

**Record Number:** 11980  
**Author, Monographic:** Perreault, L.//Haché, M.//Bobée, B.  
**Author Role:**  
**Title, Monographic:** Stationnarité et modélisation de la série d'apports énergétiques annuels agrégés : synthèse des principaux résultats et recommandations dans le cadre du projet C2  
**Translated Title:**  
**Reprint Status:**  
**Edition:**  
**Author, Subsidiary:**  
**Author Role:**  
**Place of Publication:** Québec  
**Publisher Name:** INRS-Eau  
**Date of Publication:** 1996  
**Original Publication Date:** 5 juin 1996  
**Volume Identification:**  
**Extent of Work:** ix, 37  
**Packaging Method:** pages incluant 2 annexes  
**Series Editor:**  
**Series Editor Role:**  
**Series Title:** INRS-Eau, rapport de recherche  
**Series Volume ID:** 464  
**Location/URL:**  
**ISBN:** 2-89146-427-3  
**Notes:** Rapport annuel 1996-1997  
**Abstract:** Rapport rédigé par la Chaire en hydrologie statistique.  
**Call Number:** R000464  
**Keywords:** rapport/ ok/ dl

**RAPPORT DE SYNTHÈSE - PROJET C2**

**STATIONNARITÉ ET MODÉLISATION  
DES APPORTS ÉNERGÉTIQUES  
ANNUELS AGRÉGÉS**

# **STATIONNARITÉ ET MODÉLISATION DE LA SÉRIE D'APPORTS ÉNERGÉTIQUES ANNUELS AGRÉGÉS**

## **SYNTHÈSE DES PRINCIPAUX RÉSULTATS ET RECOMMANDATIONS DANS LE CADRE DU PROJET C2**

Rapport rédigé pour

Chaire en Hydrologie statistique Hydro-Québec / CRSNG

par

**Luc Perreault  
Mario Haché  
Bernard Bobée**

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

Rapport de recherche No R-464

5 juin 1996





## **PROJET C2 : ÉQUIPE DE RECHERCHE**

---

<b>CHARGÉ DE PROJET</b>	<b>INRS-Eau</b>	LUC PERREAULT
<b>CHARGÉ DE PROJET</b>	<b>Hydro-Québec</b>	JEAN-LOUIS BISSON
<b>COLLABORATEURS</b>	<b>INRS-Eau</b>	MARIO HACHÉ MICHEL SLIVITZKY PETER RASMUSSEN
	<b>Hydro-Québec</b>	PIERRE BRUNEAU DENIS TREMBLAY NARUT KANG LAURA FAGHERAZZI
	<b>IREQ</b>	ANDRÉ TURGEON
	<b>Professeur invité</b>	JACQUES BERNIER
<b>TITULAIRE DE LA CHAIRE</b>	<b>INRS-Eau</b>	BERNARD BOBÉE

# TABLE DES MATIÈRES

---

<b>LISTE DES TABLEAUX .....</b>	<b>vii</b>
<b>LISTE DES FIGURES.....</b>	<b>ix</b>
<b>CHAPITRE 1 : Introduction.....</b>	<b>1</b>
1.1 Objectif du rapport.....	1
1.2 Séries d'apports énergétiques.....	1
1.3 Objectifs du sous-projet 2.4.....	3
<b>CHAPITRE 2 : Stationnarité de la moyenne.....</b>	<b>5</b>
2.1 Définition.....	5
2.2 Vérification de la stationnarité (série agrégée).....	6
2.3 Conclusions.....	10
<b>CHAPITRE 3 : Modélisation et prévision des apports .....</b>	<b>15</b>
3.1 Généralités concernant les modèles.....	15
3.2 Modélisation de la série d'apports annuels agrégés .....	16
3.2.1. Modèles considérés.....	17
3.2.2. Comparaison des modèles.....	19
3.3 Prévision et distribution des écarts .....	21
3.4 Conclusions.....	23
<b>RÉFÉRENCES BIBLIOGRAPHIQUES .....</b>	<b>25</b>
<b>ANNEXE A : Coefficient de Hurst.....</b>	<b>29</b>
<b>ANNEXE B : Diagramme en boîte.....</b>	<b>33</b>





## **LISTE DES TABLEAUX**

---

<b>Tableau 2.2.1.</b> Statistiques de base de la série d'apports annuels agrégés (1943-1993).....	6
<b>Tableau 2.2.2.</b> Année la plus probable d'un changement de moyenne (1943-1993) .....	9
<b>Tableau 3.2.1</b> Modèles considérés pour la simulation.....	18
<b>Tableau 3.3.1</b> Prévision (GWh), écart-type, intervalle de confiance et probabilité de surplus et de déficit pour les modèles retenus (1994-1998) .....	22



## LISTE DES FIGURES

---

<b>Figure 2.2.1</b>	Apports et écarts énergétiques annuels agrégés (1943-1993).....	7
<b>Figure 2.2.2</b>	Probabilité du changement de moyenne - Apports agrégés.....	8
<b>Figure 2.2.3</b>	Probabilité du changement de moyenne - Complexe St-Laurent.....	9
<b>Figure 2.2.4</b>	Série simulée à partir d'un modèle stationnaire AR(1).....	11
<b>Figure 3.2.1</b>	Distribution des moyennes des séries simulées.....	20
<b>Figure 3.3.1</b>	Prévisions issues des modèles retenus pour la période 1994 à 1998.....	21

# **1 INTRODUCTION**

---

## **1.1 Objectif du rapport**

Ce rapport a pour objet de présenter brièvement quelques travaux réalisés dans le cadre du projet C2 de la chaire en hydrologie statistique (Étude des séries historiques d'apports énergétiques). Ce document s'intéresse plus particulièrement au sous-projet 2.4 (Modélisation de la série agrégée d'apports énergétiques annuels) et fait suite aux questions de recherche soumises à l'INRS-Eau le 31 août 1995 en matière de planification énergétique. Trois grands thèmes ont été abordés : la stationnarité de la moyenne, la modélisation de la série d'apports énergétiques agrégés annuels (simulation et prévision) et la distribution statistique des écarts énergétiques annuels par rapport à la moyenne (déficit et surplus énergétique).

Les sections 1.2 et 1.3 sont consacrées respectivement à la description des séries de données d'apports énergétiques actuellement employées pour la planification à Hydro-Québec et aux objectifs du sous-projet 2.4. Le problème de stationnarité de la moyenne est traité au chapitre 2 du présent rapport. Le chapitre 3 est consacré à la modélisation et à la prévision des écarts énergétiques agrégés annuels. Neuf modèles ont été considérés et comparés. L'étude comparative a été réalisée d'abord en examinant leurs résidus et ensuite en étudiant par simulation la distribution d'un ensemble de caractéristiques statistiques reproduites par ces modèles. La distribution des écarts cumulés prédits  $y$  est aussi présentée.

Les détails théoriques sont omis dans le présent rapport. Une description complète des analyses effectuées dans le projet C2 est disponible dans Perreault *et al.* (1996).

## **1.2 Séries d'apports énergétiques**

Le potentiel hydroélectrique du Québec repose sur les réserves hydrauliques emmagasinées dans les réservoirs de 7 grands complexes : Outaouais, St-Maurice, Aux Outardes, Manicouagan, Bersimis, St-Laurent et La Grande. Un huitième grand complexe, Churchill-Falls, n'est pas géré par Hydro-Québec mais, puisque la majorité de sa production est dirigée vers le réseau d'Hydro-Québec, ses règles de gestion sont conformes en général aux besoins du réseau et sa production est incluse dans les simulations de gestion du réseau d'Hydro-Québec.

Afin de connaître et de prévoir avec précision l'état des réserves énergétiques, on calcule, chaque jour et pour chacun des sites du réseau, les apports naturels exprimés en  $m^3/s$ . Le cumul mensuel des apports naturels quotidiens permet de connaître les apports naturels mensuels de l'ensemble des sites. Connaissant la capacité de production de chaque centrale, on peut ensuite transformer le stock d'eau emmagasinée et les apports naturels en valeur énergétique exprimée en gigawatts-heure (GWh). L'apport énergétique mensuel  $E_n$  d'une centrale donnée est donc proportionnel à l'apport naturel mensuel  $A_n$  calculé au site. Hydro-Québec dispose de séries d'apports énergétiques mensuels calculés pour l'ensemble des centrales de son parc d'équipements. Ces séries s'échelonnent sur une période variable selon les sites (de 20 à 51 années).

L'énergie mensuelle totale produite par un complexe hydroélectrique est calculée en cumulant les apports énergétiques de chaque site du complexe de l'amont vers l'aval. Ensuite, l'énergie annuelle totale du complexe est obtenue en sommant les apports énergétiques mensuels sur toute l'année. Enfin, une série annuelle agrégée, représentant la production totale de l'ensemble du parc d'équipements d'Hydro-Québec, est construite en sommant les apports énergétiques annuels totaux de chaque complexe. Ainsi, quatre types de séries d'apports énergétiques sont obtenues:

1. Apports énergétiques mensuels de chaque centrale;
2. Apports énergétiques mensuels des huit complexes hydroélectriques (cumul spatial);
3. Apports énergétiques annuels des huit complexes hydroélectriques (cumul temporel);
4. Apports énergétiques annuels agrégés (cumul spatial).

Ces différentes étapes de calcul et d'agrégation menant à la construction de ces quatre types de séries sont précédées d'une opération préliminaire réalisée sur les apports naturels mensuels  $A_n$  aux sites. Il s'agit d'une reconstitution et d'une extension de séries permettant d'obtenir une plage commune d'au moins cinquante années de valeurs d'apports à chaque site (Bruneau *et al.*, 1984, 1986a, b, c).

Les données d'apports énergétiques sont la base du système de gestion et de planification de l'énergie produite par Hydro-Québec. Une bonne connaissance des caractéristiques statistiques des différentes séries d'apports énergétiques est nécessaire au bon fonctionnement du système de planification d'Hydro-Québec. Ainsi, les procédures de reconstitution et d'extension de séries utilisées par Hydro-Québec ont d'abord fait l'objet d'une étude critique au sous-projet 2.2. Ensuite, au sous-projet 2.3, une analyse statistique

descriptive a été effectuée pour les quatre séries d'apports énergétiques s'échelonnant sur 51 années (1943-1993). Ces deux sous-projets ont permis de faire le point sur la construction des séries d'apports énergétiques et de mettre à jour leurs caractéristiques statistiques. Les résultats et les recommandations concernant ces travaux sont donnés dans Perreault *et al.* (1996) et ne sont pas présentés ici.

### 1.3 Objectifs du sous-projet 2.4

Hydro-Québec présente périodiquement depuis 1979, un plan de développement de ses activités pour les 10 années à venir. Depuis 1990, ce plan est réactualisé tous les trois ans. Cet exercice repose sur une analyse détaillée de l'évolution du contexte économique et énergétique. Il permet, en particulier, d'établir les politiques de développement à court, moyen et long termes qu'adoptera l'entreprise. C'est pourquoi, il est important, du point de vue économique, de prévoir avec un maximum de précision la demande et l'offre hydroélectrique, puisque ces deux composantes sont déterminantes dans le choix des interventions à privilégier.

Afin de déterminer les risques de défaillance énergétique associés à différents scénarios d'exploitation et de bien planifier le parc d'équipements de production hydroélectrique, Hydro-Québec utilise différents modèles de simulation. Ces modèles reposent, entre autres, sur les caractéristiques statistiques de la série d'apports énergétiques<sup>1</sup> annuels agrégés (moyenne, écarts-type, autocorrélation) et sont basés sur l'hypothèse que les observations sont stationnaires.

Les résultats des analyses statistiques effectuées au sous-projet 2.3, en particulier ceux concernant la présence éventuelle d'un changement de moyenne dans la série des apports énergétiques annuels agrégés, ont permis de définir plus précisément les objectifs du sous-projet 2.4. Un document précisant les points sur lesquels Hydro-Québec souhaite obtenir des recommandations de l'équipe de la chaire a été rédigé et remis à l'INRS-Eau le 31 août 1995. Les questions de recherche prioritaires soumises dans ce document concernent la stationnarité de la moyenne, la modélisation de la série d'apports énergétiques agrégés

---

<sup>1</sup> Depuis le début du projet C2, nous avons considéré la série qui s'échelonne de 1943 à 1993. Dans la version finale du rapport Perreault *et al.* (1996), les années 1994 et 1995 sont prises en compte dans le chapitre 4 concernant les prévisions.

annuels (simulation et prévision) et la distribution statistique des écarts énergétiques annuels par rapport à la moyenne (déficit et surplus énergétique).

Pour répondre aux besoins d'Hydro-Québec, le sous-projet 2.4 a finalement été divisé en quatre étapes :

- Analyse approfondie de la stationnarité de la moyenne des apports énergétiques annuels agrégés
- Modélisation des écarts énergétiques annuels agrégés par rapport à la moyenne générale et prévisions
- Distribution statistique des écarts énergétiques annuels agrégés
- Sensibilité des prévisions au changement de moyenne

## 2 STATIONNARITÉ DE LA MOYENNE

---

Certains résultats de l'analyse statistique des séries d'apports énergétiques annuels (sous-projet 2.3) ont amené les membres du projet C2 à envisager sérieusement la présence, dans la série agrégée, de changements de moyennes. Puisque la moyenne est un des principaux intrants des modèles de planification (simulations et prévisions) utilisés à Hydro-Québec et que la stationnarité de cette statistique est une hypothèse de base en modélisation, la première étape du sous-projet 2.4 a été consacrée à l'étude plus approfondie de ce phénomène. Ces travaux avaient comme objectif d'étudier certaines questions de recherche soumises par Hydro-Québec (document du 31 août 1995). En particulier :

- Quelle moyenne doit être utilisée pour les simulations stochastiques ?
- Combien d'années sont nécessaires pour détecter un éventuel changement de moyenne ?
- Ce phénomène a-t-il été étudié et détecté ailleurs dans le monde ?
- Les changements de moyennes constituent-ils un cycle hydrologique ?
- Dans cette éventualité, la durée de ce cycle est-elle prévisible ?

### 2.1 Définition

Les modèles statistiques classiques permettant de représenter une série chronologique en hydrologie reposent en particulier sur l'hypothèse de stationnarité de la moyenne. En pratique, une série chronologique est dite *stationnaire* si les observations s'éloignent assez uniformément autour d'un niveau fixé. Les données peuvent toutefois s'écarter de ce niveau mais elles auront tendance à se redistribuer nécessairement autour de cette valeur. Cette déviation peut se maintenir sur une période de temps plus ou moins longue selon la corrélation temporelle entre les observations. Ce phénomène est appelé ici *persistance* et peut être mesuré à l'aide du coefficient d'autocorrélation. Plus le coefficient d'autocorrélation est élevé, plus cette déviation par rapport au niveau moyen risque de s'échelonner sur une longue période. Le niveau considéré, autour duquel les données d'une série stationnaire gravitent, est habituellement la moyenne générale de la série.

De plus en plus d'hydrologues sont maintenant soucieux du fait que l'hypothèse de stationnarité de la moyenne est susceptible d'influencer considérablement les résultats d'un modèle de série chronologique. En effet, si cette hypothèse n'est pas respectée, les



prévisions issues d'un modèle classique peuvent s'avérer erronées et par conséquent influencer grandement la prise de décision.

## 2.2 Vérification de la stationnarité (série agrégée)

Le tableau 2.2.1 donne les principales statistiques de base de la série d'apports énergétiques annuels agrégés exprimés en GWh. La figure 2.2.1 présente la courbe chronologique des séries d'apports et d'écart énergétique par rapport à la moyenne.

**Tableau 2.2.1.** Statistiques de base de la série d'apports annuels agrégés (1943-1993).

Moyenne	Médiane	Écart-type	Cv	Cs	Min.	Max.	Autocor.
183 822	184 238	19 079	0.156	0.104	145 155	234 438	0.36

Un changement de moyenne dans la série d'apports énergétiques annuels agrégés peut survenir dans trois situations distinctes :

- le changement de moyenne est engendré par la manipulation des données de base (reconstitution et extension de séries);
- le changement de moyenne est dû à une modification des appareils de mesures ou de l'aménagement du bassin;
- le changement de moyenne est lié à un phénomène physique (changement climatique, changement physiographique du bassin, etc.).

Les deux premières situations sont plus faciles à traiter puisque la date du changement éventuel est connue. Les techniques statistiques usuelles de détection du changement de moyenne sont alors applicables (par exemple, les tests statistiques de comparaison des moyennes avant et après la date d'intérêt). Des tests statistiques ont donc été utilisés pour comparer les moyennes avant et après certaines dates correspondant à des changements méthodologiques de reconstitution et d'extension de séries. Aucun changement significatif n'a été détecté à ces dates dans la série d'apports énergétiques agrégés annuels. D'ailleurs, un des critères de validation des reconstitutions et des extensions de séries effectuées par Hydro-Québec était la conservation de l'homogénéité temporelle (Bruneau *et al.*, 1984, 1986 a,b,c).

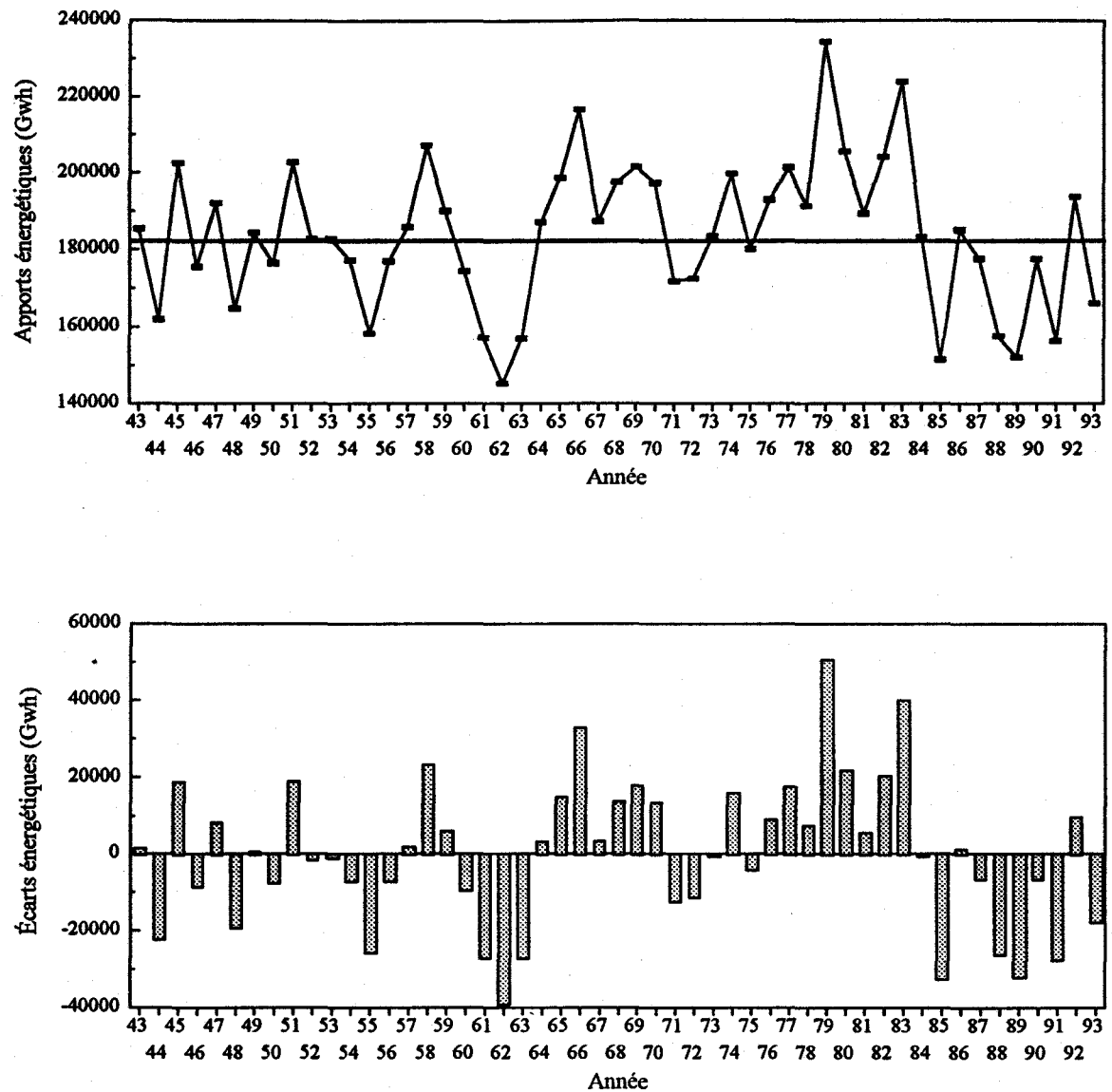
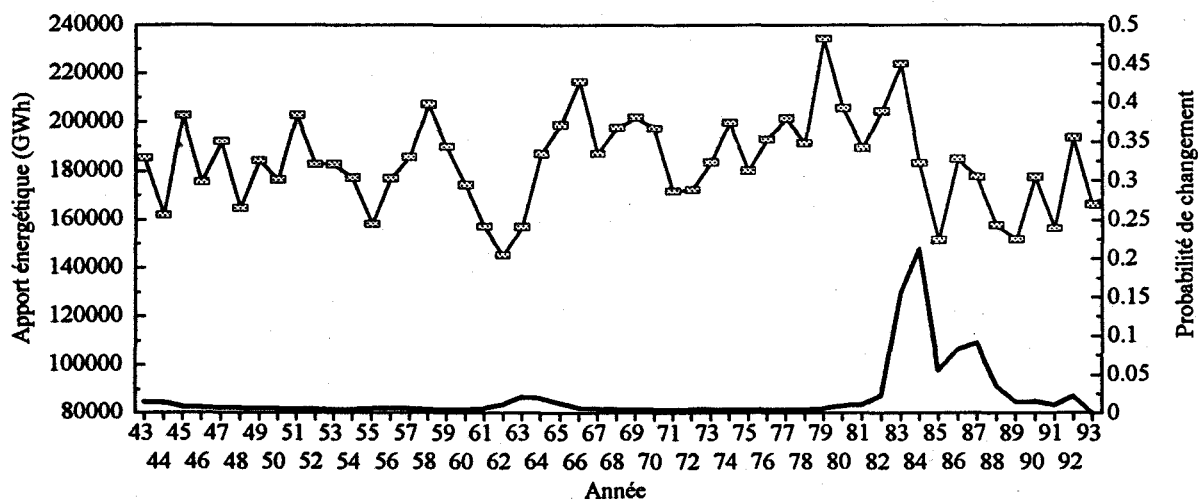


Figure 2.2.1. Apports et écarts énergétiques annuels agrégés (1943-1993).

Pour vérifier la stationnarité des apports énergétiques dans la troisième situation, les méthodes usuelles ne peuvent être employées. La procédure bayésienne de Lee et Heghinian (1977)<sup>1</sup> a donc été considérée. Cette approche permet de déterminer, sous l'hypothèse qu'un changement de moyenne soit effectivement présent dans la série, la période  $t$  où ce

<sup>1</sup> Cette procédure bayésienne a été généralisée par Bernier (1994).

changement possède la plus grande probabilité de se réaliser. La figure 2.2.2 présente les résultats de l'application de cette procédure aux données d'apports énergétiques annuels agrégés. La courbe en trait gras représente, pour chaque année, la probabilité que le changement de moyenne survienne à cette période. La procédure de Lee et Heghenian indique que, si la série admet effectivement une non-stationnarité de la moyenne, sa date la plus probable d'occurrence est l'année 1984. On observe aussi une faible probabilité que ce changement survienne en 1963. La procédure de Lee et Heghinian, appliquée au sous-échantillon 1943-1983 (c'est-à-dire, en enlevant l'effet des apports 1984-1993), confirme cette observation. Selon ces résultats, la série agrégée pourrait être non stationnaire et constituée de trois sous-séries de moyennes différentes (1943-1963, 1964-1983, 1984-1993).



**Figure 2.2.2.** Probabilités du changement de la moyenne - Apports agrégés.

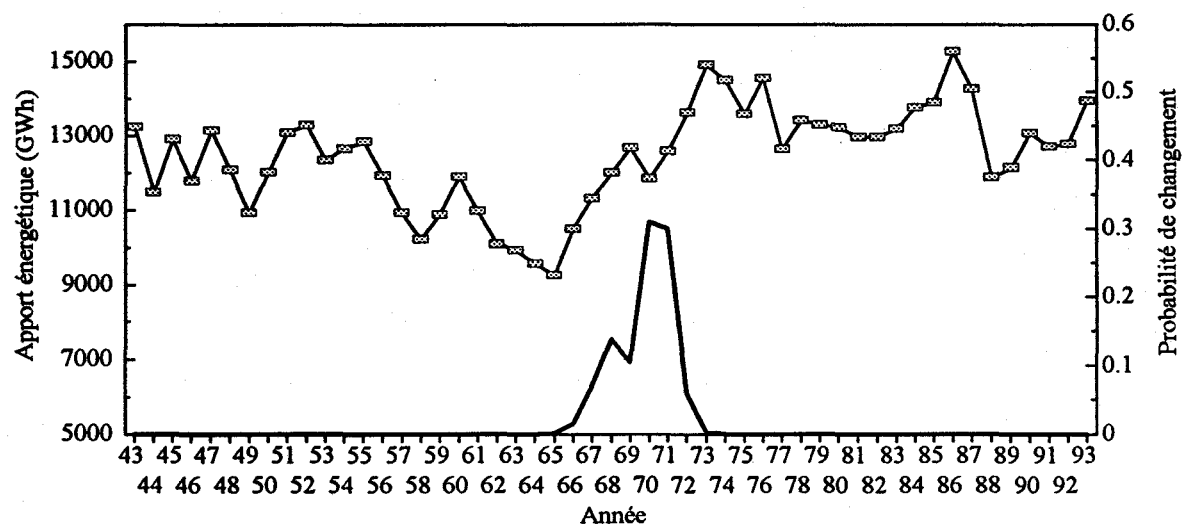
Ce résultat repose sur une information de base restreinte (51 années d'observations). Il est alors difficile de conclure à un phénomène non stationnaire uniquement à partir de cette analyse statistique. En effet, ce type de déviation par rapport à la moyenne générale de la série peut très bien se produire alors que celle-ci est stationnaire (phénomène de persistance, cf. section 2.1). De plus, la présence de changements de moyennes a des implications majeures sur la prise de décisions à Hydro-Québec. Aussi, faut-il pousser plus loin l'analyse en considérant d'autres séries reliées aux apports énergétiques annuels agrégés.

La série agrégée étant une agglomération spatiale de différentes séries, il apparaît logique d'appliquer d'abord la procédure aux apports énergétiques annuels de chacun des huit complexes. Les résultats sont présentés au tableau 2.2.2.

**Tableau 2.2.2.** Année la plus probable d'un changement de moyenne (1943-1993).

Complexes	STM	OTW	STL	LGR	MAN	OUT	BER	CHF
Année	Aucune	Aucune	1970	1984	1983	1983	1972	1984

Mentionnons que le saut de moyenne observé en 1970 (année la plus probable, cf. figure 2.2.3) pour le complexe St-Laurent est nettement plus significatif statistiquement que ceux des autres séries (cf. Perreault *et al.*, 1996). Une étude conjointe entre l'INRS-Eau et Hydro-Québec a d'ailleurs permis de détecter un changement de moyenne des apports naturels à cette période pour certains des Grands Lacs dont le comportement influence le système du St-Laurent (Mathier *et al.*, 1992). Toutefois, malgré l'intensité et la netteté de cet événement (cf. figure 2.2.3), ce saut de moyenne n'affecte pas la série agrégée pour laquelle la probabilité de changement est nulle en 1970 (cf. figure 2.2.2). La faible contribution des apports du St-Laurent aux apports totaux explique ce résultat.



**Figure 2.2.3.** Probabilités du changement de la moyenne - Complexe St-Laurent.

Finalement, la procédure de Lee et Heghinian a aussi été appliquée à des séries de précipitations annuelles (Perreault *et al.*, 1996). Les séries considérées représentent les précipitations de régions allant des plaines du sud des États-Unis aux provinces de

l'Atlantique (tirées de Findlay *et al.* (1994) pour les séries canadiennes et Karl *et al.* (1994) pour les séries américaines). La probabilité associée à l'année 1984 est négligeable pour l'ensemble des séries de précipitation contrairement à ce qui a été observé pour la série d'apports énergétiques agrégés (aussi La Grande et Churchill Falls). Ce phénomène est probablement propre au nord du Québec puisque cette date a été détectée principalement pour les complexes situés dans cette région (cf. tableau 2.2.2). Toutefois, les résultats obtenus pour les précipitations montrent une probabilité de changement non négligeable entre 1963 et 1970 pour la plupart des régions. Cela tend à montrer qu'un phénomène climatique serait effectivement survenu durant la décennie 1960-1970 et que cet événement aurait davantage influencé les apports énergétiques du St-Laurent que ceux des autres systèmes hydriques.

## 2.3 Conclusions

### *Moyenne à utiliser pour d'éventuelles simulations stochastiques*

La détection d'un changement de moyenne dans une série d'observations à partir d'une procédure statistique est relativement simple à réaliser. Toutefois, compte tenu de l'information restreinte dont on dispose ici (51 données d'apports énergétiques), on ne peut distinguer entre la non-stationnarité de la moyenne et la présence de déviations soutenues engendrées par l'autocorrélation (persistance). Pour illustrer ce propos, la figure 2.2.4 présente une série de 50 données simulées à partir d'un modèle AR(1) (cf. chapitre 3) avec les mêmes caractéristiques statistiques que celles de la série agrégée (moyenne, écart-type et autocorrélation). Ces observations, quoiqu'issues d'un modèle tout à fait stationnaire, présente des déviations persistantes d'amplitude similaire à celles observées dans la série d'apports énergétiques annuels agrégés (figure 2.2.2).

Malgré les résultats obtenus à partir des séries de précipitations, supportant la présence d'un changement de la moyenne, il est difficile de conclure "hors de tout doute raisonnable" à la non-stationnarité de la moyenne dans la série des apports énergétiques annuels agrégés. C'est pourquoi, nous recommandons l'utilisation de la moyenne générale de la série comme intrant des modèles de simulations et de prévisions. Nous considérons donc, au chapitre 3, des modèles classiques qui reposent sur l'hypothèse de stationnarité. Toutefois, une analyse de la sensibilité de ces modèles au changement de la moyenne est recommandée.

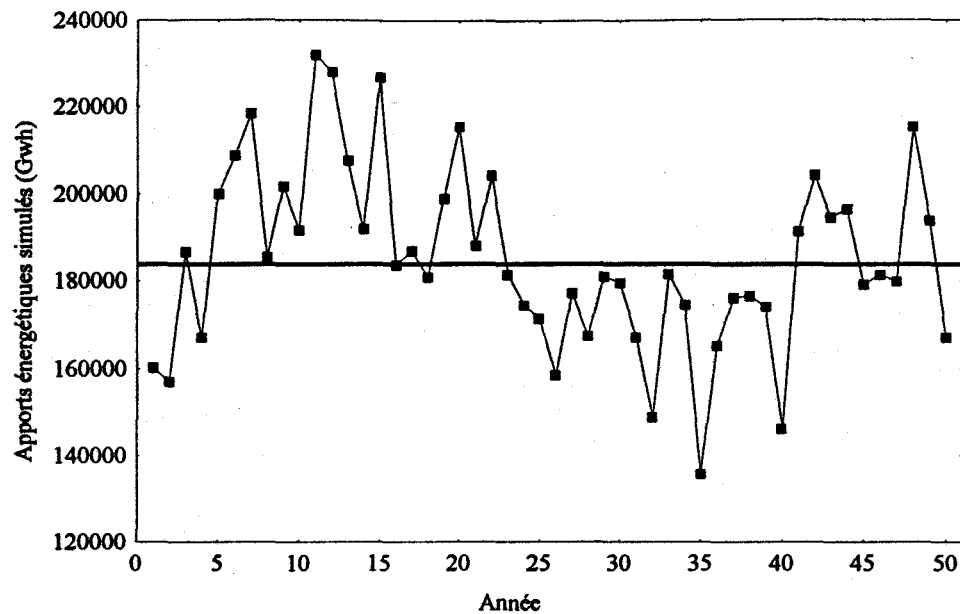


Figure 2.2.4. Série simulée à partir d'un modèle stationnaire AR(1)

#### *Nombre d'années nécessaires à la détection d'un éventuel changement de moyenne*

Mentionnons d'abord que la seule comparaison des moyennes de deux sous-échantillons ne nous permet pas de conclure à la non-stationnarité, en particulier lorsque la date de rupture est inconnue a priori (situation 3, section 2.2). L'adhésion, dans ce cas, à l'hypothèse de non-stationnarité d'un phénomène (changement climatique par exemple) nécessite plus que la simple application d'un test statistique. Il faut identifier la source du changement de moyenne et en expliquer les causes. Ainsi, on ne peut pas, compte tenu de l'information actuellement disponible, déterminer explicitement le nombre d'années nécessaires à la détection d'une non-stationnarité dans la série des apports énergétiques annuels. Toutefois, lorsqu'on dispose d'information a priori sur la date d'un éventuel changement et qu'il est possible d'expliquer cet événement, l'application d'un test statistique classique est suffisant (situations 1 et 2, section 2.2). Dans ce cas, il est possible de déterminer le nombre d'observations permettant d'appliquer la procédure avec une précision désirée (Montgomery, 1984; Cochran, 1977).

La détermination du nombre d'observations nécessaire pour détecter un changement de moyenne est intimement liée à la puissance du test statistique. La puissance d'un test est définie comme la probabilité de rejeter avec raison l'hypothèse d'égalité des moyennes et dépend à la fois de l'écart entre les moyennes et du nombre d'observations. Elle est aussi fonction du niveau de signification (probabilité de rejeter à tort l'hypothèse d'égalité des

moyennes, souvent 5%). Généralement, étant donné un écart  $\delta$  entre les moyennes, la puissance croît avec le nombre de données. Cela signifie qu'une différence fixée entre les moyennes théoriques est plus facile à détecter à l'aide d'un échantillon de plus grande taille. La fonction de puissance peut être évaluée pour la plupart des tests classiques (Student, Fisher, etc.) et permet de déterminer une taille d'échantillon optimale. Cet exercice nécessite, pour un niveau de signification donné, de choisir un écart  $\delta$  critique et de fixer la puissance désirée. L'écart critique correspond habituellement à la valeur au delà de laquelle la différence des moyennes doit être détectée avec une forte probabilité. Le choix de ces paramètres (niveau de signification, puissance et écart) n'est généralement pas effectué par le statisticien mais relève plutôt du planificateur.

### *Études concernant la non-stationnarité*

De nombreux chercheurs ont développé différentes méthodes de détection de changement de la moyenne de séries hydrologiques. La littérature sur le sujet est fort abondante. Un rapport de l'Organisation Mondiale de la Météorologie (OMM) fait une revue des diverses techniques statistiques employées en hydrologie (Cavadias, 1992).

Des travaux récents ont aussi été publiés concernant les changements climatiques qui sont susceptibles d'influencer le comportement des précipitations et par le fait même celui des apports. Citons en particulier Rodionov et Krovnin (1992), Rodionov (1994), Trenberth et Hurrell (1994), Hurrell (1995) et Hurrell et van Loon (1995). Ces auteurs ont utilisés différentes variables climatologiques (température, précipitation, pression et circulation atmosphériques) pour détecter des changements climatiques dans la région de l'Atlantique-Nord.

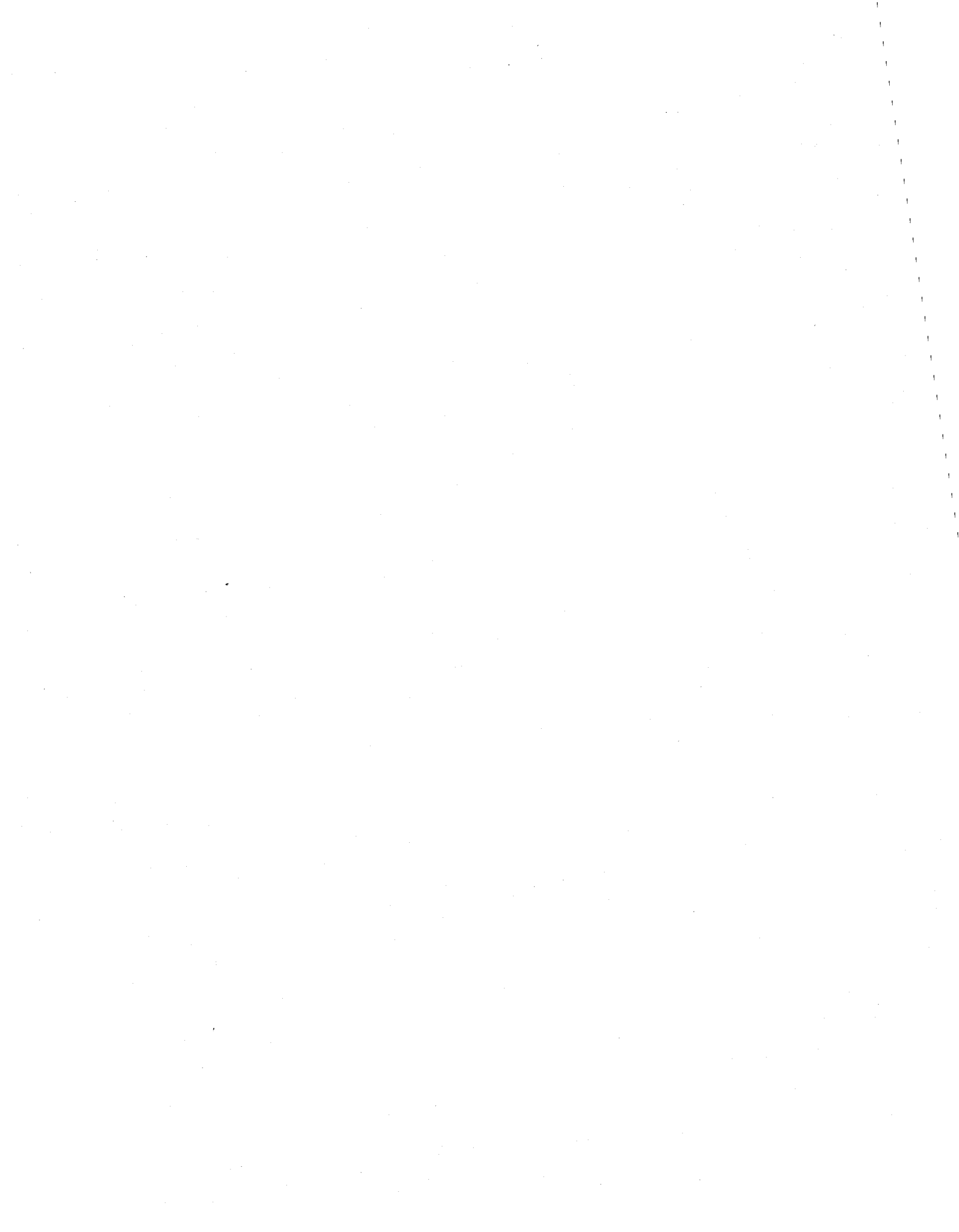
### *Cycle hydrologique*

Une série chronologique admet un cycle si les observations s'écartent de la moyenne générale selon un schéma régulier où les écarts se maintiennent sur une période donnée et changent ensuite de signe. Ce schéma est récurrent et peut par exemple être comparé à une onde sinusoïdale. Pour pouvoir affirmer statistiquement qu'un cycle est présent dans la série d'apports énergétiques annuels, il faudrait au moins observer quelques répétitions du phénomène. Or, en examinant la figure 2.2.1, il est difficile de conclure à la présence d'oscillations régulières (de durée quasi constante) des apports autour de la moyenne générale. De plus, même si la série admettait quelques cycles d'une quinzaine d'années (7 à 8 années d'écarts positifs suivis de 7 à 8 années d'écarts négatifs), le nombre d'événements

actuellement observés ne serait pas suffisant pour conclure en toute confiance. L'identification d'un cycle, dans ce cas, nécessite une justification physique ainsi qu'un nombre élevé d'observations.

Étant donné la taille limitée de l'échantillon, nous préférons conclure que la série d'apports énergétiques annuels agrégés admet plutôt des fluctuations persistantes de durée aléatoire (phénomène autocorrélé).





### **3 MODÉLISATION ET PRÉVISION DES APPORTS**

Hydro-Québec emploie actuellement un modèle AR(1) (autorégressif d'ordre 1) pour représenter la série des apports énergétiques annuels agrégés. Ce modèle est utilisé pour simuler une longue série d'écarts (jusqu'à 10 000 répétitions) par rapport à la moyenne. Les écarts cumulés sur un horizon de 1 à 10 ans sont ensuite calculés à partir de cette série. Les probabilités d'occurrence des événements ainsi simulés sont enfin déterminées pour des fins de planification énergétique. La valeur de départ de la simulation est zéro (écart initial par rapport à la moyenne générale).

Les travaux prévus au sous-projet 2.4 concernant la modélisation des apports énergétiques annuels agrégés avaient comme but l'étude de cette procédure de simulation et de certaines questions de recherche soumises par Hydro-Québec le 31 août 1995. En particulier :

- La procédure de simulation répond-elle efficacement aux objectifs de planification ?
- Le modèle AR(1) est-il adéquat ?
- Quelle est la distribution des écarts cumulés sur différents horizons ?

#### **3.1 Généralités concernant les modèles**

La modélisation de séries chronologiques à partir de modèles AR et ARMA (modèles autorégressifs et de moyenne mobile, cf. section 3.2) peut servir à des fins variées. Par exemple, l'estimation d'un modèle est un moyen économique de résumer, à partir de quelques paramètres, les propriétés statistiques de base d'une série temporelle. En effet, la modélisation permet de mieux comprendre les caractéristiques statistiques clés d'un ensemble de données. Outre cette utilité, la modélisation d'une série chronologique est employée pour deux types d'applications: la prévision et la simulation.

La prévision est une extrapolation dans le temps et vise à fournir la meilleure estimation d'une réalisation spécifique du phénomène dans le futur. Dans le cas qui nous intéresse, on pourrait par exemple prédire l'apport énergétique annuel agrégé (ou de manière équivalente l'écart par rapport à la moyenne générale) de la prochaine année. La distribution des prévisions issues d'un modèle donné (le modèle AR(1) par exemple) est entièrement connue si les résidus satisfont aux hypothèses de base (normalité, indépendance temporelle, homogénéité de la variance). Il est donc possible de déterminer analytiquement les

intervalles de confiance de ces estimations ainsi que certaines probabilités d'occurrence associées à ces prévisions (par exemple, la probabilité d'être en déficit l'année prochaine).

L'objectif général de la simulation est d'utiliser le modèle pour générer un ensemble de séries d'observations dont les propriétés statistiques reproduisent en moyenne celles de la série de données historiques. Ces séquences simulées sont aussi appelées séries synthétiques car elles ne constituent que des réalisations équiprobables pouvant éventuellement se réaliser dans le futur. La simulation est employée principalement dans trois cas:

- pour évaluer certaines caractéristiques associées aux prévisions. En effet, il est souvent difficile et même impossible de déterminer algébriquement certaines propriétés théoriques des prévisions (par exemple, la probabilité associée à une séquence de déficits consécutifs pour les prochaines années, etc.).
- pour comparer différents modèles de séries chronologiques. Dans ce cas, la simulation permet d'évaluer empiriquement les caractéristiques statistiques et ainsi d'identifier le modèle préservant le mieux, en moyenne, celles de la série originale.
- pour examiner la réaction d'un système face à une gamme de scénarios futurs possibles. On choisit généralement des scénarios simulés extrêmes. La simulation répond ici à des objectifs de planification généraux et les résultats ne peuvent être interprétés comme des indicateurs spécifiques de ce qui surviendra dans les prochaines années. Ce type d'application de la simulation est utilisée en particulier pour l'aménagement d'un nouveau bassin versant (Rasmussen *et al.*, 1996).

La prévision sert à estimer le comportement futur d'un phénomène et elle est essentielle à la planification énergétique. Elle permet d'évaluer la marge de manoeuvre à court terme qu'Hydro-Québec a pour réagir et garantir une fiabilité énergétique. La simulation est un outil complémentaire qui peut ajouter à l'information fournie par les prévisions.

### **3.2 Modélisation de la série d'apports annuels agrégés**

Que l'on effectue des prévisions ou des simulations, le modèle utilisé est le même et il doit être représentatif de la série d'apports énergétiques. Mentionnons que les modèles considérés pour représenter cette série reposent tous sur l'hypothèse de stationnarité de la moyenne.

### 3.2.1 Modèles considérés

Compte tenu de la dépendance temporelle des apports énergétiques annuels agrégés, des modèles de série chronologique ont été considérés (Box *et al.*, 1994). Après quelques analyses préliminaires, le modèle ARMA(1,1)<sup>1</sup> a été retenu pour représenter cette série. Le modèle ARMA(1,1) est un modèle mixte constitué de deux composantes: un terme autorégressif (AR) et un terme de moyenne mobile (MA). La partie autorégressive décrit comment une observation dépend directement de la donnée mesurée à la période précédente. Le terme de moyenne mobile traduit la relation existant entre une observation et l'erreur de prévision de la donnée précédente.

Pour une série chronologique  $y_1, y_2, \dots, y_t, \dots, y_n$ , le modèle ARMA(1,1) s'exprime de la manière suivante:

$$y_t = \mu + \phi \cdot (y_{t-1} - \mu) + \varepsilon_t - \theta \varepsilon_{t-1} \quad (3.1)$$

où  $\mu$  est la moyenne générale théorique du phénomène,  $\phi$  le paramètre autorégressif,  $\theta$  le paramètre de moyenne mobile et  $\varepsilon_t$  le terme d'erreur au temps  $t$  distribué selon une loi normale  $N(0, \sigma^2)$ . Le modèle AR(1), utilisé à Hydro-Québec, correspond au cas particulier ARMA(1,0) du modèle ARMA(1,1) où  $\theta = 0$ :

$$y_t = \mu + \phi \cdot (y_{t-1} - \mu) + \varepsilon_t \quad (3.2)$$

Pour appliquer un modèle ARMA(1,1), on doit estimer ses paramètres. Cela est effectué généralement à partir de la méthode classique des moindres carrés. Cette approche consiste à trouver les valeurs des paramètres  $\phi$  et  $\theta$  qui minimisent la somme des carrés des erreurs de prévision  $\varepsilon_t$ . Dans cette étude, d'autres approches ont aussi été considérées. Neuf versions différentes du modèle ARMA(1,1), correspondant à différentes valeurs de paramètres (éq. 3.1), ont été retenues pour représenter les apports énergétiques annuels agrégés. Le tableau 3.2.1 donne les valeurs des paramètres  $\phi$  et  $\theta$  correspondant à chacun des modèles.

Les méthodes d'estimations considérées (cf. tableau 3.2.1) sont décrites brièvement dans ce qui suit.

---

<sup>1</sup> Les modèles plus complexes (ARMA(2,1), etc.) ne s'ajustent pas significativement mieux aux données.

**Tableau 3.2.1** Modèles considérés pour la simulation

Méthode d'estimation	Modèle	Nomenclature	Phi	Theta
Moments modifiés	AR(1)	AR	0.410	0.00
Moindres carrés	ARMA(1,1)	ARMA1	0.498	0.156
Approche de O'Connell	ARMA(1,1)	ARMA2	0.900	0.595
Approche de O'Connell	ARMA(1,1)	ARMA3	0.920	0.507
Approche de O'Connell	ARMA(1,1)	ARMA4	0.940	0.627
Approche de O'Connell	ARMA(1,1)	ARMA5	0.960	0.667
Approche de O'Connell	ARMA(1,1)	ARMA6	0.970	0.680
Approche de O'Connell	ARMA(1,1)	ARMA7	0.980	0.696

### *Modèle AR*

Pour le cas particulier du modèle AR(1), la valeur du paramètre  $\phi$  déterminée par la méthode des moindres carrés correspond exactement au coefficient d'autocorrélation d'ordre 1 et donc à l'estimation issue de la méthode des moments (Box *et al.*, 1994). Cette valeur est ici de 0.36 pour la série d'apports énergétiques annuels agrégés (cf. tableau 2.2.1). Toutefois, il est connu que la formule usuelle du calcul de l'autocorrélation est biaisée (Quenouille, 1956). Wallis et O'Connell (1972) ont proposé une correction de biais qui, après vérification par simulation (Perreault *et al.*, 1996), donne de bons résultats. La valeur corrigée du coefficient d'autocorrélation obtenue pour la série d'apports énergétiques est de 0.41 et est utilisée comme estimation de  $\phi$  dans le modèle AR(1). Nous avons nommé cette approche "méthode des moments modifiés". Mentionnons enfin que cette correction de biais n'est valable que pour le modèle AR(1).

### *Modèle ARMA1*

Il n'existe pas de correction explicite du biais de l'autocorrélation pour le modèle ARMA(1,1). Il n'est donc pas possible de modifier analytiquement les valeurs des paramètres  $\phi$  et  $\theta$  afin de tenir compte du problème de biais dans le calcul de l'autocorrélation. La méthode classique des moindres carrés a donc été appliquée.

### *Modèles ARMA2 à ARMA7*

Le coefficient de Hurst (1951) est une statistique qui, comme l'autocorrélation, quantifie la persistance. Il est fonction du cumul des séquences de surplus et de déficits de la série (cf. Appendice A). Le coefficient de Hurst ne mesure toutefois pas la même chose que le coefficient d'autocorrélation et la communauté scientifique ne s'entend pas sur son

interprétation. En effet, le phénomène de Hurst a suscité beaucoup de controverse (Hipel et McLoed, 1980). Nous avons tout de même convenu d'inclure dans l'étude des modèles qui reproduisent cette caractéristique qui, pour la série d'apports énergétiques annuels agrégés, est de 0.81.

La méthode des moindres carrés ne permet pas d'obtenir un modèle ARMA(1,1) qui préserve le coefficient de Hurst. Cependant, O'Connell (1974) a construit, à partir de simulations, des tables permettant de déterminer des couples de paramètres  $\phi$  et  $\theta$  qui permettent de prendre en compte le coefficient de Hurst et le biais de l'autocorrélation. Les modèles ainsi obtenus préservent le coefficient de Hurst. Mentionnons que cette approche donne un poids très fort à ce coefficient et qu'il existe une infinité de couples  $\phi$  et  $\theta$  possibles. Six modèles ARMA(1,1) de ce type ont été retenus (cf. tableau 3.2.1).

### 3.2.2. Comparaison des modèles

Un modèle de prévision doit avoir des résidus  $\varepsilon_t$  qui s'apparentent à un "bruit blanc", c'est-à-dire indépendants (autocorrélation nulle) et identiquement distribués selon une loi normale de moyenne nulle. Ces hypothèses nous assurent de l'efficacité des prévisions. Les principales observations qui découlent d'une analyse complète des résidus des neuf modèles considérés sont les suivantes:

- les modèles AR et ARMA1 possèdent des résidus tout à fait compatibles avec les hypothèses;
- les six autres modèles de O'Connell (ARMA2 à ARMA7) possèdent des résidus moins compatibles avec les hypothèses. En particulier, ces résidus sont plus autocorrélés que ceux des modèles classiques (AR et ARMA1). De plus, leur moyenne s'éloigne davantage de zéro, ce qui peut engendrer un biais dans les prévisions (correspondant ici à une sous-estimation);
- l'analyse spectrale (périodogramme intégré) montre que les résidus des modèles de O'Connell s'éloignent davantage d'un bruit blanc que ceux des modèles classiques.

Nous avons mentionné, à la section 3.1, que la simulation est un outil permettant de comparer des modèles. Une telle étude comparative a été effectuée afin de vérifier si les modèles préservent adéquatement certaines caractéristiques statistiques. Les principales caractéristiques considérées sont la moyenne, l'écart-type, le coefficient d'autocorrélation et le coefficient de Hurst.

1000 séries de 50 années de données issues de chacun des modèles ont donc été générées. Les caractéristiques statistiques ont été calculées pour chaque série et la distribution de ces critères a été ensuite examinée à l'aide de diagrammes en boîte (cf. Appendice B).

Les principales conclusions sont les suivantes :

- les modèles AR et ARMA1 conduisent à des résultats similaires. Ils reproduisent en moyenne toutes les caractéristiques de la série d'apports énergétiques à l'exception du coefficient de Hurst (on obtient respectivement 0.73 et 0.72 en moyenne au lieu de 0.81);
- les modèles ARMA2 à ARMA7 donnent des résultats comparables entre eux. Ils reproduisent en moyenne toutes les caractéristiques de la série d'apports énergétiques. Toutefois, ces caractéristiques sont beaucoup plus variables que celles des modèles classiques. En particulier, certaines séries simulées admettent des apports énergétiques moyens irréalistes pour les modèles ARMA4 à ARMA7 (cf. figure 3.2.1). Ces 4 modèles ne répondent donc pas aux caractéristiques hydrologiques du réseau et ont été rejetés.

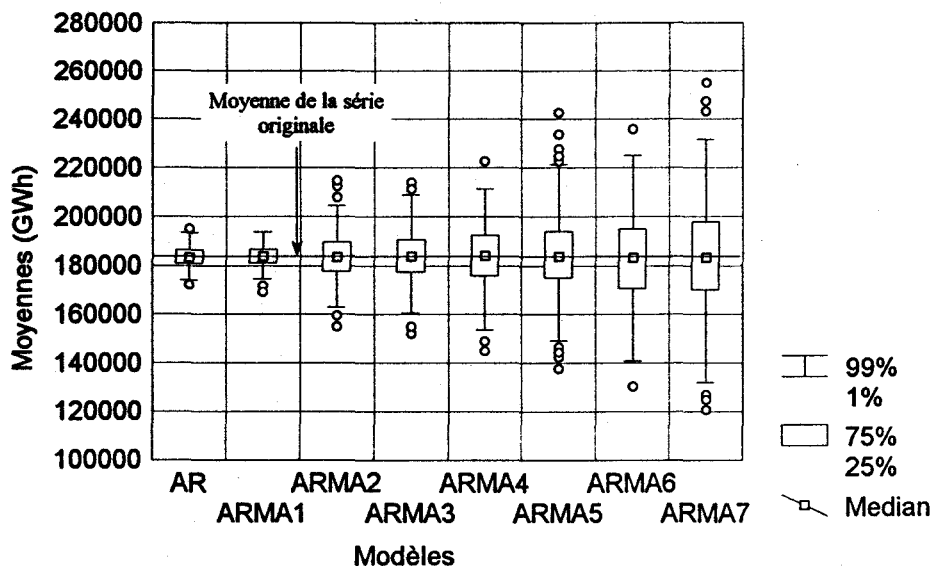
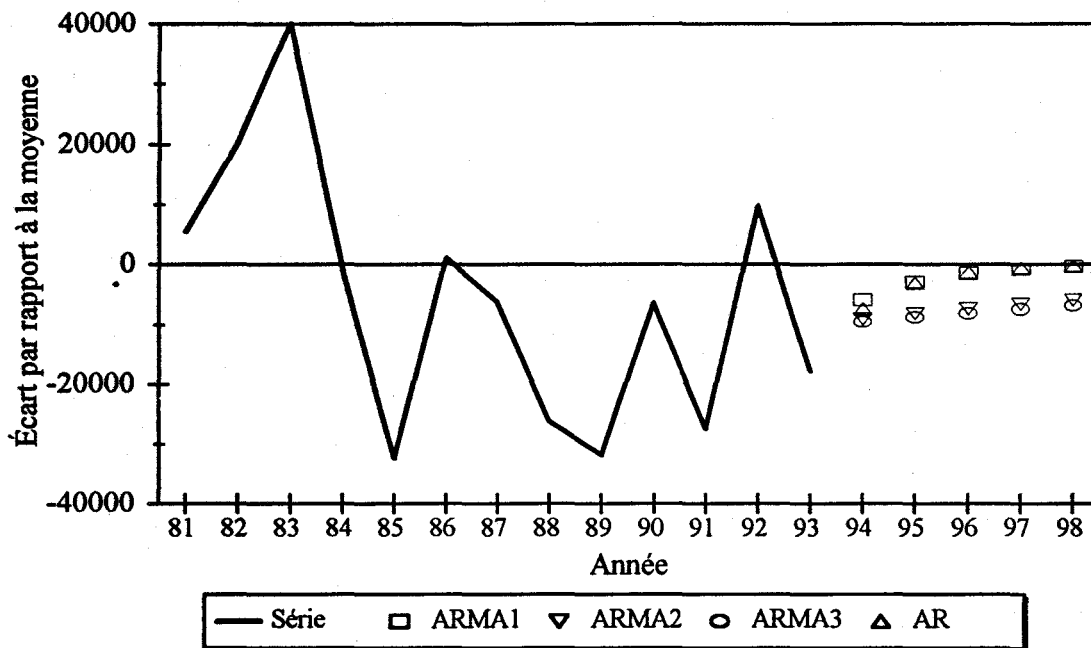


Figure 3.2.1. Distribution des moyennes des séries simulées.

Il a été convenu, dans un optique de prévision, que les modèles classiques (AR et ARMA1) sont préférables à cause du comportement de leurs résidus. Toutefois, les modèles ARMA2 et ARMA3, qui préservent le coefficient de Hurst sont aussi considérés dans la section qui suit.

### 3.3 Prévisions et distribution des écarts

Nous nous intéressons ici à la prévision ainsi qu'à la distribution des écarts énergétiques annuels agrégés par rapport à la moyenne générale. Dans la section précédente, quatre modèles ont été retenus : AR, ARMA1, ARMA2 et ARMA3. La figure 3.3.1 présente les prévisions issues de chacun des modèles pour les années 1994 à 1998. Les paramètres ont été estimés à partir de la série d'écarts énergétiques s'échelonnant de 1943 à 1993. Le tableau 3.3.1 donne les valeurs des prévisions ainsi que leur écart-type et leur intervalle de confiance au niveau de 95%.



**Figure 3.3.1.** Prévisions issues des modèles retenus pour la période 1994 à 1998 (GWh).

Comme cela a été mentionné à la section 3.1, la distribution des prévisions issues des modèles AR et ARMA est entièrement connue si les résidus satisfont aux hypothèses de base (normalité, indépendance temporelle, homogénéité de la variance). Il s'agit d'une loi normale dont la variance dépend des paramètres du modèle (Perreault *et al.*, 1996). Il est donc possible de déterminer analytiquement les probabilités d'occurrence de surplus et de déficits des prochaines années. Inversement, on peut aussi, par exemple, déterminer quel est le déficit correspondant à un risque donné (probabilité de non-dépassement dans ce cas). Le tableau 3.3.1 donne, à titre d'exemple, la probabilité d'observer un déficit et un surplus pour chaque année.



**Tableau 3.3.1.** Prévision (GWh), écart-type, intervalle de confiance et probabilité de surplus et de déficit pour les modèles retenus (1994-1998).

Modèle	Année	Prévision	Écart-type	Intervalle de confiance à 95%	Prob. de déficit	Prob. de Surplus
<b>AR</b>	1994	-7 279	17 824	[-42 214 27 656]	0.66	0.34
	1995	-2 984	19 264	[-40 742 34 773]	0.56	0.44
	1996	-1 223	19 496	[-39 435 36 988]	0.52	0.48
	1997	-502	19 534	[-38 789 37 785]	0.51	0.49
	1998	-206	19 541	[-38 505 38 094]	0.50	0.50
<b>ARMA1</b>	1994	-5 805	17 945	[-40 976 29 366]	0.63	0.37
	1995	-2 891	18 968	[-40 068 34 285]	0.56	0.44
	1996	-1 440	19 213	[-39 097 36 217]	0.53	0.47
	1997	-717	19 273	[-38 493 37 058]	0.52	0.48
	1998	-357	19 288	[-38 162 37 447]	0.51	0.49
<b>ARMA2</b>	1994	-8 967	18 604	[-45 429 27 494]	0.69	0.31
	1995	-8 077	19 449	[-46 190 30 049]	0.66	0.34
	1996	-7 263	20 108	[-46 675 32 148]	0.64	0.36
	1997	-6 537	20 627	[-46 965 33 891]	0.62	0.38
	1998	-5 884	21 037	[-47 116 35 349]	0.61	0.39
<b>ARMA3</b>	1994	-9 494	18 705	[-46 155 26 167]	0.69	0.31
	1995	-8 734	20 237	[-48 400 30 930]	0.67	0.33
	1996	-8 036	21 449	[-50 076 34 004]	0.65	0.35
	1997	-7 392	22 424	[-51 343 36 557]	0.63	0.37
	1998	-6 801	23 216	[-52 305 38 702]	0.62	0.38

Puisque la loi des prévisions est entièrement connue, il est aussi possible de déterminer la distribution des écarts cumulés des prochaines années. Il s'agit d'une loi normale dont la variance dépend des paramètres du modèles ainsi que des corrélations entre les prévisions. On peut donc calculer analytiquement les probabilités associées à ce type d'événement. La méthodologie permettant d'effectuer ces calculs est décrite en détail dans Perreault *et al.* (1996).

## 3.4 Conclusion

### *Procédure de simulation actuellement utilisée*

La procédure utilisée actuellement à Hydro-Québec pour déterminer la distribution des écarts énergétiques repose entièrement sur une simulation dont la valeur de départ est un écart nul. Telle qu'appliquée, la simulation ne donne aucune information concernant les écarts pouvant survenir dans le futur à une date spécifique. D'une part, les données simulées ne tiennent pas compte du dernier événement de la série malgré le fait qu'un écart énergétique est fortement influencé par l'événement de l'année précédente (cf. section 2.2). En effet, la valeur initiale de la procédure est fixée à zéro. D'autre part, l'approche consiste à simuler une longue série que l'on partitionne ensuite pour déterminer les probabilités. Ceci peut induire un biais dans les résultats (Hipel et McLeod, 1994).

Nous recommandons donc l'emploi de prévisions qui permettent d'estimer l'écart énergétique susceptible de se produire, avec la plus grande probabilité, à une période donnée dans le futur. L'incertitude d'échantillonnage (variance et intervalle de confiance) peut être évaluée analytiquement. Cette approche permet aussi de calculer aisément la plupart des événements ou probabilités qui intéressent Hydro-Québec (cf. section 3.3).

### *Le modèle AR(1)*

Compte-tenu des résultats présentés au chapitre 2 et aux sections 3.2 et 3.3, nous recommandons l'emploi du modèle stationnaire AR(1) pour des fins de prévision et de simulation. Ce modèle, qui donne des résultats très comparables au modèle ARMA(1,1), a l'avantage d'être très simple d'utilisation. Toutefois, nous ne rejetons pas les modèles ARMA2 et ARMA3 qui admettent une gamme de scénarios plus vaste (Perreault *et al.*, 1996). Ils permettent donc de générer des scénarios d'apports ou d'écarts énergétiques plus extrêmes qui peuvent être ensuite soumis comme intrant à des modèles de planification. Nous croyons, par contre, que l'emploi de tels modèles nécessitent une étude approfondie du phénomène de Hurst qui ne fait pas l'unanimité dans la communauté scientifique.

Nous recommandons aussi d'utiliser la formule de correction du biais (cf. section 3.2.1 et Perreault *et al.*, 1996) pour le calcul du coefficient d'autocorrélation et la détermination du paramètre du modèle AR(1). Finalement, la calibration du modèle doit être mise à jour à mesure qu'une nouvelle observation devient disponible.

***Distribution des écarts***

Les propriétés théoriques des prévisions issues des modèles considérés sont bien connues (Perreault *et al.*, 1996) ce qui permet de déterminer analytiquement les probabilités d'événements spécifiques plus complexes (cumul des écarts, par exemple). Toutefois, si on s'intéresse particulièrement aux durées des séquences consécutives de déficits ou de surplus, il n'existe pas à notre connaissance d'équations analytiques directes. On peut alors avoir recours à une simulation dont la valeur initiale doit correspondre à l'écart observé la dernière année.

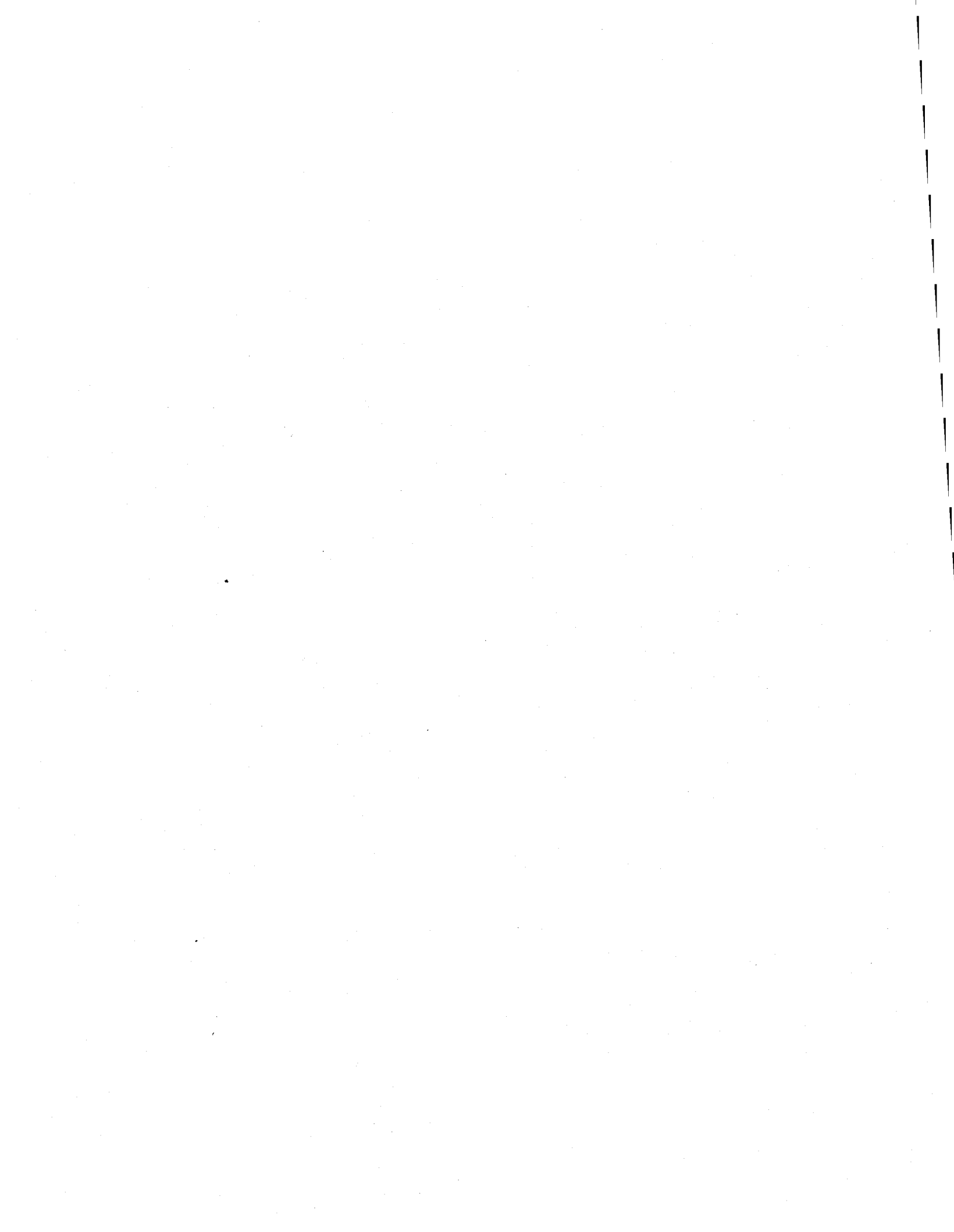
## **RÉFÉRENCES BIBLIOGRAPHIQUES**

---

- Bernier, J. 1994. Statistical detection of changes in geophysical series *dans* Engineering Risk in Natural Resources Management, 159-176, L.Duckstein and E. Parent (eds).
- Box, G.E.P., Jenkins, G.M. et Reinsel, G.C. 1994. Time Series Analysis. Prentice Hall, Englewood.
- Bruneau, P. 1984. Validation de longues séries de débits observés. Rapport interne, No 08741-RA-84/45, Hydro-Québec.
- Bruneau, P. 1986a. Étude sur la productibilité: Séries de 50 années de débits aux sites. Rapport interne, No 08741-RA-86/26, Hydro-Québec.
- Bruneau, P. 1986b. Validation de longues séries de débits observés: Phase II. Rapport interne, No 08741-RA-86/45, Hydro-Québec.
- Bruneau, P. 1986c. Validation de longues séries de débits observés: Phase II - La Grande. Rapport interne, No 08741-RA-86/45, Hydro-Québec.
- Cavadias, G. 1992. A survey of current approaches to modelling of hydrological time-series with respect to climate variability and change. World Meteorological Organization, WMO/TD-No. 534.
- Cochran, W.G. 1977. Sampling Techniques. Wiley, New York.
- Feller, W. 1951. The asymptotic distribution of the range of sums of independent random variables, *Ann. Math. Stat.*, 22, 427-432
- Findlay *et al.* 1994. Trends 93: Compendium of data on global change. ORNL.
- Hipel, K.W. et McLeod, A.I. 1980. Perspectives in stochastic hydrology présenté dans *Time Series*. Anderson ed., North-Holland, 73-102.
- Hipel, K.W. et McLeod, A.I. 1994. *Time Series Modelling of Water Resources and Environmental System*. Elsevier, Amsterdam.
- Hurrell, J.W. 1995. Decadal trends in the North Atlantic oscillation: regional temperatures and precipitation. *Science*, 269, 676-679.

- Hurrell, J.W. et van Loon, H. 1995. Analysis of low-frequency climate variations over the north atlantic using historical atmospheric data. Sixth Symposium on Global Change Studies, Dallas, Texas, 15-20.
- Hurst, H.E. 1951. Long-term storage capacity of reservoirs. Transactions of the American Society of Civil Engineers, 116, 770-808.
- Karl *et al.* 1994. Trends 93: Compendium of data on global change. ORNL.
- Lee, A.F.S. et Heghinian, S.M. 1977. A shift of the mean level in a sequence of independent normal random variables: a Bayesian approach. Technometrics, 19(4), 503-506.
- Mathier, L., Fagherazzi, L., Rassam, J.C. et Bobée, B. 1992. Great Lakes net basin supply simulation by a stochastic approach. INRS-Eau, rapport de recherche no 362.
- McLeod, A.I. et Hipel, K.W. 1978. Preservation of the rescaled adjusted range, 1, A reassessment of the Hurst phenomenon, Water Resour. Res., 14(3), 491-508.
- Montgomery, D.C. 1984. Design and Analysis of Experiments. Wiley, New York.
- Moran, P.A.R. (1959) The Theory of Storage, Methuen, London.
- O'Connell, P.E. 1974. Stochastic modelling of long-term persistence in stream flow sequences. Thèse de doctorat, Département de Génie Civil, Imperial College, Londres.
- Perreault, L., Haché, M. et Bobée, B. 1996. Rapport Final du Projet C2 - Analyse statistique des séries d'apports énergétiques. INRS-Eau, rapport de recherche en préparation.
- Quenouille, M.H. 1956. Notes in bias estimation. Biometrika, 43, 353-360.
- Rasmussen, P.F., Fagherazzi, L., Duquette, R., Legendre, P. et Perreault, L. 1996. Détermination de séries mensuelles types pour le modèle MINERVE. INRS-Eau, rapport de recherche en préparation.
- Rodionov, S.N. et Krovnin, A.S. 1992. The 1980s in the context of climatic changes in the North Atlantic region. ICES Sciences Symposium, 195, 93-102.
- Rodionov, S.N. 1994. Association between winter precipitation and water level fluctuation in the Great Lakes and atmospheric circulation patterns. American Meteorological Society, 7, 1693-1706.

- 
- Salas, J.D., Delleur, J.W., Yevjevich, V. et Lane, W.L. 1980. Applied Modeling of Hydrological Time Series. Water Resources Publications, Chelsea, Michigan.
- Trenberth, K.E. et Hurrell, J.W. 1994. Decadal atmosphere-ocean variations in the Pacific. *Climate Dynamics*, 9, 303-319.
- Wallis, J.R. et O'Connell, P.E. 1972. Small sample estimation of  $\rho_1$ . *Water Resources Research*, 8, 707-712.



# APPENDICE A : COEFFICIENT DE HURST

---

## A.1 Quelques définitions

### *Sommes partielles*

Considérons une série chronologique,  $y_1, y_2, \dots, y_n$ . La séquence des sommes partielles ajustées est définie de la manière suivante:

$$S_k^* = \sum_{i=1}^k y_i - k\bar{y}_n \quad k = 1, 2, \dots, n$$

où  $\bar{y}_n = \frac{1}{n} \sum_{i=1}^n y_i$  est la moyenne de la série. En particulier, on a que  $S_n^* = 0$ .

Pour caractériser une série donnée, on peut considérer le maximum et le minimum des  $S_k^*$ :

$$M_n^* = \max(0, S_1^*, S_2^*, \dots, S_n^*)$$

$$m_n^* = \min(0, S_1^*, S_2^*, \dots, S_n^*)$$

### *Domaine de variation*

Le domaine de variation des sommes partielles ajustées est donné par:

$$R_n^* = \max_{k \in [0, n]} S_k^* - \min_{k \in [0, n]} S_k^* = M_n^* - m_n^*$$

Cette caractéristique statistique peut être interprétée de façon physique. Par exemple,  $R_n^*$  peut représenter le stockage nécessaire d'un réservoir si l'on veut exploiter le volume total d'eau à un taux constant. La figure A.1 illustre les définitions de la somme partielle ajustée et de son domaine de variation.

### *Domaine de variation standardisé*

Le domaine de variation standardisé est défini de la façon suivante:

$$\bar{R}_n^* = \frac{R_n^*}{\hat{\sigma}_n^2}$$

où  $\hat{\sigma}_n^2 = n^{-1} \sum_{i=1}^n (X_i - \bar{X}_i)^2$  est l'estimation de la variance de la série. Les aspects théoriques de cette caractéristique statistique ont été largement étudiés en hydrologie (McLeod et Hipel, 1978; Salas et al., 1980).



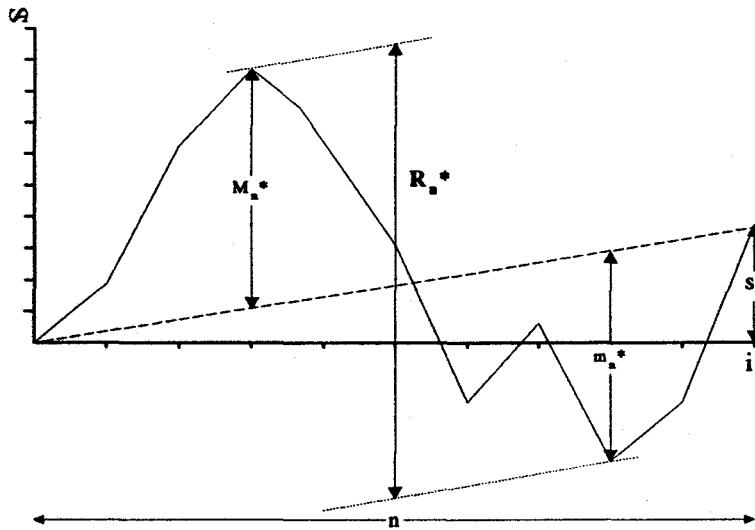


Figure A.1. Définitions des sommes partielles ajustées et du domaine de variation ajusté

## A.2 Coefficient de Hurst

Suite à des études expérimentales intensives, Hurst (1951) a identifié une relation non-linéaire entre  $\bar{R}_n^*$  et la taille d'échantillon  $n$  qu'il a approximé par:

$$\bar{R}_n^* = \frac{R_n^*}{\hat{\sigma}_n^2} = k n^h$$

où  $k$  et  $h$  sont des constantes. On peut montrer que pour  $n=2$ ,  $\bar{R}_n^* = 1$ . Pour cette raison, Hurst a simplifié son expression de la manière suivante:

$$\bar{R}_n^* = (n/2)^H$$

où  $H$  est le coefficient de Hurst. Il a trouvé que pour une série d'éléments indépendants et pour  $n$  grand, l'espérance du coefficient  $H$  est égale à 0.5. Cependant, pour les séries réelles examinées par Hurst, la moyenne obtenue du coefficient  $H$  était de 0.73. Cette observation a été appelée le phénomène de Hurst. Feller (1951) a suggéré qu'un modèle markovien pourrait expliquer le phénomène, mais Moran (1959) a pu démontrer que même les processus autocorrélés ont asymptotiquement une valeur de  $H$  égale à 0.5. En raison du lien entre  $H$  et le domaine de variation des sommes ajustées et le besoin de stockage, on a traditionnellement accordé beaucoup d'importance au coefficient de Hurst en hydrologie. En pratique, on estime  $H$  comme suit:

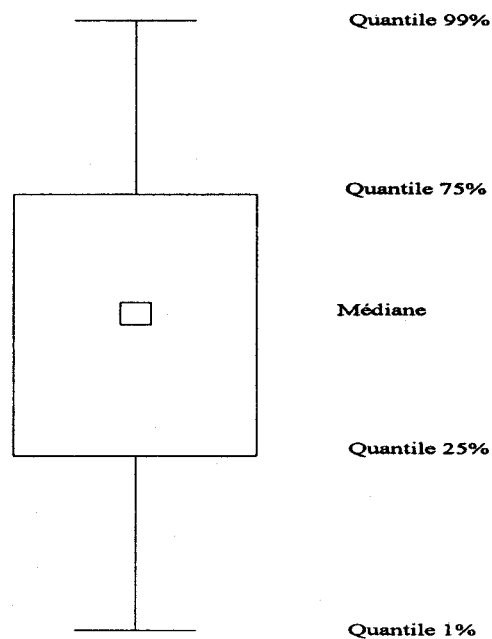
$$H = \frac{\log \bar{R}_n^*}{\log(n/2)}$$

## **APPENDICE B : DIAGRAMME EN BOÎTE**

---

Le diagramme en boîte (boxplot) est une représentation schématique de la forme de la distribution des observations. Cette méthode graphique, au même titre qu'un histogramme, est très utile pour comparer visuellement les caractéristiques statistiques de plusieurs séries. Il existe plusieurs types de diagrammes en boîte. Ceux qui sont présentés à la figure 3.1.1 du présent rapport illustrent la distribution des 1000 moyennes des séries simulées et sont construits de la manière suivante (cf. figure B.1):

- le point central représente la médiane des moyennes obtenues;
- les extrémités inférieure et supérieure de la boîte représentent respectivement le 25ème et 75ème quantiles des moyennes obtenues (valeurs correspondant à 25% et 75% de probabilité au non-dépassement). Ainsi, 50% des moyennes obtenues se situent à l'intérieur de cette boîte;
- les extrémités des tiraits représentent les quantiles de 1% et 99% de probabilité au non-dépassement.



**Figure B.1.** Diagramme en boîte.