

Analyse bayésienne rétrospective d'une rupture dans les séquences de variables aléatoires hydrologiques

Luc Perreault

(B.Sc., Statistique, Université Laval, 1987)
(M.Sc., Mathématiques, Université Laval, 1991)

Thèse en cotutelle présentée

au laboratoire de Gestion du Risque en Sciences de l'Eau (GRESE)

et à l'Institut National de la Recherche Scientifique (INRS-Eau)

pour l'obtention du grade de

Docteur de l'École Nationale du Génie Rural des Eaux et des Forêts (France)

en Statistique Appliquée aux Sciences de l'Eau

et

de Philosophiae Doctor (Ph.D.) en Sciences de l'Eau

de l'Université du Québec (Canada)

Jury d'évaluation

Président du jury	Roman Krzysztofowicz Faculty of Engineering-Systems University of Virginia (USA)
Examineur externe	Christian P. Robert Laboratoire de Statistique CREST INSEE (France)
Examineur externe	Philippe Capéraà Département de Mathématiques et Statistique Université Laval (Canada)
Examineur interne	Jacques Bernier ENGREF et INRS-Eau
Directeurs de thèse	Eric Parent ENGREF Laboratoire GRESE Bernard Bobée INRS-Eau Chaire en Hydrologie Statistique Hydro-Québec/CRSNG/Alcan

Thèse soutenue le 29 septembre 2000

Résumé

Dans ce travail, nous défendons la thèse selon laquelle la statistique bayésienne permet le développement et la mise en oeuvre d'outils statistiques adaptés et opérationnels pour l'analyse de la non-stationnarité dans les séries temporelles hydrométéorologiques. Nous proposons d'abord des modèles unidimensionnels de rupture à base de lois a priori conjuguées. Nous montrons que l'approche bayésienne permet de caractériser non seulement le changement mais aussi le développement d'une procédure de détection de rupture en ayant recours aux facteurs de Bayes. Ensuite, ces modèles locaux sont généralisés au cas multisite en considérant un modèle à structure hiérarchique dans lequel est introduit un effet aléatoire temporel. Ce modèle à trois niveaux avec variables latentes permet de prendre en compte une partie des corrélations spatiales entre les observations tout en laissant la liberté à l'instant de rupture de varier d'un site à l'autre. De structure plus complexe, la mise en oeuvre de ce modèle fait intervenir l'échantillonnage de Gibbs, outil de simulation Monte Carlo par chaînes de Markov (MCMC).

L'ensemble des modèles présentés dans cette thèse sont appliqués à des séries hydrologiques québécoises. Ces applications permettent d'attirer l'attention des hydrologues sur la possibilité qu'un changement soit survenu au Québec et donc sur la nécessité de s'adapter à ce nouveau contexte. C'est pourquoi nous présentons aussi une illustration de l'approche bayésienne décisionnelle pour le choix du débit d'équipement d'une centrale hydroélectrique sous incertitude de rupture. Non seulement les incertitudes des paramètres et des hypothèses des modèles sont prises en compte, mais aussi les conséquences des actions possibles au travers d'une fonction de coût réaliste établie avec des hydrologues experts. Ces expérimentations ont montré la souplesse et la portée des techniques présentées dans cette thèse.

Mots clés : Rupture, changement climatique, statistique bayésienne, hydrologie, analyse décisionnelle, modèle hiérarchique, variable latente, échantillonnage de Gibbs.

Remerciements

Mes remerciements vont d'abord à mon directeur de thèse de l'ENGREF, Monsieur Éric Parent. Je le remercie d'avoir accepté de diriger mes travaux le jour où j'ai enfin décidé de retourner sur les bancs d'école. Je tiens à lui exprimer ma profonde gratitude pour sa disponibilité hors du commun et pour ses vastes connaissances en statistique dont il m'a fait bénéficier. Durant mon séjour en France (juin 1997 à octobre 1999), son contact au fil des jours a été pour moi une source constante de stimulation. Je lui dois d'avoir cristallisé mon intérêt pour la statistique bayésienne et l'analyse décisionnelle. Merci Éric de m'avoir fait confiance. Merci aux amis de l'ENGREF pour leur accueil si chaleureux.

De plus, ma reconnaissance s'adresse à mon directeur de thèse de l'INRS-Eau, Monsieur Bernard Bobée. L'expérience acquise en tant qu'agent de recherche dans son équipe durant sept années m'a été d'un grand recours pour la réalisation de ces travaux. Monsieur Bobée a su, par sa riche expérience en hydrologie, ramener les statisticiens que nous sommes, Éric et moi, vers le droit chemin... Les équations c'est bien, mais il faut que ça serve !

J'éprouve une reconnaissance toute particulière pour Monsieur Jacques Bernier, statisticien et hydrologue de renom. J'ai eu la chance de pouvoir compter sur son appui scientifique. Je tiens à le remercier pour sa disponibilité, son enthousiasme et la multitude de conseils techniques qu'il m'a prodigués.

Je remercie Monsieur Michel Slivitzky, professeur émérite à l'INRS-Eau. Ces conseils d'hydrologue expert (climatologue aussi !) ont grandement orienté mes travaux. À quoi bon faire un modèle statistique s'il ne repose pas sur les connaissances physiques du phénomène étudié ?

Un grand merci à Madame Marie-Claude Roland, responsable de la formation en rédaction scientifique de l'INRA de Paris, et Madame Nathalie Nadeau, rédactrice professionnelle. Leurs conseils ont rendu la rédaction de ma thèse plus difficile, mais combien plus intéressante. Si le lecteur trouve que les titres sont informatifs, que mon exposé est clair et structuré bref, que le « message est passé », c'est grâce à elles.

Je ne peux passer sous silence toutes les discussions fructueuses que j'ai eues avec mes amis et collègues thésards de l'ENGREF, Philippe Girard et Keltoum Chaouche. C'est souvent devant un express que nous avons résolu ensemble certains problèmes techniques de nos thèses respectives. Cher cousin, chère cousine, merci et bonne chance !

Je souhaite exprimer toute ma gratitude à Guy et à Monique, mes parents. Tout au long de mes études, j'ai bénéficié de leurs encouragements et de leur soutien. Vous m'avez accompagné, je vous en remercie.

Merci Lucien, Ali, Gabriel, Nicole, Sandrine, Julie, Jocelyne et Jean-Jacques, Denise, la gang du Bistrot le Tivoli à Monpartnasse. Un GRAND merci à Martyne. Merci René, Raymond, Vincent, Mario et Taha. Merci, Sophie, Chris, Jérémy, Janine, Mariette, Gilles, Fernande, Sonya, Cathou, Marie-Ève. Si vous ne savez pas pourquoi, moi je le sais.

Pour la réalisation de cette thèse, j'ai reçu des bourses de l'ENGREF, des fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec (FCAR), du ministère des Relations Internationales du Québec (MRI) et de l'INRS-Eau. Je souhaite remercier ces organismes de leur soutien financier.

*À Nathalie
et à ce qui se cache
dans son ventre*

Table des matières

Résumé	iii
Remerciements	v
Table des figures	xv
Liste des tableaux	xviii
1 Introduction	1
2 Et si les régimes hydrologiques étaient en cours de changements?	5
2.1 Les phénomènes hydrométéorologiques peuvent être non stationnaires . . .	5
2.2 Le gestionnaire doit prendre ses décisions dans un contexte risqué	11
2.3 L'hydrologue est démuni face à l'incertitude de changement	13
2.4 Le statisticien peut construire des solutions en cohérence avec la demande du gestionnaire	17
2.4.1 Ma stratégie: adopter la perspective bayésienne	17
2.4.2 Les choix qui définissent le cadre d'étude de ma thèse	20
3 Mettre en oeuvre l'analyse bayésienne dans un modèle de rupture en variance	23
3.1 Modèle à une seule rupture	24
3.1.1 Représentation des quantités observables	24
3.1.2 Représentation des quantités non observables: lois a priori conjuguées	26
3.2 Estimation des paramètres conditionnelle à l'existence d'une rupture en variance	28
3.2.1 Changement simultané de moyenne et de variance	28
3.2.2 Changement de variance seulement	34
3.3 Application à l'analyse locale d'une rupture dans une séquence d'apports énergétiques annuels	36
3.3.1 Application au complexe de Churchill Falls	39
3.3.2 Application au complexe de l'Outaouais	41
3.3.3 Discussion	43

4	Mettre en doute l'existence et le type de la rupture	45
4.1	Détection d'une rupture et identification de son type	46
4.2	Application aux séries d'apports énergétiques annuels de Churchill Falls et de l'Outaouais	50
4.3	Discussion	54
5	Et s'il y a changement climatique, n'y aurait-il pas impact régional?	57
5.1	Modèle multisite à rupture simultanée (MM_1)	58
5.1.1	Modèle et résultats techniques	58
5.1.2	Application du modèle MM_1 à des séries d'apports naturels pour des sites du nord-est québécois	62
5.2	Échangeabilité et modèle hiérarchique de rupture	68
5.2.1	Echangeabilité des instants de rupture	69
5.2.2	Modèle hiérarchique de rupture et stratégie d'estimation	70
5.3	Modèle multisite à rupture variable sans effet de covariation spatiale (MM_2)	75
5.3.1	Modèle et notation	75
5.3.2	Lois conditionnelles complètes	77
5.3.3	Application du modèle MM_2 à des séries d'apports naturels pour des sites du nord-est québécois	79
5.4	Modèle multisite à rupture variable avec effet de covariation spatiale (MM_3)	85
5.4.1	Modèle et notation	86
5.4.2	Lois conditionnelles complètes	87
5.4.3	Application du modèle MM_3 à des séries d'apports naturels pour des sites du nord-est québécois	89
5.5	Discussion	96
6	De l'inférence à la prise de décision : l'aménagement d'une centrale hydroélectrique	97
6.1	Problématique étudiée: la centrale de Touloustouc	98
6.2	Fonctions économiques	100
6.2.1	Fonction de coût	100
6.2.2	Fonction des revenus anticipés	100
6.2.3	Fonction des bénéfices et des pertes	103
6.3	Analyse bayésienne complète	104
6.3.1	Analyse décisionnelle déterministe	104
6.3.2	Analyse décisionnelle avec prise en compte des incertitudes	105
6.4	Discussion	108
7	Conclusions et perspectives	111
	Bibliographie	116
	Annexe	122

A	Article I	123
A.1	Introduction	123
A.2	Bayesian analysis of normal sequences with an unknown change-point . . .	126
A.2.1	A single change in the mean (Model M_1)	127
A.2.2	A single change in the variance (Model M_2)	128
A.3	Implementation of the Gibbs sampler for model M_2	130
A.4	Applications	132
A.4.1	Specifying prior distributions	132
A.4.2	Churchill Falls power system	134
A.4.3	Outaouais power system	137
A.5	Discussion and conclusions	139
B	Article II	145
B.1	Introduction	145
B.2	Inference for univariate single change-point models	147
B.3	Entertaining a range of possible models for change-point analysis	149
B.3.1	Definitions	149
B.3.2	Comparison of change-point models	150
B.3.3	Accounting for model uncertainty when forecasting	152
B.4	Evaluations of predictive densities	153
B.4.1	Posterior predictive density	153
B.4.2	Prior predictive density	154
B.5	Applications assuming normal change-point models	155
B.5.1	The data	155
B.5.2	Specifying prior distributions	156
B.5.3	Inference about the existence and the configuration of a change . .	157
B.5.4	Predictive analysis	160
B.6	Discussion and conclusions	163
C	Densités de probabilité et principaux moments	171
D	Échantillonnage de Gibbs	175
D.1	Algorithme de Gibbs	175
D.2	Estimation à partir de l'échantillonnage de Gibbs	176
D.3	Éléments d'analyse de la convergence pour les modèles MM_2 et MM_3 . . .	177
E	Article III	185
E.1	Introduction	185
E.2	The multivariate normal change-point model	186
E.3	Detection of a change-point	189
E.4	Application	191
E.4.1	Specifying prior distributions	193
E.4.2	Results and discussion	194
E.5	Conclusion	196

Table des figures

2.1	Débits annuels du Nil en m^3/s	6
2.2	Niveaux moyens annuels du lac Victoria	6
2.3	Précipitations et débits annuels pour divers sites sahéliens	8
2.4	Précipitations annuelles pour les lacs Ontario et Erié	9
2.5	Valeur de l'indice NAO pour le mois de mars, de 1940 à 1995 (Kiely, 1999)	10
2.6	Apports naturels annuels observés à Churchill Falls	10
2.7	Apports naturels annuels observés au réservoir Sainte-Anne	14
2.8	Schématisation des étapes fondamentales de l'analyse bayésienne	18
3.1	Apports énergétiques annuels des huit grands complexes hydroélectriques gérés par Hydro-Québec	37
3.2	Apports énergétiques annuels à Churchill Falls, densité marginal a priori (trait pointillé) et a posteriori (trait continu) des paramètres	40
3.3	Résidus standardisés de Churchill Falls sur papier de probabilité normal . .	41
3.4	Apports énergétiques annuels de l'Outaouais, densité marginale a priori (trait pointillé) et a posteriori (trait continu) des paramètres	42
3.5	Résidus standardisés pour le complexe Outaouais sur papier de probabilité normal	43
4.1	Sensibilité du facteur de Bayes du changement contre le non changement aux variances a priori des apports énergétiques moyens	54
5.1	Positions géographiques des huit sites du nord-est québécois	63
5.2	Apports naturels moyens de février à juillet pour huit bassins du nord-est québécois ($l/(km^2 \times s)$)	64
5.3	Densité marginale a priori et a posteriori de l'année de rupture pour huit bassins du nord-est québécois	66
5.4	Intervalle de crédibilité à 95 pour cent des moyennes avant et après la rupture pour huit bassins du nord-est québécois	67
5.5	Représentation schématique d'un modèle hiérarchique de rupture	71
5.6	Détermination des lois conditionnelles complètes a posteriori pour le modèle hiérarchique de rupture	74
5.7	Représentation schématique du modèle hiérarchique de rupture MM2 . . .	77

5.8	Densité et fonction de répartition marginales a posteriori de π_{1984} pour le modèle MM2	81
5.9	Densités marginales a posteriori de ϕ_1 (trait pointillé) et ϕ_2 (trait continu) pour MM2	82
5.10	Intervalles de crédibilité à 95 pour cent des moyennes avant et après la rupture pour huit bassins du nord-est québécois (modèle MM2)	83
5.11	Densités marginales a posteriori des instants de rupture pour le modèle MM2	83
5.12	Résidus espérés du modèle MM2 sur papier de probabilité normal	84
5.13	Représentation schématique du modèle hiérarchique de rupture MM3	88
5.14	Densité et fonction de répartition marginales a posteriori de π_{1984} pour le modèle MM3	91
5.15	Densités marginales a posteriori de ϕ_1 (trait pointillé) et ϕ_2 (trait continu) pour MM3	91
5.16	Intervalles de crédibilité à 95 pour cent des moyennes avant et après la rupture pour huit bassins du nord-est québécois (modèle MM3)	92
5.17	Densités marginales a posteriori des instants de rupture pour le modèle MM3	93
5.18	Distributions a posteriori de l'autocorrélation d'ordre 1 pour chacun des bassins versants	94
5.19	Résidus espérés du modèle MM3 sur papier de probabilité normal	95
6.1	Situation géographique de la future centrale sur la rivière Toulnostouc	98
6.2	Apports naturels annuels observés au réservoir Sainte-Anne	99
6.3	Fonction de coût pour l'aménagement de la centrale sur la rivière Toulnostouc	101
6.4	Schéma de la fonction d'énergie pour l'aménagement d'une centrale	102
6.5	Production annuelle calculée pour un module de 186 m ³ /s	102
6.6	Fonction de revenu pour l'aménagement de la centrale sur la rivière Toulnostouc	103
6.7	Fonctions de bénéfice et de perte pour l'aménagement de la centrale sur la rivière Toulnostouc	103
6.8	Apports naturels du réservoir Sainte-Anne et densité marginale a posteriori des paramètres du modèle M1	106
6.9	Perte espérée en fonction du débit d'équipement	108
D.1	Simulations des paramètres régionaux des 4000 premières itérations pour le modèle MM2	178
D.2	Simulations des paramètres locaux des 4000 premières itérations pour le modèle MM2	179
D.3	Simulations des paramètres locaux des 4000 premières itérations pour le modèle MM2 (suite)	180
D.4	Simulations des paramètres régionaux des 4000 premières itérations pour le modèle MM3	181
D.5	Simulations des paramètres locaux des 4000 premières itérations pour le modèle MM3	182

D.6 Simulations des paramètres locaux des 4000 premières itérations pour le modèle MM2 (suite)	183
--	-----

Liste des tableaux

3.1	Espérances et variances a posteriori pour les paramètres du modèle de rupture simultanée en moyenne et en variance	33
3.2	Espérances et variances a priori des paramètres pour les complexes de Churchill Falls et de l'Outaouais	38
3.3	Caractéristiques a posteriori des paramètres pour le complexe de Churchill Falls	39
3.4	Caractéristiques a posteriori des paramètres pour le complexe de l'Outaouais	42
4.1	Pertes encourues pour tester l'existence d'une rupture	49
4.2	Interprétation du facteur de Bayes (tiré de Kass et Raftery (1995))	50
4.3	Facteurs de Bayes et probabilités a posteriori pour les complexes de Churchill Falls et de l'Outaouais	52
4.4	Facteurs de Bayes pour la détection d'une rupture et l'identification de son type pour les complexes de Churchill Falls et de l'Outaouais	52
5.1	Superficies des bassins versant et périodes d'échantillonnage	63
5.2	Espérances et écarts-types a posteriori des moyennes de part et d'autre de l'année de rupture pour les huit bassins versant du nord-est québécois	67
5.3	Espérances et écarts-types a posteriori des moyennes de part et d'autre de l'année de rupture pour le modèle MM2	82
5.4	Espérances a posteriori et intervalles de de crédibilité à 95 pour cent des corrélations spatiales des résidus du modèle MM2	85
5.5	Espérances et écarts-types a posteriori des moyennes de part et d'autre de l'année de rupture pour le modèle MM3	92
5.6	Espérances a posteriori et intervalles de de crédibilité à 95 pour cent des corrélations spatiales des résidus du modèle MM3	94
5.7	Intervalle de crédibilité à 95 pour cent pour le coefficient d'autocorrélation d'ordre 1	95
6.1	Coût global de construction en dollars constants de 1999 pour divers débits d'équipement	100
6.2	Production annuelle calculée pour un module de 186 m ³ /s	102
6.3	Débit d'équipement optimal en contexte incertain et pertes encourues pour l'aménagement de la centrale Tounustouc	107

D.1 Valeurs du R de Gelman et al. (1995) pour les paramètres du modèle MM3 184

Chapitre 1

Introduction

L'eau constitue une ressource énergétique majeure pour l'économie du Québec et de la France. L'énergie produite au Québec est environ à 95% d'origine hydroélectrique; en France, 75% de l'électricité est d'origine nucléaire, mais le turbinage hydroélectrique confère au système l'indispensable souplesse pour s'adapter aux variations de la demande. Sous l'effet de facteurs naturels et socio-économiques (évolution climatique, augmentation de la demande, contraintes économiques et environnementales, déréglementation des marchés de l'énergie, etc.) cette ressource devient rare et onéreuse. Une gestion efficace représente donc des enjeux considérables. Elle requiert toutefois une connaissance adéquate des observations hydrométéorologiques, notamment en ce qui concerne le caractère incertain de leur évolution temporelle. Ces données et les modèles qui les représentent conditionnent le processus décisionnel. Actuellement, ce type d'incertitude n'est pas considéré de manière rigoureuse dans le processus décisionnel pour l'aménagement de structures hydrauliques et la gestion des réserves déjà constituées : on fait généralement l'hypothèse que la structure des aléas d'aujourd'hui reste la même que celle d'hier. En effet, les hydrologues, qui interviennent à titre de conseil en matière de risque auprès du gestionnaire, s'appuient souvent sur des outils statistiques conventionnels qui supposent la stationnarité des séries de débits. Or, de nombreuses études récentes concluent de manière convergente à l'existence de changements climatiques qui peuvent avoir un impact sur la disponibilité de la ressource en eau.

Nous nous intéressons au cas particulier d'un seul changement dans les paramètres de la distribution de débits annuels qui survient subitement à un instant inconnu. Cette thèse, essentiellement méthodologique, a pour objet de développer des outils statistiques afin de quantifier l'incertitude de changement dans la structure aléatoire des séries de débits annuels et de démontrer leur intérêt opérationnel dans un contexte décisionnel. Pour ce faire, nous adoptons l'approche bayésienne. Nous verrons que la statistique bayésienne, avec l'échantillonnage de Gibbs, permet de résoudre les difficultés associées à l'estimation des paramètres de modèles de changement et à la détection de ruptures. De plus, nous montrons que les résultats obtenus peuvent ensuite être intégrés dans le processus décisionnel et conduire ainsi au développement d'une approche rationnelle complète pour tenir compte de l'effet des incertitudes de rupture lors de la prise de décision.

La suite de ce document est divisée en cinq chapitres. Le Chapitre 2 a pour objet de donner les premiers éléments de formalisation de la question de recherche. On verra que cette dernière se situe à la frontière de la statistique mathématique et de l'hydrologie opérationnelle et qu'elle est issue de l'insuffisance des méthodes employées actuellement par l'hydrologue pour analyser la non-stationnarité des séries hydrométéorologiques.

Le Chapitre 3 établit dans un premier temps le cadre statistique de la thèse. On y décrit le modèle général ainsi que les hypothèses sous-jacentes. Nous faisons également une revue de synthèse des principaux acquis concernant l'analyse bayésienne d'une rupture dans une seule séquence de variables aléatoires. Nous décrivons ensuite la mise en oeuvre de l'analyse bayésienne et l'illustrons en considérant deux types de changement dans une séquence de variables aléatoires indépendantes et distribuées selon une loi normale. Il s'agit de la rupture simultanée en moyenne et en variance (modèle M_3) et de la rupture en variance seulement (modèle M_2). Le premier modèle permet une dérivation analytique des lois a posteriori si l'on considère des lois a priori conjuguées. Le second, par contre, n'admet pas de solution explicite. Nous montrons que l'échantillonnage de Gibbs est particulièrement adapté pour résoudre ce type de problème.

Dans le Chapitre 4, nous mettons en doute le fait qu'une seule configuration de changement peut survenir dans une série d'observations hydrométéorologiques. On y développe une procédure permettant de tester l'existence d'un changement et, le cas échéant, d'en identifier le type. Notre approche repose sur la sélection bayésienne de modèles. Pour illustrer la méthode, nous confrontons les deux modèles de changement de variance avec le seul modèle actuellement considéré en hydrologie (la rupture en moyenne, modèle M_1) et l'hypothèse de stationnarité (modèle M_0). Les développements théoriques présentés dans les chapitres 3 et 4 ont fait l'objet de publications scientifiques. Ainsi, nous insistons seulement sur les aspects conceptuels et sur les principaux résultats des applications. Le lecteur intéressé par les détails techniques est invité à consulter les articles I et II qui se trouvent en Annexe A et B de la thèse.

Au Chapitre 5, nous proposons trois modèles multidimensionnels permettant l'analyse régionale multisite d'une rupture. Le premier modèle (modèle MM_1) suppose que le changement survient simultanément pour l'ensemble des séries. Il s'agit d'une simple généralisation du modèle M_1 , et seuls les idées essentielles et les principaux résultats y sont présentés. Les détails sont décrits dans un article publié qui figure en Annexe E (Article III). Les deux autres modèles régionaux (modèles MM_2 et MM_3), moins restrictifs et plus réalistes, supposent que la date de rupture varie d'une série à l'autre. De structures plus complexes, ces modèles hiérarchiques ne permettent pas d'évaluer explicitement les distributions a posteriori des paramètres. Nous avons donc recours à l'échantillonnage de Gibbs.

Au Chapitre 6, nous adoptons le point de vue du gestionnaire en montrant comment les procédures statistiques proposées peuvent être utiles pour la planification d'un projet hydroélectrique. Le cas d'étude considéré est le problème de l'aménagement d'une centrale hydroélectrique d'Hydro-Québec. Il s'agit de la centrale de Toulmoustouc dont la construction devrait débuter en 2002. Enfin, dans le dernier chapitre, nous présentons la synthèse des principaux résultats, nous discutons des limites d'application de nos méthodes et

explorons certaines pistes de recherche qui nous apparaissent prometteuses.

Chapitre 2

Et si les régimes hydrologiques étaient en cours de changements ?

L'insuffisance des méthodes statistiques employées actuellement par l'hydrologue pour mettre en évidence et caractériser la non-stationnarité des séries hydrométéorologiques a motivé ma question de recherche. En effet, l'hydrologue ne peut alors fournir au gestionnaire tous les éléments rationnels nécessaires à la prise en compte, lors du processus décisionnel, d'un éventuel changement de régime dans les observations. Ma problématique de recherche se situe donc à l'interface de la statistique mathématique et de l'hydrologie opérationnelle. Ce chapitre en donne les premiers éléments de formalisation.

2.1 Les phénomènes hydrométéorologiques peuvent être non stationnaires

Jusqu'à ce jour aucun consensus ne s'est dégagé au sein de la communauté scientifique concernant l'origine de changements dans la moyenne de séries hydrométéorologiques. Néanmoins, de nombreuses études tendent à démontrer leur existence. Dans la littérature hydrologique, les types de changements de moyenne observés dans les séries peuvent être classés en deux grandes catégories : les ruptures et les tendances.

Définition 1 *Au sens large, le terme de « rupture » correspond à une modification dans les propriétés d'un processus aléatoire qui survient subitement à l'échelle de la période d'échantillonnage, étant entendu que ces propriétés demeurent stables de part et d'autre de l'année de rupture.*

Définition 2 *Le terme « tendance » fait référence au changement dans les propriétés d'un processus aléatoire qui survient graduellement à l'échelle de la période d'échantillonnage.*

Les études de non-stationnarité les plus connues en hydrologie concernent les séquences de données africaines, comme la série du débit annuel du Nil en amont du barrage d'Assouan. À notre connaissance, Cobb (1978) a été le premier à diagnostiquer une rupture

de la moyenne de ces observations survenant en 1898. Puisque le barrage Assouan a été mis en service en 1902, cet auteur a suggéré que la rupture serait davantage due à un changement de climat qu'à l'effet de la mise en place de la structure hydraulique. La Figure 2.1 présente les débits annuels du Nil pour la période allant de 1871 à 1970.

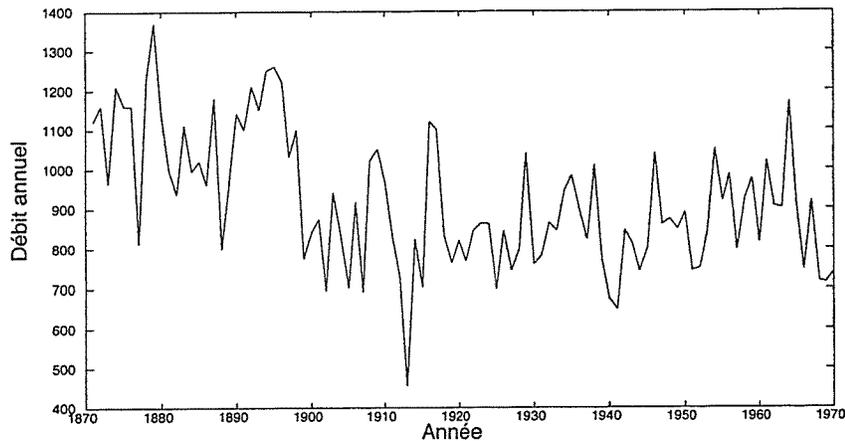


FIG. 2.1 – Débits annuels du Nil en m^3/s

Les niveaux mensuels du lac Victoria (Figure 2.2) ont eux aussi suscité de nombreuses interrogations.

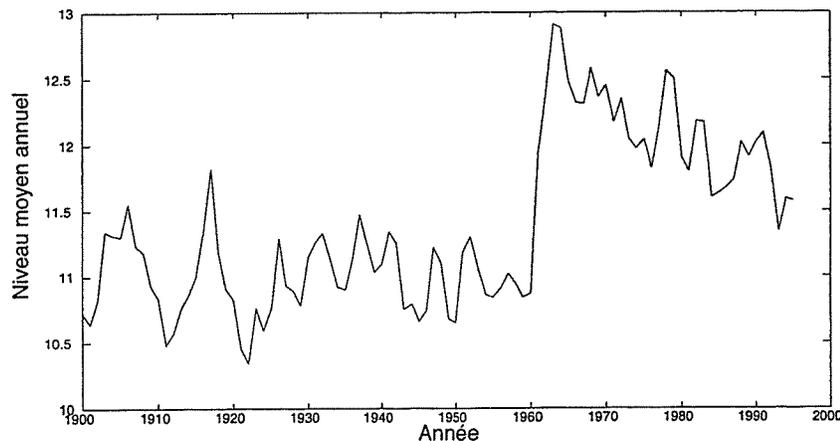


FIG. 2.2 – Niveaux moyens annuels du lac Victoria en mètres

La polémique entourant la hausse brutale du niveau moyen du lac Victoria au début des années 60 est revenue à l'ordre du jour récemment. En effet, cette rupture dans les réserves du lac Victoria, étudiée notamment par Kite (1981) et Piper et al. (1986), est maintenant au centre d'un débat concernant l'ajout d'équipement de production hydroélectrique à

son embouchure à Owens Falls. Alors que de part et d'autre de la date de changement la moyenne des débits du Nil demeure relativement stable, il semble qu'après la rupture de régime la moyenne des niveaux annuels du lac Victoria décroisse linéairement. On aurait donc ici une combinaison des deux types de changements : une rupture à la hausse suivie d'une tendance à la baisse.

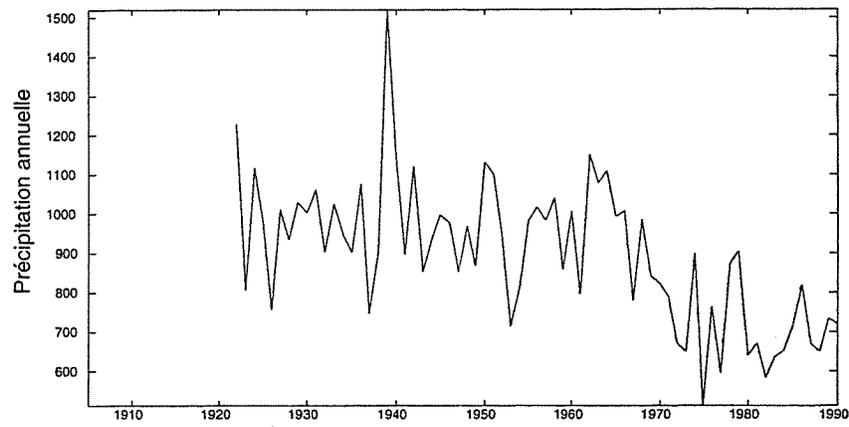
Enfin, l'étude de la stationnarité des séries hydrométéorologiques africaines a permis d'évaluer les caractéristiques de la grande sécheresse qui affecte l'Afrique de l'Ouest depuis la fin des années 60. Les premiers travaux utilisant ces données ont été menés par Hubert et Carbonnel (1987) et Hubert et al. (1989). La Figure 2.3 présente trois séries types de précipitations et de débits pour des sites de la région Sahélienne. On remarque que toutes les séquences montrent une baisse marquée du niveau moyen annuel à la fin des années 60. Par contre, contrairement aux précipitations à Dedougou, les séries de débits, et plus particulièrement celle du Niger, semblent admettre plus d'une rupture. Récemment, Servat et al. (1997) et Paturel et al. (1997) ont étendu cette étude en y incorporant plusieurs séries d'Afrique Centrale.

De nombreux auteurs ont mis en évidence des ruptures dans les séries hydrométéorologiques dans d'autres régions du globe. Yamamoto et al. (1985), par exemple, ont identifié un changement aux alentours de 1950 dans des séries de températures au Japon. On retrouve aussi une rupture à la même date dans les régions arctiques (Kelly et al. (1982)). En Europe, Carbonnel et Hubert (1994) se sont intéressés aux séries pluviométriques roumaines et bulgares pour lesquelles une baisse de moyenne au début des années 80 serait survenue.

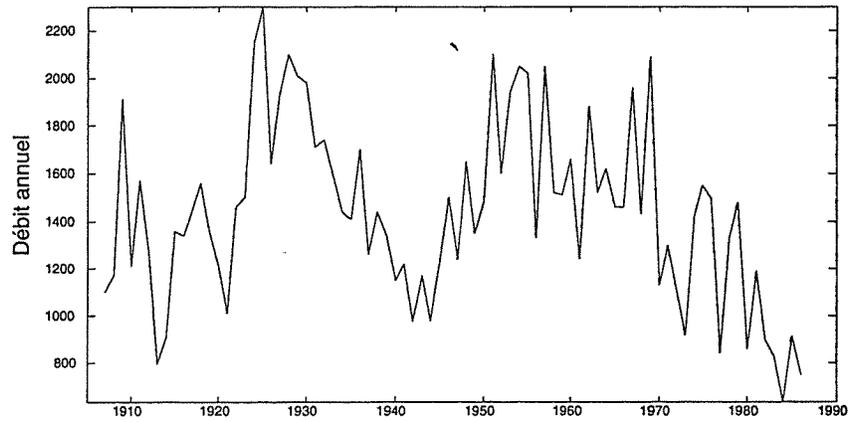
En Amérique du nord, ce sont les séries de précipitations, de débits et de niveaux des Grands Lacs qui ont suscité les études et les débats les plus importants concernant les changements depuis plus d'une décennie (Slivitzky et Mathier (1994), Chao et Hobbs (1997), Venkatesh et Hobbs (1999), Perreault et al. (1999)). La Figure 2.4 montre à titre d'exemple l'évolution des précipitations annuelles des lacs Ontario et Érié depuis le début du siècle. Les précipitations sur ces deux lacs semblent admettre non seulement une rupture de moyenne au début des années 70, mais peut-être aussi une augmentation de la variabilité depuis le milieu des années 20. Remarquons que la rupture de moyenne n'est pas aussi marquée que celle observée dans les séries africaines, à cause de la variabilité plus importante de ces observations.

Récemment, l'intérêt pour les problèmes de non-stationnarité des séries hydrométéorologiques a pris de l'ampleur à la suite de la parution du rapport de la seconde commission intergouvernementale sur les changements climatiques (*Intergovernmental Panel on Climate Change*, IPCC-95, Houghton et al. (1996)). Contrairement à la première réunion scientifique de 1990 (IPCC-90, Houghton et al. (1990)), à l'issue de laquelle la commission avait émis des doutes quant à l'existence de changements climatiques, le compte rendu IPCC-95 conclut que la Terre subit un réchauffement continu depuis la fin du 19^{ème} siècle ainsi qu'une augmentation globale de précipitations d'environ 1%. Partant, de nombreux

(a)



(b)



(c)

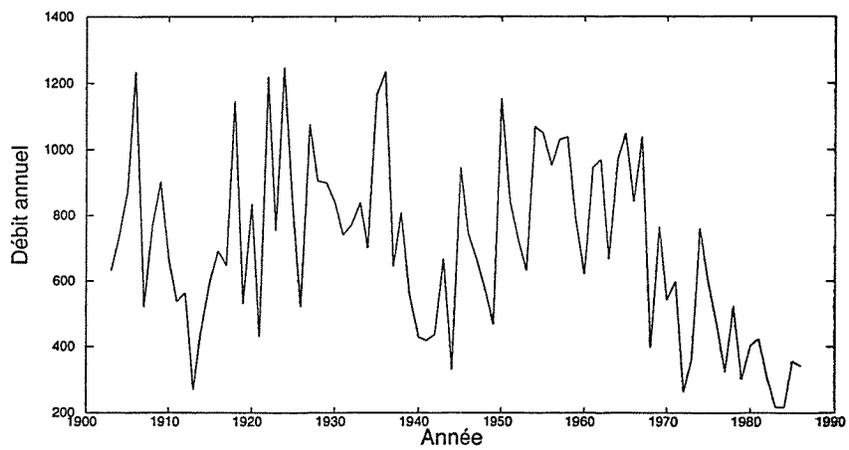


FIG. 2.3 – Précipitations (mm) et débits (m^3/s) annuels pour divers sites sahéliens : (a) station Dedougou, (b) fleuve Niger, (c) fleuve Sénégal

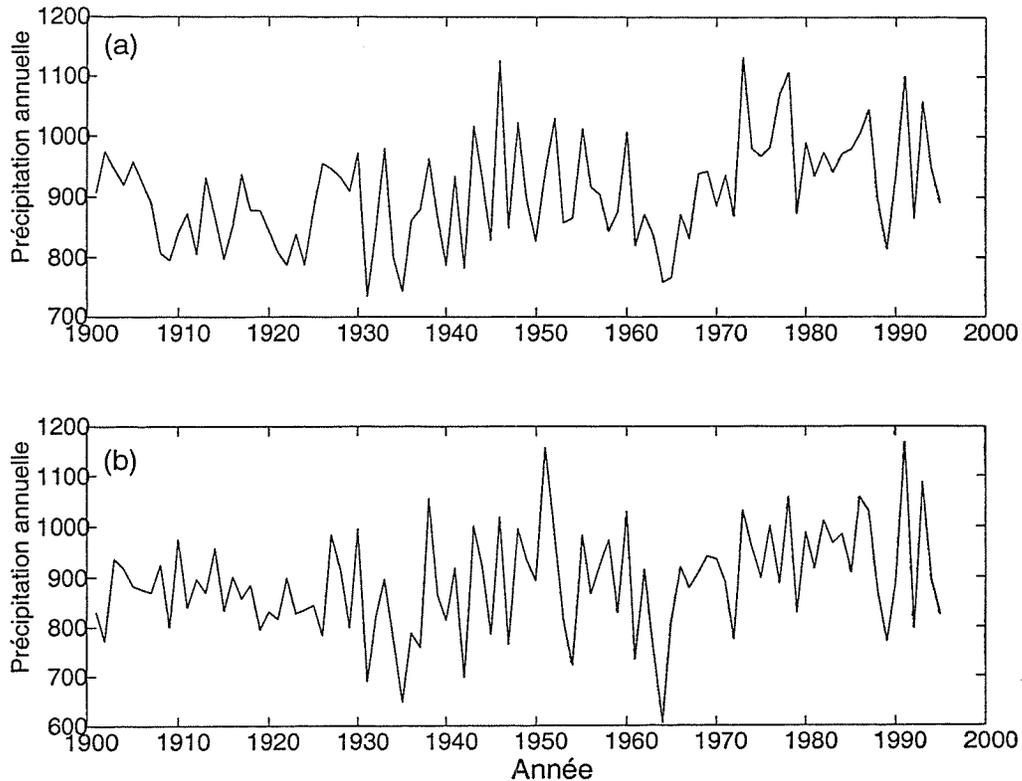


FIG. 2.4 – Précipitations (mm) annuelles pour les lacs (a) Ontario et (b) Erié

hydrologues se sont intéressés à l'analyse des changements dans les séries hydrométéorologiques. Citons en particulier Kiely (1999) qui a analysé un ensemble de séries irlandaises de débits et de précipitations. Il a identifié une tendance à la hausse qui se serait amorcée au milieu des années 70, et voit en ce phénomène les conséquences d'un changement de climat. C'est en reliant ses constatations à l'augmentation du niveau de l'indice d'oscillation Atlantique Nord (NAO), survenu aussi vers la fin des années 70 (Figure 2.5), que Kiely (1999) conclut à un changement climatique. Selon plusieurs climatologues, l'oscillation Atlantique Nord est un des facteurs prépondérant de variation climatique dans les régions comme l'Irlande, situées dans l'hémisphère nord (Kushnir (1999), Greatbatch (2000)). Ce phénomène est dû aux variations des champs de pression atmosphérique provoquées par la présence d'un anticyclone dans la région des Açores et d'un système dépressionnaire près de l'Islande. L'amplitude de l'oscillation Atlantique Nord est évaluée à partir d'indices construits en considérant la différence entre séries normalisées de pressions à deux stations: l'une située en Islande et l'autre dans la zone des hautes pressions (Açores ou Lisbonne). Une valeur très positive des indicateurs NAO correspondrait à l'entrée de masses d'air humide et tempérée sur l'Europe du nord. En Amérique du nord, à l'inverse, une telle valeur serait associée à un climat froid et sec (Hurrell (1995), Hurrell et van Loon (1997)).

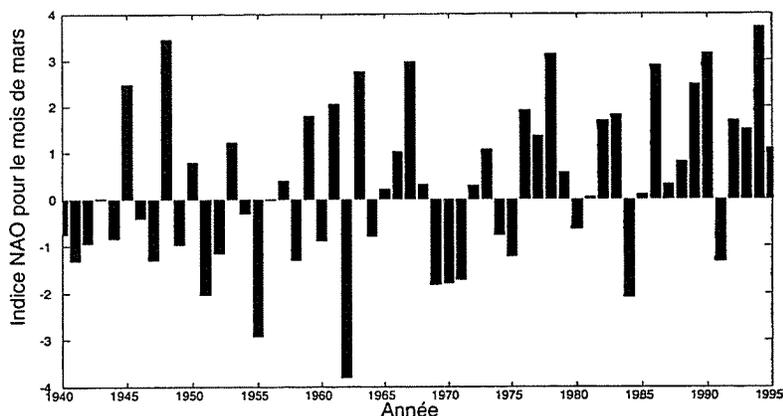


FIG. 2.5 – Valeur de l'indice NAO pour le mois de mars, de 1940 à 1995 (Kiely, 1999)

Certaines séries de débits annuels à des stations canadiennes situées au nord-est du pays présentent une baisse marquée dans leur débit annuel vers 1984, qui pourrait résulter de l'augmentation du NAO (Rasmussen et al. (1999)). A l'examen de la Figure 2.5, on remarque que depuis le milieu des années 70 les valeurs de NAO du mois de mars (l'intensité du phénomène est particulièrement prononcée de janvier à mars) sont presque toutes positives. La Figure 2.6 présente le tracé des apports naturels annuels de 1950 à 1997 observés à Churchill Falls. Ce site se trouve au Labrador dans la province de Terre-Neuve. On remarque effectivement une baisse du niveau moyen des apports annuels vers le début des années 80. Cette baisse d'hydraulicité semble aussi accompagnée d'une diminution de la variabilité.

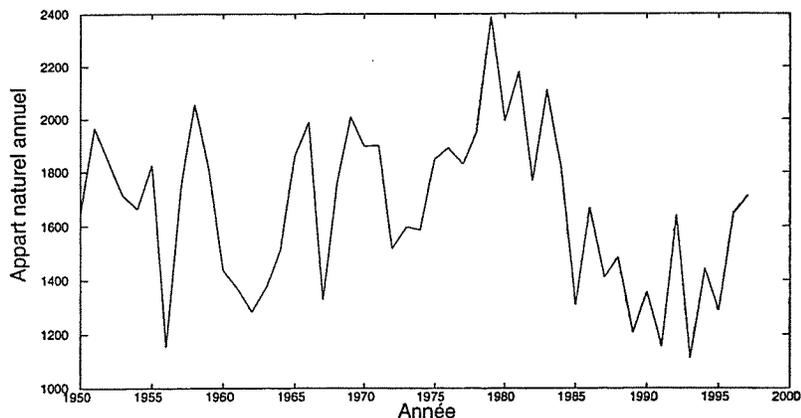


FIG. 2.6 – Apports naturels annuels (m^3/s) observés à Churchill Falls

Le caractère aléatoire des données de débit est lié à la variabilité fondamentale de l'ensemble des facteurs naturels (pluie, température, couvert du sol, etc.). Par conséquent,

ce qui peut nous sembler un changement de régime pourrait très bien n'être que le résultat d'une simple variabilité naturelle. Toutefois, compte tenu des nombreuses études sur les changements climatiques et du fait que les activités humaines peuvent modifier l'écoulement, force est d'accepter l'éventualité que les données récentes ne s'inscrivent pas dans la continuité des précédentes. Si les causes premières d'apparition de tels changements structurels demeurent à l'heure actuelle encore insuffisamment expliquées et sujet à controverses, ces phénomènes, s'ils sont présents, peuvent avoir des conséquences importantes sur la disponibilité de l'eau, et donc sur la gestion de cette ressource. Les différents intervenants dans le processus décisionnel pour l'aménagement hydroélectrique, notamment les gestionnaires et les hydrologues, doivent alors s'adapter à ce nouveau contexte.

2.2 Le gestionnaire doit prendre ses décisions dans un contexte risqué

Dans un contexte simple et sans incertitude, les décisions rationnelles s'imposent d'elles-mêmes, le meilleur outil d'aide à la décision étant alors le bon sens commun. Toutefois, la décision environnementale, et particulièrement en gestion de la ressource en eau, n'est jamais aussi simple (Bernier et al. (2000)). D'abord, parce qu'il n'y a rarement qu'un seul décideur. Le problème à résoudre intègre donc nécessairement les dimensions multiples des divers enjeux et des points de vue contradictoires. Ensuite, et surtout, puisque les ressources hydriques sont soumises à l'effet de divers facteurs naturels (aléas climatiques, etc.) et anthropiques (aménagements du territoire, déforestation, etc.), la décision se fonde souvent en hydrologie, plus que dans d'autres domaines, sur des faits incertains. Enfin, les données hydrométéorologiques dont dépendent les décisions sont rares, elles sont souvent entachées d'erreurs de mesure et sont parfois même reconstituées.

Le gestionnaire en ressource en eau doit prendre de nombreuses décisions dont les conséquences sont conditionnées par l'information disponible au moment d'arrêter son choix. En matière d'aménagement hydroélectrique, de planification énergétique et de construction d'ouvrages de protection contre les crues, les choix reposent principalement sur les séries d'observations de débit.

Exemple 1 *L'aménagement d'une centrale.* Schématiquement, on peut considérer que le dimensionnement d'une centrale hydroélectrique consiste à choisir un type d'aménagement d (la décision) parmi un ensemble d'alternatives. Le gestionnaire doit d'abord supporter un coût total d'investissement, qui ne dépend que de d . Par contre, les revenus anticipés pour un type d'aménagement donné dépendront aussi de caractéristiques du régime hydraulique réel de la rivière, comme la quantité d'eau turbinable au cours de la durée de l'installation. Cette quantité est mal connue car elle est liée à la variabilité du phénomène des pluies et à celle du climat. Souvent, le décideur fixe son choix en fonction du débit annuel moyen calculé sur tout l'historique.

Exemple 2 *La planification énergétique à Hydro-Québec.* Hydro-Québec est une entreprise de service public qui produit presque toute l'électricité consommée au Québec.

De plus, la société d'Etat vend de l'énergie à d'autres compagnies électriques du Canada et du nord-est des États-Unis. Dans un but de planification, Hydro-Québec présente périodiquement un plan de développement de ses activités pour les dix années à venir. 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'adopte l'entreprise (planification de sa production hydroélectrique, planification de ses équipements, stratégies de développement des marchés externes). Le potentiel hydroélectrique du Québec, actuellement aménagé, repose sur les réserves hydrauliques emmagasinées dans les réservoirs de huit grands complexes : Outaouais, Saint-Maurice, Aux Outardes, Manicouagan, Bersimis, Saint-Laurent, La Grande et Churchill-Falls. Pour chacun de ces complexes, les apports énergétiques annuels sont évalués. Ils sont fonction des apports naturels de chacun des réservoirs et de la capacité de production des centrales hydroélectriques. Certaines caractéristiques statistiques issues de ces séries d'apports énergétiques (moyennes, écarts-types, etc.) sont utilisées comme intrants dans des modèles de planification. Ces modèles sont employés, notamment, pour l'optimisation de l'utilisation de l'énergie à long terme, le chargement optimal des centrales et la mise en place de nouveaux équipements.

Exemple 3 *La construction d'une digue. Le gestionnaire doit dimensionner une digue en choisissant sa hauteur. S'il construit cette digue pour une hauteur h , le gestionnaire doit supporter un coût d'investissement. Par ailleurs, celui-ci voudra éviter tout dommage résultant d'un éventuel débordement. Pour chaque crue survenant après la construction et dépassant la hauteur h , il admet alors un dommage fonction du débordement, c'est-à-dire de l'écart entre l'intensité de cette crue et h . Pour construire une digue, il est d'usage d'estimer une crue de conception, c'est-à-dire le débit maximum annuel de faible probabilité au dépassement (par exemple, la crue centennale : celle qui survient en moyenne 1 fois sur 100), et de choisir la hauteur en fonction de cette valeur. Les données de références sont alors les débits maximums annuels historiques.*

Comme dans les exemples 1 et 3, le problème décisionnel est souvent ramené à un problème d'estimation statistique qui fait intervenir un ou plusieurs échantillons de débits historiques. Ces données sont soumises non seulement à l'aléa naturel (fluctuation du débit annuel due à la variabilité des pluies), mais aussi à l'aléa structurel, c'est-à-dire à une modification de leur structure aléatoire (baisse du niveau moyen des débits à la suite d'un changement climatique). Divers exemples ont été présentés à la Section 2.1. Au Québec, notamment, une partie du territoire a connu plusieurs années de précipitations déficitaires, entraînant ainsi une baisse du niveau des débits depuis 1983 (Figure 2.6). Cette baisse d'hydraulicité est-elle structurelle ou conjoncturelle? Quel est l'effet des incertitudes portant sur cette baisse d'hydraulicité sur l'aménagement optimal d'une centrale hydroélectrique? Serait-il plus rationnel d'évaluer l'apport énergétique moyen seulement à partir des observations de 1983 à nos jours, c'est-à-dire en considérant une information plus restreinte mais peut-être plus représentative de l'état actuel du système? Si dans certaines régions du globe la variabilité des débits semble avoir augmenté, est-ce

indicatif d'une plus forte occurrence des phénomènes extrêmes? Quel en est l'impact sur le choix optimal d'une hauteur de digue?

Ces trois exemples, quoique schématiques, illustrent bien toutes les difficultés du choix rationnel de décision, notamment la nécessaire prise en compte des effets des incertitudes structurelles sur ce choix. Aussi, le décideur en sciences de l'eau est-il forcément amené à prendre ses décisions en contexte risqué, c'est-à-dire à faire un pari. Or, très souvent pour un gestionnaire, la notion de risque est principalement associée à la seule intensité des conséquences dommageables éventuelles des différentes décisions qu'il doit prendre, sans référence à la notion de probabilité d'occurrence d'événements préjudiciables. Il pré-suppose généralement qu'il existe une solution hydrologique parfaitement déterministe et s'en remet à l'ingénieur-hydrologue, détenteur des informations et des connaissances techniques.

2.3 L'hydrologue est démuni face à l'incertitude de changement

L'ingénieur-hydrologue intervient directement en tant que conseil en matière de risque auprès du décideur. Il est amené à formuler certaines hypothèses pour représenter les différents types d'aléas, à les communiquer et à réaliser ou à faire exécuter ensuite les calculs d'estimation sur la base desquels le décideur règle son action. L'ingénieur-hydrologue n'a souvent qu'une culture limitée de la formalisation probabiliste nécessaire à la représentation des incertitudes. C'est pourquoi il se réfère généralement aux outils statistiques traditionnels disponibles dans les logiciels standards. Or, l'hypothèse de stationnarité, c'est-à-dire que les observations sont identiquement distribuées, est pratiquement omniprésente dans ces méthodes usuelles. La plupart des hydrologues sont démunis face à l'incertitude structurelle dans les séries de débit. Ne disposant pas d'outil adéquat pour analyser la non-stationnarité, ils se doivent actuellement d'accepter indûment l'hypothèse de stationnarité et produire des modèles éventuellement biaisés ou de rejeter celle-ci et prôner l'empirisme.

L'hypothèse de stationnarité est souvent acceptée par facilité

Même si l'hypothèse de stabilité des caractéristiques statistiques des séries hydrométéorologiques ne semble pas toujours réaliste aux yeux d'un nombre grandissant d'hydrologues, ce postulat reste pourtant omniprésent lors de l'analyse statistique des séries de débit. Par exemple, le choix du débit d'équipement et des dimensions de la galerie d'amenée de la centrale Toulmoustou au Québec est fait en fonction du débit moyen annuel (186 m³/s) calculé sur l'ensemble de l'historique des apports naturels au réservoir Sainte-Anne (Hydro-Québec (2000)). Aucune étude de sensibilité des choix optimaux vis-à-vis de l'incertitude structurelle sur cette estimation du débit n'est rapportée dans le document. Or, ce site du nord-est du Québec pourrait admettre une rupture en moyenne au milieu des années 80 (Figure 2.7). Même de faible amplitude, ce changement semble être réel. En

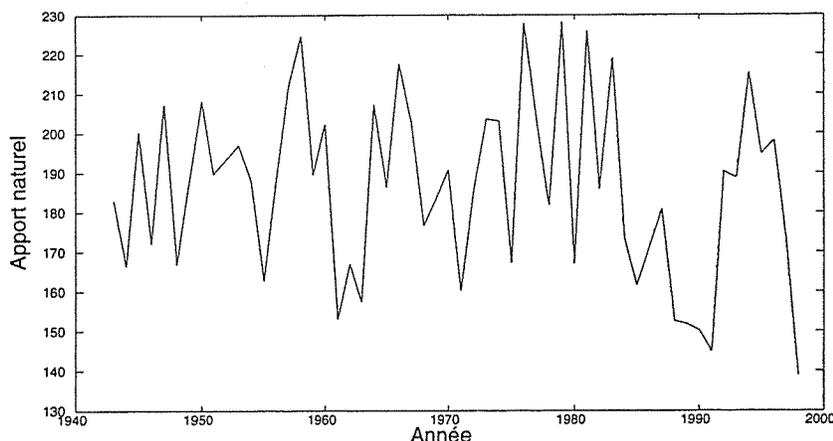


FIG. 2.7 – *Apports naturels annuels (m^3/s) observés au réservoir Sainte-Anne*

effet, de nombreux sites dans la région présentent, à peu près la même année, une baisse du niveau moyen de magnitude non négligeable.

Parce qu'ils sont accessibles et plus simples, des modèles statistiques standards, qui supposent que les observations récentes s'inscrivent dans la continuité des précédentes, sont ajustés à des séries temporelles de débits. Les modèles ainsi estimés sont ensuite employés pour simuler l'exploitation de complexes hydroélectriques (planification énergétique, conception de nouveaux ouvrages, gestion des réservoirs). Or, ne pas considérer la possibilité qu'une rupture ou une tendance soit survenue dans l'historique des débits peut être lourd de conséquences. Si en réalité les observations admettent un changement dans leur structure aléatoire, les estimations reposant sur l'hypothèse de stationnarité engendreront des conclusions erronées sur le phénomène étudié. Elles peuvent, par exemple, conduire au sous-dimensionnement des nouveaux ouvrages hydrauliques et entraîner ainsi des pertes financières importantes, voire des pertes de vies humaines.

L'étude des changements fait souvent l'objet d'une formalisation incorrecte

Nombre d'études sur la non-stationnarité des chroniques de débits emploient des techniques statistiques traditionnelles sans qu'on se soucie des hypothèses sur lesquelles elles reposent. L'utilisation du test de Student ou de Mann-Withney (Miquel (1984), Bobée et Ashkar (1991)), par exemple, suppose la connaissance a priori de la date de rupture, puisqu'il est nécessaire de diviser la série en deux sous-échantillons indépendants. Bien que ces méthodes soient adéquates pour certains types de changement se produisant à une date connue, par exemple une rupture induite par une modification d'une méthode de mesure ou le déplacement d'une jauge, elles ne le sont pas pour étudier un changement engendré par des facteurs naturels comme un changement climatique. En effet, dans ce cas, il est impossible d'identifier avec certitude la date de rupture. C'est d'ailleurs l'un des problèmes auxquels les climatologues s'intéressent : ils cherchent à identifier le moment où

le climat aurait changé et à évaluer l'incertitude qui y est associée (Barnett et al. (1999)). Si l'instant de changement n'est pas connu, pourquoi ne pas formaliser mathématiquement cette particularité du problème et construire un modèle qui inclut cette incertitude?

Une autre méthode utilisée fréquemment en hydrologie est celle de Hubert et al. (1989). Ces auteurs ont proposé une procédure originale de segmentation, couplée au test de Scheffé, pour détecter toutes les ruptures significatives dans une série. Cette méthode a été employée notamment pour les études concernant les séries africaines (Servat et al. (1997) et Paturel et al. (1997)) et celles des Grands Lacs (Slivitzky et Mathier (1994)). L'approche d'Hubert et al. (1989), qui s'apparente à une méthode de classification, repose implicitement sur un modèle où le nombre de ruptures est un paramètre inconnu. Or, ces auteurs ne posent pas le modèle permettant de représenter cette caractéristique implicite, et donc d'en quantifier l'incertitude. Sans insister sur les problèmes techniques liés à la conservation du seuil réel global de la procédure (erreur de Type I), que vaut alors la conclusion selon laquelle cette approche détecte k ruptures si, faute de formalisation mathématique, on ne peut associer à cette estimation aucune incertitude? Il faudrait au moins intégrer explicitement dans un modèle, comme paramètre inconnu, le nombre k de ruptures. La procédure de Hubert et al. (1989), pourtant fondée sur une idée novatrice, ne peut être utilisée qu'à des fins descriptives et exploratoires.

Enfin, de plus en plus, les gestionnaires responsables acceptent de considérer des scénarios réalistes, c'est-à-dire des situations types, pour lesquels il existe un certain nombre d'actions envisageables. L'estimation du risque lié à chacune d'entre elles permet alors leur classement, véritable outil d'aide à la décision. Pour simuler le phénomène, ces scénarios doivent donc reposer sur un modèle représentatif de la réalité. Or, trop souvent, l'abondance des scénarios masque la pauvreté et le manque de réflexion sur la formalisation du modèle. Comment quantifier alors le risque?

La méthode de Lee et Heghinian force l'hydrologue à considérer avec certitude l'existence d'un changement de moyenne

Même si la majorité des procédures statistiques employées par les hydrologues sont des tests paramétriques et non paramétriques qui nécessitent la connaissance a priori de la date de rupture, il existe néanmoins, dans la littérature hydrologique, quelques études récentes s'appuyant sur des modèles dont l'un des paramètres inconnus est l'instant de changement. Par exemple, de nombreux hydrologues ont employé la procédure bayésienne de Lee et Heghinian (1977). Cette approche repose sur un modèle à une seule rupture de moyenne survenant à un instant inconnu dans une suite de variables aléatoires indépendantes et distribuées selon une loi normale. Telle que proposée, la méthode de Lee et Heghinian permet d'estimer la date de rupture ainsi que son amplitude, conditionnellement au fait qu'un seul changement de moyenne soit intervenu dans la série. En effet, le modèle suppose avec certitude l'existence de ce type de rupture. Aussi, dans ce cas, l'hydrologue doit-il exclure l'hypothèse de stationnarité et de tout autre type de changement. Est-il prêt à rejeter la possibilité qu'en réalité la série soit demeurée stationnaire? Qu'en est-il des autres types de changement, notamment des ruptures en variance? Étant donné

l'incertitude sur l'existence et le type de changement, ne faut-il pas intégrer dans l'analyse diverses alternatives?

Notons que plusieurs auteurs ont employé indûment une approche conditionnelle à une hypothèse de rupture comme celle de Lee et Heghinian (1977) pour tirer des conclusions sur l'existence du changement (Bruneau et Rassam (1983), Hubert et Carbonnel (1987), Rao et Tirtotjondro (1996), Servat et al. (1997), Paturel et al. (1997), Lubès-Niel et al. (1998), Ouarda et al. (1999)). Certains justifient l'existence d'une rupture en moyenne en s'appuyant sur le fait que la distribution a posteriori de la date de changement présente un mode bien isolé et de probabilité relativement forte. D'autres construisent un intervalle de crédibilité pour l'amplitude de changement et concluent à la présence d'une rupture si l'intervalle ne contient pas la valeur zéro. Ils oublient que les calculs ont été faits en considérant l'hypothèse d'un changement certain et qu'alors l'événement « l'amplitude de rupture est nulle » n'a aucune signification claire.

L'application systématique d'une série d'analyses locales réalisées indépendamment est inadéquate pour conclure à un changement régional dans une zone homogène

La non-stationnarité de séries hydrométéorologiques peut être induite par un changement climatique. Or, un changement climatique est un phénomène global, se manifestant à une échelle au moins régionale. Il est donc naturel de considérer le caractère multidimensionnel du problème. Cette généralisation est d'autant plus justifiée, qu'en pratique, l'hydrologue suppose implicitement qu'il existe une relation structurelle entre les instants de rupture si les sites considérés forment un ensemble « hydrologiquement homogène ». En toute généralité, des bassins versants sont dits hydrologiquement homogènes si leurs régimes hydrologiques sont semblables et s'ils répondent aux facteurs climatiques de manière similaire (GREHYS (1996)). Certes, en première approche, on peut effectuer plusieurs analyses unidimensionnelles pour estimer une date globale de changement ou pour confirmer, à partir de sites voisins, l'existence d'un changement dans la structure aléatoire d'une série d'intérêt (Hubert et al. (1989), Slivitzky et Mathier (1994), Servat et al. (1997), Paturel et al. (1997), Perreault et al. (1999), Kiely (1999)). Toutefois, cette façon de procéder est répétitive et sûrement peu efficace. D'abord, cette procédure ne prend pas en compte les corrélations régionales entre les observations des différentes stations. En appliquant plusieurs analyses indépendantes, on considère en réalité que chaque station amène une information tout à fait distincte et qu'ainsi la précision des estimations correspond, en quelque sorte, à celle issue d'un échantillon de taille $p \times n$ (p étant le nombre de stations, n la taille de chaque série). Si, en fait, les observations sont corrélées d'une station à l'autre, il existe de l'information redondante et la précision réelle des estimations devrait correspondre plutôt à celle d'un seul échantillon de taille inférieure à $p \times n$. L'application systématique d'une série d'analyses locales réalisées indépendamment aura donc pour effet de sous-estimer l'incertitude des estimations. S'il existe une relation structurelle entre les instants de changement et que les observations de débits à des stations soumises aux mêmes phénomènes climatiques sont corrélées, pourquoi ne pas mobiliser

cette information dans un modèle qui prend en compte la cohérence régionale?

2.4 Le statisticien peut construire des solutions en cohérence avec la demande du gestionnaire

Compte tenu de la possibilité de changements dans les séries de débits, de la demande du gestionnaire en matière d'aménagement de structures hydrauliques et des problèmes que rencontrent les hydrologues pour la quantification des incertitudes structurelles, ma question de recherche est la suivante :

Comment peut-on prendre en compte l'incertitude de changements dans la structure aléatoire des séries de débits annuels lors de la prise de décision concernant un aménagement hydroélectrique?

Le statisticien répondra qu'il faut d'abord formaliser mathématiquement les hypothèses de l'hydrologue sous la forme d'un modèle probabiliste et développer ensuite les outils permettant de quantifier les incertitudes des paramètres et celles liées au modèle. Compte tenu du lien étroit entre les problèmes d'estimation et de choix opérationnel des décisions, j'ajouterai que les incidences des incertitudes d'estimation et de la véracité des hypothèses sur les conséquences socio-économiques des décisions doivent être envisagées et prises en compte dans l'analyse.

En matière de gestion des eaux ou d'aménagement de structures hydrauliques, il est d'usage de distinguer les problèmes d'estimation de caractéristiques hydrologiques comme le débit moyen annuel d'une rivière, et le choix opérationnel des décisions d'aménagement comme la fixation de valeurs de projet (Bernier (1990)). Or, le choix des hypothèses de travail et la précision des estimations sont déterminants pour le processus décisionnel. Reprenons le problème du choix optimal d'une hauteur de digue (Exemple 3, Section 2.2). Le critère qui consisterait, par exemple, à retenir un estimateur médian de la crue centennale, c'est-à-dire qui a autant de chance de surestimer que de sous-estimer la vraie valeur, serait fallacieux. Il ne prendrait résolument pas en compte la dissymétrie du problème : le fait qu'un surdimensionnement est moins pénalisant pour le riverain qu'un sous-dimensionnement. En effet, pour la construction d'une structure de protection contre les crues, il est clair que les conséquences dommageables (au sens large), dues à un débordement, sont bien plus importantes que les pertes encourues par un surinvestissement qui sont seulement économiques.

Dans ce qui suit, je présente d'abord ma stratégie pour répondre à la question de recherche. Ensuite, je précise les pistes de recherche que j'ai choisies d'explorer.

2.4.1 Ma stratégie : adopter la perspective bayésienne

J'ai choisi d'adopter la perspective bayésienne pour répondre de façon adéquate à la question de recherche. Ce qui distingue principalement l'approche bayésienne d'une méthode statistique classique est que, en plus de considérer l'information fournie par l'échan-

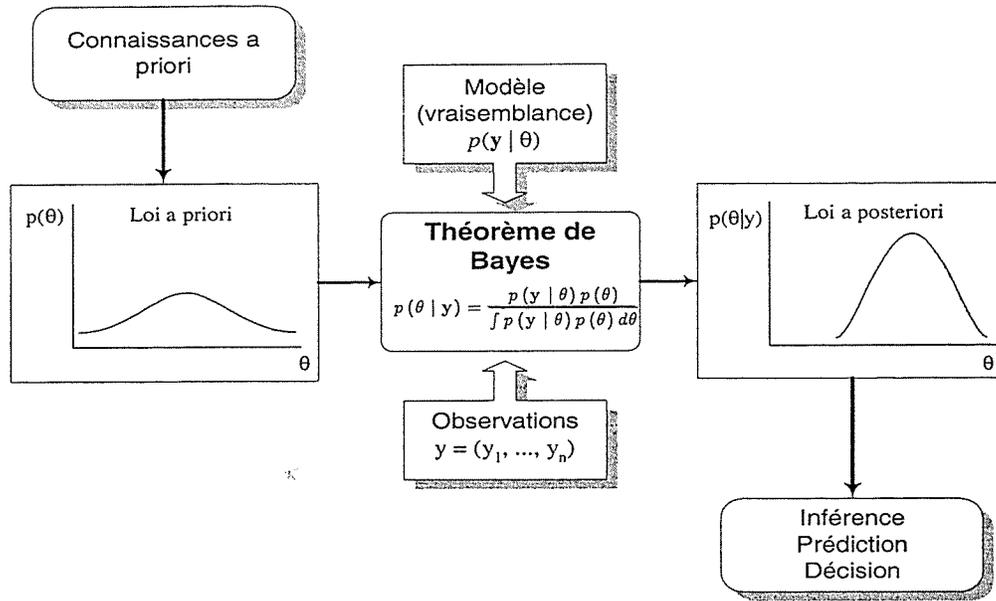


FIG. 2.8 – Schématisation des étapes fondamentales de l'analyse bayésienne

tillon de données, elle incorpore de manière formelle dans l'analyse la connaissance a priori dont on dispose concernant les paramètres d'intérêt. Cette information n'est pas issue des observations formant la série, mais provient d'autres sources (analyses statistiques antérieures, connaissances subjectives des experts ou du décideur, etc.). Les connaissances a priori sur les paramètres, souvent imprécises, sont traduites en terme de probabilités d'occurrence pour les différentes valeurs possibles et sont représentées au travers de la distribution de probabilité subjective a priori. Les observations recueillies lors d'une expérience statistique et la loi a priori forment conjointement en statistique bayésienne la base de tout jugement sur les paramètres.

Une fois les observations disponibles et les lois a priori spécifiées, la formule de Bayes est utilisée pour déterminer les lois a posteriori et assurer la cohérence du cadre inférentiel. Ces distributions représentent l'actualisation de la connaissance a priori eu égard à l'information contenue dans les observations et valorisée par le modèle. C'est à partir de ces lois a posteriori que toutes décisions et inférences concernant les paramètres du modèle sont faites. Le diagramme de la Figure 2.8 présente les étapes fondamentales du raisonnement bayésien qui conduisent à l'inférence a posteriori. Pour plus de détails sur l'analyse bayésienne, nous référons à Berger (1985) et Robert (1992).

Dans la littérature hydrologique, quelques auteurs ont adopté la perspective bayésienne durant les années 70 et au début des années 80 (voir, par exemple, Davis et al. (1972), Vincens et al. (1975) et Krzysztofowicz (1983)). L'application des concepts bayésiens aux sciences de l'eau a été très peu développée depuis, en raison fort probablement de la complexité de certains calculs et de l'absence de solutions explicites. Les avancées

récentes en matière de calcul numérique pour la mise en oeuvre des méthodes bayésiennes (méthodes de simulation Monte Carlo par chaînes de Markov, MCMC) nous permettent maintenant de contourner cet obstacle. Cela s'est traduit par une recrudescence marquée de l'emploi de l'approche bayésienne, et ce, dans de nombreux domaines d'application de la statistique (Berger (1999), voir aussi la préface de C.P. Robert dans Bernier et al. (2000)). Étonnement, l'utilisation des méthodes bayésiennes demeure encore aujourd'hui marginale en sciences de l'eau contrairement à d'autres domaines comme la médecine, la finance, l'économie et la biologie, pour ne donner que ces exemples. Les hydrologues n'ont pas su tirer partie de ces nouveaux outils et l'hydrologie-statistique accuse maintenant un sérieux retard. Pourtant, conceptuellement, la perspective bayésienne me semble bien plus proche des préoccupations du gestionnaire en sciences de l'eau que les approches dites « classiques ». En particulier, l'interprétation de la probabilité, non pas comme une fréquence limite d'une infinité de réalisations hypothétiques, mais comme un pari rationnel sur un événement dépendant d'un savoir et des informations disponibles, ouvre des perspectives immédiates pour la gestion du risque, au coeur du métier de l'ingénieur.

L'emploi de la statistique bayésienne peut être justifié formellement par le fait qu'elle s'appuie sur une base axiomatique qui permet une cohérence grâce au seul emploi du calcul des probabilités. De tels arguments sont abondamment développés dans les ouvrages de Berger (1985), Robert (1992) ou Bernardo et Smith (1994). Je me contenterai ici de souligner quelques avantages pratiques et techniques pour les problèmes hydrologiques.

- L'analyse bayésienne permet d'obtenir une loi de probabilité jointe (la distribution a posteriori) pour l'ensemble des paramètres inconnus, à partir de laquelle plusieurs conclusions peuvent être tirées simultanément. Il est possible non seulement d'obtenir des estimations pour les paramètres ou des fonctions de ces quantités, mais aussi d'évaluer directement leur précision. Notons de plus que l'inférence bayésienne ne fait référence à aucune hypothèse asymptotique.
- L'interprétation des résultats d'une analyse bayésienne est simple et directe. Les méthodes classiques sont beaucoup plus difficiles à interpréter pour un non-initié, et les gestionnaires de ressources hydriques les trouvent généralement trop abstraites pour leurs besoins opérationnels. C'est le cas notamment de l'interprétation de l'intervalle de confiance classique, qui fait intervenir la répétition d'un nombre infini d'expériences équivalentes. Nombre de praticiens interprètent d'ailleurs l'intervalle de confiance classique comme si c'était un intervalle de crédibilité bayésien (Lecoutre (1997)).
- Les hydrologues disposent souvent, avant l'étude, d'informations non quantifiées sur le phénomène étudié. Il n'y a aucune raison de ne pas tenir compte de cette expertise. L'approche bayésienne nous permet de la considérer de manière formelle à travers les lois a priori. L'ajout de ces informations à celle fournie par les données est d'autant plus utile qu'en hydrologie les échantillons sont souvent de taille restreinte. De plus, l'idée de mettre à jour la connaissance a priori des spécialistes à partir de l'information contenue dans les données semble très intéressante et rassurante pour bien des gestionnaires.

- Les décideurs sont sensibles aux aspects socio-économiques de la gestion des ressources en eau. Ils sont d’avis que les analyses statistiques et l’évaluation des conséquences socio-économiques de leurs décisions ne doivent pas être effectuées séparément (Bernier (1994a)). Or, l’analyse bayésienne permet pour le gestionnaire d’intégrer les modèles de rupture dans une démarche décisionnelle complète. Il suffit de considérer dans la procédure une fonction de coût qui quantifie la perte encourue pour les différentes conséquences. Il s’agit alors d’une analyse bayésienne dite « complète », l’objectif étant le choix de l’action optimale parmi un ensemble d’actions.
- Comme cela a été spécifié à la Section 2.3, certains gestionnaires acceptent de considérer des scénarios avant d’arrêter leur choix. De plus en plus d’hydrologues sont alors amenés à étudier les conséquences des diverses actions possibles sur une gamme de scénarios simulés. L’analyse bayésienne est particulièrement adaptée pour générer de telles séquences par l’entremise de la distribution prédictive de données futures conditionnelles aux données déjà enregistrées (Krzysztofowicz (1999)).
- Ce qui a joué, plus particulièrement, en défaveur de l’approche bayésienne est l’absence, dans certains cas, de solutions analytiques explicites pour les distributions a posteriori et la complexité des calculs numériques qui en découle. Or, les méthodes MCMC nous permettent aujourd’hui de résoudre ce problème pour des modèles de structure complexe et comportant un grand nombre de paramètres (Robert (1998), Robert et Casella (1999)).

2.4.2 Les choix qui définissent le cadre d’étude de ma thèse

J’ai choisi d’inscrire mon travail dans le cadre particulier d’une non-stationnarité induite par une seule rupture qui survient à un instant inconnu dans la structure aléatoire de séries d’observations de débit annuel. Le choix de l’hypothèse d’une seule rupture mérite quelques explications. Elle est avant tout justifiée d’un point de vue hydrométéorologique. L’objectif n’est pas celui d’une critique des données détectant les divers avatars des mesures ni celui d’analyser de manière séquentielle leurs caractéristiques statistiques comme en contrôle de la qualité, mais bien de détecter des changements significatifs de régime à une échelle climatique. Or, on ne peut pas raisonnablement admettre plus d’un changement climatique sur la très courte période où l’on dispose de mesures hydrologiques. Par ailleurs, d’un point de vue statistique, cette hypothèse est également justifiée. En effet, si on ne limite pas le nombre de ruptures, toute méthode de détection perd nécessairement de son efficacité.

Dans cette thèse, j’explore les quatre pistes de recherche suivantes :

L’analyse d’une rupture dans la variabilité des débits annuels L’estimation d’une mesure de dispersion des débits annuels joue un rôle fondamental pour la gestion des aménagements hydrauliques, car elle est essentielle pour évaluer la marge de manoeuvre. Dans la littérature hydrologique, seul le cas d’une rupture de moyenne a été considéré. Or, il semble, à l’examen de nombreuses séries, qu’une baisse (hausse) subite de la moyenne des

débites annuels puisse être accompagnée d'une diminution (augmentation) de la variance (voir par exemple la Figure 2.6). On peut aussi, dans quelques cas, identifier des situations où seule la variance change. Les précipitations au lac Érié (Figure 2.4) en sont un exemple.

Le développement d'une procédure permettant de tester l'existence d'une rupture et, le cas échéant, d'en identifier le type En hydrologie, les travaux concernant l'analyse bayésienne d'une rupture se limitent uniquement à l'estimation des paramètres conditionnellement à l'existence d'un changement. Cette approche ne permet pas d'évaluer la crédibilité d'une rupture et force l'hydrologue à ne considérer qu'une seule alternative, très souvent le changement de moyenne. Compte tenu de la variabilité naturelle des débits, des nombreux facteurs exogènes agissant sur leur évolution et de la taille généralement restreinte des séries, l'incertitude quant à l'existence et au type de rupture sera toujours présente. Il est donc important de mettre en doute les hypothèses des modèles et de proposer une méthode permettant d'évaluer la crédibilité d'un ensemble d'alternatives.

La construction d'un modèle régional multisite de rupture en moyenne et le développement des outils d'inférence Dans un récent rapport sur l'état d'avancement des travaux scientifiques concernant la détection de changements climatiques, Barnett et al. (1999) considèrent que l'une des avenues de recherche les plus urgentes est l'analyse bayésienne multidimensionnelle : « On the basis of our analysis, we identify the following (research areas) as most urgent. [...] Multivariate detection studies are required that include all available climate change indices [...] a Bayesian approach may be needed to combine all information into a single comprehensive detection and attribution analysis. » À notre connaissance, aucune approche formelle d'analyse multisite de rupture à un instant inconnu n'a été proposée jusqu'à maintenant dans la littérature hydrologique. Tel que mentionné à la Section 2.3, plusieurs études de détection de changement en hydrologie reposent sur l'hypothèse implicite qu'il existe une relation structurelle entre les dates de ruptures pour des sites voisins, mais sans la formuler explicitement dans un modèle probabiliste. De plus, il est d'usage d'employer des méthodes régionales pour améliorer les estimations de débits de conception (GREHYS (1996)). Ces approches s'appuient sur le fait que les débits annuels ou extrêmes ont un comportement similaire si les sites sont hydrologiquement semblables et s'ils sont soumis aux mêmes facteurs climatiques. Il est donc tout indiqué de construire un modèle multisite de rupture.

La mise en oeuvre de l'analyse bayésienne complète sous incertitude de rupture en moyenne L'inférence sur les paramètres d'un modèle, quoiqu'essentielle, ne peut constituer une fin en elle-même. Les procédures développées doivent être intégrées dans une démarche décisionnelle pour qu'elles soient complètement opérationnelles. Pour illustrer la mise en oeuvre d'une analyse bayésienne complète, j'ai choisi de traiter le problème de la sélection du débit d'équipement de la centrale Toulnostouc. Sa construction, à l'embouchure du réservoir Sainte-Anne (complexe hydroélectrique Manicouagan), devrait être entreprise en 2002. Au terme d'un exercice mené en 1999 par Hydro-Québec, qui repose sur la méthode du coût de revient minimal, Hydro-Québec (2000) recommande un

débit de conception de $225 \text{ m}^3/\text{s}$, et ce, en considérant comme débit de référence (appelé aussi débit module) le débit moyen annuel évalué sur tout l'historique ($186 \text{ m}^3/\text{s}$). On suppose alors implicitement, comme cela est l'usage courant, que la série est stationnaire.

Chapitre 3

Mettre en oeuvre l'analyse bayésienne dans un modèle de rupture en variance

La variabilité des données hydrologiques joue un rôle fondamental pour la gestion des aménagements hydrauliques, notamment pour le dimensionnement d'un réservoir. Cette caractéristique est aussi importante pour la planification énergétique, car la variance des séries annuelles d'apports énergétiques historiques est employée comme intrant dans des modèles de simulations de scénarios (Perreault et al. (1996)). À notre connaissance, il n'existe pas de travaux dans la littérature hydrologique traitant d'un changement à une date inconnue dans la variabilité des séries hydrométéorologiques, les hydrologues ayant concentrés leurs efforts uniquement sur le problème de rupture en moyenne. Or, comme il a été mentionné au chapitre précédent, certaines séries hydrologiques semblent admettre des changements de variance. Nous avons constaté, par exemple, qu'une baisse (hausse) subite du niveau moyen des observations est souvent accompagnée d'une diminution (augmentation) de la variance. Ce chapitre a pour objet de fournir à l'hydrologue une procédure permettant l'analyse d'une seule rupture en variance dans une série de variables aléatoires hydrologiques.

Dans la première section, nous présentons une revue de synthèse des principaux acquis concernant l'analyse bayésienne dans un modèle de rupture. Nous proposons une première formalisation probabiliste pour la représentation d'un changement brusque à un instant inconnu. Les hypothèses sous-jacentes y sont décrites et discutées à l'égard de leur validité hydrologique. En nous inspirant des travaux de Menzefricke (1981), nous montrons ensuite comment l'approche bayésienne peut être mise en oeuvre pour étudier le cas particulier d'une rupture en variance. Deux modèles sont considérés : le changement simultané de moyenne et de variance et le changement de variance seulement. À titre d'illustration, nous présentons une application aux séries d'apports énergétiques annuels des complexes hydroélectriques de Churchill Falls et de l'Outaouais. Ces travaux concernant la rupture en variance ont fait l'objet de publications scientifiques (articles I et II) présentés aux annexes A et B. Les développements techniques, dont les calculs ne font intervenir que

des manipulations algébriques standards propres à la loi normale, sont donc énoncés sous forme de propositions sans démonstration complète. De plus, seuls les principaux résultats de l'application aux données réelles sont présentés. Nous laissons le lecteur consulter les articles en annexe pour les détails concernant notamment la démarche adoptée pour spécifier les lois a priori.

3.1 Modèle à une seule rupture

Les problèmes de rupture à un instant inconnu ont suscité l'intérêt de nombreux statisticiens depuis les 40 dernières années. Habituellement, ce type de situation survient lorsqu'on soupçonne que le processus générateur d'observations recueillies chronologiquement a changé durant la période d'échantillonnage. L'objectif est alors de détecter la rupture de manière rétrospective (par opposition à séquentielle) et d'estimer, le cas échéant, l'instant de changement et son amplitude. Depuis les travaux de Page (1955) en contrôle statistique de la qualité, l'intérêt soutenu pour les problèmes de rupture est dû aux nombreuses applications possibles dans des domaines aussi différents que la théorie du signal, la médecine, l'environnement, l'économie et la finance. Diverses solutions ont été proposées, dont plusieurs s'appuient sur le paradigme bayésien.

La démarche bayésienne consiste pour l'essentiel à employer une variable aléatoire afin de décrire l'incertitude relative aux paramètres d'intérêt et de la mettre à jour au regard des données observées. Dans la perspective bayésienne, sont probabilisées non seulement les quantités observables (les données) mais aussi les quantités non observables (les paramètres).

3.1.1 Représentation des quantités observables

Un modèle paramétrique de rupture suppose généralement que la distribution d'un phénomène temporel (ou causal) voit ses paramètres changer subitement à un ou à plusieurs instants inconnus. Pour le cas d'une seule chronique d'observations de débits annuels, nous considérons le modèle à une seule rupture que l'on définit comme suit :

Définition 3 Soit $p(y | \theta)$ une densité de probabilité indexée par le vecteur-paramètres $\theta \in \Theta \subseteq \mathbb{R}^k$. Alors, une rupture survient à l'instant τ dans une séquence de n variables aléatoires indépendantes $Y_1, \dots, Y_\tau, Y_{\tau+1}, \dots, Y_n$ si

$$\begin{aligned} Y_j &\sim p(y_j | \theta_1), & j = 1, \dots, \tau, \\ Y_j &\sim p(y_j | \theta_2), & j = \tau + 1, \dots, n, \end{aligned} \tag{3.1}$$

où $\tau \in \{1, 2, \dots, n - 1\}$ et $\theta_1 \neq \theta_2$.

Les n variables aléatoires sont supposées indépendantes et distribuées selon une loi de probabilité dont la valeur des paramètres θ change à l'instant inconnu τ . Puisque l'instant de rupture τ prend ses valeurs dans l'ensemble $\{1, 2, \dots, n - 1\}$, le changement survient avec certitude dans la période d'échantillonnage. Le modèle (3.1) possède une structure

simple: connaissant τ , les réalisations sont en fait divisées en deux sous-échantillons indépendants, (y_1, \dots, y_τ) et $(y_{\tau+1}, \dots, y_n)$. La fonction de vraisemblance résultante s'écrit alors de la manière suivante :

$$\begin{aligned} p(y \mid \theta_1, \theta_2, \tau) &= p(y_1, \dots, y_\tau \mid \theta_1) p(y_{\tau+1}, \dots, y_n \mid \theta_2) \\ &= \prod_{j=1}^{\tau} p(y_j \mid \theta_1) \prod_{j=\tau+1}^n p(y_j \mid \theta_2). \end{aligned} \quad (3.2)$$

Cette classe de modèles permet la représentation de divers types de changement dans la structure aléatoire de chroniques hydrométéorologiques, notamment la possibilité d'une rupture en localisation, en échelle et en forme de la densité de probabilité. Le modèle (3.1) repose sur l'hypothèse d'indépendance entre les variables aléatoires. L'hydrologie, telle qu'elle s'est développée au cours des dernières décennies, admet généralement que les séries hydrométéorologiques annuelles observables en une station, en plus d'être stationnaires et normalement distribuées, sont constituées de réalisations successives indépendantes. Cette hypothèse est raisonnable pour les débits annuels de rivières qui ne sont pas soumises à l'effet d'un fort emmagasinement. Par contre, pour les débits du fleuve Saint-Laurent, l'hypothèse d'indépendance interannuelle ne serait pas acceptable (l'auto-corrélation d'ordre 1 est de l'ordre de 0.7). En effet, le Saint-Laurent prend sa source dans les Grands Lacs dont le pouvoir d'emmagasinement est très important. La probabilité d'obtenir un débit donné cette année est sans aucun doute influencée par l'intensité du débit de l'année dernière.

La plupart des auteurs qui ont considéré un modèle de rupture similaire à (3.1) se sont concentrés plus particulièrement sur l'estimation de l'instant de changement τ . Pour ce faire, plusieurs approches ont été considérées. Des estimateurs du maximum de vraisemblance ont été proposés notamment par Hinkley (1970), Hinkley et Hinkley (1970), Picard (1985), Joseph et Wolfson (1993) et Ghorbanzadeh (1996). D'autres se sont appuyés sur des techniques dérivées de la procédure graphique des sommes cumulées (notamment Page (1955)) et sur la méthode des moindres carrés (Hartigan (1994), Lavielle et Moulines (2000)). À notre connaissance, Chernoff et Zacks (1964) sont les premiers à avoir considéré l'approche bayésienne pour analyser une rupture dans une suite de variables aléatoires indépendantes. Leurs travaux sont à la base de la plupart des études réalisées sur le sujet dans les années 60 (voir, par exemple, Kander et Zacks (1966), Mustafi (1968), Gardner (1969)). Depuis les années 70, de nombreux statisticiens ont contribué à l'analyse bayésienne d'une seule rupture dans la distribution d'une suite de variables aléatoires. Citons en particulier Blanchet et al. (1971), Smith (1975), Sen et Srivastava (1975), Lee et Heghinian (1977), Smith (1981), Menzefricke (1981), Booth et Smith (1982), Diaz (1982), Hsu (1982), Broemeling (1985), Carlin et al. (1992), Barry et Hartigan (1993), Stephens (1994), Bernier (1994b), Chaouche et Parent (1998), Bélisle et al. (1998). Le lecteur peut consulter Shaban (1980) pour des références additionnelles d'articles publiés avant 1980 ; celui-ci propose une revue de la littérature commentée sur le sujet.

3.1.2 Représentation des quantités non observables : lois a priori conjuguées

Pour pouvoir mener à terme l'analyse bayésienne du modèle (3.1), on doit d'abord représenter la connaissance (ou l'ignorance partielle) sur les paramètres $(\theta_1, \theta_2, \tau)$ en la traduisant par une densité de probabilité dite a priori, qui dépend aussi de paramètres appelés « hyperparamètres ». Ensuite, par le théorème de Bayes (Figure 2.8), la vraisemblance (3.2) et la densité a priori $p(\theta_1, \theta_2, \tau)$ sont combinées pour obtenir la densité de probabilité a posteriori :

$$p(\theta_1, \theta_2, \tau | \mathbf{y}) \propto p(\mathbf{y} | \theta_1, \theta_2, \tau) p(\theta_1, \theta_2, \tau), \quad (3.3)$$

c'est-à-dire la mise à jour du savoir a priori au regard des observations. La loi a posteriori constitue alors le point de départ pour toute forme d'inférence ou de décision concernant les paramètres $(\theta_1, \theta_2, \tau)$.

La représentation de l'incertitude sur les paramètres du modèle de rupture (3.1) est cruciale. Le choix d'une loi de probabilité a priori fait partie intégrante de la démarche de modélisation, et ce, avec toutes les hypothèses que la construction d'un modèle implique. Cette opération n'est pas le résultat de l'examen des observations à analyser, mais des connaissances subjectives de l'hydrologue, d'informations issues d'analyses antérieures ou d'études régionales. Il est généralement difficile d'obtenir l'information nécessaire à la spécification d'une loi a priori et il faudra souvent faire un compromis entre le réalisme de la représentation, la parcimonie de la loi et la facilité avec laquelle l'analyse bayésienne peut être mise en oeuvre.

Presque tous les auteurs supposent l'indépendance entre l'instant de rupture et les autres paramètres. De plus, plusieurs d'entre eux emploient des lois a priori non informatives (notamment, Smith (1975), Lee et Heghinian (1977) et Booth et Smith (1982)). Dans cette thèse, j'ai choisi d'adopter la première hypothèse mais de rejeter la seconde. Plus précisément, je suppose a priori que :

H₁ : $p(\theta_1, \theta_2, \tau) = p(\theta_1, \theta_2) p(\tau)$. Cette hypothèse signifie, en particulier, que la connaissance de l'année de changement n'est nullement déterminée par l'intensité des débits annuels, du moins avant d'avoir pris en compte les observations. Si a priori l'hydrologue soupçonne qu'une rupture ne peut survenir que lorsque le débit atteint un certain seuil, cette hypothèse devient invalide. Enfin, la loi a priori $p(\tau)$ de l'instant de rupture est une loi discrète qui a comme support $\{1, 2, \dots, n - 1\}$.

H₂ : l'hydrologue dispose de connaissances et d'informations sur le phénomène, aussi limitées soient-elles, qui peuvent être employées pour déterminer complètement les lois a priori des paramètres du modèle de rupture. Je rejette donc la situation où l'information fait entièrement défaut et je ne considère que l'emploi de lois a priori informatives.

L'inférence sur les paramètres du modèle (3.1) requiert l'intégration de l'expression (3.3) par rapport à certains paramètres. Par exemple, pour faire de l'inférence sur l'instant de rupture τ , on devra intégrer la loi a posteriori par rapport aux paramètres θ_1 et θ_2 . On obtiendra ainsi la loi marginale a posteriori $p(\tau | \mathbf{y})$ à partir de laquelle il sera possible d'estimer la date de changement et d'en apprécier l'incertitude. Si l'on considère une loi a priori non informative pour (θ_1, θ_2) , cette opération peut généralement s'effectuer sans difficultés (Smith (1975)). Par contre, avec une loi informative, elle amène plus souvent qu'autrement des problèmes techniques importants, et ce, même si les modèles considérés sont simples. Pour faciliter la mise en oeuvre de l'analyse bayésienne, nous adoptons, à τ fixé, une loi jointe a priori conjuguée pour (θ_1, θ_2) . Ainsi s'ajoute une hypothèse quant à la représentation des connaissances a priori :

H_3 : La loi jointe a priori conjuguée, $p(\theta_1, \theta_2)$, est une distribution pour laquelle la transposition des connaissances de l'hydrologue en terme de probabilité d'occurrence peut être faite simplement. De plus, elle est assez riche de forme pour bien représenter une variété de croyances subjectives et d'informations objectives sur le phénomène de changement dans une série de débits.

La densité $p(\theta_1, \theta_2)$ aura donc la même « structure » mathématique que la fonction de vraisemblance $p(\mathbf{y} | \theta_1, \theta_2, \tau)$, lorsque celle-ci est vue comme une fonction des paramètres. En conséquence, la distribution jointe a posteriori $p(\theta_1, \theta_2 | \tau, \mathbf{y})$, conditionnelle à l'instant de rupture et aux données observées, appartiendra à la même classe de lois que $p(\theta_1, \theta_2)$. La mise à jour par le théorème de Bayes se réduira alors simplement à un changement d'hyperparamètres. Pour les modèles de rupture dont l'expression de la vraisemblance permet de séparer les paramètres θ_1 et θ_2 , il sera possible de déterminer explicitement toutes les lois marginales a posteriori : $p(\tau | \mathbf{y})$, $p(\theta_1 | \mathbf{y})$ et $p(\theta_2 | \mathbf{y})$. Pour d'autres modèles, par contre, certaines intégrales ne pourront être résolues analytiquement. L'emploi d'une loi conjuguée nous permettra toutefois de déterminer exactement les lois conditionnelles complètes $p(\tau | \theta_1, \theta_2, \mathbf{y})$, $p(\theta_1 | \theta_2, \tau, \mathbf{y})$ et $p(\theta_2 | \theta_1, \tau, \mathbf{y})$ et d'avoir recours à l'échantillonnage de Gibbs (Annexe D). On évitera ainsi l'emploi de méthodes numériques beaucoup plus difficiles à mettre en oeuvre et parfois même inefficaces (?).

Les lois a priori conjuguées ont été proposées par Raiffa et Schlaifer (1961). Ils justifient leur approche en invoquant le fait que l'information apportée par les observations sur les paramètres est limitée et ne doit pas conduire à remettre en cause la structure générale de la loi a priori, mais seulement ses paramètres. Cette approche est toutefois critiquée par les bayésiens de la dernière génération. D'une part, elle obéit à des contraintes techniques plutôt qu'à des impératifs d'adéquation à l'information a priori (Robert (1992)). D'autre part, le développement des outils MCMC, comme l'algorithme de Metropolis-Hasting (Chib et Greenberg (1995)), permettent maintenant de résoudre la plupart des problèmes d'intégration multiple et d'éviter l'emploi de méthodes numériques. Or, la « conjugaison » est régulièrement employée dans la littérature récente en statistique par ces mêmes bayésiens. Ceci est dû en particulier à l'essor des modèles hiérarchiques. Leur structure conditionnelle, combinée à l'emploi de lois a priori conjuguées,

justifie alors le recours à l'échantillonnage de Gibbs (Gelfand et Smith (1990), Casella et George (1992)), plus simple à mettre en oeuvre que l'algorithme de Metropolis-Hasting.

Il est important de spécifier que seules les lois de la famille exponentielle admettent des distributions a priori conjuguées (Robert (1992), Bernardo et Smith (1994)). L'emploi de ces lois ne constitue pas pour autant une perte de généralité très importante pour l'analyse d'une rupture dans les séries de variables hydrométéorologiques. En effet, la famille exponentielle contient en particulier les lois binomiale, de Poisson, normale, exponentielle, gamma et gaussienne inverse généralisée (ou de Halphen), toutes employées fréquemment en hydrologie statistique pour ce type de variable (intensité des débits ou des précipitations annuels, durée des étiages, nombre d'événements hydrologiques dépassant un certain seuil). Toutefois, les distributions des valeurs extrêmes, généralement considérées pour l'analyse des débits de crue, n'appartiennent pas à cette famille. En conséquence, les procédures décrites dans cette thèse, notamment l'emploi de l'algorithme de Gibbs, ne pourront être utilisées pour les lois Gumbel, Fréchet ou Weibull sans quelques remaniements. Une solution serait de les adapter en s'inspirant, par exemple, des travaux de Coles et Tawn (1996) concernant l'ajustement bayésien de la loi généralisée des valeurs extrêmes. Pour résoudre les problèmes d'estimation, ces derniers ont utilisé une méthode hybride faisant intervenir à la fois l'échantillonnage de Gibbs et l'algorithme de Metropolis-Hasting.

3.2 Estimation des paramètres conditionnelle à l'existence d'une rupture en variance

Pour illustrer la mise en oeuvre de l'analyse bayésienne dans un modèle de rupture, nous reprenons les travaux de Menzefricke (1981) et traitons le problème du changement de variance dans une suite de variables aléatoires normales. Deux modèles sont considérés : le changement simultané de moyenne et de variance et le changement de variance seulement. L'hypothèse H_1 ainsi que l'utilisation d'une loi jointe a priori conjuguée connaissant τ nous permettent d'évaluer analytiquement toutes les lois marginales a posteriori pour le premier modèle. L'inférence sur les paramètres est alors directe. Le second modèle, par contre, n'admet pas de solution explicite pour la détermination des lois a posteriori et nous aurons alors recours à l'échantillonnage de Gibbs (Annexe D). L'Annexe C contient l'expression des densités de probabilité et des principaux moments employés dans cette section ainsi que dans le reste du document.

3.2.1 Changement simultané de moyenne et de variance

Considérons une série de variables aléatoires Y_1, \dots, Y_n , par exemple les n débits annuels historiques d'une rivière. Supposons que la moyenne et la variance du régime aient changé subitement et simultanément en raison de facteurs exogènes. Un modèle de rupture traduisant le comportement d'une telle série est donné par :

$$\begin{aligned} Y_j &\sim \mathcal{N}(y_j \mid \mu_1, \sigma_1^2), \quad j = 1, \dots, \tau, \\ Y_j &\sim \mathcal{N}(y_j \mid \mu_2, \sigma_2^2), \quad j = \tau + 1, \dots, n, \end{aligned} \quad (3.4)$$

où $\mathcal{N}(y \mid \mu, \sigma^2)$ désigne la fonction de densité de probabilité (f.d.p.) de la loi normale de moyenne $\mu \in \mathbb{R}$ et de variance $\sigma^2 \in \mathbb{R}^+$. A partir du modèle (3.4) et de (3.2), on déduit la fonction de vraisemblance résultant d'une réalisation $\mathbf{y} = (y_1, y_2, \dots, y_n)$:

$$\begin{aligned} p(\mathbf{y} \mid \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \tau) &= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma_1^2}\right)^{\frac{\tau}{2}} \left(\frac{1}{\sigma_2^2}\right)^{\frac{n-\tau}{2}} \exp\left\{-\frac{\tau}{2\sigma_1^2} [s_\tau^2 + (\bar{y}_\tau - \mu_1)^2]\right\} \\ &\times \exp\left\{-\frac{n-\tau}{2\sigma_2^2} [s_{n-\tau}^2 + (\bar{y}_{n-\tau} - \mu_2)^2]\right\}, \end{aligned} \quad (3.5)$$

où $\boldsymbol{\theta}_1 = (\mu_1, \sigma_1^2)$, $\boldsymbol{\theta}_2 = (\mu_2, \sigma_2^2)$ et

$$\begin{aligned} \bar{y}_\tau &= \frac{1}{\tau} \sum_{j=1}^{\tau} y_j, & \bar{y}_{n-\tau} &= \frac{1}{n-\tau} \sum_{j=\tau+1}^n y_j, \\ s_\tau^2 &= \frac{1}{\tau} \sum_{j=1}^{\tau} (y_j - \bar{y}_\tau)^2, & s_{n-\tau}^2 &= \frac{1}{n-\tau} \sum_{j=\tau+1}^n (y_j - \bar{y}_{n-\tau})^2. \end{aligned}$$

Si à τ fixé on considère l'expression (3.5) comme une fonction des paramètres, on reconnaît, à une constante près, le produit de deux densités normales avec celle d'une loi gamma inverse. Connaissant τ , la loi jointe a priori conjuguée appartient donc à la classe des lois normale-gamma inverse (Bernardo et Smith (1994)). L'hypothèse \mathbf{H}_1 nous amène à considérer, pour représenter la connaissance a priori sur les paramètres, la distribution suivante:

$$\begin{aligned} p(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \tau) &= p(\boldsymbol{\theta}_1) p(\boldsymbol{\theta}_2) p(\tau), \\ &= p(\mu_1 \mid \sigma_1^2) p(\sigma_1^2) p(\mu_2 \mid \sigma_2^2) p(\sigma_2^2) p(\tau), \\ &= \mathcal{N}(\mu_1 \mid \phi_1, \lambda_1 \sigma_1^2) \mathcal{IG}(\sigma_1^2 \mid \alpha_1, \beta_1) \mathcal{N}(\mu_2 \mid \phi_2, \lambda_2 \sigma_2^2) \mathcal{IG}(\sigma_2^2 \mid \alpha_2, \beta_2) p(\tau), \\ &= \mathcal{NNIGIG}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \tau \mid \boldsymbol{\phi}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}) p(\tau), \end{aligned} \quad (3.6)$$

où $\boldsymbol{\phi} = (\phi_1, \phi_2)$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$, $\boldsymbol{\beta} = (\beta_1, \beta_2)$ sont les hyperparamètres et où $\mathcal{IG}(y \mid \alpha, \beta)$ désigne la f.d.p. de la loi gamma inverse de paramètres $\alpha \in \mathbb{R}^+$ et $\beta \in \mathbb{R}^+$.

La suite de cette section est consacrée à l'évaluation des lois marginales a posteriori des paramètres d'intérêt. Dans la première proposition, nous donnons l'expression de la densité marginale a posteriori de l'instant de rupture, $p(\tau \mid \mathbf{y})$.

Proposition 1 *Soit le modèle de rupture (3.4) et considérons la loi jointe a priori (3.6).*

Alors, les hyperparamètres¹ mis à jour sont donnés par :

$$\begin{aligned}\lambda'_1 &= \lambda_1 / (1 + \tau \lambda_1), & \lambda'_2 &= \lambda_2 / [1 + (n - \tau) \lambda_2], \\ \phi'_1 &= (1 - \lambda'_1 \tau) \phi_1 + \lambda'_1 \tau \bar{y}_\tau, & \phi'_2 &= [1 - \lambda'_2 (n - \tau)] \phi_2 + \lambda'_2 (n - \tau) \bar{y}_{n-\tau}, \\ \alpha'_1 &= \alpha_1 + \tau / 2, & \alpha'_2 &= \alpha_2 + (n - \tau) / 2, \\ \beta'_1 &= \frac{\tau}{2} [s_\tau^2 + (1 - \lambda'_1 \tau) (\phi_1 - \bar{y}_\tau)^2] + \beta_1, \\ \beta'_2 &= \frac{(n-\tau)}{2} [s_{n-\tau}^2 + (1 - \lambda'_2 (n - \tau)) (\phi_2 - \bar{y}_{n-\tau})^2] + \beta_2,\end{aligned}$$

et la densité marginale a posteriori de l'instant de rupture s'exprime de la manière suivante :

$$p(\tau | \mathbf{y}) \propto p(\tau) \sqrt{\lambda'_1 \lambda'_2} \frac{\Gamma(\alpha'_1) \Gamma(\alpha'_2)}{(\beta'_1)^{\alpha'_1} (\beta'_2)^{\alpha'_2}}. \quad (3.7)$$

Démonstration. La démarche consiste à calculer dans un premier temps la loi jointe a posteriori de (θ_1, θ_2) en supposant que l'instant de rupture τ est connu. Selon le théorème de Bayes (Figure 2.8), cette distribution est proportionnelle au produit suivant :

$$p(\mathbf{y} | \theta_1, \theta_2, \tau) \mathcal{N} \mathcal{N} \mathcal{I} \mathcal{G} \mathcal{I} \mathcal{G}(\theta_1, \theta_2 | \phi, \lambda, \alpha, \beta), \quad (3.8)$$

qui correspond, à une constante de normalisation près, à une loi normale-gamma inverse. En conformité avec l'hypothèse de conjugaison, la loi a posteriori est alors obtenue par une simple mise à jour des hyperparamètres de la loi jointe a priori :

$$p(\theta_1, \theta_2 | \tau, \mathbf{y}) = \mathcal{N} \mathcal{N} \mathcal{I} \mathcal{G} \mathcal{I} \mathcal{G}(\theta_1, \theta_2 | \phi', \lambda', \alpha', \beta'). \quad (3.9)$$

La détermination de $(\phi', \lambda', \alpha', \beta')$ fait intervenir en particulier des complétions de carrés. La loi marginale a posteriori de τ peut ensuite être déduite de la constante de normalisation, c'est-à-dire de la loi marginale des observations conditionnelles à τ , $p(\mathbf{y} | \tau)$. En effet, selon le théorème de Bayes,

$$p(\tau | \mathbf{y}) = \frac{p(\mathbf{y} | \tau) p(\tau)}{\sum_{\tau=1}^{n-1} p(\mathbf{y} | \tau) p(\tau)}. \quad (3.10)$$

Or, $p(\mathbf{y} | \tau)$ peut être évaluée en faisant le rapport de (3.8) par (3.9) et en éliminant les termes impliquant seulement les paramètres θ_1 et θ_2 . On obtient alors

$$p(\mathbf{y} | \tau) = \left(\frac{1}{2\pi}\right)^{n/2} \sqrt{\frac{\lambda'_1 \lambda'_2}{\lambda_1 \lambda_2} \frac{\beta_1^{\alpha_1} \beta_2^{\alpha_2}}{(\beta'_1)^{\alpha'_1} (\beta'_2)^{\alpha'_2}} \frac{\Gamma(\alpha'_1) \Gamma(\alpha'_2)}{\Gamma(\alpha_1) \Gamma(\alpha_2)}}, \quad (3.11)$$

1. Pour simplifier la notation, nous omettons de signaler la dépendance à τ et \mathbf{y} en écrivant par exemple λ'_1 au lieu de $\lambda'_1(\mathbf{y}, \tau)$.

d'où, à partir de (3.10), on déduit le résultat (3.7). ■

L'expression (3.7) fournit, à une constante près, la mise à jour de la répartition a priori des probabilités d'occurrence de rupture pour chaque instant de la série. Elle permet d'identifier, par exemple, le moment où le changement simultané de moyenne et de variance possède la plus grande probabilité d'occurrence a posteriori. La constante de normalisation étant indépendante de τ , elle peut être aisément évaluée en sommant les $n - 1$ valeurs que peut prendre l'expression (3.7).

Les lois marginales a posteriori des autres paramètres qui caractérisent le changement correspondent à des mélanges de $n - 1$ lois continues, dont les poids correspondent aux valeurs de $p(\tau | \mathbf{y})$. La proposition qui suit donne l'expression de leur densité de probabilité.

Proposition 2 *Soit le modèle de rupture (3.4) et considérons la loi jointe a priori (3.6). Les fonctions de densité de probabilité marginales a posteriori des paramètres μ_1, μ_2, σ_1^2 et σ_2^2 sont alors les suivantes :*

$$p(\mu_k | \mathbf{y}) = \sum_{\tau=1}^{n-1} ST\left(\mu_k \mid \phi'_k, \alpha'_k (\lambda'_k \beta'_k)^{-1}, 2\alpha'_k\right) p(\tau | \mathbf{y}), \quad k = 1, 2, \quad (3.12)$$

$$p(\sigma_k^2 | \mathbf{y}) = \sum_{\tau=1}^{n-1} IG(\sigma_k^2 \mid \alpha'_k, \beta'_k) p(\tau | \mathbf{y}), \quad k = 1, 2, \quad (3.13)$$

où $ST(\mu_k | a, b, c)$ désigne la f.d.p. de la loi de Student et où la densité $p(\tau | \mathbf{y})$ ainsi que les hyperparamètres $\phi'_1, \phi'_2, \alpha'_1, \alpha'_2, \beta'_1$ et β'_2 sont donnés à la Proposition 1.

Démonstration. Toutes les densités marginales de la loi jointe a posteriori (3.9) peuvent être obtenues explicitement par intégration. Après quelques manipulations algébriques, on a :

$$p(\mu_k | \tau, \mathbf{y}) = ST\left(\mu_k \mid \phi'_k, \alpha'_k (\lambda'_k \beta'_k)^{-1}, 2\alpha'_k\right), \quad k = 1, 2 \quad (3.14)$$

et

$$p(\sigma_k^2 | \tau, \mathbf{y}) = IG(\sigma_k^2 \mid \alpha'_k, \beta'_k), \quad k = 1, 2, \quad (3.15)$$

d'où l'on déduit par exemple, en vertu de $p(\mu_k, \tau | \mathbf{y}) = p(\mu_k | \tau, \mathbf{y}) p(\tau | \mathbf{y})$, que

$$\begin{aligned} p(\mu_k | \mathbf{y}) &= \sum_{\tau=1}^{n-1} p(\mu_k, \tau | \mathbf{y}) \\ &= \sum_{\tau=1}^{n-1} p(\mu_k | \tau, \mathbf{y}) p(\tau | \mathbf{y}) \\ &= \sum_{\tau=1}^{n-1} ST\left(\mu_k \mid \phi'_k, \alpha'_k (\lambda'_k \beta'_k)^{-1}, 2\alpha'_k\right) p(\tau | \mathbf{y}), \quad k = 1, 2. \end{aligned} \quad (3.16)$$

La loi marginale $p(\sigma_k^2 | \mathbf{y})$ s'obtient de la même manière. ■

Pour évaluer l'amplitude de la rupture en variance, il est naturel de définir la quantité $\eta = \sigma_2^2 / \sigma_1^2$. La densité a posteriori $p(\eta | \tau, \mathbf{y})$, conditionnelle à l'instant de rupture, peut être aisément déduite des lois conditionnelles (3.15). De manière analogue aux paramètres naturel du modèle, la distribution marginale a posteriori de η s'exprime aussi comme un mélange de lois :

$$p(\eta | \mathbf{y}) = \sum_{\tau=1}^{n-1} p(\eta | \tau, \mathbf{y}) p(\tau | \mathbf{y}).$$

La proposition qui suit donne l'expression de la densité a posteriori conditionnelle de η . Il s'agit d'une loi Beta du second ordre (Box et Tiao (1964)).

Proposition 3 *Soit le modèle (3.4) et la loi jointe a priori (3.6). Considérons le paramètre $\eta = \sigma_2^2 / \sigma_1^2$ définissant l'amplitude de la rupture en variance. Alors la densité de $(\eta | \tau, \mathbf{y})$ s'exprime de la manière suivante :*

$$p(\eta | \tau, \mathbf{y}) = \frac{(\beta'_1)^{\alpha'_1} (\beta'_2)^{\alpha'_2}}{B(\alpha'_1, \alpha'_2)} \frac{\eta^{\alpha'_1 - 1}}{(\beta'_1 \eta + \beta'_2)^{\alpha'_1 + \alpha'_2}}, \quad (3.17)$$

où $\alpha'_1, \alpha'_2, \beta'_1$ et β'_2 sont donnés à la Proposition 1 et $B(a, b) = \Gamma(a) \Gamma(b) / \Gamma(a + b)$ est la fonction beta.

Démonstration. La densité jointe a posteriori de (σ_1^2, σ_2^2) , conditionnelle à τ , s'obtient directement en intégrant (3.9) par rapport à μ_1 et μ_2 . Elle correspond au produit de deux densités gamma inverse :

$$p(\sigma_1^2, \sigma_2^2 | \tau, \mathbf{y}) = \mathcal{IG}(\sigma_1^2 | \alpha'_1, \beta'_1) \mathcal{IG}(\sigma_2^2 | \alpha'_2, \beta'_2).$$

Si on effectue le changement de variables $\theta = \sigma_2^2$ et $\eta = \sigma_2^2 / \sigma_1^2$, l'intégration par rapport à θ de la densité $p(\eta, \theta | \tau, \mathbf{y})$ ainsi obtenue mène au résultat après quelques manipulations algébriques. ■

Toutes les lois a posteriori présentées dans cette section peuvent être aisément utilisées pour estimer les paramètres d'intérêt et évaluer leur incertitude. Par exemple, si l'on considère une fonction de coût quadratique pour un paramètre θ , $C(a, \theta) = (\theta - a)^2$, il est bien connu que l'estimateur optimal \hat{a} , c'est-à-dire celui qui minimise le risque, correspond alors à l'espérance des lois marginales a posteriori (Berger (1985)). Pour l'amplitude de la rupture en variance, l'espérance a posteriori s'écrit de la manière suivante :

$$\mathbb{E}\{\eta | \mathbf{y}\} = \mathbb{E}_{\tau | \mathbf{y}} \{\mathbb{E}\{\eta | \tau, \mathbf{y}\}\} = \sum_{\tau=1}^{n-1} \mathbb{E}\{\eta | \tau, \mathbf{y}\} p(\tau | \mathbf{y}) \quad (3.18)$$

et s'exprime donc comme une moyenne pondérée des espérances conditionnelles. La fonction de risque dans ce cas correspond à la variance a posteriori de η , qui peut s'écrire comme suit :

$$\begin{aligned} \mathbb{V} \{ \eta \mid \mathbf{y} \} &= \mathbb{E}_{\tau \mid \mathbf{y}} \{ \mathbb{V} \{ \eta \mid \tau, \mathbf{y} \} \} + \mathbb{V}_{\tau \mid \mathbf{y}} \{ \mathbb{E} \{ \eta \mid \tau, \mathbf{y} \} \}, \\ &= \sum_{\tau=1}^{n-1} [\mathbb{V} \{ \eta \mid \tau, \mathbf{y} \} + \mathbb{E}^2 \{ \eta \mid \tau, \mathbf{y} \}] p(\tau \mid \mathbf{y}) - \mathbb{E}^2 \{ \eta \mid \mathbf{y} \}, \end{aligned} \quad (3.19)$$

où la notation $\mathbb{E}_{\tau \mid \mathbf{y}}$ et $\mathbb{V}_{\tau \mid \mathbf{y}}$ signifie que l'espérance et la variance sont calculées par rapport à la loi a posteriori de τ . Ces deux dernières équations s'appliquent aussi pour les paramètres naturels μ_k et σ_k^2 .

Le Tableau 3.1 donne les espérances et les variances des lois conditionnelles a posteriori pour les paramètres μ_k , σ_k^2 et η , nécessaires à l'évaluation de (3.18) et (3.19). Pour μ_k et σ_k^2 , ces expressions sont déduites directement de l'espérance et de la variance des lois de Student et gamma inverse (voir Annexe C). Les moments de η sont obtenus à partir des moments de la loi de Fisher. En effet, on montre par un simple changement de variables que $y = [\alpha'_1 \beta'_1 / (\alpha'_2 \beta'_2)] \eta$ suit une loi de Fisher $\mathcal{F}(y \mid 2\alpha'_2, 2\alpha'_1)$.

TAB. 3.1 – *Espérances et variances a posteriori pour les paramètres du modèle de rupture simultanée en moyenne et en variance*

	Espérance	Variance
μ_k	ϕ'_k	$2\lambda'_k \beta'_k (2\alpha'_k - 1)^{-1}$
σ_k^2	$\beta'_k (\alpha'_k - 1)^{-1}$	$(\beta'_k)^2 [(\alpha'_k - 1)^2 (\alpha'_k - 1)]^{-1}$
η	$\alpha'_2 \beta'_2 [(\alpha'_1 - 1) \beta'_1]^{-1}$	$(\beta'_2)^2 \alpha'_2 (\alpha'_1 + \alpha'_2 - 1) [(\beta'_1)^2 (\alpha'_1 - 1)^2 (\alpha'_2 - 1)]^{-1}$

La fonction de coût quadratique pour l'estimation des paramètres implique que les pertes encourues par une surestimation sont équivalentes à celles engendrées par une sous-estimation. Comme il a été mentionné au Chapitre 2, pour certains problèmes, une telle fonction de coût symétrique peut être inadéquate. Pour attacher des impacts différents à la surestimation et à la sous-estimation, on peut suggérer la fonction de coût suivante :

$$\begin{aligned} C(a, \theta) &= \begin{cases} c_1 (a - \theta), & \text{si } \theta \leq a, \\ c_2 (\theta - a), & \text{si } \theta > a, \end{cases} \\ &= c_1 (a - \theta) \mathbf{1}_{(\theta \leq a)}(a) + c_2 (\theta - a) \mathbf{1}_{(\theta > a)}(a). \end{aligned} \quad (3.20)$$

L'estimateur optimal \tilde{a} correspondrait alors au quantile a posteriori de probabilité au dépassement $p = c_2 / (c_1 + c_2)$ (Berger (1985)). Nous nous limitons néanmoins, jusqu'au Chapitre 6, à l'emploi d'une fonction de coût quadratique.

3.2.2 Changement de variance seulement

Dans cette section nous considérons le modèle de rupture suivant :

$$\begin{aligned} Y_i &\sim \mathcal{N}(y_i \mid \mu_1, \sigma_1^2), & i = 1, \dots, \tau, \\ Y_i &\sim \mathcal{N}(y_i \mid \mu_1, \sigma_2^2), & i = \tau + 1, \dots, n, \end{aligned} \quad (3.21)$$

où seule la variance change, la moyenne μ_1 des observations étant invariable de part et d'autre de l'instant de rupture. La fonction de vraisemblance pour la réalisation $\mathbf{y} = (y_1, y_2, \dots, y_n)$ peut s'écrire ainsi :

$$\begin{aligned} p(\mathbf{y} \mid \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \tau) &= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma_1^2}\right)^{\frac{\tau}{2}} \left(\frac{1}{\sigma_2^2}\right)^{\frac{n-\tau}{2}} \exp\left\{-\frac{\tau}{2\sigma_1^2} [s_\tau^2 + (\bar{y}_\tau - \mu_1)^2]\right\} \\ &\times \exp\left\{-\frac{(n-\tau)}{2\sigma_2^2} [s_{n-\tau}^2 + (\bar{y}_{n-\tau} - \mu_1)^2]\right\}, \end{aligned} \quad (3.22)$$

où $\boldsymbol{\theta}_1 = (\mu_1, \sigma_1^2)$, $\boldsymbol{\theta}_2 = \sigma_2^2$.

Pour traduire les connaissances de l'hydrologue sur les paramètres du modèle (3.21), nous adoptons la loi jointe a priori proposée par Menzefricke (1981) :

$$\begin{aligned} p(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \tau) &= p(\boldsymbol{\theta}_1) p(\boldsymbol{\theta}_2) p(\tau), \\ &= p(\mu_1 \mid \sigma_1^2) p(\sigma_1^2) p(\sigma_2^2) p(\tau), \\ &= \mathcal{N}(\mu_1 \mid \phi_1, \lambda_1 \sigma_1^2) \mathcal{IG}(\sigma_1^2 \mid \alpha_1, \beta_1) \mathcal{IG}(\sigma_2^2 \mid \alpha_2, \beta_2) p(\tau), \\ &= \mathcal{NIGIG}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \mid \phi_1, \lambda_1, \boldsymbol{\alpha}, \boldsymbol{\beta}) p(\tau). \end{aligned} \quad (3.23)$$

Outre la loi marginale a posteriori de l'intensité de la rupture $\eta = \sigma_2^2 / \sigma_1^2$, qu'il est possible de déterminer à une constante près (voir Article I, Annexe A), on ne peut évaluer explicitement les lois marginales a posteriori nécessaires à l'inférence sur les autres paramètres d'intérêt σ_1^2 , σ_2^2 et τ . Pour résoudre ce problème, Menzefricke (1981) recommande d'employer l'intégration numérique. Etant donné que les méthodes numériques nécessitent une bonne expertise que les hydrologues et les statisticiens appliqués n'ont généralement pas, nous nous sommes tournés vers les méthodes de simulation MCMC, beaucoup plus faciles à mettre en oeuvre. L'échantillonnage de Gibbs est tout indiqué ici. D'une part, la structure de la loi a priori (3.23) permet de déterminer explicitement les lois a posteriori conditionnelles complètes. D'autre part, ces dernières sont faciles à simuler. La mise en oeuvre de l'échantillonnage de Gibbs et son utilisation pratique est décrite brièvement en Annexe D. Pour un exposé plus général et plus détaillé, on peut consulter ?.

Pour le modèle (3.21), l'algorithme de Gibbs repose essentiellement sur des tirages successifs effectués dans les lois conditionnelles complètes de la manière suivante :

$$\begin{aligned}\mu_1^{(j)} &\sim p\left(\mu_1 \mid \sigma_1^{2(j-1)}, \sigma_2^{2(j-1)}, \tau^{(j-1)}, \mathbf{y}\right), \\ \sigma_1^{2(j)} &\sim p\left(\sigma_1^2 \mid \mu_1^{(j)}, \sigma_2^{2(j-1)}, \tau^{(j-1)}, \mathbf{y}\right), \\ \sigma_2^{2(j)} &\sim p\left(\sigma_2^2 \mid \mu_1^{(j)}, \sigma_1^{2(j)}, \tau^{(j-1)}, \mathbf{y}\right), \\ \tau^{(j)} &\sim p\left(\tau \mid \mu_1^{(j)}, \sigma_1^{2(j)}, \sigma_2^{2(j)}, \mathbf{y}\right).\end{aligned}$$

Dans ce qui suit, nous montrons que les lois conditionnelles complètes du modèle (3.21) sont standards.

Remarquons d'abord que, quelles que soient la loi de probabilité des observations et la structure de la loi a priori $p(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$, si l'instant de rupture dans un modèle du type (3.1) est indépendant des autres paramètres a priori, la loi conditionnelle complète de τ aura toujours la forme suivante :

$$p(\tau \mid \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \mathbf{y}) = \frac{p(\mathbf{y} \mid \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \tau) p(\tau)}{\sum_{\tau=1}^{n-1} p(\mathbf{y} \mid \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \tau) p(\tau)}. \quad (3.24)$$

Cette distribution étant discrète, elle peut être aisément générée. Pour le modèle (3.21), si l'on dispose d'une routine permettant d'évaluer la densité d'une loi normale multidimensionnelle, il est commode d'écrire $p(\tau \mid \mu_1, \sigma_1^2, \sigma_2^2, \mathbf{y})$ de la manière suivante :

$$p(\tau \mid \mu_1, \sigma_1^2, \sigma_2^2, \mathbf{y}) \propto \mathcal{N}_\tau(\mathbf{y}_1(\tau) \mid \mu_1 \mathbf{1}_\tau, \sigma_1^2 \mathbf{I}_\tau) \mathcal{N}_{n-\tau}(\mathbf{y}_2(\tau) \mid \mu_1 \mathbf{1}_{n-\tau}, \sigma_2^2 \mathbf{I}_{n-\tau}) p(\tau), \quad (3.25)$$

où $\mathbf{y}_1(\tau) = (y_1, \dots, y_\tau)^T$, $\mathbf{y}_2(\tau) = (y_{n-\tau}, \dots, y_n)^T$, $\mathbf{1}_p$ est un vecteur colonne unitaire à p éléments, \mathbf{I}_p est la matrice identité d'ordre p et $\mathcal{N}_p(\mathbf{y} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ désigne la densité de probabilité d'une loi normale multidimensionnelle de dimension p .

La proposition suivante donne les lois conditionnelles complètes des autres paramètres nécessaires à la mise en oeuvre de l'échantillonnage de Gibbs.

Proposition 4 *Soit le modèle de rupture (3.21) et la loi jointe a priori (3.23). Alors, les lois conditionnelles complètes des paramètres μ_1 , σ_1^2 et σ_2^2 sont les suivantes :*

$$p(\mu_1 \mid \sigma_1^2, \sigma_2^2, \tau, \mathbf{y}) = \mathcal{N}(\mu_1 \mid m, v),$$

$$p(\sigma_1^2 \mid \mu_1, \sigma_2^2, \tau, \mathbf{y}) = \mathcal{IG}(\sigma_1^2 \mid a_1, b_1),$$

$$p(\sigma_2^2 \mid \mu_1, \sigma_1^2, \tau, \mathbf{y}) = \mathcal{IG}(\sigma_2^2 \mid a_2, b_2),$$

où

$$m = \frac{1}{\lambda'_1 (n - \tau) \sigma_1^2 + \sigma_2^2} [\sigma_2^2 \phi'_1 + \lambda'_1 (n - \tau) \sigma_1^2 \bar{y}_{n-\tau}] , \quad v = \frac{\lambda'_1 \sigma_1^2 \sigma_2^2}{\lambda'_1 (n - \tau) \sigma_1^2 + \sigma_2^2}$$

$$a_1 = \alpha_1 + \frac{\tau + 1}{2}, \quad a_2 = \alpha_2 + \frac{n - \tau}{2},$$

$$b_1 = \frac{\tau}{2} \left[s_\tau^2 + (\bar{y}_\tau - \mu_1)^2 + \frac{(\mu_1 - \phi_1)^2}{\tau \lambda_1} \right] + \beta_1, \quad b_2 = \frac{n - \tau}{2} \left[s_{n-\tau}^2 + (\bar{y}_{n-\tau} - \mu_1)^2 \right] + \beta_2,$$

avec $\lambda'_1 = \lambda_1 / (1 + \tau \lambda_1)$ et $\phi'_1 = (1 - \lambda'_1 \tau) \phi_1 + \lambda'_1 \tau \bar{y}_\tau$.

Démonstration. Il suffit d'appliquer le théorème de Bayes au paramètre d'intérêt, les autres étant supposés connus. Si on élimine les termes impliquant seulement les paramètres fixés, en vertu du théorème de Bayes, on peut écrire :

$$p(\mu_1 | \sigma_1^2, \sigma_2^2, \tau, \mathbf{y}) \propto \exp \left\{ -\frac{1}{2} \left[\frac{\tau}{\sigma_1^2} (\bar{y}_\tau - \mu_1)^2 + \frac{(n - \tau)}{2\sigma_2^2} (\bar{y}_{n-\tau} - \mu_1)^2 \right] \right\} \mathcal{N}(\mu_1 | \phi_1, \lambda_1 \sigma_1^2),$$

$$p(\sigma_1^2 | \mu_1, \sigma_2^2, \tau, \mathbf{y}) \propto \left(\frac{1}{\sigma_1^2} \right)^{\frac{\tau}{2}} \exp \left\{ -\frac{\tau}{2} \left[s_\tau^2 + (\bar{y}_\tau - \mu_1)^2 \right] \frac{1}{\sigma_1^2} \right\} \mathcal{N}(\mu | \phi_1, \lambda_1 \sigma_1^2) \mathcal{IG}(\sigma_1^2 | \alpha_1, \beta_1)$$

et

$$p(\sigma_2^2 | \mu_1, \sigma_1^2, \tau, \mathbf{y}) \propto \left(\frac{1}{\sigma_2^2} \right)^{\frac{n-\tau}{2}} \exp \left\{ -\frac{(n - \tau)}{2} \left[s_{n-\tau}^2 + (\bar{y}_{n-\tau} - \mu_1)^2 \right] \frac{1}{\sigma_2^2} \right\} \mathcal{IG}(\sigma_2^2 | \alpha_2, \beta_2).$$

Pour chacun des paramètres, on reconnaît un problème de mise à jour standard impliquant une loi a priori conjuguée. Quelques manipulations algébriques élémentaires mènent alors aux résultats. ■

3.3 Application à l'analyse locale d'une rupture dans une séquence d'apports énergétiques annuels

Plusieurs décisions quant à la planification énergétique à Hydro-Québec reposent sur l'évaluation des moyennes et des variances des séries d'apports énergétiques annuels aux huit grands complexes hydroélectriques (Exemple 2, Chapitre 2). Ces séries étant construites à partir des apports naturels, elles sont sujettes aux variations structurelles qui pourraient être induites par des phénomènes climatologiques. Les apports énergétiques annuels de 1943 à 1996 pour les huit complexes sont présentés à la Figure 3.1.

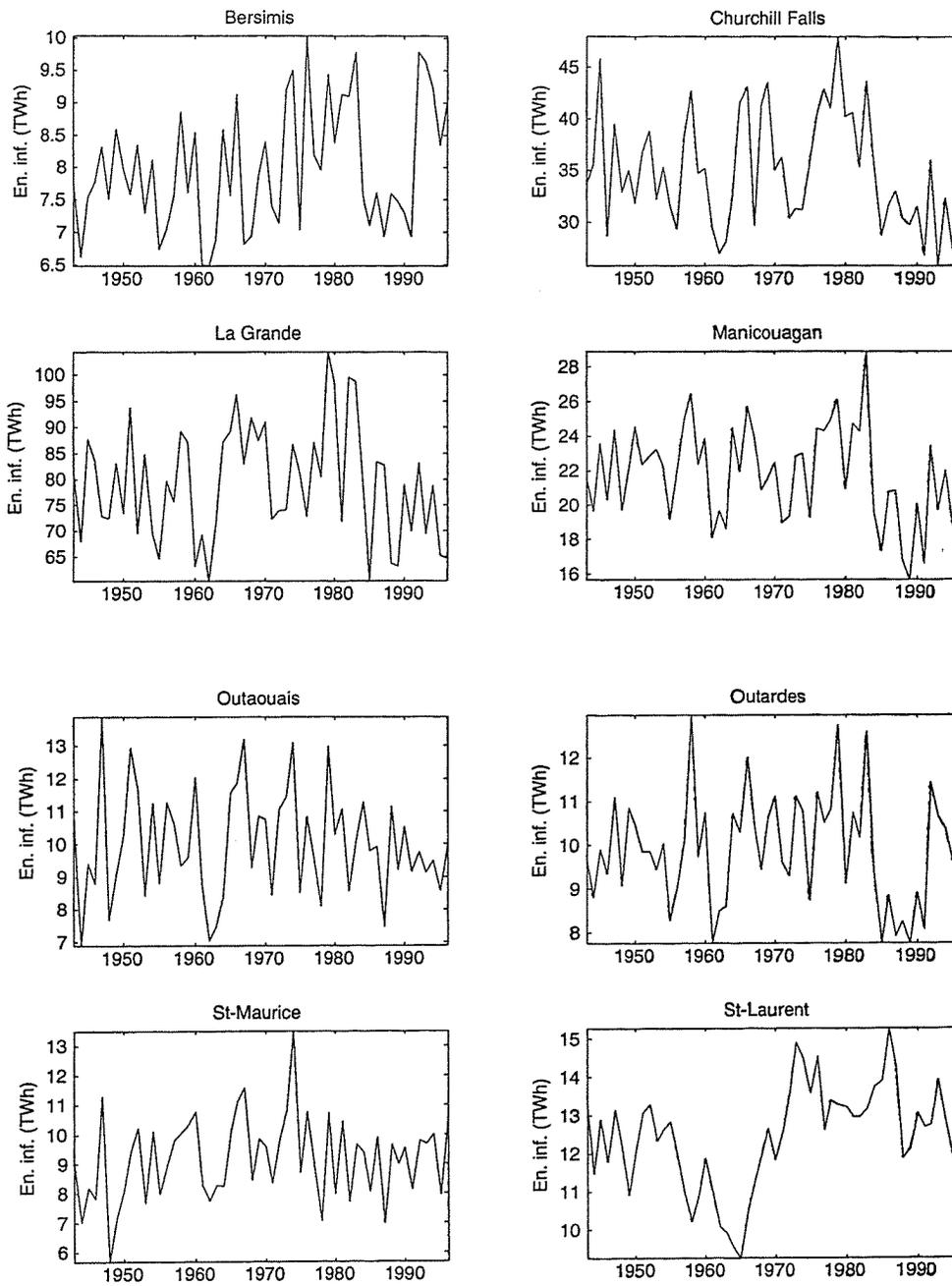


FIG. 3.1 – Apports énergétiques annuels des huit grands complexes hydroélectriques gérés par Hydro-Québec

On constate en effet que certaines de ces séries semblent admettre une rupture en moyenne et/ou en variance. C'est le cas notamment des apports énergétiques des complexes de Churchill Falls et de l'Outaouais. Depuis le milieu des années 70, la région du nord-est du Québec, où le complexe de Churchill Falls se situe, a connu plusieurs années de faible hydraulité. Cette baisse du niveau moyen des débits se reflète dans les apports énergétiques annuels. Elle semble aussi accompagnée d'une baisse de variabilité. Pour le complexe de l'Outaouais, situé à l'extrême sud-ouest de la province, on note plutôt la possibilité d'une baisse de variabilité uniquement, qui serait survenue vers la fin des années 80. Ces observations sont suffisantes pour mettre en doute la stationnarité de ces séries et justifient, du moins à titre exploratoire, d'envisager des modèles de rupture pour la représentation des apports énergétiques annuels. Les estimations des moyennes et des variances issues de tels modèles prendraient en compte de manière formelle la présence d'une rupture. Bien que les hydrologues d'Hydro-Québec soupçonnent la présence de changements dans les séries d'apports énergétiques (Perreault et al. (1996)), on évalue encore aujourd'hui les moyennes et les variances sur toute la période d'échantillonnage.

Dans ce qui suit, nous illustrons la mise en oeuvre de l'analyse bayésienne pour les modèles (3.4) et (3.21) en reprenant les principaux résultats de l'application présentée dans l'Article I (Annexe A). Le premier modèle (changement de moyenne et de variance) est employé pour représenter les apports énergétiques de Churchill Falls. Le second (changement de variance seulement) est appliqué aux observations du complexe de l'Outaouais. Nous montrons que l'analyse bayésienne d'une rupture peut être utile, notamment pour évaluer la marge de manoeuvre que possède Hydro-Québec en matière de planification énergétique. En effet, celle-ci dépend de la disponibilité en énergie, donc des apports énergétiques aux principaux complexes hydroélectriques.

Nous avons choisi de ne favoriser aucune année de rupture a priori en prenant une loi uniforme sur [1943,1995] pour $p(\tau)$. La spécification des hyperparamètres des lois a priori normale-gamma inverse (3.6) et (3.23) a été effectuée sur la base d'une transposition d'information régionale au site d'intérêt à partir des sept autres complexes hydroélectriques. Cette procédure de détermination des lois a priori est décrite en détail dans l'article susmentionné. Les espérances et variances a priori de tous les paramètres sont données au Tableau 3.2.

TAB. 3.2 – *Espérances et variances a priori des paramètres pour les complexes de Churchill Falls et de l'Outaouais*

		τ	μ_1	μ_2	σ_1^2	σ_2^2	δ^*	η^*
Churchill Falls	Espérance	1970	27.80	27.81	12.88	18.19	0.01	1.47
	Modèle (3.4) Écart-type	16	1.67	2.12	2.73	2.45	2.70	0.36
		τ	σ_1^2	σ_2^2	η			
Outaouais	Espérance	1970	1.57	2.21	1.41			
	Modèle (3.21) Écart-type	16	6.70	5.85	4.03			

* $\delta = \mu_2 - \mu_1$, $\eta = \sigma_2^2 / \sigma_1^2$

TAB. 3.3 – *Caractéristiques a posteriori des paramètres pour le complexe de Churchill Falls*

Paramètres	Mode	Espérance	Écart-type	Intervalle de crédibilité à 90%
τ	1984	1983.92	7.87	[1982 ; 1989]
μ_1	35.28	35.07	0.74	[33.65 ; 36.37]
μ_2	29.87	30.00	3.48	[27.88 ; 32.13]
σ_1^2	22.33	22.89	3.48	[16.81 ; 29.16]
σ_2^2	17.07	17.81	2.28	[13.92 ; 22.07]
δ	-5.43	-5.07	2.02	[-6.42 ; -4.19]
η	0.73	0.81	0.25	[0.53 ; 1.13]

3.3.1 Application au complexe de Churchill Falls

Nous considérons le modèle (3.4) de rupture simultanée en moyenne et en variance pour représenter les apports énergétiques annuels du complexe de Churchill Falls (Figure 3.1 et 2.6). Les expressions données dans la Section 3.2.1 sont donc utilisées ici en y incorporant les lois a priori spécifiées par considérations régionales.

Le Tableau 3.3 présente quelques caractéristiques a posteriori des paramètres (le mode, l'espérance et l'écart-type) et fournit pour chacun d'eux un intervalle de crédibilité de niveau de confiance de 90%. Les espérances et les variances a posteriori peuvent être comparées aux quantités correspondantes a priori (Tableau 3.2) pour évaluer dans quelle mesure l'information contenue dans les observations est venue modifier nos connaissances a priori sur les paramètres. La Figure 3.2 présente, avec les apports énergétiques annuels, la répartition des probabilités marginales a posteriori de l'année de rupture $p(\tau | \mathbf{y})$. Elle montre aussi les densités a posteriori des autres paramètres. Sur cette figure les lois a priori correspondantes apparaissent en traits pointillés. Le tracé des lois a posteriori permet en particulier d'apprécier l'incertitude sur les paramètres. Enfin, la Figure 3.3 montre les résidus standardisés du modèle sur papier de probabilité normal. Leur répartition sur ce graphique semble linéaire, ce qui permet de conclure que l'hypothèse de normalité est valide. D'autre part, l'autocorrélation d'ordre 1 des résidus est assez faible (0.16).

À l'examen des graphiques de la Figure 3.2 et des résultats fournis au Tableau 3.3, on retient principalement les observations suivantes :

- La prise en compte des observations vient modifier considérablement l'état des connaissances a priori sur tous les paramètres d'intérêt. Notons en particulier que la loi a posteriori de l'instant de rupture est beaucoup moins diffuse que la loi a priori (loi uniforme). Si l'on interprète les densités comme une représentation de nos degrés de croyance, ce constat signifie que nos connaissances sur l'année de rupture se sont améliorées à la suite de la prise en compte des observations.
- Le changement serait survenu en 1984 avec un écart-type de huit années.

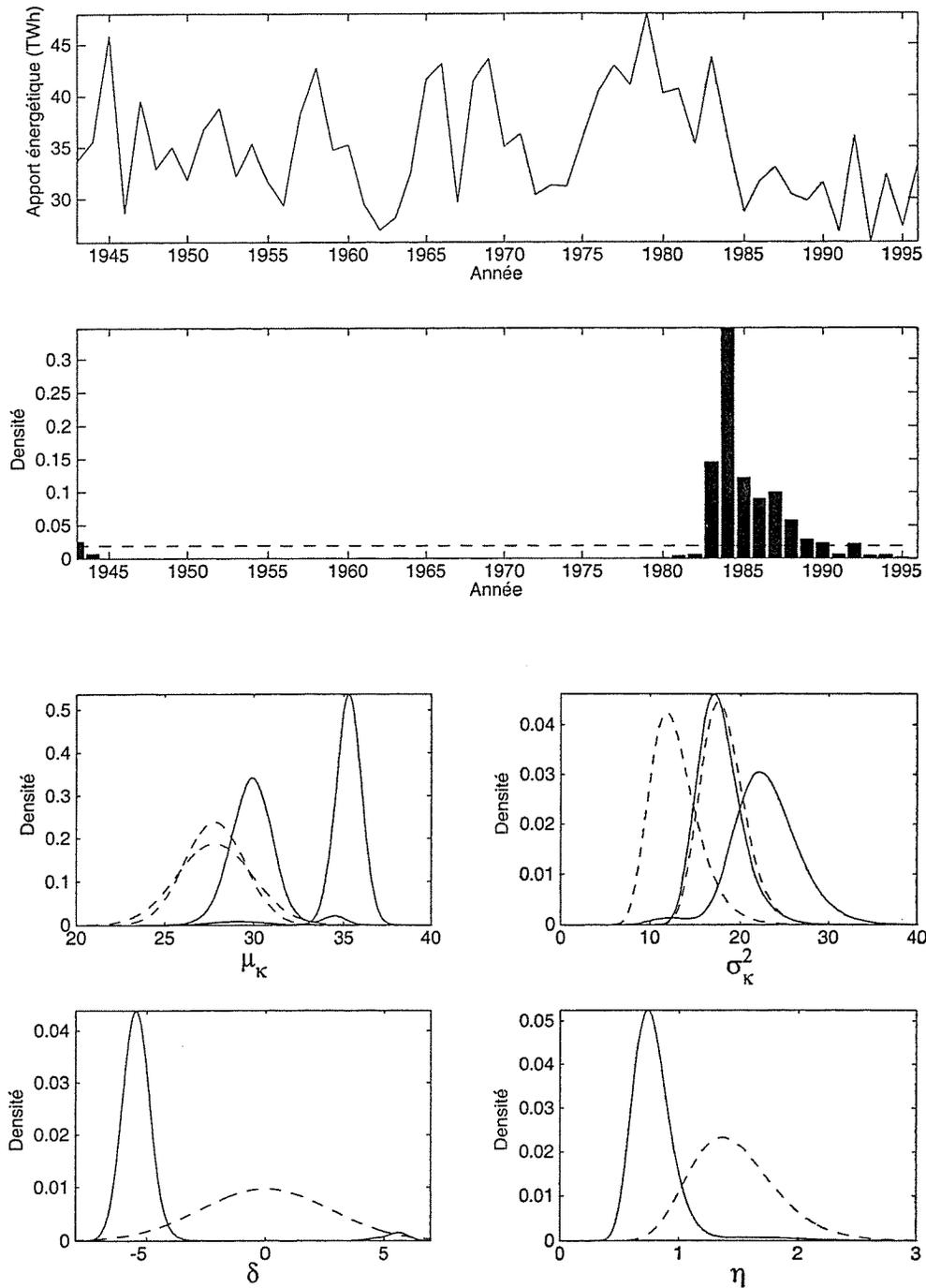


FIG. 3.2 – Apports énergétiques annuels à Churchill Falls, densité marginale a priori (trait pointillé) et a posteriori (trait continu) des paramètres

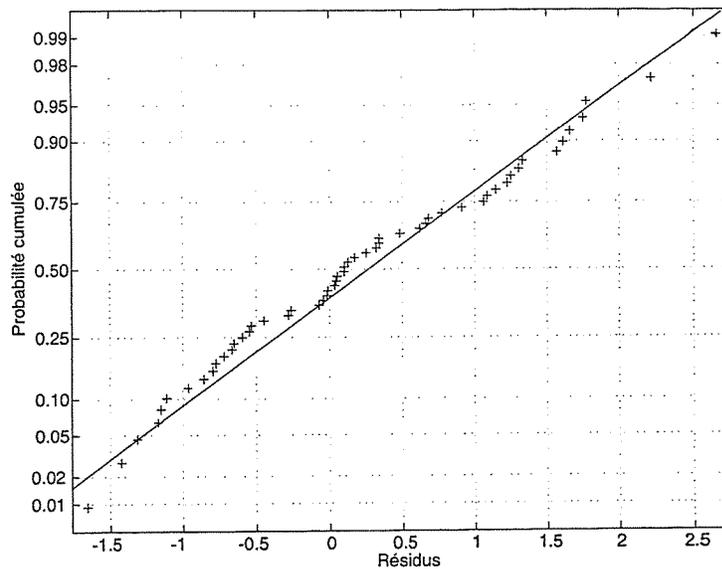


FIG. 3.3 – Résidus standardisés de Churchill Falls sur papier de probabilité normal

- La variabilité des apports énergétiques après 1984 aurait diminué, la variance passant de 23 à 18 environ.
- L'intervalle de crédibilité à 90% de η contient la valeur 1, ce qui peut suggérer une rupture en variance de faible intensité. Toutefois, la densité a posteriori est très asymétrique (Figure 3.2). Il convient donc d'évaluer la probabilité que l'amplitude du changement soit inférieure à 1. A posteriori, celle-ci est de 0.88. A priori, elle est de 0.08. Cela indique que la prise en compte des observations vient appuyer l'événement $\{\eta < 1\}$.
- Le niveau des apports énergétiques annuels semble avoir diminué de 5 TWh en moyenne avec 2 TWh d'écart-type.
- L'intervalle de crédibilité à 90% de δ indique qu'une diminution de 6.5 TWh demeure crédible. De plus, cet intervalle ne contient pas la valeur 0, ce qui suggère une baisse d'au moins 4 TWh.

3.3.2 Application au complexe de l'Outaouais

Les apports énergétiques annuels du complexe de l'Outaouais ont été analysés en supposant une rupture en variance seulement (modèle (3.21)). L'algorithme de Gibbs a été mis en oeuvre avec les lois conditionnelles complètes de la Section 3.2.2. Après quelques expérimentations, le nombre de réplicats m et le temps de chauffe t ont été fixés à 1000. Les densités marginales a posteriori ainsi que leurs caractéristiques statistiques ont été estimées avec l'estimateur « Rao-Blackwellisé » (voir Annexe D). Le Tableau 3.4 présente les caractéristiques statistiques et les intervalles de crédibilité à 90%. La Figure 3.4 montre

la série d'apports énergétiques annuels ainsi que les densités marginales a posteriori des paramètres (traits continus).

TAB. 3.4 – *Caractéristiques a posteriori des paramètres pour le complexe de l'Outaouais*

Paramètres	Mode	Espérance	Ecart-type	Intervalle de crédibilité à 90%
τ	1979	1978	12.81	[1948 ; 1991]
σ_1^2	2.77	3.03	0.77	[1.82 ; 4.48]
σ_2^2	0.90	1.43	0.52	[0.53 ; 2.80]
η	0.30	0.58	0.20	[0.14 ; 1.04]

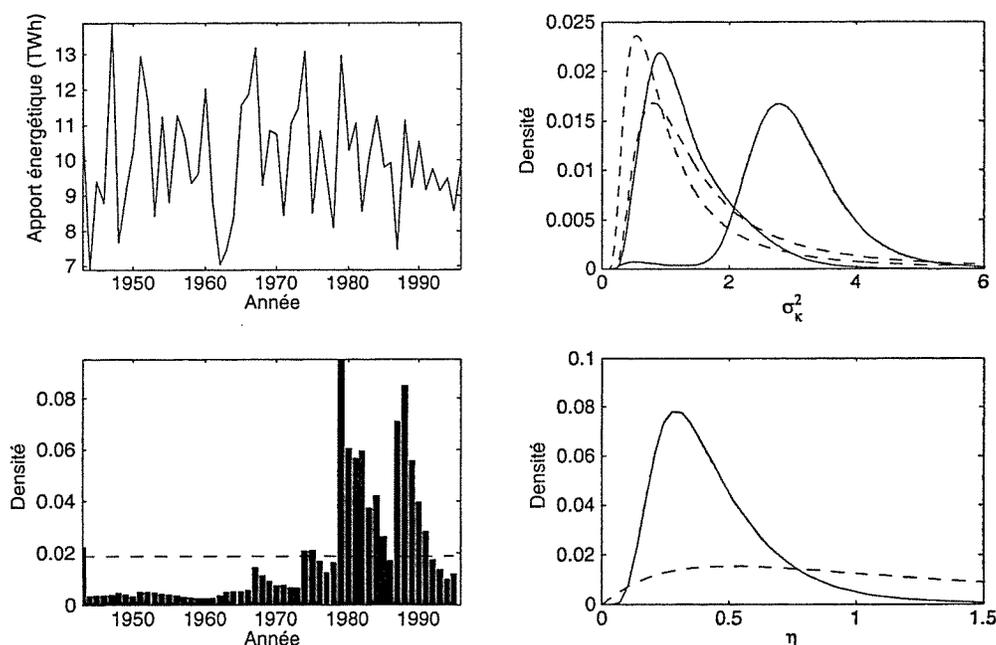


FIG. 3.4 – *Apports énergétiques annuels de l'Outaouais, densité marginale a priori (trait pointillé) et a posteriori (trait continu) des paramètres*

On tire principalement de ces résultats les conclusions suivantes :

- L'incertitude a diminué de manière assez importante en ce qui a trait à l'instant de rupture. La rupture serait survenue environ en 1979. La distribution a posteriori $p(\tau | y)$ demeure toutefois assez diffuse, révélant ainsi une forte incertitude sur cette estimation. L'écart-type est de 13 ans.
- L'ajout de l'information contenue dans les observations à celle apportée par l'analyse régionale a priori semble avoir augmenté l'incertitude sur les variances de part

et d'autre de l'instant de rupture. On note toutefois qu'en ce qui concerne l'amplitude de changement l'incertitude a diminué considérablement. Enfin, la variabilité des apports énergétiques a diminué environ de moitié après la rupture (le mode et l'espérance a posteriori de l'intensité du changement η sont respectivement 0.30 et 0.58).

- L'intervalle de crédibilité pour η contient la valeur 1, mais sa densité a posteriori est fortement asymétrique. La probabilité que l'intensité de rupture soit inférieure à l'unité est de 0.89.

Enfin, les résidus sur papier normal (Figure 3.5) sont distribués linéairement, ce qui vérifie l'hypothèse de normalité. L'autocorrélation d'ordre 1 est ici de -0.06.

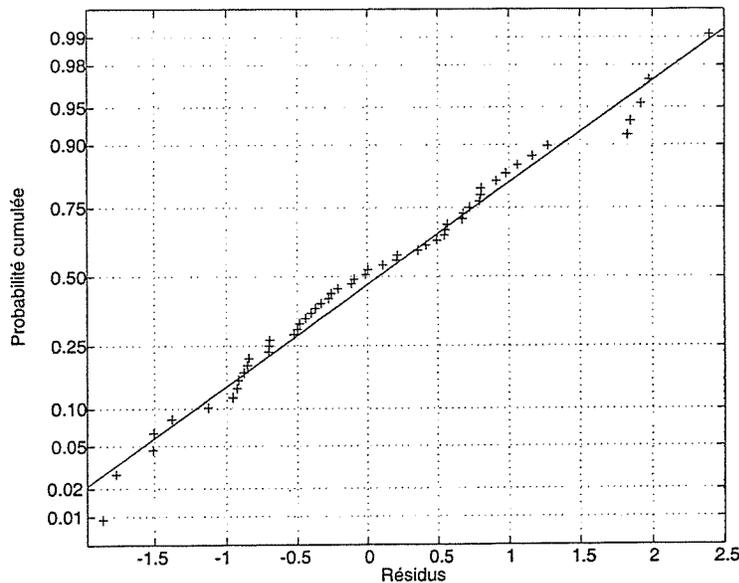


FIG. 3.5 – Résidus standardisés pour le complexe Outaouais sur papier de probabilité normal

3.3.3 Discussion

En raison de la baisse d'hydraulicité observée sur une partie du territoire (Section 3.3.1), Hydro-Québec doit effectuer une gestion rigoureuse de ses ressources. Elle doit également prendre différentes mesures afin d'assurer le maintien de l'équilibre entre l'offre et la demande en énergie pour les prochaines années. Pour accroître sa marge de manoeuvre, la société d'État peut devancer ou reporter les projets de suréquipement de centrales hydroélectriques et d'aménagement de nouveaux complexes, racheter des contrats d'exportation ou mettre en route des équipements de production thermique. La mise en

place de telles mesures peut s'avérer très onéreuse si celles-ci ne sont pas appliquées en temps opportun. C'est pourquoi il est important, du point de vue économique, de quantifier le mieux possible toutes les incertitudes relatives à l'évolution temporelle des apports énergétiques. Cette composante est en effet déterminante dans le choix des interventions à privilégier. Si effectivement une baisse importante des apports énergétiques moyens est survenue pour Churchill Falls et certains complexes de la même région (Manicouagan, par exemple), il est primordial de la caractériser et d'intégrer cette information dans le processus décisionnel. Les estimations issues des lois a posteriori présentées aux sections 3.3.1 et 3.3.2 fournissent les premiers éléments nécessaires à l'évaluation des incertitudes relatives à la marge de manoeuvre qu'Hydro-Québec possède en matière de disponibilité énergétique. Mais les changements observés dans les séries de Churchill Falls et de l'Ouataouais sont-ils réels ou seulement induits par la variabilité naturelle du phénomène? Le chapitre qui suit aborde cette question.

Chapitre 4

Mettre en doute l'existence et le type de la rupture

Le modèle général de rupture (3.1), ainsi que ses cas particuliers (3.4) et (3.21), supposent avec certitude la présence d'une rupture dans la période d'échantillonnage, puisque $\tau = 1, 2, \dots, n - 1$. L'inférence sur les paramètres de ces modèles est donc conditionnelle à l'existence d'un changement. De plus, elle ne porte que sur la structure spécifiée par le modèle. Aussi, l'hydrologue doit-il exclure l'hypothèse de stationnarité et tout autre type de rupture, ces alternatives n'étant pas formalisées. Certes essentielle, l'inférence conditionnelle sur les paramètres d'un modèle de rupture, telle que présentée au Chapitre 3, est incomplète. Et si les apports énergétiques du complexe de l'Outaouais étaient en réalité stationnaires? La rupture en moyenne dans les apports énergétiques annuels de Churchill Falls est-elle vraiment accompagnée d'un changement de variance? Comment confronter diverses alternatives de manière formelle lors de l'analyse d'une rupture dans la structure aléatoire des séries hydrologiques? Les méthodes employées actuellement en hydrologie statistique (par exemple celles de Lee et Heghinian (1977), de Hubert et al. (1989) et de Rao et Tirtotjondro (1996)) ne permettent pas de répondre formellement à ces questions. Ce chapitre propose une solution à ce type de problème. Pour ce faire, l'analyse d'une seule rupture est intégrée dans un cadre plus général, celui de la sélection bayésienne de modèles.

Le contenu de ce chapitre constitue une synthèse de l'Article II (Annexe B). Les principaux éléments conceptuels de la sélection bayésienne de modèles sont présentés dans la première section. On montre comment elle peut-être employée pour détecter un changement¹ et, si nécessaire, en identifier le type. À titre d'illustration, nous appliquons ensuite ces concepts à l'analyse des apports énergétiques annuels des complexes hydroélectriques de Churchill Falls et de l'Outaouais. Quatre modèles normaux sont mis en compétition: trois modèles de rupture et un modèle qui suppose la stationnarité. Enfin, la dernière section souligne les principaux avantages de la sélection bayésienne de modèles, notamment en ce qui a trait à la prise en compte de l'incertitude de modèles.

1. L'expression « détection d'une rupture » renvoie ici aux outils d'aide à la décision permettant de conclure à l'existence ou non d'un changement dans la série d'observations.

4.1 Détection d'une rupture et identification de son type

Considérons un ensemble de modèles possibles $\mathcal{M} = \{M_0, M_1, \dots, M_k, \dots, M_q\}$ pour les variables aléatoires $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$, où M_0 correspond au modèle représentant la stationnarité et où $M_1, \dots, M_k, \dots, M_q$ sont des modèles de rupture du type (3.1). La sélection bayésienne de modèles consiste d'abord à spécifier, pour chacune des alternatives, la probabilité a priori $p(M_k)$ que M_k soit « le vrai modèle »². En règle générale, les modèles sont considérés équiprobables a priori, c'est-à-dire que l'on prend $p(M_k) = 1/(q+1)$ pour $k = 0, \dots, q$. Dans le contexte d'une étude sur la non-stationnarité, particulièrement pour la gestion des ressources en eau, on préférera plutôt donner des poids équivalents aux hypothèses de changement et de non-changement. Si aucune connaissance a priori ne permet de favoriser l'un des modèles de rupture, un tel choix se traduira par :

$$p(M_0) = 1/2 \text{ et } p(M_1) = \dots = p(M_q) = 1/(2q) \quad (4.1)$$

Cette répartition des probabilités a priori est sans doute plus proche des degrés de crédibilité que le gestionnaire et l'hydrologue associeraient aux diverses alternatives. En effet, à moins de disposer d'information très convaincante, ce qui n'est généralement pas le cas en matière de changement dans les séries hydrologiques, ces derniers ne seront pas prêts à négliger le modèle stationnaire qu'ils utilisent depuis tant années.

Après avoir spécifié les probabilités a priori, on procède ensuite à la mise à jour de la crédibilité de chaque modèle en regard des observations $\mathbf{y} = (y_1, y_2, \dots, y_n)$. Elle s'obtient très naturellement par le théorème de Bayes :

$$p(M_k | \mathbf{y}) = \frac{p(\mathbf{y} | M_k) p(M_k)}{\sum_{j=0}^q p(\mathbf{y} | M_j) p(M_j)}, \quad k = 0, \dots, q \quad (4.2)$$

L'expression (4.2) fait intervenir la loi marginale des observations \mathbf{y} sous l'hypothèse du modèle M_k :

$$p(\mathbf{y} | M_k) = \int p(\mathbf{y} | \boldsymbol{\theta}_k, M_k) p(\boldsymbol{\theta}_k | M_k) d\boldsymbol{\theta}_k \quad (4.3)$$

Le but étant d'évaluer la vraisemblance du modèle M_k , les paramètres $\boldsymbol{\theta}_k$ ne sont que descriptifs ici. C'est pourquoi, dans (4.3), on les élimine par intégration afin de se concentrer sur le paramètre objectif M_k . De manière analogue à la fonction de vraisemblance maximisée employée en statistique classique, $p(\mathbf{y} | M_k)$ constitue la base permettant d'évaluer

2. Dans ce chapitre, puisqu'un ensemble de modèles est considéré, nous intégrons la variable indicatrice M_k dans les expressions mathématiques au même titre que des paramètres inconnus.

la compatibilité des données avec l'hypothèse a priori que les observations sont issues du modèle M_k . La proposition qui suit donne l'expression de la loi marginale des observations pour le modèle de rupture (3.1).

Proposition 5 *Considérons le modèle de rupture (3.1), noté M_k , et la loi jointe a priori de ses paramètres, $p(\theta_1, \theta_2, \tau | M_k)$. Pour une réalisation $\mathbf{y} = (y_1, y_2, \dots, y_n)$, si nous acceptons l'hypothèse H_1 , alors*

$$p(\mathbf{y} | M_k) = \sum_{\tau=1}^{n-1} p(\tau | M_k) p(\mathbf{y} | \tau, M_k).$$

Démonstration. Du fait de l'indépendance a priori entre l'instant de rupture et les autres paramètres, il vient :

$$\begin{aligned} p(\mathbf{y} | M_k) &= \sum_{\tau=1}^{n-1} \iint p(\mathbf{y} | \theta_1, \theta_2, \tau, M_k) p(\theta_1, \theta_2, \tau | M_k) d\theta_1 d\theta_2, \\ &= \sum_{\tau=1}^{n-1} p(\tau | M_k) \iint p(\mathbf{y} | \theta_1, \theta_2, \tau, M_k) p(\theta_1, \theta_2 | M_k) d\theta_1 d\theta_2, \end{aligned} \quad (4.4)$$

d'où le résultat. ■

Après avoir calculé les probabilités a posteriori (4.2), il est d'usage de sélectionner le modèle correspondant à la plus forte valeur. Toutefois, pour que (4.2) ait un sens, il faut naturellement que les probabilités a priori de chacun des modèles existent. Il est alors nécessaire de supposer que le vrai modèle appartient à l'ensemble \mathcal{M} . C'est le cadre \mathcal{M} -fermé défini par Bernardo et Smith (1994). Certes difficile à accepter au sens littéral, un modèle étant une représentation simplifiée de la réalité, cette hypothèse peut néanmoins être considérée raisonnable en hydrologie pour l'analyse d'une rupture. L'hydrologie statistique, telle qu'elle s'est développée au cours des dernières décennies, admet presque systématiquement que les séries hydrométéorologiques annuelles sont stationnaires. Le modèle M_0 a été utilisé, avec succès d'ailleurs, pour la représentation d'un large éventail de situations. Comme il a été mentionné au Chapitre 2, même si certains hydrologues mettent en doute la stationnarité des séries hydrologiques, cette hypothèse demeure aujourd'hui omniprésente pour la gestion des ressources en eau. En ce sens, la communauté hydrologique voit en M_0 une représentation assez fidèle de la réalité. Il est donc légitime, eu égard à l'état des connaissances a priori, de considérer ce modèle comme étant vrai, sinon très proche du comportement réel du système à l'étude. En conséquence, si un hydrologue soupçonne la présence d'un certain type de rupture dans une série hydrologique, et que, dans ce nouveau contexte il souhaite incorporer cette possibilité dans l'analyse par la sélection bayésienne de modèles, nous croyons pouvoir affirmer que la perspective \mathcal{M} -fermée est acceptable.

Partant de cette hypothèse, pour comparer deux modèles particuliers, M_j et M_k , il est naturel d'évaluer le ratio des cotes a posteriori sur les cotes a priori du modèle M_j contre le modèle M_k :

$$B_{jk}(\mathbf{y}) = \frac{\left[\frac{p(M_j | \mathbf{y})}{p(M_k | \mathbf{y})} \right]}{\left[\frac{p(M_j)}{p(M_k)} \right]} = \frac{p(\mathbf{y} | M_j)}{p(\mathbf{y} | M_k)}. \quad (4.5)$$

Cette mesure est appelée le facteur de Bayes (Kass et Raftery (1995)). Intuitivement, elle permet de vérifier si les observations ont augmenté ou diminué les cotes de M_j par rapport à M_k . Naturellement, $B_{jk} > 1$ signifie que le modèle M_j est maintenant plus vraisemblable que M_k au regard de \mathbf{y} ; $B_{jk} < 1$ indique que la crédibilité de M_j a diminué à la suite de la prise en compte des observations. Evidemment, le facteur de Bayes et les probabilités a posteriori des modèles jouent le même rôle : choisir un modèle sur la base du facteur de Bayes le plus grand revient à sélectionner le modèle qui possède la plus forte probabilité a posteriori. Signalons à ce propos que (4.2) peut être réécrite de manière plus intéressante sous la forme suivante :

$$p(M_k | \mathbf{y}) = \left[\sum_{j=0}^q \frac{p(M_j)}{p(M_k)} B_{jk}(\mathbf{y}) \right]^{-1}. \quad (4.6)$$

Le facteur de Bayes est particulièrement intéressant puisqu'il peut être aisément adapté pour confronter formellement différents types d'hypothèses. La proposition qui suit donne le facteur de Bayes composite qui compare un modèle particulier à l'ensemble de ses compétiteurs.

Proposition 6 *Considérons la collection de modèles $\mathcal{M} = \{M_0, M_1, \dots, M_k, \dots, M_q\}$ et leurs probabilité a priori $p(M_k)$, $k = 0, \dots, q$. Alors, le facteur de Bayes, noté $B_{j(-j)}(\mathbf{y})$, du modèle M_j contre l'ensemble des modèles $\{M_k : k = 0, \dots, q; k \neq j\}$, s'écrit de la manière suivante*

$$B_{j(-j)}(\mathbf{y}) = \left[\sum_{k \neq j} \left(\frac{p(M_k)}{1 - p(M_j)} \right) B_{kj}(\mathbf{y}) \right]^{-1}. \quad (4.7)$$

Démonstration. En vertu de (4.5) et de (4.2), on a

$$\begin{aligned} B_{j(-j)}(\mathbf{y}) &= \left(\frac{p(M_j | \mathbf{y})}{1 - p(M_j | \mathbf{y})} \right) \bigg/ \left(\frac{p(M_j)}{1 - p(M_j)} \right), \\ &= \left(\frac{1 - p(M_j)}{p(M_j)} \right) \frac{p(\mathbf{y} | M_j) p(M_j)}{\sum_{k \neq j} p(\mathbf{y} | M_k) p(M_k)}, \\ &= \left(\frac{1 - p(M_j)}{p(M_j)} \right) \left[\sum_{k \neq j} \frac{p(\mathbf{y} | M_k) p(M_k)}{p(\mathbf{y} | M_j) p(M_j)} \right]^{-1}, \end{aligned}$$

d'où le résultat. ■

Cette proposition peut être appliquée à ce qui nous intéresse plus particulièrement, la détection d'une rupture ainsi que l'identification de son type. On déduit d'abord de (4.7) que le facteur de Bayes $B_{c,nc}(\mathbf{y})$ permettant d'opposer les hypothèses de changement et de non changement est donné, en observant que $B_{jk} = B_{kj}^{-1}$, par l'expression

$$B_{c,nc}(\mathbf{y}) = \sum_{j=1}^q \left(\frac{p(M_j)}{1 - p(M_0)} \right) B_{j0}(\mathbf{y}). \quad (4.8)$$

Il s'agit d'une moyenne pondérée des facteurs de Bayes évaluant la crédibilité spécifique de chaque configuration de changement contre l'hypothèse de stationnarité M_0 . Si l'on considère la répartition des probabilités a priori donnée en (4.1), $B_{c,nc}(\mathbf{y})$ se présente alors comme une simple moyenne arithmétique. Pour comparer ensuite la crédibilité des configurations de changement entre elles, il suffit d'évaluer le facteur de Bayes composite (4.7) pour chaque alternative M_1, \dots, M_q , en excluant le modèle M_0 . On aura ainsi une gamme de mesures permettant de déterminer, sous l'hypothèse d'un changement, quel type de rupture dans \mathcal{M} est susceptible de mieux représenter les observations.

Jusqu'ici, nous nous sommes donné les outils permettant de comparer la vraisemblance de diverses hypothèses, mais sans pour autant pouvoir formellement arrêter notre choix. À partir de quelle valeur le facteur de Bayes (4.8) peut-il être considéré assez fort pour dissuader l'hydrologue et le gestionnaire de supposer la stationnarité d'une série d'observations hydrologiques? Cette question étant essentiellement un problème décisionnel, pour y répondre, il faut faire intervenir les conséquences associées au choix de chacune des alternatives. On suppose ici qu'on les prend sous la forme de coûts forfaitaires tels que ceux donnés dans le Tableau 4.1 : coûts nuls dans le cas d'une bonne décision, coûts constants positifs $l_{nc,c}$ et $l_{c,nc}$ ³ autrement.

TAB. 4.1 – Pertes encourues pour tester l'existence d'une rupture

	Hypothèses	
Décisions	Changement	Non changement
Changement	0	$l_{c,nc}$
Non changement	$l_{nc,c}$	0

La règle de décision bayésienne pour la détection d'une rupture consiste alors, avec les probabilités a priori (4.1), à rejeter la stationnarité si et seulement si

$$B_{c,nc}(\mathbf{y}) > \frac{l_{c,nc}}{l_{nc,c}}. \quad (4.9)$$

Souvent, une telle fonction de coût est difficile à établir, voire indisponible. En faisant le parallèle entre le facteur de Bayes et la statistique du rapport des vraisemblances

3. $l_{nc,c}$ désigne la perte encourue lorsqu'on conclut qu'il n'y a pas de changement alors qu'en fait une rupture est présente dans la série.

maximales, Kass et Raftery (1995) ont proposé une table pour l'interprétation du facteur de Bayes. Elle est reproduite au Tableau 4.2.

TAB. 4.2 – *Interprétation du facteur de Bayes (tiré de Kass et Raftery (1995))*

$2 \ln B_{jk}$	B_{jk}	Evidence en faveur de M_j
0 – 2	1 – 3	Ne vaut pas plus qu'une simple mention
2 – 6	3 – 20	Positive
6 – 10	20 – 150	Forte
> 10	> 150	Très forte

Remarque : *Cet outil ne constitue pas une règle de décision formelle. Ne disposant pas d'une fonction de coût réelle pour l'analyse des apports énergétiques présentés dans la section qui suit, nous utilisons tout de même la table de Kass et Raftery à titre indicatif seulement pour illustrer notre approche. Une analyse complète et réaliste nécessiterait la spécification des pertes $l_{c,nc}$ et $l_{nc,c}$ dans (4.9) à partir de considérations décisionnelles propres au problème de planification énergétique.*

Enfin, il importe de souligner que le facteur de Bayes nécessite l'emploi de lois a priori informatives pour les paramètres de chacun des modèles. En effet, si ces distributions sont non informatives, et donc impropres, le rapport des cotes a posteriori est indéterminé (voir Article III, Annexe E, pour une brève explication). Dans ce cas, il faut avoir recours à des facteurs de Bayes « modifiés », par exemple ceux de O'Hagan (1995) ou de Berger et Pericchi (1996). Ce problème ne se pose pas dans cette thèse puisque seules des lois a priori informatives sont considérées (hypothèse \mathbf{H}_2 , Section 3.1.2).

4.2 Application aux séries d'apports énergétiques annuels de Churchill Falls et de l'Outaouais

Nous reprenons l'étude des apports énergétiques annuels des complexes hydroélectriques de Churchill Falls et de l'Outaouais présentés respectivement aux sections 3.3.1 et 3.3.2. Nous nous intéressons ici au problème de détection d'une rupture et à celui de l'identification de son type. Nous considérons que les apports énergétiques annuels calculés à ces sites, Y_1, \dots, Y_n , sont issus de l'un des quatre modèles de l'ensemble $\mathcal{M} = \{M_0, M_1, M_2, M_3\}$, où

$$\begin{aligned}
\text{Modèle } M_0: & Y_i \sim \mathcal{N}(y_i | \mu_1, \sigma_1^2), \quad i = 1, \dots, n \\
\text{Modèle } M_1: & Y_i \sim \mathcal{N}(y_i | \mu_1, \sigma_1^2), \quad i = 1, \dots, \tau \\
& Y_i \sim \mathcal{N}(y_i | \mu_2, \sigma_1^2), \quad i = \tau + 1, \dots, n \\
\text{Modèle } M_2: & Y_i \sim \mathcal{N}(y_i | \mu_1, \sigma_1^2), \quad i = 1, \dots, \tau \\
& Y_i \sim \mathcal{N}(y_i | \mu_1, \sigma_2^2), \quad i = \tau + 1, \dots, n \\
\text{Modèle } M_3: & Y_i \sim \mathcal{N}(y_i | \mu_1, \sigma_1^2), \quad i = 1, \dots, \tau \\
& Y_i \sim \mathcal{N}(y_i | \mu_2, \sigma_2^2), \quad i = \tau + 1, \dots, n
\end{aligned}$$

Ainsi, trois configurations de changement ont été retenues pour cette application : la rupture en moyenne seulement (M_1), la rupture en variance seulement (M_2) et la rupture simultanée en moyenne et en variance (M_3). Remarquons qu'il y a correspondance entre les paramètres des modèles puisque ceux-ci sont imbriqués.

Nous avons adopté la répartition des probabilités a priori (4.1), c'est-à-dire celle qui donne des poids équivalents aux hypothèses de changement et de non changement :

$$p(M_0) = 1/2 \quad \text{et} \quad p(M_1) = p(M_2) = p(M_3) = 1/6 .$$

Pour les paramètres de chacun des modèles, nous faisons encore ici l'hypothèse a priori \mathbf{H}_1 et nous considérons une loi jointe a priori conjuguée. Plus précisément, on a :

$$\begin{aligned}
M_0: & p(\boldsymbol{\theta}_0) = \mathcal{N}(\mu_1 | \phi_1, \lambda_1 \sigma_1^2) \mathcal{IG}(\sigma_1^2 | \alpha_1, \beta_1), \\
M_1: & p(\boldsymbol{\theta}_1) = \mathcal{N}(\mu_1 | \phi_1, \lambda_1 \sigma_1^2) \mathcal{N}(\mu_2 | \phi_2, \lambda_2 \sigma_1^2) \mathcal{IG}(\sigma_1^2 | \alpha_1, \beta_1) p(\tau), \\
M_2: & p(\boldsymbol{\theta}_2) = \mathcal{N}(\mu_1 | \phi_1, \lambda_1 \sigma_1^2) \mathcal{IG}(\sigma_1^2 | \alpha_1, \beta_1) \mathcal{IG}(\sigma_2^2 | \alpha_2, \beta_2) p(\tau), \\
M_3: & p(\boldsymbol{\theta}_3) = \mathcal{N}(\mu_1 | \phi_1, \lambda_1 \sigma_1^2) \mathcal{IG}(\sigma_1^2 | \alpha_1, \beta_1) \mathcal{N}(\mu_2 | \phi_2, \lambda_2 \sigma_2^2) \mathcal{IG}(\sigma_2^2 | \alpha_2, \beta_2) p(\tau),
\end{aligned}$$

où $\boldsymbol{\theta}_k$ désigne l'ensemble des paramètres du modèle M_k . Les hyperparamètres de chacun des modèles ont été spécifiés à partir de l'analyse régionale décrite dans l'Article I (Annexe A). Nous rappelons que le Tableau 3.2 donne les espérances et les variances a priori des paramètres pour le modèle M_3 (complexe de Churchill Falls) et le modèle M_2 (complexe de l'Outaouais).

Pour mettre en oeuvre la sélection bayésienne de modèles telle que décrite dans la Section 4.1, on doit évaluer chacune de lois marginales $p(\mathbf{y} | M_k)$, $k = 0, \dots, 3$. Pour les modèles M_0 , M_1 et M_3 , on obtient des expressions explicites. Celles-ci sont données en annexe de l'Article II. On a en particulier, en vertu de la Proposition 5 et de l'équation (3.11) :

$$p(\mathbf{y} | M_3) = \left(\frac{1}{2\pi} \right)^{n/2} \frac{\beta_1^{\alpha_1} \beta_2^{\alpha_2}}{\sqrt{\lambda_1 \lambda_2} \Gamma(\alpha_1) \Gamma(\alpha_2)} \sum_{\tau=1}^{n-1} p(\tau | M_3) (\lambda_1' \lambda_2')^{1/2} \frac{\Gamma(\alpha_1') \Gamma(\alpha_2')}{(\beta_1')^{\alpha_1'} (\beta_2')^{\alpha_2'}}. \quad (4.10)$$

TAB. 4.3 – *Facteurs de Bayes et probabilités a posteriori pour les complexes de Churchill Falls et de l'Outaouais*

	B_{jk}	M_0	M_1	M_2	M_3	$p(M_k \mathbf{y})$
Churchill Falls	M_0	-	0.17	0.78	0.09	0.1424
	M_1	6.03	-	4.73	0.56	0.2862
	M_2	1.27	0.21	-	0.12	0.0605
	M_3	10.76	1.78	8.44	-	0.5108
	B_{jk}	M_0	M_1	M_2	M_3	$p(M_k \mathbf{y})$
Outaouais	M_0	-	3.03	0.26	0.75	0.3554
	M_1	0.33	-	0.09	0.25	0.0391
	M_2	3.78	11.46	-	2.85	0.4481
	M_3	1.33	4.02	0.35	-	0.1573

TAB. 4.4 – *Facteurs de Bayes pour la détection d'une rupture et l'identification de son type pour les complexes de Churchill Falls et de l'Outaouais*

Complexe	$B_{c,nc}(\mathbf{y})$	$B_{1(2,3)}(\mathbf{y})$	$B_{2(1,3)}(\mathbf{y})$	$B_{3(1,2)}(\mathbf{y})$
Churchill Falls	6.02	1.00	0.16	2.95
Outaouais	1.81	0.13	4.56	0.65

* $B_{i(j,k)}(\mathbf{y})$ est le facteur de Bayes de M_i contre M_j et M_k .

Les facteurs de Bayes B_{10} , B_{30} et B_{31} peuvent alors être calculés analytiquement à partir de (4.5). Toutefois, les intégrales apparaissant dans (4.4) ne peuvent être résolues analytiquement pour déterminer $p(\mathbf{y} | M_2)$. Aussi, B_{20} , B_{21} et B_{32} ne peuvent-ils être évalués explicitement. Pour résoudre ce problème, nous avons eu recours à la méthode proposée par Chib (1995), qui consiste à approcher la densité marginale des observations, $p(\mathbf{y} | M_2)$, par échantillonnage de Gibbs. La procédure astucieuse de ce dernier, qui réemploie les échantillons de paramètres générés lors de la phase d'estimation, est décrite en détails dans l'Article II. Nous conservons pour la sélection de modèles le même nombre de réplicats m et le même temps de chauffe t ($m = t = 1000$) que pour l'estimation des paramètres de M_2 effectuée à la Section 3.3.2.

Les facteurs de Bayes permettant de comparer deux à deux la crédibilité des modèles pour les complexes de Churchill Falls et de l'Outaouais sont présentés au Tableau 4.3. Les probabilités a posteriori, évaluées à partir de l'équation (4.6), y sont aussi données pour chacun des modèles. De plus, le Tableau 4.4 fournit les facteurs de Bayes composites calculés pour la détection de rupture et l'identification de son type.

L'analyse de ces tableaux nous permet d'observer, en premier lieu, que la crédibilité de l'hypothèse de stationnarité (modèle M_0) a diminué considérablement pour les apports énergétiques du complexe de Churchill Falls. Elle passe d'une chance sur deux à moins d'une chance sur sept (voir Tableau 4.3). La vraisemblance du modèle M_0 a aussi diminué au regard des observations, mais dans une moindre mesure. On note ensuite, si l'échelle de Kass et Raftery (Tableau 4.2) constitue une règle de décision acceptable, que le facteur

de Bayes comparant les hypothèses de changement et de non changement (voir Tableau 4.4) indique la présence d'une rupture dans les apports énergétiques de Churchill Falls ($B_{c,nc}(\mathbf{y}) > 3$). Par contre, pour le complexe de l'Outaouais, la valeur de $B_{c,nc}(\mathbf{y})$ nous permet difficilement de conclure à l'existence d'un changement ($B_{c,nc}(\mathbf{y}) < 3$). Rappelons qu'en théorie de la décision statistique, le choix d'accepter ou de rejeter l'hypothèse de stationnarité doit faire intervenir une fonction de coût. Aussi, pour Churchill Falls, notre conclusion n'est-elle valide que si $l_{c,nc}/l_{nc,c} < 6.02$, c'est-à-dire si les pertes associées au rejet à tort de l'hypothèse de stationnarité sont au plus 6 fois supérieures à celles encourues par le choix de l'hypothèse de stationnarité lorsque la série exhibe une rupture.

A la lumière des résultats présentés, on peut mettre en doute la validité de l'hypothèse de stationnarité pour les apports énergétiques annuels de Churchill Falls et donc le bien fondé de l'approche que privilégie Hydro-Québec en matière de planification énergétique. Rappelons que les moyennes et les variances des apports énergétiques, utilisées comme intrants dans des modèles de planification, sont évaluées sur toute la période d'échantillonnage. Il serait alors souhaitable de considérer les moyennes et les variances estimées à partir d'un modèle de rupture et d'en évaluer l'impact sur les décisions. On utilisera les estimations données au Tableau 3.3, car l'analyse indique qu'une rupture simultanée en moyenne et en variance (modèle M_3) constitue la meilleure représentation. Ce modèle possède en effet la probabilité a posteriori la plus forte (0.51) parmi les trois alternatives (Tableau 4.4). Remarquons aussi que le facteur de Bayes composite, $B_{3(1,2)}$, indique que le modèle de rupture M_3 est près de trois fois plus crédible que l'un ou l'autre de ses compétiteurs.

Il est maintenant nécessaire d'aborder la question de la sensibilité du facteur de Bayes à la spécification des lois a priori. En effet, alors que l'estimation des paramètres d'un modèle y est assez robuste, il est bien connu que le facteur de Bayes est sensible à la détermination de ces distributions (Kass et Raftery (1995)). Cette caractéristique n'est pas un handicap en soi. Elle révèle seulement que les connaissances et les informations issues de sources extérieures aux données contribuent au processus inférentiel. C'est justement l'objet de la statistique bayésienne de prendre en compte ce type d'information de manière explicite. Evidemment, pour obtenir des résultats adéquats, les lois a priori doivent bien traduire l'état des connaissances. Or, on peut toujours douter de la représentativité de ces distributions. Cela peut être le cas si les intervenants ont des opinions divergentes lorsque la spécification des lois a priori repose sur des connaissances subjectives. Pour notre problème, on peut émettre des doutes quant à l'adéquation du modèle de régression régionale décrite dans l'Article I (Annexe A). Il est donc important d'analyser la sensibilité de nos conclusions en faisant varier certaines caractéristiques des lois a priori. Il est notamment intéressant, à ce propos, d'étudier l'évolution des facteurs de Bayes $B_{10}(\mathbf{y})$, $B_{30}(\mathbf{y})$ et $B_{c,nc}(\mathbf{y})$ en fonction de la précision accordée a priori à l'espérance des apports énergétiques après la rupture.

Nous avons donc évalué ces facteurs de Bayes pour le complexe de Churchill Falls en faisant varier l'hyperparamètre λ_2 autour de la valeur obtenue à partir de considérations régionales, soit 0.35. La Figure 4.1 montre leur évolution pour $\lambda_2 \in [0.01, 2.00]$. Cette gamme de valeurs correspond à un ensemble de lois a priori relativement large et com-

patible avec la crédibilité que l'on accorde à la moyenne des apports énergétiques après la rupture. Evidemment, une grande valeur de λ_2 conduit à une loi a priori plus diffuse, accordant moins de crédibilité à l'espérance a priori issue du modèle régional. Le facteur de Bayes $B_{20}(\mathbf{y})$ n'est pas influencé par la valeur de λ_2 , puisque dans M_2 la moyenne reste invariante de part et d'autre de τ . Nous l'avons tout de même ajouté au graphique, puisqu'il intervient dans le calcul de $B_{c,nc}(\mathbf{y})$ et qu'il est obtenu par simulation. Cela nous permet de constater que les variations induites par l'approximation de Chib (1995), avec $m = t = 1000$, ne sont pas assez fortes pour avoir un impact sur les conclusions.

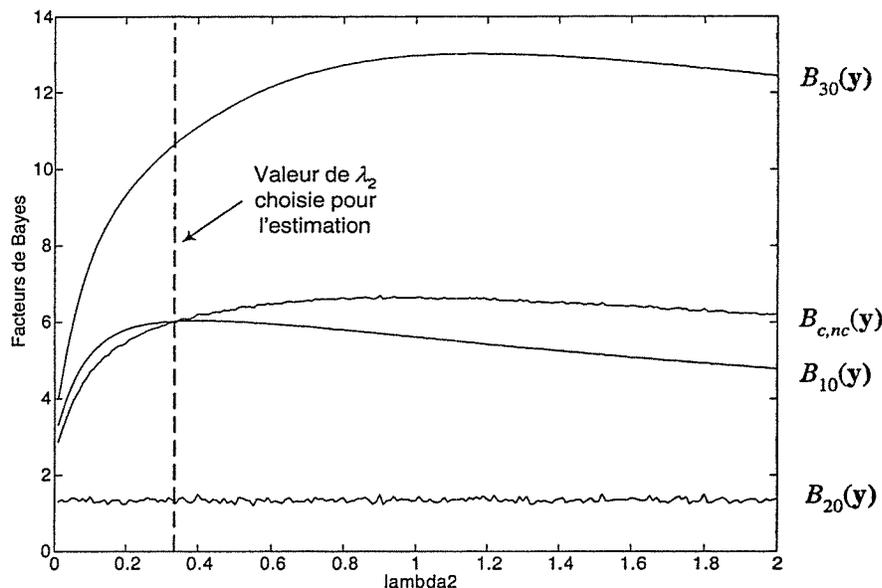


FIG. 4.1 – Sensibilité du facteur de Bayes du changement contre le non changement aux variances a priori des apports énergétiques moyens

On retient principalement de la Figure 4.1 que la crédibilité de l'hypothèse de changement croît à mesure que l'incertitude a priori sur la moyenne après la rupture augmente. Le facteur de Bayes $B_{c,nc}(\mathbf{y})$ varie entre 3 et 6.5 environ dans la gamme de lois a priori considérée. En vertu de la table de Kass et Raftery, la rupture demeure donc l'hypothèse à favoriser.

4.3 Discussion

A notre connaissance, la sélection bayésienne de modèles n'est pas employée en hydrologie statistique. Or, dans ce domaine, le problème du choix d'un modèle se pose

fréquemment : choix d'une loi de probabilité pour l'analyse des débits extrêmes, détermination de la meilleure représentation d'une série temporelles de débits, validation des apports naturels, etc. En terminant ce chapitre, nous discutons de certaines propriétés que possède cette procédure.

- En premier lieu, les résultats d'une telle analyse sont très simples à interpréter, même pour des non initiés. En effet, ceux-ci sont assez familiers avec la notion de probabilité pour pouvoir facilement interpréter une conclusion telle que « les probabilités que les modèles M_1 et M_2 soient vrais sont respectivement 0.90 et 0.10 ».
- Deuxièmement, la sélection bayésienne de modèles a l'avantage d'être « convergente » au sens où lorsque la taille d'échantillon augmente indéfiniment, on est assuré de choisir le vrai modèle, ou du moins le plus proche (Bernardo et Smith (1994)).
- Puisqu'on spécifie une loi de probabilité a priori pour les paramètres de chaque modèle, la sélection bayésienne de modèles prend automatiquement en compte le critère de parcimonie : si deux modèles sont aussi compatibles avec les données l'un que l'autre, la procédure identifiera celui des deux ayant le moins de paramètres comme meilleure représentation des observations.
- La sélection bayésienne a aussi l'avantage de pouvoir être appliquée à un grand nombre de modèles ; il n'est pas nécessaire qu'ils soient imbriqués, ni qu'ils appartiennent à la même classe.
- Enfin, cette perspective permet de prendre en compte l'incertitude de modélisation. Puisqu'à chaque alternative est associée une probabilité a posteriori, $p(M_k | \mathbf{y})$, il est possible de maintenir chaque modèle à l'intérieur du processus inférentiel concernant une quantité inconnue commune à l'ensemble des alternatives. Dans l'Article III, nous avons employée cette propriété pour prédire une nouvelle observation \tilde{y}_{n+1} en considérant l'ensemble des modèles $\mathcal{M} = \{M_0, M_1, M_2, M_3\}$. La densité prédictive correspond alors au mélange des densités prédictives $p(\tilde{y}_{n+1} | \mathbf{y}, M_k)$ de chaque modèle, affectées du poids $p(M_k | \mathbf{y})$:

$$p(\tilde{y}_{n+1} | \mathbf{y}) = \sum_{k=0}^3 p(\tilde{y}_{n+1} | \mathbf{y}, M_k) p(M_k | \mathbf{y}) \quad (4.11)$$

Cette expression montre notamment que choisir un modèle et ensuite prédire une nouvelle observation conditionnellement à celui-ci est raisonnable si l'une des probabilités a posteriori $p(M_k | \mathbf{y})$ est voisine de un. Sinon, la prévision conditionnelle ne prend pas en compte l'incertitude de modèle et sa précision risque d'être sous-estimée.

Chapitre 5

Et s'il y a changement climatique, n'y aurait-il pas impact régional?

Aux chapitres 3 et 4, nous avons proposé des solutions bayésiennes pour détecter et caractériser une rupture à un instant inconnu dans une seule séquence de variables aléatoires hydrologiques. Notre propos portera maintenant sur l'analyse d'un changement dans plusieurs séries d'observations hydrologiques à des sites soumis aux mêmes phénomènes climatiques.

On a déjà évoqué qu'un changement climatique est un phénomène global se manifestant à une échelle au moins régionale. Si les sites d'intérêt forment un ensemble hydrologiquement homogène (GREHYS (1996)), on peut alors présumer qu'il existe une certaine cohérence régionale entre les caractéristiques des ruptures présentes dans chaque série, notamment entre les instants où surviennent ces changements. D'ailleurs, il est d'usage en pratique de tirer des conclusions globales en analysant indépendamment la non-stationnarité de chacune des séries (Hubert et al. (1989), Slivitzky et Mathier (1994), Servat et al. (1997), Paturel et al. (1997), Perreault et al. (1999), Kiely (1999)). Ce type de procédures nous semble contestable. D'une part, les relations structurelles entre les caractéristiques des changements à chaque site n'y sont pas formalisées. D'autre part, les corrélations spatiales existant entre les observations ne sont pas prises en compte. Il est important de mobiliser ces informations régionales en posant un modèle qui inclut ces particularités du phénomène de rupture.

Soit l'ensemble de variables aléatoires hydrologiques suivant :

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \dots & Y_{1\tau_1} & Y_{1(\tau_1+1)} & \dots & Y_{1n} \\ Y_{21} & Y_{22} & \dots & Y_{2\tau_2} & Y_{2(\tau_2+1)} & \dots & Y_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \dots & \vdots \\ Y_{p1} & Y_{p2} & \dots & Y_{p\tau_p} & Y_{p(\tau_p+1)} & \dots & Y_{pn} \end{pmatrix}, \quad (5.1)$$

où chacune des lignes représente une chronique, par exemple les débits annuels à un site donné (il y a donc p sites et n années). Supposons qu'un changement climatique ait engendré une rupture dans chaque série à un instant inconnu τ_i , $i = 1, \dots, p$.

L'objet de ce chapitre est donc de construire des modèles multisite de rupture traduisant les manifestations régionales d'un changement climatique sur cet ensemble de données hydrologiques et de développer les outils d'estimation. Nous supposons que les variables aléatoires dans \mathbf{Y} sont normales et que seul un changement de moyenne survient dans chacune des séquences. La variance des observations ainsi que les covariances sont donc invariantes de part et d'autre des τ_i . Dans un premier temps, nous traitons du cas où la rupture survient simultanément pour l'ensemble des p sites. Puisqu'il peut exister un décalage spatial dans les phénomènes hydrologiques, nous généralisons ensuite la procédure au cas où l'instant de changement est variable d'une série à l'autre. En effet, même si les stations forment un ensemble homogène, rien ne garantit que les facteurs climatiques se manifestent exactement de la même manière dans toute la région. Deux modèles hiérarchiques sont proposés pour représenter cette situation.

5.1 Modèle multisite à rupture simultanée (MM₁)

Dans cette section, nous supposons que le changement climatique s'est manifesté exactement de la même manière pour chacune des p séquences. Pour ce faire, nous considérons la relation structurelle la plus simple entre les dates de rupture : celle qui admet que $\tau_1 = \tau_2 = \dots = \tau_p = \tau$. Ce premier modèle multisite a fait l'objet d'une publication présentée à l'Annexe E (Article III).

5.1.1 Modèle et résultats techniques

Pour $\tau_1 = \tau_2 = \dots = \tau_p = \tau$, l'ensemble de variables aléatoires (5.1) devient :

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \dots & Y_{1\tau} & Y_{1(\tau+1)} & \dots & Y_{1n} \\ Y_{21} & Y_{22} & \dots & Y_{2\tau} & Y_{2(\tau+1)} & \dots & Y_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \dots & \vdots \\ Y_{p1} & Y_{p2} & \dots & Y_{p\tau} & Y_{p(\tau+1)} & \dots & Y_{pn} \end{pmatrix} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_\tau, \dots, \mathbf{Y}_n)$$

où $\mathbf{Y}_1, \dots, \mathbf{Y}_\tau, \dots, \mathbf{Y}_n$ sont des vecteurs-colonnes aléatoires de dimension p observés à n périodes consécutives. Nous considérons comme modèle de rupture l'extension naturelle du modèle M_1 au cas multidimensionnel :

$$\begin{aligned} \mathbf{Y}_i &\sim \mathcal{N}_p(\mathbf{y}_i \mid \boldsymbol{\mu}_1, \mathbf{P}), & i = 1, \dots, \tau, \\ \mathbf{Y}_i &\sim \mathcal{N}_p(\mathbf{y}_i \mid \boldsymbol{\mu}_2, \mathbf{P}), & i = \tau + 1, \dots, n, \end{aligned} \quad (5.2)$$

où $\boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, $\tau \in \{1, \dots, n-1\}$ est l'instant de rupture, \mathbf{y}_i correspond à une réalisation de \mathbf{Y}_i et $\mathcal{N}_p(\mathbf{y} \mid \boldsymbol{\mu}, \mathbf{P})$ désigne la densité d'une loi normale multidimensionnelle. Le paramètre \mathbf{P} correspond à la matrice de précision (inverse de la matrice des covariances, $\mathbf{P} = \boldsymbol{\Sigma}^{-1}$) et représente la structure invariante de la loi de part et d'autre de τ . Si \mathbf{P} est diagonale,

on a indépendance spatiale entre les séries. Ce modèle, pour lequel le changement survient à la même position dans chacune des séries, est noté MM_1 .

A partir de (5.2), on déduit la fonction de vraisemblance résultant d'un échantillon de n observations $\mathbf{D} = (y_1, y_2, \dots, y_n)$:

$$\begin{aligned} p(\mathbf{D} \mid \mu_1, \mu_2, \tau, \mathbf{P}) &= \prod_{i=1}^{\tau} \mathcal{N}_p(y_i \mid \mu_1, \mathbf{P}) \prod_{i=\tau+1}^n \mathcal{N}_p(y_i \mid \mu_2, \mathbf{P}), \\ &= (2\pi)^{-np/2} |\mathbf{P}|^{n/2} \exp \left\{ -\frac{\tau}{2} \text{tr} \left[\mathbf{P} \left(\mathbf{S}_{\tau} + (\bar{y}_{\tau} - \mu_1) (\bar{y}_{\tau} - \mu_1)^t \right) \right] \right\} \\ &\quad \times \exp \left\{ -\frac{n-\tau}{2} \text{tr} \left[\mathbf{P} \left(\mathbf{S}_{n-\tau} + (\bar{y}_{n-\tau} - \mu_2) (\bar{y}_{n-\tau} - \mu_2)^t \right) \right] \right\}, \end{aligned} \quad (5.3)$$

où $\text{tr}(\cdot)$ est l'opérateur trace et

$$\begin{aligned} \bar{y}_{\tau} &= \frac{1}{\tau} \sum_{i=1}^{\tau} y_i, & \bar{y}_{n-\tau} &= \frac{1}{n-\tau} \sum_{i=\tau+1}^n y_i, \\ \mathbf{S}_{\tau} &= \frac{1}{\tau} \sum_{i=1}^{\tau} (y_i - \bar{y}_{\tau}) (y_i - \bar{y}_{\tau})^t, & \mathbf{S}_{n-\tau} &= \frac{1}{n-\tau} \sum_{i=\tau+1}^n (y_i - \bar{y}_{n-\tau}) (y_i - \bar{y}_{n-\tau})^t. \end{aligned}$$

Le modèle de rupture multidimensionnel MM_1 a déjà été traité par Booth et Smith (1982), mais en ne considérant que des lois a priori non informatives. En conformité avec l'hypothèse \mathbf{H}_2 , nous adoptons des lois informatives. Pour représenter la connaissance a priori sur les paramètres μ_1 , μ_2 et \mathbf{P} , nous considérons, comme aux chapitres précédents, une loi conjuguée étant donné τ . Nous admettons aussi l'hypothèse d'indépendance a priori entre l'instant de rupture et les autres paramètres (\mathbf{H}_1). Ainsi,

$$\begin{aligned} p(\mu_1, \mu_2, \mathbf{P}, \tau) &= p(\mu_1 \mid \mathbf{P}) p(\mu_2 \mid \mathbf{P}) p(\mathbf{P}) p(\tau), \\ &= \mathcal{N}_p(\mu_1 \mid \phi_1, \lambda_1 \mathbf{P}) \mathcal{N}_p(\mu_2 \mid \phi_2, \lambda_2 \mathbf{P}) \mathcal{W}_p(\mathbf{P} \mid a, \mathbf{B}) p(\tau), \\ &= \mathcal{N} \mathcal{N} \mathcal{W}_p(\mu_1, \mu_2, \mathbf{P} \mid \phi_1, \lambda_1, \phi_2, \lambda_2, a, \mathbf{B}) p(\tau), \end{aligned} \quad (5.4)$$

où $\mathcal{W}_p(\mathbf{P} \mid a, \mathbf{B})$ est la densité de la loi Wishart de dimension p et de paramètres a et \mathbf{B} . Cette distribution correspond à la généralisation de la loi gamma au cas multidimensionnel. Le produit de deux lois normales multidimensionnelles avec une loi de Wishart comme distribution a priori des paramètres μ_1 , μ_2 et \mathbf{P} , sachant τ , provient du fait que (5.3) a effectivement cette structure. Les hyperparamètres de (5.4) permettent de déterminer la forme de la loi a priori. Par exemple, la précision accordée aux connaissances a priori sur les paramètres μ_1 , μ_2 et \mathbf{P} peut être ajustée à partir de λ_1 , λ_2 et a . De faibles valeurs de λ_1 , λ_2 et a rendront les distributions a priori plus diffuses traduisant ainsi une connaissance a priori limitée sur ces paramètres. A la limite, si $\lambda_1 = \lambda_2 \rightarrow 0$ et $a \rightarrow 0$ dans (5.4), on obtiendra la loi a priori complètement non informative $p(\mu_1, \mu_2, \mathbf{P}) \propto |\mathbf{P}|^{-p/2}$. C'est cette dernière que Booth et Smith (1982) ont considérée.

Les propositions qui suivent donnent les expressions des fonctions de densité de probabilité des paramètres du modèle MM_1 . Si $p = 1$, on retrouve les expressions correspondantes pour le modèle M_1 présentées en annexe de l'Article I.

Proposition 7 *Soit le modèle de rupture (5.2) et la loi jointe a priori (5.4). Alors,*

$$p(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{P} \mid \tau, \mathbf{D}) = \mathcal{N}\mathcal{N}\mathcal{W}_p(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{P} \mid \phi'_1, \lambda'_1, \phi'_2, \lambda'_2, a', \mathbf{B}') \quad (5.5)$$

et

$$p(\tau \mid \mathbf{D}) \propto p(\tau) (\lambda'_1 \lambda'_2)^{-p/2} |\mathbf{B}'|^{-a'}, \quad (5.6)$$

où

$$\begin{aligned} \lambda'_1 &= \lambda_1 + \tau, & \lambda'_2 &= \lambda_2 + n - \tau \\ \phi'_1 &= \left(1 - \frac{\tau}{\lambda'_1}\right) \phi_1 + \frac{\tau}{\lambda'_1} \bar{\mathbf{y}}_\tau, & \phi'_2 &= \left(1 - \frac{n - \tau}{\lambda'_2}\right) \phi_2 + \frac{n - \tau}{\lambda'_2} \bar{\mathbf{y}}_{n-\tau}, \\ \mathbf{B}' &= \tau \left[\mathbf{S}_\tau + \left(1 - \frac{\tau}{\lambda'_1}\right) (\bar{\mathbf{y}}_\tau - \phi_1) (\bar{\mathbf{y}}_\tau - \phi_1)^t \right] \\ &\quad + (n - \tau) \left[\mathbf{S}_{n-\tau} + \left(1 - \frac{n - \tau}{\lambda'_2}\right) (\bar{\mathbf{y}}_{n-\tau} - \phi_2) (\bar{\mathbf{y}}_{n-\tau} - \phi_2)^t \right] + \mathbf{B}, \\ a' &= a + n/2. \end{aligned}$$

Démonstration. On procède de manière tout à fait analogue à la démonstration de la Proposition 1. Selon le théorème de Bayes, on a :

$$p(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{P} \mid \tau, \mathbf{D}) \propto p(\mathbf{D} \mid \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \tau, \mathbf{P}) \mathcal{N}\mathcal{N}\mathcal{W}_p(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{P} \mid \phi_1, \lambda_1, \phi_2, \lambda_2, a, \mathbf{B}). \quad (5.7)$$

En vertu des propriétés des lois conjuguées cette densité correspond au produit de deux lois normales multidimensionnelles avec une loi Wishart. Quelques calculs matriciels (essentiellement des complétions de carrés) mènent aux hyperparamètres. On évalue ensuite la loi marginale des données \mathbf{D} , connaissant τ , en faisant le rapport de (5.7) par (5.5). On vérifie sans peine que :

$$p(\mathbf{D} \mid \tau) = (2\pi)^{-\frac{np}{2}} \left(\frac{\lambda_1 \lambda_2}{\lambda'_1 \lambda'_2} \right)^{p/2} \prod_{k=1}^p \left\{ \frac{\Gamma\left[\frac{1}{2}(2a' + 1 - k)\right]}{\Gamma\left[\frac{1}{2}(2a + 1 - k)\right]} \right\} \frac{|\mathbf{B}|^a}{|\mathbf{B}'|^{a'}}.$$

Enfin, du fait que $p(\tau \mid \mathbf{D}) \propto p(\mathbf{D} \mid \tau) p(\tau)$, on déduit (5.6). ■

Proposition 8 *Considérons le modèle de rupture (5.2) et la loi jointe a priori (5.4). Soient les hyperparamètres mis à jour ϕ'_1 , ϕ'_2 , a' et \mathbf{B}' et la densité a posteriori $p(\tau \mid \mathbf{D})$ définis à la Proposition 7. Alors,*

$$p(\boldsymbol{\mu}_k \mid \mathbf{D}) = \sum_{\tau=1}^{n-1} \mathcal{S}\mathcal{T}_p(\boldsymbol{\mu}_k \mid \mathbf{m}_k, \mathbf{A}_k, c) p(\tau \mid \mathbf{D}), \quad k = 1, 2, \quad (5.8)$$

$$p(\mathbf{P} \mid \mathbf{D}) = \sum_{\tau=1}^{n-1} \mathcal{W}_p(\mathbf{P} \mid a', \mathbf{B}') p(\tau \mid \mathbf{D}), \quad (5.9)$$

où $\mathbf{m}_k = \phi'_k$, $\mathbf{A}_k = (2a' - p + 1) \frac{\lambda'_k}{2} \mathbf{B}'^{-1}$, $c = 2a' - p + 1$ et où $ST_p(\mathbf{x} \mid \mathbf{m}, \mathbf{A}, c)$ désigne la densité d'une loi de Student multivariée.

Démonstration. Comme pour la Proposition 2, on doit déterminer d'abord les densités conditionnelles $p(\boldsymbol{\mu}_k \mid \tau, \mathbf{D})$ et $p(\mathbf{P} \mid \tau, \mathbf{D})$. Ensuite, il suffit de déconditionner par rapport à τ . La densité $p(\mathbf{P} \mid \tau, \mathbf{D})$ est obtenue directement en intégrant (5.5) par rapport aux moyennes $\boldsymbol{\mu}_1$ et $\boldsymbol{\mu}_2$. La détermination des lois conditionnelles $p(\boldsymbol{\mu}_k \mid \tau, \mathbf{D})$ nécessitent quelques manipulations supplémentaires. A titre d'exemple, nous calculons la densité de probabilité de $(\boldsymbol{\mu}_1 \mid \tau, \mathbf{D})$.

En intégrant (5.5) par rapport à $\boldsymbol{\mu}_2$ et \mathbf{P} , on a que

$$\begin{aligned} p(\boldsymbol{\mu}_1 \mid \tau, \mathbf{D}) &= \int \int \mathcal{N}_p(\boldsymbol{\mu}_1 \mid \phi'_1, \lambda'_1 \mathbf{P}) \mathcal{N}_p(\boldsymbol{\mu}_2 \mid \phi'_2, \lambda'_2 \mathbf{P}) \mathcal{W}_p(\mathbf{P} \mid a', \mathbf{B}') d\boldsymbol{\mu}_2 d\mathbf{P}, \\ &= \int \mathcal{N}_p(\boldsymbol{\mu}_1 \mid \phi'_1, \lambda'_1 \mathbf{P}) \mathcal{W}_p(\mathbf{P} \mid a', \mathbf{B}') d\mathbf{P}, \\ &\propto \int |\mathbf{P}|^{a'+1/2-(p+1)/2} \exp \left\{ -\frac{1}{2} (\boldsymbol{\mu}_1 - \phi'_1)^t \lambda'_1 \mathbf{P} (\boldsymbol{\mu}_1 - \phi'_1) - \text{tr}(\mathbf{B}'\mathbf{P}) \right\} d\mathbf{P}. \end{aligned}$$

Puisque $(\boldsymbol{\mu}_1 - \phi'_1)^t \lambda'_1 \mathbf{P} (\boldsymbol{\mu}_1 - \phi'_1)$ est un scalaire et que \mathbf{P} est une matrice carrée, en utilisant les propriétés de l'opérateur trace (Théorèmes 9.1.20, 9.1.1, 9.1.4 dans Graybill (1983)), on peut vérifier que

$$\begin{aligned} \frac{1}{2} (\boldsymbol{\mu}_1 - \phi'_1)^t \lambda'_1 \mathbf{P} (\boldsymbol{\mu}_1 - \phi'_1) + \text{tr}(\mathbf{B}'\mathbf{P}) &= \text{tr} \left[\frac{1}{2} (\boldsymbol{\mu}_1 - \phi'_1)^t \lambda'_1 \mathbf{P} (\boldsymbol{\mu}_1 - \phi'_1) \right] + \text{tr}(\mathbf{B}'\mathbf{P}), \\ &= \text{tr} \left[\left(\frac{\lambda'_1}{2} (\boldsymbol{\mu}_1 - \phi'_1) (\boldsymbol{\mu}_1 - \phi'_1)^t + \mathbf{B}' \right) \mathbf{P} \right], \end{aligned}$$

d'où

$$p(\boldsymbol{\mu}_1 \mid \tau, \mathbf{D}) \propto \int |\mathbf{P}|^{a'+1/2-(p+1)/2} \exp \left\{ -\text{tr} \left[\left(\frac{\lambda'_1}{2} (\boldsymbol{\mu}_1 - \phi'_1) (\boldsymbol{\mu}_1 - \phi'_1)^t + \mathbf{B}' \right) \mathbf{P} \right] \right\} d\mathbf{P}.$$

Or, puisque l'intégrand correspond, à une constante près, à une densité de Wishart (Annexe C), on en déduit que

$$p(\boldsymbol{\mu}_1 \mid \tau, \mathbf{D}) \propto \left| \frac{\lambda'_1}{2} (\boldsymbol{\mu}_1 - \phi'_1) (\boldsymbol{\mu}_1 - \phi'_1)^t + \mathbf{B}' \right|^{-\left(a'+\frac{1}{2}\right)}.$$

Enfin, à partir des théorèmes 8.3.3 et 8.4.3 dans Graybill (1983), on peut écrire :

$$\begin{aligned} p(\boldsymbol{\mu}_1 | \tau, \mathbf{D}) &\propto \left| \frac{\lambda'_1}{2} (\boldsymbol{\mu}_1 - \boldsymbol{\phi}'_1) (\boldsymbol{\mu}_1 - \boldsymbol{\phi}'_1)^t + \mathbf{B}' \right|^{-(a'+\frac{1}{2})}, \\ &\propto \left| \mathbf{I}_p + \mathbf{B}'^{-1} \frac{\lambda'_1}{2} (\boldsymbol{\mu}_1 - \boldsymbol{\phi}'_1) (\boldsymbol{\mu}_1 - \boldsymbol{\phi}'_1)^t \right|^{-(a'+\frac{1}{2})}, \\ &= \left[1 + (\boldsymbol{\mu}_1 - \boldsymbol{\phi}'_1)^t \frac{\lambda'_1}{2} (\mathbf{B}'^{-1})^t (\boldsymbol{\mu}_1 - \boldsymbol{\phi}'_1) \right]^{-(a'+\frac{1}{2})}, \end{aligned}$$

ce qui correspond bien, à une constante de normalisation près, à la loi de Student multidimensionnelle (5.8). ■

On peut notamment déduire de cette dernière proposition les espérances et les covariances a posteriori des vecteurs-moyennes avant et après la rupture ($\boldsymbol{\mu}_k$, $k = 1, 2$). On obtient respectivement

$$\begin{aligned} \mathbb{E}\{\boldsymbol{\mu}_k | \mathbf{D}\} &= \mathbb{E}_{\tau|y} \{\mathbb{E}\{\boldsymbol{\mu}_k | \tau, \mathbf{D}\}\}, \\ &= \sum_{\tau=1}^{n-1} p(\tau | \mathbf{D}) \mathbb{E}\{\boldsymbol{\mu}_k | \tau, \mathbf{D}\} = \sum_{\tau=1}^{n-1} p(\tau | \mathbf{D}) \boldsymbol{\phi}'_k \end{aligned} \quad (5.10)$$

et

$$\begin{aligned} \text{Cov}\{\boldsymbol{\mu}_k | \mathbf{D}\} &= \mathbb{E}_{\tau|y} \{\text{Cov}\{\boldsymbol{\mu}_k | \tau, \mathbf{D}\}\} + \text{Cov}_{\tau|y} \{\mathbb{E}\{\boldsymbol{\mu}_k | \tau, \mathbf{D}\}\}, \\ &= \sum_{\tau=1}^{n-1} p(\tau | y) \frac{2\mathbf{B}'}{\lambda'_k (2a' - p + 1)} \\ &\quad + \sum_{\tau=1}^{n-1} p(\tau | y) (\boldsymbol{\phi}'_k - \mathbb{E}\{\boldsymbol{\mu}_k | \mathbf{D}\}) (\boldsymbol{\phi}'_k - \mathbb{E}\{\boldsymbol{\mu}_k | \mathbf{D}\})^t, \\ &= \sum_{\tau=1}^{n-1} p(\tau | y) \left[\frac{2\mathbf{B}'}{\lambda'_k (2a' - p + 1)} + (\boldsymbol{\phi}'_k - \mathbb{E}\{\boldsymbol{\mu}_k | \mathbf{D}\}) (\boldsymbol{\phi}'_k - \mathbb{E}\{\boldsymbol{\mu}_k | \mathbf{D}\})^t \right]. \end{aligned} \quad (5.11)$$

5.1.2 Application du modèle MM_1 à des séries d'apports naturels pour des sites du nord-est québécois

Les circulations océaniques influencent les conditions météorologiques au-dessus de la mer et également au-dessus des continents. Les indices climatiques, comme le NAO (Figure 2.5), mesurent l'amplitude des oscillations climatiques liées à ces circulations océaniques. Rasmussen et al. (1999) ont étudié le lien entre divers indices climatiques et les apports naturels pour un ensemble de bassins versants du Québec. L'intensité de l'oscillation climatique mesurée par ces indices est particulièrement prononcée au début de l'année et

l'effet sur la précipitation et la température au-dessus des continents est susceptible de persister plusieurs mois. Ces auteurs ont donc considéré les apports moyens sur la période de février à juillet, soit ceux qui intègrent les précipitations hivernales. Ils ont notamment trouvé des corrélations non négligeables entre l'indice NAO et les apports printaniers des bassins situés dans les régions de Churchill Falls, La Grande et de la Côte-Nord. Rappelons que cet indice mesure l'amplitude des oscillations climatiques au-dessus de l'Atlantique Nord. Rasmussen et al. (1999) ont aussi constaté la présence d'un changement de moyenne à la baisse autour de 1984 dans certaines des séries d'apports, qui semble coïncidé avec une augmentation des valeurs de l'indice NAO vers 1980. Pour ce faire, ils ont appliqué la procédure unidimensionnelle de Lee et Heghinian (1977) aux séries d'apports et à celles des indices climatiques. Compte tenu du lien entre les séries hydrologiques du nord-est québécois et l'indice NAO, on peut présumer qu'une variation du climat soit à l'origine de cette baisse d'hydraulicité.

Nous nous attachons ici à analyser les changements en moyenne dans huit séries d'apports naturels de bassins versants situés dans cette région. Nous considérons le modèle multisite à rupture simultanée MM_1 (éq. (5.2)) et reprenons l'étude effectuée dans l'Article III (Annexe E) en y ajoutant deux sites, ceux de Caniapiscou et de Ashuanipi. Le Tableau 5.1 donne les superficies de chacun des bassins versants ainsi que la période d'échantillonnage. La Figure 5.1 situe les huit stations sur la carte du Québec.

TAB. 5.1 – Superficies des bassins versant et périodes d'échantillonnage

Site	Station	Surface (km^2)	Periode
1	Romaine	13 000	1957-95
2	À la Baleine	29 800	1963-95
3	Churchill Falls	69 344	1957-95
4	Manicouagan	45 480	1957-95
5	Outardes	18 798	1957-95
6	Ste-Marguerite	4 737	1957-95
7	Caniapiscau	37 666	1963-95
8	Ashuanipi	19 000	1957-95

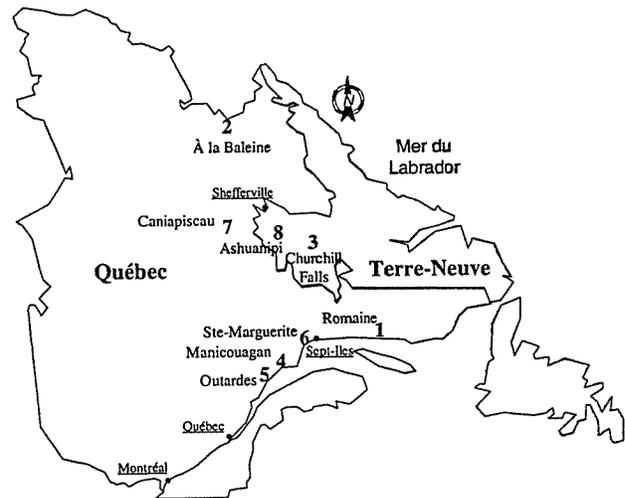


FIG. 5.1 – Positions géographiques des huit sites du nord-est québécois

Comme il est d'usage de le faire pour les études régionales, les observations ont été exprimées en lame d'eau, soit en $l/(km^2 \times s)$ plutôt qu'en m^3/s . Les valeurs sont ainsi comparables d'un site à l'autre. La Figure 5.2 présente le tracé de chacune des séquences d'observations d'apports moyens printaniers (février à juillet) pour la période concomitante de 1963 à 1995.

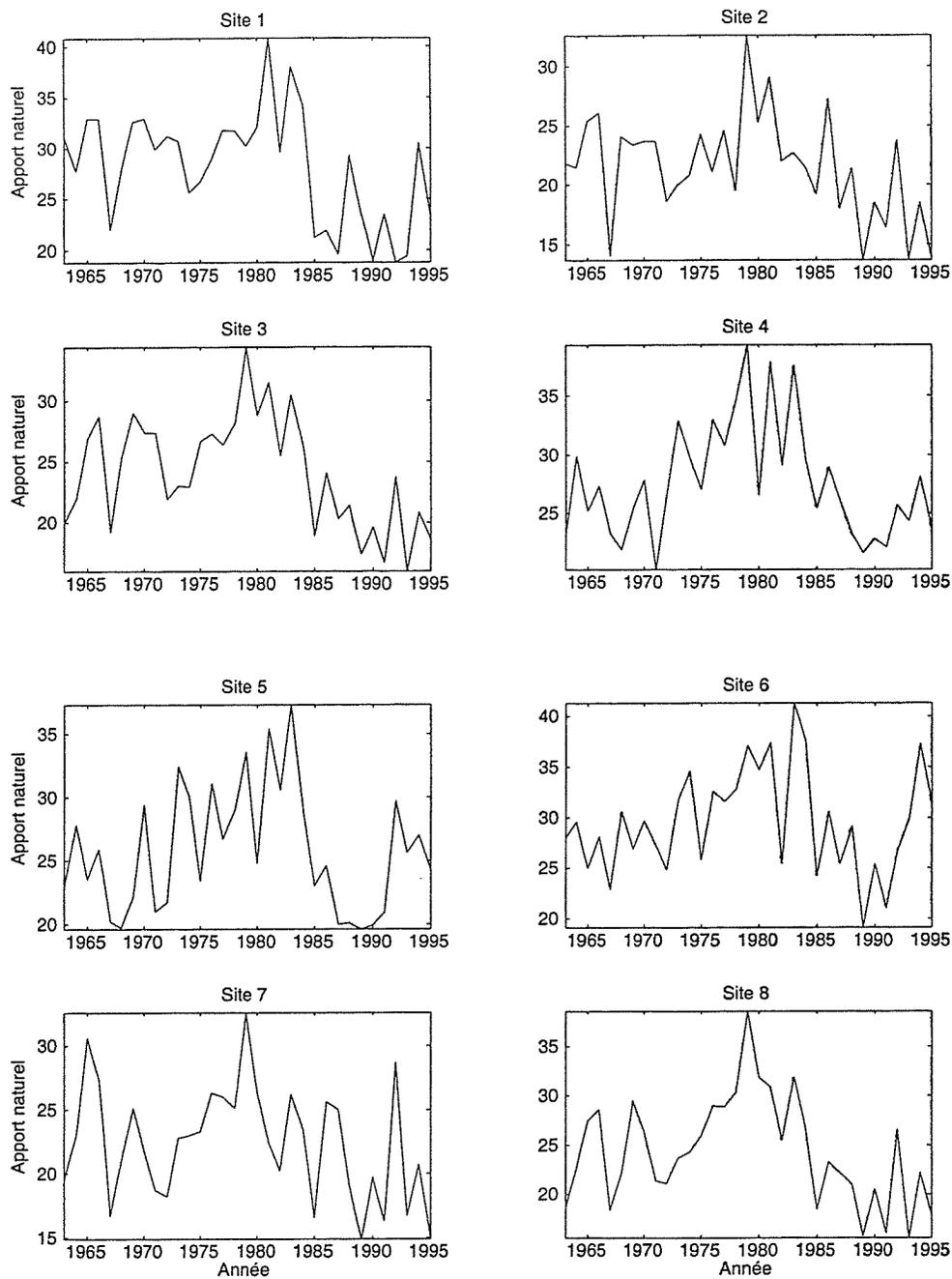


FIG. 5.2 – Apports naturels moyens de février à juillet pour huit bassins du nord-est québécois ($l/(km^2 \times s)$)

A l'examen de ces graphiques, on constate en effet la possibilité qu'une rupture en moyenne, survenue au milieu des années 80, soit présente dans chacune des séquences. L'analyse qui suit a principalement pour objet de caractériser ce changement global en estimant l'année commune de rupture ainsi que les moyennes avant et après la rupture.

Spécification des lois a priori

Nous avons adopté une loi uniforme pour la répartition annuelle des probabilité a priori de rupture $p(\tau)$. De la même manière qu'à l'Article III, les hyperparamètres ϕ_1 , λ_1 , ϕ_2 , λ_2 , a et \mathbf{B} ont été spécifiés à partir des observations de 1957 à 1962 disponibles avant la période concomitante pour six des huit bassins (Romaine, Churchill Falls, Manicouagan, Outardes, Sainte-Marguerite et Ashuanipi). Cette approche s'apparente à celle employée pour une analyse bayésienne empirique (Berger (1985)). Le vecteur-moyennes et la matrice des covariances de cet échantillon a priori sont respectivement :

$$\bar{\mathbf{y}} = \begin{pmatrix} 27.2 \\ \text{nd} \\ 23.3 \\ 27.4 \\ 25.4 \\ 24.0 \\ \text{nd} \\ 24.3 \end{pmatrix} \quad \text{et} \quad \mathbf{S}_y^2 = \begin{pmatrix} 11.70 & \text{nd} & 12.82 & 15.91 & 19.60 & 10.54 & \text{nd} & 13.94 \\ & \text{nd} \\ & & 18.70 & 19.90 & 23.16 & 15.60 & \text{nd} & 19.31 \\ & & & 26.75 & 30.70 & 17.73 & \text{nd} & 23.04 \\ & & & & 38.41 & 17.62 & \text{nd} & 28.09 \\ & & & & & 17.34 & \text{nd} & 14.68 \\ & & & & & & \text{nd} & \text{nd} \\ & & & & & & & 22.35 \end{pmatrix}$$

Pour les bassins Tête à la Baleine et Caniapiscou (sites **2** et **7**), ces valeurs ne sont pas disponibles, les stations ayant été mises en service après 1962. Les moyennes et les variances manquantes ont alors été estimées en faisant les moyennes des valeurs correspondantes des six autres sites. Cela est raisonnable puisqu'en transformant les observations en lame d'eau ($l / (km^2 \times s)$) elles sont en quelque sorte « normalisées ». Les covariances impliquant les séries **2** et **7** ont toutefois été fixées à zéro, laissant ainsi les données déterminer seules la nature de la dépendance lors de la mise à jour par le théorème de Bayes. Nous désignons respectivement par $\hat{\bar{\mathbf{y}}}$ et $\hat{\mathbf{S}}_y^2$ le vecteur-moyenne et la matrice des covariances « reconstitués » de l'échantillon a priori.

Connaissant l'année de rupture τ , on peut présumer que les observations avant le changement appartiennent à la même population que celles de l'échantillon a priori. Evidemment, cela suppose qu'elles soient stationnaires et qu'aucune rupture ne soit survenue en 1962 ou 1963. Cette hypothèse étant raisonnable, il est naturel de choisir $\phi_1 = \hat{\bar{\mathbf{y}}}$ puisque $\mathbb{E}\{\mu_1\} = \phi_1$. De manière analogue, en vertu de $\mathbb{E}\{\mathbf{P}\} = a\mathbf{B}^{-1}$, il est tout indiqué de prendre $\mathbf{B} = a\hat{\mathbf{S}}_y^2$. L'incertitude sur ces paramètres est conditionnée par la représentativité de notre ensemble de données a priori. Celle-ci est d'autant plus grande que la taille de cette séquence est restreinte. On a déjà évoqué que la précision accordée aux

connaissances a priori sur les paramètres μ_1 , μ_2 et \mathbf{P} peut être ajustée à partir de λ_1 , λ_2 et a . Ces hyperparamètres peuvent d'ailleurs être comparés à des degrés de liberté. Nous avons fixé $\lambda_1 = a = 6$. Cela signifie que l'information apportée par l'échantillon a priori sur les paramètres μ_1 and \mathbf{P} est équivalente à celle fournie par six années d'observations historiques. Par contre, l'échantillon a priori étant probablement moins informatif au sujet du vecteur-moyennes après la rupture, nous avons choisi $\lambda_2 = 1$. On a alors une loi plus diffuse pour μ_2 autour de son espérance a priori ϕ_2 . Faute d'information a priori sur ϕ_2 , la valeur \hat{y} lui fut finalement assignée.

Résultats

La Figure 5.3 présente la répartition des probabilités a posteriori $p(\tau | \mathbf{D})$ pour chaque année ainsi que la densité a priori uniforme.

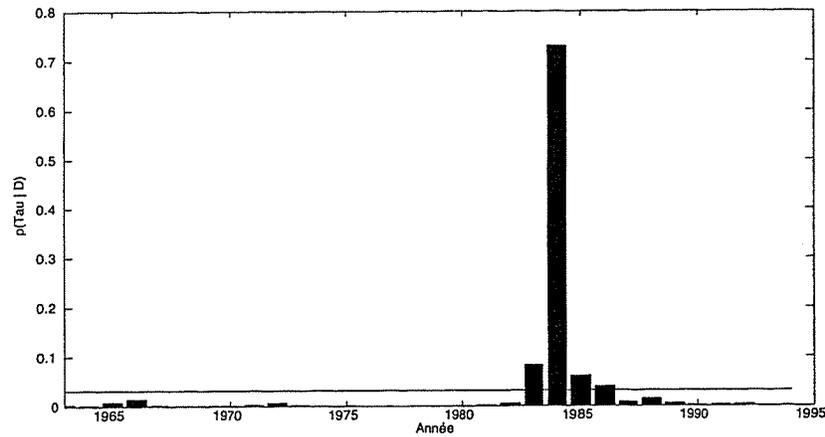


FIG. 5.3 – Densité marginale a priori et a posteriori de l'année de rupture pour huit bassins du nord-est québécois

Sous l'hypothèse que chacune des séries d'apports admette une rupture la même année, cette distribution suggère que le changement soit survenu en 1984 avec une probabilité de 73%. Le facteur de Bayes qui compare l'hypothèse de changement contre celle du non changement peut être calculé à partir de l'expression suivante (équation (E.13) de l'Article III) :

$$B_{c,nc}(\mathbf{D}) = \sum_{\tau=1}^{n-1} p(\tau) \left(\frac{(\lambda_1 + n) \lambda_2}{(\lambda_1 + \tau) (\lambda_2 + n - \tau)} \right)^{p/2} \left(\frac{|\mathbf{B}'_0|}{|\mathbf{B}'|} \right)^{a+n/2}, \quad (5.12)$$

\mathbf{B}'_0 étant la mise à jour de l'hyperparamètre \mathbf{B}_0 de la loi Wishart sous l'hypothèse de stationnarité (voir Article III). On obtient ici 47.88, ce qui tend à favoriser l'hypothèse d'une rupture dans les huit séries (forte évidence en faveur du changement selon la Table 4.2). Les résultats tendent donc à appuyer la thèse de Rasmussen et al. (1999) selon

laquelle un changement de moyenne ait pris place vers le milieu des années 80 dans les séries d'apports hydrologiques intégrant les précipitations hivernales sur la région du nord-est québécois. Aussi est-il intéressant de pousser l'analyse un peu plus loin, notamment en estimant les moyennes avant et après la rupture.

Le Tableau 5.2 donne les espérances et les écarts-types a posteriori des moyennes de part et d'autre de l'année de rupture. Les espérances a priori y sont aussi présentées. À ce propos, on note que la prise en compte des observations vient modifier considérablement l'information a priori sur les vecteurs-moyennes. La Figure 5.4 montre les intervalles de crédibilité des moyennes de part et d'autre de l'année de changement.

TAB. 5.2 – *Espérances et écarts-types a posteriori des moyennes de part et d'autre de l'année de rupture pour les huit bassins versant du nord-est québécois*

		Bassins versants							
		1	2	3	4	5	6	7	8
μ_1	Esp. a priori ϕ_1	27.2	25.3	23.3	27.4	25.4	24.0	25.3	24.3
	Esp. a posteriori	30.0	23.5	25.5	28.6	26.7	29.0	24.0	25.9
	Éc.-ty. a posteriori	(0.99)	(0.96)	(0.95)	(1.10)	(1.18)	(1.36)	(0.98)	(1.11)
μ_2	Esp. a priori ϕ_2	27.2	25.3	23.3	27.4	25.4	24.0	25.3	24.3
	Esp. a posteriori	23.4	19.2	20.3	25.1	23.5	27.2	20.4	20.6
	Éc.-ty. a posteriori	(1.65)	(1.48)	(1.49)	(1.64)	(1.75)	(1.80)	(1.48)	(1.72)

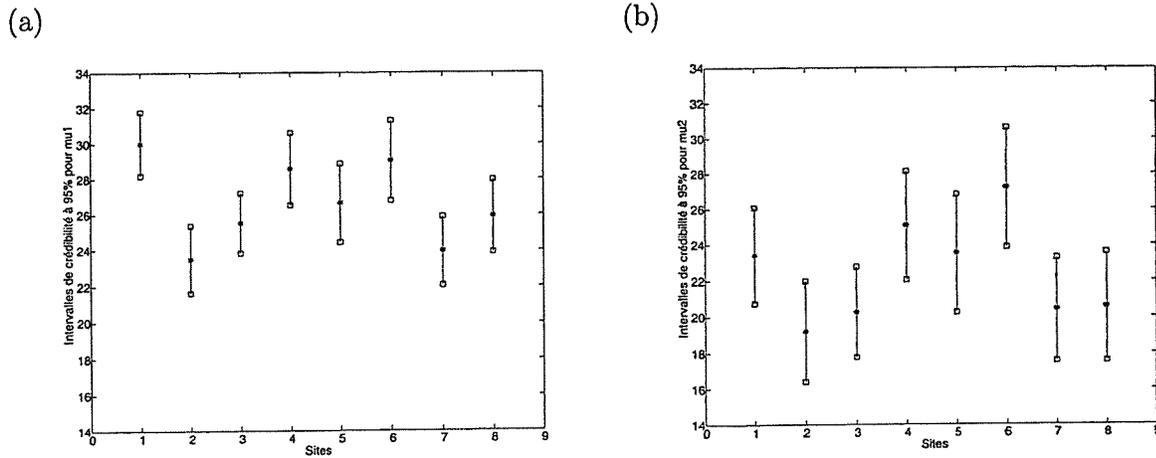


FIG. 5.4 – *Intervalles de crédibilité à 95 pour cent des moyennes avant (a) et après (b) la rupture pour huit bassins du nord-est québécois*

À la lumière des résultats, l'apport printanier moyen aurait chuté d'environ 20% pour les sites Romaine, Tête à la Baleine, Churchill Falls et Ashuanipi, de 15% pour Caniapiscau, d'environ 11% pour les bassins Outardes et Manicouagan et seulement de 6%

pour Ste-Marguerite. Pour les bassins situés plus au sud-ouest (Manicouagan, Outardes et Sainte-Marguerite), les intervalles de crédibilité des moyennes avant et après la rupture se chevauchent. Notons que pour ces trois sites, les apports naturels ont augmenté au début des années 90 contrairement aux autres bassins pour lesquels le niveau moyen après la rupture s'est maintenu jusqu'en 1995. Sans prétendre, au demeurant, à une quelconque évidence, on pourrait émettre l'hypothèse que plus les bassins sont situés à proximité de la mer du Labrador (nord-est du Québec) plus l'amplitude de la rupture est importante et plus la baisse d'hydraulicité risque de persister. Cela ne serait guère surprenant si un changement climatique avait bien eu lieu. En effet, les circulations océaniques de l'Atlantique Nord devraient avoir plus d'influence sur les conditions météorologiques de cette région. Évidemment, pour démontrer la validité d'une telle hypothèse, il faudrait disposer d'un plus grand ensemble de sites mieux répartis géographiquement.

En terminant, il convient de signaler que l'ajout des sites Caniapiscau et Ashuanipi a considérablement diminué le facteur de Bayes obtenu dans l'Article III, qui était de 184.12. De plus, les écart-types a posteriori des moyennes avant et après le changement pour les séries 1 à 6 (Tableau 5.2) sont tous sensiblement plus grands que ceux obtenus lors de la première analyse, et ce, malgré le fait qu'on dispose ici d'un échantillon de taille plus grande. Selon toute vraisemblance, la forte variabilité des apports printaniers à Caniapiscau, ajoutée au fait que la rupture à ce site semble être survenue avant 1984 (voir Figure 5.2), expliquent en grande partie cet accroissement de l'incertitude. Une analyse effectuée en ne considérant que les sept autres bassins versants tend à confirmer cela : le facteur de Bayes est alors de 112.05 et tous les écarts-types a posteriori sont inférieurs à ceux obtenus lors de l'analyse originale (Tableau 3 de l'Article III).

5.2 Échangeabilité et modèle hiérarchique de rupture

Dans la section précédente, nous avons supposé qu'il y avait une cohérence parfaite entre les instants de rupture. Selon nous, cette hypothèse est restrictive voire irréaliste. D'abord, même si les stations forment un ensemble homogène, rien ne garantit que les facteurs climatiques influencent chacun des bassins exactement de la même manière. Ensuite, la réponse hydrologique d'un bassin versant est assujettie à divers facteurs locaux (infiltration, occupation des sols, micro-climat, etc.). Ces derniers, dont la portée peut varier d'une site à l'autre, sont susceptibles d'influer sur le déclenchement du changement dans une séquence hydrologique. Il est néanmoins naturel de présumer qu'il existe une relation structurelle entre les instants de rupture puisque les changements proviennent de la même origine.

Dans cette section nous montrons qu'il est possible de construire des modèles qui supposent que l'instant de changement varie d'une séquence à l'autre, mais qui permettent toute de même de conserver une certaine cohérence régionale entre les dates de rupture. Signalons que Bélisle et al. (1998) ont proposé un modèle similaire aux nôtres dans le cadre d'une étude longitudinale sur les trains de décharges neuronales.

5.2.1 Echangeabilité des instants de rupture

Quelle hypothèse, moins restrictive que le changement simultan , peut-on faire pour laisser la libert    l'instant de rupture de varier d'une station   l'autre tout en pr servant une certaine coh rence r gionale? L'hypoth se minimale consiste   supposer que les dates de changement sont  changeables.

L' changeabilit  des quantit s al atoires $\tau_1, \tau_2, \dots, \tau_p$, dont la d finition suit, est fond e sur des arguments de sym trie en probabilit .

D finition 4 *Les instants de rupture $\tau_1, \tau_2, \dots, \tau_p$ sont dits  changeables si*

$$p(\tau_1, \tau_2, \dots, \tau_p) = p(\tau_{(1)}, \tau_{(2)}, \dots, \tau_{(p)})$$

pour toute permutation $(\tau_{(1)}, \tau_{(2)}, \dots, \tau_{(p)})$ et quel que soit p .

L'hypoth se d' changeabilit  des instants de rupture implique (*th or me de De Finetti*, voir Bernardo et Smith (1994)), lorsque p est grand, l'existence d'un param tre de conditionnement π tel que :

$$p(\tau_1, \tau_2, \dots, \tau_p) = \int \left\{ \prod_{i=1}^p p(\tau_i | \pi) \right\} p(\pi) d\pi. \quad (5.13)$$

Dans cette expression, $p(\tau_i | \pi)$ s'interpr te comme un mod le pour le param tre inconnu τ_i d termin  par un vecteur d'hyperparam tres π . Ces derniers sont inconnus et eux-m mes caract ris s par une loi de probabilit  $p(\pi)$. La loi marginale de $(\tau_1, \tau_2, \dots, \tau_p)$ est alors obtenue en int grant toute l'incertitude sur π . L'expression (5.13) signifie que les instants de changement sont ind pendants lorsque que π est connu (ind pendance conditionnelle), mais que marginalement il existe une relation structurelle entre les τ_i apport e par $p(\pi)$: les instants de ruptures pour chaque s rie appartiennent   la m me population. La loi jointe $p(\tau_1, \tau_2, \dots, \tau_p)$ traduit ainsi une certaine d pendance entre les param tres $(\tau_1, \tau_2, \dots, \tau_p)$.

Toutefois, est-on pour autant autoris    faire l'hypoth se d' changeabilit  des instants de changement? Supposons que nous ayons   notre disposition les dates de rupture d'un ensemble de p stations et que l'on puisse associer   cette r alisation une probabilit  d'occurrence. Le fait de conna tre   quelle station correspond chacun des τ_i est-il susceptible de modifier cette probabilit ? Si la r ponse est non, l'hypoth se d' changeabilit  est valide. Si aucune ou tr s peu d'information n'est disponible sur les m canismes d'un changement climatique,   l'exception du fait qu'il se manifeste de mani re similaire sur l'ensemble des sites (zone homog ne), il est opportun de supposer l' changeabilit . G n ralement, l'ignorance implique l' changeabilit  (Gelman et al. (1995)). Si aucune information ne permet de distinguer les τ_i , de les ordonner ou de les regrouper, on peut en effet supposer qu'il existe une sym trie entre eux en terme de loi de probabilit . Si toutefois nous disposions d'une telle information, cette hypoth se ne serait pas v rifi e. Par exemple, si une  tude

sur les changements climatiques avait montré que ce phénomène était causé principalement par des facteurs qui agissent de manière directionnelle (d'est en ouest, par exemple), causant ainsi un déclenchement séquentiel des ruptures, cela invaliderait l'échangeabilité des τ_i . En effet, si tel était le cas, les indices i « contiendraient » une information qui nous amènerait à rejeter l'équiprobabilité pour toute permutation des τ_i . Notons, par contre, qu'une telle information concernant la propagation des effets du changement climatique sur les dates de rupture implique une connaissance très approfondie du phénomène. Or, ce type de connaissance, qui invaliderait notre hypothèse, n'est pas encore disponible ou tout au moins ne fait pas l'unanimité chez les spécialistes en climatologie. Si le doute persiste, on pourrait quand même envisager une échangeabilité partielle en incluant dans le modèle des variables explicatives, notamment les longitudes et les latitudes.

Dans les chapitres précédents, l'hypothèse d'échangeabilité a déjà été employée implicitement pour les observations. En effet, pour tous les modèles considérés jusqu'ici, on a supposé que les observations étaient indépendantes et identiquement distribuées connaissant l'ensemble des paramètres (voir l'expression des vraisemblances). La loi marginale des données, qui correspond au dénominateur du théorème de Bayes, est d'ailleurs un mélange de lois du type de l'expression (5.13). Aussi peut-on voir dans l'échangeabilité et dans le théorème de Finetti des justifications à l'emploi d'une loi a priori pour représenter les paramètres inconnus d'un modèle. De manière analogue, si on suppose aussi l'échangeabilité de certains de ces paramètres (par exemple les τ_i), il est alors justifié d'ajouter un niveau au modèle. Ce niveau supplémentaire permet ici de formaliser la cohérence régionale qui doit exister entre les instants de rupture, et ce, sans imposer que le changement soit simultané.

5.2.2 Modèle hiérarchique de rupture et stratégie d'estimation

L'hypothèse d'échangeabilité sur les paramètres qui caractérisent les ruptures dans plusieurs séries appelle donc la construction d'un modèle à structure hiérarchique. Considérons l'ensemble de variables aléatoires (5.1) et le modèle général de rupture (3.1) pour chacune des séquences. Supposons que la cohérence régionale des changements se traduit seulement à travers les dates de changement. Une version possible d'un modèle hiérarchique de rupture prend alors la forme suivante :

$$\begin{aligned}
 \text{Niveau 1} \quad & Y_{ij} \stackrel{iid}{\sim} p(y_{ij} | \theta_{i1}), \quad j = 1, \dots, \tau_i \\
 & Y_{ij} \stackrel{iid}{\sim} p(y_{ij} | \theta_{i2}), \quad j = \tau_i + 1, \dots, n, \quad i = 1, \dots, p \\
 \text{Niveau 2} \quad & \tau_i \stackrel{iid}{\sim} p(\tau_i | \pi) \\
 & \theta_{i1} \stackrel{iid}{\sim} p(\theta_{i1} | \phi_{i1}) \\
 & \theta_{i2} \stackrel{iid}{\sim} p(\theta_{i2} | \phi_{i2}) \quad i = 1, \dots, p \\
 \text{Niveau 3} \quad & \pi \sim p(\xi)
 \end{aligned} \tag{5.14}$$

et s'interprète comme suit. Des séquences de n observations hydrologiques, $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_p$, sont disponibles en p sites différents. Ces derniers sont liés entre eux en terme de réponse

hydrologique aux manifestations climatiques : ils forment un ensemble hydrologiquement homogène. Le modèle est composé de deux ensembles de variables constitutives : les paramètres $(\tau_i, \theta_{i1}, \theta_{i2})$ qui décrivent les manifestations locales à chacune des stations et le vecteur π qui caractérise la cohérence régionale. Les composantes d'un modèle paramétrique de rupture pour les observations sont spécifiées au premier niveau de la hiérarchie. La rupture dans la séquence i se produit à l'instant τ_i et elle est caractérisée par les vecteurs-paramètres θ_{i1} et θ_{i2} (par exemple, les moyennes de part et d'autre du changement dans la série i). Dans le second niveau, nous posons les modèles a priori pour l'ensemble de ces paramètres. Les paramètres θ_{i1} et θ_{i2} ne sont pas échangeables et les hyperparamètres de leurs lois a priori, ϕ_{i1} et ϕ_{i2} , sont distincts d'une station à l'autre. Par contre, les τ_i sont considérés comme échangeables et proviennent donc de la même loi de probabilité $p(\tau_i | \pi)$. Cette distribution dépend d'un paramètre inconnu π qui conditionne la répartition spatiale des dates de rupture. Le dernier niveau de la hiérarchie spécifie enfin l'état des connaissances sur ce paramètre de cohérence régionale.

Le modèle hiérarchique (5.14) a exactement la même structure que celui de Bélisle et al. (1998), où une rupture dans le paramètre d'une loi de Poisson fut considérée au premier niveau. Il est utile de représenter ce type de modèle par un diagramme directionnel représentant la structure de dépendance entre les variables constitutives (voir Brooks (1998) et les auteurs qu'il a cités). La Figure 5.5 présente le schéma correspondant au modèle (5.14).

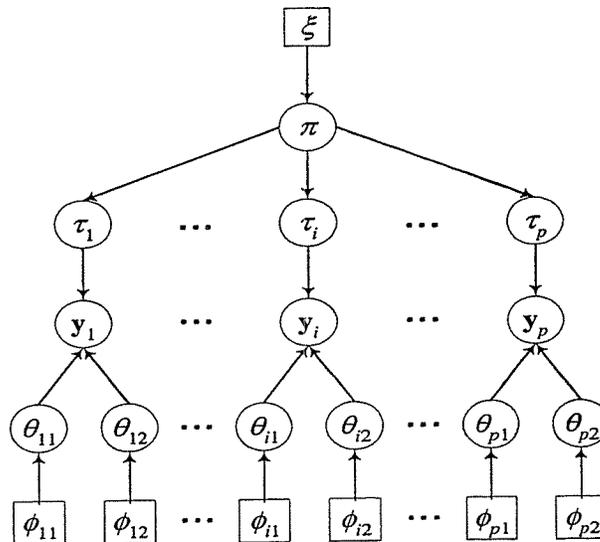


FIG. 5.5 – Représentation schématique d'un modèle hiérarchique de rupture

Dans ce diagramme, les nœuds représentent les variables. Ceux-ci prennent deux formes : soit un cercle, auquel cas la valeur de la variable est inconnue et celle-ci est sujette à estimation, soit un carré, ce qui signifie que la valeur est supposée connue. Les flèches liant deux nœuds indiquent la présence d'une relation entre les variables correspondantes ainsi

que sa direction. Dans le schéma 5.5, par exemple, le noeud ξ correspond à « l'ascendant » de π en ce sens qu'il en détermine la valeur. De manière analogue, les noeuds $\tau_1, \tau_2, \dots, \tau_p$ sont les « descendants » de π . On déduit de cette figure les expressions suivantes :

Fonction de vraisemblance. Les observations y_i de la station i dépendent à la fois de la valeur de τ_i et de celle des paramètres θ_{i1} et θ_{i2} . Toutefois, elles ne dépendent de π qu'à travers τ_i . La séquence d'observations y_i est donc conditionnellement indépendante de π sachant τ_i , ce qui nous permet d'écrire

$$p(\mathbf{D} | \Theta_1, \Theta_2, \tau, \pi) = p(\mathbf{D} | \Theta_1, \Theta_2, \tau),$$

où $\mathbf{D} = (y_1, y_2, \dots, y_n)$, $\Theta_1 = (\theta_{11}, \dots, \theta_{p1})$, $\Theta_2 = (\theta_{12}, \dots, \theta_{p2})$ et $\tau = (\tau_1, \tau_2, \dots, \tau_p)$. Or, puisque les stations ne sont liées entre elles que par le vecteur π , on en déduit :

$$\begin{aligned} p(\mathbf{D} | \Theta_1, \Theta_2, \tau, \pi) &= \prod_{i=1}^p p(y_i | \theta_{i1}, \theta_{i2}, \tau_i) \\ &= \prod_{i=1}^p \left[\prod_{j=1}^{\tau_i} p(y_{ij} | \theta_{i1}) \prod_{j=\tau_i+1}^n p(y_{ij} | \theta_{i2}) \right]. \end{aligned} \quad (5.15)$$

Loi jointe des variables constitutives. Les instants de rupture $(\tau_1, \tau_2, \dots, \tau_p)$ dépendent du vecteur-paramètres π , mais sont conditionnellement indépendants connaissant ce dernier. Les paramètres θ_{i1} et θ_{i2} ne sont pas reliés aux instants de rupture et sont indépendants entre eux. Ainsi, on a

$$\begin{aligned} p(\Theta_1, \Theta_2, \tau, \pi) &= p(\Theta_1 | \Psi_1) p(\Theta_2 | \Psi_2) p(\tau | \pi) p(\pi) \\ &= \left[\prod_{i=1}^p p(\theta_{i1} | \phi_{i1}) p(\theta_{i2} | \phi_{i2}) \right] \left[\prod_{i=1}^p p(\tau_i | \pi) \right] p(\pi), \end{aligned} \quad (5.16)$$

où $\Psi_1 = (\phi_{11}, \dots, \phi_{p1})$ et $\Psi_2 = (\phi_{12}, \dots, \phi_{p2})$.

Loi jointe a posteriori des variables constitutives. Par le théorème de Bayes, on peut combiner les expressions (5.15) et (5.16) pour obtenir la loi jointe a posteriori de l'ensemble des paramètres (manifestations locales et cohérence régionale) d'où l'inférence peut être faite :

$$\begin{aligned} p(\Theta_1, \Theta_2, \tau, \pi | \mathbf{D}) &\propto \prod_{i=1}^p \left\{ \left[\prod_{j=1}^{\tau_i} p(y_{ij} | \theta_{i1}) p(\theta_{i1} | \phi_{i1}) \right. \right. \\ &\quad \left. \left. \times \prod_{j=\tau_i+1}^n p(y_{ij} | \theta_{i2}) p(\theta_{i2} | \phi_{i2}) \right] p(\tau_i | \pi) \right\} p(\pi). \end{aligned} \quad (5.17)$$

La loi jointe a posteriori est fort complexe et n'admet généralement pas d'expressions analytiques explicites pour les densités marginales a posteriori. Néanmoins, l'échantillonnage de Gibbs s'avère particulièrement adapté pour approcher ces dernières, car la structure de dépendance du modèle hiérarchique (5.14) permet d'obtenir facilement les lois conditionnelles complètes. Il suffit d'appliquer la formule de Bayes par bloc de trois « générations » dans la Figure 5.5 (Brooks (1998)). En effet, dans le modèle (5.14), la loi conditionnelle complète a posteriori d'un noeud en particulier, v , s'exprime comme le produit

$$p(v | \text{Reste}, \mathbf{D}) \propto \left[\prod_{u \in C_v} p(u | \text{ascendants de } u) \right] p(v | \text{ascendants de } v), \quad (5.18)$$

où C_v désigne l'ensemble des descendants de v et où le terme générique « Reste » englobe tous les noeuds impliquant des paramètres inconnus. Par exemple, la loi conditionnelle complète du paramètre de cohérence régionale, π , s'écrit de la manière suivante :

$$\begin{aligned} p(\pi | \text{Reste}, \mathbf{D}) &= p(\pi | \Theta_1, \Theta_2, \tau, \mathbf{D}) \\ &\propto \left[\prod_{i=1}^p p(\tau_i | \pi) \right] p(\pi | \xi) \end{aligned} \quad (5.19)$$

On considère alors que π (le paramètre du noeud central) a pour loi a priori celle spécifiée par ξ (son ascendant) et doit être mis à jour en considérant que $\tau_1, \tau_2, \dots, \tau_p$ (ses descendants) forment un échantillon. On peut procéder ainsi, bloc par bloc, pour formuler les lois conditionnelles complètes de chaque variable constitutive à une constante de normalisation près. Plus précisément, on trouve pour les autres paramètres les densités conditionnelles complètes suivantes :

$$p(\tau_i | \text{Reste}, \mathbf{D}) \propto \left[\prod_{j=1}^{\tau_i} p(y_{ij} | \theta_{i1}) \prod_{j=\tau_i+1}^n p(y_{ij} | \theta_{i2}) \right] p(\tau_i | \pi), \quad (5.20)$$

$$p(\theta_{i1} | \text{Reste}, \mathbf{D}) \propto \left[\prod_{j=1}^{\tau_i} p(y_{ij} | \theta_{i1}) \right] p(\theta_{i1} | \phi_{i1}), \quad (5.21)$$

$$p(\theta_{i2} | \text{Reste}, \mathbf{D}) \propto \left[\prod_{j=\tau_i+1}^n p(y_{ij} | \theta_{i2}) \right] p(\theta_{i2} | \phi_{i2}). \quad (5.22)$$

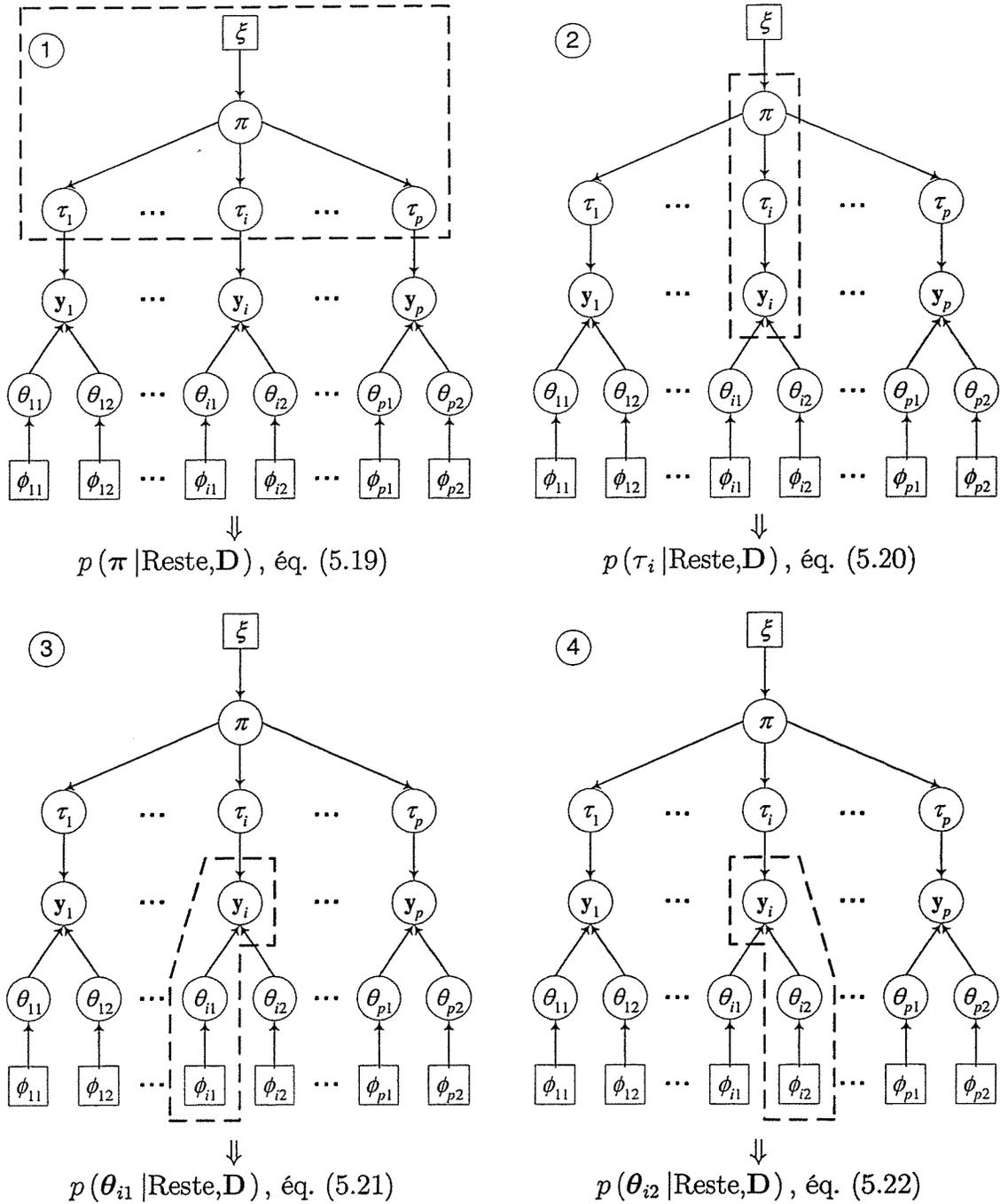


FIG. 5.6 – Détermination des lois conditionnelles complètes a posteriori pour le modèle hiérarchique de rupture

La schéma de la Figure 5.6 illustre la démarche de calcul des distributions conditionnelles complètes a posteriori en faisant intervenir pour chaque noeud ses ascendants et ses descendants selon l'expression (5.18).

Signalons que ces dernières expressions sont valables quelque soit i , puisque dans le modèle (5.14) les observations sont indépendantes d'une station à l'autre sachant π . Enfin, si pour chaque bloc de trois générations des lois a priori conjuguées sont choisies, les constantes de normalisations sont alors déterminées et on peut directement avoir recours à l'échantillonnage de Gibbs pour obtenir une approximation des densités marginales a posteriori.

5.3 Modèle multisite à rupture variable sans effet de covariation spatiale (MM₂)

Dans ce premier modèle hiérarchique (modèle MM₂), nous supposons que chacune des séquences admet une rupture en moyenne seulement et que la cohérence régionale se traduit à travers les dates de changement ainsi que leur intensité. Les observations sont considérées indépendantes d'une station à l'autre et distribuées selon une loi normale. Les covariations spatiales entre les observations à divers sites sont prises en compte dans un second modèle hiérarchique présenté à la Section 5.4.

Le modèle MM₂ diffère principalement de celui proposé par Bélisle et al. (1998) parce qu'il suppose l'échangeabilité non seulement des instants de rupture mais aussi de certains paramètres caractérisant les changements (les moyennes de part et d'autre de la rupture de chaque série).

5.3.1 Modèle et notation

Considérons l'ensemble de données (5.1) constitué de séquences de n observations hydrologiques disponibles à p stations. Supposons qu'un changement climatique ait engendré une modification du régime hydrologique se manifestant par une rupture en moyenne dans chaque séquence i à l'instant inconnu τ_i , $i = 1, \dots, p$. Le modèle de rupture traduisant ce comportement pour une série particulière i est le modèle M₁. Aussi, au premier niveau de la hiérarchie, les observations sont-elles représentées par :

$$\begin{aligned} Y_{ij} &\stackrel{iid}{\sim} \mathcal{N}(y_{ij} \mid \mu_{i1}, \sigma^2), \quad j = 1, \dots, \tau_i \\ Y_{ij} &\stackrel{iid}{\sim} \mathcal{N}(y_{ij} \mid \mu_{i2}, \sigma^2), \quad j = \tau_i + 1, \dots, n, \quad i = 1, \dots, p. \end{aligned}$$

Au second niveau, nous spécifions les lois a priori pour caractériser l'incertitude sur les paramètres décrivant les manifestations locales. Nous adoptons la représentation suivante :

$$\begin{aligned}
\tau_i &\stackrel{iid}{\sim} \mathcal{M}(t \mid \pi_1, \dots, \pi_{n-1}, 1), \quad i = 1, \dots, p, \\
\mu_{i1} &\stackrel{iid}{\sim} \mathcal{N}(\mu_{i1} \mid \phi_1, v_1), \quad i = 1, \dots, p, \\
\mu_{i2} &\stackrel{iid}{\sim} \mathcal{N}(\mu_{i2} \mid \phi_2, v_2), \quad i = 1, \dots, p, \\
1/\sigma^2 &\sim \mathcal{G}(1/\sigma^2 \mid \alpha, \beta),
\end{aligned}$$

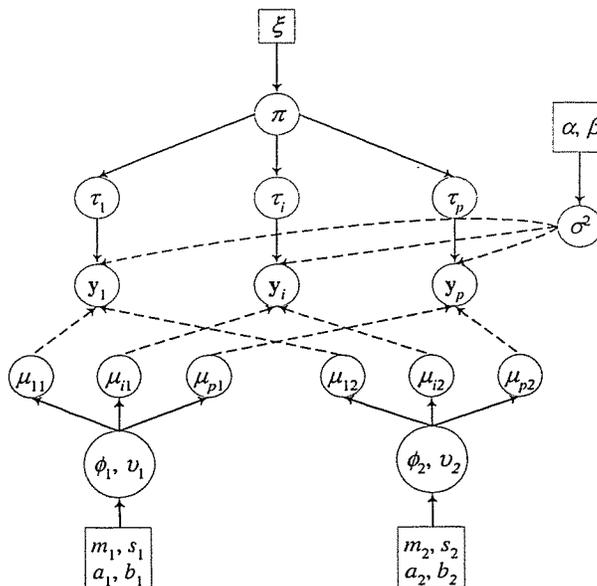
où $\mathcal{M}(t \mid \pi_1, \dots, \pi_{n-1}, 1)$, avec $\pi_h = \Pr(t = h)$, désigne la loi multinomiale à un seul tirage et où $\mathcal{G}(y \mid a, b)$ est la densité de la loi gamma. Nous supposons ici que les instants de rupture ainsi que les moyennes sont échangeables d'une station à l'autre. À l'instar des instants de rupture, il nous apparaît également justifié de présumer que les moyennes proviennent de la même population. En effet, dans la pratique, les hydrologues font implicitement cette hypothèse dans les études régionales de crues : on estime d'abord les quantiles (ou d'autres paramètres) à chaque site d'une zone homogène, ensuite on les considère comme des réalisations d'une variable dépendante dans un modèle de régression multiple faisant intervenir des facteurs physiographiques de leurs bassins versants ; enfin, on emploie ce modèle régional pour estimer le quantile à un site non-jaugé appartenant à la même zone (GREHYS (1996)).

Les paramètres de cohérence régionale sont donc ici $\boldsymbol{\pi} = (\pi_1, \dots, \pi_{n-1})$, (ϕ_1, v_1) et (ϕ_2, v_2) . Leurs lois de probabilité sont spécifiées au dernier niveau de la hiérarchie. Nous prenons :

$$\begin{aligned}
(\pi_1, \dots, \pi_{n-1}) &\sim \mathcal{D}_{n-1}(\boldsymbol{\pi} \mid \xi_1, \dots, \xi_{n-1}), \\
(\phi_1, 1/v_1) &\sim \mathcal{N}(\phi_1 \mid m_1, s_1) \mathcal{G}(1/v_1 \mid a_1, b_1), \\
(\phi_2, 1/v_2) &\sim \mathcal{N}(\phi_2 \mid m_2, s_2) \mathcal{G}(1/v_2 \mid a_2, b_2),
\end{aligned}$$

où $\mathcal{D}_{n-1}(\boldsymbol{\pi} \mid \xi_1, \dots, \xi_{n-1})$ désigne la loi de Dirichlet. Signalons enfin que par souci de parcimonie, nous avons supposé que le terme d'erreur global du modèle possède la même variance σ^2 , quelque soit la station et sur toute la période d'échantillonnage. Le diagramme directionnel de la structure de dépendance entre les variables constitutives du modèle MM_2 est présenté à la Figure 5.7.

Comme il a été montré à la Section 5.2, à chaque noeud de ce diagramme correspond une loi conditionnelle complète que l'on peut déterminer à partir de l'équation (5.18). Puisque nous avons adopté des lois conjuguées de niveau en niveau, ces distributions, dont les expressions sont données dans la section qui suit, sont complètement déterminées et de forme standard. La mise en oeuvre de l'algorithme de Gibbs pour l'estimation des paramètres est alors simple et directe.


 FIG. 5.7 – Représentation schématisée du modèle hiérarchique de rupture MM_2

5.3.2 Lois conditionnelles complètes

Étant donnée la structure de dépendance du modèle MM_2 (Figure 5.7), la loi jointe de l'ensemble des variables constitutives s'écrit de la manière suivante :

$$\begin{aligned}
 p(\Theta, D) &= \prod_{i=1}^p \left[\prod_{j=1}^{\tau_i} \mathcal{N}(y_{ij} \mid \mu_{i1}, \sigma^2) \prod_{j=\tau_i+1}^n \mathcal{N}(y_{ij} \mid \mu_{i2}, \sigma^2) \right] & (5.23) \\
 &\times \left[\prod_{i=1}^p \mathcal{N}(\mu_{i1} \mid \phi_1, v_1) \right] \left[\prod_{i=1}^p \mathcal{N}(\mu_{i2} \mid \phi_2, v_2) \right] \left[\prod_{i=1}^p \mathcal{M}(\tau_i \mid \pi, 1) \right] \\
 &\times \mathcal{G}(1/\sigma^2 \mid \alpha, \beta) \mathcal{N}(\phi_1 \mid m_1, s_1) \mathcal{G}(1/v_1 \mid a_1, b_1) \\
 &\times \mathcal{N}(\phi_2 \mid m_2, s_2) \mathcal{G}(1/v_2 \mid a_2, b_2) \mathcal{D}_{n-1}(\pi \mid \xi),
 \end{aligned}$$

où, pour simplifier la notation, Θ désigne l'ensemble des paramètres locaux et régionaux.

À partir de cette expression, on constate effectivement que chacune des lois conditionnelles complètes peuvent être obtenues en isolant trois générations dans la Figure 5.7 et en mettant à jour les hyperparamètres du noeud central. Par exemple, pour le paramètre local μ_{i1} , on a

$$\begin{aligned}
p(\mu_{i1} | \text{Reste}, \mathbf{D}) &= \frac{p(\Theta, \mathbf{D})}{\int p(\Theta, \mathbf{D}) d\mu_{i1}} \\
&\propto \left[\prod_{j=1}^{\tau_i} \mathcal{N}(y_{ij} | \mu_{i1}, \sigma^2) \right] \mathcal{N}(\mu_{i1} | \phi_1, v_1)
\end{aligned}$$

puisque tous les éléments de (5.23) qui n'impliquent pas ce dernier s'annulent. La loi conditionnelle complète est alors aisément obtenue en utilisant la conjugaison normale-normale. De manière analogue, pour le paramètre de cohérence régionale π , on trouve

$$\begin{aligned}
p(\pi | \text{Reste}, \mathbf{D}) &= \frac{p(\Theta, \mathbf{D})}{\int p(\Theta, \mathbf{D}) d\pi} \\
&\propto \left[\prod_{i=1}^p \mathcal{M}(\tau_i | \pi, 1) \right] \mathcal{D}_{n-1}(\pi | \xi)
\end{aligned}$$

Cette expression est entièrement déterminée grâce à la conjugaison multinomiale-Dirichlet. Il en va de même pour l'ensemble des paramètres du modèle MM_2 . Aussi les calculs ne font-ils appel qu'à des résultats standards bien connus concernant les lois conjuguées (voir notamment Berger (1985), Robert (1992) ou Bernardo et Smith (1994)).

Paramètres pour les manifestations locales

En exploitant la conjugaison normale-normale, nous obtenons les lois conditionnelles complètes suivantes pour les moyennes de part et d'autre de l'instant de rupture de chaque séquence $i = 1, \dots, p$:

$$p(\mu_{i1} | \text{Reste}, \mathbf{D}) = \mathcal{N} \left[\mu_{i1} \left| \left(\frac{1}{v_1^2} + \frac{\tau_i}{\sigma^2} \right)^{-1} \left(\frac{\phi_1}{v_1^2} + \sum_{j \leq \tau_i} \frac{y_{ij}}{\sigma^2} \right), \left(\frac{1}{v_1^2} + \frac{\tau_i}{\sigma^2} \right)^{-1} \right. \right], \quad (5.24)$$

$$p(\mu_{i2} | \text{Reste}, \mathbf{D}) = \mathcal{N} \left[\mu_{i2} \left| \left(\frac{1}{v_2^2} + \frac{n - \tau_i}{\sigma^2} \right)^{-1} \left(\frac{\phi_2}{v_2^2} + \sum_{j > \tau_i} \frac{y_{ij}}{\sigma^2} \right), \left(\frac{1}{v_2^2} + \frac{n - \tau_i}{\sigma^2} \right)^{-1} \right. \right]. \quad (5.25)$$

Pour la variance globale σ^2 , on tire parti de la conjugaison normale-gamma. Il vient alors :

$$p(1/\sigma^2 | \text{Reste}, \mathbf{D}) = \mathcal{G} \left(1/\sigma^2 \left| \alpha + \frac{np}{2}, \frac{1}{2} \sum_i \sum_j (y_{ij} - \mu_{ij})^2 + \beta \right. \right). \quad (5.26)$$

Enfin, la distribution conditionnelle complète pour l'instant de rupture τ_i ($i = 1, \dots, p$) est une loi discrète qui a la forme de l'expression (5.20). Elle s'écrit ici de la manière suivante pour $h = 1, \dots, n - 1$:

$$\Pr(\tau_i = h | \text{Reste}, \mathbf{D}) = \frac{\left[\prod_{j=1}^h \mathcal{N}(y_{ij} | \mu_{i1}, \sigma^2) \prod_{j=h+1}^n \mathcal{N}(y_{ij} | \mu_{i2}, \sigma^2) \right] \pi_h}{\sum_{k=1}^{n-1} \left[\prod_{j=1}^k \mathcal{N}(y_{ij} | \mu_{i1}, \sigma^2) \prod_{j=k+1}^n \mathcal{N}(y_{ij} | \mu_{i2}, \sigma^2) \right] \pi_k} \quad (5.27)$$

Paramètres de cohérence régionale

En utilisant la conjugaison normale-normale-gamma, on a directement :

$$\begin{aligned} p(\phi_1, 1/v_1^2 | \text{Reste}, \mathbf{D}) &= \mathcal{N} \left[\phi_1 \left| \left(\frac{1}{s_1^2} + \frac{p}{v_1^2} \right)^{-1} \left(\frac{m_1}{s_1^2} + \sum_i \frac{\mu_{i1}}{v_1^2} \right), \left(\frac{1}{s_1^2} + \frac{p}{v_1^2} \right)^{-1} \right. \right] \\ &\quad \times \mathcal{G} \left(1/v_1^2 \left| a_1 + \frac{p}{2}, \frac{1}{2} \sum_i (\mu_{i1} - \phi_1)^2 + b_1 \right. \right) \end{aligned} \quad (5.28)$$

et

$$\begin{aligned} p(\phi_2, 1/v_2^2 | \text{Reste}, \mathbf{D}) &= \mathcal{N} \left[\phi_2 \left| \left(\frac{1}{s_2^2} + \frac{p}{v_2^2} \right)^{-1} \left(\frac{m_2}{s_2^2} + \sum_i \frac{\mu_{i2}}{v_2^2} \right), \left(\frac{1}{s_2^2} + \frac{p}{v_2^2} \right)^{-1} \right. \right] \\ &\quad \times \mathcal{G} \left(1/v_2^2 \left| a_2 + \frac{p}{2}, \frac{1}{2} \sum_i (\mu_{i2} - \phi_2)^2 + b_2 \right. \right). \end{aligned} \quad (5.29)$$

Finalement, en vertu de la conjugaison multinomiale-Dirichlet, il vient :

$$p(\boldsymbol{\pi} | \text{Reste}, \mathbf{D}) = \mathcal{D}_{n-1}(\boldsymbol{\pi} | \xi_1 + n_1, \xi_2 + n_2, \dots, \xi_{n-1} + n_{n-1}), \quad (5.30)$$

où $n_k = \sum_{i=1}^p \mathbb{I}_{\{\tau_i=k\}}$ et où $\mathbb{I}_{\{A\}}$ est la fonction indicatrice pour l'ensemble A .

5.3.3 Application du modèle MM_2 à des séries d'apports naturels pour des sites du nord-est québécois

Nous poursuivons l'analyse de la Section 5.1.2 sur les apports naturels annuels des huit bassins versants en considérant ici le modèle hiérarchique MM_2 . Pour spécifier les lois a priori, nous avons adopté la même stratégie. D'abord, aucune année de rupture n'a été favorisée a priori. Ensuite, la valeur des hyperparamètres associés à σ^2 , ϕ_1 , v_1 , ϕ_2 et v_2

ont été fixés en s'appuyant sur les observations disponibles avant la période concomitante (1957 à 1962).

Pour obtenir un échantillon de la loi jointe a posteriori de l'ensemble des paramètres du modèle MM_2 , l'échantillonnage de Gibbs a été mis en oeuvre à partir des lois conditionnelles complètes déterminées à la Section 5.3.2. Après quelques expérimentations, le nombre de réplicats (taille de l'échantillon) pour l'estimation des densités a posteriori a été fixé à $m = 2500$ avec un temps de chauffe de $t = 2000$ itérations (voir Annexe D pour une brève analyse de la convergence). Signalons que cette procédure a requis la génération de 61 variables à chaque cycle. Au total, 2.745×10^5 quantités ont été simulées dont 1.220×10^5 ont été retenues pour l'estimation. L'exécution des 4500 itérations en MATLAB, sur un PC Pentium 600, a demandé environ 15 minutes.

Spécification des lois a priori

Afin de laisser les données parler d'elles-mêmes en ce qui a trait à l'instant du changement, nous avons pris une loi Dirichlet de paramètres $\xi_1 = \dots = \xi_{n-1} = 1/32$ pour caractériser l'incertitude sur π . En conséquence, la probabilité qu'une rupture survienne une année particulière est la même en moyenne sur toute la période d'échantillonnage. Les valeurs des autres hyperparamètres ont été assignées à partir du vecteur-moyennes $\widehat{\mathbf{y}}$ et de la matrice des covariances $\widehat{\mathbf{S}}_y^2$ reconstitués. Les lois a priori résultantes sont les suivantes :

$$\begin{aligned} 1/\sigma^2 &\sim \mathcal{G}(1/\sigma^2 \mid 8, 160), \\ (\phi_1, 1/v_1) &\sim \mathcal{N}(\phi_1 \mid 25, 2) \mathcal{G}(1/v_1 \mid 4, 8), \\ (\phi_2, 1/v_2) &\sim \mathcal{N}(\phi_2 \mid 25, 12) \mathcal{G}(1/v_2 \mid 4, 8). \end{aligned}$$

L'espérance a priori de $1/\sigma^2$ correspond alors à la moyenne arithmétique de l'inverse des variances contenues dans $\widehat{\mathbf{S}}_y^2$. L'espérance a priori de la moyenne régionale avant le changement, m_1 , est égale à la moyenne du vecteur $\widehat{\mathbf{y}}$. Sa variance est fixée à $s_1 = 2$, ce qui correspond à celle de $\widehat{\mathbf{y}}$. En conformité avec l'application de la Section 5.1.2, nous avons attribué la même espérance à la moyenne régionale après la rupture, mais lui avons affecté une incertitude six fois plus grande. Enfin, les hyperparamètres des lois de $1/v_1$ et $1/v_2$ ont été choisis de manière à ce que l'espérance a priori de ces quantités corresponde à l'inverse de la variance des éléments $\widehat{\mathbf{y}}$.

Résultats

L'échantillonnage de Gibbs produit pour π un échantillon d'une loi Dirichlet à $(n - 1)$ dimensions. Puisque cette distribution est difficile à visualiser, son analyse nécessite l'évaluation de statistiques qui la caractérisent. On peut notamment se concentrer sur ses lois marginales. Sachant que les fonctions de densité marginales d'une loi Dirichlet sont des lois beta, un estimateur « Rao-Blackwellisé » (Annexe D) peut être considéré pour les approcher. L'estimateur de la densité pour un des éléments de π correspond alors à

un mélange de densités beta qui sont évaluées à partir des 2500 générations de Gibbs conservées après le temps de chauffe.

L'espérance a posteriori de π_{1984} (probabilité que la rupture survienne en 1984) est de 0.85. Signalons que toutes les autres années ont une espérance a posteriori inférieure à 0.008. La densité et la fonction de répartition marginales a posteriori pour π_{1984} sont présentées à la Figure 5.8. Cette figure indique, entre autres, qu'il y a au moins 3 chances sur 4 que la probabilité d'un changement en 1984 soit supérieure à 80%. Ces résultats confirment ce qui a été obtenu à la Section 5.1.2 et rejoint l'hypothèse de Rasmussen et al. (1999) voulant qu'une rupture en moyenne ait pris place au milieu des années 80 dans les séries d'apports hydrologiques de la région.

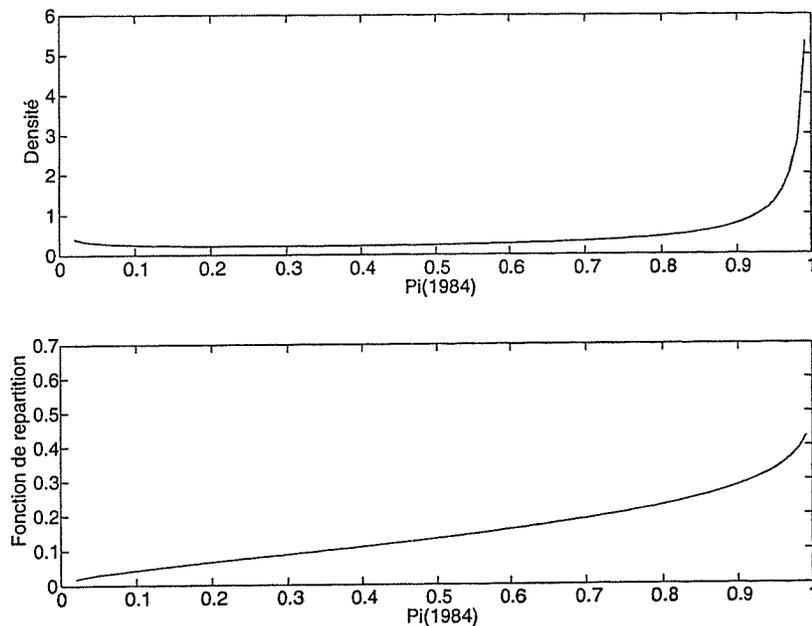


FIG. 5.8 – Densité et fonction de répartition marginales a posteriori de π_{1984} pour le modèle *MM2*

La Figure 5.9 présente les densités a posteriori des moyennes régionales ϕ_1 et ϕ_2 . Leurs espérances sont respectivement 26.7 et 22.2. Le niveau des apports naturels semble donc avoir diminué d'environ $4 \text{ l}/(\text{km}^2 \times \text{s})$ en moyenne sur l'ensemble de la région. L'intervalle de crédibilité à 95% pour l'intensité de la rupture ($[-6.73; -2.20]$) indique qu'une diminution de $6 \text{ l}/(\text{km}^2 \times \text{s})$ demeure crédible. De plus, cet intervalle ne contient pas la valeur 0, ce qui suggère qu'une baisse d'au moins $2 \text{ l}/(\text{km}^2 \times \text{s})$ est vraisemblable.

Comme l'indiquent le Tableau 5.3 et la Figure 5.10, les apports moyens de part et d'autre de l'année de rupture varient tout de même d'un bassin versant à l'autre. La

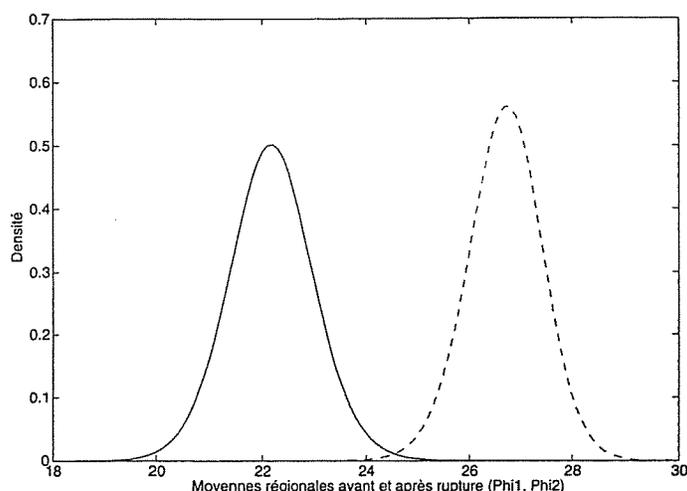


FIG. 5.9 – Densités marginales a posteriori de ϕ_1 (trait pointillé) et ϕ_2 (trait continu) pour MM2

baisse d'hydraulicité maximale est de $7.6 \text{ l}/(\text{km}^2 \times \text{s})$ (bassin de la rivière Romaine) tandis que la baisse minimale est de $3.6 \text{ l}/(\text{km}^2 \times \text{s})$ (bassin de la rivière Caniapiscau). Signalons que les apports moyens estimés ne diffèrent pas substantiellement de ceux obtenus à partir du modèle MM₁. Toutefois, les écart-types a posteriori des moyennes avant et après le changement ont diminué de manière significative pour la plupart des sites (voir Tableau 5.2). Nous pouvons observer d'ailleurs, contrairement à ce qui a été constaté à la Section 5.1.2, que les intervalles de crédibilité (Figure 5.10) ne se chevauchent plus pour les bassins des rivières Manicouagan, Outardes et Ste-Marguerite. Cette baisse d'incertitude par rapport aux résultats issus du modèle MM₁ est particulièrement remarquable lorsqu'on examine la répartition des probabilités a posteriori des années de rupture τ_i pour chacun des sites. Celles-ci sont données à la Figure 5.11. On constate qu'elles sont presque dégénérées, affichant toutes une probabilité voisine de un en 1984.

TAB. 5.3 – Espérances et écarts-types a posteriori des moyennes de part et d'autre de l'année de rupture pour le modèle MM2

		Bassins versants							
		1	2	3	4	5	6	7	8
μ_1	Esp. a posteriori	30.2	23.6	26.4	28.6	27.1	29.9	24.2	25.6
	Éc.-ty. a posteriori	(0.86)	(0.89)	(0.87)	(0.87)	(0.85)	(0.91)	(0.90)	(0.85)
μ_2	Esp. a posteriori	22.6	19.8	20.6	23.8	22.8	25.5	20.6	20.7
	Éc.-ty. a posteriori	(1.09)	(1.19)	(1.11)	(1.18)	(1.09)	(1.37)	(1.14)	(1.12)

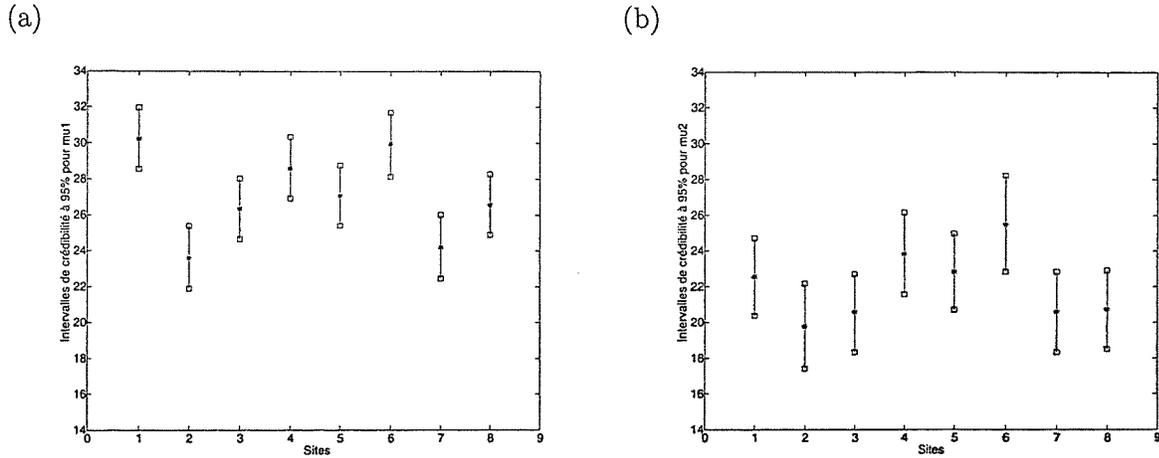


FIG. 5.10 – Intervalles de crédibilité à 95 pour cent des moyennes avant (a) et après (b) la rupture pour huit bassins du nord-est québécois (modèle MM2)

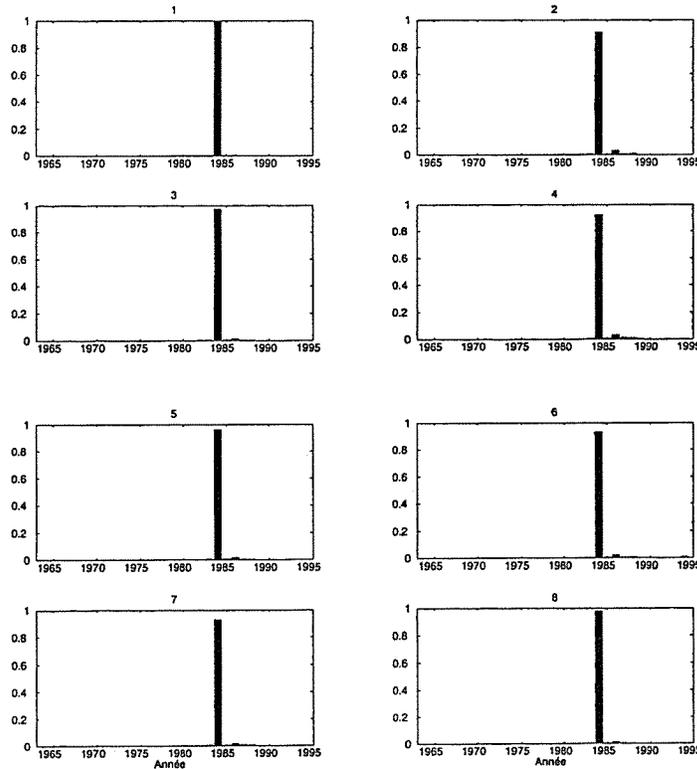


FIG. 5.11 – Densités marginales a posteriori des instants de rupture pour le modèle MM2

Nous soupçonnons ici que l'hypothèse d'indépendance spatiale a pour effet de surestimer la précision des résultats. Comme nous l'avons déjà évoqué, cette hypothèse est irréaliste et nous pouvons douter de la représentativité du modèle MM₂ pour notre jeu de données.

Analyse des résidus

Disposant d'un échantillon pour chaque paramètre du modèle obtenu par l'algorithme de Gibbs, on peut facilement en déduire un échantillon des résidus et réaliser les vérifications classiques du modèle linéaire. La Figure 5.12 présente, pour chaque site, les résidus espérés sur papier de probabilité normal. Ici encore, l'hypothèse de normalité nous semble raisonnable.

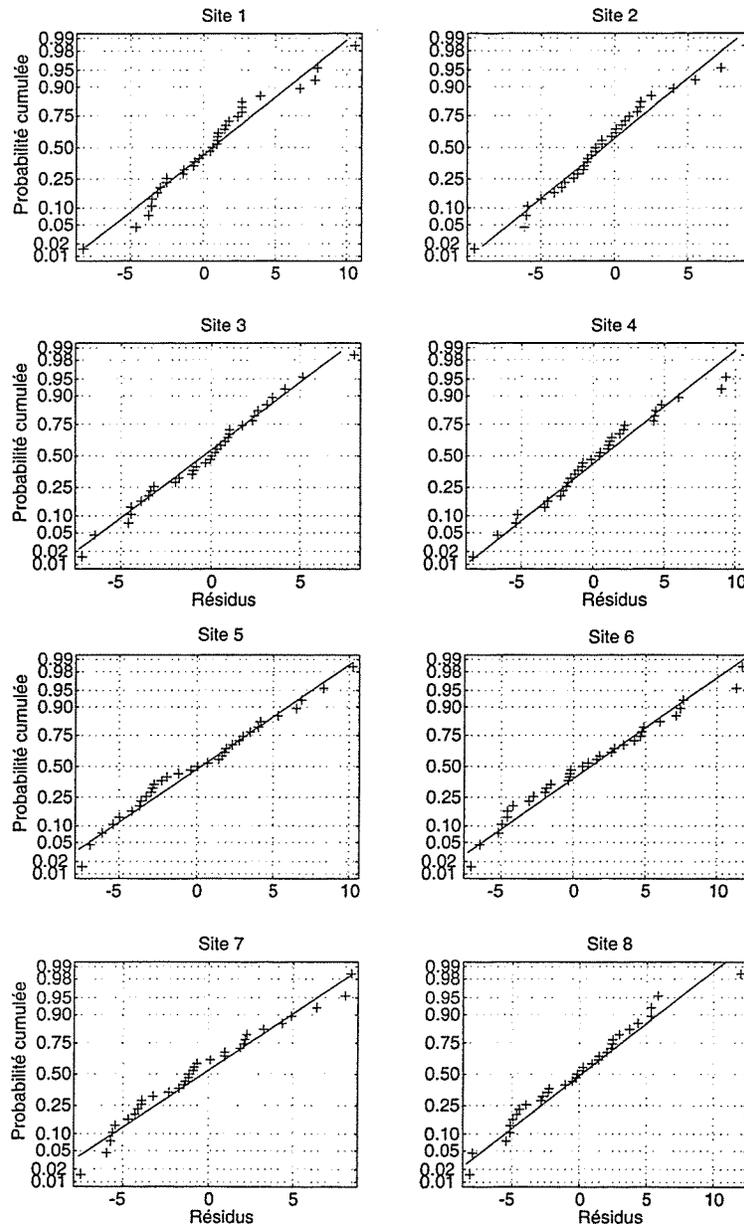


FIG. 5.12 – Résidus espérés du modèle MM2 sur papier de probabilité normal

À chaque itération de l'échantillonnage de Gibbs, la matrice des corrélations spatiales entre les résidus des huit séries d'apports a été calculée et gardée en mémoire. Après avoir éliminé les 2000 premiers cycles (temps de chauffe), nous disposons alors d'un échantillon de 2500 matrices. Ce dernier a été utilisé pour estimer la distribution a posteriori de la matrice des corrélations spatiales des termes d'erreur. Le Tableau 5.4 présente les espérances marginales obtenues ainsi que les intervalles de crédibilité à 95%. On constate en particulier que la valeur zéro n'appartient à aucun des intervalles.

L'analyse des résidus confirme donc la présence d'une structure de corrélation spatiale dans les termes d'erreur du modèle MM_2 . Cela n'est pas étonnant puisque le modèle MM_2 ne prend pas en compte la covariation spatiale. Or, les apports naturels de bassins versants appartenant au même ensemble hydrologiquement homogène sont liés entre eux. Il est donc nécessaire de modifier le modèle MM_2 pour prendre en compte les corrélations entre les apports annuels. C'est ce que nous nous proposons de faire dans la section qui suit, en introduisant dans le modèle une variable latente.

TAB. 5.4 – *Espérances a posteriori et intervalles de de crédibilité à 95 pour cent des corrélations spatiales des résidus du modèle MM_2*

	2	3	4	5	6	7	8
1	0.31 [0.22;0.38]	0.45 [0.35;0.51]	0.37 [0.29;0.42]	0.36 [0.28;0.40]	0.48 [0.38;0.53]	0.11 [0.04;0.18]	0.37 [0.30;0.42]
2		0.78 [0.66;0.83]	0.36 [0.26;0.41]	0.29 [0.21;0.35]	0.31 [0.20;0.37]	0.68 [0.59;0.72]	0.70 [0.60;0.74]
3			0.53 [0.42;0.58]	0.44 [0.35;0.49]	0.47 [0.35;0.52]	0.70 [0.58;0.74]	0.89 [0.77;0.92]
4				0.85 [0.77;0.87]	0.67 [0.58;0.70]	0.49 [0.40;0.53]	0.68 [0.58;0.71]
5					0.68 [0.60;0.71]	0.39 [0.32;0.43]	0.56 [0.48;0.59]
6						0.34 [0.25;0.39]	0.53 [0.43;0.57]
7							0.82 [0.73;0.85]

5.4 Modèle multisite à rupture variable avec effet de covariation spatiale (MM_3)

Contrairement à MM_1 , le modèle hiérarchique MM_2 ne prend pas en compte les corrélations spatiales entre les observations. L'analyse des résidus effectuée à la section précédente est révélatrice à ce sujet. D'ailleurs, ce modèle est très proche de la procédure usuelle,

contestée au Chapitre 2, qui consiste à appliquer indépendamment une série d'analyses locales.

Dans cette section, nous proposons un dernier modèle multisite, noté MM_3 , permettant de prendre en considération une partie de la structure de covariation spatiale. Pour ce faire, nous exploitons l'idée selon laquelle l'apport naturel est le résultat d'une combinaison de divers facteurs climatiques, hydrologiques et physiographiques. On peut très bien imaginer, notamment, que l'apport naturel de l'année j à la station i s'exprime de la manière suivante :

$$y_{ij} = \mu_i + \sum_{k=1}^q \theta_{ik} f_{kj} + \varepsilon_{ij}, \quad i = 1, \dots, p, \quad j = 1, \dots, n \text{ et } q \leq p, \quad (5.31)$$

où μ_i est l'apport moyen de la séquence i , f_1, \dots, f_q constituent un ensemble de facteurs, θ_{ik} correspond à la contribution du facteur k aux apports naturels du bassin i et ε_{ij} est un terme d'erreur. L'expression (5.31) correspond en fait à un modèle en composantes principales, où chaque facteur intègre, sous forme d'une combinaison linéaire particulière, l'ensemble des variables agissant sur le régime hydrologique. Ces facteurs peuvent être vus comme des quantités annuelles globales qui contribuent de manière différente à la réponse hydrologique de chaque bassin, mais qui induisent tout de même une certaine covariation spatiale entre les apports. Il est raisonnable de présumer que le facteur dominant attribue un poids élevé aux variables de précipitation qui expliquent sans aucun doute le plus la présence d'une covariation.

5.4.1 Modèle et notation

Souvent, le premier facteur (ou les toutes premières composantes principales) « explique » l'essentiel de la variabilité totale lorsqu'une analyse en composantes principales est appliquée à un ensemble de séquences d'apports annuels. C'est pourquoi nous ne considérons ici qu'un seul facteur. Aussi, le modèle MM_3 ne considère-t-il qu'une partie de la structure de covariation entre les stations.

Le premier niveau du modèle à effet de covariation suppose que les observations sont représentées par :

$$Y_{ij} \stackrel{iid}{\sim} \mathcal{N}(y_{ij} \mid \mu_{i1} + \theta_i Z_j, \sigma^2), \quad j = 1, \dots, \tau_i$$

$$Y_{ij} \stackrel{iid}{\sim} \mathcal{N}(y_{ij} \mid \mu_{i2} + \theta_i Z_j, \sigma^2), \quad j = \tau_i + 1, \dots, n, \quad i = 1, \dots, p.$$

où $Z_1, \dots, Z_n \stackrel{iid}{\sim} \mathcal{N}(z \mid 0, 1)$ et où θ_i est un nouveau paramètre local qui mesure la contribution de la variable Z_j à l'apport Y_{ij} . Puisque le facteur n'est pas observable, il est représenté ici par une variable aléatoire Z , appelée variable latente ou cachée (Tanner et Wong (1987)). Dans la mesure où le premier facteur d'une analyse en composantes principales est souvent associé au régime pluvial et neigeux, Z_j pourrait être interprétée

ici comme une quantité manquante qui caractérisait la précipitation hivernale de l'année j sur toute la région. Les variables Z_1, \dots, Z_n ont été incorporées au modèle en leur affectant une incertitude par une loi de probabilité. Ici, cette distribution est choisie en vertu de la définition d'une variable aléatoire normale multidimensionnelle : toute variable normale multidimensionnelle peut s'exprimer comme une combinaison linéaire de variables aléatoires normales centrées-réduites.

Dans le second niveau de la hiérarchie, nous conservons les mêmes lois a priori que celles spécifiées pour le modèle MM_2 , mais nous y ajoutons ici une distribution a priori pour les θ_i :

$$\begin{aligned} \tau_i &\stackrel{iid}{\sim} \mathcal{M}(t \mid \pi_1, \dots, \pi_{n-1}, 1), & i = 1, \dots, p, \\ \mu_{i1} &\stackrel{iid}{\sim} \mathcal{N}(\mu_{i1} \mid \phi_1, v_1), & i = 1, \dots, p, \\ \mu_{i2} &\stackrel{iid}{\sim} \mathcal{N}(\mu_{i2} \mid \phi_2, v_2), & i = 1, \dots, p, \\ \theta_i &\sim \mathcal{N}(\theta_i \mid \phi_{\theta_i}, v_{\theta_i}), & i = 1, \dots, p, \\ 1/\sigma^2 &\sim \mathcal{G}(1/\sigma^2 \mid \alpha, \beta). \end{aligned}$$

Nous supposons, une fois de plus, qu'ils sont indépendants et normalement distribués avec comme hyperparamètres de valeurs connues ϕ_{θ_i} et v_{θ_i} . Enfin, le dernier niveau de ce modèle hiérarchique est identique à celui du modèle MM_2 , soit

$$\begin{aligned} (\pi_1, \dots, \pi_{n-1}) &\sim \mathcal{D}_{n-1}(\boldsymbol{\pi} \mid \xi_1, \dots, \xi_{n-1}), \\ (\phi_1, 1/v_1) &\sim \mathcal{N}(\phi_1 \mid m_1, s_1) \mathcal{G}(1/v_1 \mid a_1, b_1), \\ (\phi_2, 1/v_2) &\sim \mathcal{N}(\phi_2 \mid m_2, s_2) \mathcal{G}(1/v_2 \mid a_2, b_2), \end{aligned}$$

La Figure 5.13 présente le diagramme directionnel de la structure de dépendance des variables constitutives du modèle MM_3 . Pour simplifier le schéma, seules les variables des premier et second niveaux y sont représentées. Signalons que le diagramme de la Figure 5.7 est toujours valable puisque le modèle MM_2 est un cas particulier de MM_3 lorsque $Z_1 = \dots = Z_n = 0$. Ici encore, toutes les lois conditionnelles complètes peuvent être obtenues directement en tirant partie de la structure de dépendance entre les paramètres et en exploitant les propriétés des lois conjuguées.

5.4.2 Lois conditionnelles complètes

Les lois conditionnelles complètes des paramètres locaux μ_{i1} , μ_{i2} , $1/\sigma^2$ et τ_i peuvent être déduites exactement de la même manière qu'à la Section 5.2.3. On obtient des expressions analogues à celles obtenues pour le modèle MM_2 :

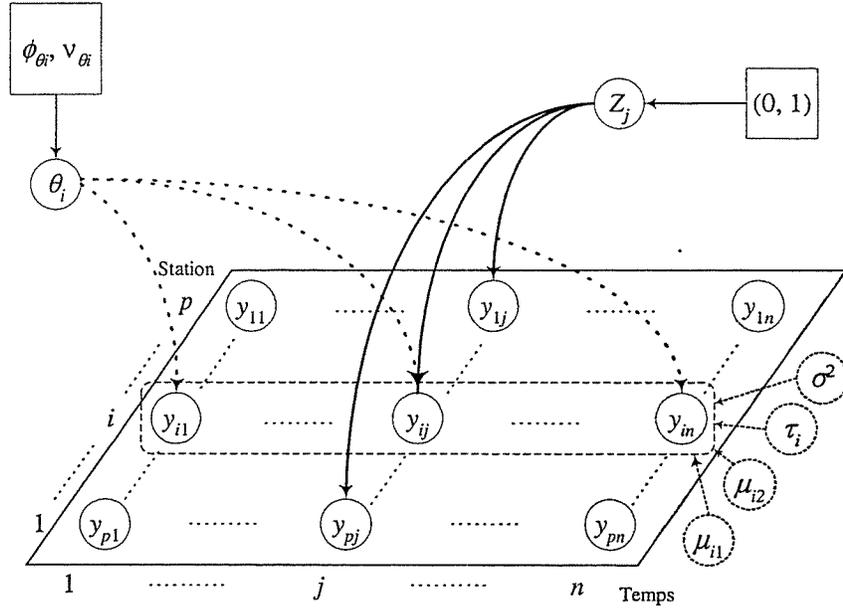


FIG. 5.13 – Représentation schématique du modèle hiérarchique de rupture MM3

$$p(\mu_{i1} | \text{Reste}, \mathbf{Z}, \mathbf{D}) = \mathcal{N} \left[\mu_{i1} \left| \left(\frac{1}{v_1^2} + \frac{\tau_i}{\sigma^2} \right)^{-1} \left(\frac{\phi_1}{v_1^2} + \sum_{j \leq \tau_i} \frac{y_{ij} - \theta_i Z_j}{\sigma^2} \right), \right. \right. \\ \left. \left. \left(\frac{1}{v_1^2} + \frac{\tau_i}{\sigma^2} \right)^{-1} \right], \quad (5.32)$$

$$p(\mu_{i2} | \text{Reste}, \mathbf{Z}, \mathbf{D}) = \mathcal{N} \left[\mu_{i2} \left| \left(\frac{1}{v_2^2} + \frac{n - \tau_i}{\sigma^2} \right)^{-1} \left(\frac{\phi_2}{v_2^2} + \sum_{j > \tau_i} \frac{y_{ij} - \theta_i Z_j}{\sigma^2} \right), \right. \right. \\ \left. \left. \left(\frac{1}{v_2^2} + \frac{n - \tau_i}{\sigma^2} \right)^{-1} \right], \quad (5.33)$$

$$p(1/\sigma^2 | \text{Reste}, \mathbf{Z}, \mathbf{D}) = \mathcal{G} \left(1/\sigma^2 \left| \alpha + \frac{np}{2}, \frac{1}{2} \sum_i \sum_j (y_{ij} - \mu_{ij} - \theta_i Z_j)^2 + \beta \right. \right), \quad (5.34)$$

et

$$\Pr(\tau_i = h | \text{Reste}, \mathbf{Z}, \mathbf{D}) = \frac{\left[\prod_{j=1}^h \mathcal{N}(y_{ij} | \mu_{i1} + \theta_i Z_j, \sigma^2) \prod_{j=h+1}^n \mathcal{N}(y_{ij} | \mu_{i2} + \theta_i Z_j, \sigma^2) \right] \pi_h}{\sum_{k=1}^{n-1} \left[\prod_{j=1}^k \mathcal{N}(y_{ij} | \mu_{i1} + \theta_i Z_j, \sigma^2) \prod_{j=k+1}^n \mathcal{N}(y_{ij} | \mu_{i2} + \theta_i Z_j, \sigma^2) \right] \pi_k} \quad (5.35)$$

où $\mathbf{Z} = (Z_1, \dots, Z_n)$. On retrouve évidemment les lois correspondantes du modèle MM_2 en posant $Z_j = 0$ pour tout j .

À partir de la conjugaison normale-normale, nous obtenons les lois conditionnelles complètes des paramètres de contributions θ_i ($i = 1, \dots, p$) et des variables latentes Z_j ($j = 1, \dots, n$):

$$p(\theta_i | \text{Reste}, \mathbf{Z}, \mathbf{D}) = \mathcal{N} \left[\theta_i \left| \left(\frac{1}{v_{\theta_i}^2} + \frac{1}{\sigma^2} \sum_{j=1}^n z_j^2 \right)^{-1} \left(\frac{\phi_{\theta_i}}{v_{\theta_i}^2} + \frac{1}{\sigma} \sum_{j=1}^n z_j d_{ij} \right), \right. \right. \\ \left. \left. \left(\frac{1}{v_{\theta_i}^2} + \frac{1}{\sigma^2} \sum_{j=1}^n z_j^2 \right)^{-1} \right] \quad (5.36)$$

$$p(Z_j | \text{Reste}, \mathbf{Z}_{(j)}, \mathbf{D}) = \mathcal{N} \left[Z_j \left| \left(1 + \frac{1}{\sigma^2} \sum_{i=1}^p \theta_i^2 \right)^{-1} \frac{1}{\sigma} \sum_{i=1}^p \theta_i d_{ij}, \right. \right. \\ \left. \left. \left(1 + \frac{1}{\sigma^2} \sum_{i=1}^p \theta_i^2 \right)^{-1} \right] \quad (5.37)$$

où $\mathbf{Z}_{(j)} = (Z_1, \dots, Z_{j-1}, Z_{j+1}, \dots, Z_n)$, $d_{ij} = (y_{ij} - \kappa_{ij}) / \sigma$ et

$$\kappa_{ij} = \begin{cases} \mu_{i1}, & \text{si } j \leq \tau_i, \\ \mu_{i2}, & \text{si } j > \tau_i. \end{cases}$$

Finalement, l'ajout de variables latentes n'affecte en rien les lois conditionnelles complètes pour les paramètres de cohérence régionale. Aussi ces dernières demeurent-elles les mêmes que celles obtenues pour le modèle MM_2 : équations (5.28), (5.29) et (5.30) respectivement pour $p(\phi_1, 1/v_1^2 | \text{Reste}, \mathbf{Z}, \mathbf{D})$, $p(\phi_2, 1/v_2^2 | \text{Reste}, \mathbf{Z}, \mathbf{D})$ et $p(\pi | \text{Reste}, \mathbf{Z}, \mathbf{D})$.

5.4.3 Application du modèle MM_3 à des séries d'apports naturels pour des sites du nord-est québécois

L'ajout d'une variable latente au modèle MM_2 a considérablement ralenti la convergence de l'algorithme de Gibbs (voir Annexe D), ce qui nous a obligé à considérer un

temps de chauffe plus grand. Le tracé des valeurs générées (figures D.4 à D.6) indique en effet la présence d'une assez forte dépendance d'une itération à l'autre. Après quelques expérimentations, le temps de chauffe a finalement été fixé à $t = 5000$ itérations. Nous avons toutefois conservé le même nombre de réplicats $m = 2500$ pour l'estimation des quantités a posteriori. L'application du modèle MM_3 a nécessité la génération de 101 variables à chaque cycle (huit paramètres de plus que MM_2 et 32 valeurs de la variable latente Z_j). Au total, 7.757×10^5 quantités ont été simulées dont 2.525×10^5 ont été retenues pour l'estimation. L'exécution des 7500 itérations a demandé moins de 30 minutes.

Spécification des lois a priori

Nous avons considéré pour π , σ^2 , ϕ_1 , ν_1 , ϕ_2 et ν_2 les mêmes lois a priori que celles employées à la Section 5.3.3. Nous avons assigné à l'ensemble des contributions θ_i une espérance nulle, supposant ainsi qu'a priori aucun effet de covariation spatiale n'était présent. Nous avons supposé que leur variance était du même ordre que celle de ϕ_1 , soit égale à 2. Ce choix, plutôt arbitraire faute d'information a priori, n'a pas fait l'objet d'une étude de sensibilité approfondie. Les quelques essais effectués ont montré que les estimations des paramètres d'intérêt sont peu sensibles à de faibles perturbations de la loi a priori des θ_i . Ce choix devrait être soumis à une analyse plus intensive. Il est donc utilisé ici à des fins d'illustration.

Résultats

L'espérance marginale a posteriori maximale de π est obtenue en 1984 avec la valeur 0.75. La seconde plus grande espérance est 0.02 et correspond à l'année 1980. La densité et la fonction de répartition a posteriori pour π_{1984} sont présentées à la Figure 5.14. Elles affichent ici une plus grande incertitude que celles obtenues à partir du modèle MM_2 (Figure 5.8). Il y a, par exemple, moins de chances que la probabilité d'un changement en 1984 soit supérieure à 80% (une sur trois). Cette augmentation de l'incertitude se traduit dans l'ensemble des distributions marginales a posteriori.

Les densités a posteriori des moyennes régionales ϕ_1 et ϕ_2 sont tracées à la Figure 5.15. Elles sont plus diffuses que celles obtenues avec le modèle MM_2 (Figure 5.9) et se chevauchent. Leurs espérances respectives sont toutefois identiques à celles calculées avec MM_2 (26.7 et 22.2). L'intervalle de crédibilité à 95%, $[-6.81; -1.43]$, pour l'amplitude régionale de la rupture est plus large. Il indique que la baisse du niveau des apports sur l'ensemble de la région peut atteindre près de $7 \text{ l}/(\text{km}^2 \times \text{s})$ en moyenne. Cet intervalle ne contient pas la valeur 0 et suggère une baisse d'au moins $1.4 \text{ l}/(\text{km}^2 \times \text{s})$. Au niveau local, la baisse d'hydraulicité est toujours maximale pour le bassin de la rivière Romaine (site 1) et minimale pour Caniapiscou (site 2) (voir Tableau 5.5 et Figure 5.16). Ici encore, les apports moyens estimés sont sensiblement les mêmes que ceux obtenus à partir du

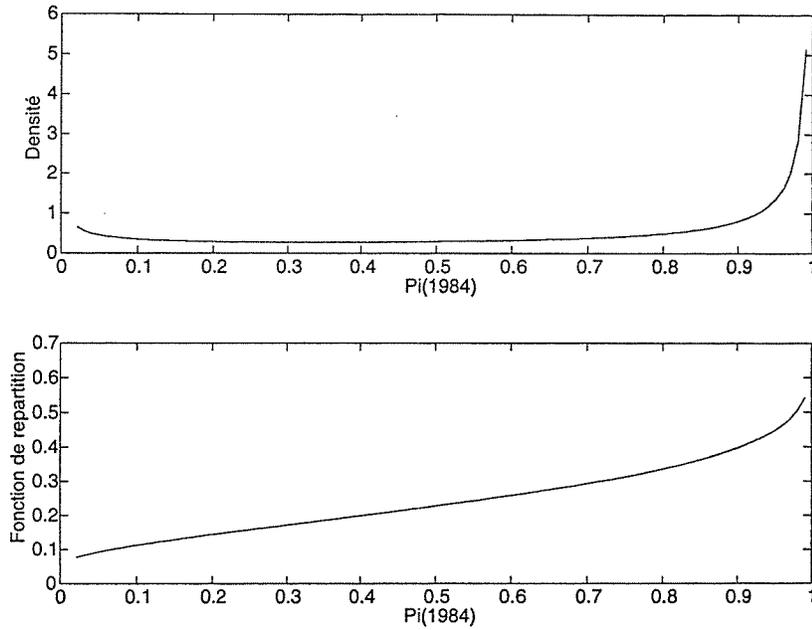


FIG. 5.14 – *Densité et fonction de répartition marginales a posteriori de π_{1984} pour le modèle MM3*

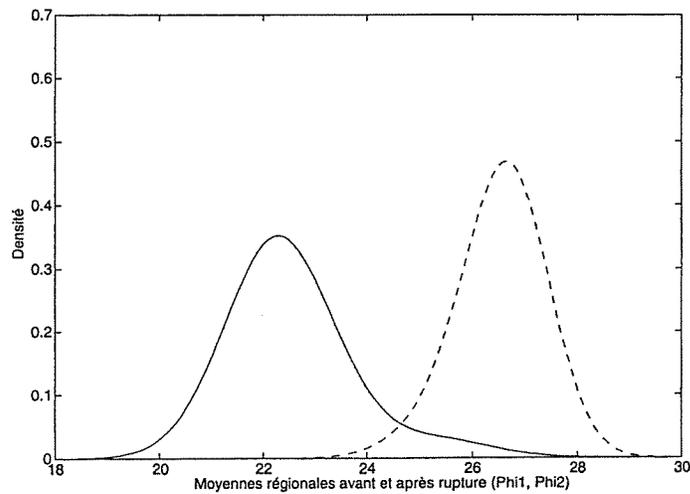


FIG. 5.15 – *Densités marginales a posteriori de ϕ_1 (trait pointillé) et ϕ_2 (trait continu) pour MM3*

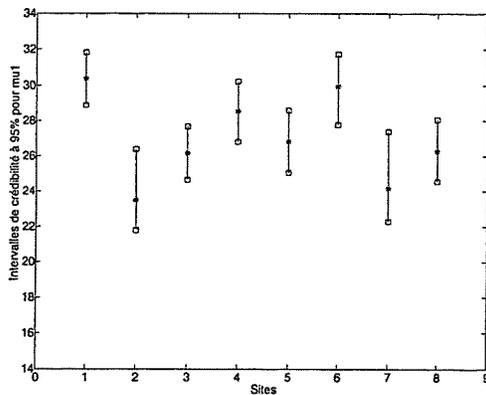
modèle MM_1 . Toutefois, les écart-types et les intervalles de crédibilités sont beaucoup plus proches de ce qui a été obtenu avec le modèle à rupture simultanée MM_1 que ceux

issus du modèle sans covariance spatiale MM_2 .

TAB. 5.5 – *Espérances et écarts-types a posteriori des moyennes de part et d'autre de l'année de rupture pour le modèle MM_3*

		Bassins versants							
		1	2	3	4	5	6	7	8
μ_1	Esp. a posteriori	30.4	23.5	26.1	28.5	26.8	30.0	24.2	26.3
	Éc.-ty. a posteriori	(0.75)	(1.11)	(0.77)	(0.88)	(0.87)	(1.02)	(1.22)	(0.88)
μ_2	Esp. a posteriori	22.8	19.5	20.4	24.6	23.3	26.7	20.6	20.7
	Éc.-ty. a posteriori	(1.06)	(1.21)	(1.12)	(1.32)	(1.31)	(1.49)	(1.18)	(1.30)

(a)



(b)

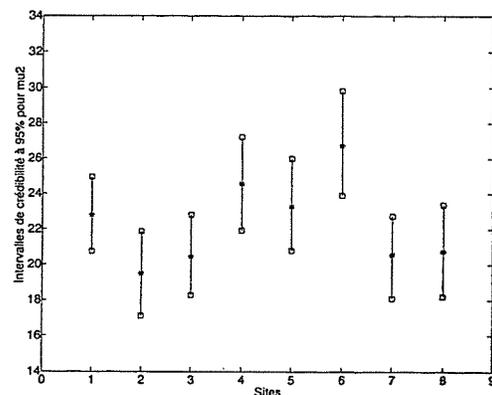


FIG. 5.16 – *Intervalles de crédibilité à 95 pour cent des moyennes avant (a) et après (b) la rupture pour huit bassins du nord-est québécois (modèle MM_3)*

La Figure 5.17 présente la répartition des probabilités des années de rupture τ_i pour chacun des sites. On remarque, en particulier, une plus forte variabilité pour les bassins versants des rivières À la Baleine (site 2) et Caniapiscou (site 7). Rappelons que nous avons déjà émis l'hypothèse que le changement de moyenne dans les apports annuel à Caniapiscou ait pu survenir au début des années 80.

Analyse des résidus

L'introduction d'une variable latente dans le modèle hiérarchique de rupture avait pour objectif de tenir compte de la covariation spatiale révélée par l'analyse des résidus

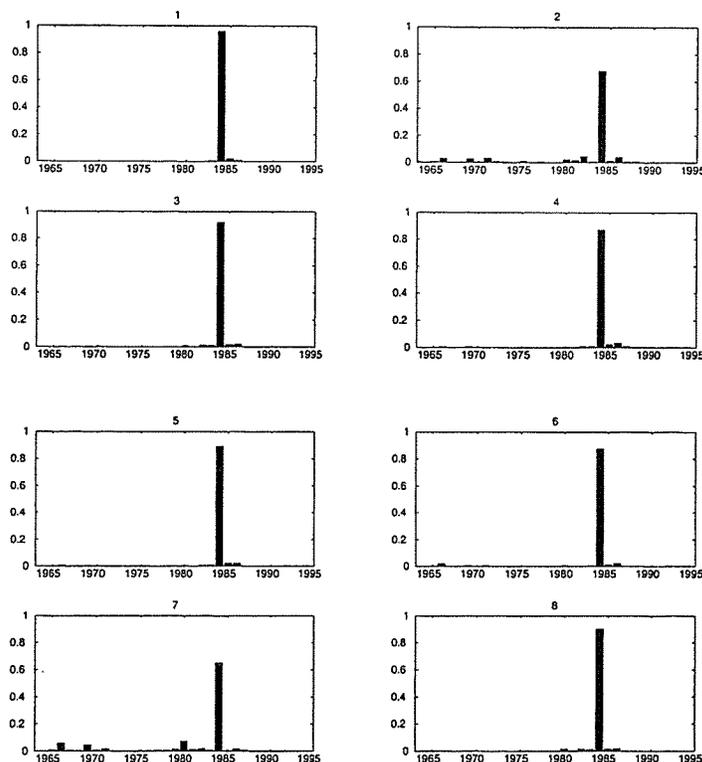


FIG. 5.17 – Densités marginales a posteriori des instants de rupture pour le modèle MM_3

du modèle MM_2 . Pour vérifier si cet ajout au modèle a permis d'extraire une partie de la structure de corrélation, nous avons étudié, comme à la Section 5.3.3, les résidus du modèle MM_3 .

Le Tableau 5.6 présente les espérances des corrélations spatiales des résidus ainsi que les intervalles de crédibilité à 95%. On note, en comparant les résultats obtenus avec ceux du Tableau 5.4, que toutes les corrélations ont diminuées. Pour certaines (près de la moitié), l'intervalle de crédibilité contient la valeur zéro. L'introduction d'une variable latente a donc considérablement amélioré le modèle. Toutefois, certaines corrélations n'ont que très peu diminué. Pour corriger ce problème, on pourrait envisager d'introduire une seconde variable latente au modèle.

Disposant d'un échantillon de 33 résidus pour chacune des séquences, il est intéressant, dans la même foulée, de vérifier l'hypothèse d'indépendance interannuelle. Nous avons donc calculé à chaque itération l'autocorrélation d'ordre 1 des résidus pour chacune des séquences. La Figure 5.18 montre les distributions empiriques et le Tableau 5.7 présente les intervalles de crédibilité à 95% pour l'autocorrélation des résidus. On constate, dans l'ensemble, que la dépendance entre résidus successifs dans le temps est faible sinon inexistante.

TAB. 5.6 – *Espérances a posteriori et intervalles de de crédibilité à 95 pour cent des corrélations spatiales des résidus du modèle MMS*

	2	3	4	5	6	7	8
1	0.00 [-0.17;0.15]	0.13 [-0.07;0.30]	-0.04 [-0.25;0.16]	-0.01 [-0.20;0.15]	0.22 [0.04;0.36]	-0.34 [-0.51;-0.15]	-0.10 [-0.33;0.13]
2		0.58 [0.36;0.72]	-0.29 [-0.51;-0.05]	-0.33 [-0.52;-0.12]	-0.25 [-0.45;-0.06]	0.42 [0.21;0.56]	0.38 [0.15;0.57]
3			-0.21 [-0.47;0.04]	-0.30 [-0.52;-0.06]	-0.19 [-0.42;0.03]	0.36 [0.13;0.53]	0.68 [0.44;0.81]
4				0.62 [0.45;0.74]	0.21 [-0.03;0.40]	-0.10 [-0.34;0.12]	0.02 [-0.25;0.27]
5					0.30 [0.12;0.46]	-0.22 [-0.44;-0.01]	-0.18 [-0.43;0.08]
6						-0.28 [-0.49;-0.08]	-0.19 [-0.44;0.05]
7							0.62 [0.41;0.77]

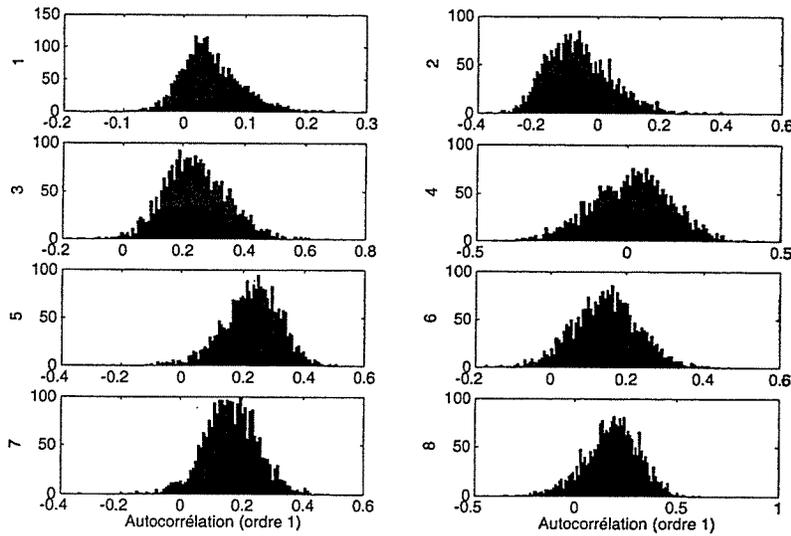


FIG. 5.18 – *Distributions a posteriori de l'autocorrélation d'ordre 1 pour chacun des bassins versants*

Enfin, la Figure 5.19 montre que les résidus espérés sur papier de probabilité normal sont distribués linéairement, ce qui supporte l'hypothèse de normalité. Mentionnons qu'on pourrait très bien appliquer notre méthode à des données transformées si cette dernière hypothèse n'était pas satisfaite.

TAB. 5.7 – Intervalle de crédibilité à 95 pour cent pour le coefficient d'autocorrélation d'ordre 1

1	2	3	4	5	6
[-0.04; 0.16]	[-0.24; 0.18]	[0.05; 0.46]	[-0.25; 0.26]	[0.03; 0.40]	[-0.03; 0.32]
7	8				
[-0.02; 0.34]	[-0.11; 0.43]				

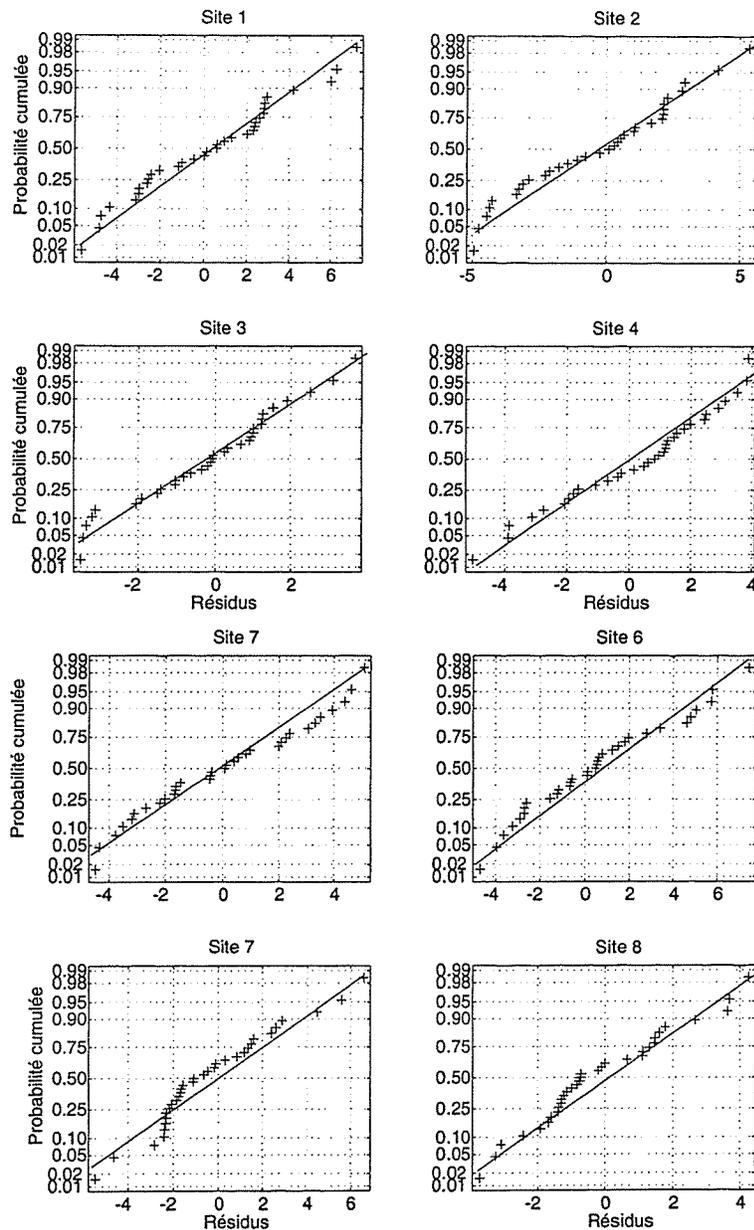


FIG. 5.19 – Résidus espérés du modèle MM3 sur papier de probabilité normal

5.5 Discussion

Dans ce chapitre, nous avons formalisé l'hypothèse courante selon laquelle il devrait exister une certaine cohérence régionale entre les caractéristiques des changements dans une zone homogène. Partant d'abord de l'hypothèse réductrice d'une rupture simultanée, la structure conditionnelle par niveau des modèles hiérarchiques a ensuite été exploitée pour poser des modèles plus réalistes. Ceux-ci laissent la possibilité à l'instant de rupture de varier d'une série à l'autre tout en conservant la cohérence régionale. Au-delà du problème de la non-stationnarité, nous avons montré que la perspective bayésienne et les outils de calcul MCMC permettent le développement et la mise en oeuvre de modèles plus représentatifs du phénomène à l'étude. Selon nous, il n'y a plus de raison de se limiter au développement de procédures « empiriques », qui reposent sur des hypothèses implicites et mal définies. Il est désormais possible, grâce à ces outils, d'intégrer dans un modèle probabiliste les particularités du phénomène d'intérêt.

Le modèle MM_3 constitue l'aboutissement des développements méthodologiques effectués jusqu'ici dans cette thèse. Bien entendu, ce dernier demeure incomplet et peut être amélioré. Les résultats de l'analyse des résidus indiquent, comme on pouvait s'y attendre, que le modèle MM_3 ne mobilise qu'une partie de la structure de corrélation spatiale qui existe entre les observations. Une seconde variable latente devrait donc être ajoutée au modèle. Celle-ci permettrait sans doute de prendre en compte l'essentiel de la covariation spatiale.

Du point de vue pratique, les applications présentées tendent à confirmer la présence d'une baisse d'hydraulicité en 1984 dans les séries d'apports du nord-est du Québec. Ce résultat soulève le problème de la planification d'aménagements futurs dans cette région et devrait attirer l'attention des hydrologues et des gestionnaires. Nous pensons plus particulièrement aux projets hydroélectriques de la rivière Churchill qui nécessiteront notamment la dérivation partielle la rivière Romaine. Rappelons que, selon nos résultats, la baisse d'hydraulicité est maximale pour le bassin de cette rivière. Est-ce qu'Hydro-Québec, chargée de dériver la rivière Romaine, prend en compte la possibilité qu'il y ait eu un changement de moyenne dans les apports? Nous pensons aussi à l'aménagement d'une nouvelle centrale sur la rivière Toulmoustouc située dans le complexe hydroélectrique de Manicouagan. Quel est l'effet des incertitudes portant sur cette baisse d'hydraulicité pour l'aménagement optimal de cette centrale? Nous abordons cette question dans le chapitre qui suit.

Chapitre 6

De l'inférence à la prise de décision : l'aménagement d'une centrale hydroélectrique

Au Chapitre 2, nous avons insisté sur l'importance des éléments décisionnels du traitement statistique de données hydrologiques. Notre question de recherche fait d'ailleurs référence aux décisions relatives à l'aménagement d'une structure hydraulique. Jusqu'ici, nous nous sommes attachés à construire des modèles probabilistes de rupture et à développer les outils permettant de quantifier les incertitudes des paramètres et des hypothèses liées aux modèles. Les méthodes proposées n'ont été illustrées que par des exemples de statistique inférentielle. L'estimation de la date et de l'amplitude du changement (Chapitres 3 et 5) et la vérification de l'existence d'une rupture (Chapitre 4) sont des problèmes types d'inférence abordés dans les chapitres précédents. Dans un contexte strictement climatique, où l'objectif principal est la détection et l'analyse d'un éventuel changement climatique, nos méthodes constituent des outils satisfaisants. Toutefois, puisque l'aménagement d'une structure hydraulique nécessite de prendre des décisions opérationnelles en aval de l'analyse des données, on ne peut se limiter à l'inférence seule. En effet, les conséquences relatives aux actions qui pourraient être entreprises sur la base de nos modèles ne sont pas prises en compte. Aussi la caractérisation et la détection d'une rupture, quoiqu'essentielles, sont-elles insuffisantes pour répondre entièrement à notre question de recherche.

Dans ce chapitre, nous montrons comment l'analyse bayésienne permet d'intégrer des modèles de rupture dans une démarche décisionnelle complète, qui prend en compte non seulement les diverses incertitudes (échantillonnage et modèles), mais aussi les conséquences de l'ensemble des actions possibles. Pour ce faire, nous considérons le problème concret de l'aménagement, par Hydro-Québec, d'une centrale sur la rivière Touloustouc, où une approche économique formelle peut être tentée. Cette illustration, bien qu'issue d'un cas réel, est schématique dans la mesure où seuls les aspects économiques liés à la production énergétique à ce site ont été intégrés. Nous ne tenons pas compte, d'une part, que l'évaluation des conséquences devrait être effectuée en considérant le potentiel de pro-

duction de l'ensemble des sites du parc d'équipement d'Hydro-Québec. Par conséquent, ce cas d'étude peut être vu comme celui de l'aménagement d'une centrale hydroélectrique privée. D'autre part, nous ne prenons pas en considération les impacts environnementaux d'un tel projet et les compensations que la compagnie pourrait être amenée à verser aux autochtones puisque cette centrale sera construite dans un territoire revendiqué par ces derniers.

6.1 Problématique étudiée : la centrale de Toulnostouc

Hydro-Québec envisage la construction d'un nouveau barrage en remblai à environ 14 km en aval du barrage Sainte-Anne (complexe hydroélectrique Manicouagan) situé sur la rivière Toulnostouc, au nord de Baie-Comeau (Figure 6.1). Une centrale hydroélectrique en surface sera construite en rive gauche de la rivière, à 14 km du futur barrage. Sa construction devrait être entreprise en 2002. La centrale de Toulnostouc devrait avoir une capacité de 440 mégawatts.

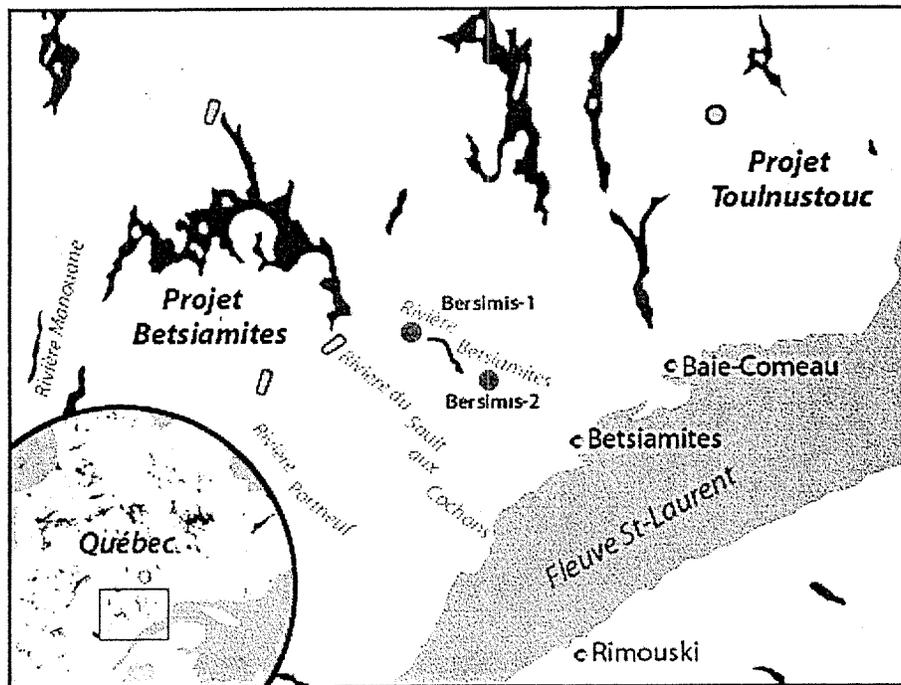


FIG. 6.1 – Situation géographique de la future centrale sur la rivière Toulnostouc

Schématiquement, le problème du dimensionnement d'une centrale hydroélectrique consiste à choisir un débit d'équipement d (la décision) auquel correspond un type d'aménagement. Le gestionnaire doit d'abord supporter un coût total d'investissement, $C(d)$,

que nous supposons indépendant du régime hydrologique. Par contre, les revenus anticipés pour un type d'aménagement donné dépendront non seulement du débit d'équipement, mais aussi de caractéristiques du régime hydrologique réel de la rivière. Le débit annuel moyen, μ , est la quantité qui a le plus d'importance dans la conception et le dimensionnement des divers ouvrages d'un complexe hydroélectrique (Hydro-Québec (1995)). Il est donc naturel de choisir le débit d'équipement d qui maximise les bénéfices $\mathcal{B}(d, \mu) = \mathcal{R}(d, \mu) - \mathcal{C}(d)$, où $\mathcal{R}(d, \mu)$ correspond au revenu anticipé pour un aménagement d avec un débit moyen annuel μ .

L'hypothèse d'un coût fixe indépendant de μ peut être contestée. Toutefois, selon les spécialistes, le coût d'investissement ne varie que très faiblement avec le débit moyen par rapport à l'énergie produite. C'est pourquoi, les aléas n'ont été intégrés que dans la fonction de revenu.

Pour dimensionner la centrale de la rivière Toulmoustouc, nous distinguons donc les trois quantités suivantes :

- Un débit d'équipement d possible. L'ensemble des actions considéré pour cette illustration est $\mathcal{A} = \{d \mid d \in [225 \text{ m}^3/\text{s} ; 360 \text{ m}^3/\text{s}]\}$.
- Le débit moyen annuel μ de la rivière, état de la nature inconnu et incertain. Cette valeur est appelée le module par les hydrologues d'Hydro-Québec. Dans la suite, nous utilisons cette terminologie.
- Le débit d'équipement d^* optimal choisi.

La Figure 6.2 présente les apports naturels annuels au réservoir Sainte-Anne. Ces données constituent l'information de base pour déterminer le débit d'équipement de la centrale. Le module $\hat{\mu}$, estimé sur toute la période d'échantillonnage comme il est d'usage à Hydro-Québec (Hydro-Québec (2000)), est de $186 \text{ m}^3/\text{s}$.

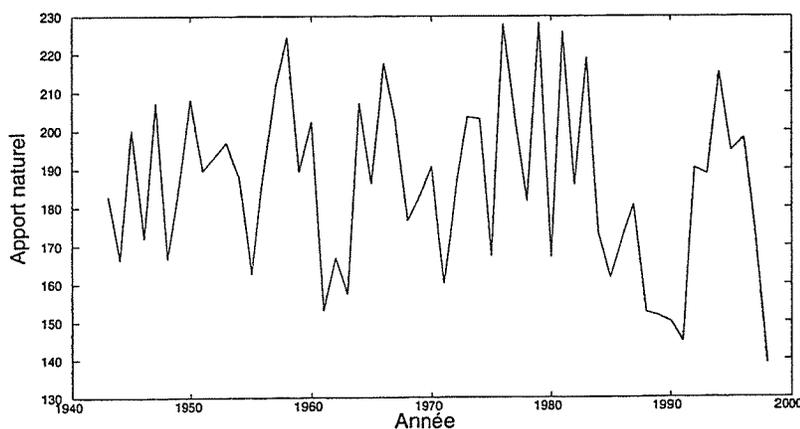


FIG. 6.2 – Apports naturels annuels (m^3/s) observés au réservoir Sainte-Anne

Ce site du nord-est du Québec pourrait admettre une rupture en moyenne au milieu des années 80. Même de faible amplitude, ce changement peut être effectivement présent.

En effet, de nombreux sites dans la région présentent, à peu près la même année, une baisse du niveau moyen de magnitude non négligeable (voir les applications du Chapitre 5).

Dans la suite de ce chapitre, nous nous concentrons d'abord sur la construction des fonctions économiques $C(\cdot)$, $\mathcal{R}(\cdot)$ et $\mathcal{B}(\cdot)$. Ensuite, nous déterminons le débit d'équipement optimal pour l'aménagement de la centrale en tenant compte de la variabilité naturelle du module et de l'éventualité d'une rupture (incertitude structurelle).

6.2 Fonctions économiques

La construction des fonctions de coût, des revenus anticipés et des bénéfices a été réalisée en collaboration avec deux spécialistes d'Hydro-Québec et en s'appuyant sur les données énergétiques et économiques tirées du rapport d'avant-projet Hydro-Québec (2000).

6.2.1 Fonction de coût

La construction d'une centrale pour un débit d'équipement d nécessite d'abord de supporter un coût d'investissement fixe $C(d)$. Compte tenu que les coûts marginaux sont décroissants, il est naturel de supposer que la fonction $C(d)$ soit concave. Dans Hydro-Québec (2000), le coût global de construction pour des débits d'équipement d de 200, 225, 250, 275 et 300 m^3/s a été évalué. De plus, les spécialistes de la société nous ont fourni les coûts relatifs à des débits d'équipement de 337 et 375 m^3/s . Le résultat de ces calculs, effectués sur la base du module estimé de 186 m^3/s , est donné au Tableau 6.1. Les valeurs sont exprimées en dollars constants de 1999.

TAB. 6.1 – Coût global de construction en dollars constants de 1999 pour divers débits d'équipement

d (m^3/s)	200	225	250	275	300	337	375
$C(d)$ (M\$)	526	520	534	541	547	558	569

La Figure 6.3 présente, d'une part, les points expérimentaux et d'autre part le résultat d'un ajustement linéaire obtenu par les moindres carrés et qui a pour équation $C(d) = 453 + 0.31d$. Comme le montre la répartition des points expérimentaux, la relation semble bien linéaire, du moins pour l'ensemble des actions considérées, à savoir $[225 \text{ m}^3/\text{s} ; 360 \text{ m}^3/\text{s}]$.

6.2.2 Fonction des revenus anticipés

A chaque débit d'équipement d correspond un revenu qui n'est pas déterministe puisqu'il est fonction de la production énergétique, elle-même dépendante du module incertain du réservoir Sainte-Anne. On suppose que le propriétaire de la centrale doit livrer à la fois

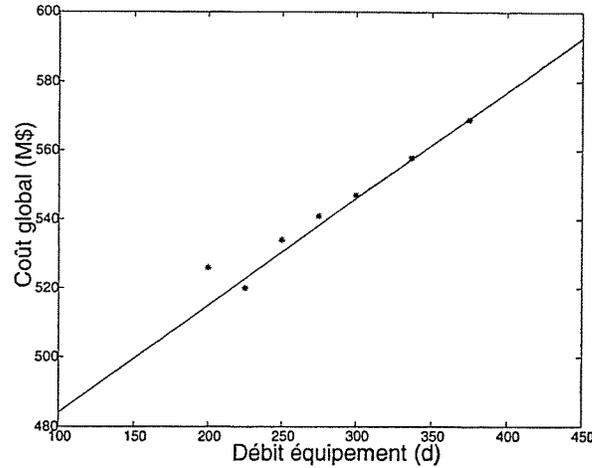


FIG. 6.3 – Fonction de coût pour l'aménagement de la centrale sur la rivière Toulmoustouc

de l'énergie pure et de la puissance. Sur la suggestion des spécialistes d'Hydro-Québec, nous avons adopté une fonction de revenu simplifiée définie par :

$$\begin{aligned}
 \mathcal{R}(d, \mu) &= (\text{Valeur actualisée de l'énergie}) \times \mathcal{E}(d, \mu) \\
 &\quad + (\text{Valeur actualisée de la puissance}) \times \mathcal{E}(d, \mu), \\
 &= (\$150000/\text{GWh}) \times \mathcal{E}(d, \mu) + (\$75000/\text{GWh}) \times \mathcal{E}(d, \mu),
 \end{aligned}
 \tag{6.1}$$

où $\mathcal{E}(d, \mu)$ désigne la production moyenne annuelle pour un module μ et un débit d'équipement d . Cette fonction de revenu, très conservatrice sur le plan commercial, respecte la situation actuelle qui prévaut à Hydro-Québec. La société d'Etat préfère sous-estimer les revenus anticipés de manière à s'assurer que la centrale soit rentable si le marché devient subitement défavorable. La fonction $\mathcal{E}(d, \mu)$ aurait les caractéristiques suivantes :

- Pour un débit d'équipement nul, l'énergie produite est nulle.
- Pour toute valeur de d inférieure à $d_1(\mu) = 0.8\mu$, l'énergie produite serait directement proportionnelle au débit d'équipement, c'est-à-dire $\mathcal{E}(d, \mu) = \alpha_1 d$. En $d_1(\mu)$, l'énergie produite est désignée par $\mathcal{E}_1(\mu)$.
- Pour les débits d'équipement supérieurs à $d_1(\mu)$, la fonction d'énergie serait faiblement concave et atteindrait une valeur maximale $\mathcal{E}_2(\mu)$ en $d_2(\mu)$. Il est raisonnable de considérer que cette portion peut être représentée par une fonction linéaire, mais de pente $\alpha_2 < \alpha_1$.
- Enfin, l'énergie maximale et le débit d'équipement pour lequel elle est atteinte varieraient en toute proportion avec l'intensité du module μ . Plus précisément, quelque soit $\mu_i \neq \mu_j$, on a

$$d_2(\mu_i) = \frac{\mu_i}{\mu_j} d_2(\mu_j), \quad \mathcal{E}_2(\mu_i) = \frac{\mu_i}{\mu_j} \mathcal{E}_2(\mu_j)$$

Le schéma de la Figure 6.4 montre l'allure de la fonction d'énergie retenue pour l'étude.

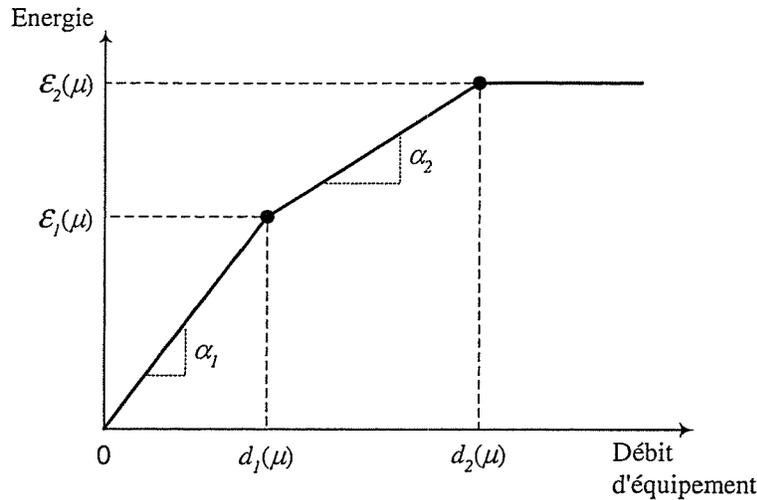


FIG. 6.4 – Schéma de la fonction d'énergie pour l'aménagement d'une centrale

En somme, pour un module μ , les installations ne suffiraient pas à turbiner toute l'eau disponible si le débit d'équipement est inférieur à $d_2(\mu)$. On devrait alors déverser, ce qui se traduirait par une perte d'énergie. Plus le débit d'équipement est faible, plus cette perte est grande. On pourrait qualifier ce seuil $d_2(\mu)$ de débit d'écêtement. Au-delà de cette valeur, la centrale serait suréquipée.

Pour définir la forme finale de la fonction d'énergie, nous disposons d'un ensemble de cinq valeurs de production annuelle relatives à des débits d'équipements de 200, 225, 250, 275, 300 m^3/s et calculées pour le module estimé $\hat{\mu} = 186 m^3/s$. Ces dernières, tirées du rapport d'avant-projet Hydro-Québec (2000), sont présentées au Tableau 6.2 et tracées en fonction des débits d'équipement à la Figure 6.5 .

TAB. 6.2 – Production annuelle calculée pour un module de $186 m^3/s$

d (m^3/s)	$\mathcal{E}(d, \hat{\mu})$ (Gwh)
200	2 352
225	2 364
250	2 376
275	2 377
300	2 379

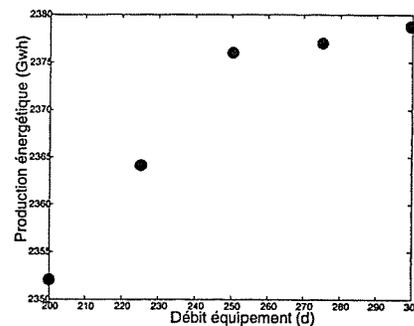


FIG. 6.5 – Production annuelle calculée pour un module de $186 m^3/s$

On note que la valeur maximale est d'environ $\mathcal{E}_2(\hat{\mu}) = 2377$ GWh et est atteinte pour un débit d'équipement $d_2(\hat{\mu}) = 250$ m³/s. Cette information, ainsi que les caractéristiques décrites plus haut, nous permettent de définir entièrement une fonction de revenu, et ce, quelque soit le module μ . La Figure 6.6 présente le tracé des fonction de revenus résultantes pour un ensemble de valeurs de μ .

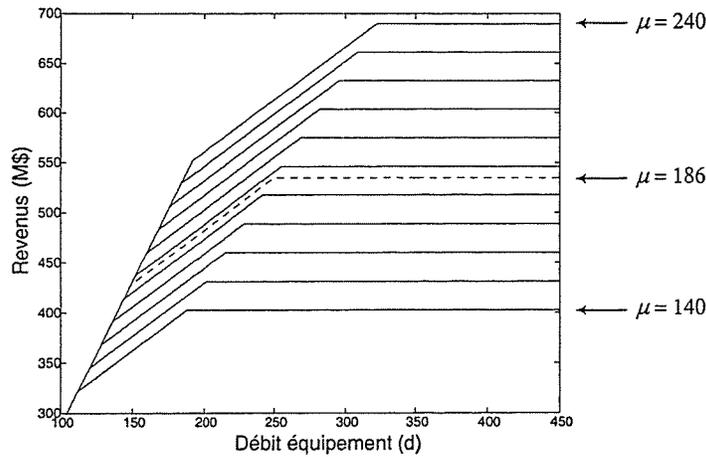


FIG. 6.6 – Fonction de revenu pour l'aménagement de la centrale sur la rivière Toulnoustouc

6.2.3 Fonction des bénéfices et des pertes

La fonction des bénéfices $\mathcal{B}(d, \mu) = \mathcal{R}(d, \mu) - \mathcal{C}(d)$ est aisément obtenue. Elle est tracée à la Figure 6.7a pour diverses valeurs de module μ . On constate que pour plusieurs modules, la centrale est déficitaire. Cela est dû à la fonction de revenu (équation 6.1) qui est conservatrice. Nous considérons donc pour la suite la fonction de perte, $\mathcal{P}(d, \mu) = -\mathcal{B}(d, \mu)$, que nous cherchons à minimiser (Figure 6.7b).

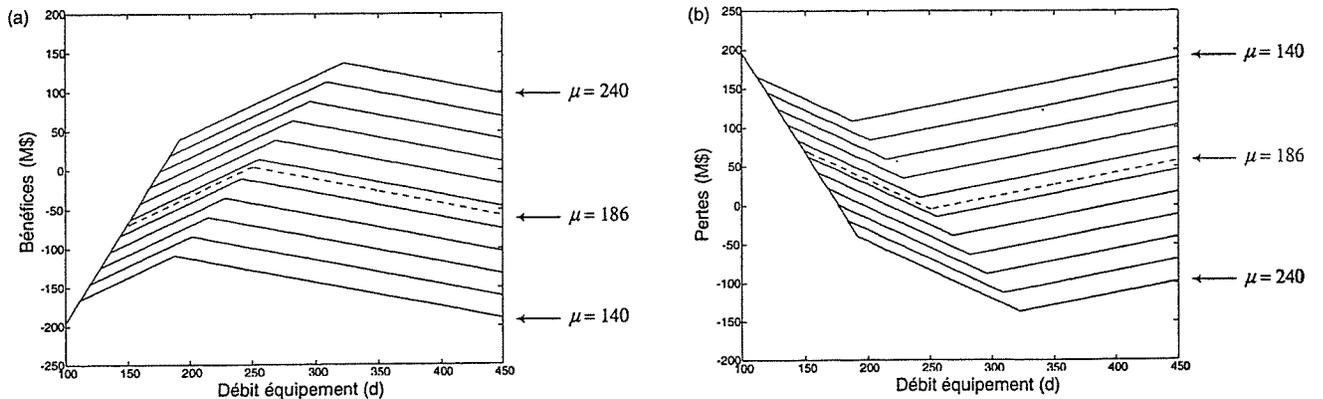


FIG. 6.7 – Fonction de bénéfice et de perte pour l'aménagement de la centrale sur la rivière Toulnoustouc

6.3 Analyse bayésienne complète

En matière d'aménagement de structures hydrauliques, il est d'usage de considérer les problèmes d'inférence et de décision opérationnelle comme censés être de natures différentes. Cette approche séparatrice implique de prendre une décision en considérant les valeurs estimées et les hypothèses comme certaines. Le traitement séparé de ces deux problèmes néglige l'aspect essentiel qui les lie : la propagation des diverses incertitudes hydrologiques de mesure, d'échantillonnage et de modélisation. Il est intéressant d'envisager ces deux problèmes dans un contexte global pour mieux apprécier l'incidence de ces incertitudes sur le choix d'une action. Nous montrons dans cette section comment l'approche bayésienne permet la prise en compte des incertitudes lors du choix d'un débit d'équipement.

6.3.1 Analyse décisionnelle déterministe

Il est clair que le processus d'occurrence des apports annuels Y_1, \dots, Y_n du réservoir Sainte-Anne est un processus aléatoire. À Hydro-Québec, il est d'usage de supposer que les apports annuels sont indépendants et identiquement distribués selon une loi normale. Nous considérons donc le modèle stationnaire M_0 pour représenter les apports annuels du réservoir lac Sainte-Anne :

$$Y_i \stackrel{iid}{\sim} \mathcal{N}(y_i \mid \mu, \sigma^2), \quad i = 1, \dots, n$$

La détermination du débit d'équipement est effectuée ici de manière déterministe, c'est-à-dire en supposant que le modèle M_0 correspond exactement à la réalité et que ses paramètres sont connus sans incertitude. Cette analyse pourrait paraître irréaliste si elle ne correspondait pas à l'approche utilisée dans le rapport Hydro-Québec (2000). Certes, les paramètres μ et σ^2 ne sont pas connus, mais l'hydrologue peut en donner des estimations et on fait ensuite comme si ces valeurs estimées correspondaient aux valeurs exactes. Dans Hydro-Québec (2000), le débit d'équipement est déterminé en supposant que le module réel correspond exactement à $\hat{\mu} = \bar{y} = 186 \text{ m}^3/\text{s}$.

Le paramètre μ étant connu, la décision déterministe optimale d_0 est telle que :

$$\mathcal{P}(d_0, \mu) = \text{Min}_d \mathcal{P}(d, \mu).$$

Pour la fonction de perte retenue dans notre étude, il est clair que le débit d'équipement optimal dans un contexte certain est de $250 \text{ m}^3/\text{s}$ (voir Figure 6.7b). En effet, il correspond au minimum de la fonction de perte pour un module de $186 \text{ m}^3/\text{s}$.

Signalons que dans l'étude d'avant-projet (Hydro-Québec (2000)) on recommande l'emploi d'un débit d'équipement de $225 \text{ m}^3/\text{s}$. Ce résultat a été obtenu en minimisant le coût global d'investissement (voir Tableau 6.1 et Figure 6.3).

6.3.2 Analyse décisionnelle avec prise en compte des incertitudes

Nous considérons que les apports annuels du réservoir Sainte-Anne, Y_1, \dots, Y_n , sont issus de l'un des deux modèles unidimensionnels suivants :

$$\begin{aligned} \text{Modèle } M_0: & \quad Y_i \sim \mathcal{N}(y_i \mid \mu_1, \sigma^2), \quad i = 1, \dots, n \\ \text{Modèle } M_1: & \quad Y_i \sim \mathcal{N}(y_i \mid \mu_1, \sigma^2), \quad i = 1, \dots, \tau \\ & \quad Y_i \sim \mathcal{N}(y_i \mid \mu_2, \sigma^2), \quad i = \tau + 1, \dots, n \end{aligned}$$

Nous avons adopté des probabilités a priori égales pour chaque modèle (poids équivalents aux hypothèses de changement et de non changement) : $p(M_0) = p(M_1) = 1/2$. En ce qui concerne leurs paramètres, nous avons utilisé les mêmes lois a priori qu'à la Section 4.2, soit des distributions de la famille normale-gamma inverse :

$$\begin{aligned} M_0: & \quad p(\theta_0) = \mathcal{N}(\mu_1 \mid \phi_1, \lambda_1 \sigma^2) \mathcal{IG}(\sigma^2 \mid \alpha, \beta), \\ M_1: & \quad p(\theta_1) = \mathcal{N}(\mu_1 \mid \phi_1, \lambda_1 \sigma^2) \mathcal{N}(\mu_2 \mid \phi_2, \lambda_2 \sigma^2) \mathcal{IG}(\sigma^2 \mid \alpha, \beta) p(\tau), \end{aligned}$$

où θ_k désigne l'ensemble des paramètres du modèle M_k , et où une loi uniforme fut considérée pour la répartition annuelle des probabilité a priori de rupture $p(\tau)$. Les hyperparamètres $(\phi_1, \lambda_1, \phi_2, \lambda_2, \alpha, \beta)$ ont été spécifiés sur la base des résultats obtenus à la Section 5.4.3 en appliquant le modèle régional MM_3 aux apports des bassins versants Romaine, À la Baleine, Churchill Falls, Manicouagan, Outardes, Outardes, Ste-Marguerite, Caniapiscou et Ashuanipi. On suppose donc que le réservoir Sainte-Anne appartient au même ensemble hydrologiquement homogène. En ramenant en m^3/s les espérances et les variances régionales a posteriori des moyennes de part et d'autre de l'année de rupture, les valeurs choisies sont telles que :

$$\mathbb{E}\{\mu_1\} = 214, \mathbb{E}\{\mu_2\} = 174, \mathbb{V}\{\mu_1\} = 252, \mathbb{V}\{\mu_2\} = 189, \mathbb{E}\{\sigma^2\} = 592$$

La Figure 6.8 présente, avec les apports annuels en m^3/s , les densités de probabilité a posteriori des paramètres d'intérêt du modèle de rupture M_1 . Le facteur de Bayes qui compare l'hypothèse de changement contre celle du non changement a aussi été calculé. On obtient ici 17.52, ce qui tend à favoriser l'hypothèse d'une rupture. Les probabilités a posteriori calculées pour chacun des deux modèles sont respectivement $p(M_0 \mid \mathbf{y}) = 0.054$ et $p(M_1 \mid \mathbf{y}) = 0.946$.

Le choix d'une action parmi un ensemble d'alternatives nécessite un critère de préférence. Dans la perspective bayésienne, le critère d'analyse des incertitudes est l'espérance mathématique a posteriori de la fonction de perte (Berger (1985), Robert (1992), Bernardo

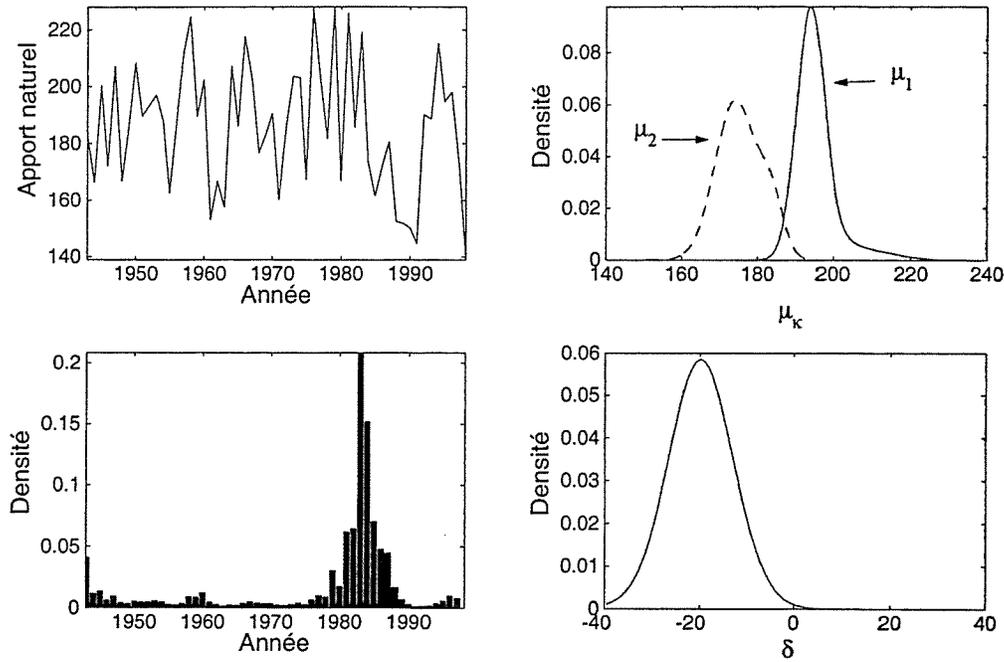


FIG. 6.8 – Apports naturels du réservoir Sainte-Anne et densité marginale a posteriori des paramètres du modèle M_1

et Smith (1994)). Pour les modèles M_0 et M_1 , elles s'expriment respectivement de la manière suivante :

$$\begin{aligned} \mathbb{E}_\mu \{ \mathcal{P}(d, \mu) | \mathbf{y}, M_0 \} &= \int \mathcal{P}(d, \mu_1) p(\mu_1 | \mathbf{y}, M_0) d\mu_1 \\ &= \int \mathcal{P}(d, \mu_1) ST \left(\mu_1 | \phi'_1, \alpha' (\lambda'_1 \beta')^{-1}, 2\alpha' \right) d\mu_1 \end{aligned} \quad (6.2)$$

et

$$\begin{aligned} \mathbb{E}_\mu \{ \mathcal{P}(d, \mu) | \mathbf{y}, M_1 \} &= \int \mathcal{P}(d, \mu_2) p(\mu_2 | \mathbf{y}, M_1) d\mu_2 \\ &= \int \mathcal{P}(d, \mu_2) \left[\sum_{\tau=1}^{n-1} p(\tau | \mathbf{y}, M_1) ST \left(\mu_2 | \phi'_2, \alpha' (\lambda'_2 \beta')^{-1}, 2\alpha' \right) \right] d\mu_2, \end{aligned} \quad (6.3)$$

où $(\phi'_1, \lambda'_1, \phi'_2, \lambda'_2, \alpha', \beta')$ sont les hyperparamètres mis à jour au regard des apports annuels observés au réservoir Sainte-Anne. Les expressions de ces hyperparamètres sont donnés en annexe de l'Article II (Annexe B).

Pour un modèle donné M_k , si donc les incertitudes sur les paramètres sont prises en compte, le gestionnaire peut choisir la décision de Bayes d^* , c'est-à-dire celle qui minimise la perte espérée a posteriori :

$$d^* = \text{Arg Min}_d \mathbb{E}_\mu \{ \mathcal{P}(d, \mu) | y, M_k \}. \quad (6.4)$$

S'il veut de plus considérer l'incertitude structurelle (incertitude de modèle), il choisira (Bernardo et Smith (1994), Chapitre 6)

$$d^* = \text{Arg Min}_d \mathbb{E}_\mu \{ \mathcal{P}(d, \mu) | y \}, \quad (6.5)$$

où

$$\mathbb{E}_\mu \{ \mathcal{P}(d, \mu) | y \} = p(M_0 | y) \mathbb{E}_\mu \{ \mathcal{P}(d, \mu) | y, M_0 \} + p(M_1 | y) \mathbb{E}_\mu \{ \mathcal{P}(d, \mu) | y, M_1 \}. \quad (6.6)$$

La Figure 6.9 représente la perte espérée comme une fonction des débits d'équipement. La courbe continue est celle qui correspond aux résultats obtenus sous l'hypothèse d'un changement de moyenne certain (expression (6.3)). Elle n'intègre donc que l'incertitude des paramètres. La courbe en pointillé considère la possibilité qu'il n'y ait pas de changement et intègre alors l'incertitude structurelle (expression (6.6)).

Les débits d'équipement optimaux d^* ainsi que les pertes encourues sous les hypothèses M_0 , M_1 et $\mathcal{M} = \{M_0, M_1\}$ sont donnés au Tableau 6.3. On y retrouve aussi les pertes relatives au débit d'équipement déterministe optimal ($d_0 = 250 \text{ m}^3/\text{s}$) et à celui qui est recommandé dans Hydro-Québec (2000), c'est-à-dire $225 \text{ m}^3/\text{s}$.

TAB. 6.3 – Débit d'équipement optimal en contexte incertain et pertes encourues pour l'aménagement de la centrale Tounustouc

Modèle(s)	d^* (m^3/s)	$\mathbb{E} \{ \mathcal{P}(d^*, \mu) y \}$ (M\$)	$\mathbb{E} \{ \mathcal{P}(250, \mu) y \}$ (M\$)	$\mathbb{E} \{ \mathcal{P}(225, \mu) y \}$ (M\$)
M_0	257	-11.57	-9.57	9.18
M_1	241	23.90	25.14	29.50
$\mathcal{M} = \{M_0, M_1\}$	242	22.43	23.28	28.40

Remarquons d'abord qu'en ne prenant en compte que l'incertitude des paramètres (et non la possibilité d'une rupture), le débit d'équipement optimal passe de $250 \text{ m}^3/\text{s}$ à $257 \text{ m}^3/\text{s}$ (hypothèse M_0). Si on prend en compte à la fois l'incertitude des paramètres et celle liée à l'existence d'une baisse d'hydraulicité dans les apports historiques du réservoir Sainte-Anne (hypothèse $\mathcal{M} = \{M_0, M_1\}$), le débit d'équipement optimal doit être de $242 \text{ m}^3/\text{s}$, soit de $8 \text{ m}^3/\text{s}$ de moins qu'en contexte certain. En terme de perte espérée, cela se traduit par une économie d'environ $850\,000.00\$$ pour l'aménagement de la centrale de Tounustouc. Par rapport à l'aménagement proposé par Hydro-Québec (2000), la décision de Bayes sous incertitude structurelle permettrait d'économiser près de $6 \text{ M\$}$.

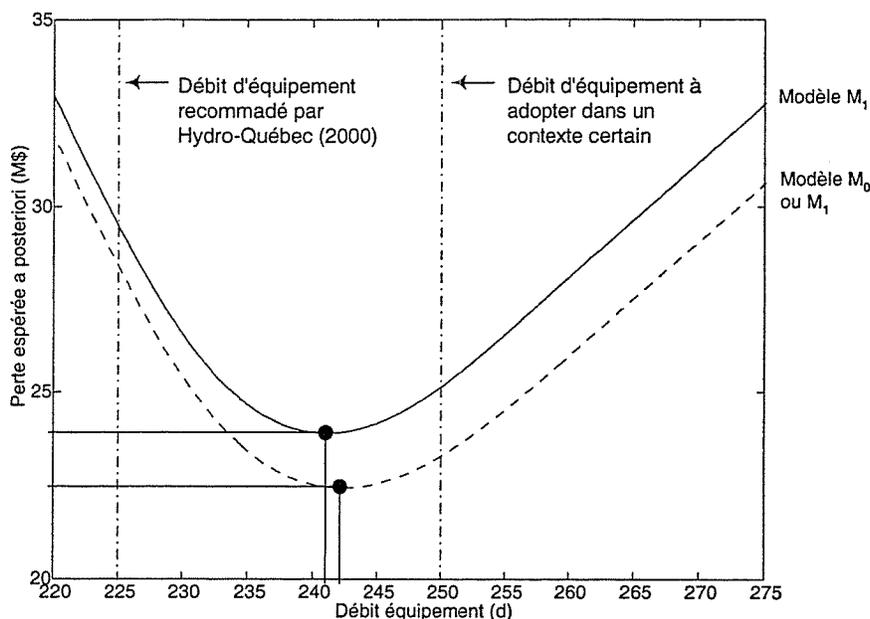


FIG. 6.9 – Perte espérée en fonction du débit d'équipement. Prise en compte de l'incertitude des paramètres sous le modèle M₁ (trait continu). Prise en compte de l'incertitude structurelle (trait pointillé)

6.4 Discussion

Nous avons présenté et développé dans ce chapitre une illustration de l'approche bayésienne complète pour le choix du débit d'équipement d'une centrale hydroélectrique. Elle est complète dans le sens où on n'isole pas la phase inférentielle de la phase décisionnelle. Ce faisant, les incertitudes des paramètres et des hypothèses des modèles ainsi que les conséquences des actions possibles sont prises en compte lors du processus décisionnel.

L'approche usuelle consiste généralement à effectuer une analyse de sensibilité en évaluant la fonction $\mathcal{P}(d, \mu)$, qui traduit les conséquences des décisions, pour différentes valeurs de μ . Par contre, l'approche bayésienne fournit ici une méthode plus systématique pour tenir compte de l'incertitude. En effet, cette démarche dépasse le cadre d'une étude de sensibilité en fournissant automatiquement une valeur de débit d'équipement optimale. Elle est aussi plus rationnelle, puisqu'elle s'appuie sur des fondements décisionnels et sur une base axiomatique qui en assure la cohérence (Savage (1954)).

Il importe toutefois de ne pas perdre de vue que cette approche intégrée introduit des difficultés supplémentaires dans les analyses. En plus des problèmes soulevés aux chapitres précédents par la détermination des lois a priori, l'approche complète impose de plus une évaluation quantitative des conséquences des décisions. C'est souvent un problème complexe nécessitant une bonne communication entre les gestionnaires, les hydrologues et les statisticiens. Il est important de signaler que la méthode bayésienne décisionnelle demeure une application d'un modèle de décision qui, comme tout modèle, présente des

écarts par rapport à la réalité. Elle n'a pour seule ambition de fournir une aide à la décision qui tient compte du contexte incertain avec lequel le gestionnaire doit composer.

Même si la fonction de perte considérée est très simplifiée, l'illustration utilisée dans ce chapitre fournit tout de même à l'hydrologue une démarche logique et intégrée à suivre pour répondre complètement à notre question de recherche. Par ailleurs, puisque la rupture à Toulousteuc peut être un phénomène régional, les résultats soulèvent le problème des prévisions d'aménagements futurs sur l'ensemble du système et devraient attirer l'attention des spécialistes d'Hydro-Québec. Enfin, la prise en compte de l'information régionale à partir d'un des modèles proposés au Chapitre 5 peut augmenter la précision des résultats pour le choix du débit d'équipement de la centrale de Toulousteuc.

Chapitre 7

Conclusions et perspectives

Je me suis intéressé dans cette thèse au développement d'outils statistiques pour quantifier l'incertitude de changement dans la structure aléatoire des séries de débits annuels et à la démonstration de leur intérêt opérationnel dans un contexte décisionnel. Pour ce faire, j'ai privilégié la perspective bayésienne, car elle permet d'intégrer les modèles dans une démarche décisionnelle complète qui prend en compte de manière formelle les conséquences des actions possibles, les connaissances a priori sur le phénomène étudié et les diverses incertitudes.

Dans ce chapitre, nous présentons les principales contributions de cette thèse, nous discutons des limites d'application de nos méthodes et expliquons en quoi certaines pistes de recherche nous apparaissent prometteuses.

Principales contributions

Motivés par l'insuffisance des méthodes employées actuellement par l'hydrologue, je me suis plus particulièrement concentré sur les aspects méthodologiques du traitement statistique de la non-stationnarité des séries d'observations hydrologiques. Les avancées qui en sont issues constituent les points forts de ce travail. D'abord, les principales contributions proviennent de la formalisation, sous forme de modèles probabilistes, de certaines caractéristiques des phénomènes de rupture observées dans les séries hydrologiques. Ensuite, elles reposent sur la conduite de l'analyse d'un changement depuis sa caractérisation jusqu'au choix opérationnel de décision.

L'hydrologue gagne de nouveaux outils d'inférence À l'examen de plusieurs séries hydrologiques, nous avons constaté que certaines d'entre elles semblaient admettre une augmentation ou une diminution dans leur variance. La variabilité des débits annuels joue un rôle fondamental pour la gestion des réserves hydrauliques. Or, dans la littérature hydrologique, seul le cas d'une rupture de moyenne a été considéré. Nous avons donc proposé des solutions au problème de rupture en variance à une date inconnue (modèles M_2 et M_3).

Ensuite, trois modèles multidimensionnels d'analyse d'une rupture en moyenne ont été développés. Ces modèles sont le résultat d'une réflexion sur les facteurs susceptibles d'engendrer une rupture régionale. Nous nous sommes attachés à formaliser le fait qu'un changement climatique risque d'avoir un impact régional en présupposant qu'il existe une certaine cohérence entre les caractéristiques d'un changement pour des sites appartenant à un même ensemble hydrologiquement homogène. Comme nous l'avons déjà évoquée, cette hypothèse est implicite dans les analyses régionales usuelles. Le premier modèle (modèle MM_1) suppose que le changement survient simultanément pour l'ensemble des séries. Les deux autres modèles multisites (modèle MM_2 sans effet de covariation spatiale et modèle MM_3 avec effet de covariation spatiale) laissent la liberté à l'instant de rupture de varier d'une station à l'autre tout en préservant une certaine cohérence régionale. Ces derniers, plus réalistes et donc plus complexes, ont pu être envisagés grâce à la disponibilité des outils de simulation MCMC, sans quoi il aurait été difficile voire impossible de les mettre en oeuvre. En effet, leur structure conditionnelle par niveau permet l'emploi de l'échantillonnage de Gibbs pour estimer les paramètres.

Par ailleurs, nous avons mis au point une approche formelle pour la détection d'une rupture en nous plaçant dans le cadre de la sélection bayésienne de modèles. Cette perspective permet également de considérer plusieurs hypothèses concurrentes lors de l'inférence, comme cela a été illustré pour la prévision d'une nouvelle observation (Article II, Annexe B). Cette approche est aussi originale en sciences de l'eau. Elle comble une lacune importante en hydrologie statistique, car jusqu'à maintenant les travaux se sont limités uniquement à l'estimation des paramètres (date et intensité du changement) conditionnellement à l'existence d'un changement.

Certaines de ces contributions ont été obtenues en utilisant ou en adaptant des techniques récentes proposées dans la littérature statistique. D'autres, par contre, constituent aussi à mon avis de réelles avancées dans des domaines plus théoriques de la statistique appliquée. C'est le cas plus particulièrement de la mise en place du modèle hiérarchique à rupture variable avec effet de covariation spatiale (modèle MM_3).

Le régime hydrologique du Québec change-t-il? Les applications présentées dans cette thèse avaient à l'origine pour seule ambition d'illustrer la mise en oeuvre de nos méthodes. Les résultats obtenus permettent néanmoins d'attirer l'attention des hydrologues sur la possibilité qu'un changement soit survenu au Québec et donc sur la nécessité de s'adapter à ce nouveau contexte. D'abord, des ruptures en moyenne et en variance dans les séries d'apports énergétiques de certains complexes d'Hydro-Québec ont été mises en évidence. Ensuite, l'emploi des modèles multisites nous a permis d'identifier une rupture régionale dans la moyenne des apports naturels de bassins versants du nord-est québécois. Dans cette région, il semble qu'une baisse d'hydraulicité non négligeable ait pris place au milieu des années 80. Ce résultat appuie l'hypothèse avancée par certains climatologues selon laquelle cette zone aurait été touchée par l'augmentation des oscillations climatiques liées aux circulations océaniques de l'Atlantique Nord, augmentation observée au début des années 80.

Aide à la décision en ingénierie hydraulique La baisse d'hydraulicité suspectée dans la région du nord-est du Québec peut avoir un impact décisionnel. Nous avons donc étudié le problème du choix d'un débit d'équipement pour l'aménagement d'une centrale hydroélectrique sur la rivière Toulnostouc. Cette rivière déverse le réservoir Sainte-Anne situé précisément dans cette région. Il a été montré par cette étude de cas que les méthodes proposées pouvaient être employées dans un tel contexte opérationnel. Ainsi, contrairement à la plupart des auteurs dans le domaine, on ne s'est pas limité dans les applications à la seule phase inférentielle.

On peut donc conclure que les travaux réalisés dans cette thèse répondent à la question de recherche formulée au Chapitre 2 :

Comment peut-on prendre en compte l'incertitude de changements dans la structure aléatoire des séries de débits annuels lors de la prise de décision concernant un aménagement hydroélectrique ?

Ces travaux apportent en effet plusieurs éléments de réponse aux préoccupations des gestionnaires concernant l'aménagement d'une structure hydroélectrique en situation risquée. D'abord, les hydrologues disposent maintenant d'outils statistiques mieux adaptés pour quantifier les diverses incertitudes inhérentes à l'analyse d'un changement de régime dans les séries hydrologiques annuelles. Ensuite, ils bénéficient désormais d'une démarche opérationnelle intégrée qui permet de choisir une action sur la base de nos modèles. Cette démarche prend en compte à la fois les conséquences des décisions et l'incertitude structurelle. Une telle approche rationnelle et complète est nécessaire à l'établissement de règles de gestion des systèmes hydriques et environnementaux.

Limites des méthodes proposées

Un modèle n'est rien d'autre qu'une représentation simplifiée de la réalité. Il repose sur des hypothèses plus ou moins fortes. Aussi les méthodes développées dans cette thèse ont-elle une portée limitée qu'il convient de discuter.

Notre démarche est rétrospective et focalisée sur une rupture unique Les modèles développés dans cette thèse ne sont pas applicables pour la détection d'une rupture en temps réel. Dans l'état actuel, ces méthodes sont rétrospectives et ne permettent pas, sans quelques remaniements, d'effectuer une analyse séquentielle de la non stationnarité. De plus, les modèles proposés supposent un seul changement de régime dans la période d'échantillonnage et aucun retour possible à l'état initial. Ainsi, on ne peut anticiper l'occurrence d'une future rupture, et à ce titre, nos méthodes n'ont qu'une portée réduite en matière de prévision.

Et si les changements étaient plutôt graduels? Il est clair que les modèles ne sont pas appropriés si l'hydrologue pense que le changement est graduel à l'échelle de la période d'échantillonnage. J'ai préféré considérer une rupture plutôt qu'une tendance, l'examen des séries d'apports observés au Québec ayant justifié principalement ce choix a priori. Cependant, les modèles proposés pourraient être aisément adaptés si on considère qu'avant un instant inconnu les observations sont stationnaires, et que celles-ci évoluent par la suite selon un modèle linéaire. La stratégie générale d'estimation adoptée dans cette thèse pourrait alors être appliquée sans modification majeure.

Les hypothèses usuelles d'indépendance et de normalité : faut-il en sortir? Les modèles proposés peuvent être employés seulement pour des observations indépendantes et distribuées selon une loi normale. L'hydrologie, telle qu'elle s'est développée au cours des dernières décennies, admet généralement ces hypothèses pour des variables hydro-météorologiques annuelles. Toutefois, pour d'autres types de caractéristiques du régime hydrologique, ces postulats peuvent s'avérer irréalistes.

La présence d'une autocorrélation pourrait être prise en compte aisément dans le modèle hiérarchique MM_3 . En effet, la structure de ce modèle permet l'ajout d'une composante sans hypothéquer la procédure d'estimation. On peut envisager, par exemple, que les résidus du modèle actuel obéissent à un modèle autorégressif (AR). Avec des lois a priori conjuguées, les distributions conditionnelles complètes associées aux paramètres du modèle AR seraient en effet disponibles (Chib (1993), Girard (2000)).

Compte tenu de l'hypothèse de normalité, l'hydrologue ne peut utiliser les procédures développées ici pour l'étude des changements dans les séquences de débits extrêmes (crues et étiages). On pourrait éventuellement appliquer nos modèles sur les données transformées. Par contre, il serait bien plus efficace d'adapter les méthodes en considérant une loi de probabilité des valeurs extrêmes (Gumbel, Fréchet ou Weibull). Comme nous l'avons évoqué à la Section 3.1.2, il faudrait alors avoir recours à un algorithme MCMC hybride qui combine les échantillonnages de Gibbs et de Metropolis-Hasting.

L'établissement des lois a priori : un problème complexe Quelques questions ont été soulevées tout au long de ce document concernant la difficulté de spécifier les lois a priori. Pour faciliter la mise en oeuvre de nos méthodes, j'ai adopté des lois a priori conjuguées. Ce choix est critiquable du fait qu'il obéit à des contraintes techniques plutôt qu'à des impératifs d'adéquation à l'information a priori (voir notamment Robert (1996)). Il se pourrait très bien que les lois a priori résultantes ne soient pas assez souples pour bien représenter les connaissances de l'hydrologue. Actuellement, la statistique bayésienne est peu employée en hydrologie. La transposition des connaissances hydrologiques en termes de probabilité est donc un problème sur lequel très peu d'auteurs ont réfléchi. C'est pourquoi, je me suis limité d'une part à ce type de distributions, et d'autre part, à la détermination de leurs hyperparamètres à partir d'observations régionales ou historiques. Nos approches dérogent donc quelque peu d'une analyse bayésienne « pure » qui s'appuie sur la transposition de connaissances subjectives. Pour traduire en terme de loi de probabilité a priori les connaissances subjectives d'un hydrologue sur le phénomène

étudié, il serait souhaitable de travailler sur les quantiles de la distribution prédictive des observations plutôt que sur les paramètres de la loi a priori. En effet, l'hydrologue est familier avec les quantiles (dans la terminologie hydrologique, la valeur associée à une période de retour donnée), il les emploie régulièrement pour l'étude des risques associés à la construction de structures de protection contre les crues. Pour ce faire, on pourrait s'inspirer des travaux récents de Coles et Tawn (1996), O'Hagan (1998), Kadane et Wolfson (1998) et Al-Awadhi et Garthwaite (1998).

La méthodologie est transposable

L'analyse de ruptures dans des séquences de variables aléatoires a de multiples applications, et les méthodes développées dans cette thèse peuvent être transposées, notamment en contrôle de la qualité, en traitement du signal, en santé, en économie et en finances.

Contrôle de la qualité Les techniques de surveillance et de maîtrise de la qualité font en général appel aux détecteurs séquentiels de changements. Ces outils sont utilisés pour détecter un dérèglement dans un contexte de production. L'estimation de l'instant et de l'amplitude d'un changement est donc d'un grand intérêt pour le responsable de la qualité.

Traitement du signal Le problème de détection de changements dans les caractéristiques d'un signal est courant en traitement d'image, en médecine, en géophysique ou encore en sismologie.

Santé Dans le domaine de la santé, on rencontre de nombreux problèmes de rupture, principalement dans des situations de surveillance au cours du temps (études longitudinales). Nous avons déjà évoqué à ce propos l'étude de Bélisle et al. (1998) sur les trains de décharges neuronales.

Économie et finances La littérature sur les changements structuraux en économétrie est abondante. Évidemment, la détection d'une rupture en économie et en finance trouve de nombreuses applications : changement de tendance du marché, crash boursier, etc.

Les perspectives en hydrologie-statistique

Les limitations évoquées plus haut constituent autant de pistes de recherche à explorer. Le développement d'une procédure de détection d'une rupture en temps réel dans les séquences de débits extrêmes est sans doute le problème le plus urgent et l'avenue de recherche la plus riche.

Par ailleurs, la diversité des modèles étudiés dans cette thèse a montré la grande souplesse de la démarche. Celle-ci n'est pas étrangère au fait que l'on ait adopté le paradigme bayésien. Les perspectives d'application prometteuses qui découlent des travaux réalisés dans cette thèse sont donc nombreuses et variées. Je retiens deux champs d'applications

où plus particulièrement les modèles hiérarchiques à variables latentes sont susceptibles de fournir, en hydrologie-statistique, des solutions plus appropriées que les approches actuelles : la régionalisation des crues et la reconstitution de données. Ces modèles permettraient en effet de formaliser la cohérence régionale qui est implicite dans la démarche que privilégient les hydrologues pour aborder ces deux problèmes.

De plus, je crois que les modèles hiérarchiques combinés à une fonction de coût seraient plus à même de traiter la question de la rationalisation des réseaux de mesures. Les méthodes traditionnelles s'en tiennent uniquement à l'inférence (voir par exemple Ouarda et al. (1996)) alors que l'élimination de stations de mesures est un problème de choix qui doit faire intervenir une fonction d'utilité. En théorie de la décision, l'évaluation de la valeur de l'information sous incertitude répond précisément à ce type de préoccupation opérationnelle. Une méthodologie très intéressante, susceptible d'apporter des solutions au problème de rationalisation des réseaux, a d'ailleurs été proposée récemment par Fortin (1997).

Finalement, si enfin la statistique bayésienne décisionnelle finit par s'imposer en hydrologie comme dans certains domaines d'application de la statistique, il serait primordial de pousser plus loin nos réflexions sur les lois a priori et les fonctions de pertes. Une transposition efficace en terme de loi de probabilité a priori et de fonction de perte des connaissances de l'ingénieur-hydrologue et du gestionnaire permettrait d'exploiter au mieux les atouts de l'analyse bayésienne. Ces problèmes ne sont pas insolubles à condition qu'un dialogue étroit soit établi entre le statisticien et les différents intervenants. Si on réussit à mieux mobiliser toute l'information disponible, qu'elle soit subjective ou objective, la précision et la crédibilité des résultats s'en trouvera alors accrue.

Bibliographie

- Al-Awadhi, S. et Garthwaite, P. H. (1998). An elicitation method for multivariate normal distributions, *Communication in Statistics - Theory Methods* **27**: 1123–1142.
- Barnett, T., Hasselmann, K., Chelliah, M., Delworth, T., Hegerl, G., Jones, P., Rasmusson, E., Roeckner, E., Ropelewski, C., Santer, B. et Tett, S. (1999). Detection and attribution of recent climate change: A status report, *Bulletin of the American Meteorological Society* **80**: 2631–2659.
- Barry, D. et Hartigan, J. A. (1993). A Bayesian analysis for change point problems, *Journal of the American Statistical Association* **88**: 309–319.
- Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York.
- Berger, J. O. (1999). Bayesian analysis: A look at today and thoughts of tomorrow, *Technical Report 99-30*, ISDS Discussion Paper, Duke University.
- Berger, J. et Pericchi, L. R. (1996). The intrinsic Bayes factor for model selection and prediction, *Journal of the American Statistical Association* **91**: 109–122.
- Bernardo, J. M. et Smith, A. F. M. (1994). *Bayesian Theory*, Wiley, New York.
- Bernier, J. (1990). Les incertitudes hydrologiques dans les problèmes de dimensionnement d'ouvrages. Valeur des informations locales et spatiales, *Revue des Sciences de l'Eau* **3**: 37–53.
- Bernier, J. (1994a). Quantitative analysis of uncertainties in water resources: Application for predicting the effects of changes, in *Engineering Risk in Natural Resources Management, Nato ASI Series E*, L. Duckstein et E. Parent (eds), Kluwer Academic Publishers, pp. 343–357.
- Bernier, J. (1994b). Statistical detection of changes in geophysical series, in *Engineering Risk in Natural Resources Management, Nato ASI Series E*, L. Duckstein et E. Parent (eds), Kluwer Academic Publishers, pp. 159–176.
- Bernier, J., Parent, E. et Boreux, J.-J. (2000). *Statistique de l'Environnement. Traitement Bayésien de l'Incertitude*, Lavoisier, Paris.
- Blanchet, M., Bouzid, M., Delhome, R., Ouakam, J.-B. et Pietri, M. (1971). Phénomène de rupture dans une série chronologique, *Revue de la Statistique Appliquée* **XIX**: 5–22.
- Bélisle, P., Joseph, L., MacGibbon, B., Wolfson, D. B. et Berger, R. (1998). Change-point analysis of neuron spike train data, *Biometrics* **54**: 113–123.
- Bobée, B. et Ashkar, F. (1991). *The Gamma Family and Derived Distributions Applied in Hydrology*, Water Resources Publications, Littleton, Colorado.

- Booth, N. et Smith, A. F. M. (1982). A Bayesian approach to retrospective identification of change-points, *Journal of Econometrics* **19**: 7–22.
- Box, G. et Tiao, G. T. (1964). A Bayesian approach to the importance assumptions applied to the comparison of variances, *Biometrika* **51**: 153–167.
- Broemeling, L. D. (1985). *Bayesian Analysis of Linear Models*, Marcel Dekker, New York.
- Brooks, S. (1998). Markov chain Monte-Carlo method and its application, *The Statistician* **47**: 69–100.
- Bruneau, P. et Rassam, J.-C. (1983). Application d'un modèle bayésien de détection de changements de moyennes dans une série, *Journal of Hydrological Sciences* **28**: 341–354.
- Carbonnel, J.-P. et Hubert, P. (1994). Études statistiques de quelques séries pluviométriques roumaines et bulgares. Incidence sur l'évolution climatique récente, *Romanian Journal of Hydrology and Water Resources* **1**: 107–113.
- Carlin, B., Gelfand, A. E. et Smith, A. F. M. (1992). Hierarchical Bayesian analysis of changepoint problems, *Applied Statistics* **41**: 389–405.
- Casella, G. et George, E. I. (1992). Explaining the Gibbs sampler, *The American Statistician* **46**: 167–174.
- Chao, P. et Hobbs, B. F. (1997). Decision analysis of shoreline protection under climate change uncertainty, *Water Resources Research* **33**: 817–829.
- Chaouche, A. et Parent, E. (1998). Sur l'apport de la statistique bayésienne au contrôle de la qualité par attribut; partie 3: contrôle rétrospectif par détection de changement dans une série de variables binomiales, *Revue de la Statistique Appliquée XLVI*: 5–29.
- Chernoff, H. et Zacks, S. (1964). Estimating the current mean of a normal distribution which is subjected to changes in time, *Annals of Mathematical Statistics* **35**: 999–1018.
- Chib, S. (1993). Bayes regression with autoregressive errors. A Gibbs sampling approach, *Journal of Econometrics* **58**: 275–294.
- Chib, S. (1995). Marginal likelihood from the Gibbs output, *Journal of the American Statistical Association* **90**: 1313–1321.
- Chib, S. et Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm, *The American Statistician* **49**: 327–335.
- Cobb, G. (1978). The problem of the Nile: conditional solution to a changepoint problem, *Biometrika* **65**: 243–251.
- Coles, S. et Tawn, J. A. (1996). A Bayesian analysis of extreme rainfall data, *Applied Statistics* **45**: 463–478.
- Davis, D., Kisiel, C. C. et Duckstein, L. (1972). Bayesian decision theory applied to design in hydrology, *Water Resources Research* **8**: 33–41.
- Diaz, J. (1982). Bayesian detection of a change of scale parameter in sequences of independent gamma random variables, *Journal of Econometrics* **19**: 23–29.
- Fortin, V. (1997). *Estimation de la valeur de l'information hydrologique à l'aide de probabilités imprécises*, Thèse de doctorat, INRS-Eau.

- Gardner, L. (1969). On detecting changes in the mean of normal variates, *The Annals of Mathematical Statistics* **40**: 116–126.
- Gelfand, A. et Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association* **85**: 398–409.
- Gelman, A., Carlin, J. B., Stern, H. S. et Rubin, D. B. (1995). *Bayesian Data Analysis*, Chapman and Hall, Londres.
- Ghorbanzadeh, D. (1996). Detection of random change point in one-parameter exponential families, *Applied Mathematics and Computation* **77**: 167–177.
- Girard, P. (2000). *Optimisation du suivi opérationnel de la qualité par la modélisation statistique de procédés agroalimentaires à partir de données recueillies sur ligne*, Thèse de doctorat, ENGREF.
- Graybill, F. A. (1983). *Matrices with Applications in Statistics*, Wadsworth, Statistics/Probability Series, Belmont, California.
- Greatbatch, R. (2000). The North Atlantic oscillation, *Stochastic Environmental Research and Risk Assessment* **14**: 213–242.
- GREHYS (1996). Presentation and review of some methods for regional flood frequency analysis, *Journal of Hydrology* **186**: 63–84.
- Hartigan, J. (1994). Linear estimators in change point problems, *The Annals of Statistics* **22**: 824–834.
- Hinkley, D. (1970). Inference about the change-point in a sequence of random variables, *Biometrika* **57**: 1–17.
- Hinkley, D. et Hinkley, E. A. (1970). Inference about the change-point in a sequence of binomial variables, *Biometrika* **57**: 477–488.
- Houghton, J. T., Filho, L. G. M., Callander, B. A., Harris, N., Kattenberg, A. et Makell, K. (1996). *The science of climate change. Contribution of Working Group 1 to the Second Scientific Assessment of the IPCC*, Cambridge University Press, Cambridge.
- Houghton, J. T., Jenkins, G. L. et Ephraums, J. J. (1990). *Climate change. The IPCC scientific assessment*, Cambridge University Press, Cambridge.
- Hsu, D. (1982). A Bayesian robust detection of shift in the risk structure of stock market returns, *Journal of the American Statistical Association* **77**: 29–39.
- Hubert, P. et Carbonnel, J.-P. (1987). Approche statistique de l'aridification de l'Afrique de l'Ouest, *Journal of Hydrology* **95**: 165–183.
- Hubert, P., Carbonnel, J.-P. et Chaouche, A. (1989). Segmentation des séries hydrométéorologiques - Application à des séries de précipitations et de débits de l'Afrique de l'Ouest, *Journal of Hydrology* **110**: 349–396.
- Hurrell, J. (1995). Decadal trends in the North Atlantic oscillation: regional temperatures and precipitations, *Science* **269**: 676–679.
- Hurrell, J. et van Loon, H. (1997). Decadal variations in climate associated with North Atlantic oscillation, *Climatic Change* **36**: 301–326.
- Hydro-Québec (1995). Plan directeur pour les relevés hydrométéorologiques, *Rapport Technique Préliminaire*, Hydro-Québec, Service Production.

- Hydro-Québec (2000). Projet Toulustouc. Choix du débit d'équipement et dimensionnement de la galerie, *Rapport Technique HG-RA-1999-0024-001*, Hydro-Québec, Service Hydraulique et Géotechnique.
- Joseph, L. et Wolfson, D. B. (1993). Maximum likelihood estimation in multi-path change-point problem, *Annals of Institut of Statistical Mathematics* **45** : 511–530.
- Kadane, J. et Wolfson, L. J. (1998). Experiences in elicitation, *The Statistician* **47** : 3–19.
- Kander, Z. et Zacks, S. (1966). Test procedures for possible changes in parameters of statistical distributions occurring at unknown time points, *Annals of Mathematical Statistics* **37** : 1196–1210.
- Kass, R. et Raftery, A. E. (1995). Bayes factors, *Journal of the American Statistical Association* **90** : 773–795.
- Kelly, P., Jones, P. D., Sear, C. B., Cherry, P. S. G. et Tavakol, R. K. (1982). Variations in surface air temperature: Part 2. Arctic region, *Monthly Weather Review* **110** : 71–83.
- Kiely, G. (1999). Climate change in Ireland from precipitation and streamflow observations, *Advances in Water Resources* **23** : 141–151.
- Kite, G. (1981). Recent changes in level of lake Victoria, *Hydrological Sciences Bulletin* **26** : 233–243.
- Krzysztofowicz, R. (1983). Why should a forecaster and a decision maker use Bayes theorem, *Water Resources Research* **19** : 327–336.
- Krzysztofowicz, R. (1999). Bayesian theory of probabilistic forecasting via deterministic hydrologic model, *Water Resources Research* **35** : 2739–2750.
- Kushnir, Y. (1999). Europe's winter prospects, *Nature* **398** : 289–291.
- Lavielle, M. et Moulines, E. (2000). Least-squares estimation of an unknown number of shifts in a time series, *Journal of Time Series Analysis* **21** : 33–59.
- Lecoutre, B. (1997). C'est bon à savoir! Et si vous étiez un bayésien qui s'ignore, *Modulad* **18** : 81–87.
- Lee, A. et Heghinian, S. M. (1977). A shift of the mean level in a sequence of independent normal random variables - a bayesian approach, *Technometrics* **19** : 503–506.
- Lehmann, E. L. (1983). *Theory of Point Estimation*, Wiley, New York.
- Lubès-Niel, H., Masson, J.-M., Paturel, J.-E. et Servat, E. (1998). Variabilité climatique et statistiques. Étude par simulation de la puissance et de la robustesse de quelques tests utilisés pour vérifier l'homogénéité de chroniques, *Revue des Sciences de l'Eau* **3** : 383–408.
- Menzeffricke, U. (1981). A Bayesian analysis of a change in the precision of a sequence of independent normal random variables at an unknown time point, *Applied Statistics* **30** : 141–146.
- Miquel, J. (1984). *Guide Pratique d'Estimation des Probabilités de Crues*, Eyrolles, Paris.
- Mustafi, C. (1968). Inference problems about parameters which are subjected to changes over time, *The Annals of Mathematical Statistics* **39** : 840–854.
- O'Hagan, A. (1995). Fractional Bayes factors for model comparisons, *Journal of the Royal Statistical Society Serie B* **57** : 99–138.

- O'Hagan, A. (1998). Eliciting expert beliefs in substantial practical applications, *The Statistician* **47**: 21–35.
- Ouarda, T., Rasmussen, P. F. et Bobée, B. (1996). Hydrometric data collection system rationalisation : Design of the networks of the future, *in International Conference on Water Resources and Environment Research : Towards the 21st Century*: 563–570.
- Ouarda, T., Rasmussen, P. F., Cantin, J.-F., Bobée, B., Laurence, R., Hoang, V. D. et Barabé, G. (1999). Identification d'un réseau hydrométrique pour le suivi des modifications climatiques dans la province de Québec, *Revue des Sciences de l'Eau* **12**: 425–448.
- Page, E. (1955). A test for a change in a parameter occurring at an unknown point, *Biometrika* **42**: 523–526.
- Paturel, J.-E., Servat, E., Kouamé, B., Lubès, H., Ouedraogo, M. et Masson, J.-M. (1997). Climatic variability in humid Africa along the Gulf of Guinea. Part II: An integrated regional approach, *Journal of Hydrology* **191**: 16–36.
- Perreault, L., Haché, M. et Bobée, B. (1996). Analyse statistique des séries d'apports énergétiques. Rapport final du projet C2, *Rapport de Recherche R-474*, INRS-Eau.
- Perreault, L., Haché, M., Slivitzky, M. et Bobée, B. (1999). Detection of changes in precipitation and runoff over Eastern Canada and U.S. using a Bayesian approach, *Stochastic Environmental Research and Risk Assessment* **13**: 201–216.
- Picard, D. (1985). Testing and estimating change-points in time series, *Advances in Applied Probability* **17**: 841–867.
- Piper, B., Plinston, D. T. et Sutcliffe, J. V. (1986). The water balance of lake Victoria, *Hydrological Sciences Bulletin* **31**: 25–37.
- Raiffa, H. et Schlaifer, R. (1961). *Applied Statistical Decision Theory*, Harvard University Press, Harvard.
- Rao, A. et Tirtotjondro, W. (1996). Investigation of changes in characteristics of hydrological time series by Bayesian methods, *Stochastic Hydrology and Hydraulics* **10**: 295–317.
- Rasmussen, P. F., Fortin, V., Slivitzky, M. et Bobée, B. (1999). Impact des oscillations climatiques à basse fréquence sur les apports des rivières québécoises. *Rapport de Recherche R-541*, INRS-Eau.
- Robert, C. P. (1992). *L'Analyse Statistique Bayésienne*, Economica, Paris.
- Robert, C. P. (1996). *Méthodes de Monte Carlo par Chaînes de Markov*, Economica, Paris.
- Robert, C. P. (1998). *Discretization and MCMC Convergence Assessment*, Springer-Verlag, New York.
- Robert, C. P. et Casella, G. (1999). *Monte Carlo Statistical Methods*, Springer-Verlag, New York.
- Savage, L. J. (1954). *The Foundations of Statistics*, Dover Publications, New York.
- Sen, A. et Srivastava, M. S. (1975). Some one-sided tests for change in level, *Technometrics* **17**: 61–64.

- Servat, E., Paturel, J.-E., Lubès, H., Kouamé, B., Ouedraogo, M. et Masson, J.-M. (1997). Climatic variability in humid Africa along the Gulf of Guinea. Part I: Detailed analysis of the phenomenon in Côte d'Ivoire, *Journal of Hydrology* **191**: 1–15.
- Shaban, S. (1980). Change point problem and two-phase regression: an annotated bibliography, *International Statistical Review* **48**: 83–93.
- Slivitzky, M. et Mathier, L. (1994). Climatic changes during the 20th century on the Laurentian Great Lakes and their impacts on hydrologic regime, in *Engineering Risk in Natural Resources Management, NATO ASI Series E*, L.Duckstein et E.Parent (eds), Kluwer Academic Publishers, pp. 235–251.
- Smith, A. (1975). A Bayesian approach to inference about a change-point in a sequence of random variables, *Biometrika* **62**: 407–416.
- Smith, A. F. M. (1981). Change-point problems: approaches and applications, in *Bayesian Statistics*, D. M. H. L. D. V. Bernardo, J. M. et A. F. M. Smith (eds), Valencia University Press, pp. 83–98.
- Stephens, D. (1994). Bayesian retrospective multiple-changepoint identification, *Applied Statistics* **43**: 159–178.
- Tanner, M. et Wong, W. H. (1987). The calculation of posterior distributions by data augmentation, *Journal of the American Statistical Association* **82**: 528–550.
- Venkatesh, B. et Hobbs, B. F. (1999). Analyzing investments for managing lake Erie levels under climate change uncertainty, *Water Resources Research* **35**: 1671–1683.
- Vincens, G., Rodriguez-Iturbe, I. et Schaake, J. C. (1975). A Bayesian framework for the use of regional information in hydrology, *Water Resources Research* **11**: 405–414.
- Yamamoto, R., Iwashima, T. et Kazadi, S. N. (1985). Climatic jump: A hypothesis in climatic diagnosis, *Journal of Meteorological Society of Japan* **63**: 1157–1160.

Annexe A

Article I

Bayesian change-point analysis in hydrometeorological time series 1. The normal model revisited

Luc Perreault^{1,2}, Jacques Bernier^{1,2}, Bernard Bobée², Éric Parent¹
¹ENGREF ²INRS-Eau

Journal of Hydrology (2000), 235 : 221-241

Research and drafting of the document have been performed at ENGREF.

Abstract

A Bayesian method is presented for the analysis of two types of sudden change at an unknown time-point in a sequence of energy inflows modeled by independent normal random variables. First, the case of a single shift in the mean level is revisited to show how such a problem can be straightforwardly addressed through the Bayesian framework. Second, a change in variability is investigated. In hydrology, to our knowledge, this problem has not been studied from a Bayesian perspective. Even if this model is quite simple, no analytic solutions for parameter inference are available, and recourse to approximations is needed. It is shown that the Gibbs sampler is particularly suitable for change-point analysis, and this Markovian updating scheme is used. Finally, a case study involving annual energy inflows of two large hydropower systems managed by Hydro-Québec is presented in which informative prior distributions are specified from regional information.

Keywords. Change-point analysis; Energy inflows; Gibbs sampling; Change in the variance; Change in the mean level

A.1 Introduction

This paper provides a Bayesian approach to characterize when and by how much a single change has occurred in a sequence of hydrometeorological random variables. Inferences herein are based on the analysis of posterior distributions and are conditional upon the fact that a change happened with certainty. The problem of testing its existence, and of identifying its type, is investigated in Perreault et al. (2000a).

Although such an hypothesis is rarely stated explicitly, the assumption that stochastic time series remain stationary plays a crucial role in water resources management. Under the assumption that tomorrow will statistically behave like yesterday, stochastic models are fitted to hydrometeorological variables such as river flow, precipitation and temperature. The

estimated models are then used for many engineering purposes, in particular for simulating the operation of hydropower systems (energy planning, design of power plants, operation of reservoirs). Consequently such models and the decisions stemming from them are based on the hypothetical stationary behavior of hydrometeorological inputs. However, some hydrometeorological time series can exhibit abrupt changes maybe caused by site-specific factors (e.g. land-use effect on water yield) or induced by a climatic change. This leads to questioning the stationarity hypothesis in hydrometeorological time series analysis.

Hydro-Québec is a public company that produces, transmits and distributes electricity throughout the province of Québec. It currently operates 54 power plants supplied by 26 large reservoirs. The sites are assembled to form 8 major hydropower systems: Saint-Laurent, Outaouais, La Grande, Saint-Maurice, Bersimis, Manicouagan, Outardes and Churchill Falls. For each of these systems, annual energy inflows are evaluated by multiplying the net basin supply of each reservoir into the system by a factor based on the production capacity of the corresponding power plant. Such annual energy inflows are therefore subject to hydrometeorological changes, if any. These data are of high importance for energy planning because some of their statistical characteristics (namely mean and variance) are used as inputs to construct scenarios or to forecast future energy availability.

Consider the time plot in Figure 1 which shows the annual energy inflow in terawatt-hour (TWh) for the hydropower system Churchill Falls. This 92 432 km² watershed is situated in northeastern Québec, in the Labrador (province of Newfoundland).

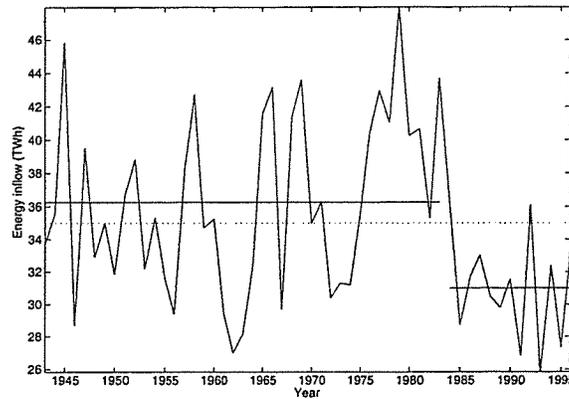


Figure 1. Annual energy inflow for Churchill Falls power system.

Examining this time series, one may suspect that an abrupt change in the mean level has occurred around 1983, and two distinct partial means could be evaluated: \bar{x}_1 from 1943 to 1983 and \bar{x}_2 from 1984 to 1996 (continuous lines). On the other hand, it can be argued that this sudden fluctuation may only be due to the natural variability of the hydrologic regime, and one would rather still consider the overall mean \bar{x} as the representative available annual amount of energy (dotted line). Let's now suppose that based on these historical observations, the construction of a power plant has to be planned for hydropower generation. To meet the energy demand and eventually export generated hydroelectricity, the hydropower company would like to rely on one and only one of the possible situations (\bar{x}_1 , \bar{x}_2) or \bar{x} . Since for this case $(\bar{x}_2 - \bar{x}_1) / \bar{x}$ may be as high as 15%, decisions for the future relying on \bar{x} can be dramatically different from anticipations based on (\bar{x}_1 , \bar{x}_2). Analyzing a change in the mean level energy inflow data is then clearly an essential step before planning hydropower systems. Other hydrological contexts such as rainfall estimation, flood analysis or climatological studies of global warming are concerned with abrupt shifts in the mean level. References can be found in Perreault et al. (1999).

As an example for a possible change of a different kind, consider the annual series of energy inflow for the Outaouais power system located in the southwestern Québec (Figure 2).

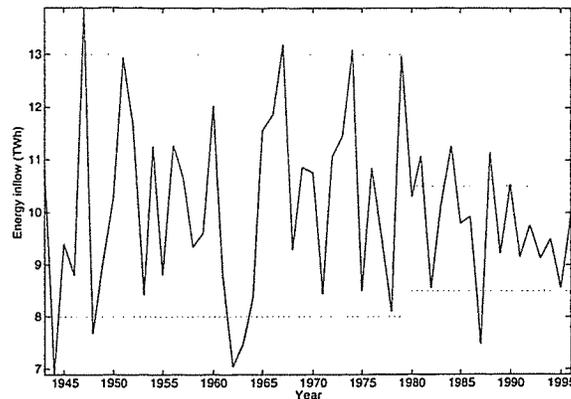


Figure 2. Annual energy inflow for Outaouais power system.

The overall variability of these observations seems to have suddenly decreased after 1980. The energy planning system of forecasts depends strongly upon energy inflow variability. For instance, a sharp knowledge of this characteristic is the key point to determine which range of scenarios is needed to make a good decision for reservoir releases and interannual storage strategy. After examining this series, an engineer will typically base his decision on one of the two possible estimates: evaluating the variance over the entire period or considering only the last observations, i.e. using more but maybe not relevant information about future realizations or taking into account less but maybe more representative information. As in the case of a shift in the mean, a change in variance may induce important decisions and have large economic consequences (it is well known for instance that the size of reservoirs will be directly related to inflow variance); inference about existence and characteristics of these changes can thus also be considered as a valuable precaution before developing management rules in water resources systems.

In this paper, interest is mainly focussed on the estimation of the unknown time-point and intensity of the change, assuming an abrupt change has in fact occurred. The aim is to outline that, in the case of a single sudden change in the mean level or variance, these problems can be straightforwardly addressed through the Bayesian framework, which in turn can be easily solved by implementation of explicit or Monte Carlo based inference techniques. We emphasize that the purpose of this paper is not to determine why such sudden changes occurred (e.g. climatic change or site-specific factors). Of course, attribution is an important and interesting research topic, but we are only concerned with developing statistical tools to infer about the point and intensity of a change to help decision making.

In Section 2, we formulate the general Bayesian change-point setting and consider two specific models: a single change in the mean and a single change in the variance of a sequence of independent normal random variables. Change in the mean level leads to solutions in closed form and results are therefore directly available. However, such is not the case for a change in the variance. The Gibbs sampler, as presented in Section 3, provides an elegant and convenient answer to this problem. In Section 4, the practitioner's point of view is taken when applying the approach to the series of Churchill Falls and Outaouais energy inflows presented above. A simple regional analysis was performed to specify the parameters of the prior distributions which are assumed to represent the prior state of belief. Finally, Section 5 offers a general discussion and conclusions.

A.2 Bayesian analysis of normal sequences with an unknown change-point

The simplest formulation of the change-point problem is the following. Let us assume that the densities $p_1(x)$ and $p_2(x)$ belong to a known parametric class of probability densities $\mathbf{P} = \{p(x | \theta); \theta \in \Theta\}$ indexed by an unknown parameter θ such that, for a sequence of n independent random variables $\mathbf{X} = (X_1, X_2, \dots, X_n)$, we have

$$\begin{aligned} X_i &\sim p_1(x) = p(x_i | \theta_1), & i = 1, \dots, \tau \\ X_i &\sim p_2(x) = p(x_i | \theta_2), & i = \tau + 1, \dots, n \end{aligned} \quad (\text{A.1})$$

where $\theta_1 \neq \theta_2$ and $\tau = 1, 2, \dots, n-1$ is an unknown parameter, called the change-point. That is, the first and second parts of the sequence of random variables are distributed as statistical distributions which belong to the same class, but with different unknown parameter θ . Since the unknown change-point τ can take values between 1 and $n-1$, this model assumes that exactly one abrupt change occurred with certainty. Indeed other types of changes such as trends can exist, and model (A.1) would not be appropriate. However, recent understanding of global climate interactions such as the El Nino/La Nina and the North Atlantic Oscillation phenomena give credence to the idea that climate may operate in two or more quasi-stationary states, and that it can rapidly switch from one state to another (Rodriguez-Iturbe et al., 1991; Kerr, 1992, 1999). Therefore, an abrupt change-point model such as (A.1) may be representative of several hydrological and climatic time series. For instance, there is evidence of such behavior in climatic and hydrological data series in the Africa (Servat et al., 1997; Paturel et al., 1997) and in the Pacific region (Kerr, 1992). As regards Hydro-Québec energy inflow series, many annual runoff series for rivers situated in Northern Québec seem to exhibit this type of sudden drastic change (Perreault et al., 2000b). Thus, model (A.1) must be considered, at least as a first methodological step. Anyhow, we think that there is a need to consider first this simple model in detail since application of Bayesian change-point analysis in hydrological literature is very limited.

The likelihood function resulting from n observations $\mathbf{x} = (x_1, x_2, \dots, x_n)$ generated by model (A.1) becomes

$$p(\mathbf{x} | \theta_1, \theta_2, \tau) = \prod_{i=1}^{\tau} p(x_i | \theta_1) \prod_{i=\tau+1}^n p(x_i | \theta_2) \quad (\text{A.2})$$

from which, for instance, maximum likelihood estimates for τ, θ_1 and θ_2 can be obtained (see for example, Hinkley, 1970). In the Bayesian perspective, a joint prior distribution $p(\theta_1, \theta_2, \tau)$ is assumed for the parameters. Bayes theorem then provides the joint posterior distribution $p(\theta_1, \theta_2, \tau | \mathbf{x})$ of θ_1, θ_2, τ given the data by normalization over \mathbf{x} of the joint density written as

$$p(\mathbf{x} | \theta_1, \theta_2, \tau) p(\theta_1, \theta_2, \tau) \quad (\text{A.3})$$

Just as the prior distribution $p(\theta_1, \theta_2, \tau)$ reflects beliefs about the parameters prior to experimentation, the posterior distribution $p(\theta_1, \theta_2, \tau | \mathbf{x})$ reflects the updated beliefs after observing the sample data. In the Bayesian framework, all statistical inference about the unknown parameters is based on the posterior distribution (Berger, 1985). The analysis of posterior distributions is often referred to as « scientific reporting ».

Interest is now on making inference about the change-point τ , the structure parameters θ_1 and θ_2 , and any suitable function $W(\theta_1, \theta_2)$ which can describe the amount of shift. To achieve this, all marginal posterior distributions must be found by integration. For example, to evaluate $p(\tau | \mathbf{x})$, θ_1 and θ_2 must be integrated out of $p(\theta_1, \theta_2, \tau | \mathbf{x})$. Using conjugate

prior distributions for fixed τ , and assuming prior independence between τ and the other parameters (θ_1, θ_2) , i.e. $p(\theta_1, \theta_2, \tau) = p(\theta_1, \theta_2) p(\tau)$, solutions in closed form can be obtained for some simple models (see Section 2.1). Likelihoods for which conjugate prior density functions exist are those corresponding to exponential family models (Bernardo and Smith, 1994). In more complex models (see Section 2.2), even with simple prior distributions and independence, integration may turn to a very difficult numerical task. However, use of Gibbs sampler, a tool particularly suitable for change-point analysis, enables a straightforward solution to such problems as developed in Section 3.

A.2.1 A single change in the mean (Model M_1)

As in most published approaches on change-point studies, we assume univariate sequences of independent normal random variables. This simple model is appropriate for Hydro-Québec large hydropower system annual energy inflows. One can invoke the central limit theorem to justify the normal assumption, since annual energy inflows for a given hydropower system are calculated as a summation over time and space of the monthly energy inflows. In addition, changes in the parameters of the normal model have a direct and simple hydrological interpretation. If needed, one may transform the data to achieve normality. However, in this case of reparametrization, interpretation of a change in the new parameters should be done very carefully. From the hydrological point of view, the independence assumption between successive values may be more questionable. It is made here for convenience as a first methodological step, and can be relaxed eventually by considering, for example, a shift in the mean-vector of an autoregressive model. Hydro-Québec annual energy inflows are proportional to the net basin supplies, which in turn are evaluated following the water balance equation. Therefore, the data considered herein are implicitly adjusted for surface and subsurface storage effect that could induce interannual correlations, and the assumption of independence seems reasonable to us. Still, it can be verified on similar annual series for which it is believed that no change has occurred.

Consider a set of random variables observed at consecutive equally spaced time points. Suppose that, due to some exogenous factors, the first and second parts of the sequence of random variables operate at two different mean levels, respectively μ_1 and μ_2 , but with the same variance σ^2 . This situation can be represented by the following model denoted by M_1 :

$$\begin{aligned} X_i &\sim \mathcal{N}(x_i | \mu_1, \sigma^2), & i = 1, \dots, \tau \\ X_i &\sim \mathcal{N}(x_i | \mu_2, \sigma^2), & i = \tau + 1, \dots, n \end{aligned} \quad (\text{A.4})$$

where $\mathcal{N}(x_i | \mu, \sigma^2)$ stands for the usual normal probability density function (p.d.f.) with parameters $\mu \in \mathfrak{R}$ and $\sigma \in \mathfrak{R}^+$.

The Bayesian framework for inference regarding model M_1 dates back to Chernoff and Zacks (1963). Main theoretical contributions can be found in Smith (1975), Lee and Heghinian (1977), Booth and Smith (1982) and Broemeling (1985). Their approaches differ mainly by the prior distributions specified to represent the unknown parameters. The method proposed by Lee and Heghinian (1977) has been used in several practical cases in hydrology, for instance by Bruneau and Rassam (1983), and more recently by Servat et al. (1997) and Paturel et al. (1997). Also, Bernier (1994) and Perreault et al. (1999) considered a Bayesian approach for model M_1 to analyse changes in hydrometeorological time series.

To illustrate the principle of Bayesian change-point analysis, the posterior distributions of the parameters of interest for model M_1 are derived in Appendix A. Conjugate prior distributions are considered and independence between the date of change τ and the model parameters is assumed. With these assumptions, the joint posterior distribution is a finite mixture of conjugate distributions, and analytical expressions are obtained for all marginals and conditionals. The assumption of independence signifies that what is known about date of change in the mean level of the annual energy inflows does not depend upon the means before and after the change. Note that this hypothesis could be questionable if for instance

one believes that a change can only occur if the mean level reaches a given threshold. These results are well known in the statistical literature, but change-point inference analysis does not belong to the standard statistical toolbox of most practitioners in hydrology.

A.2.2 A single change in the variance (Model M_2)

In this section we consider the change of variance case represented by model M_2 :

$$\begin{aligned} X_i &\sim \mathcal{N}(x_i | \mu, \sigma_1^2), & i = 1, \dots, \tau \\ X_i &\sim \mathcal{N}(x_i | \mu, \sigma_2^2), & i = \tau + 1, \dots, n \end{aligned} \quad (\text{A.5})$$

The problem of a change in the variance occurring at an unknown time-point has been less widely covered than a single shift in the mean level. In hydrology, to our knowledge, it has not been studied from a Bayesian perspective although the variance of hydrological series is an important parameter in water resources management. In the statistical literature, a change in the variance of a sequence of normal random variables was first treated within the Bayesian framework by Smith (1975) who assumed noninformative prior distributions. In this particular case, closed form for the posterior densities can be obtained. Menzefricke (1981) generalized Smith's approach by considering informative priors.

The likelihood function resulting from n observations $\mathbf{x} = (x_1, x_2, \dots, x_n)$ of model M_2 , a series of annual energy inflows such as shown in Figure 2, can be expressed as

$$\begin{aligned} p(\mathbf{x} | \mu, \sigma^2, \tau, M_2) &= \prod_{i=1}^{\tau} \mathcal{N}(x_i | \mu, \sigma_1^2) \prod_{i=\tau+1}^n \mathcal{N}(x_i | \mu, \sigma_2^2) \\ &= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma_1^2}\right)^{\frac{\tau}{2}} \left(\frac{1}{\sigma_2^2}\right)^{\frac{n-\tau}{2}} \exp\left\{-\frac{\tau}{2\sigma_1^2} \left[s_\tau^2 + (\bar{x}_\tau - \mu)^2\right]\right\} \\ &\quad \times \exp\left\{-\frac{(n-\tau)}{2\sigma_2^2} \left[s_{n-\tau}^2 + (\bar{x}_{n-\tau} - \mu)^2\right]\right\} \end{aligned} \quad (\text{A.6})$$

We assumed the same joint prior distribution as Menzefricke (1981) to represent prior knowledge about the parameters of model M_2 :

$$\begin{aligned} p(\mu, \sigma^2, \tau | M_2) &= \mathcal{N}(\mu | \phi, \lambda \sigma_1^2) \mathcal{IG}(\sigma_1^2 | \alpha_1, \beta_1) \mathcal{IG}(\sigma_2^2 | \alpha_2, \beta_2) p(\tau | M_2) \\ &= \mathcal{NIGIG}(\mu, \sigma^2 | \phi, \lambda, \alpha, \beta) p(\tau | M_2) \end{aligned} \quad (\text{A.7})$$

As usual, the prior distribution of μ is normal with precision depending on σ_1 , and priors for the variance are inverted gamma densities. However, strictly speaking, the prior distribution (A.7) although closely related is not a complete conjugate even for fixed τ . In fact, given τ , expression (A.7) assumes prior independence between σ_1^2 and σ_2^2 , while no algebraic manipulations on the likelihood function (A.6) allows us to separate these parameters in two independent expressions. Therefore, the exact conjugate should consider a fixed dependence structure between σ_1^2 and σ_2^2 . This would lead to a non-standard joint prior density for which it is difficult to specify the hyperparameters: eliciting knowledge about dependence between variances may be at least a quite complicated task if not unrealistic. We therefore adopt Menzefricke's prior assumptions letting the data account for the strength of dependence, if any.

Selection of particular values in (A.7) for the hyperparameters ϕ , λ , α and β belongs to modeler's tasks (Section 4.1). These can be chosen to give various general shapes for the joint prior distribution and take into account a variety of prior beliefs about the studied

phenomenon. The hyperparameter values may come from historical or regional information, even from subjective knowledge. When prior information about the phenomenon is limited, it may be desirable to let the prior knowledge for the unknown parameters be vague (Box and Tiao, 1973). The prior distribution can be turned to a particular form of noninformative density by letting $p(\mu, \sigma^2, \tau | M_2) \propto p(\tau | M_2) \sigma_1^{-2} \sigma_2^{-2}$, i.e. $\lambda \rightarrow \infty$, $\alpha_1 = \alpha_1 = \beta_1 = \beta_2 \rightarrow 0$ in (A.7). This corresponds to Smith's noninformative joint prior distribution for M_2 .

Using Bayes theorem to combine (A.6) and (A.7), the joint posterior distribution of (μ, σ^2, τ) , given the observed data \mathbf{x} , is proportional to

$$p(\mu, \sigma^2, \tau | \mathbf{x}, M_2) \propto \exp \left\{ -\frac{\lambda' (n - \tau) \sigma_1^2 + \sigma_2^2}{2\lambda' \sigma_1^2 \sigma_2^2} (\mu - \phi'')^2 \right\} \left(\frac{1}{\sigma_1^2} \right)^{\alpha_1' + 1} \left(\frac{1}{\sigma_2^2} \right)^{\alpha_2' + 1} \quad (\text{A.8})$$

$$\times \exp \left\{ -\left[\frac{\beta_1'}{\sigma_1^2} + \frac{\beta_2'}{\sigma_2^2} + \frac{(n - \tau) (\phi' - \bar{x}_{n-\tau})^2}{2[\lambda' (n - \tau) \sigma_1^2 + \sigma_2^2]} \right] \right\} p(\tau | M_2)$$

where

$$\begin{aligned} \lambda' &= \lambda / (1 + \tau \lambda), \quad \phi' = (1 - \lambda' \tau) \phi + \lambda' \tau \bar{x}_\tau, \\ \phi'' &= \frac{1}{\lambda' (n - \tau) \sigma_1^2 + \sigma_2^2} [\sigma_2^2 \phi' + \lambda' (n - \tau) \sigma_1^2 \bar{x}_{n-\tau}] \\ \beta_1' &= \frac{\tau}{2} [s_\tau^2 + (1 - \lambda' \tau) (\phi - \bar{x}_\tau)^2] + \beta_1, \quad \beta_2' = \frac{n - \tau}{2} s_{n-\tau}^2 + \beta_2, \\ \alpha_1' &= \alpha_1 + \frac{\tau + 1}{2}, \quad \alpha_2' = \alpha_2 + \frac{n - \tau}{2} \end{aligned}$$

It is possible to obtain an expression for the marginal posterior density of the intensity of change, defined as $\eta = \sigma_2^2 / \sigma_1^2$, which is a parameter of great interest for describing a change in variance. First, transforming to $\theta = \sigma_1^2$ and $\eta = \sigma_2^2 / \sigma_1^2$, and integrating μ and θ out from this joint posterior distribution (A.8), the joint posterior distribution of η and τ is obtained. Then, summing over τ , the marginal posterior distribution $p(\eta | \mathbf{x}, M_2)$ is seen to be proportional to

$$p(\eta | \mathbf{x}, M_2) \propto \sum_{\tau=1}^{n-1} \left[\frac{\lambda'}{\lambda' (n - \tau) + \eta} \right]^{1/2} \left(\frac{1}{\eta} \right)^{\alpha_2' + 1/2} \frac{\Gamma(\alpha_1' + \alpha_2' + 1/2)}{[g(\tau, \eta)]^{\alpha_1' + \alpha_2' + 1/2}} p(\tau | M_2), \quad (\text{A.9})$$

where

$$g(\tau, \eta) = \left\{ \beta_1' + \beta_2' \eta^{-1} + \frac{(n - \tau) (\phi' - \bar{x}_{n-\tau})^2}{2[\lambda' (n - \tau) + \eta]} \right\}$$

Writing (A.8) differently, it is also possible to find a closed form for the marginal posterior distribution of μ by integrating σ_1^2 and σ_2^2 out of this expression and summing over τ . This expression is not given here since μ is not the parameter of major interest for model M_2 . However, unlike the noninformative case or the problem of a single change in the mean level, recourse to numerical integration is necessary to make inferences about the other parameters (σ_1^2 , σ_2^2 and τ). How can we avoid performing integration to get the normalizing constants in (A.8) and (A.9)? Fortunately, the Gibbs sampler is an alternative handy tool, easy to implement for the hydrologist without numerical analysis skills, as will be showed in the next section.

A.3 Implementation of the Gibbs sampler for model M_2

The Gibbs sampler is one of a class of Markov Chain Monte Carlo (MCMC) algorithms that facilitate practical statistical problem solving (Tanner, 1992). In a Bayesian perspective, the objective is to produce posterior densities for parameters of interest. For simple models analytic calculation is possible, and results are directly available. For more complex models, recourse to numerical integration is necessary. However, in some models conventional numerical techniques are often insufficiently accurate or at least difficult to implement without being gifted in computer programming and numerical analysis. In contrast, the Gibbs sampler is generally straightforward to implement and is easily accessible to the average statistical practitioner. The Gibbs sampler was developed formally by Geman and Geman (1984) in the context of image restoration. Gelfand and Smith (1990) showed its applicability to general Bayesian computations while Carlin et al. (1992) and Stephens (1994) applied it to change-point problems. In hydrology, Lu and Berliner (1999) used the Gibbs algorithm, while other MCMC approaches were adopted by Kuczera and Parent (1998) and Campbell et al. (1999).

The idea of MCMC methods is to construct a Markov chain whose stationary and ergodic distribution is precisely the posterior distribution of interest that is intractable analytically. In the context of the general change-point model (A.1), the Gibbs sampler is an updating scheme that involves drawing t random values from each of the set of « full posterior distributions » $p(\tau | \theta_1, \theta_2, \mathbf{x})$, $p(\theta_1 | \theta_2, \tau, \mathbf{x})$ and $p(\theta_2 | \theta_1, \tau, \mathbf{x})$. A brief description in a general context of how these values should be sampled is offered in Lu and Berliner (1999). More details on theoretical properties of the algorithm can be found in Gelfand and Smith (1990). Subject to certain mild regularity conditions on the joint and conditional densities, it can be proven that the obtained sampled vector $\left\{ \left(\tau^{(i)}, \theta_1^{(i)}, \theta_2^{(i)} \right); i = 1, \dots, t \right\}$ tends in distribution to a sample from the target density $p(\tau, \theta_1, \theta_2 | \mathbf{x})$ as t tends to infinity. Hence, replicating the entire process in parallel m times and keeping the values obtained in the last iteration t provides independent and identically distributed triplets $\left(\tau_j^{(t)}, \theta_{1j}^{(t)}, \theta_{2j}^{(t)} \right)$, $j = 1, \dots, m$. These values can then easily be used for estimating different features of the posterior joint distribution, namely the marginal densities. For example, as an estimate for $p(\theta_1 | \mathbf{x})$, one can draw a histogram with $\left(\theta_{11}^{(t)}, \dots, \theta_{1m}^{(t)} \right)$ or use a more general kernel-type of estimate. But since for any marginal the corresponding full conditional has been assumed available, more efficient estimation is obtained by using this set of distributions. In this paper, we use the so-called « Rao-Blackwellized » estimate (Gelfand and Smith, 1990); for density $p(\theta_1 | \mathbf{x})$, the estimate is given by

$$\hat{p}(\theta_1 | \mathbf{x}) = \frac{1}{m} \sum_{j=1}^m p\left(\theta_1 \mid \tau_j^{(t)}, \theta_{2j}^{(t)}, \mathbf{x}\right) \quad (\text{A.10})$$

The formal argument for the higher efficiency of (A.10) is based on the Rao-Blackwell theorem (Lehmann, 1983, p.50). There may also be interest in a function of the random variables $W(\tau, \theta_1, \theta_2)$. Because each triplet $\left(\tau^{(t)}, \theta_1^{(t)}, \theta_2^{(t)} \right)$ provides an observed $W^{(t)} = W\left(\tau^{(t)}, \theta_1^{(t)}, \theta_2^{(t)}\right)$ whose marginal distribution is approximately $p(W)$, an estimate analogous to (A.10) can be obtained. In fact, if θ_1 actually appears as an argument of W , the full conditional density $p(W | \theta_2, \tau, \mathbf{x})$ can be deduced by univariate transformation of variable from $p(\theta_1 | \theta_2, \tau, \mathbf{x})$, and used in (A.10). Note that complete implementation of the Gibbs sampler requires the determination of t and, across iterations, the choice of m . These values may vary between individual problems and are thus subject of much current research (a Web site which includes a lot of information and references about convergence diagnostics

is <http://www.ensae.fr//crest/statistique/robert/McDiag>). Generally, experimentation with different settings of t and m are necessary. We do not view this as a drawback since random generation is now computationally inexpensive.

For the general change-point model (A.1), each of the full conditional distributions $p(\tau | \theta_1, \theta_2, \mathbf{x})$, $p(\theta_1 | \theta_2, \tau, \mathbf{x})$ and $p(\theta_2 | \theta_1, \tau, \mathbf{x})$ is proportional to (A.3). Moreover, whatever the prior dependence structure between the parameters θ_1 and θ_2 may be, $p(\tau | \theta_1, \theta_2, \mathbf{x})$ is exactly of the form

$$p(\tau | \theta_1, \theta_2, \mathbf{x}) = \frac{p(\mathbf{x} | \theta_1, \theta_2, \tau) p(\tau)}{\sum_{\tau=1}^{n-1} p(\mathbf{x} | \theta_1, \theta_2, \tau) p(\tau)} \quad (\text{A.11})$$

and can be easily sampled since it involves a discrete density. Suppose that θ_1 and θ_2 are assumed independent so that $p(\theta_1, \theta_2) = p(\theta_1)p(\theta_2)$. If $p(\theta_1)$ is conjugate with $p(x_i | \theta_1)$ and $p(\theta_2)$ is conjugate with $p(x_i | \theta_2)$ then $p(\theta_1 | \theta_2, \tau, \mathbf{x})$ does not depend on θ_2 and is merely the prior $p(\theta_1)$ updated by the data x_1, \dots, x_τ while $p(\theta_2 | \theta_1, \tau, \mathbf{x})$ is $p(\theta_2)$ updated by $x_{\tau+1}, \dots, x_n$. Since $p(\theta_1)$ and $p(\theta_2)$ belong to standard parametric families, it is straightforward to sample from them. More generally, this property still remains when only conditional independence upon a « nuisance » parameter, which describes the unknown but constant structure of the distribution on both sides of τ (σ^2 in M_1 and μ in M_2). Therefore, with these hypotheses, the Gibbs sampler is particularly suitable for change-point analysis. If the assumption of conjugate priors and of independence is dropped, the full conditionals may not be expressed in a simple form and we must sample from non-standardized densities of the form (A.3). In this case, more general Markov Chain Monte Carlo approaches are needed (Robert, 1996 ; 1998). The most popular tool that one can implement is undoubtedly the Metropolis-Hasting algorithm (Metropolis et al., 1953).

Because prior distribution (A.7) assumes independence between σ_1^2 and σ_2^2 , the collection of full conditional densities $p(\mu | \sigma^2, \tau, \mathbf{x}, M_2)$, $p(\sigma_1^2 | \mu, \sigma_2^2, \tau, \mathbf{x}, M_2)$ and $p(\sigma_2^2 | \mu, \sigma_1^2, \tau, \mathbf{x}, M_2)$ can readily be determined for model M_2 . Each of these distributions is proportional to (A.8) and belong to the same parametric family as their corresponding prior. They are therefore available for sampling. Finally, as it was outlined above, $p(\tau | \mu, \sigma^2, \mathbf{x}, M_2)$ is exactly of the form (A.11).

More precisely, from (A.6), (A.7) and (A.8) it is straightforward to verify that the full conditionals for M_2 are as

$$\begin{aligned} p(\mu | \sigma^2, \tau, \mathbf{x}, M_2) &= \mathcal{N}\left(\mu \mid \phi'', \frac{\lambda' \sigma_1^2 \sigma_2^2}{\lambda' (n - \tau) \sigma_1^2 + \sigma_2^2}\right) \\ p(\sigma_1^2 | \mu, \sigma_2^2, \tau, \mathbf{x}, M_2) &= \text{IG}\left(\sigma_1^2 \mid \alpha'_1, \frac{\tau}{2} \left[s_\tau^2 + (\bar{x}_\tau - \mu)^2 + \frac{(\mu - \phi)^2}{\tau \lambda} \right] + \beta_1\right) \\ p(\sigma_2^2 | \mu, \sigma_1^2, \tau, \mathbf{x}, M_2) &= \text{IG}\left(\sigma_2^2 \mid \alpha'_2, \frac{n - \tau}{2} \left[s_{n-\tau}^2 + (\bar{x}_{n-\tau} - \mu)^2 \right] + \beta_2\right) \\ p(\tau | \mu, \sigma^2, \mathbf{x}, M_2) &= \frac{p(\mathbf{x} | \mu, \sigma^2, \tau, M_2) p(\tau | M_2)}{\sum_{\tau=1}^{n-1} p(\mathbf{x} | \mu, \sigma^2, \tau, M_2) p(\tau | M_2)} \end{aligned}$$

Highly efficient estimates for the marginal posterior distributions of each parameters μ , σ_1^2 , σ_2^2 and τ from Gibbs sampler simulated values can then be obtained using a Rao-Blackwellized estimator. Following the discussion above, we can also transform σ_2^2 to $\eta \sigma_1^2$ to obtain the full conditional distribution of η as

$$p(\eta | \mu, \sigma_1^2, \tau, \mathbf{x}, M_2) = \text{IG}\left(\eta \mid \alpha'_2, \frac{n - \tau}{2 \sigma_1^2} \left[s_{n-\tau}^2 + (\bar{x}_{n-\tau} - \mu)^2 \right] + \frac{\beta_2}{\sigma_1^2}\right)$$

and use it to evaluate the Rao-Blackwellized estimate for $p(\eta | \mathbf{x}, M_2)$, i.e.

$$\hat{p}(\eta | \mathbf{x}, M_2) = \frac{1}{m} \sum_{j=1}^m p\left(\eta \mid \mu_j^{(t)}, \sigma_{1j}^{2(t)}, \tau_j^{(t)}, \mathbf{x}, M_2\right) \quad (\text{A.12})$$

A.4 Applications

Bayesian change-point analysis procedure is now applied to annual energy inflow series presented in the introduction. Model M_1 is assumed for the Churchill Falls hydropower system, while model M_2 is considered for the Outaouais hydropower system.

A.4.1 Specifying prior distributions

The first step in Bayesian analysis is to set up a *full probability model*. That is, in addition to modelling the observable quantities (model M_1 or M_2 for the annual energy inflow), we must represent the prior degree of belief about the unknowns, i.e. the parameters of the models. In Bayesian analysis, specifying a prior distribution for the parameters is an integral part of the modelling task, with all hypotheses that modelling involves. Prior elicitation is therefore a crucial component of the Bayesian approach. Generally, a compromise has to be reached between mathematical simplicity and realism. The conjugate prior distribution assumption adopted here (Section 2) facilitates the derivation of posterior distributions and may be criticized for lack of realism. Still, the family of normal-inverted gamma is flexible and allows for the representation of a wide spectrum of prior knowledge. To be fully operational, it is desirable to specify the prior distributions by eliciting knowledge of an expert and/or by using other information than the data themselves. An engineer should have valuable prior information about annual energy inflow behavior (subjective knowledge, regional information, etc.). To our knowledge, the first convincing approach to specify realistic priors in a hydrological context has been proposed by Bernier (1967) for frequency analysis with the lognormal distribution. Hyperparameter values were specified using estimated quantiles from a preliminary study. Posterior quantiles and credible intervals were then obtained using Bayes theorem. More recently, Coles and Tawn (1996) used a similar approach for the analysis of extreme rainfall data with the GEV probability distribution. To specify the prior density for the parameters of the GEV, they elicited subjective knowledge of an expert hydrologist within the quantile space, a scale with which he has familiarity. For more discussion and references on prior specification, see Berger (1985) and Bernardo and Smith (1994).

Experts in Hydro-Québec do not agree about the existence of a shift either in the mean level or in the variability of annual energy inflows. Moreover, specialists who suspect the existence of a sudden change are not willing to specify their degree of belief about the year at which it occurred without referring to the data of interest. This information is therefore biased. In fact, the use of any of these data in the prior formulation is not formally acceptable in a Bayesian analysis (Berger, 1985). Finally, if we examine the annual energy inflows for the other Hydro-Québec hydropower systems, no clear change-point can be identified. Therefore, one cannot reasonably favor any year of change, and we assume τ is distributed as an uniform discrete distribution for both model M_1 and M_2 . Note that in this case, the prior expected change-point is 1970, i.e. the mean of a discrete uniform probability distribution on the interval [1943,1995].

For the other parameters, the prior degrees of belief were assumed to be represented by normal-inverted gamma type of distributions (see (A.2) for M_1 and (A.7) for M_2). Complete specification of prior knowledge only requires the choice of the hyperparameters in these expressions. Because it is felt that the annual energy inflow of all hydropower systems are related in some way, we based the determination of these quantities on a regional analysis. The information available to establish a transfer function between the site of interest and the other systems consists, for each hydropower system, of four variables: average annual

energy inflow ($y_{\bar{e}}$), variance of annual energy inflow (y_{var}), basin area (A), average annual precipitation (\bar{P}) and installed generating capacity (C). Two data sets were considered, corresponding respectively to the values of $y_{\bar{e}}$, y_{var} and \bar{P} evaluated before and after the prior expected year of change $E\{\tau\} = 1970$. Table 1 gives the first data set.

Table 1. Data set with $y_{\bar{e}}$, y_{var} and \bar{P} calculated before the prior expected change-point 1970^a

System	$y_{\bar{e}}$ (TWh)	y_{var} (TWh ²)	A (km ²)	\bar{P} (mm)	C (TW)
Bersimis	7.65	0.56	15 695	975	0.00171
Churchill Falls	35.36	28.14	92 432	891	0.00543
La Grande	79.10	102.14	176 472	764	0.01524
Churchill Falls	22.24	4.92	45 480	1 030	0.00502
Outaouais	10.05	3.62	225 280	934	0.00188
Outardes	9.92	1.28	18 798	1 008	0.00184
Saint-Maurice	9.01	2.14	51 813	1 016	0.00164
Saint-Laurent	11.62	1.41	794 000	982	0.00226

^a $y_{\bar{e}}$: average annual energy inflow; y_{var} : variance of annual energy inflow;
 A : basin area; \bar{P} : average annual precipitation; C : installed generating capacity

The idea is to construct a simple model to predict the average energy inflow $y_{\bar{e}}$ and the variance y_{var} of the energy inflows for the site of interest, before and after 1970. These predictions, along with their standard error, are then used to elicit the mean and the variance of the prior distributions. Finally, solving a simple system of equations for the first two moments leads to estimated values for the hyperparameters based on regional information. Since a regional model is used to specify hyperparameters by transferring knowledge from the nearby sites, this approach, in some way, has much in common with an empirical hierarchical Bayesian analysis (Berger, 1985).

To illustrate the procedure, let us suppose that we want to specify the hyperparameters of a normal-inverted gamma prior distribution $\mathcal{N}(\mu | \phi, \lambda \sigma^2) \mathcal{IG}(\sigma^2 | \alpha, \beta)$. If we denote the predicted values of $y_{\bar{e}}$ and y_{var} respectively by $\hat{y}_{\bar{e}}$ and \hat{y}_{var} , and their corresponding variances by $s^2(\hat{y}_{\bar{e}})$ and $s^2(\hat{y}_{var})$, it is natural to take as prior mean and variance for μ and σ^2

$$E\{\mu\} = \hat{y}_{\bar{e}}, \quad Var\{\mu\} = s^2(\hat{y}_{\bar{e}}), \quad E\{\sigma^2\} = \hat{y}_{var}, \quad Var\{\sigma^2\} = s^2(\hat{y}_{var})$$

Using expressions for prior mean and variance of normal and inverted gamma distributions (Bernardo and Smith, 1994, p.431), and then solving this system of equations leads us to

$$\alpha = 2 + \hat{y}_{var}^2 / s^2(\hat{y}_{var}), \quad \beta = \hat{y}_{var}(\alpha - 1), \quad \phi = \hat{y}_{\bar{e}} \quad \text{and} \quad \lambda = s^2(\hat{y}_{\bar{e}}) / \hat{y}_{var}. \quad (\text{A.13})$$

Linear regression analysis was performed for each data set (the first one is given in Table 1), discarding the site of interest. A simple regression considering only the generating capacity C appeared to be the best regional models. More precisely, we have

$$y_{\bar{e}} = \gamma_{\bar{e}} C + \varepsilon \quad \text{and} \quad y_{var} = \gamma_{var} C^2 + \varepsilon.$$

In the particular case of hydropower systems in Québec, precipitation and drainage area are very poorly related to the average energy inflow or the variance compared to generating capacity C . Indeed, variable C is a good summary of what an engineer would think of the

potential annual energy output the company may derive when installing a hydropower plant. It can be used as a prior knowledge of what can be « intuitively » expected from the site, before the production started, that is in the state of knowledge before collecting a long series of observations in working conditions. The estimated coefficients and the corresponding prediction along with its variance are given in Table 2.

Table 2. Results of regression analyses for Churchill Falls and Ouataouais

Systems		1943 - 1970			1971 - 1996		
		$\hat{\gamma}$	\hat{y}	$s^2(\hat{y})$	$\hat{\gamma}$	\hat{y}	$s^2(\hat{y})$
Churchill Falls	$y_{\bar{e}}$	5120	27.80	2.80	5121	27.81	4.50
	y_{var}	436980	12.88	7.45	617000	18.19	6.00
Ouataouais	$y_{\bar{e}}$	5253	9.88	11.26	5242	9.85	11.20
	y_{var}	444950	1.57	44.83	624000	2.21	34.26

Using the results in Table 2 in (A.13) leads to the values of the hyperparameters for the two series. To summarize, the prior mean and standard deviation of each parameter are listed in Table 3 for the two cases.

Table 3. Prior specifications from regional analysis

		τ	μ_1	μ_2	δ	σ^2
		Churchill Falls (model M_1)	Exp. value	1970	27.80	27.81
	Stand. deviation	16	1.67	2.12	2.70	2.73
		τ	σ_1^2	σ_2^2	η	μ
		Outaouais (model M_2)	Exp. value	1970	1.57	2.21
	Stand. deviation	16	6.70	5.85	4.03	3.36

This table will be used in the next sections to show how prior state of belief is updated by the data.

A.4.2 Churchill Falls power system

Expressions given in Appendix A are used to analyze the annual energy inflow of Churchill Falls hydropower complex (Figure 1). Figure 3 displays the annual inflows and all marginal posterior densities of the parameters of interest together with their corresponding prior density (in dotted line). We first observe how the data modified or updated prior information by comparing the prior and the posterior distributions. Such graphs are also useful tools to interpret the parameter's uncertainty. Table 4 summarizes estimated posterior quantities (mode, mean and standard deviation). These estimates are to be compared with their corresponding prior values specified in Section 4.1 (Table 3). Table 4 also gives, for each parameter, a symmetric 95% Bayesian credible interval.

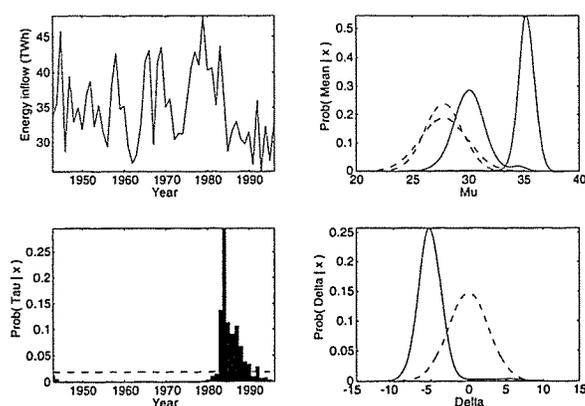


Figure 3. Marginal posterior distributions for model M_1 : Churchill Falls power system.

Table 4. Posterior moments and credible interval for Churchill Falls

Parameters	Mode	Mean	Stand. dev.	90% Credible int.
τ	1984	1984	6.99	[1981; 1990]
μ_1	35.28	35.11	0.74	[33.75; 36.28]
μ_2	30.05	30.16	1.33	[27.61; 32.76]
δ	-5.21	-4.95	1.52	[-7.86; -2.13]

This analysis indicates first, under the hypothesis of a change of type M_1 in the annual energy inflow, that the change-point occurred around 1984 with approximately a 7-year standard deviation. Also, the mean level of energy inflow after the shift seems to have decreased by an amount of 5 TWh with 1.5 TWh of standard deviation. The credible interval for the intensity of the shift shows that a decrease of almost 8 TWh is still plausible at a 90% credible level. Moreover, this interval do not contain zero which suggests a negative change of at least 2 TWh. However, this observation provides no justification for actions such as « rejecting the no change hypothesis » (the model assumes a change did occur). Under such hypothesis, the most probable intensity of the shift being -5 TWh, the hydropower company would better take action to compensate this potential loss of energy inflow and ensure balance between supply and demand for the next few years. The output of this Bayesian analysis (estimates and uncertainties) can be used to choose among different alternatives (anticipating construction of new power plants, buying electricity, using thermal power station, etc.).

One of the advantages of a Bayesian analysis is that it offers not only marginal but all joint posterior distributions, allowing for an assessment of dependencies among parameters. As an illustration, Figure 4 presents the joint posterior distribution of δ and τ . It shows in particular how, after considering the data, δ and τ are very closely related, even though we assumed prior independence between the change-point and the other parameters.

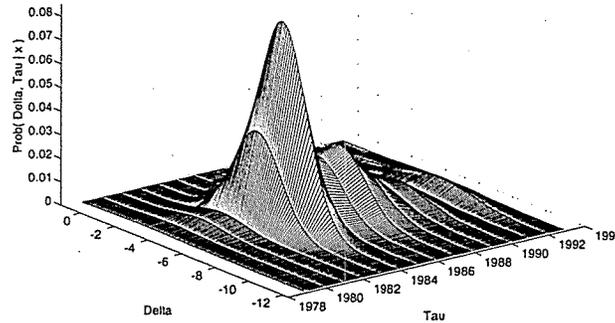


Figure 4. Joint posterior distribution of (δ, τ) for model M_1 : Churchill Falls power system.

Our approach assumes no interannual correlations on both sides of τ . Autocorrelations (lag 1 to 5) are given in Table 5 separately for the period 1943-1984 and 1985-1996.

Table 5. Autocorrelations for Churchill Falls

Lag	1	2	3	4	5
1943-1984	0.24 (0.1066)	0.17 (0.1419)	0.02 (0.2707)	-0.10 (0.3521)	-0.17 (0.3124)
1985-1996	-0.74 (0.0040)	0.49 (0.0021)	-0.22 (0.0011)	0.26 (0.0016)	-0.67 (0.0036)

Note: p -values for H_0 : « no correlation » appear in parantheses

The first subseries does not exhibit any significant interannual correlations. However, after the estimated change-point the autocorrelations are significant. One could revise the analysis by considering for example an AR model with a changing coefficient of autocorrelation. Note that the estimated correlations after the change-point are only based on ten observations. Therefore, it may not be relevant to increase the number of parameters based on such uncertain estimates.

Before analyzing the change in the variance for the annual energy inflow of the Outaouais power system using Gibbs sampler, we first outline the way that the algorithm would be typically implemented by reconsidering model M_1 . We attempt to reproduce the analytical results for the Churchill Falls example by estimating the marginal posterior densities of interest using Gibbs algorithm. Under the assumption of model M_1 and the joint prior distribution in Appendix A, the full conditional posterior distributions for each parameter are given in Appendix B. They are readily available for sampling using standard routines. Note that since $p(\tau | \mu, \sigma^2, \mathbf{x}, M_1)$ is univariate and discrete, sampling from this distribution during the Gibbs sampler cycle is straightforward by simple function inversion.

Figure 5 compares the results obtained above with those evaluated using the Gibbs sampler for the marginal posterior distributions of the change point τ (continuous line with stars at each time point) and the intensity of shift δ (dotted line). Discarding the first $t = 1000$ values of the chain, the conditional densities were averaged over $m = 1000$ replications using the Rao-Blackwellized density estimate (see (B.3) for $\hat{p}(\delta | \mathbf{x}, M_1)$). There is a clear agreement between the exact and approximate densities in this case. Further experimentation showed that the Gibbs sampler could reproduce the exact posterior distributions in considerably fewer than $t = 1000$ iterations. This is confirmed by inspection of 1000 replicates after, say, 100 iterations.

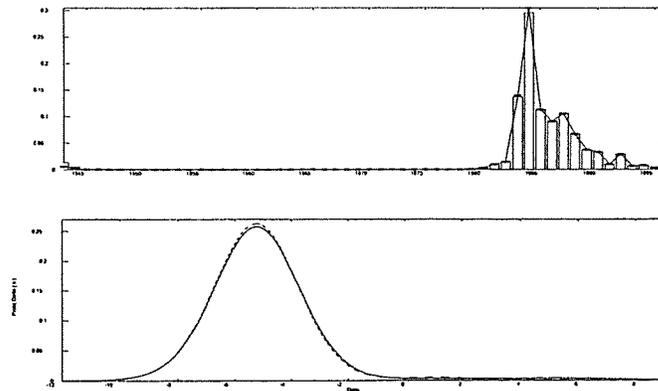


Figure 5. Comparison of exact and approximate posterior distributions ($t = 1000, m = 1000$).

A.4.3 Outaouais power system

Using the full conditionals given in Section 3 and Gibbs sampler, we analyze the annual energy inflows calculated at the Outaouais hydropower system. Inferences presented below are also based on $t = 1000$ and $m = 1000$. Figure 6 displays the annual energy inflows and all marginal posterior densities of the parameters of interest together with their corresponding prior density (in dotted line). In Table 6 we list the estimated posterior quantities and 95% credible intervals. Finally, Figure 7 presents the joint posterior distribution of η and τ .

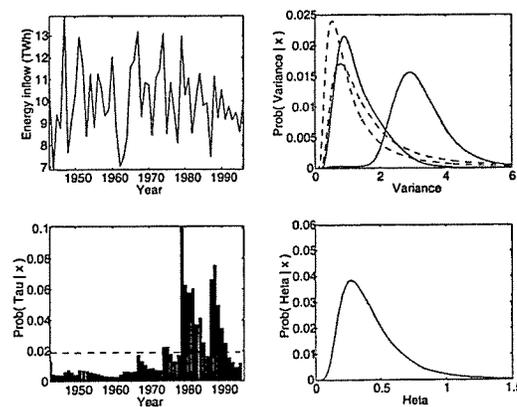


Figure 6. Marginal posterior distributions for model M_2 : Outaouais power system.

Table 6. Posterior moments and credible interval for Outaouais

Parameters	Mode	Mean	Stand. dev.	90% Credible int.
τ	1979	1978	12.81	[1948; 1991]
σ_1^2	2.77	3.03	0.77	[1.82; 4.48]
σ_2^2	0.90	1.43	0.52	[0.53; 2.80]
η	0.30	0.58	0.20	[0.15; 1.30]

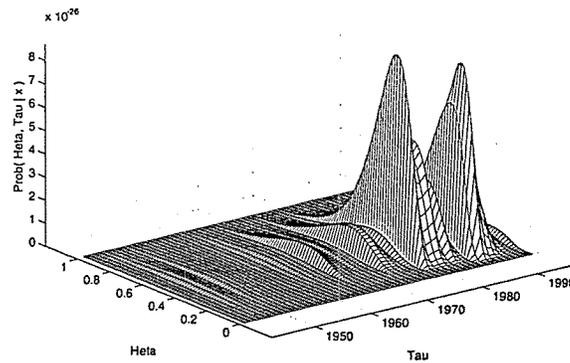


Figure 7. Joint posterior distribution of (η, τ) for model M_2 : Outaouais power system.

The analysis indicates, under the hypothesis of a change of type M_2 in the annual energy inflow, that the variance after the change is at least twice as small than it was before the change (the posterior mode and mean for η are respectively 0.30 and 0.55). The change-point seems to be 1979 with approximately 12 years of standard deviation. Comparing these values to the prior expectation (Table 3), clearly the information stemming from the data contributed to modifying considerably the prior state of belief. The credible interval for η shows that an almost sevenfold decrease is still plausible at a 90% credible level. But this interval also suggests a much smaller change, since it contains one, i.e. $\sigma_1^2 \approx \sigma_2^2$.

These results may seem statistically convincing in favor of a change, even though the procedure is not a statistical test. But, is such a decrease in the variability ($\eta \approx 0.5$) enough to change the strategy of the company's development for the future? Clearly, decisional consequences have to be considered together with the statistical analysis. If Hydro-Québec changes its policy only in light of these statistical results, the company may face costly consequences. In the particular case of a change in the variability of energy inflows, one should take into account the fact that the losses associated with the decisions, such as contracting or not new agreements to sell electricity, are probably asymmetrical. Suppose the variance is in fact one half as large as before. The uncertainty around the mean becomes smaller and the future is somewhat less variable. Therefore, the mean value, as a prediction for future observations, is more credible that it was in the past, so that new agreements to sell electricity are relatively less risky. The company can then contract for delivering more power. If, in reality, the variance has not changed as it was previously believed, future values of energy inflows may deviate from the mean much more frequently than anticipated. In this case, the company would not respect its contractual engagements, resulting in dramatic consequences. On the other hand, if the company would decide not to change its policy of development, when indeed the variance has decreased, good opportunities may be lost but catastrophic consequences would not to be encountered as in the former case.

Finally, the two subseries do not exhibit any significant interannual correlations (Table 7).

Table 7. Autocorrelations for Outaouais

Lag	1	2	3	4	5
1943-1979	-0.04 (0.7824)	-0.05 (0.9121)	-0.23 (0.4832)	-0.03 (0.6432)	0.04 (0.7643)
1980-1996	-0.35 (0.1157)	0.13 (0.2408)	-0.27 (0.2110)	0.26 (0.1831)	0.05 (0.2790)

Note: p -values for H_0 : « no correlation » appear in parantheses

A.5 Discussion and conclusions

The Bayesian method presented in this paper, can be viewed as an extension of the normal models and practitioners can perform such change-point analysis routinely using standard statistical toolboxes. This approach can be generalized to other type of normal models (a simultaneous change in the mean and variance for instance, see Perreault et al., 2000a) and even to other types of p.d.f.'s, for instance the gamma distribution. More precisely, use of probability distributions which belong to the exponential class of p.d.f.'s allows for exactly the same line of reasoning based on conjugacy.

The following conclusions have been reached in this study :

- 1) Hydrologists can take full advantage of the existence of conjugate distributions when studying a single change-point in a statistical model belonging to the exponential family. If prior independence between the epoch of change and model parameters is assumed, the joint posterior distribution is a finite mixture of conjugate distributions. This allows for easy computation of posterior odds, as illustrated by the univariate normal model under the configuration of a single change in the mean level.
- 2) It was showed that more complex change-point problems can be readily addressed by Gibbs sampling, which is easily accessible to the average statistical practitioner. This was illustrated in the case of a single change in the variance for which no explicit expressions for the posterior distributions can be found, unless one considers noninformative priors. In hydrology, the case of a single change in the variance occurring at an unknown time point was neglected although this parameter is important for water resources management. The Gibbs sampling scheme could also be used if one is interested to relax the independence assumption. For instance, change-point in AR models may be treated by generalizing the approach proposed by Chib (1993).
- 3) Two case studies involving annual energy inflow series were presented to illustrate how Bayesian change-point can be performed, from the specification of the prior distributions up to the interpretation of posterior odds. It was shown that several questions can be answered simultaneously from the posterior distributions. For instance, one can not only estimate the parameters, but can directly obtain accuracy measures for the estimate and interpret dependence between them. This is in contrast to classical statistics, for which obtaining estimates and determining their reliability are two different problems. It was also seen that interpretation of the results is easy and straightforward. The frequentist interpretation is difficult to understand for a non-statistician, and water resources managers often find it much too abstract for operational concerns. For example, interpretation of classical confidence intervals involves averaging over all possible data, while it is known which data did in fact occur. Practitioners usually interpret the results of a classical confidence interval as a Bayesian credible interval (Lecoutre, 1997). Moreover, recourse to the usual asymptotic assumption required for classical inference (see Hinkley, 1970) is not needed here, and therefore Bayesian change-point analysis can be applied to modest sample sizes.
- 4) Unlike many published Bayesian analyses, the problem of eliciting prior knowledge was not skirted by using noninformative priors or even worse, by specifying hyperparameters based on a subset of the actual sample. To ensure maximum efficiency, all observations must in fact be used to update the prior state of belief, and not to specify hyperparameters. In this paper, prior elicitation relied on expert knowledge and on external information about the studied phenomena. A simple regional analysis was performed to specify the hyperparameters and represent the prior state of belief. Effort in this direction should be developed to take full advantage of the Bayesian framework.

The problem of model uncertainty has not been addressed here. All results herein are conditional upon a given model. Both M_1 and M_2 assume with certainty that a change has occurred. This statistical reporting framework is essential but incomplete. It forces the hydrologist to choose among the different models to perform a change-point analysis, and to exclude some other kinds of changes. To take into account the « no change hypothesis » and diverse types of changes, one should assign a prior probability to the different alternatives, and consider the change-point study as a Bayesian model selection problem among the various situations that may occur. Using the results presented in this paper, this perspective

for hydrological change-point analysis is explored in Perreault et al. (2000a). The important operational issue of forecasting in changing sequences is also studied.

Acknowledgments. We wish to express our deep gratitude to Professor Lucien Duckstein for his critical review of this paper. We are also grateful to René Roy and Raymond Gauthier of Hydro-Québec for providing data. Finally, we gratefully acknowledge the constructive comments of two anonymous referees on an earlier version that considerably improved this paper. This research was supported by la Direction Générale de l'Enseignement et de la Recherche de France of the French ministry of Agriculture (DGER), les fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec (FCAR), and la Chaire en Hydrologie Statistique Hydro-Québec/CRSNG (INRS-Eau).

References

- Berger, J.O., 1985. *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York.
- Bernardo, J.M., Smith, A.F.M., 1994. *Bayesian theory*, Wiley, New York.
- Bernier, J., 1994. Statistical detection of changes in geophysical series, in *Engineering Risk in Natural Resources Management*, edited by L. Duckstein and E. Parent, NATO Advanced Studies Institute Series, Kluwer, 159-176.
- Bernier, J., 1967. Les méthodes bayésiennes en hydrologie statistique, in *Proceedings of the first Hydrology Symposium*, Fort Collins, 461-470.
- Booth, N.B., Smith, A.F.M., 1982. A Bayesian approach to retrospective identification of change-points, *Jour. of Econometrics* 19, 7-22.
- Box, G.E.P, Tiao, G.C., 1973. *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading, Massachusetts.
- Broemeling, L.D., 1985. *Bayesian Analysis of Linear Models*, Marcel Dekker Inc., Volume 60, Collection Statistics: Textbooks and Monographs.
- Bruneau, P., Rassam, J.-C., 1983. Application d'un modèle bayésien de détection de changements de moyennes dans une série, *Journal des Sciences Hydrologiques* 28, 341-354.
- Campbell, E.P., Fox, D.R., Bates, C.B., 1999. A Bayesian approach to parameter estimation and pooling in nonlinear flood event models, *Water Resour. Res* 35, 211-220.
- Carlin, B.P., Gelfand, A.E., Smith, A.F.M., 1992. Hierarchical Bayesian analysis of changepoint problems, *Appl. Statist.* 41, 389-405.
- Chernoff, H., Zacks, S., 1963. Estimating the current mean of a normal distribution which is subjected to changes in time, *Ann. Math. Statist.* 35, 999-1028.
- Chib, S., 1993. Bayes regression with autoregressive errors. A Gibbs sampling approach. *Jour. of Econometrics* 58, 275-294.
- Coles, S.G., Tawn, J.A., 1996. A Bayesian analysis of extreme rainfall data, *Appl. Statist.* 45, 463-478.
- Gelfand, A.E., Smith, A.F.M., 1990. Sampling-based approaches to calculating marginal densities, *J. Am. Stat. Assoc.* 85, 398-409.
- Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machin Intelligence* 6, 721-741.
- Hinkley, D.V., 1970. Inference about the change-point in a sequence of random variables, *Biometrika* 57, 1-16.
- Kerr, R.A., 1992. Unmasking a shifty climate system, *Science* 255, 1508-1510.
- Kerr, R.A., 1999. A new force in high-latitude climate, *Science* 284, 241-242.
- Kuczera, G., Parent, E., 1998. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm, *J. of Hydrol.* 211, 69-85.
- Lecoutre, B., 1997. C'est bon à savoir! Et si vous étiez un bayésien qui s'ignore, *Modulad* 18, 81-87.

- Lee, A.S.F., Heghinian, S.M., 1977. A shift of the mean level in a sequence of independent normal random variables - A Bayesian approach, *Technometrics* 19, 503-506.
- Lehmann, E.L., 1983. *Theory of Point Estimation*, Wiley, New York.
- Lu, Z.Q., Berliner, L.M., 1999. Markov switching time series models with application to a daily runoff series, *Water Resour. Res.* 35, 523-534
- Menzefricke, U., 1981. A Bayesian analysis of a change in the precision of a sequence of independent normal random variables at an unknown time point, *Appl. Statist.* 30, 141-146.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equations of state calculations by fast computing machines, *Journal of Chemical Physics* 21, 1087-1091.
- Paturel J.E., Servat E., Kouamé B., Lubès H., Ouedraogo M., Masson J.M., 1997. Climatic variability in humid Africa along the Gulf of Guinea Part II: an integrated regional approach, *J. of Hydrol.* 191, 16-36.
- Perreault, L., Bernier, J., Bobée, B. and Parent, E., 2000a. Bayesian change-point analysis in hydrometeorological time series 2. Comparison of change-point models and forecasting, submitted to *J. of Hydrol.*
- Perreault, L., Parent, E., Bernier, J., Bobée, B. and Slivitzky, M., 2000b. Retrospective multivariate Bayesian change-point analysis: A simultaneous single change in the mean of several hydrological sequences, submitted to *Stochastic Environmental Research an Risk Assessment*.
- Perreault, L., Haché, M., Slivitsky, M., Bobée, B., 1999. Detection of changes in precipitation and runoff over eastern Canada and U.S. using a Bayesian approach, *Stochastic Environmental Research an Risk Assessment* 13, 201-216.
- Robert, C., 1998. *Discretization and MCMC Convergence Assessment*, Lecture Notes in Statistics 135, Springer-Verlag, New York.
- Robert, C., 1996. *Méthodes de Monte Carlo par chaînes de Markov*, Economica, Paris.
- Rodriguez-Iturbe, I., Entekhabi, D., Bras, R.L., 1991. Nonlinear dynamics of soil moisture at climate scales 1. *Stochastic analysis*, *Water Resour. Res.* 27, 1899-1906.
- Servat E., Paturel J.E., Lubès H., Kouamé B., Ouedraogo M., Masson J.M., 1997. Climatic variability in humid Africa along the Gulf of Guinea Part I: detailed analysis of the phenomenon in Côte d'Ivoire, *J. of Hydrol.* 191, 1-15.
- Smith, A.F.M., 1975. A Bayesian approach to inference about change-point in sequence of random variables, *Biometrika* 62, 407-416.
- Stephens, D.A., 1994. Bayesian retrospective multiple-change-point identification, *Appl. Statist.* 43, 159-178.
- Tanner, M.H., 1992. *Tools for Statistical Inference: Observed Data and Data Augmentation Methods*, Lecture Notes in Statistics 67, Springer-Verlag, New York.

Appendix

A. Bayesian inference for model M_1

The likelihood function resulting from n observations $\mathbf{x} = (x_1, x_2, \dots, x_n)$ generated by model M_1 can be written as

$$\begin{aligned}
 p(\mathbf{x} \mid \mu, \sigma^2, \tau, M_1) &= \prod_{i=1}^{\tau} \mathcal{N}(x_i \mid \mu_1, \sigma^2) \prod_{i=\tau+1}^n \mathcal{N}(x_i \mid \mu_2, \sigma^2) & (A1A.1) \\
 &= \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} \exp \left\{ -\frac{\tau}{2\sigma^2} \left[s_{\tau}^2 + (\bar{x}_{\tau} - \mu_1)^2 \right] \right\} \\
 &\quad \times \exp \left\{ -\frac{(n-\tau)}{2\sigma^2} \left[s_{n-\tau}^2 + (\bar{x}_{n-\tau} - \mu_2)^2 \right] \right\}
 \end{aligned}$$

where $\mu = (\mu_1, \mu_2)$ and

$$\bar{x}_\tau = \sum_{i=1}^{\tau} \frac{x_i}{\tau}, \quad \bar{x}_{n-\tau} = \sum_{i=\tau+1}^n \frac{x_i}{n-\tau}, \quad s_\tau^2 = \sum_{i=1}^{\tau} \frac{(x_i - \bar{x}_\tau)^2}{\tau}, \quad s_{n-\tau}^2 = \sum_{i=\tau+1}^n \frac{(x_i - \bar{x}_{n-\tau})^2}{n-\tau}.$$

For fixed τ , the likelihood (A.1) has the structure of a product of two normal distributions with one inverted gamma distribution, which suggests an ad hoc normal-inverted gamma type of distribution to represent prior knowledge about μ and σ^2 . Assuming for model M_1 prior independence between τ and the other parameters (μ, σ^2) , and that $p(\tau | M_1)$ is any discrete distribution on the set $\{1, 2, \dots, n-1\}$, leads to the joint prior parameter p.d.f.:

$$\begin{aligned} p(\mu, \sigma^2, \tau | M_1) &= \mathcal{N}(\mu_1 | \phi_1, \lambda_1 \sigma^2) \mathcal{N}(\mu_2 | \phi_2, \lambda_2 \sigma^2) \mathcal{IG}(\sigma^2 | \alpha, \beta) p(\tau | M_1) \\ &= \mathcal{NNIG}(\mu, \sigma^2 | \phi, \lambda, \alpha, \beta) p(\tau | M_1) \end{aligned} \quad (\text{A1A.2})$$

where $\phi = (\phi_1, \phi_2)$, $\lambda = (\lambda_1, \lambda_2)$, and $\mathcal{IG}(x | \alpha, \beta)$ stands for the inverted gamma p.d.f. with parameters $\alpha \in \mathbb{R}^+$ and $\beta \in \mathbb{R}^+$ (see Bernardo and Smith, 1994, for the density).

Because of conjugate properties (Berger, 1985), under M_1 , the conditional joint posterior distribution $p(\mu, \sigma^2 | \tau, \mathbf{x}, M_1)$ given τ and the observed data \mathbf{x} also belongs to the class of normal-inverted gamma distributions, but with updated parameters $(\phi', \lambda', \alpha', \beta')$. More precisely,

$$p(\mu, \sigma^2 | \tau, \mathbf{x}, M_1) = \mathcal{NNIG}(\mu, \sigma^2 | \phi', \lambda', \alpha', \beta'), \quad (\text{A1A.3})$$

where

$$\begin{aligned} \lambda'_1 &= \lambda_1 / (1 + \tau \lambda_1), \quad \phi'_1 = (1 - \lambda'_1 \tau) \phi_1 + \lambda'_1 \tau \bar{x}_\tau, \\ \lambda'_2 &= \lambda_2 / [1 + (n - \tau) \lambda_2], \quad \phi'_2 = [1 - \lambda'_2 (n - \tau)] \phi_2 + \lambda'_2 (n - \tau) \bar{x}_{n-\tau}, \\ \beta' &= \frac{\tau}{2} \left[s_\tau^2 + (1 - \lambda'_1 \tau) (\phi_1 - \bar{x}_\tau)^2 \right] + \frac{n - \tau}{2} \left[s_{n-\tau}^2 + (1 - \lambda'_2 (n - \tau)) (\phi_2 - \bar{x}_{n-\tau})^2 \right] + \beta, \\ \alpha' &= \alpha + n/2. \end{aligned}$$

To simplify notation, dependence upon τ and \mathbf{x} was omitted by writing for example λ'_1 instead of $\lambda'_1(\mathbf{x}, \tau)$. The prior predictive density, i.e. the p.d.f. of the data \mathbf{x} only conditioned upon the change-point τ , $p(\mathbf{x} | \tau, M_1)$, which makes the Bayes theorem denominator, can be determined by dividing the joint p.d.f. $p(\mu, \sigma^2, \mathbf{x} | \tau, M_1)$ by (A.3), and cancelling factors involving μ and σ^2 . The result is

$$p(\mathbf{x} | \tau, M_1) = \left(\frac{1}{2\pi} \right)^{n/2} \sqrt{\frac{\lambda'_1 \lambda'_2}{\lambda_1 \lambda_2} \frac{\beta^\alpha}{(\beta')^{\alpha'}} \frac{\Gamma(\alpha')}{\Gamma(\alpha)}}. \quad (\text{A1A.4})$$

Integrating the appropriate parameters out of (A.3), leads to the conditional posterior distributions $p(\mu_1 | \tau, \mathbf{x}, M_1)$ and $p(\mu_2 | \tau, \mathbf{x}, M_1)$ of the means before and after the change-point. The conditional posterior density of the intensity of the shift $\delta = \mu_2 - \mu_1$, $p(\delta | \tau, \mathbf{x}, M_1)$, is then deduced by a simple univariate transformation of variable. These distributions are Student t -distributions, $\mathcal{ST}(a, b, c)$ (expression for the p.d.f. can be found in Bernardo and Smith, 1994, p.432):

$$\begin{aligned}
p(\mu_i | \tau, \mathbf{x}, M_1) &= ST(\phi'_i, \alpha' (\lambda'_i \beta')^{-1}, 2\alpha'), \quad i = 1, 2 \\
p(\delta | \tau, \mathbf{x}, M_1) &= ST(\phi'_2 - \phi'_1, \alpha' (\lambda'_1 + \lambda'_2)^{-1} (\beta')^{-1}, 2\alpha')
\end{aligned}$$

In our problem, τ is unknown and its marginal posterior distribution $p(\tau | \mathbf{x}, M_1)$ has to be derived. Using Bayes theorem and the prior predictive density (A.4), the marginal posterior density of the change-point $\tau = 1, 2, \dots, n - 1$ under model M_1 is seen to be

$$p(\tau | \mathbf{x}, M_1) = \frac{p(\mathbf{x} | \tau, M_1) p(\tau | M_1)}{\sum_{\tau=1}^{n-1} p(\mathbf{x} | \tau, M_1) p(\tau | M_1)} \propto p(\tau | M_1) \sqrt{\lambda'_1 \lambda'_2} (\beta')^{-\alpha'} \quad (\text{A1A.5})$$

This distribution is discrete and gives, for each time point, the posterior probability of shift occurrence in the mean level assuming a change occurred with certainty. Finally, to draw conclusions regarding the means before and after the change-point and the intensity of shift, the marginal posterior distributions $p(\mu_1 | \mathbf{x}, M_1)$, $p(\mu_2 | \mathbf{x}, M_1)$ and $p(\delta | \mathbf{x}, M_1)$ must be evaluated. This is done by averaging the corresponding conditional posterior distribution with respect to the posterior mass function of τ :

$$\begin{aligned}
p(\mu_i | \mathbf{x}, M_1) &= \sum_{\tau=1}^{n-1} p(\mu_i | \tau, \mathbf{x}, M_1) p(\tau | \mathbf{x}, M_1) \\
p(\delta | \mathbf{x}, M_1) &= \sum_{\tau=1}^{n-1} p(\delta | \tau, \mathbf{x}, M_1) p(\tau | \mathbf{x}, M_1)
\end{aligned}$$

where $i = 1, 2$. The marginal distributions of μ_1 , μ_2 and δ appear as finite mixtures of Student t -distributions weighted by the $n - 1$ values of (A.5).

B. Full conditionals for model M_1

The collection of full conditionals for model M_1 are all proportional to the product

$$p(\mathbf{x} | \mu, \sigma^2, \tau, M_1) p(\mu, \sigma^2, \tau | M_1), \quad (\text{A1B.1})$$

and can be seen to be

$$\begin{aligned}
p(\mu_1 | \mu_2, \sigma^2, \tau, \mathbf{x}, M_1) &= \mathcal{N}(\mu_1 | \phi'_1, \lambda'_1 \sigma^2) \\
p(\mu_2 | \mu_1, \sigma^2, \tau, \mathbf{x}, M_1) &= \mathcal{N}(\mu_2 | \phi'_2, \lambda'_2 \sigma^2) \\
p(\sigma^2 | \mu, \tau, \mathbf{x}, M_1) &= \text{IG}\left(\sigma^2 \left| \alpha' + 1, \frac{1}{2} \left[\frac{(\mu_1 - \phi'_1)^2}{\lambda'_1} + \frac{(\mu_2 - \phi'_2)^2}{\lambda'_2} \right] + \beta' \right.\right) \\
p(\tau | \mu, \sigma^2, \mathbf{x}, M_1) &\propto \frac{p(\mathbf{x} | \mu, \sigma^2, \tau, M_1) p(\tau | M_1)}{\sum_{\tau=1}^{n-1} p(\mathbf{x} | \mu, \sigma^2, \tau, M_1) p(\tau | M_1)}
\end{aligned}$$

The other parameter of interest is $\delta = \mu_2 - \mu_1$. As in Section 3, we transformed μ_2 to δ to obtain a full conditional distribution:

$$d(\delta | n_1, \sigma_2^2, \tau, \mathbf{x}, M_1) = \mathcal{N}(\delta | \phi_1^z, \chi^2 \sigma_2^2 - n_1, \chi^2 \sigma_2^2) \quad (\text{A1B.2})$$

and used it to evaluate a Rao-Blackwellized estimate for $d(\delta | \mathbf{x}, M_1)$:

$$\hat{d}(\delta | \mathbf{x}, M_1) = \sum_{m=1}^M \frac{m}{M} d(\delta | n_1^f, \sigma_2^f, \tau^f, \mathbf{x}, M_1) \quad (\text{A1B.3})$$

Annexe B

Article II

Bayesian change-point analysis in hydrometeorological time series 2. Comparison of change-point models and forecasting

Luc Perreault^{1,2}, Jacques Bernier^{1,2}, Bernard Bobée², Éric Parent¹
¹ENGREF ²INRS-Eau

Journal of Hydrology (2000), 235 : 242-263

Research and drafting of the document have been performed at ENGREF.

Abstract

This paper provides a methodology to test existence, type, and strength of changes in the distribution of a sequence of hydrometeorological random variables. Unlike most published work on change-point analysis, which consider a single structure of change occurring with certainty, it allows for the consideration in the inference process of the no change hypothesis and various possible situations that may occur. The approach is based on Bayesian model selection and is illustrated using univariate normal models. Four univariate normal models are considered : the no change hypothesis, a single change in the mean level only, a single change in the variance only, and a simultaneous change in both the mean and the variance. First, inference analysis of posterior distributions via Gibbs sampling for a given change-point model is recalled. This scientific reporting framework is then generalized to the problem of selecting among different configurations of a single change and the no change hypothesis. The important operational issue of forecasting a future observation, often neglected in the literature on change-point analysis, is also treated in the previous model selection perspective. To illustrate the approach, a case study involving annual energy inflows for eight large hydropower systems situated in Québec is detailed.

Keywords. Change-point analysis ; Energy inflows ; Bayesian model selection ; Forecasting ; Model uncertainty ; Gibbs sampling

B.1 Introduction

It is argued in this paper that, when a change in meteorological time series at an unknown epoch is suspected, hydrologists must entertain a range of possible belief models, at least one which assumes a given type of change and another which represents the no change hypothesis. When facing important decisions based on these data, decision makers should

not put all their eggs in the same basket. In fact, any individual who considers only a single model for a decision problem, for instance forecasting future observations, ignores model uncertainty, which can be a major part of overall uncertainty about quantities of interest. As hydrological studies are mainly based on samples of limited sizes, this issue becomes very important for change-point analysis since uncertainty about the existence and the configuration of a change will always remain.

Following the emergence of published studies on climate changes, a number of hydrologists have used models which describe certain types of changes to represent hydrometeorological time series. Bayesian methods to study a single unknown change-point in a sequence of random variables have been increasingly used in the hydrological literature. In most of these papers, a given type of change (often in the mean level) occurring with certainty is assumed, and focus has been put on the characterization of the change. Even if such inference is vital to help with decision making, it is not sufficient. First, conclusions stemming from this type of inference are entirely conditional upon the given configuration of change that the model assumes (for instance, in the mean only). This scientific reporting framework in turn forces the hydrologist to choose one situation among different possible models to perform the change-point analysis, and consequently to exclude all other kinds of changes. Second, these models proposed in the hydrological literature always assume a change with certainty. Therefore, the analysis of posterior distributions of parameters can not be used, in addition to inference about the characteristics of a possible change, to « test » its existence. However, although the no change hypothesis does not have any status in this type of formulation, such questionable analysis can be found in published papers for change-point models, namely in Bruneau and Rassam (1983), Rao and Tirtojondro (1996) and Lubès-Niel et al. (1998). In fact, these authors use posterior estimates stemming from Bayesian analysis under such models to reject or accept the no change hypothesis. Entertaining several models in change-point analysis allows for a clear formulation of each possible situations, and therefore to proceed adequately when diagnosis about the different alternatives is needed.

In a companion paper, Perreault et al. (2000) presented a Bayesian approach for the inference about the parameters of a single change-point model. A sudden change in the mean level and in the variance in a sequence of normal random variables were investigated. This article provides a more general formulation of change-point analysis which allows us to take into account model uncertainty when making decisions. Single change-point analysis is viewed as a model choice problem among the various possible situations that may occur. Bayesian model selection is adopted here under the perspective where the range of models under consideration is assumed to include the « true » belief model. Two different decision problems are examined within this framework: (1) the verification of the existence of a change and the identification of its type, i.e. only model choice, and (2) model choice followed by the prediction of a future observation. Although these decision problems can not be viewed in water resources management as terminal actions, they are very important steps in the decision-making process. As in Perreault et al. (2000), the purpose of this paper is not to determine why changes occurred (e.g. climatic change or site-specific factors), but is only concerned with developing statistical tools to help decision making.

In Section 2, we recall the general formulation of a single change-point model and show how marginal posterior distributions of the parameters of interest can be evaluated analytically or, if necessary, via a straightforward iterative Markov Chain Monte Carlo method, namely the Gibbs sampler. In Section 3, we show how predictive distributions can be used to test whether a change has in fact occurred, to identify its type among several alternatives, and finally to derive a prediction which takes into account the no change hypothesis and all configurations of change. Section 4 is dedicated to the evaluation of the predictive densities when no expression in closed form can be obtained. The methods are based on posterior simulation outputs by Gibbs sampling. Finally, this new perspective for change-point analysis is illustrated for a sequence of independent normal random variables in Section 5. Four models are considered: the no change hypothesis, a single change in the mean level only, a single change in the variance only, and a simultaneous change in both the mean and the variance. An application to eight series of Hydro-Québec annual energy inflows is offered.

B.2 Inference for univariate single change-point models

The general formulation for a single change-point model assumes a sequence of n independent random variables $\mathbf{X} = (X_1, X_2, \dots, X_n)$, for instance Hydro-Québec annual energy inflows, such that

$$\begin{aligned} X_i &\sim p(x_i | \theta_1), & i = 1, \dots, \tau \\ X_i &\sim p(x_i | \theta_2), & i = \tau + 1, \dots, n \end{aligned} \quad (\text{B.1})$$

where the density $p(x | \theta)$ belongs to a known class of probability densities indexed by a parameter θ taking values in the set Θ , so that $\theta_1 \neq \theta_2$. The parameter $\tau = 1, 2, \dots, n-1$ is the unknown change-point. Model (B.1) assumes that the time series exhibits an abrupt change. Indeed, in hydrometeorological time series, other types of changes such as trends can exist. However, recent understanding of global climate interactions such as the El Niño/La Niña and the North Atlantic Oscillation phenomena give credence to the idea that climate may operate in two or more quasi-stationary states, and that it can rapidly switch from one state to another (Rodríguez-Iturbe, 1991; Kerr, 1992, 1999). Hence, a sudden change may be representative for several hydrological and climatic time series. The independence assumption between successive values can also be questioned. Hydro-Québec annual energy inflows are proportional to the net basin supplies, which in turn are evaluated following the water balance equation. Therefore, the data considered herein are implicitly adjusted for surface and subsurface storage effect that could induce interannual correlations, and the assumption of independence seems reasonable to us, at least as a first methodological step. It will be relaxed in further studies.

The corresponding likelihood function for the realizations $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is

$$p(\mathbf{x} | \theta_1, \theta_2, \tau) = \prod_{i=1}^{\tau} p(x_i | \theta_1) \prod_{i=\tau+1}^n p(x_i | \theta_2) \quad (\text{B.2})$$

In the Bayesian perspective, a joint prior distribution $p(\theta_1, \theta_2, \tau)$ is assumed for the parameters. Bayes theorem provides the joint posterior distribution $p(\theta_1, \theta_2, \tau | \mathbf{x})$ of θ_1, θ_2, τ given the data \mathbf{x} . This p.d.f. is proportional to

$$p(\mathbf{x} | \theta_1, \theta_2, \tau) p(\theta_1, \theta_2, \tau) \quad (\text{B.3})$$

Evaluation of this joint density and all desired marginal posterior densities can be a very difficult task. But the Gibbs sampler enables a straightforward solution to such problem if conjugate prior distributions for fixed τ are assumed (see Carlin et al., 1992; Stephens, 1994; Perreault et al., 2000). Implementation requires sampling from the full conditional distributions $p(\tau | \theta_1, \theta_2, \mathbf{x})$, $p(\theta_1 | \theta_2, \tau, \mathbf{x})$ and $p(\theta_2 | \theta_1, \tau, \mathbf{x})$, which are all proportional to (B.3). Note that $p(\tau | \theta_1, \theta_2, \mathbf{x})$ is exactly of the form

$$p(\tau | \theta_1, \theta_2, \mathbf{x}) = \frac{p(\mathbf{x} | \theta_1, \theta_2, \tau) p(\tau)}{\sum_{\tau=1}^{n-1} p(\mathbf{x} | \theta_1, \theta_2, \tau) p(\tau)} \quad (\text{B.4})$$

and can be easily sampled since it involves a discrete density.

As an illustration for inference about the parameters of change-point models, let us consider, as in most published approaches on change-point studies, a sequence of independent

normal random variables X_1, X_2, \dots, X_n . The normality assumption is appropriate as a first step to analyze a single change in Hydro-Québec large hydropower system annual energy inflows. One can invoke the central limit theorem to justify the normal assumption, since annual energy inflows for a given hydropower system are calculated as a summation over time and space of the monthly energy inflows. Even if the normal model is rather simple, it allows for various types of changes, which are interesting in hydrology. Considering only a single change as in (B.1) along with the no change hypothesis, some basic situations with the usual parametrization of the normal model are as follows :

$$\begin{aligned} \text{Model } M_0: & \quad X_i \sim \mathcal{N}(x_i | \mu, \sigma^2), \quad i = 1, \dots, n \\ \text{Model } M_1: & \quad X_i \sim \mathcal{N}(x_i | \mu_1, \sigma^2), \quad i = 1, \dots, \tau \\ & \quad X_i \sim \mathcal{N}(x_i | \mu_2, \sigma^2), \quad i = \tau + 1, \dots, n \\ \text{Model } M_2: & \quad X_i \sim \mathcal{N}(x_i | \mu, \sigma_1^2), \quad i = 1, \dots, \tau \\ & \quad X_i \sim \mathcal{N}(x_i | \mu, \sigma_2^2), \quad i = \tau + 1, \dots, n \\ \text{Model } M_3: & \quad X_i \sim \mathcal{N}(x_i | \mu_1, \sigma_1^2), \quad i = 1, \dots, \tau \\ & \quad X_i \sim \mathcal{N}(x_i | \mu_2, \sigma_2^2), \quad i = \tau + 1, \dots, n \end{aligned}$$

where $\mathcal{N}(x_i | \mu, \sigma^2)$ stands for the normal probability density function (p.d.f.) with parameters $\mu \in \mathfrak{R}$ and $\sigma \in \mathfrak{R}^+$. Assuming conjugate prior distributions and independence between τ and the other parameters, all posterior marginals can be obtained in closed form for models M_0 , M_1 and M_3 . Inference about the parameters can therefore be made directly to characterize each of these situations. Expressions for the posterior distributions for model M_0 are well-known results and can be found, for instance, in Berger (1985). Those corresponding to the change-point model M_1 are given in Perreault et al. (2000). The general idea is to first evaluate the joint and marginal posterior distributions of μ_1 , μ_2 and σ^2 assuming the change-point τ is known. Then, the unconditional posterior densities are deduced by averaging these distributions over τ . The posterior marginals appear to be finite mixtures of Student t -distributions conditional upon τ , each weighted by the corresponding posterior mass $p(\tau | \mathbf{x})$. This last discrete density can be deduced from the Bayes theorem denominator given τ , which is easily evaluated following the conjugacy properties. In a similar way, all desired posterior marginals can be derived analytically for the more general situation of a single change in both the mean and the variance, i.e. model M_3 . In our opinion, this model is important for hydrometeorological series, since a shift in the mean level often seems to come with a change in variability. In Appendix A, the posterior distributions of the parameters of interest, assuming this type of change, are derived using conjugate priors. Results for M_0 and M_1 assuming conjugate priors are also recalled in Appendix A. Finally, a single change only in the variance (model M_2) does not lead to solutions in closed form as for M_0 , M_1 and M_3 . However, the Gibbs sampler provides an elegant and convenient answer to this problem (details can be found in Perreault et al., 2000).

Up to this point, exactly one change of a given type has been assumed, i.e. it is known that a change of this configuration has occurred. Therefore, the no change hypothesis has no status in that formulation nor any other type of change, and inference about the existence of a change or about its configuration cannot satisfactorily be made based only on a single model. For example, if only M_1 is considered, examination of the posterior density of $\delta = \mu_2 - \mu_1$ to assess the no change hypothesis is not valid. Bayesian perspective adopted in this paper has little role for the non-Bayesian concept of hypothesis test, especially where these relate to point null hypotheses such as $H_0 : \delta = \mu_2 - \mu_1 = 0$ (no change). Inferences about a specific value δ under a given hypothesis (change) in the absence of a well specified alternative (as in classical statistics) can not be made. In order for a Bayesian analysis to yield a nonzero probability for a point null hypothesis, it must begin with a nonzero prior probability for that hypothesis. However, δ is a continuous parameter and the probability for an event $\{\delta = 0\}$ is zero. Therefore, actions such as « accept or reject no change », i.e. M_0 , based on the examination of a credible interval about δ under M_1 is not justified. That is

why we cannot test the existence of a change only on the basis of a change-point model, and must consider separately the alternative M_0 in order that the no change situation becomes a possible outcome. A well known simple example which illustrates the problem is the so-called experiment on used electronic and quantum-mechanical random event generators with visual feedback. This design can be modelled by a Binomial experiment. Suppose a Bernoulli trial where $\theta = \Pr(\text{success})$, $n = 104\,490\,000$ trials and X is defined as the number of successes. Therefore X follows a $\text{Bin}(n, \theta)$. Let the realization of X be $x = 52\,263\,446$. We want to test $H_0 : \theta = 1/2$ against $H_1 : \theta \neq 1/2$. With uniform prior $p(\theta)$, we obtain as a 95% credible interval $\theta \in (0.50008, 0.50027)$ under H_1 . On the other hand, if we consider model H_0 and assume $\Pr(H_0) = \Pr(H_1) = 1/2$, we have $\Pr(H_0 | x) = 0.94$, which highly favor H_0 . Clearly, one can not do Bayesian tests using posterior distribution under a given hypothesis without considering an alternative. Therefore, the approach should be extended by including in the analysis the possibility of no change, and all other types of change likely to occur over the period of observation. Bayesian model selection is a solution to such a problem, since it allows for the consideration of a range of possible models in the inference process.

B.3 Entertaining a range of possible models for change-point analysis

In this section, Bayesian model selection is presented in the context of change-point analysis. Two different decision problems are examined within this perspective: the verification of the existence of a change and the identification of its type, if any, and the prediction of a future observation. But first, predictive analysis, on which model selection and forecasting are based, is briefly recalled. In the rest of the paper, since a range of possible models is now considered, the model indicator M_k is introduced into the list of unknown parameters.

B.3.1 Definitions

The general problem of statistical prediction may be described as that of inferring the values of unknown observable variables, say \mathbf{z} , from the current state of belief denoted H . In the Bayesian perspective, this is done through the predictive distribution, which is obtained by integrating the parameters ψ out of the joint density $p(\mathbf{z}, \psi | H)$ conditional upon the actual state of belief H :

$$p(\mathbf{z} | H) = \int p(\mathbf{z} | \psi, H) p(\psi | H) d\psi \quad (\text{B.5})$$

If the change-point model (B.1), denoted M_k , is assumed before any data $\mathbf{x} = (x_1, x_2, \dots, x_n)$ are considered, then $\mathbf{z} = \mathbf{x}$, $\psi = (\theta_1, \theta_2, \tau)$ and $H = M_k$. Therefore, assuming independence between τ and (θ_1, θ_2) , the predictive distribution of the unknown but observable \mathbf{x} is

$$\begin{aligned} p(\mathbf{x} | M_k) &= \sum_{\tau=1}^{n-1} \iint p(\mathbf{x} | \theta_1, \theta_2, \tau, M_k) p(\theta_1, \theta_2, \tau | M_k) d\theta_1 d\theta_2 \\ &= \sum_{\tau=1}^{n-1} p(\tau | M_k) p(\mathbf{x} | \tau, M_k) \end{aligned} \quad (\text{B.6})$$

This density is often called the marginal distribution of \mathbf{x} , but a more informative name is the prior predictive distribution. That is, this density is the probability of seeing the data that were actually observed assuming a given change-point model, calculated before

any data became available. This predictive density, implied by the likelihood and the prior distribution, provides a basis for assessing the compatibility of the data with our prior state of belief, i.e. model M_k .

After the data \mathbf{x} , assumed generated by model M_k , have been observed, we can predict an unknown future value y from the same process. In this case, $\mathbf{z} = y$, $\psi = (\theta_1, \theta_2, \tau)$ and $H = (\mathbf{x}, M_k)$. Using elementary probability rules (Berger, 1985), this new predictive distribution $p(y|\mathbf{x}, M_k)$ appears as finite mixtures of the predictive densities for fixed τ , weighted by the appropriate posterior probability of a shift occurring at that point:

$$\begin{aligned} p(y|\mathbf{x}, M_k) &= \sum_{\tau=1}^{n-1} \iint p(y|\theta_1, \theta_2, \tau, M_k) p(\theta_1, \theta_2, \tau|\mathbf{x}, M_k) d\theta_1 d\theta_2 \\ &= \sum_{\tau=1}^{n-1} p(\tau|\mathbf{x}, M_k) \int \int p(y|\theta_2, M_k) p(\theta_1, \theta_2|\tau, \mathbf{x}, M_k) d\theta_1 d\theta_2 \\ &= \sum_{\tau=1}^{n-1} p(\tau|\mathbf{x}, M_k) p(y|\tau, \mathbf{x}, M_k) \end{aligned} \quad (\text{B.7})$$

This distribution is called the posterior predictive distribution, posterior because it is conditional on the observed \mathbf{x} and predictive because it is a prediction for an observable.

Note that in some cases the integrals in (B.6) and (B.7) can be evaluated analytically. More often, these integrals are intractable and thus must be computed by numerical methods. In this paper, we favor an approach based on Gibbs sampling which is straightforward to implement, and easily accessible to the average statistical practitioner without numerical analysis skills (see Section 4).

B.3.2 Comparison of change-point models

Bayesian model selection consists of evaluating posterior probabilities for each model being true among a discrete set denoted \mathcal{M} . Let us assume that the random variables $\mathbf{X} = (X_1, X_2, \dots, X_n)$ have arisen from one of $q+1$ possible models in $\mathcal{M} = \{M_0, M_1, \dots, M_k, \dots, M_q\}$, where, as in Section 2, M_0 denotes the model which assumes that no change has occurred, while $M_1, \dots, M_k, \dots, M_q$ are change-point models. First, the Bayesian model selection proceeds by selecting prior probabilities $p(M_k)$ of each model being true. Often, equal prior probabilities are used, i.e. $p(M_k) = 1/(q+1)$. In the context of change-point analyses for water resources management, this choice based on a full symmetry of all situations can be criticized. In fact, the no change hypothesis will often represent the status quo situation which many decision makers favor. Moreover, the remaining hypotheses will generally induce changes in policies that could be drastic, and that may involve a lot of money. Therefore, if no prior information is available about the existence of a change in the series of interest, the decision maker will be tempted to affect at least half of the weight to the no change hypothesis. Here, a more sensible choice would give equal weight to the « change » and « no change » alternatives:

$$p(M_0) = 1/2 \quad \text{and} \quad p(M_1) = \dots = p(M_q) = 1/(2q) \quad (\text{B.8})$$

After the prior probabilities have been chosen, the analysis then proceeds by computing the posterior probabilities of each model being true. Using Bayes theorem, we have

$$p(M_k|\mathbf{x}) = \frac{p(\mathbf{x}|M_k) p(M_k)}{\sum_{j=0}^q p(\mathbf{x}|M_j) p(M_j)} \quad (\text{B.9})$$

where $p(\mathbf{x}|M_k)$ is the prior predictive density of \mathbf{x} under model M_k given in (B.6).

Let us consider only the problem involving the verification of the existence of a change and, if any, the identification of its type. This problem is essentially that of choosing a model in \mathcal{M} , without any subsequent decision, when the « state of the world » of interest is defined to be the « true » model. To proceed we first assume that the range of models under consideration includes the « true » belief model. This is the so-called \mathcal{M} -closed perspective described in Bernardo and Smith (1994). This assumption may be difficult to accept in a literal sense. However, for a change-point analysis, this hypothesis might be considered reasonable. In fact, often in hydrology, a given model, for instance model M_0 , has been extensively adopted and found to be a successful predictive device in a wide spectrum of applications. Now, suppose a hydrologist suspects that a certain type of change, for example M_1 , may have occurred, and in this new context wants to incorporate uncertainty about M_0 . Provided the hydrologist is comfortable with assigning prior weights to the alternative formulations, the \mathcal{M} -closed perspective can be exploited since M_0 is generally viewed as a proxy to the true model. If the natural zero-one utility function is used in the \mathcal{M} -closed setting, maximizing expected utility implies that the optimal model choice is the one with largest posterior probability (B.9) to be true. Therefore, these posterior probabilities can be used in a formal way to verify if a change in a hydrometeorological series is a plausible hypothesis, and if necessary what type of shift may have occurred.

For the particular situation where only two models, M_j and M_k , are to be compared, the above discussion suggests that one should use the posterior odds ratio

$$\frac{p(M_j|\mathbf{x})}{p(M_k|\mathbf{x})} = \frac{p(\mathbf{x}|M_j)p(M_j)}{p(\mathbf{x}|M_k)p(M_k)} \quad (\text{B.10})$$

Expression (B.10) reveals the key role of the prior predictive density in providing the way in which the data update the relative prior beliefs into the relative posterior belief about the true model. This leads naturally to the Bayes factor B_{jk} , defined as the ratio of posterior to prior odds on M_j against M_k ,

$$B_{jk} = \frac{p(M_j|\mathbf{x})}{p(M_k|\mathbf{x})} \bigg/ \frac{p(M_j)}{p(M_k)} = \frac{p(\mathbf{x}|M_j)}{p(\mathbf{x}|M_k)}, \quad (\text{B.11})$$

which can be interpreted as a measure of whether the data \mathbf{x} have increased or decreased the odds on M_j relative to M_k . The Bayes factor given in (B.11) is similar enough to the likelihood ratio statistic, but the parameters are eliminated by integration rather than by maximization. Naturally, $B_{jk} > 1$ signifies that M_j is now more relatively plausible than M_k in light of \mathbf{x} ; $B_{jk} < 1$ signifies that the relative plausibility of M_j has decreased. However, in many problems, such criteria for interpreting the Bayes factors can be very loose. Based on that of Jeffreys (1961), Raftery (1996) suggested more conservative guidelines for interpreting the Bayes factor, which are on the same scale as the familiar deviance and likelihood ratio test statistics. The guidelines are reproduced in Table 1.

Table 1. Raftery's scale for interpreting the Bayes factor

$2 \ln B_{jk}$	B_{jk}	Evidence for M_j
0 – 2	1 – 3	Not worth more than a bare mention
2 – 6	3 – 20	Positive
6 – 10	20 – 150	Strong
> 10	> 150	Very strong

Many questions can be answered using Bayes factors (or equivalently the posterior probability of each model) along with Table 1. For instance, if one is interested in how much the data support a given model M_k compared to its competitor alternatives in \mathcal{M} , the appropriate Bayes factor, denoted $B_{k(-k)}$, is given by

$$B_{k(-k)} = \left(\frac{p(M_k | \mathbf{x})}{1 - p(M_k | \mathbf{x})} \right) / \left(\frac{p(M_k)}{1 - p(M_k)} \right) = \left[\sum_{j=0(\neq k)}^q B_{jk} \left(\frac{p(M_j)}{1 - p(M_k)} \right) \right]^{-1} \quad (\text{B.12})$$

An important particular case is the overall assessment of change versus no change. Since $B_{jk} = B_{kj}^{-1}$, we note that testing the existence of a change in a hydrometeorological series can be based on

$$B_{c0} = B_{(-0)0} = \left(\frac{1 - p(M_0 | \mathbf{x})}{p(M_0 | \mathbf{x})} \right) / \left(\frac{1 - p(M_0)}{p(M_0)} \right) = \sum_{j=1}^q B_{j0} \left(\frac{p(M_j)}{1 - p(M_0)} \right) \quad (\text{B.13})$$

The right-hand side of (B.13) has the form of a weighted average of individual Bayes factors for specific configurations of change against no change. If, as above, equal prior probabilities are specified for the q change-point models, and equal prior weights are assigned to the « change » and « no change » alternatives, the Bayes factor B_{c0} for change against no change is just an average of the Bayes factors B_{j0} , taken over all considered change-points model. Finally, if we assume that a change has occurred, identification of its type may be carried out by evaluating the Bayes factor (B.12) for each alternative, excluding the no change model M_0 . According to the scheme in Table 1, its type is identified after examining if one of these Bayes factor provides enough evidence in favor of a given change-point model compared to its competitors.

It is important to notice that the use of Bayes factors, and therefore of (B.9), needs proper informative prior distributions for all parameters to be specified. When prior distributions are improper, the prior predictive distributions $p(\mathbf{x} | M_j)$ and $p(\mathbf{x} | M_k)$ are not defined. However, this rather embarrassing problem can be overcome by using default Bayes factors, such as the « Intrinsic Bayes factor » (Berger and Pericchi, 1996) and the « Fractional Bayes factor » (O'Hagan, 1995). In this paper, we do not address model selection with noninformative priors since we assume that some prior hydrologic expertise is available, however vague it may be.

B.3.3 Accounting for model uncertainty when forecasting

Suppose we are interested in a decision problem which involves not only model choice but also a subsequent action, such as the prediction of a future value y . Following the discussion in the previous section, if one wishes to consider a single model, the one with largest posterior probability to be true, denoted by M_k^* , will be selected for prediction. If we consider a quadratic utility function, maximizing expected utility leads to the optimal prediction \tilde{y}_k^* under M_k^* given by

$$\tilde{y}_k^* = \int y p(y | \mathbf{x}, M_k^*) dy = E \{y | \mathbf{x}, M_k^*\} \quad (\text{B.14})$$

where $p(y | \mathbf{x}, M_k^*)$ is the posterior predictive density (B.7). This prediction depends upon a given model M_k^* . Using it forces the hydrologist to select a single model and to ignore model uncertainty, which can be a major part of overall uncertainty about future values. The more general Bayesian model selection framework, presented in the previous section, allows for the evaluation of an « overall » posterior predictive distribution, which takes into account all models in \mathcal{M} . In fact, since under the \mathcal{M} -closed perspective each model has a posterior probability $p(M_k | \mathbf{x})$, one can maintain consideration of several models by weighting each conditional posterior predictive density $p(y | \mathbf{x}, M_k)$ by $p(M_k | \mathbf{x})$:

$$p(y|\mathbf{x}) = \sum_{k=0}^q p(y|\mathbf{x}, M_k) p(M_k|\mathbf{x}) \quad (\text{B.15})$$

Assuming again a quadratic utility function, it is seen that the optimal overall prediction \tilde{y} with respect to expected utility is

$$\begin{aligned} \tilde{y} &= \int y p(y|\mathbf{x}) dy = \sum_{k=0}^q E\{y|\mathbf{x}, M_k\} p(M_k|\mathbf{x}) \\ &= \sum_{k=0}^q \tilde{y}_k p(M_k|\mathbf{x}) \end{aligned} \quad (\text{B.16})$$

Equation (B.15) shows, in particular, that selecting a single model and proceeding conditionally on it may be reasonable if one of the posterior probabilities $p(M_k|\mathbf{x})$ is close to unity. If not, then prediction conditional on a single model fails to take into account all of the uncertainty about the configuration of the change, and the precision of forecasts may be overestimated.

B.4 Evaluations of predictive densities

As mentioned in Section 3.1, the prior and posterior predictive distributions, $p(\mathbf{x}|M_k)$ and $p(y|\mathbf{x}, M_k)$, may not be expressible in closed form, and recourse to numerical approximations is needed. Since estimation of the parameters for model (B.1) is done via Gibbs sampling (Section 2), it would be natural to use the simulated outputs to estimate the predictive densities. The problem is easily solved by Monte Carlo generation for posterior predictive distribution. On the other hand, the evaluation of the prior predictive distribution from Gibbs outputs is more challenging. In the following, we present how the Gibbs algorithm can be used to evaluate such integrals. The parameters of model (B.1) are denoted by $\theta = (\theta_1, \theta_2, \tau)$, and we suppose $\{\theta^{(j)}; j = 1, \dots, m\}$ are the m draws from the joint posterior density $p(\theta|\mathbf{x}, M_k)$ obtained using the Gibbs sampler, after the first t values of the chain have been discarded.

B.4.1 Posterior predictive density

The posterior predictive density can be readily approximated via the Gibbs sampling process used to estimate the posterior distributions. For each draw of the parameters $\{\theta^{(j)}; j = 1, \dots, m\}$ from the posterior distribution $p(\theta|\mathbf{x}, M_k)$, one simply has to sample one value of y from the likelihood function $p(y|\theta^{(j)}, M_k)$, which for model (B.1) is $p(y|\theta_2^{(j)}, M_k)$. The set of simulated y 's from all the $\theta^{(j)}$'s characterizes the posterior predictive density, $p(y|\mathbf{x}, M_k)$, and can be used to estimate it. For instance, as a rough estimate for $p(y|\mathbf{x}, M_k)$, one can use a histogram. A better estimate is given by

$$\hat{p}(y|\mathbf{x}, M_k) = \frac{1}{m} \sum_{j=1}^m p(y|\theta_2^{(j)}, M_k) \quad (\text{B.17})$$

Gelfand and Smith (1990) referred to this averaging technique as Rao-Blackwellization and argued that it improves on the usual histogram estimate. This approach was adopted in Perreault et al. (2000) for inference about the parameters of model M_2 .

If one is concerned with model uncertainty, an overall posterior predictive distribution can be estimated using the posterior probabilities associated with each model. If, for all models, prior and posterior predictive distributions are intractable, we have

$$\widehat{p}(y|\mathbf{x}) = \sum_{k=0}^q \widehat{p}(M_k|\mathbf{x}) \widehat{p}(y|\mathbf{x}, M_k), \quad (\text{B.18})$$

where $\widehat{p}(y|\mathbf{x}, M_k)$ is given by (B.17) and $\widehat{p}(M_k|\mathbf{x})$ is the estimated posterior probability of M_k being true obtained by substituting in (B.9) the prior predictive by an estimated value (see Section 4.2). Finally, conditional and overall point predictions for a future value, assuming a quadratic utility function, can be evaluated in a way similar to that in Section 3.3.

B.4.2 Prior predictive density

As it was mentioned, the evaluation of the prior predictive distribution $p(\mathbf{x}|M_k)$ from Gibbs outputs is a more complicated task. In fact, integration is done with respect to the prior density, whereas the Gibbs sampler produces the draws from the posterior. Therefore, the simulated outputs from the Gibbs method, or other Markovian iterative schemes, can not be used directly to evaluate prior predictive distributions as is the case for posterior predictive densities.

Several approaches, based on such posterior simulation, have been proposed to estimate prior predictive densities like $p(\mathbf{x}|M_k)$. However, many of these estimates are known to be unstable, or need large samples (see Kass and Raftery, 1995). Other more accurate estimates requires a tuning function, which can be quite difficult to determine (see Gelfand and Dey, 1994). On the other hand, Chib (1995) developed a very simple approach which is free of the problems just mentioned. This method, which is adopted here, was developed in the setting where the Gibbs sampling algorithm has been used to provide sample draws from the posterior distribution. To use it, it is necessary that all normalizing constants of the full conditional densities involved in the Gibbs sampling be known. This is usually the case when the Gibbs algorithm is implemented.

Chib (1995) exploits the fact that the prior predictive density is the normalizing constant in the Bayes theorem, and therefore can be expressed as the prior distribution times the likelihood function over the posterior density, i.e.

$$p(\mathbf{x}|M_k) = \frac{p(\mathbf{x}|\boldsymbol{\theta}, M_k) p(\boldsymbol{\theta}|M_k)}{p(\boldsymbol{\theta}|\mathbf{x}, M_k)} \quad (\text{B.19})$$

This simple identity holds for any parameter value. The numerator in the last expression can be directly evaluated at a given point, say $\boldsymbol{\theta}^*$. However, no expression in closed form is available for the denominator $p(\boldsymbol{\theta}|\mathbf{x}, M_k)$. Therefore, if a density estimate for $p(\boldsymbol{\theta}|\mathbf{x}, M_k)$, denoted $\widehat{p}(\boldsymbol{\theta}|\mathbf{x}, M_k)$, is available, an estimate for the prior predictive distribution using expression (B.19) can be derived as

$$\widehat{p}_{CH}(\mathbf{x}|M_k) = \frac{p(\mathbf{x}|\boldsymbol{\theta}^*, M_k) p(\boldsymbol{\theta}^*|M_k)}{\widehat{p}(\boldsymbol{\theta}|\mathbf{x}, M_k)}, \quad (\text{B.20})$$

where $\boldsymbol{\theta}^*$ is any parameter value. The simple solution proposed by Chib (1995) to evaluate $\widehat{p}(\boldsymbol{\theta}|\mathbf{x}, M_k)$ is to perform additional Gibbs cycles. A brief description of the implementation for our particular problem is offered in Appendix B. A scheme summarizing the approach is presented in Figure 1. See Chib (1995) for more general discussions and properties.

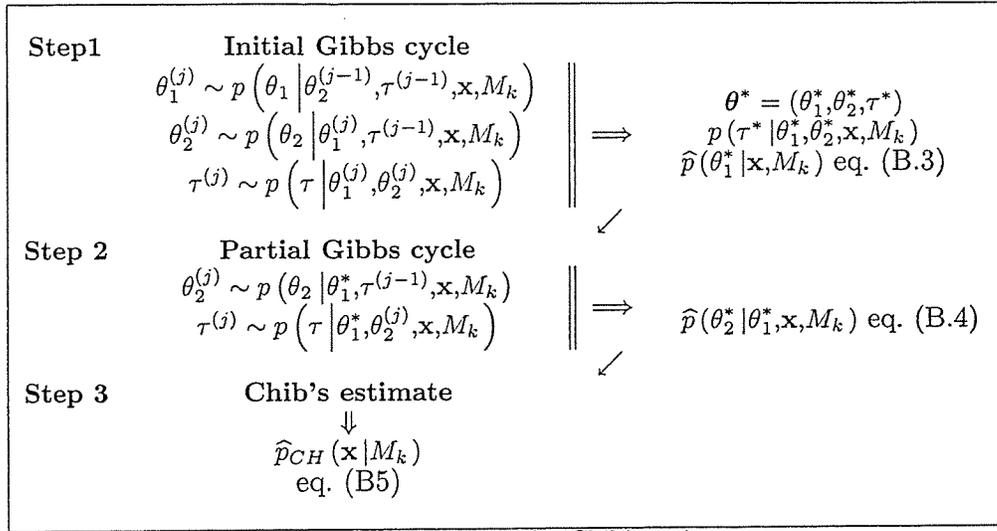


Figure 1. Basic steps for Chib's estimate.

Although Chib's procedure leads to an increase in the number of iterations, it does not require new programming and thus is straightforward to implement.

B.5 Applications assuming normal change-point models

This new perspective for change-point analysis is illustrated by applying the methods presented in the previous sections to eight series of annual energy inflows. The four normal models discussed in Section 2 are considered: the no change hypothesis, a single change in the mean level only, a single change in the variance only, and a simultaneous change in both the mean and the variance. Perreault et al. (2000) have already applied models M_1 and M_2 to annual energy inflows for two sites. However, their analysis was restricted to the estimation of the change-point and its intensity, and did not account for model uncertainty.

All calculations which require Gibbs sampling (evaluation of prior and posterior predictive distributions for model M_2) were based on $t = 1000$ iterations and $m = 1000$ replicates. To ensure maximum accuracy of Chib's approach, the prior predictive estimate was evaluated at a high density point θ^* , namely the posterior mode.

B.5.1 The data

The series of particular importance for energy planning in Québec are the annual energy inflows for the 8 major hydropower systems managed by Hydro-Québec: Saint-Laurent, Outaouais, La Grande, Saint-Maurice, Bersimis, Manicouagan, Outardes and Churchill Falls. For each of these systems, the energy inflows are evaluated by multiplying the natural inflow of each reservoir in the system by a factor based on the production capacity of the corresponding power plant. The historical annual energy inflows expressed in terawatt-hour (TWh) for all hydropower systems (1943-1996) are shown in Figure 2.

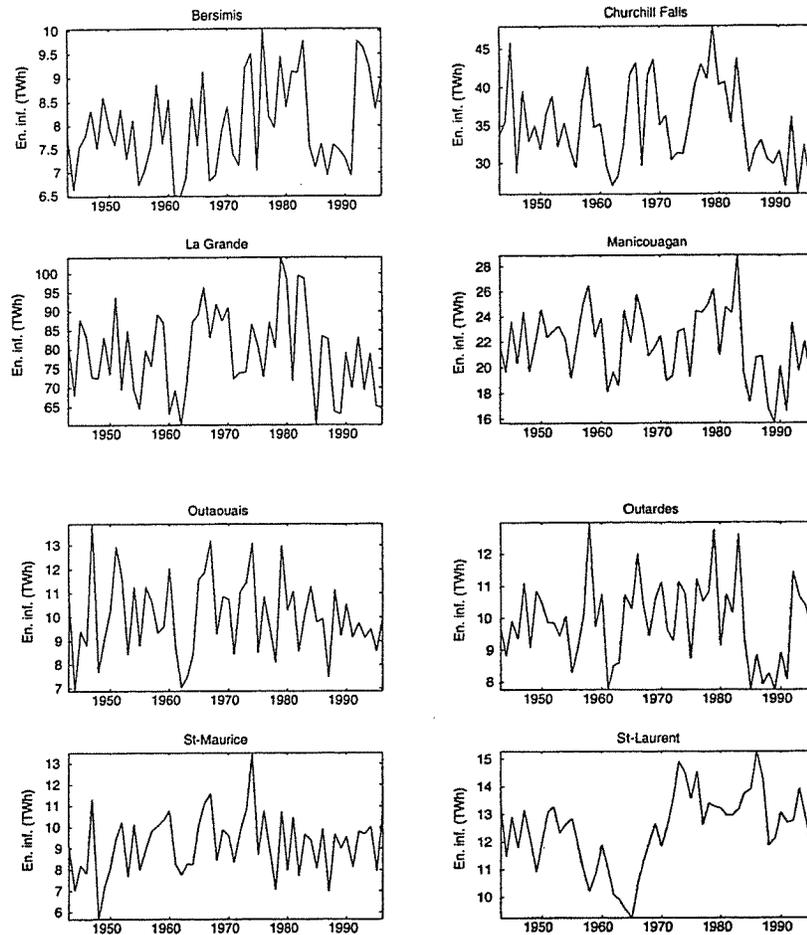


Figure 2. Annual energy inflows for 8 hydropower systems in TWh (1943-1996).

Examining these time series, an abrupt change in the mean level and/or the variance for some hydropower systems may be suspected. However, statistical characteristics for these series, such as the mean and the variance, are calculated over the entire period of observation, and then used as inputs to design scenarios or for forecasting future energy availability. Stationarity is therefore assumed, even when in fact this hypothesis could be violated. This might lead to wrong decisions regarding policies about future energy related development for the company. Therefore, in order to develop management rules for energy planning, procedures for analyzing changes in energy inflow series are needed.

B.5.2 Specifying prior distributions

Bayesian model selection using standard Bayes factors cannot be made using non-informative prior distributions. Prior knowledge must thus be elicited from expert opinion and/or from external information about the studied phenomena in order to obtain proper prior distributions. In what follows, we briefly describe how proper prior distributions were specified for our particular problem.

As mentioned in the previous section, the annual energy inflows are generally used by Hydro-Québec for energy planning under the assumption that all historical observations were obtained under the same conditions. Model M_0 , which represents the no change hypothesis, can then be considered as a proxy to the true model in a prior state of belief. Therefore, following the discussion in Section 3.2, prior probabilities for each model being

true were chosen such that « change » and « no change » alternatives have equal weight, i.e.

$$p(M_0) = 1/2 \quad \text{and} \quad p(M_1) = p(M_2) = p(M_3) = 1/6 .$$

Faced with such vague knowledge and lack of conviction about the existence of a change, one cannot reasonably favor any epoch of change for models M_1 , M_2 and M_3 . The year τ at which a change may have occurred was therefore assumed to follow a discrete uniform distribution for each of these models. Note that with this density, the prior expected change-point is 1970, i.e. the mean of a discrete uniform probability distribution on the interval [1943,1995].

Finally, for the other parameters, the prior degrees of belief were assumed to be represented by normal-inverted gamma type of distributions (see Appendix A). The complete specification of prior knowledge about these quantities required the choice of hyperparameters in these expressions. This was done by eliciting regional information, as described in Perreault et al. (2000). For each site, a regression model was used to predict the average and the variance of the energy inflows, before and after 1970. These predictions were then used, along with their standard error, to estimate the mean and the variance of the normal and inverted gamma distributions. Finally, simple systems of equations for the first two moments were solved, yielding estimated hyperparameter values for all sites. A simple regression considering only the generating capacity appeared to be the best regional model for predicting the average and the variance of the energy inflows for each hydropower system. For details, see Perreault et al. (2000).

B.5.3 Inference about the existence and the configuration of a change

We are now concerned with the problem of verifying the existence of a change in the annual energy inflow and, if any, the identification of its type among the configurations assumed by M_1 , M_2 and M_3 . As explained in Section 3.2, this is essentially the problem of choosing a model in $\mathcal{M} = \{M_0, M_1, M_2, M_3\}$, without any subsequent decision. To perform a Bayesian model selection in \mathcal{M} , each of the four prior predictive densities, $p(\mathbf{x}|M_k)$ $i = 0, \dots, 3$, has to be evaluated. This is easily done for models M_0 , M_1 and M_3 . The density $p(\mathbf{x}|M_0)$, assuming conjugate prior distributions, is a well-known result and is expressible in closed form (Berger, 1985). The prior predictive distributions $p(\mathbf{x}|M_1)$ and $p(\mathbf{x}|M_3)$ can be calculated using (B.6). Since $p(\mathbf{x}|\tau, M_1)$ and $p(\mathbf{x}|\tau, M_3)$ appearing in (B.6) can be expressed in closed form (see Appendix A), the evaluation of the prior predictive densities for models M_1 and M_3 is straightforward. More precisely, we have

$$p(\mathbf{x}|M_0) = \left(\frac{1}{2\pi}\right)^{n/2} \frac{1}{\sqrt{1+n\lambda}} \frac{\beta^\alpha}{(\beta')^{\alpha'}} \frac{\Gamma(\alpha')}{\Gamma(\alpha)} \quad (\text{B.21})$$

$$p(\mathbf{x}|M_1) = \left(\frac{1}{2\pi}\right)^{n/2} \frac{\beta^\alpha}{\sqrt{\lambda_1 \lambda_2} \Gamma(\alpha)} \sum_{\tau=1}^{n-1} p(\tau|M_1) (\lambda'_1 \lambda'_2)^{1/2} \frac{\Gamma(\alpha')}{(\beta')^{\alpha'}} \quad (\text{B.22})$$

$$p(\mathbf{x}|M_3) = \left(\frac{1}{2\pi}\right)^{n/2} \frac{\beta_1^{\alpha_1} \beta_2^{\alpha_2}}{\sqrt{\lambda_1 \lambda_2} \Gamma(\alpha_1) \Gamma(\alpha_2)} \sum_{\tau=1}^{n-1} p(\tau|M_3) (\lambda'_1 \lambda'_2)^{1/2} \frac{\Gamma(\alpha'_1) \Gamma(\alpha'_2)}{(\beta'_1)^{\alpha'_1} (\beta'_2)^{\alpha'_2}} \quad (\text{B.23})$$

where the updated hyperparameters for these models are given in Appendix A. To simplify, the same notation for the hyperparameters of each model was used. But, to stay rigorous, different notations should have been used to express the fact that hyperparameter values depend upon their specific model.

Now, since informative prior distributions were specified for each parameter, the Bayes factors B_{10} , B_{30} and B_{31} and the model posterior probabilities $p(M_k | \mathbf{x})$, $k = 0,1,3$, can be evaluated analytically using (B.11) and (B.9) with the above expressions. However, the integral needed to evaluate $p(\mathbf{x} | M_2)$ is intractable (Perreault et al., 2000). Therefore, the Bayes factors B_{20} , B_{21} and B_{32} cannot be evaluated analytically. This problem can be solved via Gibbs sampling by using Chib's method as presented in Section 4.2. The full conditional distributions needed for Chib's approach were derived in Perreault et al. (2000), and are recalled in Appendix A.

All Bayes factors for each hydropower system are presented in Table 3. These can be used to compare, two by two, all models. The posterior probabilities (B.9) of each model being true have also been evaluated, and are given in the last column of Table 2. They can be compared to the prior probabilities specified in Section 5.2 in order to see how the prior state of belief was updated by the data.

Table 2. Bayes factors and model posterior probabilities for all series

	B_{jk}	M_0	M_1	M_2	M_3	$p(M_k \mathbf{x})$
Bersimis	M_0	1	0.3865	0.2642	0.2364	0.2205
	M_1	2.5870	1	0.6835	0.6115	0.1902
	M_2	3.7850	1.4631	1	0.8947	0.2783
	M_3	4.2307	1.6354	1.1177	1	0.3110
	B_{jk}	M_0	M_1	M_2	M_3	$p(M_k \mathbf{x})$
Churchill Falls	M_0	1	0.1659	0.7848	0.0929	0.1424
	M_1	6.0284	1	4.7313	0.5603	0.2862
	M_2	1.2742	0.2114	1	0.1184	0.0605
	M_3	10.7591	1.7847	8.4441	1	0.5108
	B_{jk}	M_0	M_1	M_2	M_3	$p(M_k \mathbf{x})$
La Grande	M_0	1	0.7049	0.6445	0.9349	0.4261
	M_1	1.4187	1	0.9143	1.3264	0.2015
	M_2	1.5517	1.0937	1	1.4507	0.2204
	M_3	1.0696	0.7539	0.6893	1	0.1519
	B_{jk}	M_0	M_1	M_2	M_3	$p(M_k \mathbf{x})$
Manicouagan	M_0	1	3.3276	0.4086	1.5035	0.4678
	M_1	0.3005	1	0.1228	0.4518	0.0469
	M_2	2.4475	8.1445	1	3.6800	0.3816
	M_3	0.6651	2.2132	0.2717	1	0.1037
	B_{jk}	M_0	M_1	M_2	M_3	$p(M_k \mathbf{x})$

Table 2. (Continued)

	B_{jk}	M_0	M_1	M_2	M_3	$p(M_k \mathbf{x})$
Outaouais	M_0	1	3.0306	0.2644	0.7530	0.3554
	M_1	0.3300	1	0.0872	0.2485	0.0391
	M_2	3.7827	11.4638	1	2.8482	0.4481
	M_3	1.3281	4.0249	0.3511	1	0.1573
	B_{jk}	M_0	M_1	M_2	M_3	$p(M_k \mathbf{x})$
Outardes	M_0	1	1.7533	0.1761	1.1560	0.2966
	M_1	0.5703	1	0.1004	0.6593	0.0564
	M_2	5.6798	9.9585	1	6.5660	0.5615
	M_3	0.8650	1.5167	0.1523	1	0.0855
	B_{jk}	M_0	M_1	M_2	M_3	$p(M_k \mathbf{x})$
Saint-Maurice	M_0	1	1.1804	0.7005	1.7396	0.5129
	M_1	0.8472	1	0.5934	1.4737	0.1448
	M_2	1.4276	1.6852	1	2.4835	0.2441
	M_3	0.5748	0.6785	0.4027	1	0.0983
	B_{jk}	M_0	M_1	M_2	M_3	$p(M_k \mathbf{x})$
Saint-Laurent	M_0	1	< 0.0001	0.0727	0.0001	0.0001
	M_1	> 150	1	> 150	1.6854	0.6274
	M_2	13.7496	0.0005	1	0.0009	0.0003
	M_3	> 150	0.5933	> 150	1	0.3722
	B_{jk}	M_0	M_1	M_2	M_3	$p(M_k \mathbf{x})$

It can be seen in particular that the odds for model M_0 have decreased for every series except St-Maurice. Therefore, the data tend to contradict Hydro-Québec's prior belief that the annual energy inflows were generated from the same process. We can be particularly confident about the existence of a single change for the Saint-Laurent series, since the posterior probability for M_0 is very small (0.0001). However, for the other systems, it may be appropriate to use the Bayes factor B_{c0} , given by (B.13), along with a decision rule such as Raftery's scheme (Table 1). These Bayes factors are reported in Table 3 for each series.

Table 3. Bayes factors for change against no change and for identifying the type of the change, if any

System	B_{c0}	$B_{1(2,3)}$	$B_{2(1,3)}$	$B_{3(1,2)}$
Bersimis	3.53	0.65	1.11	1.33
Churchill Falls	6.02	1.00	0.16	2.95
La Grande	1.35	-	-	-
Manicouagan	1.14	-	-	-
Outaouais	1.81	-	-	-
Outardes	2.37	0.10	14.25	0.16
Saint-Maurice	0.95	-	-	-
Saint-Laurent	>150	3.37	0.00	1.19

According to the scheme of Table 1, the Bayes factors B_{c0} for Bersimis and Churchill Falls provide positive evidence in favor of a single change, while for Saint-Laurent, as it was anticipated, very strong evidence against no change is provided. Note that the annual energy inflows of Outardes almost exhibit a change ($B_{c0} = 2.37$). For the other systems, it is not possible to reject the no change hypothesis. Since for Bersimis, Churchill Falls, Outardes and Saint-Laurent the change hypothesis is plausible, Bayes factors to identify its type were calculated as described at the end of Section 3.2. Discarding model M_0 as a credible hypothesis, equation (B.12) was used to evaluate $B_{1(2,3)}$, $B_{2(1,3)}$ and $B_{3(1,2)}$. The Bayes factor $B_{1(2,3)}$ corresponds to the ratio of posterior to prior odds on M_1 against M_2 and M_3 , and so on. These pairwise comparisons of plausibility measures may be used to discriminate between the considered change-point models, under the hypothesis that one of these configurations has occurred. It seems that models M_2 and M_3 are equally credible

for Bersimis, while for Outardes M_2 is clearly the most plausible configuration of change, if any. The annual energy inflows observed for Churchill Falls are more likely to exhibit a simultaneous change in both the mean and the variance (M_3). Finally, it seems that the most plausible type of change for annual energy inflows for Saint-Laurent is a single change in the mean only (M_1).

Inference about the parameters of interest, assuming configuration M_1 or M_2 , can be made by using the approaches described in Perreault et al. (2000). In particular, assuming a single change in the variance only, 1960 and 1954 are seen to be the most probable years of change (mode of $p(\tau | \mathbf{x}, M_2)$) for Bersimis and Outardes respectively. In Perreault et al. (2000) a detailed analysis was performed to characterize a change in Churchill Falls annual energy inflows assuming model M_1 . This inference may be revised in light of the above results by using the posterior distributions for M_3 derived in Appendix A. It is important to note that if Raftery's scheme have to be strictly used, discrimination between the three change-point models is not possible for Bersimis and Churchill Falls. In fact, none of the three Bayes factors indicate positive evidence ($B_{i(jk)} > 3$).

B.5.4 Predictive analysis

The detection of a change and the identification of its type are important for Hydro-Québec, but prediction is often the real ultimate goal when analyzing annual energy inflows. In what follows, we derive the posterior predictive density for a future value under each of the models in the set $\mathcal{M} = \{M_0, M_1, M_2, M_3\}$ and use them to evaluate an « overall » posterior predictive distribution, as developed in Section 3.3. The forecasting of a future observation is then performed for all Hydro-Québec's series of annual energy inflows.

For models M_1 and M_3 , integration in (B.7) to evaluate $p(y | \tau, \mathbf{x}, M_k)$ is straightforward and leads to an expression in closed form. The resulting density functions $p(y | \tau, \mathbf{x}, M_1)$ and $p(y | \tau, \mathbf{x}, M_3)$ are seen to be Student t -distributions. For model M_0 , the predictive density $p(y | \mathbf{x}, M_0)$, which of course does not involve a change-point parameter τ , is also a Student distribution (Berger, 1985). The parameters of the corresponding t -distributions are presented in Table 4.

Table 4. Posterior predictive densities conditional on τ .

Density	Parameters of the Student t -distributions		
	Location a	Scale b	Degrees of freedom c
$p(y \mathbf{x}, M_0)$	ϕ'	$\alpha' [(1 + \lambda') \beta']^{-1}$	$2\alpha'$
$p(y \tau, \mathbf{x}, M_1)$	ϕ'_2	$\alpha' [(1 + \lambda'_2) \beta']^{-1}$	$2\alpha'$
$p(y \tau, \mathbf{x}, M_3)$	ϕ'_2	$\alpha'_2 [(1 + \lambda'_2) \beta'_2]^{-1}$	$2\alpha'_2$

Note: $ST(x | a, b, c) \propto [1 + c^{-1}b(x - a)^2]^{-(c+1)/2}$

Hence, for a given change-point model, say M_3 , the unconditional density $p(y | \mathbf{x}, M_3)$ is evaluated with (B.7), and appears as a finite mixture of the associated conditional distributions weighted by the $n - 1$ values of $p(\tau | \mathbf{x}, M_3)$ given by (A.5). As it was the case for the prior predictive density, the posterior predictive distribution $p(y | \mathbf{x}, M_2)$ cannot be calculated analytically. However, it can easily be approximated using the Gibbs sampling outputs (see Section 4.1).

Posterior predictive distributions for a 1-year future value, $p(y | \mathbf{x}, M_k)$, were computed for the eight series of annual energy inflows assuming the four models. The overall posterior predictive density was also evaluated using (B.18) and the posterior probabilities given in Table 3. These are shown in Figure 3 for Bersimis, Churchill Falls, Outaouais and Saint-Laurent. The dotted lines correspond to the « conditional » posterior predictive densities $p(y | \mathbf{x}, M_k)$, while the continuous line is the combined predictive mixture. This figure illustrates how the overall posterior predictive distribution behaves compared to the conditional ones. When uncertainty about the true model is quite low, i.e. one of the posterior

probabilities is near one (Saint-Laurent for instance), this distribution is very close to the predictive distributions obtained from the « optimal » model M_k^* . On the other hand, when uncertainty about the true model is considerable, i.e. all posterior probabilities are similar (Bersimis for instance), all conditional posterior predictive distributions contribute to the overall predictive distribution.

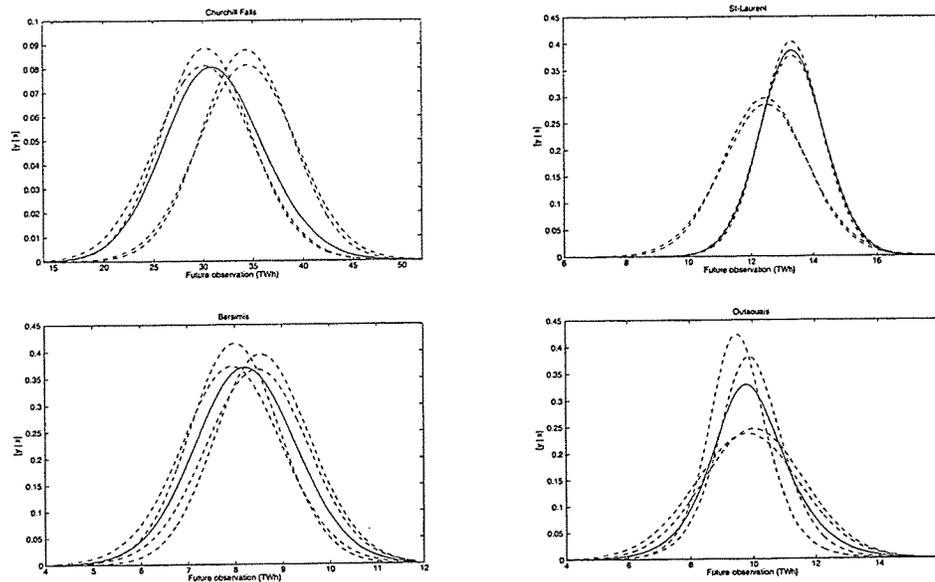


Figure 3. Conditional and overall 1-year predictive densities.

Various quantities of interest about future energy availability (quantiles, exceedance probabilities, etc.) can be derived from these densities. Table 5 reports a point prediction for each of the posterior predictive densities along with its standard deviation (in parentheses). A quadratic utility function was assumed, i.e. the 1-year forecast corresponds to the expected value (equations (B.14) and (B.16)). The 90% credible interval for the overall prediction is illustrated in Figure 4 for each hydropower system.

Table 5. Results of a 1-year prediction for all hydropower systems

System	Conditional prediction				Overall
	M_0	M_1	M_2	M_3	
Bersimis	8.00 (0.97)	8.60 (0.94)	7.93 (1.09)	8.45 (1.06)	8.23 (1.07)
Churchill Falls	34.43 (4.93)	30.16 (4.88)	34.34 (4.58)	30.44 (4.39)	31.16 (4.91)
La Grande	79.32 (11.55)	77.02 (11.64)	79.46 (13.02)	77.46 (13.37)	78.61 (12.24)
Manicouagan	21.91 (2.96)	20.88 (3.01)	22.07 (3.65)	20.91 (3.94)	21.82 (3.37)
Outaouais	10.03 (1.65)	9.82 (1.69)	9.94 (1.20)	9.60 (1.12)	9.91 (1.39)
Outardes	9.91 (1.27)	9.51 (1.26)	9.89 (1.36)	9.78 (1.35)	9.87 (1.33)
Saint-Maurice	9.19 (1.43)	9.39 (1.39)	9.20 (1.24)	9.36 (1.33)	9.24 (1.37)
Saint-Laurent	12.46 (1.36)	13.32 (1.07)	12.50 (1.40)	13.32 (1.01)	13.32 (1.05)

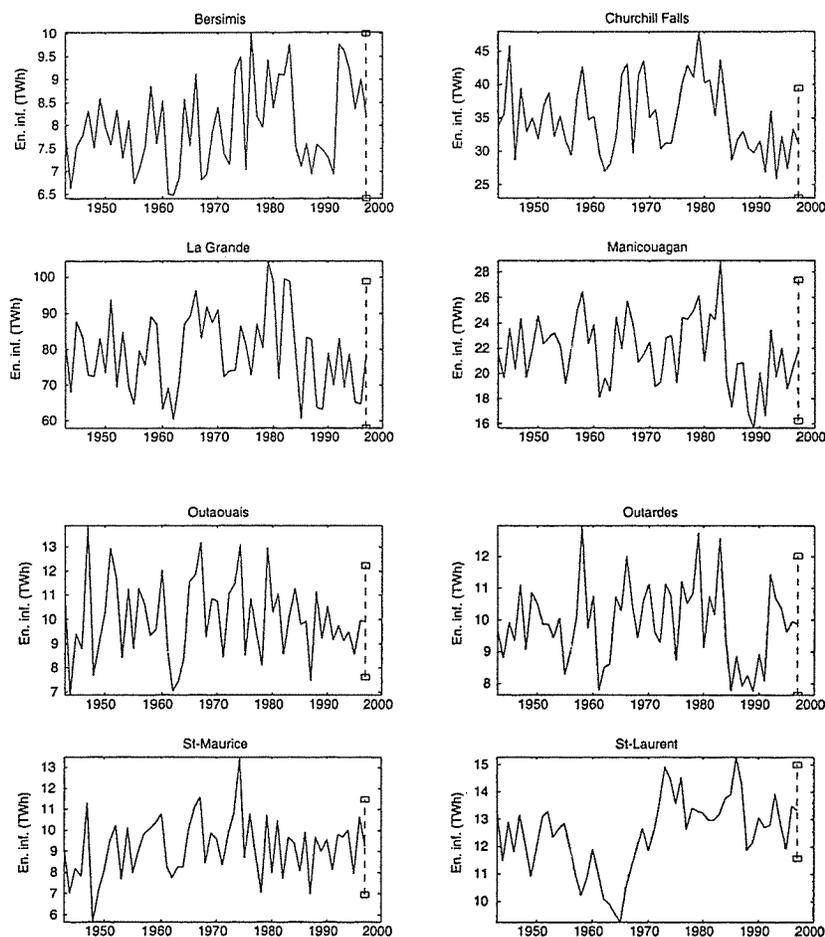


Figure 4. 90% credibility intervals for a 1-year forecast.

Among others, the main observations drawn from Table 5 and Figure 4 are as follows:

1) For 5 hydropower systems, assuming the usual stationarity hypothesis leads to 1-year higher forecasts of energy inflow than those obtained by maintaining all models in an overall point prediction. In particular, this is the case for Churchill Falls (+10%) and La Grande (+0.9%), the systems which contribute the most to the overall energy produced by Hydro-Québec. In this particular application, incorporating uncertainty about M_0 provided a more conservative prediction about future energy availability.

2) Considering all situations in an overall prediction leads to forecasts with larger standard deviations for 5 systems. This stems from the fact that, in addition to the sample uncertainty about the parameters, model uncertainty is taken into account.

3) Finally, 90% credible intervals can be very large, especially when uncertainty about the true model is important. This is the case for Bersimis, La Grande and Manicouagan. The credible interval for the overall prediction may even cover more than the range of the observed sample data (Bersimis).

It is important to mention that, in contrast with Bayesian point estimates such as the posterior mean, the Bayes factors, and therefore model selection, tend to be sensitive to the choices of priors on the model parameters. Thus, an important issue which was not discussed here is the sensitivity of the approach to prior distributions. Such analysis should be performed in practice. Some guidelines and references can be found in Kass and Raftery (1995).

B.6 Discussion and conclusions

The hydrologist is often dealing with many models that involve different assumptions about the studied phenomenon. This is the case when changes have to be analyzed in hydrometeorological series. Even if the hydrologist wishes to infer using a single model, there are necessary many choices to be made. For example, one may hesitate between a change in the mean level only and a simultaneous change in both the mean and the variance. Moreover, a hydrologist would generally not consider the no change hypothesis as an impossible situation. In estimating quantities of interest, such as future realizations, it is certainly desirable to provide an assessment of the uncertainty that accounts for the model-building process itself (Berger and Rios Insua, 1998; Kass and Raftery, 1995). As it was mentioned in Section 1, most of authors have centered their investigations on the estimation of the change-point and the intensity of the change for a given model. Few have addressed the problem of comparing various alternatives, including the no change situation. Another important issue which has been neglected in the literature is posterior predictive analysis for change-point problems. This paper has provided a new and convenient approach, based on Bayesian model selection, (1) for the overall assessment of change versus no change, and, if necessary, discrimination between different configurations of change, and (2) to maintain consideration of all models when forecasting a future observation. The approach was illustrated by its application on Hydro-Québec annual energy inflows assuming four univariate normal models (one model representing the no change hypothesis, and three change-point models).

The approaches presented herein can be used for other models such as gamma, binomial and Poisson distributions. More precisely, the use of probability distributions which belong to the exponential class of p.d.f.'s allows for exactly the same line of reasoning based on conjugacy for fixed change-point. Moreover, the independence hypothesis may be released, and change-point analysis can be performed for autoregressive models. Indeed, stationary autoregressive series with a strong coefficient of autocorrelation may exhibit the same type of behavior as series with changes in the mean level. With regards to the common use of AR models to represent hydrological persistency, it is therefore important to study which of two hypotheses (AR or M_1) is more credible. The methodology presented in this paper can answer this question by including in the set of models \mathcal{M} an AR model. It is important to note that the models considered are not appropriate to forecast in the future if another change in climate state takes place. These models assume that we will stay indefinitely in the second state. If we think that in the future the state may change again, we should generalize the approach. However, the series of interest in this paper do not indicate any return to the original state since only one change seems to have occurred. Anyhow, if one is interested to generalize the approach, useful models for forecasting if another change in climate state may take place would be the so-called hidden Markov chain models (see Robert et al. (2000) and references therein).

Finally, a complete Bayesian analysis can be achieved to cast the statistical inference of hydrometeorological changes into an operational decision-making framework with terminal actions such as planning new dams. Such investigations could develop Krzysztofowicz's results (Krzysztofowicz, 1994) who embraces this decisional aspects for nonstationary stopping-control processes, or more recent work by Hobbs et al. (1997) and Venketesh and Hobbs (1999) who studied the problem of investments for managing water levels under climate change uncertainty.

Acknowledgments. We wish to express our deep gratitude to Professor Lucien Duckstein for his critical review of this paper. We are also grateful to René Roy and Raymond Gauthier of Hydro-Québec for providing data. Finally, we gratefully acknowledge the constructive comments of two anonymous referees on an earlier version that considerably improved this paper. This research is supported by la Direction Général de l'Enseignement et de la Recherche de France (DGER), les fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec (FCAR), and l'Institut National de la Recherche Scientifique (INRS-Eau)..

References

- Berger, J.O., 1985. *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York.
- Berger, J.O., Pericchi, L.R., 1996. The intrinsic Bayes factor for model selection and prediction, *J. Amer. Statist. Assoc.* 91, 109-122.
- Berger, J.O., Rios Insua, D., 1998. Recent developments in Bayesian inference with applications in hydrology, in *Statistical and Bayesian Methods in Hydrological Sciences*, edited by E. Parent, P. Hubert, B. Bobée and J. Miquel, UNESCO Publishing, 43-62.
- Bernardo, J.M., Smith, A.F.M., 1994. *Bayesian theory*, Wiley, New York.
- Box, G.E.P, Tiao, G.C., 1973. *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading, Massachusetts.
- Bruneau, P., Rassam, J.-C., 1983. Application d'un modèle bayésien de détection de changements de moyennes dans une série, *Journal des Sciences Hydrologiques* 28, 341-354.
- Carlin, B.P., Gelfand, A.E., Smith, A.F.M., 1992. Hierarchical Bayesian analysis of changepoint problems, *Appl. Statist.* 41, 389-405.
- Chib, S., 1995. Marginal likelihood from the Gibbs output, *J. Am. Stat. Assoc.* 90, 1313-1321.
- Gelfand, A.E., Dey, D.K., 1994. Bayesian model choice: Asymptotics and exact calculations, *J. Roy. Statist. Soc. B* 56, 501-514.
- Gelfand, A.E., Smith, A.F.M., 1990. Sampling-based approaches to calculating marginal densities, *J. Am. Stat. Assoc.* 85, 398-409.
- Hobbs, B.F., Chao, P.T., Venkatesh, B.N., 1997. Using decision analysis to include climate change in water resources decision making, *Clim. Change* 37, 177-202.
- Jeffreys, H., 1961. *Theory of Probability*, Oxford, U.K., Oxford University Press.
- Kass, R.E., Raftery, A.E., 1995. Bayes factors, *J. Am. Stat. Assoc.* 90, 773-795.
- Kerr, R.A., 1992. Unmasking a shifty climate system, *Science* 255, 1508-1510.
- Kerr, R.A., 1999. A new force in high-latitude climate, *Science* 284, 241-242.
- Krzysztofowicz, R., 1994. Strategic decisions under nonstationary conditions: A stopping-control paradigm, in *Engineering Risk in Natural Resources Management*, Duckstein L. and E. Parent (eds.), NATO Advanced Studies Institute Series, Series E - Vol.275, Kluwer, 359-371.
- Lubès-Niel, H., Masson, J.M., Paturel, J.E., Servat, E., 1998. Variabilité climatique et statistiques. Etude par simulation de la puissance et de la robustesse de quelques tests utilisés pour vérifier l'homogénéité de chroniques, *Rev. Sci. Eau* 3, 383-408.
- Menzefricke, U., 1981. A Bayesian analysis of a change in the precision of a sequence of independent normal random variables at an unknown time point, *Appl. Statist.* 30, 141-146.
- O'Hagan, A., 1995. Fractional Bayes factors for model comparisons, *J. Roy. Statist. Soc. B* 57, 99-138.
- Perreault, L., Bernier, J., Bobée, B., Parent, E., 2000. Bayesian change-point analysis in hydrometeorological time series I. The normal model revisited, submitted to *J. of Hydrol.*
- Raftery, A.E., 1996. Approximate Bayes factors and accounting for model uncertainty in generalised linear models, *Biometrika* 83, 251-266.
- Rao, A.R., Tirtotjondro, W., 1996. Investigation of changes in characteristics of hydrological time series by Bayesian methods, *Stochastic Hydrology and Hydraulics* 10, 295-317.
- Robert, C.P., Rydén, T., Titterton, D.M., 2000. Bayesian inference in hidden Markov models through the reversible jump Markov chain Monte Carlo method, *J. Royal Statist. Society B* 62, 57-75.
- Rodriguez-Iturbe, I., Entekhabi, D., Bras, R.L., 1991. Nonlinear dynamics of soil moisture at climate scales I. Stochastic analysis, *Water Resour. Res.* 27, 1899-1906.
- Stephens, D.A., 1994. Bayesian retrospective multiple-changepoint identification, *Appl. Statist.* 43, 159-178.
- Venkatesh, B.N., Hobbs, B.F., 1999. Analysing investments for managing Lake Erie levels under climate change uncertainty, *Water Resour. Res.* 35, 1671-1683.

Appendix

A. Bayesian inference for normal change-point models Model M_3

Writing $\boldsymbol{\mu} = (\mu_1, \mu_2)$ and $\boldsymbol{\sigma} = (\sigma_1^2, \sigma_2^2)$, the likelihood function resulting from n observations $\mathbf{x} = (x_1, x_2, \dots, x_n)$ generated by model M_3 is given by

$$p(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\sigma}, \tau, M_3) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma_1^2}\right)^{\frac{\tau}{2}} \left(\frac{1}{\sigma_2^2}\right)^{\frac{n-\tau}{2}} \exp\left\{-\frac{\tau}{2\sigma_1^2} \left[s_\tau^2 + (\bar{x}_\tau - \mu_1)^2\right]\right\} \quad (\text{A2A.1})$$

$$\times \exp\left\{-\frac{n-\tau}{2\sigma_2^2} \left[s_{n-\tau}^2 + (\bar{x}_{n-\tau} - \mu_2)^2\right]\right\}$$

where

$$\bar{x}_\tau = \sum_{i=1}^{\tau} \frac{x_i}{\tau}, \quad \bar{x}_{n-\tau} = \sum_{i=\tau+1}^n \frac{x_i}{n-\tau}, \quad s_\tau^2 = \sum_{i=1}^{\tau} \frac{(x_i - \bar{x}_\tau)^2}{\tau}, \quad s_{n-\tau}^2 = \sum_{i=\tau+1}^n \frac{(x_i - \bar{x}_{n-\tau})^2}{n-\tau}.$$

For fixed τ , that expression has the same structure as a product of two normal distributions with two inverted gamma distributions. This suggests a normal-inverted gamma type of distribution to represent prior knowledge about $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$. Assuming prior independence between τ and the other parameters, and that $p(\tau | M_3)$ is any discrete distribution on the set $\{1, 2, \dots, n-1\}$, leads to the joint prior parameter p.d.f. :

$$p(\boldsymbol{\mu}, \boldsymbol{\sigma}, \tau | M_3) = \mathcal{N}(\mu_1 | \phi_1, \lambda_1 \sigma_1^2) \mathcal{N}(\mu_2 | \phi_2, \lambda_2 \sigma_2^2) \mathcal{IG}(\sigma_1^2 | \alpha_1, \beta_1) \quad (\text{A2A.2})$$

$$\times \mathcal{IG}(\sigma_2^2 | \alpha_2, \beta_2) p(\tau | M_3)$$

$$= \mathcal{NNIGIG}(\boldsymbol{\mu}, \boldsymbol{\sigma} | \boldsymbol{\phi}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}) p(\tau | M_3)$$

where $\boldsymbol{\phi} = (\phi_1, \phi_2)$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$, $\boldsymbol{\beta} = (\beta_1, \beta_2)$ are the hyperparameters, and $\mathcal{IG}(\sigma^2 | \boldsymbol{\alpha}, \boldsymbol{\beta})$ stands for the inverted gamma distribution with parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. Because of conjugate properties the conditional joint posterior distribution $p(\boldsymbol{\mu}, \boldsymbol{\sigma} | \tau, \mathbf{x}, M_3)$ also belongs to the class of normal-inverted-gamma distributions, but with updated parameters $(\boldsymbol{\phi}', \boldsymbol{\lambda}', \boldsymbol{\alpha}', \boldsymbol{\beta}')$. More precisely,

$$p(\boldsymbol{\mu}, \boldsymbol{\sigma} | \tau, \mathbf{x}, M_3) = \mathcal{NNIGIG}(\boldsymbol{\mu}, \boldsymbol{\sigma} | \boldsymbol{\phi}', \boldsymbol{\lambda}', \boldsymbol{\alpha}', \boldsymbol{\beta}') \quad (\text{A2A.3})$$

where

$$\lambda'_1 = \lambda_1 / (1 + \tau \lambda_1), \quad \lambda'_2 = \lambda_2 / [1 + (n - \tau) \lambda_2],$$

$$\phi'_1 = (1 - \lambda'_1 \tau) \phi_1 + \lambda'_1 \tau \bar{x}_\tau, \quad \phi'_2 = [1 - \lambda'_2 (n - \tau)] \phi_2 + \lambda'_2 (n - \tau) \bar{x}_{n-\tau},$$

$$\alpha'_1 = \alpha_1 + \tau / 2, \quad \alpha'_2 = \alpha_2 + (n - \tau) / 2,$$

$$\beta'_1 = \frac{\tau}{2} \left[s_\tau^2 + (1 - \lambda'_1 \tau) (\phi_1 - \bar{x}_\tau)^2 \right] + \beta_1,$$

$$\beta'_2 = \frac{(n-\tau)}{2} \left[s_{n-\tau}^2 + (1 - \lambda'_2 (n - \tau)) (\phi_2 - \bar{x}_{n-\tau})^2 \right] + \beta_2.$$

The prior predictive distribution $p(\mathbf{x}|\tau, M_3)$ can be determined by dividing the Bayes factor numerator ($p(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathbf{x}|\tau, M_3)$) by (A.3), and cancelling factors involving $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}^2$:

$$p(\mathbf{x}|\tau, M_3) = \left(\frac{1}{2\pi}\right)^{n/2} \sqrt{\frac{\lambda'_1 \lambda'_2}{\lambda_1 \lambda_2} \frac{\beta_1^{\alpha_1} \beta_2^{\alpha_2}}{(\beta'_1)^{\alpha_1} (\beta'_2)^{\alpha_2}} \frac{\Gamma(\alpha'_1) \Gamma(\alpha'_2)}{\Gamma(\alpha_1) \Gamma(\alpha_2)}}. \quad (\text{A2A.4})$$

Given τ and \mathbf{x} , it is then straightforward to show that μ_1 and μ_2 are conditionally distributed as Student t -distributions $ST(\mu_i | a, b, c)$, σ_1^2 and σ_2^2 as inverted gamma distributions $IG(\sigma_i^2 | \alpha, \beta)$. Integrating the appropriate parameters out of expression (A.3), leads directly to these conditional posterior distributions of the parameters before and after the change-point:

$$p(\mu_i | \tau, \mathbf{x}, M_3) = ST\left(\mu_i \mid \phi'_i, \alpha'_i (\lambda'_i \beta'_i)^{-1}, 2\alpha'_i\right), \quad i = 1, 2,$$

$$p(\sigma_i^2 | \tau, \mathbf{x}, M_3) = IG(\sigma_i^2 | \alpha'_i, \beta'_i), \quad i = 1, 2.$$

To draw conclusions regarding the intensity of shifts, it is natural to define the parameters $\delta = \mu_2 - \mu_1$ and $\eta = \sigma_2^2 / \sigma_1^2$. Their conditional posterior distributions can be deduced by simple univariate transformations of variable from the conditional distributions of the original parameters. Given τ and \mathbf{x} , η has a Beta distribution of the second kind (Menzefricke, 1981). It can also be shown that $(\alpha'_1 \beta'_1 \eta / \alpha'_2 \beta'_2)$ is distributed as a Fisher distribution with $2\alpha'_2$ and $2\alpha'_1$ degrees of freedom. The posterior conditional distribution of δ is a Behrens-Fisher distribution which cannot be evaluated in closed form. However, it can be efficiently approximated by a Student t -distribution. Details can be found in Box and Tiao (1973, p. 107).

Since, according to the Bayes theorem, the joint p.d.f. for (\mathbf{x}, τ) is just $p(\mathbf{x}|\tau, M_3)p(\tau|M_3)$, the marginal posterior density of the change-point $\tau = 1, 2, \dots, n-1$ under M_3 is readily seen to be

$$p(\tau | \mathbf{x}, M_3) = \frac{p(\mathbf{x}|\tau, M_3)p(\tau|M_3)}{\sum_{\tau=1}^{n-1} p(\mathbf{x}|\tau, M_3)p(\tau|M_3)} \quad (\text{A2A.5})$$

$$\propto p(\tau|M_3) \sqrt{\lambda'_1 \lambda'_2} \frac{\Gamma(\alpha'_1) \Gamma(\alpha'_2)}{(\beta'_1)^{\alpha_1} (\beta'_2)^{\alpha_2}}$$

This discrete distribution gives, at time point τ , the posterior probability of simultaneous shift occurrence in both the mean level and variance. To draw conclusions regarding the parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$, and the intensity of shifts $\delta = \mu_2 - \mu_1$ and $\eta = \sigma_2^2 / \sigma_1^2$, their marginal posterior distributions must be derived. The corresponding marginal distributions are finite mixtures of the associated conditional distributions weighted by the $n-1$ values of $p(\tau|\mathbf{x}, M_3)$ given by (A.5). For instance, we have

$$p(\mu_i | \mathbf{x}, M_3) = \sum_{\tau=1}^{n-1} p(\mu_i | \tau, \mathbf{x}, M_3) p(\tau | \mathbf{x}, M_3), \quad i = 1, 2$$

Model M_0

– Prior distribution

$$p(\mu, \sigma^2 | M_0) = \mathcal{N}(\mu | \phi, \lambda \sigma^2) IG(\sigma^2 | \alpha, \beta) \quad (\text{A2A.6})$$

$$= \mathcal{NIG}(\mu, \sigma^2 | \phi, \lambda, \alpha, \beta)$$

– *Posterior distribution*

$$p(\mu, \sigma^2 | \mathbf{x}, M_0) = \mathcal{NIG}(\mu, \sigma^2 | \phi', \lambda', \alpha', \beta') \quad (\text{A2A.7})$$

where

$$\alpha' = \alpha + n/2 \text{ and } \beta' = \frac{1}{2} \left[ns_n^2 + (1 - \lambda)^{-1} (\phi - \bar{x}_n)^2 \right] + \beta$$

Model M₁

– *Prior distribution*

$$\begin{aligned} p(\mu, \sigma^2, \tau | M_1) &= \mathcal{N}(\mu_1 | \phi_1, \lambda_1 \sigma^2) \mathcal{N}(\mu_2 | \phi_2, \lambda_2 \sigma^2) \mathcal{IG}(\sigma^2 | \alpha, \beta) p(\tau | M_1) \\ &= \mathcal{NNIG}(\mu, \sigma^2 | \phi, \lambda, \alpha, \beta) p(\tau | M_1) \end{aligned} \quad (\text{A2A.8})$$

– *Conditional posterior distribution*

$$p(\mu, \sigma^2 | \tau, \mathbf{x}, M_1) = \mathcal{NNIG}(\mu, \sigma^2 | \phi', \lambda', \alpha', \beta') \quad (\text{A2A.9})$$

where

$$\begin{aligned} \lambda'_1 &= \lambda_1 / (1 + \tau \lambda_1), \quad \phi'_1 = (1 - \lambda'_1 \tau) \phi_1 + \lambda'_1 \tau \bar{x}_\tau, \\ \lambda'_2 &= \lambda_2 / [1 + (n - \tau) \lambda_2], \quad \phi'_2 = [1 - \lambda'_2 (n - \tau)] \phi_2 + \lambda'_2 (n - \tau) \bar{x}_{n-\tau}, \\ \beta' &= \frac{\tau}{2} \left[s_\tau^2 + (1 - \lambda'_1 \tau) (\phi_1 - \bar{x}_\tau)^2 \right] + \frac{n - \tau}{2} \left[s_{n-\tau}^2 + (1 - \lambda'_2 (n - \tau)) (\phi_2 - \bar{x}_{n-\tau})^2 \right] + \beta, \\ \alpha' &= \alpha + n/2. \end{aligned}$$

– *Prior predictive distribution conditional on τ*

$$p(\mathbf{x} | \tau, M_1) = \left(\frac{1}{2\pi} \right)^{n/2} \sqrt{\frac{\lambda'_1 \lambda'_2}{\lambda_1 \lambda_2} \frac{\beta^\alpha}{(\beta')^{\alpha'}} \frac{\Gamma(\alpha')}{\Gamma(\alpha)}} \quad (\text{A2A.10})$$

– *Marginal posterior distribution of τ*

$$p(\tau | \mathbf{x}, M_1) \propto p(\tau | M_1) \sqrt{\lambda'_1 \lambda'_2} (\beta')^{-\alpha'} \quad (\text{A2A.11})$$

Model M₂

– *Prior distribution*

$$\begin{aligned} p(\mu, \sigma_1^2, \sigma_2^2, \tau | M_2) &= \mathcal{N}(\mu | \phi, \lambda \sigma_1^2) \mathcal{IG}(\sigma_1^2 | \alpha_1, \beta_1) \mathcal{IG}(\sigma_2^2 | \alpha_2, \beta_2) p(\tau | M_2) \\ &= \mathcal{NIGIG}(\mu, \sigma_1^2, \sigma_2^2 | \phi, \lambda, \alpha, \beta) p(\tau | M_2) \end{aligned} \quad (\text{A2A.12})$$

– *Full conditionals*

$$\begin{aligned} p(\mu | \sigma_1^2, \sigma_2^2, \tau, \mathbf{x}, M_2) &= \mathcal{N}\left(\mu \left| \phi'', \frac{\lambda' \sigma_1^2 \sigma_2^2}{\lambda' (n - \tau) \sigma_1^2 + \sigma_2^2} \right.\right) \\ p(\sigma_1^2 | \mu, \sigma_2^2, \tau, \mathbf{x}, M_2) &= \mathcal{IG}\left(\sigma_1^2 \left| \alpha'_1, \frac{\tau}{2} \left[s_\tau^2 + (\bar{x}_\tau - \mu)^2 + \frac{(\mu - \phi)^2}{\tau \lambda} \right] + \beta_1 \right.\right) \\ p(\sigma_2^2 | \mu, \sigma_1^2, \tau, \mathbf{x}, M_2) &= \mathcal{IG}\left(\sigma_2^2 \left| \alpha'_2, \frac{n - \tau}{2} \left[s_{n-\tau}^2 + (\bar{x}_{n-\tau} - \mu)^2 \right] + \beta_2 \right.\right) \\ p(\tau | \mu, \sigma_2^2, \sigma_1^2, \mathbf{x}, M_2) &\propto p(\mathbf{x} | \mu, \sigma_2^2, \sigma_1^2, \tau, M_2) p(\tau | M_2) \end{aligned}$$

where

$$\begin{aligned}\lambda' &= \lambda / (1 + \tau\lambda), \quad \phi' = (1 - \lambda'\tau) \phi + \lambda'\tau\bar{x}_\tau, \\ \phi'' &= \frac{1}{\lambda'(n-\tau)\sigma_1^2 + \sigma_2^2} [\sigma_2^2\phi' + \lambda'(n-\tau)\sigma_1^2\bar{x}_{n-\tau}] \\ \beta'_1 &= \frac{\tau}{2} [s_\tau^2 + (1 - \lambda'\tau)(\phi - \bar{x}_\tau)^2] + \beta_1, \quad \beta'_2 = \frac{n-\tau}{2} s_{n-\tau}^2 + \beta_2, \\ \alpha'_1 &= \alpha_1 + \frac{\tau+1}{2}, \quad \alpha'_2 = \alpha_2 + \frac{n-\tau}{2}\end{aligned}$$

B. Implementation of Chib's approach

In the context of the general change-point model (B.1), the full conditional distributions are $p(\tau | \theta_1, \theta_2, \mathbf{x})$, $p(\theta_1 | \theta_2, \tau, \mathbf{x})$ and $p(\theta_2 | \theta_1, \tau, \mathbf{x})$, and $\{\theta^{(j)} = (\theta_1^{(j)}, \theta_2^{(j)}, \tau^{(j)}) ; j = 1, \dots, m\}$ are the m draws from the joint posterior density $p(\theta | \mathbf{x}, M_k)$ obtained using the Gibbs sampler, after the first t values of the chain have been discarded.

Chib's estimate for $p(\mathbf{x} | M_k)$ is given by

$$\hat{p}_{CH}(\mathbf{x} | M_k) = \frac{p(\mathbf{x} | \theta^*, M_k) p(\theta^* | M_k)}{\hat{p}(\theta^* | \mathbf{x}, M_k)} \quad (\text{A2B.1})$$

where $\hat{p}(\theta^* | \mathbf{x}, M_k)$ is an estimate for $p(\theta | \mathbf{x}, M_k)$ evaluated at θ^* . Using basic probability rules, the posterior density ordinate $p(\theta^* | \mathbf{x}, M_k)$, for a given value θ^* , can be expressed as

$$p(\theta^* | \mathbf{x}, M_k) = p(\theta_1^* | \mathbf{x}, M_k) p(\theta_2^* | \theta_1^*, \mathbf{x}, M_k) p(\tau^* | \theta_1^*, \theta_2^*, \mathbf{x}, M_k) \quad (\text{A2B.2})$$

The last term of the right-hand side is nothing other than the full conditional distribution of τ evaluated at τ^* after setting $(\theta_1, \theta_2, \tau) = (\theta_1^*, \theta_2^*, \tau^*)$. Because all full conditionals are assumed to be expressible in closed form, $p(\tau^* | \theta_1^*, \theta_2^*, \mathbf{x}, M_k)$ can be evaluated exactly. For the same reason, the first term $p(\theta_1^* | \mathbf{x}, M_k)$, which is the marginal ordinate, can be estimated from the m draws of the Gibbs run using the Rao-Blackwellized estimate:

$$\hat{p}(\theta_1^* | \mathbf{x}, M_k) = \frac{1}{m} \sum_{j=1}^m p(\theta_1^* | \theta_2^{(j)}, \tau^{(j)}, \mathbf{x}, M_k) \quad (\text{A2B.3})$$

However, the partial conditional ordinate $p(\theta_2^* | \theta_1^*, \mathbf{x}, M_k)$ in (B.2) cannot be obtained directly. In fact, the draws of θ_2 from the Gibbs sampler are from the distribution $p(\theta_2 | \mathbf{x}, M_k)$ and not from $p(\theta_2 | \theta_1, \mathbf{x}, M_k)$, so that an estimate for this density cannot be calculated using these simulated values. The simple solution proposed by Chib (1995) is to continue sampling for an additional m iterations with the conditional densities

$$p(\theta_2 | \theta_1^*, \tau, \mathbf{x}, M_k) \quad \text{and} \quad p(\tau | \theta_1^*, \theta_2, \mathbf{x}, M_k)$$

where in each of these densities, θ_1 is set equal to θ_1^* . Consequently,

$$\hat{p}(\theta_2^* | \theta_1^*, \mathbf{x}, M_k) = \frac{1}{m} \sum_{j=1}^m p(\theta_2^* | \theta_1^*, \tau^{(j)}, \mathbf{x}, M_k) \quad (\text{A2B.4})$$

is a Rao-Blackwellized estimate for the second term in (B.2).

Substituting the density estimates (B.3) and (B.4) into (B.2), and then the obtained estimate $\widehat{p}(\theta^* | \mathbf{x}, M_k)$ into (B.1) yields finally to

$$\widehat{p}_{CH}(\mathbf{x} | M_k) = \frac{p(\mathbf{x} | \theta^*, M_k) p(\theta^* | M_k)}{\widehat{p}(\theta_1^* | \mathbf{x}, M_k) \widehat{p}(\theta_2^* | \theta_1^*, \mathbf{x}, M_k) p(\tau^* | \theta_1^*, \theta_2^*, \mathbf{x}, M_k)} \quad (\text{A2B.5})$$

Since expression (B.5) holds for any value of θ , the choice of θ^* is not critical. However, Chib (1995) pointed out that efficiency considerations dictate that for a given number of replicates m , the prior predictive is likely to be more accurately estimated at a high density point, where more samples are available, than at a point in the tails. The posterior mode or the maximum likelihood estimate, which can be easily approximated from the initial Gibbs output, should be used. Alternatively, one can also use the posterior mean provided there is no concern that it is a low density point. Finally, this approach can be easily generalized to more than three unknown parameters.

Annexe C

Densités de probabilité et principaux moments

Loi normale $\mu \in \mathfrak{R}, \sigma^2 \in \mathfrak{R}^+, y \in \mathfrak{R}$

$$\mathcal{N}(y | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left\{ -\frac{1}{2\sigma^2} (y - \mu)^2 \right\}$$

$$\mathbb{E}\{y\} = \mu \quad \mathbb{V}\{y\} = \sigma^2$$

Loi gamma inverse $\alpha \in \mathfrak{R}^+, \beta \in \mathfrak{R}^+, y \in \mathfrak{R}^+$

$$\mathcal{IG}(y | \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-(\alpha+1)} \exp \left\{ -\frac{\beta}{y} \right\}$$

$$\mathbb{E}\{y\} = \frac{\beta}{\alpha - 1} \quad \mathbb{V}\{y\} = \frac{\beta^2}{(\alpha - 1)^2 (\alpha - 2)}$$

Loi de Student $\mu \in \mathfrak{R}, \lambda \in \mathfrak{R}^+, \alpha \in \mathfrak{R}^+, y \in \mathfrak{R}$

$$ST(y | \mu, \lambda, \alpha) = \frac{\Gamma\left[\frac{1}{2}(\alpha + 1)\right]}{\Gamma\left(\frac{1}{2}\alpha\right)} \left(\frac{\lambda}{\alpha\pi}\right)^{1/2} \left[1 + \frac{\lambda}{\alpha}(y - \mu)^2\right]^{-(\alpha+1)/2}$$

$$\mathbb{E}\{y\} = \mu \quad \mathbb{V}\{y\} = \frac{\alpha}{\lambda(\alpha - 2)}$$

Loi de Fisher $\alpha \in \mathfrak{R}^+, \beta \in \mathfrak{R}^+, y \in \mathfrak{R}^+$

$$\mathcal{F}(y | \alpha, \beta) = \frac{\alpha^{\alpha/2} \beta^{\beta/2}}{B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} \frac{y^{\alpha/2-1}}{(\beta + \alpha y)^{(\alpha+\beta)/2}}$$

$$\mathbb{E}\{y\} = \frac{\beta}{\beta - 2} \quad \mathbb{V}\{y\} = \frac{2\beta^2(\alpha + \beta - 2)}{\alpha(\beta - 2)^2(\beta - 4)}$$

Loi normale multidimensionnelle $\mu \in \mathfrak{R}^p, \mathbf{P}$ matrice $p \times p$ sym. et déf. pos., $y \in \mathfrak{R}^p$

$$\mathcal{N}_p(y | \mu, \mathbf{P}) = (2\pi)^{-p/2} |\mathbf{P}|^{1/2} \exp\left\{-\frac{1}{2}(y - \mu)^t \mathbf{P} (y - \mu)\right\}$$

$$\mathbb{E}\{y\} = \mu \quad \text{Cov}\{y\} = \mathbf{P}^{-1}$$

Loi de Wishart $2a > p - 1, \mathbf{B}$ matrice $p \times p$ sym. non sing., \mathbf{P} sym. et déf. pos.

$$\mathcal{W}_p(\mathbf{P} | a, \mathbf{B}) = \frac{\pi^{-p(p-1)/4} |\mathbf{B}|^a}{\prod_{k=1}^p \Gamma\left[\frac{1}{2}(2a + 1 - k)\right]} |\mathbf{P}|^{a-(p+1)/2} \exp\{-tr(\mathbf{BP})\}$$

$$\mathbb{E}\{\mathbf{P}\} = a\mathbf{B}^{-1} \quad \mathbb{E}\{\mathbf{P}^{-1}\} = \left(a - \frac{p+1}{2}\right)^{-1} \mathbf{B}$$

Loi de Student multi. $\mathbf{m} \in \mathfrak{R}^p, c \in \mathfrak{R}^+, \mathbf{A}$ matrice $p \times p$ sym. et déf. pos., $y \in \mathfrak{R}^p$

$$ST_p(y | \mathbf{m}, \mathbf{A}, c) = \frac{\Gamma\left[\frac{1}{2}(c+p)\right]}{\Gamma\left(\frac{1}{2}c\right) (c\pi)^{p/2}} |\mathbf{A}|^{1/2} \left[1 + \frac{1}{c}(y - \mathbf{m})^t \mathbf{A} (y - \mathbf{m})\right]^{-(c+p)/2}$$

$$\mathbb{E}\{\mathbf{y}\} = \mathbf{m} \quad \text{Cov}\{\mathbf{y}\} = \left(\frac{c}{c-2}\right) \mathbf{A}^{-1}$$

Loi de Dirichlet $\xi_h > 0, h = 1, \dots, n-1, \xi_0 = \sum_{h=1}^{n-1} \xi_h$

$$\mathcal{D}_{n-1}(\boldsymbol{\pi} \mid \xi_1, \dots, \xi_{n-1}) = \frac{\Gamma(\xi_0)}{\prod_{h=1}^{n-1} \Gamma(\xi_h)} \prod_{h=1}^{n-1} \pi_h^{\xi_h-1}$$

$$\mathbb{E}\{\pi_j\} = \frac{\xi_j}{\xi_0} \quad \mathbb{V}\{\pi_j\} = \frac{(\xi_0 - \xi_j) \xi_j}{\xi_0^2 (\xi_0 + 1)} \quad \text{Cov}\{\pi_i, \pi_j\} = \frac{-\xi_i \xi_j}{\xi_0^2 (\xi_0 + 1)}$$

Loi beta $\alpha \in \mathfrak{R}^+, \beta \in \mathfrak{R}^+, y \in (0,1)$

$$\mathcal{B}e(y \mid \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\beta)} y^{\alpha-1} (1-y)^{\beta-1}$$

$$\mathbb{E}\{y\} = \frac{\alpha}{\alpha + \beta} \quad \mathbb{V}\{y\} = \frac{\alpha\beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$$

Annexe D

Échantillonnage de Gibbs

L'échantillonnage de Gibbs est un outil de simulation Monte Carlo par chaînes de Markov (MCMC) qui permet d'obtenir des réalisations provenant approximativement d'une loi jointe d'un ensemble de variables aléatoires. Ces réalisations peuvent être employées pour évaluer diverses caractéristiques de cette distribution, notamment les densités marginales.

D.1 Algorithme de Gibbs

Étant donné un ensemble de p variables aléatoires U_1, \dots, U_p , nous souhaitons évaluer par échantillonnage de Gibbs les distributions marginales $p(U_i)$, $i = 1, \dots, p$, ainsi que certaines de leurs caractéristiques. Supposons que la loi jointe $p(U_1, \dots, U_p)$ est entièrement et uniquement déterminée par les lois conditionnelles complètes que nous notons $p(U_i | U_j, j \neq i)$, $i = 1, \dots, p$. Partant d'une valeur initiale arbitraire $\mathbf{U}^{(0)} = (U_1^{(0)}, \dots, U_p^{(0)})$, l'algorithme de Gibbs consiste à tirer successivement

- $U_1^{(1)}$ à partir de $p(U_1 | U_2^{(0)}, \dots, U_p^{(0)})$
- $U_2^{(1)}$ à partir de $p(U_2 | U_1^{(1)}, U_3^{(0)}, \dots, U_p^{(0)})$,
- $U_3^{(1)}$ à partir de $p(U_3 | U_1^{(1)}, U_2^{(1)}, U_4^{(0)}, \dots, U_p^{(0)})$,
- et ainsi de suite jusqu'à $U_p^{(1)}$ à partir de $p(U_p | U_1^{(1)}, \dots, U_{p-1}^{(1)})$

Ce cycle est répété t fois (le temps de chauffe) pour obtenir $\mathbf{U}^{(t)} = (U_1^{(t)}, \dots, U_p^{(t)})$.

On peut montrer que le vecteur $\mathbf{U}^{(t)}$ ainsi obtenu converge en loi vers un échantillon provenant de la distribution $p(U_1, \dots, U_p)$ lorsque t devient grand. Les conditions de régularité sous lesquelles l'échantillonnage de Gibbs converge sont quasi-inexistantes (Gelfand et Smith (1990), Robert et Casella (1999)). Ainsi, si on répète tout le processus m fois

en parallèle (les réplicats), on obtiendra un échantillon de m vecteurs $U_1^{(t)}, \dots, U_m^{(t)}$ indépendamment et identiquement distribués selon la loi jointe $p(U_1, \dots, U_p)$. L'échantillon $U_1^{(t)}, \dots, U_m^{(t)}$ peut être utilisé pour approcher les densités marginales ou toute autre caractéristique de la loi jointe (espérances, intervalles de crédibilité, quantiles, etc.) par de simples sommes finies d'expressions connues.

Cette manière de procéder nous oblige à éliminer $(t-1) \times m$ tirages. Nous lui avons donc préféré la stratégie qui consiste à générer une seule chaîne de taille $t+m$, où seules les m dernières valeurs sont conservées pour l'estimation. Cette procédure est beaucoup plus économique en terme de temps de simulation. L'échantillon $U_1^{(t)}, \dots, U_m^{(t)}$ résultant provient toujours de la distribution $p(U_1, \dots, U_p)$. Toutefois, ces valeurs ne sont pas indépendantes. Cela ne cause pas de problème pour l'estimation des densités et de leurs espérances. En revanche, si nous souhaitons estimer leur variance, il est préférable d'employer des techniques issues du domaine de l'analyse des séries chronologiques.

D.2 Estimation à partir de l'échantillonnage de Gibbs

On peut envisager d'estimer, par exemple, la densité marginale $p(U_i)$ par l'histogramme des valeurs simulées $(U_{i1}^{(t)}, \dots, U_{im}^{(t)})$ et son espérance par la moyenne $(1/m) \sum_j U_{ij}^{(t)}$. Mais ces estimations se révèlent rudimentaires puisqu'elles n'exploitent pas le fait que les lois conditionnelles complètes sont connues. Des estimateurs pour la densité et son espérance, qui mobilisent cette information, sont respectivement :

$$\hat{p}(U_i) = \frac{1}{m} \sum_{j=1}^m p(U_i | U_{ij}^{(t)}, j \neq i) \quad (\text{D.1})$$

et

$$\hat{\mathbb{E}}(U_i) = \frac{1}{m} \sum_{j=1}^m \mathbb{E}\{U_i | U_{ij}^{(t)}, j \neq i\} \quad (\text{D.2})$$

L'argument formel qui justifie l'emploi de ces expressions repose sur le théorème de Rao-Blackwell (Lehmann (1983)). C'est pourquoi, dans la littérature, on les retrouvent souvent sous le nom d'estimateurs « Rao-Blackwellisés ».

Par ailleurs, il est aussi possible d'approcher les caractéristiques d'une fonction des variables aléatoires $W(U_1, \dots, U_p)$. Pour chaque vecteur $(U_{1j}^{(t)}, \dots, U_{pj}^{(t)})$, on peut évaluer $W_j^{(t)} = W(U_{1j}^{(t)}, \dots, U_{pj}^{(t)})$ dont la loi marginale est approximativement $p(W)$. Des estimateurs « Rao-Blackwellisés » peuvent être obtenus. En effet, si U_i apparaît comme argument dans la fonction W , la loi conditionnelle complète $p(W | U_j, j \neq i)$ peut être déterminée à partir de $p(U_i | U_j, j \neq i)$ par un simple changement de variables.

D.3 Éléments d'analyse de la convergence pour les modèles MM_2 et MM_3

La mise en oeuvre de l'algorithme de Gibbs requiert le choix d'un temps de chauffe t et du nombre de répliqués m . Dans ce qui suit nous présentons quelques graphiques permettant d'examiner la convergence de l'algorithme pour les applications du Chapitre 5. Évidemment, cette étude de la convergence est très sommaire. Pour plus de détails concernant les problèmes de convergence des outils MCMC, le lecteur est invité à consulter Robert (1998) et Robert et Casella (1999).

Après quelques expérimentations, le temps de chauffe t a été fixé à 2000 itérations pour le modèle MM_2 et 5000 pour le MM_3 . Cette analyse a consisté principalement à étudier la sensibilité des conditions initiales sur l'inférence. Nous avons comparé les histogrammes des densités marginales de chaque paramètre pour plusieurs valeurs t . Nous avons arrêté notre choix lorsque l'augmentation du temps de chauffe ne semblait plus avoir d'effet sur la forme de ces densités. Le tracé des 4000 premières valeurs générées est présenté aux figures D.1 à D.3 pour le modèle MM_2 et aux figures D.4 à D.6 pour le modèle MM_3 . On remarque en particulier la présence d'une plus forte dépendance d'une itération à l'autre pour ce dernier modèle. Cela révèle une plus lente convergence vers la loi jointe. C'est pourquoi nous avons considéré un temps de chauffe plus grand pour MM_3 que pour MM_2 .

Le nombre de répliqués a été fixé à 2500 pour les deux modèles. Ici encore, ce choix repose sur des analyses graphiques en comparant les estimations « Rao-Blackwellisées » pour diverses valeurs de m . La statistique de Gelman et al. (1995), $\sqrt{\widehat{R}}$, dont la définition est donnée en page 331 de leur ouvrage, a aussi été calculée pour différentes valeurs de m . Les résultats pour le modèle MM_3 sont donnés au Tableau D.1. Les auteurs recommandent de choisir le nombre de répliqués qui correspond à une valeur de $\sqrt{\widehat{R}}$ inférieure à 1.2.

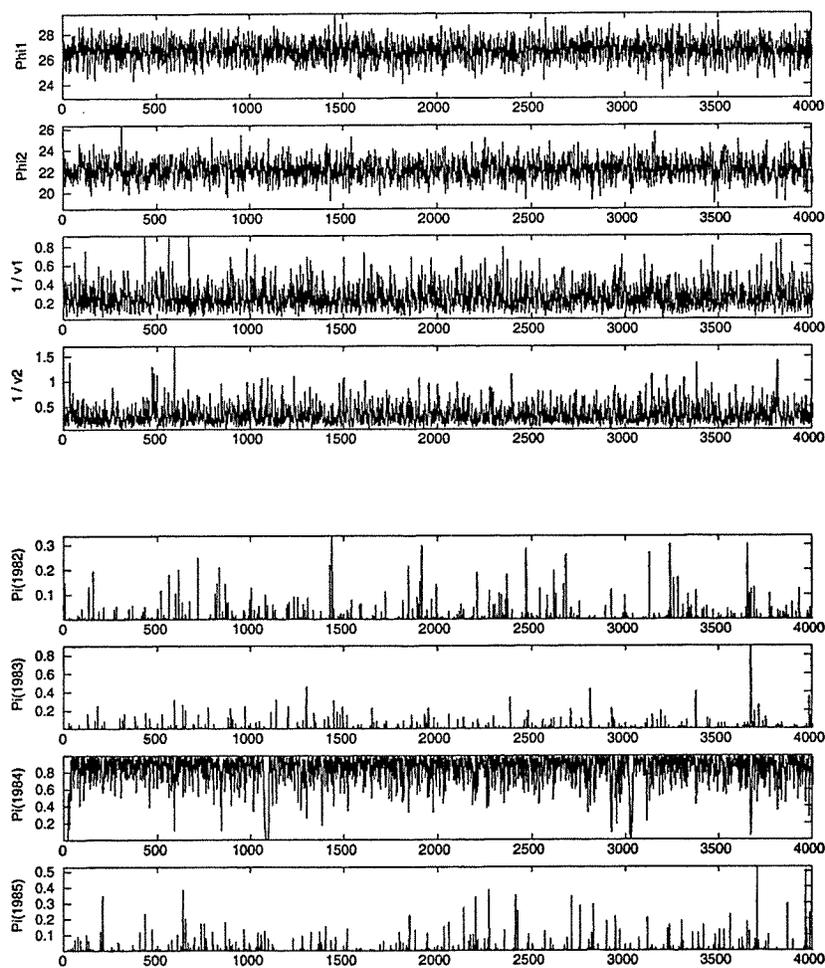


FIG. D.1 – Simulations des paramètres régionaux des 4000 premières itérations pour le modèle MM2

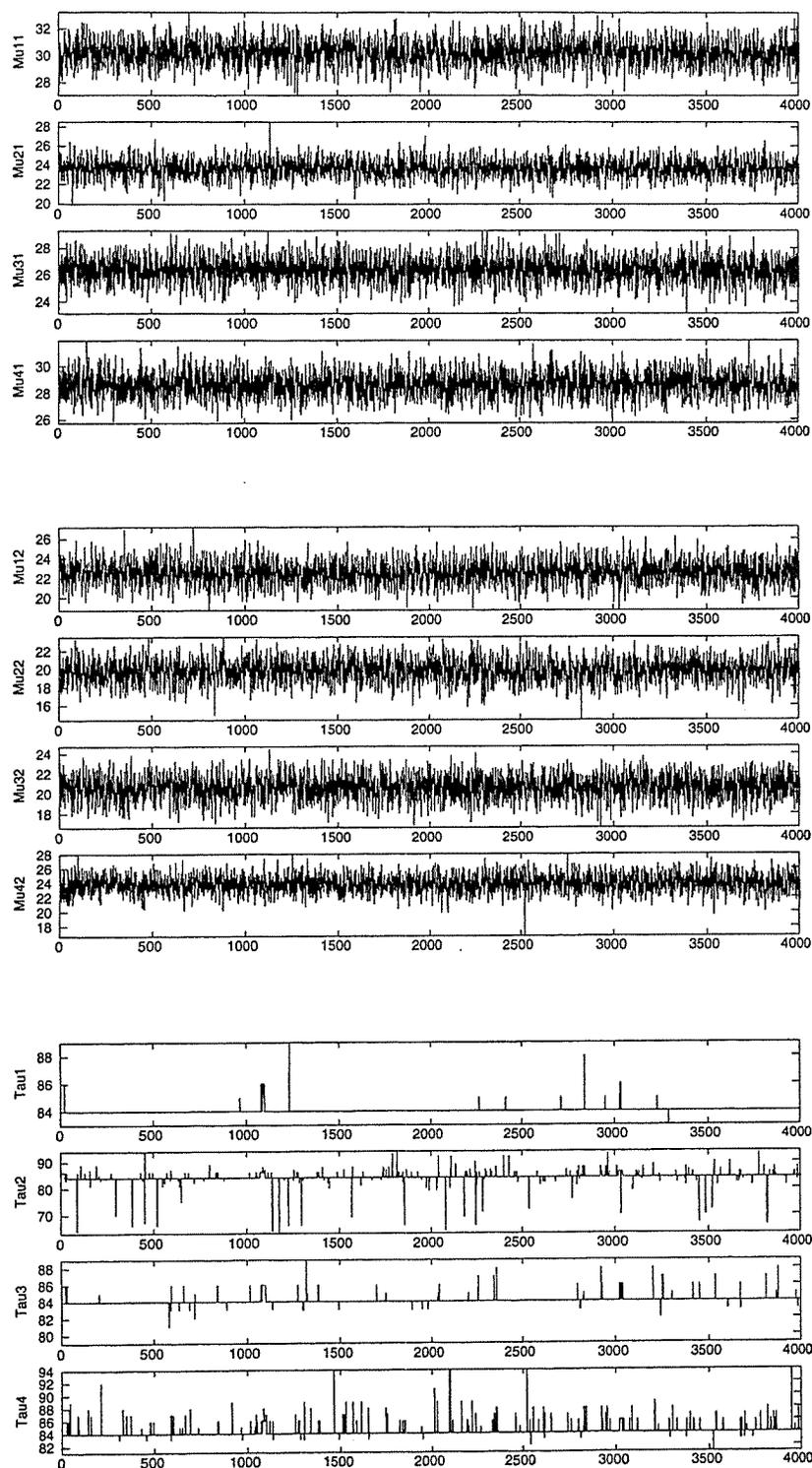


FIG. D.2 – Simulations des paramètres locaux des 4000 premières itérations pour le modèle MM2

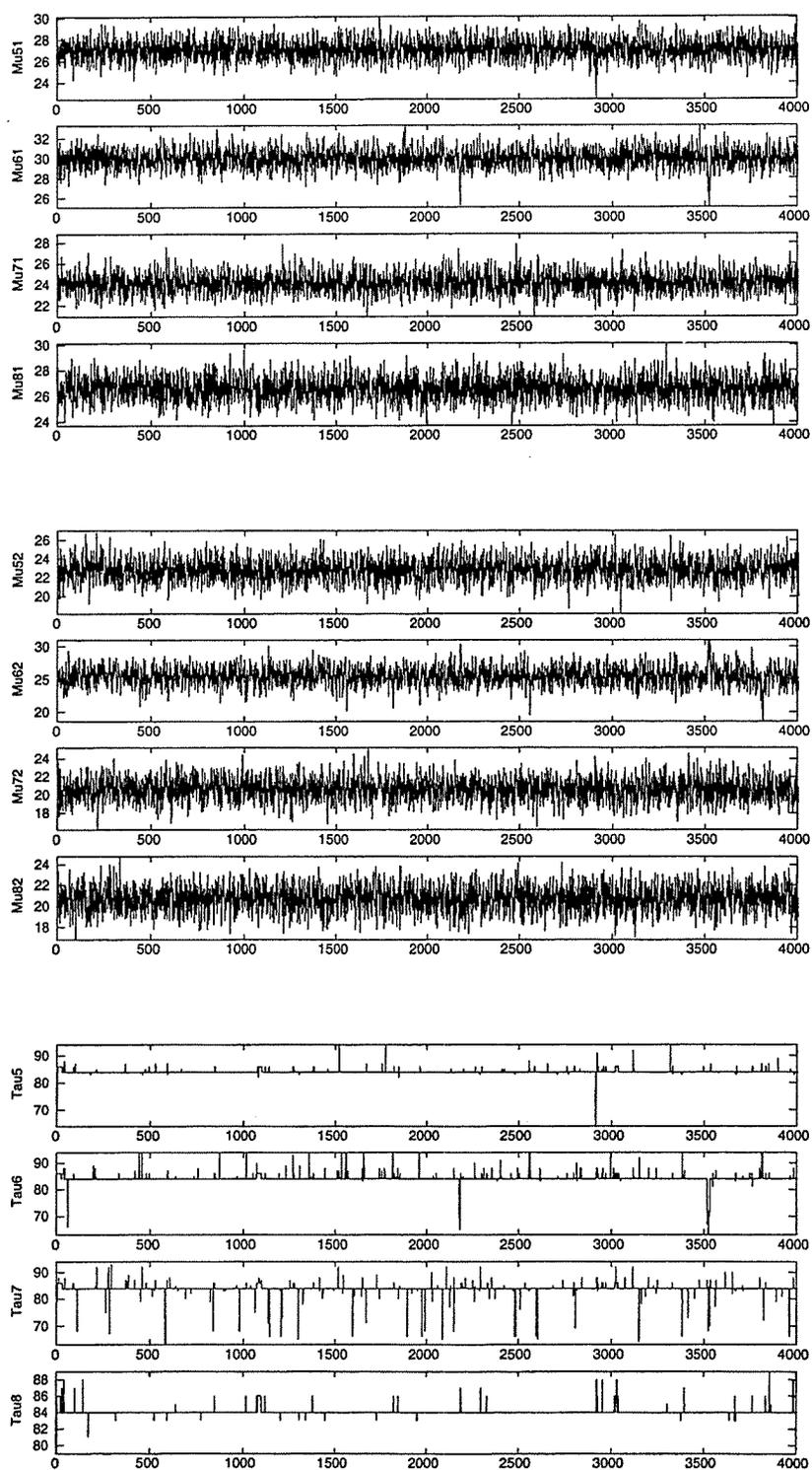


FIG. D.3 – Simulations des paramètres locaux des 4000 premières itérations pour le modèle MM2 (suite)

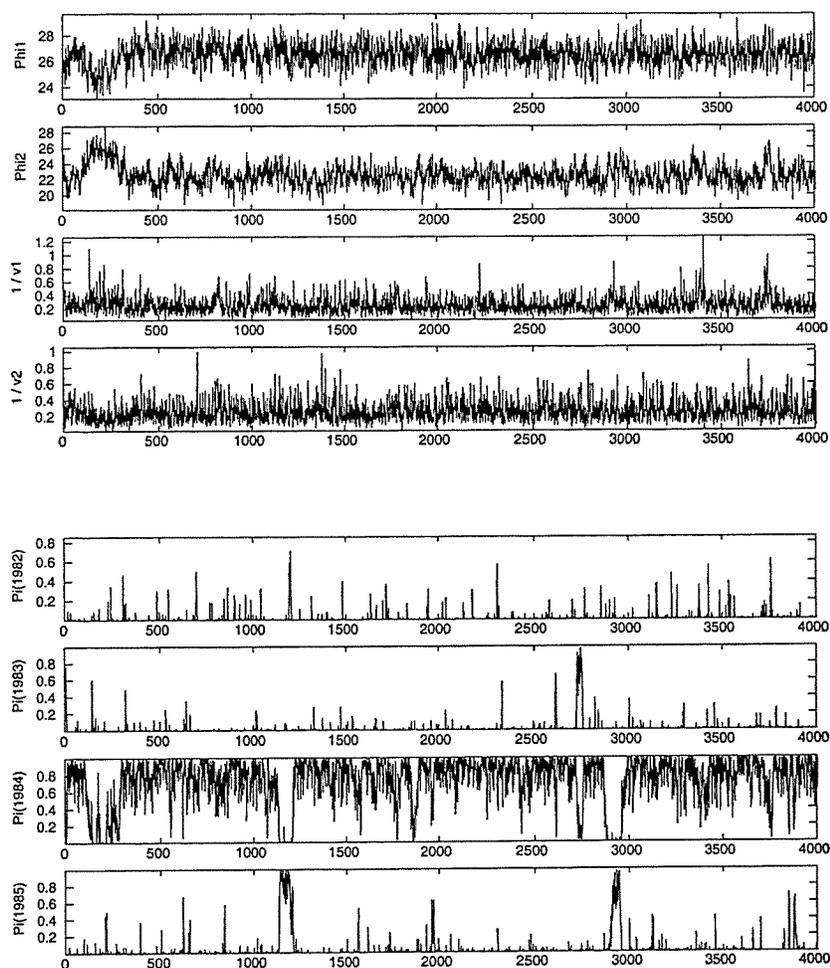


FIG. D.4 – Simulations des paramètres régionaux des 4000 premières itérations pour le modèle MM3

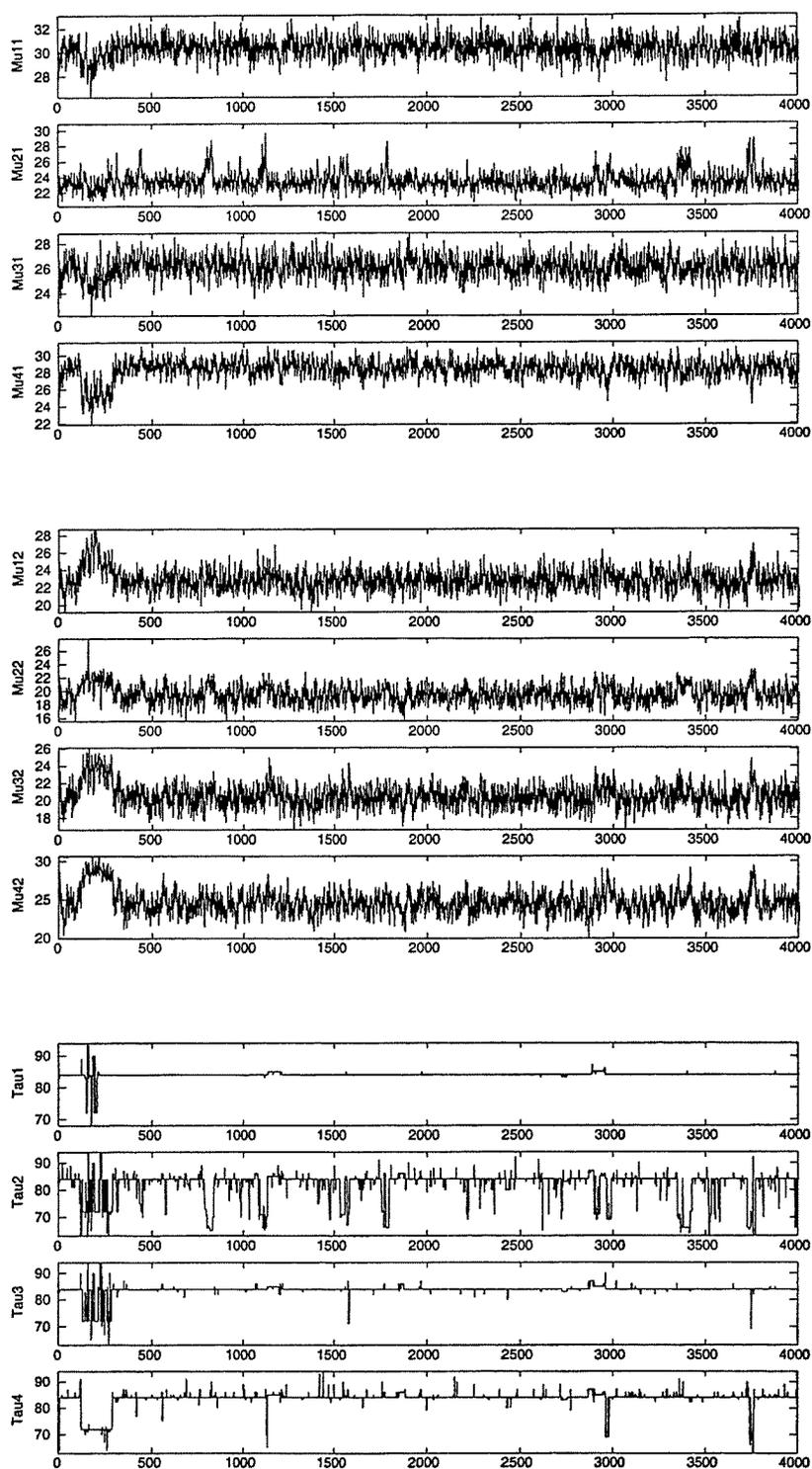


FIG. D.5 – Simulations des paramètres locaux des 4000 premières itérations pour le modèle MM3

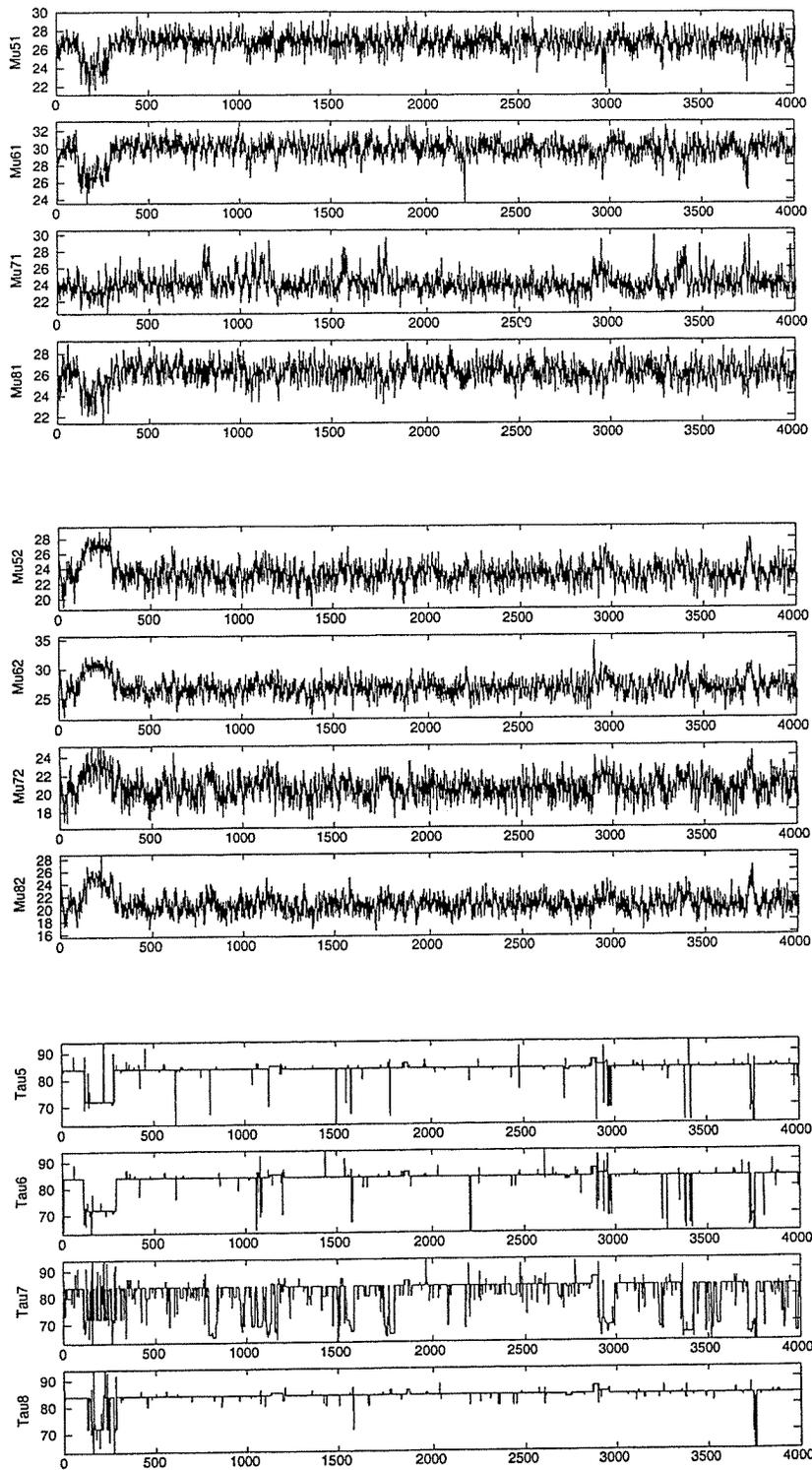


FIG. D.6 – Simulations des paramètres locaux des 4000 premières itérations pour le modèle MM2 (suite)

TAB. D.1 – Valeurs du R de Gelman et al. (1995) pour les paramètres du modèle $MM3$

Itérations	μ_{11}	μ_{21}	μ_{31}	μ_{41}	μ_{51}	μ_{61}	μ_{71}	μ_{81}
50	1.067	1.126	1.082	1.118	1.060	1.130	1.075	1.162
250	1.012	1.016	1.002	1.013	1.020	1.019	1.005	1.011
500	1.007	1.012	1.004	1.005	1.006	1.009	1.002	1.006
1000	1.009	1.001	1.002	1.012	1.011	1.011	1.000	1.008
2500	1.003	1.000	1.002	1.006	1.006	1.006	1.001	1.005
5000	1.001	1.001	1.001	1.003	1.002	1.002	1.001	1.002
	μ_{12}	μ_{22}	μ_{32}	μ_{42}	μ_{52}	μ_{62}	μ_{72}	μ_{82}
50	1.053	1.060	1.061	1.112	1.167	1.264	1.061	1.096
250	1.003	1.007	1.000	0.999	1.000	1.004	1.005	1.000
500	1.006	1.007	1.003	1.002	1.001	1.002	1.008	1.004
1000	1.006	1.006	1.007	1.010	1.009	1.010	1.010	1.010
2500	1.004	1.003	1.003	1.004	1.004	1.004	1.004	1.006
5000	1.004	1.003	1.004	1.006	1.005	1.005	1.003	1.006
	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	τ_7	τ_8
50	1.021	1.373	1.112	1.007	1.058	1.155	1.135	1.146
250	1.168	1.026	1.045	1.013	1.025	1.033	1.002	1.055
500	1.112	1.018	1.027	1.006	1.012	1.015	1.007	1.029
1000	1.052	1.012	1.017	1.021	1.021	1.016	1.010	1.023
2500	1.010	1.004	1.001	1.003	1.003	1.002	1.004	1.002
5000	1.002	1.004	1.004	1.007	1.007	1.002	1.003	1.004
	$1/\sigma^2$	ϕ_1	ϕ_2	$1/v_1^2$	$1/v_2^2$	π_{1984}		
50	1.029	1.047	1.047	1.055	1.055	1.019		
250	1.006	1.003	1.003	1.013	1.013	1.003		
500	1.004	1.002	1.002	1.005	1.005	1.001		
1000	1.001	1.002	1.002	1.004	1.004	1.001		
2500	1.000	1.001	1.001	1.001	1.001	1.004		
5000	1.000	1.000	1.000	1.000	1.000	1.002		

Annexe E

Article III

Retrospective multivariate Bayesian change-point analysis: A simultaneous single change in the mean of several hydrological sequences

Luc Perreault^{1,2}, Éric Parent¹, Jacques Bernier^{1,2}, Bernard Bobée², Michel Slivitzky²

¹ENGREF ²INRS-Eau

Stochastic Environmental Research and Risk Assessment (2000), 14: 243-261
Research has been performed at ENGREF and drafting of the document at INRS-Eau.

Abstract. Although hydrological time series for different sites in a given region are usually correlated and that climate changes should have a regional impact on water resources, very little has appeared in the literature about multivariate change-point analysis. This paper generalizes the univariate Bayesian approach for the detection of a single shift in the mean level to study a change in the mean-vector of a sequence of multivariate normal vectors. Two different problems are considered: the first one is the estimation of the unknown regional change-point under the hypothesis that a shift occurred, while the second one is the overall assessment of change versus no change. This method is illustrated by an application to streamflow data series for six rivers situated in the Northern Québec Labrador region.

Keywords. Bayesian analysis; Multivariate change-point analysis; Streamflow data; Regional analysis; Climate change; Change in the mean-vector

E.1 Introduction

One important step in climate change studies is the estimation of the unknown change-point in hydrometeorological time series. Since climate change should have a global impact, it is natural to think that it may induce a common change-point for all series which are subject to the same climatic phenomena. In number of studies the assumption of a common date of change is assumed implicitly. In fact, conclusions about a global change-point are generally drawn from estimates obtained by applying independently as much univariate analyses as hydrometeorological times series are available for a given region (Bruneau and Rassam, 1983; Hubert et al., 1989; Servat et al., 1997; Paturel et al., 1997; Lubès-Niel et al., 1998; Perreault et al., 1999; Kiely, 1999; and references therein). Even though this type of approach gives interesting insights about regional impact of climate change, the model behind it considers a change-point parameter for each sequence and does not formally

represent the multivariate nature of the problem. Consequently, there is no way but empirical or *ad hoc* calculations to evaluate uncertainty about the global estimate, and correlations that may exist between the series are not taken into account. Assuming a common change-point in one hand, and on the other hand that there is no relation between the observations from site to site is quite inconsistent. While the univariate change-point problem has been extensively studied in the literature, multivariate change-point analysis has been less widely covered. In stochastic hydrology and climatology, to our knowledge, Bayesian multivariate estimation of an unknown change-point has not been considered. In a recent status report on detection of climate change, Barnett et al. (1999) identified multivariate analysis as one of the most urgent research areas to be developed. These authors also outlined some of the advantages of Bayesian analysis compared to conventional methods for the detection of climate changes.

In this paper, the Bayesian approach proposed by Perreault et al. (1999) is generalized to the case of a single change in the mean-vector of a sequence of multivariate normal vectors. The model assumes a simultaneous shift in the mean level for each sequence, and constant covariance matrix before and after the change-point. Interest is mainly focussed on the estimation of the global unknown change-point. The aim is to outline that this problem can be straightforwardly addressed through the Bayesian framework, which in turn can be easily solved by implementation of explicit inference technics when conjugate prior distributions are assumed.

In Section 2, we formulate the multivariate normal change-point model. To illustrate the principle, we derive in Section 3 the posterior distribution of the change-point using conjugate prior distributions. It is seen that this density is expressible in closed form and therefore results to make inference are directly available under the hypothesis that a change has occurred. In Section 4, we briefly discuss the problem of the assessment of change against no change. Finally, Section 5 offers an application to streamflow data series of six rivers situated in Northern Québec.

E.2 The multivariate normal change-point model

In the following, we consider that the data are under a form of a $p \times n$ array

$$\mathbf{X} = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1\tau} & X_{1(\tau+1)} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2\tau} & X_{2(\tau+1)} & \dots & X_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \dots & \vdots \\ X_{p1} & X_{p2} & \dots & X_{p\tau} & X_{p(\tau+1)} & \dots & X_{pn} \end{pmatrix} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_\tau, \dots, \mathbf{X}_n)$$

Each row represents observations over time. Then, the columns $\mathbf{X}_1, \dots, \mathbf{X}_\tau, \dots, \mathbf{X}_n$ are p -dimensional random vectors observed at n consecutive equally spaced time points (say, the annual runoff measured at p neighboring sites). Suppose that, due to some exogenous factors (e.g. a climatic change), the first and second parts of the sequences of random variables operate at two different mean-vectors, respectively μ_1 and μ_2 , but with same covariance matrix Σ . This situation can be represented by the following model:

$$\begin{aligned} \mathbf{X}_i &\sim \mathcal{N}_p(\mathbf{x}_i \mid \mu_1, \mathbf{P}), & i = 1, \dots, \tau \\ \mathbf{X}_i &\sim \mathcal{N}_p(\mathbf{x}_i \mid \mu_2, \mathbf{P}), & i = \tau + 1, \dots, n \end{aligned} \quad (\text{E.1})$$

where $\mu_1 \neq \mu_2$, τ is the unknown change-point, \mathbf{x}_i a realization of the random vector \mathbf{X}_i , and $\mathcal{N}_p(\mathbf{x} \mid \mu, \mathbf{P})$ stands for the multivariate normal probability density function (p.d.f.):

$$\mathcal{N}_p(\mathbf{x} \mid \mu, \mathbf{P}) = (2\pi)^{-p/2} |\mathbf{P}|^{1/2} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mu)^t \mathbf{P} (\mathbf{x} - \mu) \right\}, \quad \mathbf{x} \in \mathbb{R}^p, \quad (\text{E.2})$$

$|\cdot|$ denoting the determinant of a matrix. That is, the shift in the means occurs at the same position τ in each sequence. The parameter \mathbf{P} labels the precision matrix (inverse of the matrix of covariances, $\mathbf{P} = \Sigma^{-1}$) and represents the unknown but constant structure of the distribution on both sides of τ . As usual, for such a model, the unknown change-point τ can take values between 1 and $n - 1$. Therefore, we assume that a change occurred with certainty.

The likelihood function resulting from a sample of size n , $\mathbf{D} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$, generated by this model can be written as

$$\begin{aligned} p(\mathbf{D} \mid \mu_1, \mu_2, \tau, \mathbf{P}) &= \prod_{i=1}^{\tau} \mathcal{N}_p(\mathbf{x}_i \mid \mu_1, \mathbf{P}) \prod_{i=\tau+1}^n \mathcal{N}_p(\mathbf{x}_i \mid \mu_2, \mathbf{P}) \quad (\text{E.3}) \\ &= (2\pi)^{-np/2} |\mathbf{P}|^{n/2} \exp \left\{ -\frac{\tau}{2} \text{tr} \left[\mathbf{P} \left(\mathbf{S}_{\tau} + (\bar{\mathbf{x}}_{\tau} - \mu_1) (\bar{\mathbf{x}}_{\tau} - \mu_1)^t \right) \right] \right\} \\ &\quad \times \exp \left\{ -\frac{n-\tau}{2} \text{tr} \left[\mathbf{P} \left(\mathbf{S}_{n-\tau} + (\bar{\mathbf{x}}_{n-\tau} - \mu_2) (\bar{\mathbf{x}}_{n-\tau} - \mu_2)^t \right) \right] \right\} \end{aligned}$$

where

$$\begin{aligned} \bar{\mathbf{x}}_{\tau} &= \frac{1}{\tau} \sum_{i=1}^{\tau} \mathbf{x}_i, & \bar{\mathbf{x}}_{n-\tau} &= \frac{1}{n-\tau} \sum_{i=\tau+1}^n \mathbf{x}_i, \\ \mathbf{S}_{\tau} &= \frac{1}{\tau} \sum_{i=1}^{\tau} (\mathbf{x}_i - \bar{\mathbf{x}}_{\tau}) (\mathbf{x}_i - \bar{\mathbf{x}}_{\tau})^t, & \mathbf{S}_{n-\tau} &= \frac{1}{n-\tau} \sum_{i=\tau+1}^n (\mathbf{x}_i - \bar{\mathbf{x}}_{n-\tau}) (\mathbf{x}_i - \bar{\mathbf{x}}_{n-\tau})^t, \end{aligned}$$

and $\text{tr}(\cdot)$ is the trace-matrix operator.

For instance, maximum likelihood estimates for τ, μ_1 and μ_2 could be derived from (E.3). In the Bayesian perspective, a joint prior distribution is assumed for the unknown parameters $(\mu_1, \mu_2, \tau, \mathbf{P})$. For fixed τ , expression (E.3) has the structure of a product of two multivariate normal distributions with a Wishart distribution. Assuming conjugacy suggests therefore to represent prior knowledge about μ_1, μ_2 and \mathbf{P} by:

$$\begin{aligned} p(\mu_1, \mu_2, \mathbf{P}) &= p(\mu_1 \mid \mathbf{P}) p(\mu_2 \mid \mathbf{P}) p(\mathbf{P}) \quad (\text{E.4}) \\ &= \mathcal{N}_p(\mu_1 \mid \phi_1, \lambda_1 \mathbf{P}) \mathcal{N}_p(\mu_2 \mid \phi_2, \lambda_2 \mathbf{P}) \mathcal{W}_p(\mathbf{P} \mid a, \mathbf{B}) \\ &= \mathcal{N} \mathcal{N} \mathcal{W}_p(\mu_1, \mu_2, \mathbf{P} \mid \phi_1, \lambda_1, \phi_2, \lambda_2, a, \mathbf{B}) \end{aligned}$$

where $\mathcal{W}_p(\mathbf{P} \mid a, \mathbf{B})$ stands for the Wishart p.d.f. of dimension p with parameters a and \mathbf{B} ($2a > p - 1$ and \mathbf{B} is a $p \times p$ symmetric, non-singular matrix):

$$\mathcal{W}_p(\mathbf{P} \mid a, \mathbf{B}) = \frac{\pi^{-p(p-1)/4} |\mathbf{B}|^a}{\prod_{k=1}^p \Gamma\left[\frac{1}{2}(2a+1-k)\right]} |\mathbf{P}|^{a-(p+1)/2} \exp\{-\text{tr}(\mathbf{B}\mathbf{P})\}. \quad (\text{E.5})$$

This distribution is the multivariate generalization of the gamma distribution (Bernardo and Smith, 1994, pp.138-139). If $p = 1$, so that \mathbf{B} is a scalar, then $\mathcal{W}_p(\mathbf{P} \mid a, \mathbf{B})$ reduces to the gamma density.

The joint distribution (E.4) is sufficiently rich to enable a wide variety of prior opinions to be represented. Selecting particular values for the hyperparameters $(\phi_1, \lambda_1, \phi_2, \lambda_2, a, \mathbf{B})$ leads to a different joint prior distribution and allows to take into account the prior beliefs about the studied phenomenon. The hyperparameter values may come from historical or regional information, even from subjective knowledge. The precision of the prior knowledge

about μ_1, μ_2 and \mathbf{P} can be adjusted respectively through the values of λ_1, λ_2 and a . In fact, these hyperparameters can be interpreted as degrees of freedom, and small λ_1, λ_2 and a will represent less prior knowledge in a certain sense. When prior information is very limited, it may be desirable to let the prior knowledge for the unknown parameters be extremely vague. The joint prior distribution can be turned to a particular form of noninformative density by letting $p(\mu_1, \mu_2, \mathbf{P}) \propto |\mathbf{P}|^{-p/2}$, i.e. $\lambda_k \rightarrow 0$ and $a \rightarrow 0$ in (E.4). This improper p.d.f., i.e. specified only up to an undefined multiplicative constant, corresponds to the joint prior distribution considered by Booth and Smith (1982) for model (E.1).

Finally, assuming prior independence between τ and the other parameters $(\mu_1, \mu_2, \mathbf{P})$, and that $p(\tau)$ is any discrete distribution on the set $\{1, 2, \dots, n-1\}$, leads from (E.4) to

$$p(\mu_1, \mu_2, \mathbf{P}, \tau) = \mathcal{N}\mathcal{N}\mathcal{W}_p(\mu_1, \mu_2, \mathbf{P} \mid \phi_1, \lambda_1, \phi_2, \lambda_2, a, \mathbf{B}) p(\tau) \quad (\text{E.6})$$

This last assumption signifies that the date of change in the mean-vector is independent of the means before and after the shift. This hypothesis seems reasonable, and proves to be convenient for computations.

Inference about the unknown change-point

Since our main interest in this paper is the unknown change-point, we focus on the determination of the marginal posterior distribution of τ given the observations \mathbf{D} , $p(\tau \mid \mathbf{D})$. The reader who is interested in a complete analysis can find in Appendix A expressions for the marginal posterior p.d.f.'s of μ_1, μ_2 and \mathbf{P} . All results parallels the ones obtained in the univariate case (Perreault et al., 1999).

By Bayes's theorem (Berger, 1985), the joint posterior density $p(\mu_1, \mu_2, \mathbf{P} \mid \tau, \mathbf{D})$ for fixed τ is proportional to the product of the likelihood function (E.3) with the joint prior density (E.4):

$$p(\mu_1, \mu_2, \mathbf{P} \mid \tau, \mathbf{D}) \propto p(\mathbf{D} \mid \mu_1, \mu_2, \tau, \mathbf{P}) \mathcal{N}\mathcal{N}\mathcal{W}_p(\mu_1, \mu_2, \mathbf{P} \mid \phi_1, \lambda_1, \phi_2, \lambda_2, a, \mathbf{B}) \quad (\text{E.7})$$

The normalizing constant is the p.d.f. of the data \mathbf{D} only conditioned upon the change-point τ , $p(\mathbf{D} \mid \tau)$. After some algebraic manipulations (essentially completing squares as in the standard normal case, e.g. see Gelman et al. (1995, p.80-81)), the factors involving μ_1, μ_2 and \mathbf{P} in this last expression are clearly recognizable, up to the normalizing constant $p(\mathbf{D} \mid \tau)$, as belonging to a normal-Wishart distribution with updated parameters

$$\begin{aligned} \lambda'_1 &= \lambda_1 + \tau, & \lambda'_2 &= \lambda_2 + n - \tau \\ \phi'_1 &= \left(1 - \frac{\tau}{\lambda'_1}\right) \phi_1 + \frac{\tau}{\lambda'_1} \bar{\mathbf{x}}_\tau, & \phi'_2 &= \left(1 - \frac{n-\tau}{\lambda'_2}\right) \phi_2 + \frac{n-\tau}{\lambda'_2} \bar{\mathbf{x}}_{n-\tau}, \\ \mathbf{B}' &= \tau \left[\mathbf{S}_\tau + \left(1 - \frac{\tau}{\lambda'_1}\right) (\bar{\mathbf{x}}_\tau - \phi_1) (\bar{\mathbf{x}}_\tau - \phi_1)^t \right] \\ &\quad + (n-\tau) \left[\mathbf{S}_{n-\tau} + \left(1 - \frac{n-\tau}{\lambda'_2}\right) (\bar{\mathbf{x}}_{n-\tau} - \phi_2) (\bar{\mathbf{x}}_{n-\tau} - \phi_2)^t \right] + \mathbf{B}, \\ a' &= a + n/2. \end{aligned}$$

where, to simplify notation, dependence upon τ and \mathbf{D} was omitted by writing for example λ'_1 instead of $\lambda'_1(\mathbf{D}, \tau)$. Therefore,

$$p(\mu_1, \mu_2, \mathbf{P} \mid \tau, \mathbf{D}) = \mathcal{N}\mathcal{N}\mathcal{W}_p(\mu_1, \mu_2, \mathbf{P} \mid \phi'_1, \lambda'_1, \phi'_2, \lambda'_2, a', \mathbf{B}') \quad (\text{E.8})$$

The joint posterior distribution is also a normal-Wishart density because of the conjugate properties. The updated parameters of this posterior distribution combine the prior knowledge and the information conveyed by the data. For example, the posterior expected value

of μ_1, ϕ_1' , is a weighted average of the prior mean ϕ_1 and the sample mean before the change $\bar{\mathbf{x}}_\tau$, with weights determined by the relative precision of the two types of information.

Now, according to the Bayes theorem, the marginal posterior distribution for the change-point writes

$$p(\tau | \mathbf{D}) = \frac{p(\mathbf{D} | \tau) p(\tau)}{\sum_{\tau=1}^{n-1} p(\mathbf{D} | \tau) p(\tau)} \quad (\text{E.9})$$

where $p(\mathbf{D} | \tau)$ is the normalizing constant which makes the Bayes theorem denominator omitted in expression (E.7). Evaluation of this denominator usually involves numerical integration, but since conjugate priors were used it can be easily determined by dividing (E.7) by (E.8), and cancelling factors involving μ_1, μ_2 and \mathbf{P} :

$$p(\mathbf{D} | \tau) = (2\pi)^{-\frac{np}{2}} \left(\frac{\lambda_1 \lambda_2}{\lambda_1' \lambda_2'} \right)^{p/2} \prod_{k=1}^p \left\{ \frac{\Gamma[\frac{1}{2}(2a' + 1 - k)]}{\Gamma[\frac{1}{2}(2a + 1 - k)]} \right\} \frac{|\mathbf{B}|^a}{|\mathbf{B}'|^{a'}}. \quad (\text{E.10})$$

Therefore, from (E.9) and (E.10) the marginal posterior density of the change-point $\tau = 1, 2, \dots, n-1$ is seen to be proportional to

$$p(\tau | \mathbf{D}) \propto p(\tau) (\lambda_1' \lambda_2')^{-p/2} |\mathbf{B}'|^{-a'} \quad (\text{E.11})$$

This distribution is discrete and gives, for each time point, the posterior probability of shift occurrence in the mean-vector, assuming a change occurred with certainty. The whole posterior distribution for the global change-point is therefore available to evaluate point estimates for τ (e.g. posterior mean, median or mode) and any quantity that can characterize the uncertainty about the global change-point (e.g. posterior variance, credible intervals, quantiles, etc.). Observe that no asymptotic assumption was used in the evaluation of the posterior distribution $p(\tau | \mathbf{D})$ and those given in Appendix A. Therefore, the approach is valid for limited sample sizes, which is often the case in hydrology.

E.3 Detection of a change-point

Up to this point, a change in the mean vector has been assumed with certainty. Therefore, the no change hypothesis has no status in that formulation and the posterior distributions derived in the previous section and those in Appendix A can not be used to infer about the existence of a shift. Let us now consider the problem involving the verification of the existence of a change, which is essentially a problem of testing the statistical hypotheses « change » against « no change ». This is equivalent to compare the alternatives models:

$$\begin{aligned} \text{No change, model } H_0: & \quad \mathbf{X}_i \sim \mathcal{N}_p(\mathbf{x}_i | \mu_1, \mathbf{P}), \quad i = 1, \dots, n \\ \text{Change, model } H_1: & \quad \begin{cases} \mathbf{X}_i \sim \mathcal{N}_p(\mathbf{x}_i | \mu_1, \mathbf{P}), & i = 1, \dots, \tau \\ \mathbf{X}_i \sim \mathcal{N}_p(\mathbf{x}_i | \mu_2, \mathbf{P}), & i = \tau + 1, \dots, n \end{cases} \end{aligned}$$

for data \mathbf{D} . The main tool to perform such an analysis is the Bayes factor B_{10} defined as the ratio of posterior to prior odds on H_1 against H_0 ,

$$\begin{aligned}
B_{10}(\mathbf{D}) &= \frac{p(H_1|\mathbf{D})}{p(H_0|\mathbf{D})} \bigg/ \frac{p(H_1)}{p(H_0)} = \frac{p(\mathbf{D}|H_1)}{p(\mathbf{D}|H_0)} \\
&= \frac{\int p(\mathbf{D} | \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \tau, \mathbf{P}) p(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{P}) p(\tau)}{\int p(\mathbf{D} | \boldsymbol{\mu}_1, \mathbf{P}) p(\boldsymbol{\mu}_1, \mathbf{P})}
\end{aligned} \tag{E.12}$$

The p.d.f.'s $p(\mathbf{D} | H_k)$, $k = 0, 1$, appearing in (E.12) are obtained by integrating all parameters out. Density $p(\mathbf{D} | H_k)$ is often called the marginal distribution of \mathbf{D} given model H_k , but in the Bayesian paradigm a more informative name is the prior predictive distribution. That is, this p.d.f. is the probability of seeing the data that were actually observed assuming model H_k , calculated before any data became available. This predictive density, implied by the likelihood and the prior distribution, provides a basis for assessing the compatibility of the data with our prior state of belief, i.e. model H_k . Therefore, the Bayes factor can be interpreted as a measure of whether the data \mathbf{D} have increased or decreased the odds on the change situation relative to no change. The Bayes factor is similar enough to a likelihood ratio statistic used in the classical framework, but the parameters are eliminated by integration rather than by maximization.

If we assume conjugate prior distributions for the parameters $(\boldsymbol{\mu}_1, \mathbf{P})$ of model H_0 , i.e. a normal-Wishart $\mathcal{NW}_p(\boldsymbol{\mu}_1, \mathbf{P} | \boldsymbol{\phi}_1, \lambda_1, a, \mathbf{B})$, it is shown in Appendix B that the Bayes factor for the assessment of change against no change is given by:

$$B_{10}(\mathbf{D}) = \sum_{\tau=1}^{n-1} p(\tau) \left(\frac{(\lambda_1 + n) \lambda_2}{(\lambda_1 + \tau)(\lambda_2 + n - \tau)} \right)^{p/2} \left(\frac{|\mathbf{B}'_0|}{|\mathbf{B}'|} \right)^{a+n/2}. \tag{E.13}$$

where \mathbf{B}'_0 , the updated parameter \mathbf{B} under model H_0 , reads

$$\mathbf{B}'_0 = \frac{n}{2} \left[\mathbf{S}_n + \left(1 - \frac{n}{(\lambda_1 + n)} \right) (\bar{\mathbf{x}}_n - \boldsymbol{\phi}_1) (\bar{\mathbf{x}}_n - \boldsymbol{\phi}_1)^t \right] + \mathbf{B}.$$

The decision rule which defines the Bayes test to choose among the two alternatives is to prefer the change situation to the no change hypothesis if and only if

$$B_{10}(\mathbf{D}) > \frac{l_{10} p(H_0)}{l_{01} p(H_1)}, \tag{E.14}$$

where l_{kj} is the loss resulting in choosing H_k when the true situation is in fact H_j , and no losses for correct conclusions are assumed ($l_{kk} = l_{jj} = 0$). The critical value of the Bayes test takes into account the prior probabilities of each alternatives, and the relative seriousness of the two possible ways of selecting the wrong hypothesis. If one wishes not to favor any of the two hypotheses in the prior state of belief, and if the losses in choosing the wrong situation are considered to be equivalent, the change situation is to be selected when $B_{10}(\mathbf{D}) > 1$.

It is important to recall that the use of Bayes factor (E.13) needs informative proper prior distributions. In fact, noninformative p.d.f.'s depend on unspecified constants, and since the alternatives models do not have the same number of parameters (one mean-vector for model H_0 versus two different mean-vectors under H_1), all constants are not eliminated when evaluating (E.13) in the case of such improper priors. For example, if we consider Booth and Smith's noninformative priors, i.e. $p(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{P}) \propto |\mathbf{P}|^{-p/2}$ or equivalently

$$\begin{aligned} \text{Model } H_0: \quad p(\mu_1 | \mathbf{P}) &= c_1 |\mathbf{P}|^{1/2} \\ p(\mathbf{P}) &= c |\mathbf{P}|^{-(p+1)/2} \end{aligned}$$

$$\begin{aligned} \text{Model } H_1: \quad p(\mu_1 | \mathbf{P}) &= c_1 |\mathbf{P}|^{1/2} \\ p(\mu_2 | \mathbf{P}) &= c_2 |\mathbf{P}|^{1/2} \\ p(\mathbf{P}) &= c |\mathbf{P}|^{-(p+1)/2} \end{aligned}$$

the Bayes factor for change against no change becomes (Booth and Smith, 1982)

$$B_{10}(\mathbf{D}) = c_2 \sum_{\tau=1}^{n-1} p(\tau) \left(\frac{n}{\tau(n-\tau)} \right)^{p/2} \left(\frac{|n\mathbf{S}_n|}{|\tau\mathbf{S}_\tau + (n-\tau)\mathbf{S}_{n-\tau}|} \right)^{n/2} \quad (\text{E.15})$$

and depends upon the unspecified constant c_2 . Arbitrariness necessarily enters into such a Bayes factor. In Perreault et al. (1999), where a similar Bayes test was considered for the univariate change-point problem, diffuse but proper priors were used to illustrate their approach. While, in their particular application, subsequent analyses with more informative prior distributions did not contradict their conclusions, the authors want to warn the reader against the use of noninformative or extremely vague priors in evaluating Bayes factors. This problem can be overcome by using default Bayes factors, such as the « Intrinsic Bayes factor » (Berger and Pericchi, 1996) and the « Fractional Bayes factor » (O'Hagan, 1995). In this paper, we do not address Bayes tests with noninformative priors since we assume that some prior hydrologic expertise is available about the mean-vectors and the precision matrix, however vague it may be. More details about Bayes tests can be found in Berger (1985), Bernardo and Smith (1994) and Kass and Raftery (1995).

E.4 Application

A number of rivers in the Northern Québec Labrador region, north of the 50th parallel, show a marked decrease in average streamflow since the early eighties, which is mostly manifest in spring (January to June runoff), while the annual series of the average streamflow from July to December looks stationary for the whole period. Figure 1 shows the spring average streamflow series expressed in $\text{liter}/(\text{km}^2 \times \text{s})$ for six rivers situated in this region. A brief description of each river is given in Table 1. Figure 2 shows their location.

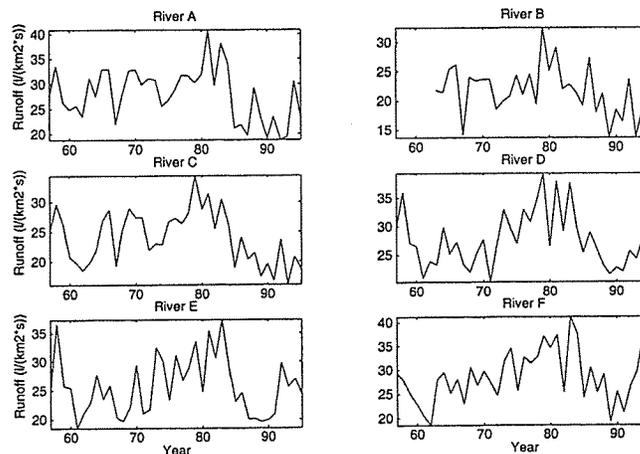


Figure 1. Annual average runoff from January to June for the six rivers.

Table 1. Description of each river

Site	River	Basin area (km^2)	Period	Average runoff (m^3/s)	Average runoff ($l/(km^2 \times s)$)
A	Romaine	13 000	1957-1995	364	28.2
B	À la Baleine	29 800	1963-1995	641	21.5
C	Churchill Falls	69 344	1957-1995	1 668	24.1
D	Manicouagan	45 480	1957-1995	1 252	27.6
E	Outardes	18 798	1957-1995	482	25.8
F	Sainte-Marguerite	4 737	1957-1995	137	29.6



Figure 2. Location of the six rivers.

While it seems possible that an abrupt change in the mean level may have occurred in the mid-eighties for each sequence, we can not evaluate the uncertainty of such a visual estimate. Moreover, as mentioned in Section 1, application of six independent univariate analyses do not furnish any global estimate, neither any measure of uncertainty about a regional change-point. Finally, since significant spatial correlations between these series can be anticipated, a multivariate change-point analysis is certainly more suitable for such a problem than independent univariate analyses.

A first step in that direction is to consider model (E.1), i.e. a simultaneous single change in the means occurring at an unknown time point in a sequence of independent normal random vectors. The normality assumption is appropriate for the average streamflow as well as for a large range of other hydrological series. Since the data in Figure 1 are obtained from at least a daily summation over time, the central limit theorem can be invoked. From the hydrological point of view however, the independence assumption between successive values may be more questionable. It is made here for convenience as a first methodological step, and can be relaxed eventually by considering, for example, a shift in the mean-vector of a multivariate autoregressive model. But still, for the rivers considered herein, no information (e.g. ground water storage and input, control structures) indicates possible interannual correlations.

In what follows, the multivariate change-point analysis is applied to the six series for the common period 1963-1993. The remaining available six observations for five of the six rivers (1957-1962) are considered as a prior sample and used to elicit the hyperparameters.

E.4.1 Specifying prior distributions

Prior knowledge about the parameters of model (E.1) must be elicited from expert opinion and/or from external information about the studied phenomena in order to specify the prior distributions. Obtaining the necessary information can be difficult, but it is worth it since it allows for taking full advantage of the Bayesian framework, i.e. including all types of knowledge in the analysis in a formal way. In what follows, we briefly describe how prior distributions were specified for our particular problem. As mentioned in Section 4, we considered informative prior distributions for μ_1, μ_2 and \mathbf{P} .

Experts did not agree about the possibility that a climatic change took place in the last four decades, and had an impact on water resources in Northern Québec. Therefore, prior probabilities for each hypothesis (« change » and « no change ») were given equal weight, i.e. $p(H_1) = p(H_0) = 1/2$. Moreover, the ones who suspected the existence of a sudden change in those series were not willing to specify their degree of belief about the year at which it occurred without referring to the data itself. This information was therefore biased. In fact, the use of any of these data in the prior formulation is not formally acceptable in a Bayesian analysis (Berger, 1985). Then, we could not reasonably favor any year of change, and τ was assumed to be distributed as an uniform discrete distribution for model H_1 . Note that with this density, the prior expected change-point and standard deviation were respectively 1978 and 9.2 years (the mean and standard deviation of a discrete uniform probability distribution on the interval [1963,1995]).

For the other parameters, the prior degrees of belief were assumed to be represented by normal-Wishart type of distributions (equations (E.4) and (B.3)). The complete specification of prior knowledge about these quantities required the choice of hyperparameters in these expressions, i.e. $(\phi_1, \lambda_1, \phi_2, \lambda_2, a, \mathbf{B})$. They were elicited by using the six prior observations (1957-1962). The mean-vector and the corresponding covariance matrix for the prior data set are

$$\bar{y} = \begin{pmatrix} 26.9 \\ \text{na} \\ 23.3 \\ 27.4 \\ 25.4 \\ 24.0 \end{pmatrix} \quad \text{and} \quad \mathbf{S}_y^2 = \begin{pmatrix} 12.40 & \text{na} & 13.59 & 16.11 & 19.40 & 10.76 \\ & \text{na} & \text{na} & \text{na} & \text{na} & \text{na} \\ & & 18.74 & 20.06 & 23.31 & 15.56 \\ & & & 27.00 & 30.83 & 17.82 \\ & & & & 38.42 & 17.67 \\ & & & & & 17.42 \end{pmatrix}$$

showing missing prior information since data for river B was not available for this period (see Table 1). The missing mean and variance were estimated by averaging the corresponding available values. This led us to a mean and a variance for river B respectively given by $25.4 \text{ l}/(\text{km}^2 \times \text{s})$ and 22.75. The remaining missing prior information, i.e. the covariances, were set to zero because such statistical characteristics are very difficult to specify. The resulting « reconstituted » prior mean-vector and covariance matrix are denoted $\hat{\bar{y}}$ and $\hat{\mathbf{S}}_y^2$.

The prior information about $(\mu_1, \mu_2, \mathbf{P})$ consisted therefore of a sample of streamflow data from 1957 to 1962, which a portion had been estimated. It was felt these data behaved like the runoffs before the change in the mean-vector had occurred. Since we have a priori $E\{\mu_1\} = \phi_1$, and that $\hat{\bar{y}}$ is a natural estimate for this expected value, we took $\phi_1 = \hat{\bar{y}}$. In a similar way, since $E\{\mathbf{P}\} = a\mathbf{B}^{-1}$ (Bernardo and Smith, 1994), we assigned $\mathbf{B} = a\hat{\mathbf{S}}_y^2$. As mentioned in Section 3, the hyperparameters λ_1 and a can be interpreted as degrees of freedom. Because an almost complete prior sample of six observations was available, we considered $\lambda_1 = a = 6$. That is, we thought that our prior information about μ_1 and \mathbf{P} was equivalent as if we had 6 years of complete measured historical data. Finally, prior information about the mean-vector after the possible change was felt to be quite vague. Moreover, we were not very confident about the information conveyed by the prior sample about the mean-vector after the change. Thus, the prior degree of freedom corresponding to μ_2 was taken to be $\lambda_2 = 1$, and the prior expected value after the shift was taken to be the same as that before the change, i.e. $\phi_2 = \hat{\bar{y}}$.

E.4.2 Results and discussion

Figure 3 displays the marginal posterior p.d.f. of the global change-point τ obtained with the hyperparameters specified in the previous section. The corresponding prior density (the continuous line) is also plotted. We first observe how the data modified or updated prior information by comparing visually the prior and the posterior densities. Such a graph is especially a useful tool to interpret the uncertainty about the change-point. Table 2 summarizes estimated posterior quantities (mode, median, mean, standard deviation and 90% credible interval). The mean and the standard deviation can be compared to their corresponding prior values, respectively 1978 and 9.2 (Section 5.1).

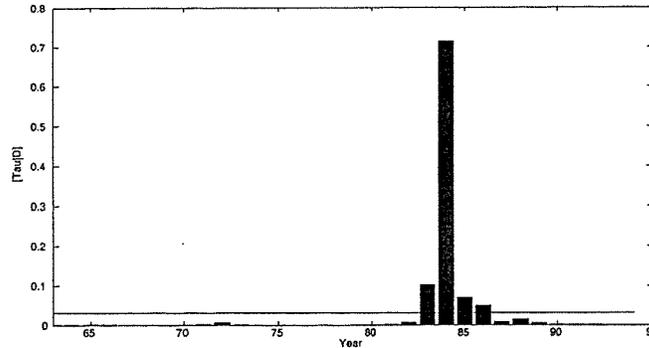


Fig. 3. Marginal posterior probability density function for the change-point, $p(\tau | \mathbf{D})$.

Table 2. Posterior characteristics for $p(\tau | \mathbf{D})$

Mode	Median	Mean	St. dev.	90% Cred. Int.
1984	1983.5	1983.9	2.33	[1982,1985]

This analysis indicates, under the hypothesis of a change in the mean-vector of streamflows, that a regional change-point occurred around 1984 with approximately a 2.5-year standard deviation. The corresponding Bayes factor for the assessment of overall change against no change, $B_{10} = 184.12$, indicates that the data have considerably increased the odds on the change hypothesis relative to the no change situation. If for example losses in choosing the wrong model were considered to be equivalent, this result would lead us to a very significant Bayesian decision in favor of an abrupt shift in the mean vector of the runoffs. It can therefore be interesting to go further in the analysis by evaluating posterior estimates for the other parameters, namely the mean-vectors before and after the change. Using expressions given in Appendix A, posterior mean and standard deviation for μ_1 and μ_2 were computed. These are given in Table 3.

Table 3. Prior expected values and posterior quantities for μ_1 and μ_2

		Rivers					
		A	B	C	D	E	F
μ_1	Prior exp. val. ϕ_1	26.9	25.4	23.3	27.4	25.4	24.0
	Posterior mean	29.9	23.5	25.6	28.6	26.7	29.1
	Posterior stand. dev.	(0.92)	(0.90)	(0.85)	(1.02)	(1.10)	(1.18)
μ_2	Prior exp. val. ϕ_2	26.9	25.4	23.3	27.4	25.4	24.0
	Posterior mean	23.3	19.1	20.2	25.0	23.5	27.2
	Posterior stand. dev.	(1.50)	(1.42)	(1.34)	(1.56)	(1.67)	(1.72)

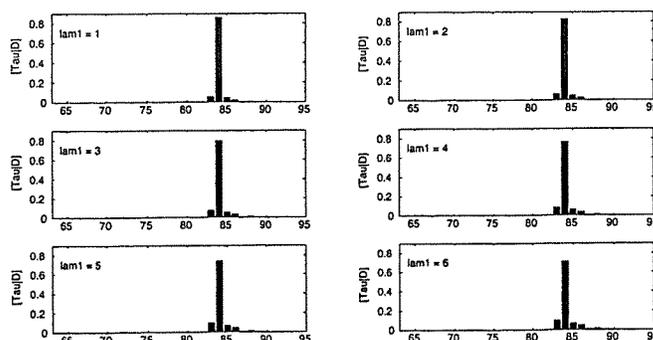
Our approach assumes constant precision matrix, or equivalently constant correlation matrix, on both sides of τ . The correlations are given in Table 4 separately for the period before and after the estimated change-point 1984.

Table 4. Between sites correlations.

	1963-1983						1984-1995					
	A	B	C	D	E	F	A	B	C	D	E	F
A	1	0.52	0.61	0.47	0.51	0.50	1	0.12	0.43	0.39	0.26	0.63
B		1	0.80	0.35	0.31	0.44		1	0.83	0.64	0.36	0.30
C			1	0.55	0.48	0.58			1	0.77	0.55	0.57
D				1	0.92	0.76				1	0.68	0.75
E					1	0.77					1	0.67
F						1						1

Even if the two correlation matrix seem globally similar, one could still questioned this hypothesis. Such an uncertainty could be taken into account by considering a model which makes the precision matrix change simultaneously with the mean-vector, as it was done by Perreault et al. (2000b) for an univariate change in both the mean and the variance. However, since the sampling error on estimated correlation coefficients is generally high for such a limited sample size, the assumption of equal precision matrix was considered acceptable to us for this particular case study.

Finally, one could doubt about the accuracy of the prior sample to well represent our prior state of belief. Thus, an important issue is the sensitivity of these results to the choices of hyperparameter values. We thought it would be interesting to evaluate sensitivity of the above analysis when confidence on the information conveyed by the prior observations decreases. All posterior densities, and the Bayes factor B_{10} , were therefore recomputed with $\lambda_1 = \{6, \dots, 1\}$ and $a = \{6, \dots, 3\}$, that is for increasing uncertainty about the elicited expected values of μ_1 and \mathbf{P} (by definition of a Wishart distribution, $2a > p - 1 \iff a > 2.5$). This sensitivity analysis showed that the posterior distributions and especially posterior estimates are quite stable. As an example, $p(\tau | \mathbf{D})$, $\lambda_1 = 1, \dots, 6$, are plotted in Figure 4 for $a = 6$ (very confident about the elicited parameters for $p(\mathbf{P})$ from the prior sample), and in Figure 5 for $a = 3$ (less confident). Apart for a slight increase in the variability when the degree of freedom λ_1 increases, the posterior densities are similar and posterior quantities do not differ significantly.

Figure 4. Sensitivity analysis: marginal posterior densities $p(\tau | \mathbf{D})$ for $a = 6$ and $\lambda_1 = 1, \dots, 6$.

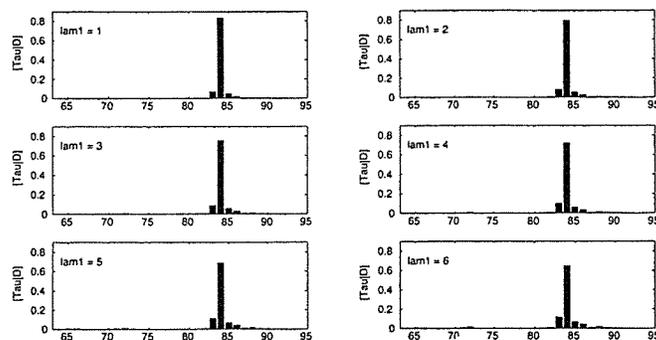


Figure 5. Sensitivity analysis: marginal posterior densities $p(\tau | \mathbf{D})$ for $a = 3$ and $\lambda_1 = 1, \dots, 6$.

As it could be anticipated (Kass and Raftery, 1995), in contrast with Bayesian point estimate such as the posterior mode, mean and median, the Bayes factor B_{10} does tend to be sensitive to the choices of priors on the model parameters. Figure 6 presents the computed Bayes factors for all situations considered.

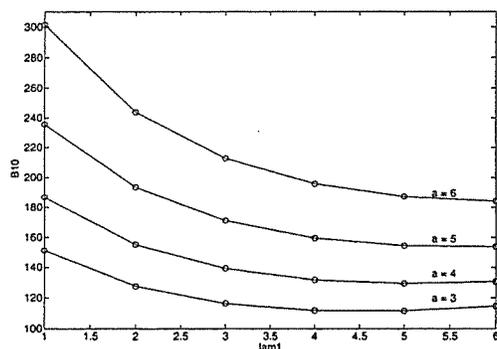


Figure 6. Sensitivity analysis: Bayes factors B_{10} computed for $a = 3, \dots, 6$ and $\lambda_1 = 1, \dots, 6$.

It is seen that when the degrees of freedom involved in the prior distribution of μ_1 increases, i.e. when we are more confident about the elicited expected value for μ_1 , strength in favor of the change hypothesis declines. On the other hand, the Bayes factor increases when a gets larger. The minimum value obtained over all priors considered is $B_{10} = 111.54$, and still show considerable evidence of an abrupt change in the mean-vector for these six rivers situated in the Northern Québec Labrador region.

E.5 Conclusion

As mentioned in Section 1, the current literature deals with detection of change-points only in the context of the observation of a single sequence. Few have addressed change-point analysis for more than one sequence even though this problem is of considerable practical importance in hydrology and climatology (regional analysis, multi-reservoirs management, global warming, analysis with several climate change indices, etc.). This paper has provided a Bayesian approach based on the multivariate normal model, (1) for the estimation of a regional unknown date of change in the means of several hydrological sequences, which take into account correlations between sites, and (2) for the overall assessment of change versus no change. The paper has also provided expressions to characterize the shift, namely Bayesian estimates for the mean-vector before and after the change. It was showed that by using conjugate prior distributions, all expressions are in closed form, and inference can be

made straightforwardly. The Bayesian method presented in this paper, can be viewed as an extension of the multivariate normal model, and practitioners can perform such change-point analysis routinely using standard statistical toolboxes. Note that even if nonconjugate priors were used, this procedure could still be performed without major technical difficulties. Indeed, no explicit solutions would be available, but this problem could be efficiently solved by using Markov Chain Monte-Carlo simulation schemes which are easy to implement (see Gelfand and Smith (1990), Robert and Casella (1999) and Kuzera and Parent (1998) for general discussions and applications, and Perreault et al. (2000a,b) for applications to change-point analysis in hydrology).

This multi-site Bayesian approach was illustrated by its application on a set of spring average streamflow for six rivers situated in the Northern Québec Labrador region. A sensitivity analysis was performed to investigate how prior distributions have an impact on the inference process. It was seen that while the estimation of the parameters are stable, the Bayes factor tends to be sensitive to the choice of priors on the model parameters. However, the minimum Bayes factor obtained for this case study still showed strong evidence in favor of an overall decrease in average runoff in the early eighties.

Sustained work to improve the approach in several directions is needed. One important issue is to take into account the fact that the position of the change-point in hydrological time series may vary from sequence to sequence according to a structural relationship. In fact, nothing guaranty that climate change has exactly the same effect on each sequence, and a delay in the occurrence of a shift may exists from site to site. Constructing such a model with spatially correlated data is not obvious, but is certainly more realistic than the strict assumption of a common change-point. The authors are now working in that direction, and will eventually report on this problem.

Acknowledgments. The authors are highly grateful to Vincent Fortin of IREQ for providing the data. This research was supported by la Direction Générale de l'Enseignement et de la Recherche de France of the French ministry of Agriculture (DGER), les fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec (FCAR), la Chaire en Hydrologie Statistique Hydro-Québec/CRSNG (INRS-Eau) and le Ministère des Relations internationales du Québec (MRI).

References

- Barnett TP, Hasselmann K, Chelliah M, Delworth T, Hegerl G, Jones P, Rasmusson E, Roeckner E, Ropelewski C, Santer B, Tett S (1999) Detection and attribution of recent climate change: A status report. *Bulletin of the American Meteorological Society* 80: 2631-2659.
- Berger JO (1985) *Statistical decision theory and Bayesian analysis*. New York: Springer-Verlag
- Berger JO, Pericchi LR (1996) The intrinsic Bayes factor for model selection and prediction. *J. Amer. Statist. Assoc.* 91: 109-122
- Bernardo JM, Smith AFM (1994) *Bayesian theory*. New York: Wiley
- Booth NB, Smith AFM (1982) A Bayesian approach to retrospective identification of change-points. *Jour. of Econometrics* 19: 7-22
- Bruneau P, Rassam J-C (1983) Application d'un modèle bayésien de détection de changements de moyennes dans une série. *Journal des Sciences Hydrologiques* 28: 341-354
- Gelfand AE, Smith AFM (1990) Sampling-based approaches to calculating marginal densities. *J. Am. Stat. Assoc.* 85: 398-409
- Gelman A, Carlin JB, Stern HS, Rubin DB (1995) *Bayesian Data Analysis*. New York: Chapman & Hall
- Hubert P, Carbonnel JP, Chaouche A (1989) Segmentation des séries hydrométéorologiques - Applications à des séries de précipitations et de débits de l'Afrique de l'Ouest. *J. of Hydrol.* 110: 349-367

- Kass RE, Raftery AE (1995) Bayes factors. *J. Am. Stat. Assoc.* 90 : 773-795
- Kiely G (1999) Climate change in Ireland from precipitation and streamflow observations. *Advances in Water Resources* 23 : 141-151.
- Kuczera G, Parent E (1998) Monte Carlo assessment of parameter uncertainty in conceptual catchment models : the Metropolis algorithm. *J. of Hydrol.* 211 : 69-85.
- Lubès-Niel H, Masson JM, Paturel JE, Servat E (1998) Variabilité climatique et statistiques. Etude par simulation de la puissance et de la robustesse de quelques tests utilisés pour vérifier l'homogénéité de chroniques. *Rev. Sci. Eau* 3 : 383-408
- O'Hagan A (1995) Fractional Bayes factors for model comparisons. *J. Roy. Statist. Soc. B* 57 : 99-138
- Paturel JE, Servat E, Kouamé B, Lubès H, Ouedraogo M, Masson JM (1997) Climatic variability in humid Africa along the Gulf of Guinea Part II: an integrated regional approach. *J. of Hydrol.* 191 : 16-36.
- Perreault L, Haché M, Slivitsky M, Bobée B (1999) Detection of changes in precipitation and runoff over eastern Canada and U.S. using a Bayesian approach. *Stochastic Environmental Research and Risk Assessment* 13 : 201-216
- Perreault L, Bernier J, Bobée B, Parent E (2000a) Bayesian change-point analysis in hydrometeorological time series 1: The normal model revisited. Submitted to *J. of Hydrol.*
- Perreault L, Bernier J, Bobée B, Parent E (2000b) Bayesian change-point analysis in hydrometeorological time series 2: Comparison of change-point models and forecasting. Submitted to *J. of Hydrol.*
- Robert CP, Casella, G (1999) *Monte Carlo Statistical Methods*. New York: Springer-Verlag.
- Servat E, Paturel JE, Lubès H, Kouamé B, Ouedraogo M, Masson JM (1997) Climatic variability in humid Africa along the Gulf of Guinea Part I: detailed analysis of the phenomenon in Côte d'Ivoire. *J. of Hydrol.* 191 : 1-15.

Appendix

A. Inference about μ_1 , μ_2 and P

Given τ and \mathbf{D} , it is straightforward to show that μ_1 and μ_2 are conditionally distributed as p -variate Student t -distributions, and \mathbf{P} as a Wishart distribution. Integrating the appropriate parameters out of expression (E.8), leads directly to these conditional posterior distributions. More precisely, we have

$$\begin{aligned} p(\mu_k | \tau, \mathbf{D}) &= St_p(\mu_k | \mathbf{m}_k, \mathbf{A}_k, c) \quad k = 1, 2 \\ p(\mathbf{P} | \tau, \mathbf{D}) &= \mathcal{W}_p(\mathbf{P} | a', \mathbf{B}') \end{aligned} \quad (\text{A3A1})$$

where $\mathbf{m}_k = \phi'_k, \mathbf{A}_k = (2a' - p + 1) \lambda'_k (2\mathbf{B}')^{-1}, c = 2a' - p + 1$, and

$$St_p(\mathbf{x} | \mathbf{m}, \mathbf{A}, c) \propto |\mathbf{A}|^{1/2} \left[1 + \frac{1}{c} (\mathbf{x} - \mathbf{m})^t \mathbf{A} (\mathbf{x} - \mathbf{m}) \right],$$

$(\phi'_1, \lambda'_1, \phi'_2, \lambda'_2, a', \mathbf{B}')$ being the updated hyperparameters given in Section 2. To draw conclusions regarding these parameters, their unconditional marginal posterior distributions must be evaluated. Since $p(\mu_k | \tau, \mathbf{D}) = p(\mu_k, \tau | \mathbf{D}) / p(\tau | \mathbf{D})$, this is done by averaging the corresponding conditional posterior distribution with respect to the posterior mass function of τ (E.11):

$$p(\mu_k | \mathbf{D}) = \sum_{\tau=1}^{n-1} p(\mu_k | \tau, \mathbf{D}) p(\tau | \mathbf{D}), \quad k = 1, 2, \quad (\text{A3A.2})$$

which are finite mixtures of p -variate Student t -distributions. Note that the moments of these posterior distributions are easily derived from the moments of the conditional distributions. For instance,

$$\begin{aligned} E\{\mu_i | \mathbf{D}\} &= \sum_{\tau=1}^{n-1} E\{\mu_i | \tau, \mathbf{D}\} p(\tau | \mathbf{D}) \\ &= \sum_{\tau=1}^{n-1} p(\tau | \mathbf{D}) \phi'_i \end{aligned} \quad (\text{A3A.3})$$

and this expression can be used as an estimate for the mean-vector before and after the change. In the same way, the unconditional posterior p.d.f. of the precision matrix can be written as

$$p(\mathbf{P} | \mathbf{D}) = \sum_{\tau=1}^{n-1} p(\mathbf{P} | \tau, \mathbf{D}) p(\tau | \mathbf{D}) \quad (\text{A3A.4})$$

which is a finite mixture of Wishart distributions. Since for a Wishart distribution $\mathcal{W}_p(\mathbf{P} | a, \mathbf{B})$

$$E\{\mathbf{P}^{-1}\} = \left(a - \frac{p+1}{2}\right)^{-1} \mathbf{B},$$

then an estimate for the covariances matrix Σ , which take into account the shift in the mean-vector, is given by

$$E\{\Sigma | \mathbf{D}\} = \left(a' - \frac{p+1}{2}\right)^{-1} \sum_{\tau=1}^{n-1} p(\tau | \mathbf{D}) \mathbf{B}'. \quad (\text{A3A.5})$$

If a change in the mean level is suspected, this estimator will be a more efficient than the usual sample covariance matrix. For example, it could improve the results of a principal component analysis.

B. Bayes factor for change against no change

To perform the Bayes test, the densities $p(\mathbf{D} | H_1)$ and $p(\mathbf{D} | H_0)$ in (E.12) have to be evaluated. Assuming the normal-Wishart prior distribution (E.4), the first one can be deduced from (E.10):

$$\begin{aligned} p(\mathbf{D} | H_1) &= \sum_{\tau=1}^{n-1} p(\mathbf{D}, \tau | H_1) \\ &= \sum_{\tau=1}^{n-1} p(\mathbf{D} | \tau, H_1) p(\tau) \\ &= (2\pi)^{-\frac{np}{2}} (\lambda_1 \lambda_2) |\mathbf{B}|^a \prod_{k=1}^p \left\{ \frac{\Gamma\left[\frac{1}{2}(2a' + 1 - k)\right]}{\Gamma\left[\frac{1}{2}(2a + 1 - k)\right]} \right\} p(\tau) \sum_{\tau=1}^{n-1} (\lambda'_1 \lambda'_2)^{-p/2} |\mathbf{B}'|^{-a'} \end{aligned} \quad (\text{A3B.1})$$

The likelihood function resulting from a sample of size n , $\mathbf{D} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$, generated by model H_0 in Section 4 can be written as

$$\begin{aligned} p(\mathbf{D} | \mu_1, \mathbf{P}) &= \prod_{i=1}^n \mathcal{N}_p(\mathbf{x}_i | \mu_1, \mathbf{P}) \\ &= (2\pi)^{-np/2} |\mathbf{P}|^{n/2} \exp \left\{ -\frac{n}{2} \text{tr} \left[\mathbf{P} \left(\mathbf{S}_n + (\bar{\mathbf{x}}_n - \mu_1) (\bar{\mathbf{x}}_n - \mu_1)^t \right) \right] \right\} \end{aligned} \quad (\text{A3B.2})$$

Let us consider a normal-Wishart as the joint prior distribution

$$\begin{aligned} p(\mu_1, \mathbf{P}) &= \mathcal{N}_p(\mu_1 | \phi_0, \lambda_0 \mathbf{P}) \mathcal{W}_p(\mathbf{P} | a_0, \mathbf{B}_0) \\ &= \mathcal{NNW}_p(\mu_1, \mathbf{P} | \phi_0, \lambda_0, a_0, \mathbf{B}_0) \end{aligned} \quad (\text{A3B.3})$$

By Bayes's theorem, combining the likelihood function (A3B.2) with (A3B.3) results in the joint posterior density of (μ_1, \mathbf{P}) , which is also a normal-Wishart distribution but with updated hyperparameters:

$$\begin{aligned} \lambda'_0 &= \lambda_0 + n \\ \phi'_0 &= \left(1 - \frac{n}{\lambda'_0} \right) \phi_0 + \frac{n}{\lambda'_0} \bar{\mathbf{x}}_n \\ \mathbf{B}'_0 &= \frac{n}{2} \left[\mathbf{S}_n + \left(1 - \frac{n}{\lambda'_0} \right) (\bar{\mathbf{x}}_n - \phi_0) (\bar{\mathbf{x}}_n - \phi_0)^t \right] + \mathbf{B}_0 \\ a'_0 &= a_0 + n/2. \end{aligned}$$

The density $p(\mathbf{D} | H_0)$ is the normalizing constant of the Bayes theorem. It can therefore be expressed as

$$\begin{aligned} p(\mathbf{D} | H_0) &= \frac{p(\mathbf{D} | \mu_1, \mathbf{P}) p(\mu_1, \mathbf{P})}{p(\mu_1, \mathbf{P} | \mathbf{D})} \\ &= \frac{p(\mathbf{D} | \mu_1, \mathbf{P}) \mathcal{NNW}_p(\mu_1, \mathbf{P} | \phi_0, \lambda_0, a_0, \mathbf{B}_0)}{\mathcal{NNW}_p(\mu_1, \mathbf{P} | \phi'_0, \lambda'_0, a'_0, \mathbf{B}'_0)} \end{aligned} \quad (\text{A3B.4})$$

Cancelling factors involving μ_1 and \mathbf{P} , we obtain:

$$p(\mathbf{D} | H_0) = (2\pi)^{-\frac{np}{2}} \left(\frac{\lambda_0}{\lambda'_0} \right)^{p/2} \prod_{k=1}^p \left\{ \frac{\Gamma \left[\frac{1}{2} (2a'_0 + 1 - k) \right]}{\Gamma \left[\frac{1}{2} (2a_0 + 1 - k) \right]} \right\} \frac{|\mathbf{B}_0|^{a_0}}{|\mathbf{B}'_0|^{a'_0}}. \quad (\text{A3B.5})$$

Finally, dividing (A3B.1) by (A3B.5) and taking $\phi_0 = \phi_1, \lambda_0 = \lambda_1, a_0 = a, \mathbf{B}_0 = \mathbf{B}$ leads to (E.13).

