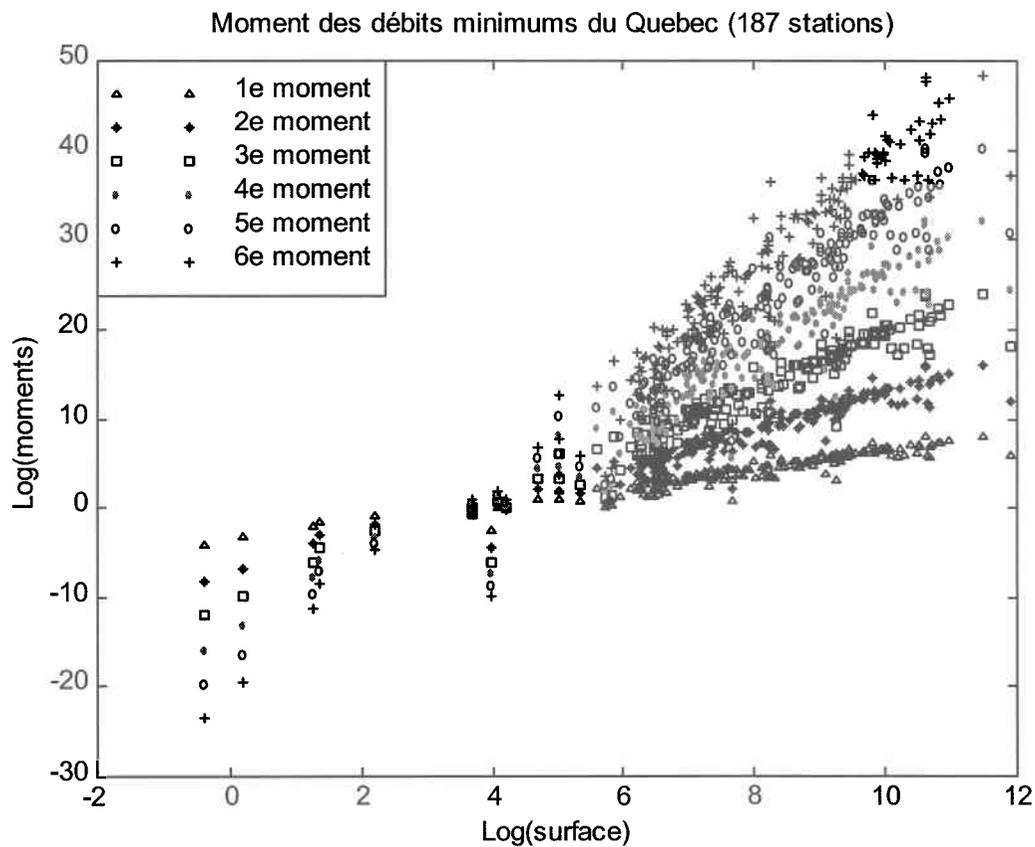
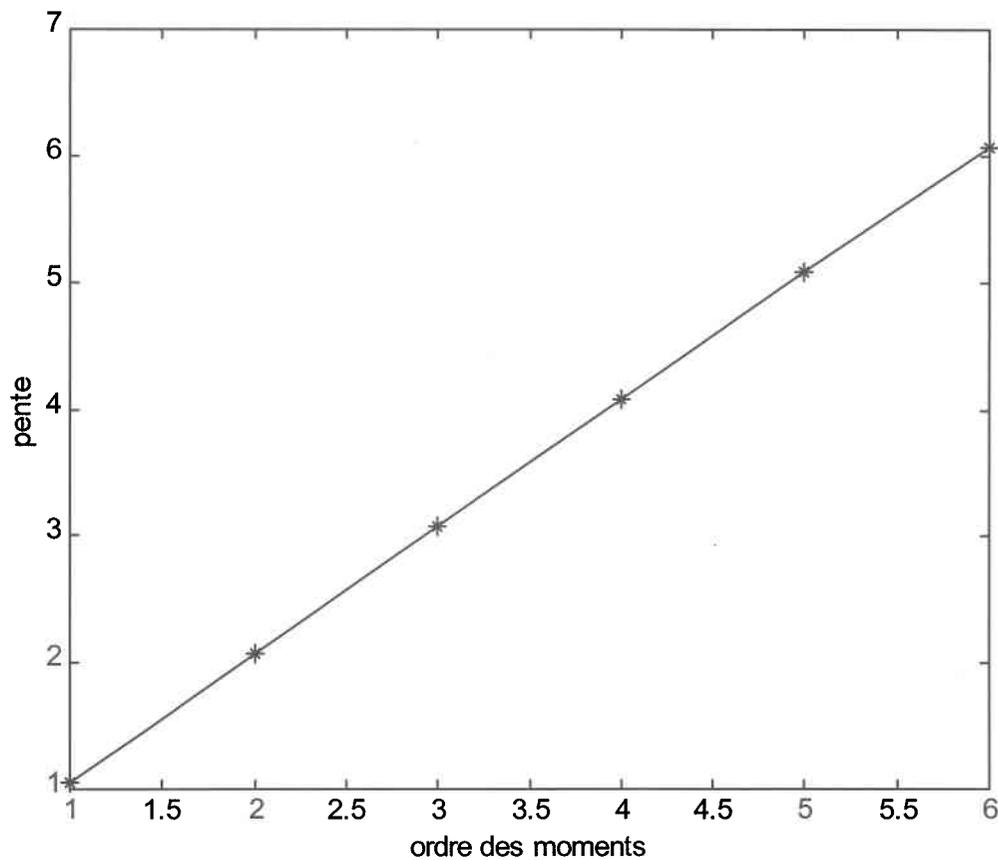


Figure 4.2 Tracé des moments statistiques des débits d'étiage en fonction de la surface de drainage (187 sites).

Dans ce cas, l'adéquation avec les hypothèses de la similitude d'échelle simple paraît acceptable. En effet, on remarque de manière globale, une linéarité entre le logarithme des moments statistiques des débits minimums et le logarithme des surfaces des bassins versants, la première hypothèse de l'invariance d'échelle simple (I.E.S) est donc vérifiée. (cf. 3.2.1)

La figure 4.3 qui traduit l'évolution de la pente avec l'ordre des moments, montre que les pentes pour les six premiers moments sont proportionnelles aux ordres des moments correspondants, la seconde hypothèse relative à l'I.E.S est aussi satisfaite.



**Figure 4.3** Ligne de régression des pentes (187 sites).

Les estimateurs des coefficients de régression pour les six premiers MNC sont déduits à partir de l'équation (3.8a), et sont reportés dans le tableau 4.1.

Québec		
Ordre des MNC	$\alpha(k)$	$\beta(k)$
1	1.0448	- 4.1899
2	2.0695	- 8.1134
3	3.0815	- 11.8603
4	4.0844	- 15.4748
5	5.0800	- 18.9830
6	6.0698	- 22.4027

**Tableau 4.1 Estimation des paramètres représentant la relation entre les moments non centraux (MNC) des étiages et la surface du bassin versant.**

Finalement, ces premiers résultats empiriques indiquent clairement que l'approche de la similitude d'échelle simple peut-être appliquée pour la modélisation des débits d'étiage du Québec.

### 4.3 Délimitation des régions homogènes

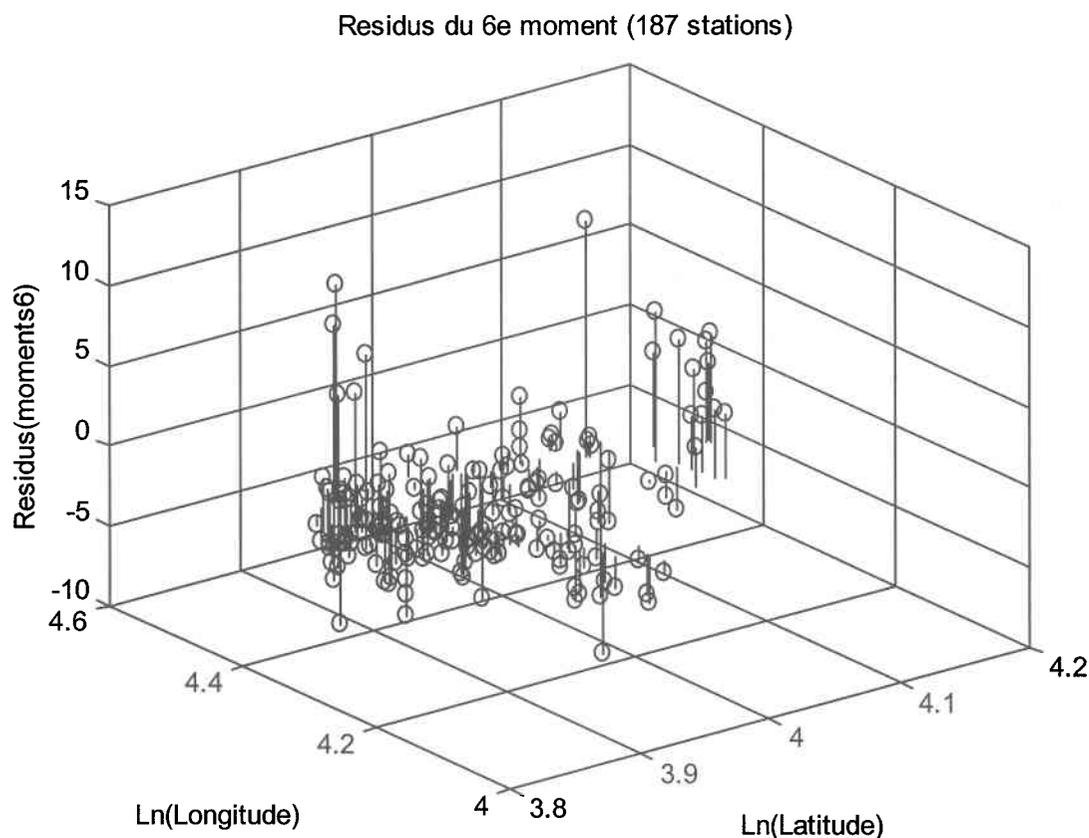
La détermination des régions homogènes consiste à regrouper, pour une variable donnée (où une combinaison de variables données), les stations ayant le même comportement hydrologique. Conscient de l'influence des caractéristiques physiographiques, géologiques et hydrologiques locales sur le régime des basses eaux, on a tenté une première approche de régionalisation en divisant le territoire Québécois en zone d'étiage dite homogène. La délimitation de zones d'étiage homogènes repose sur trois critères mesurables. Le premier,

le coefficient de variation des débits interannuel, qui exprime les conditions de variabilité dans l'espace et qui caractérise la dispersion des valeurs de la série autour de la moyenne, le second, l'altitude moyenne de la station de jaugeage, qui est un facteur physique important, enfin, le dernier qui est tout aussi important comme facteur physique, la surface de drainage. En général, une région est considérée comme homogène, si le coefficient de variation est constant dans cette région. En considérant cette condition ainsi que les paramètres décrits précédemment, nous avons essayé de déterminer des régions où les stations avaient des caractéristiques comparables ; malheureusement, il a été impossible de procéder à une délimitation objective des régions homogènes et nous nous sommes vite rendu compte que ce mode de regroupement basé sur l'homogénéité physique, rend les résultats mathématiques moins précis, compte tenu de l'imprécision de la mesure des paramètres, et par conséquent affecte la justesse des relations trouvées.

Gupta et al (1994) ont critiqué la notion d'homogénéité fondée sur la constance du coefficient de variation dans une région, comme étant « ad hoc et pas très utile » pour le développement d'une théorie statistique des événements extrêmes régionaux, qui repose sur des bases physiques. Nguyen et Pandey (1996), dans leur étude sur l'estimation régionale des débits de crue au Québec, ont défini une région homogène, comme étant une région où toutes les séries de débits de crue possèdent des propriétés d'échelle similaires. En d'autres termes, la relation entre les moments non centraux et la surface des bassins versants (équation 3.8a du sous chapitre invariance d'échelle simple), pour différents bassins versants contenus dans une même région homogène peut-être représentée par une ligne droite sur un graphique logarithmique, donnant ainsi un ensemble unique des paramètres  $\alpha(k)$  et  $\beta(k)$ . On peut donc regrouper les bassins similaires à partir d'un plan (log(moments) en fonction de log(surface)) en sous groupes basés sur l'égalité des coefficients  $\alpha(k)$  et  $\beta(k)$ .

En nous basant en partie sur le concept de délimitation des régions homogènes proposé par Nguyen et Pandey, nous nous sommes intéressés aux résidus des moments calculés à partir des moments.

La figure 4.4 montre le tracé, sur un graphique à trois dimensions, des résidus du sixième moment en fonction de la longitude et la latitude.



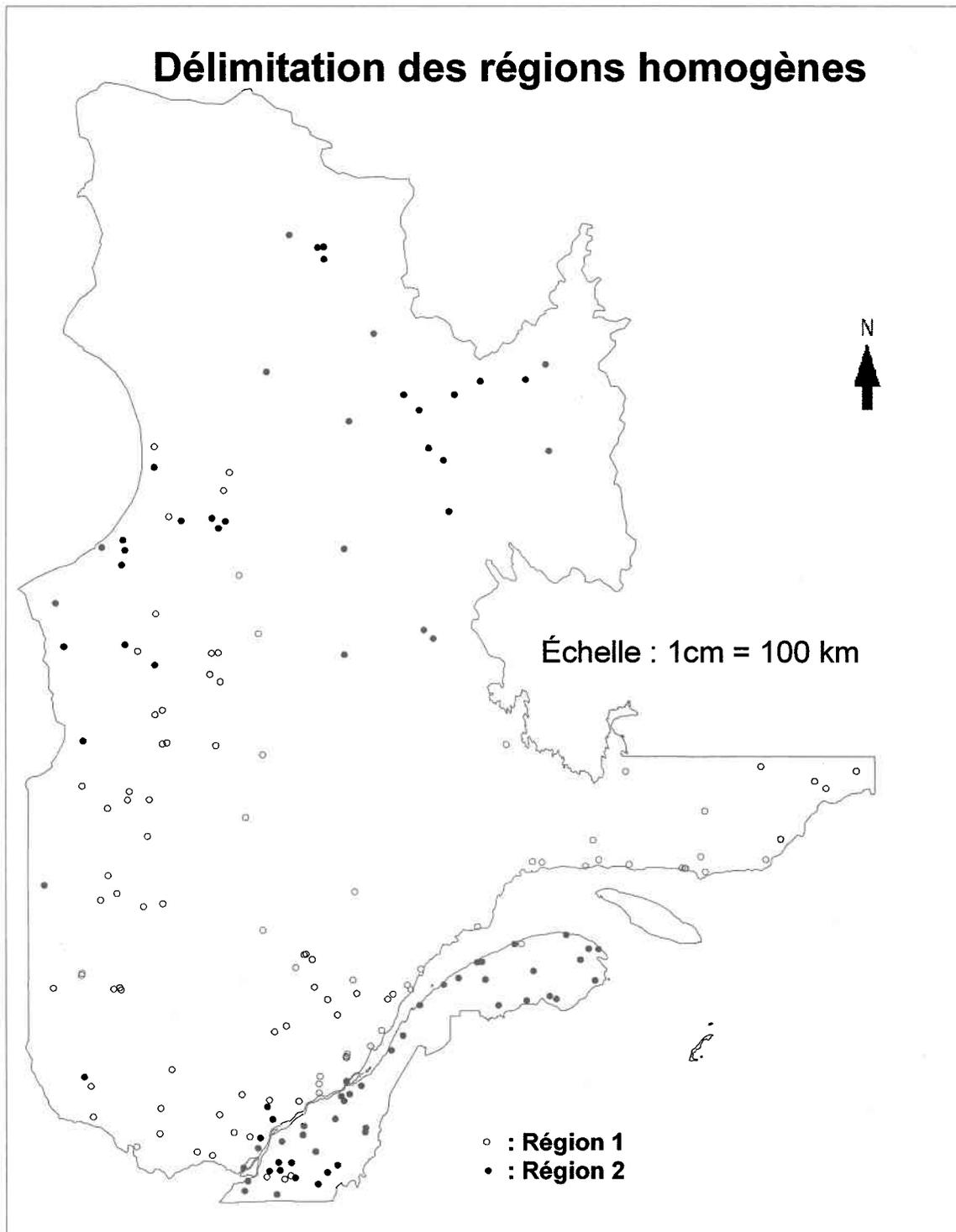
**Figure 4.4 Répartition des résidus sur une base géographique**

Cette approche a été adoptée afin de montrer la répartition des erreurs sur une base géographique, et tenter une première détermination de régions homogènes. On peut observer à travers cette figure que bien que cette distribution géographique paraisse hétérogène, il est cependant possible de distinguer deux zones où le comportement des stations est différent. La première région située plus au centre, dont les résidus sont minimums et généralement négatifs. La deuxième région constituée de deux sous groupes de stations situées respectivement au nord et au sud de la province, dont les résidus sont

positifs et maximums. A partir de cette constatation nous avons scindé la province en deux groupes comme le montre la figure 4.5. La validité de toute méthode de délimitation de régions homogènes doit être appuyée par quelques évidences physiques, telle que la proximité géographique des stations ainsi que la similarité des caractéristiques climatiques. À cette effet, la localisation géographique des stations pour chaque sous groupe homogène est illustrée par la figure 4.5. On peut voir que, sans tenir compte des dimensions des bassins versants, la nouvelle délimitation des groupes homogènes est relativement bien définie dans un contexte statistique, car la délimitation se base surtout sur la distribution des résidus des moments sur un plan géographique. Par contre, si on tient compte des évidences physiques, on notera que la région 2 se divise en deux sous groupes montrant des caractéristiques physiques différentes. Un groupe situé au Nord de la limite des arbres dont le climat est sub-artique, alors que le deuxième groupe situé plus au Sud (Appalaches) dont le climat est tempéré. En effets, en tenant compte du contexte physique, on peut scinder la province de Québec en trois groupes homogènes. Mais dans le présent contexte on se limitera à la délimitation statistique des régions homogènes.

Notons que quelques bassins versants situés à la limite entre deux groupes, ont été considérés comme appartenant à l'un ou l'autre des régions, selon les valeurs des résidus des moments correspondants.

Finalement, la région 1 contient 108 stations et la région 2 en contient 79.

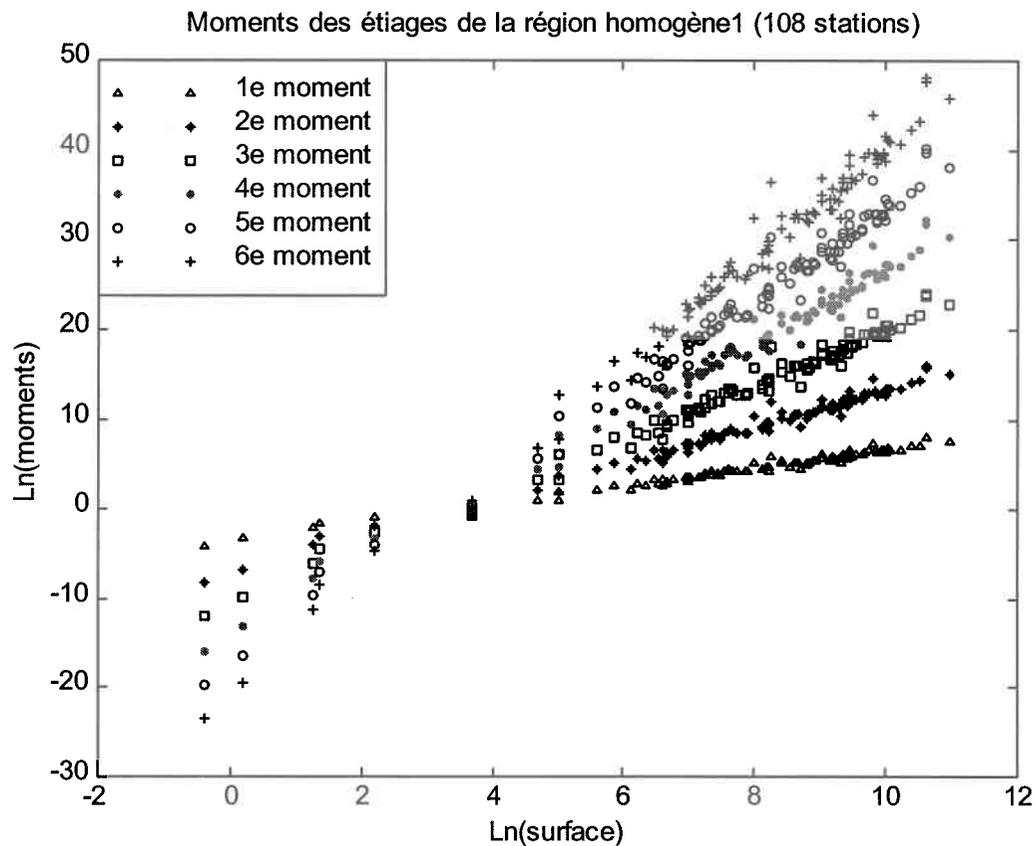


**Figure 4.5** Localisation géographique de 187 sites jaugés de la province de Québec.

Après avoir déterminé les régions homogènes, nous avons appliqué l'approche de la méthode d'invariance d'échelle simple aux deux régions respectivement.

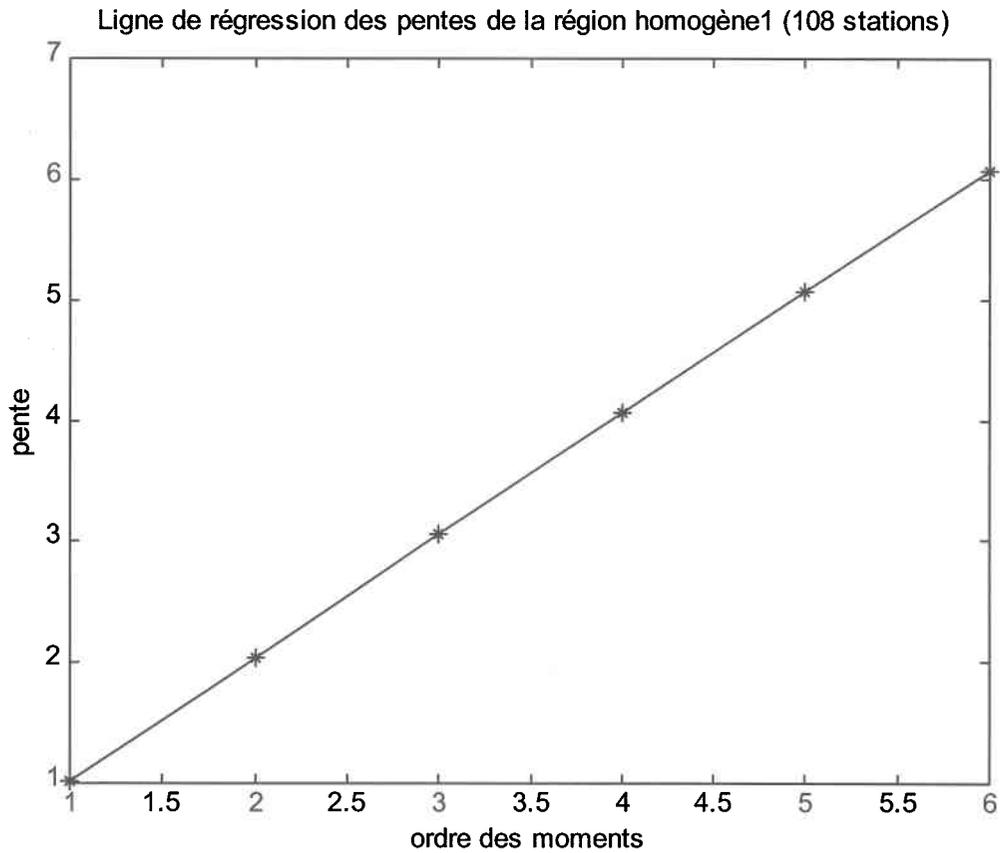
On remarque clairement sur la figure 4.6 représentant les résultats des données de la première région homogène, que les conditions relatives à une similitude d'échelle simple sont respectés :

Linéarité du logarithme des six premiers moments non centraux en fonction du logarithme de la superficie des bassins versants (figure 4.6a) ; et



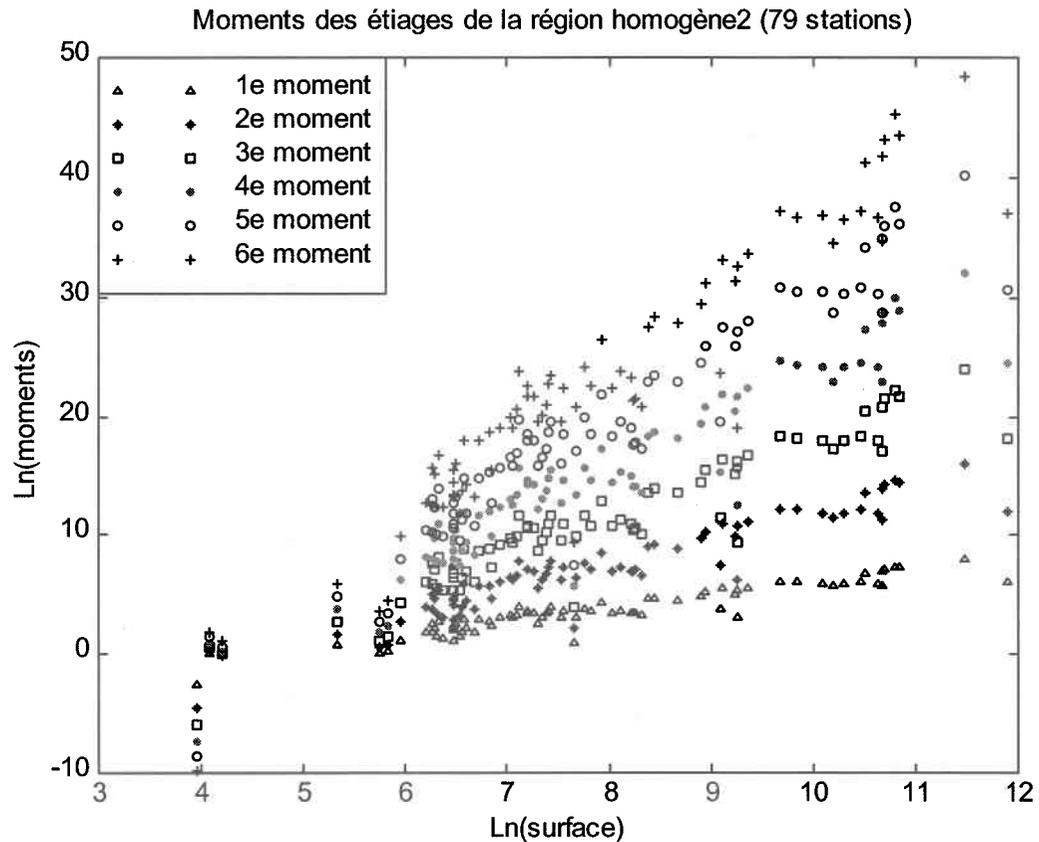
**Figure 4.6a** Tracé des moments statistiques des débits d'étiage en fonction de la surface de drainage pour la région homogène1 (108 sites).

Une croissance linéaire de la pente en fonction de l'ordre des moments (figure 4.6b).



**Figure 4.6b** Ligne de régression des pentes correspondant à la région1 (108 sites).

Les résultats obtenus pour la région homogène 2 illustrés par la figure 4.7 exhibent aussi des caractéristiques d'invariance d'échelle simple. Cette figure montre que les deux conditions à la base de la méthode d'invariance d'échelle simple sont respectées.

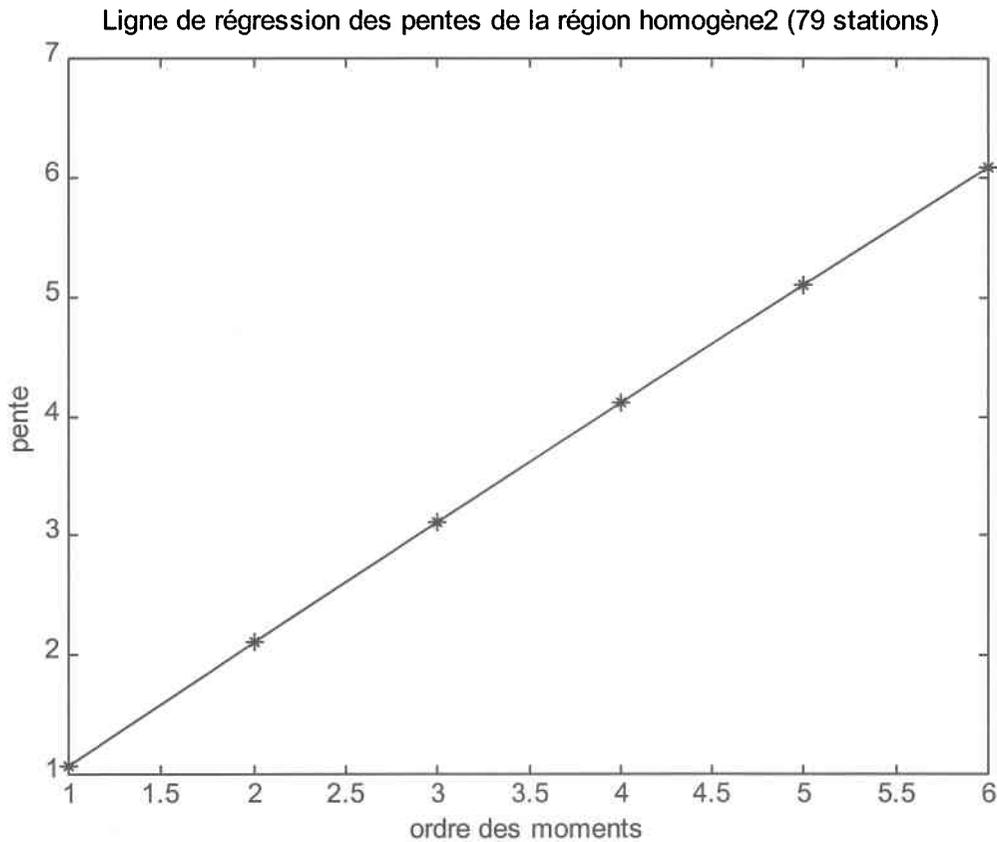


**Figure 4.7a** Tracé des moments statistiques des débits d'étiage en fonction de la surface de drainage pour la région homogène 2 (79 sites).

On observe :

La log-log linéarité des six premiers moments non centraux des débits d'étiage en fonction de la surface de drainage (figure 4.7a) ; et

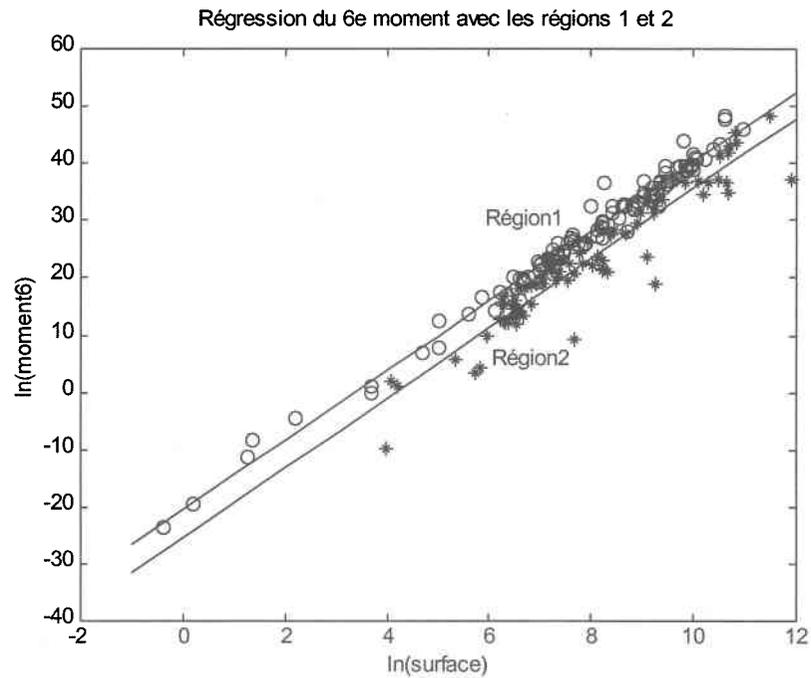
La proportionnalité entre les pentes des moments et l'ordre de ces derniers (figure 4.7b).



**Figure 4.7b** Ligne de régression des pentes correspondant à la région 2 (79 sites).

L'hypothèse d'invariance d'échelle simple est donc vérifiée pour les données de débits d'étiage des deux régions.

La figure 4.8 représente la régression des logarithmes des moments d'ordre six en fonction des logarithmes des superficies de drainage avec respectivement la région1 et la région2. Cette figure montre qu'il est possible de distinguer deux régions différentes.



**Figure 4.8 Relations entre les moments et les surfaces de drainage pour les deux régions homogènes.**

La figure 4.9, représente la régression des logarithmes des moments d'ordre six versus les logarithmes des superficies de drainage avec respectivement la région 1 (figure 4.9a) et la région 2 (figure 4.9b), considérées indépendamment.

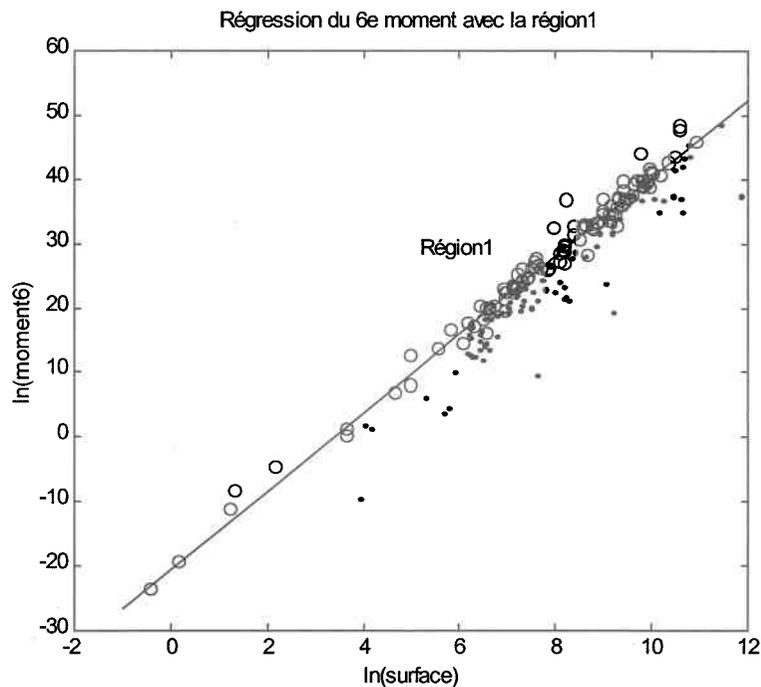


Figure 4.9a Relation entre le sixième moment de la région 1 et la surface de drainage

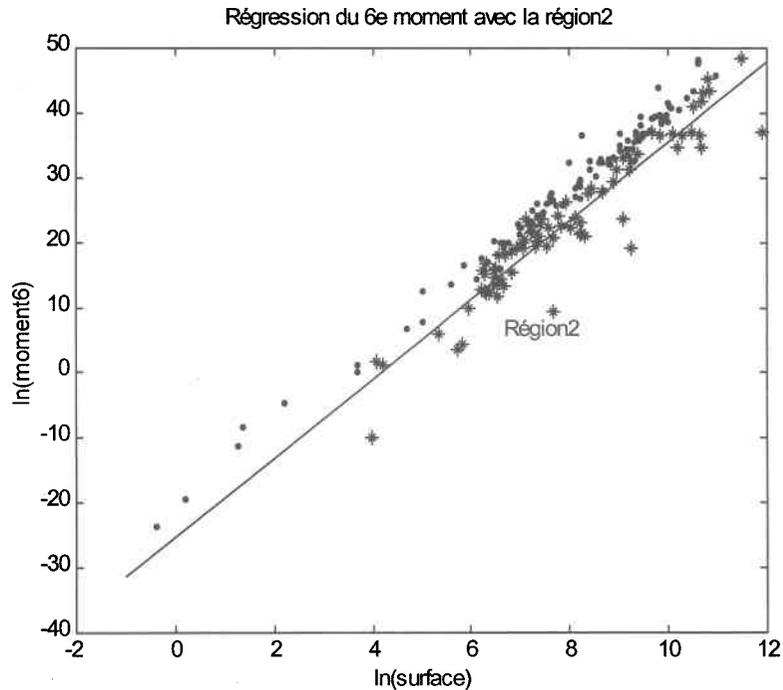


Figure 4.9b Relation entre le sixième moment de la région 2 et la surface de drainage

## 4.4 Application du modèle de queue conditionnel des distributions Weibull et Lognormale

Comme on l'a montré au chapitre 3, l'application des modèles de queue nécessite la détermination *a priori* d'un seuil  $u$ . À cet effet, nous proposons dans ce qui suit l'approche que nous avons adoptée pour la définition de ce seuil.

### 4.4.1 Détermination du seuil $u$

La détermination du seuil  $u$  a été développée sur la base d'une comparaison de différents rapports correspondants à différents percentiles. Le rapport choisi est  $Q_{10cen} / Q_{10emp}$ , où  $Q_{10cen}$  représente le débit de période de retour 10 ans, estimé à partir de l'ajustement d'une distribution (dans le présent contexte, les distributions Weibull et Lognormale) aux données de débits d'étiage qui sont inférieures ou égales au seuil  $u$  choisi dans un premier temps de manière arbitraire.  $Q_{10emp}$  est le quantile de période de retour 10 ans estimé à partir du graphique de la probabilité empirique. Cette dernière correspond à une probabilité au non dépassement  $P_i$  et elle est obtenue à partir de la formule de Cunnane dont l'équation s'écrit :

$$P_i = \frac{(i - 0.4)}{(n - 0.2)}$$

Cette probabilité est attribuée aux débits minimums classés en ordre croissant. Pour une période de retour  $T$  fixée, on détermine la probabilité au non dépassement  $P = \frac{1}{T}$  et on déduit la valeur  $i$  telle que  $P_i < P < P_{i+1}$ . Le débit  $Q(T)$  est alors interpolé entre les valeurs  $Q_i$  et  $Q_{i+1}$  correspondant respectivement aux probabilités au non dépassement  $P_i$  et  $P_{i+1}$ , et peut-être calculé graphiquement sur du papier normal ou directement avec la relation :

T étant fixe et  $P = 1/T$ , on cherche  $i$  tel que  $P_i < P < P_{i+1}$

$$Q(T) = Q_i + (U - U_i) \frac{Q_i - Q_{i+1}}{U_i - U_{i+1}} \quad (4.1)$$

où  $U, U_i$  et  $U_{i+1}$  correspondent aux variables normales dont les probabilités au non dépassement sont respectivement  $P, P_i$  et  $P_{i+1}$ .

Les percentiles 25%, 50%, 60% et 75% ont été choisis pour l'estimation de  $Q_{10cen}$  afin de procéder à l'analyse des rapports correspondants. Le choix de ces percentiles est fait de manière arbitraire.

#### 4.4.1.1 Estimation de $u$ (avec la distribution Weibull censurée)

La figure 4.10 (a), est un graphique « box plot » que nous avons nommé « boîtes à ailes » représentant le ratio  $(Q_{10cen} / Q_{10emp})$ . Voici une brève description de ce genre de graphique.

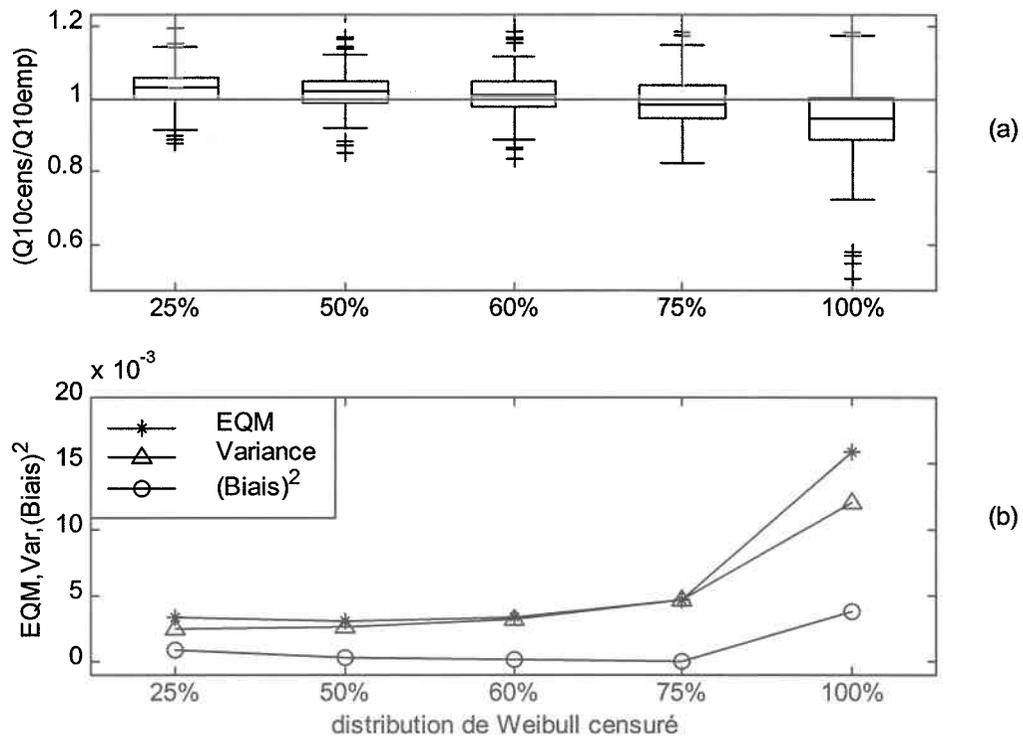
La boîte centrale contient 50% des valeurs situées entre le troisième quartile (75%) et le premier quartile (25%). La distance entre le premier et le troisième quartile est appelée distance interquartile (D.I.Q).

La ligne centrale représente la médiane.

Les « ailes » s'étendent jusqu'à la valeur la plus éloignée à l'intérieur de la distance  $1,5D.I.Q$  du premier quartile et du troisième quartile.

Les (+) sont des valeurs « extrêmes » situées à l'extérieur des « ailes » à une distance supérieure à  $3.0D.I.Q$ .

Sur cette figure, l'axe vertical représente le rapport de l'estimateur déterminé avec la méthode du modèle de queue conditionnel de la distribution Weibull en fonction de l'estimateur empirique.



**Figure 4.10(a) Comparaison du rapport  $Q_{10cens} / Q_{10emp}$  pour les différents percentiles**

**Figure 4.10(b) Erreur quadratique moyenne, Variance et Biais<sup>2</sup> pour les différents percentiles**

Le calcul de ces rapports a été effectué pour chacun des 187 sites et pour chacun des quatre percentiles considérés. À partir de la figure 4.10 (a) le percentile 60% apparaît comme un seuil de censure adéquat pour l'échantillon régional. Néanmoins, cette information est incomplète car l'estimation est basée uniquement sur l'évaluation du biais. Ce dernier représente, dans notre cas, la différence entre la médiane et la valeur définissant le rapport unitaire, c'est une valeur centrée sur 1. En effet, il serait donc plus approprié d'introduire une approche plus rigoureuse basée sur l'évaluation de l'erreur quadratique moyenne (EMQ) pour l'estimation du seuil  $u$ . L'EMQ est un paramètre qui définit non seulement le

biais, mais aussi la dispersion des valeurs autour de la moyenne qui est la variance ( $EMQ = Variance + Bais^2$ ). La figure 4.10 (b) représente l'EMQ en fonction des différents percentiles. On remarque bien sur ce graphique que l'EMQ minimum est associée au percentile 50% et 60%. Une analyse de sensibilité a été effectuée entre l'EMQ du percentile 50% et celle du percentile 60%, qui a permis de trouver une différence de 0.2%, non significative. Finalement, nous avons choisi comme seuil de censure le percentile 60%

#### 4.4.1.2 Estimation de $u$ (avec la distribution Lognormale censurée)

Afin d'avoir une vue d'ensemble, et dans le but de choisir un seuil de censure  $u$  standard, nous avons comparé les rapports de quantiles ( $Q_{10cen} / Q_{10emp}$ ) pour les mêmes percentiles que ceux choisis dans la section précédente. L'estimation du quantile  $Q_{10cen}$  a été effectuée avec la distribution Lognormale.

La figure 4.11 montre que les percentiles 60% et 75% sont plus appropriés comme seuil de censure des données, car leur erreur moyenne quadratique est minimum. En combinant les résultats obtenus pour l'estimation de  $u$  avec les distributions Weibull censurée et Lognormale censurée, nous avons opté pour le choix d'un seuil de censure correspondant au 60<sup>ème</sup> percentile.

On remarque sur ce graphique, que l'estimation des débits d'étiage de la province de Québec en considérant l'échantillon au complet (100% des données) est moins performante que celle qui considère l'échantillon censuré, pour différents niveaux de censure (25%, 50%, 60% et 75%). L'erreur moyenne quadratique du percentile 100% est plus grande que celle des niveaux de censure 50%, 60% et 75%. La même constatation se retrouve sur le graphique 4.10(a) et 4.10(b), qui représente l'estimation du seuil  $u$  avec la distribution de Weibull.

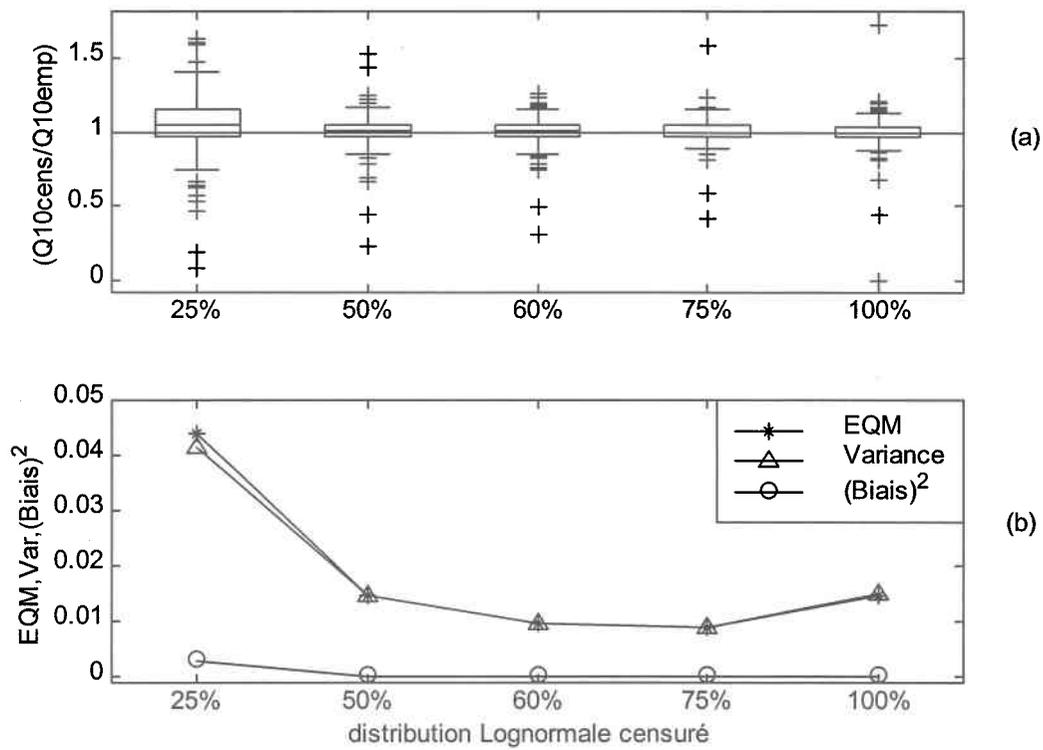
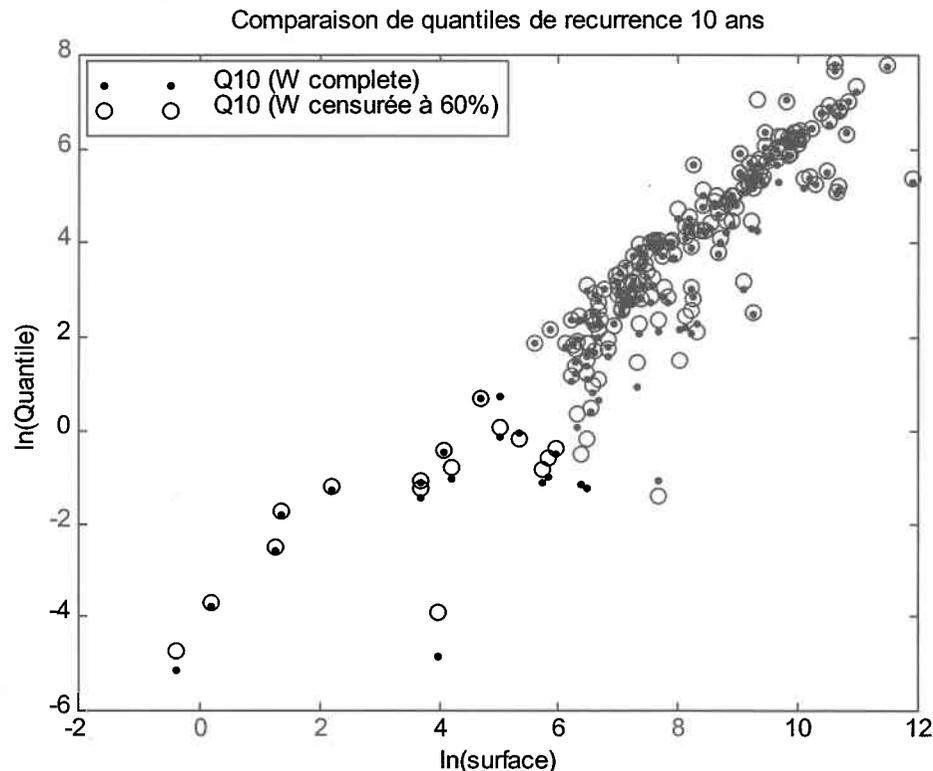


Figure 4.11(a) Comparaison du rapport  $Q_{10cens} / Q_{10emp}$  pour les différents percentiles

Figure 4.11(b) Erreur quadratique moyenne, Variance et Biais<sup>2</sup> pour les différents percentiles.

### Performance des modèles de queue

Afin de tester la performance des modèles de queue, nous avons décidé de comparer les quantiles de période de 10 ans estimés à partir de la distribution de Weibull censurée, avec un seuil correspondant au 60<sup>ème</sup> percentile, à ceux obtenus par la distribution Weibull complète. Les résultats de cette comparaison sont illustrés par la figure 4.12.



**Figure 4.12** Comparaison entre  $Q_{10}$  estimé par le modèle de queue de Weibull et  $Q_{10}$  estimé par le modèle de Weibull complet.

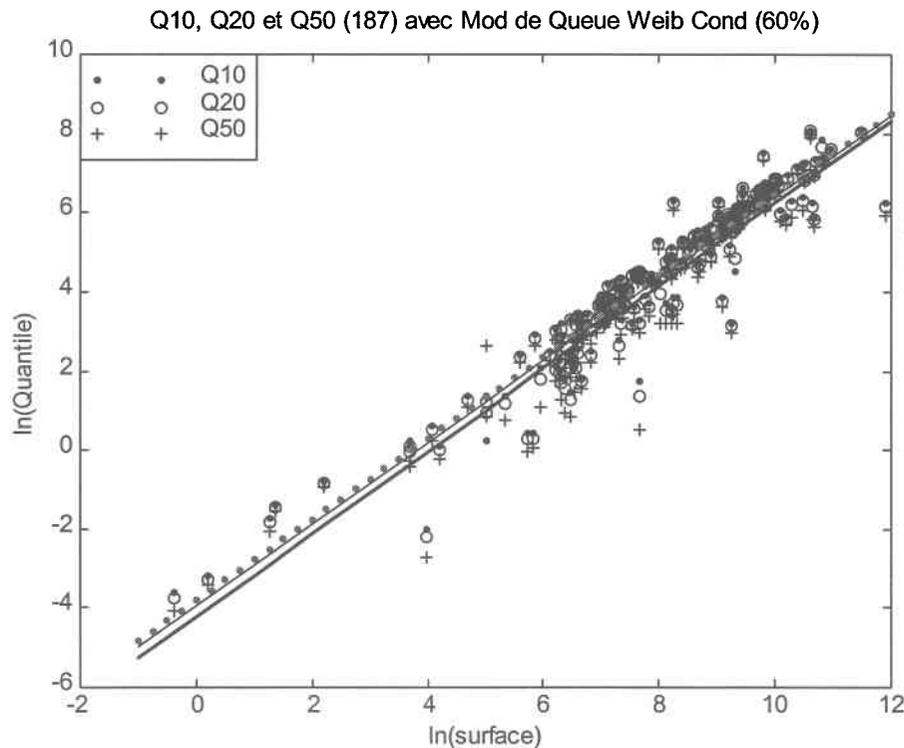
L'examen de cette figure montre que les quantiles estimés à partir d'une distribution de Weibull censurée à 60% sont en général comparables, à quelques exceptions, à ceux obtenus par la distribution de Weibull complète. Nous pouvons donc avancer l'hypothèse que le modèle de queue conditionnel de la distribution de Weibull peut-être substitué au modèle de Weibull complet pour l'estimation régionale des débits minimums.

#### **Définition alternative de la similitude d'échelle simple**

Une manière différente de définir une invariance d'échelle simple serait d'estimer, pour différentes période de retour, les débits correspondants pour chacun des 187 sites, et de faire par la suite une régression sur les différents quantiles. Si les droites résultants des

régressions sont parallèles entre elles, on peut en conclure que les débits d'étiage considérés suivent une invariance d'échelle simple.

La figure 4.13 montre la régression des logarithmes des quantiles de périodes de retour 10 ans, 20 ans et 50 ans en fonction des logarithmes des surfaces de drainage en considérant les 187 sites comme faisant partie d'une région homogène.



**Figure 4.13** comparaison des droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$

La visibilité des droites de régression est réduite à cause du nuage de données des quantiles. Afin de bien les distinguer, nous avons ôté les points du graphique, les résultats détaillés se retrouvent à l'annexe A. Les quantiles ont été estimés à partir d'une distribution Weibull censurée à 60% ; la même analyse a été effectuée en utilisant la distribution Lognormale censurée à 60%, ainsi que la distribution de Weibull complète. Les régions 1 et 2 ont aussi été explorées, avec respectivement les trois modèles cités précédemment. On

retrouvera les résultats obtenus à l'annexe A. Notons que la comparaison des droites de régression a été effectuée de manière subjective, après un examen visuel du graphique.

#### **4.4.2 Estimation de $X_T$ à un site non jaugé**

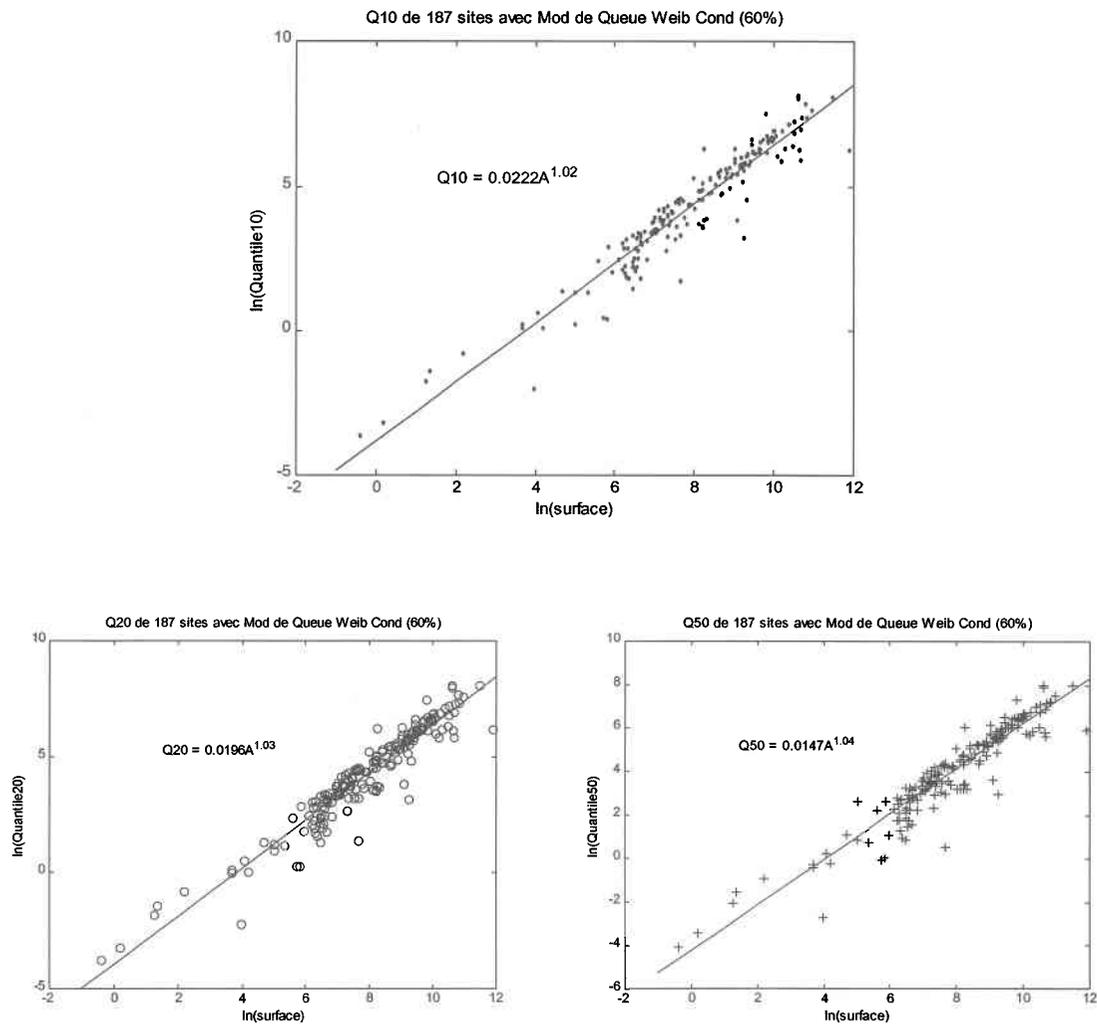
Lorsque l'on veut estimer le débit  $X_T$  régionalement pour un site non jaugé, nous devons l'affecter à un des deux groupes de stations homogènes pour lesquelles nous disposons d'une série de débits minimums annuels mesurés. Les régions obtenues sont en partie délimitées géographiquement, il suffirait donc de situer la rivière non jaugée sur la carte représentée par la figure 4.5 et de l'affecter à son groupe. Par la suite, connaissant la superficie de drainage des stations non jaugées et l'équation résultant du développement des modèles de prédiction régionaux à l'aide des stations jaugées, on peut estimer  $X_T$  pour chacune des stations non jaugée. Afin d'étudier les relations entre  $X_T$  et la surface du bassin versant, pour différentes périodes de retour (dans le présent cas, 10 ans, 20 ans et 50 ans), nous avons dans un premier temps effectué une transformation logarithmique sur les données originales (pour  $X_T$  et la surface de drainage). Cette transformation a pour but de linéariser la relation entre ces deux variables (hypothèse de base en régression). Plusieurs distributions de probabilité ont été suggérées comme étant adéquates pour la modélisation des débits d'étiage. Riggs (1972), entre autres, a trouvé que les distributions Weibull, Lognormale et Pearson type3 sont des distributions adéquates pour l'estimation régionale des débits minimums. À cet effet, nous avons considéré les distributions Weibull et Lognormale dont l'ajustement des paramètres est effectué en considérant des échantillons censurés. Pour des fins de comparaison, nous avons examiné l'estimation régionale, avec la distribution de Weibull en considérant l'échantillon complet. L'élaboration des modèles régionaux a été effectuée d'abord en considérant les deux classes de stations homogènes. La partition des 187 stations en deux groupes de stations permet une meilleure estimation régionale de  $X_T$  sur une rivière non jaugée si les stations des deux groupes ont des comportements hydrologiquement différents. Afin de vérifier cette condition, nous avons considéré, lors de l'estimation régionale, l'ensemble de la région comme étant homogène.

Une telle analyse permet d'évaluer l'impact de la subdivision, c'est à dire, si on obtient un raffinement substantiel de la définition d'une région homogène qui se traduit par l'amélioration dans l'estimation des quantiles.

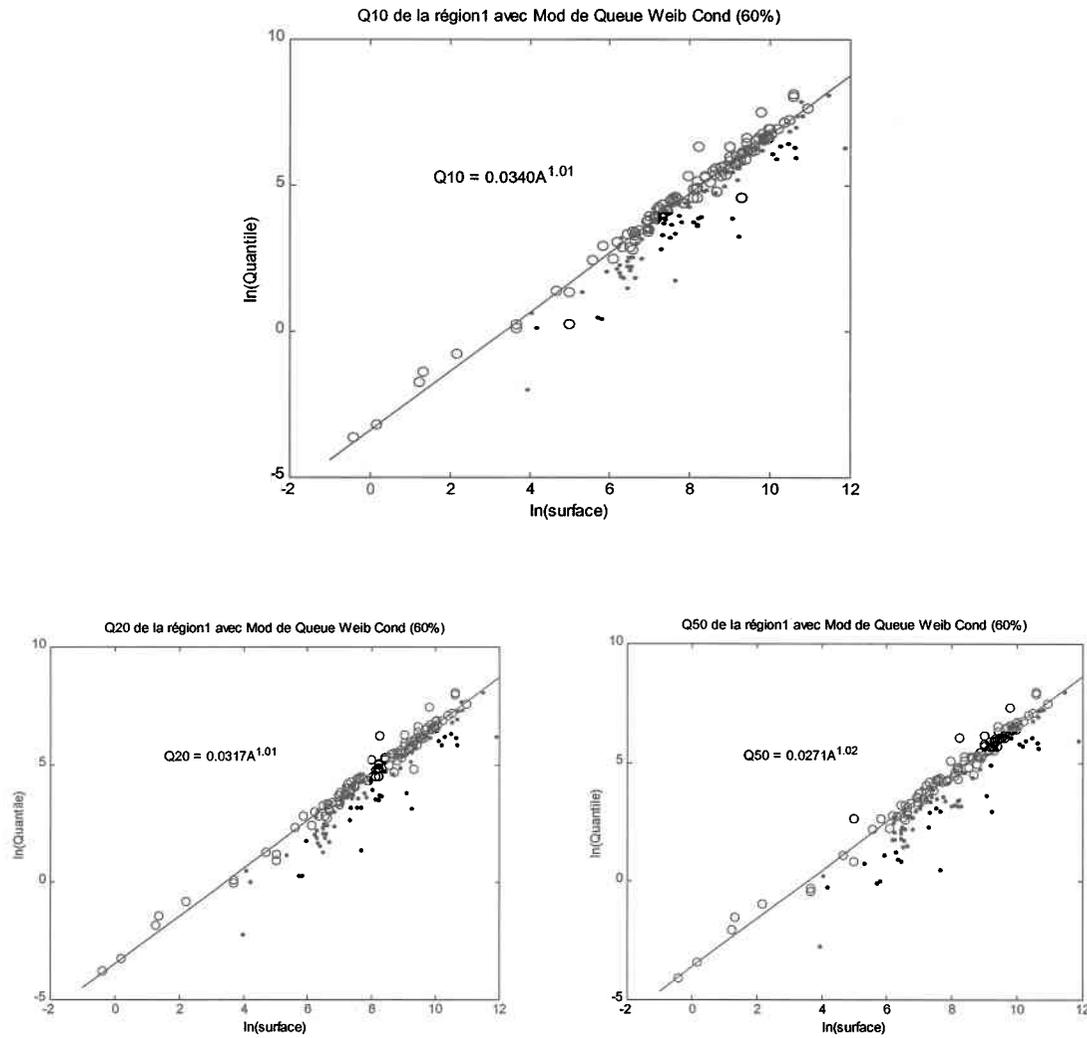
La régression a donc été effectuée en considérant les 187 sites comme faisant partie d'une région homogène. Nous proposons dans ce qui suit les résultats des différents modèles qu'il est possible d'appliquer pour l'estimation régionale de débits d'étiage correspondant aux périodes de retour 10 ans, 20 ans et 50 ans respectivement.

### Modèle de queue conditionnel de la distribution Weibull censurée à 60%

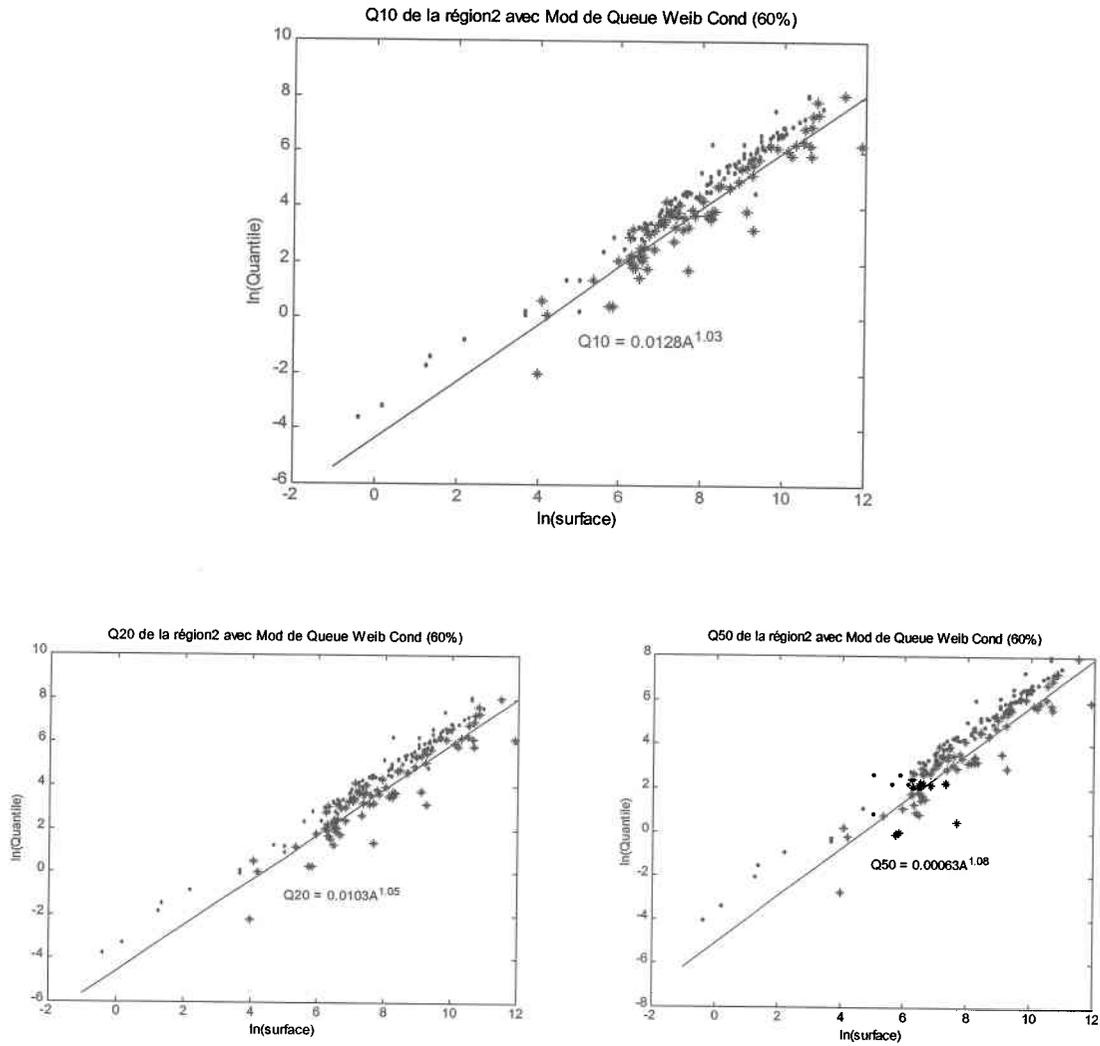
i) Toute la région est considérée comme homogène



**Figure 4.14** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  (187 sites)  
avec Weibull censurée à 60%

ii) Région 1

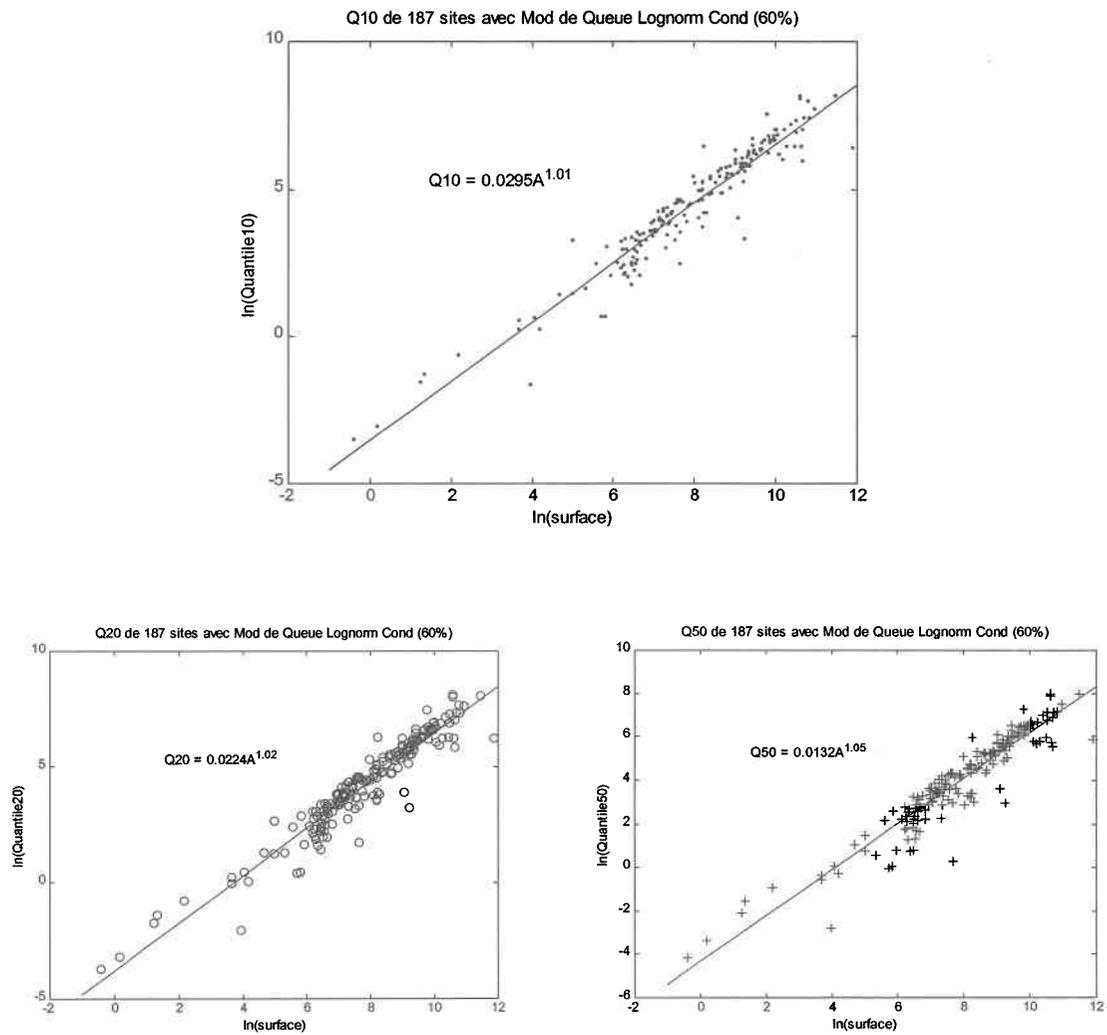
**Figure 4.15** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 1 (108 sites) avec Weibull censurée à 60%

iii) Région 2

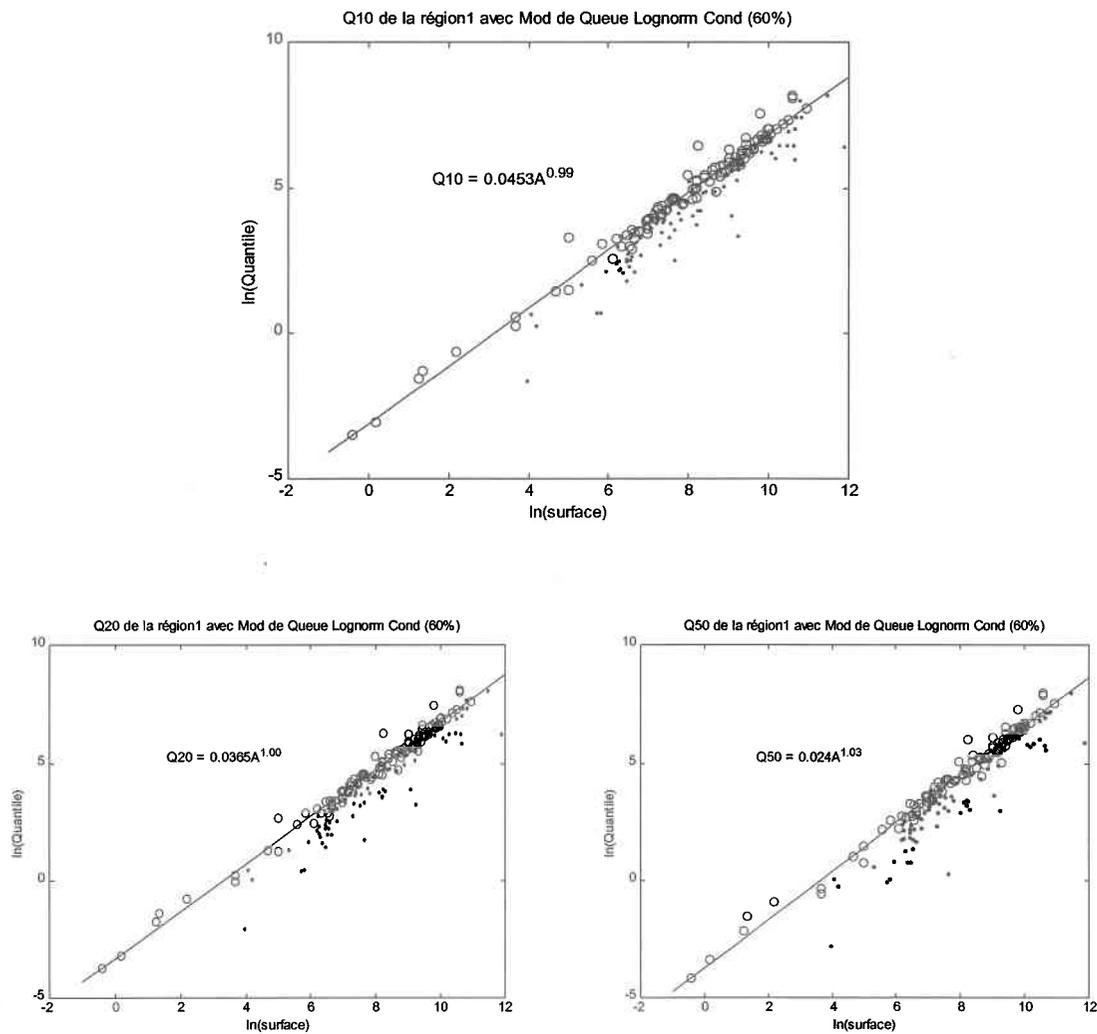
**Figure 4.16** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 2 (79 sites) avec Weibull censurée à 60%

**Modèle de queue conditionnel de la distribution Lognormale censurée à 60%**

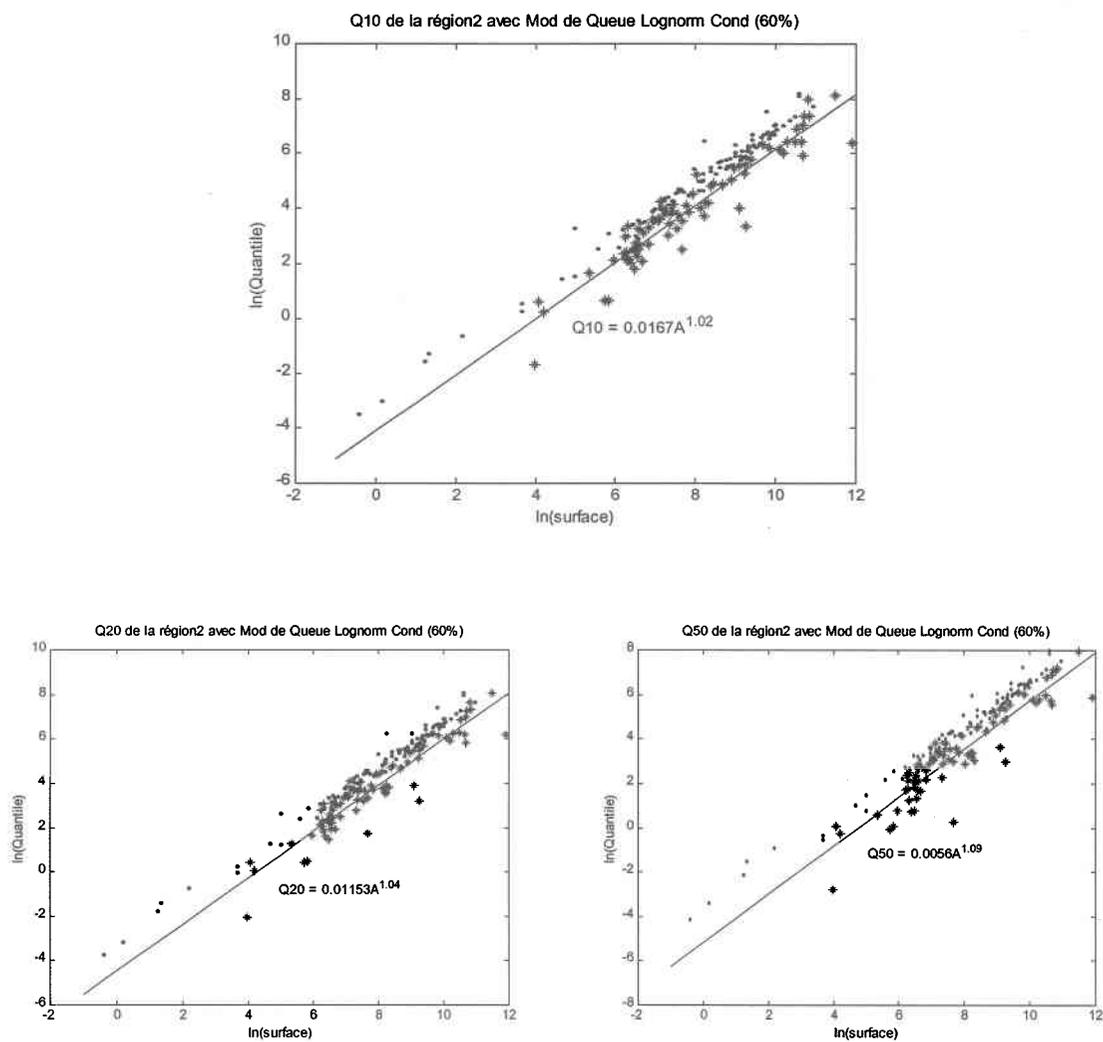
i) Toute la région est considérée comme homogène



**Figure 4.17 Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  (187 sites) avec Lognormale censurée à 60%**

ii) Région 1

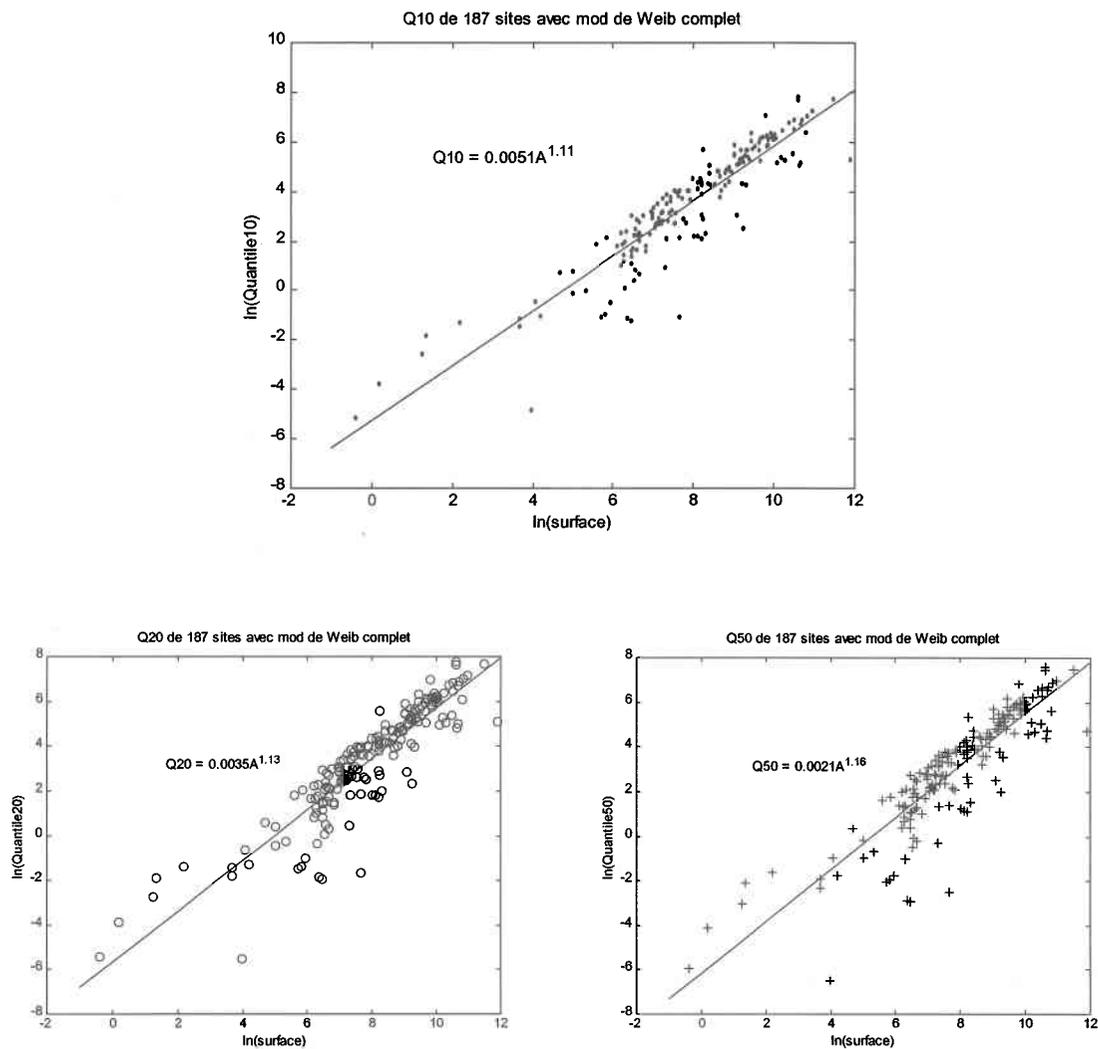
**Figure 4.18** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 1 (108 sites) avec Lognormale censurée à 60%

iii) Région 2

**Figure 4.19** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 2 (79 sites) avec Lognormale censurée à 60%

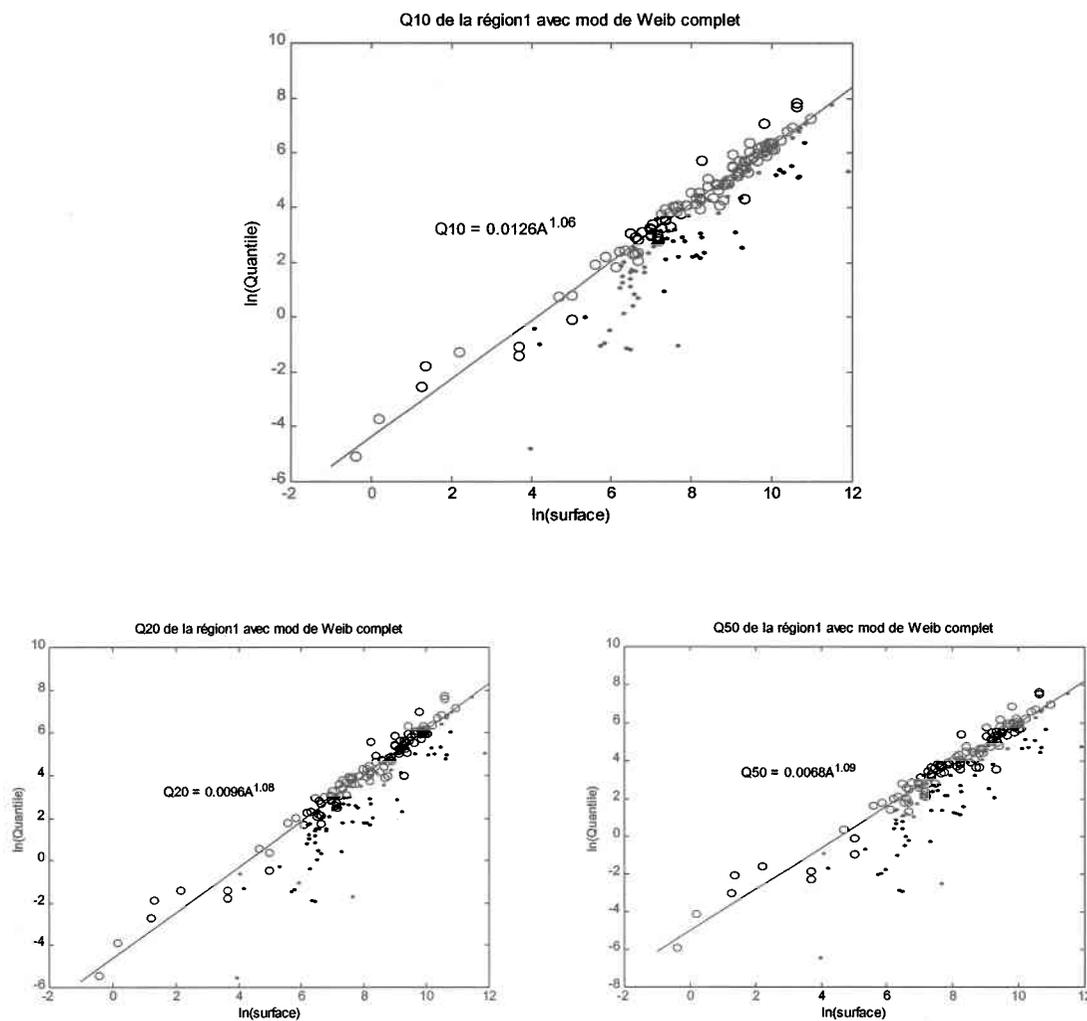
**Modèle de la distribution Weibull complète**

i) Toute la région est considérée comme homogène

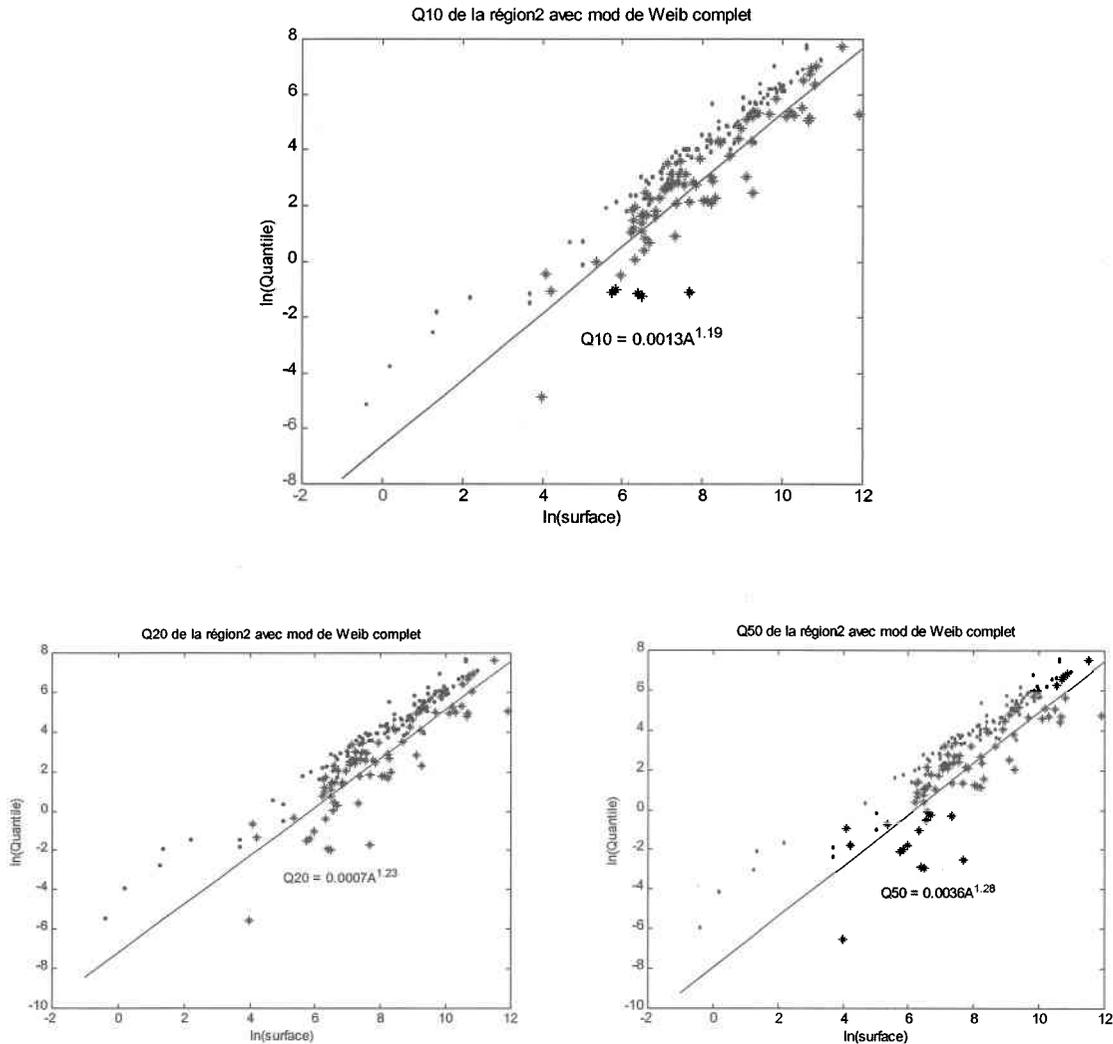


**Figure 4.20** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  (187 sites) avec Weibull complète

## ii) Région 1



**Figure 4.21** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 1 (108 sites) avec Weibull complète

iii) Région 2

**Figure 4.22** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 2 (79 sites) avec Weibull complète

Dans ce qui suit nous présentons le résumé des équations régionales obtenues en considérant les différentes régions homogènes (toute la province, région 1 et région 2).

*Modèle de queue Weibull censuré à 60%*

<b>Toute la province</b>	
Equations	Coefficients de corrélation
$Q_{10} = 0.0222A^{1.02}$	$R^2 = 0.9111$
$Q_{20} = 0.0196A^{1.03}$	$R^2 = 0.9077$
$Q_{50} = 0.0147A^{1.04}$	$R^2 = 0.8904$

*Modèle de queue Lognormal censuré à 60%*

<b>Toute la province</b>	
Equations	Coefficients de corrélation
$Q_{10} = 0.0295A^{1.01}$	$R^2 = 0.9176$
$Q_{20} = 0.0224A^{1.02}$	$R^2 = 0.9114$
$Q_{50} = 0.0132A^{1.05}$	$R^2 = 0.8901$

*Modèle Weibull complet*

<b>Toute la province</b>	
Equations	Coefficients de corrélation
$Q_{10} = 0.0051A^{1.11}$	$R^2 = 0.8394$
$Q_{20} = 0.0035A^{1.13}$	$R^2 = 0.8171$
$Q_{50} = 0.0021A^{1.16}$	$R^2 = 0.7879$

**Tableau 4.2** Sommaire des équations de régressions pour toute la province

*Modèle de queue Weibull censuré à 60%*

<b>Région 1</b>	
Equations	Coefficients de corrélation
$Q_{10} = 0.0340A^{1.01}$	$R^2 = 0.9753$
$Q_{20} = 0.0317A^{1.01}$	$R^2 = 0.9801$
$Q_{50} = 0.0271A^{1.02}$	$R^2 = 0.9794$

*Modèle de queue Lognormal censuré à 60%*

<b>Région 1</b>	
Equations	Coefficients de corrélation
$Q_{10} = 0.0453A^{0.99}$	$R^2 = 0.9790$
$Q_{20} = 0.0365A^{1.00}$	$R^2 = 0.9811$
$Q_{50} = 0.0240A^{1.03}$	$R^2 = 0.9802$

*Modèle Weibull complet*

<b>Région 1</b>	
Equations	Coefficients de corrélation
$Q_{10} = 0.0126A^{1.06}$	$R^2 = 0.9692$
$Q_{20} = 0.0096A^{1.08}$	$R^2 = 0.9628$
$Q_{50} = 0.0068A^{1.09}$	$R^2 = 0.9532$

**Tableau 4.3** Sommaire des équations de régressions pour région 1

*Modèle de queue Weibull censuré à 60%*

<b>Région 2</b>	
Equations	Coefficients de corrélation
$Q_{10} = 0.0128A^{1.03}$	$R^2 = 0.8818$
$Q_{20} = 0.0103A^{1.05}$	$R^2 = 0.8775$
$Q_{50} = 0.0063A^{1.08}$	$R^2 = 0.8624$

*Modèle de queue Lognormal censuré à 60%*

<b>Région 2</b>	
Equations	Coefficients de corrélation
$Q_{10} = 0.0167A^{1.02}$	$R^2 = 0.8955$
$Q_{20} = 0.0115A^{1.04}$	$R^2 = 0.8903$
$Q_{50} = 0.0056A^{1.09}$	$R^2 = 0.8628$

*Modèle Weibull complet*

<b>Région 2</b>	
Equations	Coefficients de corrélation
$Q_{10} = 0.00136A^{1.19}$	$R^2 = 0.8025$
$Q_{20} = 0.00075A^{1.23}$	$R^2 = 0.7778$
$Q_{50} = 0.00036A^{1.28}$	$R^2 = 0.7467$

**Tableau 4.4** Sommaire des équations de régressions pour la région 2

On remarque que les équations ci-dessus présente une certaine cohérence relativement à une invariance d'échelle simple. Notamment une constance de l'exposant d'échelle  $\theta$ , et une décroissance du paramètre  $\beta(K)$ , représentant l'abscisse à l'origine, pour les quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$ . Ceci pour les différents modèles utilisés (Modèle de queue de la distribution Weibull censurée à 60%, Modèle de queue de la distribution Lognormale censurée à 60% et la distribution de Weibull en considérant l'échantillon complet). En considérant tous les sites de la province comme échantillon homogène, on note une constance de l'exposant d'échelle et une diminution du paramètre  $\beta(K)$ , cette remarque est valable pour les différents modèles utilisés. Le coefficient de corrélation varie de 0.7879 à 0.9176. En considérant les régions séparément, on constate que la région 1 présente les mêmes caractéristiques que la province au complet, c'est à dire un exposant d'échelle  $\theta$  relativement constant et le paramètre  $\beta(K)$  décroît, pour les différents modèles. Le coefficient de corrélation varie de 0.9532 à 0.9811. La région 2 présente un exposant d'échelle sensiblement constant et un  $\beta(K)$  qui diminue, ceci pour les différents modèles utilisés. Le coefficient de corrélation varie de 0.7467 à 0.8955. On peut donc avancer l'hypothèse suivante : L'hypothèse d'une invariance d'échelle simple s'applique bien respectivement à la région 1, la région 2 et toute la province.

D'après les résultats ci-dessus, et en se référant aux graphiques 4.10(a), 4.10(b) et 4.11(a), 4.11(b) de la section 4.41 traitant de la détermination du seuil de censure  $u$ , on peut avancer la conclusion suivante : Le modèle le plus performant pour les échantillons de débits d'étiage de la province Québec, est le modèle de queue de la distribution Lognormale, avec un seuil de censure de 60%, ainsi que le modèle de queue de la distribution Weibull, avec un seuil de censure de 60%.

Nous avons présentés les coefficients de corrélation des équations régionales à titre indicatif, afin de montrer qu'il y a un gain substantiel en terme de statistiques, Riggs(1972).

## 5. CONCLUSIONS

---

Les ressources en eau superficielles et souterraines mondiales sont de plus en plus sollicitées du fait de l'essor industriel, du développement agricole, mais surtout en raison de l'accroissement rapide de la population. La connaissance des caractéristiques statistiques extrêmes des rivières, et plus particulièrement des étiages, est essentielle pour la planification et la conception des politiques de gestion et de contrôle des ressources hydriques. Contrairement aux crues, qui sont un phénomène de très courte durée pouvant être représenté par un débit maximal instantané, la préhension du phénomène d'étiage est beaucoup plus délicate. Son impact est dû surtout à la durabilité des périodes de faibles débits ; ce qui rend le choix de la variable d'étiage difficile.

L'analyse régionale s'avère être un outil essentiel pour l'estimation des débits, volumes ou durées d'événements hydrologiques d'étiage en des sites où l'on dispose de peu ou d'aucune information hydrométrique. La revue de littérature, présentée au chapitre 2, a permis de montrer que plusieurs méthodes, depuis l'approche classique proposée par Dalrymple (1949, 1960), tentent de combler le manque d'information hydrométrique en lui substituant, selon différentes méthodes, de l'information physiographique, climatologique, morphométrique ou géomorphométrique. À travers ces contributions, on a démontré que l'analyse régionale des caractéristiques d'étiage a évolué dans une direction qui devait accorder davantage d'attention à la station non jaugée et aux processus associés aux événements hydrologiques pouvant survenir à ce site.

La contribution la plus importante de ce travail réside en l'application d'une méthodologie qui s'inscrit dans l'axe des contributions récentes à l'analyse régionale d'événements extrêmes. Cette procédure repose sur la combinaison des propriétés de la méthode de

l'invariance d'échelle simple des séries régionales des débits d'étiage et de l'analyse de fréquence des débits minimums avec l'approche par modèles de queue conditionnels.

La première étape a consisté à vérifier les hypothèses de base d'une similitude d'échelle simple que doivent respecter les échantillons de débits minimums. L'analyse des caractéristiques des débits d'étiage de 187 sites répartis dans la province de Québec a montré que les moments non centraux d'ordre un à six suivent une invariance d'échelle simple par rapport à la superficie du bassin versant. Suite à ce résultat nous avons déterminé les bassins hydrologiquement semblables à partir d'une analyse des résidus du sixième moment ; cette délimitation correspond à des régions statistiquement homogènes. Le comportement des débits minimums des différentes régions a montré qu'eux aussi suivaient une invariance d'échelle simple.

Les distributions empiriques des débits minimums annuels sont souvent complexes, ce qui rend difficile le choix et l'ajustement d'une distribution statistique adéquate pour la modélisation des échantillons. À cet effet, la seconde étape de notre travail a porté sur l'application des modèles de queue des distributions Weibull et Lognormale pour l'estimation des débits d'étiage en des sites non jaugés. L'avantage de ces modèles réside en leur facilité d'utilisation, car on réduit le nombre de paramètres à estimer et par la même on diminue la variance des estimateurs de quantiles. Cette méthode suppose la détermination *a priori* d'un seuil de censure  $u$ . L'estimation de ce seuil telle que présentée dans la littérature est assez subjective, car basée essentiellement sur un examen visuel de graphiques des probabilités empiriques. Une procédure originale basée sur la comparaison des rapports de quantiles calculés pour différents percentiles a été développée. Cette approche a mené au choix d'un seuil  $u$  standard pour tous les échantillons. Ce seuil est représenté par le percentile 60%. Nous avons ensuite testé la performance des modèles de queue en comparant leurs équations de régression régionale respectives à celles résultant du modèle régional de la distribution Weibull et Lognormale basée sur l'échantillon complet. Nous avons trouvé que le modèle de queue des distributions Lognormale et Weibull, avec un

seuil de censure de 60%, est plus performant relativement au reste des modèles, voir figure 4.10(a), 4.10(b), 4.11(a) et 4.11(b).

La variable d'étiage  $Q_{7,10}$ , a été choisie *a priori*, car c'est la variable la plus couramment utilisée en Amérique du Nord. Dans le but de tester davantage la performance de la méthodologie développée, nous avons choisi de l'appliquer à une autre variable qui est  $Q_{30,5}$ . Cette variable est utilisée fréquemment par le Ministère de l'Environnement du Québec ; les résultats de cette application se trouvent dans l'annexe B. Tout comme la variable  $Q_{7,10}$ , l'analyse des caractéristiques des débits d'étiage des 187 sites pour la variable d'étiage  $Q_{30,5}$  a montré que les moments non centraux d'ordre un à six suivaient une invariance d'échelle simple par rapport à la surface du bassin versant. L'analyse des résidus du sixième moment a permis de distinguer deux régions homogènes : la première contenant 117 sites et la seconde région qui en contient 70. L'étude des débits minimums des différentes régions a montré qu'elles aussi suivaient une similitude d'échelle simple. Nous avons ensuite appliqué la théorie des modèles de queue des distributions Weibull et Lognormale pour l'estimation régionale des débits d'étiage. L'application de la procédure présentée au chapitre 4 pour la détermination d'un seuil de censure  $u$  pour la variable  $Q_{30,5}$ , a donné des résultats différents, relativement aux distributions Lognormale et Weibull. En effet, l'analyse des résultats montre que pour la distribution Lognormale (figure B.6), la considération de l'échantillon au complet (seuil = 100%) offre la meilleure estimation. Le modèle régional Lognormal non censuré est plus performant que les modèles censurés. Cette observation est corroborée par la figure B.6(b) où l'erreur quadratique moyenne est minimum pour le percentile 100%. Par contre, l'analyse de la figure B.5, représentant la détermination du seuil de censure  $u$  avec la distribution Weibull, montre que le choix du percentile 50% est plus approprié, le modèle de queue régional de la distribution de Weibull censurée à 50% est plus performant que les autres modèles. Dans ce cas, l'estimation régionale avec la distribution de Weibull non censurée ne performe pas de manière satisfaisante, car l'erreur quadratique moyenne qui lui est associée figure B.5(b) est la plus

grande. On peut donc conclure que la détermination du seuil  $u$ , dépend du choix de la variable d'étiage à considérer.

Le problème de l'estimation régionale des débits d'étiage demeure ouvert, c'est à dire que des travaux restent encore à faire. Les résultats obtenus indiquent que cette approche possède un potentiel important. Plusieurs voies de recherche se présentent comme suite à ce travail, notamment en ce qui concerne le choix de la variable d'étiage à retenir. En effet, une étude du choix de la variable basée sur la durée de l'étiage paraît nécessaire. L'analyse du minimum annuel du débit moyen sur  $m$  jours consécutifs, où  $m$  peut représenter 7, 10, 30, 60, 90, etc., permet d'adopter comme débit critique un débit de  $m$  jours ayant une période de retour de  $T$  années. Pour les pays nordiques comme le Canada (et pour le Québec en particulier), il serait intéressant de tester l'applicabilité de cette méthodologie en considérant une approche saisonnière. En effet, dans ces régions, on peut distinguer des étiages d'hiver et des étiages d'été. Ces deux phénomènes sont très distincts. La saison d'hiver est caractérisée par des températures froides et un gel des cours d'eau. En été le niveau des rivières baisse dans certains cas considérablement, car l'apport en précipitations est moindre, réduisant ainsi le débit dans la rivière à son minimum. Une telle approche permettrait d'associer des types de modèles régionaux à des classes de rivières caractérisées par la présence soit d'un étiage d'été ou d'un étiage d'hiver. Le développement de nouveaux modèles de queue permettrait de tester et de déterminer si un gain substantiel est obtenu par rapport aux modèles déjà existant. Ce gain peut se manifester par une réduction de la variance des estimateurs des quantile et par un ajustement plus précis.

## 6. RÉFÉRENCES

---

Bernier, J. (1964) « La prévision statistique des bas débits », Association Internationale d'Hydrologie, No.63.

Benson, M.A. (1962a) « Evolution of methods for evaluating the occurrence of flood », USGS Water Supply, article 1580-A.

Benson, M.A. (1962b) « Factors influencing the occurrence of floods in a humid region of diverse terrain », USGS Water Supply, article 1580-B.

Bobée, B (1994) « Atelier sur l'estimation régionale des crues - méthodologie d'estimation régionale des débits de crues : application à la région Québec-Ontario », Comptes rendus 10<sup>e</sup> congrès régional de l'Est du Canada, de l'association Canadienne de la qualité de l'eau, Montréal, Québec, Canada, 15-17 novembre : pp.428-433.

Bobée, B et F.Ashkar (1988) « The generalised method of moments applied to the log-Pearson type3 distribution », Journal of Hydraulic Engineering, ASCE, Vol.114, No.8 : pp.899-909.

Bobée, B. et F.Ashkar (1991) « The gamma family and derived distributions applied in hydrology », Water Resources Publications, Littleton, Colo.

Bobée, B., Perreault, L. et F. Ashkar (1993) « Two kinds of moment ratio diagrams and their applications in hydrology », Stochastic Hydrology and Hydraulics, Vol.7 : pp.41-65.

Burn, D.H (1989) « Cluster analysis as applied to regional flood frequency », Journal of Water Resources Planning and Management, ASCE, Vol.115, No.5 : pp.567-582.

Cadavid, E.E (1988) « Hydraulic geometry of channel networks : tests of scaling invariance », M.S.Thesis, Department of Civil Engineering, University of Mississippi, Oxford.

Cavadias, G.S (1989) « Regional flood estimation by canonical correlation », Article présenté au 1989 Annual Conference of the Canadian Society for Civil Engineering, St-John's, Terre Neuve.

Condie, R., et G.A.Nix (1975) « Modelling of low-flow frequency distributions and parameter estimation », Comptes rendus du International Water Resources Symposium, Water for Arid Lands, Iran.

- Cunnane, C (1978) « Unbiased plotting positions - A review », *Journal of Hydrology*, No.37 : pp.205-222.
- Cunnane, C (1989) « Methods and merits of regional flood frequency analysis », *Journal of Hydrology*, Vol.100 : pp.269-290.
- Dalrymple, T. (1949) « Regional flood frequency », *Comptes rendus du 29<sup>th</sup> annual meeting of the Highway Research Board* : 22 pages.
- Dalrymple, T. (1960) « Flood frequency analysis », *Geological Survey Water Supply*, article 1543-A : 80 pages.
- Dawdy, D.R et V.K.Gupta (1995) « Multiscaling and skew separation in regional floods », *Water Resources Research*, Vol.31, No.11 : pp.2761-2767.
- Durrans, S.R., (1996) « Low-flow analysis with a conditional Weibull tail model », *Water Resources Research*, Vol.32, No.6 : pp.1749-1760.
- Durrans, S.R. et S.Tomic (1996) « Regionalisation of low-flow frequency estimates : an Alabama case study », *Water Resources Bulletin* 32 : pp.23-37.
- Furey, P.R et V.K.Gupta (1995) « Scaling of saturated soil depth using a wetness index », presentation at 1995 meeting, American Geophysical Union, San Francisco.
- Greenwood, J.A., Landwerh, J.M., Matalas, N.C et J.R.Wallis (1979) « Probability weighted moments : definition and relation to parameters of several distributions expressible in inverse form », *Water Resources Research*, Vol.15 : pp.1049-1054.
- GREHYS (1996a) « Presentation and review of some methods for regional flood frequency analysis », *Journal of Hydrology*, Vol.186 : pp.63-84.
- GREHYS (1996b) « Inter-comparaison of regional flood frequency procedures for Canadian rivers », *Journal of Hydrology*, Vol.186 : pp.85-103.
- Gumbel, E.J (1954) « Statistical theory of drought », *Comptes rendus du American Society of Civil Engineers*, Vol.80, Separate 439.
- Gupta, V.K et E.Waymire (1990) « Multiscaling properties of spatial rainfall and river flow distribution », *Journal of Geophysical Research*, Vol.95 (D3) : pp.1999-2009.
- Gupta, V.K, O.J.Mesa et D.R.Dawdy (1994) « Multiscaling theory of flood peaks :Regional quantile analysis », *Water Resources Research*, Vol.30, No.12 : pp.3405-3421.

Hoang, V.D et R.Tremblay (1976) « Estimation des débits d'étiage des rivières du Québec Meridional » Ministère des Richesses Naturelles, Direction générale des eaux, H.P. : 37 pages.

Hoang, V.D (1978) « Meilleur ajustement statistique aux débits journaliers de crue et d'étiage des rivières du Québec » Ministère des Richesses Naturelles, Direction générale des eaux, H.P. : 47 pages.

Hoerl, A.E et R.W.Kennard (1970) « Ridge regression : applications to nonorthogonal problems », *Technometrics*, Vol.12, No.1 : pp.96-82.

Hosking, J.R.M, J.R.Wallis et E.F.Wood (1985) « Estimation of the generalised extreme value distribution by the method of probability-weighted moments », *Technometrics*, Vol.27, No.3 : pp.251-261.

Johnson, N.L et Kotz, S (1970) « Continuous univariate distribution-1 », Wiley, New York.

Joseph, E.S (1970) « Probability distribution of annual drought », *Journal of the Irrigation and Drainage Division, ASCE*, Vol.96, IR4 : pp.461-471.

Loganathan, G.V., P.Mattejat, C.Y.Kuo, et M.H. Diskin (1986), « Frequency analysis on low-flows : Hypothetical Distribution methods and a physically based approach », *Nordic Hydrology*, Vol.17 : pp.129-150.

Matalas, N.C (1963) « Probability distribution of low-flows », *United States Geological Survey*, Article 434-A.

Miller, M., (1996) « Regionalization of low-flow estimates using quantile analysis », *Comptes rendus du 16<sup>th</sup> annual AGU Hydrology Days*, Fort Collins, Colorado : pp.317-328.

Nathan, R.J., et T.A.McMahon (1990) « Practical aspects of low-flow frequency analysis », *Water Resources Research*, Vol.26, no.9 : pp.2135-2141.

Neter, J, W.Wasserman et M.H.Kutner (1989) « Applied linear regression models », 2<sup>ème</sup> Edition, Irwin, Burr Ridge, Il.

Nguyen, V.T.V. et G.R.Pandey (1996) « A new approach to regional estimation of floods in Quebec », *Comptes rendus de la conférence nationale sur le développement rationnel et durable des ressources en eau*, Association Canadienne de la ressource hydrique, Québec : pp.587-596.

Ouarda, T.B.M.J., G.Boucher, P.F.Rasmussen et B.Bobée (1997) « Regionalisation of floods by canonical correlation analysis », Comptes rendus de European Water Resour.Ass. Conference, Operational Water Management, Balkema, Rotterdam : pp.297-302.

Ouarda, T.B.M.J., B.Bobée, J.Bernier et P.F.Rasmussen (1997) « Comment on ``Multiscaling and skew separation in regional floods`` by David R.Dawdy and Vijay K.Gupta », Water Resources Research Vol.33, no.1 : pp.271-272.

Ouarda, T.B.M.J., M.Lang, B.Bobée, J.Bernier et P.Bois (1999) « Synthèse de modèles régionaux d'estimation de crue utilisés en France et au Québec », Revue des Sciences de l'Eau, Vol.12, No.12 : pp.155-182.

Perreault, L., R.Roy, B.bobée et L.Mathier (1991) « Estimation régionale des débits maximums annuels  $X_T$  de période de retour T », INRS-Eau, Rapport scientifique No.305.

Perreault, L et B.Bobée (1992) « Loi de Weibull à deux paramètres » INRS-Eau, Rapport Scientifique No.351.

Pilon, P.J., (1990), « The weibull distribution applied to regional low-flow frequency analysis », Regionalization in Hydrology (Proceedings of the Ljubljana Symposium). IAHS Publ.No.191 : pp.227-237.

Potter, K.W. et D.P.Lettenmaier (1990) « A comparaison of regional flood frequency estimation methods using a resampling method », Water Resources Research, Vol.16, No.3 : pp.415-424.

Rasmussen, P.F (1994) « The index flood method for regional estimation of T-year events », INRS-Eau Internal Report, NSERC Strategic grant no.STR0118482, 15 pages.

Ribeiro, J et J.Rousselle (1996) « Robust simple scaling analysis of flood peaks series », Candian Journal of Civil Engineering, No.23 : pp.1139-1145.

Riggs, H.C., (1972) « Low-flows investigations », Livre 4, chap.B1, U.S.Geol.Sur., Tech. of Water Resour.Invest., Washington, D.C : pp.1-17.

Riggs, H.C., (1980) « Characteristics of low-flows », Journal of the Hydraulics Division, ASCE, Vol.106 : pp.717-731.

Rodriguez-Iturbe, I et J.B Valdés (1979) « The geomorphologic structure of hydrologic response », Water Resources Research, Vol.15, No.6 : pp.1409-1420.

Seber, G.A.F (1977) « Linear regression analysis », John Wiley & sons, New York.

- Smith, J.A (1989) « Representation of basin scale in flood peak distributions », *Water Resources Research*, Vol.28, No.11 : pp.2993-2999.
- Smith, J.A (1992) « Regional flood frequency analysis using extreme order statistics of the annual peak record », *Water Resources Research*, Vol.25 : pp.311-317.
- Sparth, H (1982) « Mathematical algorithms for linear regression », Academic Press, San Diego.
- Stedinger, J.R et G.D.Tasker (1985) « Regional hydrologic analysis ; ordinary, weighted, and generalised least squares compared », *Water Resources Research*, Vol.21, No.9 : pp.1421-1432
- Tasker, G.D (1987) « A comparison of methods for estimating low-flow characteristics of streams », *Water Resources Publications*, Vol.23, no.6, pp.1077-1038.
- Tomic, S (1994) « Regional low-flow frequency analysis for Alabama Rivers », M.S.Thesis, Department of Civil Engineering, the University of Alabama, Tuscaloosa
- Viessman, W, Jr., Lewis, G.I., Knapp, J.W., (1989) « Introduction to hydrology », 3<sup>ème</sup> edition, Harper&Row Publishers, New York : pp.678-680.
- Vogel, R.M et C.N.Kroll (1989) « Low-flow frequency analysis using probability-plot correlation coefficients », *Journal of Water Resources Planning and Management*, ASCE, Vol.115, No.3 : pp.338-357.
- Vogel, R.M et C.N.Kroll (1990) « Generalised low-flow frequency relationships for ungaged sites in Massachusetts », *Water Resources Bulletin*, American Water Resources.Association., Vol.26, No2.
- Wallis, J.R (1980) « Risk and uncertainties in the evaluation of flood events for the design of hydraulic structures », E.Guggino, G.Ross, et E.Todinni (eds), Peine e Siccita Fondazione Politecnica del Mediterraneo Gatanie, Erice, Italie : pp.3-36.
- Waylen, P.R et M.Woo (1987) « Annual low-flows generated by mixed processes », *Hydrological Sciences Journal*, Vol.32 : pp.371-383.
- Wiltshire, S.E (1985) « Grouping basins for regional flood frequency analysis », *Hydrological Sciences Journal*, Vol.30, No.1 : pp.151-159.



## 7. APPENDICES

---

### ANNEXE A

RÉSULTATS DE LA VARIABLE D'ÉTIAGE :

$Q_{7,10}$

**Tableau A1 Quelques caractéristiques physiographiques et hydrologiques des 187 sites, avec la correspondance code provinciale vs code fédérale.**

Numéro	Code prov	Code Féd	Altitude (m)	CoefVar	Surface (km2)	Taille
1	11507	01bd002	43	0,309	2760	24
2	11508	01bd008	157	0,377	558	13
3	11201	01bf001	15	0,336	1140	31
4	11003	01bg005	140	0,175	1680	17
5	10902	01bg008	27	0,232	1340	14
6	10802	01bg009	40	0,275	1910	12
7	20602	01bh005	18	0,362	645	26
8	10101	01bh007	50	0,403	534	21
9	20404	01bh010	104	0,398	659	15
10	43012	02jb013	274	0,132	2590	26
11	41301	02kg001	137	0,307	5150	64
12	41903	02kj003	320	0,204	2110	28
13	41902	02kj004	183	0,316	3760	30
14	42607	02kj007	320	0,876	2110	28
15	40129	02lc027	360	0,229	109	14
16	40204	02lc029	122	0,257	5460	29
17	40212	02lc043	372	0,486	39,9	27
18	40406	02ld005	162	0,417	1330	26
19	40830	02lg005	232	0,211	6840	20
20	40810	02lh002	180	0,347	1780	18
21	40814	02lh004	168	0,385	1290	24
22	50144	02nd003	243	0,166	2640	13
23	50135	02ne011	183	0,211	1570	30
24	50119	02nf003	366	0,301	1390	60
25	43202	02oa035	23	0,37	146000	36
26	30905	02oa054		0,407	2490	25

27	30907	02oa057	37	1,032	643	21
28	52219	02ob008	53	0,412	1340	26
29	52212	02ob017	152	0,409	1270	30
30	52233	02ob037	17	0,457	647	16
31	52601	02oc002	76	0,377	1030	62
32	52805	02oc004	183	0,46	774	30
33	52806	02oc011	287	1	52,8	22
34	30101	02od001	152	0,684	549	65
35	30103	02od003	64	0,477	1540	29
36	30215	02oe022	191	0,338	521	32
37	30234	02oe027	198	0,406	642	42
38	30282	02oe032	274	0,317	738	21
39	30309	02og007	122	0,553	153	24
40	30335	02og011	197	0,458	67,1	16
41	30304	02og019	32	0,745	1490	29
42	30316	02og026	12	0,604	342	26
43	30339	02og041	75	1,04	583	14
44	30340	02og042	80	0,199	271	12
45	30341	02og043	22	0,642	3310	12
46	30420	02oh008	534	0,758	387	14
47	30401	02oj007	24	0,331	22000	59
48	30415	02oj024	23	0,602	309	22
49	50304	02pa007	98	0,215	4480	28
50	50409	02pb006	162	0,176	642	30
51	50408	02pb019	130	0,196	1550	30
52	50701	02pc009	31	0,278	355	28
53	51001	02pd002	61	0,257	1100	43
54	51002	02pd012	732	0,201	3,94	30
55	51003	02pd013	564	0,24	9,17	29
56	51004	02pd014	762	0,249	1,22	28
57	51007	02pd015	732	0,305	3,57	24
58	51011	02pd018	780	0,501	0,69	14
59	51301	02pe009	15	0,156	865	28

60	51502	02pf001	31	0,16	1700	28
61	22507	02pg006	126	0,357	525	17
62	22704	02pg022	70	0,565	795	13
63	23106	02ph010	84	0,311	821	29
64	23303	02ph011	78	0,278	1160	15
65	23402	02pj005	114	0,36	5820	63
66	23401	02pj007	91	0,543	709	70
67	23429	02pj014	161	0,586	3070	16
68	23422	02pj030	207	0,565	691	27
69	24003	02pl005	180	0,499	919	26
70	24007	02pl007	35	0,449	2330	23
71	22301	02qa001	38	0,378	932	72
72	22003	02qa002	24	0,306	1610	33
73	21915	02qa017	78	0,429	492	21
74	21601	02qb001	18	0,29	1650	68
75	21702	02qb005	31	0,424	208	29
76	21502	02qb011	15	0,245	721	29
77	20802	02qc001	40	0,309	1200	39
78	21407	02qc009	43	0,215	774	22
79	20502	02qc014	12	0,33	59,8	14
80	62209	02rb004	241	0,167	3600	16
81	61801	02rc011	104	0,159	1090	20
82	62101	02rd002	116	0,155	9320	34
83	62102	02rd003	117	0,174	9870	33
84	61901	02rf001	111	0,137	15300	33
85	61905	02rf002	305	0,138	11100	33
86	61502	02rg005	220	0,271	2280	18
87	61022	02rh027	555	0,267	495	22
88	61020	02rh035	360	0,189	1110	24
89	62701	02rh045	99	0,217	746	19
90	62802	02rh047	18	0,265	1100	17
91	60601	02rh048	155	0,215	572	19
92	60101	02rh049	43	0,208	736	21

93	70201	02sc001	31	0,361	793	19
94	70401	02sc002	66	0,249	2010	21
95	70202	02sc003	115	0,544	40,7	11
96	71401	02ua003	3	0,246	1570	21
97	72301	02uc002	31	0,278	19000	29
98	72302	02uc003	503	0,289	3390	18
99	72502	02uc005	122	0,269	458	11
100	73303	02va001	31	0,307	684	34
101	73502	02vb003	91	0,289	7590	15
102	73503	02vb004	210	0,26	7230	17
103	73801	02vc001	15	0,211	13000	38
104	73802	02vc002		0,304	6550	12
105	74601	02wa001	15	0,329	2060	27
106	74701	02wa002	15	0,276	5590	21
107	74702	02wa003	15	0,531	151	10
108	74901	02wb001	31	0,185	16000	11
109	74902	02wb002	152	0,238	11600	18
110	74903	02wb003	43	0,262	15600	14
111	75601	02wc001	76	0,352	2950	20
112	75701	02xa006	61	0,235	19100	13
113	75702	02xa007	275	0,304	12100	13
114	76101	02xb001	61	0,355	5750	16
115	76201	02xb002	72	0,273	1060	12
116	76601	02xc001	31	0,425	6630	25
117	80702	03ab001	244	0,253	22500	14
118	80718	03ab002	290	0,155	31900	27
119	80704	03ab003		0,188	18700	21
120	80703	03ac001	305	0,192	2010	38
121	80705	03ac002	335	0,169	8310	28
122	80707	03ac004	250	0,177	22200	32
123	80717	03ac007	305	0,144	8310	19
124	80701	03ad001	244	0,194	57500	23
125	81007	03bb002	366	0,162	18100	25

126	81008	03bb003	258	0,205	4490	10
127	81001	03bc001	232	0,139	40900	11
128	81002	03bc002	220	0,151	40900	32
129	80809	03bd002	253	0,143	9820	20
130	80801	03be001	209	0,194	17100	22
131	81101	03bf001	67	0,285	6090	20
132	90610	03ca001	366	0,297	11600	18
133	90602	03cb001	213	0,173	27700	21
134	90605	03cb003	244	0,166	1870	19
135	90613	03cb004	305	0,197	21400	16
136	90601	03cc001	34	0,151	44300	22
137	90608	03cd002	244	0,264	3700	17
138	90607	03ce001	244	0,105	3390	20
139	92722	03da002	472	0,126	12800	21
140	92706	03db003	122	0,139	18600	17
141	92723	03db004	175	0,277	9320	14
142	92703	03dc002	244	0,172	37000	22
143	92717	03dc003	335	0,189	10200	17
144	92715	03dd002	366	0,179	13200	31
145	92708	03dd003	244	0,148	19100	22
146	92711	03dd005	335	0,117	13800	17
147	92707	03de001	152	0,138	10900	19
148	92709	03de002	213	0,233	1860	14
149	92704	03df001	31	0,166	96600	20
150	93806	03ea001		0,172	21000	27
151	93808	03ea004	412	0,217	7720	12
152	93804	03ec001	152	0,231	4660	33
153	93801	03ed001	152	0,174	36300	29
154	93803	03ed002	31	0,129	43200	11
155	93805	03ed004	183	0,244	7280	27
156	93302	03ee003	56	0,167	4370	13
157	95002	03fa002		0,173	12500	18
158	95003	03fa003		0,133	8390	21

<b>159</b>	94502	03fb001		0,219	3830	17
<b>160</b>	94701	03fb002	91	0,347	1410	16
<b>161</b>	94201	03fc001	244	0,212	1390	17
<b>162</b>	94203	03fc003	274	0,151	11700	15
<b>163</b>	94204	03fc005	274	0,224	1220	16
<b>164</b>	94207	03fc007	335	0,311	10400	29
<b>165</b>	94206	03fc008	305	0,139	10400	31
<b>166</b>	102001	03ha001	31	0,203	26900	22
<b>167</b>	102004	03ha009	55	0,482	4120	18
<b>168</b>	102018	03ha012	205	0,255	3760	15
<b>169</b>	102010	03hb004	93	0,355	10200	17
<b>170</b>	102704	03ja001	207	0,336	3860	10
<b>171</b>	102701	03jb001	15	0,429	41700	25
<b>172</b>	103603	03ka001	122	0,366	8830	25
<b>173</b>	103605	03kc004	27	0,289	42700	29
<b>174</b>	103714	03la001	533	0,295	23600	10
<b>175</b>	103717	03la002	671	0,108	20300	10
<b>176</b>	103715	03ld004	232	0,207	8990	22
<b>177</b>	103704	03ld006	77	0,415	15900	11
<b>178</b>	103703	03le002	156	0,125	50800	19
<b>179</b>	103702	03lf002	61	0,478	48500	33
<b>180</b>	103801	03mb001	119	0,495	2140	18
<b>181</b>	104001	03mb002	152	0,363	29800	33
<b>182</b>	104401	03mc001	153	0,614	3680	21
<b>183</b>	104803	03md001	282	0,379	24200	18
<b>184</b>	104801	03me002	153	0,305	35200	15
<b>185</b>	80101	03na001	305	0,211	3680	63
<b>186</b>	80102	03na002	305	0,289	3680	20
<b>187</b>	80104	03nb001	232	0,468	11200	20

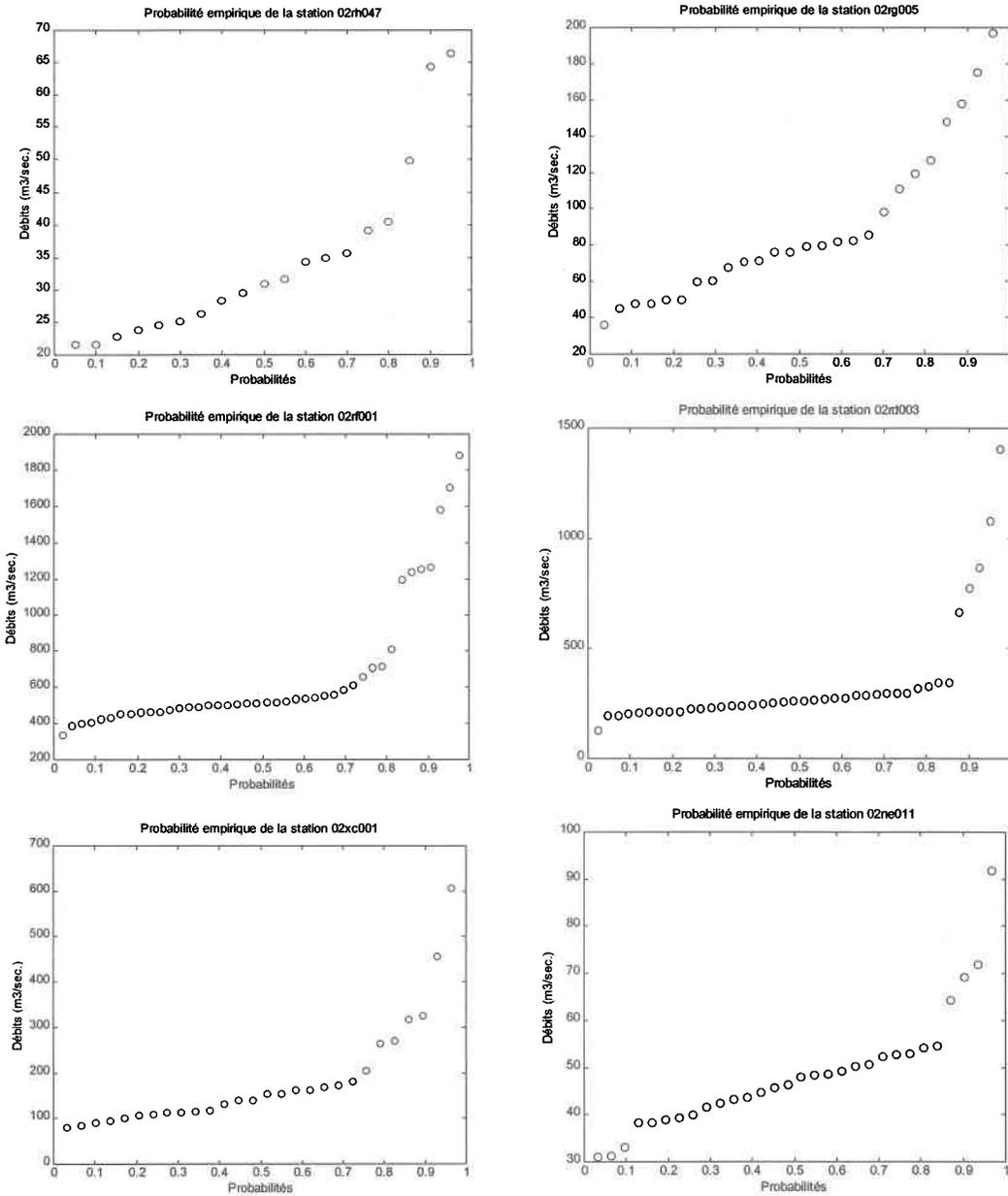
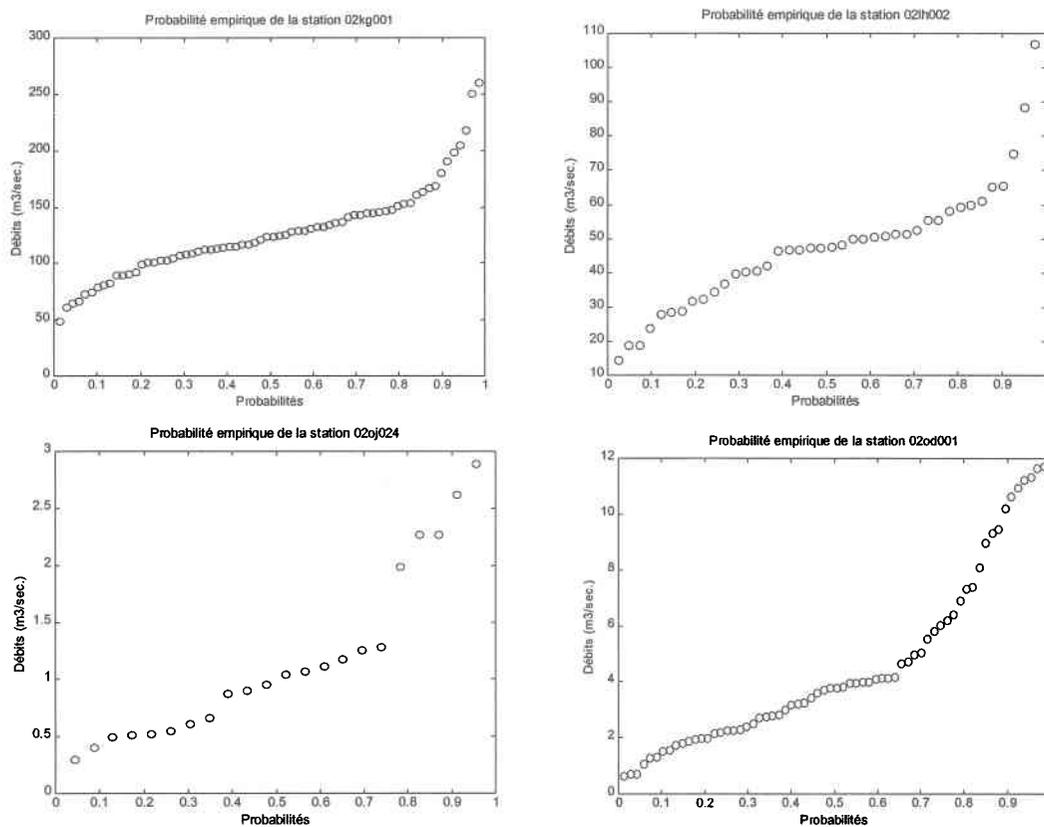
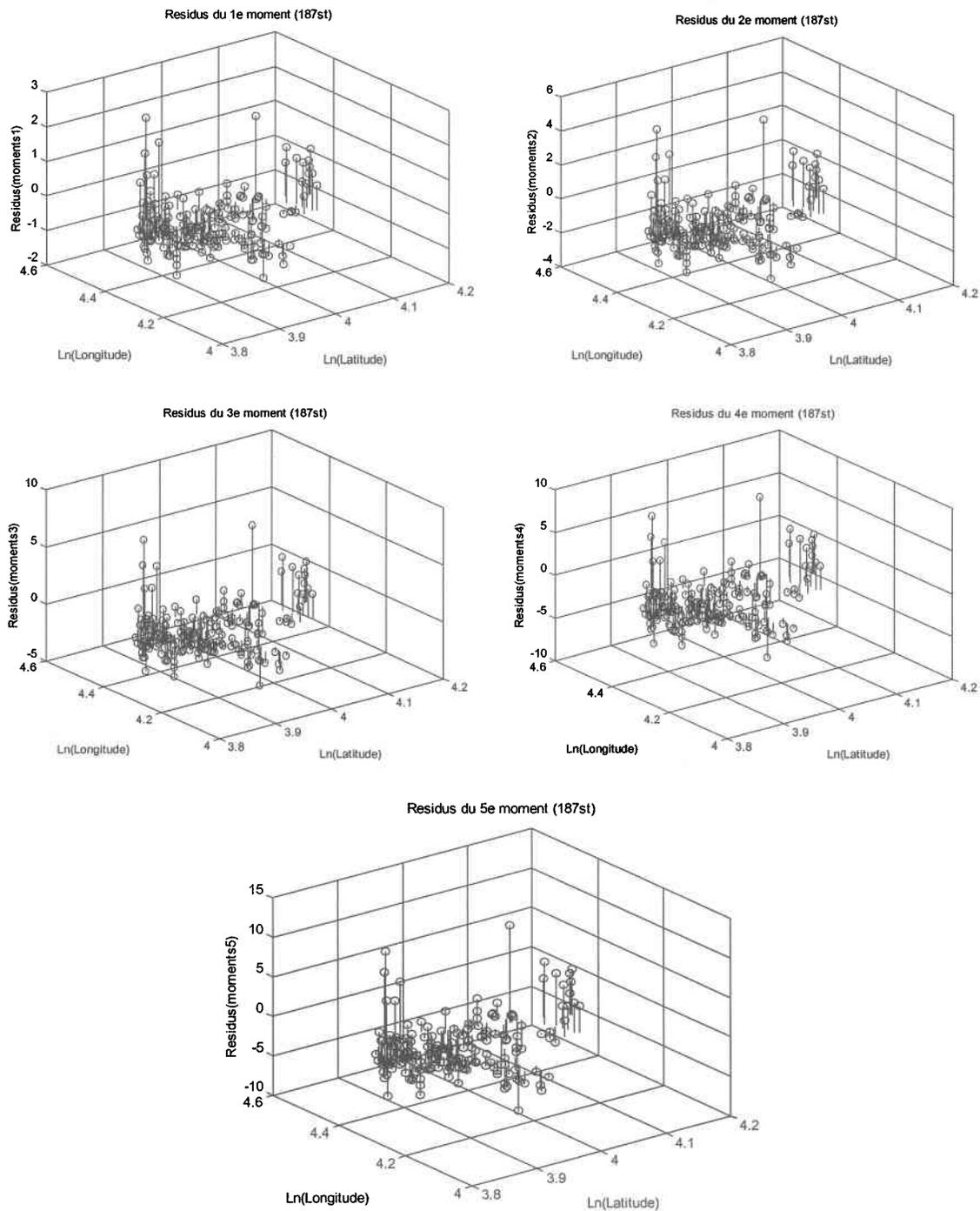


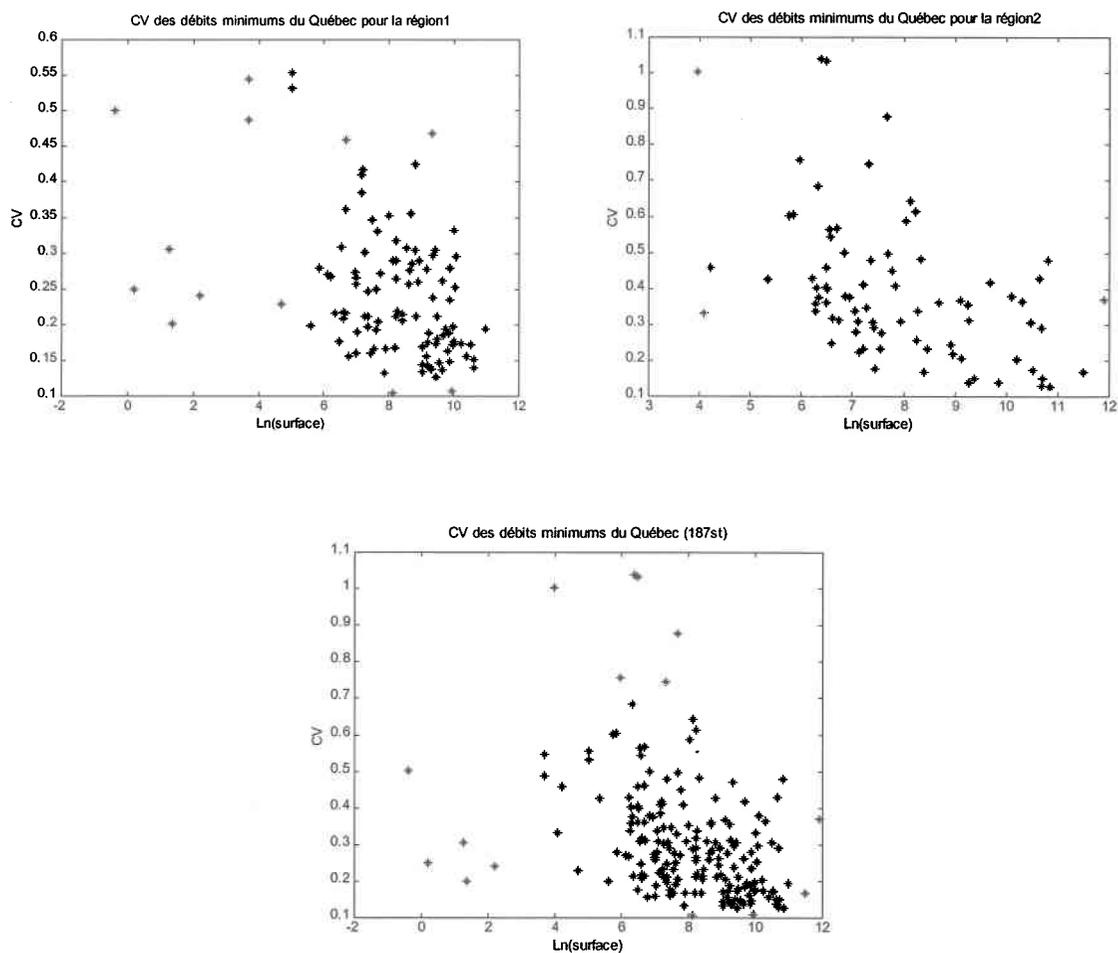
Figure A1 Débits annuels  $Q_7$  de quelques sites de la Province de Québec.



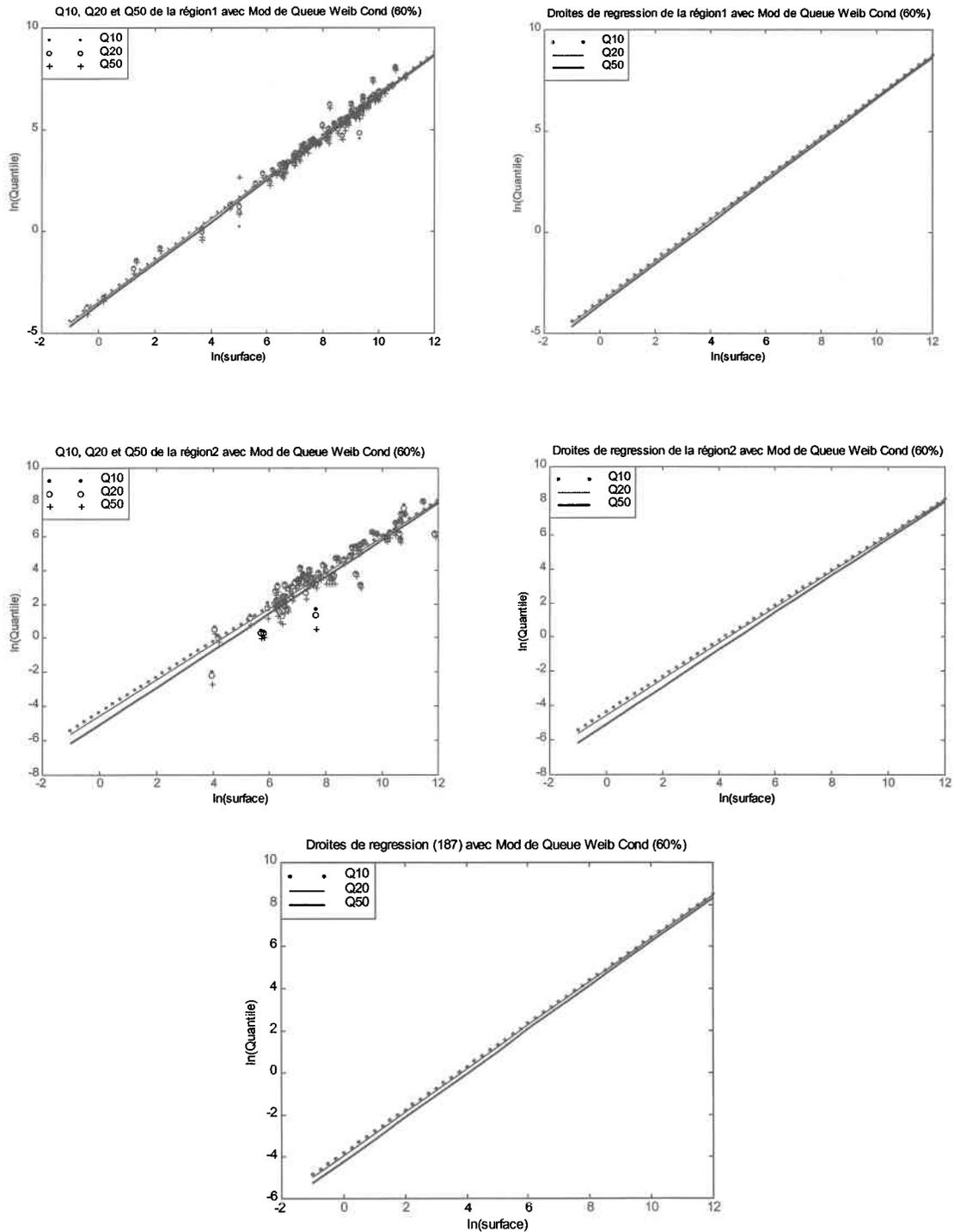
**Figure A2** Débits annuels  $Q_7$  de quelques sites de la Province de Québec (suite).



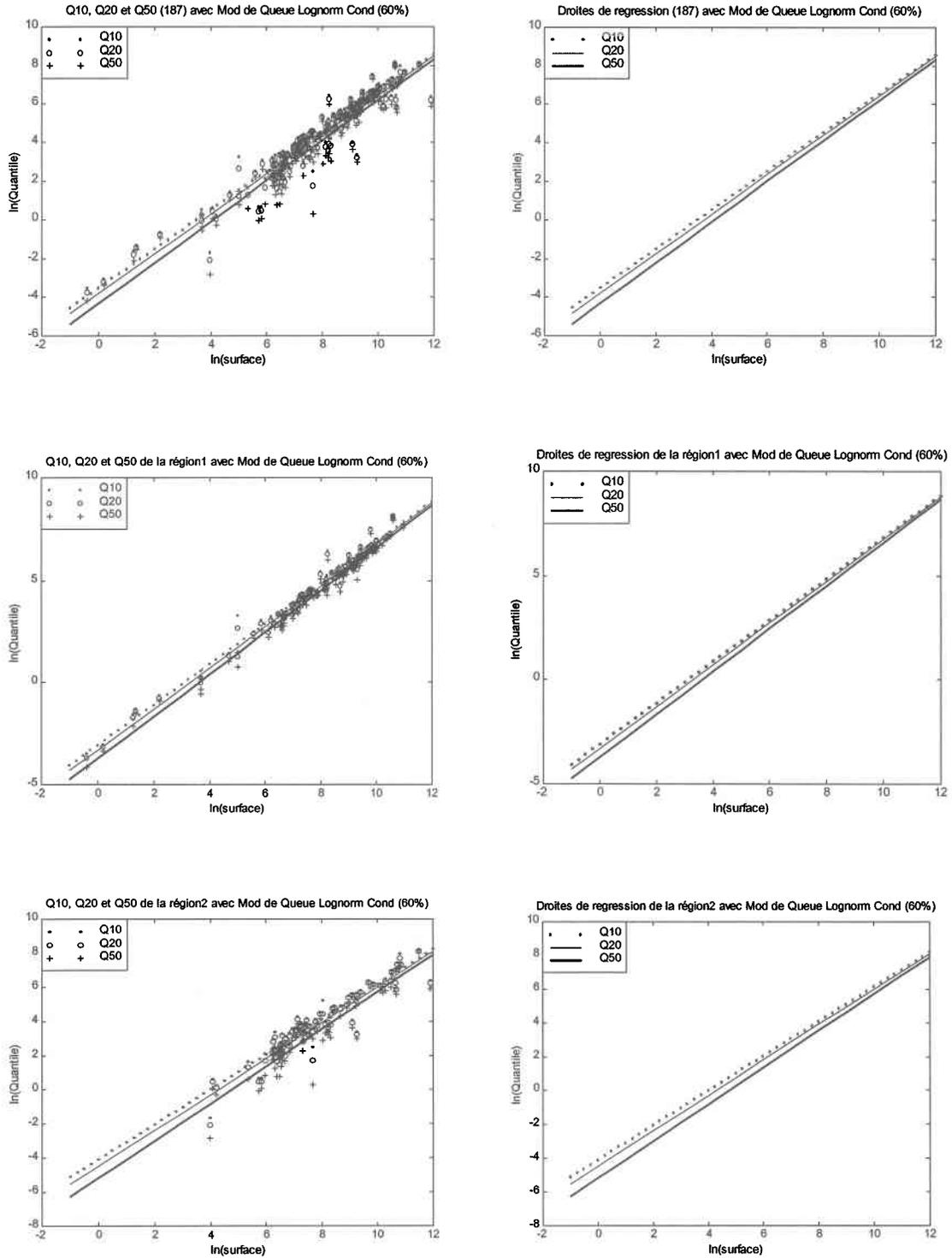
**Figure A3 Répartition des résidus des moments 1, 2, 3, 4, 5 sur un plan géographique.**



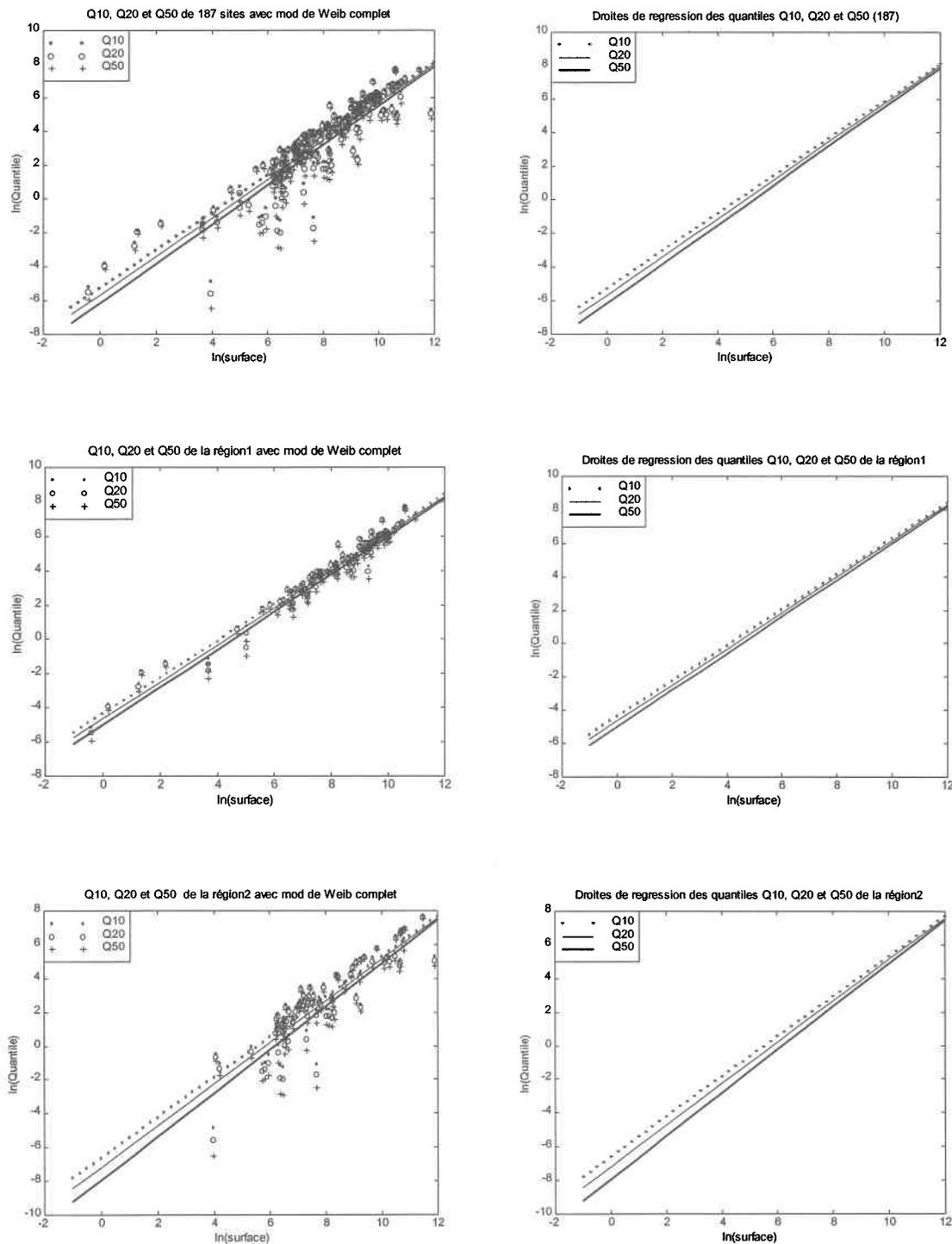
**Figure A4** Répartition du coefficient de variation des différentes régions considérées pour les 187 sites.



**Figure A5** Droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$  pour les différentes régions considérées (Weibull censurée à 60%).



**Figure A6** Droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$  pour les différentes régions considérées (Lognormale censurée à 60%).

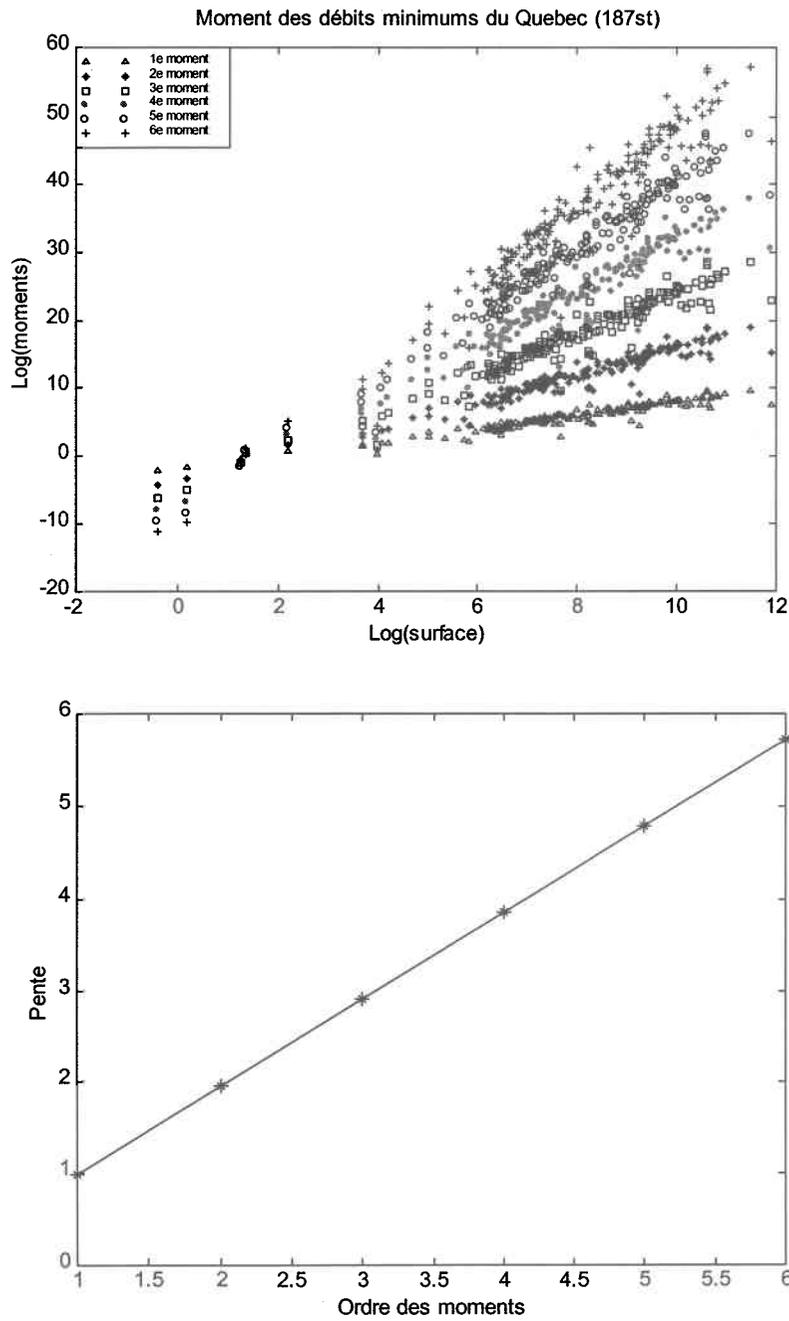


**Figure A7** Droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$  pour les différentes régions considérées (Weibull complète).

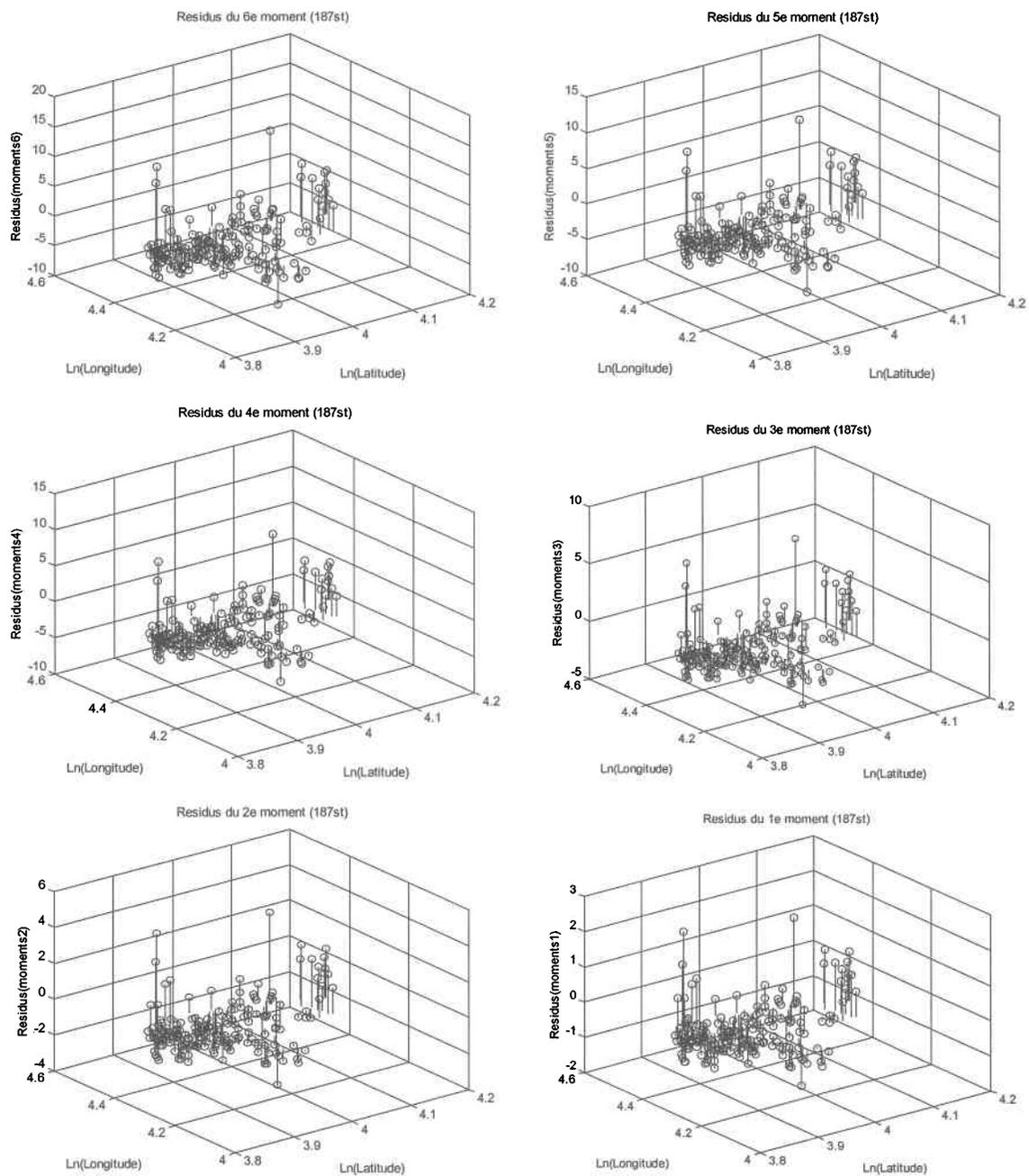
## **ANNEXE B**

**RÉSULTATS DE LA VARIABLE D'ÉTIAGE :**

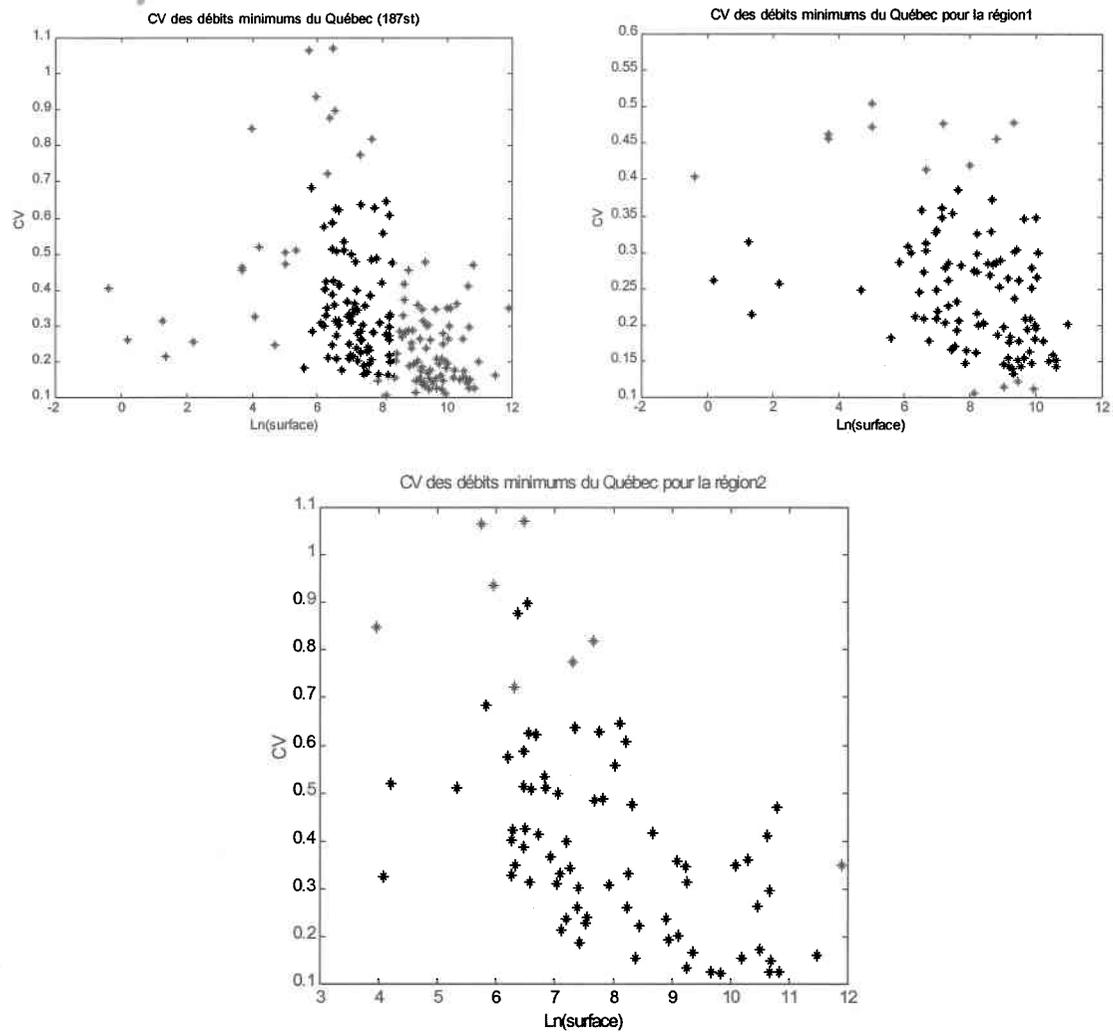
**Q<sub>30,5</sub>**



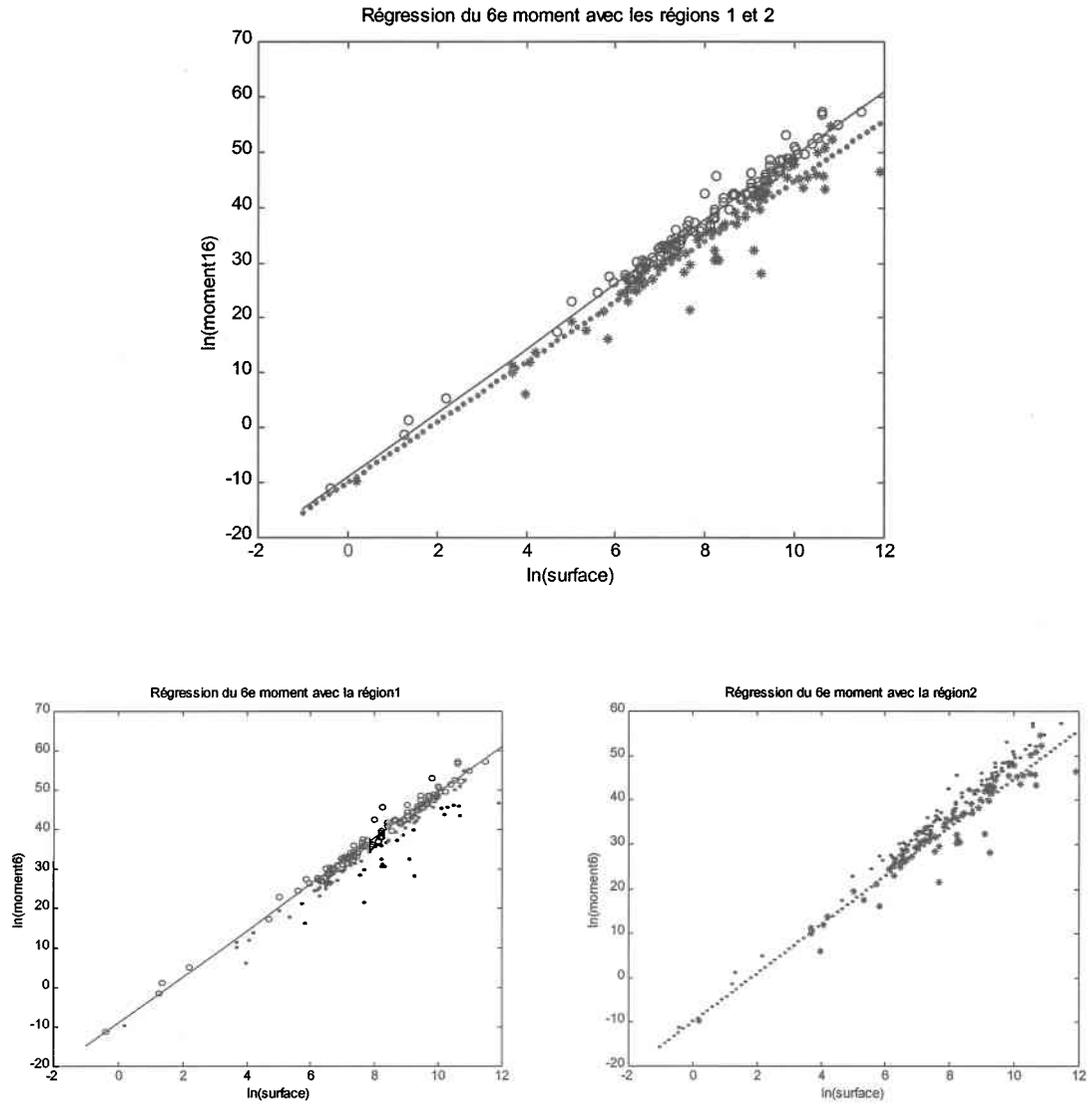
**Figure B1** Caractéristiques de l'invariance d'échelle simple en considérant les 187 sites comme échantillon homogène.



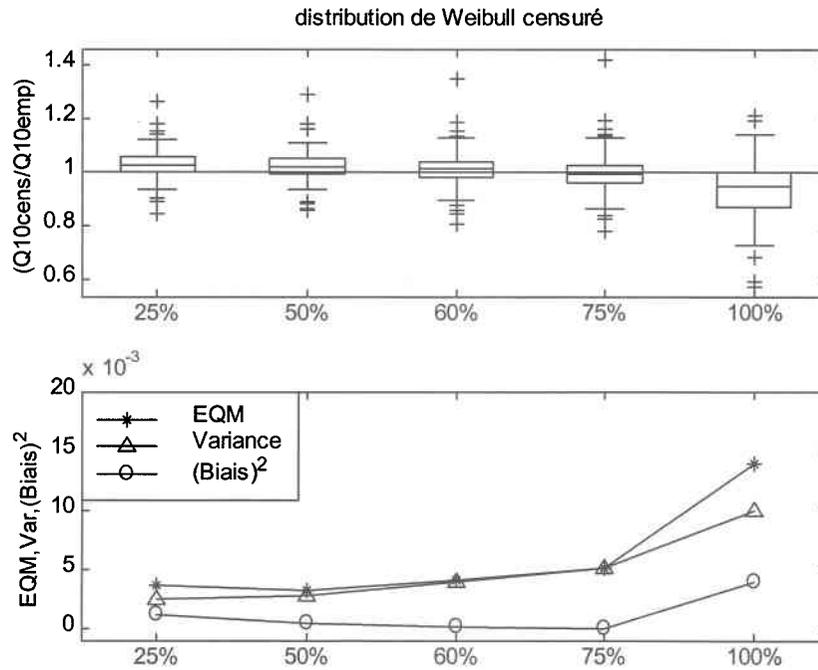
**Figure B2 Répartition des résidus des moments 1, 2, 3, 4, 5 et 6 sur un plan géographique.**



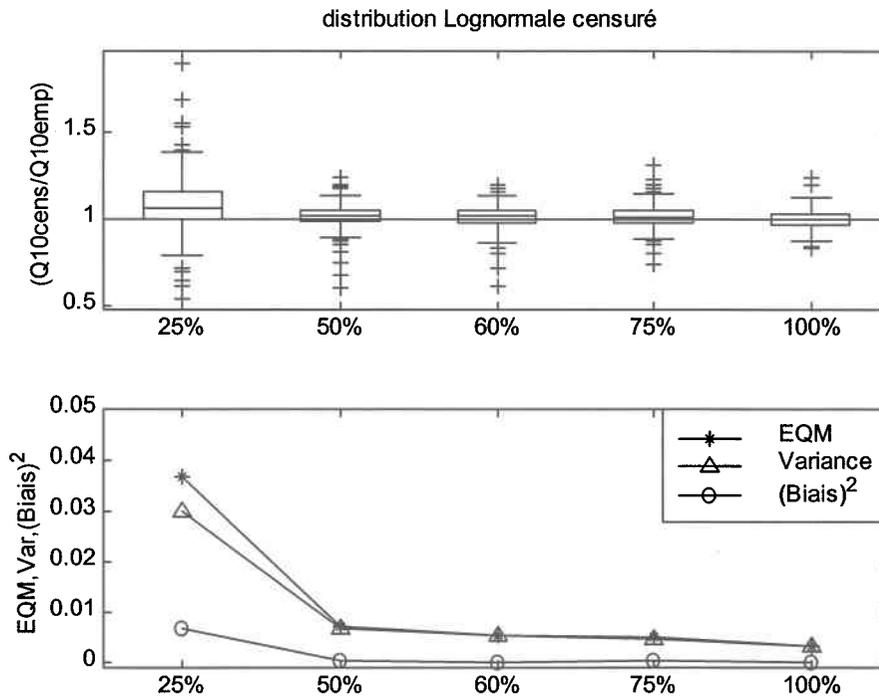
**Figure B3** Répartition du coefficient de variation des différentes régions considérées pour les 187 sites.



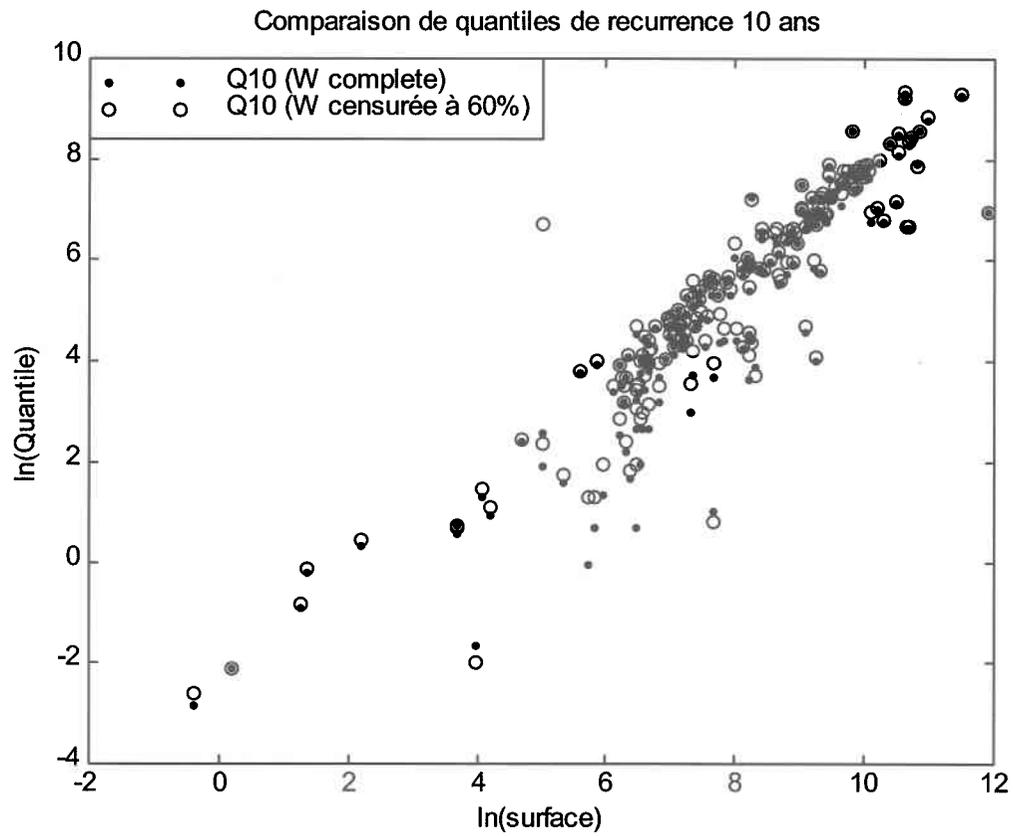
**Figure B4** Distinction entre les différentes régions sur une base régressive.



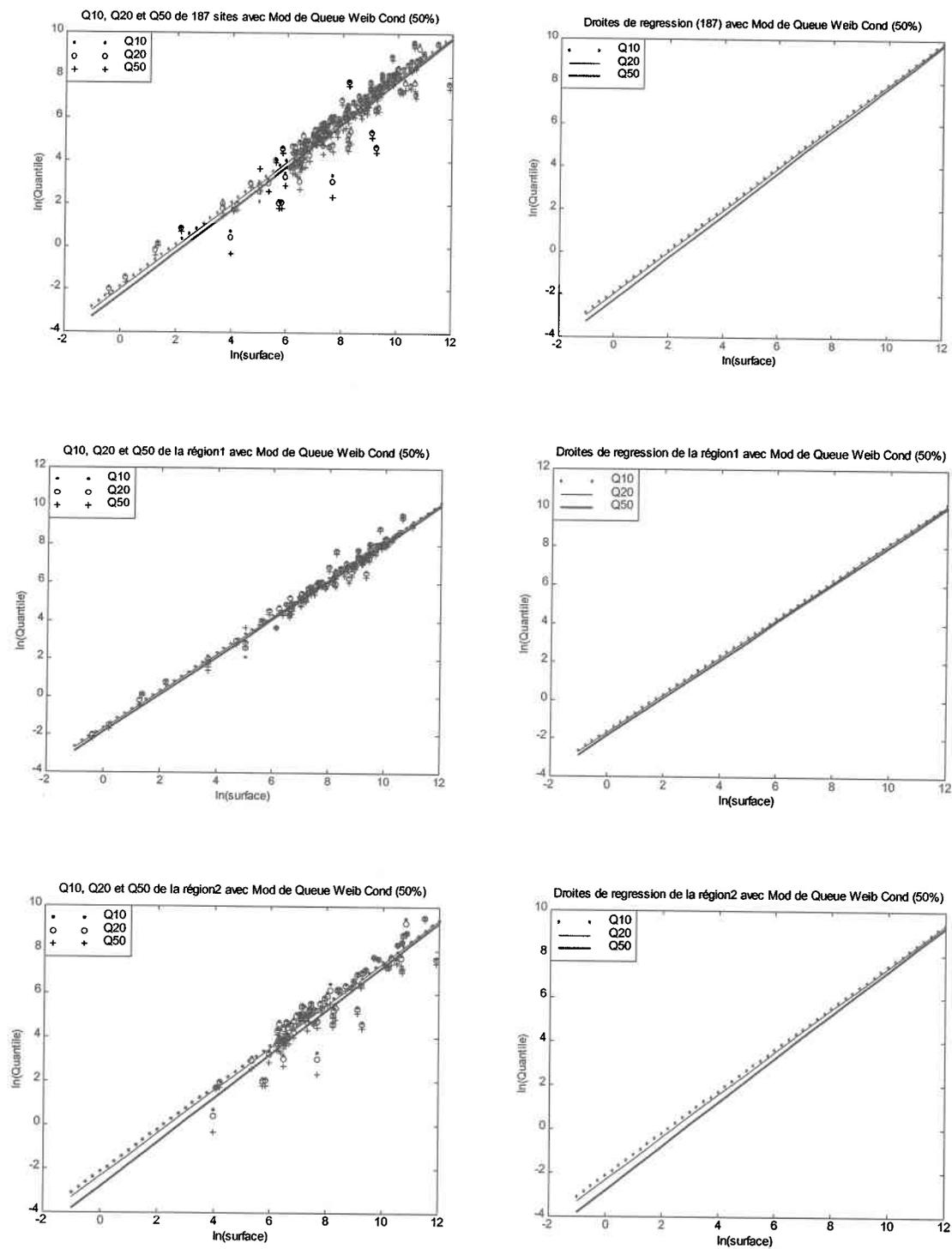
**Figure B5 Estimation du seuil de censure u avec la distribution de Weibull.**



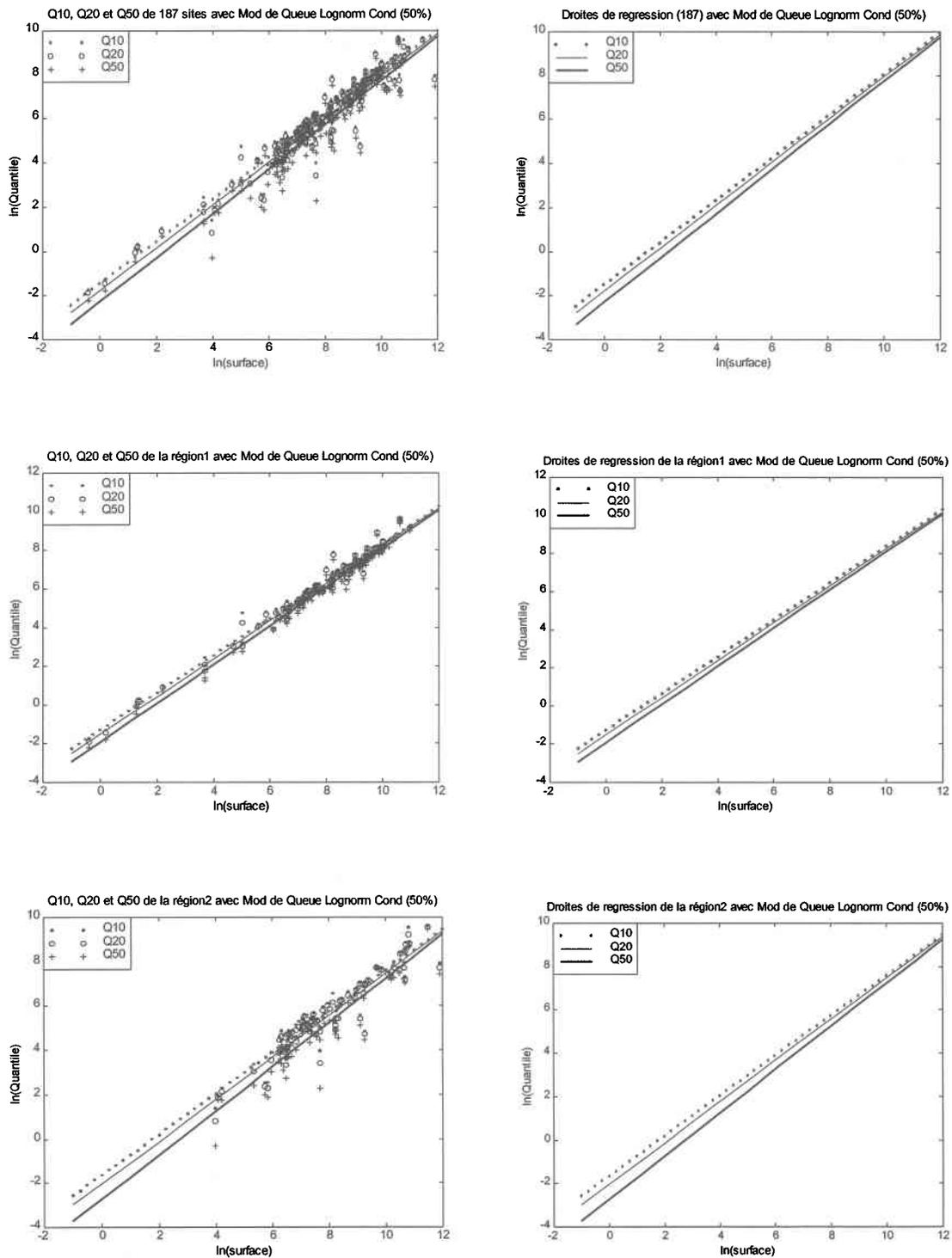
**Figure B6 Estimation du seuil de censure u avec la distribution de Lognormale.**



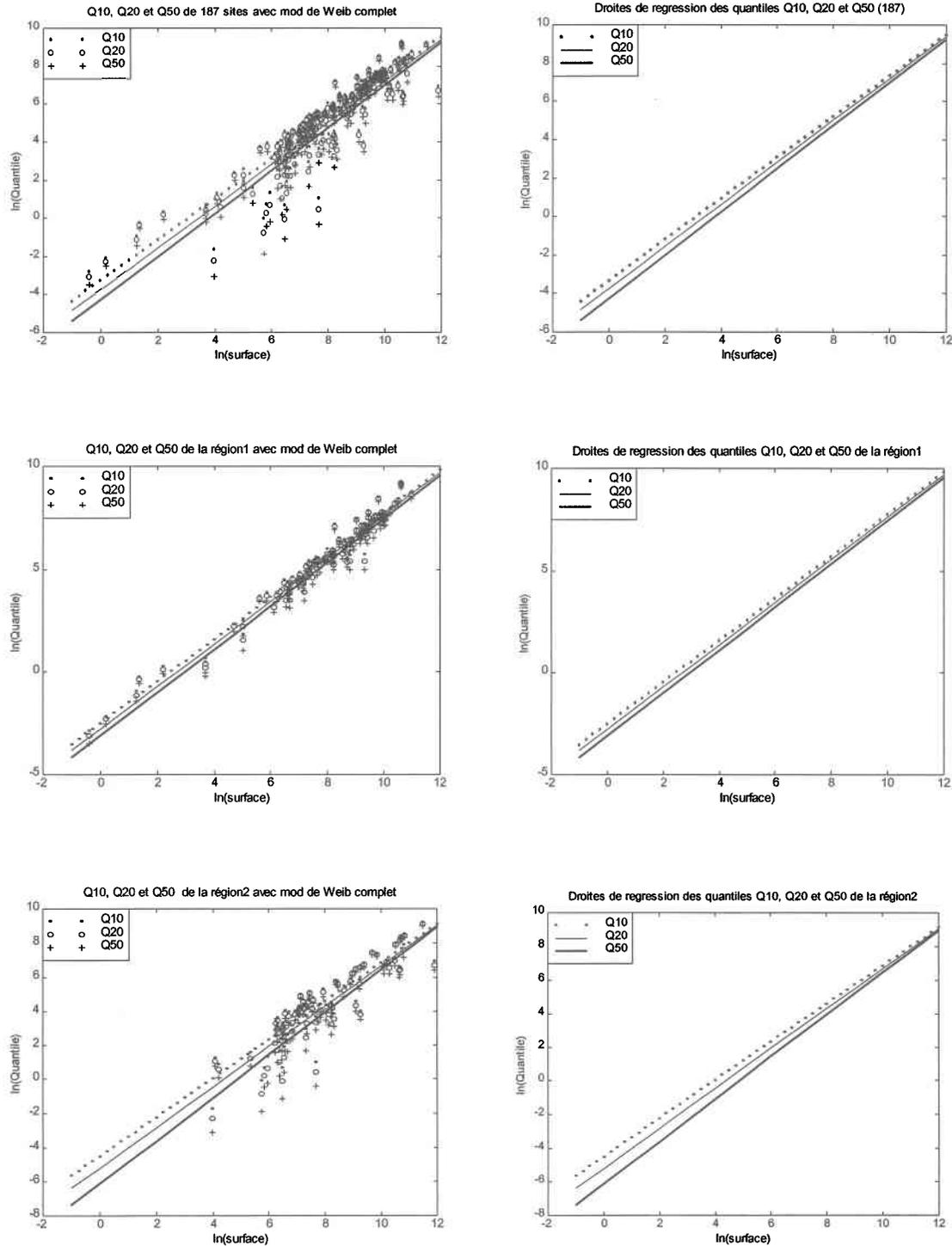
**Figure B7** Comparaison entre  $Q_{10}$  estimé par le modèle de queue de Weibull et  $Q_{10}$  estimé par le modèle de Weibull complet.



**Figure B8** Droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$  pour les différentes régions considérées (Weibull censurée à 50%).



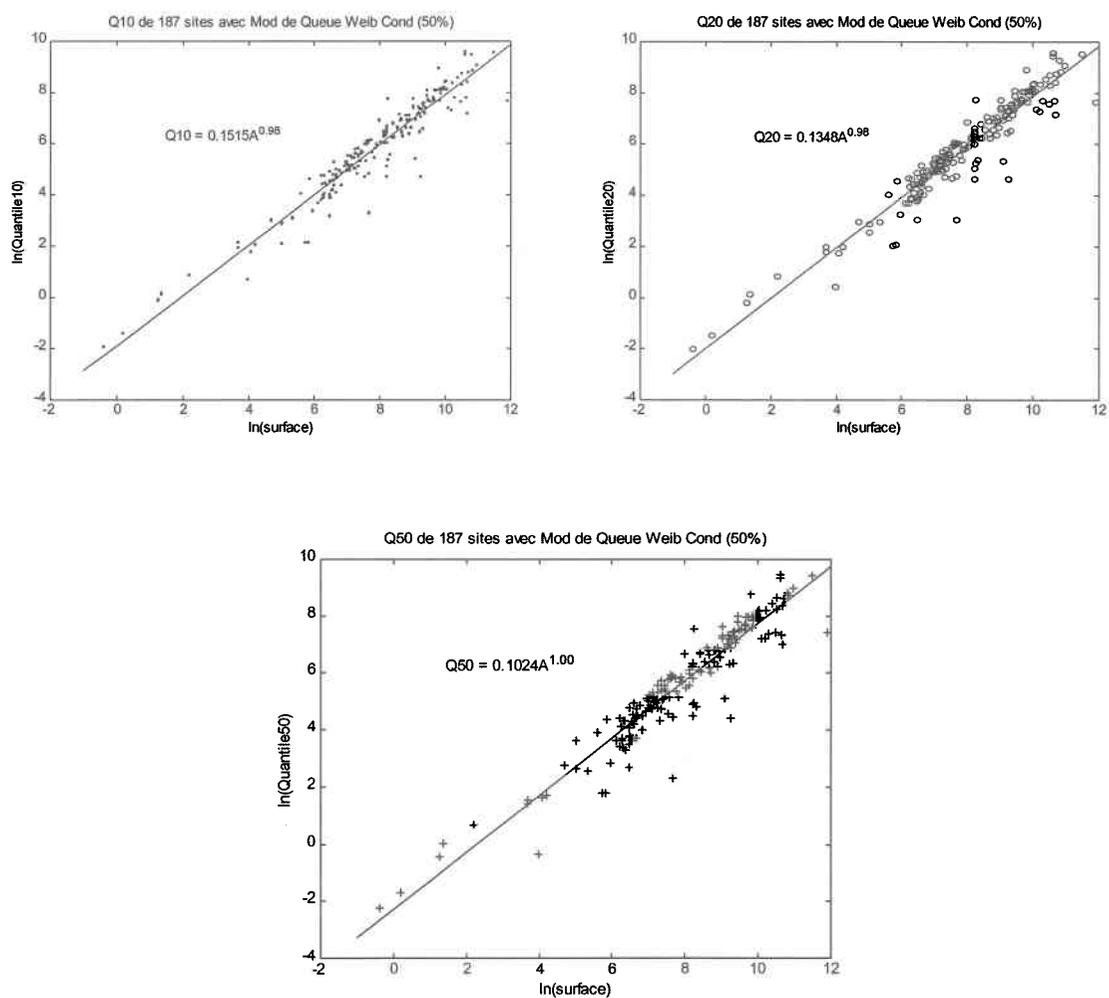
**Figure B9** Droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$  pour les différentes régions considérées (Lognormale censurée à 50%).



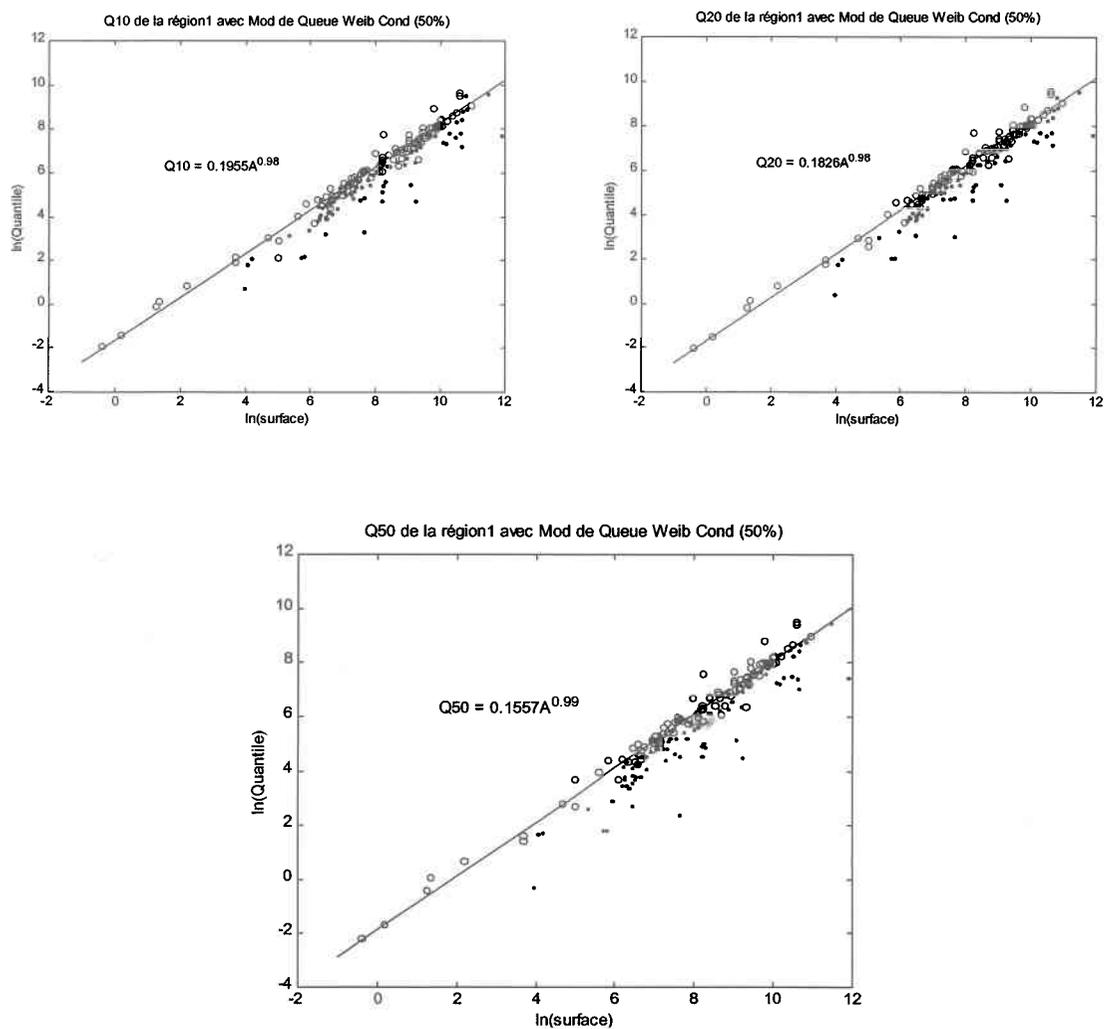
**Figure B10** Droites de régression des quantiles  $Q_{10}$ ,  $Q_{20}$  et  $Q_{50}$  pour les différentes régions considérées (Weibull complète).

**Modèle de queue conditionnel de la distribution de Weibull censurée à 50%**

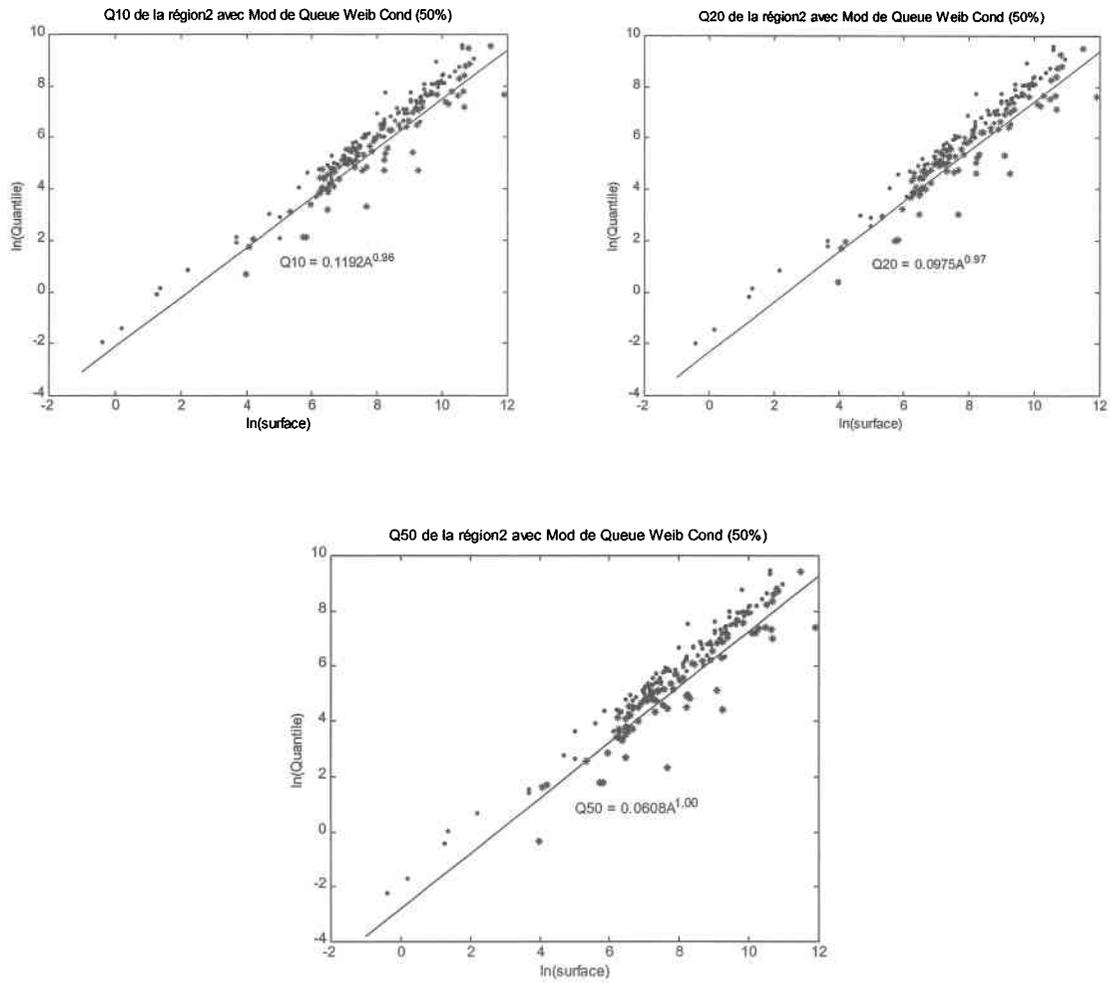
i) Toute la région est considérée comme homogène



**Figure B11** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  (187 sites) avec Weibull censurée à 50%.

ii) Région1

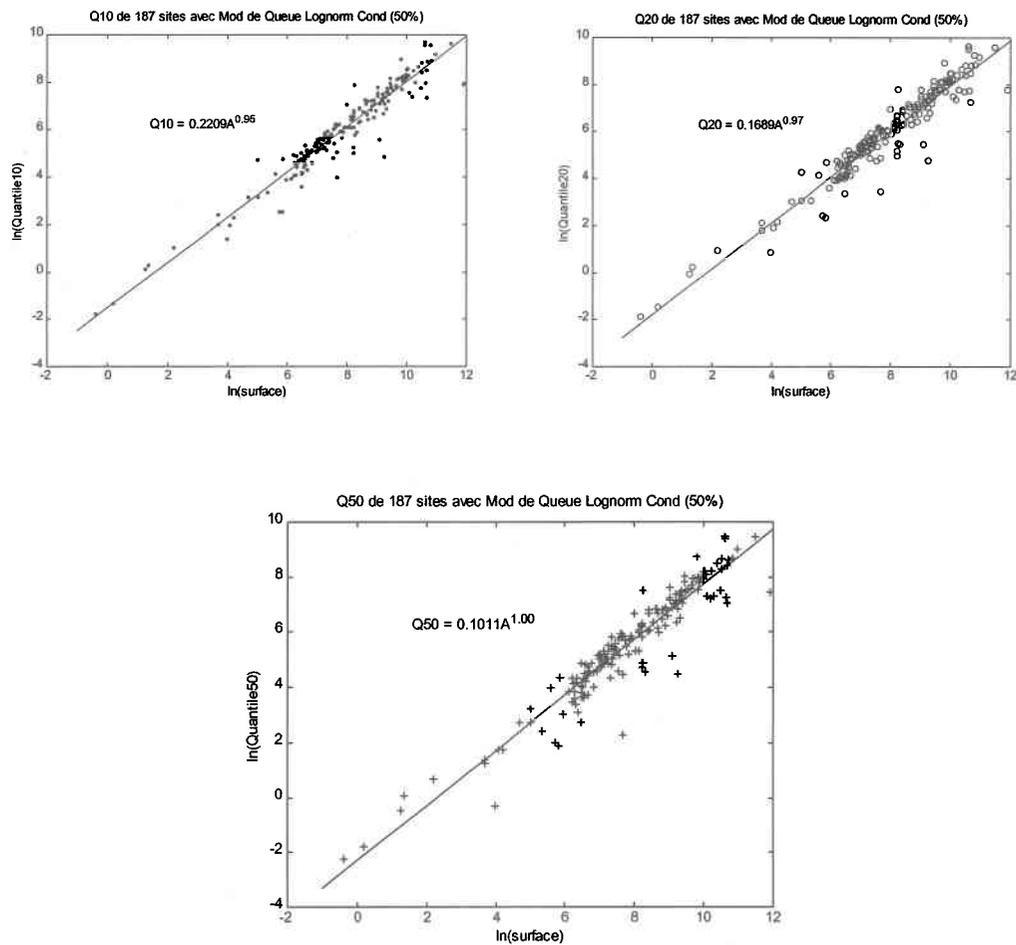
**Figure B12** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 1 (117 sites) avec Weibull censurée à 50%.

iii) Région2

**Figure B13** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 2 (70 sites) avec Weibull censurée à 50%.

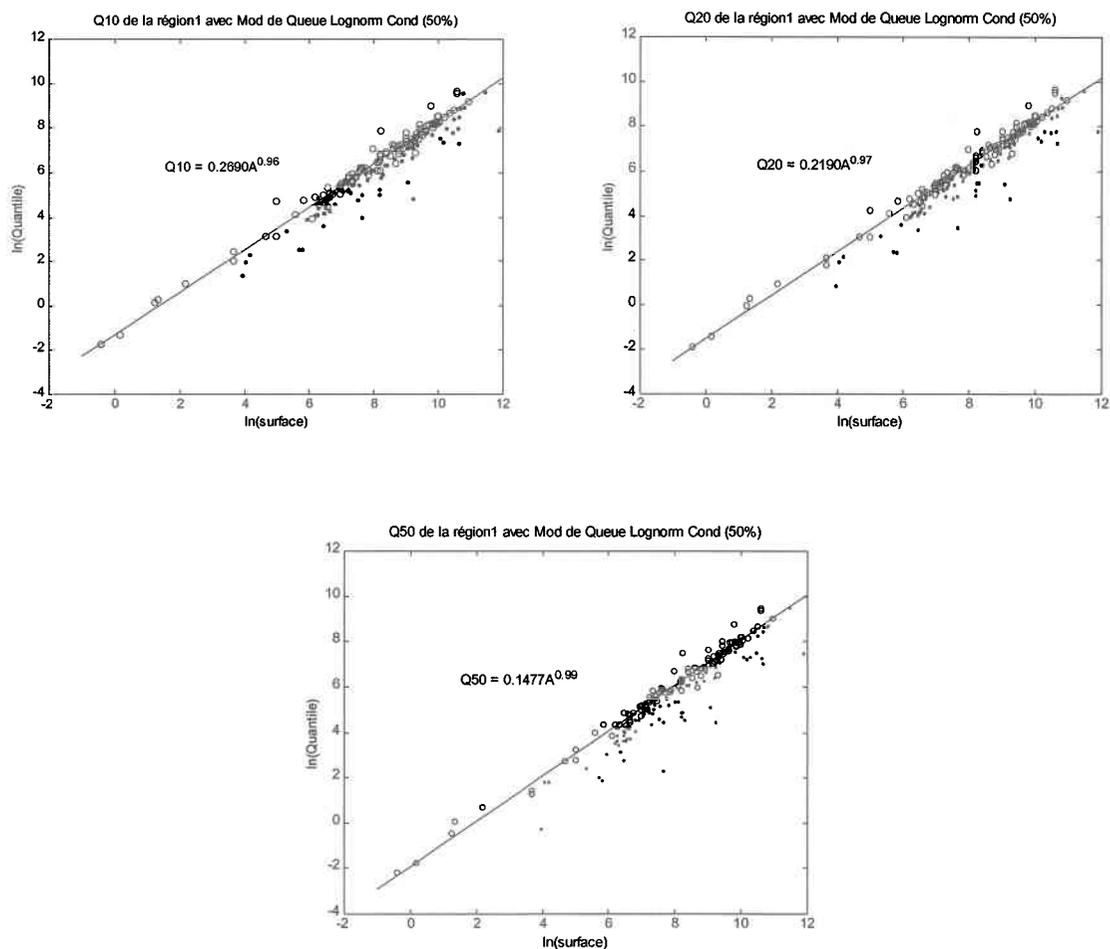
### Modèle de queue conditionnel de la distribution Lognormale censurée à 50%

i) Toute la région est considérée comme homogène



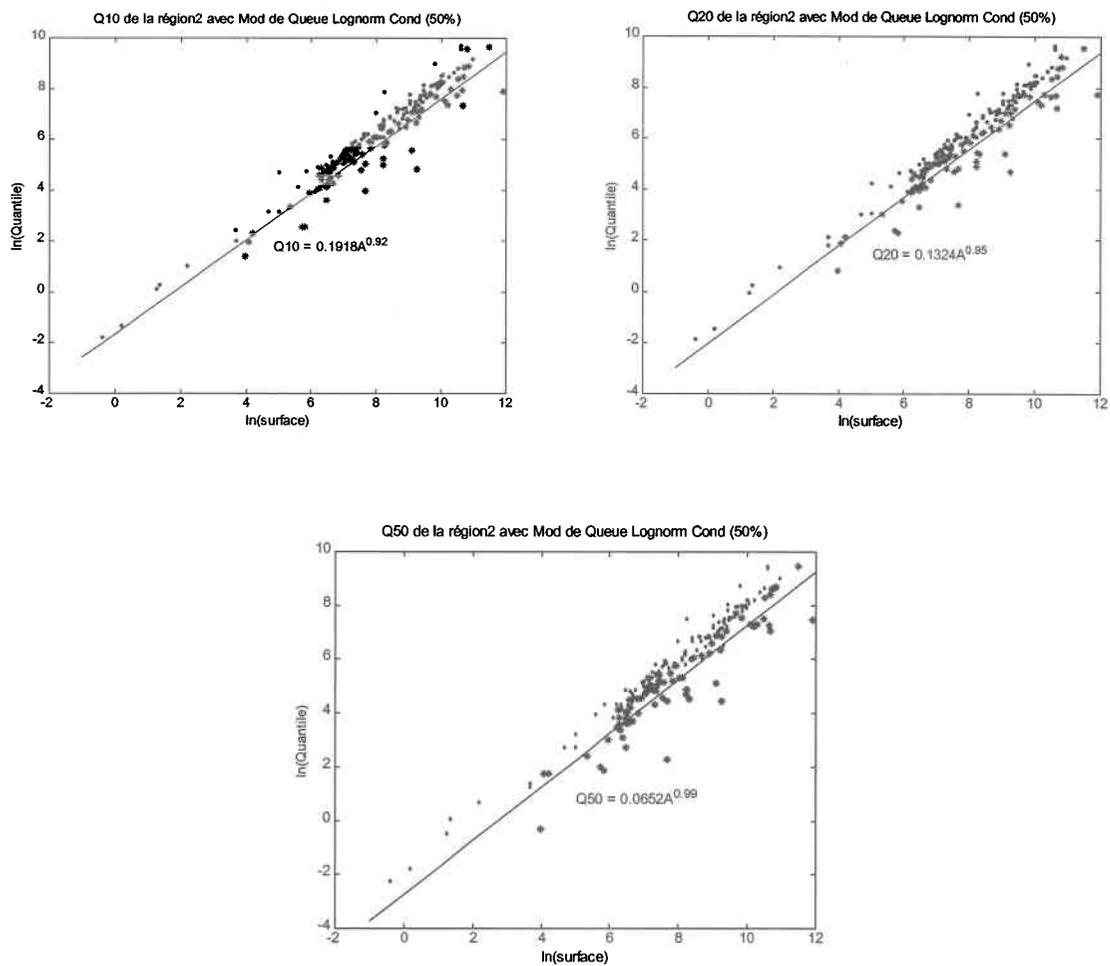
**Figure B14** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  (187 sites) avec Lognormale censurée à 50%.

## ii) Région1



**Figure B15** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 1 (117 sites) avec Lognormale censurée à 50%.

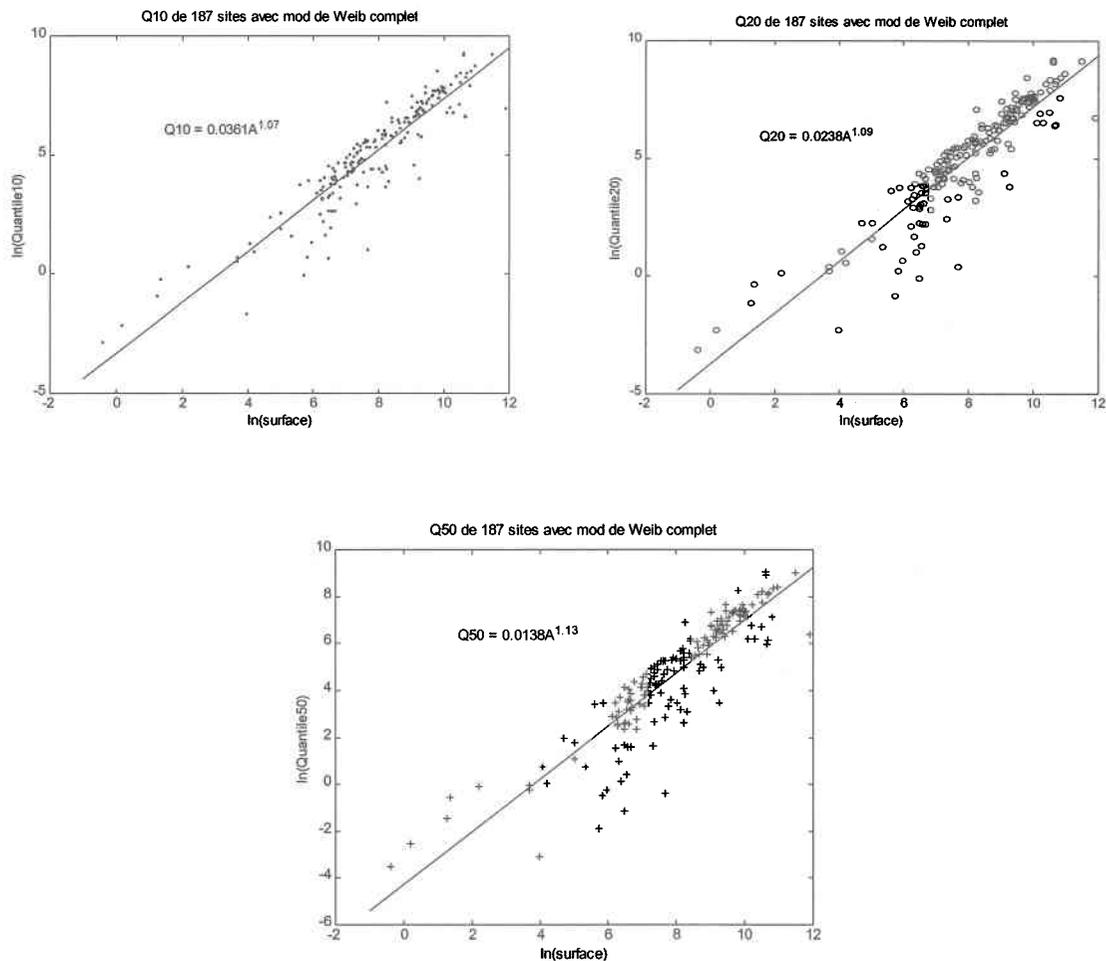
## iii) Région2



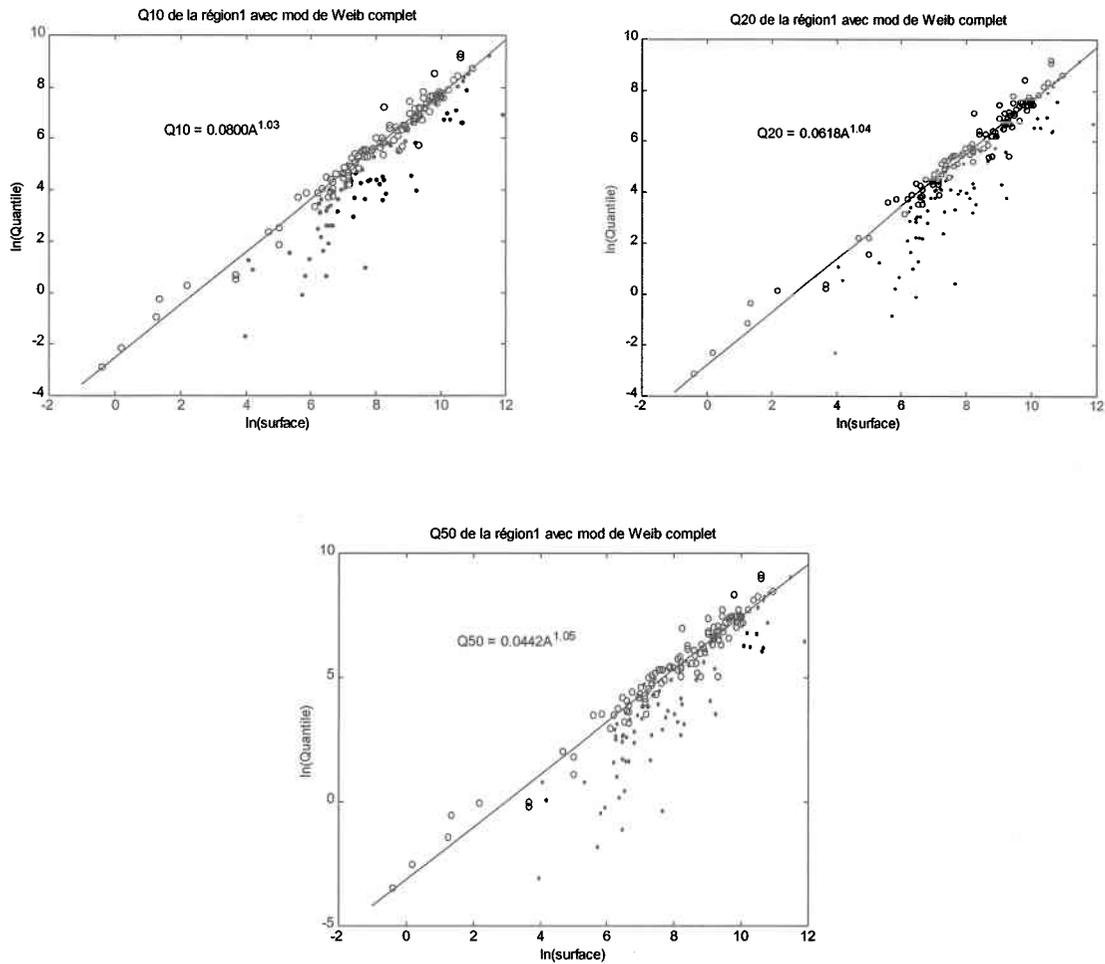
**Figure B16** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 2 (70 sites) avec Lognormale censurée à 50%.

**Modèle de la distribution de Weibull complète**

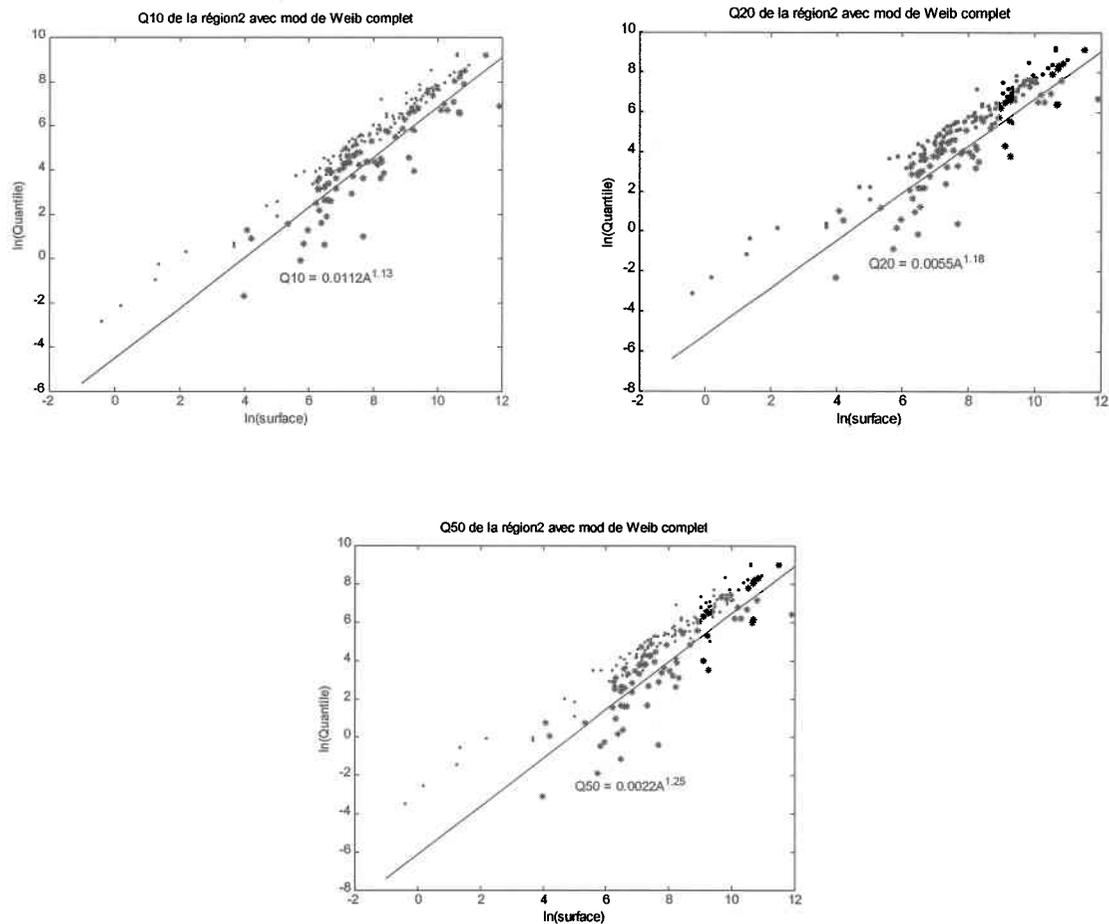
i) Toute la région est considérée comme homogène



**Figure B17** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  (187 sites) avec Weibull complète.

ii) Région1

**Figure B18** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 1 (117 sites) avec Weibull complète.

iii) Région2

**Figure B19** Equations régionales  $Q_{10}$ ,  $Q_{20}$ ,  $Q_{50}$  pour la région 2 (79 sites) avec Weibull complète.

Dans ce qui suit nous présentons le résumé des équations régionales obtenues en considérant toute la province comme échantillon homogène, et aussi en considérant les différentes régions séparément.

**Modèle de queue Weibull censuré à 50%**

<b>Toute la province</b>
Equations
$Q_{10} = 0.1515A^{0.98}$
$Q_{20} = 0.1348A^{0.98}$
$Q_{50} = 0.1024A^{1.00}$

**Modèle de queue Lognormal censuré à 50%**

<b>Toute la province</b>
Equations
$Q_{10} = 0.2209A^{0.95}$
$Q_{20} = 0.1689A^{0.97}$
$Q_{50} = 0.1011A^{1.00}$

**Modèle Weibull complet**

<b>Toute la province</b>
Equations
$Q_{10} = 0.0361A^{1.07}$
$Q_{20} = 0.0238A^{1.09}$
$Q_{50} = 0.0138A^{1.13}$

**Tableau B1 Sommaire des équations de régressions pour toute la province.**

**Modèle de queue Weibull censuré à 50%**

<b>Région 1</b>
Equations
$Q_{10} = 0.1955A^{0.98}$
$Q_{20} = 0.1826A^{0.98}$
$Q_{50} = 0.1557A^{0.99}$

**Modèle de queue Lognormal censuré à 50%**

<b>Région 1</b>
Equations
$Q_{10} = 0.2690A^{0.96}$
$Q_{20} = 0.2190A^{0.97}$
$Q_{50} = 0.1477A^{0.99}$

**Modèle Weibull complet**

<b>Région 1</b>
Equations
$Q_{10} = 0.0800A^{1.03}$
$Q_{20} = 0.0618A^{1.04}$
$Q_{50} = 0.0442A^{1.05}$

**Tableau B2 Sommaire des équations de régressions pour région 1.**

**Modèle de queue Weibull censuré à 50%**

<b>Région 2</b>
Equations
$Q_{10} = 0.1192A^{0.96}$
$Q_{20} = 0.0975A^{0.97}$
$Q_{50} = 0.0608A^{1.00}$

**Modèle de queue Lognormal censuré à 50%**

<b>Région 2</b>
Equations
$Q_{10} = 0.1918A^{0.92}$
$Q_{20} = 0.1324A^{0.95}$
$Q_{50} = 0.0652A^{0.99}$

**Modèle Weibull complet**

<b>Région 2</b>
Equations
$Q_{10} = 0.0112A^{1.13}$
$Q_{20} = 0.0055A^{1.28}$
$Q_{50} = 0.0022A^{1.25}$

**Tableau B3 Sommaire des équations de régressions pour la région 2.**

## **ANNEXE C**

### **APPROCHE FRÉQUENTIELLE**

Différentes méthodes d'estimation des paramètres d'une distribution  
statistique.

Exemples d'estimation régionale (approches classiques)

# ANNEXE C

---

## C.1 Approche fréquentielle

L'objectif de cette section est de décrire les principales caractéristiques des distributions statistiques et des méthodes d'estimation des paramètres. Nous insisterons ensuite sur la distribution Weibull ainsi que la distribution Lognormale à deux paramètres.

### C.1.1 Définition d'une distribution statistique

Il est important de définir en terme mathématique ce qu'est une distribution statistique. Soit une variable aléatoire continue  $X$ . La fonction  $f(x)$  est une fonction de densité de probabilité (f.d.p) si :

$$1/ f(x) \geq 0 \text{ pour } x \text{ réel} \quad (\text{C.1})$$

$$2/ \int_{-\infty}^{+\infty} f(x) = 1 \quad (\text{C.2})$$

$$3/ P[a \leq x \leq b] = \int_a^b f(x)dx \text{ pour } a \text{ et } b \text{ réel} \quad (\text{C.3})$$

La fonction de distribution  $F(x)$  est définie par :

$$F(x) = P[X \leq x] \quad \text{pour } x \text{ réel ; est la probabilité au non dépassement}$$

Elle est reliée à la fonction de densité de probabilité par la relation :

$$F(x) = \int_{-\infty}^x f(t)dt \quad (\text{C.4})$$

### C.1.2 Moments d'une distribution

La forme d'une distribution est déterminée par la valeur de ses paramètres. Elle est également caractérisée par ses moments. En calculant les moments empiriques d'un échantillon et en les comparant avec les moments théoriques de diverses distributions, il est possible de sélectionner à priori certaines distributions pour représenter l'échantillon et d'en écarter d'autres. Le moment non-centré d'ordre  $r$  d'une distribution  $F(x)$  est défini par :

$$\mu_r = E[X^r] = \int_{-\infty}^{\infty} X^r f(x)dx \quad (\text{C.5})$$

$$\mu_1 = E[X] \quad \text{est la moyenne}$$

et le moment centré par rapport à la moyenne  $\mu_1$  s'écrit :

$$\mu_r = E[(X - \mu_1)^r] = \int_{-\infty}^{\infty} (X - \mu_1)^r f(x)dx \quad (\text{C.6})$$

$$\mu_2 = E[(X - \mu_1)^2] \quad \text{est la variance} \quad (\text{C.7})$$

### C.1.3 Moments d'un échantillon

Soit un échantillon de valeurs observées  $x_1, x_2, \dots, x_n$ . Le moment non-centré et le moment centré par rapport à la moyenne d'ordre  $r$  de l'échantillon sont définis respectivement par :

$$m'_r = \frac{1}{n} \sum_{i=1}^n x_i^r \quad (\text{C.8})$$

$$m_r = \frac{1}{n} \sum_{i=1}^n (x_i - m'_1)^r \quad (\text{C.9})$$

Particulièrement,  $m'_1$  est la moyenne de l'échantillon (souvent notée  $\bar{x}$ ) et  $m_2$  est la variance biaisée de l'échantillon. La variance non biaisée est obtenue par :

$$s^2 = \frac{n}{n-1} m_2 \quad (\text{C.10})$$

### C.1.4 Méthodes d'estimation des paramètres

Il existe plusieurs méthodes pour estimer les paramètres d'une distribution statistique. Tout d'abord, nous verrons les méthodes bien connues et largement utilisées du maximum de vraisemblance et des moments. Ensuite, nous présenterons de manière succincte la méthode des moments pondérées et la méthode généralisée des moments.

#### C.1.4.1 Méthode du maximum de vraisemblance (MV)

La méthode du maximum de vraisemblance consiste pour un échantillon donné, à maximiser la fonction de vraisemblance (fonction de densité jointe) par rapport aux paramètres. Soit une distribution  $F(X, \theta_1, \theta_2, \dots, \theta_p)$ . Lorsqu'ils existent, les estimateurs

obtenus par la méthode du maximum de vraisemblance sont les solutions  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_p$  du système de p équations :

$$\frac{\partial \mathcal{L}(\theta_1, \theta_2, \dots, \theta_p)}{\partial \theta_r} = 0 \quad r = 1, 2, \dots, p \quad (\text{C.11})$$

La fonction de vraisemblance est définie par :

$$L(\theta_1, \theta_2, \dots, \theta_p) = \prod_{i=1}^n f(x_i, \theta_1, \theta_2, \dots, \theta_p) \quad (\text{C.12})$$

Il est souvent plus simple de maximiser le logarithme naturel de la fonction de vraisemblance que la vraisemblance elle-même. L'une ou l'autre des méthodes conduit au même maximum car la fonction logarithme est une fonction monotone croissante.

#### C.1.4.2 Méthode des moments (MM)

La méthode des moments appliquée à une loi à p paramètres, consiste à évaluer p moments théoriques indépendants de la distribution aux p moments correspondants de l'échantillon. Cette méthode est généralement utilisée parce qu'elle est très simple à appliquer. Elle permet de fournir des estimations pour les paramètres lorsque la méthode du maximum de vraisemblance ne converge pas. De plus, la méthode des moments peut-être utilisée pour initialiser les paramètres dans le cadre du processus itératif du maximum de vraisemblance. Soit un échantillon  $x_1, x_2, \dots, x_n$  tiré d'une distribution  $F(X, \theta_1, \theta_2, \dots, \theta_p)$ . Les estimateurs obtenus par la méthode des moments sont les solutions  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_p$  du système de p équations :

$$\mu_r = m'_r \quad r = 1, 2, 3, \dots, p \quad (\text{C.13})$$

ou

$$\mu_2 = m_2$$

où  $\mu_r$  et  $m_r$  sont les moments non-centrés d'ordre  $r$  de la distribution et de l'échantillon définis dans le sous-chapitre précédent. Les estimations obtenues par la méthode des moments sont convergents, mais biaisées et généralement non efficaces.

#### C.1.4.3 Méthodes des moments pondérés (MMP) (Greenwood et al., 1979)

La méthode des moments pondérés présente un intérêt pour les lois pouvant être exprimées explicitement sous forme inverse  $X = x(F)$ . Soit un échantillon  $x_1, x_2, \dots, x_n$ , tiré d'une distribution  $F(X; \theta_1, \theta_2, \dots, \theta_p)$ . La méthode des moments pondérés conduit au système de  $p$  équations :

$$\beta_{r=b_r} \quad r = 1, 2, \dots, p \quad (C.14)$$

avec les moments pondérés de la distribution et de l'échantillon qui s'expriment respectivement par :

$$\beta_r = E(XF^r) = \int_0^1 X(F)F^r dF \quad (C.15)$$

$$b_r = \sum_{i=1}^n \frac{(i-1)(i-2)\dots(i-r)}{(n-1)(n-2)\dots(n-r)} X_{(i)} \quad (C.16)$$

ou  $X_{(j)}$  est le  $j^{\text{ème}}$  élément de l'échantillon classé en ordre croissant.

#### C.1.4.4 Méthode généralisée des moments (Bobée et Ashkar, 1988)

La différence fondamentale qui existe entre la méthode classique des moments et la méthode généralisée des moments réside dans le fait que cette dernière n'utilise pas nécessairement les  $p$  premiers moments mais plutôt des moments d'ordre  $r_1, r_2, r_3$ . (Ashkar et Ouarda, 1996)

#### C.1.5 Ajustement d'une distribution statistique

Pour plusieurs applications en hydrologie, et notamment pour l'estimation de débits d'étiage ou de crues de conception, on se doit de connaître la probabilité qu'un débit dépasse ou égale une valeur préalablement définie ou encore le débit associé à une probabilité au dépassement donnée. Pour résoudre ce problème on a recours aux notions de fréquences ou de probabilité. Rappelons que la fréquence d'un événement est définie par le nombre d'occurrences de cet événement divisé par le nombre total d'observations de l'échantillon. Considérons un échantillon de  $N$  observations  $x$  ( $x_i ; i = 1, 2, 3, j, \dots, N$ ). La fréquence  $f_i$  de chaque événement est définie par :  $f_i = n_i / N$  ou  $n_i$  est le nombre d'événement à l'intérieur de l'intervalle  $i$  et  $j$  le nombre d'intervalles.

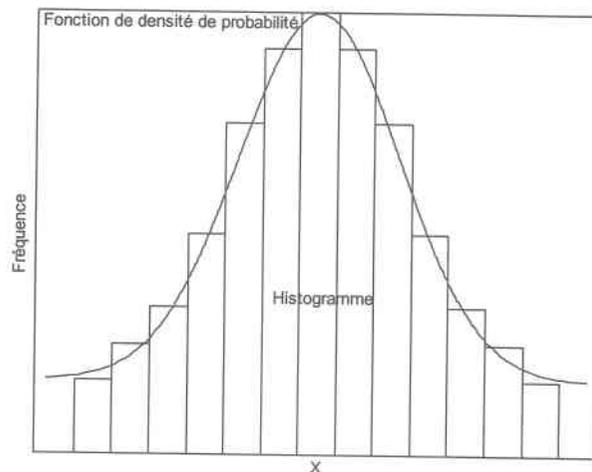


Figure C1 histogramme et fonction de densité de probabilité.

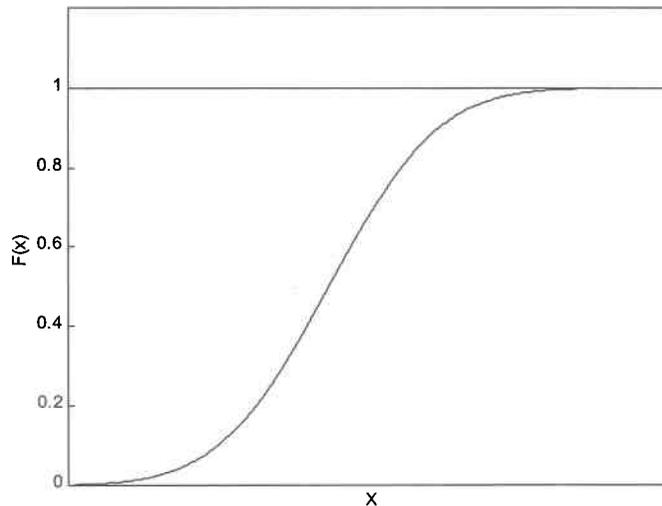
Si  $N$  tend vers l'infini,  $\Delta x$  tend vers 0, l'histogramme prend la forme d'une courbe que l'on appelle fonction de densité de probabilité (f.d.p), et définie par :

$$f(x)dx = P [x \leq X \leq x + dx] \quad (\text{C.17})$$

l'intégrale de la f.d.p de  $-\infty$  à  $x$  en fonction de  $x$  nous permet d'obtenir la fonction de distribution cumulée (f.d.c) ou probabilité au non dépassement, définie par :

$$F(x) = P [X \leq x] = \int_{-\infty}^x f(x)dx \quad (\text{C.18})$$

Une illustration de la f.d.c est donnée par la figure suivante :



**Figure C2** Fonction de distribution cumulée.

Cette fonction s'appelle également la fonction de non dépassement et son complément la fonction de dépassement est donnée par :

---

$$1 - F(x) = P [X > x] \quad (C.19)$$

L'analyse de fréquence permet donc d'estimer les valeurs de débit critique  $X_T$  correspondant à une certaine période de retour, définie comme étant le temps moyen (calculé sur une longue période) entre deux occurrences d'étiage et donnée par :

$$T = \frac{1}{P [X > x]} = \frac{1}{1 - F(x)} \quad (C.20)$$

La probabilité qu'un événement ne se produise pas durant N années  $P_N$  est donnée par :

$$P_N = \left(1 - \frac{1}{T}\right)^N \quad (C.21)$$

On peut alors définir le risque hydrologique (R) comme la probabilité que durant N années un débit de période de retour T dépasse au moins une fois (Viessman et al, 1989) par :

$$R = 1 - \left(1 - \frac{1}{T}\right)^N \quad (C.22)$$

## C.2 Exemples d'estimation régionale classique

### C.2.1 Méthode régressive (exemple)

Vogel et Kroll (1990), proposent une procédure de régionalisation basée sur le développement d'équations de régression régionale généralisée pour l'estimation du quantile d'étiage  $Q_{d,T}$  aux sites non jaugés. ( $d = 3, 7, 14$  et  $30$  jours consécutifs). La méthode utilise comme modèle de départ, celui proposé par Benson (1962a) et qui est :

$$Q_{d,T} = b_0 X_1^{b_1} X_2^{b_2} X_3^{b_3} \dots X_m^{b_m} \varepsilon \quad (\text{C.23})$$

L'estimation des paramètres est effectuée avec la méthode du maximum de vraisemblance, pour les différents sites jaugés et sont obtenus par :

$$Q_{d,T} = \exp(\hat{\mu}_y(d) + Z_T \hat{\sigma}_y(d)) \quad (\text{C.24})$$

où

$$\hat{\mu}_y(d) = \frac{1}{n} \sum_{i=1}^n Y_i(d) \quad (\text{C.25})$$

$$\hat{\sigma}_y^2(d) = \frac{1}{n-1} \sum_{i=1}^n (Y_i(d) - \hat{\mu}_y(d))^2 \quad (\text{C.26})$$

$$Y_i(d) = \ln[q_i(d)] \quad (\text{C.27})$$

$q_i(d)$  : débit minimum moyen de  $d$  jours dans l'année  $i$

$n$  : nombre d'années d'enregistrement

$Z_T$  : variable aléatoire normale standardisée pour la période de retour  $T$ , dont une approximation a été proposée par Tukey (1960) et qui est :

$$Z_T = 4.91 \left[ \left( \frac{1}{T} \right)^{0.14} - \left( 1 - \frac{1}{T} \right)^{0.14} \right], \quad 1.01 \leq T \leq 100 \text{ans} \quad (\text{C.28})$$

Le choix des caractéristiques physiographiques  $X_m$ , est effectué sur la base d'une étude faite par Vogel et Kroll. (1989). Ces auteurs proposent un modèle régional des débits d'étiage qui décrit le temps de réponse du débit d'un bassin versant durant les périodes sèches, et qui s'écrit :

$$Q_t = 4KA(Hd)^2 K_b^t \quad (\text{C.29})$$

A : Surface de drainage

H : Relief du bassin versant

K : conductivité hydraulique

$K_b$  : constante adimensionnelle de récession de l'écoulement de base

d : Densité de drainage

t : temps

(Hd) représente une approximation de la pente du bassin versant

Une estimation précise de K et  $K_b$  étant difficile à obtenir pour un site non jaugé, par conséquent ces deux caractéristiques seront éliminées du modèle. La densité de drainage, d, sera elle aussi ignorée car son utilisation n'améliore pas l'estimation des statistiques des débits minimums aux sites non jaugés. Finalement, les variables indépendantes qui seront retenues pour être utilisées dans le modèle sont, la surface du drainage A et le relief du bassin versant H. L'ajustement du modèle par la procédure de régression des moindres carrés ordinaire (OLS), permet d'obtenir les équations de régression régionales pour la moyenne,  $\hat{\mu}_q(d)$ , et l'écart type,  $\hat{\sigma}_q(d)$ , des séries de débit minimums annuels journaliers

$$\ln[\hat{\mu}_q(d)] = \ln[b_0] + b_1 \ln[A] + b_2 \ln[H] + \varepsilon_\mu \quad (\text{C.30})$$

$$\ln[\hat{\sigma}_q(d)] = \ln[b_0] + b_1 \ln[A] + \varepsilon_\mu \quad (\text{C.31})$$

En passant à l'espace de la variable réelle, et en opérant une transformation on obtient :

$$\hat{\mu}_q(d) = \exp(\hat{\mu}_y(d) + \frac{1}{2} \hat{\sigma}_q^2(d)) \quad (\text{C.32})$$

$$\hat{\sigma}_q(d) = \exp(2\hat{\mu}_y(d) + \hat{\sigma}_y^2(d) [\exp(\hat{\sigma}_y^2(d)) - 1]) \quad (\text{C.33})$$

$\hat{\mu}_q(d)$  et  $\hat{\sigma}_q(d)$  sont obtenues en utilisant les estimateurs trouvée par l'équation (C.24) ; les estimateurs des moments ainsi déduits de la régression régionale dans l'espace des réels  $(\hat{\mu}_q'(d), \hat{\sigma}_q'(d))$  s'écrivent :

$$\hat{\mu}_q'(d) = b_0 A^{b_1} H^{b_2} \quad (\text{C.34})$$

$$\hat{\sigma}_q'(d) = b_0 A^{b_1} \quad (\text{C.35})$$

Les relations de fréquence régionale généralisées des débits d'étiage sont obtenues par :

$$Q_{d,T} = \exp(\hat{\mu}_y'(d) + z_T \hat{\sigma}_y'(d)) \quad (\text{C.36})$$

où les estimateurs des moments dans un espace logarithmique sont obtenus en utilisant les relations suivantes :

$$\hat{\mu}'_y(d) = \ln \left[ \frac{\hat{\mu}'_q(d)}{[1 + (\hat{\sigma}'_q(d) / \hat{\mu}'_q(d))^2]^{1/2}} \right] \quad (\text{C.37})$$

$$\hat{\sigma}'^2_y(d) = \ln \left[ 1 + \left( \frac{\hat{\sigma}'_q(d)}{\hat{\mu}'_q(d)} \right)^2 \right] \quad (\text{C.38})$$

Où  $\hat{\mu}'_y(d)$  et  $\hat{\sigma}'^2_y(d)$  sont données par l'équation (C.34) et (C.35), et  $Z_T$  par l'équation (C.28).

Depuis les travaux de Benson (1962a), plusieurs analyses régionales ont été effectuées à travers le monde, et pour chacune d'entre elles plusieurs caractéristiques physiographiques et climatologiques différentes ont été utilisées comme variables indépendantes dans des modèles régressifs régionaux où on cherche à estimer la variable dépendante  $X_T$ . Un problème souvent commun à ces études réside dans le fait que les variables indépendantes retenues pour les modèles régressifs sont corrélées entre elles. Ce problème de multicollinéarité peut-être atténué en utilisant la régression généralisée des moindres carrés (Steidinger et Tasker, 1985) ou la «ridge regression» (Hoerl et Kenward, 1970 ).

### C.2.2 Indice d'étiage (Exemple)

Pilon (1990), a proposé d'utiliser la méthode de l'indice de crue et de l'adapter aux débits d'étiage, d'où le nom "indice d'étiage". Cette procédure est basée sur l'utilisation d'une distribution théorique largement utilisée dans l'analyse de fréquence des étiages, qui est la distribution Weibull à trois paramètres (W3) dont la densité de probabilité peut s'écrire sous la forme :

$$\phi(x) = \frac{a}{u-e} \left[ \frac{x-e}{u-e} \right]^{a-1} \exp \left[ - \left[ \frac{x-e}{u-e} \right]^a \right] \quad (\text{C.39})$$

$e$  : paramètre représentant la limite inférieure

$u$  : caractéristique de sécheresse

$a$  : paramètre de forme

La fonction de densité cumulée est donnée par :

$$F(x) = 1 - \exp\left[-\left[\frac{x-e}{u-e}\right]^a\right] \quad (\text{C.40})$$

$F(x)$  est la probabilité de non dépassement  $P(X \leq x)$  et correspond à l'inverse de la période de retour. On peut obtenir  $x$  en réarrangeant l'équation (C.40)

$$x = e + (u - e) \left\{ -\ln[1 - F(x)] \right\}^{1/a} \quad (\text{C.41})$$

L'analyse de fréquence des étiages est effectuée pour chaque station en considérant la distribution (W3) ; la distribution adimensionnelle est obtenue en divisant  $X_T$  par un indice d'étiage (dans notre cas  $Q_2$  = débit de période de retour 2 ans). Dans une région homogène, pour chaque station  $i$ , la courbe de fréquence adimensionnelle est considérée comme échantillon aléatoire. La courbe de fréquence régionale adimensionnelle est obtenue en faisant la médiane des courbes adimensionnelles pour toutes les stations de la région. Le résultat est applicable à travers toute la région où les conditions d'homogénéité sont rencontrées. À partir de l'équation (C.41), les paramètres  $a$ ,  $e$  et  $u$  sont obtenus en résolvant le système d'équations à trois inconnus ( $a$ ,  $e$  et  $u$ ) obtenu en choisissant trois débits d'étiage avec leurs probabilités au non dépassement respectives. Si par exemple, les indices d'étiage  $Q_2$ ,  $Q_{12.488}$  et  $Q_{100}$  sont sélectionnés, on devra trouver les valeurs médianes des débits d'étiage adimensionnelles 12.488 et 100 ans, dans la région. Les valeurs médianes seront ensuite substituées dans les expressions suivantes pour obtenir  $a$ ,  $e$  et  $u$  :

---

$$e = (Q_{100} - Q_{12.488}^2) / (1 + Q_{100} - 2Q_{12.488}) \quad (\text{C.42})$$

$$a = 4.23464 / \ln[(1 - e) / (Q_{100} - e)] \quad (\text{C.43})$$

$$u = \left[ (1 - e) / 0.69315^{\frac{1}{a}} \right] + e \quad (\text{C.44})$$

$Q_{100}$  et  $Q_{12.488}$  représentent les valeurs médianes des indices d'étiage 12.488 et 100 ans. Les débits d'étiage régionaux adimensionnels sont ainsi déduit de ces paramètres et de l'équation (C.41). Un test d'homogénéité est ensuite appliqué, ce dernier est similaire à celui décrit par Dalrymple (1960). Par la suite, Pilon (1990) a appliqué cette méthode à une région située au sud de l'Ontario qui a donné des résultats satisfaisants.