UNIVERSITÉ DE MONTRÉAL

MÉTHODE RÉGIONALE D'ESTIMATION DES DÉBITS D'ÉTIAGE POUR LA PROVINCE DE QUÉBEC

ESTHER ROY
DÉPARTEMENT DES GÉNIES CIVIL, GÉOLOGIQUE ET DES MINES
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE CIVIL)
MAI 2000

© Esther Roy, 2000.

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-57426-1

Canada

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

MÉTHODE RÉGIONALE D'ESTIMATION DES DÉBITS D'ÉTIAGE POUR LA PROVINCE DE QUÉBEC

présenté par: ROY Esther

en vue de l'obtention du diplôme de: Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

M. CLÉMENT Bernard, Ph. D., président

M. ROUSSELLE Jean, Ph. D., membre et directeur de recherche

M. HOANG Van Diem, Ph. D., membre et codirecteur de recherche

M. EVORA Noël, Ph. D., membre

REMERCIEMENTS

Cette étude n'aurait pu se réaliser sans la confiance et le constant soutien de monsieur Jean Rousselle, directeur de ce mémoire, que je remercie sincèrement. Je tiens aussi à remercier monsieur Van Diem Hoang, codirecteur de ce mémoire, pour ses idées de projet de recherche et ses précieux conseils d'orientation.

Je désire également offrir mes remerciements aux membres du jury, messieurs Bernard Clément et Noël Evora, pour leurs commentaires constructifs.

Enfin, je remercie sincèrement mon conjoint Charles pour qui j'aurai toujours une grande reconnaissance.

RÉSUMÉ

Le débit annuel d'étiage est un paramètre hydrologique important, autant au niveau de l'activité économique humaine, qu'au point de vue de la préservation des écosystèmes aquatiques. Le phénomène d'étiage se produit généralement à l'été, période de l'année où le niveau de précipitation atteint son minimum. Dans certaines circonstances, l'étiage annuel peut aussi être observé en hiver. L'analyse statistique des données historiques de 119 stations de jaugeage a été effectuée afin d'éliminer les stations ayant des séries de débits non-stationnaires et/ou dépendantes. Cette analyse a permis d'identifier 16 stations non-conformes. L'ensemble des 103 stations restantes a été utilisé dans une méthode d'analyse en composantes principales visant à identifier des blocs de stations homogènes selon les paramètres géographiques et physiques déterminant le phénomème d'étiage. L'analyse a révélé que trois nouvelles variables orthogonales, combinaisons linéaires des variables géophysiques, suffisent à expliquer 80% de la variance observée entre les stations jaugées. Ces variables, lorsque portées en graphique, permettent d'isoler six régions homogènes géographiquement contigües. Des modèles linéaires estimant les étiages 7 jour, périodes de retour de 2 ans et 10 ans (7Q2 et 7Q10), 1 jour avec retour de 10 ans (1Q10) et 30 jours avec retour de 5 ans (30Q5) ont été obtenus par régression sur les variables physiques et sur les trois composantes principales de celles-ci. Il a été trouvé que le modèle optimal, en terme de quantité de variables indépendantes et de qualité d'ajustement, est le modèle linéaire appliqué aux composantes principales sur une base régionale.

Dans le but de généraliser l'application des modèles linéaires obtenus du processus de régionalisation, une étude visant à identifier la fonction de distribution des débits d'étiages a été effectuée. Dans un premier temps, il a été montré que l'ensemble des données disponibles pouvaient être normalisées par la moyenne observée à chaque station de jaugeage afin d'obtenir une courbe de distribution de débits adimensionnels commune à toutes les régions. La réduction des données a aussi réuni en une seule courbe

maîtresse les étiages d'un, sept et trente jours. Dans un deuxième temps, plusieurs fonctions de distribution théoriques ont été ajustées à cette courbe par la méthode du maximum de vraisemblance. La distribution de Pearson III a été identifiée comme étant la mieux adaptée à la courbe expérimentale. La validité de l'ajustement est cependant limitée aux périodes de retour inférieures à cent ans. Tel que démontré par quelques exemples, la modélisation effectuée permet de prédire les étiages d'un, sept et trente jours inférieurs à cent ans pour tout bassin versant de propriétés géographiques et physiques connues.

ABSTRACT

The annual low flow of a river is an important hydrological factor for both economical activities and aquatic ecosystems preservation. These lowest flow levels are usually observed during the summer season, when the level of rainfall is also at its minimum. In special circumstances, the annual low flow may be observed at wintertime, when snow cover hinders the supply of fresh water to rivers and streams. A statistical analysis of historical data coming from 119 gauged stations within the province of Québec has been conducted in order to eliminate those exhibiting behaviors such as non-stationary and year-to-year dependency. The remaining 103 stations were used in a principal components analysis with the aim of identifying homogeneous groups of stations with respect to their physical and geographical properties known to have an effect on the low flow phenomenon. This analysis showed that only three new orthogonal variables, created by a linear combination of the eight original variables, are needed to explain 80% of the sampled population variance. These variables, when plotted in a parametric graph, identified six homogeneous and contiguous regions. Within each of these new regions, a linear regression analysis was performed to obtain predicting models of variables 7Q10, 7Q2, 1Q10 and 30Q5. It was found that the best series of models were those based on principal components variables rather than physical variables.

In order to extend the use of the linear models obtained by the regression procedure, it was necessary to study the distribution of low flow data. This was first undertaken by standardizing the set of data by the average observation at each gauged station. It was found that a single non-dimensional mastercurve common to all regions could be plotted. This data reduction method also showed that 1, 7 and 30 days low flow fell on the same curve, thus enabling the search for a single theoretical distribution for all of these three experimental variables. A number of distribution functions were fitted to this curve by

using the maximum likelihood procedure. The Pearson Type III distribution was found to be the most appropriate to depict the experimental behavior for a return period below a hundred years. As shown with a few calculated examples, the modelling process allows the evaluation of daily, weekly and monthly low flow given that the physical and geographical properties of the area of interest are known.

TABLE DES MATIÈRES

REMERCIEMENTS ..... **iv**

RÉSUMÉ ..... **v**

ABSTRACT ..... **vii**

TABLE DES MATIÈRES ..... **ix**

LISTE DES TABLEAUX ..... **xi**

LISTE DES FIGURES ..... **xii**

LISTE DES SIGLES ET ABRÉVIATIONS ..... **xiv**

LISTE DES ANNEXES ..... **xvi**

CHAPITRE 1: INTRODUCTION **1**

1.1 Contexte 1

1.2 Objectifs 2

CHAPITRE 2: REVUE DE LA LITTÉRATURE **4**

2.1 Classification des débits d'étiage 4

2.2 Régionalisation 7

2.2.1 Identification de régions homogènes 8

2.2.2 Fonction de distribution régionale 10

2.2.3 Estimation des débits 11

2.3 Situation canadienne et études québécoises 13

CHAPITRE 3: STRATÉGIE ET MÉTHODES D'ANALYSE **15**

3.1 Tests statistiques 17

3.1.1 Test d'homogénéité 17

3.1.2 Test d'indépendance 18

3.2 Analyse en composantes principales 19

3.3 Modèles de régression linéaires 21

3.4 Distributions statistiques 24

3.5 Calcul des débits d'étiage 31

CHAPITRE 4: ÉTUDE DES DÉBITS D'ÉTIAGE 34

4.1 Description des données 34

4.2 Caractérisation statistique 37

4.2.1 Homogénéité des séries 37

4.2.2 Test de stationnarité 43

4.2.3 Indépendance des observations 43

4.2.4 Sélection définitive des stations 44

4.3 Régionalisation 45

4.3.1 Description des caractéristiques physiographiques 47

4.3.2 Application de l'ACP 49

4.3.3 Établissement des régions 51

4.4 Modèles de régression linéaires 56

4.4.1 Modèle linéaire multivarié sur variables physiques 57

4.4.2 Régression sur les composantes principales 67

4.5 Modèles de distribution et ajustement 76

4.6 Application 83

4.6.1 Variable hydrologique connue. 83

4.6.2 Aucune variable hydrologique connue 85

4.7 Analyse générale des résultats 88

CHAPITRE 5: CONCLUSION 92

RÉFÉRENCES 94

ANNEXES 102

LISTE DES TABLEAUX

Tableau 4.1: Stations ne satisfaisant pas le test d'homogénéité 38
Tableau 4.2: Stations ayant des séries non stationnaires 43
Tableau 4.3: Stations dépendantes (α=5%) 44
Tableau 4.4: Compilation des stations suite aux tests 45
Tableau 4.5: Analyse en composantes principales sur les variables physiques 50
Tableau 4.6: Composition des nouvelles régions 52
Tableau 4.7: Corrrélation des variables géophysiques des stations 58
Tableau 4.8: Modèle linéaire multivarié global 7Q10 59
Tableau 4.9: Modèle linéaire multivarié global 30Q5 60
Tableau 4.10: Modèle linéaire multivarié global 7Q2 61
Tableau 4.11: Modèle linéaire multivarié global 1Q10 62
Tableau 4.12: Coefficients des modèles linéaires 68
Tableau 4.13: Erreur moyenne associée aux modèles de régression 76
Tableau 4.14: Sommaire de l'écart-type ponctuel des distributions 77
Tableau 4.15: Résultats de l'analyse statistique par le test de Smirnov 79
Tableau 4.16: Sommaire de l'ajustement par maximum de vraisemblance 83
Tableau 4.17: Calcul des propriétés physiques normalisées de la station 080701 86
Tableau A.1: Sommaire des stations étudiées 103
Tableau B.1: Statistiques des débits minimums (annuel, 1 jour) 112
Tableau D.1: Variables 1Q10, 7Q2, 7Q10 et 30Q5 124
Tableau E.1 Variables physiques brutes 128
Tableau E.2 Variables physiques standardisées 131

LISTE DES FIGURES

Figure 3.1: Stratégie de modélisation 16
Figure 3.2: Calcul d'une variable hydrologique de type 7QT à une station "S" 33
Figure 4.1: Répartition du nombre d'années d'enregistrement des stations 35
Figure 4.2: Distribution des minima annuels selon la période d'étiage 36
Figure 4.3: Régions hydrographiques définies par le MEQ. 37
Figure 4.4: Test d'homogénéité-région 01 39
Figure 4.5: Test d'homogénéité-région 02 39
Figure 4.6: Test d'homogénéité-région 03 40
Figure 4.7: Test d'homogénéité-région 04 40
Figure 4.8: Test d'homogénéité-région 05 41
Figure 4.9: Test d'homogénéité-région 06 41
Figure 4.10: Test d'homogénéité-région 07 42
Figure 4.11: Test d'homogénéité-région 08 42
Figure 4.12: Représentation graphique: composantes principales régions A et B 53
Figure 4.13: Représentation graphique: composantes principales régions C et D 54
Figure 4.14: Représentation graphique: composantes principales régions E et F 55
Figure 4.15: Limites géographiques des nouvelles régions 56
Figure 4.16: Variable 7Q10 estimée par modèle linéaire: superficie et précipitation 63
Figure 4.17: Variable 30Q5 estimée par modèle linéaire: superficie et précipitation 64
Figure 4.18: Variable 7Q2 estimée par modèle linéaire: superficie et précipitation 65
Figure 4.19: Variable 1Q10 estimée par modèle linéaire: superficie et précipitation 66
Figure 4.20: Estimation de 1Q10, 7Q10, 7Q2, et 30Q5 pour la région A 69
Figure 4.21: Estimation de 1Q10, 7Q10, 7Q2, et 30Q5 pour la région B 69
Figure 4.22: Estimation de 1Q10, 7Q10, 7Q2, et 30Q5 pour la région C 70
Figure 4.23: Estimation de 1Q10, 7Q10, 7Q2, et 30Q5 pour la région D 70
Figure 4.24: Estimation de 1Q10, 7Q10, 7Q2, et 30Q5 pour la région E 71
Figure 4.25: Estimation de 1Q10, 7Q10, 7Q2, et 30Q5 pour la région F 71

Figure 4.26: Variable 7Q10 estimée par modèle linéaire: PC1, PC2 et PC3 72
Figure 4.27: Variable 30Q5 estimée par modèle linéaire: PC1, PC2 et PC3 73
Figure 4.28: Variable 1Q10 estimée par modèle linéaire: PC1, PC2 et PC3 74
Figure 4.29: Variable 7Q2 estimée par modèle linéaire: PC1, PC2 et PC3 75
Figure 4.30: Distribution cumulative des minima annuels normalisés par μ 78
Figure 4.31: Distribution cumulative des minima annuels normalisés μ et σ 78
Figure 4.32: Distribution globale des débits normalisés –méthode #1 80
Figure 4.33: Ajustement de la distribution Weibull aux données observées..... 80
Figure 4.34: Ajustement de la distribution gamma aux données observées..... 81
Figure 4.35: Ajustement de la distribution log-normale aux données observées 81
Figure 4.36: Ajustement de la distribution Pearson III aux données observées 82
Figure C.1 Stationnarité du débit minimum 7 jours pour la station 010901 117
Figure C.2 Stationnarité du débit minimum 7 jours pour la station 021915 117
Figure C.3 Stationnarité du débit minimum 7 jours pour la station 022504 118
Figure C.4 Stationnarité du débit minimum 7 jours pour la station 030215 118
Figure C.5 Stationnarité du débit minimum 7 jours pour la station 030242 119
Figure C.6 Stationnarité du débit minimum 7 jours pour la station 030309 119
Figure C.7 Stationnarité du débit minimum 7 jours pour la station 030401 120
Figure C.8 Stationnarité du débit minimum 7 jours pour la station 050118 120
Figure C.9 Stationnarité du débit minimum 7 jours pour la station 051006 121
Figure C.10 Stationnarité du débit minimum 7 jours pour la station 072301 121
Figure C.11 Stationnarité du débit minimum 7 jours pour la station 076601 122

LISTE DES SIGLES ET ABRÉVIATIONS

1Q10	Débit d'étiage sur 1 jour avec une période de retour de 10 ans
7Q2	Débit d'étiage sur 7 jours avec une période de retour de 2 ans
7Q10	Débit d'étiage sur 7 jours avec une période de retour de 10 ans
30Q5	Débit d'étiage sur 30 jours avec une période de retour de 5 ans
ACP	Analyse en composantes principales
$B_0,..B_i$	Coefficients de régression
C1	Coefficient de composante principale
CV	Coefficient de variation
DAM	Débit annuel moyen
DCE_{355}	Débit journalier moyen dépassé pendant 355 jours par année
F_0	Statistique de Fisher
H	Hypothèse
K	Facteur d'uniformité
L	Moment
L	Fonction de vraisemblance
MSE	Moyenne des carrés de l'erreur
PCi	Composante principale i
Q_{50}	Débit mensuel médian (débit minimum dépassé 50% du temps à chaque mois)
Q_{90}	Débit minimum dépassé 90% du temps à chaque mois
Q_k	Débit minimum d'une période de retour k
R^2	Coefficient de corrélation multiple
$R_{W\&W}$	Statistique d'indépendance de Wold et Wolfowitz
SSE	Somme des carrés de l'erreur
SSR	Somme des carrés de l'erreur expliquée par le modèle
T_1	Statistique de Smirnov
μ	Moyenne
β	Coefficient de vecteur de régression

α	Degré de confiance
ε	Résidu d'une équation de régression
σ	Écart-type

LISTE DES ANNEXES

Annexe A: Sommaire des stations étudiées 102
Annexe B: Statistiques des débits minimums 111
Annexe C: Séries de données non stationnaires 116
Annexe D: Variables 1Q10, 7Q2, 7Q10 et 30Q5 123
Annexe E: Données physiques brutes et standardisées 127

CHAPITRE 1

INTRODUCTION

1.1 Contexte

L'eau est une ressource de plus en plus convoitée, ce qui nécessite une législation pour en assurer sa disponibilité et sa qualité. Les inondations et les sécheresses sont des événements extrêmes qui ont de lourdes conséquences sur l'activité humaine. Ces événements sont issus de phénomènes hydrologiques et météorologiques complexes. La présente étude hydrologique ne concerne que les débits minimums. L'étude des débits d'étiage et la connaissance de leurs valeurs extrêmes permet de mieux gérer les besoins d'approvisionnement en eau des populations et des industries, de contrôler la pollution causée par les rejets de contaminants dans l'eau et de protéger les écosystèmes aquatiques.

Le ministère de l'Environnement du Québec est à la recherche d'une méthode de détermination des débits d'étiage afin de faciliter le processus d'analyse des dossiers environnementaux . Ces dossiers sont plus précisément l'analyse des rejets des pisciculteurs et des industries dans le but de protéger les habitats de poissons. Aussi, les bassins hydrologiques concernés sont principalement localisés en régions urbaines ou semi-urbaines du Québec méridional. La méthode de détermination recherchée doit prendre en considération l'étendue du territoire provincial et la diversité des conditions climatiques qu'on y retrouve tout en restant d'utilisation pratique. Ces conditions peuvent souvent être remplies à la suite d'un exercice de régionalisation visant à regrouper, au sein d'un même territoire, des stations de jaugeage statistiquement similaires. Les territoires ainsi délimités doivent être géographiquement continus et contigüs afin de faciliter l'utilisation des statistiques obtenues.

Les phénomènes hydrologiques extrêmes sont souvent caractérisés par un comportement statistique complexe. Ceci est particulièrement vrai dans le cas des débits d'étiage, un événement essentiellement attribuable à un manque de précipitation (étiage estival) ou à une déficience du processus de ruissellement (étiage hivernal). Avant même de considérer l'étude statistique de ces débits selon une perspective régionale, il est nécessaire de vérifier la validité des données historiques disponibles sur une base ponctuelle à chaque site de jaugeage.

1.2 Objectifs

Selon les besoins exprimés par un représentant du ministère, l'étude comprend trois objectifs principaux. Dans un premier temps, il s'agit d'identifer des régions hydrologiques homogènes du point de vue de leur comportement d'étiage. Cet objectif sous-entend que les variables physiographiques et hydrologiques les plus siginificatives, c'est-à-dire celles qui expliquent le plus les variations des débits minimums, ont été préalablement identifiées. Ce sont ces variables qui, conjuguées à des techniques d'analyse de variance, telle l'analyse en composantes principales, seront utilisées pour définir ces nouvelles régions basées sur des critères de débits minimums. Dans un deuxième temps, on tentera d'établir des modèles régionaux permettant d'estimer le débit d'étiage en fonction d'un petit nombre de caractéristiques physiographiques. Ces modèles seront obtenus au moyen d'analyse de régressions linéaires effectuées sur des variables hydrologiques particulières tel que l'étiage de 7 jours avec période de récurrence de 10 ans (7Q10). Enfin, dans un troisième temps, on effectuera l'ajustement de distributions de probabilité théoriques à l'ensemble des débits d'étiage normalisés disponibles. Ceci permettra d'être en mesure de généraliser l'application des modèles linéaires obtenus à tous les types d'étiage pour un plus large éventail de périodes de retour.

Les chapitres qui suivent comprennent la revue de la littérature, portant principalement sur les débits d'étiage et la régionalisation, suivie de la stratégie globale et des méthodes utilisées dans l'étude. Par la suite, les résultats des analyses sur les débits d'étiage sont présentés. Enfin, les principaux résultats et points de discussion sont résumés dans la conclusion.

CHAPITRE 2

REVUE DE LA LITTÉRATURE

Les débits extrêmes, qu'ils soient des inondations ou des étiages, ont des conséquences très néfastes sur les populations. Bien que les domaines d'études reliés à ces types d'événements soient nombreux, cette recherche est limitée aux débits minimums. Ce chapitre traite principalement des méthodes de régionalisation et de détermination de débits d'étiage utilisées en hydrologie. Ces techniques ont fait l'objet de plusieurs études de comparaison et de mesure d'efficacité. Dans un premier temps, les types de débits d'étiage sont définis, car la sélection du débit à étudier constitue une étape préliminaire à la régionalisation. La deuxième partie de cette revue résume les approches les plus utilisées pour délimiter des régions homogènes du point de vue de leur comportement statistique et des méthodes de détermination des débits d'étiage. En dernier lieu, un bref aperçu est donné spécifiquement sur les études canadiennes qui ont été menées sur les débits minimums.

2.1 Classification des débits d'étiage

La complexité des méthodes de détermination des débits d'étiage dépend des objectifs de protection demandés (Beecher, 1990). Loar et Sale (1981) proposent de classer les approches en trois grandes catégories: les méthodes d'habitat préférentiel, les méthodes hydrauliques et hydrologiques.

Les analyses d'habitat préférentiel sont les méthodes les plus complexes de détermination du débit réservé. En plus de considérer des paramètres hydrauliques et hydrologiques, elles comprennent des variables biologiques reliées aux espèces étudiées (Collings, 1972; Bovee, 1982). Ces méthodes nécessitent une grande quantité de données de terrain. Les

méthodes hydrauliques sont plus simples et requièrent moins de relevés des sites. Ces techniques établissent une relation directe entre les caractéristiques hydrauliques d'un cours d'eau, comme l'aire et le périmètre mouillé et l'habitat des poissons (Hamilton et Kosakoski, 1982). Leur inconvénient est qu'elles ne peuvent établir un lien entre les niveaux de débits et la convenabilité des habitats ou les cycles de vie de différentes espèces de poissons. Enfin, les méthodes hydrologiques sont basées sur l'analyse de séries de débits. Ces méthodes sont les plus utilisées, car elles sont simples, faciles à mettre en application et ne nécessitent pas de relevés de terrain. Ces méthodologies sont donc moins coûteuses à développer et à mettre en application. Dans cette recherche, des données hydrologiques, soient les débits minimums, seront utilisées considérant le fait qu'une étude à l'échelle de la province devait être effectuée.

Les méthodes hydrologiques peuvent être adaptées pour des études de régionalisation en utilisant des analyses de régressions multiples qui permettent de déterminer un débit à un site non jaugé. Ce type d'approche est généralement nécessaire pour des études à grande échelle. Les méthodes hydrologiques incluent les méthodes qui déterminent les débits réservés à partir d'un pourcentage du débit moyen. La méthode de Tennant établi des classes de débits réservés correspondant à un pourcentage du débit moyen annuel (DAM). Par exemple, un débit inférieur à 10% du DAM est rélié à une "détérioration grave" des habitats aquatiques (Belzile *et al.*, 1997). Dans une revue des différentes méthodes de quantification des débits d'étiage, Estes et Orsborn (1986) ont comparé la méthode de Tennant à trois autres méthodes, dont la méthode conjointe "Instream Flow Incremental Methodology (IFIM) – PhysiHabitat Simulation (PHABSIM)", développée par le "US Fish and Wildlife Service ". La méthode de Tennant suppose que des rivières ayant un même DAM auront des conditions de vie similaires de leur habitat aquatique. Les méthodes hydrologiques comprennent également des statistiques décrivant des conditions d'étiage comme le débit minimum consécutif sur 7 jours de période de retour de 10 ans (7Q10).

Une étude comparative des méthodes de détermination des débits d'étiage a été réalisée par Caissie et El-Jabi (1995). Cinq méthodes hydrologiques, les plus utilisées au Canada et aux États-Unis, ont été appliquées sur 70 rivières des provinces atlantiques: la méthode de Tennant, décrite précédemment, 25% du débit annuel moyen (DAM de 25%), le débit mensuel médian (Q_{50}), le débit dépassé 90% du temps sur la courbe de durée du débit mensuel (Q_{90}) et le faible débit de 7 jours d'une récurrence de 10 ans (7Q10). La méthode du 25% de débit annuel moyen est une variante simplifiée de la méthode de Tennant où l'on juge que 25% du DAM suffise à préserver la qualité de l'écosystème aquatique peu importe la saison ou l'espèce considérée. La méthode du débit mensuel médian (Q_{50}) établi que le débit minimum présent 50% du temps à chaque mois est nécessaire à la survie de l'habitat aquatique. La méthode du Q_{90} est similaire la celle du Q_{50} sauf qu'elle utilise un critère de rang centile plus sévère soit 90%. Enfin, la méthode du 7Q10 fait appel à une analyse statistique des données quotidiennes historiques afin d'établir le débit minimum pour sept jours consécutifs pour une période de retour de dix ans. Les résultats des méthodes ont été comparés au débit annuel moyen de 25% qui est couramment utilisé dans les provinces maritimes. Les prévisions de débits issues des méthodes Q_{90} et 7Q10 étaient très faibles; jusqu'à -90% du DAM de 25%. Les méthodes de Tennant, Q_{50} et DAM de 25% ont été recommandées.

Par ailleurs, il existe d'autres méthodes d'estimation des débits d'étiage qui sont moins connues au Canada: la méthode de l'Utah, la New England Flow Method, la méthode de Hoppe et les méthodes européennes (Belzile *et al.*, 1997). La méthode de l'Utah a été conçue pour les rivières à salmonidés de l'Utah et nécessite des données historiques. Elle s'applique en divisant l'année en deux périodes, hivernale et estivale. Pour chaque période, le débit moyen journalier recommandé s'obtient par la moyenne des débits mensuels minimums pour chaque mois de la période considérée. La New England Flow Method préconise de laisser dans les rivières aménagées un débit minimum correspondant au débit mensuel du mois d'août, soit le débit réservé de base. Le mois d'août est choisi comme débit réservé de base parce qu'il correspond à la période de faible

hydraulicité dans la région de la Nouvelle-Angleterre. La méthode de Hoppe est une variante de la méthode Q_{90}, discutée précédemment, où l'on utilise des rangs centiles différents selon le cycle de vie de l'habitat d'intérêt. Ainsi, un rang centile de 80% suffit à assurer des abris et de la nourriture aux poissons alors que 40% de rang centile est nécessaire pour la reproduction. Finalement les méthodes européennes sont fondées sur des corrélations mathématiques utilisant des valeurs de débits observés durant les périodes d'étiage. Souvent le débit d'étiage y est défini comme le débit journalier moyen dépassé pendant 355 jours par année (DCE_{355}). Le débit réservé est calculé à partir de DCE_{355} selon une formule adaptée propre à la région ou l'organisme considéré.

2.2 Régionalisation

La régionalisation est un moyen souvent utilisé en hydrologie pour faciliter l'extrapolation des données des sites jaugés à des sites où les données sont inexistantes ou trop peu nombreuses pour effectuer des analyses concluantes. En divisant une région en sous-régions ayant des comportements similaires, l'extrapolation de données devient plus précise. Aussi, il a été démontré qu'une distribution de probabilité régionale a de meilleurs résultats qu'une distribution au site, car elle attribue un meilleur poids aux événements extrêmes comme pour les crues (Anctil *et al.*, 1998; Potter, 1987). L'approche régionale est encore plus importante dans la détermination des débits d'étiage. Ces débits sont grandement dépendants de la géologie et du phénomène d'évapotranspiration. Étant donné que ces paramètres sont difficiles à décrire par des indices, la régionalisation permet de faire des approximations acceptables des caractéristiques d'étiage (Hoang et Tremblay, 1976).

L'analyse régionale se divise généralement en trois étapes: l'identification de régions homogènes, le choix d'une fonction de distribution régionale et l'estimation des

paramètres de la fonction. De plus, différentes méthodes de calcul de débit se greffent au processus.

2.2.1 Identification de régions homogènes

Une des premières méthodes utilisées fut le regroupement des régions selon leur proximité géographique (Dalrymple, 1960; Mosley, 1981; Hawley et McCuen, 1982). Cependant des régions définies par des similarités hydrologiques ou physiographiques ne sont pas nécessairement dans le même secteur géographique. D'autres méthodologies ont également été développées telle que l'analyse par régions d'influence (Acreman et Wiltshire, 1987; Burn, 1990), l'analyse par comparaison d'échelles (Nguyen et Pandey, 1996) et l'analyse spatiale des résidus d'équations de régression (Tasker, 1982). Dans l'analyse par région d'influence, chaque site est défini comme étant le centre de sa propre région. Ensuite, le rayon d'influence de ce site sur ces voisins est calculé selon une distance euclidienne dans un hyperespace formé des variables hydrographiques du bassin. D'autre part, l'analyse spatiale des résidus de régression est en fait une procédure de classification fondée sur les variances observées à chaque station.

Cependant, les approches les plus utilisées pour identifier des régions homogènes sont les analyses multivariées. Les techniques multivariées comprennent l'analyse d'agrégation ou par regroupement (Tasker, 1982; Acreman et Sinclair, 1986; Burn, 1989); l'analyse canonique (Ribeiro-Corréa *et al.*, 1995) et l'analyse en composantes principales (Nathan et McMahon, 1990). L'analyse canonique est une technique d'analyse statistique multivariée qui permet de réduire les problèmes de dépendence linéaire entre deux groupes de variables, dans ce cas-ci des variables hydrologiques et physiographiques. Plusieurs auteurs ont utilisé l'analyse en composantes principales à des fins hydrologiques (Morin *et al.*, 1979; Tessier, 1983; Siew-Yan-Yu *et al.*, 1985; Boucher *et al.*, 1986). Ils ont employé cette méthode pour identifier des stations prédominantes afin

de rationaliser le réseau de mesure ou pour déterminer les caractéristiques physiographiques de bassins les plus explicatives.

Une étude, effectuée par le Groupe de recherche en hydrologie statistique (GREHYS, 1996a), compare des combinaisons de méthodes de détermination de régions et de méthodes d'estimation régionale de débits de crues. Les méthodes de délimitation géographique étudiées sont l'analyse par région d'influence, l'analyse canonique, l'analyse par correspondance et la méthode des moments ("L-moments"). L'analyse par correspondance, une variante de l'analyse en composantes principales, vise à identifier au sein d'une base de données, des groupes similaires selon un certain nombre de variables considérées. En fait, c'est une représentation multidimensionnelle de la dépendance entre les rangées et les colonnes d'une base de données. La méthode des moments L vise à vérifier si les observations de débits à différents sites dans une région homogène peuvent être considérées comme provenant d'une distribution régionale commune.

Suite à une première étape de détermination de régions, plusieurs techniques de raffinement de l'homogénéité de ces régions peuvent être appliquées. Ces techniques comprennent principalement des tests statistiques, tel que le test de discordance (Hosking et Wallis, 1993) et les tests d'homogénéité (Hosking et Walllis, 1993; Fill et Stedinger, 1995). Le test de discordance utilise les moments *L* (*L* CV, *L* asymétrie et *L* curtosis) comme coordonnées d'un espace tri-dimensionnel pour comparer la dispersion des sites d'échantillonnage. Les tests d'hétérogénéité sont effectués à partir des moments ou en utilisant des variables hydrologiques estimées de l'échantillon local (1Q10, 1Q2.33 etc..) Suite à ces tests, certaines stations ou éléments peuvent être rejetés. La méthode des moments est également couramment utilisée pour vérifier l'homogénéité des régions proposées (Hosking et Wallis, 1993; GREHYS, 1996b).

Certains problèmes peuvent être associés à la régionalisation, principalement lors de l'utilisation de l'analyse par regroupement et l'analyse en composantes principales. Par exemple, des sous-régions peuvent être similaires selon certaines variables (ex: pourcentage de forestation) mais pas selon d'autres (ex: géologie du sol). Aussi, l'ordre de grandeur des variables mesurées peut influencer le poids de ces variables (ex: la hauteur de précipitation en millimètres ou en mètres). Les variables mesurées par de petites unités (grand nombre), vont davantage contribuer à la similarité entre les sous-régions. Pour diminuer l'influence de ces facteurs, Nathan et McMahon (1990) proposent d'utiliser une combinaison des méthodes citées précédemment. En premier lieu, ils effectuent une régression multiple pour sélectionner les variables significatives. Par la suite, ils appliquent une analyse d'agrégation pour déterminer des régions préliminaires. Finalement, l'hétérogénéité des régions préliminaires est investiguée et minimisée à l'aide d'une technique graphique multidimensionnelle (Andrews, 1972).

2.2.2 Fonction de distribution régionale

Dans le processus de régionalisation, l'analyse de fréquences comprend le choix et l'ajustement d'une fonction de distribution pour modéliser les données observées. Généralement, il est préférable d'approximer les séries de débits d'étiage par une fonction à trois paramètres ou plus. Les distributions les plus connues pour représenter les débits d'étiage sont les distributions de Weibull, Gumbel et du log-Pearson III. La fonction la plus utilisée et utilisant trois paramètres, est la fonction log-Pearson III. Les lois de distributions doivent autant que possible représenter fidèlement les séries de données. Une revue de l'analyse de fréquences appliquée aux débits d'étiage a été effectuée par Loganathan *et al.* (1985). Des techniques de transformations de type "SMEMAX" (Small, MEdian, MAXimum) y sont discutées ainsi que les distributions de Weibull (deux ou trois paramètres) et log-Pearson III. Les fonctions peuvent être ajustées

aux données par diverses techniques dont la méthode du maximum de vraisemblance et la méthode des moments (Haan, 1977; Bobée et Ashkar, 1988; Hosking, 1990).

Lorsque les échantillons sont de petite taille, (20 données et moins), il est préférable de régionaliser une fonction de distribution (Nathan et McMahon, 1990). Dans le cadre d'une étude de comparaison de méthodes d'estimation de paramètres d'une distribution, Nathan et McMahon (1990) ont démontré qu'il y avait de grands écarts entre les quantiles des différentes approches. Ils ont donc recommandé d'utiliser plus d'une méthode pour déterminer les débits d'étiage extrêmes principalement dans les cas de petits échantillons. Les travaux de Bobée et Ashkar (1991) corroborent également cet énoncé. Ribeiro-Correa et Rousselle (1993) ont développé une approche régionale pour la fonction Pearson III contrant ainsi l'inconvénient précédemment expliqué. Les auteurs ont utilisé une méthode hiérarchique-empirique Bayes pour la fonction Pearson III (HP3) .

Suite aux méthodes d'ajustement des fonctions, on peut effectuer des tests de vérification du niveau ou de la qualité de l'ajustement d'une fonction. Le test de Kolmogorov-Smirnov en est un exemple (Llamas, 1993). Plusieurs autres auteurs décrivent ces types de tests dont Hosking et Wallis(1993), et Nathan et McMahon (1990).

2.2.3 Estimation des débits

Lorsque les régions sont bien définies et que les fonctions de distribution représentent adéquatement les données, les méthodes d'estimation de débits peuvent être appliquées. Une compilation des méthodes existantes a été effectuée pour l'analyse régionale des crues (Cunnane, 1988). Les techniques d'estimation sont principalement les régressions linéaires, la méthode par index ("index flood method"), la méthode des moments et l'analyse régionale non-paramétrique. La méthode par index fait l'hypothèse que le débit

de crue à chaque site d'une région homogène est identiquement distribué si normalisé avec un facteur d'échelle fonction des caractéristiques physiographiques du bassin. Cette fonction de distribution doit être choisie pour l'ensemble des sites d'une région. L'analyse régionale non-paramétrique est une alternative aux méthodes traditionnelles. Elle se résume par la représentation d'une distribution de fréquences sous forme d'histogramme. L'étude conclue en une distinction entre les méthodes de régression et les méthodes par index. La méthode par index, qui se base sur des paramètres régionaux, exige une détermination de régions très homogènes. Si le critère d'homogénéité est plus ou moins satisfaisant, les modèles de régressions sont à privilégier. Pour les sites jaugés, la méthode par index combinée avec une ou l'autre des méthodes de détermination de régions présentent les meilleurs modèles régionnaux. Pour les sites non-jaugés, toutes les méthodes d'estimation combinées à l'analyse de corrélation canonique semblent donner de bons résultats. Cette étude, réalisée à l'aide de données des provinces de l'Ontario et du Québec, met en évidence qu'il n'y a pas de de modèle de régionalisation universel. Ainsi, pour une région donnée, différentes analyses sont requises.

Dans l'analyse régionale des débits d'étiage, les techniques les plus connues sont les équations de régression utilisant des variables physiographiques ou hydrologiques. Cependant, Thomas et Benson (1970) ont démontré que les erreurs moyennes de prédiction des débits d'étiage issues des modèles de régression sont beaucoup plus significatives que celles des débits de crues. Les erreurs de prédiction peuvent être cependant atténuées en additionnant plus de variables dans les équations de régression (Vogel et Fennessey, 1993). Une autre façon de déterminer des débits est de régionaliser les paramètres de distribution d'une fonction ou de régionaliser d'autres statistiques comme le coefficient d'asymétrie (Durrans et Tomic, 1996). Plus récemment, Liu *et al.* (1996) ont proposé un modèle bivarié d'estimation des débits d'étiage pour 7 jours avec une récurrence de 2 et 10 ans (7Q2 et 7Q10).

2.3 Situation canadienne et études québécoises

Les travaux visant à caractériser les événements hydrologiques extrêmes au Canada et dans la province de Québec ont avant tout porté sur le phénomène des inondations. Souvent, les études effectuées visaient à choisir une distribution décrivant adéquatement des données historiques locales ou régionales. Un guide à cet effet a été publié par Watt *et al.* (1989). Dans le cas des débits d'étiage, Lawal et Watt (1996a, 1996b) et Chow et Watt, (1992) se sont aussi intéressés à établir la pertinence d'utiliser une distribution théorique à trois paramètres. Ils ont montré que, pour un échantillonnage provenant de plus de 51 stations de jaugeage au Canada, il était préférable de choisir une distribution à deux paramètres, i.e. faire l'hypothèse que l'étiage minimum est nul.

Plusieurs études de régionalisation ont été effectuées pour la province de Québec (Desforges et Tremblay, 1974; Hoang et Tremblay, 1976; Rasmussen *et al.*, 1995; Belzile *et al.*, 1997 et Anctil *et al.*, 1998). Toutes ces études ont mené à la détermination de régions hydrologiques. Le travail de Anctil *et al.* (1998) visait à effectuer une analyse régionale des crues journalières de la province de Québec. Ils ont regroupé, à l'aide de tests de discordance et d'homogénéité, 136 rivières en trois régions homogènes. Seules les délimitations des zones homogènes des études récentes, telle celle de Anctil *et al.* (2000), reposent sur des critères de débits d'étiage.

Dans leurs travaux, Hoang et Tremblay (1976) ont développé différentes méthodes de calculs pour étudier les faibles débits en saison estivale sur les rivières du Québec méridional. Ils ont en premier lieu déterminé 11 régions homogènes du point de vue de leurs caractéristiques de débits d'étiage. Ces régions ont été finalement regroupées en 4 zones selon leur géologie, météorologie et physiographie. Ces mêmes auteurs ont développé trois méthodes d'estimation de débits d'étiage aux stations non-jaugées:

1- équation de régression multiple dont les paramètres sont des caractéristiques physiques du bassin;
2- courbe de fréquence régionale et transposition du débit moyen annuel d'étiage par régression;
3- régionalisation des moments d'une loi statistique.

La première méthode, par équation de régression multiple, a déjà été discutée. Pour obtenir le débit d'étiage de différentes périodes de retour, la deuxième méthode consiste à multiplier le débit moyen annuel d'étiage par un quotient régional. Le débit moyen est trouvé par une équation de régression basée sur les caractéristiques physiographiques. Le quotient régional, étant le ratio du débit d'étiage de récurrence "T" sur le débit moyen annuel d'étiage, est déduit de la courbe de fréquence régionale. La troisième méthode utilise des équations de régression régionales, basées sur les caractéristiques physiques de bassins, pour trouver les paramètres statistiques suivants: la moyenne, l'écart-type et le coefficient d'asymétrie. Ce dernier utilise celui de la région. Ces paramètres sont par la suite utilisés dans l'équation de la loi log Pearson III pour calculer le débit d'étiage. Enfin, toutes les équations de régression et les courbes de récurrence ont été développées pour des débits d'étiage de 1 jour, 7 et 30 jours consécutifs.

Dans leur analyse récente sur les étiages 7-jours de la province de Québec, Anctil *et al.* (2000) utilisent une méthodologie similaire à leur approche antérieure pour les débits de crues (Anctil *et al.*, 1998). A partir du débit observé à 119 stations de jaugeage du Québec, ils ont délimité au moyen des tests de discordance et d'homogénéité 6 régions homogènes. Les auteurs ont effectué cette régionalisation selon une période hivernale et estivale, selon l'hypothèse que les facteurs gouvernant l'étiage estival sont différents de ceux affectant l'étiage hivernal. Pour chacune des régions, une distribution régionale normalisée a été ajustée à partir de la méthode des moments. Ils ont conclu que la méthode proposée était suffisamment robuste pour permettre l'estimation de l'étiage d'une rivière non-jaugée

CHAPITRE 3

STRATÉGIE ET MÉTHODES D'ANALYSE

L'étude des phénomènes d'étiage nécessite l'utilisation de plusieurs techniques d'analyses statistiques, mathématiques et même géographiques afin de caractériser adéquatement les données hydrologiques recueillies périodiquement dans les régions d'intérêt. Ces méthodes sont aussi requises afin d'estimer la récurrence d'événements extrêmes et d'identifier les variables hydrologiques ou physiographiques significatives. Ce chapitre présente la stratégie globale retenue pour effectuer le traitement des données de débits et la modélisation à réaliser afin d'estimer régionalement, pour la province de Québec, le phénomène d'étiage. Il décrit aussi chacune des méthodes utilisées ultérieurement dans ce travail.

L'approche retenue, afin d'arriver à une description quantitative adéquate des débits d'étiage, est résumée dans un arbre de décision présenté à la figure 3.1. Elle fait appel aux observations de 119 stations et à huit régions hydrologiques déjà définies par le ministère de l'Environnement du Québec. Dans un premier temps, un test d'homogénéité sera effectué sur les régions existantes. Si, tel qu'anticipé, plusieurs régions renferment des stations non-homogènes par rapport au groupe, la procédure de régionalisation sera initiée en éliminant, des 119 stations originales, celles qui ne satisfont pas aux critères d'indépendance et de stationnarité. Dans un deuxième temps, une analyse en composantes principales (ACP) sera conduite à partir des stations restantes. Cette analyse s'effectuera en fonction des caractéristiques physiques de chacune des stations retenues après les tests de sélection initiaux. L'ACP vise à transformer les caractéristiques physiques ayant un effet sur le comportement hydrologique d'un bassin versant en un nombre réduit de variables orthogonales.

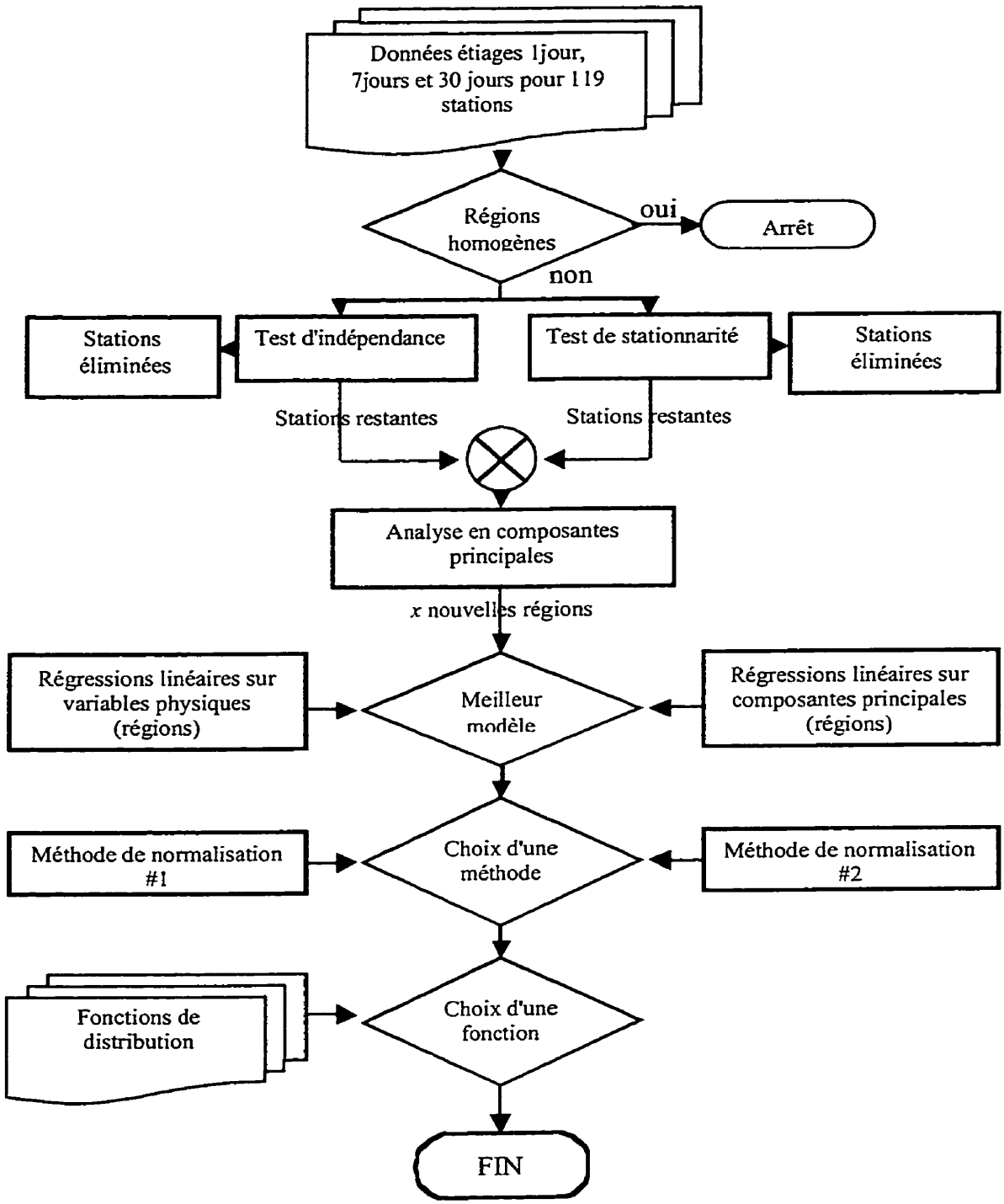

Figure 3.1: Stratégie de modélisation

La régionalisation sera complétée par la comparaison graphique des valeurs des nouvelles variables orthogonales à chacune des stations. Ceci permettra d'identifier *n* nouvelles régions. Dans un troisième temps, on appliquera aux nouvelles régions des modèles de régressions linéaires reliant le débit d'étiage à des caractéristiques physiques ou à des combinaisons linéaires de celles-ci, obtenues par l'ACP. Dans un quatrième temps, une normalisation des données sera effectuée afin d'obtenir des courbes de distribution expérimentales communes pour toutes les régions. Enfin, une fonction de distribution sera ajustée à cette distribution normalisée. Ceci, conjointement avec les modèles linéaires obtenus précédemment, permettra d'estimer le débit d'étiage pour une période de retour quelconque à une station non-jaugée. La procédure de calcul nécessaire est discutée à la dernière section de ce chapitre.

3.1 Tests statistiques

3.1.1 Test d'homogénéité

La méthode utilisée afin d'évaluer l'homogénéité d'un ensemble de données est celle décrite par Llamas (1993). Sommairement, il s'agit d'abord de tracer les courbes limites selon les formules de la référence citée. Puis, pour chaque station, il faut calculer le débit suivant:

$$Q_K = KQ_{2,33} \tag{3.1}$$

où

$$K = \frac{1}{N}\sum^{N}\left(\frac{Q_{10}}{Q_{2,33}}\right) \tag{3.2}$$

avec:

K: facteur d'uniformité

Q_{10}: débit minimum d'une période de retour de 10 ans

$Q_{2,33}$: débit minimum d'une période de retour de 2,33 ans

N: nombre de stations

Sur chacun des N graphiques de distribution, on lit la période de retour correspondant à Q_k. Enfin, les différentes périodes de retour avec leur période d'enregistrement correspondante ("n" années) sont portées en graphique. Les stations non homogènes sont celles situées à l'extérieur des bandes de confiance du test.

3.1.2 Test d'indépendance

Une série de données est temporellement indépendante si aucune observation dans la série n'a une influence sur les observations suivantes (Tomic, 1994). Dans le cas d'une série de débits d'étiage annuels, on émet donc l'hypothèse que le débit d'une année donnée n'est pas influencé par le débit de l'année précédente. Cette condition est importante lorsqu'on désire associer une série d'observations à une fonction de distribution. Le test d'indépendance retenu pour ce travail fait appel à la méthode de Wald et Wolfowitz (Bobée et Ashkar, 1991), car cette méthode est la plus utilisée dans les différents logiciels de statistique. Toutefois, dans le cadre de ce projet, un code de calcul spécifique a été écrit afin d'utiliser cette technique. Un niveau de confiance de 95% a été employé. Cette méthode est fondée sur la statistique:

$$R_{w\&w} = \sum_{i=1}^{n-1} x_i x_{i+1} + x_1 x_n \tag{3.3}$$

où x_i, i=1...n, sont les éléments de l'échantillon

Si les éléments de la série de données sont indépendants, $R_{w\&w}$ est normalement distribué avec une variance et une moyenne donnée par:

$$\overline{R}_{w\&w} = \frac{s_1^2 - s_2}{n-1} \tag{3.4}$$

$$Var(R_{w\&w}) = \frac{s_2^2 - s_4}{n-1} - \overline{R}_{w\&w}^2 + \frac{s_1^4 - 4s_1^2 s_2 + 4s_1 s_2 + s_2^2 - 2s_4}{(n-1)(n-2)} \tag{3.5}$$

où

$$s_r = nm'_r \tag{3.6}$$

$$m'_r = \frac{1}{n}\sum_{i=1}^{n} x_i^r \tag{3.7}$$

$$u = \frac{R_{w\&w} - \overline{R}_{w\&w}}{\sqrt{Var\left(R_{w\&w}\right)}} \tag{3.8}$$

Puisque u est normalement distribué avec une moyenne nulle et un écart-type unitaire, on peut effectuer un test d'inférence afin de vérifier l'hypothèse d'indépendance à un niveau de confiance α.

3.2 Analyse en composantes principales

L'analyse en composantes principales est une technique statistique qui permet de réduire le nombre de variables à celles qui sont les plus significatives parmi un ensemble de variables. Dans cette étude, l'analyse en composantes principales est utilisée dans le but de trouver un lien entre les stations et de les regrouper par régions homogènes du point de vue des débits d'étiage.

La méthode théorique de calcul des composantes est décrite en détail dans plusieurs ouvrages. Ceux qui ont été consultés pour ce travail sont les suivants: Beaudoin (1980), Haan (1977) et Yevjevich (1972). En bref, cette méthode construit de nouvelles variables non corrélées ou indépendantes à partir de variables corrélées. Ces nouvelles variables se nomment les composantes principales. Chacune d'elles représente une combinaison linéaire associée à un maximum de variance des variables observées. Une matrice de données contenant **N** observations et les **n** variables est définie. Les éléments de cette matrice sont représentés géométriquement par des points dans un espace à **N** dimensions. La première composante C1 correspond à la droite qui traverse l'ensemble des données dans la direction de la plus grande dispersion des points de façon à maximiser la variance. Ainsi, la première composante principale est la relation linéaire qui explique le mieux les variables et par le fait même les données utilisées. Si on l'appliquait à un ensemble de débits d'étiage, l'équation linéaire obtenue par composantes principales s'exprimerait de la façon suivante:

$$C1 = a_1X_1 + a_2X_2 + ... + a_nX_n \tag{3.9}$$

où X_1 : débits minimums de la station 010101

X_2 : débits minimums de la station 010801

X_n : débits minimums de la station 081101

a_1, a_2, a_n : coefficients

La deuxième composante principale passe également par la plus grande dispersion des points, mais est orthogonale à la première composante. Il en est de même pour toutes les autres composantes. Cette orthogonalité permet d'assurer que chacune des composantes demeure indépendante de celles déterminées précédemment. En conséquence, la variance totale expliquée par ces composantes, augmente très rapidement avec leur nombre. Généralement, on conservera un nombre suffisant de composantes pour atteindre une variance cumulée d'au moins 80%.

L'un des buts de l'analyse en composantes principales est de tirer l'information significative d'une matrice de données et, souvent, de fournir une représentation graphique des données afin de faciliter leur analyse. La procédure mathématique d'analyse en composantes principales est en fait une méthode statistique multivariée devant être adaptée aux données à traiter. Dans le cadre de ce travail, on a fait appel à la matrice de corrélation et les composantes ont été établies avec et sans rotation des axes orthogonaux. La méthode de rotation retenue, de type "equamax", permet d'obtenir des composantes principales à peu près équivalentes du point de vue du pourcentage de variance expliquée par la nouvelle variable. Tel que discuté ultérieurement, ceci s'avère un avantage certain lors du regroupement des observations (stations) selon les composantes retenues. Cet algorithme de rotation fait partie des fonctionnalités du logiciel Minitab™ (Minitab Inc., 1995) utilisé pour effectuer l'ACP.

3.3 Modèles de régression linéaires

Le modèle linéaire de régressions multiples s'exprime de la façon suivante:

$$Y = b_0 + b_1X_1 + b_2X_2 + \ldots + b_nX_n + \varepsilon \tag{3.10}$$

où Y : variable dépendante

X_1, X_2 et X_n : variables indépendantes

b_0 : terme constant

b_1, b_2 et b_n : coefficients de régression

ε : résidu

Le degré de dépendance linéaire entre la variable Y et les variables X_i est mesuré par le carré du coefficient de corrélation multiple R^2 qui s'exprime sous la forme suivante:

$$R^2 = \frac{SS_R}{S_{yy}} = \frac{\sum_{j=1}^{n}\left(\hat{y}_j - \bar{y}\right)^2}{\sum_{j=1}^{n}\left(y_j - \bar{y}\right)^2} \tag{3.11}$$

où SS_R est la variance attribuable au modèle et S_{YY} la variance associée à l'échantillon.

L'utilisation de modèles de régression linéaires peut mener à l'établissement de relations mathématiques qui, bien que décrivant adéquatement les données à ajuster, sont dénuées de sens physique. Généralement, on retiendra le modèle utilisant le moins de paramètres pour obtenir le degré d'ajustement désiré. Afin d'effectuer la sélection des variables significatives, une procédure en deux étapes est souhaitable. Premièrement, l'ensemble des variables géophysiques doit être défini et l'indépendance de celles-ci vérifiée sommairement à l'aide de la matrice de corrélation. Cette analyse initiale permet de réduire la quantité de variables à étudier en éliminant celles fortement corrélées (donc non-indépendantes). Deuxièmement, les variables restantes sont introduites une à une dans le modèle de régression linéaire. A chaque variable ajoutée, il faut s'assurer que le coefficient supplémentaire intégré au modèle est significatif. En fait, l'addition d'une variable aura toujours pour effet d'augmenter la somme des carrés de l'erreur expliquée par le modèle et donc de diminuer celle attribuable à la variance de l'échantillon. Le nombre de degrés de liberté est aussi réduit par l'ajout de chaque variable de régression.

La procédure retenue pour l'optimisation des modèles de regression est fondée sur l'approche proposée par Montgomery (Montgomery, 1984). L'équation (3.10) peut s'écrire sous forme vectorielle:

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon \tag{3.12}$$

et le vecteur de coefficients β peut être lui-même partitionné:

$$\beta = \left[\frac{\beta_1}{\beta_2}\right] \tag{3.13}$$

où β_1 est un vecteur de taille (r x 1), r étant le nombre de variables ajoutées au modèle et β_2 est de dimension {(p –r) x 1}. Afin de vérifier la pertinence des nouveaux paramètres, on veut effectuer le test d'hypothèses suivant:

$$H_0: \beta_1 = 0$$
$$H_1: \beta_1 \neq 0 \tag{3.14}$$

Pour le modèle complet, les carrés de l'erreur expliquée par le modèle de régression sont donnés, pour un échantillon de taille n , par:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \tag{3.15}$$

$$SS_R(\beta) = \hat{\beta}'\mathbf{X}'\mathbf{y} \quad (p \text{ degrés de liberté}) \tag{3.16}$$

$$MS_E = \frac{\sum_{j=1}^{n} y_j^2 - \hat{\beta}'\mathbf{X}'\mathbf{y}}{n-p} \tag{3.17}$$

Afin d'isoler la contribution du ou des nouveaux paramètres, il faut connaître l'erreur de régression pour le modèle original (sans β_1), soit

$$\hat{\beta}_2 = \left(\mathbf{X}'_2\mathbf{X}_2\right)^{-1}\mathbf{X}'_2\mathbf{y} \tag{3.18}$$

$$SS_R(\beta_2) = \hat{\beta}'_2\mathbf{X}'_2\mathbf{y} \quad \text{(p-r degrés de liberté)} \tag{3.19}$$

et donc la somme des carrés de l'erreur attribuable à β_1 en présence de β_2 est donnée par

$$SS_R(\beta_1|\beta_2) = SS_R(\beta) - SS_R(\beta_2) \quad \text{(r degré de libertés)} \tag{3.20}$$

Puisque $SS_R(\beta_1|\beta_2)$ est indépendant de MS_E, l'hypothèse nulle β_1=0 peut être vérifiée par la statistique:

$$F_0 = \frac{SS_R(\beta_1|\beta_2)/r}{MS_E} \tag{3.21}$$

Si $F_0 > F_{\alpha, r, n-p'}$, obtenu de la distribution de Fisher à un niveau de confiance α de 95%,on rejette H_0 et on conclut qu'au moins un des paramètres de β_1 est non-nul.

Cette procédure se généralise pour plusieurs paramètres par:

$$SS_R(\beta_i|\beta_0,\beta_0,\ldots,\beta_{i-1},\beta_{i+1},\ldots,\beta_k) \tag{3.22}$$

3.4 Distributions statistiques

L'utilisation de modèles de régression permet d'établir, à des fins de prévision, des relations linéaires entre les paramètres physiques d'un bassin et des variables hydrologiques telles 7Q2, 7Q10 ou 30Q5. Cependant, ces modèles de régression demeurent spécifiques à un ensemble limité de variables. Chaque variable dépendante doit être associée à un groupe de coefficients β_i différents. Puisque la variable

hydrologique d'intérêt à une situation donnée peut varier considérablement (i.e. 7Q50, 7Q100, 1Q5, etc.), l'approche par modèles de régression est peu pratique si elle n'est pas généralisée à l'aide de fonctions de distribution statistiques ajustées aux données expérimentales normalisées. La procédure de normalisation vise d'abord à transformer les débits expérimentaux en quantités adimensionnelles de manière à obtenir une courbe de distribution unique pour tous les bassins pour une période d'observation donnée (1 jour, 7 jours et 30 jours). Par la suite, on vérifiera si cette normalisation peut être effectuée de façon telle que les courbes de distribution des différentes périodes d'observation soient confondues. Dans un tel cas, la distribution ajustée pourrait être applicable à toute période d'observation comprise entre un et trente jours. Le modèle ainsi obtenu serait particulièrement versatile puisqu'il permettrait d'estimer le débit d'étiage d'un bassin quelconque uniquement à partir des données physiques du site.

La transformation des données brutes en valeurs adimensionnelles sera effectuée selon deux méthodes différentes. La première méthode consiste à diviser chaque donnée annuelle (1 jour, 7 jours et 30 jours) d'une station donnée, par la moyenne globale à cette station pour ces variables:

Méthode de standardisation #1:

$$x' = \frac{x_{station}}{\bar{x}_{station}} \tag{3.23}$$

Cette méthode est particulièrement avantageuse puisqu'elle permet d'évaluer un débit d'étiage quelconque avec un seul paramètre de normalisation ($\bar{x}_{station}$) pour autant que la fonction de distribution de x' des variables Min_{1j}, Min_{7j} et Min_{30j} soit commune à toutes les stations. Il sera montré que $\bar{x}_{station}$ peut être évalué directement à partir du modèle de régression et de la fonction de distribution de la variable normalisée $f(x')$.

D'autre part, compte tenu de la nature même des variables Min_{1j}, Min_{7j} et Min_{30j}, il est vraisemblable que les variances de leur population soient différentes entre elles. Aussi, une seconde méthode de normalisation sera évaluée. La méthode de normalisation #2 est une technique de standardisation conventionnelle permettant l'obtention d'une distribution de moyenne nulle avec écart-type unitaire pour un échantillon normalement distribué:

Méthode de standardisation #2

$$x' = \frac{(x_{station} - \bar{x}_{station})}{\sigma_{x,station}} \tag{3.24}$$

L'inspection visuelle des courbes de distribution obtenues pour chacune des méthodes de normalisation ne peut suffir à justifier la fusion des trois variables standardisées en une seule fonction de distribution. Aussi, cette décision sera basée sur l'analyse, point à point, de l'écart entre les trois courbes de distribution et sur un test statistique non-paramétrique, soit le test de Smirnov.

Dans le cas de l'analyse fondée sur la statistique de Smirnov, les tests doivent être effectués par comparaison pairée de chacun des trois couples possibles (Min_{1j} vs Min_{7j}, Min_{1j} vs Min_{30j} et Min_{7j} vs Min_{30j}). Le test non-paramétrique de Smirnov cherche à vérifier l'hypothèse suivante (Conover 1980):

H_0: $F(x)=G(x)$ pour tout x
H_1: $F(x)\neq G(x)$ pour au moins un x

Soient deux fonctions de distribution empiriques, $S_1(x)$ et $S_2(x)$, basées sur des échantillons respectifs $X_1, X_2, ...X_n$ et $Y_1, Y_2, ... Y_m$. La statistique de Smirnov, T_1, est définie par la plus grande distance verticale entre les deux distributions:

$$T_1 = \sup_x |S_1(x) - S_2(x)| \tag{3.25}$$

La règle de décision est telle qu'on rejettera l'hypothèse H_0 au niveau de confiance α si:

$$T_1 >\approx \frac{k}{\sqrt{n}}, \; k = f(1-\alpha), \; pour\, n > 20 \tag{3.26}$$

où k=2.3 pour α=0.99

Une fois la distribution expérimentale normalisée établie, il est possible de procéder à l'ajustement d'une fonction de distribution théorique à l'ensemble de données. Les paramètres de ces distributions sont généralement obtenus par la méthode des moments ou la méthode du maximum de vraisemblance. La première technique est mathématiquement plus simple d'utilisation mais s'avère souvent moins efficace du point de vue de la qualité de l'ajustement. Par ailleurs, la méthode du maximum de vraisemblance est maintenant la plus employée compte tenu des facilités de calcul contemporaines. Soit une fonction de distribution de probabilité donnée par:

$$p(y) = f(y, \theta) \tag{3.27}$$

où θ est un vecteur contenant quelques paramètres. Si chacune des observations y_t sont statistiquement indépendantes, la probabilité d'obtenir une observation y_t est:

$$f(y_1, \theta) \cdot f(y_2, \theta) \cdot .... \cdot f(y_N, \theta) dy_1 \cdot dy_2 \cdot .. dy_N \tag{3.28}$$

pour autant que y_1, y_2, ...y_N soit un échantillon aléatoire d'une population. Pour une observation y_t, la fonction de vraisemblance s'exprime par:

$$L(\theta; y_1, \ldots, y_N) = \prod_{t=1}^{N} f(y, \theta) \tag{3.29}$$

Si on prend le logarithme de cette fonction, on obtient:

$$l(\theta) = \sum_{t=1}^{N} \ln f(y, \theta) \tag{3.30}$$

Généralement, le vecteur optimal θ, correspondant au maximum de la fonction de vrainsemblance, est identifié par différentiation et résolution du système d'équations:

$$\frac{\partial l}{\partial \theta} = \frac{\partial \ln L}{\partial \theta} = L^{-1} \frac{\partial L}{\partial \theta} = 0 \tag{3.31}$$

La résolution du système nécessite parfois l'utilisation de méthodes numériques itératives. Dans un tel cas, les estimés de départ peuvent être obtenus par la méthode des moments (Clark 1994). Bien que plusieurs logiciels commerciaux ou partagiciels permettent la résolution de ce type de problème, des routines particulières, avec codes sources en FORTRAN ont été écrites pour la réalisation de ce travail. La performance de ces routines a été établie à partir d'ensemble de données avec distribution connue (Haan 1977).

L'identification de la fonction de distribution, associée à la distribution expérimentale, sera effectuée par la comparaison de l'ajustement obtenu pour quatre fonctions différentes. La première de ces distributions est la distribution de Weibull, reconnue pour être particulièrement adaptée aux minima tels les débits d'étiage. La fonction de densité de probabilité est donnée par:

$$p(y)=\left(\frac{k}{\alpha}\right)\left(\frac{y}{\alpha}\right)^{k-1}\exp\left(-\left(\frac{y}{\alpha}\right)^{k}\right),\ y\geq 0\ ;\ \alpha,k\geq 0 \tag{3.32}$$

La fonction cumulative de probabilité ou fonction de répartition s'obtient par intégration de (3.32):

$$P(y)=1-\exp\left(-\left(\frac{y}{\alpha}\right)^{k}\right) \tag{3.33}$$

Enfin, lorsqu'on applique la relation (3.30) à la fonction de densité de probabilité, il est possible d'évaluer la fonction de vraisemblance par:

$$l(k,\alpha)=N\ln(k/\alpha)+(k-1)\ln(y/\alpha)-\sum_{t=1}^{N}\left(\frac{y}{\alpha}\right)^{k} \tag{3.34}$$

Les paramètres optimum k et α s'obtiennent par différentiation de (3.34) tel que décrit par l'équation (3.31) et par la résolution du système d'équations à l'aide de la méthode de Newton-Raphson.

La seconde distribution étudiée est la distribution gamma avec deux paramètres ajustables, soit k et u. Les fonctions de densité de probabilité et de répartition de la loi Gamma sont données par les équations (3.35) et (3.36) respectivement:

$$p(y)=\left(\frac{1}{u}\right)^{k}y^{k-1}\frac{\exp\left(-\frac{y}{u}\right)}{\Gamma(k)} \tag{3.35}$$

$$P(y)=\int_0^y\left(\frac{1}{u}\right)^k y^{k-1}\frac{\exp\left(-\frac{y}{u}\right)}{\Gamma(k)}dy \tag{3.36}$$

La fonction de vraisemblance s'obtient par les relations (3.30) et (3.35):

$$l(u,k)=Nk\ln(k/u)+(k-1)\sum\ln y_t-(k/u)\sum y_t-N\ln\Gamma(k) \tag{3.37}$$

Et les paramètres optimaux sont évalués par différentiation de (3.37) d'où (3.38) et (3.39):

$$u=\bar{y} \tag{3.38}$$

$$\ln(k)-\frac{\partial\ln\Gamma(k)}{\partial k}=\ln(\bar{y})-\overline{\ln(y)} \tag{3.39}$$

La distribution log-normale est la troisième distribution considérée afin d'ajuster les données de débits minimums. Cette fonction comprenant deux paramètres ajustables est exprimée par:

$$p(y,u,\sigma^2)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(y-u)^2}{2\sigma}\right),\quad y=\ln x \tag{3.40}$$

La fonction de répartition est simplement l'intégrale de la relation (3.40) alors que la fonction de vraisemblance est donnée par:

$$l(u,\sigma^2)=-N\ln\sigma-\frac{N}{2}\ln 2\pi-\sum\ln y_t-\sum\frac{(\ln y-u)^2}{2\sigma^2} \tag{3.41}$$

Dans le cas de cette distribution, les paramètres optimaux associés au maximum de vraisemblance s'obtiennent analytiquement par:

$$u = \sum \frac{\ln y_i}{N} \tag{3.42}$$

$$\sigma^2 = \frac{\sum (\ln y - u)^2}{N} \tag{3.43}$$

La dernière distribution considérée est celle de Pearson III. Cette distribution est en fait une variante de la distribution de gamma à deux paramètres dans laquelle on effectue une transformation des données par:

$$y' = y - a \tag{3.44}$$

Et donc la fonction de densité de probabilité devient:

$$p(y) = \left(\frac{1}{u}\right)^k (y-a)^{k-1} \frac{\exp\left(-\frac{(y-a)}{u}\right)}{\Gamma(k)} \tag{3.45}$$

La détermination des paramètres k,u et a s'effectue de la même façon que dans le cas de la fonction gamma après la transformation des données selon la relation (3.44) .

3.5 Calcul des débits d'étiage

Le calcul des débits d'étiage pour une période de retour *T*, suite à la procédure de modélisation et de régionalisation effectuée, fera appel à la fonction de densité de probabilité sélectionnée et aux modèles régionaux obtenus par régression linéaire. La figure 3.2 montre un arbre de décision et d'opération qui résume les étapes du calcul (ici,

on assume que la méthode de normalisation #1, équation (3.23) est utilisée). Essentiellement, la stratégie consiste à déterminer le facteur de normalisation, propre à une station donnée ($\bar{x}$), à partir des modèles de régression linéaires. S'il s'agit d'une station non jaugée, on estime d'abord une variable hydrologique telle que 7Q10 en fonction des caractéristiques physiques (après transformation linéaire en composantes principales) d'une station donnée. Dans le cas d'une station jaugée, 7Q10 peut aussi être évalué directement des observations historiques. Connaissant la distribution théorique ajustée, on peut trouver le x correspondant à une probabilité de 1/10 (période de retour de 10 ans) et donc déterminer le facteur de normalisation $\bar{x}$. Par la suite, il suffit d'utiliser la même distribution pour estimer x pour une probabilité de 1/T et finalement de calculer 7QT à l'aide du facteur de normalisation déjà estimé. Donc, si une fonction de densité de probabilité théorique est disponible, il est possible de calculer les variables 1QT, 7QT ou 30QT à n'importe laquelle des stations, jaugée ou non, à l'intérieur des régions géographiques identifiées.

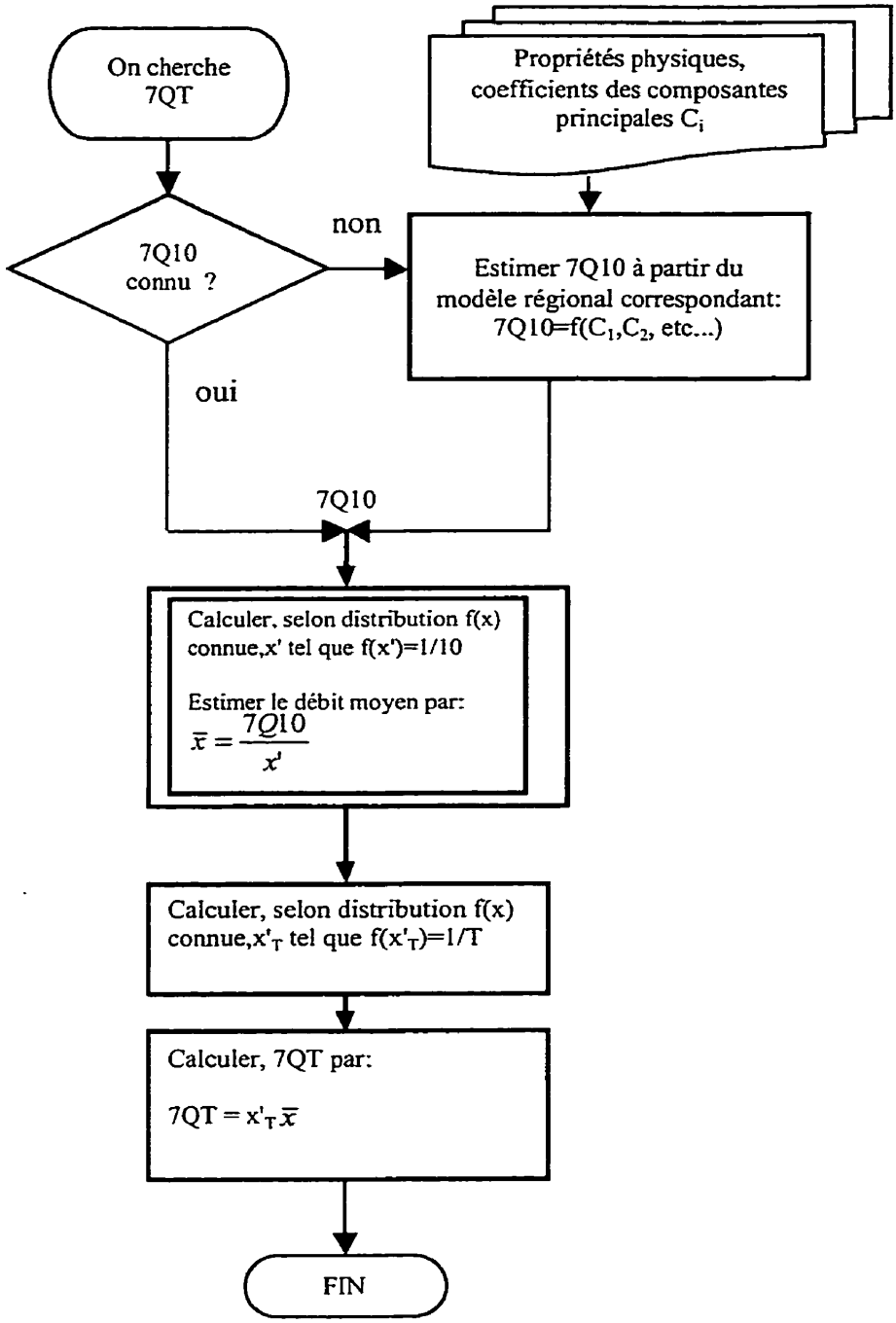

Figure 3.2: Calcul d'une variable hydrologique de type 7QT à une station "S"

CHAPITRE 4

ÉTUDE DES DÉBITS D'ÉTIAGE

4.1 Description des données

Les données utilisées proviennent de huit régions hydrographiques de la province de Québec. Les régions situées plus au nord (région de la Baie d'Hudson - 9 et région de la Baie d'Ungava-10) ne sont pas considérées dans la présente étude puisqu'elles ne cadrent pas avec les objectifs du projet de recherche. En effet, ces territoires renferment peu d'activités économiques pouvant être affectés par la variation annuelle des débits d'étiage.

Les 119 stations de débit ont été sélectionnées selon deux critères. D'abord, leur cours d'eau ne sont pas régularisés, donc leur écoulement est naturel et deuxièmement leur nombre d'années d'enregistrement est égal ou supérieur à 15. La répartition du nombre d'années des stations est illustrée à la figure 4.1. Il est à noter que, selon la planification initiale de ce projet de recherche, seules les stations couvrant un bassin inférieur à 100 km^2 devaient être étudiées. Par contre, suite à l'application des deux critères cités ci-haut, il n'aurait resté que quelques dizaines de stations. Donc, ce critère de superficie n'a pas été retenu afin d'obtenir un ensemble de données de taille suffisante.

La liste des stations est présentée à l'annexe A. Elle comprend le nom, le numéro des stations, la superficie du bassin versant en km^2, la période d'observation et le nombre d'années d'enregistrement. Il faut souligner que la liste comprend des stations fermées, malgré que la majorité d'entre elles soient encore actives.

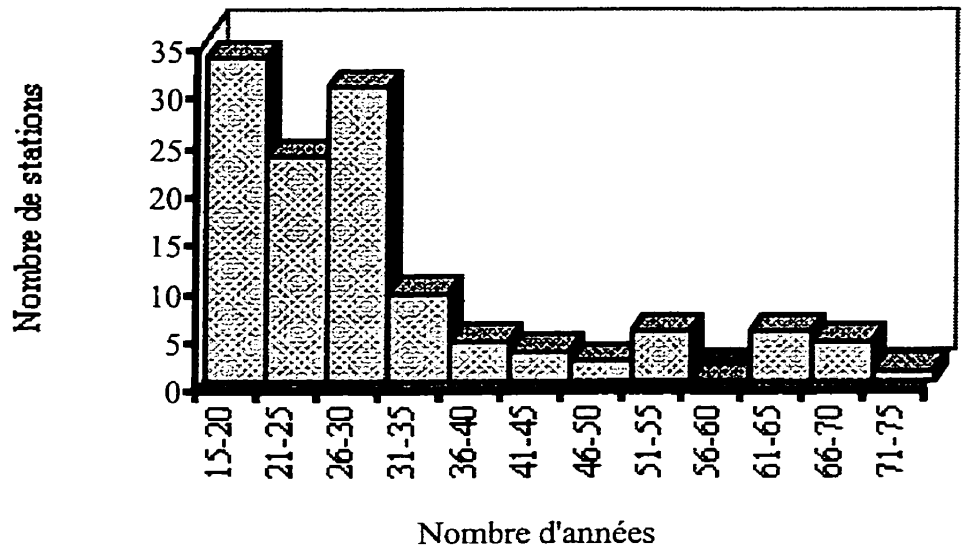

Figure 4.1: Répartition du nombre d'années d'enregistrement des stations

Les données utilisées sont des débits minimums annuels. Ces débits d'étiage ont été tirés des valeurs journalières. Les débits sont également classés selon la période hivernale (1er janvier au 31 mai), la période estivale (1er juin au 31 octobre) et la période annuelle (1er janvier au 31 décembre) . Les débits minimums annuels 7 et 30 jours consécutifs sont aussi disponibles. Il est à noter que toutes les données utilisées dans cette étude proviennent du ministère de l'Environnement du Québec.

Le graphique ci-dessous (figure 4.2) montre la répartition des périodes d'étiage par région et selon la période estivale, hivernale et autre. Il est à noter que la période autre sur le graphique représente les mois de novembre et décembre seulement. Les années sont classées selon la période à laquelle se sont produits les étiages. L'analyse des

données montre que les régions 2, 3 et 4 ont un débit minimum qui se produit, dans la majorité des cas, durant la période estivale. Ces régions sont illustrées à la figure 4.3. Il en est autrement pour les territoires 1,5,6,7 et 8 qui sont localisés plus au nord et plus à l'est dans le cas du Bas Saint-Laurent et de la Gaspésie (région 01)

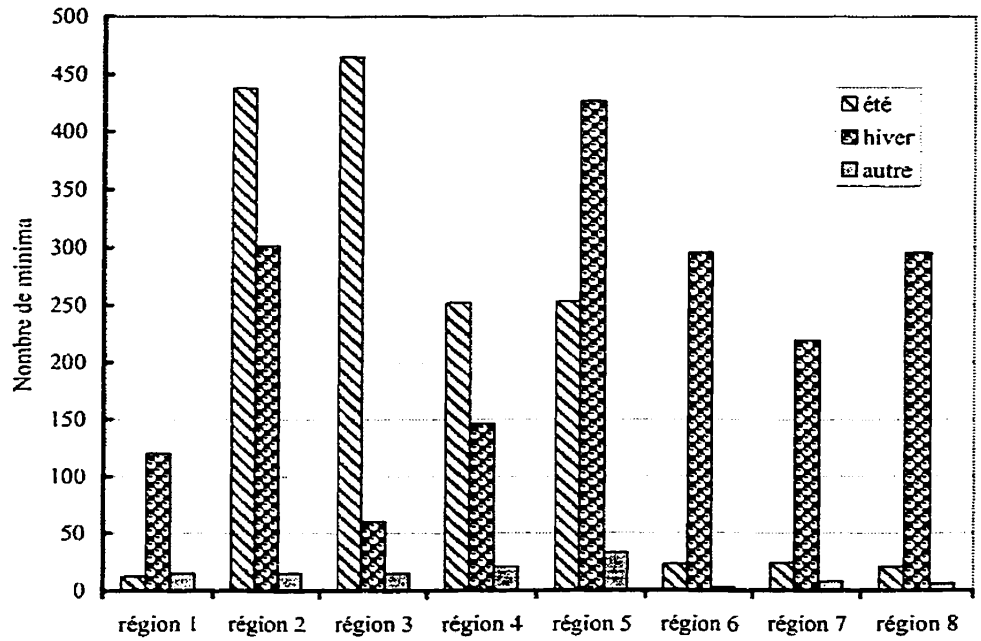

Figure 4.2: Distribution des minima annuels selon la période d'étiage.

Le tableau à l'annexe B présente les statistiques des débits minimums pour chaque station. L'ordre de grandeur des débits varie beaucoup entre les régions. La région 08, située au nord-ouest du Québec, possède de grands réservoirs tandis que les régions 02 et 03, situées au sud, possèdent de plus petits cours d'eau. Le maximum des débits minimums est de 502 m^3/s (station 081002) alors que le minimum est de 0,002 m^3/s pour la station 030420. De plus, les plus grands coefficients de variation sont observés dans la région 03.

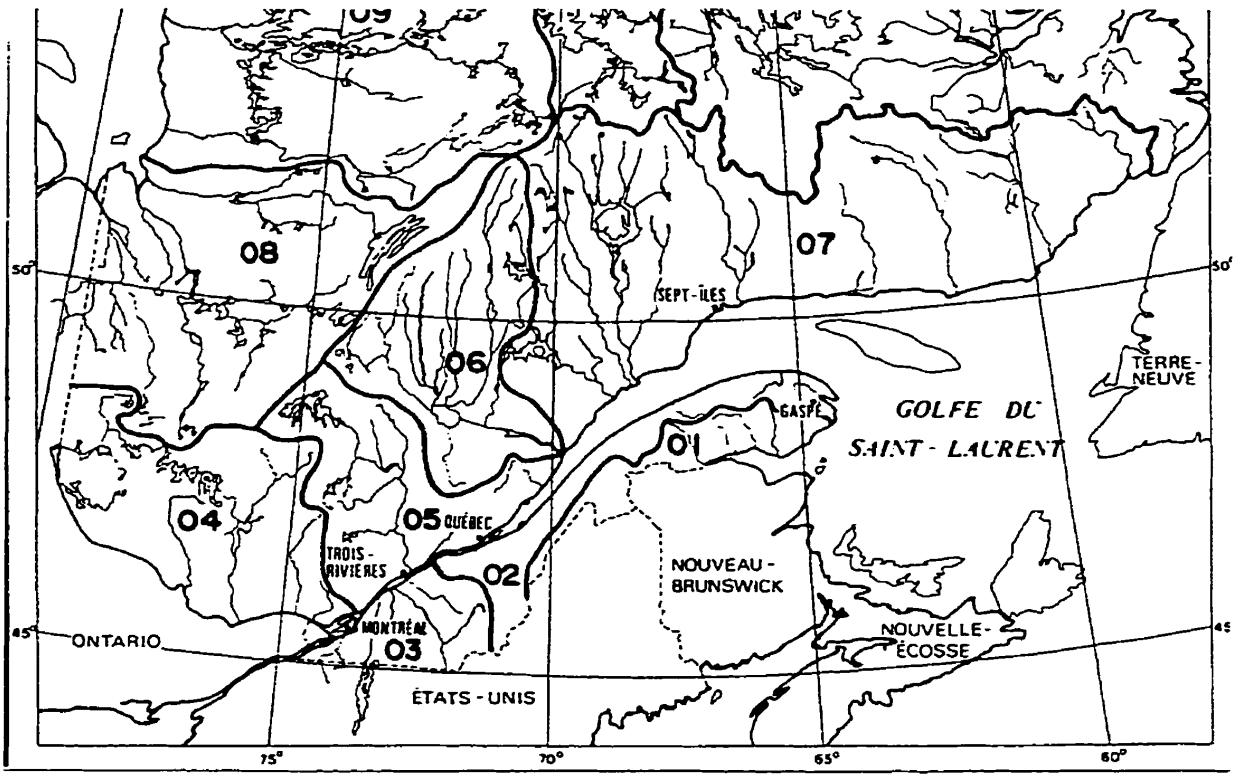

Figure 4.3: Régions hydrographiques définies par le MEQ.

4.2 Caractérisation statistique

4.2.1 Homogénéité des séries

Un test d'homogénéité, tel que décrit à la section 3.1.1 a été effectué sur l'ensemble des stations sélectionnées dans le but de faire ressortir les stations ayant des statistiques très différentes. Cependant, selon Rousselle *et al.* (1990), cette méthode suppose qu'on ait

identifié préalablement des régions homogènes au point de vue de leurs caractéristiques physiographiques, de leur climat, etc. Par la suite, le test est appliqué à chacune de ces régions et les stations ayant des écarts statistiques trop grands, par rapport à ceux de la région, sont éliminées.

Pour les fins d'application de ce test, les régions hydrographiques déjà définies par le ministère de l'Environnement du Québec (figure 4.3), ont été utilisées avec un niveau de confiance de 95%. Le tableau 4.1 contient la liste des 17 stations qui possèdent une tendance statistique différente des autres stations alors que les figures 4.4 à 4.11 illustrent les résultats du test région par région. L'analyse d'homogénéité n'a pas nécessairement pour but d'éliminer des stations pour les étapes ultérieures du processus d'étude des débits d'étiage. Il s'agit plutôt d'un outil permettant d'évaluer le degré d'homogénéité des régions actuelles et donc de justifier, si besoin est, une réorganisation des stations selon un nouvel ordre régional. Ainsi, les résultats du tableau 4.1 montrent que les régions 03 et 05 contiennent un nombre élevé de stations qui ne passent pas le test d'homogénéité. Ainsi, on peut donc assumer que ces stations pourraient être regroupées à d'autres régions géographiquement voisines afin de former des sous-ensembles plus homogènes. Ce processus de "régionalisation" est l'objet de la section 4.3.

Tableau 4.1: Stations ne satisfaisant pas le test d'homogénéité

Région 01	**Région 02**	**Région 03**	**Région 04**	**Région 05**	**Région 06**
010801	020802	030234	043012	050304	061905
011003	021502	030282		050408	061906
	021601	030401		050409	
		031401		051301	
				051502	

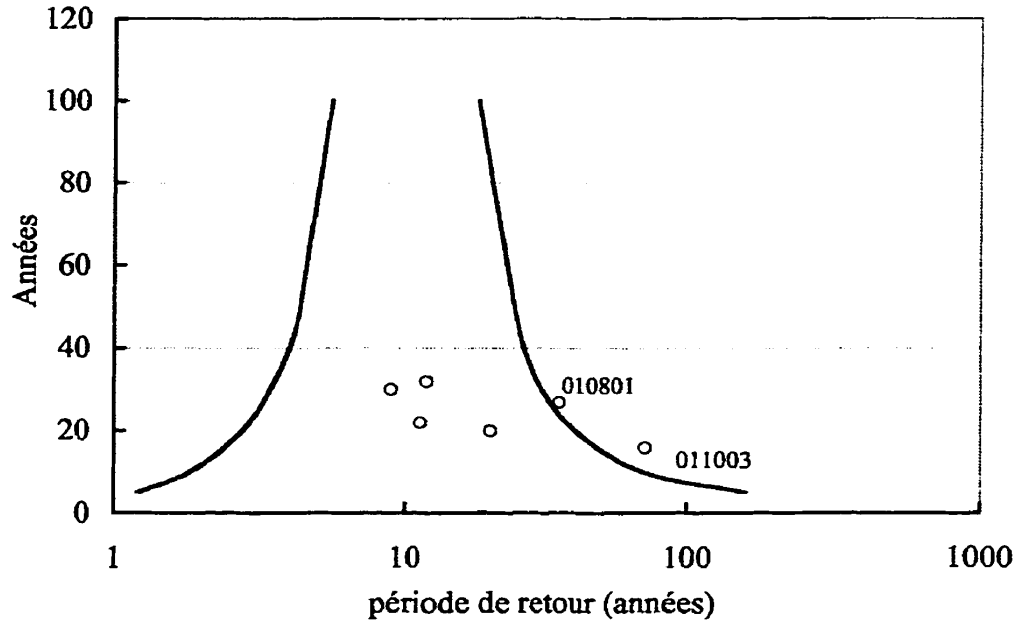

Figure 4.4: Test d'homogénéité-région 01

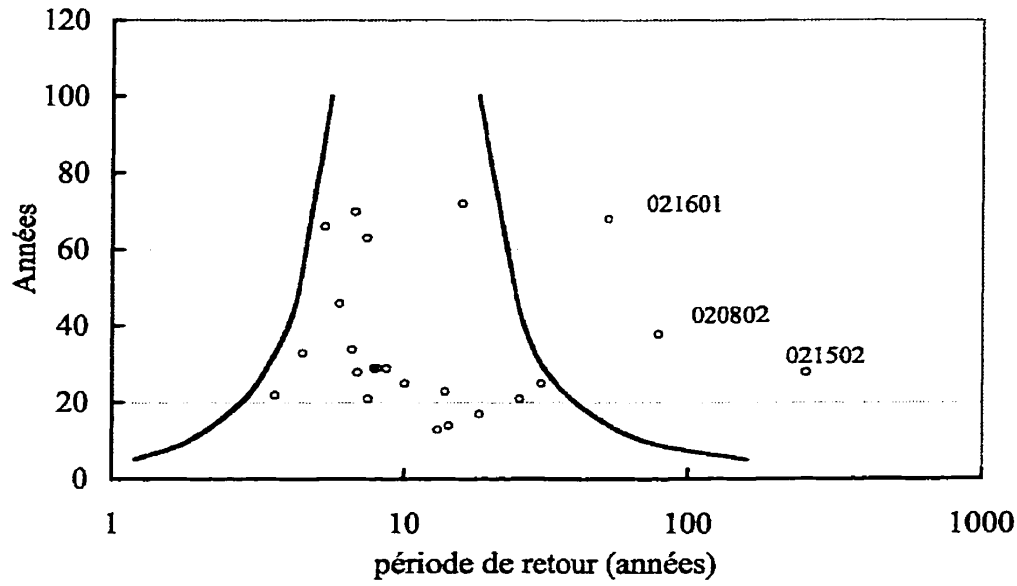

Figure 4.5: Test d'homogénéité-région 02

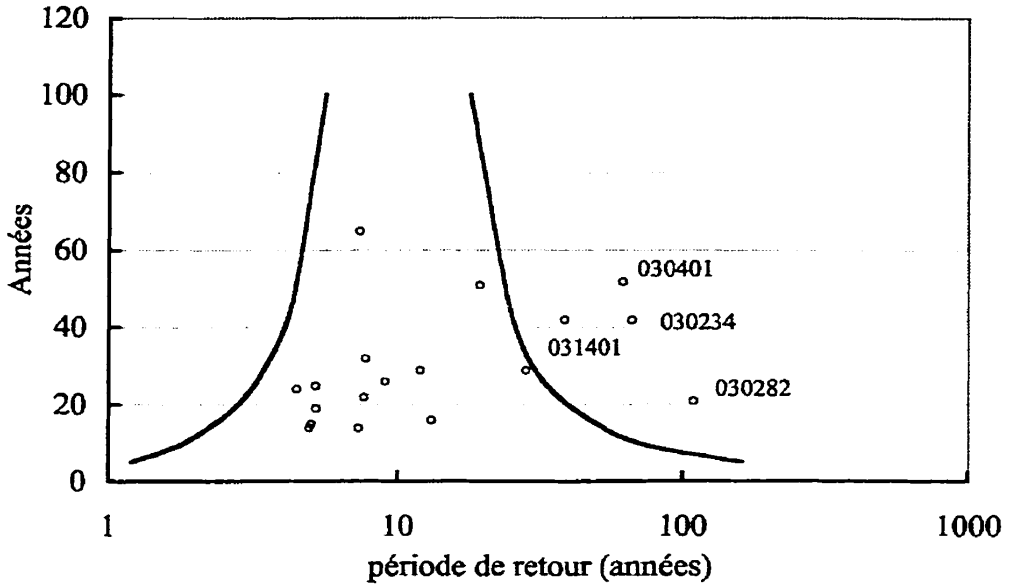

Figure 4.6: Test d'homogénéité-région 03

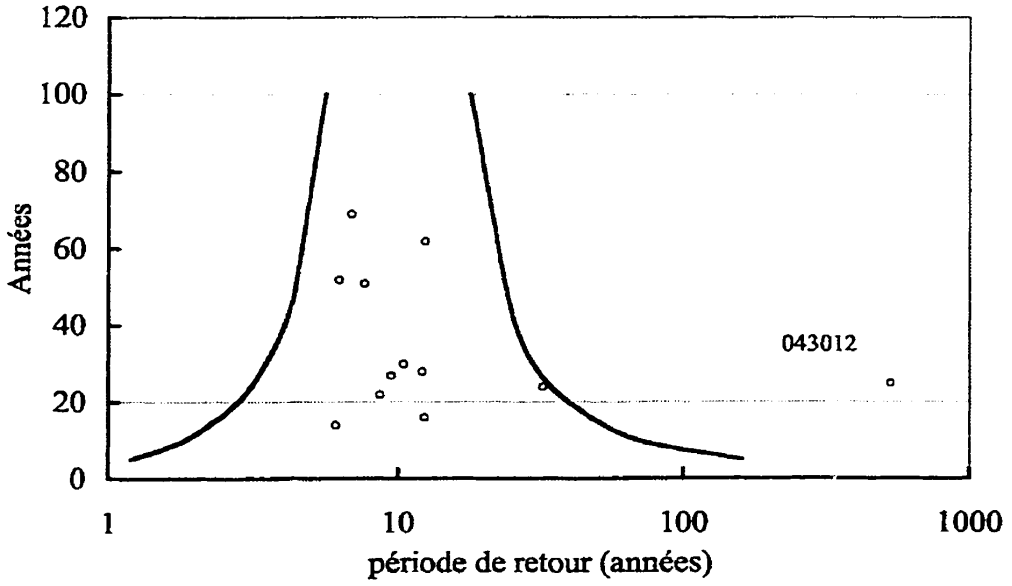

Figure 4.7: Test d'homogénéité-région 04

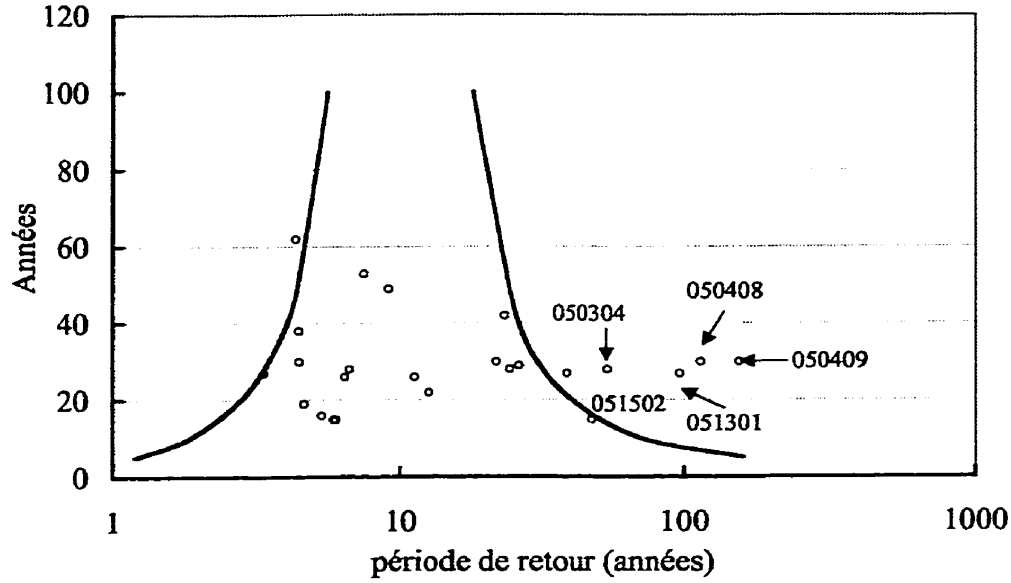

Figure 4.8: Test d'homogénéité-région 05

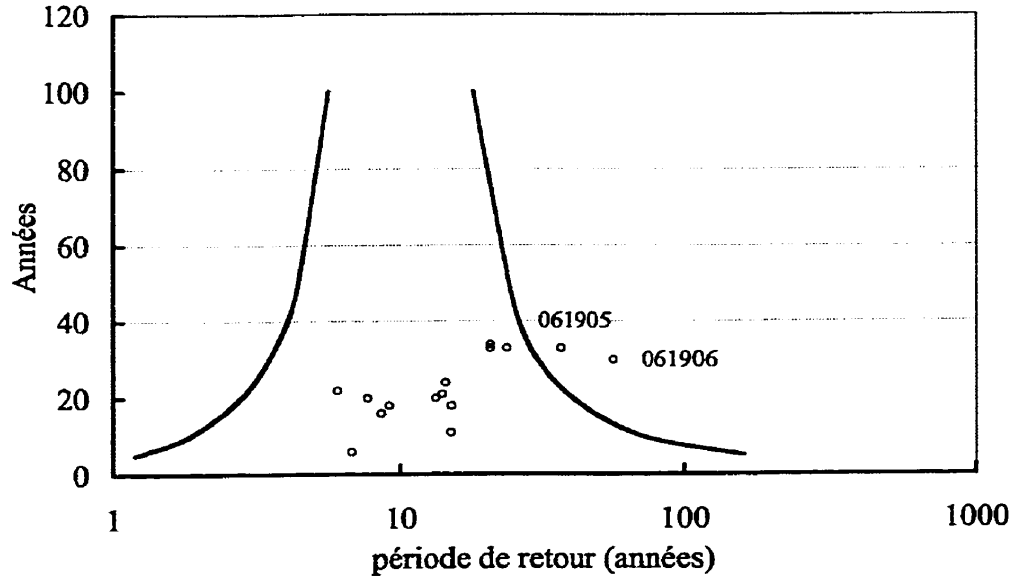

Figure 4.9: Test d'homogénéité-région 06

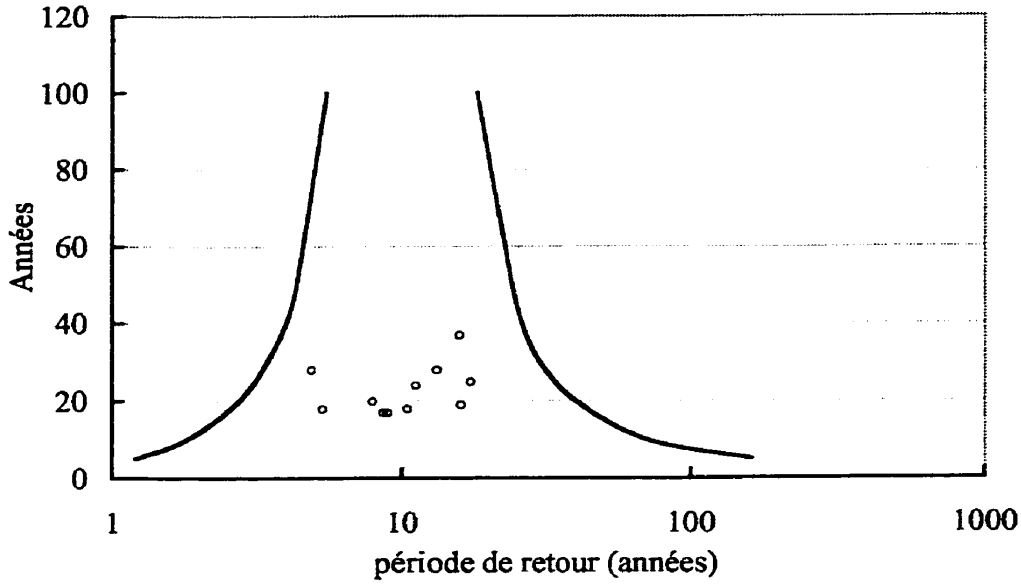

Figure 4.10: Test d'homogénéité-région 07

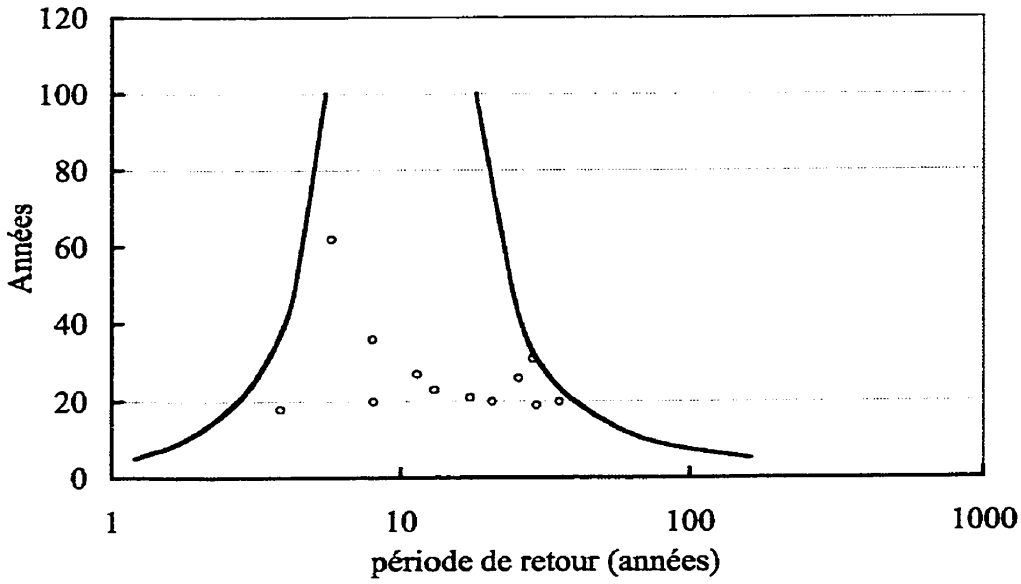

Figure 4.11: Test d'homogénéité-région 08

4.2.2 Test de stationnarité

La stationnarité signifie qu'une série de données est invariable dans le temps, excluant les fluctuations des événements extrêmes (Durrans et Tomic, 1996). Il existe trois types de non-stationnarité: les tendances, les sauts et les cycles. Ces irrégularités ont été décelées suite à une inspection visuelle des séries de données.

Aussi, l'inspection visuelle permet de déceler divers types d'erreurs. Celles-ci peuvent être causées par une mauvaise lecture des instruments, un malfonctionnement ou un mauvais positionnement des instruments et par la manipulation des données (transfert d'unité pi^3/s en m^3/s, erreur de retranscription, etc).

Le tableau 4.2 contient les stations qui possèdent des séries non stationnaires et les graphiques de ces mêmes séries irrégulières sont présentés à l'annexe C.

Tableau 4.2: Stations ayant des séries non stationnaires

Région 01	**Région 02**	**Région 03**	**Région 05**	**Région 07**
010901	021915	030215	050118	072301
	022504	030242	051006	076601
		030309		
		030401		

4.2.3 Indépendance des observations

En utilisant la méthode décrite à la section 3.1.2, l'indépendance des débits annuels a été vérifiée pour chacune des 119 stations retenues au départ. Puisque u est normalement distribué avec une moyenne nulle et un écart-type unitaire, on peut effectuer un test d'inférence afin de vérifier l'hypothèse d'indépendance à un niveau de confiance α.. Bien

que les valeurs utilisées soient généralement des minima annuels aléatoires, certaines séries présentent une dépendance entre leurs données. Le tableau 4.3 indique les 10 stations associées à ces séries.

Tableau 4.3: Stations dépendantes (α=5%)

Région 02	**Région 03**	**Région 04**	**Région 05**	**Région 07**
022504	030242	040202	050120	076601
	030309		051001	
	030335		051006	
			052601	

4.2.4 Sélection définitive des stations

Suite à l'analyse des données disponibles, un certain nombre de stations qui ne satisfont pas aux critères d'homogénéité, de stationnarité et d'indépendance ont été identifiées. Le sommaire de ces stations se résume au tableau 4.4 pour chacune des régions.

Des 38 stations identifiées, six ne répondent pas aux critères d'au moins deux tests différents. Lorsqu'on élimine cette redondance, il reste encore 32 stations à rejeter. Des 119 stations initiales, il en resterait 87 pour fin d'études. Ce nombre est insuffisant pour faire des analyses. Il a donc été décidé de conserver les 17 stations ne passant pas le test d'homogénéité puisque de toute façon de nouvelles régions homogènes seront identifiées à une étape ultérieure du projet. En d'autres termes, des 119 stations initiales, on doit soustraire les 11 stations du test de stationnarité et les 10 stations du test d'indépendance Il en résulte ainsi 98 stations. Cependant, 5 stations ont été rejetées deux fois, c'est-à-dire qu'elles ont été excluent à la fois par le test de stationnarité et d'indépendance. Ces stations sont identifiées dans la note du tableau 4.4. En conséquence, les travaux de modélisation qui suivent furent effectués sur un total de 103 stations.

Tableau 4.4: Compilation des stations suite aux tests

Région	nb. total (avant tests)	Homogénéité (par région)	Stationnarité	Indépendance	Stations redondantes*	nb.total (après tests)
01	6	2	1	-	-	3
02	23	3	2	1	1	18
03	18	4	4	3	3	10
04	12	1	-	1	-	10
05	24	5	2	4	1	14
06	13	2	-	-	-	11
07	11	-	2	1	1	9
08	12	-	-	-	-	12
Total	119	17	11	10	6	87

*Stations redondantes: stations que l'on retrouve dans deux tests:
Région 02: 022504 (Stationnarité et Indépendance)
Région 03: 030401 (Homogénéité et Stationnarité)
030401, 030309 (Stationnarité et Indépendance)
Région 05: 051006 (Stationnarité et Indépendance)
Région 07: 076601 (Stationnarité et Indépendance)

4.3 Régionalisation

Le processus de régionalisation vise à assembler les différentes stations d'échantillonnage de débits en des groupes ayant les mêmes comportements hydrologiques. D'un point de vue statistique, ce regroupement peut s'effectuer par l'étude de la variance entre les stations d'un certain nombre de variables les caractérisant. C'est l'objet de la méthode d'analyse en composantes principales présentée à la section 3.2. L'analyse en composantes principales peut s'effectuer selon deux approches. Selon la première

méthode, les stations sont les **n** variables utilisées alors que les N observations correspondent à une série de débits minimums annuels pour chacune de ces stations. C'est l'approche déjà utilisée par Morin *et al.* (1979). Pour être concluants, les résultats de l'analyse doivent cerner un nombre limité de composantes principales, chaque composante étant obtenue par contribution linéaire des données de quelques stations. L'examen des coefficients de contribution de ces stations obtenus par l'analyse multivariée permet ensuite de les regrouper par région. Malheureusement, cette approche demande généralement un ensemble de données consécutives sur plusieurs décades pour un large éventail de stations. Dans le cadre de cette étude, la base de données disponibles était trop incomplète pour effectuer cette analyse. Ainsi, l'analyse par composantes principales appliquée aux débits minimums n'a pas été concluante. Chacune des stations ont leur série de données situées dans différentes périodes dans le temps. L'intersection de toutes les séries donne un ensemble de 13 années consécutives et ce pour 21 stations. Si on augmente le nombre de stations, le nombre d'années diminue, d'où l'impossibilité d'appliquer la méthode efficacement.

La deuxième méthode d'analyse par composantes principales peut aussi s'appliquer à des données comme les caractéristiques physiographiques des bassins. Dans ce cas, la matrice des valeurs est continue puisque chaque station possède une valeur pour chaque variable. Cette deuxième approche consiste donc à appliquer l'analyse en composantes principales sur les variables physiques affectant le débit d'étiage (Nathan et McMahon 1990). Les observations pour ces huit variables sont disponibles pour chacune des 103 stations non préalablement rejetée par l'analyse statistique préliminaire. Les résultats de l'ACP permettent de regrouper ces variables physiques en un ensemble réduit de nouvelles variables parfaitement indépendantes. Une méthode graphique permet ensuite de regrouper les stations en régions, sur la base de la valeur des composantes principales pour chacune d'entre elles.

4.3.1 Description des caractéristiques physiographiques

L'analyse en composantes principales est effectuée sur un ensemble de variables physiographiques explicatives du comportement hydrologique d'un bassin. Les facteurs qui sont les plus susceptibles d'influencer les débits d'étiage sont la grandeur de la superficie de drainage, le relief du bassin, le climat et l'utilisation du sol. Selon Tessier (1983) et Boucher *et al.* (1986), ces variables expliquent convenablement le comportement hydrologique des bassins. A partir de ces facteurs, les caractéristiques physiques et climatiques suivantes ont été étudiées: la superficie et la longueur du cours d'eau, la pente, la précipitation, le pourcentage de la superficie couverte par la forêt, les lacs et les marais, la latitude et la longitude. Les données utilisées proviennent principalement de l'étude des débits réservés pour la protection des habitats de poissons dans les rivières du Québec (Belzile *et al.* 1997). Les données manquantes, soient pour une dizaine de stations, ont été mesurées de la même façon et avec les mêmes outils que celles de l'étude citée précédemment. Les commentaires explicatifs suivants s'appliquent:

Superficie (km^2) :

La superficie est une information publiée dans les annuaires hydrologiques du ministère de l'Environnement du Québec.

Longueur du cours d'eau (km):

La longueur du cours d'eau se définie généralement comme étant la distance entre l'embouchure du cours d'eau et le point le plus éloigné à la limite du bassin. Cette droite fictive suit l'axe longitudinal du bassin c'est-à-dire qu'elle suit le cours d'eau principal. Cette distance a été mesurée sur les cartes topographiques à l'échelle 1:50 000.

Pente moyenne (%):

Dans l'étude de Belzile *et al.* (1997), la pente moyenne a été calculée selon la méthode préconisée par Benson (1959). Selon cet auteur, la pente moyenne d'un bassin est celle calculée entre les points situés à 0,85L et 0,1L où L est la longueur totale du cours d'eau expliquée précédemment.

Précipitations (mm):

Les précipitations totales moyennes annuelles sont tirées d'un document de climatologie (Proulx *et al.*, 1987). Celles-ci sont déduites des courbes d'isohyètes. Pour la plupart des stations, il faut interpoler la précipitation entre 2 isohyètes.

Pourcentage de la superficie couverte par la forêt, les lacs et les marais:

Plusieurs études ont utilisées ce type d'information. Les ouvrages consultés sont: Belzile *et al.* (1997), Desforges et Tremblay (1974), Hoang et Tremblay (1976).

Position géographique:

Considérant la superficie importante du territoire considéré, il est plausible que le régime de précipitations des stations étudiées varie dans la direction nord-sud ou même est-ouest. Ici, deux variables différentes sont utilisées pour exprimer la variation de la position géographique d'une station, soient la longitude et la latitude.

L'application de la méthode en composantes principales demande que les données physiques et géographiques de chacune des stations soient normalisées afin d'obtenir des observations individuelles adimensionnelles. Sinon, la pondération mathématique d'une variable par rapport à la variance totale du système pourrait facilement être biaisée. Par exemple, la contribution de la superficie à la variance expliquée, avec des valeurs de superficie comprises entre 1 km^2 et 57498 km^2, serait démesurée en comparaison de

l'effet de la pente, comprise entre 0.2 et 23.6 % Cette normalisation s'obtient de façon usuelle par:

$$z = \frac{x - \bar{x}}{s} \tag{4.1}$$

où $\bar{x}$ est la moyenne, s l'écart-type, z la donnée normalisée et x la donnée d'origine.

L'annexe E fait état des variables physiques brutes et normalisées pour chacune des 103 stations utilisées pour l'analyse en composantes principales. Dans le cas particulier de la longitude et la latitude, la normalisation s'est effectuée en transformant ces données en variables adimensionnelles par:

$$long^* = \frac{long - long_{\min}}{long_{\max} - long_{\min}} \times 100 \ , \ lat^* = \frac{lat - lat_{\min}}{lat_{\max} - lat_{\min}} \times 100 \tag{4.2}$$

où les variables $long_{max}$, $long_{min}$, lat_{max} et lat_{min} sont respectivements les longitudes et latitudes maximales et minimales de l'ensembles du territoire couvert par cette étude.

4.3.2 Application de l'ACP

Les résultats de l'analyse en composantes principales effectuée sur les variables physiques à l'aide du logiciel Minitab™ sont rapportés au tableau 4.5. En accord avec le critère de variance cumulée, les contributions de chacune des variables physiques aux composantes orthogonales sont présentées pour les trois premières d'entre elles, pour un total de 77.6 %. Ce pourcentage est le même, selon qu'on considère ou non le cas avec rotation "Equamax". La rotation Equamax est un algorithme par lequel l'ordre des variables dans la matrice initiale est successivement modifié de manière à obtenir un ensemble de composantes principales expliquant chacune la même proportion de la variance totale.

Tableau 4.5: Analyse en composantes principales sur les variables physiques

	sans rotation			rotation Equamax		
	C1	C2	C3	C1	C2	C3
Superficie	0.254	0.023	-0.130	0.224	-0.057	-0.170
Longueur	0.117	0.186	-0.143	0.242	-0.092	0.040
Pente	0.328	-0.647	-0.214	-0.160	-0.174	-0.718
Précip.	0.034	-0.346	0.010	-0.214	-0.011	-0.274
Forêt	0.097	-0.307	-0.471	-0.026	-0.454	-0.345
Eau et marais	0.426	-0.124	-0.044	0.221	0.058	-0.383
LAT	-0.015	0.102	-0.428	0.164	-0.407	0.034
LONG	0.614	-0.350	0.280	0.118	0.400	-0.636
%VAR	0.462	0.217	0.097	0.28	0.266	0.23
% Total	0.462	0.679	0.776	0.28	0.546	0.776

Par ailleurs, l'analyse sommaire des coefficients ne permet pas d'associer une composante principale à un groupe de variables physiques données. Par exemple, il n'apparaît pas de composante principale décrivant le facteur "dimension" du bassin et qui serait principalement fonction des variables originales "superficie" et "longueur". Toutefois, on peut montrer que ces huit variables de départ peuvent être remplacées par trois composantes non corrélées, obtenues après rotation, et dès lors appelées PC1, PC2 et PC3. Les valeurs numériques de ces nouvelles variables sont calculées pour chaque station par le calcul matriciel suivant:

$$\begin{pmatrix} Superficie_1 & Longueur_1 & .... & Longitude_1 \\ Superficie_2 & Longueur_2 & .... & Longitude_2 \\ \vdots & \vdots & \vdots & \vdots \\ Superficie_{103} & Longueur_{103} & .... & Longitude_{103} \end{pmatrix} \times \begin{pmatrix} C1_1 & C2_1 & C3_1 \\ C1_2 & C2_2 & C3_2 \\ \vdots & \vdots & \vdots \\ C1_8 & C2_8 & C3_8 \end{pmatrix} = \begin{pmatrix} PC1_1 & PC2_1 & PC3_1 \\ PC1_2 & PC2_2 & PC3_2 \\ \vdots & \vdots & \vdots \\ PC1_{103} & PC2_{103} & PC3_{103} \end{pmatrix} \tag{4.3}$$

4.3.3 Établissement des régions

Les nouvelles variables obtenues par l'analyse en composantes principales sont des combinaisons linéaires des paramètres physiques originaux. Dans le cadre d'un exercice de régionalisation, il est souhaité que les valeurs calculées de PC1, PC2 et PC3 pour chacune des stations permettent le regroupement de ces dernières en un certain nombre de régions homogènes du point de vue de ces variables. Ce regroupement peut généralement s'effectuer selon différentes techniques graphiques ou numériques, lesquelles comportent toujours un aspect subjectif. La méthode graphique retenue ici fut employée par Nathan et McMahon (1990) lors d'une étude sur les débits d'étiages en Australie. Elle origine du travail d'Andrews (1972). Elle consiste à tracer des fonctions sinusoïdales (une par station) dont les paramètres sont les valeurs des composantes principales pour chaque station. L'inspection visuelle des courbes ainsi obtenues permet de regrouper les stations par région, selon la similitude des traces calculées par:

$$f_i(t) = \frac{PC1_i}{\sqrt{2}} + PC2_i \sin(t) + PC3_i \cos(t) \tag{4.4}$$

Les figures 4.12, 4.13 et 4.14 illustrent les $f_i(t)$ calculées pour chacune des 103 stations. Bien qu'il existe un certain nombre de traces non-apparentées, il est facile de discerner des groupes de courbes similaires. En considérant la similarité des traces et la nécessité d'éviter le fractionnement géographique des régions, les 103 stations utilisées pour cette étude ont été regroupées en 6 régions distinctes, en comparaison avec les 8 régions initiales, telles que déterminées par le ministère de l'Environnement du Québec. Les limites géographiques de ces 6 régions sont illustrées à la figure 4.15. Le tableau 4.6 rapporte le sommaire des stations regroupées selon ces régions intitulées A,B,C,D,E et F.

Tableau 4.6: Composition des nouvelles régions

Région	Nombre	Liste des stations
A	24	10101, 10801, 11003, 11201, 11507, 20401, 20404, 20502, 20602, 20802, 21407, 21502, 21601, 21702, 22003, 22301, 51502, 61502, 62209, 62701, 62802,70201, 70401, 71401
B	17	23106, 23301, 23401, 23402, 24001, 24003, 24007, 30101, 30103, 30304, 30314, 30316, 30339, 30415, 30420, 30901, 30905
C	16	22507, 22703, 50402, 50408, 50409, 50904, 51002, 51003, 51004, 51005, 51007, 51301, 52805, 60101,61020, 61022
D	22	23422, 30234, 30282, 31401 40129, 40201, 40212, 40401, 40810, 40814, 41301, 41902, 41903, 42103, 43012, 50116, 50135, 50304, 50305, 50701, 52219, 52233
E	9	61801, 62101, 62102, 72201, 73501, 73801, 74601, 74701, 75601
F	15	61901, 61905, 61906, 80101, 80104, 80701, 80703, 80704, 80705, 80718, 80801, 80809, 81002, 81007, 81101
total	103	

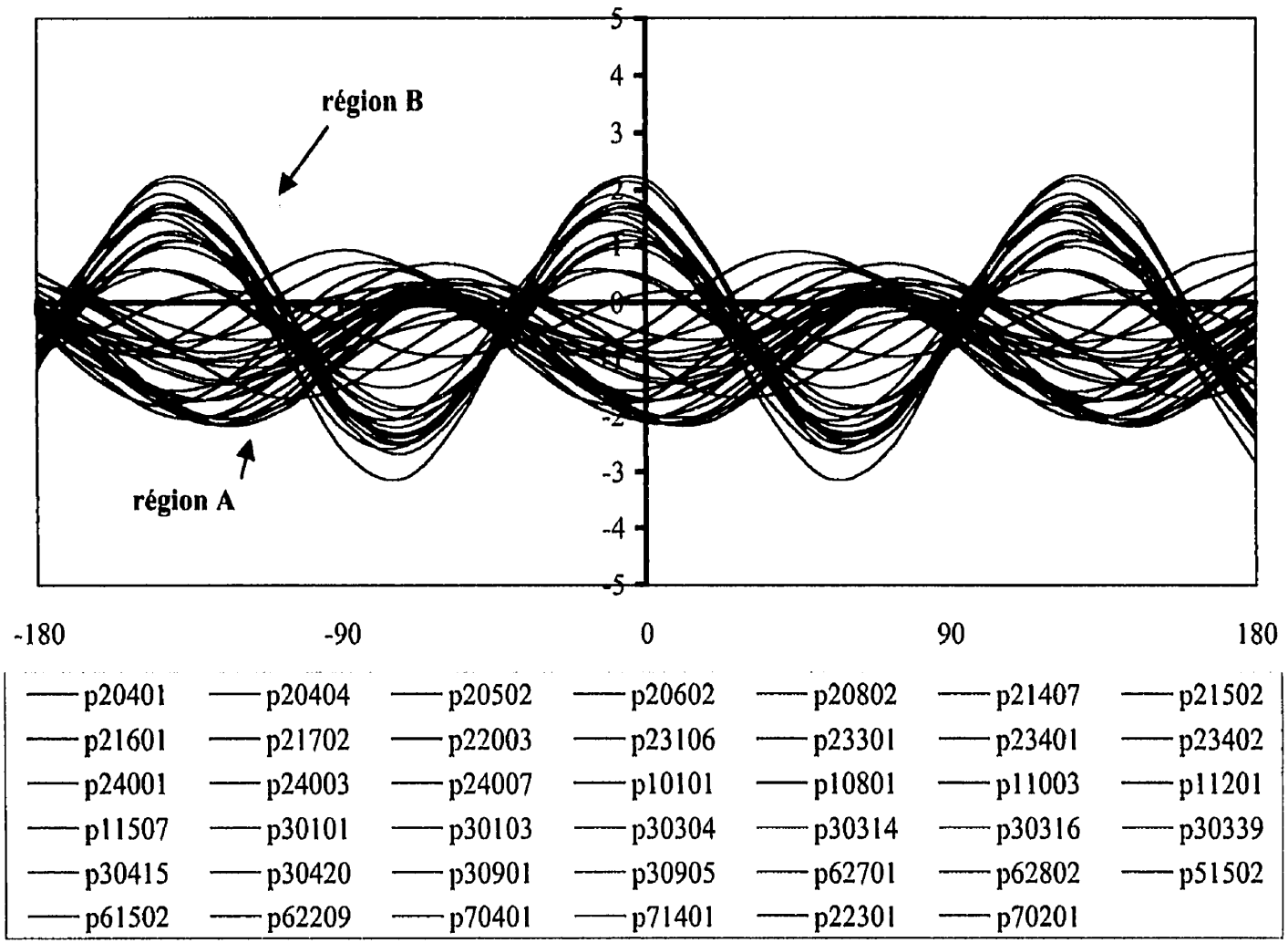

Figure 4.12: Représentation graphique: composantes principales régions A et B

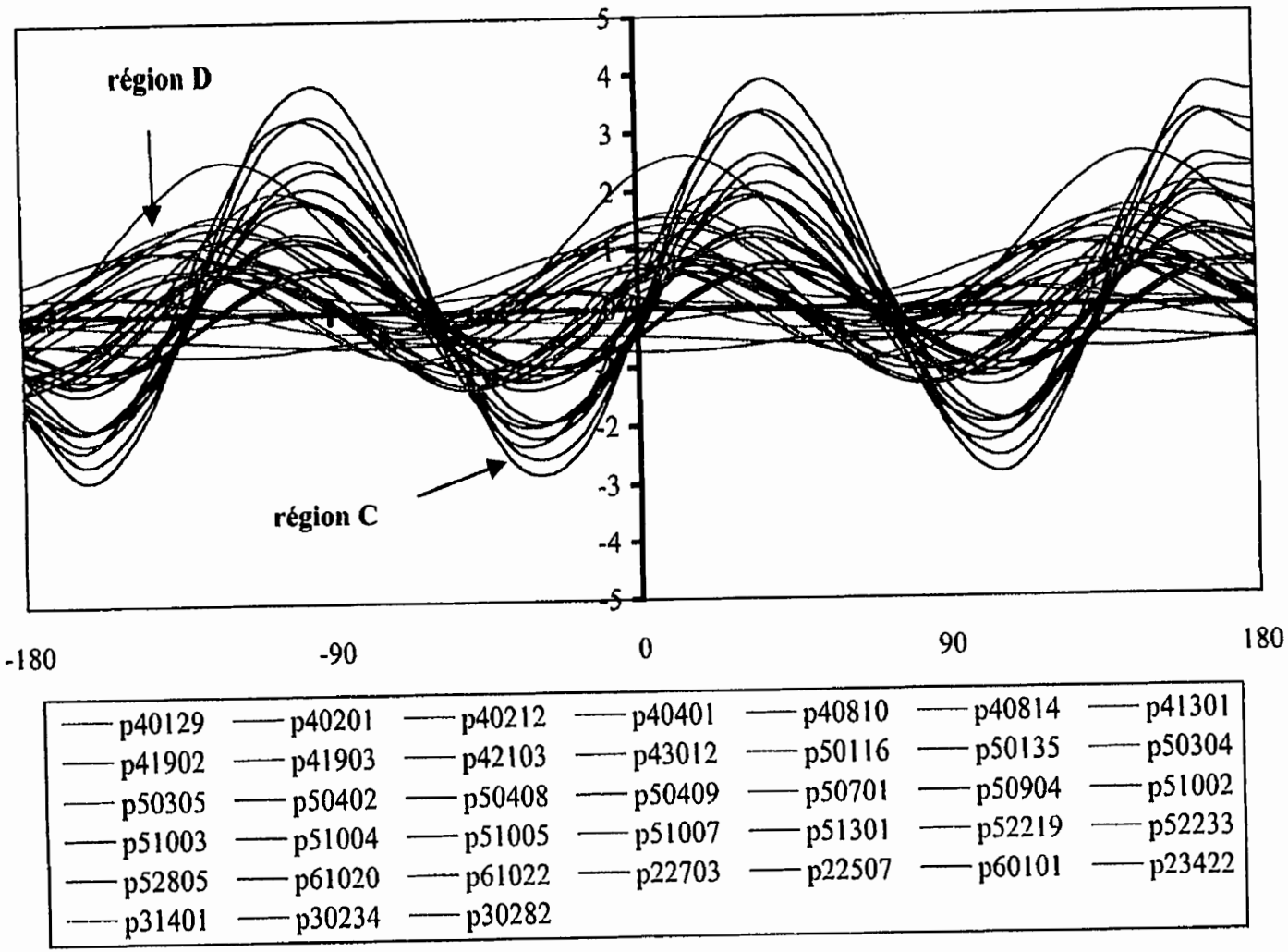

Figure 4.13: Représentation graphique: composantes principales régions C et D

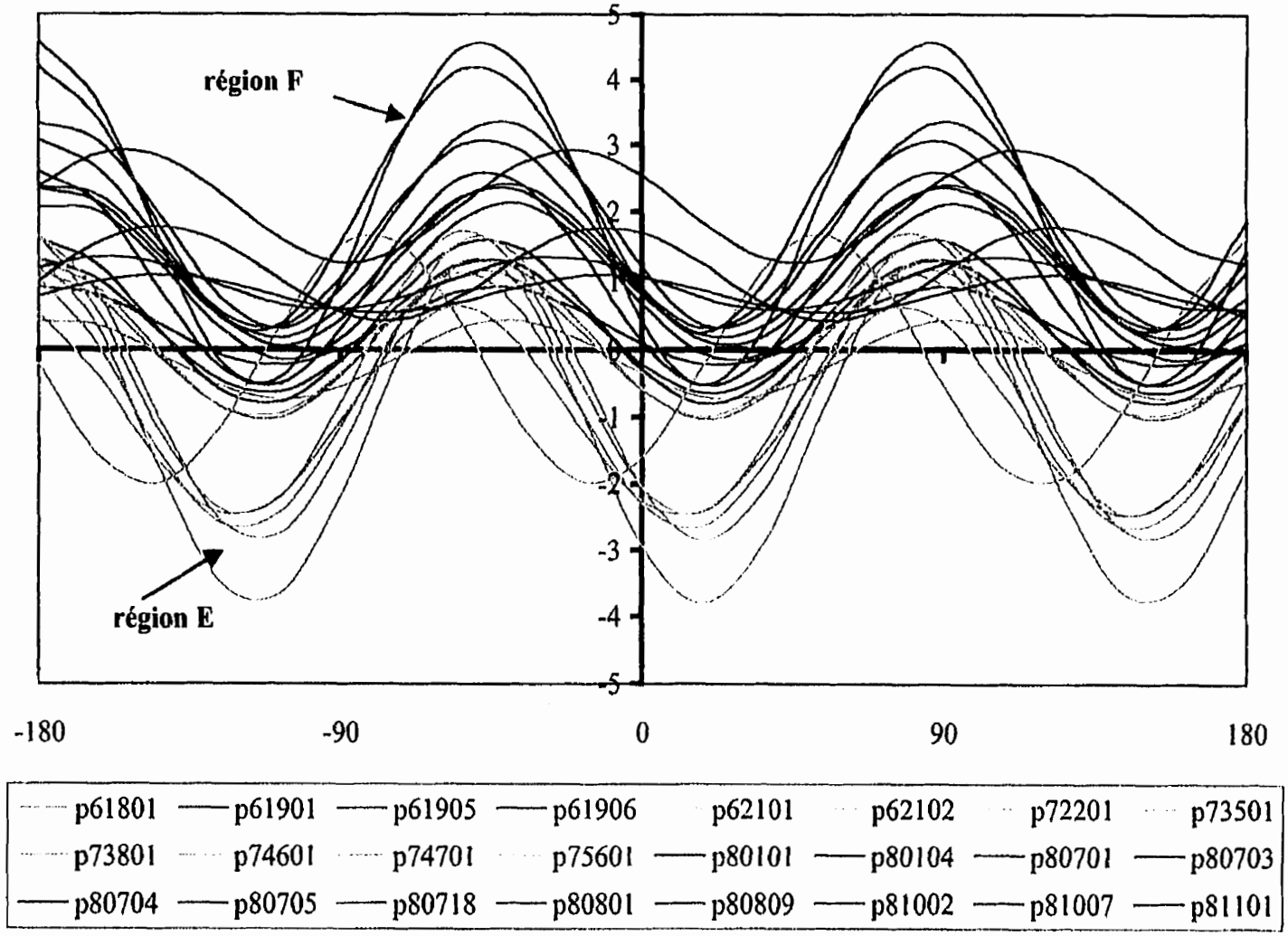

Figure 4.14: Représentation graphique: composantes principales régions E et F

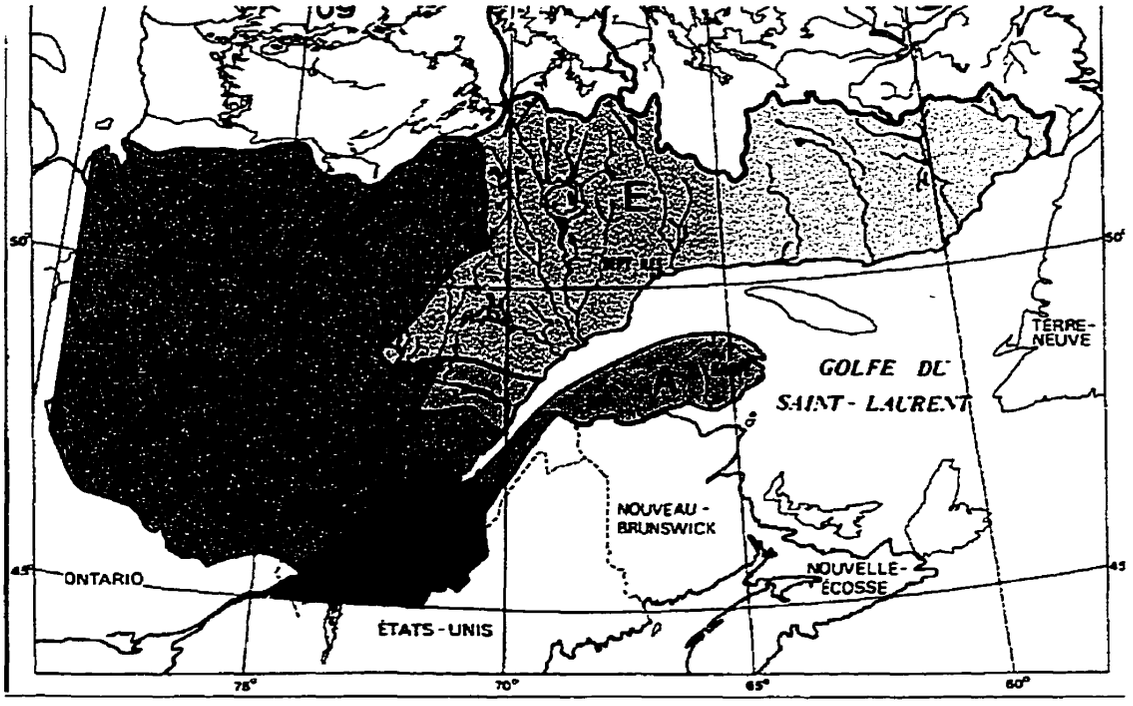

Figure 4.15: Limites géographiques des nouvelles régions

4.4 Modèles de régression linéaires

L'utilisation de modèles linéaires vise d'abord à prédire la réponse d'une variable. Dans le cas présent, il s'agit du débit d'étiage d'une période de retour donnée, à partir de variables indépendantes. Lorsque plusieurs variables indépendantes sont considérées, le modèle sera d'autant plus performant que la corrélation entre les variables indépendantes sera faible. Par ailleurs, la capacité d'ajustement d'un modèle linéaire étant relativement limitée, il peut être nécessaire de fractionner le domaine d'étude des variables considérées en plusieurs parties (ici des régions) afin d'obtenir la qualité de prévision désirée.

La stratégie adoptée dans le cadre de ce travail a été de comparer les modèles linéaires obtenus à partir des variables physiques initiales avec les modèles utilisant les variables orthogonales définies par l'analyse en composantes principales. Cette comparaison s'effectue à la fois pour un domaine global ou pour un ensemble de régions.

4.4.1 Modèle linéaire multivarié sur variables physiques

En accord avec la méthode décrite à la section 3.3, un modèle linéaire a été établi pour les variables hydrologiques 7Q10, 30Q5, 7Q2 et 1Q10. Ces variables, tel le débit 7 jours consécutifs et de période de retour de 10 ans (7Q10), sont les plus utilisées dans l'étude des débits d'étiage. L'analyse de la corrélation entre les variables physiques montre, au tableau 4.7, la relation étroite entre la longueur et la superficie du bassin associé à une station. Pour cette raison, la variable *longueur* a été éliminée du modèle de régression. Les résultats complets de cette modélisation sont rapportés respectivement aux tableaux 4.8, 4.9, 4.10 et 4.11. Dans le premier cas, soit pour la variable 7Q10, pour un modèle à une variable, on constate que tous les facteurs, sauf le pourcentage de forêt, peuvent être utilisés à des niveaux de confiance supérieurs à 95% (On peut rejeter avec 95% de confiance l'hypothèse que le coefficient est nul). La variable la plus significative est la superficie. Le deuxième paramètre ajouté à cette variable est les précipitations totales moyennes annuelles, bien que, statistiquement parlant, la longitude ait une importance légèrement meilleure. Ce choix s'explique par le fait qu'à un niveau de confiance équivalent, ou presque, l'interprétation physique de l'effet de la variable "précipitation" est plus facile à concevoir que celui de la variable "longitude". L'ajout d'une troisième variable s'est avéré statistiquement injustifié. Par ailleurs, les régressions ont été effectuées en imposant une valeur nulle au paramètre constant du modèle linéaire.

Tableau 4.7: Corrrélation des variables géophysiques des stations

	S (km^2)	L (km)	Pente (m/km)	Précip (mm)	Forêt (%)	Eau et marais (%)	LAT	LONG
S	1.00							
L	0.86	1.00						
Pente	-0.32	-0.48	1.00					
Précip.	-0.52	-0.59	0.47	1.00				
Forêt	-0.03	0.01	0.21	0.14	1.00			
Eau et marais	0.53	0.62	-0.40	-0.57	-0.11	1.00		
LAT	0.48	0.60	-0.15	-0.50	0.39	0.50	1.00	
LONG	0.31	0.24	-0.28	-0.28	-0.33	0.54	-0.23	1.00

Les modèles linéaires obtenus avec les variables "superficie" et "précipitation" ont été utilisés afin de calculer 7Q10, 30Q5, 1Q10 et 7Q2 à chacune des 103 stations conservées pour cette étude. Ces résultats sont présentés aux figures 4.16 à 4.19. A titre indicatif, les points expérimentaux sont regroupés selon les huit régions hydrographiques initiales (MEQ). En plus de la modélisation globale, i.e. en considérant les toutes les stations en un seul ensemble de données, huit modèles linéaires régionaux ont aussi été calculés. L'analyse des courbes obtenues montre que le modèle global répond généralement bien aux tendances obervées dans les données expérimentales. Cependant, la capacité du modèle à prédire quantitativement les variables 7Q10, 30Q5, 1Q10 et 7Q2 de façon ponctuelle pour une station donnée est parfois insuffisante, même lorsque celui-ci est appliqué sur une base régionale. Dans plusieurs cas, des valeurs négatives sont obtenues. Aussi, le processus de modélisation et régionalisation doit donc être redéfini afin d'espérer un meilleur ajustement des données. Pour ce faire, les variables résultantes de l'analyse en composantes principales seront utilisées.

Tableau 4.8: Modèle linéaire multivarié global 7Q10

1 variable	Superficie	Pente	Préc.	Latitude	Longitude	% forêt	%marais
SSR	157278	13213	42918	41398	13803	28	40504
SSE	27093	171159	141454	142973	170569	184344	143868
F_0	592.11	7.80	30.64	29.24	8.17	0.02	28.43
α	5.00E-44	6.26E-03	2.46E-07	4.29E-07	0.00517	9.02E-01	5.94E-07
β_0	0.000	26.042	192.296	-17.888	-17.78967	17.692	-1.957
β_1	0.005	-2.301	-0.174	0.787	0.548284	-0.029	2.634
Superficie +1 variable	**Pente**	**Préc.**	**Latitude**	**Longitude**	**% forêt**	**%marais**	
SSR(Superficie)	157278	157278	157278	157278	157278	157278	
SSR	157522	158507	157774	158541	158346	158018	
SSR(B2/Superfice)	244	1229	495	1263	1068	739	
SSE	26850	25865	26598	25831	26026	26354	
MSE	265.8	256.1	263.3	255.8	257.7	260.9	
F_0	0.92	4.80	1.88	4.94	4.14	2.83	
α	0.341	0.031	0.173	0.029	0.044	0.095	
β_0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
β_1	0.0047	0.0048	0.0048	0.0049	0.0048	0.0049	
β_2	-0.2277	-0.0037	-0.0557	-0.0635	-0.0422	-0.3511	
Superficie + Précipitation+1 variable	**Pente**	**Latitude**	**Longitude**	**% forêt**	**%marais**		
SSR(Super+ préc)	158507	158507	158507	158507	158507		
SSR	158760	158600	158577	158538	158628		
SSR(B3/Super+ préc.)	253	93	70	31	121		
SSE	25611	25771	25795	25834	25743		
MSE	256.1	257.7	257.9	258.3	257.4		
F_0	0.99	0.36	0.27	0.12	0.47		
α	0.322	0.549	0.604	0.729	0.494		
β_0	0.00000	0.00000	0.00000	0.00000	0.00000		
β_1	0.00488	0.00475	0.00487	0.00480	0.00491		
β_2	-0.00553	-0.00501	-0.00160	-0.00601	-0.00299		
β_3	0.34923	0.04032	-0.03827	0.02989	-0.16471		

Tableau 4.9: Modèle linéaire multivarié global 30Q5

1 variable	Superficie	Pente	Préc.	Latitude	Longitude	% forêt	%marais
SSR	192311	16826	53080	51022	17048	41	50588
SSE	32499	207984	171730	173788	207762	224769	174223
F_0	603.58	8.17	31.22	29.65	8.29	0.02	29.33
α	2.13E-44	5.17E-03	1.96E-07	3.64E-07	0.00487	8.93E-01	4.15E-07
β_0	0.000	29.757	214.475	-19.208	-19.13844	20.531	5.425
β_1	0.005	-2.597	-0.194	0.874	0.609337	-0.035	0.543
Superficie +1 variable	**Pente**	**Préc.**	**Latitude**	**Longitude**	**% forêt**	**%marais**	
SSR(Superficie)	192311	192311	192311	192311	192311	192311	
SSR	192518	193349	192645	193378	193191	192875	
SSR(B2/Superfice)	207	1038	334	1067	880	564	
SSE	32292	31461	32165	31432	31619	31935	
MSE	319.7	311.5	318.5	311.2	313.1	316.2	
F_0	0.65	3.33	1.05	3.43	2.81	1.78	
α	0.423	0.071	0.308	0.067	0.097	0.185	
β_0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
β_1	0.0052	0.0053	0.0053	0.0054	0.0053	0.0054	
β_2	-0.2097	-0.0034	-0.0457	-0.0584	-0.0384	-0.3067	
Superficie + Précipitation+1 variable	**Pente**	**Latitude**	**Longitude**	**% forêt**	**%marais**		
SSR(Super+ préc)	193349	193349	193349	193349	193349		
SSR	193561	193506	193408	193393	193424		
SSR(B3/Super+ préc.)	212	157	59	44	76		
SSE	31249	31304	31402	31417	31386		
MSE	312.5	313.0	314.0	314.2	313.9		
F_0	0.68	0.50	0.19	0.14	0.24		
α	0.412	0.481	0.665	0.710	0.625		
β_0	0.00000	0.00000	0.00000	0.00000	0.00000		
β_1	0.00538	0.00523	0.00537	0.00531	0.00540		
β_2	-0.00508	-0.00511	-0.00147	-0.00614	-0.00283		
β_3	0.31987	0.05226	-0.03519	0.03545	-0.12997		

Tableau 4.10: Modèle linéaire multivarié global 7Q2

1 variable	Superficie	Pente	Préc.	Latitude	Longitude	% forêt	%marais
SSR	235120	21521	67701	65150	21436	60	66616
SSE	41550	255150	208969	211521	255235	276611	210054
F_0	577.18	8.52	32.72	31.11	8.48	0.02	32.03
α	1.47E-43	4.33E-03	1.09E-07	2.05E-07	0.00441	8.83E-01	1.43E-07
β_0	0.000	33.636	242.229	-21.707	-21.30909	23.443	-2.204
β_1	0.006	-2.937	-0.219	0.987	0.683273	-0.043	3.377
Superficie +1 variable	**Pente**	**Préc.**	**Latitude**	**Longitude**	**% forêt**	**%marais**	
SSR(Superficie)	235120	235120	235120	235120	235120	235120	
SSR	235358	236165	235296	236116	235965	235335	
SSR(B2/Superfice)	238	1045	176	996	845	215	
SSE	41312	40506	41375	40555	40706	41336	
MSE	409.0	401.0	409.7	401.5	403.0	409.3	
F_0	0.58	2.60	0.43	2.48	2.10	0.52	
α	0.447	0.110	0.514	0.118	0.151	0.471	
β_0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
β_1	0.0058	0.0059	0.0059	0.0060	0.0059	0.0059	
β_2	-0.2252	-0.0034	-0.0332	-0.0564	-0.0376	-0.1892	
Superficie + Précipitation+1 variable	**Pente**	**Latitude**	**Longitude**	**% forêt**	**%marais**		
SSR(Super+ préc)	236165	236165	236165	236165	236165		
SSR	236336	236607	236186	236256	236168		
SSR(B3/Super+ préc.)	172	442	21	91	4		
SSE	40334	40064	40485	40415	40502		
MSE	403.3	400.6	404.8	404.1	405.0		
F_0	0.43	1.10	0.05	0.23	0.01		
α	0.516	0.296	0.819	0.636	0.925		
β_0	0.00000	0.00000	0.00000	0.00000	0.00000		
β_1	0.00594	0.00574	0.00592	0.00587	0.00588		
β_2	-0.00491	-0.00631	-0.00224	-0.00739	-0.00349		
β_3	0.28747	0.08774	-0.02111	0.05115	0.02852		

Tableau 4.11: Modèle linéaire multivarié global 1Q10

1 variable	Superficie	Pente	Préc.	Latitude	Longitude	% forêt	%marais
SSR	156327	13009	42614	41309	13653	22	40225
SSE	27261	170580	140974	142280	169936	183567	143364
F_0	584.91	7.70	30.53	29.32	8.11	0.01	28.34
α	8.30E-44	6.57E-03	2.57E-07	4.16E-07	0.00532	9.13E-01	6.17E-07
β_0	0.000	25.723	191.435	-18.087	-17.84452	17.196	-2.132
β_1	0.005	-2.284	-0.174	0.786	0.545304	-0.026	2.624
Superficie +1 variable	**Pente**	**Préc.**	**Latitude**	**Longitude**	**% forêt**	**%marais**	
SSR(Superficie)	156327	156327	156327	156327	156327	156327	
SSR	156597	157675	156878	157713	157496	157125	
SSR(B2/Superfice)	270	1347	550	1385	1168	798	
SSE	26992	25914	-156327	25876	26093	26463	
MSE	267.2	256.6	264.5	256.2	258.3	262.0	
F_0	1.01	5.25	2.08	5.41	4.52	3.05	
α	0.318	0.024	0.152	0.022	0.036	0.084	
β_0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
β_1	0.0047	0.0048	0.0048	0.0049	0.0048	0.0049	
β_2	-0.2396	-0.0038	-0.0587	-0.0666	-0.0442	-0.3647	
Superficie + Précipitation+1 variable	**Pente**	**Latitude**	**Longitude**	**% forêt**	**%marais**		
SSR(Super+ préc)	157675	157675	157675	157675	157675		
SSR	157948	157772	157752	157710	157802		
SSR(B3/Super+ préc.)	274	97	77	36	127		
SSE	25640	25817	25837	25879	25787		
MSE	256.4	258.2	258.4	258.8	257.9		
F_0	1.07	0.38	0.30	0.14	0.49		
α	0.304	0.541	0.586	0.712	0.484		
β_0	0.00000	0.00000	0.00000	0.00000	0.00000		
β_1	0.00487	0.00473	0.00486	0.00479	0.00490		
β_2	-0.00578	-0.00521	-0.00167	-0.00634	-0.00315		
β_3	0.36307	0.04113	-0.04024	0.03195	-0.16865		

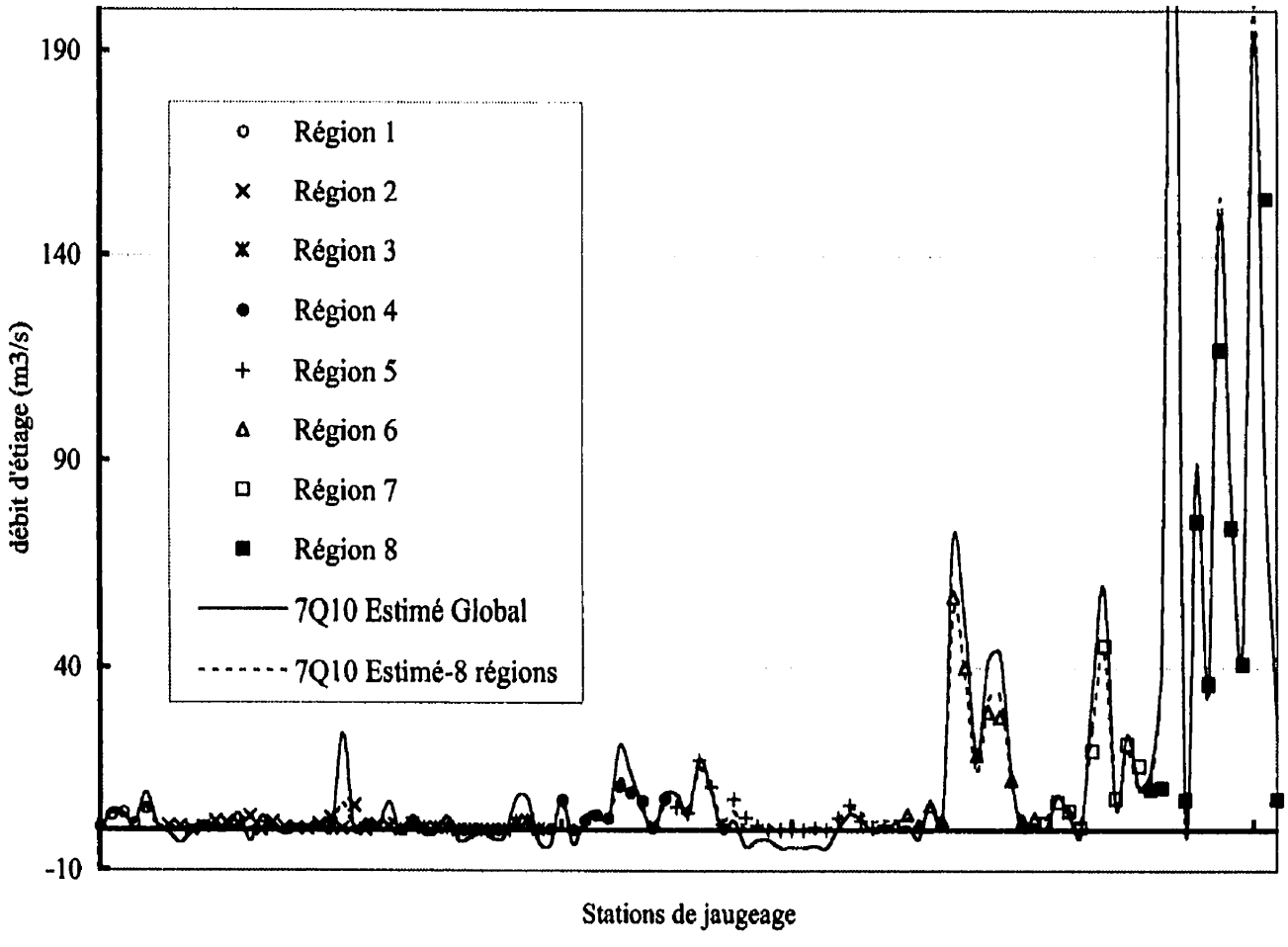

Figure 4.16: Variable 7Q10 estimée par modèle linéaire: superficie et précipitation

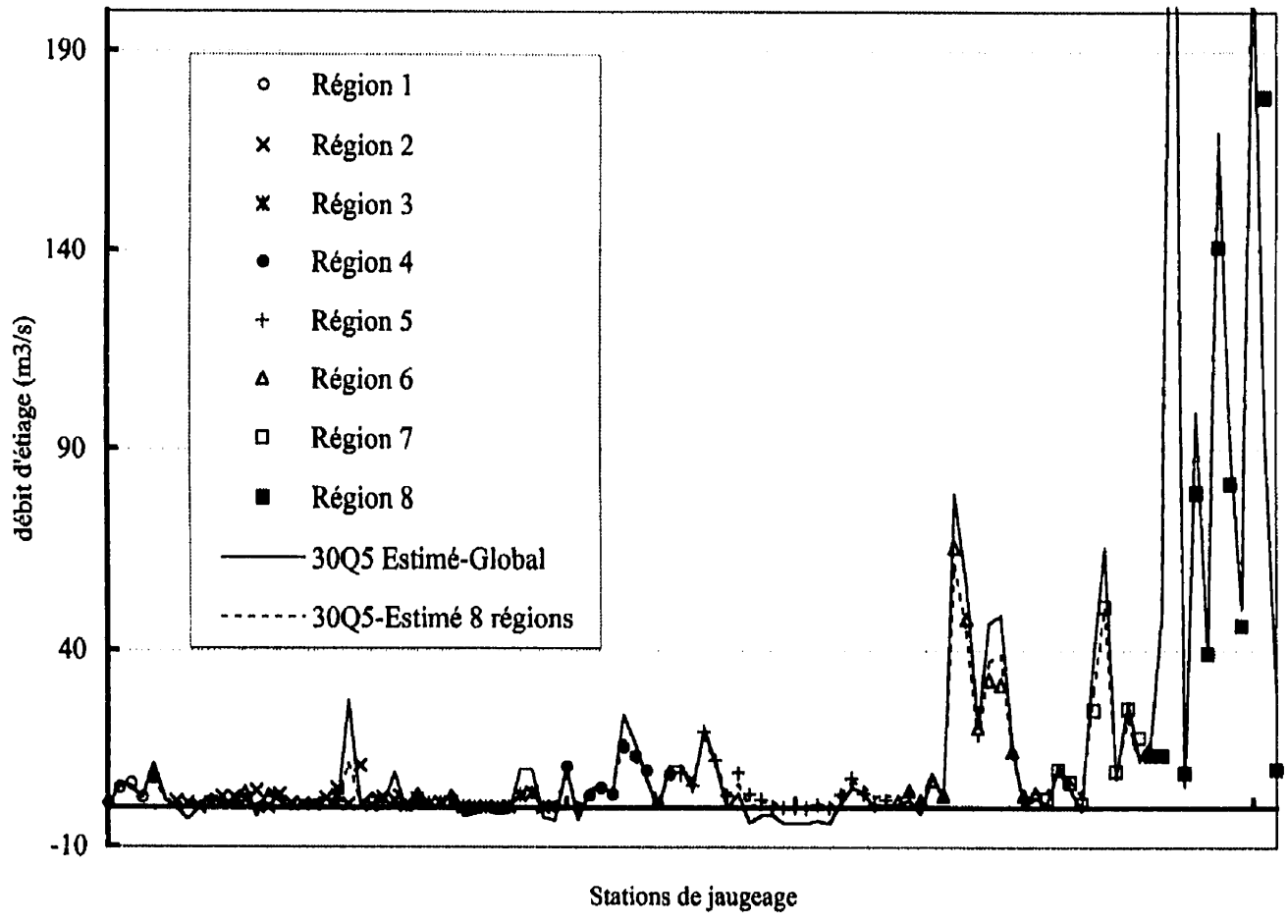

Figure 4.17: Variable 30Q5 estimée par modèle linéaire: superficie et précipitation

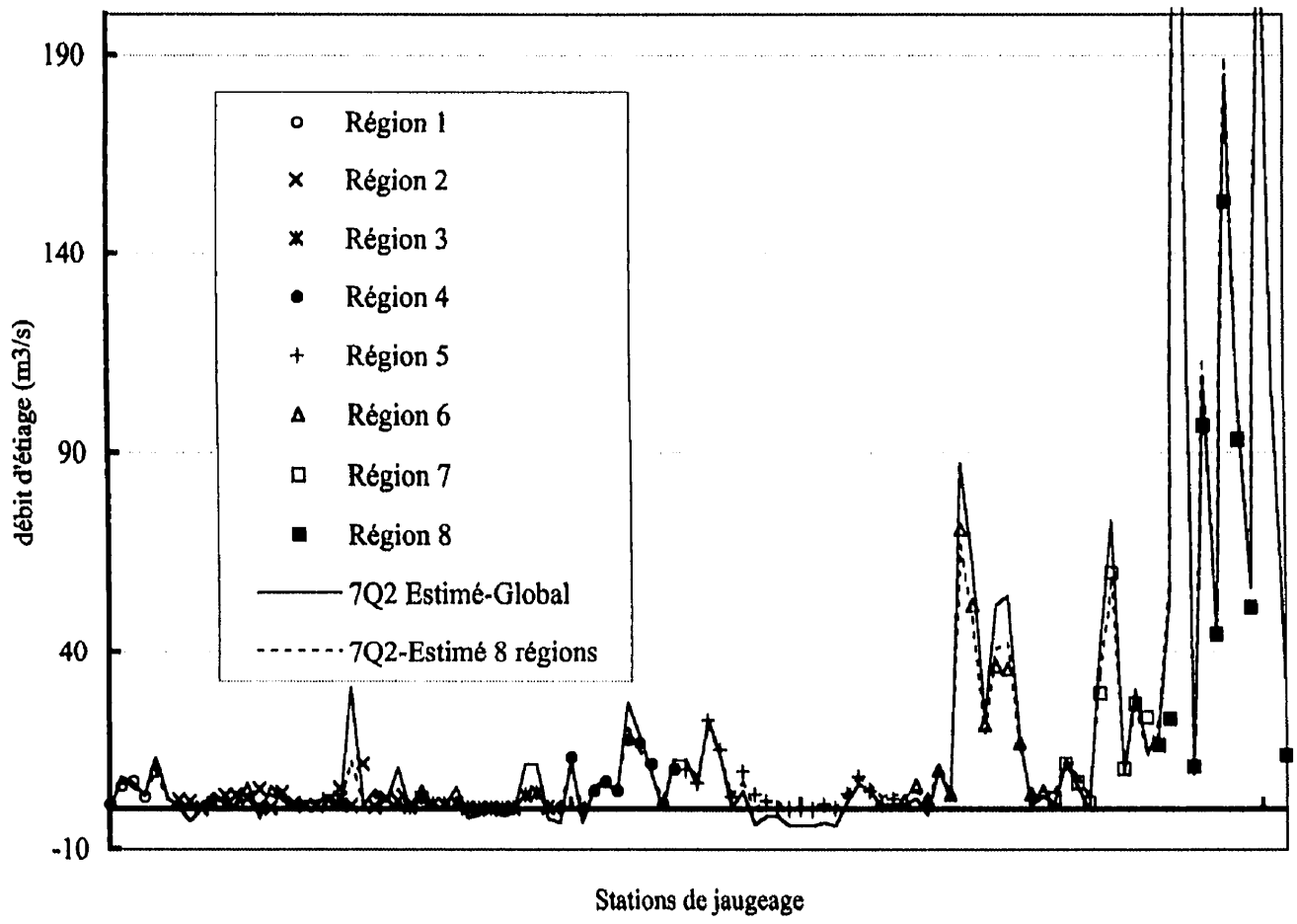

Figure 4.18: Variable 7Q2 estimée par modèle linéaire: superficie et précipitation

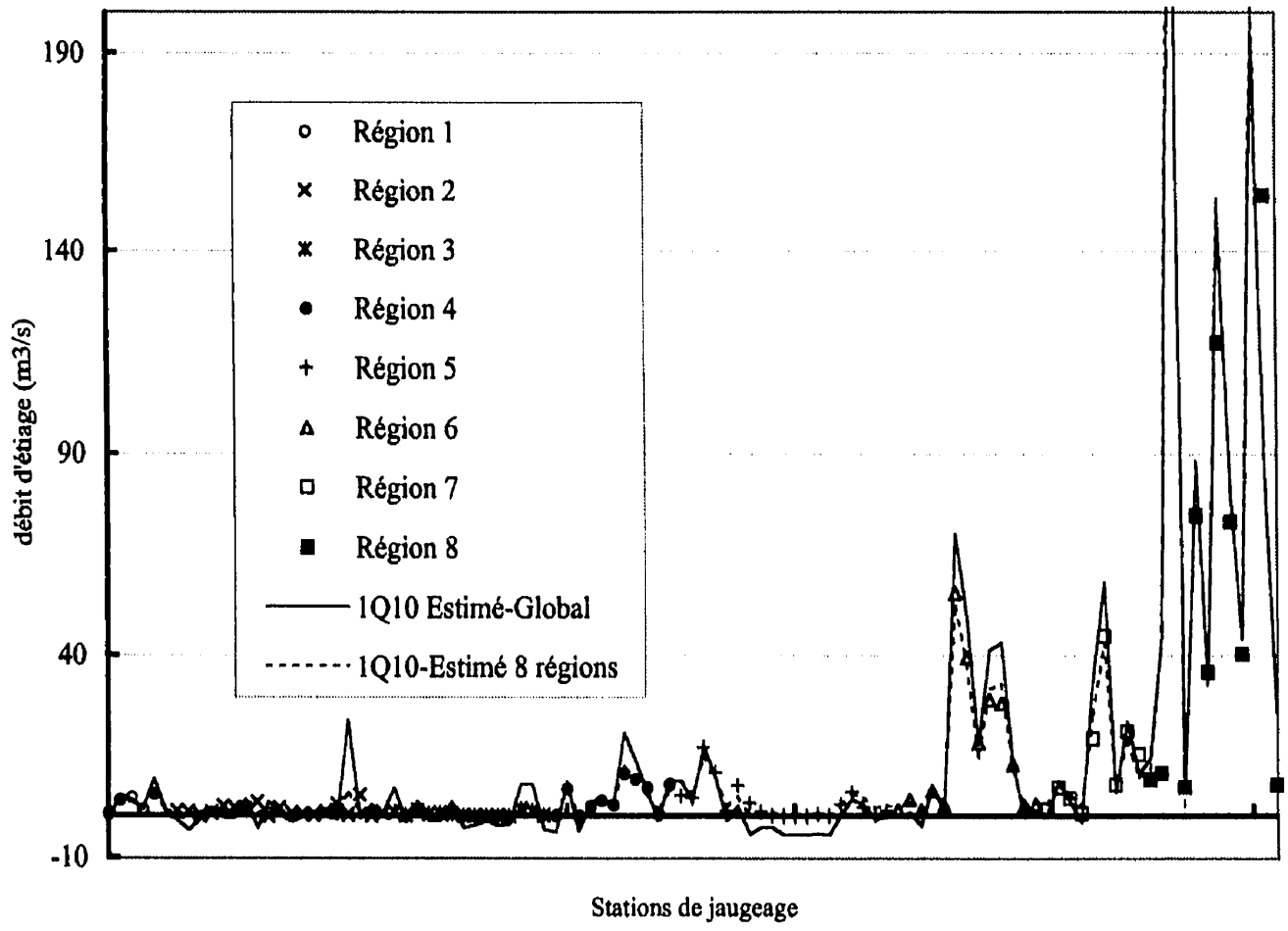

Figure 4.19: Variable 1Q10 estimée par modèle linéaire: superficie et précipitation

4.4.2 Régression sur les composantes principales

Lors de l'analyse en composantes principales effectuée à la section 4.3, il a été possible d'identifier six nouvelles régions et d'attribuer à chacune d'entre elles un ensemble de coefficients correspondant à trois composantes orthogonales. Lorsque les coefficients sont utilisés selon l'équation 4.3, ceux-ci permettent de transformer les variables physiques en trois nouvelles variables indépendantes ici appellées PC1, PC2, et PC3. Les régressions linéaires ont été effectuées à partir de ces variables et les modèles obtenus sont présentés au tableau 4.12. Les courbes correspondantes sont rapportées aux figures 4.20 à 4.25 pour les régions A-F alors que les figures 4.26 à 4.29 font état de l'application d'un modèle linéaire sur les variables PC1, PC2 et PC3 pour l'ensemble des 103 stations (modèle global). On constate, dans le cas du modèle régional, que l'ajustement est généralement satisfaisant, particulièrement pour les régions B, C, E et F. Dans le cas de plusieurs stations des régions A et D, la valeur estimée de la variable hydrologique est suffisamment différente de la valeur expérimentale pour n'être utile qu'à des fins d'analyse de tendances. Le tableau 4.13 présente l'erreur moyenne associée aux différents modèles de régression.

Tableau 4.12: Coefficients des modèles linéaires

Région	Variable	B_0	B_1	B_2	B_3
A	7Q10	5.509	6.049	0.006	-0.656
	30Q5	6.958	7.133	0.179	-0.729
	1Q10	5.255	5.773	0.035	-0.585
	7Q2	8.328	8.252	0.591	-0.648
B	7Q10	6.864	8.155	-1.278	-0.357
	30Q5	11.360	13.333	-2.324	-0.366
	1Q10	5.457	6.410	-1.234	-0.016
	7Q2	12.861	14.877	-2.252	-1.456
C	7Q10	3.863	2.226	1.951	-0.351
	30Q5	4.802	2.556	2.749	-0.413
	1Q10	3.669	2.093	1.883	-0.331
	7Q2	5.463	3.191	2.792	-0.489
D	7Q10	8.576	6.032	-4.284	0.208
	30Q5	10.605	7.728	-5.249	-0.031
	1Q10	8.237	5.774	-4.125	0.212
	7Q2	12.554	9.921	-6.401	0.023
E	7Q10	-4.365	33.382	-13.340	-9.145
	30Q5	-5.202	37.811	-15.418	-10.102
	1Q10	-4.275	33.227	-13.201	-9.213
	7Q2	-7.147	43.795	-18.484	-11.557
F	7Q10	6.085	29.101	-87.659	-65.789
	30Q5	8.535	32.418	-96.955	-69.944
	1Q10	5.519	29.212	-87.080	-65.594
	7Q2	-3.383	36.154	-148.769	-89.252

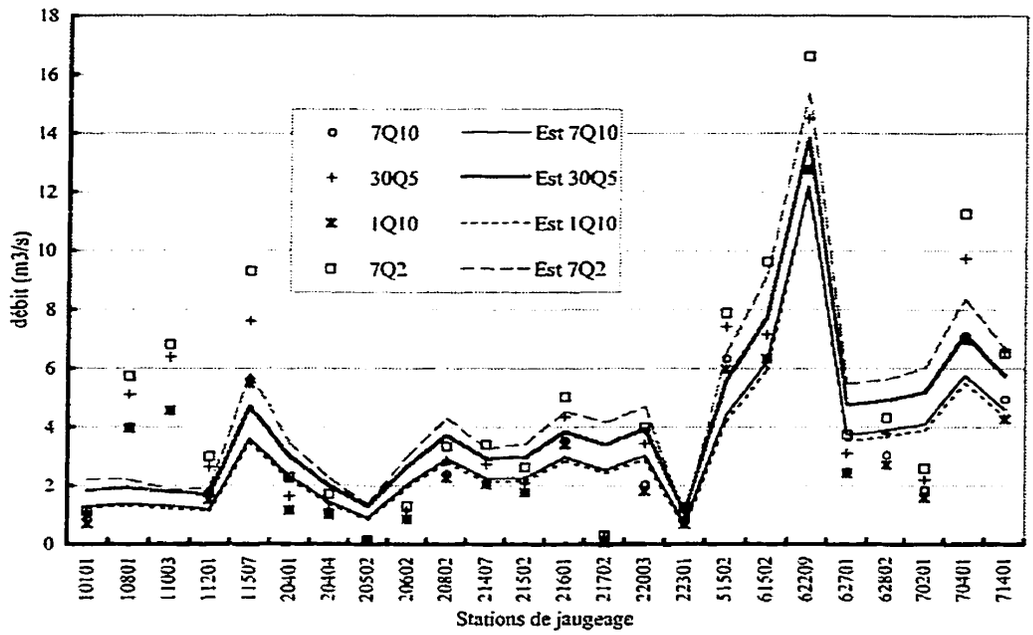

Figure 4.20: Estimation de 1Q10, 7Q10, 7Q2, et 30Q5 pour la région A

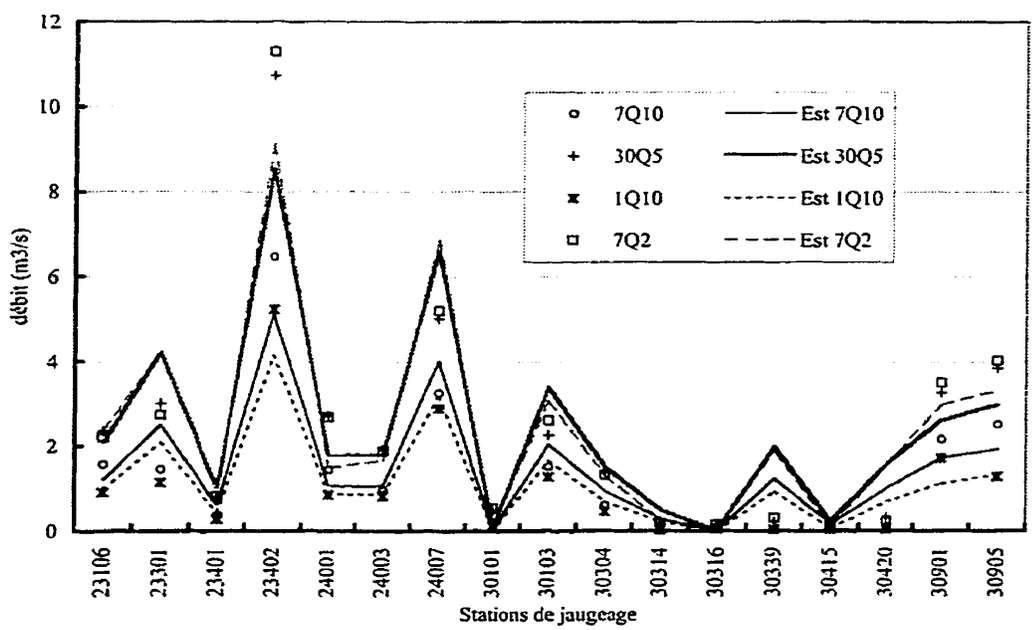

Figure 4.21: Estimation de 1Q10, 7Q10, 7Q2, et 30Q5 pour la région B

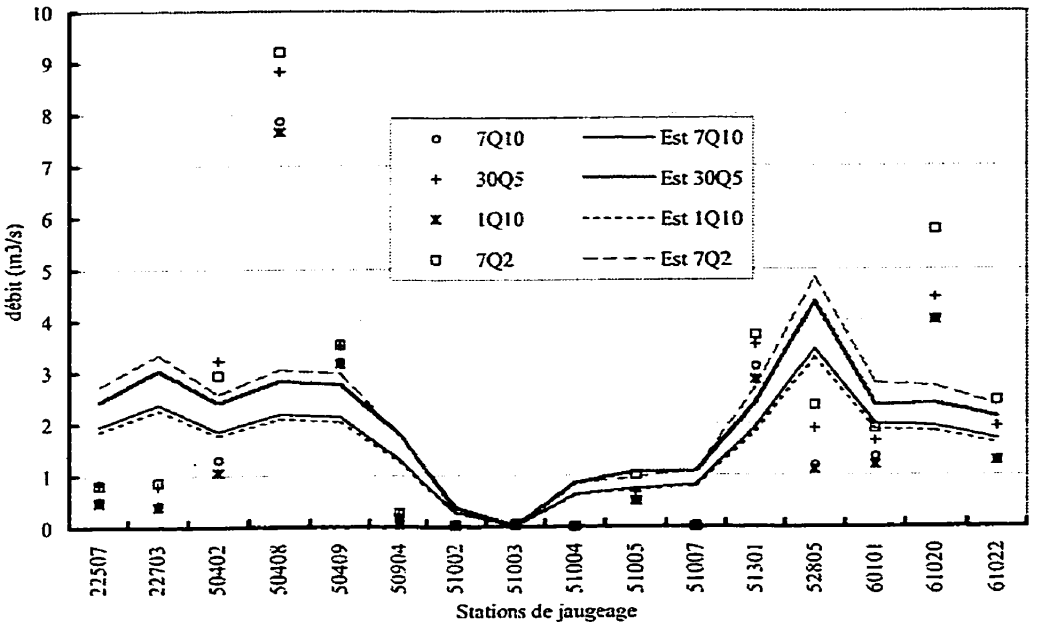

Figure 4.22: Estimation de 1Q10, 7Q10, 7Q2, et 30Q5 pour la région C

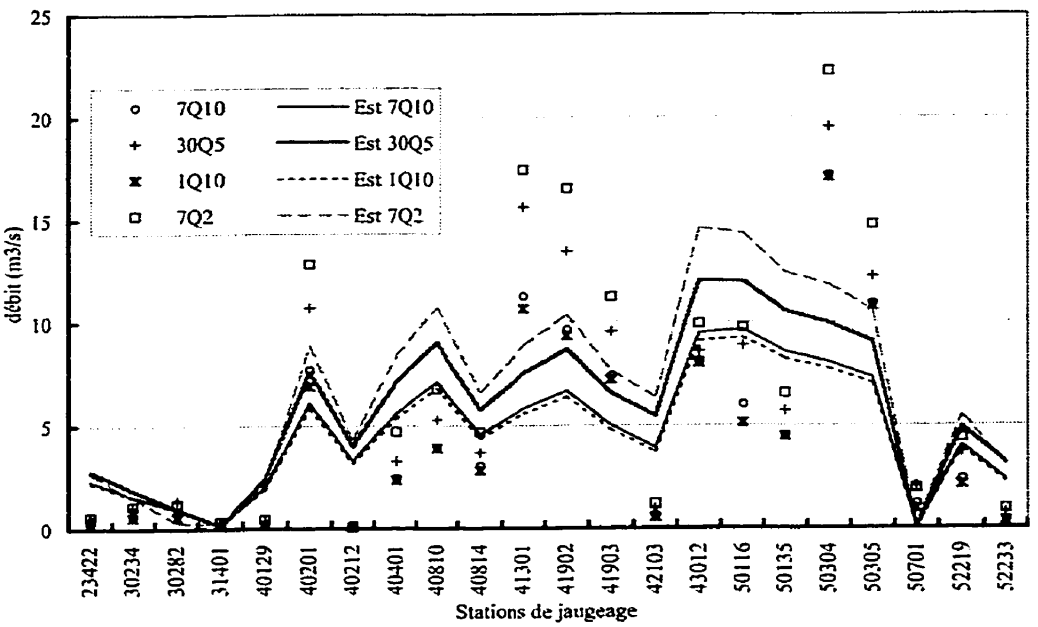

Figure 4.23: Estimation de 1Q10, 7Q10, 7Q2, et 30Q5 pour la région D

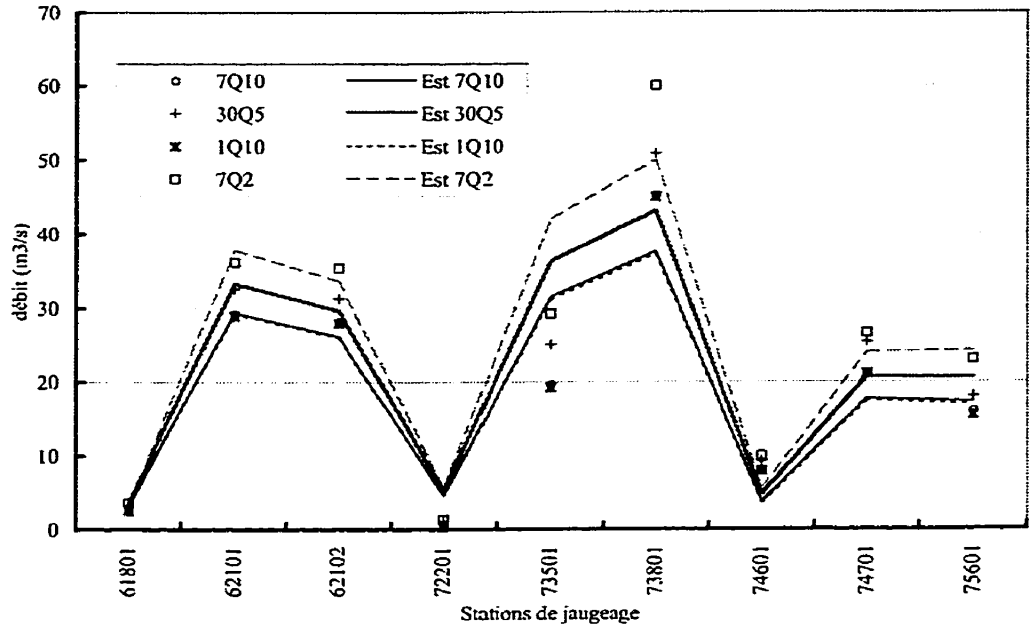

Figure 4.24: Estimation de 1Q10, 7Q10, 7Q2, et 30Q5 pour la région E

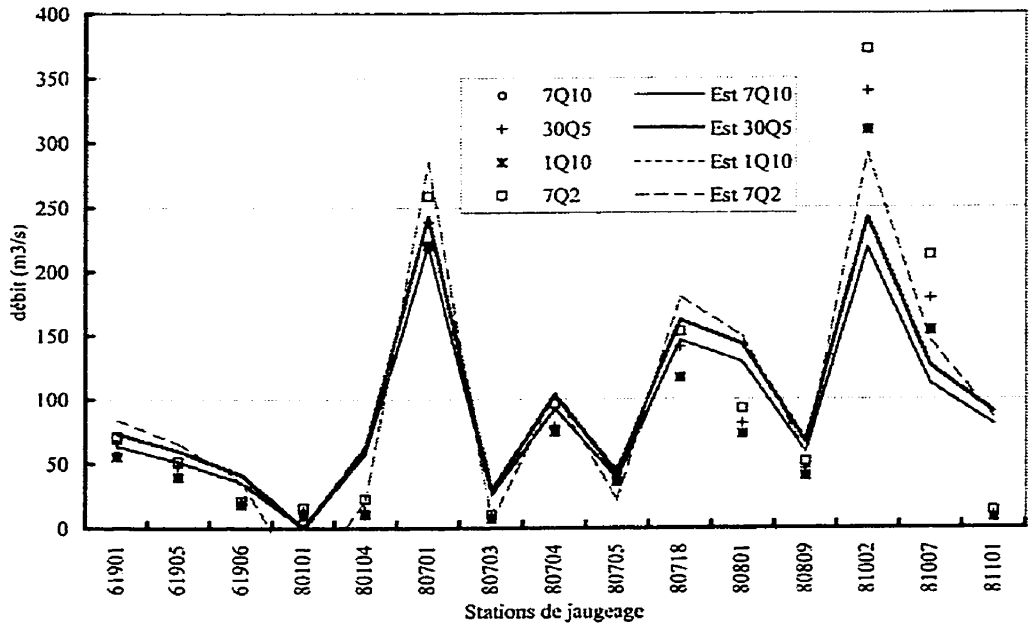

Figure 4.25: Estimation de 1Q10, 7Q10, 7Q2, et 30Q5 pour la région F

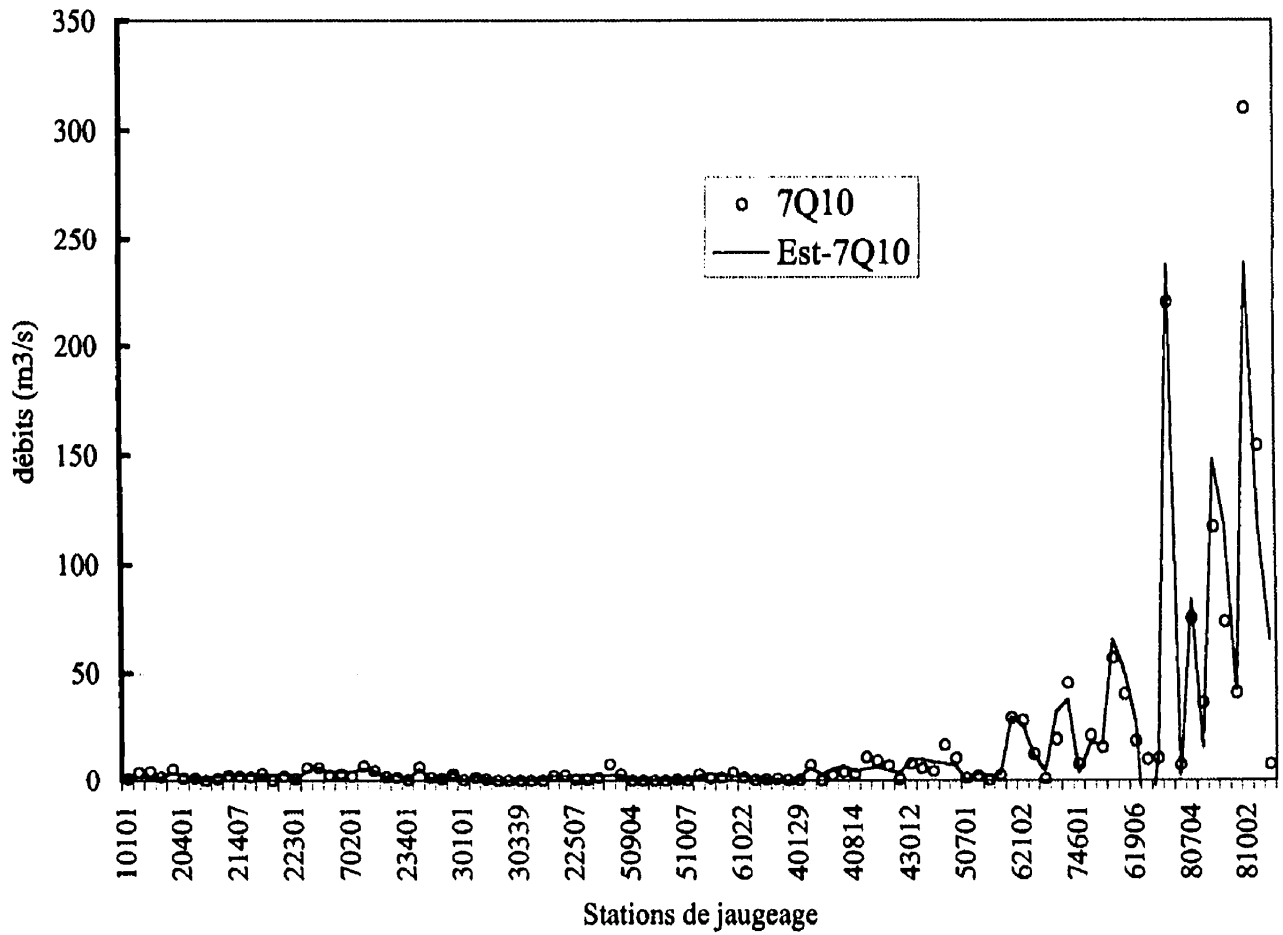

Figure 4.26: Variable 7Q10 estimée par modèle linéaire: PC1, PC2 et PC3

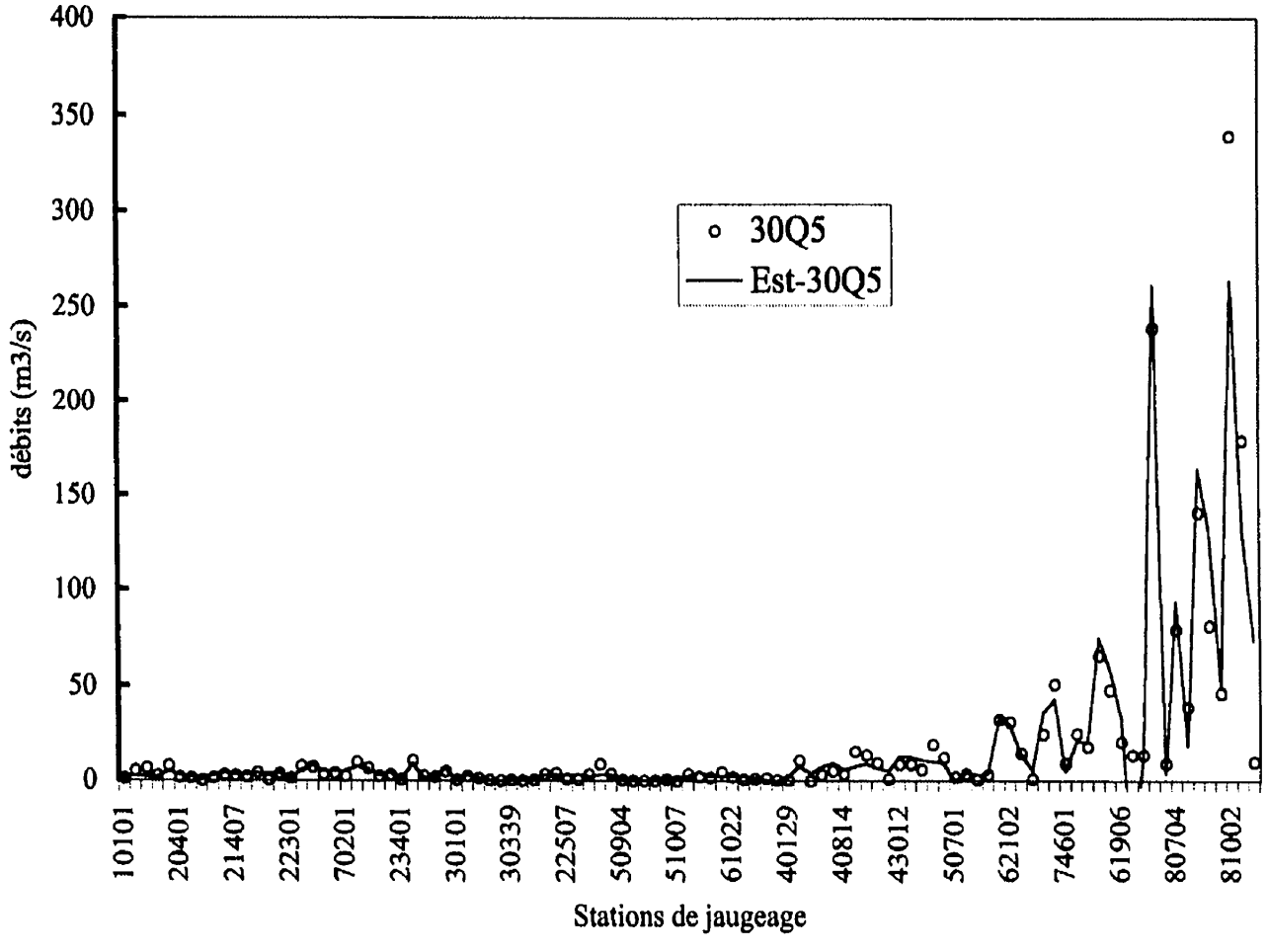

Figure 4.27: Variable 30Q5 estimée par modèle linéaire: PC1, PC2 et PC3

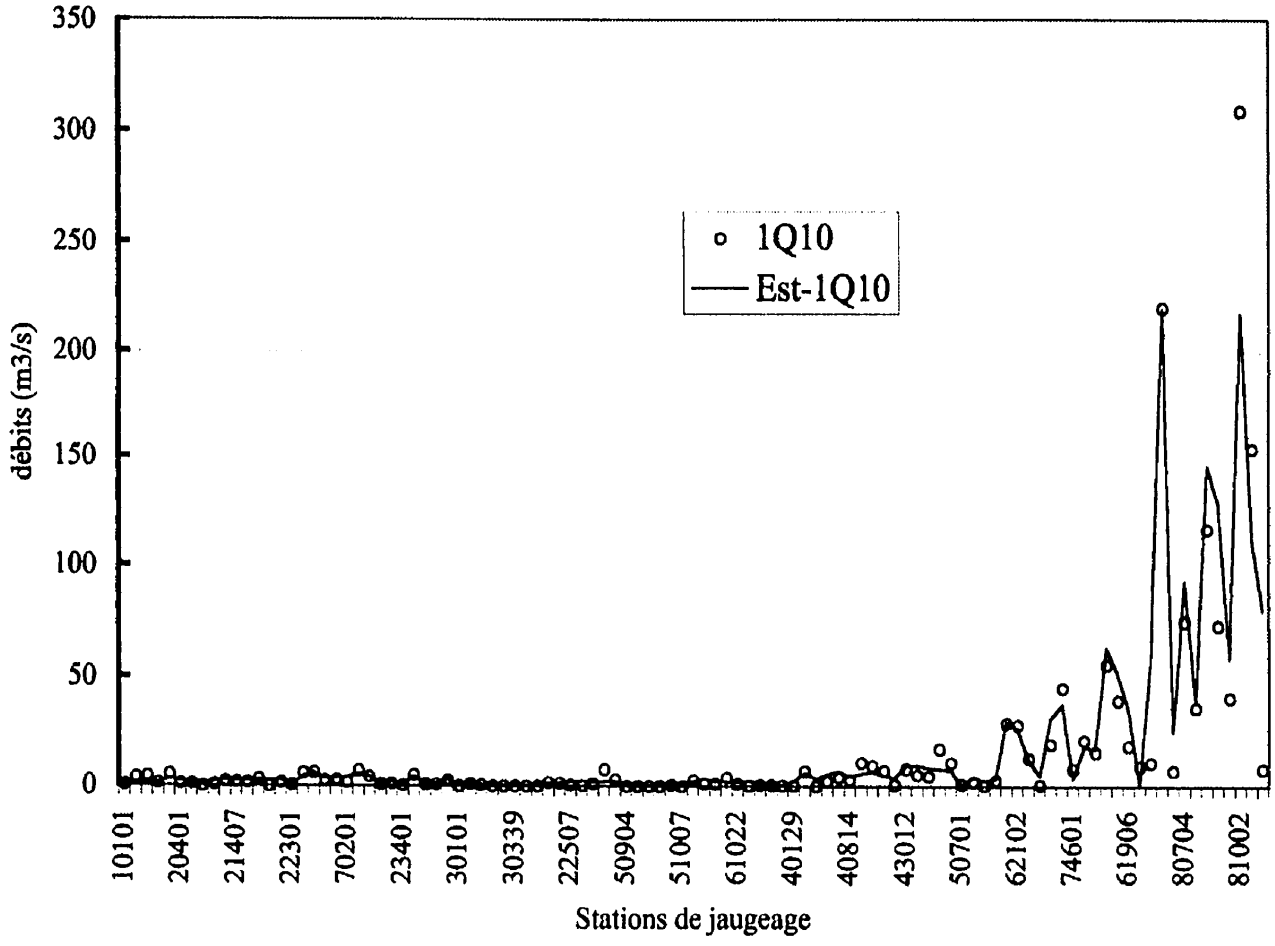

Figure 4.28: Variable 1Q10 estimée par modèle linéaire: PC1, PC2 et PC3

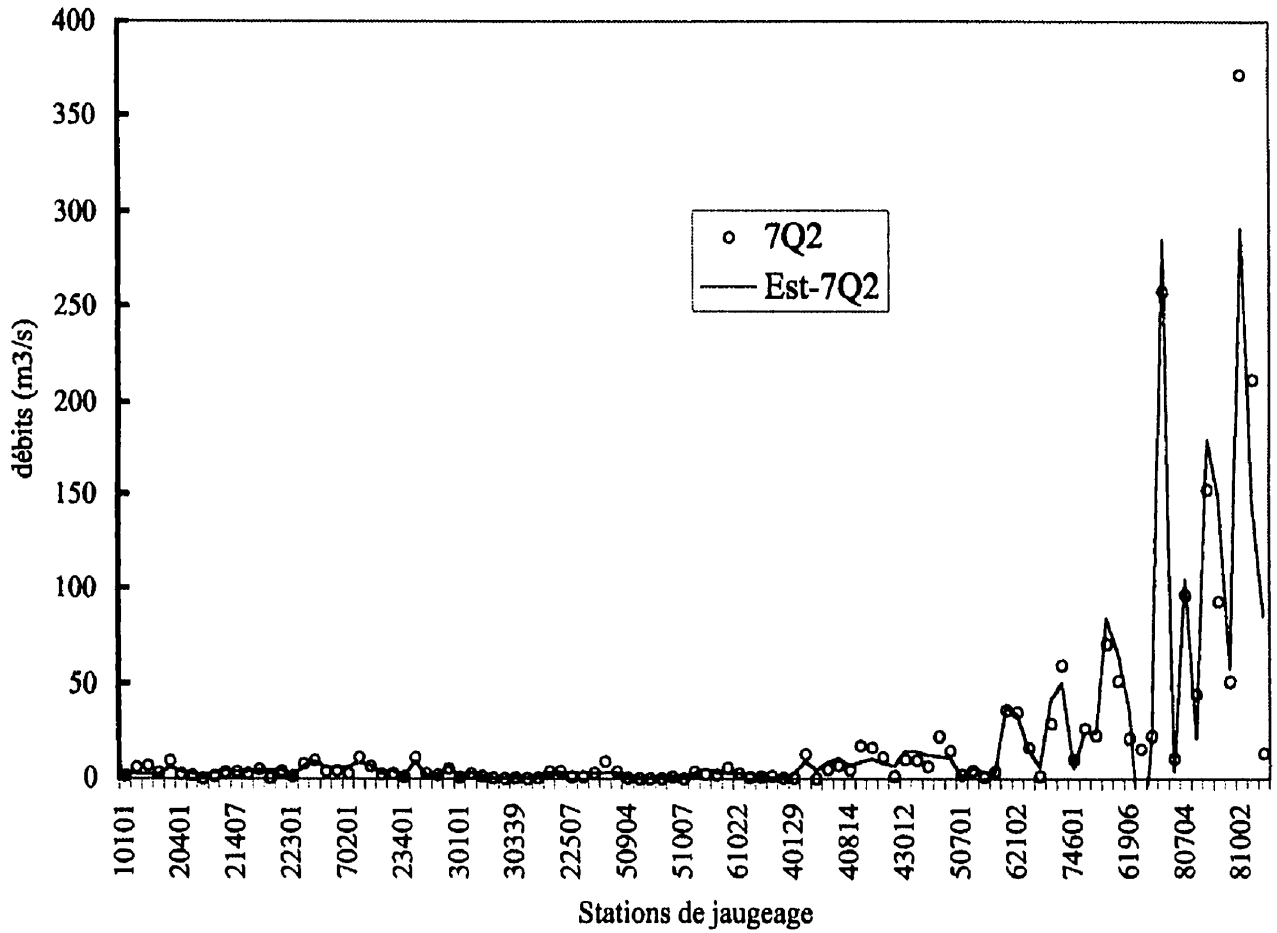

Figure 4.29: Variable 7Q2 estimée par modèle linéaire: PC1, PC2 et PC3

Tableau 4.13: Erreur moyenne associée aux modèles de régression

	Variables physiques		Composantes principales	
	Global-Physiques	8 Régions	Global	6 Régions
7Q10	25%	16%	15%	13%
30Q5	15%	6%	10%	7%
1Q10	13%	8%	9%	3%
7Q2	28%	17%	21%	18%
moyenne	20%	12%	14%	11%

4.5 Modèles de distribution et ajustement

L'exercice de modélisation, présenté à la section précédente 4.4, a permis d'identifier des relations linéaires entre les paramètres physiques d'un bassin et les variables hydrologiques 7Q2, 7Q10 et 30Q5. Pour être utilisées à des fins de prévision, ces variables doivent être associées à des fonctions de densité de probabilité normalisées permettant ainsi l'évaluation des étiages pour toute période de retour. Les données des 103 stations sélectionnées ont été normalisées selon les méthodes #1 et #2, c'est-à-dire par les relations (3.23) et (3.24), et les fonctions de répartition obtenues sont présentées aux figures 4.30 et 4.31 respectivement. Les moyennes, écart-types et coefficients de variation utilisés pour chacune des stations sont rapportées à l'annexe B .

L'inspection visuelle des courbes de distribution démontre une très forte similitude entre les 3 variables. En fait, pour les deux méthodes de transformation, les valeurs des débits normalisés sont virtuellement confondus pour une probabilité cumulative supérieure à 0.1. Par ailleurs, aux valeurs extrêmes (x minimum) une certaine divergence est observée.

La comparaison directe des deux figures est trompeuse puisque leur abcisse possède des échelles différentes. Ainsi, bien que la figure 4.31 montre une dispersion spatiale plus importante des points dans la portion de la courbe associée aux faibles débits, cette région implique uniquement 3 ou 4 points. A l'opposée, la divergence entre les variables est plus faible en valeur absolue, mais elle inclut un nombre important de données puisque cette tendance est observable à partir de $F(x')<0.07$.

Afin de justifier la fusion de ces trois variables standardisées en une seule fonction de répartition, il faut utiliser des critères plus objectifs que la simple inspection visuelle. Ici, cette décision est basée sur l'analyse, point à point, de l'écart-type entre les trois courbes de distribution et une statistique non-paramétrique, soit le test de Smirnov. Les résultats de la première analyse sont résumés au tableau 4.14 . On y constate que l'écart-type global entre les courbes est plus faible pour la méthode #1 de normalisation. Cependant, c'est aussi selon cette méthode qu'on trouve le grand écart-type ponctuel. Il est donc difficile de conclure uniquement à partir de cette analyse. Dans le cas de l'analyse fondée sur la statistique de Smirnov, les tests furent effectués par comparaison pairée de chacun des trois couples possibles (Min_{1j} vs Min_{7j}, Min_{1j} vs Min_{30j} et Min_{7j} vs Min_{30j}).

Tableau 4.14: Sommaire de l'écart-type ponctuel des distributions

	Méthode #1	Méthode #2
total	41.4	100.9
minimum	6.5×10^{-5}	7.8×10^{-4}
maximum	7.4×10^{-1}	3.6×10^{-1}

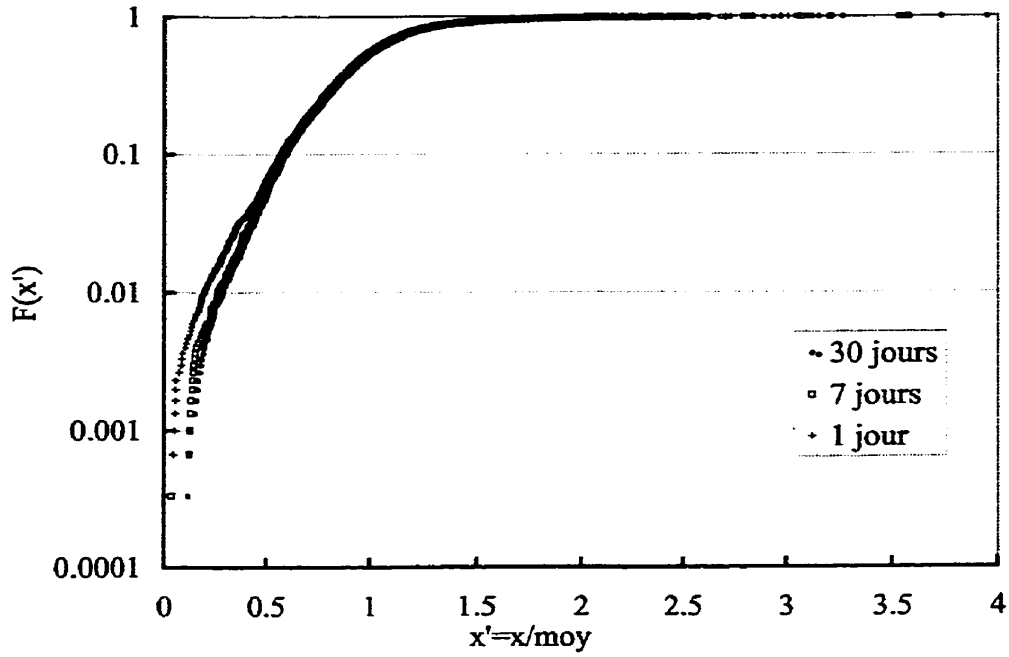

Figure 4.30: Distribution cumulative des minima annuels normalisés par μ

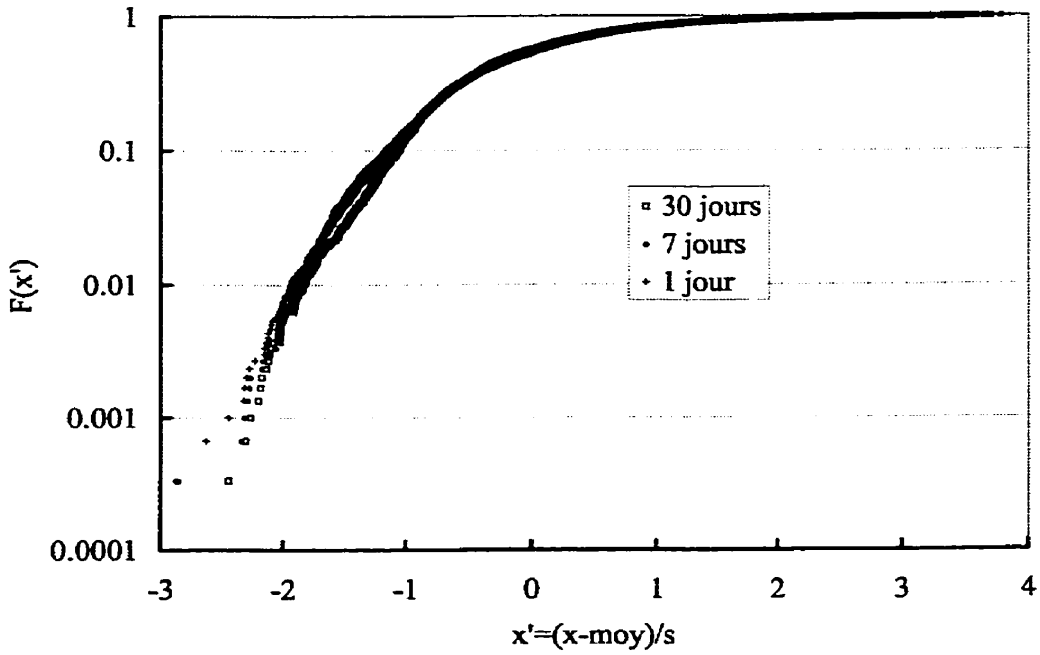

Figure 4.31: Distribution cumulative des minima annuels normalisés μ et σ

Le test a été appliqué sur les 3 distributions d'intérêt transformées par les méthodes #1 et #2. Les résultats de cette analyse, rapportés au tableau 4.15 confirme qu'on ne peut rejeter H_0 (les distributions sont identiques) à un niveau de 99%, et ce pour les deux types de méthodes de normalisation. En fait, le test de Smirnov montre que la méthode #1 devrait être favorisée par rapport à la méthode #2. Suite à ces résultats et compte tenu de l'analyse de l'écart-type entre les courbes, c'est la première méthode de normalisation qui a été retenue pour la modélisation de la distribution statistique des débits d'étiage. Ce choix peut paraître surprenant puisqu'il équivaut à dire que l'écart-type des variables Min1j, Min7j et Min30j est distribué de façon identique à leur moyenne. Cependant, dans le processus d'estimation des débits d'étiage pour des bassins sans données hydrologiques (donc à partir des variables physiques), la méthode de normalisation #1 permet d'effectuer le calcul à partir d'une seule corrélation, ce qui constitue un avantage important. La courbe de distribution globale ainsi obtenue est présentée à la figure 4.32.

Tableau 4.15: Résultats de l'analyse statistique par le test de Smirnov

	Méthode #1		Méthode #2		
	T_1	Accepte H_0	T_1	Accepte H_0	$T_{1,\ 1-\alpha}=0.99$
Min_{1j}-Min_{7j}	0.00498339	oui	0.0166113	oui	0.04192225
Min_{1j}-Min_{30j}	0.00531561	oui	0.03388704	oui	0.04192225
Min_{7j}-Min_{30j}	0.00531561	oui	0.02624585	oui	0.04192225

L'identification de la fonction de répartition associée à la distribution présentée à la figure 4.32 s'est effectuée par la comparaison de l'ajustement obtenu pour 4 fonctions différentes. Ces fonctions, présentées à la section 3.4, sont les distributions de Weibull, Pearson III, log-normale et gamma. Les courbes théoriques résultantes sont rapportées aux figures 4.33 à 4.36 ainsi que les données observées.

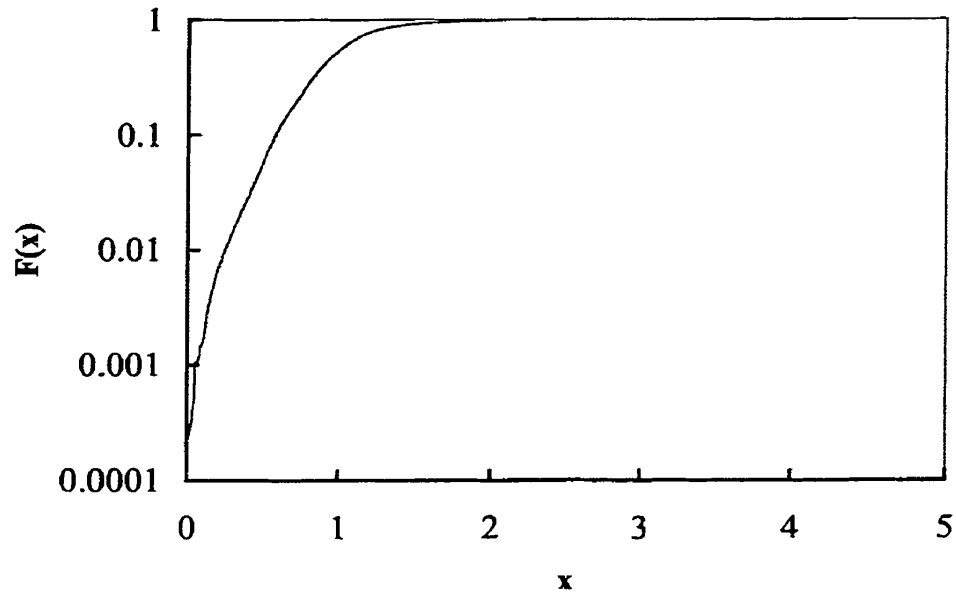

Figure 4.32: Distribution globale des débits normalisés –méthode #1

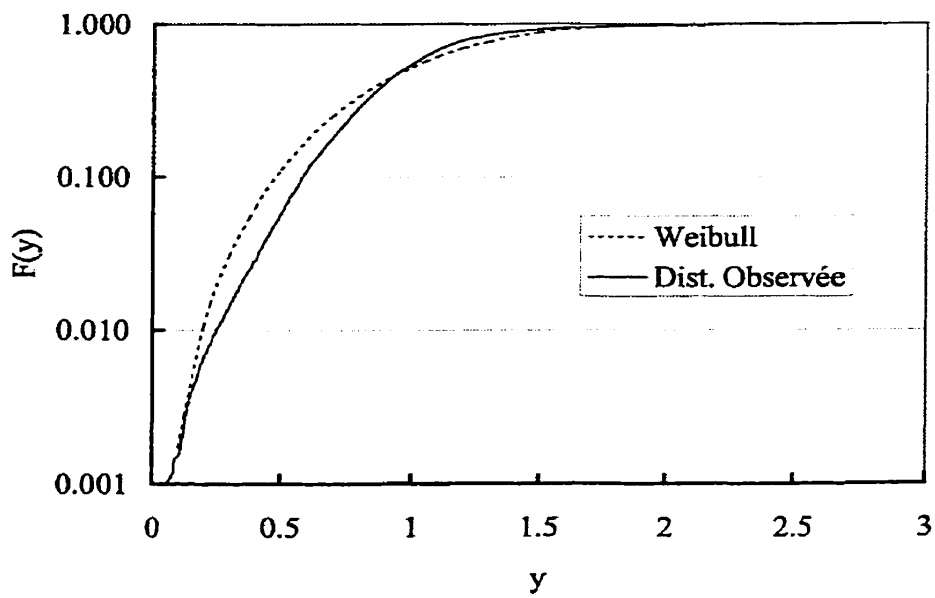

Figure 4.33: Ajustement de la distribution Weibull aux données observées

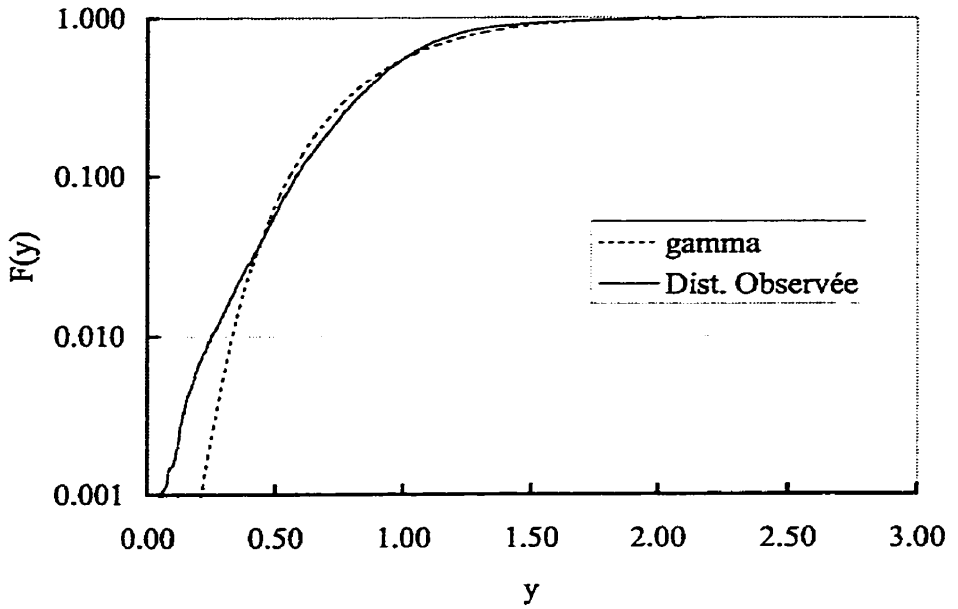

Figure 4.34: Ajustement de la distribution gamma aux données observées

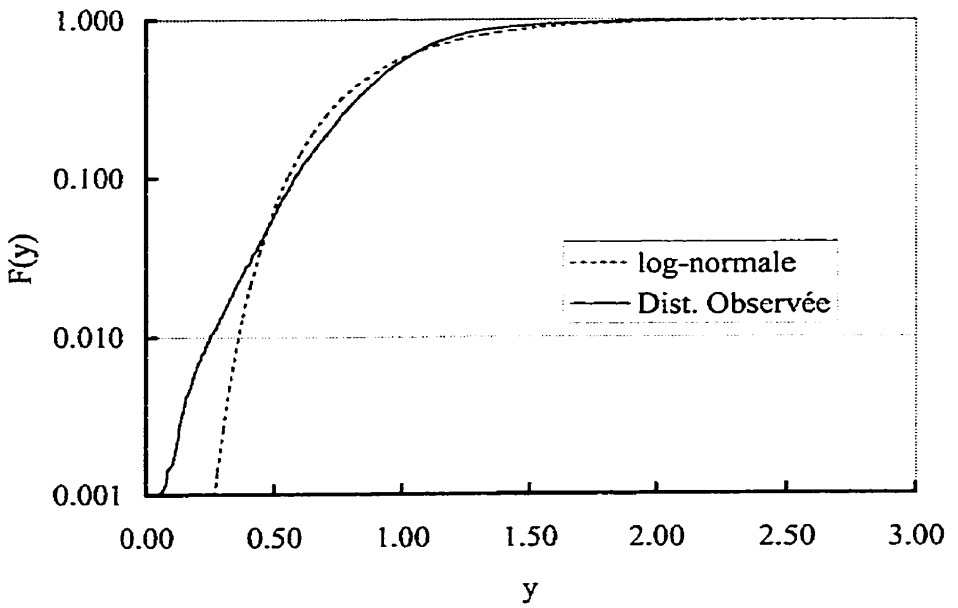

Figure 4.35: Ajustement de la distribution log-normale aux données observées

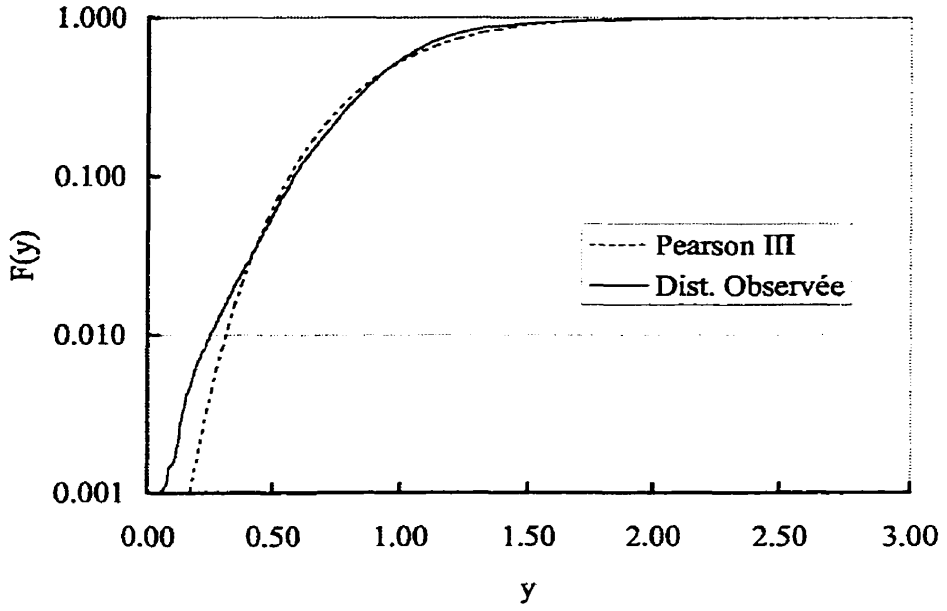

Figure 4.36: Ajustement de la distribution Pearson III aux données observées

L'analyse de ces figures permet d'identifier assez facilement la fonction la plus appropriée. Le tableau 4.16 rapporte la valeur des différents paramètres ajustés ainsi que le maximum de la fonction de vraisemblance obtenu pour chacune des quatre distributions. On y constate que la fonction la mieux ajustée est la distribution de Pearson III, suivie de près par la distribution gamma. Puisque la distribution de Pearson III renferme un paramètre supplémentaire, on pourrait être tenté d'utiliser la fonction gamma. Cependant, l'inspection visuelle des courbes obtenues montre que la distribution de Pearson III est nettement mieux ajustée que la distribution gamma dans la plage des périodes de retour comprises entre 10 et 100 ans. C'est donc la première qui doit être retenue comme outil de calcul principal à des fins prévisionnelles. Dans le cas de périodes de retour supérieures à 100 ans, la distribution de Weibull est la plus appropriée. Malheureusement, le fait qu'aucune des distributions étudiées ne s'ajuste de façon vraiment satisfaisante, rend difficile l'énoncé d'un modèle fiable permettant

l'extrapolation en dehors de la plage des observations. Toutefois, la similitude entre les données expérimentales et la distribution calculée pourrait peut-être s'améliorer par l'utilisation de fonctions de densité de probabilité généralisées pour valeurs extrêmes. Cependant, ceci se fera au profit de la simplicité puisque ces fonctions comprennent normalement un minimum de trois ou quatre paramètres ajustables. De plus, la différence apparente entre ces fonctions calculées et observées aux faibles valeurs de probabilités est amplifiée par l'échelle logarithmique.

Tableau 4.16: Sommaire de l'ajustement par maximum de vraisemblance

Distribution	paramètres de la distribution			fonction de vraisemblance
Weibull	*u* 1.11969	*k* 2.63		l(u,k) -4064.1
Gamma2	*u* 1	*k* 6.94753	*k/u* 6.94753	l(u,k) -3599.06
Pearson III	*u* 1.198	*k* 10.4667	a -0.1927	l(u,k,a) -3499.68
log-normale	*u* -0.0734849	*k* 0.164506		l(u,k) -4000

4.6 Application

4.6.1 Variable hydrologique connue.

Dans le cas d'un emplacement où une variable hydrologique telle 7Q10 ou 30Q5 est disponible, l'application directe des fonctions de densité de probabilité présentées dans la section 4.5 permet d'évaluer un débit d'étiage pour une période de retour désirée. Ces variables sont rapportées à l'annexe D. Ainsi, prenons à titre d'exemple la station 041301 située dans la région D. La valeur de 7Q10 à cette station est de:

$$7Q10\big|_{station\,041301} = 11.291 \text{ m}^3/\text{s} \tag{4.5}$$

On désire obtenir le débit d'étiage de 7 jours consécutifs avec une période de retour de 20 ans (7Q20). Dans un tel cas, on utilise d'abord la densité de probabilité de Pearson III (équation (3.45)) avec les paramètres ajustés rapportés au tableau 4.16. Ainsi il faut trouver y tel que:

$$P(y) = \int_0^y \left(\frac{k}{u}\right)^k (x-a)^{k-1} \frac{\exp\left(-\left(\frac{k(x-a)}{u}\right)\right)}{\Gamma(k)} dx \tag{4.6}$$

avec

k= 10.467

u=1.198

a=-0.1927

$P(y)$ = 1/20 = 0.05

$$P(y) = \int_0^y 6987.15(x+0.1927)^{9.467} e^{(-8.737(x+0.1927))} dx = 0.05 \tag{4.7}$$

l'équation intégrale se résoud numériquement et on trouve:

$$y|_{P(y)=0.05} = 0.4677 \tag{4.8}$$

Cependant, tel que discuté précédemment à la section 3.5, les données ajustées à la fonction de répartition sont normalisées selon l'équation (3.23). Ainsi on a :

$$y|_{P(y)=0.05} = 0.4677 = \frac{7Q20}{\bar{y}|_{station041301}} \tag{4.9}$$

Afin de déterminer $\bar{y}\big|_{station041301}$ on utilise la valeur de 7Q10 et la valeur de P(y) pour une période de retour de 10 ans:

$$y\big|_{P(y)=0.1} = 0.560 = \frac{7Q10}{\bar{y}\big|_{station041301}} = \frac{11.291\ \mathrm{m^3/s}}{\bar{y}\big|_{station041301}} \tag{4.10}$$

d'où

$$\bar{y}\big|_{station041301} = 19.873 \tag{4.11}$$

et enfin par (4.9) on calcule directement

$$7Q20\big|_{station041301} = 0.4677 \times 19.873 = 9.29\,\mathrm{m^3/s} \tag{4.12}$$

4.6.2 Aucune variable hydrologique connue

Lorsqu'on désire obtenir un estimé d'un débit d'étiage pour un emplacement où aucun enregistrement de données d'écoulement est disponible, il faut utiliser les corrélations linéaires développées aux sections 4.3 et 4.4 afin d'estimer une variable hydrologique telle 7Q10, 30Q5, 1Q10 ou 7Q2. Ici cette procédure est illustrée pour la station 080701, en faisant l'hypothèse qu'aucune donnée historique n'est disponible. Le but recherché est d'évaluer 7Q25 à cette station.

Dans un premier temps, il faut assembler les variables physiques à cette station. On les retrouve à la deuxième colonne du tableau 4.17. Dans le cas de la latitude et de la

longitude, les données brutes sont d'abord transformées en une forme adimensionnelle avant d'être normalisées, selon la relation (4.2).

$$long^* = \frac{long - long_{min}}{long_{max} - long_{min}} \times 100 = \frac{77:25:00 - 59:53:00}{79:05:00 - 59:53:00} \times 100 = 91.32 \qquad (4.13)$$

$$lat^* = \frac{lat - lat_{min}}{lat_{max} - lat_{min}} \times 100 = \frac{50:08:00 - 45:05:00}{51:31:00 - 45:05:00} \times 100 = 78.5 \qquad (4.14)$$

Par la suite, en utilisant les moyennes et écart-types globaux pour chacune de ces propriétés physiques (deuxième et troisième colonnes du tableau 4.17) on normalise ces valeurs selon l'équation (4.1). Les valeurs normalisées se retrouvent à la dernière colonne du tableau 4.17 ainsi que dans le tableau E.2 pour cette station.

Tableau 4.17: Calcul des propriétés physiques normalisées de la station 080701

	Valeur brute	moyenne	écart-type	Valeur normalisée
Superficie (km^2)	57498	3956	81[illegible]27	6.62
Longueur (km)	482.8	122.5	11[illegible]2.5	3.22
Pente (m/km)	0.2	4.7	4[illegible].9	-0.913
Précipitation (mm)	800	1014	1[illegible]18	-1.813
% Forêt	81	82.2	17.9	-0.07
% Eau et marais	17	6.6	7[illegible].6	1.374
Latitude	78.5 (50:08:00)	42.17	25.6	1.371
Longitude	91.32 (77:25:00)	60.35	21[illegible].22	1.47

A partir des variables physiques standardisées, il faut ensuite calculer les valeurs correspondantes des variables orthogonales PC1, PC2 et PC3. En utilisant les coefficients rapportés au tableau 4.5 pour l'analyse avec rotation equamax, l'application directe de l'équation 4.3 se traduit par:

$$\begin{pmatrix} 6.62 & 3.22 & -0.913 & -1.813 & -0.070 & 1.374 & 1.371 & 1.470 \end{pmatrix} \times \begin{pmatrix} 0.224 & -0.057 & -0.170 \\ 0.242 & -0.092 & 0.040 \\ -0.160 & -0.174 & -0.718 \\ -0.214 & -0.011 & -0.274 \\ -0.026 & -0.454 & -0.345 \\ 0.221 & 0.058 & -0.383 \\ 0.164 & -0.407 & 0.034 \\ 0.118 & 0.400 & -0.636 \end{pmatrix}$$

$$= \begin{pmatrix} 3.50 & -0.356 & -1.22 \end{pmatrix} \tag{4.15}$$

Avec les variables orthogonales, on peut maintenant évaluer à l'aide des modèles de régression, une variable hydrologique qui permettra ensuite d'utiliser la distribution théorique normalisée afin de calculer 7Q25. Pour cet exemple, la variable 7Q2 sera utilisée. En accord avec le tableau 4.12 (région F), le modèle linéaire correspondant s'exprime par:

$$\begin{aligned} 7Q2 &= \beta_0 + \beta_1 \text{ PC1} + \beta_2 \text{ PC2} + \beta_3 \text{ PC3} \\ 7Q2 &= -3.383 + 36.154 \times 3.50 - 148.769 \times -0.356 - 89.252 \times -1.22 = 286.15 \text{ m}^3/\text{s} \end{aligned} \tag{4.16}$$

A titre comparatif, notons que la valeur expérimentale de 7Q2 à la station 080701 est de 258.6 m^3/s soit une différence d'environ 10%. Puisqu'on recherche le débit d'étiage de 7 jours consécutifs avec une période de retour de 25 ans , on utilise maintenant l'équation (4.6) avec:

$P(y) = 1/25 = 0.04$

$$P(y) = \int_0^y 6987.15(x+0.1927)^{9.467} e^{(-8.737(x+0.1927))} dx = 0.04 \tag{4.17}$$

l'équation intégrale se résoud numériquement et on trouve:

$$y\Big|_{P(y)=0.04} = 0.4471 = \frac{7Q25}{\bar{y}\Big|_{station080701}} \tag{4.18}$$

Afin de déterminer $\bar{y}\Big|_{station80701}$ on utilise la valeur de 7Q2 et la valeur de $P(y)$ pour une période de retour de 2 ans:

$$y\Big|_{P(y)=0.5} = 0.9720 = \frac{7Q2}{\bar{y}\Big|_{station080701}} = \frac{286.15\ \mathrm{m^3/s}}{\bar{y}\Big|_{station080701}} \tag{4.19}$$

d'où

$$\bar{y}\Big|_{station080701} = 294.39\,\mathrm{m^3/s} \tag{4.20}$$

et enfin par (4.18) on calcule directement

$$7Q25\Big|_{station080701} = 0.4471 \times 294.39 = 131.62\ \mathrm{m^3/s} \tag{4.21}$$

4.7 Analyse générale des résultats

Les résultats dont ce chapitre fait état, doivent être analysés selon la perspective globale de ce travail, soit la modélisation régionale du phénomène d'étiage sur le territoire exploité de la province de Québec. A cette fin, il convient d'effectuer cet exercice selon deux perspectives: la régionalisation et la modélisation. A titre comparatif, les travaux

récents de Anctil *et al.* (2000) seront aussi considérés, particulièrement au niveau du processus de régionalisation.

L'exercice de régionalisation vise à regrouper, en un nombre limité de régions géographiques, des basssins versants ayant des variables hydrologiques statistiquement homogènes. Plus l'homogénéité de ces régions sera grande, meilleures seront les chances d'obtenir un modèle prédictif satisfaisant sur l'ensemble du territoire de la région. Bien qu'il soit évident qu'un partage du territoire en plusieurs régions favorise l'homogénéité de celles-ci, un trop grand nombre de régions se traduit aussi par un nombre équivalent de modèles de régression. Il en résulte donc une complication inutile du processus de modélisation ce qui limite l'intérêt pratique des modèles ainsi obtenus. L'évaluation de la modélisation proposée doit tenir compte de ces considérations. Le nombre de régions créées est volontairement contraint à une plage comprise entre un et huit, soit une seule région pour la province ou huit régions définies par le MEQ.

Lorsque la modélisation linéaire est effectuée globalement sur tout le territoire (une seule région) on peut comparer directement l'efficacité des modèles obtenus à partir des variables physiques (figures 4.16 à 4.19) et l'efficacité fondés sur les trois composantes principales de celles-ci (figures 4.26 à 4.29). Dans le cas de la régression effectuée sur les variables physiques, de nombreuses valeurs négatives sont obtenues pour les stations à très faible débit. Cependant, l'utilisation des composantes principales améliore significativement la qualité de l'ajustement et la plupart des estimés négatifs sont éliminés. Cette amélioration est aussi observée dans le sommaire des écarts rapportés au tableau 4.14. Par ailleurs, lorsqu'on poursuit cette comparaison entre les deux types de modélisation au niveau d'une approche régionale, on constate que la qualité de l'ajustement obtenu par la combinaison "huit régions (MEQ) – variables physiques" et l'ensemble "six régions (ce travail) – composantes principales" est pratiquement identique (voir tableau 4.13). Malgré les apparences, il ne faut pas prétendre pour autant que l'utilisation de trois composantes principales, par rapport à deux variables physiques

partiellement corrélées, ne comporte pas d'avantages. Si c'était le cas, la même tendance serait aussi observée dans l'approche globale. Il faut plutôt chercher l'explication dans la façon dont les régions furent définies.

Ainsi, l'étude des figures 4.20 à 4.25 nous apprend que les régions où la qualité de l'ajustement est particulièrement bon, telles les régions B et E sont aussi des régions dont les traces paramètriques présentées aux figures 4.12 à 4.14 sont aussi particulièrement homogènes. Cependant, ces même figures montrent aussi que certaines régions furent formées à partir de groupes de traces légèrement dispersées. Une contrainte importante imposée lors de l'établissement des nouvelles régions fut la continuité géographique de celles-ci. Aussi, dans certains cas, il a fallu inclure à une région des stations dont la trace était nettement éloignée du reste du groupe. Bien que visant à faciliter l'utilisation des modèles obtenus, cette contrainte est aussi un facteur qui réduit l'homogénéité des régions formées. Cela pourrait même éventuellement se traduire par la nécessité de fractionner le territoire en un plus grand nombre de régions afin d'obtenir des capacités prédictives adéquates pour les modèles établis.

Un autre facteur à considérer est le moment de l'année auquel le débit d'étiage est observé. Selon la position des stations de jaugeage sur le territoire, l'étiage annuel peut se produire en hiver ou en été. Dans le cadre de ce travail, nous avons considéré des variables physiques connues pour avoir un effet hydrologique important. Cependant, cet effet a été étudié en fonction des écoulements maximaux, rarement observés en période hivernale. Dans le cas d'un étiage, il est facile de concevoir que l'effet d'une variable physique, telle la pente moyenne d'un bassin versant, puisse être différent selon qu'on considère des conditions hivernales ou estivales. Or, on constate qu'effectivement pour la région D, où la qualité de l'ajustement du modèle linéaire avec composantes principales est particulièrement faible, on retrouve un mélange de stations ayant un étiage estival ou hivernal. On constate aussi des conditions similaires pour la région A.

L'influence saisonnière observée dans les résultats de cette étude est cohérente avec les travaux récents de Anctil *et al.* (2000) sur les étiages 7-jours de la province de Québec. Ces auteurs ont effectué une régionalisation des débits d'étiage à partir d'un test de discordance et d'un test d'hétérogénéité. Ils ont identifié six régions homogènes pour les étiages hivernaux et six autres régions pour les étiages estivaux. Contrairement à ce travail, ils ont observé qu'il était impossible d'obtenir des régions homogènes pour les débits d'étiage 1 jour et 30 jours. De plus, ils ont aussi ajusté une fonction de densité de probabilité à chacune des régions homogènes, alors que notre approche a montré qu'une fois normalisés, les étiages 1,7 et 30 jours étaient distribués de façon similaire pour l'ensemble des stations du territoire provincial.

CHAPITRE 5

CONCLUSION

Ce travail avait pour but d'effectuer la régionalisation des débits d'étiage pour la province de Québec. L'approche retenue a fait appel à huit propriétés physiques des bassins étudiés pour établir des régions homogènes. Il a été montré qu'une analyse en composantes principales permet de réduire le nombre de variables indépendantes nécessaires pour expliquer la variance observée. Une telle réduction est souhaitable afin de permettre l'utilisation de courbes paramétriques permettant l'identification de groupes de traces homogènes. Cet outil s'est avéré efficace puisque six nouvelles régions homogènes ont pu être établies à partir des huit régions administratives déterminées par le MEQ. Dans chacune de ces régions, les étiages 1 jour, 7 jours et 30 jours sont maintenant estimables à partir de modèles linéaires applicables aux trois premières composantes principales des variables physiques. Il a aussi été trouvé qu'une fois normalisés, les débits d'étiage sont distribués selon une loi Pearson III commune à toutes les régions, pour une période de retour maximale de cent ans.

Lors de la planification du processus de régionalisation, il avait été décidé que les nouvelles régions ne devaient pas être morcelées afin de faciliter l'utilisation des modèles obtenus. L'analyse des résultats mène à la conclusion, qu'en plus de cette contrainte géographique, on devrait aussi ajouter un critère assurant que toutes les rivières d'une même région ont leur étiage à la même période. En l'absence d'une telle mesure, un modèle linéaire peut ne pas être suffisamment adéquat pour permettre une estimation quantitative des débits d'étiages. Ce problème n'affecte pas les sites avec des débits historiques puisque dans ce cas, l'étiage moyen observé à la station permet, avec la loi de densité de probabilité, d'évaluer directement le débit d'étiage à une période de retour quelconque.

Malgré les limitations discutées au paragraphe précédent, les résultats de ce travail apportent à l'hydrologue praticien un moyen immédiat d'évaluer le débit d'étiage de la plupart des rivières du Québec ayant un intérêt économique important. Dans le cas où l'usage recherché demande un niveau de précision élevé, la fonction de densité de probabilité obtenue devrait, dans les régions A et D, être utilisée en conjonction avec des données directes plutôt qu'avec les modèles de régressions actuels.

RÉFÉRENCES

ACREMAN, M.C. et SINCLAIR, C.D. (1986). Classification of drainage basins according to their physical characteristics: an application for flood frequency analysis in Scotland. Journal of Hydrology, 84, 365-380.

ACREMAN, M.C. et WILTSHIRE, S.E. (1987). Identification of regions for regional flood frequency analysis, Eos Transactions, 68 (44), 1262.

ANCTIL, F., MARTEL, N. et HOANG, V.D. (1998). Analyse régionale des crues journalières de la province de Québec, Canadian Journal of Civil Engineering, 25, 360-369.

ANCTIL, F., LAROUCHE, W. et HOANG, V.D. (2000). Analyse régionale des étiages 7-jours de la province de Québec, Water Quality and Resources Journal of Canada, 35(1), 125-146.

ANDREWS, D.F. (1972) Plots of high dimensional data, Biometrics, 28, 125-136.

BEAUDOIN, P. (1980). Analyse des précipitations journalières dans la région de Montréal, Mémoire de maîtrise, Département de génie civil, École Polytechnique, Montréal, 180 pages.

BEECHER, H.A. (1990). Standards for instream flows, Rivers, 1(2) , 97-109.

BELZILE, L., BÉRUBÉ, P., HOANG, V.D. et LECLERC, M. (1997). Méthode échohydrologique de détermination des débits réservés pour la protection des habitats du poisson dans les rivières du Québec, Rapport préliminaire présenté par l'INRS-Eau et le

Groupe-conseil Génivar inc. au ministère de l'Environnement et de la Faune et au ministère des Pêches et des Océans Canada, 82 pages.

BOBÉE, B et ASHKAR, F. (1988). Generalized method of moments applied to LP3 distribution, Journal of Hydraulic Engineering, ASCE, 114, 899-909.

BOBÉE, B et ASHKAR, F. (1991). The gamma family and derived distributions applied in hydrology, Water Resources Publications, Fort Collins, Colorado, 203 pages.

BOUCHER, S., SPARKS, D. et ROUSSELLE, J. (1986). Analyses statistiques multidimensionnelles des caractéristiques physiographiques des bassins versants de l'Estrie et des Bois-Francs, École Polytechnique de Montréal, département de génie civil, 81 pages.

BOVEE, K.D. (1982). A guide to stream habitat analysis using the incremental methodology, U.S. Fish and Wildlife Service, Washington D.C., Instream Flow Information, Paper No.12 (FWS/OBS-82/26).

BURN, D.H. (1989). Cluster analysis as applied to regional flood frequency, Journal of Water Resources Planning and Management, 115(5), 567-582.

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

CAISSIE, D. et EL-JABI, N. (1995). Comparison and regionalization of hydrologically based instream flow techniques in Atlantic Canada, Canadian Journal of Civil Engineering, 22, 235-246.

CHOW, K. C. A. et WATT, W. E. (1992) Use of Akaike information criterion for selection of flood frequency distribution, Canadian Journal of Civil Engineering, 19, 616-626

CLARK, R. T. (1994) Statistical modelling in hydrology, first edition, John Wiley and Sons, New York, United States, 412 pages.

COLLINGS, M.R. (1972). A methodology for determining instream flow for fish, Proceedings of Instream Flow Methodology Workshop, Washington Department of Ecology, Olympia, Wash., p.72-86.

CONOVER, W.J. (1980) Practical nonparametric statistics, second edition, John Wiley and Sons, USA, 493 pages.

CUNNANE, C. (1988). Methods and merits of regional flood frequency analysis, Journal of Hydrology, 100, 269-290.

DALRYMPLE, T. (1960). Flood frequency methods, US Geological Survey, Water Supply Paper, 1543-A, 11-51.

DESFORGES, P. et TREMBLAY, R. (1974). Analyse de la fréquence des crues pour le Québec, Ministère des Richesses naturelles, Direction générale des eaux, Rapport H.P.-33, 46 pages.

DURRANS, S.R. et TOMIC, S. (1996). Regionalization of low-flow frequency estimates: an Alabama case study, Water Resources Bulletin, 32(1), 23-37.

ESTES, C. C. et ORSBORN, J. F (1986). Review and analysis of methods for quantifying instream flow requirements, Water Resources Bulletin, 22(3), 389-398.

FILL. H.D. et STEDINGER, J.R. (1995). Homogeneity tests based upon Gumbel distribution and a critical appraisal of Dalrymple's test, Journal of Hydrology, 166, 81-105.

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

GREHYS, (1996b). Inter-comparison of regional flood frequency procedures for Canadian rivers, Journal of Hydrology, 186, 85-103.

HAAN, C.T. (1977). Statistical Methods in hydrology, first edition,The Iowa State University Press, United States, 378 pages.

HAMILTON, R.E. et KOSAKOSKI, G.T. (1982). Water requirements for fisheries of the Englishman River, Vancouver Island, B.C. Canadian Manuscript Report of Fisheries and Aquatic Sciences No. 1676, 38 pages.

HAWLEY, M.E. et McCUEN, R.H. (1982). Water yield estimation in western United-States. Journal of Irrigation and Drainage Division, ASCE, 108, 25-34.

HOANG, V.D. et TREMBLAY, R. (1976). Estimation des débits d'étiage d'été des rivières du Québec méridional, Service de l'hydrométrie, Ministère des Richesses naturelles, Direction générale des eaux, 63 pages.

HOSKING, J.M.R. (1990). L-moments: analysis and estimation of distributions using linear combination of order statistics, Journal of the Royal Statistical Society, 52, 105-124.

HOSKING, J.M.R. et WALLIS, J.R. (1993). Some statistics useful in regional frequency analysis, Water Resources Research, 29(2), 271-281.

LAWAL, S., A. et WATT, W., E. (1996a) Non-zero lower limit in flow frequency analysis?, Water Resources Bulletin, 32 (6), 1159-1166

LAWAL, S., A. et WATT, W., E. (1996b) Frequency analysis of low flows using the Akaike information criterion, Canadian Journal of Civil Engineering, 23, 1180-1189

LIU, S., LU, J.C. et UNAL, C. (1996). Analysis of bivariate censored low flows, Journal of Hydraulic Engineering, 122(2), 97-103.

LLAMAS, J. (1993). Hydrologie générale, principes et applications, 2 ième édition, Gaëtan Morin éditeur, 656 pages.

LOGANATHAN, G. V., KUO, C. Y., et McCORMIK, T. C. (1985) Frequency analysis of low flows, Nordic Hydrology, 16, 105-128

LOAR J.M. et SALE, M.J. (1981). Analysis of environmental issues related to small-scale hydroelectric development. V. Instream flow needs for fishery resources. Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tenn., Publication No. 1829, 54 pages.

MINITAB Inc. (1995). Minitab reference manual – Release 10 Xtra, Minitab Inc., USA .

MONTGOMERY, D. C. (1984) Design and analysis of experiments, 2ed, John Wiley and Sons, New York., 418 pages.

MORIN, G., FORTIN, J.P., SOCHANSKA, W., LARDEAU, J.P. et CHARBONNEAU, R. (1979). Use of principal component analysis to identify homogeneous precipitation stations for optimal interpolation, Water Resources Research, 15(6), 1841-1850.

MOSLEY, M.P. (1981). Delimitation of New Zealand hydrologic regions, Journal of hydrology, 49, 173-192.

NATHAN, R.J. et McMAHON, T.A. (1990). Identification of homogeneous regions for the purposes of regionalisation, Journal of Hydrology, 121, 217-238.

NGUYEN, V. et PANDEY, G.R. (1996). A new approach to regional estimation of floods in Québec, Rapport de la 49 ième conférence annuelle, ACRH, 587-596.

POTTER, K.W. (1987). Research on flood frequency analysis: 1983-1986, Review of Geophysics, 25(2), 113-118.

PROULX, H., JACQUES, G., LAMOTHE, A. M., et LITYNSKI, J., (1987). Climatologie du Québec méridional. Environnement Québec, Québec, rapport MP-65, 72 pages

RASMUSSEN, P., OUARDA, T. et BOBÉE, B. (1995). Méthodologie de rationalisation du réseau hydrométrique du Québec, Rapport de recherche de l'INRS-Eau, no R456, 99 pages.

RIBEIRO-CORREA, J. et ROUSSELLE, J. (1993). A hierarchical and empirical bayes approach for the regional Pearson type III distribution, Water Resources Research, 29(2), 435-444.

RIBEIRO-CORREA, J., CAVADIAS, G.S., CLÉMENT, B. et ROUSSELLE, J. (1995). Identification of hydrological neighbourhoods using canonical correlation analysis, Journal of Hydrology, 173, 71-89.

ROUSSELLE, J., WATT, W. E., LATHEM, K. W., NEILL, C. R. et RICHARDS, T. L. (1990). Hydrologie des crues au Canada, Guide de planification et de conception, Conseil national de recherches Canada, 277 pages.

SIEW-YAN-YU, T.O., ROUSSELLE, J., JACQUES, G. et NGUYEN, V.T.V. (1985). Régionalisation du régime des précipitations dans la région des Bois-Francs et de l'Estrie par l'analyse en composantes principales, Revue canadienne de génie civil, 25(6), 1050-1058.

TASKER, G.D. (1982). Comparing methods of hydrological regionalisation, Water Resources Bulletin, 18(6), 965-970.

TESSIER, D. (1983). Analyses statistiques des caractéristiques physiographiques de petits bassins versants en relation avec les débits de crue, Mémoire de Maîtrise, Département de génie civil, École Polytechnique, Montréal, 189 pages.

THOMAS, D.M. et BENSON, M.A. (1970). Basin characteristics, U.S. Geological Survey Water Supply Paper 1975, 71 pages

TOMIC, S. (1994). Regional low-flow frequency analysis for Alabama rivers, MS Thesis, Department of Civil Engineering, The University of Alabama, Tuscaloosa, Alabama, 131 pages.

VOGEL, R.M., et FENNESSEY, N.M. (1993). L-moment diagrams should replace product moment diagrams, Water Resources Research, 29, 1745-1752.

WATT, W. E., LATHEM, K. W., NEILL, C. R., RICHARDS, T. L. et ROUSSELLE, J. (1989). Hydrology of Floods in Canada: A guide to Planning and Design. National Reasearch Council, Ottawa, pp. 45-74

YEVYEVICH, V.C. (1972). Probability and statistics in hydrology, Water Resources Publications, FT. Collins, Colorado, 302 pages.

ANNEXE A: SOMMAIRE DES STATIONS ETUDIEES

Tableau A.1: Sommaire des stations étudiées

Description	station	Superficie (km2)	Période d'observation	Années d'observation
GRANDE RIVIERE A 2,3 KM EN AMONT DE LA BRANCHE OUEST	010101	534	1973-1979 1982-1994	20
BONAVENTURE A 9,7 KM EN AMONT DU PONT-ROUTE 132	010801	2130	1966-1967 1969-1982 1984-1994	27
PETITE RIVIERE CASCAPEDIA A 3,9 KM EN AMONT DU PONT DU CN	010901	1410	1962-1980 1982-1994	32
CASCAPEDIA A 2,5 KM EN AVAL DU RUISSEAU BERRY	011003	1680	1979-1994	16
NOUVELLE AU PONT-ROUTE 132 A NOUVELLE	011201	1140	1965-1994	30
MATAPEDIA A 1,0 KM EN AMONT DE L'ASSEMETQUAGAN	011507	2760	1970-1991	22
YORK A 4,3 KM EN AMONT DU PONT-ROUTE 132 A SUNNY BANK	020401	1010	1946-1979	34
YORK A 1,4 KM EN AVAL DU RUISSEAU DINNER ISLAND	020404	659	1981-1994	17
AU RENARD A 4,7 KM DE L'EMBOUCHURE	020502	59.8	1978-1979 1982-1986 1989-1996	15
DARTMOUTH A 1,6 KM EN AMONT DU RUISSEAU DU PAS DE DAME	020602	645	1970-1994	25
MADELEINE A 1,5 KM EN AMONT DU PONT-ROUTE 132 A RIVIERE-LA-MADELEINE	020802	1200	1955-1966 1968-1990 1992-1994	38
SAINTE-ANNE A 9,7 KM EN AMONT DU PONT-ROUTE 132	021407	774	1974-1994	21
DU CAP CHAT A 5,3 KM EN AMONT DU PONT-ROUTE 132 A CAP-CHAT	021502	721	1967-1994	28
MATANE A 8,5 KM EN AMONT DU PONT-ROUTE 132 A MATANE	021601	1650	1927-1939 1941-1995	68
BLANCHE A 3,5 KM EN AMONT DU PONT ROUTE 132	021702	208	1967-1995	29
NEIGETTE A 3,7 KM DE LA MITIS	021915	494	1975-1995	21
RIMOUSKI A 3,7 KM EN AMONT DU PONT-ROUTE 132	022003	1610	1963-1995	33
DES TROIS PISTOLES A 2,1 KM EN AMONT DU PONT-ROUTE 132	022301	932	1924-1995	72
DU LOUP A 10,8 KM EN AMONT DU PONT-ROUTE 289	022504	500	1957-1978	22

Tableau A.1: Sommaire des stations étudiées...suite

Description	station	Superficie (km2)	Période d'observation	Années d'observation
DU LOUP AU PONT-ROUTE A SAINT-JOSEPH-DE-KAMOURASKA	022507	525	1979-1995	17
OUELLE AU PONT-ROUTE A SAINT-GABRIEL-DE-KAMOURASKA	022703	783	1971-1980 1983-1995	23
DU SUD A 1,0 KM EN AMONT DU PONT-ROUTE A ARTHURVILLE	023106	821	1967-1995	29
ETCHEMIN A 8,7 KM EN AMONT DU PONT-ROUTE 173 A SAINT-HENRI-DE-LEVIS	023301	1130	1928-1978 1981-1995	66
BEAURIVAGE A 1,1 KM EN AVAL DU PONT-ROUTE 171 A SAINT-ETIENNE	023401	709	1926-1995	70
CHAUDIERE AU PONT-ROUTE 218 A ST-LAMBERT	023402	5820	1921-1921 1923-1925 1937-1995	63
FAMINE A 6,3 KM EN AMONT DU PONT-ROUTE 173 A SAINT-GEORGES	023422	691	1965-1990 1993-1995	29
BECANCOUR A 0,8 KM EN AVAL DU PONT-ROUTE 116 A LYSTER	024001	1410	1923-1968	46
BECANCOUR A 2,1 KM EN AMONT DE LA PALMER	024003	919	1967-1968 1970-1995	28
BECANCOUR AU PONT-ROUTE PRES DE SAINT-SYLVERE	024007	2330	1970-1992 1994-1995	25
NICOLET SUD-OUEST AU PONT-ROUTE 255 EN AVAL D'ASBESTOS	030101	549	1930-1991 1993-1995	65
NICOLET A 5,8 KM EN AVAL DE LA BULSTRODE	030103	1540	1967-1995	29
COATICOOK A 1,3 KM EN AVAL DU PONT-ROUTE 143 A WATERVILLE	030215	519	1960-1970 1972-1980 1982-1992 1995	32
EATON A 0,6 KM DE LA SAINT-FRANCOIS	030234	642	1954-1995	42
EATON EN AVAL DU PONT-ROUTE 210 A SAWYERVILLE	030242	190	1967-1981	15
AU SAUMON A 1,9 KM EN AMONT DE LA MOFFAT	030282	738	1975-1995	21

Tableau A.1: Sommaire des stations étudiées ...suite

Description	station	Superficie (km2)	Période d'observation	Années d'observation
NOIRE A 7,6 KM DE LA YAMASKA	030304	1490	1966-1973 1975-1995	29
YAMASKA NORD A 0,2 KM EN AMONT DU PONT-ROUTE A SHEFFORD-VALE	O30309	139	1969-1971 1975-1995	24
YAMASKA SUD-EST AU PONT-ROUTE 202 A COWANSVILLE	030314	210	1969-1970 1979-1995	19
DAVID AU PONT-ROUTE A SAINT-DAVID	030316	316	1970-1995	26
YAMASKA NORD A 0,1 KM EN AVAL DU PONT-ROUTE 243 PRES DE WARDEN	030335	59.3	1979-1982 1984-1995	16
NOIRE A 4,1 KM EN AMONT DU RUISSEAU RUNNELS	030339	595	1982-1996	15
RICHELIEU AUX RAPIDES FRYERS (2OJ7)	030401	22000	1938-1986 1988-1990	52
DES HURONS AU PONT-ROUTE 229 EN AVAL DU RUISSEAU SAINT-LOUIS	030415	309	1974-1995	22
AUX BROCHETS A 0,7 KM EN AVAL DU RUISSEAU GROAT	030420	404	1982-1996	15
CHATEAUGUAY A 0,3 KM EN AVAL DU PONT-ROUTE 205	030901	2440	1921-1971	51
CHATEAUGUAY A 4,3 KM EN AMONT DU PONT-ROUTE 138	030905	2490	1971-1995	25
RUISSEAU HALL A 0,4 KM DE LA CONNECTICUT (20E18)	031401	219	1949-1990	42
DONCASTER A 0,3 KM EN AVAL DE LA DECHARGE DE LAC ELEVE	040129	109	1982-1996	15
ROUGE AU PONT DU CP PRES DE LA MACAZA	040201	2570	1931-1981	51
ROUGE A 0,2 KM EN AVAL DE LA CENTRALE DE LA CHUTE BELL	040202	5440	1919-1930 1945-1947 1950-1957 1966-1988 1990-1995	52
RUISSEAU SAINT-LOUIS A 0,3 KM DE LA RIVIERE DU DIABLE	040212	39.9	1969-1995	27
DE LA PETITE NATION AU PONT-ROUTE PRES DE COTE-SAINT-PIERRE	040401	1300	1926-1971	69

Tableau A.1: Sommaire des stations étudiées...suite

Description	station	Superficie (km2)	Période d'observation	Années d'observation
			1973-1995	
DESERT A 4,0 KM EN AMONT DE LA RIVIERE DE L'AIGLE	040810	1780	1927-1930 1960-1971	16
PICANOC A 0,5 KM EN AMONT DU PONT-ROUTE 105 PRES DE WRIGHT	040814	1290	1927-1930	24
			1976-1995	
COULONGE A 8,4 KM EN AMONT DU PONT-ROUTE 148 PRES DE FORT-COULONGE	041301	5150	1927-1930	62
			1933	
			1936-1992	
DUMOINE A LA SORTIE DU LAC ROBINSON	041902	3760	1966-1995	30
DUMOINE A LA SORTIE DU LAC DUMOINE	041903	2110	1968-1995	28
MAGANASIPI A LA SORTIE DU LAC JOHNSON	042103	575	1971-1989	22
			1992	
			1994-1995	
KINOJEVIS A 0,2 KM EN AMONT DU PONT-ROUTE A CLERICY	043012	2590	1968	25
			1971-1992	
			1994-1995	
VERMILLON A 2,3 KM DU SAINT-MAURICE	050116	2670	1929-1954	53
			1956-1965	
			1967-1983	
MEKINAC A 7,1 KM DU SAINT-MAURICE	050118	997	1929-1941	38
			1943-1966	
			1971-1971	
MATAWIN A 0,3 KM EN AVAL DU BARRAGE MATAWIN	050120	4070	1940-1979	49
			1984-1986	
			1988-1988	

Tableau A.1: Sommaire des stations étudiées ...suite

Description	station	Superficie (km2)	Période d'observation	Années d'observation
			1990-1994	
CROCHE A 2,6 KM EN AVAL DU RUISSEAU CHANGY	050135	1570	1966-1995	30
BATISCAN A 3,4 KM EN AVAL DE LA RIVIERE DES ENVIES	050304	4480	1968-1995	28
BATISCAN A 9,7 KM EN AMONT DE LA RIVIERE A PIERRE	050305	2870	1968-1982	15
SAINTE-ANNE A 4,8 KM EN AMONT DE LA SAINTE-ANNE (BRAS DU NORD)	050402	751	1952-1960	19
			1962-1971	
SAINTE-ANNE A 1,3 KM EN AVAL DU BARRAGE A CHUTE-PANET	050408	1550	1966-1995	30
SAINTE-ANNE (BRAS DU NORD) A 4,5 KM EN AMONT DE LA MAUVAISE	050409	642	1966-1995	30
PORTNEUF A 2,9 KM DU SAINT-LAURENT PRES DE PORTNEUF	050701	355	1967-1982	28
			1984-1995	
SAINT-CHARLES A 0,8 KM EN AMONT DE LA LORETTE	050904	357	1969-1995	27
MONTMORENCY A 0,6 KM EN AVAL DU BARRAGE DES MARCHES NATURELLES	051001	1100	1926-1938	42
			1965-1980	
			1982-1994	
RUISSEAU DES EAUX VOLEES A 0,2 KM EN AMONT DU CHEMIN DU BELVEDERE	051002	3.94	1966-1994	29
RUISSEAU DES EAUX VOLEES A 0,2 KM DE LA MONTMORENCY	051003	9.17	1967-1994	28
RUISSEAU DES AULNAIES OUEST A 0,2 KM EN AMONT DU CHEMIN BELVEDERE	051004	1.22	1967-1982	26
			1984-1993	
MONTMORENCY A 0,3 KM EN AMONT DE LA BLANCHE	051005	269	1967-1981	15
NOIRE A 5,5 KM DE LA MONMORENCY	051006	58.8	1967-1981	15
RUISSEAU DES AULNAIES A 30 M DU RUISSEAU DES EAUX VOLEES	051007	3.57	1972-1993	22
DU GOUFFRE A 1,6 KM EN AMONT DU PONT-ROUTE 362 A BAIE-SAINT-PAUL	051301	865	1968-1994	27
MALBAIE A 0,3 KM EN AVAL DU PONT-ROUTE A CLERMONT	051502	1700	1968-1994	27
L'ASSOMPTION AU PONT-ROUTE 50 A JOLIETTE	052219	1340	1970-1995	26
DE L'ACHIGAN AU PONT-ROUTE 341 A L'EPIPHANIE	052233	647	1980-1995	16
MASKINONGE AU PONT DU C.N. PRES DE SAINT-URSULE	052601	1030	1926-1971	62
			1980-1995	

Tableau A.1: Sommaire des stations étudiées ...suite

Description	station	Superficie (km2)	Période d'observation	Années d'observation
DU LOUP A 0,3 KM EN AVAL DU RUISSEAU CARUFEL	052805	774	1966-1995	30
PETIT SAGUENAY A 3,1 KM EN AMONT DU PONT-ROUTE 170	060101	736	1975-1994	20
AUX ECORCES A 0,6 KM EN AMONT DU PONT-ROUTE 169	061020	1110	1972-1195	24
PIKAUBA A 8,2 KM EN AMONT DE L'UPICA	061022	495	1970-1982 1985-1985 1988-1995	22
METABETCHOUANE A 1,8 KM EN AMONT DE LA CENTRALE S.R.P.C.	061502	2280	1978-1995	18
PETITE PERIBONCA A 0,2 KM EN AMONT DU PONT-ROUTE 169	061801	1090	1976-1995	20
CHAMOUCHOUANE A LA TETE DE LA CHUTE AUX SAUMONS	061901	15300	1963-1995	33
CHAMOUCHOUANE A 8,5 KM EN AVAL DE LA DU CHEF	061905	11100	1963-1995	33
CHAMOUCHOUANE A 0,2 KM EN AMONT DU PONT-ROUTE 167	061906	4330	1963-1992	30
MISTASSIBI A 1,6 KM DE LA MISTASSINI	062101	9320	1961 1963-1995	34
MISTASSINI A 3,4 KM EN AMONT DE LA MISTASSIBI	062102	9870	1963-1995	33
MANOUANE A LA SORTIE DU LAC DUHAMEL	062209	3600	1980-1995	16
VALIN A 3,5 KM DE L'EMBOUCHURE	062701	746	1975-1995	21
SAINTE-MARGUERITE NORD-EST A 1,8 KM DE LA SAINTE-MARGUERITE	062802	1100	1977-1994	18
DES ESCOUMINS A 3,1 KM EN AMONT DU PONT-ROUTE 138 A ESCOUMINS	070201	793	1955-1968 1970-1970 1982-1983	17
PORTNEUF A L'AMONT DES CHUTES PHILIAS	070401	2010	1974-1986 1988-1992	18
GODBOUT A 1,6 KM EN AMONT DU PONT-ROUTE 138	071401	1570	1975-1994	20
DES RAPIDES A 3,1 KM DE L'EMBOUCHURE	072201	557	1955-1982	28
MOISIE A 5,1 KM EN AMONT DU PONT DU Q.N.S.L.R.	072301	19000	1966-1966 1968-1994	28

Tableau A.1: Sommaire des stations étudiées ...suite

Description	station	Superficie (km2)	Période d'observation	Années d'observation
MAGPIE A 2,3 KM DE L'EMBOUCHURE	073501	7610	1955-1958 1966-1978	17
ROMAINE AU PONT DE LA Q.I.T. A 14,8 KM DE L'EMBOUCHURE	073801	13000	1957-1959 1961-1994	37
NABISIPI A 2,4 KM DE L'EMBOUCHURE	074601	2060	1963-1987	25
AGUANUS A 1,4 KM EN AMONT DE LA DANY	074701	5590	1963-1981	19
ETAMAMIOU A 17,5 KM DE L'EMBOUCHURE	075601	2950	1975-1992	18
SAINT-PAUL EN FACE DU RUISSEAU CHANION	076601	6630	1968-1971 1973-1980 1982-1989 1991-1994	24
HARRICANA A 3,4 KM EN AVAL DU PONT-ROUTE 111 A AMOS	080101	3680	1934-1995	62
TURGEON A 14,3 KM DE L'HARRICANA	080104	11200	1972-1974 1976-1989 1993-1993	18
NOTTAWAY A LA TETE DU LAC SOSCUMICA	080701	57500	1961-1981	21
BELL A 5,0 KM EN AMONT DU PONT DU C.N. A SENNETERRE	080703	2010	1928-1963	36
WASWANIPI A 0,5 KM EN AVAL DE L'OPAWICA	080704	18700	1963-1981	19
MEGISCANE A 2,3 KM EN AVAL DU PONT DU C.N. PRES DE MEGISCANE	080705	8310	1925-1940 1955-1963 1965-1965	26
WASWANIPI A LA TETE DE LA CHUTE ROUGE	080718	31900	1968-1989 1991-1995	27
BROADBACK A 1,6 KM EN AVAL DE LA OUASOUAGAMI	080801	17100	1962-1981	20
BROADBACK A LA SORTIE DU LAC KENONISKA	080809	9820	1973-1978 1981-1991	20

Tableau A.1: Sommaire des stations étudiées ...suite

Description	station	Superficie (km2)	Période d'observation	Années d'observation
DE RUPERT A 11,4 KM EN AVAL DU LAC NEMISCAU	081002	40900	1993-1995 1964-1993 1995-1995	31
DE RUPERT ET CHENAL CHIPASTOUC AUX SORTIES DU LAC MISTASSINI	081007	18100	1970-1992	23
PONTAX A 60,4 KM DE L'EMBOUCHURE	081101	6090	1976-1995	20

ANNEXE B: STATISTIQUES DES DEBITS MINIMUMS

Tableau B.1: Statistiques des débits minimums (annuel, 1 jour)

No de la station	Nombre d'années d'observation	Valeur minimale	Valeur maximale	Moyenne	Écart-type	Coefficient de variation
010101	20	0.64	2.08	1.21	0.43	0.35
010801	27	3.51	10.10	6.12	1.84	0.30
010901	32	0.86	6.24	4.15	1.23	0.30
011003	16	4.50	8.05	6.66	1.12	0.17
011201	30	0.96	5.03	3.06	1.06	0.35
011507	22	4.90	16.70	9.25	2.84	0.31
020401	34	0.57	3.60	2.01	0.68	0.34
020404	14	0.75	2.78	1.63	0.56	0.34
020502	15	0.08	0.21	0.15	0.04	0.30
020602	25	0.71	2.89	1.27	0.43	0.34
020802	38	1.98	5.64	3.44	1.06	0.31
021407	21	1.90	3.96	3.14	0.64	0.20
021502	28	1.60	3.72	2.46	0.57	0.23
021601	68	2.33	9.54	5.05	1.43	0.28
021702	29	0.02	0.48	0.26	0.12	0.46
021915	21	0.21	1.65	0.82	0.36	0.44
022003	33	1.61	5.79	3.64	1.24	0.34
022301	72	0.34	2.97	1.21	0.49	0.41
022504	22	0.11	1.11	0.59	0.31	0.52
022507	17	0.44	1.53	0.83	0.31	0.38
022703	23	0.27	2.46	0.94	0.59	0.63
023106	29	0.10	4.08	1.93	0.88	0.45
023301	66	0.65	5.80	2.36	0.96	0.41
023401	70	0.09	2.04	0.75	0.44	0.59
023402	63	3.00	20.40	10.15	3.60	0.35
023422	29	0.04	1.08	0.49	0.25	0.52
024001	46	0.57	5.07	2.23	1.00	0.45
024003	28	0.68	4.34	1.97	1.04	0.53
024007	25	1.63	11.70	5.25	2.41	0.46
030101	65	0.03	1.53	0.51	0.38	0.74
030103	29	0.85	5.35	2.49	1.20	0.48
030215	32	0.30	2.69	1.14	0.55	0.48
030234	42	0.46	2.27	0.98	0.38	0.39
030242	15	0.02	0.51	0.21	0.14	0.68
030282	21	0.53	1.95	1.11	0.37	0.33
030304	29	0.27	6.09	1.46	1.22	0.84
030309	24	0.01	1.42	0.77	0.44	0.57

Tableau B.1: Statistiques des débits minimums...suite

No de la station	Nombre d'années d'observation	Valeur minimale	Valeur maximale	Moyenne	Écart-type	Coefficient de variation
030314	19	0.01	0.35	0.16	0.11	0.66
030316	26	0.01	0.38	0.16	0.10	0.61
030335	16	0.03	0.20	0.10	0.05	0.51
030339	15	0.01	1.05	0.30	0.32	1.06
030401	52	39.90	235.00	98.55	39.09	0.40
030415	22	0.03	0.27	0.11	0.06	0.58
030420	15	0.00	0.62	0.26	0.20	0.75
030901	51	0.74	6.94	3.41	1.56	0.46
030905	25	0.71	6.75	3.49	1.70	0.49
031401	42	0.11	0.69	0.31	0.14	0.46
040129	15	0.22	0.65	0.38	0.11	0.29
040201	51	5.52	22.00	12.13	3.69	0.30
040202	52	3.34	48.40	22.31	9.42	0.42
040212	27	0.01	0.22	0.10	0.05	0.47
040401	69	0.34	10.40	4.68	2.04	0.44
040810	16	3.06	12.60	6.52	2.45	0.38
040814	24	2.63	10.80	4.87	2.06	0.42
041301	62	6.20	37.10	17.47	5.60	0.32
041902	30	6.68	27.10	16.27	5.33	0.33
041903	28	5.73	15.40	10.61	2.41	0.23
042103	22	0.29	2.59	1.32	0.63	0.48
043012	25	7.80	12.20	9.74	1.27	0.13
050116	53	4.25	18.10	9.54	3.01	0.32
050118	38	0.59	5.83	1.97	1.21	0.61
050120	49	0.63	7.00	2.13	1.03	0.48
050135	30	3.51	9.91	6.48	1.39	0.21
050304	28	14.40	36.20	21.76	4.71	0.22
050305	15	10.40	23.40	15.11	3.96	0.26
050402	19	0.91	4.25	2.65	1.05	0.40
050408	30	6.97	15.20	9.40	1.72	0.18
050409	30	2.97	5.66	3.70	0.57	0.15
050701	28	0.52	2.66	1.70	0.51	0.30
050904	27	0.04	1.95	0.36	0.41	1.15
051001	42	2.58	7.22	4.89	1.28	0.26
051002	29	0.02	0.05	0.03	0.01	0.20
051003	28	0.03	0.09	0.05	0.01	0.23

Tableau B.1: Statistiques des débits minimums ...suite

No de la station	Nombre d'années d'observation	Valeur minimale	Valeur maximale	Moyenne	Écart-type	Coefficient de variation
051004	26	0.00	0.01	0.00	0.00	0.24
051005	15	0.48	1.98	1.00	0.38	0.38
051006	15	0.11	0.36	0.23	0.08	0.33
051007	22	0.01	0.03	0.02	0.01	0.31
051301	27	2.48	5.15	3.57	0.61	0.17
051502	27	5.13	9.32	7.47	1.09	0.15
052219	26	1.75	7.87	4.01	1.61	0.40
052233	16	0.33	1.55	0.93	0.41	0.44
052601	62	0.51	4.90	2.51	1.04	0.41
052805	30	0.49	6.46	2.39	1.17	0.49
060101	20	0.96	2.41	1.79	0.40	0.22
061020	24	3.91	7.31	5.49	1.06	0.19
061022	22	1.10	3.57	2.35	0.65	0.28
061502	18	5.09	13.90	9.20	2.36	0.26
061801	20	2.27	4.23	3.38	0.52	0.15
061901	33	47.30	92.30	69.41	9.68	0.14
061905	33	37.40	62.80	51.62	7.31	0.14
061906	30	16.80	26.60	21.58	2.50	0.12
062101	34	22.70	47.20	35.85	5.66	0.16
062102	33	17.70	49.00	34.92	5.98	0.17
062209	16	12.10	21.90	16.79	2.93	0.17
062701	21	2.17	4.92	3.54	0.76	0.21
062802	18	2.65	6.58	4.10	1.02	0.25
070201	17	1.57	5.35	2.47	0.86	0.35
070401	18	6.51	16.50	11.23	2.91	0.26
071401	20	1.69	10.30	6.53	1.97	0.30
072201	28	0.52	2.37	1.35	0.50	0.37
072301	28	49.80	125.00	78.80	21.60	0.27
073501	17	17.00	66.50	32.22	11.79	0.37
073801	37	35.80	84.80	59.67	12.26	0.21
074601	25	5.50	18.50	10.89	3.34	0.31
074701	19	16.60	46.00	27.09	6.98	0.26
075601	18	11.60	39.00	23.28	7.64	0.33
076601	24	11.10	45.60	20.92	8.55	0.41
080101	62	7.70	23.00	14.30	3.24	0.23
080104	18	10.50	41.10	23.94	11.13	0.46

Tableau B.1: Statistiques des débits minimums ...suite

No de la station	Nombre d'années d'observation	Valeur minimale	Valeur maximale	Moyenne	Écart-type	Coefficient de variation
080701	21	175.00	401.00	266.48	50.84	0.19
080703	36	4.39	15.40	10.12	2.18	0.22
080704	19	72.20	125.00	96.79	18.04	0.19
080705	26	31.10	58.90	44.66	7.64	0.17
080718	27	108.00	221.00	154.93	25.84	0.17
080801	20	68.20	144.00	94.07	18.29	0.19
080809	20	39.00	66.20	51.70	8.01	0.16
081002	31	266.00	502.00	379.71	56.84	0.15
081007	23	138.00	260.00	206.96	33.60	0.16
081101	20	7.09	20.00	12.99	3.83	0.29

ANNEXE C: SERIES DE DONNEES NON STATIONNAIRES

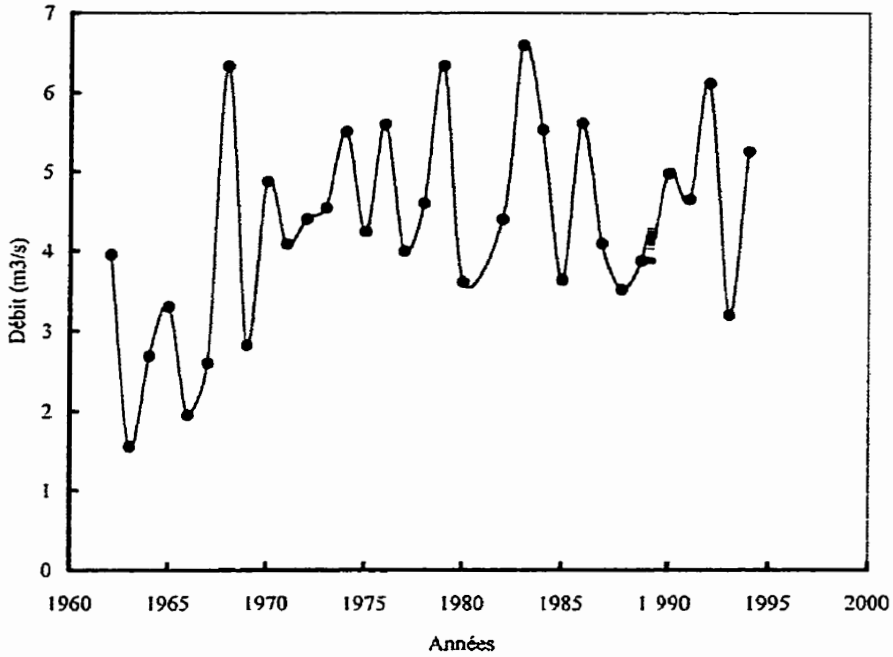

Figure C.1 Stationnarité du débit minimum 7 jours pour la station 010901

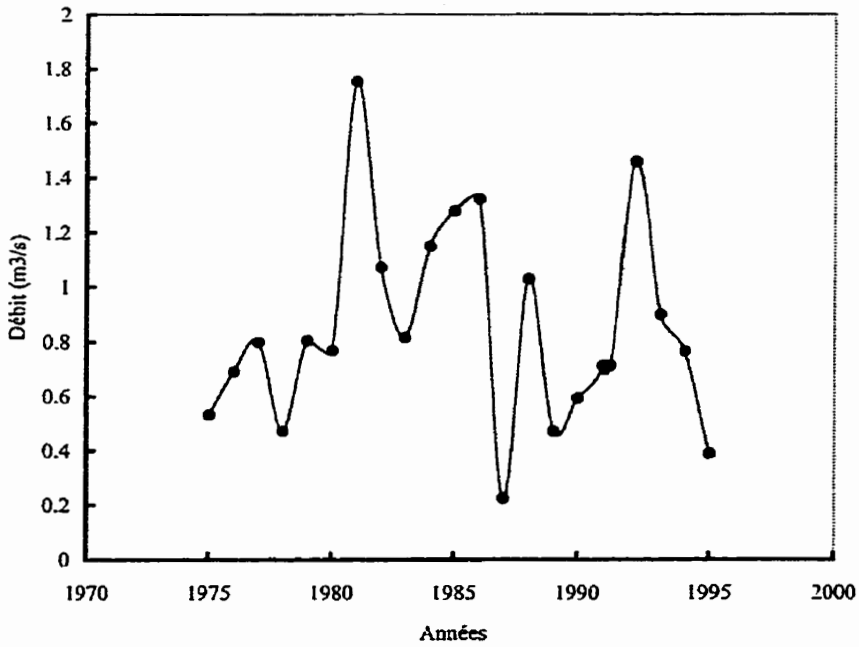

Figure C.2 Stationnarité du débit minimum 7 jours pour la station 021915

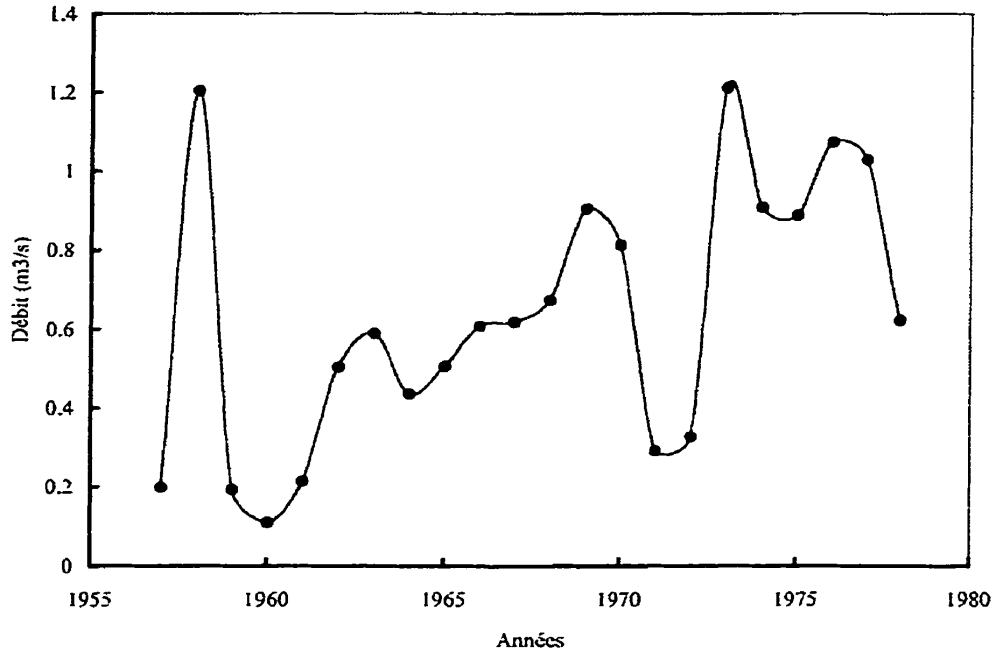

Figure C.3 Stationnarité du débit minimum 7 jours pour la station 022504

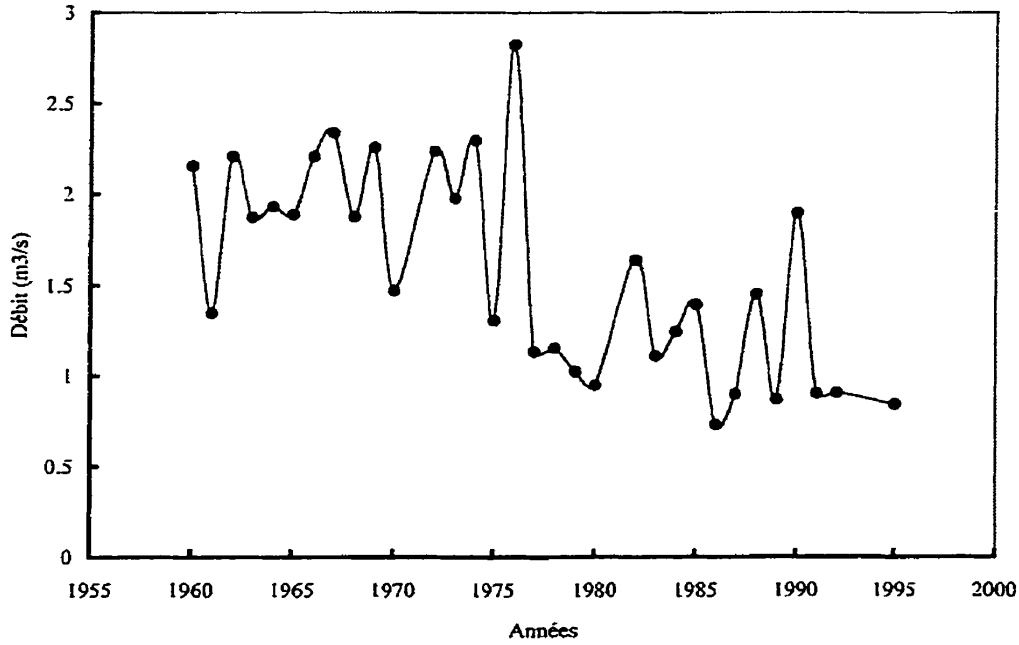

Figure C.4 Stationnarité du débit minimum 7 jours pour la station 030215

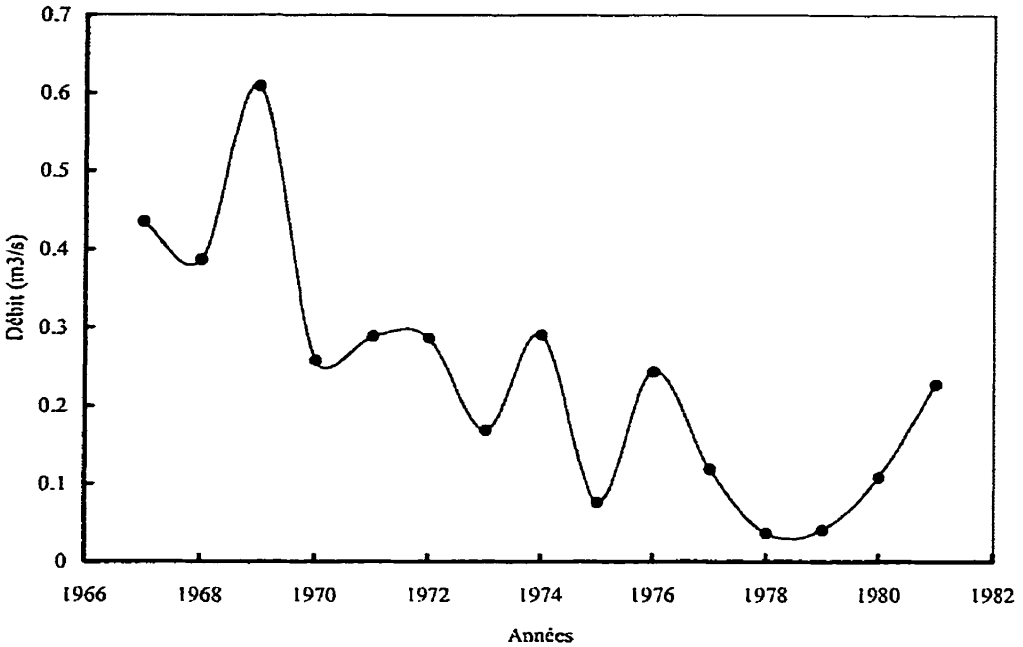

Figure C.5 Stationnarité du débit minimum 7 jours pour la station 030242

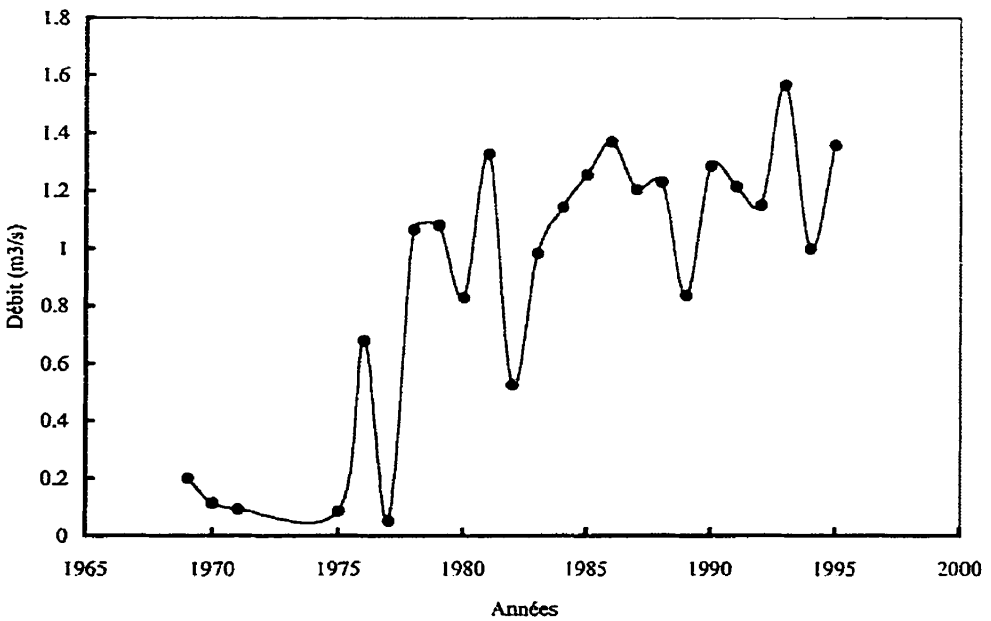

Figure C.6 Stationnarité du débit minimum 7 jours pour la station 030309

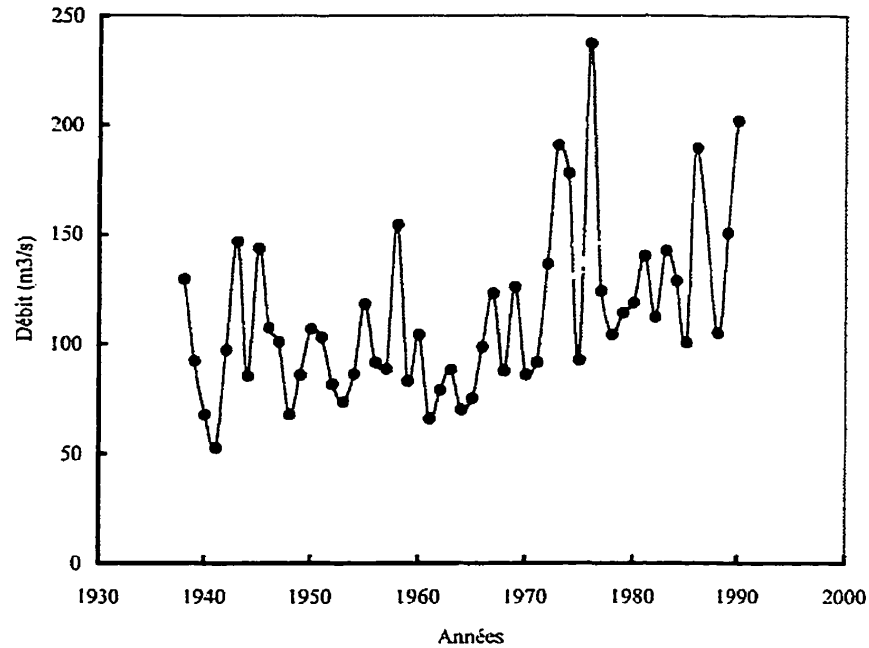

Figure C.7 Stationnarité du débit minimum 7 jours pour la station 030401

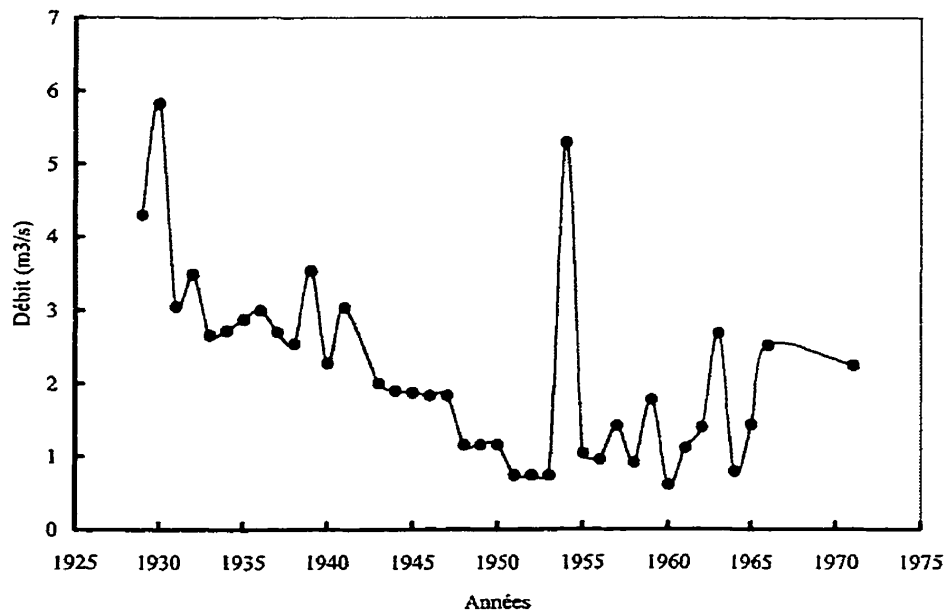

Figure C.8 Stationnarité du débit minimum 7 jours pour la station 050118

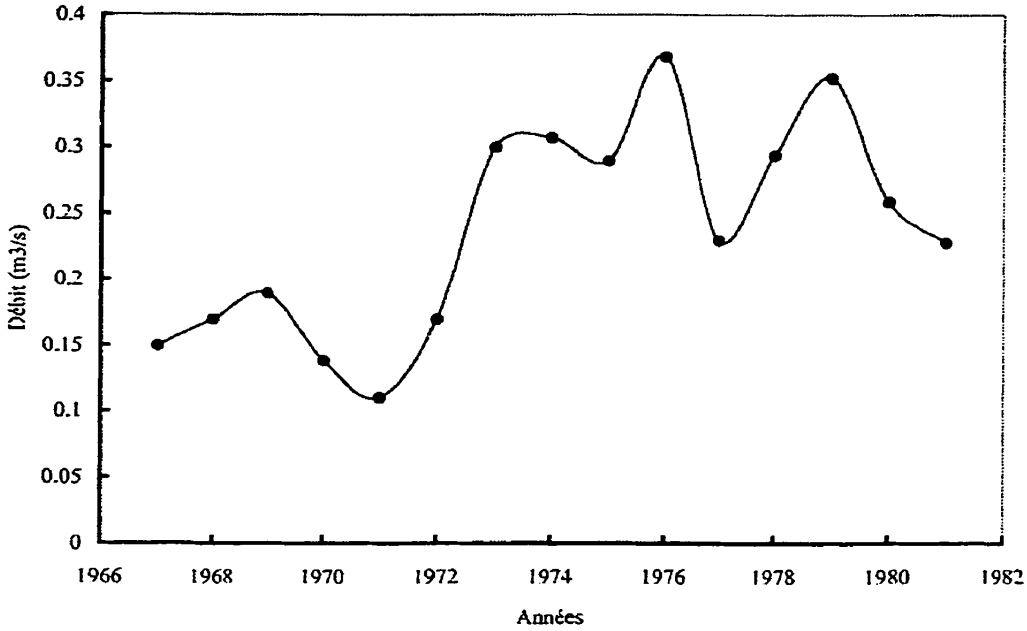

Figure C.9 Stationnarité du débit minimum 7 jours pour la station 051006

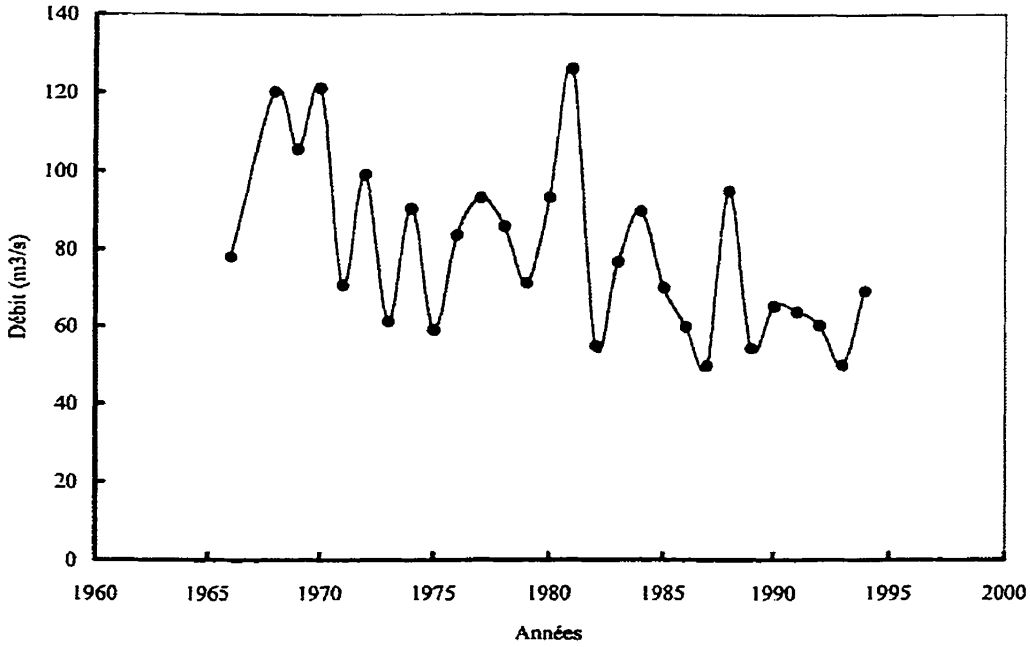

Figure C.10 Stationnarité du débit minimum 7 jours pour la station 072301

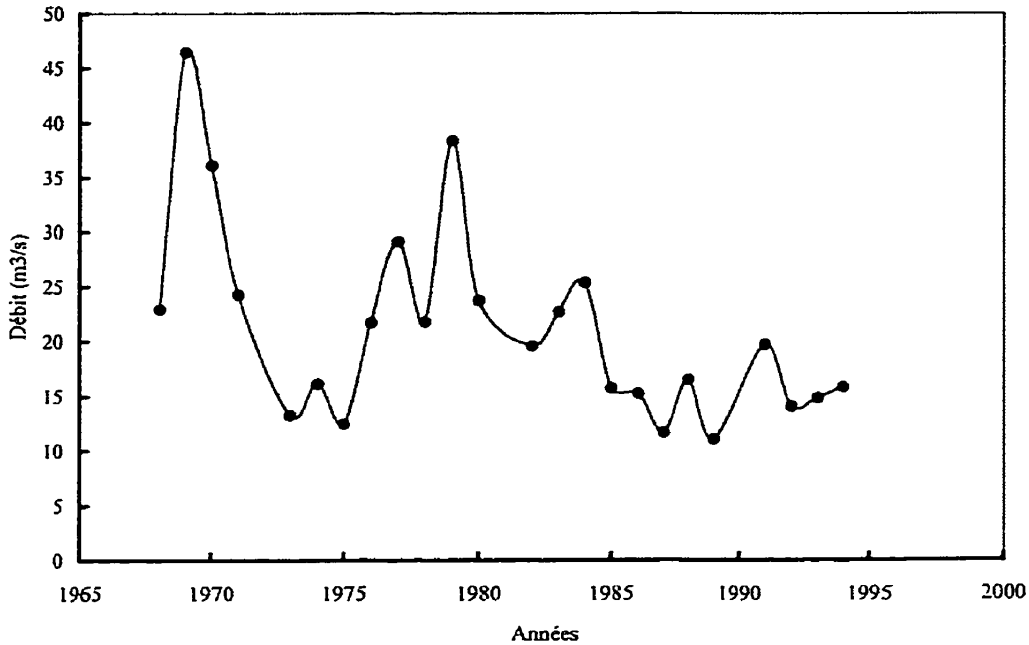

Figure C.11 Stationnarité du débit minimum 7 jours pour la station 076601

ANNEXE D: VARIABLES 1Q10, 7Q2, 7Q10 ET 30Q5

Tableau D.1: Variables 1Q10, 7Q2, 7Q10 et 30Q5

No de la station	**1Q10** (m^3/s)	**7Q10** (m^3/s)	**7Q2** (m^3/s)	**30Q5** (m^3/s)
10101	0.700	0.842	1.137	1.093
10801	3.956	4.003	5.744	5.103
11003	4.566	4.590	6.809	6.397
11201	1.542	1.788	3.003	2.637
11507	5.493	5.618	9.293	7.608
20401	1.160	1.170	2.287	1.650
20404	0.999	1.008	1.711	1.191
20502	0.094	0.100	0.136	0.138
20602	0.874	0.877	1.319	1.148
20802	2.263	2.412	3.343	2.862
21407	2.049	2.089	3.406	2.743
21502	1.769	1.775	2.626	2.096
21601	3.400	3.539	5.016	4.339
21702	0.062	0.082	0.307	0.240
22003	1.808	2.045	3.989	3.440
22301	0.673	0.787	1.253	1.240
22507	0.472	0.514	0.811	0.839
22703	0.392	0.428	0.865	0.785
23106	0.900	1.580	2.204	2.348
23301	1.145	1.449	2.734	2.998
23401	0.283	0.384	0.818	0.788
23402	5.219	6.472	11.300	10.731
23422	0.225	0.253	0.585	0.553
24001	0.850	1.443	2.672	2.711
24003	0.814	0.958	1.876	1.913
24007	2.878	3.253	5.191	5.003
30101	0.126	0.206	0.537	0.484
30103	1.270	1.521	2.611	2.260
30234	0.553	0.656	1.100	1.052
30282	0.580	0.745	1.210	1.368
30304	0.466	0.624	1.319	1.385
30314	0.025	0.033	0.197	0.172
30316	0.051	0.089	0.150	0.137
30339	0.033	0.049	0.304	0.214
30415	0.039	0.061	0.132	0.126
30420	0.048	0.097	0.239	0.315
30901	1.718	2.165	3.510	3.263
30905	1.287	2.520	4.030	3.848

Tableau D.1: Variables 1Q10, 7Q2, 7Q10 et 30Q5...suite

No de la station	**1Q10** (m^3/s)	**7Q10** (m^3/s)	**7Q2** (m^3/s)	**30Q5** (m^3/s)
31401	0.170	0.199	0.333	0.321
40129	0.228	0.253	0.447	0.379
40201	6.926	7.743	12.857	10.721
40212	0.051	0.054	0.113	0.088
40401	2.350	2.474	4.707	3.261
40810	3.900	3.959	6.781	5.276
40814	2.816	3.072	4.668	3.675
41301	10.667	11.291	17.425	15.620
41902	9.346	9.657	16.520	13.455
41903	7.227	7.446	11.251	9.536
42103	0.527	0.651	1.211	0.969
43012	8.037	8.140	9.940	8.587
50116	5.132	6.029	9.740	8.861
50135	4.478	4.504	6.564	5.722
50304	17.069	17.173	22.231	19.510
50305	10.775	10.918	14.760	12.240
50402	1.050	1.306	2.934	3.221
50408	7.650	7.863	9.202	8.826
50409	3.162	3.201	3.532	3.516
50701	0.936	1.212	1.948	2.021
50904	0.059	0.109	0.267	0.239
51002	0.024	0.025	0.031	0.029
51003	0.038	0.042	0.057	0.051
51004	0.003	0.003	0.005	0.004
51005	0.510	0.512	1.004	0.676
51007	0.011	0.012	0.018	0.016
51301	2.847	3.115	3.710	3.534
51502	5.980	6.337	7.889	7.433
52219	2.094	2.376	4.370	3.685
52233	0.335	0.388	0.930	0.745
52805	1.107	1.193	2.350	1.900
60101	1.193	1.358	1.899	1.666
61020	4.008	4.030	5.761	4.446
61022	1.275	1.289	2.436	1.933
61502	6.331	6.333	9.610	7.150
61801	2.540	2.548	3.576	3.161
61901	55.541	56.791	71.000	65.738
61905	39.477	40.047	51.830	47.829
61906	18.365	18.670	21.227	20.619

Tableau D.1: Variables 1Q10, 7Q2, 7Q10 et 30Q5...suite

No de la station	**1Q10** (m^3/s)	**7Q10** (m^3/s)	**7Q2** (m^3/s)	**30Q5** (m^3/s)
62101	28.943	29.221	36.168	32.570
62102	28.012	28.102	35.430	31.339
62209	12.759	12.784	16.616	14.497
62701	2.431	2.495	3.730	3.095
62802	2.707	3.064	4.304	3.782
70201	1.588	1.861	2.594	2.204
70401	6.974	7.114	11.251	9.743
71401	4.264	4.948	6.517	6.583
72201	0.662	0.719	1.392	1.035
73501	19.222	19.657	29.200	25.073
73801	45.000	45.250	60.000	50.885
74601	7.865	8.003	9.910	9.210
74701	21.100	21.240	26.600	25.360
75601	15.484	16.053	23.062	18.033
80101	9.137	10.328	15.922	13.550
80104	10.690	10.730	22.737	13.552
80701	220.000	220.627	258.600	238.545
80703	7.145	7.715	10.535	9.009
80704	74.800	75.270	96.700	79.190
80705	35.944	36.137	44.576	39.023
80718	117.000	117.271	153.000	140.843
80801	73.243	73.633	93.190	81.314
80809	40.500	40.869	51.467	46.155
81002	309.500	310.350	372.400	339.525
81007	154.000	154.150	212.400	178.750
81101	7.857	7.880	13.455	9.885

ANNEXE E: DONNEES PHYSIQUES BRUTES ET STANDARDISEES

Tableau E.1 Variables physiques brutes

station	S (km^2)	L (km)	Pente (m/km)	Pluie (mm)	Forêt (%)	Eau (%)	Lat.	Long.
10101	534	13.8	4.8	950	99.2	0.5	52.59	24.13
10801	1930	109.1	4.6	1100	99.8	0.2	46.89	28.99
11003	1681	66.8	9.6	1050	99.4	0.5	54.92	32.73
11201	1138	68.6	7.2	1050	98.2	0.2	47.67	33.59
11507	2734	104.9	1.6	1000	83.6	2.6	46.63	37.59
20401	1015	106.1	5.0	1000	98.3	0.4	58.29	24.65
20404	662	80.2	5.2	1050	99.4	0.3	57.77	26.22
20502	60	5.4	8.3	975	99.0	1.0	60.36	23.70
20602	630	61.1	6.4	950	99.8	0.2	60.36	25.00
20802	1217	97.4	3.4	1000	99.0	1.0	63.99	28.13
21407	749	64.0	7.3	1000	99.0	1.0	61.66	34.38
21502	722	56.6	5.4	1000	97.8	0.4	61.66	35.33
21601	1647	77.2	4.6	1000	94.0	1.0	57.25	39.84
21702	208	33.6	4.9	1000	83.0	3.0	57.25	40.45
22003	1586	111.8	3.7	1100	92.0	3.0	51.55	45.14
22301	966	58.7	23.6	1000	74.0	1.0	46.63	48.44
22507	516	51.9	4.1	1000	93.0	0.4	39.12	50.78
22703	783	46.7	4.4	1000	88.0	1.0	35.23	52.34
23106	824	53.5	6.4	1050	74.0	1.0	26.94	56.60
23301	1130	85.3	4.1	1000	63.0	1.0	24.09	58.07
23401	707	56.8	4.9	1100	44.0	4.0	24.35	59.38
23402	5820	196.3	1.4	1100	69.0	2.0	23.32	58.94
23422	691	43.2	9.0	1000	81.0	3.0	7.51	57.29
24001	1412	96.3	2.2	1100	46.0	2.0	17.10	60.68
24003	924	76.4	2.4	1100	53.9	2.8	18.91	60.24
24007	2317	162.7	1.3	1050	47.0	4.0	17.10	64.58
30101	544	45.3	2.5	1050	49.0	1.0	10.88	62.93
30103	1540	93.3	1.9	1000	40.0	1.0	15.03	64.67
30234	642	48.9	5.6	1100	84.0	4.0	5.96	61.28
30282	738	53.4	4.1	1200	89.4	0.3	7.51	59.90
30304	1490	68.0	2.8	1000	42.0	1.0	6.22	67.80
30314	209	25.2	1.5	1000	66.5	0.6	1.81	66.93
30316	342	26.0	1.8	1000	29.0	0.2	13.47	67.53
30339	595	33.6	1.8	1000	47.0	4.0	8.29	66.23
30415	309	17.1	1.8	1000	42.0	2.0	6.22	69.27
30420	404	25.6	1.2	1000	51.0	5.0	0.52	68.23

Tableau E.1: Variables physiques brutes ...suite

station	S (km^2)	L (km)	Pente (m/km)	Pluie (mm)	Forêt (%)	Eau (%)	Lat.	Long.
30901	2463	114.0	4.3	1000	31.0	3.0	0.00	73.61
30905	2463	114.0	4.3	1000	31.0	3.0	3.63	72.22
31401	218	29.5	6.9	1100	84.0	1.0	1.30	60.50
40129	110	19.9	2.4	1150	97.6	1.8	15.54	74.13
40201	2574	127.3	2.3	1000	92.0	5.0	19.69	77.52
40212	40	2.6	1.4	1000	89.0	4.0	19.95	76.13
40401	1404	95.4	1.2	1000	81.0	15.0	19.43	78.91
40810	1691	122.7	1.5	900	88.5	9.6	21.50	84.11
40814	1272	95.8	1.3	1000	88.9	7.1	15.28	84.29
41301	5154	193.9	1.2	1100	90.0	10.0	12.18	87.50
41902	3729	165.3	1.4	1000	88.0	12.0	19.43	93.32
41903	2111	96.7	0.5	1100	87.0	13.0	26.94	93.66
42103	623	57.3	2.8	1050	76.0	24.0	19.17	96.09
43012	2575	123.8	0.3	900	88.1	8.2	50.78	98.78
50116	2614	151.3	1.5	1000	90.9	9.1	39.90	68.06
50135	1579	110.3	3.1	1000	96.0	3.0	41.71	66.93
50304	4400	172.0	1.9	1000	90.0	5.0	23.32	65.19
50305	2863	102.0	2.1	1050	94.0	6.0	30.31	64.50
50402	737	75.2	8.4	1200	95.7	1.8	27.98	62.24
50408	1438	84.2	7.9	1200	93.4	2.7	27.46	62.41
50409	41	74.0	6.8	1200	96.0	4.0	29.27	62.33
50701	355	36.1	10.3	1150	58.0	5.0	25.13	62.41
50904	357	17.8	23.0	1150	76.0	2.0	26.68	59.55
51002	4	0.2	20.0	1200	95.0	2.0	33.94	58.68
51003	9	0.6	23.0	1200	95.0	2.0	33.94	58.59
51004	1	0.1	16.0	1200	95.0	2.0	34.20	58.68
51005	282	43.8	7.1	1450	97.0	3.0	33.68	58.59
51007	4	0.3	13.0	1200	97.0	2.0	33.94	58.68
51301	862	66.0	10.1	1000	91.0	1.0	36.53	55.30
51502	1631	156.1	6.0	1000	96.0	3.0	40.41	53.82
52219	1340	109.4	3.3	1000	80.0	4.0	14.25	70.49
52233	647	53.9	14.0	1000	85.0	4.0	11.66	70.83
52805	774	59.2	12.0	1000	92.0	8.0	23.58	69.27
60101	754	70.9	6.7	1000	95.7	3.9	48.19	52.95
61020	1119	100.6	5.0	1200	96.8	3.2	47.93	61.20
61022	487	77.3	4.2	1250	97.0	3.0	44.30	59.81
61502	2190	149.5	2.5	950	96.0	3.5	51.04	63.02

Tableau E.1: Variables physiques brutes...suite

station	S (km^2)	L (km)	Pente (m/km)	Pluie (mm)	Forêt (%)	Eau (%)	Lat.	Long.
61801	1001	105.9	2.0	900	89.0	6.7	57.77	63.28
61901	15333	321.9	0.9	900	92.0	8.0	55.96	65.63
61905	11163	220.5	0.5	900	91.0	9.0	65.03	70.14
61906	4429	141.6	1.4	900	87.0	13.0	64.77	71.96
62101	9350	303.6	1.3	900	94.0	5.0	59.07	64.15
62102	9712	265.5	1.6	900	95.0	4.0	59.07	64.41
62209	3600	122.0	3.5	1000	92.0	7.0	74.35	57.47
62701	746	25.0	9.0	1000	91.0	7.0	52.85	57.73
62802	1101	100.9	8.5	1000	95.1	4.5	49.48	52.17
70201	791	90.9	6.4	1000	95.0	5.0	53.63	50.95
70401	2586	195.1	2.8	1050	94.1	5.8	55.18	48.35
71401	1558	109.4	3.1	1050	93.3	6.7	65.80	40.45
72201	554	48.1	13.6	1100	91.0	8.0	80.83	34.20
73501	7610	233.0	2.5	800	89.0	10.0	82.12	23.61
73801	13028	370.1	1.3	1100	88.0	12.0	81.09	19.44
74601	2080	161.3	3.4	1100	86.0	10.0	80.31	12.07
74701	5750	272.6	2.0	1100	87.0	11.0	80.05	11.63
75601	2950	242.0	1.7	900	90.0	9.0	84.20	0.00
80101	3704	144.8	0.3	900	72.0	23.0	54.66	94.88
80104	9691	206.2	0.3	800	55.0	43.0	75.91	100.00
80701	57498	482.8	0.2	800	81.0	17.0	78.50	91.32
80703	2000	118.9	0.4	950	81.4	18.1	63.47	90.97
80704	18803	324.2	0.3	850	79.0	21.0	71.50	83.85
80705	7977	250.7	0.4	950	83.0	17.0	50.52	88.28
80718	31611	441.1	0.3	850	80.0	19.0	74.09	90.10
80801	17366	417.0	0.5	750	68.0	32.0	94.82	91.32
80809	10003	247.8	0.6	800	73.0	20.0	88.08	85.94
81002	40964	664.4	0.3	750	78.0	20.0	98.70	88.45
81007	18120	287.2	0.8	800	78.0	22.0	92.49	72.48
81101	6051	217.9	0.6	700	75.3	24.7	100.00	94.88
moyenne	3956	122.5	4.7	1014	82.2	6.6	42.17	60.35
ecart-type	8127	112.5	4.9	118	17.9	7.6	25.60	21.22

Tableau E.2 Variables physiques standardisées

station	S	L	Pente	Pluie	Forêt	Eau	Lat.	Long.
10101	-0.424	-0.969	0.023	-0.549	0.949	-0.809	0.378	-1.700
10801	-0.251	-0.118	-0.020	0.734	0.983	-0.849	0.160	-1.471
11003	-0.282	-0.496	0.976	0.306	0.960	-0.809	0.467	-1.295
11201	-0.349	-0.480	0.506	0.306	0.893	-0.849	0.189	-1.254
11507	-0.152	-0.155	-0.632	-0.121	0.076	-0.531	0.150	-1.065
20401	-0.364	-0.145	0.057	-0.121	0.899	-0.823	0.596	-1.675
20404	-0.408	-0.376	0.101	0.306	0.960	-0.836	0.577	-1.602
20502	-0.482	-1.044	0.724	-0.335	0.938	-0.743	0.676	-1.720
20602	-0.412	-0.547	0.348	-0.549	0.983	-0.849	0.676	-1.659
20802	-0.339	-0.222	-0.262	-0.121	0.938	-0.743	0.815	-1.512
21407	-0.397	-0.521	0.522	-0.121	0.938	-0.743	0.725	-1.217
21502	-0.401	-0.587	0.138	-0.121	0.871	-0.823	0.725	-1.172
21601	-0.286	-0.403	-0.020	-0.121	0.658	-0.743	0.557	-0.959
21702	-0.464	-0.792	0.045	-0.121	0.042	-0.478	0.557	-0.930
22003	-0.294	-0.094	-0.212	0.734	0.546	-0.478	0.338	-0.709
22301	-0.370	-0.568	3.815	-0.121	-0.462	-0.743	0.150	-0.553
22507	-0.426	-0.629	-0.119	-0.121	0.602	-0.823	-0.138	-0.443
22703	-0.393	-0.675	-0.070	-0.121	0.322	-0.743	-0.287	-0.369
23106	-0.388	-0.614	0.336	0.306	-0.462	-0.743	-0.605	-0.168
23301	-0.350	-0.331	-0.135	-0.121	-1.078	-0.743	-0.714	-0.099
23401	-0.402	-0.585	0.027	0.734	-2.142	-0.346	-0.704	-0.037
23402	0.230	0.661	-0.662	0.734	-0.742	-0.611	-0.744	-0.058
23422	-0.404	-0.706	0.865	-0.121	-0.070	-0.478	-1.349	-0.136
24001	-0.315	-0.232	-0.505	0.734	-2.030	-0.611	-0.982	0.024
24003	-0.376	-0.410	-0.472	0.734	-1.587	-0.505	-0.913	0.004
24007	-0.203	0.361	-0.691	0.306	-1.974	-0.346	-0.982	0.208
30101	-0.423	-0.688	-0.446	0.306	-1.862	-0.743	-1.220	0.131
30103	-0.299	-0.259	-0.569	-0.121	-2.366	-0.743	-1.061	0.212
30234	-0.411	-0.656	0.178	0.734	0.098	-0.346	-1.409	0.053
30282	-0.399	-0.615	-0.131	1.588	0.400	-0.836	-1.349	-0.013
30304	-0.306	-0.485	-0.398	-0.121	-2.254	-0.743	-1.399	0.360
30314	-0.464	-0.867	-0.658	-0.121	-0.882	-0.796	-1.568	0.319
30316	-0.448	-0.861	-0.600	-0.121	-2.981	-0.849	-1.121	0.348

Tableau E.2: Variables physiques standardisées...suite

station	S	L	Pente	Pluie	Forêt	Eau	Lat.	Long.
30339	-0.416	-0.792	-0.590	-0.121	-1.974	-0.346	-1.320	0.286
30415	-0.452	-0.940	-0.590	-0.121	-2.254	-0.611	-1.399	0.429
30420	-0.440	-0.864	-0.711	-0.121	-1.750	-0.214	-1.617	0.380
30901	-0.185	-0.074	-0.084	-0.121	-2.869	-0.478	-1.637	0.634
30905	-0.185	-0.074	-0.084	-0.121	-2.869	-0.478	-1.498	0.569
31401	-0.463	-0.829	0.441	0.734	0.098	-0.743	-1.588	0.016
40129	-0.476	-0.915	-0.476	1.161	0.859	-0.637	-1.042	0.659
40201	-0.172	0.045	-0.485	-0.121	0.546	-0.214	-0.883	0.818
40212	-0.485	-1.069	-0.679	-0.121	0.378	-0.346	-0.873	0.753
40401	-0.316	-0.240	-0.703	-0.121	-0.070	1.110	-0.893	0.884
40810	-0.281	0.004	-0.648	-0.976	0.350	0.395	-0.813	1.130
40814	-0.333	-0.237	-0.689	-0.121	0.372	0.064	-1.052	1.138
41301	0.147	0.639	-0.711	0.734	0.434	0.448	-1.171	1.289
41902	-0.029	0.384	-0.664	-0.121	0.322	0.713	-0.893	1.564
41903	-0.229	-0.229	-0.860	0.734	0.266	0.845	-0.605	1.580
42103	-0.413	-0.581	-0.382	0.306	-0.350	2.301	-0.903	1.695
43012	-0.171	0.013	-0.895	-0.976	0.327	0.210	0.309	1.822
50116	-0.167	0.259	-0.658	-0.121	0.484	0.329	-0.108	0.372
50135	-0.295	-0.107	-0.335	-0.121	0.770	-0.478	-0.039	0.319
50304	0.054	0.444	-0.569	-0.121	0.434	-0.214	-0.744	0.237
50305	-0.136	-0.181	-0.529	0.306	0.658	-0.081	-0.476	0.204
50402	-0.399	-0.421	0.750	1.588	0.753	-0.637	-0.565	0.098
50408	-0.312	-0.340	0.633	1.588	0.624	-0.518	-0.585	0.106
50409	-0.485	-0.431	0.429	1.588	0.770	-0.346	-0.515	0.102
50701	-0.446	-0.770	1.128	1.161	-1.358	-0.214	-0.674	0.106
50904	-0.446	-0.933	3.694	1.161	-0.350	-0.611	-0.615	-0.029
51002	-0.489	-1.090	3.088	1.588	0.714	-0.611	-0.337	-0.070
51003	-0.489	-1.087	3.694	1.588	0.714	-0.611	-0.337	-0.074
51004	-0.490	-1.091	2.280	1.588	0.714	-0.611	-0.327	-0.070
51005	-0.455	-0.701	0.481	3.725	0.826	-0.478	-0.347	-0.074
51007	-0.489	-1.090	1.673	1.588	0.826	-0.611	-0.337	-0.070
51301	-0.383	-0.503	1.092	-0.121	0.490	-0.743	-0.237	-0.230
51502	-0.288	0.302	0.249	-0.121	0.770	-0.478	-0.089	-0.299
52219	-0.324	-0.115	-0.280	-0.121	-0.126	-0.346	-1.091	0.487
52233	-0.410	-0.611	1.876	-0.121	0.154	-0.346	-1.191	0.503
52805	-0.394	-0.564	1.471	-0.121	0.546	0.183	-0.734	0.429

Tableau E.2: Variables physiques standardisées...suite

station	S	L	Pente	Pluie	Forêt	Eau	Lat.	Long.
60101	-0.397	-0.459	0.396	-0.121	0.753	-0.359	0.209	-0.340
61020	-0.352	-0.194	0.053	1.588	0.815	-0.452	0.199	0.049
61022	-0.430	-0.402	-0.101	2.016	0.826	-0.478	0.060	-0.017
61502	-0.219	0.243	-0.458	-0.549	0.770	-0.412	0.318	0.135
61801	-0.366	-0.146	-0.547	-0.976	0.378	0.011	0.577	0.147
61901	1.406	1.783	-0.771	-0.976	0.546	0.183	0.507	0.257
61905	0.890	0.877	-0.846	-0.976	0.490	0.316	0.855	0.470
61906	0.058	0.172	-0.664	-0.976	0.266	0.845	0.845	0.556
62101	0.666	1.619	-0.689	-0.976	0.658	-0.214	0.626	0.188
62102	0.711	1.279	-0.630	-0.976	0.714	-0.346	0.626	0.200
62209	-0.045	-0.003	-0.246	-0.121	0.546	0.051	1.212	-0.127
62701	-0.398	-0.869	0.865	-0.121	0.490	0.051	0.388	-0.115
62802	-0.354	-0.191	0.768	-0.121	0.719	-0.280	0.259	-0.377
70201	-0.392	-0.280	0.342	-0.121	0.714	-0.214	0.418	-0.435
70401	-0.170	0.650	-0.394	0.306	0.663	-0.108	0.477	-0.557
71401	-0.297	-0.115	-0.319	0.306	0.619	0.011	0.884	-0.930
72201	-0.421	-0.663	1.789	0.734	0.490	0.183	1.460	-1.225
73501	0.451	0.989	-0.456	-1.831	0.378	0.448	1.510	-1.724
73801	1.121	2.213	-0.691	0.734	0.322	0.713	1.470	-1.921
74601	-0.233	0.348	-0.276	0.734	0.210	0.448	1.440	-2.269
74701	0.221	1.342	-0.541	0.734	0.266	0.580	1.430	-2.290
75601	-0.125	1.069	-0.610	-0.976	0.434	0.316	1.589	-2.838
80101	-0.032	0.201	-0.895	-0.976	-0.574	2.168	0.457	1.637
80104	0.708	0.749	-0.895	-1.831	-1.526	4.815	1.271	1.879
80701	6.620	3.220	-0.913	-1.831	-0.070	1.374	1.371	1.470
80703	-0.243	-0.030	-0.870	-0.549	-0.048	1.520	0.795	1.453
80704	1.835	1.803	-0.903	-1.403	-0.182	1.904	1.103	1.117
80705	0.496	1.147	-0.873	-0.549	0.042	1.374	0.299	1.326
80718	3.419	2.847	-0.885	-1.403	-0.126	1.639	1.202	1.412
80801	1.657	2.632	-0.860	-2.258	-0.798	3.359	1.996	1.470
80809	0.747	1.121	-0.840	-1.831	-0.518	1.771	1.738	1.216
81002	4.575	4.842	-0.887	-2.258	-0.238	1.771	2.145	1.334
81007	1.751	1.473	-0.802	-1.831	-0.238	2.036	1.907	0.581
81101	0.258	0.854	-0.832	-2.686	-0.389	2.393	2.195	1.637
71102	0.076	-0.226	1.067	-0.121	0.210	0.845	2.195	-0.904