

Record Number: 14630
Author, Monographic: Rasmussen, P. F.//Haché, M.//Fagherazzi, L.//Duquette, R.//Legendre, P.
Author Role:
Title, Monographic: Détermination de séries mensuelles types pour le modèle MINERVE (phase 2)
Translated Title:
Reprint Status:
Edition:
Author, Subsidiary:
Author Role:
Place of Publication: Québec
Publisher Name: INRS-Eau
Date of Publication: 1998
Original Publication Date: Mars 1998
Volume Identification:
Extent of Work: iv, 47
Packaging Method: pages
Series Editor:
Series Editor Role:
Series Title: INRS-Eau, rapport de recherche
Series Volume ID: 506
Location/URL:
ISBN: 2-89146-389-7
Notes: Rapport annuel 1997-1998
Abstract: ISBN double avec le R000506 b
chaire en hydrologie statistique
Call Number: R000506
Keywords: rapport/ ok/ dl

**DÉTERMINATION DE SÉRIES
MENSUELLES TYPES POUR
LE MODÈLE MINERVE
(Phase 2)**

**Détermination de séries mensuelles types
pour le modèle MINERVE
(Phase 2)**

RAPPORT PRÉPARÉ PAR

**Peter F. Rasmussen
Mario Haché**

En collaboration avec

**Laura Fagherazzi
Robin Duquette
Pierre Legendre**

Chaire Hydro-Québec / CRSNG en hydrologie statistique
Institut national de la recherche scientifique, INRS-Eau
2800, rue Einstein, Case postale 7500, Sainte-Foy (Québec), G1V 4C7

Rapport de recherche No R-506
Mars 1998

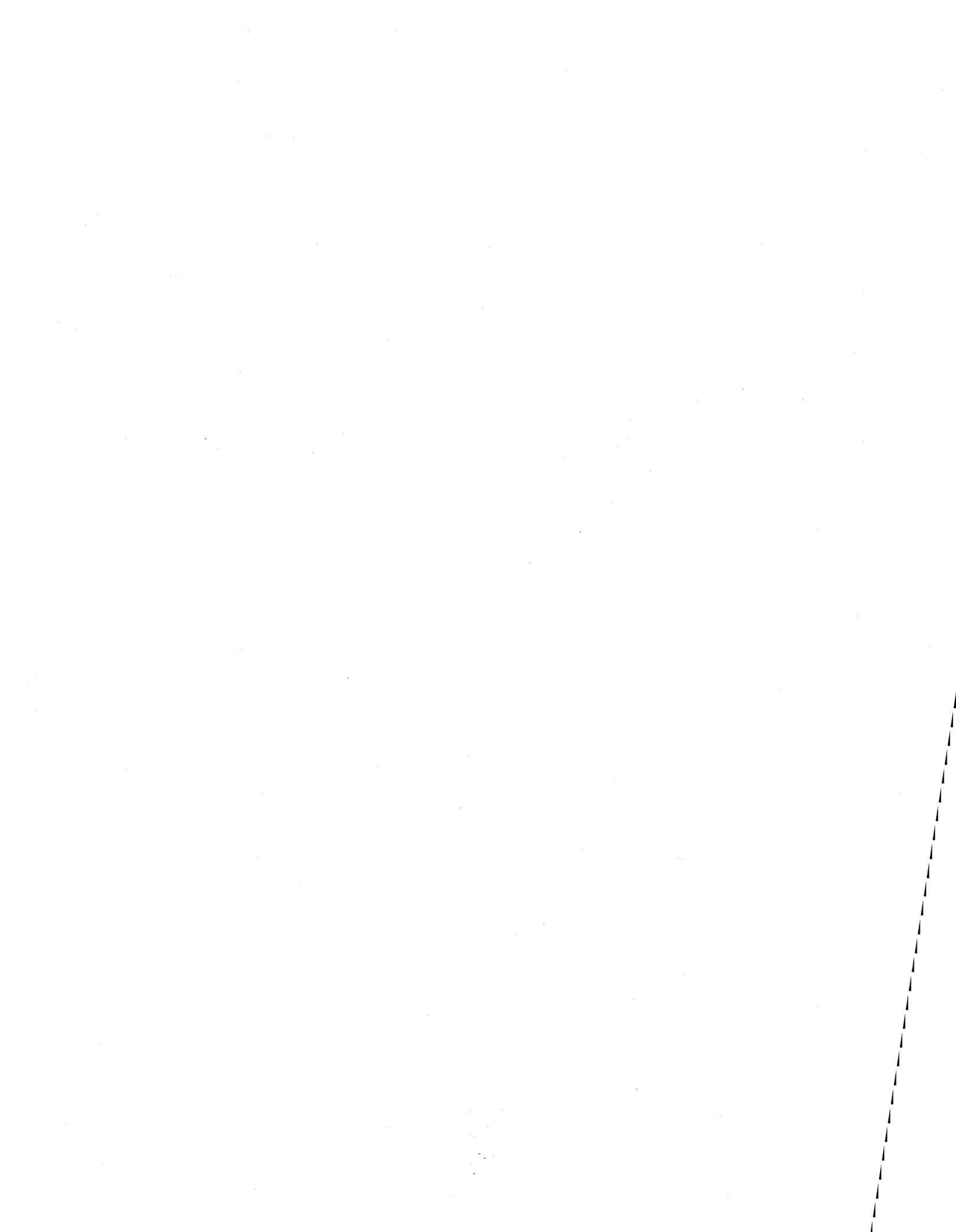
Avant-propos

Le présent rapport résume les résultats de la deuxième phase du projet intitulé "Détermination de séries mensuelles types pour le modèle MINERVE". Les résultats découlant de la première phase du projet sont présentés dans le Rapport 1 (1996). Ensemble, ces deux rapports décrivent les détails de chacune des étapes que l'on doit considérer pour effectuer un choix rationnel des séries types d'apport mensuel à considérer dans la nouvelle version de MINERVE.

La recherche présentée dans ce rapport a été réalisée par Peter Rasmussen et Mario Haché, membres de la Chaire en hydrologie statistique à l'INRS-Eau. Le travail représente néanmoins un effort collectif, car différentes personnes d'Hydro-Québec, en particulier Pierre Legendre, Laura Fagherazzi et Robin Duquette, ont contribué de manière significative à préciser les objectifs spécifiques du projet et à discuter les différents éléments de la méthodologie proposée.

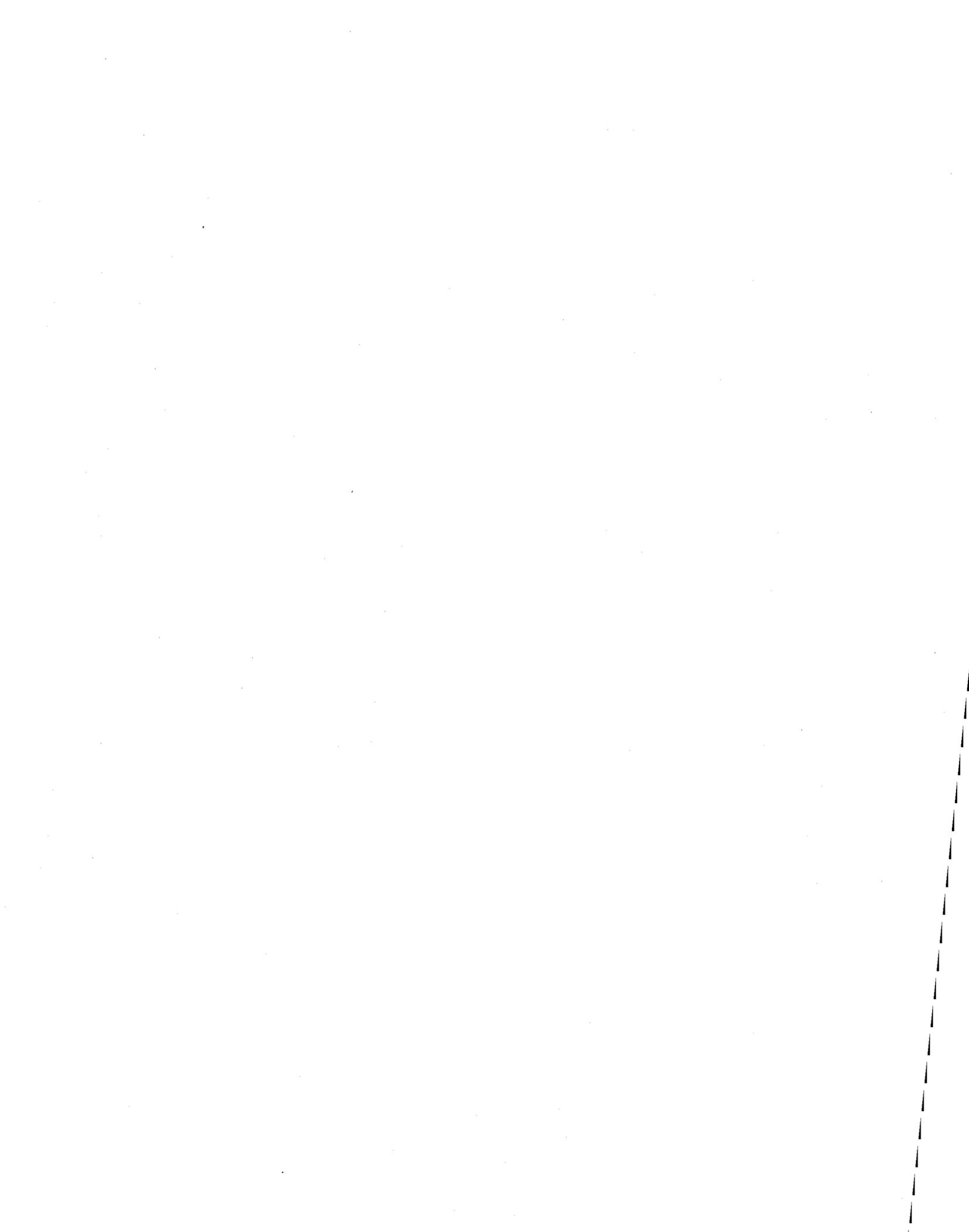
Table des matières

1 INTRODUCTION	1
2 FONCTION OBJECTIF DE MINERVE96	5
3 DÉTERMINATION DE LA RELATION FACTEUR/COÛT	9
3.1 INTRODUCTION	9
3.2 RÉDUCTION DE LA DIMENSIONNALITÉ DES APPORTS	9
3.3 ANALYSE DE SENSIBILITÉ DES FACTEURS	12
4 EXEMPLE D'ANALYSE DE SENSIBILITÉ	19
4.1 DESCRIPTION DU SYSTÈME HYDRIQUE DE LA RIVIÈRE GRANDE BALEINE.....	19
4.2 GÉNÉRATION DE SÉRIES SYNTHÉTIQUES	19
4.2.1 <i>Sélection des critères</i>	20
4.2.2 <i>Choix du type de séries</i>	20
4.2.3 <i>Génération - phase I</i>	21
4.2.4 <i>Analyse en composantes principales</i>	23
4.2.5 <i>Génération - phase II</i>	27
4.2.6 <i>Désagrégation des séries d'apport annuel agrégé</i>	27
4.3 RELATION ENTRE LES FACTEURS ET LE COÛT D'EXPLOITATION	28
4.3.1 <i>Détermination du coût d'exploitation</i>	28
4.3.2 <i>Relation coût/facteur</i>	29
5 EXEMPLE D'APPLICATION DE MINERVE96	33
5.1 CHOIX DES SCÉNARIOS TYPES POUR MINERVE96	33
5.2 RÉSULTATS DU MODÈLE MINERVE96	33
6 CONCLUSION	37
7 BIBLIOGRAPHIE	39



Liste des figures

3.1	Sélection des scénarios pour l'analyse de sensibilité.....	14
3.2	Relation entre coût optimal en fonction des scénarios et coût associé à un aménagement donné.....	15
3.3	Exemples de la variation du coût optimal en fonction de trois facteur.....	16
3.4	Illustration du choix des scénarios pour l'analyse du sensibilité.....	17
4.1	Illustration schématique du complexe de Grande Baleine	19
4.2	Position des critères dans l'espace des deux premiers facteurs pour les séries d'apport naturel agrégé (type 1) et les séries d'apport énergétique agrégé (type 3).....	22
4.3	Position des critères dans l'espace des deux premiers facteurs pour les séries d'apport naturel annuel agrégé (type 1) et les séries d'apport naturel à chacun des sites (type 2)	22
4.4	Position des critères dans l'espace des quatre premiers facteurs.....	25-26
4.5a	Coût d'exploitation en fonction du facteur 1	30
4.5b	Coût d'exploitation en fonction du facteur 2.....	31
4.5c	Coût d'exploitation en fonction du facteur 3	31
4.5d	Coût d'exploitation en fonction du facteur 4.....	32
5.1	Position des séries synthétiques choisies (+) dans l'espace des facteurs 1 et 4.....	34



Liste des tableaux

4.1	Caractéristiques statistiques des critères pour les séries synthétiques et pour la série historique	23
4.2	Matrice de corrélation des critères	24
4.3	Coordonnées des critères sur les axes factoriels	24
4.4	Coûts d'exploitation (en millions) en fonction des valeurs initiales et des facteurs	29
5.1	Coût d'exploitation (en millions) pour les neuf séries types avec le modèle déterministe	34
5.2	Coût d'exploitation déterminé à partir de MINERVE96	35

1 Introduction

Le présent rapport fait suite aux travaux de Rasmussen *et al.* (1996) visant la détermination de séries mensuelles types à utiliser dans la nouvelle version du logiciel MINERVE¹.

On rappelle que MINERVE est un logiciel d'optimisation qui permet de déterminer l'aménagement hydroélectrique optimal d'une nouvelle vallée, c'est-à-dire la localisation des centrales, la taille des réservoirs, les hauteurs de chute, etc. L'objectif principal des modifications apportées au logiciel MINERVE au cours de la période 1995-96 était d'introduire l'aspect stochastique dans l'optimisation. Dans la version originale de MINERVE, seuls les apports historiques étaient utilisés dans l'optimisation (Duquette et al., 1992). De plus, certaines hypothèses reliées à la gestion des réservoirs pouvaient comporter des lacunes. Dans la nouvelle version, la plupart de ces problèmes ont été réglés.

L'aspect stochastique est introduit dans MINERVE96 par la considération de plusieurs scénarios d'apport mensuel dans la fonction objectif. De manière générale, on cherche un aménagement qui est optimal "en moyenne", c'est-à-dire qui assure une bonne rentabilité pour une gamme de scénarios d'apports futurs. Ce principe d'optimisation est appelé "optimisation par scénarios". Il est évidemment important que les scénarios d'apport considérés soient réalistes du point de vue hydrologique. Rasmussen *et al.* (1996) proposent de générer des apports synthétiques à l'aide d'un modèle stochastique. Le modèle stochastique doit être calibré de façon à préserver les principales caractéristiques statistiques des données historiques. Ainsi, on peut s'assurer que les apports générés sont réalistes du point de vue hydrologique et statistique. À l'aide du modèle stochastique, on peut ensuite générer un grand nombre de séries d'apports synthétiques. Cependant, en raison du temps de calcul, on ne peut considérer qu'un nombre limité de scénarios (environ dix scénarios) dans MINERVE96. Il faut donc choisir, parmi les scénarios générés à l'aide du modèle stochastique, quelques-uns qui représentent une gamme raisonnable d'apports futurs incluant des

¹ Dans ce rapport, la nouvelle version de MINERVE sera appelée MINERVE96.

scénarios "pessimistes" et des scénarios "optimistes". Le choix de scénarios types pourrait être très important pour l'aménagement optimal. Il importe donc de les choisir de manière rationnelle.

Le problème du choix de scénarios est généralement très complexe, car souvent il y a plusieurs sites d'intérêt et il faut alors considérer des apports multivariés correspondant aux différents sites. De plus, on s'intéresse généralement à des apports mensuels ou trimestriels. Le modèle stochastique devrait préserver l'autocorrélation ainsi que la corrélation croisée entre les différents sites ce qui assure la cohérence temporelle et spatiale des apports. Cependant, pour la discrimination entre les scénarios, il est difficile, voire impossible, de tenir compte de tous les aspects des apports. Dépendant du cas d'intérêt, il existe plusieurs façons de simplifier l'analyse de discrimination. Par exemple, dans le cas où on s'intéresse à des réservoirs interannuels, on propose, pour chaque scénario généré, d'effectuer une agrégation temporelle et spatiale des apports et de considérer uniquement la série agrégée lors de la discrimination des scénarios. Rasmussen *et al.* (1996) décrivent plusieurs façons d'obtenir une série annuelle agrégée spatialement qui est représentative du scénario d'apport associé. C'est à partir des séries annuelles agrégées que la discrimination et le choix des scénarios se font. Cependant, lors de l'optimisation dans MINERVE, ce sont les apports saisonniers (mensuels ou trimestriels) et spatiaux qui sont considérés.

Pour permettre la discrimination, c'est-à-dire l'identification des scénarios pessimistes et optimistes, il faut nécessairement adopter des critères. Dans le contexte actuel, les critères doivent être des caractéristiques statistiques des séries d'apport, par exemple la moyenne sur une période donnée, le minimum, le déficit accumulé, etc. Les critères doivent être choisis en fonction de l'objectif de l'étude, c'est-à-dire l'aménagement hydroélectrique optimal d'une vallée. Or, il peut être difficile de savoir *a priori* quels sont les critères pertinents pour la discrimination des scénarios et encore plus difficile de les classer *a priori* selon leur degré de pertinence. De plus, les différentes caractéristiques statistiques retenues pour l'étude peuvent être corrélées, ce qui pourrait fausser l'interprétation de leur importance. Dans le rapport 1, on propose de commencer l'étude avec tous les critères qui pourraient avoir un intérêt pour le problème d'aménagement optimal. À l'aide d'une analyse en composantes principales, ces critères peuvent être réduits en un nombre limité de facteurs indépendants qui reflètent différents aspects des séries annuelles agrégées. L'analyse factorielle permet également de visualiser les corrélations qui existent entre les différents critères.

Dans une application préliminaire, nous avons proposé de diviser l'espace factoriel en classes équiprobables et de choisir un scénario représentatif dans chaque classe. Les scénarios d'apport

ainsi sélectionnés constituent des séries types d'apport mensuel multivarié. Cependant, le choix de classes dans l'application préliminaire était assez arbitraire. On a divisé l'espace des deux premiers facteurs en neuf carreaux équiprobables, mais ce choix n'est pas nécessairement optimal. Les facteurs obtenus par l'analyse en composantes principales reflètent la corrélation entre les critères; cependant, on ne peut pas *a priori* associé une pertinence aux facteurs vis-à-vis le problème d'optimisation de l'aménagement hydroélectrique. Pour ce faire, il faut effectuer une analyse de sensibilité.

Dans le présent rapport, nous présentons les étapes requises pour choisir d'une manière rationnelle des scénarios types. On insiste sur l'aspect méthodologique, car chaque étude est unique et demande ses propres considérations. Cependant, pour illustrer la méthodologie, on considère un exemple d'application avec l'aménagement de la rivière Grande Baleine.

2 Fonction objectif de MINERVE96

Avant de décrire la méthodologie d'analyse de sensibilité dont le but est de permettre un choix rationnel de séries types, il sera utile de rappeler les principaux termes de la fonction objectif de MINERVE96. Soit \mathbf{A} un vecteur ayant comme éléments les dimensions particulières d'un aménagement donné (taille des réservoirs, hauteur des chutes, puissance installée, etc.). Soit $\xi(\mathbf{A}, S)$ la fonction de coût, c'est-à-dire le coût associé à l'aménagement \mathbf{A} , si le scénario d'apport S se produit dans le futur. Plus particulièrement, on peut écrire la fonction de coût comme:

$$\xi(\mathbf{A}, S) = \text{Coût}(\mathbf{A}) + \sum_{t \in T} (cD_{t,S} - rE_{t,S}) \quad (1)$$

Le premier terme du membre droit de cette expression est le coût de *construction* associé à l'aménagement \mathbf{A} . Le deuxième terme du membre droit de l'équation est le coût d'*exploitation*. La quantité $D_{t,S}$ est le déficit en énergie au cours de la période t (mois, trimestre) correspondant au scénario S et c est le prix marginal d'achat d'énergie. Le terme $cD_{t,S}$ est donc le coût total d'achat d'énergie au cours de la période t . La quantité $E_{t,S}$ représente le surplus d'énergie au cours de la période t (mois, trimestre) correspondant au scénario S . Ce surplus pourrait être vendu à un prix marginal de r et $rE_{t,S}$ représente donc le revenu total de la vente réalisée au cours de la période t . Le coût d'achat et le revenu de vente sont agrégés par la sommation de toutes les périodes considérées.

On peut remarquer que la permission de vendre et d'acheter de l'énergie constitue une nouveauté majeure par rapport à l'ancienne version de MINERVE où le prix d'achat est considéré infini et le prix de vente est considéré nul. L'introduction de la possibilité d'acheter et de vendre est nécessaire dans MINERVE96, car certains scénarios extrêmes pourraient ne pas permettre de satisfaire la demande prioritaire. Bien que la probabilité que des scénarios "extrêmes" se produisent est faible, ce sont des scénarios possibles et il est important de pouvoir les considérer dans le modèle. Les prix marginaux d'achat et de vente doivent être les plus réalistes possibles et leur détermination nécessite généralement une étude approfondie du marché énergétique futur. On peut s'attendre à ce que les prix marginaux dépendent du scénario. Par exemple, le prix marginal associé à un scénario

de très faible hydraulicité qui nécessite l'achat d'une quantité d'énergie élevée pourrait être nettement supérieur à un autre scénario où seulement une faible quantité doit être achetée. En effet, plus l'achat d'énergie est grande, plus le prix augmente. Inversement, plus la vente d'énergie est grande, plus le prix diminue. MINERVE96 permet d'associer différents prix marginaux aux différents scénarios considérés (Duquette, 1996). Alternativement, on peut choisir de considérer des prix marginaux qui ne dépendent pas des scénarios en décourageant l'achat d'énergie par un prix marginal élevé et la vente d'énergie par un faible prix marginal de vente. Dans le dernier cas, le choix de prix marginaux sera généralement subjectif et pourrait influencer grandement l'aménagement optimal. Cependant, cette procédure constitue une alternative simple et sera utile dans l'analyse de sensibilité puisque la quantité d'énergie qui doit être achetée ou qui peut être vendue n'est pas connue *a priori*.

Dans l'ancienne version de MINERVE, on détermine l'aménagement A qui minimise $\xi(A, S_{his})$ où S_{his} est le scénario historique, c'est-à-dire un vecteur contenant les apports observés. Dans la nouvelle version, on minimise plutôt le coût espéré:

$$E[\xi(A)] = \int_S \xi(A, S) f(S) dS \quad (2)$$

où $E[]$ signifie l'espérance mathématique et $f(S)$ est la fonction de probabilité de S . L'intégration se fait sur l'ensemble S des scénarios d'apport possibles. Soulignons la multidimensionnalité associée à un scénario S . Par exemple, si dans l'étude de la rivière Grande Baleine, on considère les apports mensuels à cinq sites sur un horizon de 30 ans, alors la dimension de S est 1800 ($5 \times 12 \times 30$). Il est évidemment impossible de calculer une intégrale de cette dimension. On pourrait dans ce cas considérer une approximation par une sommation finie:

$$E[\xi(A)] \approx \sum \xi(A, S_i) p_i \quad (3)$$

où S_i est un scénario donné et p_i est la probabilité discrète que l'on associe au scénario S_i . L'approximation d'un intégral par une sommation finie correspond à l'approximation d'une distribution de probabilité continue par un histogramme. On souligne d'ailleurs que les intervalles de l'histogramme ne sont pas tenus à être équiprobables. En effet, aux endroits où la distribution change beaucoup, il est préférable de considérer une graduation fine des classes alors que là où la distribution ne change que peu, on peut considérer des classes plus grossières. A la limite, pour une distribution uniforme, on peut considérer une seule classe. Si le but est de déterminer l'espérance d'une fonction de la variable aléatoire considérée, il faut également tenir compte du

comportement de la fonction. En effet, aux endroits où la fonction change beaucoup, on doit considérer une graduation fine de l'espace de la variable aléatoire.

Le but de la recherche effectuée dans le cadre du présent projet est de déterminer les classes de scénarios de manière à ce que l'équation (3) fournisse une bonne approximation à l'équation (2) pour une gamme d'aménagements possibles. Dans la première partie de l'étude (Rapport 1), nous avons proposé de représenter un scénario d'apport S par un ensemble des facteurs indépendants, $\mathbf{F} = (F_1, F_2, \dots, F_k)$ qui sont supposés contenir toute l'information pertinente pour l'aménagement optimal (nous faisons un bref rappel de cette approche au début du chapitre 3 du présent rapport). Dans l'espace réduit des facteurs, on peut définir des classes qui servent ensuite à choisir des scénarios types. Cependant, la définition des classes ne doit pas être basée sur un choix subjectif qui pourrait s'avérer non optimal. Pour déterminer de manière rationnelle quels aspects d'un scénario ont une influence sur l'aménagement optimal et les coûts d'exploitation, on doit effectuer une analyse de sensibilité. Dans le présent contexte, cette analyse a deux buts:

- pour chaque facteur d'intérêt potentiel, examiner la variation du coût en fonction du facteur;
- classer les facteurs selon leur importance pour le coût total.

Si l'on trouve que le coût varie peu ou pas du tout en fonction d'un facteur donné, ce facteur peut être éliminé de l'étude car il fournit aucune information pertinente pour la détermination du coût espéré. En effet, si l'on remplace S par \mathbf{F} dans l'équation (2) et divise le vecteur \mathbf{F} en $(\mathbf{F}^{(P)}, \mathbf{F}^{(NP)})$ où $\mathbf{F}^{(P)}$ et $\mathbf{F}^{(NP)}$ signifient respectivement les facteurs pertinents et non pertinents (i.e. qui n'ont aucune influence sur la fonction du coût), alors on trouve

$$\begin{aligned}
 E[\xi(\mathbf{A})] &= \int_{\mathbf{s}^{(P)}} \int_{\mathbf{s}^{(NP)}} \xi(\mathbf{A}, \mathbf{F}^{(P)}) f((\mathbf{F}^{(P)}, \mathbf{F}^{(NP)})) d(\mathbf{F}^{(P)}, \mathbf{F}^{(NP)}) \\
 &= \int_{\mathbf{s}^{(P)}} \xi(\mathbf{A}, \mathbf{F}^{(P)}) \int_{\mathbf{s}^{(NP)}} f((\mathbf{F}^{(P)}, \mathbf{F}^{(NP)})) d(\mathbf{F}^{(P)}, \mathbf{F}^{(NP)}) \\
 &= \int_{\mathbf{s}^{(P)}} \xi(\mathbf{A}, \mathbf{F}^{(P)}) f(\mathbf{F}^{(P)}) d\mathbf{F}^{(P)}
 \end{aligned} \tag{4}$$

où $f(\mathbf{F}^{(P)})$ est la fonction de probabilité marginale de $\mathbf{F}^{(P)}$. On doit donc chercher les facteurs pour lesquelles une variation conduit à une variation significative du coût. Soulignons que si l'on veut calculer l'espérance mathématique d'une fonction g d'une variable aléatoire X par une approximation discrète de la distribution de probabilité f de X , il importe de bien considérer la variation de g en fonction de X lors de la discrétisation de X .

Dans l'analyse de sensibilité nous devons alors examiner la variation de $\xi(\mathbf{A}, \mathbf{F})$ en fonction de \mathbf{F} , ou plutôt des éléments de \mathbf{F} et choisir la discrétisation de l'espace de \mathbf{F} de façon à obtenir une bonne approximation du coût espéré. On note ici une difficulté pratique puisque la fonction du coût dépend également de l'aménagement particulier, \mathbf{A} . Il faut donc décider quel aménagement considérer dans l'analyse de sensibilité. Nous proposons deux approches:

- l'aménagement optimal correspondant aux apports historiques; et
- l'aménagement optimal correspondant au scénario considéré.

Le premier cas est l'analyse de sensibilité classique. En considérant les apports historiques, on détermine l'aménagement correspondant et on examine ensuite la variation du coût total en fonction des facteurs. L'avantage de cette approche est que l'on doit exécuter MINERVE une seule fois. Le désavantage est que si l'aménagement optimal obtenu par l'approche stochastique est loin de l'aménagement optimal déterminé à partir des données historiques, le résultat de l'analyse de sensibilité peut être peu représentatif. Dans la deuxième approche ci-dessus, on doit exécuter MINERVE pour chaque scénario considéré. Cependant, en considérant un seul scénario d'apport à la fois, MINERVE peut être exécuté en version déterministe ce qui représente un gain de temps important. La deuxième procédure permettra donc d'établir le lien entre un scénario donné et le coût total minimal si ce scénario se réalise. Dans l'exemple présenté dans la section 4, nous avons retenu la deuxième approche.

La fonction objectif considérée dans MINERVE96 est donnée par

$$\overline{\xi(\mathbf{A})} + \lambda \sum_{i \in I_F} [\xi(\mathbf{A}, F_i) - \overline{\xi(\mathbf{A})}]^2 p_i \quad (5)$$

où $\overline{\xi(\mathbf{A})} = \sum_{i \in I_F} \xi(\mathbf{A}, F_i) p_i$ est le coût espéré associé à l'aménagement \mathbf{A} , déterminé de façon approximative par une discrétisation de l'espace factoriel. Le paramètre p_i signifie la probabilité associée au facteur F_i . Le deuxième terme dans (5) est la variance du coût de l'aménagement \mathbf{A} pondérée avec un coefficient λ . Ce coefficient, qui doit généralement être situé entre zéro et un, détermine le niveau de risque que l'on est prêt à accepter. Un choix conservateur consiste à utiliser une valeur de λ élevée de façon à pénaliser un aménagement dont la variation des coûts d'exploitation est grande. Si $\lambda > 0$, alors pour deux aménagements différents dont les coûts espérés sont identiques, on préférera celui dont le coût est le moins variable. Si par contre $\lambda = 0$, alors le choix d'aménagement est basé uniquement sur le coût espéré. Le choix de λ est une décision politique et économique qui revient à l'usager. Dans le chapitre 5, nous étudions l'impact du choix de λ .

3 Détermination de la relation facteur/coût

3.1 Introduction

La méthodologie proposée ici pour déterminer les relations entre les scénarios d'apport et le coût total du projet fait suite aux travaux réalisés dans le cadre de la première phase du projet (Rasmussen *et al.*, 1996). Nous ferons dans la section suivante un rappel rapide des principaux résultats de la première phase de l'étude. Ensuite, nous présentons une méthodologie rationnelle permettant l'identification des scénarios les plus pertinents à retenir pour le modèle MINERVE96.

3.2 Réduction de la dimensionnalité des apports

Dans la première partie du projet, nous avons proposé d'effectuer d'abord une analyse statistique approfondie des apports observés (ou reconstitués) aux sites où l'on prévoit l'aménagement de nouvelles installations. Pour ce type d'étude, on s'intéresse généralement aux apports mensuels ou trimestriels. Suite à l'examen des caractéristiques statistiques et stochastiques des apports observés, on doit sélectionner et estimer un modèle stochastique permettant de générer des apports synthétiques multidimensionnels. Différents modèles adaptés à ce problème sont décrits dans le Rapport 1 (1996). Il est important que les scénarios synthétiques d'apport soient réalistes du point de vue hydrologique, c'est-à-dire qu'ils aient des caractéristiques que l'on pourrait réellement observer dans le futur. Nous avons déjà, dans le chapitre précédent, souligné la multidimensionnalité des apports. Une première approche pour réduire la dimensionnalité consiste à effectuer une agrégation spatiale et/ou temporelle des apports. En considérant les apports mensuels ou trimestriel plutôt que des apports journaliers, on a déjà obtenu une forte réduction de la dimensionnalité. Pour déterminer l'aménagement optimal et les coûts d'exploitation associée, il est généralement important de considérer les apports saisonniers (mensuel/trimestriel) aux différents sites du système hydrique. Cependant, pour la discrimination des scénarios, il peut être suffisant de considérer les apports à une échelle plus grossière, par exemple des apports régionaux et/ou des apports annuels. Le choix du niveau d'agrégation doit être effectué en fonction des

corrélations spatiales et temporelles des apports. Par exemple, s'il y a une forte corrélation entre les apports aux différents sites d'une vallée, alors il peut être avantageux d'effectuer une agrégation spatiale. De la même manière, si la répartition de l'apport annuel à l'intérieur d'une année est relativement stable, il sera avantageux d'effectuer une agrégation temporelle des apports. L'envergure des installations doit également jouer un rôle dans le choix d'agrégation. Par exemple, si les réservoirs sont conçus pour le stockage interannuel, on doit préconiser une agrégation temporelle des apports, par exemple en considérant les apports annuels. Soulignons que pour la discrimination entre scénarios, il est avantageux d'agréger les apports le plus possible, sous la contrainte, bien entendu, de ne pas éliminer de l'information pertinente. L'agrégation des apports permet de réduire significativement la dimensionnalité des variables d'intérêt. Puisque seulement un nombre limité de scénarios peut être considéré dans MINERVE96, il est essentiel d'accentuer l'étude de discrimination sur les aspects des apports qui sont vraiment importants pour la détermination de l'aménagement optimal.

Les apports peuvent être agrégés de plusieurs manières. Par exemple, dans le Rapport 1 (1996) on propose trois façons de définir des séries annuelles, spatialement agrégées:

1. apport agrégé du bassin versant;
2. apport agrégé à chaque site du bassin; et
3. apport agrégé à chaque site du bassin multiplié avec un facteur relié à les hauteurs de chutes.

La première variable est tout simplement l'apport total du bassin versant. La deuxième variable est la sommation des apports annuels observés à chaque sites d'intérêt, en général les endroits où l'on prévoit la construction des centrales. Cette variable tient donc compte du fait que l'apport intermédiaire à un site en amont est plus précieuse que l'apport intermédiaire qui arrive à un site en aval. La troisième variable est similaire à la précédente, mais introduit en plus la notion de la valeur de l'eau en terme d'énergie. Pour obtenir cette variable, on doit multiplier l'apport à chaque site avec la hauteur moyenne des chutes. Les hauteurs des chutes ne sont généralement pas connues à l'avance (un des buts de l'optimisation étant justement de les déterminer), mais souvent on a une idée approximative de leur hauteur. Même si les hauteurs ne sont pas connues exactement, on peut obtenir de bons résultats avec des valeurs approximatives.

Lorsque la variable d'intérêt a été définie (on l'appellera l'apport agrégé dans ce qui suit), on procède à la deuxième étape de la réduction de la dimensionnalité, car même la série agrégée a généralement trop de dimensions. Dans cette phase, on doit énumérer tous les aspects des séries d'apports agrégés pouvant avoir une importance pour l'aménagement optimal et pour le coût associé à l'exploitation. Encore là, il est important de bien comprendre la problématique de

l'optimisation et, en particulier, les différentes contraintes qui peuvent influencer les coûts d'exploitation pour le cas d'étude. On a intérêt à considérer uniquement les aspects qui ont une importance significative sur le coût d'exploitation. On doit souligner que les aspects importants peuvent varier d'une application à une autre. Dans la première partie de l'étude, nous avons proposé de retenir les statistiques suivants des apports agrégés (annuels) pour la rivière Grande Baleine:

-
- Moyenne
 - Écart-type
 - Longueur maximum des séquences de déficit par rapport à la moyenne historique
 - Longueur moyenne des séquences de déficit par rapport à la moyenne historique
 - Écart-type de la longueur des séquences de déficit par rapport à la moyenne historique
 - Maximum des déficits accumulés par rapport à la moyenne historique
 - Moyenne des déficits accumulés par rapport à la moyenne historique
 - Écart-type des déficits accumulés par rapport à la moyenne historique
 - Domaine de variation ajusté
 - Domaine de variation ajusté standardisé
 - Coefficient de Hurst
-

Une description détaillée de ces caractéristiques statistiques est fournie dans Rapport1 (1996). Ces critères qui nous semblent tous pertinents pour le coût d'exploitation ont été utilisés à titre d'exemple; on n'exclut pas que d'autres caractéristiques pourraient avoir une importance significative (on rappelle que le but de notre travail est de développer une méthodologie et non pas l'étude spécifique d'un aménagement donné). Cependant, à ce point, notre hypothèse est que la majorité de l'information pertinente est contenue dans l'ensemble des caractéristiques statistiques choisies.

Même si la représentation des scénarios d'apports par un ensemble de caractéristiques statistiques conduit à une forte réduction de la dimensionnalité, il est nécessaire de la réduire davantage. On rappelle que seulement 10-20 scénarios peuvent être considérés dans MINERVE96, ce qui nous amène à proposer que la discrimination doit être faite sur la base d'un maximum de deux à trois critères. Puisque certains critères peuvent être fortement corrélés entre eux, la prochaine étape de la méthodologie consiste à éliminer toute information redondante engendrée par cette corrélation. L'analyse en composantes principales est une méthode statistique multivariée qui permet de synthétiser un tableau de données en quelques facteurs. Les détails techniques de cette méthode et

son application pour le problème en question sont présentés dans le Rapport 1 (1996). Si l'on suppose que M critères statistiques sont considérées, l'analyse en composantes principales donne comme résultat les poids factoriels qui permettent de calculer les facteurs associés à un scénario donné. Si les M critères sont corrélés entre eux, certains facteurs ne seront pas significatifs. Le principe de l'analyse en composantes principales est de retenir seulement les K facteurs significatifs. À part de synthétiser les scénarios de manière plus parcimonieuse, on peut noter que l'analyse en composantes principales a comme avantage que les K facteurs retenus sont indépendants, ce qui facilite leur interprétation et leur utilisation dans l'analyse qui suit.

Il nous semble important de rappeler que le fait qu'un facteur donné est identifié comme étant significatif dans l'analyse en composantes principales n'implique nullement que ce facteur est important pour l'aménagement optimal et les coûts d'exploitation. Les facteurs significatifs décrivent différents aspects statistiques des scénarios, mais ces aspects ne sont pas nécessairement importants pour le problème d'aménagement optimal. Une analyse de sensibilité facteur-coût permettra par contre de répondre à cette question.

3.3 Analyse de sensibilité des facteurs

Ayant synthétisé chacun des scénarios d'apports en K facteurs indépendants, nous pouvons poser le problème auquel on s'adresse dans la présente étude de la façon suivante:

- parmi l'ensemble des K facteurs, y a-t-il certains qui sont non pertinents pour le coût d'exploitation;
- quels sont les facteurs qui affectent le plus les coûts d'exploitation; et
- quelle discrétisation de l'espace factoriel permet la meilleure approximation de l'espérance et de la variance des coûts d'exploitation.

Pour répondre adéquatement à ces questions, il faut effectuer une analyse de sensibilité dont le but est d'établir la relation entre le coût d'exploitation et chaque facteur. On peut montrer qu'en général les facteurs sont distribués approximativement comme une loi normale à K dimensions. Puisque les variables ont été standardisées, cette loi normale a une vecteur moyenne nulle et une matrice de covariance I , où I est la matrice unitaire. Idéalement, lors de l'analyse de sensibilité, on souhaiterait fixer $(K-1)$ facteurs et faire varier le K 'ième, afin de déterminer la variation des coûts d'exploitation en fonction de cet K 'ième facteur. Ceci pose un problème pratique, car on ne peut fixer les facteurs et ensuite revenir aux apports. Les facteurs sont calculés à partir des apports et non pas non pas le contraire. Pour cette raison, on propose de procéder selon les étapes suivantes:

Etape 1: Génération - phase 1

On doit d'abord générer quelques scénarios synthétiques (100-500) permettant d'étudier les corrélations entre les critères (moyenne, écart-type, séquence, minimum, etc.). Plus particulièrement, on effectue une analyse en composantes principales qui permet de déterminer les poids factoriels. Ces poids sont utilisés dans ce qui suit pour associer des facteurs à chaque scénario d'apport. Les outils pour effectuer ces analyses sont décrits dans le Rapport 1 (1996).

Etape 2: Sélection des intervalles de tolérance

On doit à cette étape sélectionner un intervalle de tolérance pour chaque facteur. Si un facteur se retrouve à l'intérieur de son intervalle de tolérance, il est considéré quasi-identique à sa valeur moyenne (historique). L'intervalle de tolérance pour le facteur k peut être choisi comme un intervalle de probabilité. Par exemple, on peut définir l'intervalle comme $[\Phi^{-1}(0.45); \Phi^{-1}(0.55)]$ où Φ signifie la fonction de répartition d'une loi normale réduite. Dans ce cas, environ 10% des scénarios auront une valeur du facteur k suffisamment proche de sa valeur historique pour être considérée quasi-identique à celle-ci. Pour des raisons pratiques, le choix des intervalles de tolérance va dépendre du nombre de facteurs retenus.

Etape 3: Génération - phase 2

On génère ensuite un nombre élevé de scénarios et calcule leurs coordonnées dans l'espace factoriel en utilisant les coefficients déterminés à l'étape 1. On s'intéresse notamment aux scénarios pour lesquels $(K-1)$ facteurs se retrouvent à l'intérieur de leur intervalle de tolérance alors que le K ième varie. Le nombre de scénarios à générer va dépendre du nombre de facteurs et des intervalles de tolérance sélectionnés. Par exemple, si quatre facteurs sont retenus et tous les intervalles sont définis comme des intervalles de probabilité de 10%, alors la probabilité qu'un scénario donné a trois facteurs proches de leur valeur historique est $(0.10)^3 = 0.001$. Ou, par exemple, si l'on désire de disposer de 100 scénarios pour lesquels le facteur 1 varie alors que les autres sont proches de leurs valeurs historiques, il faut générer environ 100,000 scénarios. Notons que si les critères sont calculés à partir des séries agrégées d'apport, on pourrait gagner du temps en générant directement les séries agrégées. Lors de l'analyse de sensibilité, on désagrègerait uniquement les scénarios retenus.

Etape 4: Sélection des scénarios pour l'analyse de sensibilité

Pour étudier la sensibilité du coût de construction et d'exploitation vis-à-vis le facteur k , on sélectionne des scénarios pour lesquelles tous les facteurs à part le numéro k sont à l'intérieur de leur intervalle de tolérance. Les scénarios sont choisis de manière à ce que le k ième facteur couvre

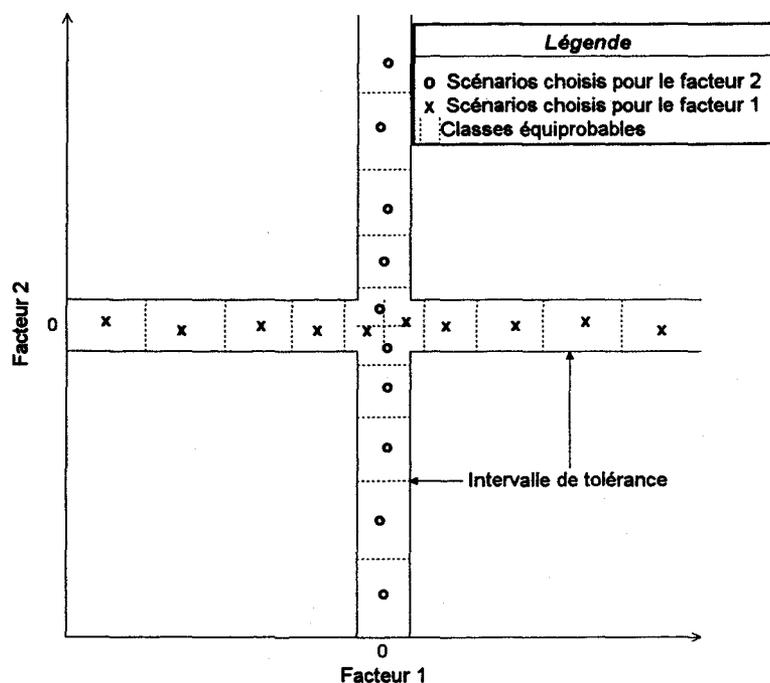


Figure 3.1 Sélection des scénarios pour l'analyse de sensibilité

une plage intéressante de valeurs (figure 3.1). La procédure est répétée pour chacun des facteurs. Le nombre de scénarios à considérer pour chaque facteur va dépendre uniquement du temps que l'utilisateur est prêt à consacrer à l'étude. Si l'on a généré seulement l'apport agrégé dans l'étape précédente, il faut désagréger les apports à ce stade afin qu'ils puissent être considérés dans MINERVE96.

Etape 5: Détermination de la relation facteur/coût

À ce point, on doit déterminer le coût associé à chaque scénario d'apport. Or, ce coût pourrait dépendre fortement de l'aménagement choisi. Dans le chapitre 2, nous avons proposé deux procédures pour déterminer ces coûts:

- Détermination du coût si l'aménagement est fixé. On pourrait considérer l'aménagement qui est optimal par rapport au scénario historique;
- Détermination du coût si l'aménagement est optimal par rapport au scénario en question.

Dans le premier cas, MINERVE96 doit être exécuté en version déterministe une seule fois avec les apports historiques. Ensuite, on utilise l'équation 1 pour déterminer le coût. Cette approche sera appropriée si la variation des coûts est relativement peu influencés par l'aménagement. Dans le deuxième cas, on doit exécuter MINERVE96 pour chacun des scénarios. De cette façon, on détermine le coût minimal possible, si le scénario en question se réalise. Alors que dans le premier cas, la variation du coût total est reliée uniquement à l'achat et la vente d'énergie, le deuxième tient

aussi compte du coût d'aménagement. Bien que l'intérêt principal soit la variation des coûts d'exploitation pour un aménagement donné, plus particulièrement l'aménagement optimal, le problème est que cet aménagement n'est pas connu *a priori*. De manière heuristique, on peut montrer un certain lien entre le coût optimal en fonction des scénarios et la variation du coût d'exploitation pour un aménagement fixe. La figure 3.2 illustre cette relation.

Dans la figure 3.2, la ligne droite illustre le coût optimal en fonction du scénario, S . Logiquement, plus le scénario S est optimiste, plus le coût optimal est faible. La deuxième courbe de la figure, illustre le coût associé à l'aménagement qui est optimal pour le scénario S_0 . On peut faire les remarques suivantes:

- la courbe correspondant à $\xi(A_{opt}(S_0), S)$ est au-dessus de la courbe $\xi(A_{opt}(S), S)$; et
- pour un scénario S_1 qui est plus optimiste que S_0 , on a $\xi(A_{opt}(S_0), S_1) < \xi(A_{opt}(S_0), S_0)$.

Ces observations permettent de stipuler qu'une variation de $\xi(A_{opt}(S), S)$ implique généralement une variation de $\xi(A_{opt}(S_0), S)$. Alors un facteur qui a une influence sur l'aménagement optimal devrait également avoir une influence sur la variation des coûts associés à un aménagement fixe (coûts d'exploitation).

Pour l'analyse de sensibilité, nous donnerons une préférence à l'approche basée sur la variation du coût optimal en fonction des facteurs, car elle nous semble de mieux prendre en compte le coût de construction.

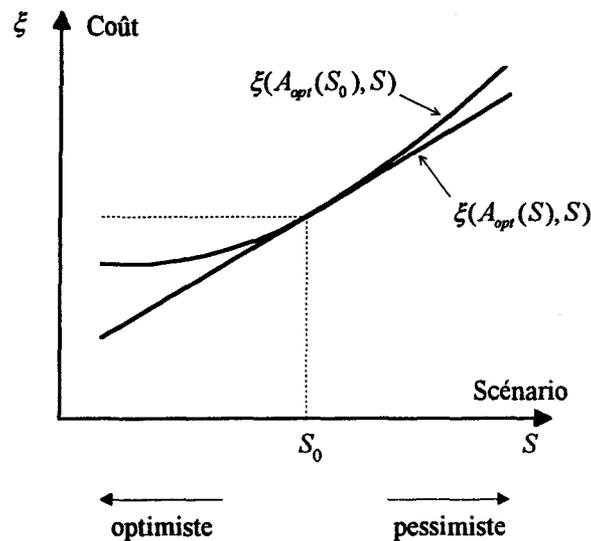


Figure 3.2 Relation entre coût optimal en fonction des scénarios et coût associé à un aménagement donné.

Etape 6: Interprétation des résultats de l'analyse de sensibilité

Pour faciliter l'interprétation des résultats obtenus au cours de l'étape précédente, nous suggérons de mettre sur graphique la relation entre les coûts d'exploitation et les scénarios d'apports. Ces relations seront généralement caractérisées par un certain bruit qui pourrait être causé par:

- la variation des autres facteurs à l'intérieur de leur intervalle de tolérance. Si cela est le cas et que l'ampleur du bruit est considérée non acceptable, on pourrait diminuer les intervalles de tolérance.
- les facteurs non considérés. On a fait l'hypothèse que toute information pertinente est synthétisée dans les K facteurs. Si l'on soupçonne que d'autres facteurs pertinents ont été oubliés, il faudrait les identifier et refaire les étapes 1-5.
- des solutions non optimales produites par MINERVE96. Puisqu'il s'agit d'un problème d'optimisation non convexe, il est recommandé, pour un scénario d'apports donné, d'exécuter MINERVE96 avec différentes valeurs initiales et de choisir l'aménagement ayant le coût le plus faible. Différentes approximations numériques pourraient également introduire une certaine variation au niveau des coûts.

En raison du bruit qui sera inévitablement présent dans l'estimation du coût, on doit se limiter à dégager les grandes tendances des relations coût/facteur. Plus particulièrement, on doit chercher à identifier les facteurs qui ont le plus grand impact sur le coût associé à un aménagement optimal. La figure 3.3 illustre, à titre d'exemple, des relations qui pourraient exister entre le coût optimal et trois facteurs différents. Dans cet exemple, on voit que le facteur 1 a une forte influence sur le coût minimal, au moins si les facteurs 2 et 3 sont proches de leur valeurs historiques (figure 3.3a). Le facteur 2, par contre, affecte peu les coûts associés à l'aménagement optimal. Ce facteur pourrait donc être éliminé de l'analyse. Dans le cas du facteur 3, on voit que les coûts sont peu affectés si la valeur du facteur est faible. Par contre, lorsque la valeur du facteur dépasse un certain niveau, les coûts augmentent. On doit évidemment considérer le facteur 3 dans l'analyse, mais la discrétisation de l'espace pour le facteur 3 pourrait être différente de celle du facteur 1. De manière

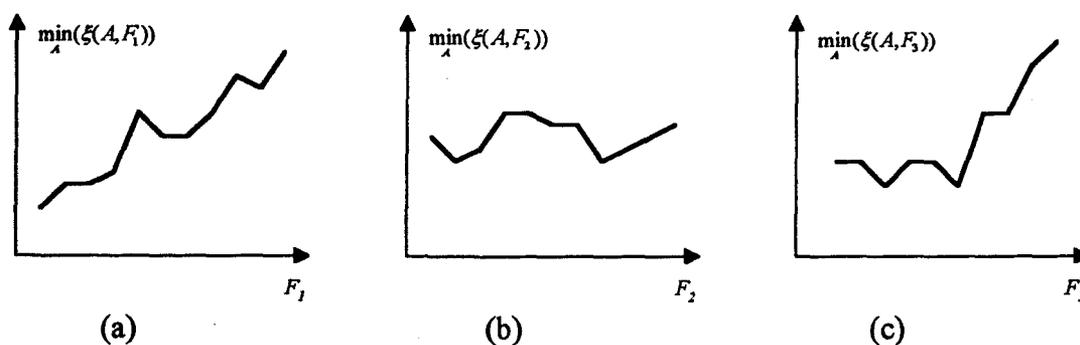


Figure 3.3 Exemples de la variation du coût optimal en fonction de trois facteur

générale, il s'agit de classer les facteurs selon leur importance pour le coût minimal et d'identifier les plages où le coût minimal varie de manière significative. Tout facteur non pertinent peut être éliminé de l'étude.

On souligne d'ailleurs que l'analyse de sensibilité, telle que proposée ici, est basée sur l'hypothèse que $(K-1)$ facteurs sont proches de leurs valeurs historiques et que seulement un des K facteurs varie. Cette approche ne permet donc pas d'étudier l'importance de l'interaction des facteurs sur le coût.

Etape 7: Division de l'espace factoriel et détermination des poids

Lorsque l'analyse de sensibilité a été complétée, on doit procéder à une discrétisation de l'espace factoriel. Cette discrétisation permet d'identifier les scénarios les plus pertinents à considérer dans MINERVE96.

Une approche simple consiste à choisir des classes équiprobables. Les bornes de chaque classe peuvent être déterminées analytiquement en faisant l'hypothèse que les facteurs sont distribués selon un loi normale de moyenne nulle et de variance unitaire. Si par exemple deux facteurs sont

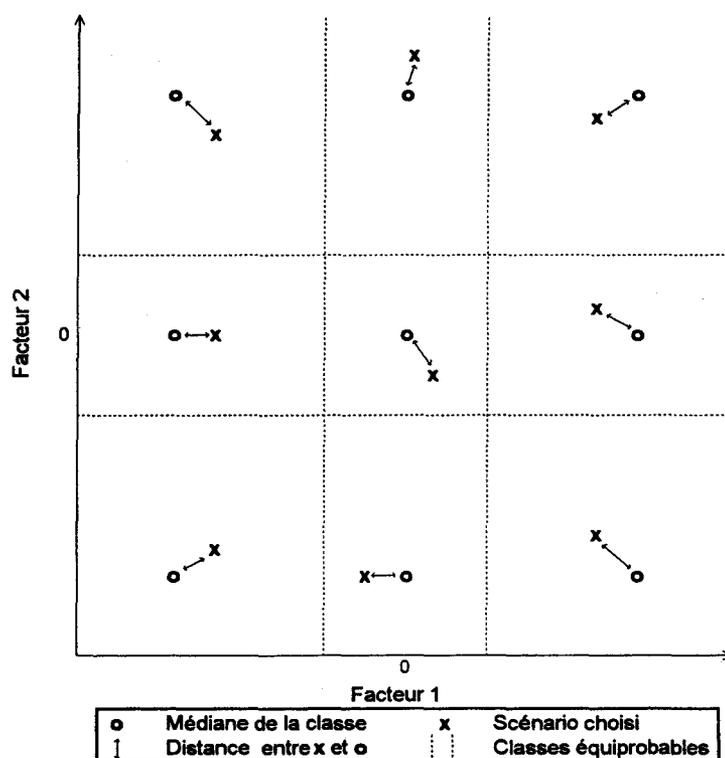


Figure 3.4 Illustration du choix des scénarios pour l'analyse du sensibilité

considérés importants pour l'aménagement, on peut choisir de diviser chaque facteur en trois classes équiprobables. Tel que illustré dans la figure 3.4, ce choix donne lieu à 9 classes dans l'espace des deux facteurs. Il n'est cependant pas nécessaire que les classes soient équiprobables. Par exemple, pour un facteur donné, on pourrait choisir une graduation plus fine dans la partie où la variation du coût est particulièrement forte.

Le poids associé à une classe donnée peut être déterminé comme la probabilité qu'un scénario se trouve dans la classe. Cette probabilité peut être obtenue facilement en utilisant l'hypothèse que les facteurs sont distribués selon une loi normale multivariée de vecteur moyenne nulle et de matrice de covariances égale à la matrice unitaire. Ceci est illustré dans la figure 3.4.

Etape 8: Choix de scénarios pour MINERVE96

A cette dernière étape, il faut effectuer le choix final des séries à considérer dans MINERVE96. Il s'agit plus particulièrement de choisir un scénario représentatif pour chaque classe définie à l'étape précédente. On propose de choisir des scénarios qui sont proches du centre des classes. On peut définir le centre de plusieurs manières, par exemple

- le centre de gravité, i.e. la moyenne de la classe; ou
- le médiane de la classe.

La première approche nécessite l'intégration numérique, alors que dans le deuxième cas le centre peut être déterminé analytiquement. Dans le cas général, les deux méthodes donneront approximativement le même résultat. Une fois le centre de chaque classe déterminé, on peut utiliser une distance euclidienne pour déterminer quel scénario synthétique se trouve le plus proche d'un centre donné (voir figure 3.4).

Le choix des scénarios étant fait, on peut procéder à l'application de MINERVE96. Dans les chapitres suivants, nous présentons un exemple qui illustre l'utilisation de la méthodologie proposée.

4 Exemple d'analyse de sensibilité

Dans ce chapitre, nous présentons un exemple des étapes 1 à 6 décrites à la section 3.3 qui permettent d'effectuer l'analyse de sensibilité du modèle MINERVE96 aux différents facteurs.

4.1 Description du système hydrique de la rivière Grande Baleine

Le système hydrique de la rivière Grande Baleine a été décrit en détail dans le Rapport 1 (1996). Les stations des sous-bassins et la localisation des centrales sont données de manière schématique à la figure 4.1. Dans l'exemple de ce chapitre, les données observées (ou reconstituées) des stations en rivière naturelle (IVR00B1, I93891, I93892, I93893 et I93894) ont été considérées pour la période de 1960 à 1991.

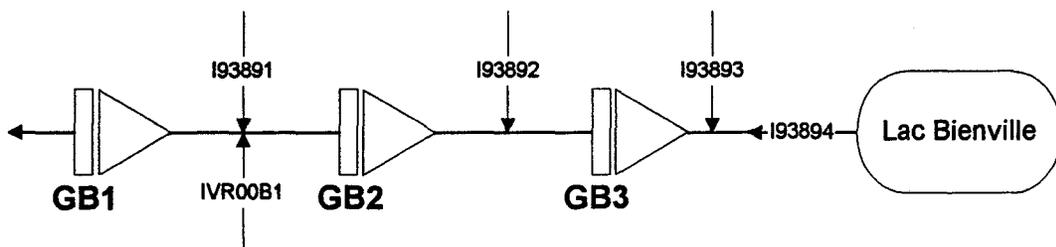


Figure 4.1 Illustration schématique du complexe de Grande Baleine.

4.2 Génération de séries synthétiques.

L'objectif de cette section est de montrer le cheminement menant à l'identification de séries synthétiques pour évaluer la relation entre les facteurs de l'espace factoriel et le coût de construction des aménagements d'une nouvelle vallée (étapes 1 à 4 de la section 3.3).

4.2.1 Sélection des critères

Les douze critères suivants sont considérés afin de discriminer les séries simulées:

-
- Moyenne (E)
 - Écart-type (S)
 - Valeur minimum (MIN)
 - Longueur maximum des séquences de déficit par rapport à la moyenne de la série simulée (MAX_RL)
 - Longueur moyenne des séquences de déficit par rapport à la moyenne de la série simulée (E_RL)
 - Écart-type de la longueur des séquences de déficit par rapport à la moyenne de la série simulée (S_RL)
 - Maximum des déficits cumulés par rapport à la moyenne de la série simulée (MAX_RS)
 - Moyenne des déficits cumulés par rapport à la moyenne de la série simulée (E_RS)
 - Écart-type des déficits cumulés par rapport à la moyenne de la série simulée (S_RS)
 - Domaine de variation ajusté (AR)
 - Domaine de variation ajusté standardisé (RAR)
 - Coefficient de Hurst (H)
-

Les critères retenus pour cette étude nous semblent tous pertinents au problème étudié. Cependant, certains critères peuvent décrire sensiblement la même caractéristique d'une série. Un des objectifs de l'analyse en composantes principales est de réduire le nombre de critères en quelques facteurs moins nombreux en éliminant les critères superflus.

Par rapport à l'étude précédente (Rapport 1, 1996), le critère de la valeur minimum a été ajouté. En effet, cette caractéristique des séries est intéressante, car elle représente une valeur critique minimum en apport pour laquelle la demande en électricité doit tout de même être comblée. Il apparaît donc que cette variable peut avoir une influence sur le coût de l'exploitation des installations. On note également un changement concernant le calcul des déficits. Dans cette étude, les déficits sont établis par rapport à la moyenne de la série simulée et non pas par rapport à la moyenne de la série historique.

4.2.2 Choix du type de séries

Avant d'effectuer la génération des séries pour l'analyse de sensibilité, il faut choisir le type de série que l'on veut considérer. Quatre types de séries ont été retenus pour notre étude:

- 1) Série d'apport naturel agrégé (1 série);
- 2) Séries d'apport naturel agrégé à chacun des sites (3 séries);
- 3) Série d'apport énergétique agrégé (1 série) ;
- 4) Séries d'apport énergétique agrégé à chacun des sites (3 séries).

Le calcul pour obtenir ces séries est décrit à la section 2.2 et dans le Rapport 1 (1996).

Afin de procéder à une élimination des types de séries, une simulation préliminaire de 100 séries a été effectuée et les différents critères énumérés à la section précédente ont été calculés pour chaque type de séries. Ensuite, une analyse en composantes principales nous a permis d'identifier la corrélation entre les critères des quatre types de séries. Dans le cas du complexe de la Grande Baleine et des caractéristiques statistiques choisies, il n'y a pas de différence importante entre une ou l'autre des types de séries. Les figures 4.2 et 4.3 montrent le poids factoriel des différentes variables dans le plan des deux premiers facteurs pour, respectivement, les séries du type 1 et 3 et les séries du type 1 et 2. Il est clair, après examen de ces deux graphiques, que les différentes séries contiennent la même information car les mêmes caractéristiques des différents types de séries se retrouvent dans les mêmes régions.

Puisqu'il n'est pas nécessaire d'utiliser tous les types de séries, la série des apports agrégés du bassin versant (type 1) a été sélectionnée pour la suite de l'étude. Il est à noter que le choix d'un autre type de séries aurait été tout aussi adéquat.

4.2.3 Génération - phase I

La plupart des modèles de génération de séries synthétiques reposent sur l'hypothèse que les données suivent une loi normale. En réalité, cela est rarement le cas et il faut alors procéder à une transformation des données. Le choix de la meilleure transformation à utiliser pour normaliser les données peut être basé sur la statistique de Filliben qui est en fait le coefficient de corrélation entre les données transformées ordonnées et leur espérance mathématique (Filliben, 1975). La transformation log-normale à deux paramètres est la plus adéquate dans le cas de la série historique des apports naturels agrégés.

L'analyse des données historiques d'apport naturel agrégé n'indique pas une autocorrélation d'ordre 1 significative ($\rho_1 = 0.12$). En fait, cela démontre qu'il n'y a pas de lien entre l'apport agrégé d'une année et celui de l'année suivante. Pour cette raison, le modèle autorégressif d'ordre 0 [AR(0)] a été utilisé pour la génération des séries annuelles. Il n'est pas nécessaire à cette étape d'effectuer

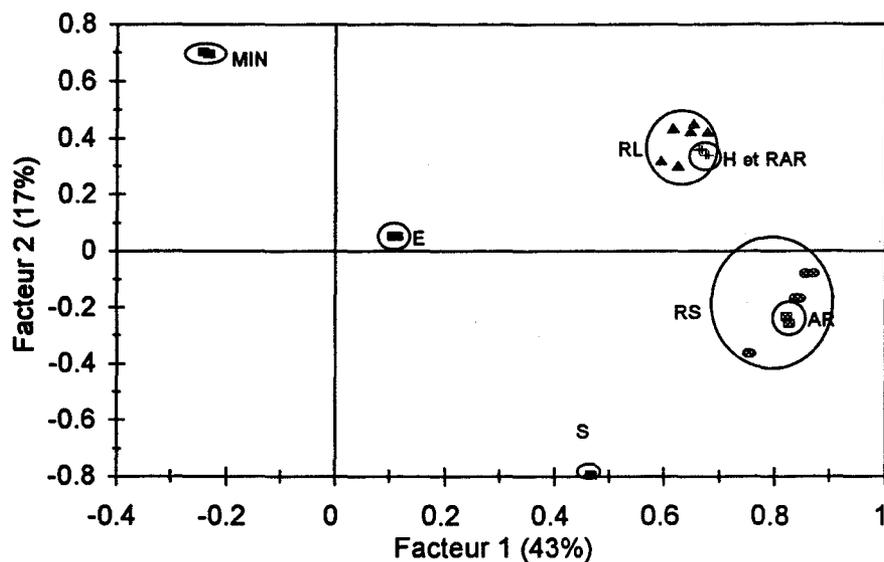


Figure 4.2 Position des critères dans l'espace des deux premiers facteurs pour les séries d'apport naturel agrégé (type 1) et les séries d'apport énergétique agrégé (type 3).

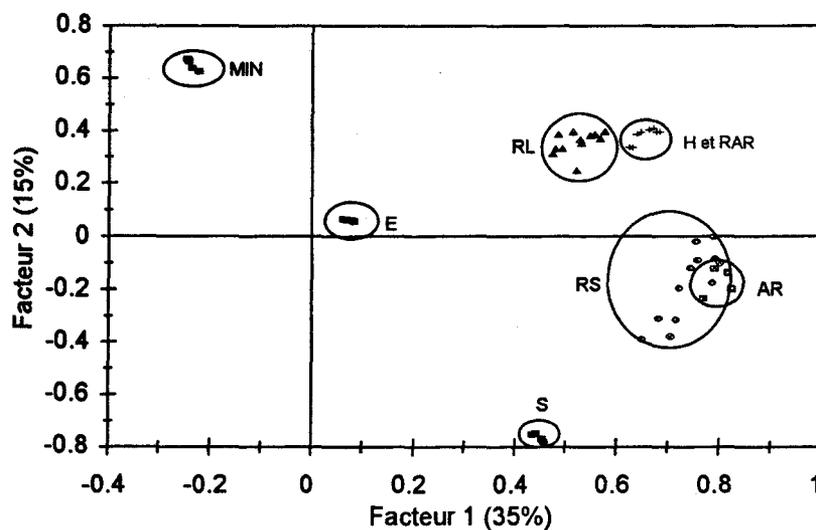


Figure 4.3 Position des critères dans l'espace des deux premiers facteurs pour les séries d'apport naturel annuel agrégé (type 1) et les séries d'apport naturel à chacun des sites (type 2).

une désagrégation temporelle et spatiale des séries, car toutes les caractéristiques peuvent être calculées à partir de la série agrégée.

À titre d'illustration de l'approche proposée, 1000 séries synthétiques d'apport naturel annuel agrégé ont été générées afin de déterminer les facteurs de l'analyse en composantes principales. Un nombre moindre de séries aurait été suffisant, car l'analyse des poids factoriels montre que ces derniers se stabilisent après une centaine de séries.

4.2.4 Analyse en composantes principales

Pour chaque série synthétique d'apport naturel agrégé, les caractéristiques décrite à la section 4.2.1 ont été calculées. Le tableau 4.1 donne quelques statistiques de base de chaque caractéristique et la valeur correspondante pour la série historique. Le tableau 4.2 montre la corrélation entre les différents critères considérés. On remarque que la corrélation entre certains critères (par exemple, AR, RAR et H) est grande, ce qui indique qu'il y a de l'information redondante.

La matrice de corrélation des critères a quatre valeurs propres supérieures à un. Les deux, trois et quatre premières composantes principales expliquent respectivement 61%, 75% et 85% de la variance totale. Le tableau 4.3 donne les coordonnées des douze critères sur les quatre premiers axes factoriels et les figures 4.4a à 4.4f montrent la position des critères selon ces mêmes axes. Les

Tableau 4.1 Caractéristiques statistiques des critères pour les séries synthétiques et pour la série historique

Variable	Historique	Séries synthétiques					
		Moyenne	Médiane	Minimum	Maximum	Écart-type	Cs
MAX_RL	4.00	4.50	4.00	2.00	13.00	1.52	1.08
E_RL	1.75	2.04	2.00	1.09	4.50	0.44	1.01
S_RL	1.39	1.32	1.19	0.30	5.17	0.58	1.45
MAX_RS	2696	2812	2649	1138	7133	921	1.16
E_RS	1051	1128	1084	540	2651	266	0.96
S_RS	971	911	843	290	2739	350	1.35
E	8731	8732	8727	8344	9131	130	0.15
S	728	727	721	446	1075	94	0.21
MIN	7277	7335	7362	6261	8095	294	-0.43
AR	4774	4262	4117	1849	8621	1104	0.63
RAR	6.55	5.86	5.73	2.95	11.50	1.29	0.47
H	0.69	0.64	0.64	0.39	0.89	0.08	-0.11

Tableau 4.2 Matrice de corrélation des critères

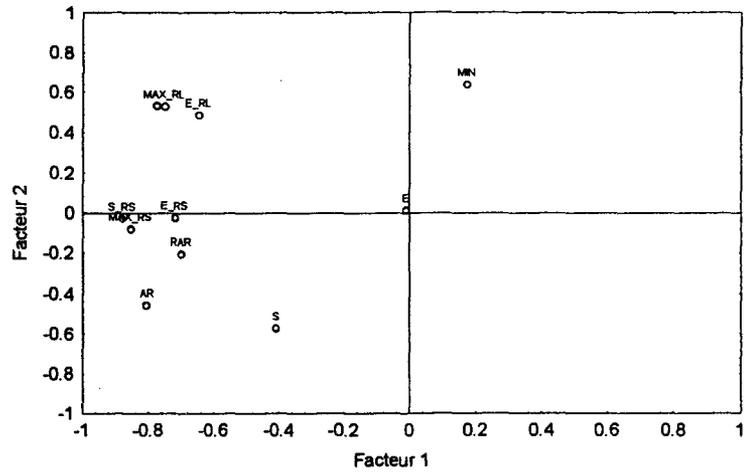
Variable	MAX_RL	E_RL	S_RL	MAX_RS	E_RS	S_RS	E	S	MIN	AR	RAR	H
MAX_RL	1.00	0.62	0.96	0.62	0.42	0.63	0.04	0.10	0.07	0.35	0.36	0.35
E_RL	0.62	1.00	0.68	0.42	0.72	0.48	0.06	0.11	0.11	0.34	0.34	0.34
S_RL	0.96	0.68	1.00	0.65	0.50	0.68	0.04	0.11	0.05	0.38	0.39	0.38
MAX_RS	0.62	0.42	0.65	1.00	0.60	0.95	0.07	0.42	-0.20	0.61	0.47	0.47
E_RS	0.42	0.72	0.50	0.60	1.00	0.61	0.11	0.58	-0.18	0.57	0.34	0.32
S_RS	0.63	0.48	0.68	0.95	0.61	1.00	0.06	0.38	-0.21	0.62	0.50	0.49
E	0.04	0.06	0.04	0.07	0.11	0.06	1.00	0.15	0.37	0.11	0.04	0.04
S	0.10	0.11	0.11	0.42	0.58	0.38	0.15	1.00	-0.44	0.53	0.06	0.05
MIN	0.07	0.11	0.05	-0.20	-0.18	-0.21	0.37	-0.44	1.00	-0.21	0.00	0.01
AR	0.35	0.34	0.38	0.61	0.57	0.62	0.11	0.53	-0.21	1.00	0.87	0.86
RAR	0.36	0.34	0.39	0.47	0.34	0.50	0.04	0.06	0.00	0.87	1.00	0.99
H	0.35	0.34	0.38	0.47	0.32	0.49	0.04	0.05	0.01	0.86	0.99	1.00

En gras : significativement différent de 0 à 5%

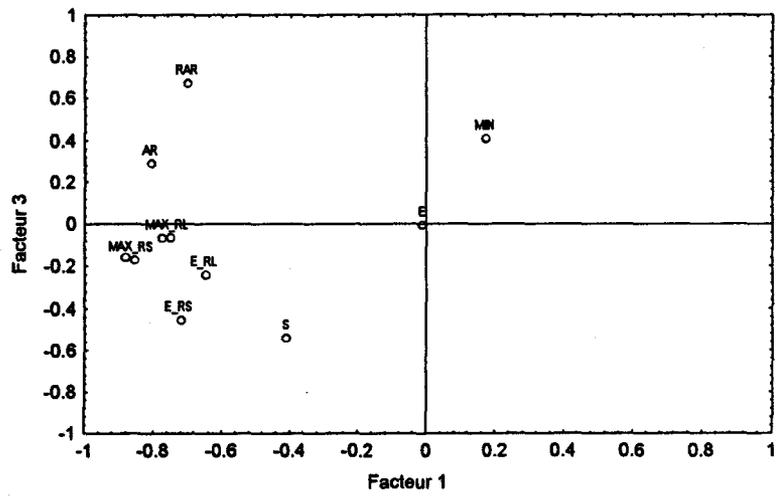
coordonnées supérieures à 0.70 sont indiquées en gras. Pour le premier facteur, les coordonnées des statistiques reliées aux séquences de déficit (RS et RL) et les variables AR, RAR et H sont élevées et de même signe. Ce facteur est donc relié aux séquences des déficits et au domaine de variation ou au besoin de stockage. Sur le deuxième axe factoriel, les variables MIN et S ont les coordonnées les plus élevées en valeur absolue et de signe opposé. Ce résultat n'est pas surprenant, car pour une même moyenne, lorsque l'écart-type d'un échantillon augmente la valeur minimum est susceptible de diminuer (corrélation négative). En ce qui concerne le troisième facteur, les caractéristiques H et RAR ont les poids les plus importants. Il est à noter que pour les axes 2 et 3 aucune coordonnée ne dépasse 0.70. Enfin, le quatrième facteur est clairement associé à la moyenne de la série avec une coordonnée de 0.94.

Tableau 4.3 Coordonnées des critères sur les axes factoriels

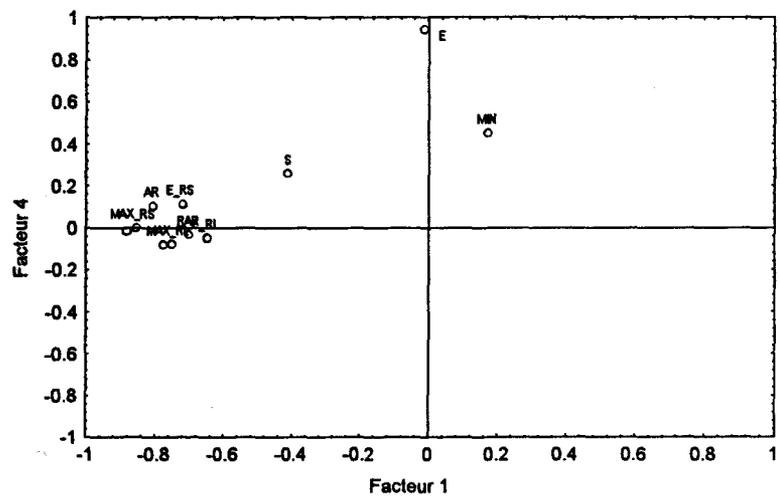
Variable	Facteur 1	Facteur 2	Facteur 3	Facteur 4
MAX_RL	-0.7504	0.5269	-0.0675	-0.0787
E_RL	-0.6470	0.4844	-0.2422	-0.0477
S_RL	-0.7751	0.5307	-0.0688	-0.0810
MAX_RS	-0.8544	-0.0830	-0.1730	0.0009
E_RS	-0.7199	-0.0264	-0.4556	0.1144
S_RS	-0.8800	-0.0306	-0.1598	-0.0153
E	-0.0118	0.0089	-0.0077	0.9452
S	-0.4119	-0.5743	-0.5408	0.2587
MIN	0.1736	0.6348	0.4071	0.4513
AR	-0.8055	-0.4605	0.2894	0.1047
RAR	-0.6994	-0.2090	0.6703	-0.0312
H	-0.6893	-0.2023	0.6799	-0.0273
Valeur propre	5.418	1.832	1.785	1.205



(a)

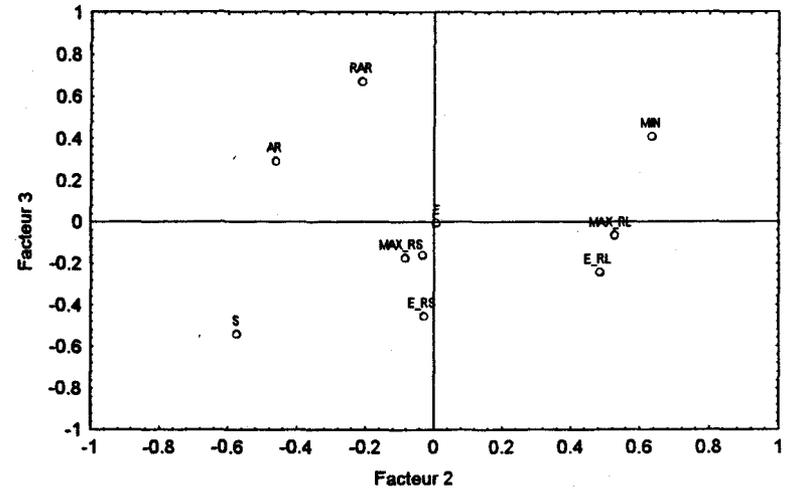


(b)

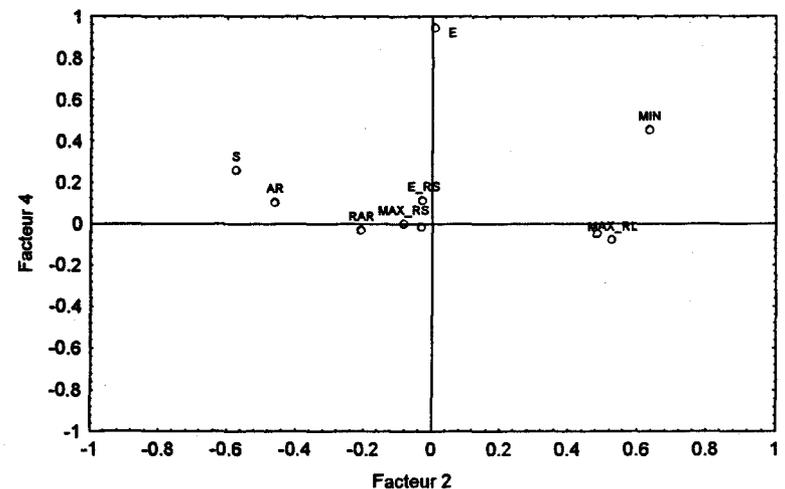


(c)

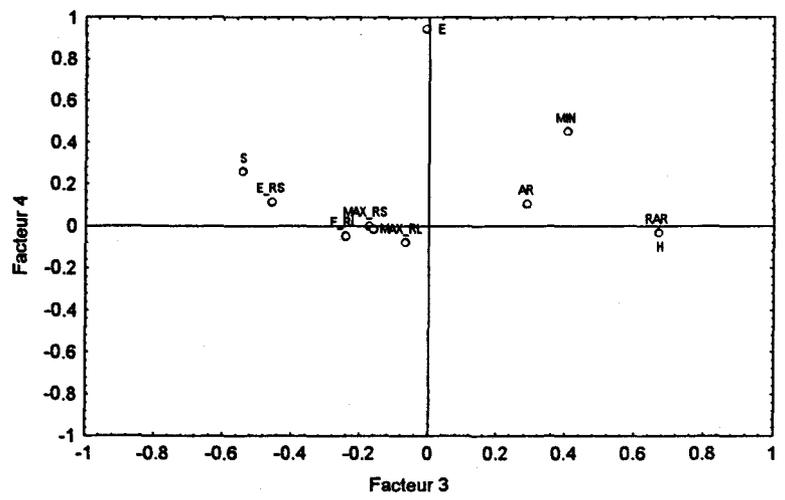
Figure 4.4 Position des critères dans l'espace des quatre premiers facteurs



(d)



(e)



(f)

Figure 4.4 Position des critères dans l'espace des quatre premiers facteurs

Les quatre premiers facteurs résument bien les aspects des séries qui nous intéressent: la répartition de l'eau dans le temps, la persistance dans la série, la quantité d'eau moyenne et la valeur minimale. Les quatre facteurs sont donc retenues pour la deuxième génération de séries synthétiques.

4.2.5 Génération - phase II

Une fois les poids factoriels déterminés et les facteurs à considérer choisis, la deuxième génération conduisant aux choix des séries pour déterminer la relation facteur/coût est effectuée. Rappelons que les quatre premiers facteurs sont retenues dans cette étude et qu'un grand nombre de scénarios doivent être générés dans ce cas. Une grande quantité de scénarios est donc générée et seulement ceux pour lesquels leur position dans l'espace factoriel est à l'intérieur de l'intervalle de tolérance pour trois des quatre axes sont retenues (voir étape 2, section 3.3). Ici, en supposant une distribution multinormale des scénarios dans l'espace factoriel, nous avons considéré que tous les scénarios compris dans une bande de probabilité 0.05 autour de zéro, c'est-à-dire entre les valeurs -0.063 et 0.063, étaient raisonnablement près de l'historique. Un nombre total de 2400 scénarios ont été générés et retenus de cette manière. Cela représente un nombre moyen de 600 scénarios par facteur variable.

Pour déterminer la relation facteur/coût, un nombre réduit de scénarios doivent être choisis pour ensuite être utilisés comme intrants du modèle MINERVE. Dans cette étude, 10 scénarios par facteurs (40 scénarios au total) ont été utilisés. Chaque facteur a été divisé en 10 classes équiprobables. Il y a donc environ 60 scénarios par classe. Dans chacune de ces classes, la série synthétique se trouvant la plus près de la médiane de la classe est sélectionnée, voir la figure 3.3 dans le chapitre 3.

4.2.6 Désagrégation des séries d'apport annuel agrégé

Les intrants du logiciel MINERVE sont les apports naturels mensuels à chacun des sites. Le logiciel SPIGOT (Grygier et Stedinger, 1990, 1991) a été utilisé afin, dans un premier temps, de désagréger les apports agrégés du bassin en apports mensuels du bassin et, dans un deuxième temps, les apports mensuels du bassin en apports mensuels à chacun des sites. Cette procédure permet de conserver l'autocorrélation d'un mois à l'autre et la corrélation spatiale d'un site à l'autre. La théorie concernant la désagrégation des apports a été présentée dans le Rapport 1 (1996) et n'est pas répétée ici.

4.3 Relation entre les facteurs et le coût d'exploitation

Cette section est consacrée à l'analyse de sensibilité du modèle MINERVE, c'est-à-dire la variation du coût en fonction des facteurs (étapes 5 à 6 de la section 3.3).

4.3.1 Détermination du coût d'exploitation

Le coût d'exploitation comprend à la fois le coût de construction, le coût d'achat d'énergie et le revenu de vente d'énergie. Nous avons souligné précédemment que les prix marginaux de vente et d'achat doivent être déterminés de façon la plus réaliste possible, ce qui nécessite généralement une étude approfondie du marché énergétique. Cependant, une telle étude ne fait pas partie de notre recherche. Nous avons exécuté MINERVE96 avec différents prix marginaux d'achat et de vente d'énergie. Il semble que le coût d'exploitation et le choix de l'aménagement en général sont très sensibles aux prix marginaux d'achat et de vente. Pour cette raison, dans cet exemple nous avons exécuté le modèle avec un prix d'achat très élevé et un prix de vente nul. Essentiellement, ceci revient à effectuer l'analyse de sensibilité avec l'ancienne version du modèle.

Le logiciel MINERVE offre la possibilité de résoudre le problème d'optimisation du coût à partir de trois valeurs initiales. La minimisation du coût étant un problème non convexe, les méthodes d'optimisation conventionnelles ne garantissent pas de trouver une solution optimale globale. En réalité, la solution trouvée est une solution optimale locale qui peut très bien ne pas être la solution optimale globale (pour plus de détail à ce sujet, voir Duquette et Lafond, 1992). L'utilisation de trois valeurs initiales augmente donc les chances de trouver la solution optimale globale.

Le logiciel MINERVE a été exécuté avec les 40 séries spécifiées à la section 4.1.5 et trois valeurs initiales différentes (un totale de 120 exécutions). Des trois solutions optimales obtenues pour un scénario synthétique donné avec les trois valeurs initiales, la plus petite a été retenue pour déterminer la relation facteur/coût. Le tableau 4.4 donnent les coûts obtenus en fonction des valeurs initiales et des facteurs. Il est important d'indiquer que les trois solutions optimales pour un même scénario peuvent varier très significativement dans certains cas. En fait, la variation du coût en fonction de la valeur initiale peut être aussi importante que celle en fonction des différents scénarios d'un facteur donné. L'interprétation de l'effet d'un facteur sur le coût devient alors difficile, car les variations observées peuvent provenir autant des différents scénarios que du choix de la valeur initiale.

Tableau 4.4 Coûts d'exploitation (en millions) en fonction des valeurs initiales et des facteurs

Scénario	Facteur 1			Facteur 2		
	Val. Initiale 1	Val. initiale 2	Val. initiale 3	Val. Initiale 1	Val. initiale 2	Val. initiale 3
1	4112	4112	4112	4116	4116	4234
2	4119	4120	4236	4232	4097	4098
3	4104	4104	4111	4118	4092	4091
4	4078	4097	4079	4065	4065	4190
5	4084	4084	4084	4132	4132	4245
6	4159	4088	4207	4248	4092	4217
7	4105	4100	4105	4079	4079	4085
8	4086	4078	4200	4210	4085	4093
9	4070	4055	4074	4088	4088	4088
10	4067	4067	4188	4207	4092	4206

	Facteur 3			Facteur 4		
	Val. initiale 1	Val. initiale 2	Val. initiale 3	Val. initiale 1	Val. initiale 2	Val. initiale 3
1	4090	4102	4202	-	-	-
2	4228	4099	4229	-	4122	4122
3	-	4073	4201	-	4092	-
4	4093	4093	4100	4093	4093	4217
5	4064	-	4199	4089	4089	4103
6	4086	4070	-	4203	4089	4089
7	4205	4089	4206	4210	4077	4197
8	4211	4075	4231	4102	4097	4215
9	4200	4079	4194	4176	4063	4222
10	4108	4108	4232	4080	4067	4079

En gras: valeur minimum

Certains scénarios n'ont pas conduit à de solution pour les facteurs 3 et 4 avec certaines valeurs initiales. En particulier, le scénario 1 du facteur 4 n'a pas donné de résultat pour aucune des trois valeurs initiales. Ces cas représentent des scénarios pour lesquels il n'est pas possible de satisfaire la demande en énergie.

4.3.2 Relation coût/facteur

Les graphiques 4.5a à 4.5d donnent respectivement la relation entre le coût d'exploitation et les facteurs 1, 2, 3 et 4.

Le graphique 4.5a montre qu'il y a une tendance en fonction des scénarios. Rappelons que le facteur 1 est relié aux caractéristiques des déficits. Les scénarios optimistes se trouvent à droite du graphique alors que les scénarios pessimistes sont à gauche.

La relation entre le coût et le facteur 2 est donnée à la figure 4.6b. Une baisse du coût en fonction des scénarios est quelque peu notable, mais aucune tendance nette n'est évidente dans le cas du

facteur 2 qui est relié à l'écart-type et à la valeur minimum. Le scénario 4 conduit à un coût relativement faible en comparaison des autres scénarios.

La figure 4.5c illustre le lien entre le coût et le facteur 3 identifié au coefficient de Hurst. Un comportement aléatoire prévaut dans ce cas-ci. Aucune relation ne semble donc exister entre les différents scénarios et le coût pour le facteur 3.

Enfin, le coût en fonction du facteur 4 est donné à la figure 4.5d. Le facteur 4 qui est relié à la moyenne a un effet sur le coût. Les scénarios optimistes sont situés à droite du graphique alors que les scénarios pessimistes se retrouvent à gauche. Mentionnons que le scénario 1 n'a pas donné de solution.

Les deux facteurs qui influent le plus le coût de construction sont les facteurs 1 et 4. Le premier de ces facteurs est identifié aux séquences des déficits et le second à la quantité d'eau. Ces deux facteurs sont retenus pour identifier les scénarios types qui seront utilisés dans la version stochastique de MINERVE (chapitre 5).

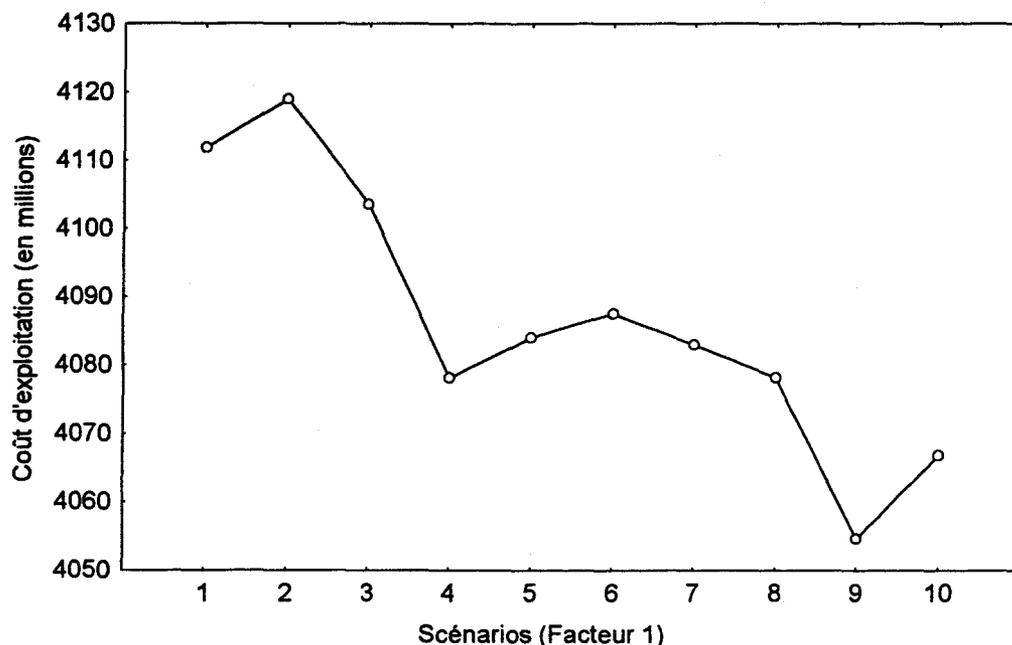


Figure 4.5a Coût d'exploitation en fonction du facteur 1

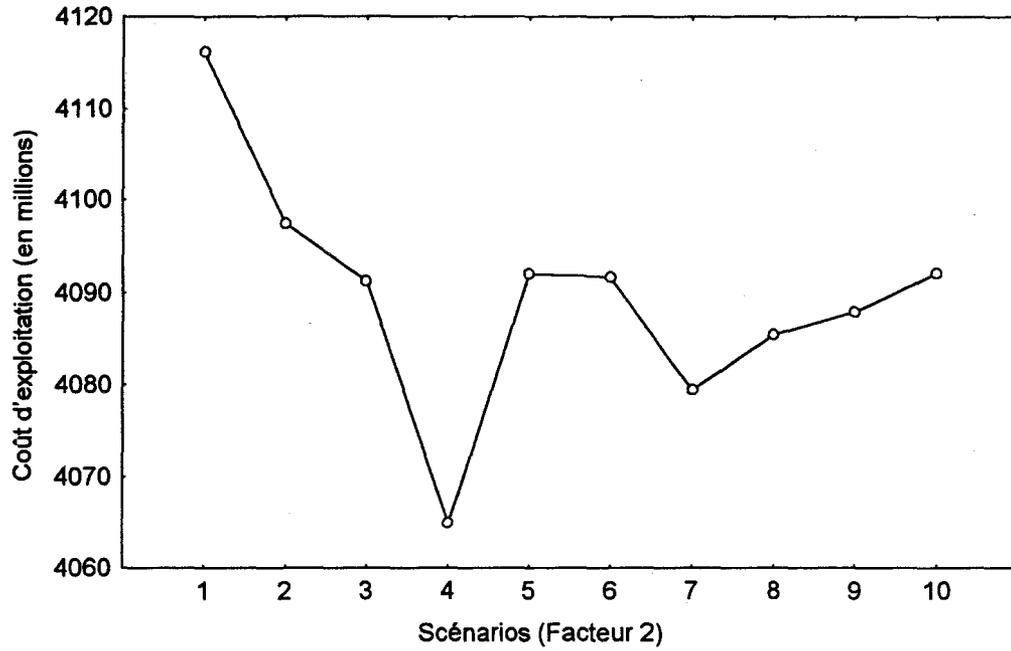


Figure 4.5b Coût d'exploitation en fonction du facteur 2

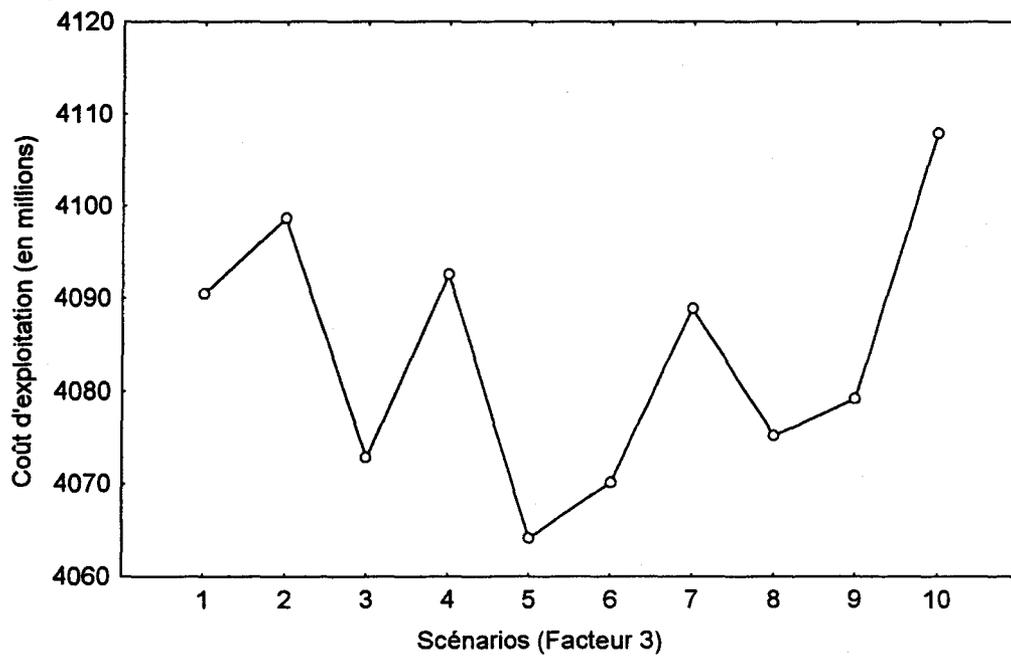


Figure 4.5c Coût d'exploitation en fonction du facteur 3

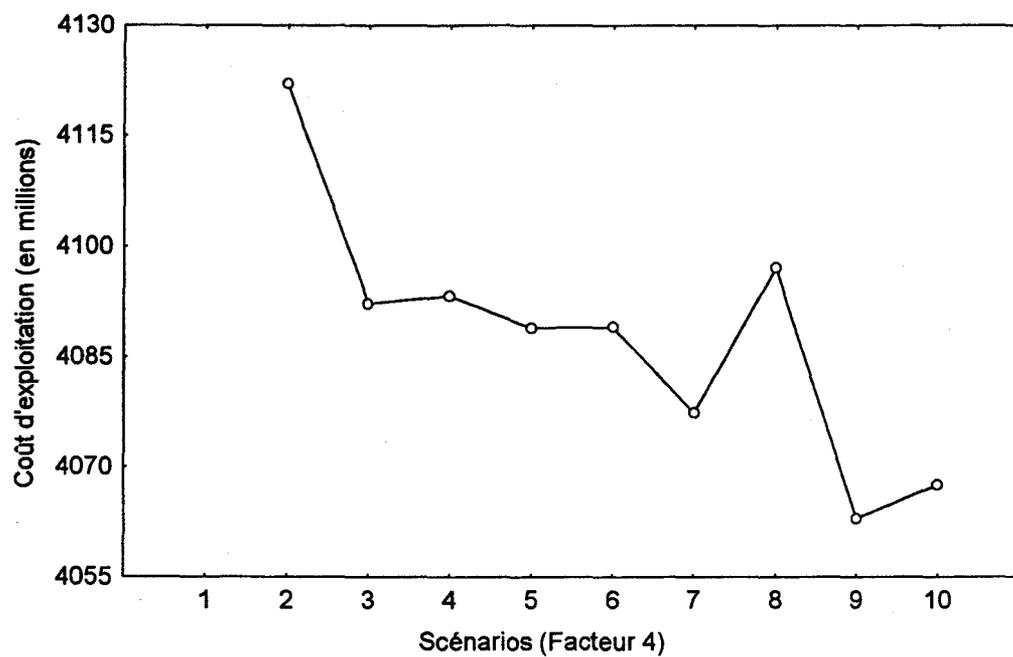


Figure 4.5d Coût d'exploitation en fonction du facteur 4

5 Exemple d'application de MINERVE96

Ce chapitre a pour but de donner un exemple d'application du modèle stochastique MINERVE96. Dans une première partie, des scénarios types sont choisis et, dans une deuxième partie, le modèle MINERVE96 est exécuté avec différentes valeurs du terme λ (équation 5) et avec différentes fonctions pour le coût d'achat d'énergie et de pour le revenu total de la vente d'énergie (équation 1).

5.1 Choix des scénarios types pour MINERVE96

Dans le chapitre 4, les facteurs 1 et 4 ont été identifiés comme étant les facteurs ayant un lien avec le coût d'exploitation d'une nouvelle vallée. La figure 5.1 montre 1000 séries synthétiques d'apports dans l'espace factoriel de ces deux facteurs. Par souci d'économie de temps, neuf scénarios ont été sélectionnés pour être employés comme intrants au modèle MINERVE96. Chacun des deux facteurs indépendants a été divisé en trois classes équiprobables selon l'hypothèse d'une distribution normale réduite. La probabilité associée à chacune des séries est donc 1/9. Ainsi, il y a environ 110 scénarios dans chacune des classes. Les lignes horizontales et verticales de la figure 5.1 illustrent ce classement. La médiane de chacune des classes a ensuite été déterminée. La série synthétique la plus proche du centre a été choisie comme scénario représentatif de cette classe. Les séries types sélectionnées pour chacune des classes sont indiquées par le symbole + dans la figure 5.1.

5.2 Résultats du modèle MINERVE96

Les neuf séries choisies à la section précédente ont servi d'intrants au modèle déterministe pour vérifier si elle conduisait à une solution réalisable. Les trois valeurs initiales ont été utilisées pour

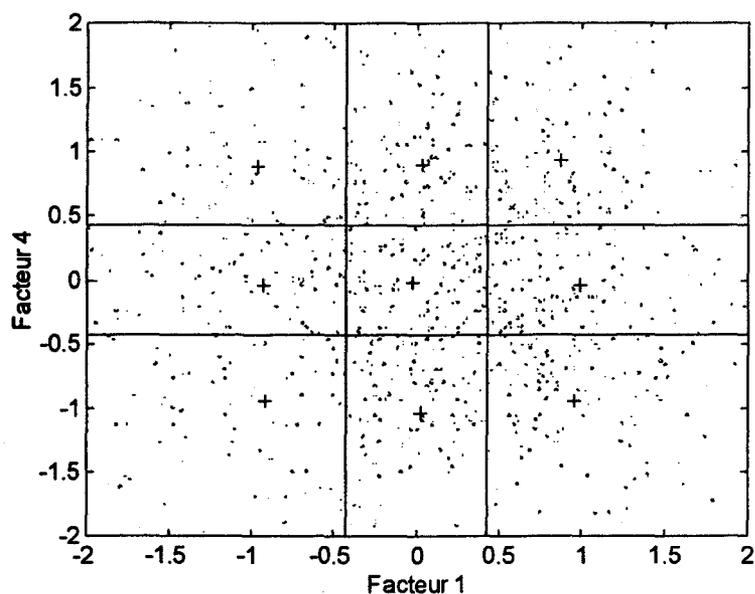


Figure 5.1 Position des séries synthétiques choisies (+) dans l'espace des facteurs 1 et 4

chacun des scénarios et la solution minimum a été retenue. Tous les neuf scénarios ont donné une solution. Le tableau 5.1 donne les résultats du modèle déterministe pour les neuf scénarios types.

Le modèle stochastique MINERVE96 a ensuite pu être utilisé pour déterminer le coût d'exploitation. Tout comme pour la version déterministe, le modèle a été premièrement exécuté sous la contrainte de ne permettre aucun achat d'énergie ni aucune vente. Dans ce cas-ci, la valeur du paramètre λ n'a aucune influence, car le terme de variance de l'équation (5) est nul. Ensuite, des fonctions simples de coût d'achat et de revenu de vente d'énergie ont été fixées afin d'examiner l'effet de différentes valeurs de lambda sur le coût d'exploitation. La fonction de coût d'achat d'énergie, CD, utilisé est:

$$CD = 2 C E$$

Tableau 5.1 Coût d'exploitation (en millions) pour les neuf séries types avec le modèle déterministe

Scénarios	Coût d'exploitation	Scénarios	Coût d'exploitation
1	4112	6	4066
2	4072	7	4103
3	4080	8	4075
4	4073	9	4069
5	4092		

où C est le coût moyen de la production en énergie du complexe déterminé à partir de l'historique et E est l'énergie. La fonction de revenu de la vente de l'énergie, RE , est:

$$RE = \frac{1}{2} C E$$

Les fonctions ainsi définies n'encouragent pas particulièrement l'achat d'énergie (l'énergie coûte 2 fois plus que la production) ni la vente d'énergie (l'énergie vaut deux fois moins que la production).

Trois valeurs de λ ont été considérées dans cette étude. Il s'agit des valeurs 0, 0.5 et 1. Lorsque la valeur de λ est égale à 1, un poids maximum est donné au terme de variance de l'équation (5). Lorsque la valeur de λ est nulle, le terme de variance est également nul. La valeur 0.5 est une valeur intermédiaire entre les deux extrêmes. Le tableau 5.2 donne les coûts d'exploitations pour chacune des valeurs de λ avec permission d'acheter et de vendre de l'énergie. Le coût d'exploitation sans permissions d'acheter et de vendre y est également donné. À titre comparatif, le coût d'exploitation de la série historique avec le modèle déterministe est également inclus.

L'approche stochastique sans achat et vente d'énergie conduit à un coût d'exploitation plus grand que l'approche déterministe avec la série historique. Dans le modèle stochastique, l'aménagement obtenu est optimum pour un ensemble de neuf scénarios. Ces scénarios constituent une gamme réaliste d'apport naturel qui sont susceptibles de se produire dans l'avenir. Certains de scénarios sont pessimistes par rapport à l'historique alors que d'autres sont optimistes. Il est donc normal que la prise en compte de ces divers scénarios dans le modèle stochastique donne un coût plus élevé que le modèle déterministe avec la série historique. En effet, l'aménagement obtenu avec le modèle stochastique va satisfaire la demande pour une grande variété de scénarios d'apport.

Tableau 5.2 Coût d'exploitation (en millions)
déterminé à partir de MINERVE96

Modèle	Achat et vente d'énergie	λ	Coût d'exploitation
Déterministe (historique)	Interdit	-	4107
Stochastique	Interdit	-	4218
Stochastique	Permis	0.0	3097
Stochastique	Permis	0.5	3211
Stochastique	Permis	1.0	3211

Les coûts d'exploitation obtenus avec permission d'acheter et de vendre sont beaucoup plus faibles que ceux sans permission d'acheter et de vendre. Dans la première situation, le modèle MINERVE96 conduit à l'aménagement de la centrale GB1 et des réservoirs GB1 et Lac Bienville seulement. Naturellement, avec ces installations, il n'est pas possible de satisfaire la demande. La différence est comblée par l'achat d'électricité sur le marché extérieure d'énergie. Dans la deuxième situation, les trois centrales GB1, GB2 et GB3 et les quatre réservoirs GB1, GB2, GB3 et Lac Bienville sont construits pour satisfaire la demande, ce qui conduit à un coût plus élevé. Les fonctions qui déterminent l'achat et la vente de l'énergie ont donc une influence très grande sur l'aménagement et le coût d'exploitation.

Le paramètre λ qui pondère le terme de variance de l'équation (5) a un effet sur le coût d'exploitation. Lorsque la valeur de λ est égale à 1 (grand poids à la variance) le coût d'exploitation est plus élevé que lorsqu'elle est égale à 0 (aucun poids à la variance). Les valeurs 1 et 0.5 du paramètre λ conduisent à la même solution optimale. Le terme de moyenne de l'équation (5) semble donc être de moindre amplitude que le terme de variance. Ainsi, un poids de 0.5 donne une grande importance au terme de variance aux dépens du terme de moyenne qui devient négligeable. Cela explique le fait que le coût soit identique pour $\lambda = 0.5$ et $\lambda = 1$. En conclusion, tout comme les fonctions d'achat et de vente, le paramètre λ a un impact important sur l'aménagement et le coût d'exploitation.

6 Conclusion

Dans le présent rapport, nous avons développé davantage la méthodologie proposée par Rasmussen *et al.* 1996 pour le choix des séries types d'apport à considérer dans la nouvelle version du logiciel MINERVE. La contribution de la deuxième phase du projet concerne plus particulièrement une procédure d'analyse de sensibilité permettant d'établir le lien entre le coût d'exploitation et différents aspects des scénarios d'apport. Puisqu'il est nécessaire de réduire la dimensionnalité des apports à 1-3 critères (ou facteurs), il est essentiel que ceux-ci soient choisis de manière optimale et rationnelle. La méthodologie proposée répond à notre avis très bien à ces exigences.

Nous avons effectué plusieurs exécutions de la nouvelle version de MINERVE en considérant le cas de l'aménagement du complexe Grande Baleine. Nous nous sommes rendu compte que certains paramètres ont une influence très importante sur l'aménagement optimal, par exemple les prix marginaux d'achat et de vente d'énergie. Dans le cadre du présent projet, on n'a pas pu déterminer ces prix de façon réaliste, et pour cette raison les résultats obtenus avec MINERVE96 sont présentés à titre d'illustration uniquement. Une application pratique de MINERVE96 requiert une étude approfondie de chacun des paramètres présents dans le fichier d'intrant au modèle.

7 Bibliographie

Duquette, R. (1996) Description de la nouvelle modélisation retenue pour le problème lié à l'aménagement optimal d'un nouveau bassin, Rapport IREQ-96-351.

Duquette, R. et Lafond, L. (1992) Une nouvelle approche d'aménagement optimal d'un bassin hydraulique pour fin de production d'électricité, Rapport IREQ-92-171.

Filliben, J.J. (1975) The probability plot correlation coefficient test for normality, *Technometrics*, 17(1): 111-117.

Grygier, J.C. et Stedinger, J.R. (1990) SPIGOT. A Synthetic Streamflow Generation Software Package, Technical Description, Version 2.6, Cornell University, Ithaca, New York.

Grygier, J.C. et Stedinger, J.R. (1991) SPIGOT. A Synthetic Streamflow Generation Software Package, User's Manual, Version 2.6, Cornell University, Ithaca, New York..

Rapport 1 (1996) Même que Rasmussen *et al.* (1996).

Rasmussen, P.F., Fagherazzi, L., Duquette, R., Legendre, P. et Perreault, L. (1996) Détermination de séries mensuelles types pour le modèle MINERVE - Phase 1, Rapport de recherche No. R-480, INRS-Eau, Université du Québec.