

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE  
UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À  
L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

COMME EXIGENCE PARTIELLE  
À L'OBTENTION DE LA  
MAÎTRISE EN GÉNIE DE LA CONSTRUCTION  
M. Ing.

PAR  
CLAUDE LEMAIRE

IMPACT DES ERREURS SUR L'ESTIMATION  
DES QUANTILES DE CRUE

MONTRÉAL, LE 22 OCTOBRE 2004

@ droits réservés de Claude Lemaire

CE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

M. Saad Bennis, directeur de mémoire  
Département de génie de la construction à l'École de technologie supérieure

M. Daniel Perraton, président du jury  
Département de génie de la construction à l'École de technologie supérieure

M. Marc Barbet, membre externe  
Hydro-Québec

IL A FAIT L'OBJET D'UNE PRÉSENTATION DEVANT JURY ET PUBLIC

LE 22 OCTOBRE 2004

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

# **IMPACT DES ERREURS SUR L'ESTIMATION DES QUANTILES DE CRUE**

Claude Lemaire

## **SOMMAIRE**

L'objectif général de notre travail était de déterminer l'impact des erreurs dans les données hydrométriques mesurées sur les sites. Ces données comportent toujours un degré d'erreurs provenant de différentes sources et pouvant survenir à différentes étapes : lors de la prise des mesures, de la transmission des mesures, du stockage ou encore durant le traitement des données.

L'impact des erreurs a été déterminé par la comparaison de quantiles calculés à l'aide d'une méthode statistique soit la loi log normal. Les données utilisées pour notre étude sont les débits de pointe de la crue printanière. Ces débits sont souvent les plus importants de l'année.

Notre recherche a montré, entre autres, que les erreurs sur les débits mesurés ont un impact pouvant être important sur l'estimation des quantiles de crue. Ces erreurs sur les débits de pointe auront des conséquences sociales, économiques et environnementales lors de la gestion ou de la conception des ouvrages d'art. Ainsi, on retrouve des impacts sur l'environnement, sur le contrôle des inondations, sur les coûts de construction et d'opération des ouvrages.

Par conséquent, nous recommandons, entre autres, l'amélioration des instruments, des techniques de mesure et du traitement des données, ainsi que la mise sur pied d'une banque de données centralisée regroupant les données, leur historique et leurs paramètres de traitement et en permettant l'accessibilité.

# **IMPACT OF THE ERRORS ON THE ESTIMATE OF THE FRACTILES OF FLOOD**

Claude Lemaire

## **ABSTRACT**

The general object of this study was to determine the impact of errors in hydrometric data measured on site. This data is never perfect and contains errors, which can occur from various sources and at various steps of the process. These errors may happen during measurement, transmission of the data, storage and even during the analysis of the data.

The estimate of the fractiles of flood was calculated using statistical methods related to the discharges of river for the design of spillways. The data of our study are the peak output of the spring rising. These flows are often most significant.

The results show that simulation of errors carried out on the peak output of the spring flood had a significant impact for spillways. The errors on fractiles of flood rates affecting the environment, the control of floods, the construction costs and operating cost.

We therefore recommend, amongst other things, to improve the measuring instruments and techniques, and to better overview the treatment of the data. We also recommend the creation of a centralized database for all flood data, including their history and treatment parameters. This database should be accessible to all who desire it.

## **REMERCIEMENTS**

Un grand merci et toute ma reconnaissance à monsieur Saad Bennis, ing., Ph.D. qui à titre de directeur de recherche, m'a accueilli au sein de son équipe de recherche et m'a apporté le support nécessaire et accordé de son temps et qui, de par son expérience a su me prodiguer de précieux conseils.

Un merci tout spécial à monsieur Marc Barbet, ing., M.Sc.A pour Hydro-Québec, pour sa collaboration, ses conseils éclairés, sa grande expérience et les documents de travail qu'il m'a fournis.

Je ne voudrais passer sous silence la collaboration des membres de l'équipe de recherche, messieurs Sylvain Côté et Éric Crobeddu, qui m'ont apporté un support logistique apprécié.

## TABLE DES MATIÈRES

	Page
SOMMAIRE .....	i
ABSTRACT.....	ii
REMERCIEMENTS .....	iii
TABLE DES MATIÈRES .....	iv
LISTE DES TABLEAUX.....	vii
LISTE DES FIGURES.....	viii
LISTE DES ABRÉVIATIONS ET DES SIGLES .....	ix
INTRODUCTION.....	1
CHAPITRE 1 LES DONNÉES HYDROMÉTRIQUES.....	4
1.1 Généralités.....	4
1.2 Utilité des données .....	5
1.3 Instrumentation des ouvrages hydrauliques .....	6
1.4 Types de données .....	8
1.4.1 Calcul des débits de rivières.....	8
1.4.2 Mesure des niveaux.....	10
1.4.3 Calcul des apports naturels.....	10
1.4.4 Calcul du débit turbiné.....	12
1.4.5 Calcul du débit déversé .....	12
1.5 Erreurs dans les données .....	13
1.6 Conséquences des erreurs dans les mesures.....	14
1.7 Chaînes d'acquisition des mesures.....	15
1.7.1 Acquisition des données.....	16
1.7.2 Traitement primaire des données .....	16
1.7.3 Contrôle des données .....	16
1.7.4 Organisation des données.....	17
1.7.5 Diffusion et publication.....	17
1.8 Principaux utilisateurs des données .....	18
CHAPITRE 2 LA VALIDATION .....	20
2.1 Revue de la documentation .....	20
2.2 But de la validation .....	22
2.3 Objectifs de la validation .....	22
2.4 Logiciel de validation.....	25

2.5	Traitement des données.....	28
2.5.1	Estimation des données manquantes.....	29
2.5.1.1	Techniques univariées.....	29
2.5.2	Méthodes de filtrage des données.....	30
2.5.2.1	Validation par filtrage univarié.....	31
2.5.2.2	Validation par filtrage multivarié.....	35
CHAPITRE 3 MODÈLE DE PRÉVISION.....		38
3.1	Introduction.....	38
3.2	Modèles mathématiques.....	39
3.3	Analyse des crues.....	40
3.4	Méthode stochastique.....	41
3.5	Méthode déterministe-statistique.....	41
3.6	Méthode statistique.....	42
3.7	Bassin versant.....	45
3.8	Cadre théorique.....	46
3.8.1	Méthodologie de la loi log normale.....	47
CHAPITRE 4 APPROCHE MÉTHODOLOGIQUE.....		53
4.1	Introduction.....	53
4.2	Impacts des erreurs.....	55
4.3	Étapes de calcul.....	56
4.4	Analyse pour les débits historiques.....	57
4.5	Analyse pour les débits modifiés.....	60
4.5.1	Modification des débits.....	61
4.5.2	Calcul des quantiles.....	63
4.5.3	Calcul de l'impact pour les débits modifiés.....	65
4.6	Analyse pour les débits validés.....	67
4.6.1	Validation des données.....	67
4.6.2	Application d'un filtre médian.....	68
4.6.2.1	Débits validés d'un filtre médian.....	68
4.6.2.2	Calcul des quantiles pour les débits validés avec le filtre médian.....	73
4.6.2.3	Calcul de l'impact pour les débits validés avec le filtre médian.....	75
4.6.3	Application d'un filtre winsoriseur.....	76
4.6.3.1	Débits validés d'un filtre winsoriseur.....	77
4.6.3.2	Calcul des quantiles pour les débits validés avec le filtre winsoriseur.....	79
4.6.3.3	Calcul de l'impact pour les débits validés avec le filtre winsoriseur.....	81

DISCUSSION ET INTERPRÉTATION DES RÉSULTATS .....	83
CONCLUSION .....	89
RECOMMANDATIONS.....	91
BIBLIOGRAPHIE .....	93

## LISTE DES TABLEAUX

	Page
Tableau I	Lois d'extrapolation des crues les plus couramment utilisées ..... 44
Tableau II	Récurrences des crues les plus couramment utilisées ..... 45
Tableau III	Débits de pointe des crues printanières des données historiques . 58
Tableau IV	Quantiles de crues des débits historiques..... 60
Tableau V	Débits (m <sup>3</sup> /s) historiques modifiés..... 62
Tableau VI	Statistiques pour les scénarios 1, 2 et 3..... 62
Tableau VII	Quantiles de crue (m <sup>3</sup> /s) pour les scénarios 1, 2 et 3..... 64
Tableau VIII	Pourcentages de variations des quantiles pour les 3 scénarios... 65
Tableau IX	Débits (m <sup>3</sup> /s) modifiés et validés – Filtre médian..... 68
Tableau X	Statistiques pour les scénarios 1, 2 et 3 – Filtre médian ..... 73
Tableau XI	Quantiles de crue (m <sup>3</sup> /s) pour les 3 scénarios – Filtre médian ..... 74
Tableau XII	Variations des quantiles – scénarios 1, 2 et 3 – Filtre médian ..... 75
Tableau XIII	Débits (m <sup>3</sup> /s) modifiés et validés – Filtre winsoriseur..... 77
Tableau XIV	Statistiques pour le scénario 1 – Filtre winsoriseur ..... 79
Tableau XV	Quantiles de crue (m <sup>3</sup> /s) – scénarios 1, 2 et 3 – Filtre winsoriseur..... 80
Tableau XVI	Variations des quantiles – scénarios 1, 2 et 3 – Filtre winsoriseur..... 81
Tableau XVII	Synthèse des impacts – Débits modifiés ..... 85
Tableau XVIII	Synthèse des impacts – Débits validés..... 87
Tableau XIX	Comparaison entre les 2 filtres..... 87

## LISTE DES FIGURES

	Page
Figure 1 Exemple d'une courbe de tarage .....	9
Figure 2 Méthodologie du calcul des apports naturels .....	11
Figure 3 Chaîne d'acquisition et de traitement des données .....	18
Figure 4 Erreur sur une mesure de niveau dans les données .....	24
Figure 5 Comparaison des apports naturels historiques et validés .....	25
Figure 6 Menu de pré-validation des données de PREVAL.....	27
Figure 7 Menu de validation et de calcul de PREVAL .....	28
Figure 8 Méthodes de validation .....	37
Figure 9 Les différentes approches de modélisation .....	39
Figure 10 Localisation de la station 073801 sur la rivière Romaine .....	54
Figure 11 Courbe des débits de la rivière Romaine pour l'année 1981.....	55
Figure 12 Validation avec 5 % de variation – Filtre médian – 1981 .....	69
Figure 13 Validation avec 20 % de variation – Filtre médian – 1981 .....	70
Figure 14 Validation avec 50 % de variation – Filtre médian – 1981 .....	70
Figure 15 Validation avec -5 % de variation – Filtre médian – 1981.....	71
Figure 16 Validation avec -20 % de variation – Filtre médian – 1981.....	71
Figure 17 Validation avec -50 % de variation – Filtre médian – 1981.....	72
Figure 18 Validation avec 5 % de variation – Filtre winsorizeur – 1981.....	78
Figure 19 Validation avec 20 % de variation – Filtre winsorizeur – 1981.....	78
Figure 20 Impact d'une modification sur la série de données .....	88

## LISTE DES ABRÉVIATIONS ET DES SIGLES

a	Constante propre à une section de rivière
g	Gravité 9,81 m / s <sup>2</sup>
h	Profondeur d'écoulement, m
h <sub>o</sub>	Point bas du lit de la rivière ou correction pour obtenir le niveau réel, m
i	Indice
m	Mètre
n	Constante propre à une section de rivière
s	Seconde
t	Temps, s
hm <sup>3</sup>	Hectomètre cube
h <sub>o</sub>	Point bas du lit de la rivière
ρ	Densité de l'eau
σ	Écart-type
η	Rendement du groupe turbine alternateur
p	Probabilité au non-dépassement
dV	Variation du volume, m <sup>3</sup>
dt	Variation du temps, s
AN	Apport naturel, m <sup>3</sup> / s
CERIU	Centre d'expertise et de recherche en infrastructures urbaines
CRSNG	Conseil de recherche en sciences naturelles et génie du Canada
Cs	Coefficient d'asymétrie
DebEst	Logiciel d'estimation des débits
DOS	Disk operating system
FCAR	Formation des Chercheurs & Aide à la Recherche
F <sub>Z</sub>	Probabilité de non-dépassement
H <sub>nette</sub>	Hauteur d'eau nette, m

K	Coefficient de distribution statistique
K <sub>T</sub>	Paramètre de fréquence propre à chaque fonction de distribution de probabilité
N	Nombre d'élément dans une série
P	Puissance générée, Watt
PREVAL	Logiciel de prévision et de validation des débits et des concentrations de polluants
Q	Débit, m <sup>3</sup> / s
Q <sub>entrée</sub>	Débit apport naturel, m <sup>3</sup> / s
Q <sub>sortie</sub>	Débit turbiné + débit évacué, m <sup>3</sup> / s
Q <sub>i</sub>	Débit pour l'indice i, m <sup>3</sup> / s
Q=f(h)	Q est une fonction de h
Q <sub>T</sub>	Débit pour la période de retour, m <sup>3</sup> / s
T	Période de retour, année
VALIDEB	Logiciel de validation des débits
V	Volume, m <sup>3</sup>

À Laure, mon épouse

et

à Fabien, mon fils

de la reconnaissance

pour leur compréhension

et leur patience.

## INTRODUCTION

La gestion de l'eau douce apparaît comme l'un des problèmes les plus critiques qui se pose à l'humanité en matière de ressources naturelles. Aujourd'hui, la population mondiale connaît une augmentation rapide. Or, le cycle hydrologique de notre planète qui comprend de grandes étendues d'eaux que sont les mers, les océans et les lacs, les eaux atmosphériques, les eaux de surface et les eaux souterraines, ne produit pas plus d'eau maintenant qu'il y a 2 000 ans. Cependant, à cette époque la population était beaucoup moins importante qu'à l'heure actuelle. Dans le processus du cycle hydrologique, le passage d'une étendue à l'autre s'effectue via les précipitations, l'évaporation et l'évapotranspiration, le ruissellement de surface, l'infiltration et l'écoulement souterrain. Ainsi, on peut dire que le cycle hydrologique est un processus naturel, en perpétuel changement avec ses apports et ses pertes, qui n'a pas de fin en soi.

Dans le but de satisfaire les différents besoins, le cycle hydrologique est modifié par les interventions humaines soit directement par les aménagements hydrauliques, les prélèvements, les rejets, et l'urbanisation ou indirectement par la déforestation, le défrichage, le brûlis et l'érosion. Pour contrer les effets nuisibles de ces interventions, l'humain doit se doter de mécanismes de contrôle et de gestion hydrologiques pour minimiser les impacts de son intervention sur l'environnement et sur le cycle de l'eau, comme par exemple l'analyse des données hydrologiques obtenues par des relevés sur les sites. Cependant, ces données ne sont pas toujours représentatives compte tenu des erreurs et imprécisions qui surviennent lors du processus de relevés ou de traitement. Il est donc nécessaire de corriger ces erreurs par la validation des données.

Le défi principal de la validation des données est d'arriver à concilier différents objectifs techniques mais aussi de permettre aux gestionnaires et concepteurs de mieux prévoir et planifier les impacts sociaux et économiques des décisions liées aux ouvrages hydrauliques. Cependant, cette étape n'est pas toujours reconnue à sa juste valeur par les

concepteurs et les gestionnaires. Ceux-ci peuvent concevoir et prendre des décisions en analysant les événements hydrologiques avec les données brutes.

L'objectif de notre étude est de déterminer l'impact des erreurs sur l'estimation des quantiles de crue. La conception et la gestion des ouvrages hydrauliques ont besoin d'une information primordiale qui est le débit de conception. Ce débit est obtenu par le calcul des quantiles de crue via le traitement des données hydrométriques mesurées sur les sites. On mettra donc en relation les principes utilisés pour la conception et la gestion des ouvrages d'art avec les impacts sociaux économiques.

L'aspect social a trait à la gestion des ouvrages, des inondations, de l'alimentation d'eau potable, de l'irrigation, de l'agriculture, des loisirs, etc. Quant à l'aspect économique, il se situe au niveau des coûts de conception d'ouvrages et par conséquent de gestion.

Cette étude est réalisée à partir de l'analyse de séries chronologiques de données hydrométriques fournies par Hydro-Québec et le Ministère de l'Environnement du Québec sur la rivière Romaine située dans le canton de Duplessis dans la province de Québec.

Nous avons débuté notre étude par une revue de littérature traitant de la validation des données, de l'hydrologie en général et de la conception et la gestion des ouvrages d'art. Nous y avons trouvé des principes et des méthodes de calcul nécessaires à notre méthodologie de travail.

La méthodologie est basée sur le calcul et la comparaison des quantiles de crue, quantiles des données historiques avec les quantiles des données modifiées, quantiles des données historiques avec les quantiles des données validées. À partir des données historiques fournies par Hydro-Québec, nous avons faussé ces données, d'une faible variation jusqu'à une variation aberrante, pour simuler les erreurs et imprécisions dans le

but d'analyser et de comparer les résultats. Ensuite, nous avons validé les données modifiées dans le but d'effectuer d'autres comparaisons.

La comparaison des différents résultats obtenus nous a permis de dégager l'impact des erreurs sur les quantiles de crue. Cet impact aura une répercussion sociale et économique lors de la conception et de la gestion des ouvrages.

## CHAPITRE 1

### LES DONNÉES HYDROMÉTRIQUES

#### 1.1 Généralités

Les ressources en eau douce représentent un des éléments importants de l'évolution des civilisations, pour son apport en irrigation agricole, pour le transport maritime ou fluvial, pour la consommation et pour la production d'énergie. La majorité de ces ressources disponibles est constituée principalement des précipitations atmosphériques de pluie ou de neige, des eaux des rivières, des lacs, des retenues des barrages et des eaux souterraines.

De nos jours, avec un nombre croissant d'utilisateurs, la pression exercée sur les ressources en eau est de plus en plus importante afin de satisfaire leurs besoins essentiels, de permettre le développement économique, de maintenir le milieu naturel et de pratiquer des activités récréo-touristiques. Cette pression liée aux incertitudes sur l'évolution climatologique nécessite d'apporter une attention particulière à l'observation et l'analyse des disponibilités et des variabilités des ressources en eau à l'aide de données hydrométriques denses, homogènes et fiables.

Les données hydrométriques permettent une meilleure compréhension des processus physiques régissant le cycle hydrologique dans le temps et l'espace, et ainsi garantir le maintien d'un niveau de qualité environnemental acceptable pour la société. De plus, les données constituent les seuls indicateurs fiables de l'état actuel et futur des réserves d'eau de surface, elles jouent donc un rôle indispensable dans le processus de prise de décisions concernant les ressources hydriques [1].

La plupart des secteurs de l'économie d'un pays utilisent les données hydrométriques pour satisfaire leurs exigences en matière de planification et de développement. Par conséquent, la gestion des eaux est capitale. Cependant, cette gestion est rendue complexe car elle doit être vue dans un ensemble de données mesurables et qualitatives, nécessitant de plus en plus une expertise qui doit tenir compte de plusieurs écosystèmes. Aussi, une bonne compréhension de l'hydrologie permettra une meilleure gestion des ressources en eau et ainsi minimiser les impacts sur l'environnement et mieux planifier les besoins socio-économiques.

Dans ce mémoire portant sur l'impact des erreurs sur l'estimation des quantiles de crue, nous nous sommes plus particulièrement intéressé aux données relatives aux débits de rivière. Une évaluation précise de ces débits est primordiale en raison des risques importants associés à une mauvaise connaissance de cette variable [2].

## **1.2 Utilité des données**

L'utilité des données hydrométriques est de fournir des outils pour atteindre une gestion efficace et durable des ressources en eau. Parmi les objectifs les plus importants, on retrouve la prévention des inondations, l'étude des effets des ressources hydriques sur l'environnement, la conception d'ouvrages hydrauliques ou la promotion des activités touristiques. Par conséquent, une connaissance et une gestion adéquates des ressources hydriques assureront une meilleure prise de décision pour atteindre ces objectifs.

Les inondations peuvent être prévues en caractérisant et numérisant les écoulements dans les cours d'eaux, en utilisant des débits de différentes récurrences calculés statistiquement dans le but de définir les niveaux de crue. Ces niveaux servent en particulier pour déterminer les zones inondables utiles pour les aménagements des municipalités riveraines.

L'étude des effets des ressources hydriques sur l'environnement permettra de minimiser les conséquences nuisibles sur la qualité et la quantité d'eau, telles que : la modification des zones de sédimentation, l'érosion des berges, les changements biologiques, les impacts sur la faune ainsi que les changements sur l'écologie végétale et animale.

La conception d'ouvrages hydrauliques et la gestion de leurs réservoirs nécessitent les connaissances sur les quantités d'eau arrivant au site afin de régulariser les débits des cours d'eau. Ces ouvrages d'art servent à satisfaire les besoins essentiels des populations et à assurer leur développement économique. On retrouve entre autres comme besoins : le contrôle des inondations, la production électrique et les loisirs. Tout en permettant le cycle naturel de croissance et de migration de poissons, la navigation, l'alimentation en eau domestique et industrielle, l'irrigation des sols et enfin la sécurité des ouvrages.

En terme de développement économique, on retrouve les activités touristiques qui requièrent des ressources hydriques en grande quantité et de bonne qualité. Pour certaines régions, ces activités sont cruciales pour leur économie et leur survie.

### **1.3 Instrumentation des ouvrages hydrauliques**

Au cours des dernières années, les problèmes liés à la gestion de l'eau et à la conservation de l'environnement se sont accrus considérablement, mais aujourd'hui ils sont plus complexes et ont souvent des répercussions internationales [1], d'où la nécessité de disposer d'un réseau de mesurage hydrologique en conséquence.

Au Canada, il y a environ 2 900 stations hydrométriques mesurant les niveaux d'eau et les débits. Elles sont exploitées soit par le gouvernement fédéral, les provinces, les territoires ou des compagnies privées ou para publiques. Il y a un nombre additionnel de 5 300 stations hydrométriques qui ne sont plus en activité, mais leurs données sont

conservées avec celles des stations en activité dans la banque nationale de données sur les eaux de surface. De plus chaque organisme possède sa propre base de données [1].

La plupart des stations sont localisées dans le sud du pays, c'est-à-dire où vivent les populations et où on retrouve donc des activités économiques importantes. Elles sont situées sur des lacs, des rivières, des fleuves et des cours d'eau de toutes dimensions, allant de bassins versants de quelques hectares à d'autres beaucoup plus importants.

Au Québec, une grande part de la connaissance hydrologique et hydraulique provient d'environ 250 stations hydrométriques réparties sur l'ensemble du territoire; elles sont gérées par le Centre d'expertise hydrique du Québec [3].

Les stations de mesure, appelées stations de jaugeage, mesurent les niveaux de rivière qui permettent de déterminer le débit à un point précis d'une rivière à tout instant. Ces stations opèrent à l'aide d'appareils d'enregistrement analogiques et d'appareils d'enregistrement numériques plus modernes et environ 45 % des stations du réseau fonctionnent par télémétrie permettant le traitement et la transmission de données en temps réel.

Les éléments d'une station de jaugeage sont la section de contrôle, le limnimètre qui est l'instrument de mesure de base comportant des dispositifs de lecture et d'enregistrement du niveau de l'eau, le puits, les prises d'eau, la guérite de protection des appareils et les appareils enregistreurs. À l'aide du limnimètre, les données relevées fournissent de l'information sur la hauteur d'eau d'un site (rivière, lac, ...) selon un procédé automatisé ou non, ce qui aura une influence sur le type d'erreurs que l'on peut retrouver dans les données mesurées.

## **1.4 Types de données**

Dans notre étude de l'impact des erreurs sur l'estimation des quantiles de crue nous avons utilisé comme données des débits de rivières.

Les mesures des niveaux d'eau constituent la base des données hydrométriques car elles peuvent être utilisées telles quelles ou servir au calcul de débit. Ainsi les débits de rivières sont calculés à l'aide de la courbe de tarage, les apports naturels sont calculés à l'aide de l'équation du bilan hydrique, les débits turbinés sont calculés en fonction de la puissance obtenue des groupes turbines alternateurs et les débits déversés sont calculés en fonction de l'ouverture des portes d'évacuation. Ces débits et niveaux sont nécessaires pour évaluer les ressources en eau naturelles disponibles, ou qui restent mobilisables, après utilisation en amont tout en respectant les besoins en aval [7].

### **1.4.1 Calcul des débits de rivières**

Les débits des cours d'eau sont déterminés par la relation entre le débit de la rivière et la hauteur d'eau ( $h$ ) mesurée en différents points dans une section de contrôle de la rivière. Le résultat reporté sur un graphique à l'échelle nous permet de tracer la courbe de tarage qui est une expression de la fonction  $Q = f(h)$ , voir figure 1.

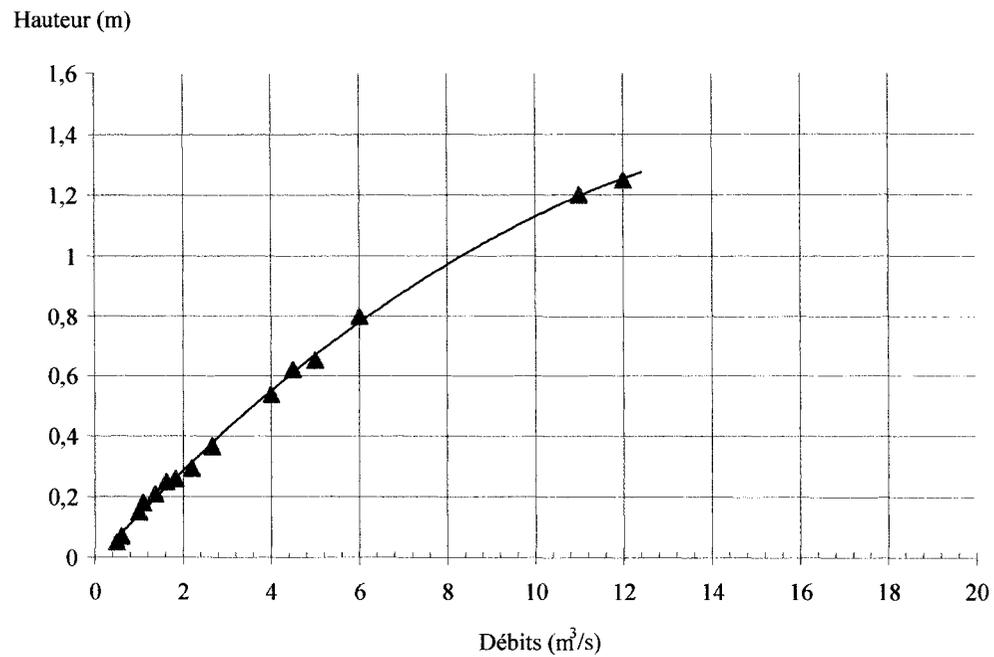


Figure 1 Exemple d'une courbe de tarage

Les données mesurées sont utilisées pour étalonner une relation (1.1) de la forme suivante :

$$Q = a (h + h_0)^n \quad (1.1)$$

Où :

$Q$  = débit, en  $m^3/s$

$a$  et  $n$  = constantes propres à la section de la rivière

$h_0$  = point bas du lit de la rivière ou correction pour obtenir le niveau réel

$h$  = profondeur d'écoulement, en m

Cette équation comporte systématiquement deux sources d'erreurs; la première, une erreur de modélisation qui lie  $Q$  et  $h$  et la deuxième, une erreur de mesure sur le débit ( $Q$ ) et sur la profondeur ( $h$ ) [4].

#### 1.4.2 Mesure des niveaux

Les niveaux de réservoir des barrages sont obtenus à partir de mesures brutes relevées directement sur les sites. Ces mesures servent à connaître la quantité d'eau emmagasinée et disponible ainsi que de calculer les apports naturels qui contribuent à une gestion optimale des ressources disponibles.

#### 1.4.3 Calcul des apports naturels

Les apports naturels sont déterminés par l'équation (1.2) du bilan hydrique qui considère les quantités d'eau qui entrent et qui sortent du réservoir étudié ainsi que les variations du niveau d'eau emmagasinée. Les apports naturels peuvent être calculés selon les besoins avec un pas de temps de 1, 3, 12 ou 24 heures. Toutefois, la mesure la plus courante est la mesure journalière. Les apports naturels sont utilisés comme donnée d'entrée dans différents modèles de prévisions à court, moyen et long termes. La reconstitution des apports naturels passe par la connaissance des débits de sortie, de la variation de niveau du réservoir et de la courbe d'emmagasinement du réservoir [5], voir figure 2.

$$\frac{dV(t)}{dt} = Q_{\text{entrée}}(t) - Q_{\text{sortie}}(t) + AN \quad (1.2)$$

Où :

$Q_{\text{entrée}}$  = débit apport naturel, en  $\text{m}^3/\text{s}$

$Q_{\text{sortie}}$  = débit turbiné + débit évacué, en  $\text{m}^3/\text{s}$

AN = apport naturel, en  $\text{m}^3/\text{s}$

$dV$  = variation du volume, en  $\text{m}^3$

$t$  = temps, en s

$dt$  = variation du temps

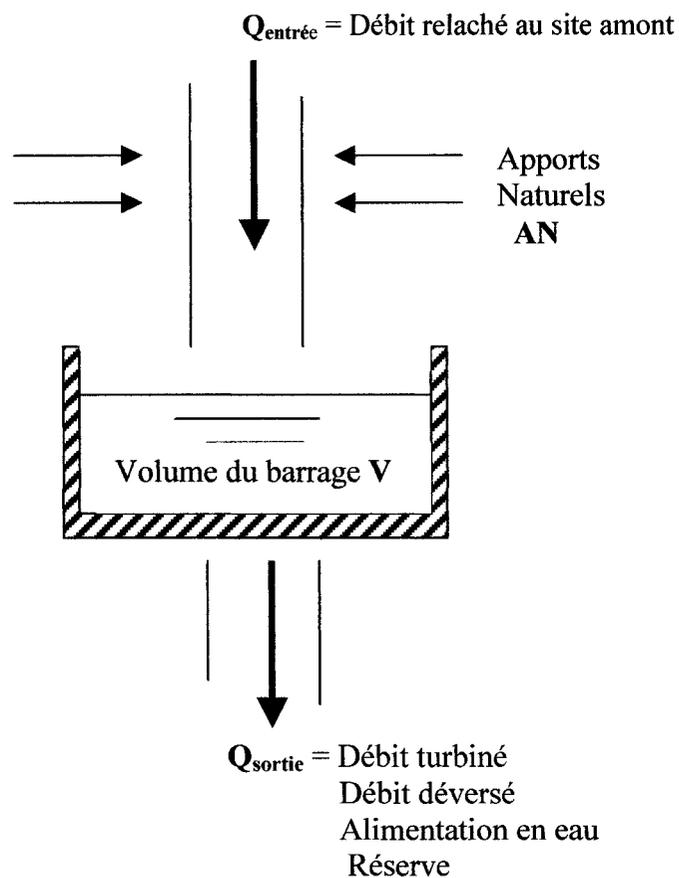


Figure 2 Méthodologie du calcul des apports naturels

#### 1.4.4 Calcul du débit turbiné

La puissance générée (P) d'un groupe turbine alternateur est calculée à partir de l'équation (1.3) qui est fonction de la hauteur de chute et du débit turbiné. Pour une puissance générée donnée et un niveau mesuré il est possible d'en calculer le débit [6].

$$P = \rho * g * Q * \eta * H_{\text{nette}} \quad (1.3)$$

Où :

P	=	puissance, en Watt
$\rho$	=	densité de l'eau
g	=	gravité 9,81 m/s <sup>2</sup>
Q	=	débit, en m <sup>3</sup> /s
H <sub>nette</sub>	=	hauteur d'eau nette, en m
$\eta$	=	rendement du groupe turbine alternateur

#### 1.4.5 Calcul du débit déversé

Le débit déversé est déterminé en fonction d'une équation tenant compte de la hauteur d'ouverture des portes d'évacuation, du type de portes, de la largeur de la passe, de la forme de l'évacuateur (pilier, type de déversoir, tablier du déversoir, etc.), du niveau d'eau amont et du niveau d'eau aval.

## 1.5 Erreurs dans les données

Comme nous l'avons vu précédemment, les données hydrologiques sont généralement difficiles à évaluer avec précision car elles dépendent de nombreux paramètres et de diverses variations dans le temps et dans l'espace et ce, principalement dû à divers éléments physiques comme les écoulements en surface, les processus hydrodynamiques dans les sols, les aménagements hydroélectriques (ouvertures et fermetures de vannes, etc.) et les activités humaines [7].

Le relevé et la transformation de données hydrométriques comportent toujours un degré d'erreurs provenant de différentes sources et pouvant survenir à différentes étapes : lors de la prise des mesures, de la transmission des mesures, du stockage ou encore durant le traitement des données.

Nous parlerons ici des erreurs de mesures car notre étude évalue l'impact de ce type d'erreurs. Une erreur de mesure est définie comme étant la différence entre la vraie valeur qui est l'idéal recherché, mais qui n'est en principe et malheureusement jamais connue, et la valeur mesurée [8]. Plusieurs facteurs naturels ou techniques peuvent fausser les composantes mesurées ; on les nomme les erreurs aléatoires et les erreurs systématiques. Chacune de ces erreurs a une importance et une pertinence à considérer pour la suite des travaux.

Les erreurs aléatoires affectent la précision des données. Leur origine est le plus souvent liée aux conditions extérieures comme : le vent, la seiche, les précipitations de pluie et de neige, les couvertures de glace et l'exploitation des ouvrages hydrauliques. Ce type d'erreurs présente un degré de retraçabilité en permettant une estimation plus précise de ces erreurs. Toutefois, une des difficultés qui se pose dans les séries de données, est la localisation et l'estimation de ces erreurs.

Les erreurs systématiques affectent quant à elles, la fiabilité des données. Leur origine est souvent liée au calibrage imparfait de l'appareil de mesure ou à un phénomène extérieur qui perturbe la mesure comme : les erreurs d'appareillage, les changements d'observateur, les bris d'instrument, la déformation du plan d'eau, la non représentativité des courbes servant aux calculs (courbe d'emménagement, courbe de rendement et courbe de caractéristique des réservoirs), les paramètres de laminage représentant le temps de cheminement de l'eau ou encore des problèmes de transmission des données [8]. Ce type d'erreurs caractérisé par des données largement aberrantes ou manquantes est relativement facile à localiser et à corriger par des méthodes de validation parce que les données ne respectent pas la tendance de variation de la série.

## **1.6 Conséquences des erreurs dans les mesures**

Étant donné que les mesures servent à la conception d'ouvrages hydrauliques et à la gestion des ressources hydriques, toute erreur sur les mesures peut fausser les résultats et avoir une influence sur la prise de décision avec comme conséquence un impact d'ordre social, économique et environnemental. L'impact sera plus ou moins grand selon la nature et l'importance de l'erreur.

Au plan social et économique, la prévision des phénomènes naturels liés au cycle de l'eau est primordiale pour permettre une gestion efficace des événements météorologiques. Citons en exemple un événement extrême vécu au Québec, les inondations de 1996 au Saguenay suite à des pluies diluviennes ayant eu des dégâts matériels importants ainsi que des pertes de vie humaine.

Les erreurs peuvent avoir des impacts environnementaux importants; par exemple, une augmentation du débit d'un cours d'eau par rapport aux conditions initiales provoque l'érosion des rives, alors qu'une diminution du débit entraîne la création de nouvelles

zones de sédimentation. Aussi, la fluctuation du niveau d'eau dans les réservoirs peut occasionner une érosion accrue des berges.

L'utilisation irrationnelle de la consommation des ressources hydriques est problématiques. Elle entraîne une dégradation des sols, une pollution des ressources naturelles par rejet et menacent l'approvisionnement en eau potable pouvant engendrer des pertes économiques considérables dues à des traitements coûteux.

En conséquence, il faut se poser les questions suivantes : s'il pleut sur une région donnée, à quelle quantité d'eau devrait-on s'attendre au barrage situé à plusieurs kilomètres en aval ? et comment être en mesure de prévoir les précipitations exactes au-delà de 2 ou 3 jours puisqu'une fois ces valeurs connues il est relativement facile de déterminer les apports naturels entrant dans un réservoir ? Ainsi, les gestionnaires se retrouvent face aux problèmes et aux décisions suivantes : soit les apports naturels sont sous-estimés et il y a danger de dépasser le niveau maximal déterminé lors de la conception du barrage, soit ils sont surestimés et l'eau sera déversée ou turbinée à moindre coût.

Il est donc indispensable, avant d'utiliser des séries de données, de se préoccuper de leur qualité et de leur représentativité en utilisant diverses techniques de validation. Une saine gestion et de bonnes décisions en dépendent. Des études démontrent que les pertes dues à la mauvaise gestion imputée à la qualité des données se chiffrent en millions de dollars par année [5]. De même, qu'une mauvaise utilisation de données aussi bonnes soient-elles entraînera des erreurs conséquentes et coûteuses.

### **1.7 Chaînes d'acquisition des mesures**

Avant d'être utilisées par les différents utilisateurs, les données doivent passer par les étapes suivantes : acquisition, traitement, contrôle et validation, organisation, diffusion et publication [8].

### **1.7.1 Acquisition des données**

L'acquisition de données consiste à procéder, par le biais d'un instrument de mesure, à acquérir de l'information (par exemple : hauteur d'eau d'une station limnimétrique, etc...). Le procédé peut être automatisé ou non. Ceci aura une influence sur le type d'erreurs que l'on peut retrouver.

### **1.7.2 Traitement primaire des données**

La donnée acquise nécessite souvent un traitement primaire afin de la rendre pertinente et exploitable. Il s'agit de la conversion de la mesure effectuée en une grandeur qui soit hydrologiquement significative (par exemple : transformation d'impulsions électriques en intensités de précipitations, de hauteurs d'eau en débits).

Le traitement des données inclut aussi le contrôle primaire des données qui comprend les contrôles de cohérence à l'exclusion de tous traitements statistiques. Dans la situation où l'on procède à l'acquisition de données de précipitations et de débits, on vérifie encore la cohérence temporelle des données acquises, à savoir par exemple qu'une crue est bien la conséquence d'un épisode pluvieux.

### **1.7.3 Contrôle des données**

Avant d'exploiter les données, il importe de contrôler leur fiabilité et leur précision. Le contrôle permet de les valider avant leur organisation dans une banque de données pour leur mise à disposition à des fins opérationnelles. Lors de cette opération, on introduit des indices de qualité de la donnée ainsi que des indices indiquant que celle-ci est reconstituée, calculée ou manquante.

#### **1.7.4 Organisation des données**

Au vu de l'importance quantitative et qualitative des données, il importe de les organiser avec soin. Ceci se fait à partir d'unité de stockage électronique et constitue les archives accessibles à un personnel spécifique. La traduction des archives sous la forme de fichiers de base génère les « fichiers en l'état » et fournit une indication sur la provenance de la donnée (mesure, calcul, copie, etc.) ainsi que sur sa qualité (fiable, complète ou non) et sa précision. Enfin, on constitue un « fichier de travail » provisoire permettant une visualisation des données et permettant de procéder aux différents tests de qualité et de précision. L'élément ultime de cette chaîne opératoire est la constitution des « fichiers opérationnels » avec indices de qualité pouvant être publiés et distribués auprès des utilisateurs.

#### **1.7.5 Diffusion et publication**

La diffusion des données en vue de leur utilisation opérationnelle peut se faire selon deux méthodes :

- 1) publication papier ou émission de bulletins;
- 2) diffusion informatique via le réseau Internet.

La figure 3 présente les différentes étapes de la chaîne d'acquisition et de traitement des données :

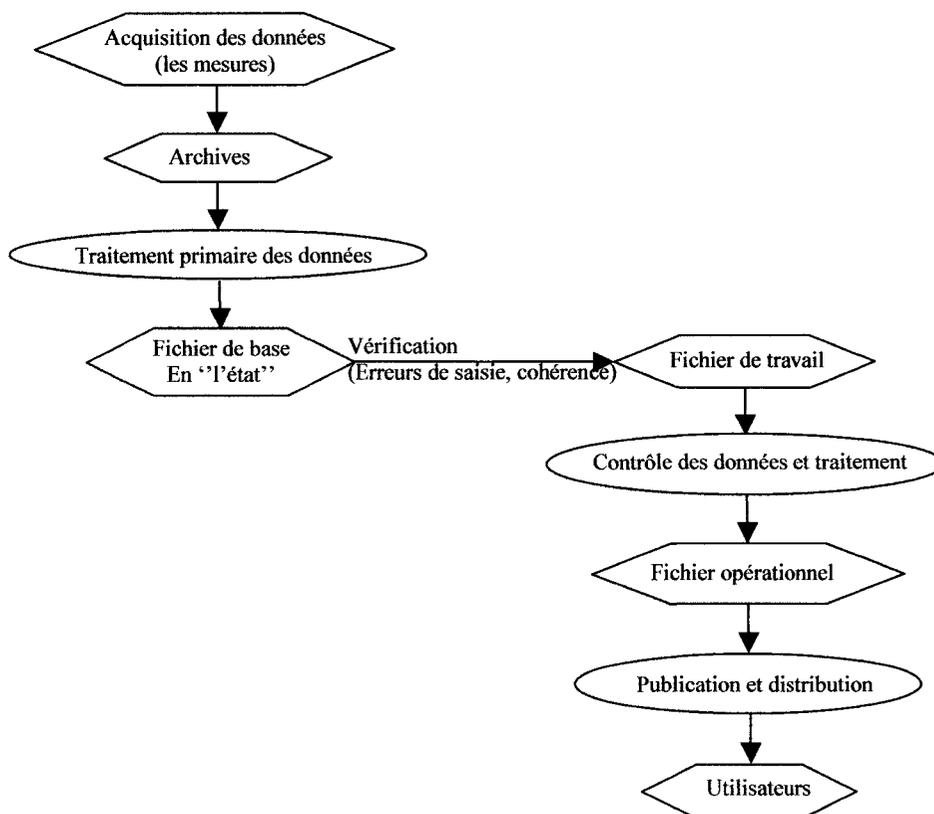


Figure 3 Chaîne d'acquisition et de traitement des données

### 1.8 Principaux utilisateurs des données

Parmi les principaux utilisateurs de données hydrométriques, on retrouve entre autres : les gouvernement fédéral et provinciaux via les ministères concernés (environnement, sécurité publique), et les centres d'expertise hydrologique, les centres de recherche sur l'hydrologie (privés et publics), les compagnies hydroélectriques, le secteur manufacturier et le secteur de l'agriculture.

Notre étude sur l'impact des erreurs sur l'estimation des quantiles de crue porte sur des données hydrométriques pouvant être utilisées par des compagnies hydroélectriques. Ces dernières se basent sur la prévision des apports en eau aux réservoirs lorsqu'il s'agit de la protection contre les inondations, de l'optimisation de la production d'électricité ou de la construction de nouvelles installations. Il est important de pouvoir estimer les apports futurs avec le plus de précision possible. Cela est toutefois difficile principalement en raison de l'imprécision des prévisions météorologiques, et de l'imprécision dans la modélisation du processus pluie-débit. Même avec un modèle « parfait », la qualité des prévisions d'apport est limitée par la qualité des prévisions météorologiques. De plus, l'échelle de temps d'intérêt diffère selon l'objectif de la prévision [9].

Au Québec, la compagnie Hydro-Québec est une des plus importantes compagnies de production d'électricité en Amérique du nord. Ses spécialistes contrôlent des réservoirs d'une capacité de 200 000 hm<sup>3</sup>, qui produisent 95 % de l'énergie consommée au Québec. Afin d'optimiser la production hydroélectrique, les spécialistes exigent une évaluation précise de l'énergie d'eau stockée dans les réservoirs. Ces données sont alors employées pour projeter et contrôler les usines hydroélectriques selon quatre horizons de temps : très court terme, court terme, moyen terme, et long terme [10].

Par conséquent, pour les compagnies de production hydroélectrique, il est important de connaître les apports des rivières sur lesquelles sont situés les barrages et les centrales dans le but de faire une gestion optimale des installations existantes. Ces connaissances permettent l'évaluation des différents risques associés à ces activités et assurent une gestion optimale et sécuritaires des réservoirs : De plus, elles sont utilisées pour le dimensionnement de nouvelles installations hydroélectriques.

## CHAPITRE 2

### LA VALIDATION

#### 2.1 Revue de la documentation

La littérature est peu fournie dans le domaine de la validation des données en général et des données hydrométriques en particulier. Cette lacune peut s'expliquer par deux raisons :

- 1) dans le passé, le contexte de surplus énergétique rendait l'impact de la qualité médiocre des données moins évident sur la productibilité et la fiabilité des systèmes hydroélectriques;
- 2) la généralisation et l'automatisation des procédures de contrôle de la qualité des données sont des tâches difficiles. On fait souvent face à des situations aberrantes plutôt qu'à des valeurs isolées [11].

Nous savons par l'étude de Gagnon [11] portant sur le filtrage des séries hydrométriques que plusieurs auteurs ont apporté des améliorations dans le traitement des données hydrologiques, à savoir :

- 1) Charbonneau et Bérubé (1987) ont développé une méthodologie faisant appel aux séries de Fourier pour corriger les mesures d'apports naturels historiques. Il s'agit en fait d'utiliser un filtre passe-bas ayant une fréquence de coupure prédéterminée sur la série que l'on désire valider. Cette approche diminue les fluctuations improbables observées dans les séries d'apports naturels historiques;

- 2) Perreault et Al (1991) ont développé un modèle de validation et d'estimation des apports naturels journaliers en temps réel. Leur approche est basée sur deux corrections des apports non filtrés. Une première effectuée a priori élimine les mesures tout à fait improbables, et la deuxième qui considère les apports filtrés par les séries de Fourier comme variable dépendante dans une équation de régression;
- 3) Bennis et Bruneau (1993) ont utilisé le filtre de Kalman pour améliorer l'estimation de débits journaliers. Le choix de ce filtre a été motivé par la précision que l'on peut obtenir sur l'estimation de données manquantes et la production de résidus indépendants.

Durant les dernières années, Bennis (1994) a développé plusieurs techniques de traitement des séries hydrométéorologiques. Il s'est consacré, entre autres, à la validation des débits journaliers mesurés directement en rivière, à l'estimation des données manquantes et au calcul des apports naturels à l'aide des niveaux d'eau. Deux approches, selon le nombre de limnimètres implantés sur le réservoir, ont été proposées pour corriger les séries chronologiques des niveaux. Le cas le plus favorable est celui où plusieurs stations limnimétriques existent sur le même réservoir. Dans ce cas, on adopte un filtrage multivarié pour valider les niveaux enregistrés par chacune des stations. Cette méthode est appelée « multifiltre » dont le principe consiste à comparer les écarts entre la valeur prévue par un modèle autorégressif, la valeur historique mesurée et la valeur estimée par un modèle régressif sur les stations avoisinantes. Parmi les valeurs mesurées et estimées par le modèle de régression sur les stations avoisinantes, celle qui se rapproche le plus de la valeur prévue est retenue. Le cas où une seule station de mesure existe sur un réservoir constitue le cas le plus défavorable. On procède alors à un filtrage univarié sur les niveaux de réservoirs. Cette méthode est appelée « unifiltre » [11].

Dans le cadre de notre étude, nous avons utilisé la méthode unifiltre pour valider des débits de rivières.

## **2.2 But de la validation**

La validation des données brutes représente une étape capitale dans le processus de traitement des données. Elle permet de vérifier les données même si celles-ci ont été mesurées à l'aide de moyens technologiques sophistiqués et ce, dans le but d'obtenir des résultats fiables et reflétant bien la réalité pour permettre de prendre des décisions appropriées. Il est à noter que la validation des données permet de limiter les erreurs mais non de les éliminer [12]. Toutefois; il n'existe aucun moyen de s'assurer de l'exactitude des données mesurées.

## **2.3 Objectifs de la validation**

Toute validation de données commence par l'assurance de travailler avec les bonnes données c'est-à-dire celles qui permettront d'atteindre les objectifs recherchés pour une gestion optimale du système hydrique. À partir de ces données, leur validation servira à améliorer la qualité des données mesurées en épurant les données aberrantes, en complétant les données manquantes et finalement en filtrant toutes les données historiques de la série pour les rendre appropriées et utilisables.

Le traitement de la validation des données peut être effectué manuellement, notamment pour les données aberrantes et les données manquantes, ou à l'aide d'outils informatiques pour l'ensemble de la validation. Selon la littérature parcourue, nous constatons que la validation des données fait encore l'objet de nos jours de recherches et de développements afin d'améliorer les méthodes de validation et les résultats.

En définitive, l'objectif principal de la validation des données consiste à outiller les décideurs et les concepteurs pour leur permettre une prise de décision la plus juste possible en fonction de la valeur et de la quantité de l'information utilisée [12].

La validation des données permet de déterminer la tendance d'une série de données historiques et d'effectuer un filtrage de l'historique à partir de filtres adéquats. Elle repose sur le fait qu'une série temporelle historique brute peut se décomposer en une somme d'une tendance, d'un bruit et d'une composante saisonnière. Étant donné que l'on travaille avec des données journalières, cette dernière composante peut être incluse dans la tendance. En effet, avec un pas de temps journalier, l'effet saisonnier se traduit par une variation lente et on peut supposer qu'autour d'un point la tendance est consistante. Il faut donc faire l'estimation de la tendance en effectuant un filtrage de l'historique. Un filtre peut s'interpréter comme une sorte de boîte noire qui discrimine, selon les attributs de ce qu'on lui présente à l'entrée, ce qui doit demeurer à la sortie. Il faut choisir un filtre qui gardera intacte la tendance tout en atténuant considérablement la composante aléatoire [11].

À titre d'exemple de validation, nous présentons à la figure 4 , une erreur relevée sur une série de données historiques de niveaux d'eau pour un très grand réservoir d'emmagasinement d'eau. Le niveau de la journée 20 montre une baisse du niveau d'eau d'environ 1,6 m ainsi qu'une remontée du niveau de la même valeur de 1,6 m, ce qui en pratique est impossible dans un laps de temps aussi court puisque cela impliquerait des débits d'eau considérables.

De plus, comme mentionné précédent, la tendance de la courbe des niveaux historiques n'est pas respectée. Après l'analyse des débits turbinés et évacués de l'ouvrage en amont ainsi que des débits turbinés de l'ouvrage considéré, cela a confirmé que cette variation n'est pas possible. Cette erreur peut être détectée par une pré-validation visuelle de la

série chronologique. La correction de cette erreur pourrait être faite en remplaçant le niveau erroné par la moyenne des niveaux des journées précédentes et suivantes.

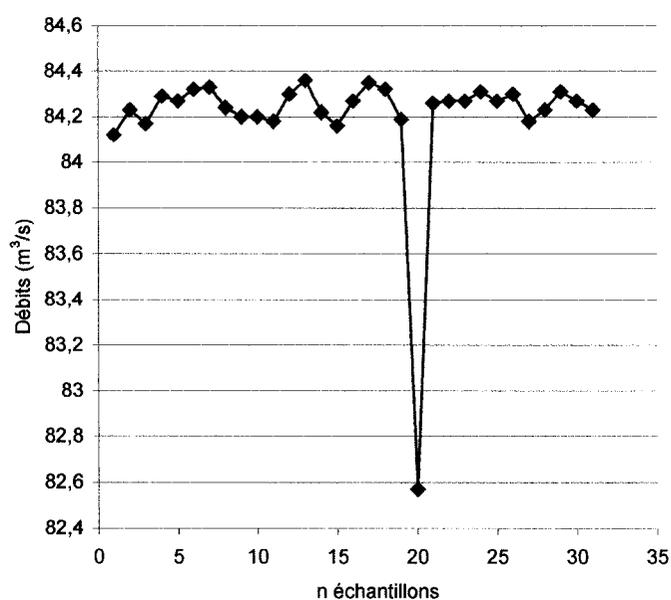


Figure 4 Erreur sur une mesure de niveau dans les données

Un autre exemple à la figure 5, représente une validation que nous avons effectuée sur des apports naturels historiques pour un réservoir situé au Québec. Pour ce faire, nous avons procédé, dans un premier temps, au calcul des apports naturels historiques par la méthode du bilan hydrique. Par la suite, nous avons appliqué une méthode de filtrage sur les apports naturels historiques.

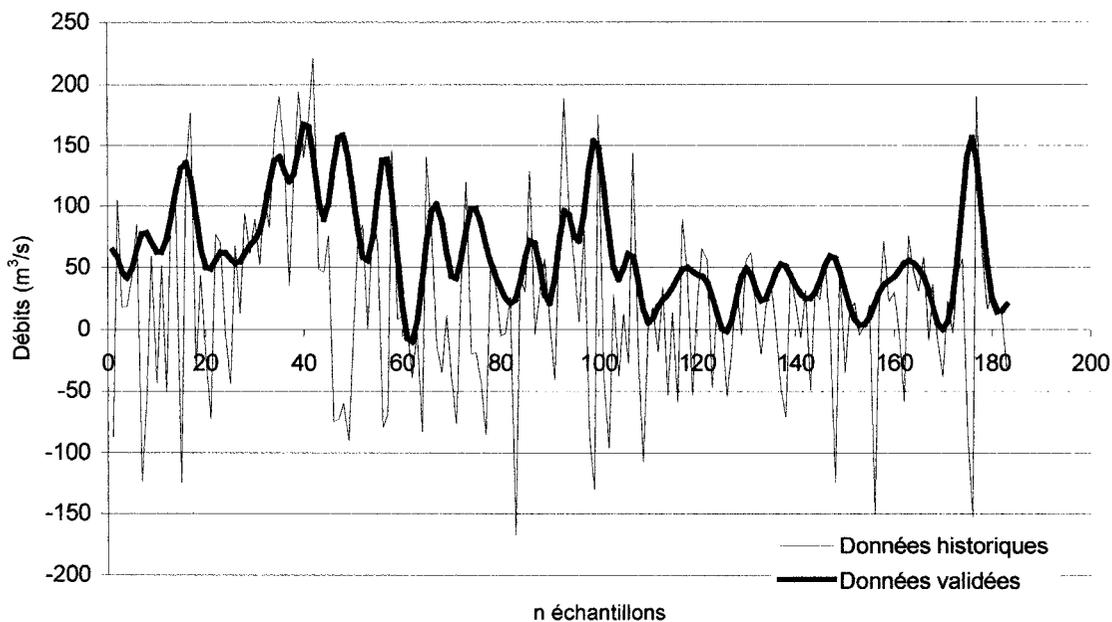


Figure 5 Comparaison des apports naturels historiques et validés

Comme on le voit à la figure 5, l'utilisation sans filtrage préalable des données conduit souvent à un calcul erroné où l'on peut voir plusieurs débits négatifs. Ce qui au Québec, est impossible étant donné le climat. Après validation des apports naturels historiques, les valeurs négatives ont pour ainsi dire toutes disparues et la tendance de variation des données a été respectée.

#### 2.4 Logiciel de validation

Dans le cadre de notre étude, nous avons utilisé le logiciel PREVAL pour valider nos données. Ce logiciel fut développé en trois étapes par le professeur Bennis (2004). Il avait comme première version, le logiciel DebEst qui visait l'estimation des valeurs manquantes de débits journaliers; son développement qui a eu lieu entre 1990 et 1992 était financé par Hydro-Québec et le CRSNG. Par la suite, avec des changements

apportés en 1995, le logiciel devint en deuxième version VALIDEB développé en Turbopascal pour un environnement DOS et visait le développement de méthodologies de validation des débits journaliers, d'apports naturels journaliers et de niveaux journaliers, en partenariat avec Hydro-Québec, le CRSNG et le FCAR.

Durant la période 2000-2004, Bennis (2004) dans le cadre du *projet de recherche et développement de méthodologies et conception d'un logiciel de validation et de prévision des débits et des concentrations de polluants en milieux urbain et rural*, le rapport final a été soumis au CERIU, la troisième version PREVAL actuellement en cours de développement vise des objectifs fort différents selon les contextes rural et urbain. En milieu rural, les objectifs poursuivis sont :

- 1) transférer les algorithmes de calcul du langage Turbopascal à DÉLPHI;
- 2) généraliser les méthodes de validation et d'estimation des données manquantes à des pas de temps allant à la minute;
- 3) intégrer les modèles de laminage de crues pour tenir compte des délais de propagation et d'atténuation des pointes;
- 4) développer des méthodes de pré-validation des données.

Par ailleurs, en milieu urbain, en plus des quatre objectifs énoncés précédemment, PRÉVAL permet de calibrer des modèles hydrologiques et de faire les bilans des déversements des volumes et des charges polluantes en temps de pluie. Les partenaires qui financent PRÉVAL sont : Hydro-Québec, Ville de Montréal, Ville de Laval, Arrondissement de Verdun, CERIU, Environnement ESA, et le CRSNG.

Ce logiciel fonctionne sur micro-ordinateur sous environnement Windows avec interface graphique. On accède aux différentes fonctions par l'intermédiaire de menus déroulants avec la possibilité de visualiser simultanément toutes les séries analysées et validées. Il permet entre autres la pré-validation et la validation des séries de mesure en analysant les

données. Ces analyses sont effectuées en utilisant diverses méthodes de filtrage et en jugeant les résultats et leurs effets sur chacune des composantes du bilan hydrique. Pour la validation des données, on peut avoir recours à des stations de référence situées soit dans le même bassin soit dans un bassin voisin.

La figure 6 présente les fonctions de pré-validation (logiciel Préval, 2004) qui nous permettent d'imposer une valeur minimale, maximale ou un gradient à la série chronologique de données à analyser. Cette étape nous permet entre autres de supprimer les erreurs aberrantes dans notre série, par exemple les apports naturels négatifs, les variations de niveau trop importantes, etc.

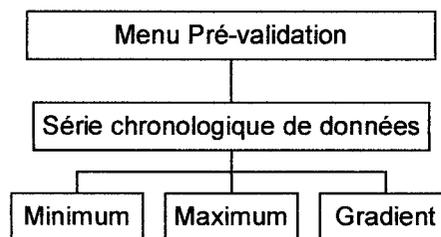


Figure 6 Menu de pré-validation des données de PREVAL

Le menu de validation de la figure 7 (logiciel Préval, 2004) permet de sélectionner le filtre de validation le plus adapté au cas étudié les filtres sont : le filtre fréquentiel, le filtre Fraser, le filtre à régression linéaire, le filtre à régression quadratique, le filtre médian et les filtres winsorizeur (a priori ou a posteriori). D'autres fonctions sont disponibles pour la validation des données telles que des fonctions d'estimation des données manquantes et des fonctions rattachées au calcul du bilan hydrique.

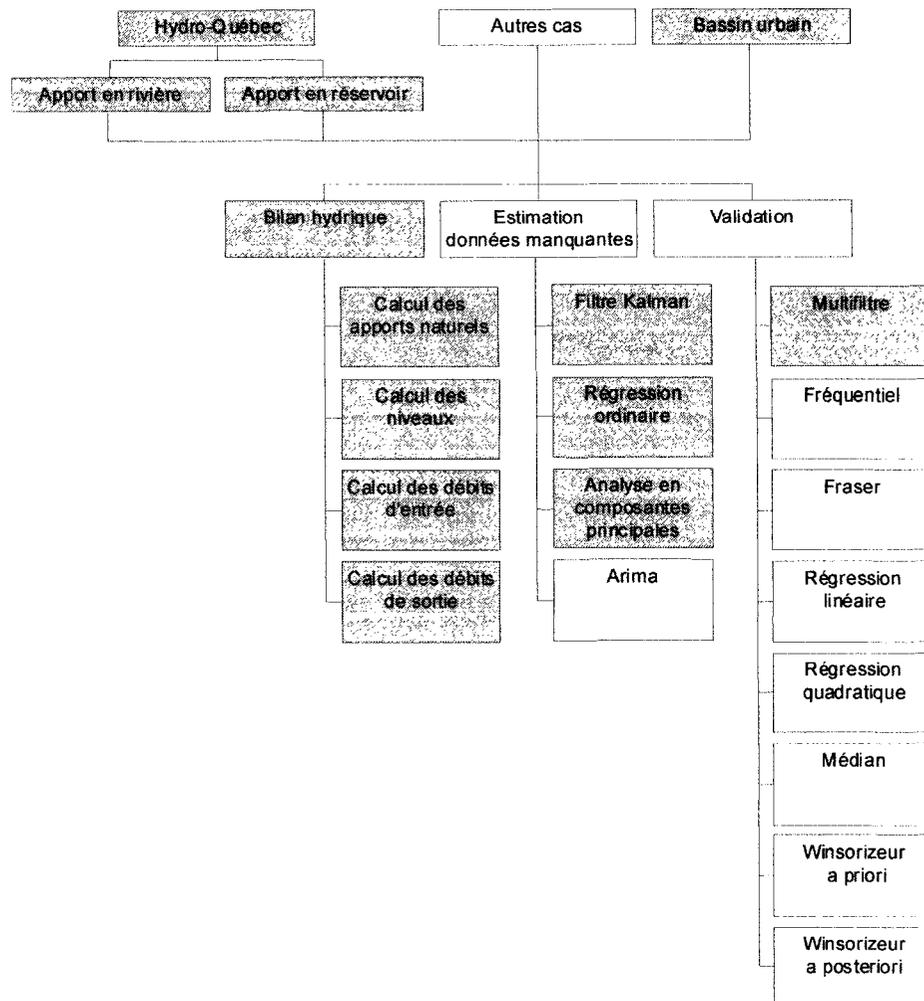


Figure 7 Menu de validation et de calcul de PREVAL

Les cases ombragées sont des fonctions de validation de données qui lors de notre étude étaient en cours de développement pour l'environnement WINDOWS.

## 2.5 Traitement des données

La recherche d'un modèle hydrologique ou son application en vue de faire de la prévision, de la simulation ou de la prédétermination, nécessite au préalable une analyse

et un contrôle minutieux et critique des données. De ce contrôle dépendent toutes les étapes subséquentes de l'étude. Gagnon [11] conclut que trente à cinquante pour cent du travail devrait être alloué à la critique des séries de données qui seront utilisées. Autrement, une erreur sur les données météorologiques enregistrées ou prévues se reflétera par une erreur équivalente sur la conception, la planification et la gestion d'un système hydrique.

Pendant le traitement des données, un des problèmes que l'on peut rencontrer sont les données manquantes. Celles-ci peuvent être corrigées lors de la validation.

### **2.5.1 Estimation des données manquantes**

L'estimation des données manquantes peut se faire par plusieurs méthodes telles que : le filtre de Kalman, la régression multiple et pseudo-orthogonale, l'analyse en composantes principales et la méthode ARIMA.

Toutes ces méthodes d'estimation sont adaptées à des circonstances précises selon que l'on dispose d'une seule ou de plusieurs séries de données et selon le nombre de données dont on dispose avant et après le point où se fera l'estimation.

#### **2.5.1.1 Technique univariée**

Cette technique adopte la méthodologie de deux modèles autorégressifs pour l'estimation des données manquantes. Le premier modèle autorégressif qui fonctionne dans le sens normal du temps est exprimé comme suit :

$$X_t^A = \theta_{t/t-1} \times X_{t-1}^A + W_{t-1}^A \quad (2.1)$$

Où :

A = désigne la direction aller dans le temps

La grandeur qu'on cherche à estimer à l'instant t ( $X_t$ ) se calcule en fonction de la même grandeur estimée ou mesurée à l'instant t-1,  $X_{t-1}$  moyennant la matrice de transition

$\Phi_{t/t-1}$

$W_{t-1}$  est un terme aléatoire qui représente les bruits de mesure et de modélisation.

Le deuxième modèle autorégressif qui fonctionne dans le sens rétrograde du temps est exprimé comme suit :

$$X_t^R = \theta_{t+1/t} \times X_{t+1}^R + W_{t+1}^R \quad (2.2)$$

Où :

R = la direction retour dans le temps

Chaque donnée manquante est estimée deux fois. Une première fois par extrapolation du modèle (2.1) dans le sens croissant du temps et une seconde fois par extrapolation du modèle (2.2) dans le sens rétrograde du temps. Afin de comprendre le fonctionnement des modèles, on considère une série de données mesurées à partir de l'instant « 0 » à « t », une succession de données manquantes de « t + 1 » à « t + M » et une série de données mesurées disponibles de « t + M + 1 » à « t<sub>f</sub> ». L'estimation aller se fera de « t + 1 » à « t + M » et l'estimation retour de « t<sub>f</sub> » à « t + M + 1 ».

## 2.5.2 Méthodes de filtrage des données

Le choix de la méthode de validation dépend du type d'erreur à corriger et du nombre de séries de données dont on dispose. La validation permettra d'éliminer et de reconstituer

au mieux les valeurs aberrantes dans les séries chronologiques. On distingue deux méthodes de validation : les méthodes de filtrage univarié utilisées dans le cas où on dispose d'une seule série de données et les méthodes de filtrage multivarié utilisables lorsqu'on se retrouve en présence de plusieurs séries de données. Toutefois, quelle que soit la méthode de filtrage utilisée, elle doit être en mesure de trouver un compromis entre les deux objectifs essentiels suivants :

- 1) éliminer au maximum les bruits de mesure;
- 2) déformer au minimum le signal de base.

#### **2.5.2.1 Validation par filtrage univarié**

Il existe plusieurs types de filtres pour la validation univariée. Il s'agit du :

- 1) lissage de type fenêtre mobile;
- 2) lissage exponentiel;
- 3) filtrage par fréquence.

##### ***2.5.2.1.1 Lissage de type fenêtre mobile***

Ce lissage permet l'ajustement d'un polynôme à toutes les observations d'une série. Cela peut conduire à retenir un degré plus élevé. L'estimation des coefficients peut dans ces conditions s'avérer peu précise. De ce fait, plutôt que de retenir un ajustement global, l'estimation se fait en un point à l'aide des observations entourant ce dernier par une droite (régression linéaire), une parabole (régression quadratique), une médiane. Les résidus d'ajustement étant supposés centrés, il en résulte une atténuation de leur importance et par conséquent leur lissage.

**a) Filtre à régression linéaire**

Il a l'avantage d'approximer par une droite, la tendance de  $(2m + 1)$  points de mesures successives représentant la largeur de la fenêtre;  $m$  est un entier positif choisi a priori. Ainsi, pour l'estimation de la mesure à un instant  $t$ , la tendance sur cette fenêtre est un segment de droite d'équation :

$$\text{tendance}(k) = \alpha k + \beta \quad \text{avec } -m \leq k \leq m \quad (2.3)$$

Où :

$\alpha$  et  $\beta$  sont déterminés par la méthode des moindres carrés qui minimise la somme quadratique donnée par :

$$S(\alpha, \beta) = \sum_{k=-m}^m (\alpha k + \beta - x_k)^2 \quad (2.4)$$

**b) Filtre à régression quadratique**

Ce filtre permet d'effectuer un filtrage en approximant la tendance sur  $(2m + 1)$  points (où  $m$  représente le nombre de données d'une série) par un polynôme de second degré de la forme de :

$$\text{tendance}(k) = \alpha + \beta k + \gamma k^2 \quad \text{avec } -m \leq k \leq m \quad (2.5)$$

Où :

$\alpha$ ,  $\beta$ , et  $\gamma$  sont déterminés par la méthode des moindres carrés :

$$S(\alpha, \beta, \gamma) = \sum_{k=-m}^m (\alpha + \beta k + \gamma k^2 - X_k)^2 \quad (2.6)$$

Ce filtre se comporte mieux pour une courbe en dents de scie. On note également que pour un même nombre de points le filtre à régression quadratique reproduit bien les pointes.

**c) Filtre médian**

Contrairement aux deux méthodes précédentes, le filtre médian n'est pas paramétrique. Il est le mieux adapté à l'élimination des valeurs aberrantes et consiste simplement à prendre la médiane et non la moyenne pour estimer le paramètre central d'une population. Sur une fenêtre de largeur  $(2m + 1)$ , avec  $m$  entier positif choisi a priori, le filtre s'exprime par :

$$X_t = \text{médiane}\{X_{t-m} + X_{t-m+1} + \dots + X_t + \dots + X_{t+m}\} \quad (2.7)$$

La médiane d'une série continue ordonnée de la façon suivante  $(x_1 < x_2 < \dots x_p \dots < x_n)$  est donnée par :

$$\begin{aligned} \text{Md}\{x_1, x_2, \dots, x_p, \dots, x_n\} &= x_p & \text{si } n &= 2p+1 \\ &= 1/2(x_p + x_{p+1}) & \text{si } n &= 2p \end{aligned} \quad (2.8)$$

Ce filtre a l'avantage d'être plus robuste que les deux méthodes vues précédemment. Car une observation aberrante dans les deux précédentes méthodes affectera de façon importante le lissage des données. Donc, dans les séries où l'on suspecte la présence d'observations aberrantes, il est souhaitable d'utiliser un filtre médian.

**d) Filtre « ébarbeur » et « winsorizeur »**

Ce filtre constitue un compromis entre le filtre médian et les filtres moyenniers. Il consiste à exclure les observations les plus éloignées de la médiane avant de calculer la moyenne. Cette opération est connue sous le nom d'ébarbage. Les valeurs exclues

sont ensuite reconstituées à partir des valeurs les plus proches conservées; c'est le winsorizing. Ce type de filtre donne un filtrage assez régulier en adoucissant les pentes abruptes et en enlevant les irrégularités.

### 2.5.2.1.2 *Lissage exponentiel*

#### a) **Lissage exponentiel simple**

C'est un type de lissage défini par la récurrence suivante :

$$Y_t = \alpha X_t + (1 - \alpha) Y_{t-1} \quad \text{avec } 0 < \alpha < 1 \quad (2.9)$$

Où :

$Y_t$  = la valeur lissée

$X_t$  = la valeur brute

$\alpha$  = constante de lissage appelée aussi facteur d'oubli

Ce filtre peut aussi s'écrire sous forme d'une combinaison linéaire infinie d'observations du passé avec une pondération  $\omega_k = \alpha(1-\alpha)^k$  qui décroît géométriquement lorsqu'on s'éloigne de ce dernier. Cet estimateur est sans biais pour une série chronologique stationnaire ( Bosq et Lecoutre, 1992) et la variance de la série lissée est moindre que celle de la série originale.

#### b) **Lissage exponentiel double**

Le lissage exponentiel double est en fait un lissage exponentiel simple appliqué deux fois. Les équations qui le caractérisent sont les suivantes

$$a0_t = \lambda x_t + (1 - \lambda) [a0_{t-1} + a1_{t-1}] \quad (2.10)$$

$$a1_t = \mu [a0_t - a0_{t-1}] + (1 - \mu) a1_{t-1} \quad (2.11)$$

avec

$$k = \alpha - (1 - \alpha)^2 \quad (2.12)$$

$$\mu = \alpha / (2 - \alpha) \quad (2.13)$$

Ce type de filtre qui se prête beaucoup mieux à la prévision à court terme est conçu de façon à répondre rapidement à une variation brusque de l'historique en ne sous estimant pas les valeurs futures.

### **2.5.2.1.3 Filtrage par fréquence**

Cette technique consiste à appliquer à la série originale de données un filtre qui enlèvera les composantes fréquentielles non désirables. Les filtres les plus utilisés sont les filtres passe-bas, passe-haut et passe-bande. Pour ce type de filtre, les phénomènes mis en cause ont des périodes très longues et inversement des fréquences très basses ce qui fait que le filtre passe-bas est le plus adéquat pour l'élimination de fluctuations indésirables.

### **2.5.2.2 Validation par filtrage multivarié**

On a recours à ce type de filtrage quand on se retrouve dans un contexte où plusieurs capteurs mesurent la même grandeur au même endroit. Dans la validation multivariée, trois grandeurs sont à considérer :

- 1)  $G_m$  : la grandeur mesurée par le capteur qui est la valeur à valider;
- 2)  $G_R$  : l'estimation de la grandeur à partir d'un modèle de régression multiple dont les paramètres sont estimés par le filtre de Kalman;

- 3)  $G_{AR}$ : l'estimation de la grandeur à partir d'un modèle auto régressif dont les paramètres sont estimés par le filtre de Kalman.

L'algorithme de validation est itératif et il est appliqué dans les sens aller et retour du temps. La valeur mesurée n'est jugée aberrante que si elle a été éliminée dans les deux sens.

Lors de l'utilisation d'une méthode de validation, il est important de connaître les objectifs recherchés car chaque méthode possède ses avantages et ses inconvénients et présente une efficacité différente suivant le type de données à traiter. Le choix d'une méthode de filtrage dépend :

- 1) de l'importance des aberrations dans les séries;
- 2) de la régularité dans la variation de la série;
- 3) du nombre de valeurs aberrantes présentes dans le signal;
- 4) du degré de lissage désiré;
- 5) du but recherché : conservation des pointes, du volume, etc.

La figure 8 illustre un exemple d'application des techniques univariées de filtrage sur une série de débits mesurés.

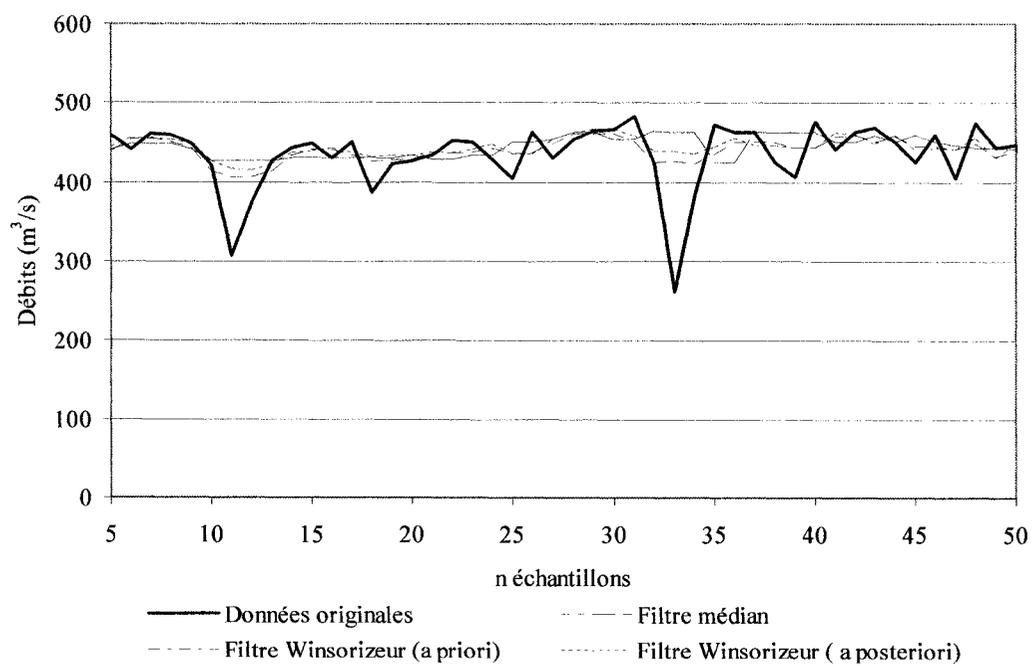
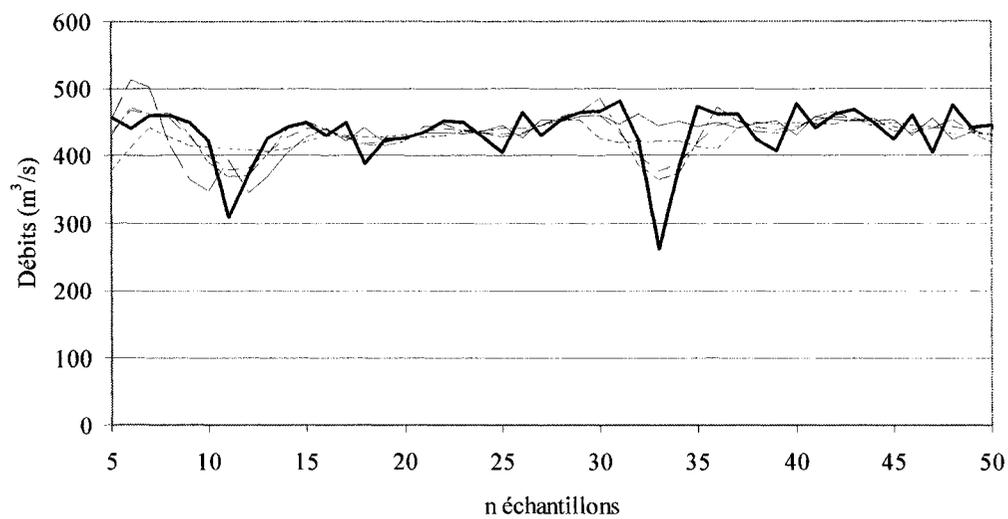


Figure 8 Méthodes de validation

## CHAPITRE 3

### MODÈLE DE PRÉVISION

#### 3.1 Introduction

L'utilité principale des modèles de prévision est de fournir des informations pour permettre une conception appropriée des ouvrages et une gestion efficace et durable des ressources en eau [13]. Cette gestion efficace passe par la connaissance des débits des cours d'eau.

Les débits importants à connaître pour la conception des ouvrages sont ceux de la crue maximale probable, la crue de printemps et la crue d'été-automne pouvant couler dans un cours d'eau et arrivant à un réservoir. Celles-ci sont déterminées par la combinaison d'événements météorologiques extrêmes (précipitation, couvert nival, température) et par l'analyse des crues au site considéré, afin de générer des hydrogrammes de crues résultant de ces événements. Ces hydrogrammes serviront à prévoir, à planifier, à gérer les ressources en eau naturelles pour concevoir des ouvrages d'art ou encore pour délimiter des zones inondables.

L'analyse des crues est réalisée à partir de modèles mathématiques basés sur des lois mathématiques ou statistiques. Ils utilisent entre autres comme informations les caractéristiques physiques du bassin versant et les séries chronologiques de débit de rivière ou de niveaux de réservoir mesurés sur les sites.

### 3.2 Modèles mathématiques

Les événements hydrologiques qui influencent l'écoulement de l'eau sur la surface du bassin et dans le sol ainsi que dans les cours d'eau sont représentés par des modèles mathématiques. Ces modèles simulent schématiquement la réaction d'un bassin versant en fonction des événements météorologiques [14]. La modélisation mathématique est donc utilisée comme outil de recherche pour interpréter les données mesurées. C'est aussi un outil de prévision pour anticiper des évolutions futures du débit d'un cours d'eau et également un outil d'extrapolation pour reconstituer les séries de données manquantes. La figure 9 représente les différentes approches de modélisation en distinguant de couleur grise les plus couramment employées en modélisation pluie-débit [15].

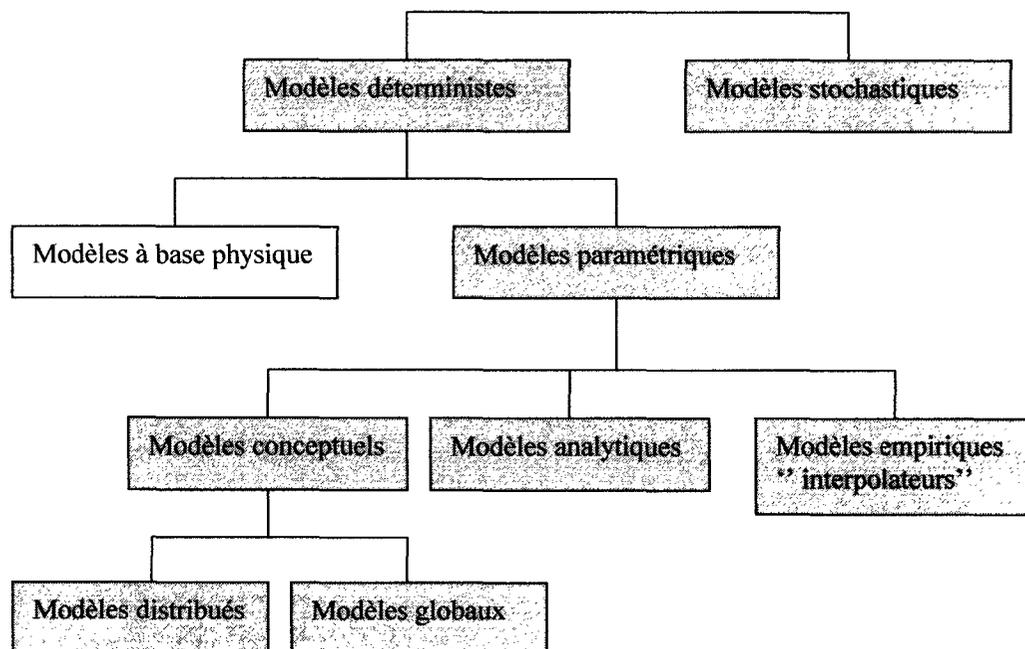


Figure 9 Les différentes approches de modélisation

Pour effectuer des prévisions hydrologiques, on se sert principalement de trois types de modèle : les modèles déterministes, les modèles conceptuels et les modèles empiriques [9].

Les modèles déterministes sont basés sur des connaissances physiques du bassin versant en utilisant les équations de continuité, d'énergie et de moment pour présenter le déplacement de l'eau dans les bassins versants.

Les modèles conceptuels se caractérisent par une représentation simplifiée mais plausible des principales composantes du processus pluie-débit.

Les modèles empiriques, quant à eux, mettent l'accent sur l'identification d'une relation entre les précipitations et les débits sans se préoccuper explicitement des mécanismes physiques de transformation des précipitations en débits [9].

Selon l'étude à effectuer, sommaire, préliminaire ou de faisabilité, le concepteur ou le gestionnaire aura recours à ces modèles pour effectuer les analyses de crue. L'analyse et le modèle utilisés peuvent aussi dépendre des données disponibles ou du type de résultats recherchés.

### **3.3 Analyse des crues**

Les crues sont des événements hydrologiques avec comme particularité une montée plus ou moins rapide du débit d'un cours d'eau, associée à une augmentation du débit. Le niveau d'eau maximal ou le débit maximal atteint lors d'une crue redescend à son niveau normal plus lentement. En situation de crue exceptionnelle, un débordement du cours d'eau hors de son lit peut survenir et causer des dommages importants soit par le niveau atteint, soit par la durée des inondations ou encore par la vitesse des courants. Par

conséquent, pour des raisons de sécurité il est très important d'être en mesure de prévoir ces événements exceptionnels. Les ingénieurs analysent les crues en fonction de l'étude à entreprendre et selon différentes méthodes [14], on retrouve :

- 1) la méthode stochastique pour l'analyse du risque lors du choix de la crue de conception des évacuateurs de crue pour les systèmes hydriques d'envergure en raison de leur taille, du nombre des réservoirs, des aspects juridiques, etc.;
- 2) la méthode mixte déterministe-statistique de la crue maximale probable, pour l'analyse du risque lors du choix de la crue de conception des évacuateurs de crue;
- 3) la méthode statistique pour l'analyse du risque lors du choix de la crue de conception pour les évacuateurs de crue ainsi que pour les études relatives aux ouvrages de dérivation provisoires et pour des études liées à l'environnement.

### **3.4 Méthode stochastique**

Cette méthode consiste à produire des séries chronologiques artificielles d'apports d'eau pour plusieurs sites appartenant à un même système hydrique tout en considérant la relation spatiale entre les différents sous-bassins. L'objectif est d'obtenir des données statistiques identiques aux apports historiques [14].

### **3.5 Méthode déterministe-statistique**

Cette méthode permet de calculer et de simuler la crue maximale probable en utilisant un modèle numérique de terrain. Cette crue est souvent considérée pour la gestion et la conception d'ouvrages hydrauliques car elle cumule les conditions climatiques et hydrologiques extrêmes pouvant survenir dans une région donnée. Pour la déterminer, il faut simuler le comportement hydrologique du bassin étudié selon diverses conditions

météorologiques à l'aide d'un modèle numérique. Les modèles hydrologiques de terrain les plus couramment utilisés sont le modèle conceptuel HSAMI, le modèle HYDROTEL et le modèle SSARR [14].

Ces différents modèles de calcul permettent entre autres de reconstituer des débits, de simuler la propagation des débits, de simuler la régularisation des retenues, de tester diverses variantes d'ouvrages hydrauliques, etc. La méthodologie de calcul de la crue maximale probable s'ajuste avec le temps grâce aux recherches qui apportent des connaissances hydrologiques et météorologiques nouvelles (évaluation plus précise des phénomènes météorologiques) [14].

### **3.6 Méthode statistique**

La méthode statistique est utilisée pour effectuer des prévisions sur les débits d'apports d'eau futurs. Dans le contexte d'une gestion en temps réel d'un réservoir, l'ingénieur fait appel aux notions développées dans l'hydrologie déterministe. Il prévoit ainsi les apports d'eau pour les prochaines heures à partir des données météorologiques mesurées ou prévues. Dans l'étape de conception de l'ouvrage, il n'est pas possible de prévoir les apports d'eau d'une manière déterministe. À titre d'exemple l'ingénieur doit a priori fixer le débit journalier maximal que les évacuateurs de crues sont capables d'évacuer. Le dimensionnement des évacuateurs de crue doit être prévu en fonction des apports d'eau pour toute la durée de vie que l'ouvrage peut atteindre. Ainsi, pour déterminer une crue qui a une période de retour de 100 ans, les données météorologiques ne peuvent pas être connues pour cette durée, il faut donc déterminer les apports d'eau pour cette durée de vie en utilisant une méthode purement statistique [4].

Par ailleurs, on utilise cette méthode pour déterminer les crues de conception pour les différents ouvrages d'un aménagement hydroélectrique. Ces crues varient généralement entre la pointe de la crue décennale (1 / 10 000 ans) et la crue 1 / 20 ans; elles sont,

par conséquent, calculées statistiquement étant donné parfois l'absence de données sur ces périodes de temps [14].

Pour cette méthode d'analyse, les concepteurs considèrent deux types de crues, soit :

- 1) la crue de printemps qui est souvent la plus importante due à la combinaison entre la fonte des neiges et les précipitations printanières. Celle-ci s'accumule en grande partie ou en totalité dans les grands réservoirs dont le niveau est le plus bas de l'année (le niveau est bas à cause de l'exploitation hivernale et des faibles apports naturels);
- 2) la crue d'été-automne qui survient alors que les réservoirs sont pleins (dépend encore du mode de gestion des réservoirs).

C'est en tenant compte des crues qui sont observées aux stations que les concepteurs établissent des fréquences de retour des débits en utilisant les différentes lois d'extrapolation connues. Ces lois permettent de calculer et d'analyser les fréquences de retour des crues dites périodes de récurrence des crues. Elles permettent de déterminer les pointes et les volumes, d'établir les paramètres de distribution des débits de pointe et d'en déduire des relations permettant d'estimer ces variables pour différentes périodes de récurrence, toujours dans le but de concevoir des ouvrages hydroélectriques. Il existe différentes lois d'extrapolation des crues. Le tableau I énumère quelques lois utilisées au Québec [14].

Tableau I

Lois d'extrapolation des crues les plus couramment utilisées

<b>Lois d'extrapolation</b>	<b>Applicabilités aux volumes</b>	<b>Applicabilités aux pointes</b>
Normale	Très rarement	Jamais
Log normale	Parfois	Rarement
Pearson III	Donne parfois un bon résultat	Donne parfois un bon résultat
Gûmbel	Peu fréquemment	Souvent choisie pour les séries de données de moins de 25 années
Log Pearson	Souvent	Souvent la meilleure, surtout si les séries de données ont plus de 25 années

Des outils informatiques spécialisés tels que HYFRAN, SAS et MATLAB sont employés pour utiliser ces différentes lois d'extrapolation étant donné la quantité et la complexité des données à traiter. Le tableau II indique l'utilisation des périodes de récurrence pour la conception d'ouvrage d'art [14].

Tableau II

Récurrences des crues les plus couramment utilisées

<b>Récurrences (ans)</b>	<b>Utilisations</b>
20	Conception d'un ouvrage de dérivation temporaire servant pour un an maximum.
40	Conception d'un ouvrage de dérivation temporaire servant pour deux ans au maximum.
60	Conception d'un ouvrage de dérivation temporaire servant pour trois ans au maximum.
100	Seuil minimal pour la conception d'un ouvrage d'évacuation permanent selon une étude de risque.
200 500 1 000	} Intrans à une étude de risque faite au cours d'une étude de faisabilité.
10 000	

Les études de crue doivent toujours être faites à l'échelle du bassin versant considéré, quelle que soit la méthode adoptée, étant donné que tous les ouvrages en amont ont des impacts sur la conception de ceux en aval [14].

### 3.7 Bassin versant

Un bassin versant représente la superficie de terrain qui est drainée pour alimenter un cours d'eau lors de pluies. Il faut délimiter les superficies dans le but de mettre en

relation les précipitations et les débits d'un cours d'eau qui dépendent de plusieurs caractéristiques [15].

Les caractéristiques physiques d'un bassin versant qui doivent être étudiées pour permettre le meilleur aménagement possible d'un cours d'eau pour fin de production hydroélectrique sont :

- 1) la superficie du bassin de drainage;
- 2) la régularisation naturelle effectuée par les lacs et les marécages;
- 3) la pente du bassin;
- 4) le type et la répartition de la végétation;
- 5) les caractéristiques hydrométriques telles que les précipitations, le taux d'infiltration et le taux d'évaporation.

Toutefois, l'élément qui affecte le plus la variabilité des crues est la superficie du bassin versant, dont l'estimation est primordiale.

Dans le cadre de notre étude, nous déterminons des crues de conception par la méthode statistique qui est, comme nous l'avons vu, la plus appropriée. Nous la présentons dans le cadre théorique ci-après.

### **3.8 Cadre théorique**

La détermination des crues de conception par la méthode statistique consiste à utiliser des fonctions de probabilité comme la loi normale ou la loi log normale. Ces fonctions font appel à certains paramètres comme : la moyenne, l'écart-type et le coefficient d'asymétrie. La loi normale, caractérisée par la moyenne et l'écart-type, est très utilisée en hydrologie pour analyser les valeurs moyennes comme les précipitations par exemple. Elle n'est cependant pas appropriée pour le calcul des débits extrêmes de crue

printanière dans les régions nordiques car ils dépendent du phénomène régulier et important qu'est la fonte des neiges. Donc, nous utiliserons la loi log normale qui est plus adaptée pour le cas d'événements extrêmes. La procédure d'utilisation de la loi log normale est identique à la loi normale à une transformation logarithmique près.

Pour mener à bien l'étude statistique, il faut suivre les étapes générales suivantes :

- 1) rassembler le maximum de données existantes sur le problème traité : critiquer, classer, mettre en ordre ces données. Par la suite, il faut réduire le tableau de données à un volume de chiffres beaucoup moins importants et plus maniable;
- 2) analyser les résultats afin de leur adapter le modèle probabiliste le plus adéquat. On condense donc la masse d'informations dont on dispose dans une équation mathématique qu'on appelle modèle mathématique;
- 3) utiliser le modèle mathématique obtenu précédemment pour faire du design, de la prévision ou de la reconstitution de données manquantes.

Le cheminement de la méthode statistique que nous allons suivre est celle de la loi log normale proposée par Bennis [4].

### **3.8.1 Méthodologie de la loi log normale**

Les étapes de calcul des débits de conception (quantiles de crues) consistent à :

- 1) extraire les débits de pointe de la crue de printemps pour chaque année relevée;
- 2) trier les débits de pointe en ordre décroissant et calculer leurs logarithme;
- 3) calculer la moyenne, l'écart-type et le coefficient d'asymétrie du logarithme des débits de pointe avec les équations (3.1), (3.2) et (3.3);

- 4) déterminer la probabilité de dépassement et la probabilité de non-dépassement, et la période de retour du logarithme des débits de pointe avec les équations (3.4), (3.5), et (3.6);
- 5) calculer les quantiles de crue pour les différentes périodes de retour avec les équations (3.7), (3.9) et (3.10).

En statistique, on travaille presque toujours à l'aide d'échantillons choisis de façon à représenter la population dans son ensemble. Comme la population est en pratique inconnue, on suppose a priori qu'elle peut être générée par une fonction de probabilité. Les paramètres de cette fonction (moyenne, écart-type, coefficient d'asymétrie) sont supposés égaux aux paramètres de l'échantillon, appelé en statistique la méthode des moments.

Les différentes équations utilisées pour calculer les quantiles de crue sont les suivantes :

Équation pour le calcul de la moyenne :

La moyenne est le premier moment centré par rapport à l'origine. C'est la valeur la plus utilisée en pratique pour caractériser l'ordre de grandeur des observations :

$$\bar{Q} = \frac{1}{N} \sum_{i=1}^N Q_i \quad (3.1)$$

Où :

- $\bar{Q}$  = moyenne des débits, en m<sup>3</sup>/s  
 N = nombre d'éléments dans la série  
 Q<sub>i</sub> = débit de la série pour l'échantillon considéré, en m<sup>3</sup>/s  
 i = échantillon considéré

Équation pour le calcul de l'écart-type :

La variance  $\sigma^2$  est le second moment centré par rapport à la moyenne. L'écart-type qui est la racine carré de la variance, indique la dispersion des observations autour de la moyenne.

$$\sigma = \left[ \frac{1}{N-1} \sum_{i=1}^N (Q_i - \bar{Q})^2 \right]^{1/2} \quad (3.2)$$

Où :

$\bar{Q}$	=	moyenne des débits, en m <sup>3</sup> /s
N	=	nombre d'éléments dans la série
Q <sub>i</sub>	=	débit de la série pour l'échantillon considéré, en m <sup>3</sup> /s
i	=	échantillon considéré
$\sigma$	=	écart-type

Équation pour le calcul du coefficient d'asymétrie :

Ce coefficient est le troisième moment centré autour de la moyenne. Comme son nom l'indique, ce coefficient mesure la symétrie de la distribution par rapport à la moyenne :

$$C_s = \frac{N \sum_{i=1}^N (Q_i - \bar{Q})^3}{(N-1)(N-2)\sigma^3} \quad (3.3)$$

Où :

$\bar{Q}$	=	moyenne des débits, en m <sup>3</sup> /s
-----------	---	--

N	=	nombre d'éléments dans la série
Q <sub>i</sub>	=	débit de la série pour l'échantillon considéré, en m <sup>3</sup> /s
i	=	échantillon considéré
σ	=	écart-type
C <sub>s</sub>	=	coefficient d'asymétrie

Équation de probabilité de dépassement et de non-dépassement :

C'est la valeur du débit qui peut causer une inondation ou un dépassement de la capacité d'un évacuateur :

$$p = P(Q \geq Q_T) = \frac{1}{T} \quad (3.4)$$

L'ingénieur doit considérer deux facteurs conflictuels dans la conception d'une structure hydraulique : le facteur économique et le facteur de risque. Lorsque les vies humaines sont en jeu, l'analyse économique n'est plus valable et le facteur social devient prépondérant. La probabilité  $p$  qui quantifie la vulnérabilité d'un système hydrique est appelée risque hydrologique. La fiabilité est la probabilité complémentaire de non-dépassement d'une valeur critique  $Q_T$ .

$$p = P(Q < Q_T) = 1 - \frac{1}{T} \quad (3.5)$$

Où :

T	=	période de retour, en année (équation 3.6)
Q <sub>T</sub>	=	débites pour la période de retour, en m <sup>3</sup> /s
Q	=	débit, en m <sup>3</sup> /s
$p$	=	probabilité de dépassement

Équation pour le calcul de la période de retour :

La période de retour est une moyenne à long terme du temps ou du nombre d'années séparant un événement de grandeur donnée d'un second événement d'une grandeur égale ou supérieure (glossaire international d'hydrologie, 1992).

$$T = \frac{1}{[1 - F(Z)]} \quad (3.6)$$

Où :

$T$  = période de retour, en année

$F(Z)$  = probabilité de non-dépassement (selon les tables de distribution)

Équation pour l'estimation des quantiles de crue :

Le quantile de crue est le débit extrême correspondant à un risque donné (c'est-à-dire à une probabilité au dépassement donnée). L'estimation des quantiles dépend directement de la loi de probabilité et de la méthode d'estimation des paramètres utilisées.

La valeur du débit de conception  $Q_T$  est déterminée à partir de la période moyenne tolérée,  $T$  en année, qui sépare deux défaillances. Lorsque l'ouvrage ne remplit pas la mission pour laquelle il a été conçu on parle alors de défaillance. La durée moyenne séparant deux défaillances de l'ouvrage s'appelle l'intervalle de récurrence.

Pour déterminer le débit  $Q_T$  correspondant à une période de retour  $T$ , on utilise la formulation générale suivante (Chow, 1951).

$$Q_T = \bar{Q} + K_T \sigma \quad (3.7)$$

Pour déterminer la période de retour  $T$  correspondante à un débit donné  $Q_T$  on utilise l'équation (3.7) sous la forme suivante :

$$K_T = \frac{Q_T - \bar{Q}}{\sigma} \quad (3.8)$$

Où :

- $\bar{Q}$  = moyenne des débits, en  $m^3/s$
- $Q_T$  = débit de période de retour  $T$ , en  $m^3/s$
- $\sigma$  = écart-type
- $K_T$  = paramètre de fréquence propre à chaque fonction de distribution de probabilité

Une fois les étapes de calcul avec les logarithmes des débits complétés, il est nécessaire d'effectuer la transformation logarithmique suivante pour déterminer les quantiles de crue.

$$\text{Log}(Q_x) = \bar{Q} + K \sigma \quad (3.9)$$

$$Q_R = 10^{\text{log}(Q_x)} \quad (3.10)$$

Où :

- $Q_R$  = débit, en  $m^3/s$
- $\bar{Q}$  = moyenne logarithmique des débits, en  $m^3/s$
- $K$  = paramètre de fréquence propre à chaque fonction de distribution de probabilité
- $\sigma$  = écart-type
- $Q_x$  = débit associé à une période de retour  $X$ , en  $m^3/s$

## CHAPITRE 4

### APPROCHE MÉTHODOLOGIQUE

#### 4.1 Introduction

Dans ce chapitre, nous étudions l'impact des erreurs dans les données mesurées sur l'estimation des quantiles de crue. Notre méthodologie de travail pour évaluer l'impact a consisté à calculer et à comparer des quantiles de crue.

On s'intéresse dans ce mémoire aux valeurs extrêmes du débit de pointe de crue printanière observé sur une période de plusieurs dizaines d'années. Pour étudier ces valeurs extrêmes, on applique des lois mathématiques ou statistiques. Parmi les lois les plus appropriées, la loi log normale est la plus utilisée en Amérique du nord [4]. Nous l'avons donc choisie pour effectuer nos calculs et déterminer les quantiles de crue.

C'est à l'aide de deux comparaisons de quantiles de crue que nous démontrerons l'impact des erreurs sur les données mesurées. La première analyse a comparé des quantiles déterminés avec des données historiques et des quantiles déterminés avec des données modifiées. Quant à la deuxième analyse, nous avons comparé des quantiles déterminés avec des données historiques et des quantiles déterminés avec des données modifiées et validées. Les différences obtenues, exprimées en pourcentage, représentent l'impact des erreurs sur l'estimation des quantiles.

Nos données historiques de travail sont des données hydrométriques fournies par Hydro-Québec. Ces données concernent les débits journaliers de la rivière Romaine mesurés à la station de jaugeage no 073801. Elle est située dans le canton de Duplessis au Québec, voir figure 10 (Hydro-Québec).

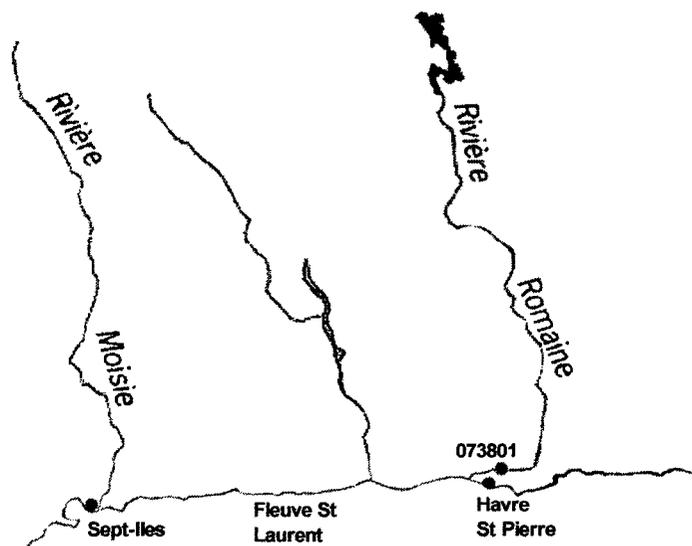


Figure 10 Localisation de la station 073801 sur la rivière Romaine

La série de données commence le 1er janvier 1956 et se termine le 30 septembre 2000. Pour uniformiser notre analyse, nous utilisons les années complètes de données commençant en 1957 et se terminant en 1999.

Les données historiques sont déjà validées par Hydro-Québec. Par conséquent, il n'est pas nécessaire d'effectuer une prévalidation ni une validation. Nous choisissons de représenter la courbe des débits de la rivière Romaine à la station 073801 pour l'année 1981 à la figure 11.

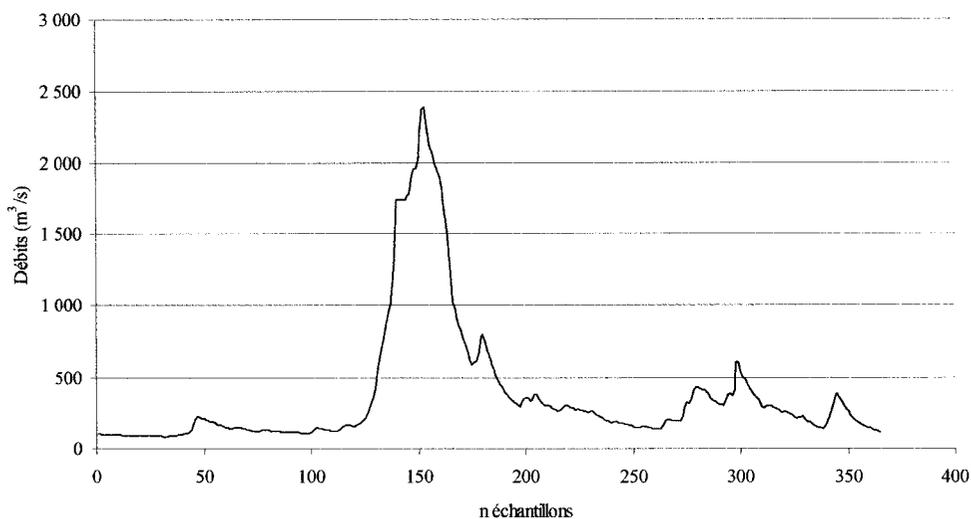


Figure 11 Courbe des débits de la rivière Romaine pour l'année 1981

Nous voyons sur cette figure que la crue de printemps a lieu autour du 150<sup>e</sup> jour de l'année 1981 soit le 2 juin 1981.

#### 4.2 Impacts des erreurs

Les impacts relatifs aux erreurs dans la mesure des niveaux ou des débits sont de natures diverses et se résument entre autres à :

- 1) surestimation du débit de pointe : effets sur les dimensions des évacuateurs de crue, des canaux de dérivations, sur la cote du barrage, etc.;
- 2) surestimation ou sous-estimation du niveau d'opération du réservoir : le niveau d'exploitation d'un réservoir est déterminé en fonction de multiples facteurs, dont entre autres les apports naturels au site. Les superficies inondées dépendent donc de la mesure des niveaux ou des débits dans les cours d'eau;

- 3) disparition plus importante que nécessaire de la faune et de la flore : due à un déboisement excessif de la zone inondée incluant le déplacement de la population humaine;
- 4) difficulté à maintenir le débit réservé : si les débits réels d'apport sont inférieurs à ceux estimés lors de la conception;
- 5) érosion des berges du réservoir plus grande : causée par une variation du niveau d'eau plus fréquente;
- 6) production hydroélectrique dépendante des apports naturels globaux : l'estimation des pointes de crue ont une influence directe sur la gestion des réservoirs;
- 7) problème d'alimentation en eau : pour l'agriculture, pour la consommation, pour les activités récréatives, etc., si le volume d'eau disponible est inférieur aux estimations;
- 8) augmentation des coûts de construction de l'ouvrage et des équipements : turbines, alternateurs, canal d'aménage, prise d'eau, etc.
- 9) augmentation des coûts environnementaux associés au programme de mise en valeur, aux mesures d'atténuation, au suivi environnemental pendant et après les travaux, et à la relocalisation humaine et animale;
- 10) augmentation des risques d'inondation;
- 11) augmentation des risques de bris de l'ouvrage (selon le choix de la crue de sécurité);
- 12) coûts reliés à la gestion des risques et des conséquences.

### **4.3 Étapes de calcul**

Pour déterminer l'impact des erreurs sur l'estimation des quantiles de crue nous utilisons les débits de la crue printanière. Celle-ci est souvent la plus importante. Les étapes chronologiques suivies sont :

- 1) calculs des quantiles des données historiques. Ces quantiles servent de référence et nous supposons ces débits calculés avec des données exempts d'erreurs;
- 2) modification d'un certain nombre de débits de pointe des crues printanières pour simuler des erreurs dans les mesures;
- 3) calculs des quantiles pour les débits de pointe modifiés. Nous supposons que ces quantiles sont calculés avec des données comportant des erreurs;
- 4) comparaison entre les quantiles des données historiques et ceux des données modifiées;
- 5) validation avec le logiciel PREVAL des données modifiées;
- 6) calculs des quantiles des données validées;
- 7) comparaison entre les quantiles des données historiques et ceux des données validées.

#### **4.4 Analyse pour les débits historiques**

Pour les quantiles de crue calculés avec les débits historiques, nous avons extrait 43 débits de pointe de crue printanière des données historiques. Ces débits de pointe représentent les débits de base utilisés tout au long de notre analyse. Les calculs des quantiles avec la loi log normale étant répétitifs, nous avons utilisé le logiciel Excel sous Windows pour les effectuer.

Les débits de pointe de la crue printanière extraits de la série de données historiques et les périodes de retour correspondantes sont présentés en ordre décroissant de débit au tableau III.

Tableau III

Débits de pointe des crues printanières des données historiques

Années	Débits Q (m <sup>3</sup> /s)	Log Q	Probabilités de non- dépassement (z)	Probabilités de dépassement (1-z)	Périodes de retour ( 1/(1-z) )
<b>02-juin-81</b>	<b>2 390</b>	3,378	0,935	0,065	15,34
<b>12-mai-99</b>	2 329	3,367	0,924	0,076	13,12
<b>06-juin-66</b>	2 300	3,362	0,918	0,082	12,19
<b>09-juin-69</b>	2 270	3,356	0,911	0,089	11,30
<b>06-juin-72</b>	2 250	3,352	0,907	0,093	10,74
<b>04-mai-83</b>	2 060	3,314	0,852	0,148	6,74
<b>01-juin-61</b>	2 050	3,312	0,848	0,152	6,58
<b>04-juin-82</b>	2 050	3,312	0,848	0,152	6,58
<b>24-mai-76</b>	2 010	3,303	0,833	0,167	5,98
<b>26-mai-77</b>	1 960	3,292	0,812	0,188	5,32
<b>09-juin-65</b>	1 950	3,290	0,808	0,192	5,20
<b>21-mai-78</b>	1 780	3,250	0,720	0,280	3,57
<b>26-mai-84</b>	1 760	3,246	0,707	0,293	3,42
<b>23-mai-80</b>	1 750	3,243	0,701	0,299	3,35
<b>28-mai-63</b>	1 740	3,241	0,695	0,305	3,28
<b>05-juin-57</b>	1 700	3,230	0,669	0,331	3,02
<b>17-mai-98</b>	1 675	3,224	0,652	0,348	2,87
<b>25-mai-73</b>	1 650	3,217	0,634	0,366	2,73
<b>24-mai-88</b>	1 650	3,217	0,634	0,366	2,73
<b>30-mai-59</b>	1 610	3,207	0,604	0,396	2,53
<b>06-juin-74</b>	1 600	3,204	0,597	0,403	2,48
<b>22-mai-94</b>	1 540	3,188	0,549	0,451	2,22
<b>22-mai-70</b>	1 500	3,176	0,516	0,484	2,07
<b>19-mai-60</b>	1 460	3,164	0,482	0,518	1,93
<b>15-mai-71</b>	1 450	3,161	0,474	0,526	1,90
<b>02-juin-62</b>	1 380	3,140	0,412	0,588	1,70
<b>10-juin-67</b>	1 360	3,134	0,394	0,606	1,65
<b>25-mai-95</b>	1 360	3,134	0,394	0,606	1,65

Tableau III (suite)

<b>24-mai-97</b>	1 339	3,127	0,375	0,625	1,60
<b>04-juin-75</b>	1 300	3,114	0,340	0,660	1,52
<b>21-mai-89</b>	1 240	3,093	0,288	0,712	1,40
<b>31-mai-64</b>	1 230	3,090	0,279	0,721	1,39
<b>25-mai-96</b>	1 210	3,083	0,262	0,738	1,35
<b>23-mai-93</b>	1 160	3,064	0,220	0,780	1,28
<b>07-mai-79</b>	1 150	3,061	0,212	0,788	1,27
<b>05-juin-68</b>	1 100	3,041	0,174	0,826	1,21
<b>03-juin-85</b>	1 100	3,041	0,174	0,826	1,21
<b>27-mai-91</b>	1 100	3,041	0,174	0,826	1,21
<b>31-mai-58</b>	1 030	3,013	0,126	0,874	1,14
<b>16-mai-92</b>	872	2,941	0,047	0,953	1,05
<b>17-mai-86</b>	855	2,932	0,041	0,959	1,04
<b>13-mai-90</b>	778	2,891	0,021	0,979	1,02

À l'aide des équations de la loi log normale, nous obtenons les statistiques suivantes pour les débits de pointe historiques :

Moyenne des débits (Q)	=	1 550	m <sup>3</sup> /s
Écart-type ( $\sigma$ )	=	451,96	
Coefficient d'asymétrie	=	0,0619	
Moyenne des log Q (Q <sub>m</sub> )	=	3,170	m <sup>3</sup> /s
Écart-type des log Q ( $\sigma$ )	=	0,1374	
Coefficient d'asymétrie des log Q	=	-0,6495	

Le coefficient d'asymétrie étant négatif, la distribution est asymétrique et présenterait une courbe plus allongée vers la gauche. Les quantiles de crue correspondant aux différentes périodes de retour sont donnés au tableau IV.

Tableau IV

Quantiles de crue des débits historiques

<b>Périodes de retour</b>	<b>Probabilités de non- dépassement</b>	<b>k</b>	<b>Log (Q<sub>x</sub>) (Q<sub>m</sub>+k σ)</b>	<b>Quantiles de crue (m<sup>3</sup>/s)</b>
<b>10</b>	0,9000	1,282	3,3466	2 221
<b>20</b>	0,9500	1,645	3,3966	2 492
<b>50</b>	0,9800	2,054	3,4528	2 836
<b>100</b>	0,9900	2,326	3,4902	3 092
<b>200</b>	0,9950	2,576	3,5245	3 346
<b>500</b>	0,9980	2,878	3,5661	3 682
<b>1 000</b>	0,9990	3,090	3,5952	3 938
<b>10 000</b>	0,9999	3,719	3,6817	4 805

Nous constatons un écart important des quantiles entre la période de retour 1 / 10 ans et celle de 1 / 10 000 ans. Par ailleurs, nous savons que le choix *a priori* d'une distribution de données s'accompagne systématiquement d'une incertitude sur l'estimation des quantiles qui pourrait se valider à l'aide de tests statistiques pour juger de l'adéquation de la fonction de distribution de données utilisée [4].

#### 4.5 Analyse pour les débits modifiés

Dans le but d'évaluer les impacts des erreurs sur les données mesurées sur les sites, les débits de pointe de la crue printanière pour les données historiques ont été modifiés pour simuler ces erreurs.

#### 4.5.1 Modification des débits

Pour simuler les erreurs, nous modifions certaines valeurs maximales des débits de pointe de la crue printanière soit les débits des années 1966, 1969, 1981 et 1999. Nous procédons alors à l'aide des 3 scénarios suivants :

- 1) Scénario 1 : le débit de 1981 est modifié de 50 %, 20 %, 5 %, -5 %, -20 % et -50 %;
- 2) Scénario 2 : les débits de 1981 et de 1999 sont modifiés de 50 %, 20 %, 5 %, -5 %, -20 % et -50 %;
- 3) Scénario 3 : les débits de 1981, 1999, 1966 et de 1969 sont modifiés de 50 %, 20 %, 5 %, -5 %, -20 % et -50 %.

Ces scénarios permettent entre autres de :

- 1) connaître les débits erronés;
- 2) connaître l'importance de l'erreur sur la mesure du débit;
- 3) obtenir une variation maximale des quantiles de crue;
- 4) considérer les cas amplifiés de 50 % et diminués de -50 % comme des erreurs aberrantes.

Le tableau V présente les différentes valeurs des débits de pointe modifiés pour les 3 scénarios.

Tableau V

Débits (m<sup>3</sup>/s) historiques modifiés

Années	Débits historiques (m <sup>3</sup> /s)	Pourcentages de modification des débits					
		50 %	20 %	5 %	-5 %	-20 %	-50 %
<b>1981</b>	<b>2 390</b>	3 585	2 868	2 510	2 271	1 912	1 195
<b>1999</b>	<b>2 329</b>	3 494	2 795	2 445	2 213	1 863	1 165
<b>1966</b>	<b>2 300</b>	3 450	2 760	2 415	2 185	1 840	1 150
<b>1969</b>	<b>2 270</b>	3 405	2 724	2 384	2 157	1 816	1 135

Les débits de pointe historiques pour les autres années de la série ne sont pas modifiés.

Nous présentons au tableau VI les statistiques obtenues pour les 3 scénarios.

Tableau VI

Statistiques pour les scénarios 1, 2 et 3

Scénarios	% d'erreur	Débits moyens Q (m <sup>3</sup> /s)	Log Q	Écart-type (Log Q)	Coefficients d'asymétrie (Log Q)
<b>1</b>	<b>50 %</b>	1 578	3,1746	0,1461	-0,2474
	<b>20 %</b>	1 561	3,1723	0,1408	-0,5142
	<b>5 %</b>	1 553	3,1710	0,1382	-0,6208
	<b>-5 %</b>	1 547	3,1700	0,1367	-0,6738
	<b>-20 %</b>	1 539	3,1682	0,1347	-0,7148
	<b>-50 %</b>	1 522	3,1635	0,1342	-0,6341

Tableau VI (suite)

2	50 %	1 605	3,1787	0,1539	-0,0013
	20 %	1 561	3,1723	0,1406	-0,5297
	5 %	1 556	3,1715	0,1390	-0,5966
	-5 %	1 545	3,1695	0,1360	-0,6957
	-20 %	1 528	3,1660	0,1321	-0,7783
	-50 %	1 495	3,1565	0,1312	-0,6152
3	50 %	1 658	3,1869	0,1677	-0,2623
	20 %	1 593	3,1779	0,1496	-0,2634
	5 %	1 561	3,1725	0,1404	-0,5557
	-5 %	1 539	3,1684	0,1346	-0,7370
	-20 %	1 507	3,1615	0,1269	-0,9207
	-50 %	1 442	3,1425	0,1246	-0,6022
<b>Données historiques</b>		1 550	3,1700	0,1374	-0,6495

Les moyennes, les écart-type et les coefficients d'asymétrie varient dans le même sens que la modification des débits. Toutefois, pour la modification du débit de -50 % le coefficient d'asymétrie augmente. Ceci s'explique par le débit modifié qui devient inférieur au débit moyen, ce qui change la distribution des données.

#### 4.5.2 Calcul des quantiles

Après avoir calculé les statistiques pour les débits modifiés, nous présentons au tableau VII les résultats obtenus pour les 3 scénarios.

Tableau VII

Quantiles de crue (m<sup>3</sup>/s) pour les scénarios 1, 2 et 3

Scénarios	Périodes de retour	Pourcentages d'erreur dans les débits					
		50 %	20 %	5 %	-5 %	-20 %	-50 %
<b>1</b>	<b>10</b>	2 301	2 253	2 229	2 214	2 192	2 165
	<b>20</b>	2 600	2 535	2 503	2 482	2 454	2 423
	<b>50</b>	2 983	2 894	2 850	2 823	2 786	2 749
	<b>100</b>	3 270	3 161	3 109	3 076	3 031	2 991
	<b>200</b>	3 556	3 428	3 366	3 327	3 275	3 230
	<b>500</b>	3 937	3 780	3 706	3 659	3 597	3 547
	<b>1 000</b>	4 229	4 050	3 965	3 912	3 842	3 787
	<b>10 000</b>	5 226	4 966	4 844	4 768	4 670	4 600
<b>2</b>	<b>10</b>	2 377	2 251	2 237	2 206	2 164	2 112
	<b>20</b>	2 703	2 532	2 513	2 472	2 417	2 356
	<b>50</b>	3 125	2 890	2 864	2 810	2 737	2 666
	<b>100</b>	3 442	3 157	3 125	3 060	2 974	2 895
	<b>200</b>	3 760	3 422	3 385	3 309	3 208	3 122
	<b>500</b>	4 185	3 774	3 729	3 637	3 517	3 420
	<b>1 000</b>	4 512	4 042	3 990	3 887	3 752	3 646
	<b>10 000</b>	5 639	4 955	4 881	4 733	4 543	4 410
<b>3</b>	<b>10</b>	2 522	2 342	2 251	2 192	2 109	2 005
	<b>20</b>	2 902	2 654	2 532	2 453	2 345	2 225
	<b>50</b>	3 398	3 056	2 890	2 785	2 643	2 502
	<b>100</b>	3 776	3 356	3 156	3 030	2 862	2 706
	<b>200</b>	4 157	3 658	3 421	3 274	3 078	2 907
	<b>500</b>	4 672	4 059	3 772	3 595	3 362	3 170
	<b>1 000</b>	5 071	4 367	4 040	3 839	3 577	3 369
	<b>10 000</b>	6 466	5 424	4 952	4 666	4 299	4 035

L'équation (3.7) montre que le quantile de crue est fonction de la moyenne et de l'écart-type. Comme mentionné précédemment la moyenne et l'écart-type varient dans le sens de la modification des débits. Par conséquent, les quantiles de crue varient dans le même sens de la modification des débits.

#### 4.5.3 Calcul de l'impact pour les débits modifiés

Pour déterminer l'impact des erreurs sur les quantiles de crue nous comparons les quantiles calculés avec les débits historiques et ceux calculés avec les débits modifiés. Les écarts obtenus représentent l'impact des erreurs sur les débits mesurés. Les résultats pour chaque scénario sont présentés au tableau VIII.

Tableau VIII

Pourcentages de variations des quantiles pour les 3 scénarios

Scénarios	Périodes de retour	Pourcentages d'erreur dans les débits					
		50 %	20 %	5 %	-5 %	-20 %	-50 %
1	10	3,6 %	1,4 %	0,3 %	-0,3 %	-1,3 %	-2,5 %
	20	4,3 %	1,7 %	0,4 %	-0,4 %	-1,5 %	-2,8 %
	50	5,2 %	2,0 %	0,5 %	-0,5 %	-1,8 %	-3,1 %
	100	5,8 %	2,2 %	0,5 %	-0,5 %	-2,0 %	-3,3 %
	200	6,3 %	2,4 %	0,6 %	-0,6 %	-2,1 %	-3,5 %
	500	6,9 %	2,7 %	0,6 %	-0,6 %	-2,3 %	-3,7 %
	1 000	7,4 %	2,8 %	0,7 %	-0,7 %	-2,4 %	-3,8 %
	10 000	8,7 %	3,3 %	0,8 %	-0,8 %	-2,8 %	-4,3 %

Tableau VIII (suite)

<b>2</b>	<b>10</b>	7,0 %	1,3 %	0,7 %	-0,7 %	-2,6 %	-4,9 %
	<b>20</b>	8,5 %	1,6 %	0,8 %	-0,8 %	-3,0 %	-5,4 %
	<b>50</b>	10,2 %	1,9 %	1,0 %	-0,9 %	-3,5 %	-6,0 %
	<b>100</b>	11,3 %	2,1 %	1,1 %	-1,0 %	-3,8 %	-6,4 %
	<b>200</b>	12,4 %	2,3 %	1,2 %	-1,1 %	-4,1 %	-6,7 %
	<b>500</b>	13,7 %	2,5 %	1,3 %	-1,2 %	-4,5 %	-7,1 %
	<b>1 000</b>	14,6 %	2,7 %	1,3 %	-1,3 %	-4,7 %	-7,4 %
	<b>10 000</b>	17,3 %	3,1 %	1,6 %	-1,5 %	-5,5 %	-8,2 %
<b>3</b>	<b>10</b>	13,5 %	5,4 %	1,3 %	-1,3 %	-5,1 %	-9,7 %
	<b>20</b>	16,4 %	6,5 %	1,6 %	-1,6 %	-5,9 %	-10,7 %
	<b>50</b>	19,8 %	7,7 %	1,9 %	-1,8 %	-6,8 %	-11,8 %
	<b>100</b>	22,1 %	8,6 %	2,1 %	-2,0 %	-7,4 %	-12,5 %
	<b>200</b>	24,3 %	9,3 %	2,2 %	-2,2 %	-8,0 %	-13,1 %
	<b>500</b>	26,9 %	10,2 %	2,5 %	-2,4 %	-8,7 %	-13,9 %
	<b>1 000</b>	28,8 %	10,9 %	2,6 %	-2,5 %	-9,2 %	-14,5 %
	<b>10 000</b>	34,6 %	12,9 %	3,0 %	-2,9 %	-10,5 %	-16,0 %

À titre d'exemple de résultats, nous constatons :

- 1) pour le scénario 1, que la variation des quantiles de crue est faible quelle que soit l'erreur du débit et de la période de retour. Ce pourcentage se situe entre +8,7 % et -4,3 %.
- 2) pour le scénario 2, que la variation des quantiles de crue se situe entre +17,3 % et -8,2 %.
- 3) pour le scénario 3, la variation des quantiles de crue peut devenir importante, elle se situe entre +34,6 % et -16,0 %.

De l'analyse précédente nous en tirons les conclusions suivantes :

- 1) l'impact des erreurs est minime pour les faibles modifications de débit;
- 2) l'impact des erreurs est important pour les modifications aberrantes;
- 3) le pourcentage de variation des erreurs n'est pas symétrique par rapport aux données historiques;
- 4) le pourcentage de variation n'est pas proportionnel au pourcentage d'erreur;
- 5) l'importance de l'erreur est fonction du nombre de données erronées et du pourcentage d'erreurs sur les données.

#### **4.6 Analyse pour les débits validés**

À cette étape nous évaluons l'impact des erreurs en considérant des données validées. Dans un premier temps nous modifions les données suivant les 3 scénarios. Ensuite, nous insérons les débits modifiés dans la série historique puis nous validons les données. Et pour finir, nous calculons à nouveau les quantiles pour les données validées.

##### **4.6.1 Validation des données**

Le type et le nombre de données à notre disposition nous orientent dans le choix d'une méthode de validation univariée. Nous utiliserons deux filtres, soit le filtre médian et le filtre winsoriseur. Ces filtres qui sont bien adaptés pour l'élimination des valeurs aberrantes sont implantés dans le logiciel PREVAL. Après avoir configuré les données dans le logiciel nous avons validé les données.

#### 4.6.2 Application du filtre médian

Nous avons vu au chapitre 2 que le filtre médian est un filtre à fenêtre mobile. Il est bien adapté pour l'élimination des valeurs aberrantes. Dans notre cas cette méthode éliminera ces valeurs que nous introduirons volontairement dans les données.

Le paramètre principal de validation requis par PREVAL concerne le nombre de données dans la fenêtre, nous utilisons 3 comme valeur. Celle-ci permet de respecter la tendance de variation des données.

##### 4.6.2.1 Débits validés d'un filtre médian

Les débits de pointe de la crue printanière validés pour les différents scénarios, sont présentés au tableau IX.

Tableau IX

Débits (m<sup>3</sup>/s) modifiés et validés – Filtre médian

Années	Débits historiques		Pourcentages de modification des débits					
			50 %	20 %	5 %	-5 %	-20 %	-50 %
1981	2 390	modifié	3 585	2 868	2 510	2 271	1 912	1 195
		validé	2 370	2 370	2 370	2 271	2 270	2 270
1999	2 329	modifié	3 494	2 975	2 445	2 213	1 863	1 165
		validé	2 326	2 326	2 326	2 213	2 078	2 078
1966	2 300	modifié	3 450	2 760	2 415	2 185	1 840	1 150
		validé	2 210	2 210	2 210	2 185	2 100	2 100
1969	2 270	modifié	3 405	2 724	2 384	2 157	1 840	1 150
		validé	2 210	2 210	2 210	2 157	2 150	2 150

La validation avec le filtre médian donne pour une même période une uniformisation des débits on retrouve par exemple pour l'année 1981 :

- 1) la valeur des débits validés se situés entre 2 270 m<sup>3</sup>/s et 2 370 m<sup>3</sup>/s comparativement à des débits situés entre 1 95 m<sup>3</sup>/s et 3 585 m<sup>3</sup>/s sans validation;
- 2) un débit modifié de plus de 5 % donnera un débit validé de 2 370 m<sup>3</sup>/s;
- 3) un débit modifié de moins de -5 % donnera un débit validé de 2 270 m<sup>3</sup>/s;
- 4) les données aberrantes sont ramenées à des valeurs plus près de la tendance de variation;
- 5) nous observons le même constat pour les autres années soit 1966, 1969 et 1999.

Les figures 12, 13 et 14 présentent respectivement la validation pour une variation des débits de 5 %, 20 % et de 50 % pour 1981.

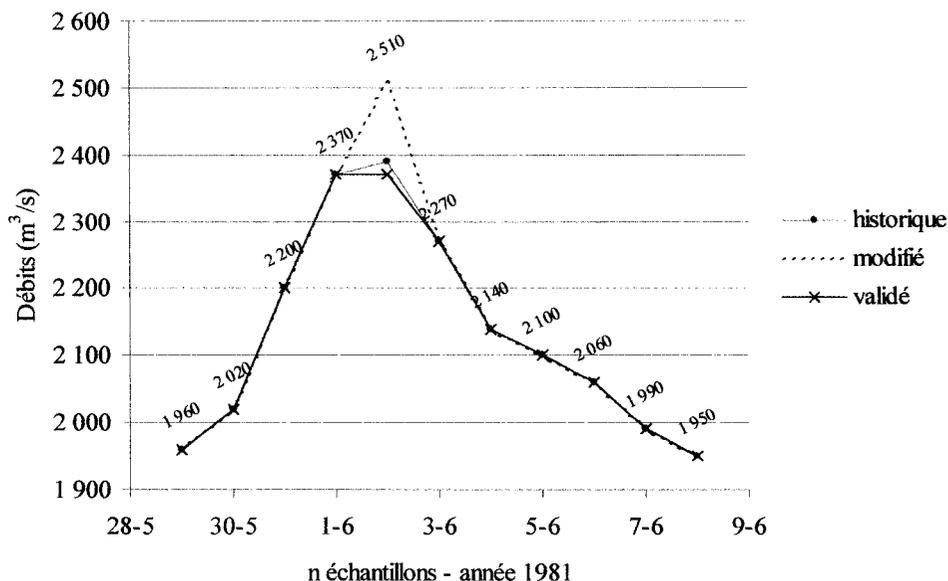


Figure 12 Validation avec 5 % de variation – Filtre médian – 1981

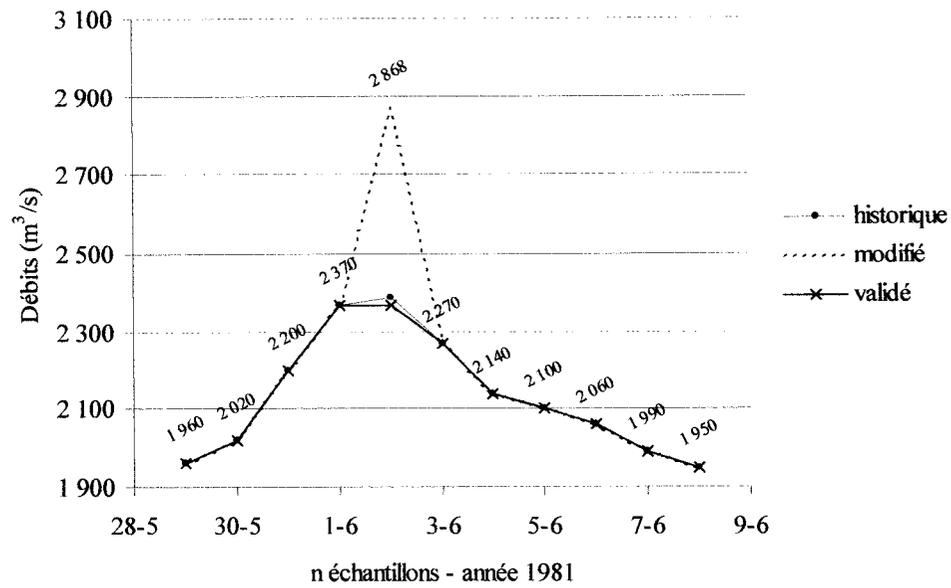


Figure 13 Validation avec 20 % de variation – Filtre médian – 1981

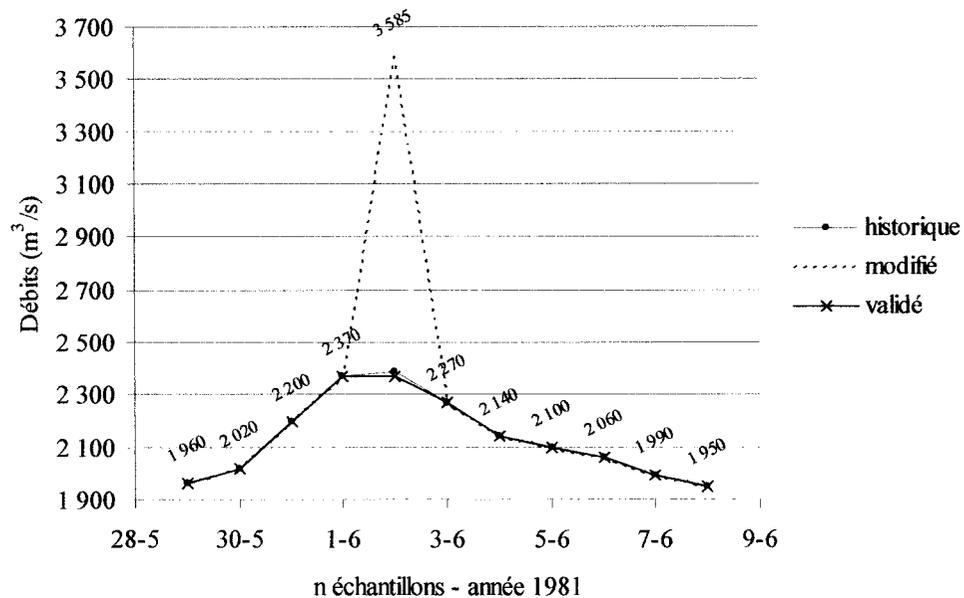


Figure 14 Validation avec 50 % de variation – Filtre médian – 1981

Nous observons sur ces figures que le débit validé est le même soit  $2\,370\text{ m}^3/\text{s}$ . La validation pour les autres pourcentages de variations pour l'année 1981 est présentée aux figures 15, 16 et 17.

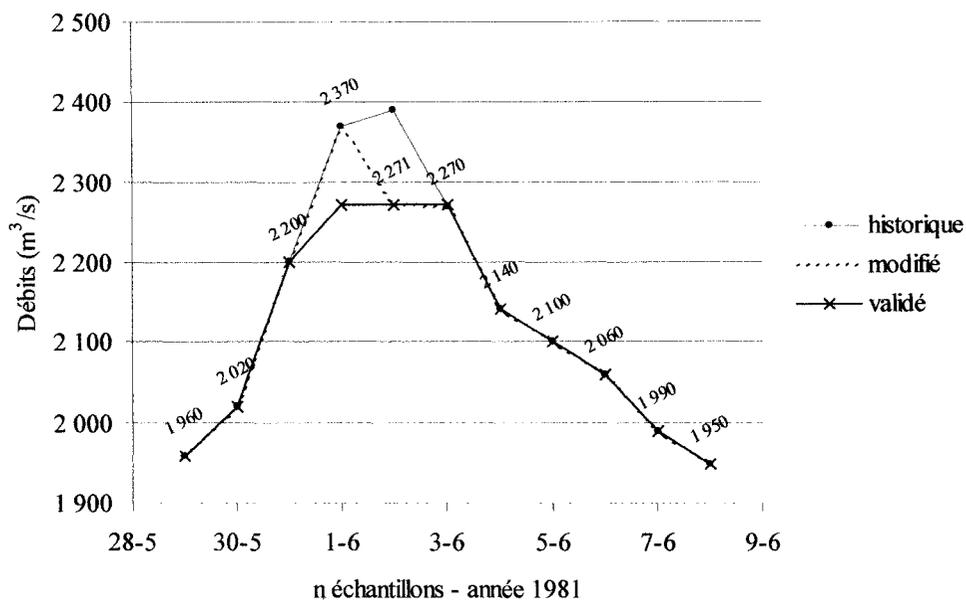


Figure 15 Validation avec -5 % de variation – Filtre médian – 1981

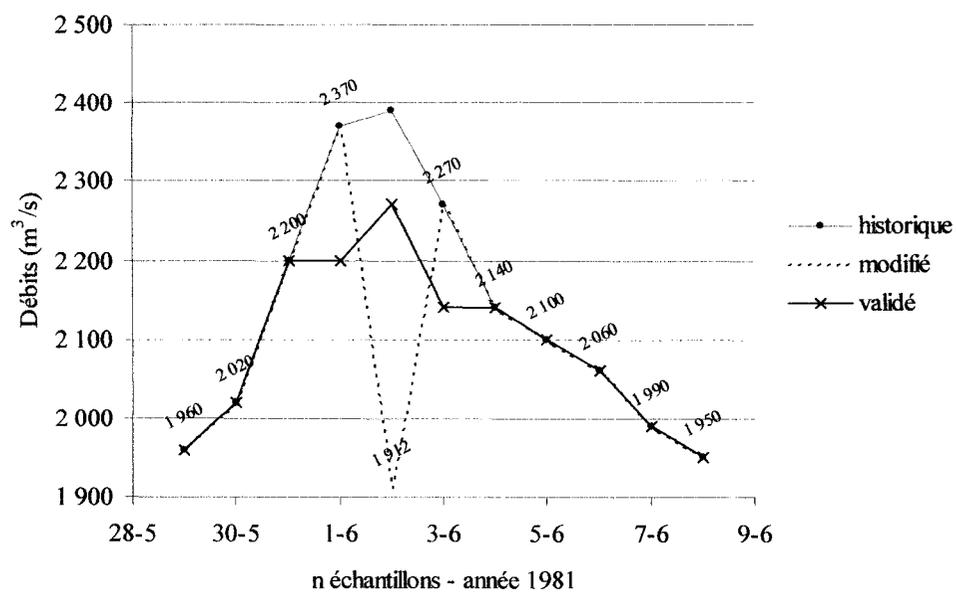


Figure 16 Validation avec -20 % de variation – Filtre médian – 1981

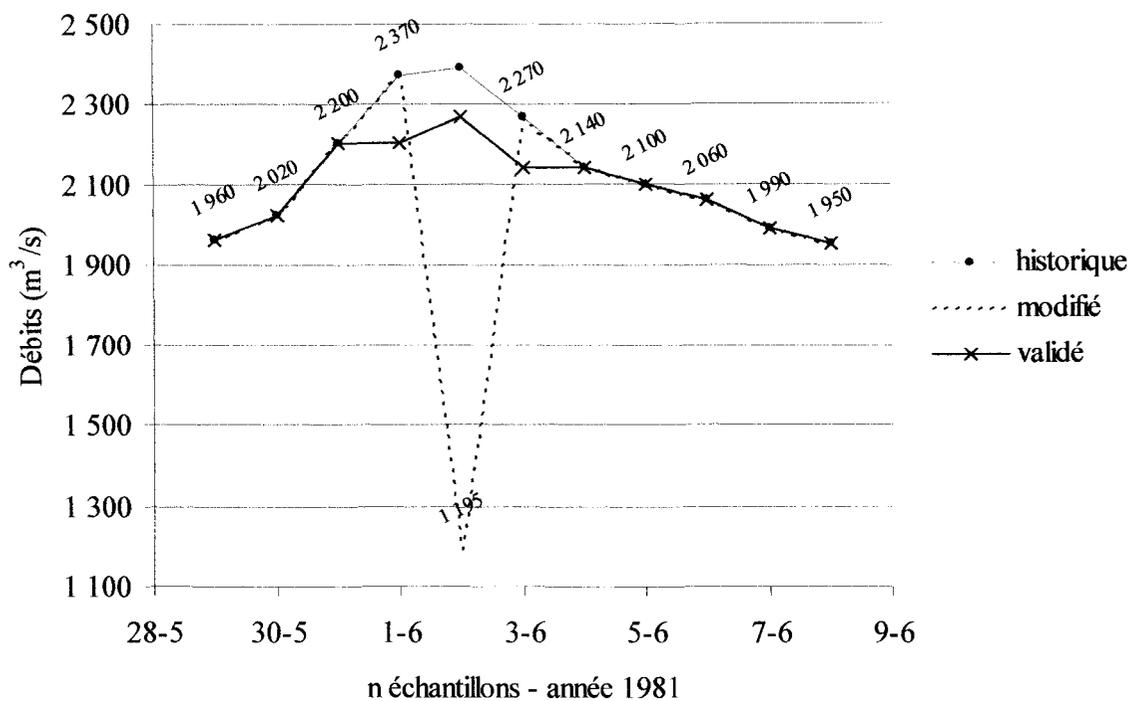


Figure 17 Validation avec -50 % de variation – Filtre médian – 1981

Nous constatons sur ces figures la même uniformisation de débit validé soit  $2\,270\text{ m}^3/\text{s}$ .

Après la validation, nous avons déterminé les quantiles de crue pour les débits validés avec le filtre médian. Les statistiques pour les 3 scénarios sont présentées au tableau X.

Tableau X

Statistiques pour les scénarios 1, 2 et 3 – Filtre médian

Scénarios	% d'erreur	Débits moyens Q (m <sup>3</sup> /s)	Log Q	Écart-type (Log Q)	Coefficients d'asymétrie (Log Q)
1	50 %	1 423	3,1339	0,1347	-0,6050
	20 %	1 423	3,1339	0,1347	-0,6050
	5 %	1 423	3,1339	0,1347	-0,6050
	-5 %	1 421	3,1335	0,1340	-0,6362
	-20 %	1 420	3,1335	0,1340	-0,6365
	-50 %	1 420	3,1335	0,1340	-0,6365
2	50 %	1 434	3,1362	0,1379	-0,5201
	20 %	1 434	3,1362	0,1379	-0,5201
	5 %	1 434	3,1362	0,1379	-0,5201
	-5 %	1 429	3,1352	0,1363	-0,5764
	-20 %	1 426	3,1346	0,1354	-0,6058
	-50 %	1 426	3,1346	0,1354	-0,6058
3	50 %	1 440	3,1376	0,1398	-0,4695
	20 %	1 440	3,1376	0,1398	-0,4695
	5 %	1 440	3,1376	0,1398	-0,4695
	-5 %	1 434	3,1363	0,1377	-0,5369
	-20 %	1 428	3,1352	0,1362	-0,5830
	-50 %	1 428	3,1352	0,1362	-0,5830
<b>Données historiques</b>		1 550	3,1700	0,1374	-0,6495

Les statistiques présentent une uniformisation des résultats. Ce qui est normal puisque l'on retrouve les mêmes débits validés selon les différentes modifications de débit.

#### 4.6.2.2 Calcul des quantiles pour les débits validés avec le filtre médian

Les résultats du calcul des quantiles de crue effectués avec les débits validés sont présentés au tableau XI.

Tableau XI

Quantiles de crue (m<sup>3</sup>/s) pour les 3 scénarios – Filtre médian

Scénarios	Périodes de retour	Pourcentages d'erreur dans les débits					
		50 %	20 %	5 %	-5 %	-20 %	-50 %
1	10	2 026	2 026	2 026	2 019	2 019	2 019
	20	2 268	2 268	2 268	2 259	2 259	2 259
	50	2 574	2 574	2 574	2 562	2 562	2 562
	100	2 801	2 801	2 801	2 787	2 787	2 787
	200	3 027	3 027	3 027	3 010	3 010	3 010
	500	3 325	3 325	3 325	3 305	3 304	3 304
	1 000	3 551	3 551	3 551	3 528	3 528	3 528
	10 000	4 316	4 316	4 316	4 284	4 284	4 284
2	10	2 055	2 055	2 055	2 041	2 033	2 033
	20	2 306	2 306	2 306	2 288	2 276	2 276
	50	2 626	2 626	2 626	2 601	2 586	2 586
	100	2 864	2 864	2 864	2 834	2 815	2 815
	200	3 100	3 100	3 100	3 064	3 043	3 043
	500	3 412	3 412	3 412	3 369	3 343	3 343
	1 000	3 649	3 649	3 649	3 601	3 572	3 572
	10 000	4 456	4 456	4 456	4 387	4 346	4 346
3	10	2 074	2 074	2 074	2 055	2 040	2 040
	20	2 331	2 331	2 331	2 306	2 286	2 286
	50	2 659	2 659	2 659	2 625	2 599	2 599
	100	2 903	2 903	2 903	2 862	2 831	2 831
	200	3 145	3 145	3 145	3 098	3 061	3 061
	500	3 467	3 467	3 467	3 409	3 366	3 366
	1 000	3 712	3 712	3 712	3 647	3 597	3 597
	10 000	4 545	4 545	4 545	4 452	4 382	4 382

Comme pour les statistiques les quantiles sont uniformisés. Par conséquent, on peut dire que la validation uniformise les quantiles. Par exemple, pour une période de retour de 1 / 10 ans quel que soit le scénario, nous obtenons un quantile situé entre 2 019 m<sup>3</sup>/s et 2 074 m<sup>3</sup>/s. Nous observons le même constat pour les autres périodes de retour.

#### 4.6.2.3 Calcul de l'impact pour les débits validés avec le filtre médian

Les écarts obtenus par la comparaison des quantiles des débits validés avec les quantiles des débits historiques pour les 3 scénarios sont présentés au tableau XII.

Tableau XII

Variations des quantiles – scénarios 1, 2 et 3 – Filtre médian

Scénarios	Périodes de retour	Pourcentages d'erreur dans les débits					
		50 %	20 %	5 %	-5 %	-20 %	-50 %
1	10	-8,8 %	-8,8 %	-8,8 %	-9,1 %	-9,1 %	-9,1 %
	20	-9,0 %	-9,0 %	-9,0 %	-9,4 %	-9,4 %	-9,4 %
	50	-9,2 %	-9,2 %	-9,2 %	-9,7 %	-9,7 %	-9,7 %
	100	-9,4 %	-9,4 %	-9,4 %	-9,9 %	-9,9 %	-9,9 %
	200	-9,5 %	-9,5 %	-9,5 %	-10,0 %	-10,0 %	-10,0 %
	500	-9,7 %	-9,7 %	-9,7 %	-10,2 %	-10,3 %	-10,3 %
	1 000	-9,8 %	-9,8 %	-9,8 %	-10,4 %	-10,4 %	-10,4 %
	10 000	-10,2 %	-10,2 %	-10,2 %	-10,8 %	-10,9 %	-10,9 %
2	10	-7,5 %	-7,5 %	-7,5 %	-8,1 %	-8,5 %	-8,5 %
	20	-7,4 %	-7,4 %	-7,4 %	-8,2 %	-8,7 %	-8,7 %
	50	-7,4 %	-7,4 %	-7,4 %	-8,3 %	-8,8 %	-8,8 %
	100	-7,4 %	-7,4 %	-7,4 %	-8,4 %	-9,0 %	-9,0 %
	200	-7,4 %	-7,4 %	-7,4 %	-8,4 %	-9,1 %	-9,1 %
	500	-7,3 %	-7,3 %	-7,3 %	-8,5 %	-9,2 %	-9,2 %
	1 000	-7,3 %	-7,3 %	-7,3 %	-8,5 %	-9,3 %	-9,3 %
	10 000	-7,3 %	-7,3 %	-7,3 %	-8,7 %	-9,6 %	-9,6 %

Tableau XII (suite)

<b>3</b>	<b>10</b>	-6,7 %	-6,7 %	-6,7 %	-7,5 %	-8,2 %	-8,2 %
	<b>20</b>	-6,5 %	-6,5 %	-6,5 %	-7,5 %	-8,3 %	-8,3 %
	<b>50</b>	-6,3 %	-6,3 %	-6,3 %	-7,5 %	-8,4 %	-8,4 %
	<b>100</b>	-6,1 %	-6,1 %	-6,1 %	-7,4 %	-8,4 %	-8,4 %
	<b>200</b>	-6,0 %	-6,0 %	-6,0 %	-7,4 %	-8,5 %	-8,5 %
	<b>500</b>	-5,8 %	-5,8 %	-5,8 %	-7,4 %	-8,6 %	-8,6 %
	<b>1 000</b>	-5,7 %	-5,7 %	-5,7 %	-7,4 %	-8,6 %	-8,6 %
	<b>10 000</b>	-5,4 %	-5,4 %	-5,4 %	-7,4 %	-8,8 %	-8,8 %

De l'analyse précédente nous en tirons les conclusions suivantes :

- 1) l'impact des erreurs varie peu quel que soit la modification de débit;
- 2) l'impact des erreurs varie peu quel que soit la période de retour;
- 3) le pourcentage de variation n'est pas proportionnel au pourcentage d'erreur;
- 4) l'impact des erreurs est toujours négatif.

#### 4.6.3 Application d'un filtre winsorizeur

Le filtre winsorizeur est un filtre qui exclut les observations les plus éloignées de la médiane avant de calculer la moyenne. Les valeurs exclues sont ensuite reconstituées à partir des valeurs les plus proches conservées. Ce type de filtre donne un filtrage assez régulier en adoucissant les pentes abruptes et en enlevant les irrégularités. Les paramètres principaux de validation requis par PREVAL concernent le nombre de données dans la fenêtre, nous utilisons 5 et comme nombre de données à éliminer, nous utilisons 2. Ces valeurs sont celles qui respectent le mieux la tendance de variation des données.

#### 4.6.3.1 Débits validés d'un filtre winsoriseur

Après validation des débits de pointe de la crue printanière modifiés pour les différents scénarios, nous obtenons des nouveaux débits de pointe validés. Ces débits sont présentés au tableau XIII.

Tableau XIII

Débits (m<sup>3</sup>/s) modifiés et validés – Filtre winsoriseur

Années	Débits historiques (m <sup>3</sup> /s)		Pourcentages d'erreur dans les débits					
			50 %	20 %	5 %	-5 %	-20 %	-50 %
1981	2 390	modifié	3 585	2 868	2 510	2 271	1 912	1 195
		validé	2 280	2 280	2 280	2 247	2 203	2 203
1999	2 329	modifié	3 494	2 975	2 445	2 213	1 863	1 165
		validé	2 087	2 087	2 087	2 070	1 953	1 922
1966	2 300	modifié	3 450	2 760	2 415	2 185	1 840	1 150
		validé	2 113	2 113	2 113	2 113	2 020	2 017
1969	2 270	modifié	3 405	2 724	2 384	2 157	1 840	1 150
		validé	2 150	2 150	2 150	2 150	2 097	2 070

Nous constatons que la validation avec le filtre winsoriseur présente une uniformisation similaire à celle obtenue avec le filtre médian. De plus, comme pour le filtre médian, les débits obtenus avec le filtre winsoriseur atteignent des valeurs maximales quelle que soit la modification du débit. Le filtre analyse bien aussi les données aberrantes puisque celles-ci sont ramenées à des valeurs plus près de la tendance de variation des données. Aux figures 18 et 19 nous présentons un exemple de la validation pour une variation des débits de 5 % et de 20 % pour 1981.

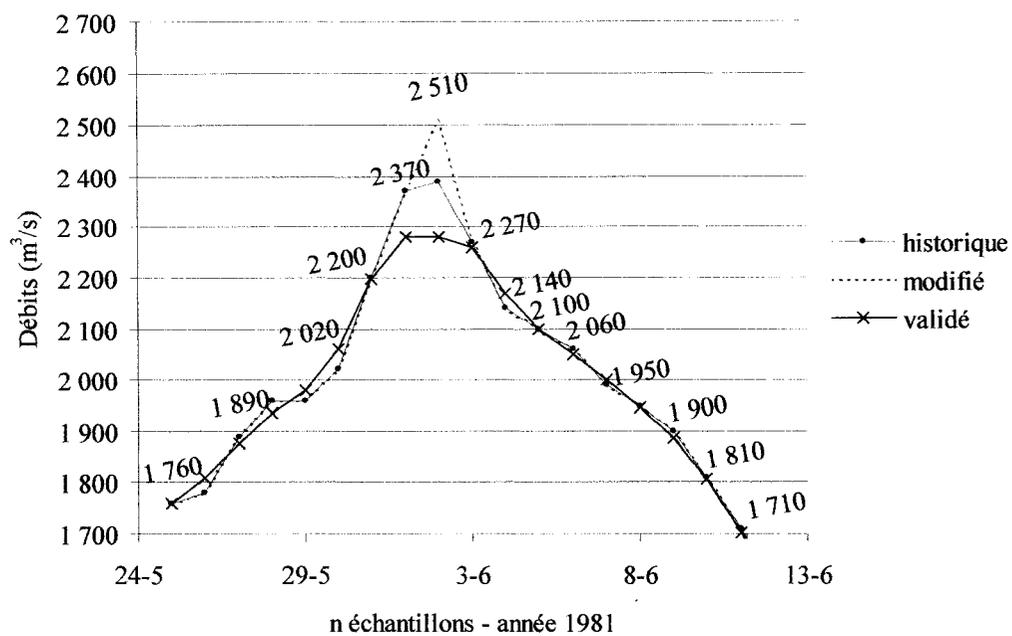


Figure 18 Validation avec 5 % de variation – Filtre winsoriseur – 1981

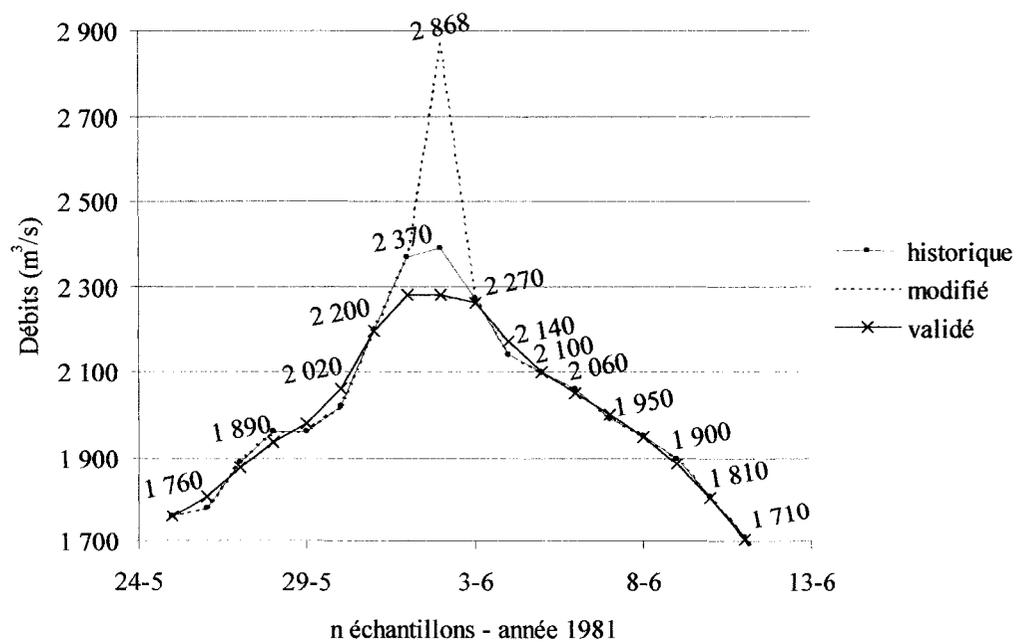


Figure 19 Validation avec 20 % de variation – Filtre winsoriseur – 1981

Nous observons sur ces figures que le débit validé est le même soit 2 280 m<sup>3</sup>/s. Nous constatons une uniformisation des débits validés. La validation pour les autres périodes soit 1966, 1969 et 1999 nous a donné aussi une uniformisation des débits.

Après la validation, nous avons déterminé les quantiles de crue pour les débits validés. Les statistiques pour le scénario 1 sont présentées au tableau XIV.

Tableau XIV

Statistiques pour le scénario 1 – Filtre winsorizeur

<b>% d'erreur</b>	<b>Débits moyens (m<sup>3</sup>/s)</b>	<b>Log (Q<sub>x</sub>)</b>	<b>Écart-type (Log Q)</b>	<b>Coefficient d'asymétrie (Log Q)</b>
<b>50 %</b>	1 466	3,1469	0,1355	-0,6620
<b>20 %</b>	1 466	3,1469	0,1355	-0,6620
<b>5 %</b>	1 466	3,1469	0,1355	-0,6620
<b>-5 %</b>	1 465	3,1467	0,1352	-0,6703
<b>-20 %</b>	1 464	3,1465	0,1349	-0,6806
<b>-50 %</b>	1 464	3,1465	0,1349	-0,6806

Comme pour le filtre médian que nous avons traité précédemment, nous avons aussi une uniformisation des statistiques puisque les débits sont identiques.

#### **4.6.3.2 Calcul des quantiles pour les débits validés avec le filtre winsorizeur**

Les résultats du calcul des quantiles de crue effectués avec les débits validés sont présentés au tableau XV .

Tableau XV

Quantiles de crue (m<sup>3</sup>/s) – scénarios 1, 2 et 3 – Filtre winsoriseur

Scénarios	Périodes de retour	Pourcentages d'erreur dans les débits					
		50 %	20 %	5 %	-5 %	-20 %	-50 %
1	10	2 092	2 092	2 092	2 089	2 087	2 087
	20	2 343	2 343	2 343	2 340	2 336	2 336
	50	2 661	2 661	2 661	2 657	2 652	2 652
	100	2 898	2 898	2 898	2 893	2 887	2 887
	200	3 132	3 132	3 132	3 127	3 120	3 120
	500	3 442	3 442	3 442	3 435	3 427	3 427
	1 000	3 677	3 677	3 677	3 670	3 660	3 660
	10 000	4 475	4 475	4 475	4 464	4 451	4 451
2	10	2 092	2 092	2 092	2 088	2 078	2 077
	20	2 343	2 343	2 343	2 338	2 325	2 323
	50	2 661	2 661	2 661	2 656	2 638	2 635
	100	2 898	2 898	2 898	2 891	2 870	2 866
	200	3 132	3 132	3 132	3 124	3 100	3 096
	500	3 442	3 442	3 442	3 432	3 403	3 398
	1 000	3 677	3 677	3 677	3 666	3 634	3 628
	10 000	4 475	4 475	4 475	4 459	4 413	4 405
3	10	2 092	2 092	2 092	2 088	2 069	2 066
	20	2 343	2 343	2 343	2 338	2 313	2 308
	50	2 661	2 661	2 661	2 656	2 622	2 616
	100	2 898	2 897	2 898	2 891	2 851	2 843
	200	3 132	3 132	3 132	3 124	3 078	3 069
	500	3 442	3 442	3 442	3 432	3 377	3 366
	1 000	3 677	3 677	3 677	3 666	3 603	3 591
	10 000	4 475	4 475	4 475	4 459	4 370	4 354

Comme pour le filtre médian que nous avons traité précédemment, nous avons aussi une uniformisation des quantiles de crue puisque les débits sont identiques.

#### 4.6.3.3 Calcul de l'impact pour les débits validés avec le filtre winsoriseur

Les écarts obtenus par la comparaison des quantiles des débits validés avec les quantiles des débits historiques représentent l'impact de la validation. Les résultats sont présentés au tableau XVI.

Tableau XVI

Variations des quantiles – scénarios 1, 2 et 3 – Filtre winsoriseur

Scénarios	Périodes de retour	Pourcentages d'erreur dans les débits					
		50 %	20 %	5 %	-5 %	-20 %	-50 %
1	10	-5,8 %	-5,8 %	-5,8 %	-5,9 %	-6,1 %	-6,1 %
	20	-6,0 %	-6,0 %	-6,0 %	-6,1 %	-6,3 %	-6,3 %
	50	-6,2 %	-6,2 %	-6,2 %	-6,3 %	-6,5 %	-6,5 %
	100	-6,3 %	-6,3 %	-6,3 %	-6,4 %	-6,6 %	-6,6 %
	200	-6,4 %	-6,4 %	-6,4 %	-6,6 %	-6,8 %	-6,8 %
	500	-6,5 %	-6,5 %	-6,5 %	-6,7 %	-6,9 %	-6,9 %
	1 000	-6,6 %	-6,6 %	-6,6 %	-6,8 %	-7,0 %	-7,0 %
	10 000	-6,9 %	-6,9 %	-6,9 %	-7,1 %	-7,4 %	-7,4 %

Tableau XVI (suite)

<b>2</b>	<b>10</b>	-5,8 %	-5,8 %	-5,8 %	-6,0 %	-6,4 %	-6,5 %
	<b>20</b>	-6,0 %	-6,0 %	-6,0 %	-6,2 %	-6,7 %	-6,8 %
	<b>50</b>	-6,2 %	-6,2 %	-6,2 %	-6,4 %	-7,0 %	-7,1 %
	<b>100</b>	-6,3 %	-6,3 %	-6,3 %	-6,5 %	-7,2 %	-7,3 %
	<b>200</b>	-6,4 %	-6,4 %	-6,4 %	-6,6 %	-7,4 %	-7,5 %
	<b>500</b>	-6,5 %	-6,5 %	-6,5 %	-6,8 %	-7,6 %	-7,7 %
	<b>1 000</b>	-6,6 %	-6,6 %	-6,6 %	-6,9 %	-7,7 %	-7,9 %
	<b>10 000</b>	-6,9 %	-6,9 %	-6,9 %	-7,2 %	-8,2 %	-8,3 %
<b>3</b>	<b>10</b>	-5,8 %	-5,8 %	-5,8 %	-6,0 %	-6,9 %	-7,0 %
	<b>20</b>	-6,0 %	-6,0 %	-6,0 %	-6,2 %	-7,2 %	-7,4 %
	<b>50</b>	-6,2 %	-6,2 %	-6,2 %	-6,4 %	-7,6 %	-7,8 %
	<b>100</b>	-6,3 %	-6,3 %	-6,3 %	-6,5 %	-7,8 %	-8,0 %
	<b>200</b>	-6,4 %	-6,4 %	-6,4 %	-6,6 %	-8,0 %	-8,3 %
	<b>500</b>	-6,5 %	-6,5 %	-6,5 %	-6,8 %	-8,3 %	-8,6 %
	<b>1 000</b>	-6,6 %	-6,6 %	-6,6 %	-6,9 %	-8,5 %	-8,8 %
	<b>10 000</b>	-6,9 %	-6,9 %	-6,9 %	-7,2 %	-9,1 %	-9,4 %

Nous constatons que quel que soit le scénario ou la période de retour, la validation donne toujours sensiblement le même pourcentage de variations par rapport aux quantiles de crues des débits historiques soit de -5,8 % à -9,4 %. Ce pourcentage est toujours négatif.

## DISCUSSION ET INTERPRÉTATION DES RÉSULTATS

Avant d'entreprendre les démarches pour déterminer l'impact des erreurs des données mesurées sur l'estimation des quantiles de crue, nous avons situé les données hydrologiques par rapport aux mesures, aux traitements des données et aux erreurs qui en découlent. Par la suite, nous avons expliqué leur utilisation dans le cadre de la conception et de la gestion des ouvrages hydrauliques.

Le postulat de base de notre étude énonce que les données historiques comportent toujours des erreurs qu'il est difficile d'identifier car ce sont des données passées. Par conséquent, l'hypothèse de notre travail avance qu'il faut valider les données car les erreurs entraînent des risques et des coûts socio-environnementaux importants.

Dans une série de données, il est difficile de savoir quels sont les débits qui sont erronés lorsque ceux-ci respectent la tendance de variation. Les débits aberrants et manquants sont plus faciles à détecter et à corriger.

### Simulation des erreurs

Notre méthode pour déterminer l'impact des erreurs a consisté à simuler des erreurs sur des débits de pointe de la crue printanière déjà validés que nous avons considérés comme des données historiques. Cette simulation s'est réalisée en 2 étapes :

- 1) pour simuler la quantité des débits erronés, nous utilisons 3 scénarios ou nous sélectionnons dans les débits de pointe historiques 1, 2 et 4 débits. Ces débits sont constitués des valeurs maximales des débits de pointe de crue printanière;
- 2) nous modifions les débits sélectionnés de : 50 %, 20 %, 5 %, -5 %, -20 %, et -50 % pour simuler les erreurs sur les débits mesurés.

Les résultats confirment que l'impact des erreurs peut être très important selon l'erreur sur les débits. Les erreurs peuvent causer une surestimation ou une sous-estimation des quantiles de crue entraînant des risques et des coûts importants. Nous prenons comme exemple la conception d'un ouvrage d'évacuation. Nous utilisons une période de retour de 1 / 1 000 ans ou de 1 / 10 000 ans selon le cas et les objectifs à atteindre.

Ainsi, comme exemple, pour une période de retour de 1 / 1 000 ans pour le scénario 1 avec le débit surestimé de 5 %, on obtient un quantile de crue de 3 965 m<sup>3</sup>/s soit une différence de 0,7 %. Et pour le même scénario avec un débit sous-estimé de -5 %, on obtient un débit de 3 912 m<sup>3</sup>/s soit une différence de -0,7 %.

Le tableau XVII présente la synthèse des résultats pour les 3 scénarios avec une période de retour de 1 / 1 000 ans et de 1 / 10 000 ans pour un ouvrage d'évacuation.

Tableau XVII

## Synthèse des impacts – Débits modifiés

% erreur	Période de retour 1 / 1 000 ans			Période de retour 1 / 10 000 ans		
	Scénarios			Scénarios		
	1	2	3	1	2	3
<b>50 %</b>	7,4 %	14,6 %	28,8 %	8,7 %	17,3 %	34,6 %
<b>20 %</b>	2,8 %	2,7 %	10,9 %	3,3 %	3,1 %	12,9 %
<b>5 %</b>	0,7 %	1,3 %	2,6 %	0,8 %	1,6 %	3,0 %
<b>-5 %</b>	-0,7 %	-1,3 %	-2,5 %	-0,8 %	-1,5 %	-2,9 %
<b>-20 %</b>	-2,4 %	-4,7 %	-9,2 %	-2,8 %	-5,5 %	-10,5 %
<b>-50 %</b>	-3,8 %	-7,4 %	-14,5 %	-4,3 %	-8,2 %	-16,0 %

Nous constatons que l'impact des erreurs est faible, soit 0,7 % ou -0,7 % pour les petites variations avec 1 / 1000 ans. Cependant il peut devenir important lors des variations plus grandes soit 34,6 % et -16,0 % avec 1 / 10 000 ans. Nous remarquons que les variations ne montrent pas de grandes différences entre les 2 périodes de retour.

### Validation des données

Notre méthode pour déterminer l'impact des erreurs a consisté à simuler des erreurs sur des données historiques puis à les valider. Cette simulation s'est réalisée en utilisant le filtre médian et le filtre winsoriseur avec le logiciel PREVAL.

Par exemple, pour la validation avec le filtre médian et une période de 1 / 1 000 ans pour le scénario 1, avec le débit surestimé de 5 % on obtient un quantile de crue de 3 551 m<sup>3</sup>/s soit une différence de -9,8 %. Pour le débit sous-estimé de -5 % on obtient un quantile de crue de 3 528 m<sup>3</sup>/s soit une différence de -10,4 %.

Le tableau XVIII présente la synthèse des variations des débits pour les périodes de 1 / 1 000 ans et de 1 / 10 000 ans pour les 3 scénarios validés avec le filtre médian pour 1981.

Tableau XVIII

Synthèse des impacts – Débits validés

%	Période de retour 1 / 1 000 ans			Période de retour 1 / 10 000 ans		
	Scénarios			Scénarios		
	1	2	3	1	2	3
<b>50 %</b>	-9,8 %	-7,3 %	-5,7 %	-10,2 %	-7,3 %	-5,4 %
<b>20 %</b>	-9,8 %	-7,3 %	-5,7 %	-10,2 %	-7,3 %	-5,4 %
<b>5 %</b>	-9,8 %	-7,3 %	-5,7 %	-10,2 %	-7,3 %	-5,4 %
<b>-5 %</b>	-10,4 %	-8,5 %	-7,4 %	-10,8 %	-8,7 %	-7,4 %
<b>-20 %</b>	-10,4 %	-9,3 %	-8,6 %	-10,9 %	-9,6 %	-8,8 %
<b>-50 %</b>	-10,4 %	-9,3 %	-8,6 %	-10,9 %	-9,6 %	-8,8 %

De ce tableau, nous en tirons les observations suivantes :

- 1) les variations sont toutes négatives;
- 2) les variations se situent entre -8,8 % et -10,9 %;
- 3) les quantiles sont uniformisés. Par exemple, pour le scénario 1, on obtient un impact de -9,8 % pour des débits erronés de 5 %, 20 % et 50 %.

Notre travail s'est limité à utiliser des données validées, sans connaître la méthode de validation utilisée. Dans le cadre de notre étude, cette information n'est pas conséquente car tous nos calculs se basent sur la même série historique, il y a donc cohérence dans les résultats. Les calculs se limitent à un nombre restreint de combinaisons de variantes. Les calculs des quantiles de crue ont été effectués avec la loi log normale, qui est une des lois les plus utilisées en Amérique du nord.

Nous comparons au tableau XIX les résultats obtenus avec le filtre médian et ceux obtenus avec le filtre winsoriseur. Les paramètres retenus sont : une période de retour de 1 / 1 000 ans et des débits modifiés avec un taux d'erreurs simulés de +5 % et de -5 %.

Tableau XIX

Comparaison entre les 2 filtres

Filtres	% erreur	Scénarios pour 1 / 1 000 ans		
		1	2	3
<b>médian</b>	<b>5 %</b>	-9,8 %	-7,3 %	-5,7 %
<b>winsoriseur</b>	<b>5 %</b>	-6,6 %	-6,6 %	-6,6 %
<b>médian</b>	<b>-5 %</b>	-10,4 %	-8,5 %	-7,4 %
<b>winsoriseur</b>	<b>-5 %</b>	-6,8 %	-6,9 %	-6,9 %

Nous constatons que les résultats avec les deux filtres sont sensiblement identiques. Que le filtre médian est plus conservateur que le filtre winsoriseur pour l'exemple que nous avons analysé.

La validation donne toujours des variations négatives. En effet, les nouveaux débits estimés sont une moyenne des débits des journées précédentes et suivantes. Par conséquent, le nouveau débit validé est toujours inférieur au débit historique. De plus, dans la réalité nous ne connaissons pas le vrai débit ce qui ne nous permet pas d'affirmer que la validation sous-estime systématiquement les débits.

Analyse des modifications sur les débits

Lorsque nous modifions des débits, il est important de vérifier l'impact des modifications sur l'ensemble de la série. Pour ce faire, nous avons comparé à l'aide d'un graphique les débits historiques, les débits modifiés et les débits validés avec le filtre médian. La figure 20 représente une modification effectuée sur un débit pour 1981.

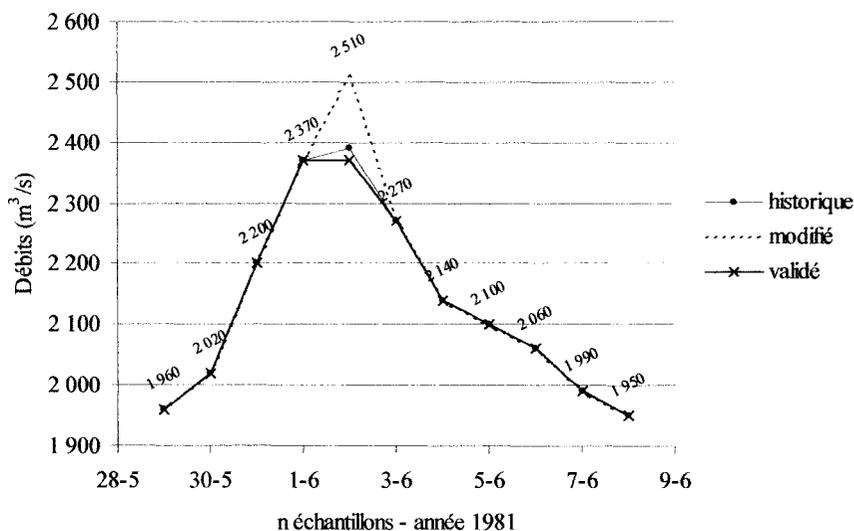


Figure 20 Impact d'une modification sur la série de données

On remarque que la modification a un impact uniquement sur les valeurs voisines du débit modifié. On peut voir que les courbes historiques, modifiées et validées se superposent avant et après la modification. Alors, la modification d'un débit à une période donnée n'a pas d'impact sur les débits des autres périodes.

## CONCLUSION

Cette étude a permis de déterminer l'impact des erreurs sur l'estimation des quantiles de crue qui peut avoir une répercussion sur les décisions des concepteurs et des gestionnaires d'ouvrages d'art.

L'impact des erreurs a été déterminé par la comparaison des quantiles calculés à l'aide d'une méthode statistique de la loi log normale. Nous avons effectué deux comparaisons : les quantiles des données historiques versus les quantiles des données modifiées et les quantiles des données historiques versus les quantiles des données validées. Pour ce faire, nous avons faussé les données historiques pour simuler des erreurs, puis nous avons validé les données modifiées.

Les résultats obtenus ont permis de déterminer l'impact des erreurs qui peut avoir des effets à court, moyen et long terme sur la conception ou la gestion des ouvrages. Cet impact est fonction des erreurs sur les débits de pointe de la crue printanière.

Selon l'importance des erreurs, il y aura des risques avec plusieurs impacts socio-économiques. D'une part, on peut obtenir une surestimation des quantiles de crue avec comme impacts des coûts supplémentaires de construction ainsi que des effets plus ou moins importants sur l'environnement. D'autre part, une sous-estimation des quantiles de crue aura pour effet d'augmenter les risques d'inondation et des bris d'ouvrage (selon la crue de sécurité) et d'engendrer des coûts liés à la gestion des risques et des conséquences avec des répercussions importantes sur la société.

Cependant, l'impact des erreurs est difficile à évaluer avec exactitude puisque de nombreux paramètres interviennent dans les mesures de données, dans le traitement des données ou dans la méthodologie de calcul.

Ainsi, pour diminuer des erreurs dans les données mesurées, plusieurs actions peuvent être posées. À savoir, retracer et corriger les erreurs en observant les résultats passés, en étudiant les données historiques, en les comparant avec les données actuelles ou en utilisant des données de sites comparables ou proches. Finalement, pour limiter les erreurs, il est toujours possible d'améliorer les instruments et les techniques de mesure et de traitement des données.

## RECOMMANDATIONS

- 1) Encourager et généraliser la validation des données auprès des concepteurs et des gestionnaires.
- 2) Choisir parmi toutes les méthodes existantes, les méthodes d'analyse et de validation les plus appropriées en fonction des données mesurées et des objectifs à atteindre.
- 3) Utiliser un outil informatique spécialisé dans le domaine, comme le logiciel PREVAL, afin d'accélérer et de simplifier le traitement des données.
- 4) Favoriser la mise sur pied d'une banque de données centralisée afin de regrouper toutes les données, leur historique et leurs paramètres de traitement. Permettre l'accessibilité et le partage entre les différents intéressés.
- 5) Standardiser les méthodes de validation en fonction des données et objectifs à atteindre.
- 6) Améliorer les instruments, les techniques de mesure et de traitement pour réduire l'impact des erreurs.

Dans la continuité de notre étude, il serait intéressant d'explorer et d'élargir d'autres pistes d'étude, à savoir :

- 1) Analyser les impacts en utilisant la méthode stochastique et la méthode mixte déterministe-statistique pour comparer les résultats des trois méthodes.

- 2) Déterminer une plus-value monétaire pour évaluer la relation entre les calculs des quantiles de crue et ses impacts sur les coûts de construction.
- 3) Évaluer l'impact des erreurs de la reconstitution des données manquantes.
- 4) Évaluer l'impact des erreurs en utilisant les crues d'étiages.
- 5) Faire des tests avec des ouvrages existants ou des essais en laboratoire sur l'évaluation des quantiles de crues.
- 6) Déterminer une méthodologie pour contrôler la validation du débit de pointe.

## BIBLIOGRAPHIE

- [1] Environnement Canada, relevés hydrologiques du Canada, *Programme hydrométrique*. [En ligne], [http://www.msc.ec.gc.ca/wsc/hydrometric/main\\_f.cfm?cname=hydrometric\\_f.cf](http://www.msc.ec.gc.ca/wsc/hydrometric/main_f.cfm?cname=hydrometric_f.cf), (page consultée le 8 mai 2003).
- [2] Chaire en hydrologie statistique de l'INRS-Eau, terre et environnement, *Régionalisation des événements hydrologiques extrêmes*, projet C5.
- [3] Site du Centre d'expertise hydrique du Québec, *La gestion des barrages publics*, [En ligne], <http://www.menv.gouv.qc.ca/cehq/index.htm>, (page consultée le 20 mai 2003).
- [4] Saad Bennis (2003), *Hydraulique et hydrologie*, chapitre 10, Université du Québec, École de Technologie Supérieure.
- [5] Chaire en hydrologie statistique de l'INRS-Eau, terre et environnement, *Validation des apports naturels historiques*, projet C1.
- [6] Monition L., Le Nir M., Roux J., *Les microcentrales hydroélectriques*, édition MASSON.
- [7] Marc Morell (1998), *Le système méditerranéen d'observation du cycle hydrologique*, Congrès international de Kaslik.
- [8] Site Hydram, *L'organisation et le contrôle des données*, chapitre 8, [En ligne] <http://hydram.epfl.ch/labo/hydr/labo.htm>. (page consultée le 28 août 2003).
- [9] P. Coulibaly, T. Ouarda, B. Bobée, Water news profile volume 20 issue 2 – juin 2001, *Comment améliorer la précision des prévisions hydrologiques*, Association canadienne des ressources hydriques.
- [10] Saad Bennis, Narut Kang, *Multivariate technique for validating historical hydrometric data with redundant measurements*, *Nordic hydrology*, 31 (2), 2000, 107-126
- [11] Luc Gagnon (1994), *Filtrage des séries hydrométéorologiques, projet d'application d'un rapport de projet de maîtrise*, Université du Québec, École de Technologie Supérieure.

- [12] Chaire en hydrologie statistique de l'INRS-Eau, terre et environnement, *Acquisition et validation des données hydrométéorologiques*, Projet T1.
- [13] Brochure d'information de l'Agence Internationale de l'énergie Atomique, *Gestion des ressources en eau au moyen de l'hydrologie isotopique*, 02-01579/FS série 2/03/F.
- [14] Hydro-Québec (janvier 2000), *Guide de conception des aménagements hydroélectriques*, Direction ingénierie, Groupe Ingénierie, Approvisionnement et construction.
- [15] Eric Gaume (2000), *Hydrologie de bassins versants et modélisation pluie-débits*, cours d'hydrologie, chapitres 2 et 3, École nationale des ponts et chaussées.
- [16] Sylvain Côté, Saad Bennis, Narut Kang (mars 1995), *Manuel d'utilisation du logiciel de validation des données hydrométriques historiques journalières*, Université du Québec, École de Technologie Supérieure
- [17] Gérard Baillargeon (1989), *Probabilités statistiques et techniques de régression*, éditions SMG.
- [18] Publication d'Environnement Québec (mai 2001), *La chronique environnementale, Les domaines d'intervention du centre d'expertise hydrique du Québec*.