

École Nationale du Génie Rural
des Eaux et des Forêts (ENGREF)
Laboratoire GRESE

Université du Québec
INRS-Eau

**LA MODÉLISATION STOCHASTIQUE DES ÉTIAGES ET DE LEURS DURÉES
EN VUE DE L'ANALYSE DU RISQUE**

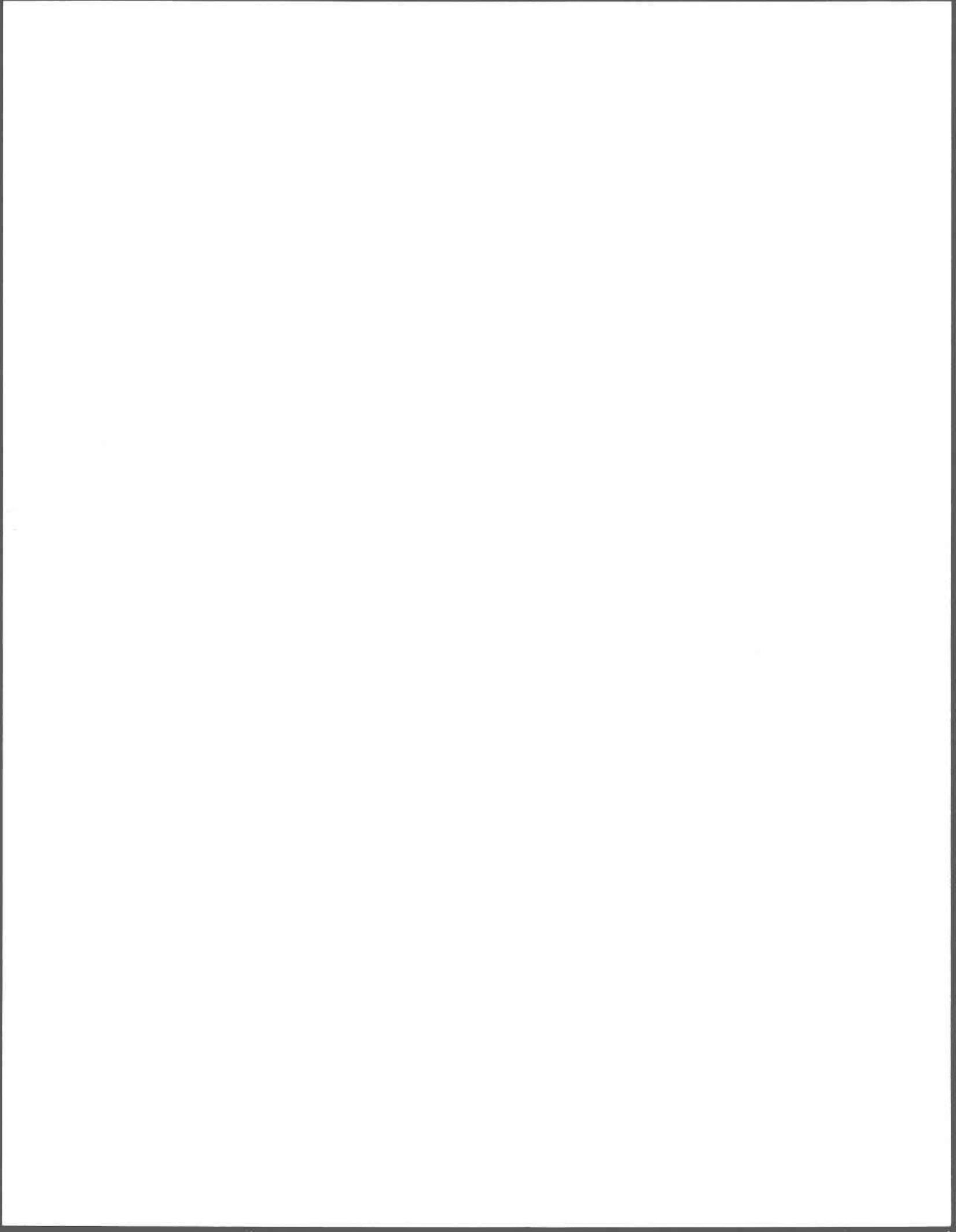
Par
Irène Abi-Zeid
(M. Sc. Mathématiques)

Thèse
présentée pour l'obtention
du grade Philosophiae Doctor (Ph. D.) en Sciences de l'Eau
et celui de Docteur de l'ENGREF en Statistiques Appliquées aux Sciences de l'Eau

Jury d'évaluation

Président du jury et examinateur externe	Christian ONOF Département du Génie Civil Imperial College of Science technology and medicine
Examineur externe	Philippe BOIS Laboratoire LTHE ENSHMG
Examineur externe	Christian GENEST Département de Mathématiques et de Statistique Université Laval
Examineur interne	Jacques BERNIER INRS-Eau Université du Québec
Directeur en cotutelle	Eric PARENT Laboratoire GRESE ENGREF
Directeur en cotutelle	Bernard BOBÉE INRS-Eau Université du Québec

Décembre, 1997



REMERCIEMENTS

Je remercie vivement le professeur Bernard Bobée de l'INRS-Eau, titulaire de la chaire en Hydrologie Statistique - Hydro Québec, d'avoir accepté de diriger cette thèse. Ses conseils et sa confiance tout au long de ce travail m'ont permis de persévérer. Je remercie également le professeur Éric Parent, directeur du laboratoire GRESE à l'ENGREF, d'avoir accepté de diriger cette thèse en cotutelle à l'ENGREF, et de m'avoir accueillie dans son laboratoire pour un stage d'un an où j'ai pu, grâce à nos nombreuses discussions, progresser dans mon travail.

Je remercie les professeurs Philippe Bois du laboratoire d'étude des transferts en hydrologie et environnement de L'ENSHMG, Christian Onof du département de génie civil du *Imperial College of Science, Technology and Medicine* qui a agi à titre de président du jury, et Christian Genest du département de mathématiques et de statistique de l'université Laval, d'avoir bien voulu participer au jury.

Je tiens à exprimer ma profonde gratitude à Monsieur Bernier, professeur invité à l'INRS-Eau et codirecteur de cette thèse, pour son intérêt et son enthousiasme. Il a grandement contribué à faire avancer ce travail par ses judicieux conseils et ses suggestions.

Les commentaires de tous les membres du jury furent très pertinents et très appréciés, et je les en remercie.

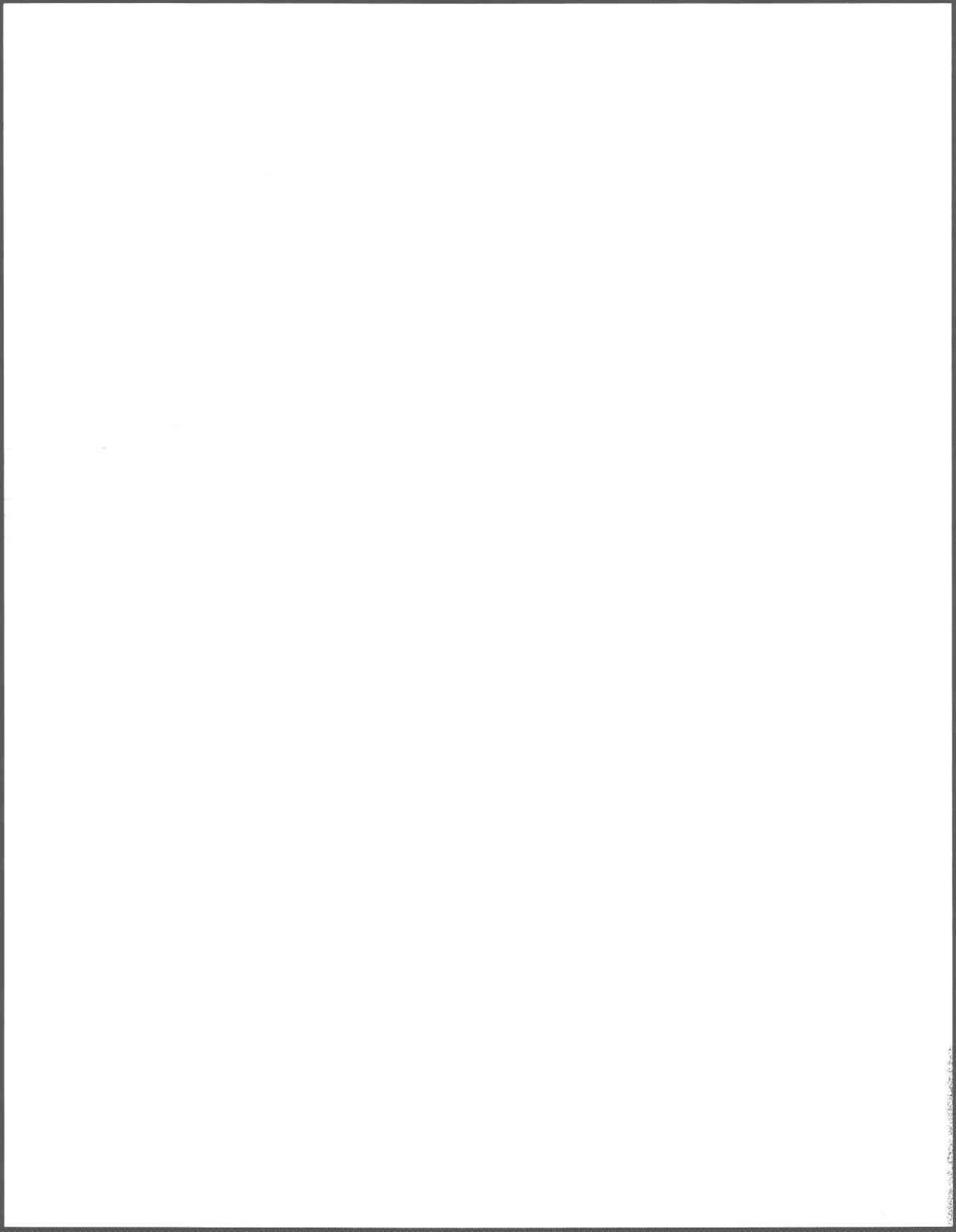
Je désire remercier Dr Paul Pilon d'Environnement Canada de m'avoir fourni les données que j'ai traitées dans ce travail.

Je remercie mes parents de m'avoir appris par l'exemple, l'optimisme, le courage et la persévérance.

Finalement, je suis infiniment reconnaissante envers mon mari Luc Lamontagne, qui m'a encouragée et soutenue tout au long de cette thèse, et sans qui ce travail n'aurait pas été possible. Merci aussi à Mathieu, qui est arrivé un peu plus tard dans le processus, mais qui m'a inspirée et m'a donné la volonté d'aller jusqu'au bout.

Québec, décembre 1997

Le financement de cette thèse a été assuré par FCAR, et par l'INRS-Eau. Le stage à l'ENGREF a été financé en partie par le gouvernement français.



RÉSUMÉ

Les objectifs de ce travail intitulé "La modélisation stochastique des étiages et de leurs durées en vue de l'analyse du risque" sont les suivants:

- la modélisation, selon l'approche des processus ponctuels non stationnaires, des occurrences et des durées des surplus et des étiages;
- la construction et l'évaluation de ces modèles sur différentes rivières; et
- la définition et la quantification d'indices de risque basés sur les durées d'étiage.

Ce document comprend 9 chapitres. Après l'introduction du chapitre 1, nous effectuons au chapitre 2 une revue bibliographique exhaustive des travaux reliés à la modélisation des étiages. Au chapitre 3, nous résumons les éléments de la théorie des processus ponctuels nécessaires à notre recherche. Les processus ponctuels retenus sont le processus de Poisson non homogène (PNH), le processus de renouvellement non homogène (RNH), le processus de Poisson non homogène conditionnel (PNHC), et le processus autoexcité (AE).

Nous développons au chapitre 4 une méthode d'estimation bayésienne semi-paramétrique de la fonction d'intensité d'un processus de Poisson non homogène. Nous comparons cette méthode avec la méthode paramétrique du maximum de vraisemblance (MV) sur des données simulées ainsi que sur des données de précipitations observées sur la station française de Bar-sur-Seine. La méthode bayésienne donne des résultats comparables à la méthode du MV sans avoir à spécifier a priori la forme paramétrée de la fonction d'intensité.

Au chapitre 5, nous développons une méthodologie pour la modélisation des séries alternées de surplus et d'étiage. Cette méthodologie est suffisamment générale pour s'appliquer à toute série alternée de deux états. Plus particulièrement, nous proposons des modèles basés sur les PNHC qui incorporent les précipitations et les températures journalières dans la formulation des fonctions d'intensité.

Au chapitre 6, nous construisons, appliquons, et évaluons les processus non stationnaires de PNH, RNH et PNHC sur 10 rivières canadiennes et une rivière française¹. Les étiages sont définis à partir de deux seuils correspondant aux 10^{ème} et 25^{ème} quantiles, c'est-à-dire les

¹ Nous utilisons les données journalières des rivières suivantes: Pigeon, Magnetawan Nord, Black, Missinaibi, et Turtle en Ontario, Upper Humber à Terre Neuve, Harricana au Québec, Bow et Athabasca en Alberta, Sprague au Manitoba, et Bar-sur-Seine en France. Les débits moyens journaliers de ces rivières, de superficie entre 321 km² et 74600 km², varient entre 1.6 m³/s et 442 m³/s.

débats dépassés 90% et 75% du temps respectivement. Nous constatons que les modèles simples de PNH, RNH, et PNHC sont adéquats pour modéliser les occurrences de surplus et d'étiage dans la plupart des cas.

Au chapitre 7, Nous avons développé pour les PNH, RNH et PNHC les indices de risque suivants: $I_1(x,u)$, la probabilité qu'un étiage qui a débuté à l'instant u dure plus de x jours; $I_2(x,u)$, la probabilité qu'un étiage qui a débuté à l'instant u se termine étant donné qu'il a déjà duré x jours; $I_3(y,u)$, la probabilité d'entrer en étiage sur l'intervalle $(u,u+y)$ étant donné qu'on est entré en surplus à l'instant u ; $I_4(z,u)$, la probabilité d'entrer à nouveau en étiage sur $(u, u+z)$ étant donné qu'on est entré en étiage à l'instant u ; et $I_5(x,u)$, la probabilité d'être en étiage au temps $t=x+u$ étant donné que le dernier étiage a débuté à l'instant u . Nous avons quantifié ces indices pour des données réelles afin d'illustrer leur applicabilité. Les résultats présentés au chapitre 7 sont plutôt de nature exploratoire.

Nous concluons au chapitre 8 en résumant les motivations de ce travail, les objectifs, la démarche suivie, les résultats obtenus, les contributions ainsi que des avenues de recherche pour des travaux futurs.

TABLE DES MATIÈRES

1. INTRODUCTION	1
1.1 Problématique de recherche	1
1.1.1 La modélisation statistique des étiages	1
1.1.2 L'analyse du risque	2
1.2 Objectifs de recherche	3
1.3 Plan de la thèse	4
2. REVUE BIBLIOGRAPHIQUE.....	7
2.1 La sécheresse	7
2.1.1 Une définition méthodologique de la sécheresse	11
2.2 Les étiages	12
2.3 La modélisation des sécheresses hydrologiques et des étiages	13
2.3.1 L'analyse fréquentielle des faibles débits.....	14
2.3.2 L'analyse des durées et les processus stochastiques	18
2.4 Les indicateurs de sécheresse	29
2.5 Les indicateurs d'étiage.....	31
2.6 Le risque et l'incertitude	33
2.6.1 Les définitions de l'incertitude	34
2.6.2 Les définitions du risque	35
2.7 L'analyse du risque	36
2.7.1 La modélisation de l'incertitude	37
2.7.2 La quantification du risque	38
2.8 En résumé.....	43
3. LES PROCESSUS PONCTUELS.....	45
3.1 Définitions générales.....	45
3.2 Définitions de quelques processus ponctuels.....	50
3.2.1 Le processus de Poisson homogène	51
3.2.2 Le processus de Poisson non homogène	52
3.2.3 Le processus de renouvellement homogène.....	52
3.2.4 Le processus de Poisson doublement stochastique	53
3.2.5 Le processus autoexcité	53
3.3 La stationnarité des processus ponctuels.....	57
3.3.1 L'homogénéisation d'un processus ponctuel non homogène	58

3.3.2 L'homogénéisation du processus de Poisson non homogène	58
3.4 L'analyse des séries d'événements	60
3.4.1 L'analyse des tendances et de la périodicité	61
3.4.2 L'indépendance des intervalles - Tests de renouvellement.....	66
3.4.3 Tests du processus de Poisson	68
4. L'ESTIMATION BAYÉSIENNE DE L'INTENSITÉ D'UN PROCESSUS DE POISSON NON HOMOGÈNE	71
4.1 Le principe de superposition	71
4.2 L'estimation bayésienne de l'intensité d'un PNH.....	72
4.2.1 Les hypothèses	73
4.2.2 La méthodologie et l'algorithme d'estimation.....	75
4.2.3 Quantification de l'incertitude sur l'intensité estimée par la méthode BSP	81
4.3 Méthode du maximum de vraisemblance	82
4.3.1 Quantification de l'incertitude sur l'intensité estimée par la méthode du MV.....	83
4.4 Applications	83
4.4.1 Méthode BSP appliquée à un processus simulé.....	84
4.4.2 Comparaison de la méthode du MV et la méthode BSP, $\bar{\mu}$ inconnue, pour les précipitations sur Bar-sur-Seine.....	87
4.5 Discussion	89
4.5.1 Processus simulé	89
4.5.2 Processus de précipitations (Bar-sur-Seine)	90
4.5.3 Conclusions.....	90
5. MÉTHODOLOGIE DE CONSTRUCTION ET D'ÉVALUATION DES MODÈLES D'OCCURRENCES	91
5.1 La définition des événements.....	92
5.2 La définition du processus alterné	93
5.2.1 Le processus de Poisson non homogène (PNH)	96
5.2.2 Le processus de renouvellement non homogène (RNH)	97
5.2.3 Le processus de Poisson non homogène conditionnel (PNHC)	99
5.3 Estimation	101
5.3.1 Le processus de Poisson non homogène (PNH)	101
5.3.2 Le processus de renouvellement non homogène (RNH)	104
5.3.3 Le processus de Poisson non homogène conditionnel (PNHC)	105
5.3.4 Les variances des paramètres des fonctions d'intensité des PNH et RNH	106

5.3.5 Les variances des paramètres des fonctions d'intensité du PNHC.....	106
5.4 Vérification de l'adéquation des modèles.....	107
5.4.1 L'homogénéisation du processus observé	107
5.4.2 L'application de tests graphiques	108
5.4.3 L'application de tests statistiques.....	108
5.4.4 Simulation	112
5.4.5 Analyse de sensibilité.....	112
6. APPLICATIONS ET RÉSULTATS	115
6.1 Présentation des données.....	115
6.2 Estimation de $\lambda^i(t)$ (PNH, RNH).....	118
6.3 Vérification de l'adéquation des modèles (PNH, RNH).....	121
6.3.1 Les étages: résultats du PNH	122
6.3.2 Les surplus: résultats du PNH.....	131
6.3.3 Vérification de l'adéquation du RNH.....	135
6.3.4 Les étages: résultats du RNH.....	135
6.3.5 Les surplus: résultats du RNH	136
6.4 Résumé des résultats du PNH et du RNH.....	137
6.5 Vérification de l'adéquation du PNHC.....	137
6.5.1 Les étages: résultats du PNHC.....	138
6.5.2 Les surplus: résultats du PNHC	138
6.6 Analyse de sensibilité et simulation.....	144
6.7 Conclusions	151
7. QUANTIFICATION DU RISQUE	153
7.1 La loi des durées d'étiage, $I_1(x,u)$	155
7.2 La probabilité qu'un étiage se termine, $I_2(x,u)$	158
7.3 La probabilité d'entrer en étiage, $I_3(y,u)$	161
7.4 La probabilité d'un nouvel étiage, $I_4(z,u)$	164
7.4.1 La densité des intervalles entre les débuts d'étiages	165
7.5 La probabilité d'être en étiage, $I_5(x,u)$	169
7.6 Conclusions	172
8. CONCLUSION	175

8.1 Motivation	175
8.2 Objectifs	175
8.3 Démarche	176
8.4 Résumé des résultats	178
8.5 Contributions.....	179
8.6 Travaux futurs	180
9. BIBLIOGRAPHIE	183
10. ANNEXE 1: ESTIMATION BAYÉSIENNE DE L'INTENSITÉ D'UN PNH.....	191
10.1 La matrice de variance-covariance V.....	191
10.2 L'estimation par la méthode MV des paramètres λ , k , et θ	194
11. ANNEXE 2: LES LOIS CITÉES DANS LE TEXTE.....	203
12. ANNEXE 3: LES VARIANCES DES PARAMÈTRES DES PNH ET RNH.....	207
13. ANNEXE 4: ALGORITHMES DE SIMULATION.....	221
14. ANNEXE 5: MODÈLES DE $\rho^1(t;H_t)$ POUR LE PNHC	225
15. ANNEXE 6: LES VARIANCES DES PARAMÈTRES DU PNHC	227
16. ANNEXE 7: LES DÉBITS MOYENS MENSUELS.....	235
17. ANNEXE 8: LES FONCTIONS ESTIMÉES $\lambda^i(t)$	241
18. ANNEXE 9: RÉSULTATS À $\alpha=.01$	265

19. ANNEXE 10: LES COURBES DÉBITS-DURÉES	273
--	-----

LISTE DES FIGURES

Figure 2.1: La définition des variables décrivant la sécheresse.....	9
Figure 2.2: Définition de Y_t à partir de l'offre et de la demande.....	10
Figure 2.3: Classification des événements hydrologiques (adaptée de Dracup <i>et al.</i> , 1980a).....	12
Figure 2.4: Exemple d'une courbe débit-durée.....	18
Figure 2.5: La définition de X_t à partir de Y_t	20
Figure 2.6: Schéma d'un processus ponctuel.....	25
Figure 2.7: Définition du processus ponctuel alterné de surplus et d'étiage.....	26
Figure 2.8: Indicateur d'étiage à débit fixé (Miquel et Phien Bou Pha, 1978).....	32
Figure 2.9: Analyse du risque (adaptée de Rowe, 1988).....	37
Figure 2.10: Illustration des concepts liés à la sécheresse régionale (Correia <i>et al.</i> 1991).....	41
Figure 3.1: La réalisation d'un processus stochastique ponctuel.....	45
Figure 3.2: La définition des intervalles dans un processus stochastique ponctuel.....	46
Figure 3.3: Le nombre d'événements dans une union d'intervalles.....	46
Figure 3.4: Le lien entre le nombre d'événements et les intervalles.....	47
Figure 3.5: Une fonction d'intensité typique (adaptée de Snyder, 1975).....	48
Figure 3.6: Illustration des variables liées à la probabilité de survie.....	49
Figure 3.7: Un exemple d'une réalisation d'un processus ponctuel (adaptée de Snyder, 1975).....	50
Figure 3.8: Schéma illustrant les relations entre certains processus ponctuels.....	54
Figure 3.9: Un processus débutant à un temps arbitraire ou à un événement arbitraire.....	57
Figure 3.10: La définition de la fonction $\tau^{-1}(t)$ (adaptée de North, 1981).....	59
Figure 3.11: L'homogénéisation d'un PNH (adaptée de North, 1981).....	60
Figure 3.12: Les précipitations sur Bar-sur-Seine, 1951-1955.....	61
Figure 3.13: Le graphique cumulatif des précipitations sur Bar-sur-Seine.....	62
Figure 3.14: La fonction de survie logarithmique.....	62
Figure 3.15: Le taux moyen journalier d'occurrence de précipitation, $\tau=30$	63
Figure 3.16: Le spectre des précipitations sur Bar-sur-Seine.....	64
Figure 4.1: La superposition de processus.....	72
Figure 4.2: L'algorithme d'estimation BSP de la fonction d'intensité d'un PNH.....	80
Figure 4.3: La fonction d'intensité réelle.....	84

Figure 4.4: Comparaison de la fonction d'intensité réelle et son estimation bayésienne, période de génération 15 ans, $\bar{\mu}$ connue	85
Figure 4.5: Comparaison de la fonction d'intensité réelle et son estimation bayésienne, période de génération 55 ans, $\bar{\mu}$ connue	85
Figure 4.6: Comparaison de la fonction d'intensité réelle et son estimation bayésienne, période de génération 15 ans, $\bar{\mu}$ inconnue	86
Figure 4.7: Comparaison de la fonction d'intensité réelle et son estimation bayésienne, période de génération 55 ans, $\bar{\mu}$ inconnue	86
Figure 4.8: Estimateur MV de l'intensité des précipitations (Bar-sur-Seine).....	88
Figure 4.9: Estimateur BSP de l'intensité des précipitations (Bar-sur-Seine), 12 intervalles	88
Figure 4.10: Estimateur BSP de l'intensité des précipitations (Bar-sur-Seine), 30 intervalles	89
Figure 5.1: Définition des temps d'occurrence des événements	93
Figure 5.2: Le processus alterné de surplus et d'étiage	94
Figure 5.3: Le processus de type 1 (rupture de l'état surplus)	95
Figure 5.4: Le processus de type 2 (rupture de l'état étiage)	95
Figure 5.5: Un processus alterné de surplus et d'étiage observé.....	102
Figure 5.6: La vérification de l'adéquation des modèles.....	111
Figure 6.1: Les stations canadiennes (Anderson <i>et al.</i> , 1991)	116
Figure 6.2: Le bassin versant de la Seine à Bar-sur-Seine (Bernier, 1964).....	116
Figure 6.3: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St08(25).....	124
Figure 6.4: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St14(25).....	124
Figure 6.5: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St18(25).....	125
Figure 6.6: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St19(25).....	125
Figure 6.7: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St08(10).....	126
Figure 6.8: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St15(10).....	126

Figure 6.9: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St08(25)	127
Figure 6.10: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St14(25)	127
Figure 6.11: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St18(25)	128
Figure 6.12: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St19(25)	128
Figure 6.13: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St08(10)	129
Figure 6.14: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St15(10)	129
Figure 6.15: Les fonctions de survie logarithmiques des durées de surplus empiriques non homogénéisées et des durées théoriques d'un PNH, St14(25)	133
Figure 6.16: Les fonctions de survie logarithmiques des durées de surplus empiriques non homogénéisées et des durées théoriques d'un PNH, St22(10)	133
Figure 6.17: Les fonctions de survie logarithmiques des durées de surplus empiriques homogénéisées et des durées théoriques d'un PH, St14(25)	134
Figure 6.18: Les fonctions de survie logarithmiques des durées empiriques de surplus homogénéisées et des durées théoriques d'un PH, St22(10)	134
Figure 6.19: Temps d'arrivée dans l'année des événements de rupture de surplus - St19(25)	140
Figure 6.20: Temps d'arrivée dans l'année des événements de rupture de surplus - Stse(25)	141
Figure 6.21: Temps d'arrivée dans l'année des événements de rupture de surplus - St08(10)	141
Figure 6.22: Temps d'arrivée dans l'année des événements de rupture de surplus - St18(25)	142
Figure 6.23: Temps d'arrivée dans l'année des événements de rupture de surplus - St14(10)	142
Figure 6.24: Temps d'arrivée dans l'année des événements de rupture de surplus - St20(10)	143
Figure 6.25: Des séries de surplus/étiage simulées et observées à St07(25) (PNH/PNH)	147
Figure 6.26: Des séries de surplus/étiage simulées et observées à St08(25) (PNH/RNH)	147
Figure 6.27: Des séries de surplus/étiage simulées et observées à St14(25) (PNHC/PNH)	148
Figure 6.28: Des séries de surplus/étiage simulées et observées à St15(25) (RNH/RNH)	148
Figure 6.29: Histogramme des durées d'étiage observées et simulées à St06(25)- PNH	149
Figure 6.30: Histogramme des durées d'étiage observées et simulées à St08(25)- RNH	149
Figure 6.31: Histogramme des durées d'étiage observées et simulées à St14(10)- PNHC	150

Figure 6.32: Histogramme des durées de surplus observées et simulées à St14(10)- PNHC.....	150
Figure 7.1: $I_1(x, 2 \text{ janvier}) = P(X > x T = 2 \text{ janvier})$ en fonction de x - PNH (St07(25)).....	155
Figure 7.2: $I_1(x, 30 \text{ avril}) = P(X > x T = 30 \text{ avril})$ en fonction de x - RNH (St15(25)).....	156
Figure 7.3: $p_{\text{norm}}(t, t = 7/12/43 + x)$ St18	157
Figure 7.4: $I_1(x, 7/12/43) = P(X > x T = 7/12/43)$, en fonction de x - PNHC (St18(25)).....	158
Figure 7.5: $I_2(x, 2 \text{ janvier}) = P(X = x X \geq x, T = 2 \text{ janvier})$ en fonction de x - PNH (St07(25))	159
Figure 7.6: $I_2(x, 30 \text{ avril}) = P(X = x X \geq x, T = 30 \text{ avril})$ en fonction de x - RNH (St15(25)).....	160
Figure 7.7: $I_2 = P(X = x X \geq x, T = 7/12/43)$ en fonction de x - PNHC (St18(25)).....	161
Figure 7.8: $I_3(y, u) = P(Y \leq y T = 2 \text{ janvier})$ en fonction de y - PNH (St07(25)).....	162
Figure 7.9: $I_3(y, 30 \text{ avril}) = P(Y \leq y T = 30 \text{ avril})$ en fonction de y - RNH (St15(25)).....	163
Figure 7.10: $I_3(y, 7/12/43) = P(Y \leq y T = 7/12/43)$ en fonction de y - PNHC (St18(25)).....	164
Figure 7.11: $n_{\text{joursec}}(t, t = 7/12/43 + y, \text{St18})$	164
Figure 7.12: Z , l'intervalle entre 2 débuts d'étiage.....	165
Figure 7.13: $I_4(z, 2 \text{ janvier}) = D(z T = 2 \text{ janvier})$ en fonction de z - PNH (St07(25)).....	168
Figure 7.14: $I_4(z, 30 \text{ avril}) = D(z T = 30 \text{ avril})$ en fonction de z - RNH (St15(25)).....	168
Figure 7.15: $I_4(z, 7/12/43) = D(z T = 7/12/43)$ en fonction de z - PNHC St18(25))	169
Figure 7.16: $I_5(x, 2 \text{ janvier}) = P_2(x + 2 \text{ janvier} T = 2 \text{ janvier})$ en fonction de x - PNH (St07(25))	171
Figure 7.17: $I_5(x, 30 \text{ avril}) = P_2(x + 30 \text{ avril}, T = 30 \text{ avril})$ en fonction de x - RNH (St15(25)).....	171
Figure 7.18: $I_5(x, 7/12/43) = P_2(x + 7/12/43, T = 7/12/43)$ en fonction de x - PNHC (St18(25)).....	172
Figure 16.1: Les débits moyens mensuels - St06.....	235
Figure 16.2: Les débits moyens mensuels - St07.....	235
Figure 16.3: Les débits moyens mensuels - St08.....	236
Figure 16.4: Les débits moyens mensuels - St14.....	236
Figure 16.5: Les débits moyens mensuels - St15.....	237
Figure 16.6: Les débits moyens mensuels - St16.....	237
Figure 16.7: Les débits moyens mensuels - St18.....	238
Figure 16.8: Les débits moyens mensuels - St19.....	238
Figure 16.9: Les débits moyens mensuels - St20.....	239
Figure 16.10: Les débits moyens mensuels - St22.....	239
Figure 16.11: Les débits moyens mensuels - Stse	240

Figure 17.1: Les fonctions estimées par le mv $\lambda^2(t)$ et $\lambda^1(t)$ - St06(25).....	241
Figure 17.2: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St06(25).....	242
Figure 17.3: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St07(25).....	243
Figure 17.4: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St07(25).....	243
Figure 17.5: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St08(25).....	244
Figure 17.6: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St08(25).....	244
Figure 17.7: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St14(25).....	245
Figure 17.8: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St14(25).....	245
Figure 17.9: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St15(25).....	246
Figure 17.10: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St15(25).....	246
Figure 17.11: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St16(25).....	247
Figure 17.12: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St16(25).....	247
Figure 17.13: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St18(25).....	248
Figure 17.14: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St18(25).....	248
Figure 17.15: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St19(25).....	249
Figure 17.16: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St19(25).....	249
Figure 17.17: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St20(25).....	250
Figure 17.18: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St20(25).....	250
Figure 17.19: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St22(25).....	251
Figure 17.20: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St22(25).....	251
Figure 17.21: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - Stse(25).....	252
Figure 17.22: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - Stse(25).....	252
Figure 17.23: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St06(10).....	253
Figure 17.24: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St06(10).....	253
Figure 17.25: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St07(10).....	254
Figure 17.26: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St07(10).....	254
Figure 17.27: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St08(10).....	255
Figure 17.28: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St08(10).....	255
Figure 17.29: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St14(10).....	256
Figure 17.30: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St14(10).....	256
Figure 17.31: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St15(10).....	257
Figure 17.32: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St15(10).....	257
Figure 17.33: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St16(10).....	258

Figure 17.34: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St16(10)	258
Figure 17.35: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St18(10)	259
Figure 17.36: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St18(10)	259
Figure 17.37: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St19(10)	260
Figure 17.38: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St19(10)	260
Figure 17.39: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St20(10)	261
Figure 17.40: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St20(10)	261
Figure 17.41: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St22(10)	262
Figure 17.42: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St22(10)	262
Figure 17.43: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - Stse(10).....	263
Figure 17.44: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - Stse(10).....	263
Figure 18.1: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St22(25).....	265
Figure 18.2: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St22(25)	266
Figure 18.3: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St14(10).....	266
Figure 18.4: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St14(10)	267
Figure 18.5: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St18(10).....	267
Figure 18.6: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St18(10)	268
Figure 18.7: Les fonctions de survie logarithmiques des durées de surplus empiriques non homogénéisées et des durées théoriques d'un PNH, St18(10).....	268
Figure 18.8: Les fonctions de survie logarithmiques des durées de surplus empiriques homogénéisées et des durées théoriques d'un PH, St18(10)	269
Figure 18.9: Les fonctions de survie logarithmiques des durées de surplus empiriques homogénéisées (RNH) et des durées théoriques d'un RH, St14(25).....	270
Figure 18.10: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées (PNHC) et des durées théoriques d'un PH, Stse(10).....	270

Figure 19.1: La courbe débit-durée - St06	273
Figure 19.2: La courbe débit-durée - St07	273
Figure 19.3: La courbe débit-durée - St08	274
Figure 19.4: La courbe débit-durée - St14	274
Figure 19.5: La courbe débit-durée - St15	275
Figure 19.6: La courbe débit-durée - St16	275
Figure 19.7: La courbe débit-durée - St18	276
Figure 19.8: La courbe débit-durée - St19	276
Figure 19.9: La courbe débit-durée - St20	277
Figure 19.10: La courbe débit-durée - St22	277
Figure 19.11: La courbe débit-durée - Stse.....	278

LISTE DES TABLEAUX

Tableau 2.1: Résumé des lois appliquées aux débits d'étiage	17
Tableau 2.2: Résumé des travaux d'analyse de séquences	25
Tableau 2.3: Résumé des processus ponctuels appliqués aux sécheresses et aux étiages	29
Tableau 2.4: Quelques indicateurs d'étiages (adapté de Miquel et Phien Bou Pha, 1979).....	31
Tableau 2.5: Incertitude, imprécision et erreur	33
Tableau 2.6: Résumé des incertitudes	35
Tableau 3.1: Certaines définitions reliées aux processus ponctuels PH, PNH et RH.....	55
Tableau 3.2: Certaines définitions reliées aux processus ponctuels AE et PNHC.....	56
Tableau 3.3: Les valeurs critiques des statistiques de K-S et d'Anderson	68
Tableau 4.1: Les estimateurs du MV pour les précipitations.....	87
Tableau 4.2: Les variances moyennes de $\hat{\lambda}_i$	89
Tableau 5.1: Étapes de la modélisation d'une série d'événements	92
Tableau 5.2: Étapes de la définition du processus alterné	96
Tableau 5.3: Les hypothèses pour l'application d'un PNH.....	97
Tableau 5.4: Les hypothèses pour l'application du RNH.....	98
Tableau 5.5: Les hypothèses pour l'application du PNHC.....	101
Tableau 5.6: Étapes de la vérification de l'adéquation des modèles	107
Tableau 5.7: L'application des tests aux divers modèles.....	111
Tableau 5.8: Analyse de sensibilité.....	113
Tableau 6.1: Description des stations de mesure de débit	117
Tableau 6.2: Description des événements aux stations correspondant aux seuils Q_{25} et Q_{10}	118
Tableau 6.3: Les paramètres estimés de $\hat{\lambda}^2(t)$, seuil Q_{25} (Étiage).....	119
Tableau 6.4: Les paramètres estimés de $\hat{\lambda}^1(t)$, seuil Q_{25} (Surplus).....	119
Tableau 6.5: Les paramètres estimés de $\hat{\lambda}^2(t)$, seuil Q_{10} (Étiage).....	120
Tableau 6.6: Les paramètres estimés de $\hat{\lambda}^1(t)$, seuil Q_{10} (Surplus).....	120
Tableau 6.7: Tests du PNH - $i=2$ (Étiage).....	122
Tableau 6.8: Tests du PNH - $i=1$ (Surplus).....	131
Tableau 6.9: Test d'adéquation de la loi Gamma aux durées homogénéisées - $i=2$ (Étiage).....	135
Tableau 6.10: Test d'adéquation de la loi Gamma aux durées homogénéisées - $i=1$ (Surplus).....	136
Tableau 6.11: Stations canadiennes utilisées pour l'évaluation des PNHC.....	137
Tableau 6.12: Les paramètres estimés de $\hat{\rho}^2(t, H_1)$ - PNHC	138

Tableau 6.13: Test du PNHC - $i=2$ (Étiage).....	138
Tableau 6.14: Les paramètres estimés de $\hat{\rho}^1(t, H_t)$ - PNHC.....	139
Tableau 6.15: Tests du PNHC - $i=1$ (Surplus).....	139
Tableau 6.16: Synthèse des résultats.....	144
Tableau 6.17: Analyse de sensibilité - $i=2$ (Étiage).....	145
Tableau 6.18: Analyse de sensibilité - $i=1$ (Surplus).....	146
Tableau 7.1: Stations pour lesquelles les indices de risque sont calculés.....	154

1. INTRODUCTION

1.1 Problématique de recherche

La croissance continue de la population mondiale et l'augmentation du niveau de vie dans certaines parties de la planète exercent une pression de plus en plus forte sur la demande quantitative et qualitative de la ressource hydrique, nécessitant ainsi une gestion plus adéquate.

Afin d'assurer une gestion efficace des ressources, il est nécessaire d'avoir des outils (1) de modélisation statistique et de simulation des étiages (faibles débits) qui tiennent compte de leur caractère non stationnaire, et (2) qui permettent la quantification des risques par des indices de performance dans un cadre d'analyse du risque.

1.1.1 La modélisation statistique des étiages

La connaissance des propriétés statistiques des étiages est essentielle pour la conception des politiques de gestion et de contrôle des ressources hydriques. L'information concernant les faibles débits est utilisée pour déterminer les débits réservés, planifier les débits de régulation, analyser les impacts environnementaux et économiques, modéliser la qualité, et améliorer le niveau général de compréhension des systèmes naturels et régularisés. Pourtant, d'après Dracup (1986), l'étude des sécheresses en termes de durée, d'intensité (volume de déficit moyen), et de sévérité (volume de déficit total) est l'un des aspects les plus négligés en hydrologie, et les articles scientifiques reliés aux crues sont 6 fois plus nombreux que ceux reliés aux sécheresses.

La majorité des études de modélisation statistique des étiages ont porté sur la distribution des valeurs minimales annuelles des débits moyens journaliers, ou sur le minimum annuel du débit moyen sur m jours, Q_m . Pourtant, à l'inverse d'une crue qui est un phénomène spectaculaire de durée souvent négligeable et dont l'impact est dû surtout à son intensité, un étiage constitue un phénomène qui perdure et dont l'impact est dû surtout à la durée des faibles débits. Il s'avère alors impératif de développer des modèles d'étiages dont la variable principale est la durée du déficit. Ces modèles devront tenir compte de la non-stationnarité du phénomène, et en particulier de son effet saisonnier.

La plupart des techniques de modélisation des étiages, telles que l'analyse des débits, ne tiennent pas compte des durées et émettent des hypothèses trop simplifiées de stationnarité. L'analyse des séquences permet d'obtenir les lois des durées pour des processus de débits très simples Y_t , où les variables Y_t sont indépendantes, distribuées selon une loi normale, ou au plus un processus autorégressif simple. Peu de distributions ont été obtenues pour des modèles Y_t plus réalistes, tels que des processus stochastiques non stationnaires périodiques. La modélisation des états surplus-étiage par des processus markoviens n'est pas appropriée pour les dépendances à long terme, et la prise en compte de l'effet saisonnier passe par une subdivision de l'année, ce qui augmente le nombre des paramètres à estimer et implique un certain arbitraire dans la division.

L'approche des processus ponctuels est prometteuse car elle permet la caractérisation et la simulation des étiages tout en tenant compte de la non-stationnarité du phénomène. En général, les processus de type renouvellement sont appropriés lorsqu'on veut étudier les intervalles entre les événements. Toutefois, ils ne peuvent être utilisés s'il existe une structure de dépendance complexe entre ces intervalles. Nous émettons et voulons vérifier l'hypothèse qu'un processus de Poisson non homogène conditionnel pourrait s'avérer un modèle pertinent puisque les étiages sont conditionnés par des processus stochastiques externes tels que les précipitations et que leur structure de dépendance a une origine physique. En pratique, nous faisons l'hypothèse que les processus stochastiques externes sont connus.

1.1.2 L'analyse du risque

L'analyse du risque peut aider à la prise de décision en avenir incertain en fournissant des critères à optimiser dans un problème de décision. On identifie deux volets: le premier est l'évaluation du risque qui consiste à identifier les facteurs de risque et à les estimer, et le deuxième consiste à l'accepter et à le gérer. Le risque a été défini classiquement comme la probabilité de dépassement d'un seuil critique. Toutefois, cette définition du risque n'est pas suffisante pour bien décrire l'événement (étiage) et en cerner les impacts possibles, et il est préférable de définir le risque par des indices de risque. Ces indices peuvent servir à juger et à comparer plusieurs politiques de gestion dans un contexte de prise de décision multicritère puisque les projets de gestion des ressources en eau visent généralement à combler plusieurs

objectifs souvent non commensurables et conflictuels, certains à caractère quantitatif, d'autres à caractère qualitatif.

1.2 Objectifs de recherche

Le but de cette recherche est l'évaluation des processus ponctuels pour modéliser les occurrences et les durées des étiages et des surplus afin de caractériser et quantifier les risques qui en découlent, tout en tenant compte de la non-stationnarité du phénomène. Dans ce travail, un étiage n'est pas interprété au sens physique de période où le ruissellement est nul (Dub et Dzubak, 1960), mais plutôt comme un déficit défini par rapport à un seuil. Un cours d'eau peut alors être considéré comme un système à deux états, un état satisfaisant (surplus), et un état non-satisfaisant (déficit, étiage); l'incident indésirable (passage à l'état de déficit) se produit lorsque le débit descend en dessous d'un seuil critique donné.

L'étude des étiages peut être abordée soit comme un événement aléatoire dont on cherche à connaître la loi de probabilité, soit comme un événement naturel dont on cherche à connaître le processus physique; la première approche est adoptée dans ce travail. Trois étapes principales sont identifiées:

- la modélisation, selon l'approche des processus ponctuels non stationnaires, des occurrences et des durées des surplus et des étiages;
- la construction et l'évaluation de ces modèles sur différentes rivières; et
- la définition et la quantification d'indices de risque basés sur les durées d'étiage.

Notre démarche d'analyse du risque s'inscrit dans le cadre de l'évaluation du risque. La gestion du risque n'est pas abordée directement dans ce travail. Nous cherchons plutôt à faire un lien entre l'hydrologie statistique qui consiste à évaluer des probabilités d'occurrence d'événements tels que les crues ou les sécheresses (risque hydrologique) et leur gestion, et ce, en fournissant des mesures de risque pour quantifier les impacts de ces événements. Ces mesures pourraient être optimisées lors de l'élaboration de politique de gestion des ressources hydriques.

Les objectifs plus particuliers de cette recherche sont les suivants:

- la présentation d'une revue des travaux reliés à la modélisation des étiages et des sécheresses;
- le développement d'une méthode d'estimation bayésienne de l'intensité d'un processus de Poisson non homogène;
- l'évaluation des processus de Poisson non homogènes (PNH) et des processus de renouvellement non homogènes (RNH) comme modèles d'occurrence de déficit et de surplus sur des rivières présentant des étiages d'hiver et dont les bassins versants sont de différentes dimensions. Cette évaluation est effectuée à partir des débits journaliers de plusieurs rivières canadiennes et d'une rivière française. Cette étude permettra de juger si les modèles développés par North (1981) pour des rivières françaises sont applicables au Canada, et ce compte tenu des conditions climatiques et géologiques du pays;
- la construction de modèles d'occurrence d'étiage et de surplus ainsi que leurs durées à l'aide de processus de Poisson non homogènes conditionnels (PNHC) qui font intervenir des variables explicatives telles que les précipitations et la température;
- l'évaluation des modèles PNHC sur les rivières pour lesquelles ni le PNH ni le RNH ne sont adéquats; et
- la définition et la quantification d'indices de risque reliés aux durées des étiages pour les processus PNH, RNH et PNHC.

1.3 Plan de la thèse

En premier lieu, cette thèse se veut une synthèse des travaux effectués sur la modélisation des sécheresses et des étiages. Une revue bibliographique permet de présenter au chapitre 2

un résumé des travaux pertinents à notre étude et de faire le point sur les diverses approches à l'étude des sécheresses et des étiages. Les principaux outils mathématiques utilisés provenant de la théorie des processus stochastiques sont présentés au chapitre 3. Puisque le PNH est souvent utilisé comme modèle d'occurrence des précipitations, des crues, et des étiages, une méthode d'estimation bayésienne de la fonction d'intensité d'un PNH est développée au chapitre 4. Elle est comparée à la méthode d'estimation par le maximum de vraisemblance pour un PNH simulé et pour des données réelles de précipitations sur la station française de Bar-sur-Seine. La méthodologie de construction et d'évaluation des modèles d'occurrence préconisée dans notre étude est présentée au chapitre 5. Ce chapitre décrit la définition d'un événement, les processus appliqués, les méthodes d'estimation des paramètres et d'homogénéisation du processus, le calcul de l'incertitude, ainsi que les tests de séries d'événements utilisés afin de vérifier l'adéquation des processus ponctuels proposés. Le chapitre 6 présente les résultats de l'application de ces processus à un ensemble de rivières canadiennes et à une rivière française ainsi que l'analyse de sensibilité. Le chapitre 7, de nature plutôt exploratoire, propose des indices de risque reliés aux durées des étiages. Ces indices sont quantifiés analytiquement pour les PNH, RNH, et PNHC, et évalués sur des données réelles de débits. Le chapitre 8 présente la conclusion et les perspectives futures de recherche.

2. REVUE BIBLIOGRAPHIQUE

Cette revue bibliographique comporte deux volets principaux: le premier est relié à la modélisation des étiages, et le deuxième à l'analyse du risque.

Avant d'entreprendre une revue des travaux en modélisation des étiages, il est important de faire le point sur la différence entre la sécheresse (*drought*), et les étiages (*low flows*) car ces termes sont souvent confondus dans les publications, et il existe un manque d'uniformité au niveau de la terminologie. Le thème de la sécheresse est abordé dans la section 2.1 et celui des étiages dans la section 2.2. Par la suite, nous présentons les principales approches de modélisation des étiages (section 2.3), les indicateurs de sécheresse (2.4), les indicateurs d'étiage (2.5), le risque et l'incertitude (2.6), et finalement l'analyse du risque (2.7).

2.1 La sécheresse

La sécheresse diffère des autres sources de risques naturels en plusieurs points. Premièrement, c'est un phénomène qui s'installe lentement et ses effets s'accumulent pendant une longue période de temps. Deuxièmement, ses impacts sont moins spectaculaires que les dommages résultant d'un tremblement de terre ou d'une inondation par exemple. Troisièmement, l'absence d'une définition précise et universellement acceptée ajoute à la confusion reliée à l'existence ou non d'une sécheresse, et lorsqu'elle existe, à sa sévérité.

Comment définir une sécheresse? La sécheresse est un concept très vaste qui touche plusieurs disciplines et dont le domaine de recherche inclut les sciences sociales, les sciences naturelles et le génie. Il n'existe pas de définition universellement acceptée de la sécheresse. Elle dépend du point de vue du domaine affecté (météorologie, hydrologie, géographie, développement des ressources, agriculture, industrie, production hydroélectrique, alimentation en eau, navigation, activités récréatives, etc.). Un gestionnaire des ressources hydriques conçoit la sécheresse comme un problème d'offre et de demande. Un géophysicien a un point de vue incluant la climatologie, la météorologie, l'hydrologie et la limnologie. Pour un fermier, la sécheresse est une fonction de l'humidité du sol et du type

spécifique de culture. D'autres perçoivent la sécheresse selon son impact sur les activités humaines et institutionnelles.

Beaucoup de désaccord existe autour des occurrences de sécheresse: Wilhite et Glantz (1987) décrivent la situation au nord-est du Brésil au milieu des années 1980, où les scientifiques et les gestionnaires suggéraient que la région avait été victime d'une sécheresse de cinq ans, assertion réfutée par les météorologues brésiliens pour qui seulement deux des cinq années pouvaient être classées comme sèches. Des conflits similaires eurent lieu en Australie en 1984 entre le bureau de la météorologie et le département d'agriculture. Les définitions des événements de sécheresse sont fortement reliées à la climatologie et à la géologie de l'endroit. Par exemple, à Bali, une sécheresse est une absence de pluie pendant six jours consécutifs, tandis qu'en Libye c'est l'absence de pluie pendant deux ans.

La sécheresse est un phénomène de régions de fortes précipitations pluvieuses aussi bien que de régions à faibles précipitations. La sévérité est difficile à déterminer, car elle dépend de la durée, de l'intensité, de l'étendue géographique ainsi que de la pression des activités humaines et de la végétation sur la ressource disponible de la région. Il est important de distinguer la sécheresse et l'aridité, bien que les deux soient caractérisées par un manque d'eau. L'aridité est une condition climatique permanente. La sécheresse est une condition extrême spatio-temporelle. L'occurrence des sécheresses est due aux variations inhérentes à la circulation atmosphérique, et peut dépendre de facteurs tels que le transport de poussières volcaniques ou de poussières dans l'air.

Plusieurs définitions spécifiques ont été avancées pour identifier, caractériser et classer les sécheresses (Dracup, 1991). Ces définitions sont basées sur les conditions hydrologiques et climatologiques (Russel *et al.*, 1970), les effets sur l'humidité du sol (Palmer, 1965), la production agricole (Bidwell, 1972), et l'activité économique (Millan, 1972). Wilhite (1982) définit la sécheresse comme une singularité climatique impliquant un manque de précipitations suffisant pour affecter la production des cultures. Kebaili-Bargaoui (1990) présente douze définitions de sécheresse météorologique répertoriées dans trois travaux. En résumé, on peut parler de sécheresse météorologique lorsque la précipitation est sous la normale pour une période prolongée; la sécheresse hydrologique est reliée à une période prolongée de bas débits, de bas niveaux des eaux de surface ou encore des eaux

souterraines; la sécheresse agricole a lieu lorsque l'humidité du sol est déficiente et insuffisante pour supporter les cultures; la sécheresse écologique est un manque d'eau causant un stress aux écosystèmes et affectant la faune et la flore. Dans un contexte économique et social, on tient compte de l'offre et de la demande et de l'impact de la sécheresse sur l'activité économique. Beran et Rodier (1985) ont décrit six aspects possibles de la sécheresse hydrologique exprimés en termes du volume total ruisselé, du débit minimum instantané, et sur une période donnée, du niveau maximum de la rivière, et du niveau minimum de l'aquifère. En général, la sécheresse est synonyme de pénurie ou de rareté d'eau pendant une période de temps prolongée sur un espace géographique significatif. Elle est caractérisée par des précipitations faibles qui persistent dans le temps, et/ou des températures élevées causant une évapotranspiration élevée. Toutefois, l'usage excessif de la ressource dans un contexte de précipitations normales peut aussi causer un déficit hydrique.

Parmi toutes les définitions de sécheresse, une définition objective et assez flexible ayant émergé est celle avancée par Yevjevich (1967) et reliée à la théorie des séquences, une branche de l'analyse statistique des séries chronologiques. À partir de la série temporelle $\{Y_t\}$ d'une variable aléatoire définissant la sécheresse telle que la précipitation, la fonte de neige, l'infiltration, l'humidité du sol, les niveaux de la nappe phréatique, on étudie les variables aléatoires décrivant la sécheresse, définies à l'aide d'un seuil x_0 , telles que la durée n du déficit (RL) et le volume du déficit (RS_n). Ces variables sont illustrées sur la figure 2.1.

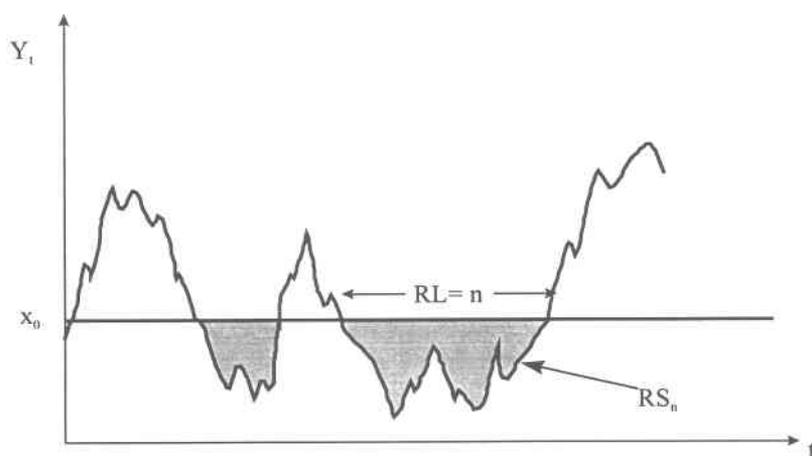


Figure 2.1: La définition des variables décrivant la sécheresse

Au lieu d'utiliser un seuil fixe x_0 afin de définir les déficits et les surplus, Y_t peut être aussi définie comme la différence entre la série temporelle de l'offre et la série temporelle de la demande, telle que présentée sur la figure 2.2.

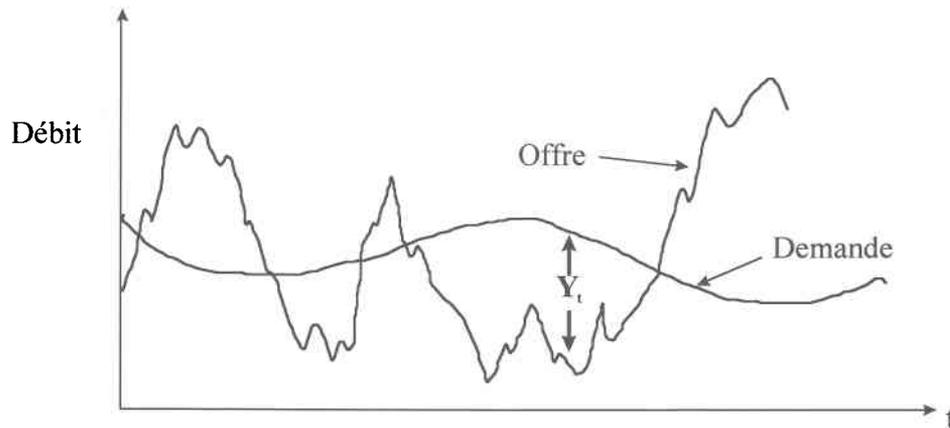


Figure 2.2: Définition de Y_t à partir de l'offre et de la demande

Les sécheresses diffèrent entre elle selon quatre caractéristiques essentielles: l'origine, l'intensité, la durée et l'étendue spatiale. Les conditions de sécheresse peuvent varier substantiellement dans l'espace et dans le temps en fonction de l'irrégularité spatio-temporelle, de la distribution des pluies, et de l'hétérogénéité des réponses hydrologiques des bassins affectés. Il est essentiel de définir ce qu'on veut exactement étudier à propos des sécheresses et les caractéristiques climatologiques et hydrologiques auxquelles on veut accorder une importance. Dans ce contexte, la sécheresse n'est pas une entité définissable mais plutôt un événement qui a des attributs ou des conséquences (Beran et Rodier, 1985). Contrairement aux crues qui peuvent être quantifiées et mesurées, elles ont plutôt des connotations qualitatives.

Pour ce qui a trait à la prédiction des sécheresses, on peut citer deux approches majeures:

- le développement de relations statistiques entre les variables prédites et certaines variables explicatives par le biais de méthodes telles que la corrélation et la régression, l'analyse en composantes principales, l'analyse factorielle; et

- l'utilisation de l'information contenue dans des données historiques afin de prédire le futur, par des processus markoviens par exemple.

2.1.1 Une définition méthodologique de la sécheresse

Afin de définir le type de sécheresse étudié, Dracup *et al.* (1980a) suggèrent de considérer un ensemble de définitions en réponse aux questions suivantes qui affectent la taille de l'échantillon, la corrélation, la sévérité, la magnitude, et la durée:

- Est-ce que l'intérêt principal est dans les précipitations (sécheresse météorologique), les débits (sécheresse hydrologique), ou l'humidité du sol (sécheresse agricole)?
- Quelle est l'unité de temps considérée (jour, mois, saison, année)?
- Comment distinguer analytiquement les sécheresses des autres événements de la série chronologique?
- Comment tient-on compte de l'aspect régional des sécheresses?

2.1.1.1 Le type de déficit étudié

Le déficit peut être relié au débit, à la précipitation, à l'humidité du sol ou à toute combinaison de ces trois aspects. La nature du cycle hydrologique indique que les périodes de faibles précipitations, de faibles débits et de faible humidité du sol sont intimement reliées. La précipitation est un des intrants du système et les débits constituent un des extrants.

2.1.1.2 L'unité de temps

La deuxième étape est de déterminer l'unité de temps à considérer. Ce choix dépend du but pour lequel l'étude a été entreprise. La taille des données historiques disponibles joue un rôle dans ce choix, et la corrélation des événements est différente selon le pas de temps choisi. La plupart des études ont utilisé des données annuelles.

2.1.1.3 Le seuil de dépassement

Une méthode par laquelle les événements de sécheresse sont identifiés est nécessaire. Le concept de seuil est très présent lors de l'analyse des séquences. Il existe plusieurs façons de choisir ce seuil, mais en général, c'est une mesure de la tendance centrale de l'échantillon.

Les seuils utilisés sont souvent la moyenne, la médiane, ou encore des quantiles entre .1 et .5. Ce seuil peut aussi être le maximum des minima annuels, la moyenne des minima annuels, ou encore défini selon un critère économique d'offre et de demande (Ozga-Zielinska, 1989). En combinant la notion de seuil de dépassement avec la notion d'unité de temps du paragraphe précédent, Dracup *et al.* (1980a) ont défini quatre catégories d'événements hydrologiques tels qu'illustrées sur la figure 2.3.

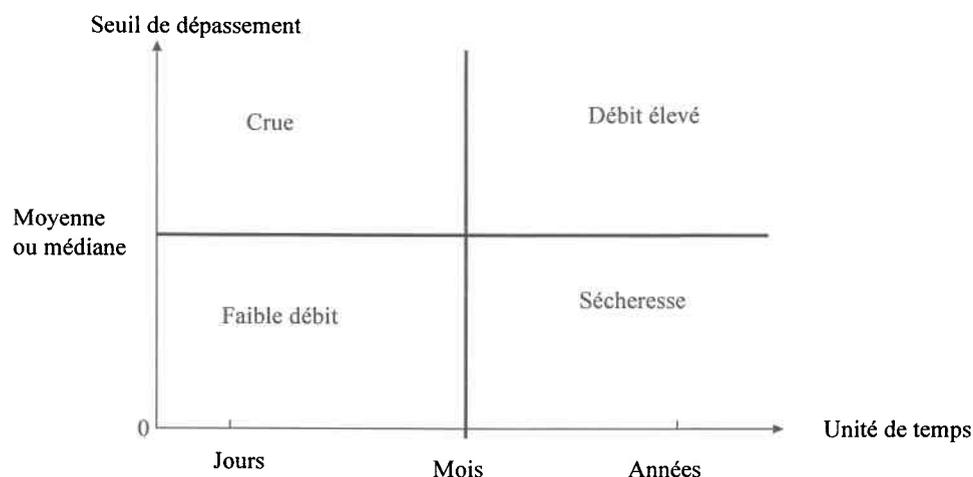


Figure 2.3: Classification des événements hydrologiques (adaptée de Dracup *et al.*, 1980a)

2.1.1.4 La régionalisation

L'analyse régionale est souhaitable car les sécheresses sont foncièrement un phénomène régional. Leur étendue spatiale devrait être considérée, surtout en raison de la petite taille des échantillons généralement disponibles. Cet aspect n'est pas traité en détail dans cette revue bibliographique. Nous présentons toutefois à la section 2.7.2.1 les travaux de Correia *et al.* (1991) reliés aux indices de risque régionaux.

2.2 Les étiages

La sécheresse des cours d'eau est une faiblesse des débits de rivières, dites à l'étiage. Cette sécheresse est due soit à la sécheresse atmosphérique, soit à la sécheresse du sol, soit à la sécheresse souterraine.

Un étiage peut être la conséquence d'une sécheresse hydrologique, mais à l'inverse de celle-ci, il n'est pas généralement considéré comme un événement exceptionnel. On dit qu'un cours d'eau est à l'étiage lorsqu'il y a faiblesse des débits. L'étiage est une notion aux multiples aspects: il peut être une période de l'année, un volume d'eau, un faible débit, ou une durée. Dans la pratique, chacun de ces aspects est décrit par un indicateur précis. Les étiages sont définis sur une base saisonnière et ils sont reliés au cycle solaire annuel et à ses effets climatiques locaux. Un étiage dure habituellement quelques journées ou quelques semaines.

L'étude des faibles débits est reliée au traitement statistique et à la compréhension du développement physique des débits à un point d'une rivière à court terme, ce qui contraste avec l'étude des sécheresses qui inclut la description des précipitations, des débits de rivières, de l'humidité du sol et des eaux souterraines pendant une saison, plusieurs saisons ou plusieurs années et ce, en tenant compte de l'étendue spatiale du phénomène. Les principales caractéristiques numériques attachées aux faibles débits sont: le débit minimum moyen sur m jours consécutifs Q_m , les dates des occurrences, et la fréquence du phénomène. Cette dernière peut être déterminée de plusieurs façons, soit par un tri simple, soit par une probabilité empirique de non-dépassement, soit par une analyse statistique plus rigoureuse de tous les minima annuels, soit par des méthodes alternatives qui expriment les probabilités par des périodes de retour (Beran et Rodier, 1985). Une autre caractéristique des étiages est le taux auquel le débit tarit en l'absence de pluie, exprimé par la courbe de tarissement. Cette courbe peut être difficile à déterminer dans le cas d'un cours d'eau régulé. La courbe de tarissement varie avec les saisons, et sa définition dans les pays nordiques est difficile à cause des effets du gel et du dégel.

2.3 La modélisation des sécheresses hydrologiques et des étiages

Il existe deux approches à la modélisation des sécheresses hydrologiques et des étiages, l'approche déterministe et l'approche stochastique. Cette dernière approche est réaliste vue la nature aléatoire des débits des rivières et des facteurs météorologiques qui les conditionnent (Bernier, 1965), et elle est adoptée dans ce travail. On peut identifier deux grandes classes d'approches stochastiques:

- l'étude fréquentielle du minimum annuel des débits journaliers et du minimum annuel de la moyenne mobile sur m jours des débits journaliers, Q_m par les lois statistiques; et
- l'analyse des durées par les processus stochastiques.

Cette revue de littérature présente un survol des deux classes ainsi que des indicateurs de sécheresse et des indices de performance reliés au risque. La distinction n'étant pas faite entre étiage et sécheresse dans les travaux résumés, nous reprendrons ici les expressions utilisées par les divers auteurs pour décrire le déficit hydrique.

2.3.1 L'analyse fréquentielle des faibles débits

La plupart des études d'analyse de fréquence ont pour objet d'assigner à des valeurs de débit une fréquence cumulative exprimée en quantiles, soit directement par des fonctions de répartition empiriques, soit après ajustement par des distributions théoriques. Ceci permet le calcul des débits d'étiage critiques X_T correspondant à une certaine période de retour T , définie comme étant le temps moyen (calculé sur une très longue période) entre deux occurrences d'étiages. Ces travaux ont surtout porté sur l'étude du minimum annuel du débit journalier où les données hydrologiques sont supposées indépendantes et identiquement distribuées. Il est aussi d'usage d'analyser le minimum annuel du débit moyen sur m jours consécutifs, m étant souvent 7, 10, 30, 60, 90, ou 120 jours, et d'adopter comme débit critique un débit m jours ayant une période de retour de T ans. Le débit critique le plus utilisé en Amérique du Nord est le $Q_{7,10}$ qui correspond au minimum annuel lissé sur 7 jours, ayant une période de retour de 10 ans. Riggs (1972) décrit la procédure à suivre afin d'obtenir les caractéristiques fréquentielles des débits d'étiage. Plusieurs lois combinées à diverses méthodes d'estimation des paramètres ont été utilisées en hydrologie statistique.

Loganathan *et al.* (1986) classent ces approches en quatre catégories. Les trois premières catégories appartiennent aux modèles dits de boîte noire et n'incorporent pas l'aspect physique du processus hydrologique: (1) la méthode de transformation, où les données originales sont normalisées et la loi normale est appliquée, (2) la distribution hypothétique, où une loi est ajustée directement aux données, (3) la distribution empirique, où les données sont ordonnées et tracées selon une formule de probabilité empirique (*plotting position*) qui

est utilisée pour estimer les probabilités de dépassement, et (4) les méthodes de probabilité à base physique qui utilisent les relations hydrologiques et des distributions hypothétiques des variables aléatoires hydrologiques indépendantes afin d'obtenir les distributions des variables hydrologiques dépendantes. Bernier (1964) a ainsi obtenu la loi de probabilité du débit minimum annuel en utilisant les précipitations et la courbe de tarissement. L'expression de cette loi est formellement celle de la théorie des valeurs extrêmes de Gumbel.

Gumbel (1954) a été le premier à appliquer la théorie des valeurs extrêmes aux événements de sécheresse et en particulier la loi de Weibull au débit minimum annuel. Cette approche ne tient pas compte des durées ni des volumes des déficits. Dans ce cas, on fait l'hypothèse que la distribution initiale des débits est limitée à gauche. La loi de Weibull est à trois paramètres: un paramètre de forme, un paramètre d'échelle, et un paramètre de position ε . L'estimation doit être telle que $\hat{\varepsilon}$ est inférieur au plus faible débit minimum observé.

Matalas (1963) a effectué une étude de comparaison de quatre distributions théoriques asymétriques à trois paramètres, ayant une limite inférieure: la loi de Gumbel, la loi lognormale à trois paramètres (LN3), la loi de Pearson type 3, et la loi de Pearson type 5. Il en a conclu que la loi de Gumbel et la loi de Pearson type 3 s'appliquent mieux aux débits minima annuels de 34 cours d'eau étudiés lorsque la méthode d'estimation est la méthode des moments, et ce, selon deux critères: la limite inférieure estimée ne doit pas être inférieure à la limite observée, et la relation entre le coefficient d'asymétrie et le coefficient d'aplatissement de la distribution théorique doit ressembler à cette même relation présente dans les données observées.

Joseph (1970) a comparé cinq distributions: la loi Gamma, la loi lognormale à deux paramètres (LN2), la loi normale, et la loi de Weibull appliquées au minimum annuel de la moyenne mobile sur 14 jours des débits journaliers et ce, pour 37 stations de mesure dont les données sont supposées indépendantes. L'année hydrologique commence le 1^{er} avril, la méthode d'estimation est la méthode du maximum de vraisemblance (MV). Il en a conclu que la loi Gamma, suivie de la loi de Weibull, et de la loi LN2 sont les plus adéquates. Les critères de comparaison utilisés par cet auteur sont le test d'ajustement du Khi-deux, et la comparaison des relations des coefficients d'asymétrie et d'aplatissement.

Condie et Nix (1975) ont appliqué la loi de Pearson type 3, la loi de Pearson type 5, la loi de Weibull, et la loi LN3 à l'analyse des faibles débits de 38 rivières canadiennes. Les débits journaliers minimaux sont utilisés, et afin d'assurer l'indépendance, l'année débute le 1^{er} juin. Les paramètres sont estimés par la méthode du MV et la méthode des moments. La loi de Weibull est la plus adéquate (estimation par le MV), suivie par la loi de Weibull (estimation par la méthode du plus faible débit observé). La loi est rejetée si la limite inférieure estimée est supérieure au plus faible débit minimum observé. Un avantage de la loi de Weibull est qu'elle peut représenter des données qui ont un coefficient d'asymétrie négatif.

Au Québec, Hoang (1978) a trouvé que la loi log Pearson type 3 (LP3) s'applique le mieux aux minima annuels des débits journaliers pour l'ensemble de toutes les stations de mesure de la province de Québec. Raynal-Villasenor et Douriet-Cardenas (1987) ont recommandé l'usage de la loi GEV pour l'analyse de fréquence des débits minimaux annuels dans le nord-ouest du Mexique.

Tasker (1987) a utilisé la méthode du *bootstrap* (rééchantillonnage dans les distributions empiriques observées) afin d'estimer Q_7 pour les périodes de retour de 10 ans et de 20 ans. Les méthodes basées sur une distribution hypothétique telle que la loi de Weibull ou la loi LP3, donnent une erreur quadratique moyenne plus faible que les méthodes basées sur une distribution empirique.

Nathan et McMahon (1990) ont étudié les méthodes d'estimation des paramètres de la loi de Weibull pour le débit minimum annuel ainsi que pour les débits Q_m , $m = 1, 7, 15, 30, 60, 120, 183, 284$. Ils ont recommandé l'utilisation de plus d'une méthode surtout pour les petits échantillons. Récemment, Adamczyk (1992) a effectué une analyse de fréquence non paramétrique des débits Q_m , où $m = 3, 14, 21, 28$.

Le tableau 2.1 résume les travaux reliés à l'analyse de fréquence des faibles débits. Les lois utilisées sont présentées à l'annexe 2.

Loi	Variable	Référence
Weibull	Minimum annuel	Gumbel, 1954
Gumbel	Minimum annuel	Matalas, 1963
Approche physique (Gumbel)	Minimum annuel	Bernier, 1964
Gamma, Weibull, LN2	Q_{14}	Joseph, 1970
Weibull	Minimum annuel	Condie et Nix, 1975
LP3	Minimum annuel	Hoang, 1978
LP3, Weibull	Q_7	Tasker, 1987
GEV	Minimum annuel	Villasenor et Douriet, 1987
Weibull	$Q_m, m = 10,30,60,90,120$	Nathan et McMahan, 1990
Approche non paramétrique	$Q_m, m = 3,14,21,28$	Adamczyk, 1992

Tableau 2.1: Résumé des lois appliquées aux débits d'étiage

La plupart des études décrites dans cette section utilisent des séries de minima annuels et n'utilisent donc qu'une valeur de débit faible par année. Ainsi, des débits plus faibles que certains minima annuels ne sont pas inclus dans la série car ils correspondent au deuxième débit plus faible de leur année même s'ils sont plus bas que certains minima provenant d'autres années. Afin de pallier ce problème, et pour tenter d'augmenter l'information utilisable dans une série, l'approche de la série partielle est utilisée, où tous les débits en deçà d'un certain seuil, habituellement le maximum des débits minima annuels, sont conservés. Loganathan *et al.* (1986) ont appliqué la loi de Weibull à la série partielle des minima annuels de débit.

Pour certains cours d'eau, les bas débits annuels sont corrélés parce que les niveaux souterrains sont dépendants de la recharge de l'année précédente. Dans ces conditions, la courbe de fréquence des bas débits est une description correcte de la distribution des bas débits annuels, mais la dépendance présente dans les données hydrologiques analysées a un effet sur l'évaluation des périodes de retour et crée un biais qui exagère ces périodes de retour lorsque le coefficient d'autocorrélation d'ordre 1 est supérieur à 0.2. Cette exagération dépend de la structure de corrélation, de la taille de l'échantillon et de l'ordre de grandeur de la dépendance, car la corrélation a un effet sur le tracé de courbe (Beran, 1979).

2.3.2 L'analyse des durées et les processus stochastiques

Toute technique d'analyse de fréquence pour caractériser les bas débits devrait inclure leur durée. Bien que le débit moyen sur une durée soit important, il est souvent nécessaire de déterminer aussi la loi des durées pendant lesquelles le débit demeure faible. Le critère 7 jours 10 ans n'est pas pertinent pour toutes les applications. Dans une région donnée, ce débit peut aller de moins de 0.1% à 15% du débit moyen. Il faut donc utiliser d'autres critères pour tenir compte des durées. Un outil couramment utilisé est la courbe de fréquence cumulative débit-durée (figure 2.4), qui montre le pourcentage de temps durant lequel certains débits sont dépassés pendant la période couverte par l'échantillon, sans toutefois refléter si les périodes de faibles débits sont consécutives ou non. Cette courbe correspond à la rivière identifiée au chapitre 6 comme étant St06. Cette courbe de durée ne peut être considérée comme une courbe de probabilité à cause des dépendances en série et de la variation saisonnière.

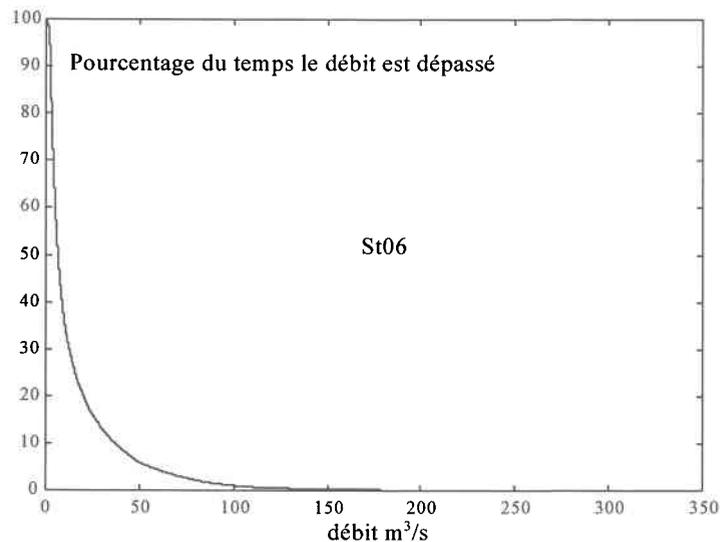


Figure 2.4: Exemple d'une courbe débit-durée

Une autre approche à l'analyse des durées des bas débits est basée sur les processus stochastiques. Dans le cadre des processus stochastiques, on peut parler d'approche directe (analytique) et d'approche indirecte (expérimentale) (Yevjevich *et al.*, 1983). L'approche indirecte (qui n'est pas décrite en détail dans cette revue bibliographique) consiste à obtenir

un modèle des débits, à générer des données synthétiques et à étudier d'une façon empirique certaines variables d'intérêt à partir des données synthétiques générées. L'approche directe consiste à modéliser les événements de déficit ou les variables d'intérêt sans utiliser de modèles de débit. Les processus aléatoires utilisés dans l'approche directe peuvent être classés en deux grandes catégories basées sur une approche par seuil, selon la façon d'introduire la composante aléatoire dans le modèle:

- le concept d'essai aléatoire est assigné à chaque intervalle unitaire de temps Δt , et la modélisation est alors dite par états. Le résultat de l'essai correspond à un des états du processus stochastique $\{X_t\}$ défini à partir des observations hydrologiques $\{Y_t\}$ par le biais de seuils et d'intervalles. Le nombre des états de $\{X_t\}$ est fini et la théorie des séquences est souvent utilisée pour analyser les paramètres de sécheresses lorsque $\{X_t\}$ est à deux états, surplus ou déficit; et
- le concept d'essai aléatoire est assigné à l'occurrence de chaque événement de sécheresse. On applique alors la théorie des processus stochastiques ponctuels, où la sécheresse est considérée comme un événement rare, caractérisé par son temps d'occurrence.

2.3.2.1 La théorie des séquences

Plusieurs définitions d'une séquence ont été proposées, mais la définition de Mood (1940) est la plus utilisée et la plus appropriée: une séquence est une suite d'événements du même type précédée et suivie par des événements différents (Guerrero-Salazar et Yevjevich, 1975). À partir d'un seuil x_0 qui peut être fixe ou variable, une séquence négative (positive) est une séquence d'observations en dessous (au-dessus) du seuil. Un processus à deux états $\{X_t\}$ peut ainsi être défini par une variable de Bernoulli, où l'état 1 correspond au surplus, et l'état 0 au déficit. La figure 2.5 illustre la définition de $\{X_t\}$ à partir de $\{Y_t\}$. Lorsque le processus $\{Y_t\}$ est indépendant, il en résulte que le processus des états $\{X_t\}$ est aussi indépendant. Lorsque $\{Y_t\}$ n'est pas indépendant, on fait souvent l'hypothèse que $\{X_t\}$ est un processus markovien d'ordre 1.

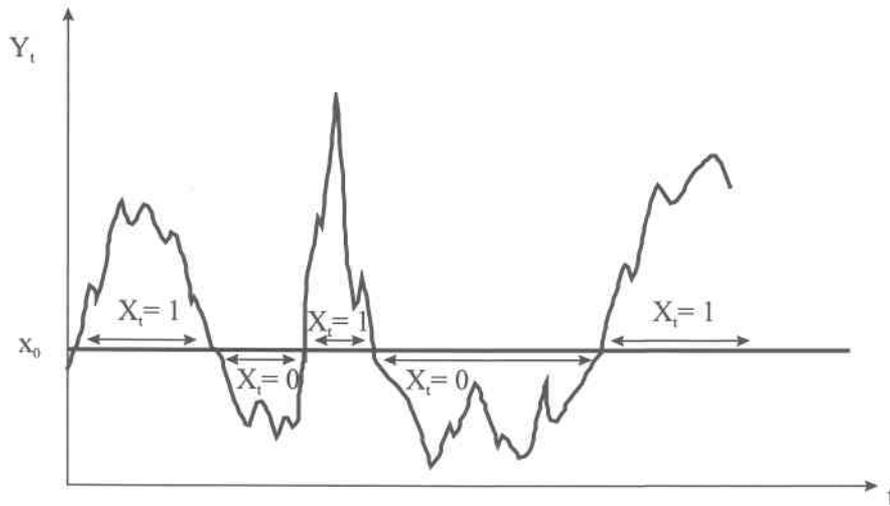


Figure 2.5: La définition de X_t à partir de Y_t

Trois variables sont importantes pour étudier les sécheresses (figure 2.1):

- la durée définie par la longueur d'une séquence négative RL (variable discrète);
- la sévérité qui est la somme des déviations du seuil correspondant à une séquence négative de longueur n , RS_n (variable continue); et
- l'intensité I_n qui est le rapport RS_n/RL .

L'objectif de l'analyse des séquences est de déterminer les distributions des variables RL , RS_n , et I_n et celles de leurs valeurs extrêmes. Le calcul analytique est possible uniquement pour des modèles de débits $\{Y_t\}$ simples stationnaires supposés indépendants ou autorégressifs d'ordre 1. La chaîne de Markov à temps discret et à deux états a souvent été choisie pour représenter le processus $\{X_t\}$ des états de déficit ou de surplus hydrologique, définis à partir de la série continue de débits $\{Y_t\}$ à l'aide d'un seuil. D'après Millan (1973), l'application de la chaîne de Markov discrète à deux états est acceptable pour des débits générés par un processus autorégressif d'ordre 1 qui a un coefficient d'autocorrélation d'ordre 1 inférieur à 0.5.

Mathier *et al.* (1992) ont étudié les débits mensuels en définissant les états de déficit et de surplus à partir d'un seuil qui varie en fonction des mois. Le processus des états est une

chaîne de Markov homogène discrète à deux états, les durées sont distribuées selon une loi géométrique, et la sévérité du déficit suit une loi exponentielle. Les auteurs ont estimé les quantiles et obtenu des intervalles de confiance pour la durée du déficit et pour sa sévérité.

Dracup *et al.* (1980b) ont étudié les caractéristiques statistiques des paramètres d'intensité, de sévérité et de durée. Ils ont examiné la stationnarité, l'autocorrélation, la corrélation entre les paramètres d'un même événement, et la corrélation croisée entre les paramètres de déficit et les paramètres de surplus. Les données annuelles proviennent de sept cours d'eau californiens aux climats semi-aride à humide, et le seuil de dépassement est le débit moyen. Qualitativement, ils ont observé qu'une sévérité qui augmente implique une durée qui diminue, qu'il n'y a pas de non-stationnarité (tendance linéaire) significative dans la sévérité, et que l'intensité et la durée semblent non stationnaires. Ils ont constaté que la durée et l'intensité sont les variables les moins corrélées et la durée et la sévérité le plus. Ces caractéristiques de dépendance offrent un choix de méthodes de modélisation, soit en utilisant l'hypothèse d'indépendance de la durée et de l'intensité, soit en utilisant l'hypothèse de dépendance de la durée et de la sévérité. La première approche fut utilisée par Jackson (1975) qui appliqua une chaîne de Markov à la durée et un modèle autorégressif à l'intensité.

Chang et Stenson (1990) ont montré qu'il existe une forte corrélation entre la durée RL et la sévérité RS_n . Ces résultats sont confirmés par Bonacci (1993) qui a constaté que la corrélation entre la durée et l'intensité est non significative, et que la corrélation entre l'intensité et la sévérité existe, mais qu'elle est quand même plus faible que celle entre la durée et la sévérité.

Deux approches sont utilisées lors de l'application de la théorie des séquences aux problèmes hydrologiques, l'approche d'intégration et l'approche combinatoire. L'approche d'intégration est reliée à des séquences d'une population infinie. L'approche combinatoire traite les séquences dans un échantillon de longueur finie. Les solutions analytiques exactes sont possibles lorsque les processus sont simples, et des approximations sont obtenues lorsque les processus sont complexes. La distribution des durées est facile à obtenir lorsque le processus de débits $\{Y_t\}$ est stationnaire, indépendant ou lorsque la structure de dépendance est simple; l'estimation des durées est non fiable dans le cas non stationnaire, et

on procède alors par simulation. Les durées ne dépendent pas de la distribution de $\{Y_t\}$ lorsque les $\{Y_t\}$ sont indépendants. La distribution des déficits dépend de la distribution de $\{Y_t\}$ et elle n'a pas de solution analytique générale.

2.3.2.1.1 L'approche d'intégration

Llamas et Siddiqui (1969) ont présenté les expressions pour la moyenne, la variance, et la covariance des variables RL , RS_n et I_n pour deux processus d'états $\{X_t\}$: stationnaire indépendant et markovien homogène dont le seuil x_0 est zéro. Ils ont aussi inclus une approximation des distributions de RS_n et de I_n pour un processus indépendant lorsque les débits $\{Y_t\}$ suivent une loi normale, ainsi que les moments de RS_n et de I_n lorsque les $\{Y_t\}$ suivent une loi Gamma.

Saldarriga et Yevjevich (1970) ont présenté les distributions du nombre de séquences, et la probabilité qu'une séquence soit d'une longueur donnée pour un processus de débit indépendant. La contribution principale de leur étude est l'expression générale de $P(k^-, j^+)$ la probabilité des j premières séquences positives et des k premières séquences négatives, sous la forme d'une intégrale multinormale, lorsque la série hydrologique $\{Y_t\}$ est stationnaire, indépendante ou autorégressive normale d'ordre 1 avec un coefficient d'autocorrélation inférieur à 0.4. Cette intégrale n'a pas de solution explicite dans le cas dépendant, mais des approximations ont été obtenues, dans le cas où Y_t et Y_{t+1} suivent une loi normale bivariée, par Chander *et al.* (1981).

Sen (1977) a obtenu, par une approche différente de celle de Downer *et al.* (1967), les deux premiers moments de RS_n pour un processus $\{Y_t\}$ indépendant, normal et lognormal, ainsi que pour un processus $\{Y_t\}$ normal autorégressif d'ordre 1. Il a constaté que la corrélation entre les observations successives augmente la moyenne et la variance des volumes de surplus et de déficit.

2.3.2.1.2 L'approche combinatoire

Les distributions de la plus longue séquence RL_{\max} (la période de sécheresse critique), et du plus gros déficit correspondant RS_{\max} dans un échantillon de longueur n sont difficiles à obtenir. Millan et Yevjevich (1971) ont obtenu les distributions de ces deux paramètres par la simulation de processus de débits $\{Y_t\}$ stationnaires indépendants et autorégressifs

d'ordre 1 normaux et non normaux. Les propriétés de la longueur des séquences des processus stationnaires sont indépendantes de la moyenne et de l'écart type, mais elles dépendent du seuil x_0 et de l'asymétrie de la distribution des débits. Dans le cas des déficits, leur ordre de grandeur est directement proportionnel à l'écart type des débits. La loi lognormale s'applique le mieux aux fréquences empiriques de ces deux paramètres, obtenues par simulation.

Millan (1972) a obtenu la distribution exacte de la plus longue séquence dans un échantillon de taille N , lorsque le processus des états $\{X_t\}$ suit une chaîne de Markov stationnaire.

Sen (1980a) a obtenu par énumération la probabilité de la plus longue durée dans le cas stationnaire, ce qui donne une expression exacte pour un processus d'états $\{X_t\}$ indépendant et une expression approximative pour un processus d'états markovien d'ordre 1. Sen (1980c) a obtenu cette même expression selon une approche basée sur la théorie des extrêmes d'un nombre aléatoire de variables aléatoires. Il a aussi présenté une expression approximative pour la fonction de densité du déficit maximum basé sur les débits annuels, où les durées et les déficits sont supposés indépendants.

Sen (1980b) a obtenu par énumération la distribution exacte de la plus longue durée dans un échantillon de taille N , lorsque le processus de débits $\{Y_t\}$ est périodique, et lorsqu'il correspond à un processus d'états $\{X_t\}$ markovien d'ordre 1 non homogène périodique.

Güven (1983) a présenté une approche relativement simple pour l'estimation des probabilités de la plus longue durée de sécheresse et du maximum de la somme des déficits sur une période de longueur m . Des débits annuels sont utilisés pour définir une chaîne de Markov discrète stationnaire $\{X_t\}$ à deux états. L'auteur a fait l'hypothèse que la somme des déficits suit une loi Gamma à deux paramètres.

Moyé *et al.* (1988) ont appliqué la théorie des séquences à la prédiction des événements de sécheresse pour un processus indépendant. Ils ont obtenu la probabilité d'avoir exactement i séquences négatives de longueur k dans un échantillon de taille N , et pour une probabilité de défaillance q , ils ont obtenu la moyenne du nombre de ces séquences et leur variance, ainsi que l'estimation de la longueur moyenne d'une séquence. Ainsi, il est possible de

déterminer l'espérance du nombre de sécheresses et l'espérance de la durée d'une sécheresse dans les N prochaines années.

Sen (1990) a obtenu par énumération l'expression analytique de la probabilité de la plus longue durée de sécheresse pour un nombre d'années, pour un seuil et pour des probabilités de transition donnés, et ce pour une chaîne de Markov $\{X_i\}$ discrète homogène d'ordre 2 à deux états. La plus longue durée de déficit et le plus grand déficit d'une sécheresse historique pour une probabilité q que la demande dépasse l'offre, dépend de la distribution de Y_i , de la taille de l'échantillon N , et de la variation de l'échantillon des valeurs extrêmes des variables décrivant la sécheresse.

Récemment, Moyé et Kapadia (1995) ont présenté la probabilité que les séquences négatives soient de longueur comprise entre K et L , pour un échantillon de longueur N , et pour une probabilité q de ressources inadéquates. Ils ont obtenu la loi des durées des sécheresses extrêmes sur une échelle annuelle pour un processus stationnaire indépendant.

Les études par l'approche de la théorie des séquences présentées ci-dessus mettent l'accent sur les paramètres de sécheresse extrême, et elles supposent que q , la probabilité d'être dans un état déficitaire est connue. Elles font souvent intervenir des processus stationnaires valables surtout pour des données annuelles, négligeant ainsi la modélisation des données journalières. Le tableau 2.2 présente un résumé des études des paramètres de sécheresse par l'analyse des séquences.

Processus	Variable	Référence
X_t Stationn., indép., Markov 1	Moments de RL, RS_n, I_n (PI)	Llamas et Siddiqui, 1969
Y_t Stationnaire, indép., AR(1)	$P(k, j^+)$ (PI)	Saldarriga et Yevjevich, 1970
Y_t Stationnaire, indép., AR(1)	Moments RS_n (PI)	Sen, 1977
Y_t Stationnaire, indép., AR(1)	RL_{max}, RS_{max} (EF)(sim.)	Millan et Yevjevich, 1971
X_t Stationnaire, Markov 1	RL_{max} (EF)	Millan, 1972
X_t Périodique, Markov 1	RL_{max} (EF)	Sen, 1980a
X_t Stationn., indép, Markov 1	RL_{max}, RS_{max} appr. (EF)	Sen, 1980b
X_t Stationn., indép, Markov 1	RL_{max} (EF)	Sen, 1980c
X_t Stationnaire, Markov 1	RL_{max}, RS_{max} (EF)	Güven, 1983
X_t Stationnaire, indép.	nbr séq. de longueur K (EF)	Moyé <i>et al.</i> , 1988
X_t Stationnaire, Markov 2	RL_{max} (EF)	Sen, 1990
X_t Stationnaire, Markov 1	RL, RS_n (PI)	Mathier <i>et al.</i> , 1992
X_t Stationnaire, indép.	RL (EF)	Moyé et Kapadia., 1995

EF: Échantillon Fini, approche d'intégration

PI: Population Infinie, approche combinatoire

Tableau 2.2: Résumé des travaux d'analyse de séquences

2.3.2.2 Les processus ponctuels

Les processus ponctuels sont basés sur la notion d'occurrence d'un événement. Dans le cas des débits, un événement correspond par exemple, au changement d'état (de surplus à déficit ou vice versa). Un processus ponctuel, obtenu à partir du processus continu de débits et d'un seuil donné, est défini par les temps d'occurrence t_j d'un événement e_j tel que présenté sur la figure 2.6.

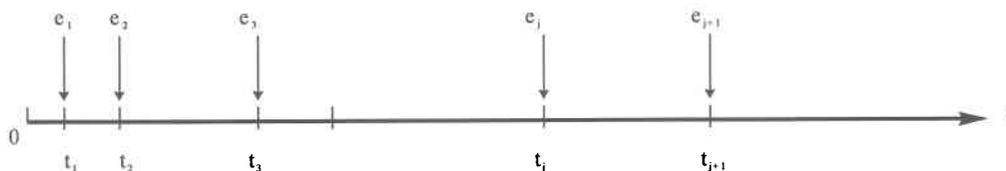
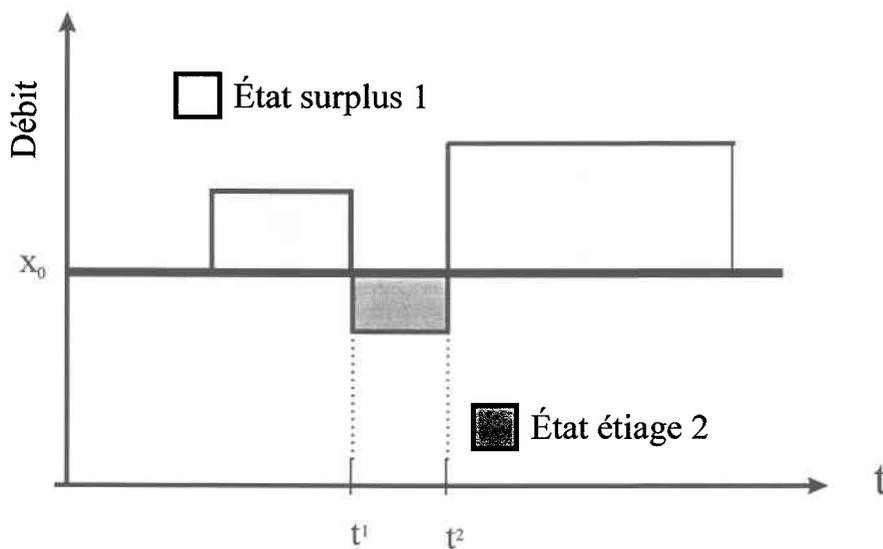


Figure 2.6: Schéma d'un processus ponctuel

Ainsi, le processus ponctuel des déficits est défini par le temps de passage de l'état de surplus à l'état de déficit. L'avantage de cette approche par rapport à l'analyse des séquences est qu'elle permet d'étudier des processus complexes, dépendants et non stationnaires. De plus, elle permet la génération synthétique des temps d'occurrence d'événements. Les processus ponctuels utilisés dans ce travail sont décrits au chapitre 3.

La plupart des méthodes statistiques proposées pour l'analyse des durées sont purement empiriques, et North (1981) a été le premier à effectuer une analyse théorique des durées d'étiage à partir de données journalières. Son approche pour modéliser les étiages sous forme de processus intermittents tient compte de façon continue de l'effet saisonnier et de la dépendance du temps (mémoire). L'étiage est défini comme un débit en dessous d'un seuil fixe x_0 , et les modèles sont à deux états, étiage et surplus. Le processus ponctuel alterné de surplus (type 1) et d'étiage (type 2) est défini par les temps d'occurrence des deux événements, tel qu'illustré sur la figure 2.7. Pour reprendre la notation de North (1981), un événement de type i correspond à la rupture de l'état i .



t^1 : temps d'occurrence de l'événement 1
 t^2 : temps d'occurrence de l'événement 2
 Événement i : rupture de l'état i ; $i = 1, 2$

Figure 2.7: Définition du processus ponctuel alterné de surplus et d'étiage

Deux modèles de ce processus alterné sont proposés par North (1981): le premier correspond à deux processus de Poisson non homogènes sans mémoire obtenus lorsque l'intensité $\lambda^i(t)$ est une fonction périodique, c'est-à-dire lorsque les taux d'occurrence d'événement dépendent uniquement du temps:

$$\lambda^i(t) = \lambda_i \exp\left[K_i \sin(\omega t + \theta_i)\right], \quad \omega = \frac{2\pi}{365}, \quad i = 1, 2 \quad (2.1)$$

Le deuxième modèle est composé de deux processus de renouvellement non homogènes avec mémoire limitée, où l'occurrence d'un événement à l'instant t dépend du temps écoulé depuis le dernier événement qui eut lieu à l'instant u . L'intensité est $\varphi^i(t-u, t)$, où $\lambda^i(t)$ est la fonction périodique de l'équation 2.1, et $\varphi^i(t-u)$ est l'intensité d'un processus de renouvellement homogène:

$$\varphi^i(t-u, t) = \varphi^i(t-u)\lambda^i(t), \quad i = 1, 2 \quad (2.2)$$

Après avoir estimé les fonctions d'intensité pour les deux types de modèles, l'auteur a obtenu la loi des durées dans un état, la loi conjointe de la durée et du temps de déclenchement, ainsi que la loi de la durée sur un intervalle donné. Ces deux modèles sont décrits en détail au chapitre 5 et évalués sur des rivières canadiennes et une rivière française au chapitre 6.

Lee *et al.* (1986) ont effectué une analyse de fréquence des durées des sécheresses multiannuelles à l'aide des processus ponctuels de renouvellement. Une sécheresse a lieu lorsque le débit moyen annuel est inférieur à un seuil donné. Leur but est de trouver la loi de la durée de la sécheresse multiannuelle $f(t)$. Considérant ces événements de sécheresse comme un processus de renouvellement, ils ont utilisé les notions suivantes: t est la durée, $f(t)$ est la fonction de densité de la durée, $R(t)$ est la probabilité que la durée dépasse t , et $h(t)=f(t)/R(t)$ est le taux de rupture. On a alors:

$$R(t) = \exp\left[-\int_0^t h(s)ds\right] = 1 - F(t) \quad (2.3)$$

$$f(t) = h(t) \exp\left[-\int_0^t h(s)ds\right] \quad (2.4)$$

Ils ont cherché à ajuster une fonction $h(t)$ à $P[D = t | d \geq t]$, la probabilité qu'une sécheresse se termine après t années étant donné qu'elle a déjà duré au moins t années (le taux de rupture d'une sécheresse de t années). Ce taux empirique est estimé à partir des données par le nombre de sécheresses qui durent t années divisé par le nombre de sécheresses qui durent plus de t années. Les auteurs ont conclu que la forme la plus adéquate pour représenter le taux de rupture des sécheresses est la fonction logistique: $h(t) = \frac{1}{1 + p \exp(-qt)}$.

Zelenhasic et Salvai (1987) ont analysé le processus intermittent des sécheresses de durée inférieure à un an, à partir des débits journaliers et selon la théorie des valeurs extrêmes d'un nombre aléatoire de variables aléatoires présentées par Todorovic (1970). Un événement de sécheresse est défini pour chaque année à partir des débits journaliers et d'un seuil. Une sécheresse qui débute une année et se termine l'année suivante est associée à la première année. Afin d'assurer l'indépendance des événements, des conditions d'écart au seuil et de durée minimale entre événements sont imposées. Ces auteurs ont appliqué ainsi un processus de Poisson non homogène à l'occurrence des événements de sécheresse. La loi Gamma est alors la loi du temps d'occurrence de la $i^{\text{ème}}$ sécheresse par rapport à un temps d'origine arbitraire t_0 , et la loi de Poisson est la loi du nombre de sécheresses dans un intervalle de temps $[0, t]$. Les auteurs ont aussi déduit la distribution du plus gros déficit sur un intervalle de temps lorsque les volumes de déficit sont indépendants et distribués selon une loi de la famille Gamma. Cette approche est valable pour les 90^{ème} et 95^{ème} quantiles.

Chang (1989) a émis l'hypothèse que les sécheresses définies par le biais d'un seuil ont lieu selon un processus de Poisson non homogène, et il a estimé une fonction d'intensité ayant une forme périodique. L'auteur a constaté que les durées et les déficits sont fortement corrélés pour tous les seuils de dépassement, cette corrélation étant plus forte pour les faibles seuils.

Récemment, Madsen et Rosbjerg (1995) ont utilisé un processus de Poisson homogène pour modéliser trois processus d'occurrence d'étiage. Le premier est défini à partir des débits journaliers et d'un seuil correspondant au 90^{ème} quantile; le deuxième est défini à partir des débits mensuels et d'un seuil variable correspondant au débit mensuel moyen; et le troisième est défini à partir des débits annuels et d'un seuil correspondant au débit annuel médian.

Dans le cas journalier, un événement est défini (comme par Zelenhasic et Salvai (1987)) de façon à assurer l'indépendance de ces événements. La loi géométrique est ajustée aux durées. Le tableau 2.3 résume l'application des processus ponctuels aux étiages et sécheresses.

Processus	Variables	Référence
Poisson alterné, et renouvellement alterné non homogènes	Durée, durée et temps de déclenchement, durée à l'intérieur d'un intervalle	North, 1981
Renouvellement homogène	Durée	Lee <i>et al.</i> , 1986
Poisson non homogène	Nombre de sécheresses, et temps de la $i^{\text{ème}}$ dans $[0,t]$	Zelenhasic et Salvai, 1987
Poisson non homogène	-	Chang, 1989
Poisson homogène	Durée, volume de déficit	Madsen et Rosbjerg , 1995

Tableau 2.3: Résumé des processus ponctuels appliqués aux sécheresses et aux étiages

2.4 Les indicateurs de sécheresse

Plusieurs indicateurs de sécheresse, surtout de type météorologique, ont été développés et utilisés (Yevjevich *et al.*, 1983). Un indicateur de sécheresse est toute variable hydrométéorologique ou toute fonction de ces variables dont les fluctuations peuvent révéler la présence plus ou moins marquée de sécheresse. Un indicateur adéquat sert à résumer et à transmettre l'information concernant le phénomène et permet d'établir une base quantitative pour évaluer l'intensité et la sévérité de la sécheresse. Nous présentons dans les paragraphes qui suivent quelques indicateurs de sécheresse hydrologique.

Les principaux outils utilisés aux États-Unis sont l'indice de sévérité de sécheresse de Palmer (Palmer, 1965), (*Palmer Index*, PI) (long terme) et l'indice d'humidité des cultures (Palmer, 1968), (*Culture Humidity Index*, CMI) (court terme). Le PI combine l'information sur les précipitations et sur la température et mesure la différence cumulative entre les conditions normales et les conditions observées. C'est essentiellement un indice météorologique mais qui est également utilisé comme mesure de la sécheresse

hydrologique. Le PI a été souvent critiqué au cours des années et ne s'applique pas aux régions irriguées et aux régions où le ruissellement dû à la fonte de neige est prédominant.

Riggs (1979) a défini un indicateur de sécheresse hydrologique pour des sécheresses saisonnières ainsi que des sécheresses pluriannuelles, basé sur la période de retour d'une sécheresse. Cet indicateur a une forme régionale qui permet de caractériser plus globalement les sécheresses dans des grandes régions.

Dezman *et al.* (1982) et Garen (1993) ont présenté un indice de sévérité de sécheresse pour les eaux de surface du Colorado, le *Surface Water-Supply Index* (SWSI), qui incorpore plusieurs composantes hydrologiques et climatologiques, et qui est complémentaire au PI. Le SWSI est un indicateur de la disponibilité en eau sur un bassin donné, et il est formulé en termes de la probabilité conjointe de non-dépassement du volume disponible et du débit prédit sur une période de temps. Il prend des valeurs entre -4 et +4, un SWSI égal à -4 étant un indicateur d'une sécheresse extrême. Le PI est valide au Colorado dans les zones indépendantes des eaux des montagnes, tandis que le SWSI est utile dans les régions qui dépendent de la fonte des neiges pour la production d'eau et tient compte des probabilités de non-dépassement du couvert neigeux, des précipitations et des composants des réservoirs.

Kendall et Dracup (1990) ont présenté un indice hydrologique de sécheresse (*Hydrological Drought Index*, HDI) qui est formulé pour tenir compte de la durée et de l'intensité d'une sécheresse. Pour une sécheresse qui dure d années, cet indice correspond à l'écart pondéré entre le volume de déficit pendant d années et l'espérance de ce volume calculée à partir des données historiques.

Un indice simple est l'hydraulicité d'une période donnée qui constitue le pourcentage de débits par rapport à une moyenne à long terme de la même période. On peut ainsi diviser l'hydraulicité en cinq classes: très sèche, sèche, normale, humide, très humide (Matalas, 1991).

L'intensité et la durée peuvent être combinées en un indice en utilisant la longueur des séquences, la somme des séquences, le volume du déficit, et les courbes de masse résiduelles. D'autres indices développés par divers auteurs sont présentés par Dracup (1991),

qui conclut qu'il faudrait développer un nouvel indice général qui serait applicable partout aux États-Unis.

2.5 Les indicateurs d'étiage

Un indicateur d'étiage est un couple débit-durée (Q,D) qui caractérise le processus des débits journaliers, l'un des termes étant fixé et l'autre variant dans le temps et pouvant être probabilisé par l'ajustement d'une loi à ses valeurs observées. Il existe de nombreux indicateurs d'étiage dont ceux présentés au tableau 2.4 parmi lesquels on choisit l'indicateur le plus approprié aux besoins de l'utilisateur.

	Indicateurs	Définitions
Durée fixée	• Débits classés (Q,D)	Débit dépassé D jours dans l'année
	• Débits moyens mensuels (Q,i)	Débit moyen du ième mois
	• Débits minima mensuels (Q,i)	Débit minimum du ième mois
	• Débits moyen mobiles minima (Q,D)	Moyenne mobile des débits de D jours consécutifs la plus faible de l'année
Débit fixé	• Durées classées annuelles (Q,D)	Durée totale dans l'année où le débit est inférieur à Q fixé
	• Durées individualisées (Q,D)	Durée pendant laquelle le débit reste sans interruption inférieur à Q fixé
	• Durées classées mensuelles	Durées classées au niveau de chaque mois

Tableau 2.4: Quelques indicateurs d'étiages (adapté de Miquel et Phien Bou Pha, 1979)

Miquel et Phien Bou Pha (1979) ont appliqué les lois normale, lognormale, et de Gumbel aux indicateurs à durée fixée. Les indicateurs à débit fixés sont moins répandus et ces

auteurs leur ont appliqué la loi de Pearson type 3. La figure 2.8 présente un exemple d'indicateur d'étiage à débit fixé.

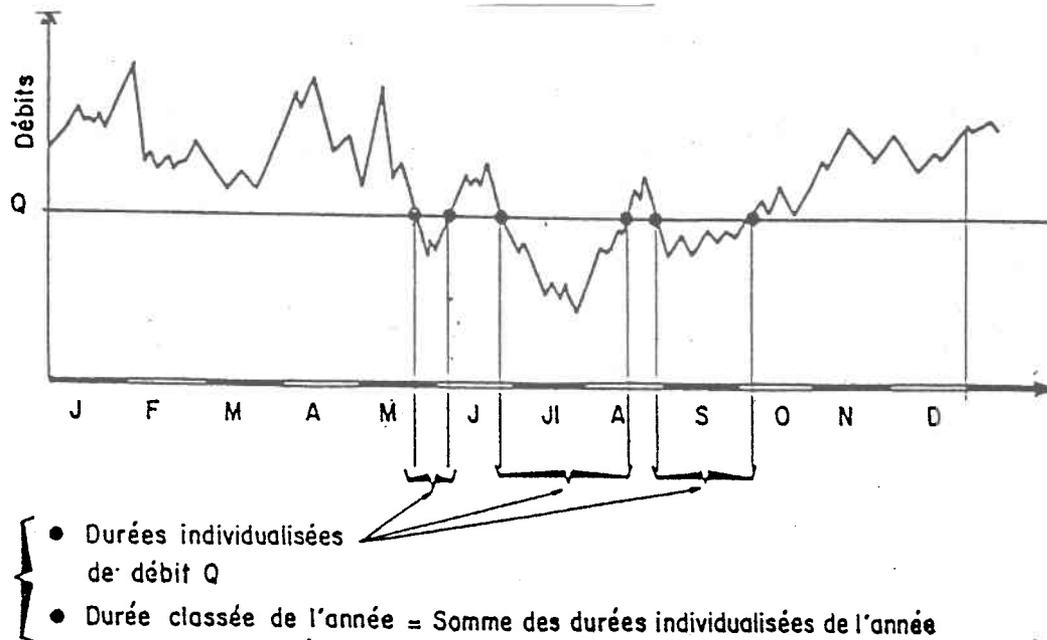


Figure 2.8: Indicateur d'étiage à débit fixé (Miquel et Phien Bou Pha, 1978)

D'autres indicateurs d'étiage sont les suivants (C.T.G.R.E.F., 1979):

- QME: débit du mois d'étiage, i.e. débit moyen du mois dont la moyenne interannuelle est la plus faible, cette notion fixe la date de l'étiage dans l'année;
- QMNA: débit mensuel minimal annuel, i.e. minimum annuel des débits moyens mensuels; ce débit peut correspondre, suivant les années, à des mois différents; et
- QCN_n : seuil minimal en dessous duquel les débits sont restés pendant n jours consécutifs.

2.6 Le risque et l'incertitude

Chacun de nous est assujéti à plusieurs types de risques quotidiennement, à la fois au niveau individuel et comme membre de la société. Certains sont naturels comme par exemple les étiages ou les ouragans, d'autres sont causés par les humains; certains sont hors de notre contrôle et d'autres sont contrôlables.

Il existe une grande confusion dans le domaine des ressources de l'eau concernant la différence entre l'incertitude, le risque, l'imprécision, l'erreur, le processus aléatoire et stochastique. Tous ces aspects sont souvent regroupés sous le terme incertitude.

L'incertitude existe en l'absence d'information complète concernant le passé, le présent et le futur. L'imprécision fait appel à la notion de "vague" où l'affirmation n'est pas en termes exacts. L'erreur est une déviation d'une valeur correcte. Pour bien illustrer ces définitions, considérons l'exemple suivant: supposons que la température à l'extérieur soit 25°C. L'énoncé "il fait 25°C à l'extérieur" ne contient ni imprécision, ni incertitude ni erreur; l'énoncé "il fait entre 24°C et 26°C à l'extérieur" contient de l'imprécision mais pas d'incertitude ni d'erreur; l'énoncé "je suis sûre à 75% qu'il fait 25°C à l'extérieur" contient de l'incertitude mais pas d'imprécision ni d'erreur. Ces notions sont résumées au tableau 2.5.

Énoncé dans le cas où il fait 25° C	Incetitude	Imprécision	Erreur
Il fait 25°C	Non	Non	Non
Il fait 22°C	Non	Non	Oui
Je suis sûre à 75% qu'il fait 22°C	Oui	Non	Oui
Il fait entre 24°C et 26°C	Non	Oui	Non
Je suis sûre à 75% qu'il fait 25°C	Oui	Non	Non
Je suis sûre à 75% qu'il fait entre 24°C et 26°C	Oui	Oui	Non
Il fait entre 22°C et 23°C	Non	Oui	Oui
Je suis sûre à 75% qu'il fait entre 22°C et 23°C	Oui	Oui	Oui

Tableau 2.5: Incertitude, imprécision et erreur

L'aléatoire est une variation temporelle et/ou spatiale imprévisible. Un processus est aléatoire si, étant donnée l'information complète concernant le passé, il n'est pas possible de prédire exactement le futur (Tumeo, 1993). Si un phénomène est aléatoire et complètement imprévisible, il est chaotique. Par contre, si le phénomène est aléatoire de façon à ce qu'on puisse identifier les probabilités de différents événements futurs, il est stochastique. La stochasticité est un sous-ensemble de l'aléa. Ainsi, l'incertitude peut découler de l'aléatoire, du stochastique, de l'erreur, de l'imprécision, mais n'est pas synonyme d'aucun de ces termes.

2.6.1 Les définitions de l'incertitude

L'incertitude est le résultat de tout facteur causant la connaissance non parfaite ou partielle; c'est une condition nécessaire mais non suffisante au risque. Plusieurs types d'incertitude concernant un système peuvent être identifiés, dont un ou plusieurs seront prédominants selon le problème étudié: les objectifs pouvant changer à long terme, les erreurs d'estimation de paramètres, l'impossibilité de prédire le futur, la variation naturelle du phénomène, la difficulté de quantifier certains objectifs, l'incertitude reliée au choix du modèle, et l'incertitude qui découle de notre ignorance.

D'une façon générale, on peut classer l'incertitude dans le domaine des ressources en eau, en trois grandes classes (Chow, 1979; Duckstein, 1979; Yen, 1989;): hydrologique, technologique et stratégique. L'incertitude hydrologique inclut l'incertitude naturelle (incertitude de base, incertitude géophysique) reliée à la nature aléatoire du phénomène, l'incertitude du modèle associée avec le cas où le modèle ne représente pas nécessairement la réalité, l'incertitude informationnelle qui comprend deux volets: l'incertitude des paramètres (échantillonnage), et l'incertitude économique reliée aux coûts réels, aux bénéfices, à la démographie, à la demande, et à d'autres considérations économiques. L'incertitude technologique est reliée aux difficultés de prédire l'état d'avancement futur des sciences et technologies incluant les aspects techniques, structurels, et opérationnels. Elle est due au futur incertain de nos calculs hydrologiques, géologiques et économiques. L'incertitude stratégique inclut les considérations sociologiques, institutionnelles, écologiques, et environnementales. Elle est due à notre incapacité de prédire les états futurs du monde, par exemple, les attitudes culturelles face à l'écologie, l'incertitude politique et institutionnelle. Ces notions sont classées au tableau 2.6.

Incertitudes hydrologiques	Incertitudes technologiques	Incertitudes stratégiques
Naturelle, de mesure, du modèle, des paramètres, économique	État futur des sciences et technologies	Considérations sociologiques, institutionnelles, écologiques, environnementales

Tableau 2.6: Résumé des incertitudes

2.6.2 Les définitions du risque

Le risque a deux composantes principales: L'existence possible de conséquences indésirables et l'incertitude quant à l'occurrence de ces conséquences exprimée sous forme de probabilité d'occurrence (Colson, 1985). Une décision est risquée parce que le résultat est incertain. Plusieurs définitions du risque existent. Classiquement, il a été défini comme la probabilité de dépassement d'un critère de conception ou d'un indice de performance. Toutefois, nous retenons la définition de Kaplan et Garrick (1981):

$$\text{Risque} = \text{Incertitude} + \text{Dommage} \quad (2.5)$$

Le risque est alors une mesure de l'incertitude et de la sévérité des effets négatifs. C'est à la fois la probabilité et la conséquence.

D'une façon générale, la différence entre le risque et l'incertitude est exprimée ainsi: le risque inclut les situations où les résultats possibles peuvent être décrits par des distributions de probabilité, tandis que l'incertitude englobe les situations où ces distributions ne peuvent pas être élaborées. Cette définition est celle qu'on retrouve souvent en économie, où le terme risque sous-entend une probabilité d'occurrence inférieure à 1, pour laquelle il existe une estimation numérique, et l'incertitude sous-entend une probabilité d'occurrence inférieure à 1 pour laquelle on n'a pas d'estimation. Ceci implique que le risque peut être incorporé dans la modélisation mathématique des objectifs et du comportement du système, tandis que l'incertitude est plutôt incorporée dans les modèles de jugement (interprétation, raisonnement). Plusieurs statisticiens affirment que la différence entre le risque et l'incertitude est au niveau d'événements qu'on peut répéter (risque) et ceux qu'on ne peut pas

répéter (incertitude) (Moore, 1983). Dans un contexte bayésien de probabilité subjective, cette différence ne tient plus.

2.7 L'analyse du risque

L'analyse du risque est un outil d'aide à la décision en avenir incertain. Traditionnellement, l'analyse du risque est l'analyse de la probabilité de défaillance, où une défaillance survient lorsque la performance d'un système n'est plus satisfaisante. Plus généralement, c'est le processus qui établit une relation entre une action ou une situation et son impact, c'est une alternative aux normes et standards traditionnellement utilisés. Le but de cette approche n'est pas de savoir si un nombre critique a été dépassé, mais plutôt de découvrir quelles sont les sources de risque et où se trouvent les plus grandes incertitudes. L'analyse du risque et sa quantification peuvent fournir des données d'entrée à un problème de décision qui inclut non seulement des risques mais aussi d'autres formes de coûts et de bénéfices.

L'analyse du risque est composée de deux volets principaux: le premier volet est l'évaluation du risque qui consiste en l'identification du risque et son estimation. Le deuxième volet est la gestion du risque qui comprend le contrôle du risque et son acceptation.

L'évaluation du risque implique l'identification de la source de péril, l'évaluation de la source, l'évaluation de l'exposition à la source, l'évaluation des effets, et la politique d'évaluation du risque. La gestion du risque tient compte des politiques publiques, des mécanismes de réaction institutionnels ainsi que des solutions techniques conçues pour réduire les risques à la sécurité, les risques économiques, et les risques à la santé. Lors de l'évaluation du risque, l'analyste tente de répondre aux questions suivantes: quel événement négatif peut survenir? Quelle est la probabilité qu'il survienne? Et, quelles en sont les conséquences? Les réponses à ces questions aident l'analyste du risque à identifier, mesurer, quantifier et évaluer les risques, leurs conséquences et leurs impacts (Haimès, 1991). La gestion du risque inclut la caractérisation du risque, l'évaluation du coût, l'évaluation coût-bénéfice ainsi que la considération de valeurs subjectives inhérentes à la perception du risque, les contraintes légales et sociales, ainsi que la réalité politique (Diewald, 1989). La gestion du risque tente de répondre aux questions suivantes: qu'est-ce qui peut être fait? Quelles sont les options disponibles et les compromis associés en terme de coûts, risques et

bénéfices? Et quels sont les impacts des décisions actuelles sur les options futures? (Haimès, 1991). La figure 2.9 décrit les deux volets de l'analyse du risque.

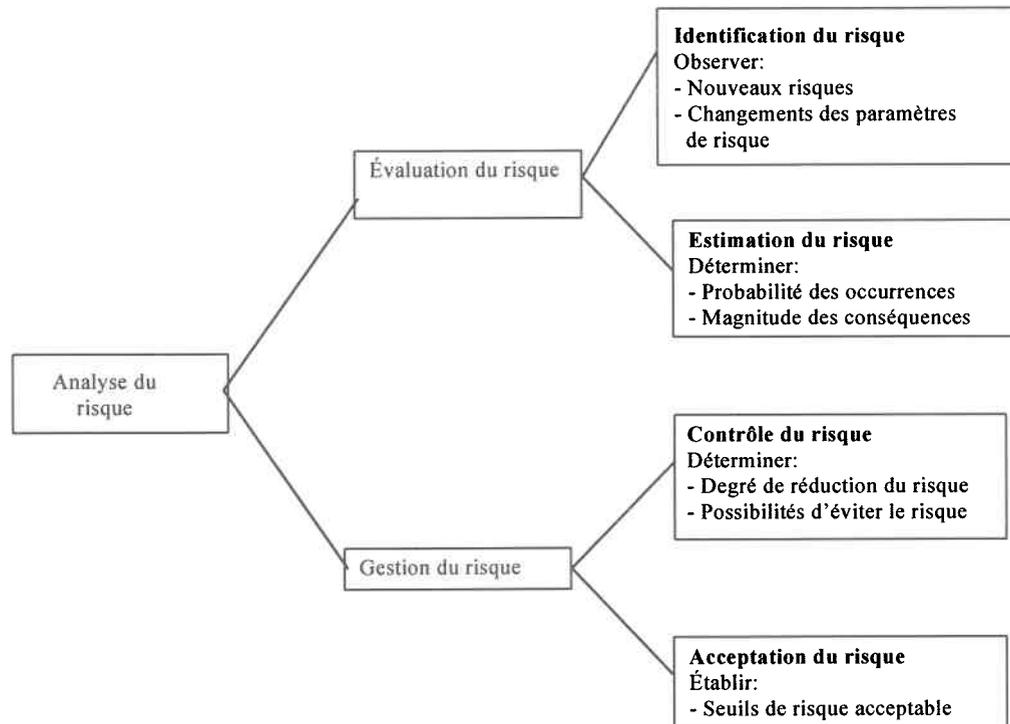


Figure 2.9: Analyse du risque (adaptée de Rowe, 1988)

2.7.1 La modélisation de l'incertitude

Dans un contexte décisionnel, il est important de modéliser l'incertitude liée aux états de la nature avant d'appliquer les diverses méthodes d'optimisation. Les approches principales à ce problème sont l'approche probabiliste (French, 1988), l'approche de Dempster-Shafer (Dempster, 1968) et l'approche des ensembles flous (Bellman et Zadeh, 1970). Dans l'approche probabiliste, l'incertitude modélisée par une distribution de probabilité, est quantifiée par une valeur ponctuelle (nombre net). Dans la théorie de Dempster-Shafer, l'incertitude est un degré de croyance associé à un intervalle. Dans la théorie des ensembles flous, l'incertitude est un degré d'appartenance associé à un ensemble. Une autre approche liée à l'incertitude des données incomplètes est l'approche des ensembles grossiers (*rough sets*, Szladow and Ziarko, 1993), un outil d'analyse de données ambiguës et imprécises, qui

est souvent confondu avec les ensembles flous. L'approche considérée dans ce travail est l'approche probabiliste.

2.7.1.1 L'approche probabiliste

Il existe plusieurs interprétations d'une probabilité, mais les deux principales sont la probabilité fréquentielle (objective) et la probabilité subjective. L'approche fréquentielle préconise qu'une probabilité ne peut avoir de sens que dans le contexte d'une expérience qui peut être répétée. La probabilité d'un événement est alors sa fréquence d'occurrence relative.

L'approche subjective associe la probabilité avec la personne qui observe un système et non avec le système lui-même. La probabilité représente alors le degré de croyance de l'observateur que le système adoptera un certain état. L'approche bayésienne, qui consiste à réviser des probabilités a priori pour formuler des probabilités a posteriori en utilisant le théorème de Bayes et de l'information additionnelle est généralement associée avec la probabilité subjective. L'effet de l'incertitude peut être étudié par une analyse de sensibilité où on fait varier les états de la nature pour ensuite examiner l'impact de ces paramètres sur la décision finale prise. Cette approche est adoptée dans ce travail afin d'analyser les effets de l'incertitude sur l'adéquation des modèles d'occurrence d'étiage et de surplus.

2.7.2 La quantification du risque

Une des étapes de l'analyse du risque est la quantification du risque par des mesures. Le risque hydrologique a été défini classiquement comme la probabilité d'occurrence d'événements naturels extrêmes (Bobée *et al.*, 1991). Toutefois, cette mesure n'est pas suffisante pour tenir compte des diverses facettes du risque. D'autres mesures de performance telles que la fiabilité, la résilience, et la vulnérabilité peuvent être utilisés explicitement comme objectifs dans un processus de prise de décision multicritère. La fiabilité est le complément du risque, la résilience est une estimation du temps moyen au bout duquel le système retourne à un état satisfaisant, et la vulnérabilité mesure l'importance moyenne d'un incident en termes des conséquences. Nous présentons des indices de performance provenant dans les travaux de Abi-Zeid et Bobée (1994); Parent *et al.*, 1991; Duckstein *et al.* 1991; et Hashimoto *et al.*, 1982. Correia *et al.* (1991) ont appliqué certains de ces indices au risque de sécheresse. Nous résumons leurs travaux dans le paragraphe 2.7.2.1

Quelques indices reliés au risque et à la fiabilité (Duckstein *et al.* 1991)

- la fonction de survie $R(t)$: $P\{\text{système dans un état satisfaisant pendant } (0,t) | \text{ système dans un état satisfaisant à } t=0\}$;
- la disponibilité ponctuelle $PA(t)$: cette quantité correspond à ce qui est communément appelé la fiabilité: $P\{\text{système dans un état satisfaisant au temps } t\}$;
- la fiabilité d'intervalle $IR(t, t+\theta)$: $P\{\text{système dans un état satisfaisant pendant } (t, t+\theta)\}$;
- la disponibilité moyenne $A_{\text{moy}}(t)$: la fraction de temps moyenne que le système sera dans un état satisfaisant:

$$A_{\text{moy}}(t) = \frac{1}{t} \int_0^t PA(x) dx \quad (2.6)$$

- la disponibilité conjointe $JA(t, t+\theta)$: $P\{\text{système dans un état satisfaisant à } t \text{ et à } t+\theta\}$;
- le nombre moyen d'incidents dans (t^*, t) ; et
- la fiabilité de service $S(x, t)$: $P\{\text{système dans un état satisfaisant à } t, \text{ ou de le devenir entre } t \text{ et } t+x\}$.

Quelques indices reliés à la résilience

- le temps moyen entre les incidents;
- la durée moyenne de séjour dans un état non satisfaisant; et
- la durée moyenne de séjour dans un état satisfaisant.

Quelques indices reliés à la vulnérabilité

Hashimoto (1980) a défini les indices de vulnérabilité suivants:

- l'espérance de vulnérabilité maximum, V_1 :

$$V_1 = \sum_i s^i \varepsilon^i \quad (2.7)$$

- l'espérance de vulnérabilité moyenne, V_2 :

$$V_2 = \frac{\sum_i s^i \pi^i}{\sum_i \pi^i} \quad (2.8)$$

où s^i correspond aux impacts de chaque état non satisfaisant i , ε^i est la probabilité que s^i soit le pire impact dans l'état i , et π^i est la probabilité que le système soit dans l'état non satisfaisant i .

Dans le cas d'un système représenté par une chaîne de Markov discrète stationnaire à deux états, Hashimoto (1980) a déterminé les expressions de vulnérabilité V_1 et V_2 , de la résilience comme étant l'inverse du temps moyen de séjour dans un état non satisfaisant, et de la fiabilité comme étant la probabilité d'être dans un état satisfaisant.

2.7.2.1 Indices de performance reliés aux sécheresses régionales

Correia *et al.* (1991) ont défini des indices de risque, de vulnérabilité et de résilience reliés aux sécheresses hydrologiques dans un contexte local ainsi que régional, et ce en utilisant des débits moyens annuels obtenus par génération de données synthétiques. Nous résumons leur travail dans ce qui suit. Deux quantités sont importantes: le seuil critique de débit $f(t)$, et la surface critique a_c . On dit qu'il y a sécheresse à une station k si le débit est inférieur au seuil critique $f(t)$. $K(t) = \{k \mid X_k(t) \leq f(t)\}$ est l'ensemble de ces stations.

La surface de déficit instantanée $A_s(t)$ est la surface affectée par la sécheresse. C'est la somme des surfaces correspondant aux stations affectées. Pour une région donnée, ayant p stations de mesure:

$$A_s(t) = \sum_{k=1}^p a_k I_{K(t)}(k) \quad (2.9)$$

où a_k est la surface associée à la station k et

$$I_{K(t)}(k) = \begin{cases} 1 & \text{si le débit de la station } k \text{ est inférieur au seuil critique } f(t) \\ 0 & \text{autrement} \end{cases}$$

$D_s(t)$, le déficit de surface, correspond à la somme pondérée des déficits associés aux stations déficitaires, un déficit étant la différence entre le débit annuel $X_k(t)$ et le seuil critique $f_k(t)$:

$$D_s(t) = \sum_{k=1}^p a_k I_{K(t)}(k) [f_k(t) - X_k(t)] \quad (2.10)$$

La durée d'une sécheresse L correspond à la durée pendant laquelle la surface de déficit instantanée est supérieure à a_c , une surface critique donnée:

$$L = t_e - t_0 + 1, \text{ telle que } A_s(t_0) \geq a_c, A_s(t_0 - 1) < a_c, A_s(t_e) \geq a_c, \text{ et } A_s(t_e + 1) < a_c \quad (2.11)$$

D, le déficit de surface total, correspond à la somme des $D_s(t)$ intégrée sur la durée de la sécheresse L, et l'intensité de sécheresse I est le rapport D/L . Les auteurs calculent les indices décrits dans ce qui suit.

Indice de risque

Les auteurs ont adopté le risque hydrologique défini par la probabilité que le débit soit inférieur au seuil critique. Ils ont quantifié le risque d'une sécheresse par la courbe de fréquence des débits en dessous des seuils critiques pour diverses surfaces critiques.

Indice de résilience

Un excédent $e(t)$ est la différence entre un débit supérieur au seuil critique de débit $f(t)$ et un niveau de récupération $r(t)$. L'excédent cumulé $E(t')$, calculé à partir de t_e , est la somme des excédents $e(t)$ dans $[t_e, t']$. La résilience est définie comme étant le temps de récupération t_r :

$$t_r = \text{MIN}(t - t_e) \text{ telle que } E(t)/D \geq \alpha_r \quad (2.12)$$

où α_r est un taux de récupération fixe représentant la fraction du déficit total qui devra être comblée. Pour α_r , a_c , $f(t)$ et $r(t)$ donnés, il est alors possible d'étudier les propriétés statistiques de la résilience t_r et les auteurs ont présenté les distributions empiriques de t_r .

Indice de vulnérabilité

Pour $f(t)$ et a_c donnés, la vulnérabilité peut être exprimée soit en terme du déficit D, soit en terme de l'intensité I et les auteurs ont présenté les distributions empiriques de ces quantités.

2.8 En résumé

Nous avons présenté dans cette revue bibliographique, avec différents degrés de détail, les points suivants:

- les définitions des sécheresses et les définitions des étiage (sections 2.1 et 2.2);
- les techniques de modélisation des sécheresses et des étiages incluant l'analyse fréquentielle, la théorie des séquences et les processus ponctuels (section 2.3);
- les indicateurs utilisés pour décrire quantitativement les sécheresses et les étiages (sections 2.4 et 2.5);
- les définitions du risque et de l'incertitude (section 2.6); et
- l'analyse du risque incluant la modélisation de l'incertitude, la quantification du risque par des indices ainsi qu'une application de ces indices aux sécheresses (section 2.7).

3. LES PROCESSUS PONCTUELS

Nous présentons dans ce chapitre quelques notions de base reliées aux processus stochastiques ponctuels ainsi que des tests relatifs à l'analyse des séries d'événements. Ces notions sont nécessaires à la bonne compréhension des résultats obtenus aux chapitres 4, 5, 6 et 7. Les références principales utilisées dans ce chapitre sont: Cox et Lewis (1966), Snyder (1975), et Cox et Isham (1980). Nous commençons par définir à la section 3.1 les processus stochastiques, les processus ponctuels, la fonction d'intensité complète, la fonction d'intensité conditionnelle, la probabilité de survie, la densité conjointe des temps d'occurrence, la densité de l'échantillon, et la densité conditionnelle des temps d'occurrence. Nous incluons ensuite à la section 3.2 des définitions reliées aux processus de Poisson homogène, de Poisson non homogène, de renouvellement homogène, doublement stochastique et autoexcité. Nous définissons à la section 3.3 la stationnarité et l'homogénéisation de processus. Finalement, nous décrivons à la section 3.4 des tests statistiques appliqués aux séries d'événements.

3.1 Définitions générales

Un processus stochastique est une collection de variables aléatoires $\{Z_t\}$, où le paramètre t est défini sur un espace paramétrique, T . Le paramètre t représente généralement le temps¹; il est défini sur l'échelle des réels non négatifs. Un processus ponctuel PP est un processus stochastique dont la réalisation est une collection de points, chaque événement (point) e_j ayant une position t_j bien définie dans le temps tel qu'illustré sur la figure 3.1. On dénote par H_t l'historique du processus au temps t , c'est-à-dire les positions de tous les points dans l'intervalle $(0,t]$. Nous utiliserons la notation A pour désigner un intervalle $(t',t'']$.

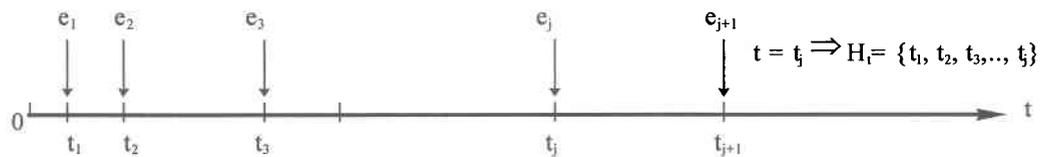


Figure 3.1: La réalisation d'un processus stochastique ponctuel

¹ Pour les fins de cette étude, le paramètre t représentera le temps.

Soient T_1, T_2, \dots, T_j , les variables aléatoires correspondant aux temps d'occurrence d'un événement. Les intervalles entre les événements sont dénotés par X_1, X_2, \dots, X_j où $X_j = T_j - T_{j-1}$, $j > 1$. L'intervalle entre l'origine et le premier événement est X_1 (figure 3.2).



Figure 3.2: La définition des intervalles dans un processus stochastique ponctuel

N est une mesure de dénombrement aléatoire et N_t est un entier non négatif représentant le nombre d'événements dans l'intervalle $(0, t]$ tel qu'illustré sur la figure 3.3. $N(A)$ est le nombre d'événements dans l'intervalle A . Si $A = \bigcup_r A_r$, et si les intervalles A_r sont disjoints, alors $N(A) = \sum_r N(A_r)$.

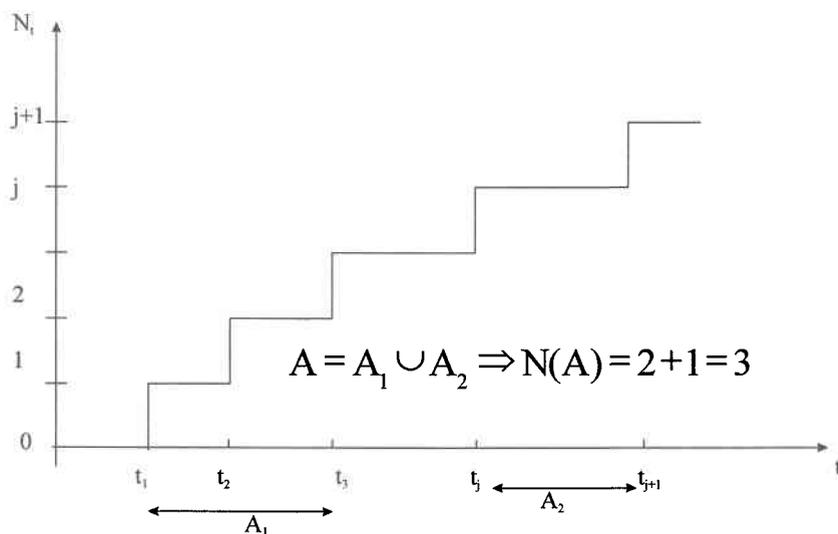


Figure 3.3: Le nombre d'événements dans une union d'intervalles

Le lien entre le nombre d'événements et les intervalles s'exprime par la relation suivante:

$$P(N_s > r) = P(X_{r+1} \leq s) \tag{3.1}$$

Ainsi, nous ne retrouvons plus de r événements dans l'intervalle $(0, s]$ que si et seulement si le $(r+1)^{\text{ème}}$ événement a lieu au plus tard au temps s (figure 3.4):

$$N_s < r \Leftrightarrow X_1 + X_2 + \dots + X_r > s \quad (3.2)$$

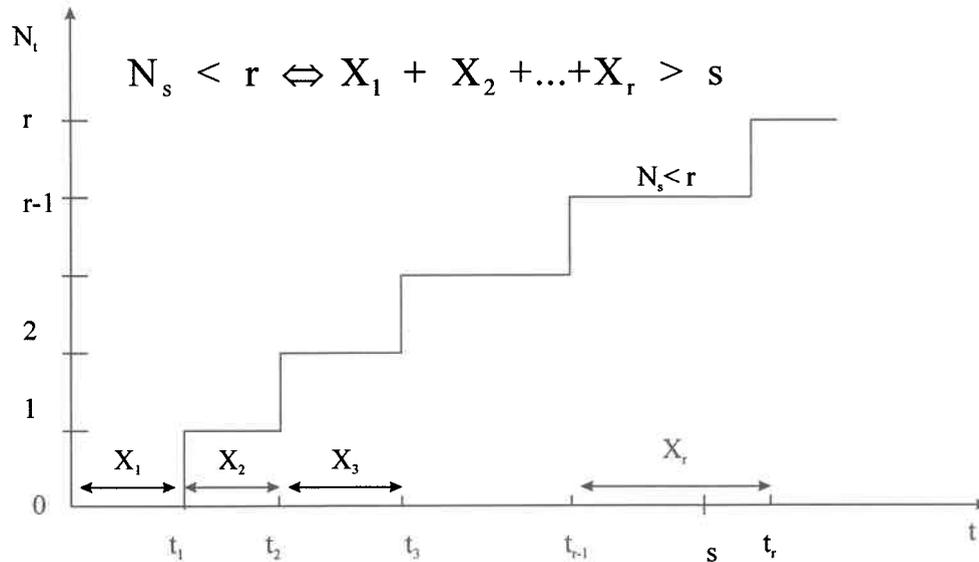


Figure 3.4: Le lien entre le nombre d'événements et les intervalles

Un processus ponctuel, défini sur l'ensemble des réels R^+ non négatifs, et ne permettant pas l'occurrence d'événements multiples simultanés, peut être spécifié mathématiquement de trois façons:

- en fonction de la distribution conjointe du nombre des points $N(A)$ sur un intervalle A ;
- en fonction de la distribution conjointe des intervalles $\{X_j\}$ entre les points successifs; ou
- par le biais d'une fonction d'intensité complète, $\rho(t; H_t)$.

La frontière entre les processus stochastiques ponctuels et d'autres types de processus stochastiques, tel que le processus de Markov n'est pas bien délimitée. Tout processus stochastique ayant un nombre fini d'états peut être associé à un processus ponctuel, où un événement de type i peut correspondre à la sortie de l'état i . Nous présentons dans les

paragraphes qui suivent, certaines quantités importantes pour étudier les processus ponctuels.

Définition 1: la fonction d'intensité complète

La fonction d'intensité complète $\rho(t;H_t)$ est définie par:

$$\rho(t;H_t) = \lim_{\Delta t \rightarrow 0^+} \frac{P(N_{t+\Delta t} = 1 | H_t)}{\Delta t} \quad (3.3)$$

Cette quantité correspond au taux moyen instantané d'occurrence d'un événement. En d'autres mots, $\rho(t;H_t)\Delta t$ est la probabilité instantanée d'un événement dans l'intervalle $(t,t+\Delta t)$. Une fonction d'intensité typique est présentée sur la figure 3.5.

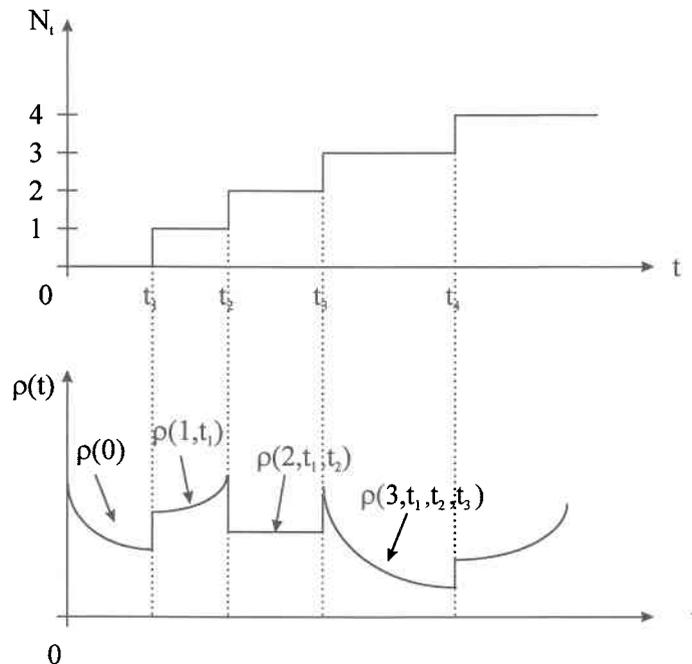


Figure 3.5: Une fonction d'intensité typique (adaptée de Snyder, 1975)

Définition 2: la fonction d'intensité conditionnelle

Dans certaines situations, il est utile d'examiner la probabilité instantanée d'un événement dans l'intervalle $(t,t+\Delta t]$, étant donné qu'un événement a eu lieu à l'origine. Cette probabilité est la fonction d'intensité conditionnelle $h(t)$:

$$h(t) = \lim_{\delta, \Delta t \rightarrow 0^+} \frac{P(N_{t, t+\Delta t} > 0 | N_{-\delta, 0} > 0)}{\Delta t} \quad (3.4)$$

Définition 3: la probabilité de survie

La probabilité de survie $P(X_{n+1} \geq x | t_1, t_2, \dots, t_n)$, illustrée sur la figure 3.6, est la probabilité conditionnelle que X_{n+1} , l'intervalle entre le $n^{\text{ème}}$ et le $n+1^{\text{ème}}$ événement, soit plus grand ou égal à x étant donnés t_j , les temps d'occurrence des événements $j, j=1, 2, \dots, n$.

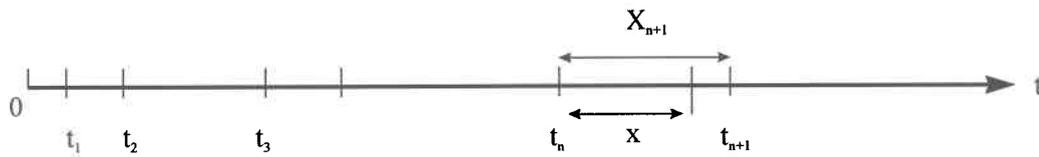


Figure 3.6: Illustration des variables liées à la probabilité de survie

Définition 4: la densité conjointe des temps d'occurrence

$P_{\vec{t}}^{(n)}(\vec{t}) = P_{t_1, t_2, \dots, t_n}^{(n)}(t_1, t_2, \dots, t_n)$ est la fonction de densité conjointe des n premiers temps d'occurrence des événements donnés par $\vec{t} = (t_1, t_2, \dots, t_n)$.

Définition 5: la densité de l'échantillon

La densité de l'échantillon $P(N_\sigma; 0 \leq \sigma < t)$ détermine la probabilité d'obtenir une réalisation particulière d'un processus de n points sur $[0, t)$, situés à $T_j = t_j, j=1, \dots, n$:

$$P(N_\sigma; 0 \leq \sigma < t) \cong P_{\vec{t}}(\vec{t}, N_t = n) \quad (3.5)$$

où:

$$P_{\vec{t}}(\vec{t}, N_t = n) \cong P(N_t = n | T_1 = t_1, T_2 = t_2, \dots, T_n = t_n) P_{\vec{t}}^{(n)}(\vec{t}) \quad (3.6)$$

Cette quantité joue un rôle important pour les problèmes d'inférence. Le symbole \cong signifie "est approximativement égal à". La figure 3.7 présente un exemple d'une réalisation d'un processus ponctuel.

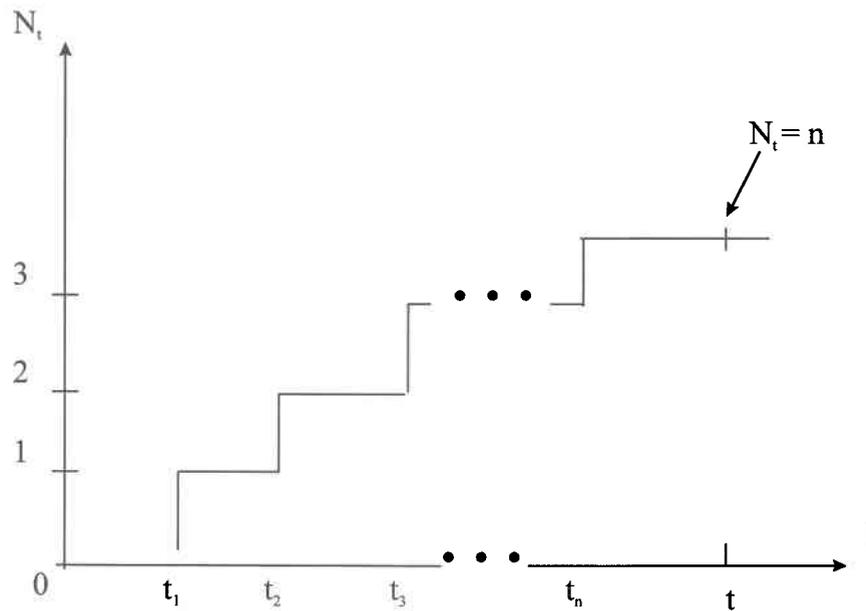


Figure 3.7: Un exemple d'une réalisation d'un processus ponctuel (adaptée de Snyder, 1975)

Définition 6: la densité conditionnelle des temps d'occurrence

La densité conditionnelle des temps d'occurrence $P_{\bar{t}}^{(n)}(\bar{t}|N_t = n)$ est définie par:

$$P_{\bar{t}}^{(n)}(\bar{t}|N_t = n) = \frac{P(N_{\sigma}; 0 \leq \sigma < t)}{P(N_t = n)} \quad (3.7)$$

3.2 Définitions de quelques processus ponctuels

Un processus ponctuel PP est spécifié selon deux approches: soit à l'aide du nombre de points N_t dans un intervalle de temps $(0, t]$, soit à l'aide des intervalles X_j entre les points. Pour un type de processus donné, une approche est habituellement plus simple que l'autre. Nous présentons ci-dessous quelques processus ponctuels utiles dans la suite de ce travail: le processus de Poisson homogène (PH), le processus de Poisson non homogène (PNH), le processus de renouvellement homogène (RH), le processus de Poisson doublement stochastique (PDS), et le processus autoexcité (AE).

3.2.1 Le processus de Poisson homogène

Le processus de Poisson homogène (PH) est un choix classique en raison de sa simplicité. Il est caractérisé par une fonction d'intensité constante $\rho(t;H_t)=h(t)=\rho$ et par les hypothèses de stationnarité, de non-multiplicité, et d'indépendance suivantes:

- la stationnarité implique que la probabilité d'occurrence d'un événement est proportionnelle à la longueur d'un petit intervalle et indépendante du temps d'occurrence de l'événement lui-même;
- la non-multiplicité implique que la probabilité de deux ou plusieurs événements dans un petit intervalle $(t, t+\Delta t]$ est négligeable;
- l'indépendance implique que le nombre d'événements dans un intervalle donné est indépendant du nombre d'événements dans d'autres intervalles disjoints. Ainsi, les probabilités ne dépendent pas de H_t , c'est-à-dire que la probabilité de trouver un point dans l'intervalle $(t, t+\Delta t]$ ne dépend pas du nombre de points avant t ni du fait qu'il y ait eu ou non un point à t .

Un processus ponctuel PP est un processus de Poisson homogène de taux ρ constant si les conditions suivantes sont vérifiées, $\forall t, \Delta t \rightarrow 0^+$:

$$\begin{aligned} P\{N_{t,t+\Delta t}=1|H_t\} &= \rho\Delta t + o(\Delta t) \\ P\{N_{t,t+\Delta t}>1|H_t\} &= o(\Delta t) \\ \text{donc } P\{N_{t,t+\Delta t}=0|H_t\} &= 1 - \rho\Delta t + o(\Delta t) \end{aligned} \quad (3.8)$$

Le nombre de points dans un intervalle de longueur t est distribué selon une loi de Poisson de moyenne ρt . Les intervalles X_1, X_2, \dots, X_j entre les points successifs sont indépendants et distribués selon une loi exponentielle à paramètre ρ : $f_X(x) = \rho e^{-\rho x}$. $T_j = X_1 + X_2 + \dots + X_j$ suit une loi Erlang, dont la densité est:

$$g(t) = \frac{\rho(\rho t)^{j-1} e^{-\rho t}}{(j-1)!} \quad (3.9)$$

3.2.2 Le processus de Poisson non homogène

Le processus de Poisson non homogène (PNH) est une extension du processus de Poisson homogène, où la fonction d'intensité $\rho(t)$ est une fonction du temps. On a alors $\rho(t)=h(t)$. Les hypothèses d'indépendance et de non-multiplicité des événements sont maintenues tandis que l'hypothèse de stationnarité ne l'est plus. Le nombre d'événements N_t est distribué selon une loi de Poisson de paramètre $\int_0^t \rho(u)du$. Les intervalles X_j entre les points du PNH ne sont plus indépendants, l'indépendance est plutôt conditionnelle à t_j . Nous présentons au chapitre 4 une méthode d'estimation bayésienne semi-paramétrique de la fonction d'intensité $\rho(t)$ d'un PNH.

3.2.3 Le processus de renouvellement homogène

Un processus de Poisson dont les intervalles sont indépendants et distribués selon une loi exponentielle, est généralisé à un processus de renouvellement homogène (RH) lorsque les intervalles $\{X_1, X_2, \dots, X_j\}$ sont indépendants et distribués selon une loi commune arbitraire f . L'intensité $\rho(t, H_t)$ d'un RH dépend uniquement de l'instant d'occurrence du dernier événement, donc $\rho(t, H_t) = \rho(t, t_{N_t})$. On définit pour un RH la densité de renouvellement $m(t) = \frac{d}{dt} E(N(t))$. C'est la probabilité d'un événement dans l'intervalle $(t, t+\Delta t]$ (pas nécessairement le premier depuis l'origine). L'intensité conditionnelle $h(t)$ correspond à $m(t)$ pour un RH ayant un événement à l'origine.

Une quantité importante pour le RH est le taux d'avarie (*hazard rate*) $\varphi(x)$. $\varphi(x)\Delta t$ est la probabilité d'occurrence d'un événement étant donné que le temps écoulé depuis le dernier événement est x :

$$\varphi(x) = \lim_{\Delta x \rightarrow 0^+} \frac{P(x < X \leq x + \Delta x | x < X)}{\Delta x} \quad (3.10)$$

$\varphi(x)$ est défini à partir de la densité des intervalles f comme suit:

$$\varphi(x) = \frac{f(x)}{1 - F(x)} \quad (3.11)$$

où $F(x) = \int_0^x f(a) da$. Le produit $\varphi(t-t_{N_t})\Delta t$ est la probabilité conditionnelle qu'un événement survienne dans l'intervalle $(t, t+\Delta t]$ étant donné que le dernier événement est survenu à l'instant t_{N_t} . La relation entre l'intensité et le taux d'avarie est: $\rho(t, t_{N_t}) = \varphi(t-t_{N_t})$. Le taux de rupture d'un PH, est $\varphi(x) = \rho$. Pour un PNH, c'est $\varphi(x, t) = \rho(t)$. $F_n = F * F_{n-1}$ est la convolution de F et F_{n-1} telle que:

$$F_2 = F * F(a) = \int_{-\infty}^{\infty} F(a-y)F(y)dy \quad (3.12)$$

3.2.4 Le processus de Poisson doublement stochastique

Le processus de Poisson doublement stochastique (PDS) est un processus dont l'intensité est un processus aléatoire qui peut dépendre d'un processus d'information externe, lui-même aléatoire, $\{Y_t\}$. Étant donné le processus ponctuel, le processus d'intensité $\rho(t, Y_t)$ n'est pas connu exactement. Toutefois, étant donné le processus externe $\{Y_t = y; t \geq 0\}$, le PDS devient un processus de Poisson non homogène conditionnel (PNHC) de fonction d'intensité $\rho(t, y_t)$. C'est ce cas spécial du PDS que nous utiliserons dans notre travail. On définit:

$$\hat{\lambda}_t = E\left(\rho(t, Y_t) \mid N_t; t_1, t_2, \dots, t_{N_t}\right), N_t \geq 1 \text{ et } \hat{\lambda}_t(n) = E\left(\hat{\lambda}_t \mid N_t = n\right) \quad (3.13)$$

Il est en général très difficile, voire impossible de calculer $\hat{\lambda}_t(n)$ et $\hat{\lambda}_t$.

3.2.5 Le processus autoexcité

Le processus autoexcité (AE) est le processus le plus général présenté dans ce chapitre. Il est obtenu à partir d'un processus de Poisson en permettant à l'intensité de dépendre du processus ponctuel lui-même. Ainsi, le nombre d'événements N_t passés et leurs temps d'occurrence $(t_1, t_2, \dots, t_{N_t})$ peuvent influencer le nombre d'événements futurs et leurs temps d'occurrence.

L'intensité d'un AE $\rho(t, N_t; t_1, t_2, \dots, t_{N_t}), t_{N_t} < t \leq t_{N_t+1}$, est un processus aléatoire ayant des chemins connus, étant donnée l'historique du processus ponctuel H_t . Le processus a une mémoire m si le processus d'intensité dépend des m derniers événements. On définit $\hat{\rho}_t(N_t; t_1, t_2, \dots, t_{N_t}) = E(\rho(t, N_t; t_1, t_2, \dots, t_{N_t}) \mid N_t)$.

Un PH a une mémoire d'ordre 0 et un RH a une mémoire d'ordre 1. Un processus a une mémoire d'ordre m si et seulement si les temps d'occurrence des événements forment une chaîne de Markov d'ordre m . Un PDS est caractérisé comme un processus autoexcité dont le processus d'intensité est $\{\hat{\lambda}_t; t \geq t_0\}$, tel que défini dans la section 3.2.4. Le RNH est un AE non homogène de mémoire d'ordre 1, et dont l'intensité dépend uniquement de l'instant t et de l'instant d'occurrence du dernier événement t_{N_t} , tel que $\rho(t, N_t; t_1, t_2, \dots, t_{N_t}) = \rho(t, t_{N_t})$, $t_{N_t} < t \leq t_{N_t+1}$.

La figure 3.8 résume les relations entre ces processus.

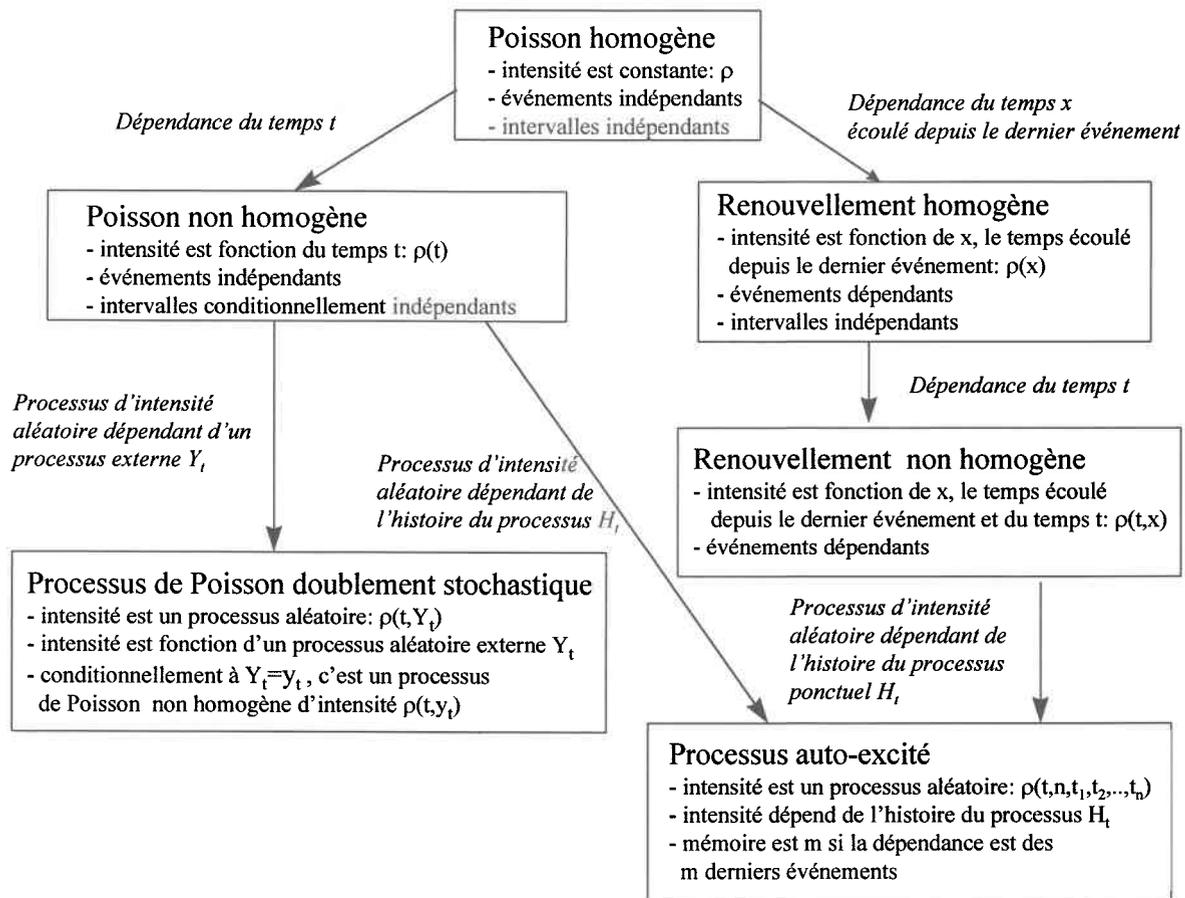


Figure 3.8: Schéma illustrant les relations entre certains processus ponctuels

Les tableaux 3.1 et 3.2 présentent les quantités définies à la section 3.1 pour chacun des processus. La variable σ est une variable d'intégration correspondant aux temps d'occurrence t_1, \dots, t_n .

	PH	PNH	RH
$P(N_t) = n$	$\frac{e^{-\rho t} (\rho t)^n}{n!}$	$\frac{\exp\left(-\int_0^t \rho(u) du\right) \left(\int_0^t \rho(u) du\right)^n}{n!}$	$F_n(t) - F_{n+1}(t)$
$f_x(x)$ la densité des intervalles X	$\rho e^{-\rho x}$	$\rho(t_0 + x) \exp\left(-\int_{t_0}^{t_0+x} \rho(u) du\right)$	$\varphi(x) \exp\left(-\int_0^x \varphi(u) du\right)$
La probabilité de survie $P(X_{n+1} \geq x t_1, t_2, \dots, t_n)$	$e^{-\rho x}$	$\exp\left(-\int_{t_n}^{t_n+x} \rho(u) du\right)$	$\exp\left(-\int_0^x \varphi(u) du\right)$
La densité conjointe des temps d'occurrence $P_{\vec{t}}^{(n)}(\vec{t})$	$\rho^n \exp(-\rho t_n)$	$\prod_{j=1}^n \rho(t_j) \exp\left(-\int_0^{t_n} \rho(u) du\right)$	$\rho(t_1, 0) \left(\prod_{j=2}^n \rho(t_j, t_{j-1})\right) \exp\left(-\int_0^{t_1} \rho(0) d\sigma - \sum_{i=2}^n \int_{t_{i-1}}^{t_i} \rho(u, t_{i-1}) du\right)$
la densité de l'échantillon $P(N_\sigma; t_0 \leq \sigma < t)$	$\rho^n \exp(-\rho t)$	$\prod_{j=1}^n \rho(t_j) \exp\left(-\int_0^t \rho(u) du\right)$	$(1 - F_x(t)) \varphi(t) \varphi(t_2 - t_1) \dots \varphi(t_n - t_{n-1})$
La densité conditionnelle des temps d'occurrence $P_{\vec{t}}(\vec{t} N_t = n)$	$\frac{n!}{t^n}$	$n! \prod_{j=1}^n \frac{\rho(t_j)}{\int_0^t \rho(u) du}$	$\frac{P(N_\sigma; t_0 \leq \sigma < t)}{P(N_t = n)}$

Tableau 3.1: Certaines définitions liées aux processus ponctuels PH, PNH et RH

	AE	PDS
$P(N_t) = n$	$\int_{t_0}^t \hat{\rho}_\tau (n-1) P(N_\tau = n-1) \exp\left(-\int_{\tau}^t \hat{\rho}_\sigma (n) d\sigma\right) d\tau$	$P[N_t = n] = \int_{t_0}^t \hat{\lambda}_\tau (n-1) P[N_\tau = n-1] \exp\left[-\int_{\tau}^t \hat{\lambda}_\sigma (n) d\sigma\right] d\tau$
$f_X(x)$ la densité des intervalles X	$\rho(t, n; t_1, t_2, \dots, t_n)^* \exp\left(-\int_{t_n}^t \rho(t, n; t_1, t_2, \dots, t_n) d\sigma\right)$	-
La probabilité de survie $P(X_{n+1} \geq x t_1, t_2, \dots, t_n)$	$\exp\left(-\int_{t_n}^t \rho(t, n; t_1, t_2, \dots, t_n) d\sigma\right)$	$\exp\left(-\int_{t_n}^t \hat{\lambda}_\sigma d\sigma\right)$
La densité conjointe des temps d'occurrence $P_{\vec{T}}^{(n)}(\vec{t})$	$\rho(t_1, 0) \left\{ \prod_{j=2}^n \rho(t_j, j-1; t_1, t_2, \dots, t_{j-1}) \right\} \exp\left[-\int_{t_0}^{t_1} \rho(\sigma, 0) d\sigma - \sum_{j=2}^n \int_{t_{j-1}}^{t_j} \rho(\sigma, j-1; t_1, t_2, \dots, t_{j-1}) d\sigma\right]$	$\exp\left(-\int_0^{t_0} \hat{\lambda}_\sigma d\sigma + \int_0^{t_n} \ln \hat{\lambda}_\sigma dN_\sigma\right)$
la densité de l'échantillon $P(N_\sigma; t_0 \leq \sigma < t)$	$\rho(t_1, 0) \left\{ \prod_{j=2}^n \rho(t_j, j-1; t_1, t_2, \dots, t_{j-1}) \right\} \exp\left[-\int_{t_0}^{t_1} \rho(\sigma, 0) d\sigma - \sum_{j=2}^n \int_{t_{j-1}}^{t_j} \rho(\sigma, j-1; t_1, t_2, \dots, t_{j-1}) d\sigma - \int_{t_n}^t \rho(\sigma, j-1; t_1, t_2, \dots, t_n) d\sigma\right]$	$\exp\left(-\int_{t_0}^t \hat{\lambda}_\sigma d\sigma + \int_{t_0}^t \ln \hat{\lambda}_\sigma dN_\sigma\right)$
La densité conditionnelle des temps d'occurrence $P_{\vec{T}}(\vec{t} N_t = n)$	$\frac{P(N_\sigma; t_0 \leq \sigma < t)}{P(N_t = n)}$	$\frac{P(N_\sigma; t_0 \leq \sigma < t)}{P(N_t = n)}$

Tableau 3.2: Certaines définitions liées aux processus ponctuels AE et PDS

3.3 La stationnarité des processus ponctuels

La stationnarité d'un processus ponctuel est définie soit en fonction des événements, soit en fonction des intervalles. Ces deux définitions ne sont pas équivalentes.

Définition 7: la stationnarité en fonction des événements

La stationnarité en fonction des événements implique que la distribution du nombre d'événements dans un intervalle fixe $(t_1, t_2]$ est invariante suite à une translation, elle est la même que dans $(t_{1+s}, t_{2+s}] \forall s$. La distribution conjointe des nombres d'événements dans les intervalles $(t_1, t_2], (t_3, t_4]$ est invariante suite à une translation.

Une conséquence immédiate est que l'espérance du nombre d'événements dans un intervalle ne dépend que de la longueur de l'intervalle et qu'elle est proportionnelle à sa longueur.

Définition 8: la stationnarité en fonction des intervalles

La stationnarité en fonction des intervalles implique que les intervalles X_j entre les événements ont la même distribution marginale $F_X(x)$. La distribution conjointe de k intervalles entre événements est alors invariante suite à une translation.

Un processus peut soit débiter à un temps arbitraire, soit débiter à un événement arbitraire. La différence entre les deux types de processus est illustrée sur la figure 3.9.

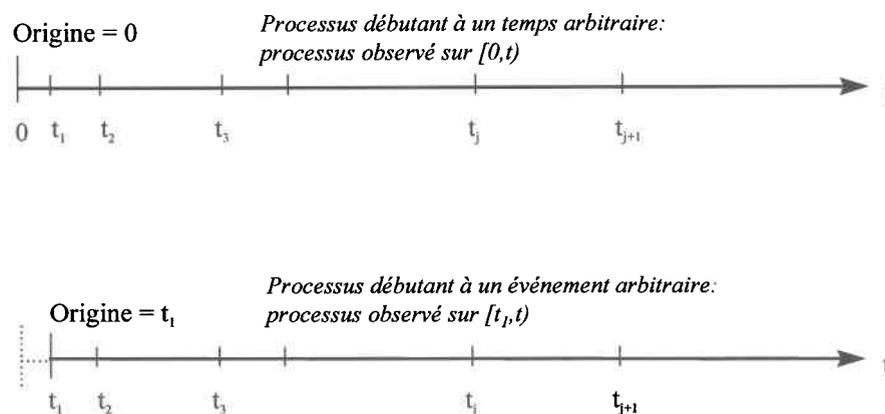


Figure 3.9: Un processus débutant à un temps arbitraire ou à un événement arbitraire

La série d'événements commençant à un événement arbitraire n'est généralement pas stationnaire dans le sens de la définition 7. De plus, les intervalles $\{X_j\}$ peuvent être stationnaires sans que la série d'événements le soit, comme par exemple dans un processus de renouvellement débutant à un temps arbitraire où les intervalles sont indépendants et identiquement distribués par définition. Le nombre d'événements par intervalle dépend alors de la distance entre cet intervalle et l'origine (Cox et Lewis, 1966).

3.3.1 L'homogénéisation d'un processus ponctuel non homogène

En général, on peut homogénéiser un processus ponctuel PP d'intensité $\rho(t;H_t)$, par la transformation suivante:

$$t' = \tau(t, H_t) = \int_0^t \rho(u, H_u) du \quad (3.14)$$

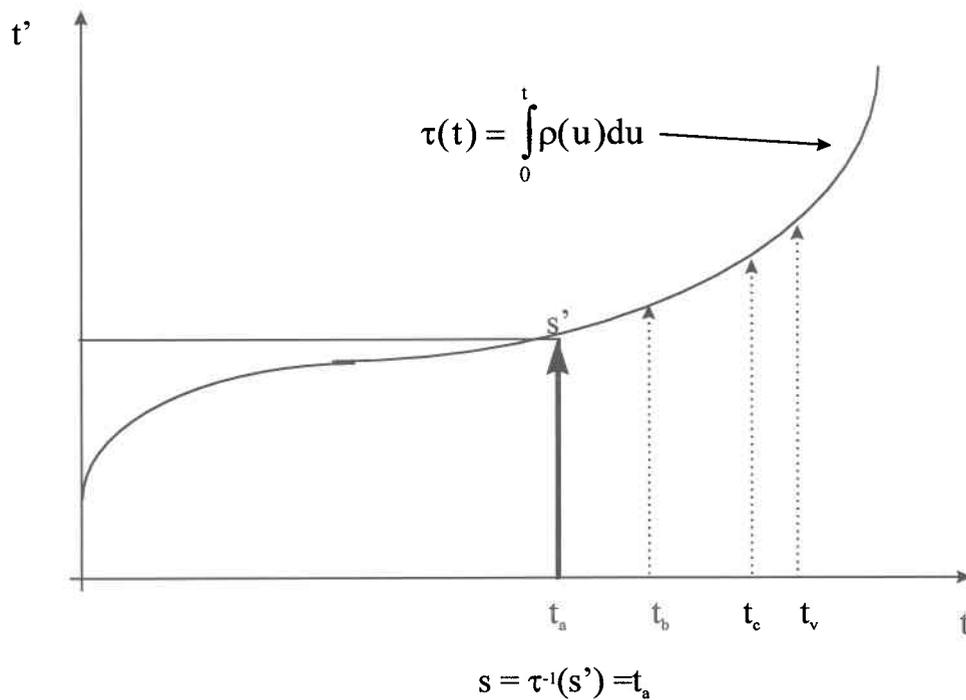
Le processus obtenu PP', défini par les temps d'occurrence t'_j , est alors un processus de Poisson homogène d'intensité 1. Toutefois, τ est une fonction déterministe du temps si et seulement si PP est un processus de Poisson non homogène. Dans les autres cas, $\rho(t;H_t)$ est une fonction de H_t et la transformation est alors aléatoire puisqu'elle dépend de la réalisation particulière du processus (Cox et Isham, 1980).

3.3.2 L'homogénéisation du processus de Poisson non homogène

Supposons que PP soit un PNH. PP est complètement spécifié par l'intensité $\rho(t)$ où $E(N_t) = \int_0^t \rho(u) du$. Définissons une nouvelle variable $t' = \tau(t)$ appelée le temps opérationnel telle que:

$$t' = \tau(t) = \int_0^t \rho(u) du, \text{ alors } \tau(t) = E(N_t) \quad (3.15)$$

Puisque $\tau(t)$ est continue et non décroissante, on peut définir sa fonction inverse $\tau^{-1}(t')$ comme suit: pour $s' = \tau(s) > 0$, $s = \tau^{-1}(s') = \min(t_a, t_b, \dots, t_v)$ telle que $\tau(s) \geq s'$, tel qu'illustrée sur la figure 3.10.

Figure 3.10: La définition de la fonction $\tau^{-1}(t')$ (adaptée de North, 1981)

Le processus de dénombrement $\{M_{t'}, t' \geq 0\}$ défini par $M_{t'} = N_{\tau^{-1}(t')}$ est un processus de Poisson PP' dont l'espérance est:

$$E(M(t')) = E[N_{\tau^{-1}(t')}] = \tau(\tau^{-1}(t')) = t'. \quad (3.16)$$

PP' est donc un processus de Poisson homogène d'intensité 1 en fonction de la nouvelle variable de temps t' , puisque cette transformation n'affecte pas la propriété d'indépendance des événements.

Pour homogénéiser en pratique, on constate que:

$$\frac{d\tau(t)}{d\tau} = \frac{dt}{d\tau} \frac{d}{dt} \int_0^t \rho(u) du, \text{ donc } 1 = \frac{dt}{d\tau} \rho(t) \quad (3.17)$$

et pour de petits intervalles, on écrit $\Delta\tau = \rho(t)\Delta t$. Les intervalles x_j dans PP sont soit compressés soit étirés afin de devenir x'_j dans PP', tels qu'illustrés sur la figure 3.11:

$$x'_j = \int_{t_{j+1}}^{t_j} \rho(u) du \quad (3.18)$$

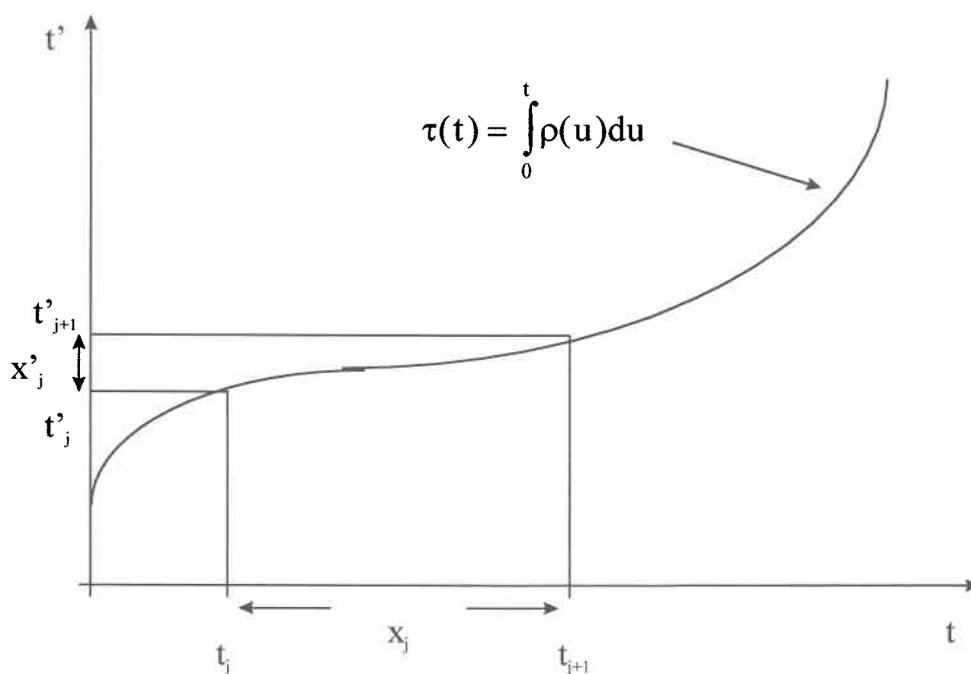


Figure 3.11: L'homogénéisation d'un PNH (adaptée de North, 1981)

3.4 L'analyse des séries d'événements

Une série d'événements correspond à une réalisation d'un processus ponctuel. L'analyse des séries d'événements permet de choisir un modèle (processus ponctuel) approprié pour le phénomène observé. Elle permet de dégager les propriétés des intervalles et des dénombrements avant de construire le modèle, et une fois le modèle choisi, elle permet de tester son adéquation. Les techniques d'analyse des propriétés des dénombrements et des intervalles présentées ici sont basées sur les travaux de Cox et Lewis (1966), et de Kavvas (1975). Ces méthodes graphiques et ces tests statistiques permettent de détecter les non-homogénéités, et de vérifier les hypothèses d'un processus de renouvellement et d'un processus de Poisson. Les précipitations enregistrées à la station française de Bar-sur-Seine (1951-1977) sont utilisées afin d'illustrer l'usage des méthodes graphiques. La figure 3.12 présente un sous-ensemble de ces précipitations (1951-1955). La série d'événements analysée est définie par les temps d'occurrence des précipitations t_j ou d'une façon équivalente, par les intervalles entre les événements x_j .

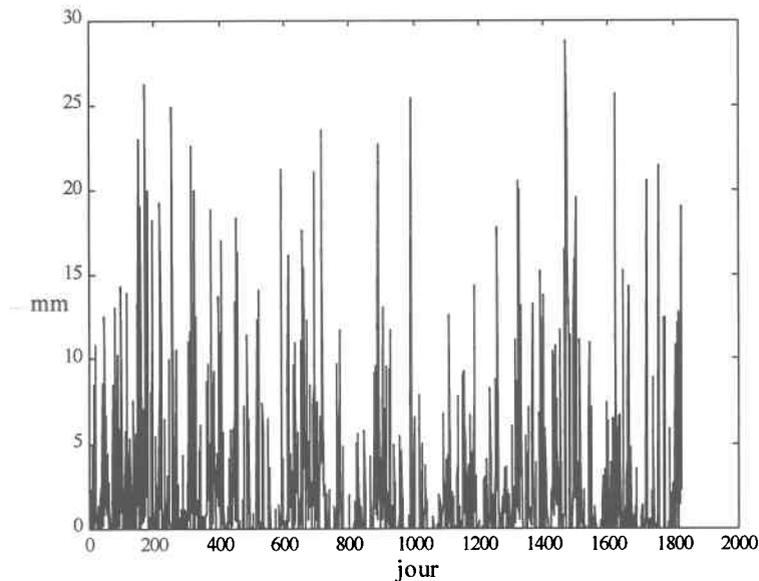


Figure 3.12: Les précipitations sur Bar-sur-Seine, 1951-1955

3.4.1 L'analyse des tendances et de la périodicité

Nous commençons par présenter les méthodes d'analyse des séries d'événements pour détecter les non-homogénéités.

3.4.1.1 Le graphique cumulatif

Le graphique cumulatif est le nombre de jours depuis le début de la période d'observation jusqu'au $j^{\text{ème}}$ événement en fonction de j . Sa pente correspond à l'inverse du taux moyen d'occurrence. L'hypothèse d'homogénéité du processus est rejetée lorsque cette pente n'est pas constante. La figure 3.13 illustre ce graphique pour les précipitations sur Bar-sur-Seine. On en déduit que les données ne semblent pas homogènes puisque la pente de la droite n'est pas constante.

3.4.1.2 La fonction de survie logarithmique empirique

Ce graphique est basé sur les intervalles ordonnés en ordre croissant, x_i . Pour un processus de Poisson homogène, les x_i suivent une loi exponentielle, et le graphique de $\ln\left\{1 - \frac{i}{n_0 + 1}\right\}$

en fonction de x_i , où n_0 est le nombre total d'événements, est une droite. Ceci ne semble pas être le cas pour les précipitations sur Bar-sur-Seine (figure 3.14).

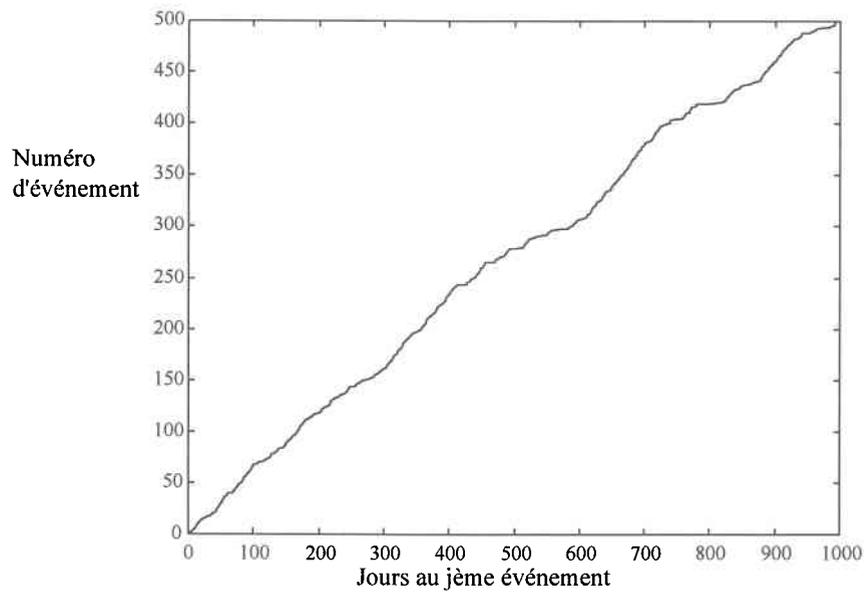


Figure 3.13: Le graphique cumulatif des précipitations sur Bar-sur-Seine

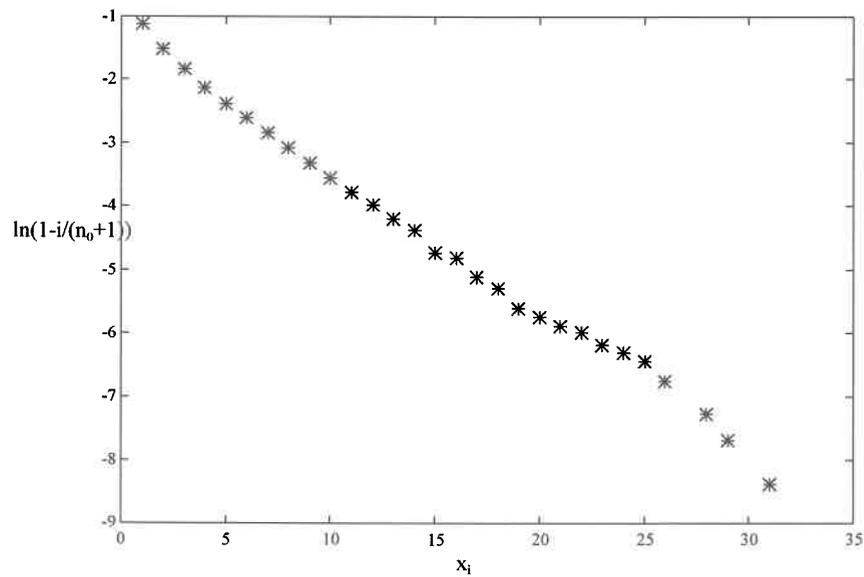


Figure 3.14: La fonction de survie logarithmique

3.4.1.3 Le taux moyen journalier d'occurrence

Le taux moyen d'occurrence $\rho(t)$ est estimé par $\lambda_\tau(t)$, tel que:

$$\lambda_\tau(t) = \frac{n(t, t + \tau)}{\tau} \quad (3.19)$$

En commençant à un temps arbitraire et en prenant des intervalles disjoints de longueur τ , $n(t, t + \tau)$ est le nombre d'événements dans l'intervalle $(t, t + \tau]$. $\lambda_\tau(t)$ est non biaisé lorsque la stationnarité est vérifiée à l'intérieur de cet intervalle. Le graphique de $\lambda_\tau(t)$ en fonction du temps permet de déceler les cyclicités et/ou les tendances de ce taux. La figure 3.15 ne semble pas montrer de tendance évidente mais on pourrait toutefois soupçonner l'existence d'une saisonnalité annuelle.

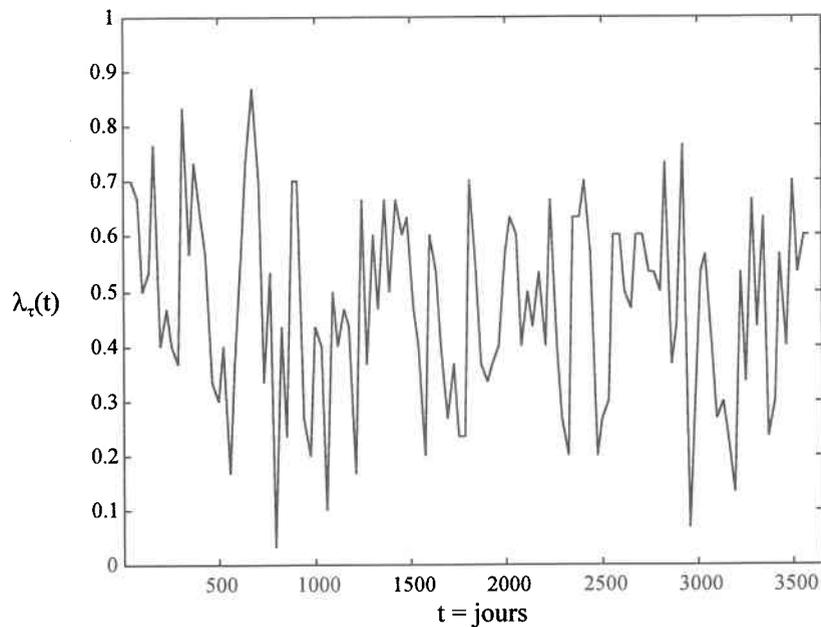


Figure 3.15: Le taux moyen journalier d'occurrence de précipitation, $\tau=30$

3.4.1.4 Le spectre des dénombrements

Un spectre de dénombrements est défini par:

$$g_+(\omega) = \int_0^{\infty} \exp(i\omega t) \gamma(t) dt \quad (3.20)$$

$$\text{où: } \gamma(t) = \lim_{s \rightarrow 0} \text{cov} \left(\frac{N_{t+u+s} - N_{u+t}}{s}, \frac{N_{u+s} - N_u}{s} \right) \quad (3.21)$$

Le périodogramme est l'estimateur du spectre des dénombrements défini par:

$$\hat{g}_+(\omega_p) = \frac{\pi T}{n} \frac{1}{\pi T} I_{\omega_p}, \text{ où } I_{\omega_p} = [A_T^2(\omega_p) + B_T^2(\omega_p)] \quad (3.22)$$

$$A_T(\omega_p) = \sum_{j=1}^n \cos(t_j \omega_p), \text{ et } B_T(\omega_p) = \sum_{j=1}^n \sin(t_j \omega_p) \quad (3.23)$$

où $\omega_p = \frac{p2\pi}{T}$, $p = 1, 2, \dots, n/2$. Ici p , est dit l'indice de fréquence, T est l'horizon d'observation, et t_j est le temps d'arrivée de l'événement j .

Un spectre des dénombrements $g_+(\omega)$ non plat peut révéler des périodicités et une structure de dépendance des dénombrements. Toutefois cette théorie est basée sur les processus stationnaires. Une déviation d'un spectre plat (spectre d'un PH) peut donc être une déviation d'un processus de Poisson. Le spectre d'un processus de Poisson d'intensité ρ , normé par π/ρ lissé sur k périodes, et multiplié par $2k$, est distribué selon une loi de χ^2 à $2k$ degrés de liberté. On voit sur la figure 3.16 que le spectre des précipitations diffère de celui d'un processus de Poisson homogène.

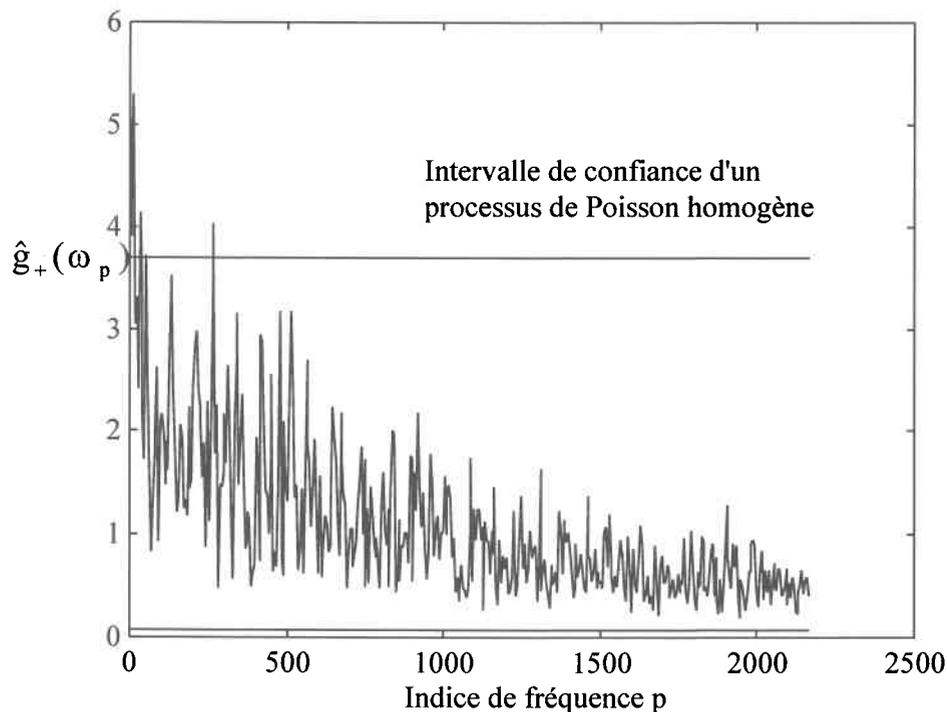


Figure 3.16: Le spectre des précipitations sur Bar-sur-Seine

3.4.1.5 Le test de Cramér

En faisant l'hypothèse que le processus est poissonien d'intensité $\rho(t)=e^{a+bt}$, le test de Cramér (1946) permet de détecter la présence d'une tendance dans l'intensité.

Hypothèses:

H_0 : $b=0$ (intensité constante)

H_1 : $b \neq 0$ (intensité dépend du temps)

La statistique de Cramér définie par $U = \frac{S - \frac{t_n}{2}}{\frac{t_n}{\sqrt{12n}}}$, où $S = \frac{\sum_{j=1}^n t_j}{n}$, est asymptotiquement

distribuée selon une loi $N(0,1)$. H_0 est rejetée à 5% si $|U| > 1.96$.

3.4.1.6 L'homogénéité de la variance

Le test de l'homogénéité de la variance (test de Bartlett) (Kendall et Stuart, 1961) permet de vérifier la stationnarité des intervalles, lorsqu'on suppose qu'ils proviennent de populations normales. On groupe les intervalles entre les événements X_j en k groupes de taille n_i . σ_k^2 est la variance du $k^{\text{ème}}$ groupe d'intervalles.

Hypothèses:

H_0 : $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$

H_1 : variance non homogène

H_0 est rejetée au niveau α si $K > \chi_\alpha^2(k-1)$, où $K = \frac{-2 \log l^*}{1 + \frac{1}{3k-3} \left[\sum_{i=1}^k \frac{1}{v_i} - \frac{1}{v} \right]} \sim \chi^2(k-1)$,

$$v_i = n_i - 1, \quad -2 \log l^* = v \log s^2 - \sum_{i=1}^k v_i \log s_i^2 \quad (3.24)$$

$$\text{et: } s_i^2 = \frac{1}{v_i} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2, \quad s^2 = \frac{1}{v} \sum_{i=1}^k v_i s_i^2, \quad v = \sum_{i=1}^k v_i \quad (3.25)$$

où x_{ij} est le $j^{\text{ème}}$ intervalle du groupe i .

3.4.1.7 Le spectre des intervalles - Test de tendance dans le périodogramme

On peut tester l'hypothèse d'homogénéité du processus (absence de tendance dans le temps)

en utilisant l'estimateur du spectre des intervalles $I_n(\omega_p) = \frac{1}{2\pi n} \left| \sum_{j=1}^n X_j e^{ij\omega_p} \right|$, où n est le nombre d'intervalles.

Hypothèses:

H_0 : $I_n(\omega_p)$ sont asymptotiquement indépendants et distribués selon une loi exponentielle de moyenne $\frac{\sigma^2}{2\pi}$ (σ^2 est la variance des intervalles X_j)

H_1 : $I_n(\omega_p)$ présentent des tendances.

Posons $S = \sum_{i=1}^L \sum_{j=1}^i \frac{I_n(\omega_j)}{L}$, $U = \frac{S - \frac{1}{2}r_L}{r_L/\sqrt{12L}}$, où $r_i = \sum_{j=1}^i I_n(\omega_j)$, $i=1,2,\dots,L$, L étant la partie entière de $(n-1)/2$. U est alors distribuée selon une loi $N(0,1)$ et H_0 est rejetée à 5% si $|U| > 1.96$.

3.4.2 L'indépendance des intervalles - Tests de renouvellement

3.4.2.1 La fonction d'autocorrélation

L'examen de la fonction d'autocorrélation des intervalles permet de déterminer si ces derniers sont indépendants, donc si un processus de renouvellement est adéquat.

Hypothèses:

H_0 : les intervalles sont indépendants (processus de renouvellement)

H_1 : les intervalles ne sont pas indépendants (le processus de renouvellement est rejeté)

La fonction d'autocorrélation est estimée par $\hat{r}_j = \frac{n}{n-j} \frac{\sum_{i=1}^{n-j} (X_i - \bar{X})(X_{i+j} - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2}$, où

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

$\hat{r}_j \sqrt{n - J}$ est distribuée asymptotiquement selon une loi $N(0,1)$ et H_0 est rejetée à 5% si $|\hat{r}_j \sqrt{n - J}| > 1.96$. Ce test est surtout valable pour un processus où la distribution des intervalles est une loi normale.

3.4.2.2 Le périodogramme des intervalles - Tests ne dépendant pas de la distribution

Le périodogramme des intervalles normalisé peut être utilisé pour tester l'hypothèse d'indépendance des intervalles, et donc d'un processus de renouvellement.

Hypothèses:

H_0 : $I_n(\omega_p)$ sont indépendants - un processus de renouvellement

H_1 : $I_n(\omega_p)$ sont dépendants - le processus de renouvellement est rejeté.

Il s'agit de vérifier l'adéquation de la loi uniforme sur $(0,1)$ aux variables $U(i)$ définies par

$$U(i) = \frac{\sum_{j=1}^i I_n(\omega_j)}{\sum_{j=1}^L I_n(\omega_j)}, \quad i = 1, 2, \dots, L$$

par les test de Kolmogorov-Smirnov et d'Anderson-Darling

définis plus bas.

3.4.2.2.1 La statistique de Kolmogorov-Smirnov

Supposons qu'on ait un échantillon de taille n d'une variable aléatoire X dont la distribution continue est $F_X(x)$. On veut tester l'adéquation d'une loi F_0 donnée.

Hypothèses:

H_0 : $F_X = F_0$

H_1 : $F_X \neq F_0$

Considérons les variables $U_1 = F_X(X_1)$, $U_2 = F_X(X_2)$, ..., $U_n = F_X(X_n)$, telles que $X_1 \leq X_2 \leq \dots \leq X_n$. U_i sont distribuées uniformément sur $(0,1)$. Le problème est alors de tester si les points observés $u_1 = F_0(x_1)$, $u_2 = F_0(x_2)$, ..., $u_n = F_0(x_n)$ divisent l'intervalle unitaire de façon aléatoire. On calcule alors la statistique de Kolmogorov-Smirnov $D_n = \max(D_n^+, D_n^-)$ où:

$$D_n^+ = \max_{1 \leq i < n} \left(\frac{i}{n} - F_0(X_i) \right) \quad \text{et} \quad D_n^- = \max_{1 \leq i < n} \left(F_0(X_i) - \frac{i-1}{n} \right) \quad (3.26)$$

On rejette l'hypothèse H_0 au niveau α si la valeur observée D_n est plus grande que $d_{n,\alpha}$ telle que $P[D_n > d_{n,\alpha}] = \alpha$. Les valeurs de $d_{n,\alpha}$ sont données par des tables pour différents n et α .

3.4.2.2 la statistique d'Anderson

Supposons qu'on ait un échantillon de taille n d'une variable aléatoire X dont la distribution continue est $F_X(x)$. On veut tester l'adéquation d'une loi F_0 donnée.

Hypothèses:

$$H_0: F_X = F_0$$

$$H_1: F_X \neq F_0$$

La statistique d'Anderson W_n^2 est définie par:

$$W_n^2 = -n - \frac{1}{n} \sum_{i=1}^n \{(2i-1) \log U_i + (2n-2i+1) \log(1-U_i)\} \quad (3.27)$$

Sous l'hypothèse nulle H_0 , cette statistique a une moyenne de 0 et une variance de 1. La distribution asymptotique de cette statistique a été obtenue par Anderson et Darling (1952). On rejette l'hypothèse H_0 au niveau α si la statistique calculée W_n^2 est plus grande que $w_{n,\alpha}^2$ telle que $P[w_{n,\alpha}^2 > W_n^2] = \alpha$.

Les valeurs critiques des statistiques de Kolmogorov-Smirnov et d'Anderson sont présentées au tableau 3.3 pour n très grand.

	Kolmogorov-Smirnov	Anderson
$\alpha=.05$	1.358	2.492
$\alpha=.1$	1.628	3.857

Tableau 3.3: Les valeurs critiques des statistiques de K-S et d'Anderson

3.4.3 Tests du processus de Poisson

3.4.3.1 Le test conditionnel uniforme

On peut tester l'hypothèse d'un PH en utilisant les variables $U_j = t_j/T$. Considérons t_j , les temps d'arrivée des événements, et supposons que le processus est PH. Étant donné n événements sur l'intervalle $(0, T)$, tels que $t_1 \leq t_2 \leq \dots \leq t_n \leq t_{n+1}$, $U_j = t_j/T$ sont alors indépendants et distribués

selon une loi uniforme sur $(0,1)$. On peut alors utiliser le test de Kolmogorov-Smirnov et la statistique d'Anderson afin de vérifier l'hypothèse d'un processus de Poisson homogène, en examinant l'adéquation de la loi uniforme pour représenter la distribution des U_j .

Hypothèses:

H_0 : U_j sont indépendants distribués selon une loi uniforme sur $(0,1)$ - un PH

H_1 : le PH est rejeté.

3.4.3.2 Le test d'adéquation de la loi exponentielle aux intervalles

Les intervalles X_j d'un PH sont distribués selon une loi exponentielle. Le test de Kolmogorov-Smirnov et la statistique d'Anderson peuvent être utilisés afin de vérifier cette hypothèse.

Hypothèses:

H_0 : X_j suivent une loi exponentielle - un PH

H_1 : le PH est rejeté.

De plus, un coefficient de variation des intervalles (le rapport de l'écart type à la moyenne) différent de 1 permet de conclure que les intervalles ne sont pas distribués selon une loi exponentielle, dont l'écart type est égal à la moyenne.

4. L'ESTIMATION BAYÉSIENNE DE L'INTENSITÉ D'UN PROCESSUS DE POISSON NON HOMOGENÈME

Le processus de Poisson non homogène (PNH) est souvent utilisé pour la modélisation des sécheresses et des étiages (North, 1981; Zelenhasic et Salvai, 1987; Chang, 1989; Madsen et Rosbjerg, 1995). L'estimation de sa fonction d'intensité se fait habituellement par la méthode du maximum de vraisemblance (MV) (Lewis, 1972; Kavvas, 1975; North, 1981; Konecny et Nachnebel, 1985; Nachtnebel et Konecny, 1987). Toutefois, cette méthode d'estimation présente l'inconvénient que la forme de la fonction d'intensité doit être postulée. Les auteurs supposent alors que la fonction d'intensité a la forme suivante:

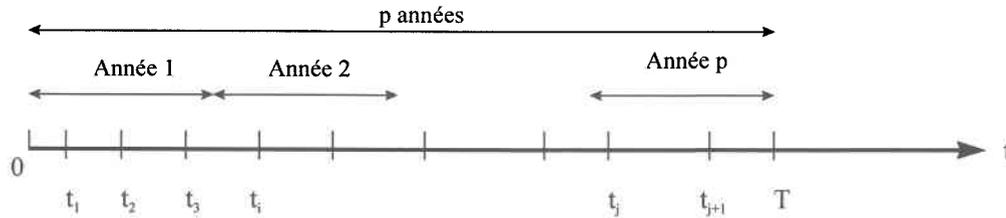
$$\lambda(t) = \beta_0 \exp \left[\sum_{j=1}^q \beta_j \sin(j\omega t + \beta_{q+j}) \right], \quad \omega = \frac{2\pi}{365} \quad (4.1)$$

Or, d'après Bernier (1981a), cette forme analytique n'est pas toujours réaliste, en hydrologie notamment. Il propose alors d'estimer cette fonction selon une approche bayésienne semi-paramétrique (BSP). Dans ce chapitre, nous présentons cette méthode, et nous développons certains résultats analytiques que nous avons obtenus lorsque la structure de dépendance des intensités moyennes est supposée circulaire. Nous décrivons ensuite la méthode du MV ainsi que les calculs nécessaires pour obtenir les variances des estimateurs. Nous comparons les résultats obtenus par ces deux méthodes en considérant un PNH simulé ainsi que des données réelles de précipitation enregistrées à la station française de Bar-sur-Seine. Nous commençons par expliquer le principe de superposition qui est à la base de ces deux méthodes d'estimation.

4.1 Le principe de superposition

L'estimation bayésienne de l'intensité d'un PNH est basée sur la construction d'un processus superposé: étant donné un processus de base PB observé pendant p périodes τ (années), le processus superposé PS est obtenu en ramenant les p flux de durées τ (1 an) sur un même intervalle de temps de durée τ . La figure 4.1 illustre le concept d'un processus superposé.

Le processus de base (PB) observé sur $(0, T = p\tau)$, $\tau = 1$ année, d'intensité $\lambda(t)$



Le processus superposé (PS) sur $(0, TS = \tau)$, $\tau = 1$ année, d'intensité $p\lambda(t)$

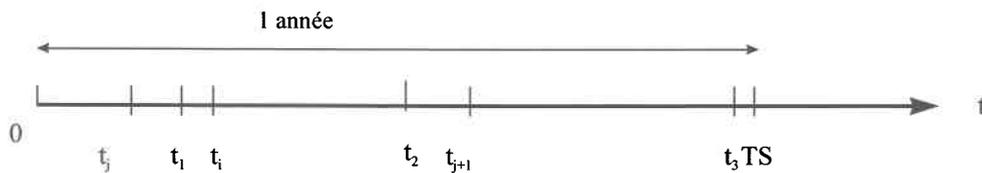


Figure 4.1: La superposition de processus

Si le processus de base PB est poissonien d'intensité $\lambda(t)$, périodique de période τ , alors le processus superposé PS sera poissonien d'intensité $p\lambda(t)$. Sinon, pour $p \rightarrow \infty$, le processus PS résultant de la superposition sera quand même poissonien (Çinlar, 1973; Bernier, 1981a).

4.2 L'estimation bayésienne de l'intensité d'un PNH

Le principe fondamental de la statistique bayésienne est formulé par Bernier (1981b) comme suit:

Toutes les données nécessaires pour porter un jugement opérationnel sur un paramètre θ sont contenues dans $\beta_y(\theta)$ la densité de probabilité a posteriori de θ :

$$\beta_y(\theta) = \frac{f(y|\theta)b(\theta)}{\int_{\Theta} f(y|\theta)b(\theta)d\theta} \quad (4.2)$$

où $b(\theta)$ est la densité a priori de θ . Cette densité a priori est estimée préalablement à et indépendamment de la prise en compte des observations y .

$f(y|\theta)$ est la densité conditionnelle des observations. Elle est identique à la vraisemblance de l'échantillon $L(n,\theta)$.

La fonction de vraisemblance

Nous commençons par présenter la fonction de vraisemblance de l'échantillon, $L(n, \Lambda)$. Considérons l'intervalle $[0, \tau]_K$ séparé en K intervalles disjoints de longueur Δ_i tels que $\sum_i \Delta_i = \tau$. Soit $\bar{\lambda}_i = \frac{1}{\Delta_i} \int_{\Delta_i} \lambda(u) du$, l'intensité moyenne sur Δ_i , où $\lambda(u)$ est l'intensité du processus de base PB.

Soit n_i le nombre d'événements superposés sur Δ_i . Chaque n_i suit indépendamment des autres n_j une loi de Poisson de paramètre $p\bar{\lambda}_i$ telle que la vraisemblance conjointe pour les K intervalles s'écrit:

$$L(n, \Lambda) = \exp\left(-\sum_1^K p\bar{\lambda}_i\right) \prod_{i=1}^K \frac{(p\bar{\lambda}_i)^{n_i}}{n_i!}, \quad \Lambda = (\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_K) \tag{4.3}$$

Posons $\theta_i = \text{Log}(p\bar{\lambda}_i)$, et c une constante ne dépendant pas des θ_i . Nous avons alors:

$$L(n, \Theta) = c \exp\left(-\sum_1^K \exp(\theta_i) + \sum_1^K n_i \theta_i\right) \tag{4.4}$$

L'estimation de $\lambda(u)$ revient à estimer $\Theta = (\theta_1, \theta_2, \dots, \theta_K)$. Dans ce travail, nous développons la méthode d'estimation pour une matrice de covariance V circulaire (définie dans 4.2.1), et nous appliquons cette approche à l'estimation de l'intensité dans les deux cas: la moyenne de Θ , $\bar{\mu} = (\mu_1, \mu_2, \dots, \mu_K)$, connue et inconnue. Les hypothèses que nous posons sont développées à la section 4.2.1.

4.2.1 Les hypothèses

L'estimation de $\Theta = (\theta_1, \theta_2, \dots, \theta_K)$ se fera selon l'approche bayésienne décrite dans Bernier (1981a), basée sur les travaux de Leonard (1973, 1978) qui propose d'estimer les θ_i par les valeurs modales de la distribution a posteriori $\beta_n(\Theta)$. Les θ_i sont alors la solution du système $\frac{\partial \beta_n(\Theta)}{\partial \theta_i} = 0, i = 1..K$.

Hypothèse 1: normalité de la densité a priori des paramètres θ_i

Nous supposons que $b(\Theta)$ suit une loi $N(\bar{\mu}, V)$, où $C=V^{-1}$ est l'inverse de la matrice de variance-covariance des θ_i . La loi $b(\Theta)$ est donnée par:

$$b(\Theta) = r \frac{\exp\left(-\frac{1}{2}(\Theta - \bar{\mu})^T C(\Theta - \bar{\mu})\right)}{\sqrt{\det(V)}} \quad (4.5)$$

où r est une constante. Une conséquence de cette hypothèse est que la densité a posteriori $\beta_n(\Theta)$ est:

$$\beta_n(\Theta) = s \exp\left(-\sum_1^K \exp(\theta_i) + \sum_1^K n_i \theta_i\right) * b(\Theta) \quad (4.6)$$

où la constante s est choisie telle que $\int_{\Theta} \beta_n(\Theta) d\Theta = 1$.

Hypothèse 2: modèle de dépendance circulaire

Afin d'estimer les valeurs de θ_i , la matrice V doit être spécifiée pour tenir compte de la structure de dépendance des θ_i . Pour l'application aux processus périodiques, une hypothèse naturelle est de choisir la distribution des θ_i circulaire, c'est-à-dire, qu'il existe une dépendance directe entre θ_K et θ_1 en plus de la corrélation d'ordre 1 entre θ_i et θ_{i+1} , $i=1, K-1$. Posons $\varphi_i = \theta_i - \mu_i$. Un modèle compatible avec l'hypothèse de dépendance circulaire est (Bernier, 1981a):

$$\begin{aligned} \varphi_i &= a(\varphi_{i-1} + \varphi_{i+1}) + \varepsilon_i, \quad i=2, \dots, K-1 \\ \varphi_1 &= a(\varphi_K + \varphi_2) + \varepsilon_1, \quad \varphi_K = a(\varphi_{K-1} + \varphi_1) + \varepsilon_K \end{aligned} \quad (4.7)$$

où les K variables ε_i sont des variables aléatoires indépendantes normales d'espérance nulle et de variance σ^2 constante. Posons $h=1/\sigma^2$, $\Phi=[\varphi_1, \varphi_2, \dots, \varphi_K]$, $E=[\varepsilon_1, \varepsilon_2, \dots, \varepsilon_K]$. Il faut alors estimer a et h afin de connaître θ_i .

Notre approche diffère de celle de Vernos (1982) qui a utilisé cette méthode d'estimation afin d'estimer l'intensité d'un processus d'occurrence de crue. Il a supposé que V possédait la propriété suivante: $v_{ij} = v^2 \rho^{|j-i|}$, $-1 < \rho < 1$.

Hypothèse 3: $\bar{\mu}$ est connue

On suppose que la moyenne a priori de Θ , $\bar{\mu} = (\mu_1, \mu_2, \dots, \mu_K)$ est connue.

Hypothèse 3': $\bar{\mu}$ est inconnue

On suppose que la moyenne a priori de Θ , $\bar{\mu} = (\mu_1, \mu_2, \dots, \mu_K)$ est inconnue, et que $\mu_i = \mu_0$, où la loi a priori de μ_0 est $N(m, v^2)$. Les hypothèses 3 et 3' sont mutuellement exclusives.

Hypothèse 4: indépendance de a et de h et leurs densités

On suppose que a et h sont indépendants et que leurs densités a priori sont non informatives, c'est-à-dire qu'elles ne contiennent pas d'information sur a et h, telles que $b(a|h)=1$ et $b(h)=1/h$ (Berger, 1985). Notons que ces lois sont impropres.

4.2.2 La méthodologie et l'algorithme d'estimation

La méthodologie d'estimation suivie s'appuie sur les étapes de calcul suivantes:

1. spécifier la matrice de variance-covariance V, et sa matrice inverse C;
2. estimer a, h, et Θ lorsque $\bar{\mu}$ est connue; et
- 2'. estimer a, h, et Θ lorsque $\bar{\mu}$ est inconnue.

Ces étapes sont décrites dans les paragraphes qui suivent.

Étape 1: matrice de variance-covariance V

Il s'agit en premier lieu de spécifier V, la matrice de variance-covariance du système $\varphi_i = a(\varphi_{i-1} + \varphi_{i+1}) + \varepsilon_i$, ainsi que son déterminant et sa matrice inverse C. Le système d'équations (4.7) implique que:

$$\text{Var}(\varphi_i) = V_0, \text{ et que } \text{Cov}(\varphi_i, \varphi_{i+j \bmod K}) = V_j, \forall i, \forall j, i, j = 1..K. \quad (4.8)$$

L'annexe 1 présente les calculs menant à la matrice V et à $C = h * A^2$ où:

$$A^2 = \begin{bmatrix} 2a^2 + 1 & -2a & a^2 & 0 & 0 & \dots & 0 & a^2 & -2a \\ -2a & 2a^2 + 1 & -2a & a^2 & 0 & \dots & \dots & 0 & a^2 \\ a^2 & -2a & 2a^2 + 1 & -2a & a^2 & 0 & \dots & \dots & 0 \\ 0 & a^2 & -2a & \dots & \dots & 0 & \dots & \dots & \dots \\ 0 & 0 & a^2 & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & a^2 & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 & \dots & \dots & -2a & a^2 \\ a^2 & 0 & \dots & \dots & \dots & 0 & a^2 & -2a & 2a^2 + 1 \\ -2a & a^2 & 0 & \dots & \dots & \dots & 0 & a^2 & -2a & 2a^2 + 1 \end{bmatrix} \quad (4.9)$$

Le déterminant de la matrice V est (annexe 1):

$$\det(V) = \frac{1}{\det(C)} = \frac{1}{h^K (\det A)^2} \quad (4.10)$$

$$\text{où } \det(A) = \left\{ -2a^k + \prod_{m=1}^{K-1} \left(1 + 2a \cos\left(\frac{m\pi}{K}\right) \right) - 2a^2 \prod_{m=1}^{K-2} \left(1 + 2a \cos\left(\frac{m\pi}{K-1}\right) \right) \right\} \quad (4.11)$$

Étape 2: estimation de a, h, et Θ lorsque $\bar{\mu}$ est connue

Nous développons dans cette section les estimateurs de a, h, et Θ lorsque $\bar{\mu}$ est connue.

Suite aux hypothèses 1, 2 et 3, la densité conjointe a priori $b_n(a, h, \Theta)$ de a, h, et Θ lorsque $\bar{\mu}$ connue est:

$$b_n(a, h, \Theta) \propto b(\Theta | a, h) * b(a|h) * b(h) \quad (4.12)$$

et la densité a posteriori correspondante $\beta_n(a, h, \Theta)$ est:

$$\beta_n(a, h, \Theta) \propto b(\Theta | a, h) * b(a|h) * b(h) * L(n, \Theta) \quad (4.13)$$

où $b(\Theta | a, h)$ est la loi normale multivariée de moyenne $\bar{\mu} = (\mu_1, \mu_2, \dots, \mu_K)$ connue et de matrice de variance-covariance V. Ceci implique que $b(\Phi | a, h)$ suit une loi $N(0, V)$, où Q est une constante et $C=V^{-1}$:

$$b(\Phi | a, h) = Q \frac{\exp\left(-\frac{1}{2} \Phi^T C \Phi\right)}{\sqrt{\det(V)}} \quad (4.14)$$

et que la densité a posteriori de Φ est:

$$\beta_n(a, h, \Phi) \propto b(\Phi, | a, h) \frac{1}{h} \exp\left(-\sum_{i=1}^K \exp(\varphi_i + \mu_i) + \sum_{i=1}^K n_i (\varphi_i + \mu_i)\right) \quad (4.15)$$

Afin d'estimer θ_i par la solution de $\frac{\partial \beta_n(a, h, \Theta)}{\partial \theta_i} = 0$, a par la solution de $\frac{\partial \beta_n(a, h, \Phi)}{\partial a} = 0$, et

h par la solution de $\frac{\partial \beta_n(a, h, \Phi)}{\partial h} = 0$, nous utilisons le résultat suivant (Vernos, 1982):

La coordonnée en θ_1 du mode conjoint de (θ_1, θ_2) est égale à la valeur du mode de la distribution conditionnelle de θ_1 pour θ_2 fixé égal à la valeur en θ_2 du mode conjoint.

Étape 2.1: estimation de Θ

Les estimateurs $\hat{\theta}_i$ sont les solutions du système $\frac{\partial \beta_n(\hat{a}, \hat{h}, \Theta)}{\partial \theta_i} = 0$, pour a et h fixés, ce qui correspond au système: (Bernier, 1981a)

$$e^{\hat{\theta}_i} = n_i \sum_{j=1}^K \hat{C}_{ij} (\hat{\theta}_j - \mu_j) \quad (4.16)$$

Une approximation linéaire de la solution itérative de (4.16) à l'étape r est:

$$e^{\hat{\theta}_i^{r+1}} = e^{\hat{\theta}_i^r} \left[1 + [\hat{\theta}_i^{r+1} - \hat{\theta}_i^r] \right] \quad (4.17)$$

ce qui donne en reportant (4.16) dans (4.17):

$$\left(e^{\hat{\theta}_i^r} + \hat{C}_{ii}^r \right) \hat{\theta}_i^{r+1} + \sum_{j \neq i} \hat{C}_{ij}^r \hat{\theta}_j^{r+1} = e^{\hat{\theta}_i^r} (\hat{\theta}_j^r - 1) + n_i + \sum_{j=1}^K \hat{C}_{ij}^r \mu_j \quad (4.18)$$

Le système d'équations (4.18) doit alors être résolu itérativement afin d'obtenir $\hat{\theta}_i$.

Étape 2.2: estimation de a et de h

Posons $W = \det(A)$; nous avons alors $\det(V) = h^{-K} W^2$, et:

$$\beta_n(a, h, \Phi) \propto \frac{\exp\left(-\frac{1}{2} \Phi^T C \Phi\right)}{h \sqrt{\det(V)}} \Rightarrow \beta_n(a, h, \Phi) \propto \exp\left(-\frac{1}{2} \Phi^T C \Phi\right) h^{\frac{K}{2}-1} W \quad (4.19)$$

où φ_j correspond à $\varphi_{j \bmod K}$ et:

$$\Phi^T C \Phi = h \left[\left(2a^2 + 1 \right) \sum_{j=1}^K \varphi_j^2 \right] - 4a \sum_{j=1}^K \varphi_j \varphi_{j+1} + 2a^2 \sum_{j=1}^K \varphi_j \varphi_{j+2} \quad (4.20)$$

Posons $P = (\Phi^T C \Phi)/h$; nous avons alors:

$$\beta_n(a, h, \Phi) \propto \exp\left(-\frac{hP}{2}\right) h^{\frac{K}{2}-1} W \quad (4.21)$$

L'estimation de h pour a et Φ fixés est obtenue par:

$$\frac{\partial \beta_n(\hat{a}, h, \hat{\Phi})}{\partial h} = \frac{K-2}{2} \left(h^{\frac{K-4}{2}} \exp\left(-\frac{h\hat{P}}{2}\right) \hat{W} \right) - \frac{1}{2} \left(h^{\frac{K-2}{2}} \exp\left(-\frac{h\hat{P}}{2}\right) \hat{W}\hat{P} \right) \quad (4.22)$$

$$\text{où } \hat{P} = \left[\left((2\hat{a}^2 + 1) \sum_{j=1}^K \hat{\phi}_j^2 \right) - 4\hat{a} \sum_{j=1}^K \hat{\phi}_j \hat{\phi}_{j+1} + 2\hat{a}^2 \sum_{j=1}^K \hat{\phi}_j \hat{\phi}_{j+2} \right]$$

$$\text{et } \hat{W} = \left\{ -2\hat{a}^k + \prod_{m=1}^{K-1} \left(1 + 2\hat{a} \cos\left(\frac{m\pi}{K}\right) \right) - 2\hat{a}^2 \prod_{m=1}^{K-2} \left(1 + 2\hat{a} \cos\left(\frac{m\pi}{K-1}\right) \right) \right\}$$

$\frac{\partial \beta_n(\hat{a}, h, \hat{\Phi})}{\partial h} = 0$, et $\hat{W} \neq 0, h \neq 0$ impliquent que:

$$\hat{h} = \frac{K-2}{\left[\left((2\hat{a}^2 + 1) \sum_{j=1}^K \hat{\phi}_j^2 \right) - 4\hat{a} \sum_{j=1}^K \hat{\phi}_j \hat{\phi}_{j+1} + 2\hat{a}^2 \sum_{j=1}^K \hat{\phi}_j \hat{\phi}_{j+2} \right]} \quad (4.23)$$

L'estimation de a pour h et Φ fixés est donnée par:

$$\frac{\partial \beta_n(a, \hat{h}, \hat{\Phi})}{\partial a} = \left[\frac{\partial P}{\partial a} \left(-\frac{\hat{h}^{\frac{K}{2}}}{2} \exp\left(-\frac{h\hat{P}}{2}\right) W \right) \right] + \left(\hat{h}^{\frac{K-2}{2}} \exp\left(-\frac{h\hat{P}}{2}\right) \frac{\partial W}{\partial a} \right) \quad (4.24)$$

$$\text{où } \frac{\partial P}{\partial a} = \left[\left(4a \sum_{j=1}^K \hat{\phi}_j^2 \right) - 4 \sum_{j=1}^K \hat{\phi}_j \hat{\phi}_{j+1} + 4a \sum_{j=1}^K \hat{\phi}_j \hat{\phi}_{j+2} \right]$$

$\frac{\partial \beta_n(a, \hat{h}, \hat{\Phi})}{\partial a} = 0$, et $W \neq 0, \hat{h} \neq 0$ impliquent que:

$$-\frac{\hat{h}}{2} \left[4a \sum_{j=1}^K \hat{\varphi}_j^2 - 4 \sum_{j=1}^K \hat{\varphi}_j \hat{\varphi}_{j+1} + 4a \sum_{j=1}^K \hat{\varphi}_j \hat{\varphi}_{j+2} \right] + \frac{\partial W}{\partial a} = 0 \quad (4.25)$$

Cette dernière équation doit être résolue pour a .

Étape 2.2': estimation de a , h , et Θ lorsque $\bar{\mu}$ est inconnue

Lorsqu'on suppose $\bar{\mu}$ inconnue, la densité conjointe a priori est:

$$b_n(a, h, \Theta, \mu_0) \propto b(\Theta | a, h, \mu_0) * b(a|h, \mu_0) * b(h | \mu_0) * b(\mu_0) \quad (4.26)$$

et la densité a posteriori correspondante $\beta_n(a, h, \Theta, \mu_0)$ est:

$$\beta_n(a, h, \Theta, \mu_0) \propto b(\Theta | a, h, \mu_0) * b(a|h, \mu_0) * b(h | \mu_0) * b(\mu_0) \quad (4.27)$$

Les valeurs modales $\hat{\theta}_i$ pour a , h , et μ_0 fixés, sont les solutions du système $\frac{\partial \beta_n(\hat{a}, \hat{h}, \hat{\mu}_0, \Theta)}{\partial \theta_i} = 0$, ce qui correspond à l'équation (4.17).

L'estimateur de μ_0 tel que donné par la solution modale de $\frac{\partial \beta_n(\hat{a}, \hat{h}, \hat{\Theta}, \mu_0)}{\partial \mu_0} = 0$, $v^2 \rightarrow \infty$ est à l'étape r (Bernier, 1981a):

$$\hat{\mu}_0^{r+1} = \frac{\sum_{i=1}^K \hat{C}_i^r \hat{\theta}_i^{r+1}}{\hat{C}_-^r}, \text{ où } C_i = \sum_{j=1}^K C_{ij}, C_j = \sum_{i=1}^K C_{ij}, C_- = \sum_{j=1}^K \sum_{i=1}^K C_{ij} \quad (4.28)$$

En reportant l'équation (4.28) dans l'équation (4.18), on obtient pour a , h et μ_0 fixés:

$$\left[e^{\hat{\theta}_i^r} + \frac{\hat{C}_-^r \hat{C}_{ii}^r - \hat{C}_i^{r2}}{\hat{C}_-^r} \right] \hat{\theta}_i^{r+1} + \sum_{j \neq i} \frac{\hat{C}_{ij}^r \hat{C}_-^r - \hat{C}_j^{r2}}{\hat{C}_-^r} \hat{\theta}_j^{r+1} = e^{\hat{\theta}_i^r} (\hat{\theta}_i^r - 1) + n_i \quad (4.29)$$

les estimateurs de h et de a sont donnés par les équations (4.23) et (4.25).

L'algorithme d'estimation

L'algorithme d'estimation selon l'approche bayésienne semi-paramétrique de la fonction d'intensité présenté sur la figure 4.2 se rapporte aux étapes 2 et 2':

1. Initialiser Θ (et μ_0 pour le cas où μ est inconnue)

2. À l'étape r :

a. calculer $\hat{\phi}_i^r = \hat{\theta}_i^r - \hat{\mu}_0^r$

b. Résoudre l'équation (4.25) suivante pour obtenir \hat{a}^r :

$$-\frac{\hat{h}}{2} \left(4a \sum_{j=1}^K \hat{\phi}_j^2 - 4 \sum_{j=1}^K \hat{\phi}_j \hat{\phi}_{j+1} + 4a \sum_{j=1}^K \hat{\phi}_j \hat{\phi}_{j+2} \right) + \frac{\partial W}{\partial a} = 0,$$

$$\text{où } W = \frac{1}{\sqrt{h^K \det C}}$$

c. Obtenir \hat{h}^r par l'équation (4.23) suivante:

$$\hat{h}^r = \frac{K-2}{(2\hat{a}^2 + 1) \sum_{j=1}^K \hat{\phi}_j^2 - 4\hat{a} \sum_{j=1}^K \hat{\phi}_j \hat{\phi}_{j+1} + 2\hat{a}^2 \sum_{j=1}^K \hat{\phi}_j \hat{\phi}_{j+2}}$$

d. Calculer $\hat{C}^r = \hat{h} \hat{A}^2$

e. Si μ est connue, résoudre le système d'équation (4.18) suivant:

$$\left(e^{\hat{\theta}_i^r} + \hat{C}_{ii}^r \right) \hat{\theta}_i^{r+1} + \sum_{j \neq i} \hat{C}_{ij}^r \hat{\theta}_j^{r+1} = e^{\hat{\theta}_i^r} (\hat{\theta}_i^r - 1) + n_i + \sum_{j=1}^K \hat{C}_{ij}^r \mu_j$$

e'. Si μ est inconnue, résoudre le système d'équation (4.29) suivant:

$$\left(e^{\hat{\theta}_i^r} + \frac{\hat{C}_{ii}^r \hat{C}_{ii}^r - \hat{C}_{i.}^r{}^2}{\hat{C}_{..}^r} \right) \hat{\theta}_i^{r+1} + \sum_{j \neq i} \frac{\hat{C}_{.j}^r \hat{C}_{ij}^r - \hat{C}_{j.}^r{}^2}{\hat{C}_{..}^r} \hat{\theta}_j^{r+1} = e^{\hat{\theta}_i^r} (\hat{\theta}_i^r - 1) + n_i$$

et calculer $\hat{\mu}_0^{r+1}$ par l'équation (4.28) suivante:

$$\hat{\mu}_0^{r+1} = \frac{\sum_{i=1}^K \hat{C}_i^r \hat{\theta}_i^{r+1}}{\hat{C}_{..}^r}, \text{ où } C_{i.} = \sum_{j=1}^K C_{ij}, C_{.j} = \sum_{i=1}^K C_{ij}, C_{..} = \sum_{i=1}^K \sum_{j=1}^K C_{ij}$$

3. Répéter 1 et 2 jusqu'à ce que l'estimation soit jugée satisfaisante.

Figure 4.2: L'algorithme d'estimation BSP de la fonction d'intensité d'un PNH

4.2.3 Quantification de l'incertitude sur l'intensité estimée par la méthode BSP

Variance de $\hat{\theta}_i$

Sous certaines conditions de dérivabilité de $b(\Theta)$ et de $L(n, \Theta)$ (fonctions positives et deux fois dérivables, Berger, 1985), la loi a posteriori β_n de Θ est approximativement $N(\hat{\Theta}, [\hat{I}(n)]^{-1})$, où les éléments (i, j) de $\hat{I}(n)$, la matrice d'information de Fisher, sont définis par:

$$\hat{I}_{ij}(n) = \left[- \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(L(n, \Theta)) \right]_{\Theta = \hat{\Theta}} \quad (4.30)$$

Nous avons donc:

$$\hat{I}_{ii}(n) = \exp(-\hat{\theta}_i), \hat{I}_{ij}(n) = 0, i \neq j, \text{ et } \text{Var}(\hat{\theta}_i) = \exp(-\hat{\theta}_i) \quad (4.31)$$

Variance de $\hat{\lambda}_i$

La matrice d'information de Fisher de $\hat{\lambda}_i$, $I^*(\hat{\lambda}_i)$, est obtenue à partir de celle de $\hat{\theta}_i$. Puisque $\hat{\theta}_i = h(\hat{\lambda}_i) = \log(p \hat{\lambda}_i)$ nous avons (Lehman, 1983):

$$I^*(\hat{\lambda}_i) = I(h(\hat{\lambda}_i))(h'(\hat{\lambda}_i))^2 \quad (4.32)$$

$$\Rightarrow I^*(\hat{\lambda}_i) = \exp(\log(p \hat{\lambda}_i)) \left(\frac{p}{p \hat{\lambda}_i} \right)^2 \quad (4.33)$$

$$\Rightarrow I^*(\hat{\lambda}_i) = \exp(\hat{\theta}_i) \left(\frac{p}{\frac{p \exp(\hat{\theta}_i)}{p}} \right)^2 \quad (4.34)$$

$$\Rightarrow I^*(\hat{\lambda}_i) = \frac{p^2}{\exp(\hat{\theta}_i)} \quad (4.35)$$

$$\Rightarrow \text{var}(\hat{\lambda}_i) = I^{*-1}(\hat{\lambda}_i) = \frac{\exp(\hat{\theta}_i)}{p^2} = \frac{\hat{\lambda}_i}{p} \quad (4.36)$$

4.3 Méthode du maximum de vraisemblance

Nous présentons maintenant la méthode du maximum de vraisemblance (MV) pour l'estimation de la fonction d'intensité $\lambda(t)$ d'un PNH. Considérons la fonction d'intensité $\lambda(t) = \lambda \exp\left(k \sin\left(\frac{2\pi t}{365} + \theta\right)\right)$ d'un PNH défini par les temps d'occurrence $\vec{t} = (t_1, t_2, \dots, t_n)$ sur l'horizon $(0, T)$. Nous voulons estimer les paramètres, λ , k , et θ par la méthode du MV. Cette méthode consiste à optimiser la fonction de vraisemblance suivante (Lewis, 1972):

$$L(t_j, \lambda, k, \theta) = \prod_{j=1}^n \lambda(t_j) \exp\left(-\int_0^T \lambda(u) du\right) \quad (4.37)$$

$$L(t_j, \lambda, k, \theta) = \prod_{j=1}^n \lambda \exp\left(k \sin\left(\frac{2\pi t_j}{365} + \theta\right)\right) \exp\left(-\int_0^T \lambda \exp\left(k \sin\left(\frac{2\pi u}{365} + \theta\right)\right) du\right) \quad (4.38)$$

Maximiser $L(t_j, \lambda, k, \theta)$ revient à maximiser $LL(t_j, \lambda, k, \theta) = \log(L(t_j, \lambda, k, \theta))$ telle que:

$$LL(t_j, \lambda, k, \theta) = n \ln(\lambda) + \sum_{j=1}^n k \sin\left(\frac{2\pi t_j}{365} + \theta\right) - \lambda \int_0^T \exp\left(k \sin\left(\frac{2\pi u}{365} + \theta\right)\right) du \quad (4.39)$$

Les estimateurs sont alors (voir annexe 1):

$$\hat{\lambda} = \frac{n}{\int_0^T \exp\left(\hat{k} \sin\left(\frac{2\pi u}{365} + \hat{\theta}\right)\right) du} \quad (4.40)$$

$$\hat{\theta} = \tan^{-1}\left(\frac{A}{B}\right), \quad A = \sum_{j=1}^n \cos\left(\frac{2\pi t_j}{365}\right), \quad \text{et } B = \sum_{j=1}^n \sin\left(\frac{2\pi t_j}{365}\right) \quad (4.41)$$

et \hat{k} est la solution de l'équation suivante:

$$\frac{1}{n} \sqrt{A^2 + B^2} = \frac{I_1(k)}{I_0(k)} \quad (4.42)$$

$$\text{où: } I_0(k) = \int_0^T \exp\left(k \sin\left(\frac{2\pi u}{365} + \hat{\theta}\right)\right) du \quad (4.43)$$

$$I_1(k) = \frac{\partial I_0(k)}{\partial k} = \int_0^T \exp\left(k \sin\left(\frac{2\pi u}{365} + \hat{\theta}\right)\right) \sin\left(\frac{2\pi u}{365} + \hat{\theta}\right) du \quad (4.44)$$

4.3.1 Quantification de l'incertitude sur l'intensité estimée par la méthode du MV

Posons $\vec{p} = (p_1, p_2, p_3)$ où $p_1 = \lambda, p_2 = k, p_3 = \theta$. La loi asymptotique des estimateurs du MV, $\hat{\vec{p}} = (\hat{\lambda}, \hat{k}, \hat{\theta})$ est approximativement $N(\hat{\vec{p}}, \hat{I}^{-1})$, où les éléments (r,s) de $\hat{I}(\hat{\lambda}, \hat{k}, \hat{\theta})$ sont: $\hat{I}_{rs} = \left[-\frac{\partial^2}{\partial p_r \partial p_s} \log(L(t_j, \vec{p})) \right]_{\vec{p}=\hat{\vec{p}}}$ (Lehman, 1983). La matrice de variance-covariance de $\hat{\vec{p}} = (\hat{\lambda}, \hat{k}, \hat{\theta})$ est $\hat{I}^{-1}(\hat{\lambda}, \hat{k}, \hat{\theta})$ telle que présentée à l'annexe 1.

La variance de $\hat{\lambda}(t) = \hat{\lambda} \exp(\hat{k} \sin(\frac{2\pi t}{365} + \hat{\theta}))$ est $\text{var}(\hat{\lambda}(t)) = I^{-1}_{1,1} * (\bar{\xi})$ telle que présentée à l'annexe 1.

4.4 Applications

Nous commençons par comparer à la section 4.4.1, les méthodes d'estimation BSP et du MV sur un PNH généré par simulation selon l'algorithme de Ross (1990) décrit à l'annexe 4. Par la suite, nous comparons ces méthodes sur des données réelles de précipitation enregistrées sur la station française Bar-sur-Seine (4.4.2). Nous procéderons selon les deux hypothèses 3 et 3', c'est-à-dire, pour les cas où $\bar{\mu}$ est connue et $\bar{\mu}$ est inconnue.

La comparaison sur un processus simulé fournit une évaluation plus juste des méthodes, car les données simulées proviennent exactement par définition d'un PNH, ce qui n'est pas nécessairement le cas des données observées, pour lesquelles le PNH est au mieux un modèle approximatif.

4.4.1 Méthode BSP appliquée à un processus simulé

La fonction d'intensité du PNH simulé utilisée pour l'étude est la suivante:

$$\lambda(t) = .01 \exp[2 \sin(\omega t + 1)], \quad \omega = \frac{2\pi}{365} \quad (4.45)$$

Cette fonction est présentée sur la figure 4.3.

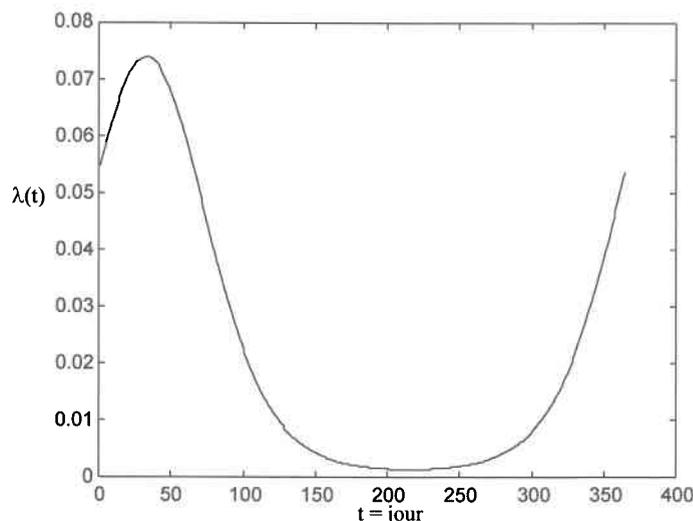


Figure 4.3: La fonction d'intensité réelle

Le PNH est généré pour des périodes d'observation égales à 15 ans et à 55 ans, ce qui correspond à 130 et à 460 événements respectivement. Les algorithmes de génération et d'estimation pour $\bar{\mu}$ connue et $\bar{\mu}$ inconnue sont programmés en MAPLE.

Cas où $\bar{\mu}$ est connue

Dans ce cas, tous les θ_i sont initialisés à la même valeur correspondant au log du nombre moyen d'événements observés sur l'horizon de génération du processus (15 ans ou 55 ans);

μ_i est fixée égale à sa vraie valeur, c'est-à-dire que $\mu_i = \frac{1}{\Delta_i} \int_{\Delta_i} \lambda(t) dt$, où $\lambda(t)$ est la fonction

d'intensité réelle de l'équation (4.45), et où Δ_i correspond à l'intervalle i . Les figures 4.4 et 4.5 présentent la comparaison entre la fonction d'intensité réelle et la fonction d'intensité estimée en utilisant une discrétisation de l'année en 12 et 26 intervalles, et ce à partir de 15 ans et de 55 ans d'observations, pour $\bar{\mu}$ connue.

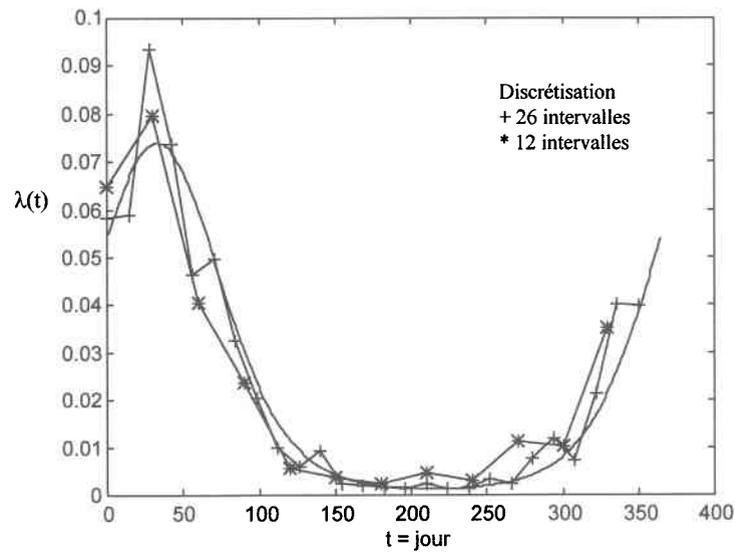


Figure 4.4: Comparaison de la fonction d'intensité réelle et son estimation bayésienne, période de génération 15 ans, $\bar{\mu}$ connue

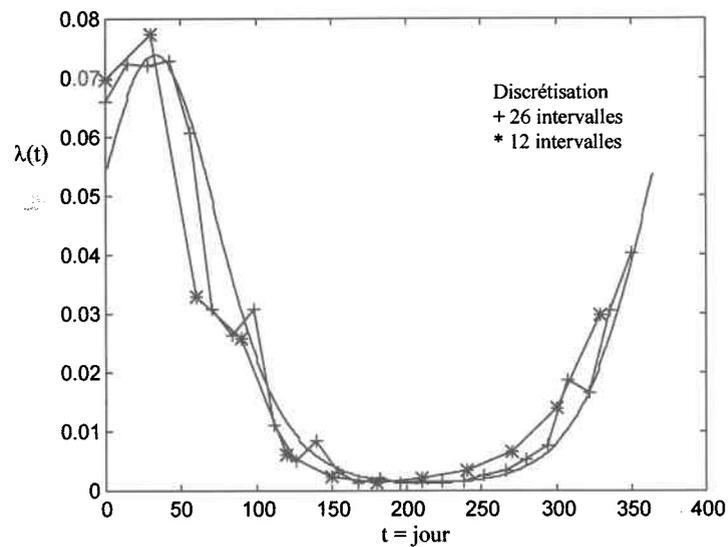


Figure 4.5: Comparaison de la fonction d'intensité réelle et son estimation bayésienne, période de génération 55 ans, $\bar{\mu}$ connue

Cas où $\bar{\mu}$ est inconnue

Dans ce cas, θ_i est initialisé au log du nombre d'événements observés sur l'intervalle i ou à 0; m , la moyenne a priori de μ_0 , est égale au log du nombre moyen d'événements observés. Les figures 4.6 et 4.7 présentent la comparaison entre la fonction d'intensité réelle et la fonction d'intensité estimée en utilisant une discrétisation de l'année en 12 et 26 intervalles, et ce à partir de 15 ans et de 55 ans d'observations, pour $\bar{\mu}$ inconnue.

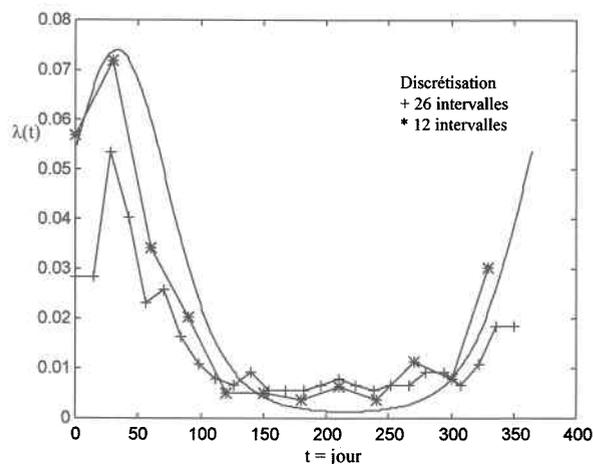


Figure 4.6: Comparaison de la fonction d'intensité réelle et son estimation bayésienne, période de génération 15 ans, $\bar{\mu}$ inconnue

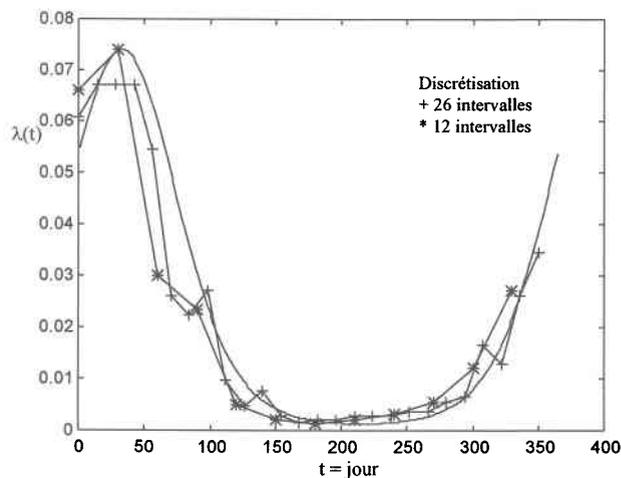


Figure 4.7: Comparaison de la fonction d'intensité réelle et son estimation bayésienne, période de génération 55 ans, $\bar{\mu}$ inconnue

4.4.2 Comparaison de la méthode du MV et la méthode BSP, $\bar{\mu}$ inconnue, pour les précipitations sur Bar-sur-Seine

Les deux méthodes d'estimation sont comparées en utilisant le processus d'occurrence des précipitations enregistrées sur la station française de Bar-sur-Seine présentées au chapitre 3, et pour lesquelles 27 années de données journalières sont disponibles. Puisqu'on ne peut connaître avec certitude le processus sous-jacent, on suppose que dans le cas du BSP, $\bar{\mu}$ est inconnue. Les résultats de l'estimation par le MV sont présentés au tableau 4.1.

$\hat{\lambda}$	$\text{var}(\hat{\lambda})$	\hat{k}	$\text{var}(\hat{k})$	$\hat{\theta}$	$\text{var}(\hat{\theta})$
.44	4×10^{-5}	.05	4.8×10^{-4}	1.56	.18

Tableau 4.1: Les estimateurs du MV pour les précipitations

La fonction d'intensité estimée est donc $\hat{\lambda}(t) = .44 \exp(.05 \sin(\frac{2\pi t}{365} + 1.56))$. En utilisant le résultat de l'annexe 4 nous avons:

$$\text{var}(\hat{\lambda}(t)) = \frac{-11828.2}{\left(\exp(-\hat{k} \sin(\omega t + \theta))\right)^2 \left(-2.65 \cdot 10^{-10} - .001 \hat{\lambda} \sin(\omega t + \theta) + 10^{-6} \hat{\lambda}^2 \hat{k} \sin(\omega t + \theta) \cos(\omega t + \theta) - 10^{-7} \hat{\lambda} \hat{k} \cos(\omega t + \theta)\right)} \quad (4.46)$$

$$\text{où } \omega = \frac{2\pi}{465}, t=1..365$$

Les calculs montrent que pour $t=1 \dots 365$, $\text{var}(\hat{\lambda}(t))$ est comprise entre .000041 et .000049. On peut donc dire que cette variance est constante autour de 4×10^{-5} . Les résultats de l'estimation par le MV sont présentés sur la figure 4.8.

La méthode BSP avec $\bar{\mu}$ inconnue est ensuite utilisée pour estimer la fonction d'intensité du processus de précipitations sur Bar-sur-Seine. L'année est découpée en 12 et en 30 intervalles. Dans les deux cas, la variance de $\hat{\lambda}_i$ est à peu près constante à .016. Les résultats sont présentés sur les figures 4.9 et 4.10.

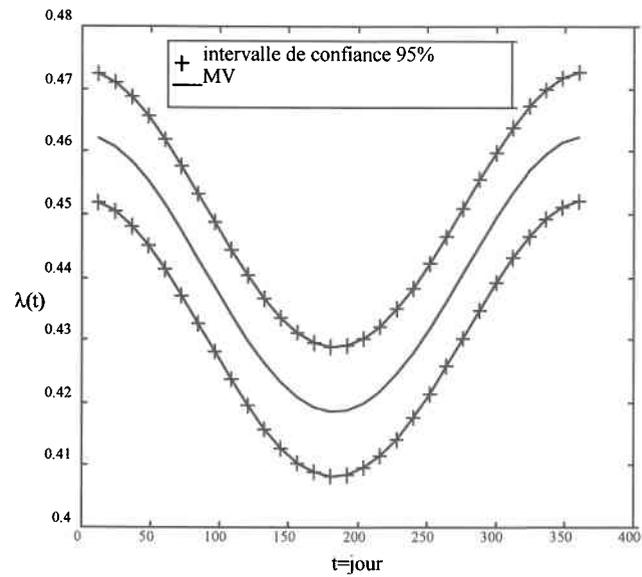


Figure 4.8: Estimateur MV de l'intensité des précipitations (Bar-sur-Seine)

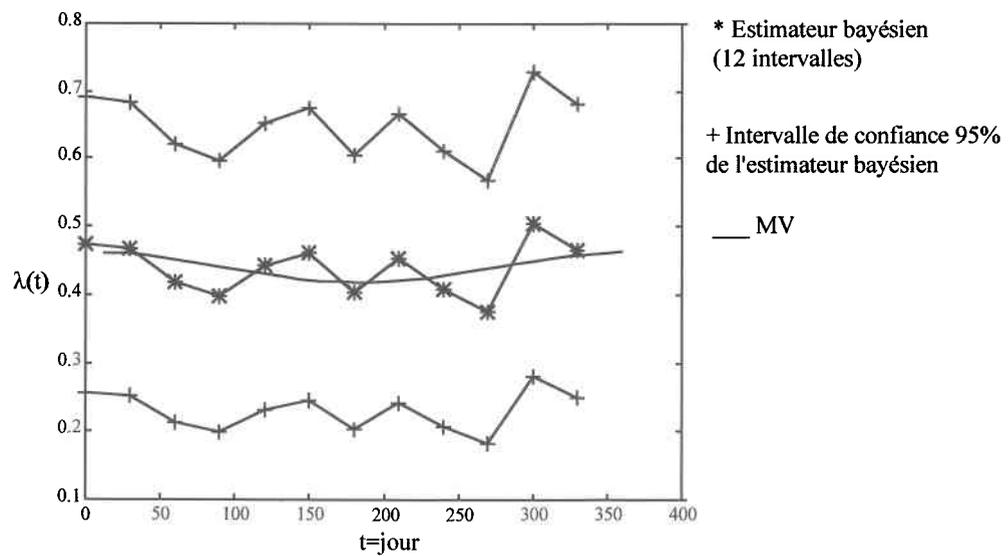


Figure 4.9: Estimateur BSP de l'intensité des précipitations (Bar-sur-Seine), 12 intervalles

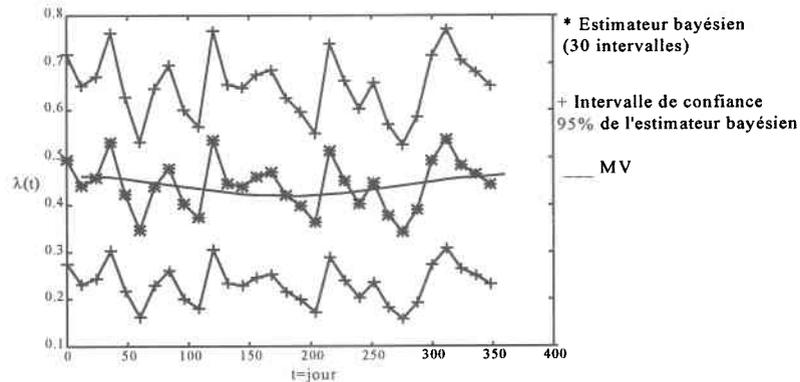


Figure 4.10: Estimateur BSP de l'intensité des précipitations (Bar-sur-Seine), 30 intervalles

4.5 Discussion

4.5.1 Processus simulé

D'après les figures 4.4 à 4.7 qui présentent les estimateurs par la méthode BSP et la fonction d'intensité réelle, il apparaît que la performance de cette méthode est satisfaisante. En comparant les figures 4.4 et 4.6, 4.5 et 4.7, on vérifie que l'estimation est légèrement meilleure lorsque $\bar{\mu}$ est supposée connue. De plus, il n'est pas surprenant de constater que l'estimation est meilleure lorsqu'on dispose de 55 ans d'observations plutôt que de 15 ans. Les résultats obtenus ont montré que les deux approches convergent après quatre à cinq itérations. De plus, nous avons remarqué que les méthodes d'estimation sont peu sensibles aux valeurs initiales de $\bar{\mu}$ (μ_0 , pour $\bar{\mu}$ inconnue). Quant à la discrétisation, l'utilisation de 26 intervalles ne conduit pas nécessairement à une meilleure estimation que 12 intervalles, ce qui est probablement dû au faible nombre d'événements par intervalle, lorsque l'année est divisée en 26 intervalles. Le tableau 4.2 présente les variances moyennes des estimateurs de $\hat{\lambda}_i$. Ni le nombre d'intervalles considéré, ni l'hypothèse de la connaissance de la moyenne ne semblent avoir d'incidence majeure sur la variance de l'estimateur. Par contre, la longueur de la période de simulation a une influence très significative.

Variance moyenne de $\hat{\lambda}_i$ par la méthode bayésienne				
	$\bar{\mu}$ connue		$\bar{\mu}$ inconnue	
	12 intervalles	26 intervalles	12 intervalles	26 intervalles
15 ans	1.5×10^{-3}	1.5×10^{-3}	1.4×10^{-3}	1.0×10^{-3}
55 ans	4.1×10^{-4}	4.1×10^{-4}	3.8×10^{-4}	3.8×10^{-4}

Tableau 4.2: Les variances moyennes de $\hat{\lambda}_i$

4.5.2 Processus de précipitations (Bar-sur-Seine)

La comparaison de la méthode du MV et de la méthode BSP sur les données réelles de précipitations sur la station Bar-sur-Seine est moins aisée que pour le processus simulé, car l'hypothèse d'un PNH peut ne pas être exacte. D'après les figures 4.8 à 4.10, on voit que la méthode du MV donne une intensité presque constante autour de .44 tandis que la méthode BSP est beaucoup moins lisse et reflète davantage le nombre d'événements observé. On constate encore une fois pour l'estimation par la méthode BSP que le nombre d'intervalles n'a pas tellement d'importance sur les résultats. La variance du MV (4×10^{-5}) est beaucoup plus faible que celle de la méthode BSP (1.6×10^{-2}). Cette différence dans les variances des deux estimateurs peut paraître surprenante. Toutefois elle s'explique par le fait que l'estimateur du MV ne contient pas d'incertitude sur le modèle, puisqu'on estime les paramètres du modèle postulé et donc connu (la forme paramétrée de $\lambda(t)$). Dans ce cas, l'incertitude provient principalement des données. Tandis que dans le cas de l'estimateur du BSP, on doit tenir compte et de l'incertitude sur le modèle et de l'incertitude sur les données. On ne compare donc pas deux méthodes d'estimation différentes mais bel et bien deux modèles différents.

4.5.3 Conclusions

En conclusion, l'avantage principal de la méthode BSP, par rapport à la méthode du MV, est qu'il n'est pas nécessaire de supposer, a priori, une forme complètement paramétrée de la fonction d'intensité. Par contre, cette approche présente le désavantage d'avoir à découper le temps en intervalles, n'estimant ainsi que des valeurs moyennes d'intensité. Son avantage

par rapport à la méthode conventionnelle non paramétrique où $\hat{\lambda}_i = \frac{\exp\left(\frac{1}{n_i}\right)}{p}$ et

$\text{var}(\hat{\lambda}_i) = \frac{\hat{\lambda}_i}{p}$ concerne le cas où il y a des données manquantes, c'est-à-dire lorsque le

nombre d'événements est 0 pour un intervalle donné. Pour ce qui a trait à l'évaluation par simulation de cette méthode, la répétition de l'expérience (réplicats) permettrait d'avoir une idée de l'erreur d'échantillonnage, et un critère de proximité entre la fonction d'intensité estimée et la courbe théorique serait utile pour étudier avec plus de rigueur l'impact du choix des valeurs initiales et du nombre d'intervalles.

5. MÉTHODOLOGIE DE CONSTRUCTION ET D'ÉVALUATION DES MODÈLES D'OCCURRENCE

Nous avons présenté au chapitre 4 une méthode d'estimation de l'intensité d'un PNH. L'estimation, bien qu'importante, n'est qu'un aspect de la modélisation d'une série d'événements. Nous développons dans ce chapitre la méthodologie globale que nous avons suivie afin de construire et d'évaluer des modèles d'occurrence d'étiage et de surplus.

Une série d'événements est définie par les temps d'occurrence des événements et modélisée par un processus ponctuel qui préserve les caractéristiques statistiques des données observées. Afin d'obtenir ce modèle, on doit bien identifier le type d'événement dont on veut modéliser l'occurrence, choisir un type de processus ponctuel, une fonction d'intensité, estimer les paramètres, et vérifier l'adéquation du modèle construit. Toutefois, il est important de se rappeler que ce modèle n'est qu'une abstraction mathématique qui vise à donner une approximation et qu'il ne décrit pas nécessairement le mécanisme physique sous-jacent. Le but est d'obtenir un modèle représentatif, mais assez simple (parcimonieux et facile à estimer) pour permettre l'analyse de systèmes dépendant du processus physique.

Le choix de méthodes statistiques pour évaluer les processus ponctuels complexes est restreint à cause de la difficulté de calcul des propriétés théoriques. La discrimination empirique entre divers processus proposés peut s'avérer difficile lorsque leurs réalisations sont semblables. Au lieu de baser notre choix sur des tests spécifiques, nous tendrons à favoriser un type de processus qui aura les propriétés observées dans les données, qui sera le plus parcimonieux et le plus simple à construire, et qui aura une justification théorique suffisante. Nous vérifierons ensuite l'adéquation de ce processus par des tests statistiques et nous effectuerons une analyse de sensibilité. La méthodologie suivie dans ce travail est présentée au tableau 5.1 et décrite dans ce qui suit.

Il est très important de souligner que la discussion aux chapitres 3 et 4 a porté sur les processus ponctuels d'un seul type d'événement, et que dans ce qui suit, nous considérons deux types d'événements: étiage et surplus.

Étapes de la modélisation d'une série d'événements

1. *La définition d'un événement;*
2. *Le choix du type de processus et le choix de la fonction d'intensité;*
3. *L'estimation des paramètres de l'intensité, à partir des données qu'on veut modéliser;*
4. *La vérification par des tests statistiques de l'adéquation du modèle aux données;*
5. *L'analyse de sensibilité.*

Tableau 5.1: Étapes de la modélisation d'une série d'événements

5.1 La définition des événements

La première étape de l'analyse d'une série d'événements est l'identification et la définition de l'événement en question. Nous avons dans le cadre de cette étude deux types d'événements définis à partir du débit journalier et d'un seuil x_0 : l'événement de type surplus (1) est la rupture de l'état de surplus (passage d'au-dessus du seuil en dessous du seuil), et l'événement de type étiage (2) est la rupture de l'état d'étiage (passage d'en dessous du seuil au-dessus du seuil). L'événement de type i , $i=1,2$ correspond donc à la sortie de l'état i , et à l'entrée dans l'état k , $i \neq k$. x_j^i est le $j^{\text{ème}}$ temps de séjour (durée) dans l'état i . t_j^i est le temps d'arrivée du $j^{\text{ème}}$ événement de type i . Les temps d'occurrence des événements $t_1^1, t_1^2, t_2^1, t_2^2, \dots, t_j^1, t_j^2$, définissent le processus alterné de surplus et d'étiage. Ce processus, illustré sur la figure 5.1, est défini d'une manière équivalente par les durées dans chaque état: $x_1^2, x_1^1, x_2^2, x_2^1, \dots, x_j^2, x_j^1$.

Quant au seuil x_0 , celui-ci dépend de l'usage envisagé. Il ne peut être défini d'une façon absolue: il convient donc, pour les buts de cette étude de construire plusieurs séries d'événements à partir de divers seuils. Nous utilisons en pratique deux seuils: le 25^{ème} quantile, et le 10^{ème} quantile, c'est-à-dire, les débits dépassés 75% et 90% du temps respectivement. À noter que nous utilisons aussi bien le terme étiage que le terme déficit.

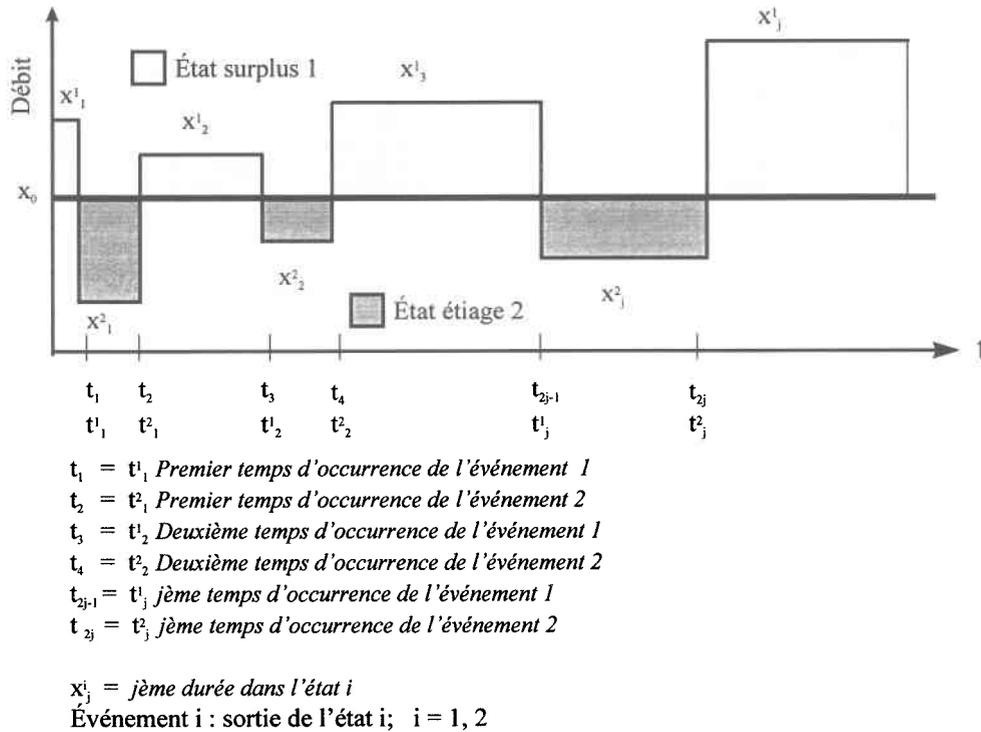


Figure 5.1: Définition des temps d'occurrence des événements

Cette approche pour définir un événement ne permet pas de tenir compte des erreurs de mesure associées aux données. On observe ainsi un grand nombre d'événements qui génèrent des durées d'un jour, et dont l'écart du débit par rapport au seuil est très faible et même zéro. Afin de pallier cette situation, nous acceptons comme événement les changements d'état dont l'écart du débit par rapport au seuil est d'au moins 10%.

5.2 La définition du processus alterné

L'existence de deux types de processus (surplus et étiages) qui alternent, peut être envisagée de deux façons:

- nous avons un processus alterné de surplus et d'étiage défini par les temps d'occurrence $t^1_1, t^2_1, t^1_2, t^2_2, \dots, t^1_j, t^2_j$ ou par les durées $x^2_1, x^1_1, x^2_2, x^1_2, \dots, x^2_j, x^1_j$. Il est composé de deux processus qui alternent, PA^k $k=1,2$. Ces processus

sont définis par leurs durées respectives x_j^k (figure 5.2). La fonction d'intensité correspondant au taux de rupture de l'état k est $\rho^k(t, H_t)$. Elle est conditionnelle à ce qu'on soit dans l'état k au temps t ; ou

- nous avons deux processus ponctuels séparés:
 - * le processus de rupture des surplus P^1 , défini par les temps d'occurrence de l'événement 1, t_1^1, t_2^1, t_3^1 etc., et les intervalles entre ces événements z_1^1, z_2^1 , etc. (figure 5.3); et
 - * le processus de rupture des étiages P^2 , défini par, les temps d'occurrence de l'événement 2, t_1^2, t_2^2, t_j^2 etc., et les intervalles entre ces événements z_1^2, z_2^2 , etc. (figure 5.4).

Le processus alterné de surplus et d'étiage composé du processus de surplus PA^1 et du processus d'étiage PA^2

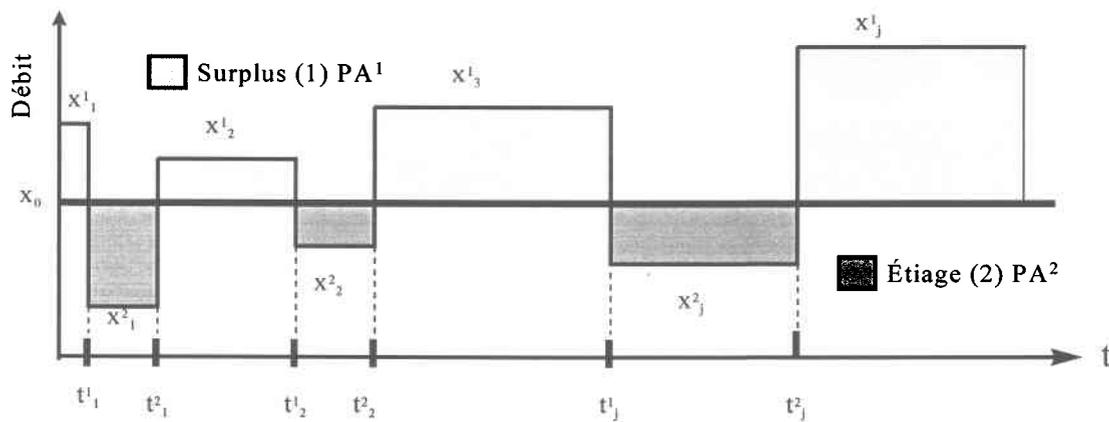


Figure 5.2: Le processus alterné de surplus et d'étiage

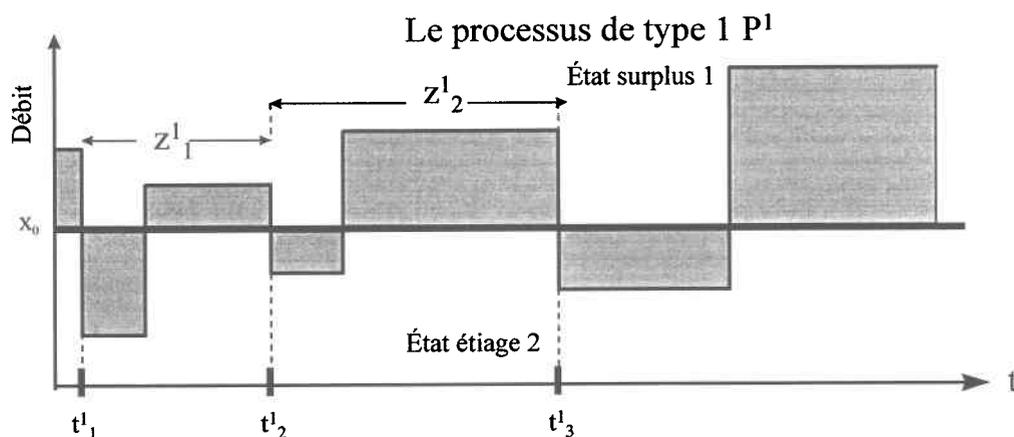


Figure 5.3: Le processus de type 1 (rupture de l'état surplus)

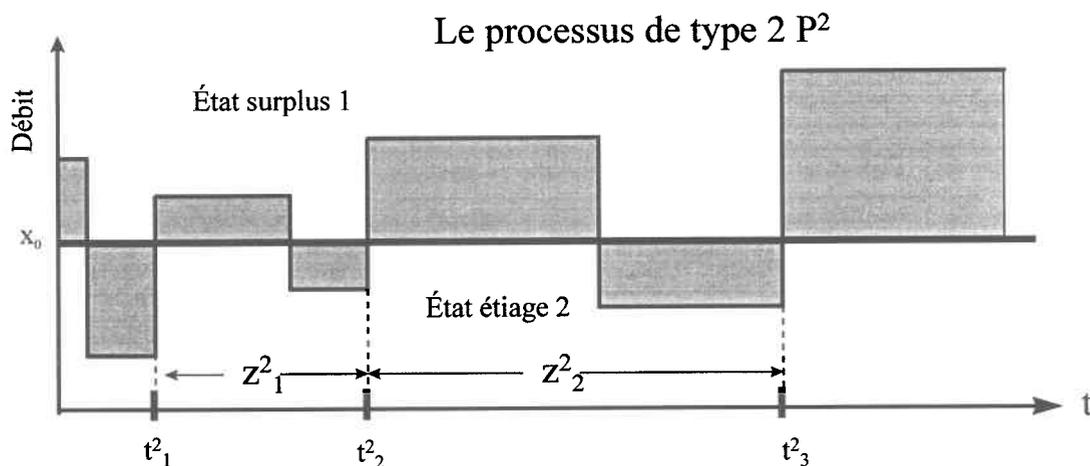


Figure 5.4: Le processus de type 2 (rupture de l'état étiage)

L'intensité du processus d'un seul type P^k correspondant au taux d'occurrence d'une rupture de type k est $\rho^k(t, H_t)$. Elle est liée à l'intensité conditionnelle $\rho^k(t, H_t)$ par:

$$\rho^k(t, H_t) = \frac{\rho^k(t, H_t)}{p^k(t)} \quad (\text{North, 1981}), \text{ où } p^k(t) \text{ est la probabilité d'être dans l'état k.}$$

Nous adoptons dans ce travail la première approche qui consiste à définir un modèle pour le processus alterné à partir des modèles des processus PA^k. Le but de la modélisation de la série d'étiage et de surplus par un processus alterné est d'obtenir:

- la loi de N_t^k , le nombre d'événements de type k dans $(0,t]$; ou d'une façon équivalente,
- la loi du temps de séjour X_j^k dans l'état k , $k \neq i$ car:

$$P\left[\left(N_{t+x}^k - N_t^k\right) > 0 \mid T = t_j^i\right] = P\left[X_j^k \leq x \mid T = t_j^i\right] \quad (5.1)$$

où T est l'instant du dernier événement.

Afin d'obtenir ce modèle, on doit (tableau 5.2):

- spécifier si le processus alterné est composé de deux processus PA^k de Poisson non homogène (PNH), de renouvellement non homogène (RNH), de Poisson non homogène conditionnel (PNHC), ou de tout autre type/combinaison de processus ponctuels; et
- spécifier la forme de l'intensité du processus d'événements de type k , $\rho^k(t; H_t)$, telle que $P[1 \text{ événement de type } k \text{ en } (t, t+\Delta t) \mid \text{processus est en l'état } i \text{ au temps } t, k \neq i] = \rho^k(t; H_t) \Delta t$.

Étapes de la définition du processus alterné

1. Le choix des types de processus ponctuel de PA^1 et PA^2 ;
2. Le choix des fonctions d'intensité $\rho^1(t, H_t)$ et $\rho^2(t, H_t)$.

Tableau 5.2: Étapes de la définition du processus alterné

Dans ce qui suit, nous décrivons les types de processus ponctuels qui sont appliqués et évalués dans le cadre de ce travail.

5.2.1 Le processus de Poisson non homogène (PNH)

Le modèle PNH est le plus simple qu'on puisse imaginer pour des séries d'événements non homogènes. Il a été utilisé par North (1981) afin de modéliser l'occurrence des événements de rupture des étiages et des surplus. Il est sans mémoire, et il suppose que les événements sont indépendants. Afin de tenir compte des saisonnalités, North (1981) propose l'intensité $\rho^k(t; H_t) = \lambda^k(t)$, $k=1,2$ suivante:

$$\lambda^k(t) = \exp[\beta_k + K_k \sin(\omega t + \theta_k)], \quad \omega = \frac{2\pi}{365} \quad (5.2)$$

On a alors:

$$P[(N_{t+x}^k - N_t^k)] = 1 - \exp\left[-\int_t^{t+x} \lambda^k(u) du\right] \quad (5.3)$$

Le tableau 5.3 présente les hypothèses pour l'application d'un PNH.

Les hypothèses pour l'application d'un PNH

1. Absence de mémoire;
2. Indépendance des événements;
3. Les intensités sont périodiques;
4. Les fonctions d'intensité ont la forme suivante, $k=1,2$:

$$\rho^k(t; H_t) = \lambda^k(t) = \exp[\beta_k + K_k \sin(\omega t + \theta_k)], \quad \omega = \frac{2\pi}{365}$$

Tableau 5.3: Les hypothèses pour l'application d'un PNH

Une clarification quant à la définition d'un processus alterné s'impose: un processus de Poisson alterné ne veut pas dire que P^1 et P^2 sont des processus de Poisson mais que PA^1 et PA^2 le sont. Les durées x_j^1 et x_j^2 de PA^1 et PA^2 , correspondent aux durées de deux processus de Poisson dont les fonctions d'intensité sont respectivement $\lambda^1(t)$ et $\lambda^2(t)$. Le processus P^1 dont l'intensité est $\lambda^1(t)$ (P^2 d'intensité $\lambda^2(t)$) ne peut être poissonien car les intervalles entre les événements de type 1 (2) $t_{j+1}^1 - t_j^1$ ($t_{j+1}^2 - t_j^2$), ne suivent pas une loi exponentielle puisqu'ils correspondent à la somme de deux variables aléatoires $x_{j+1}^1 + x_j^2$ ($x_{j+1}^1 + x_{j+1}^2$).

5.2.2 Le processus de renouvellement non homogène (RNH)

Ce modèle est un peu plus complexe que le précédent, car il tient compte d'une mémoire limitée. Il a été utilisé par North (1981) afin de modéliser les étiages et les surplus. Il suppose que l'occurrence d'un événement au temps t dépend non seulement de t mais aussi du temps écoulé $x = t - u$ depuis le dernier événement (au temps u). On a alors, si le dernier événement est de type $i \neq k$:

$$P[(N_{t+x}^k - N_t^k) | U = u] = 1 - \exp\left[-\int_u^{u+x} \phi^k(t - u, t) dt\right] \quad (5.4)$$

Le modèle d'intensité $\rho^k(t;H_t) = \varphi^k(t-u,t)$ proposé par North (1981) est un modèle multiplicatif qui a la forme suivante:

$$\rho^k(t;H_t) = \varphi^k(t-u,t) = \varphi^k(t-u)\lambda^k(t) \quad (5.5)$$

$$\text{où } \lambda^k(t) = \exp\left[\beta_k + K_k \sin(\omega t + \theta_k)\right], \quad \omega = \frac{2\pi}{365}$$

On fait l'hypothèse que le coefficient de variation des durées dans l'état k est constant (indépendant de l'effet saisonnier), et qu'il n'y a qu'un changement d'échelle avec la saison. Ceci revient à dire que l'intensité $\varphi^k(t-u,t)$ est égale au produit de l'intensité d'un processus de renouvellement homogène $\varphi^k(t-u)$ et d'une fonction périodique $\lambda^k(t)$ représentant le paramètre de changement d'échelle telle que:

$$\varphi^k(t-u) = \varphi^k(x'^k), \quad x'^k = \int_u^t \lambda^k(t) dt \quad (5.6)$$

où x'^k est la durée $t-u$ homogénéisée et $\varphi^k(x'^k)$ est le taux de rupture du processus homogénéisé. On suppose aussi que la loi des durées homogénéisées x'^k est une loi Gamma de moyenne unitaire égale à $\frac{\hat{\alpha}_k}{\hat{\rho}_k}$ et de variance égale à $\frac{\hat{\alpha}_k}{\hat{\rho}_k^2}$ où $\hat{\alpha}_k = \hat{\rho}_k = \frac{1}{\hat{\sigma}_k^2}$, $\hat{\sigma}_k^2$ étant la variance estimée des durées homogénéisées x'^k . Le tableau 5.4 présente les hypothèses pour l'application d'un RNH.

Les hypothèses pour l'application du RNH

1. Une mémoire limitée;
2. L'occurrence d'un événement dépend du temps écoulé depuis le dernier événement;
3. Les coefficients de variation des durées sont constants dans le temps;
4. Il y a seulement un changement d'échelle avec la saison; et donc
5. Les fonctions d'intensité sont le produit de l'intensité d'un processus homogénéisé $\varphi^k(t)$ et d'une fonction de changement d'échelle $\lambda^k(t)$:

$$\rho^k(t;H_t) = \varphi^k(t-u,t) = \varphi^k(t-u)\lambda^k(t)$$

$$\lambda^k(t) = \exp\left[\beta_k + K_k \sin(\omega t + \theta_k)\right], \quad \omega = \frac{2\pi}{365};$$

6. La loi des durées homogénéisées est une loi Gamma de moyenne unitaire.

Tableau 5.4: Les hypothèses pour l'application du RNH

5.2.3 Le processus de Poisson non homogène conditionnel (PNHC)

Les processus présentés aux sections 5.2.1 et 5.2.2 sont plutôt de type boîte noire et ne tiennent pas compte du processus physique sous-jacent. Sans aller jusqu'à construire un modèle physique, il serait intéressant de tenir compte de manière explicite, dans la formulation du processus ponctuel, de certaines variables explicatives qui influencent le processus de débit et donc les processus de rupture de surplus et d'étiage. Nous proposons d'utiliser un processus de Poisson non homogène conditionnel (PNHC) dont les processus d'intensité $\rho^1(t;H_t)$ et $\rho^2(t;H_t)$ sont fonction de certaines variables explicatives supposées connues, telle que les précipitations et les températures.

Plusieurs facteurs physiques influencent l'écoulement tels que la capacité d'infiltration du sol ainsi que l'épaisseur, la texture et la structure de la nappe aquifère, la géomorphologie, et le degré d'incision du lit de la rivière (Boisvert, 1969). Toutefois, les variables les plus importantes sont les précipitations, la température, et le pouvoir d'évaporation et d'évapotranspiration. Nous avons retenu comme variables explicatives les précipitations et les températures journalières. Puisque le gel de l'hiver supprime tous les apports sur le bassin, nous tenons indirectement compte des températures, en ramenant les précipitations à zéro lorsque la température est inférieure à -2°C .

Nous avons évalué dans le cadre de ce travail plusieurs fonctions d'intensité d'un PNHC, que nous avons jugées insatisfaisantes suite à la démarche d'examen de leur adéquation qui est présentée à la section 5.4. Ces fonctions sont décrites à l'annexe 5. Nous avons toutefois retenu un modèle pour $\rho^1(t;H_t)$ et un modèle pour $\rho^2(t;H_t)$ qui font en sorte que le PNHC représente généralement bien les données observées.

Modèle proposé pour $\rho^1(t;H_t)$ du processus de rupture des surplus

L'étiage d'été se produit en l'absence de précipitations significatives durant une période prolongée, alors que la rivière est alimentée par les ressources souterraines. En hiver, l'étiage se produit lorsque les précipitations sont sous forme de neige. Il est donc raisonnable de supposer une fonction d'intensité pour les événements de rupture de surplus (entrée en déficit) qui soit fonction du nombre de jours sans précipitations.

Dans le cadre de ce travail, nous définissons une journée sans précipitations par une absence de précipitations ou par une température inférieure à -2°C . Nous proposons la forme suivante pour $\rho^1(t;H_t)$:

$$\rho^1(t;H_t) = \exp\left[\beta_1 + K_1 \text{nbjoursec}(t) + \theta_1 \sin\left(\frac{2\pi t}{365}\right)\right] \quad (5.7)$$

où $\omega = \frac{2\pi}{365}$, et $\text{nbjoursec}(t)$ est le nombre de jours sans précipitations précédant t divisé par 365.

Modèle proposé pour $\rho^2(t;H_t)$ du processus de rupture des étiages

L'alimentation des cours d'eau se fait instantanément par les chutes de pluie sur les surfaces d'eau libre, par le ruissellement de surface avec un décalage, et par l'écoulement souterrain avec un temps de réponse supérieur à celui des deux premières sources d'alimentation. Les quantités de précipitations pluvieuses influencent donc directement le processus de sortie d'un état de déficit. Nous proposons la forme suivante pour $\rho^2(t;H_t)$:

$$\rho^2(t, H_t) = \exp(K_2 \text{pnorm}(t) + \theta_2) \quad (5.8)$$

$\text{pnorm}(t)$ est la hauteur de pluie tombée pendant les 90 jours précédant t normée par cette

quantité maximale observée: $\frac{\sum_{kk=t-90-1}^{kk=t} P(kk)}{\max_t \left(\sum_{kk=t-90-1}^{kk=t} P(kk) \right)}$. Le tableau 5.5 présente les hypothèses pour

l'application d'un PNHC.

Les hypothèses pour l'application du PNHC

1. Les processus d'intensité sont fonction de variables explicatives exogènes;
2. Les variables explicatives (supposées connues) sont les précipitations $P(t)$ et la température $T(t)$ telles que:

- si $T(t) < -2$ alors $P(t)=0$;

- $nbjoursec(t)$ est le nombre des jours $s < t$ tels que $P(s)=0$, divisé par 365;

$$- pnorm(t) = \frac{\sum_{kk=t-90-1}^{kk=t} P(kk)}{\max_t \left(\sum_{kk=t-90-1}^{kk=t} P(kk) \right)};$$

3. Les processus d'intensité ont la forme suivante:

$$\rho^1(t, H_t) = \exp \left[\beta_1 + K_1 nbjour sec(t) + \theta_1 \sin \left(\frac{2\pi t}{365} \right) \right];$$

$$\rho^2(t, H_t) = \exp(K_2 pnorm(t) + \theta_2).$$

Tableau 5.5: Les hypothèses pour l'application du PNHC

5.3 Estimation

5.3.1 Le processus de Poisson non homogène (PNH)

L'estimation de $\lambda^1(t)$ et $\lambda^2(t)$ peut s'effectuer soit directement à partir des temps d'occurrence du processus alterné P , $t_1^1, t_1^2, t_2^1, t_2^2, t_j^1, t_j^2$, soit par l'estimation de $\rho^1(t, H_t) = \lambda^1(t)$ à partir de t_1^1, t_2^1, t_j^1 et de $\rho^2(t, H_t) = \lambda^2(t)$ à partir de t_1^2, t_2^2, t_j^2 , et ce après la décomposition de P en deux processus d'un seul type, P^1 et P^2 . Dans le deuxième cas, on fait l'hypothèse inexacte mais asymptotiquement vraie (Grigelionis, 1963) que les processus d'événements d'un seul type P^1 et P^2 sont poissoniens. On peut ensuite utiliser la méthode bayésienne semi-paramétrique (chapitre 4) ou la méthode du maximum de vraisemblance. Toutefois, l'hypothèse de P^k poissonien est inexacte car les intervalles entre les événements de type k ne suivent plus une loi de Poisson, mais plutôt une loi Gamma, puisqu'un intervalle correspond à la somme de deux variables distribuées selon une loi exponentielle. En ce qui a trait à nos données, nous avons confirmé le rejet de cette hypothèse à l'aide de tests d'adéquation dont les résultats ne sont pas présentés ici.

Nous procédons dans ce travail à l'estimation de la fonction $\lambda^k(t)$, $k=1,2$, directement par la méthode du maximum de vraisemblance. La figure 5.5 décrit un processus alterné de surplus et d'étiage observé.

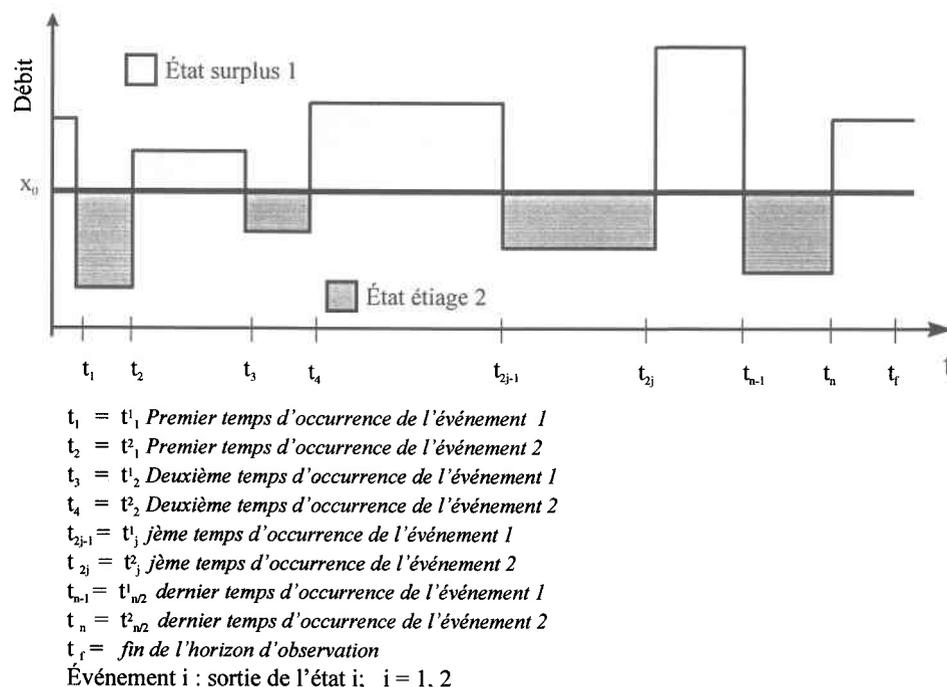


Figure 5.5: Un processus alterné de surplus et d'étiage observé

Le logarithme de la vraisemblance du processus alterné de Poisson non homogène qu'on doit optimiser, étant donnés des formes paramétriques pour $\lambda^k(t)$, $k=1,2$, les temps d'arrivées observés $t_1=t_1^1$, $t_2=t_1^2$, $t_3=t_2^1$, $t_4=t_2^2$, $t_{n-1}=t_{n/2}^1$, $t_n=t_{n/2}^2$, et en supposant que le processus débute et se termine dans l'état 1 s'écrit comme suit (Lewis, 1972; Kavvas, 1975; North, 1981):

$$\ln L(t_s; s = 1, \dots, n) = \sum_{s=1}^n \ln \lambda^k(t_s) - \sum_{s=1}^n \int_{t_{s-1}}^{t_s} \lambda^k(u) du - \int_{t_n}^{t_r} \lambda^{k'}(u) du \quad (5.9)$$

où $k=i$, si l'on a en t_s une rupture de l'état i et $k' \neq i$ si l'on a en t_n une rupture de l'état i , $i=1,2$.

L'optimisation pouvant s'effectuer séparément pour $i=1$ et $i=2$, nous avons alors:

$$\ln L^1 = \sum_{j=2}^{n/2} \ln \lambda^1(t_{2j-1}) - \sum_{j=2}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} \lambda^1(u) du - \int_{t_n}^{t_r} \lambda^1(u) du \quad (5.10)$$

$$\ln L^2 = \sum_{j=1}^{n/2} \ln \lambda^2(t_{2j}) - \sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \lambda^2(u) du \quad (5.11)$$

où $\lambda^k(t) = \exp[\beta_k + K_k \sin(\omega t + \theta_k)]$, $\omega = \frac{2\pi}{365}$, ce qui donne:

$$\begin{aligned} \ln L^1 &= \sum_{j=2}^{n/2} \ln \left(\exp[\beta_1 + K_1 \sin(\omega t_{2j-1} + \theta_1)] \right) \\ &- \sum_{j=2}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} \exp[\beta_1 + K_1 \sin(\omega u + \theta_1)] du - \int_{t_n}^{t_r} \exp[\beta_1 + K_1 \sin(\omega u + \theta_1)] du \end{aligned} \quad (5.12)$$

$$\ln L^2 = \sum_{j=1}^{n/2} \ln \left(\exp[\beta_2 + K_2 \sin(\omega t_{2j} + \theta_2)] \right) - \sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \exp[\beta_2 + K_2 \sin(\omega u + \theta_2)] du \quad (5.13)$$

Il s'agit donc de trouver les valeurs des paramètres de β_1 , K_1 , θ_1 , et de β_2 , K_2 , θ_2 , qui maximisent $\ln L^1$ et $\ln L^2$ respectivement. Notons que nous n'avons pas inclus t_1 dans la formulation de $\ln L^1$, ce qui veut dire que nous n'avons pas tenu compte de l'intervalle entre l'origine et le premier événement de type 1, et ce, parce que l'origine a été spécifiée d'une façon arbitraire en fonction de la disponibilité des données. En dérivant $\ln L^1$ par rapport à β_1 , K_1 , θ_1 , nous obtenons le système d'équations suivant:

$$\begin{aligned} \frac{\partial \ln L^1}{\partial \beta_1} &= \frac{n}{2} - \exp(\beta_1) \sum_{j=2}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} \exp \left(K_1 \sin \left(\frac{2\pi t}{365} + \theta_1 \right) \right) dt - \exp(\beta_1) \int_{t_n}^{t_r} \exp \left(K_1 \sin \left(\frac{2\pi t}{365} + \theta_1 \right) \right) dt \\ \frac{\partial \ln L^1}{\partial K_1} &= \sum_{j=2}^{n/2} \sin \left(\frac{2\pi t_{2j-1}}{365} + \theta_1 \right) - \sum_{j=2}^{n/2} \exp(\beta_1) \int_{t_{2j-2}}^{t_{2j-1}} \sin \left(\frac{2\pi t}{365} + \theta_1 \right) \exp \left(K_1 \sin \left(\frac{2\pi t}{365} + \theta_1 \right) \right) dt \\ &- \exp(\beta_1) \int_{t_n}^{t_r} \sin \left(\frac{2\pi t}{365} + \theta_1 \right) \exp \left(K_1 \sin \left(\frac{2\pi t}{365} + \theta_1 \right) \right) dt \end{aligned} \quad (5.14)$$

$$\frac{\partial \ln L^1}{\partial \theta_1} = K_1 \sum_{j=2}^{n/2} \cos \left(\frac{2\pi t_{2j-1}}{365} + \theta_1 \right) - \frac{365 \exp(\beta_1)}{2\pi} \left[\sum_{j=2}^{n/2} \exp \left(K_1 \sin \left(\frac{2\pi t_{2j-1}}{365} + \theta_1 \right) \right) - \exp \left(K_1 \sin \left(\frac{2\pi t_{2j-2}}{365} + \theta_1 \right) \right) + \exp \left(K_1 \sin \left(\frac{2\pi t_r}{365} + \theta_1 \right) \right) - \exp \left(K_1 \sin \left(\frac{2\pi t_n}{365} + \theta_1 \right) \right) \right]$$

En dérivant $\ln L^2$ par rapport à β_2 , K_2 , θ_2 nous obtenons le système d'équations suivant :

$$\frac{\partial \ln L^2}{\partial \beta_2} = \frac{n}{2} - \exp(\beta_2) \sum_{j=1}^{\frac{n}{2}} \int_{t_{2j-1}}^{t_{2j}} \exp\left(K_2 \sin\left(\frac{2\pi t}{365} + \theta_2\right)\right) dt$$

$$\frac{\partial \ln L^2}{\partial K_2} = \sum_{j=1}^{\frac{n}{2}} \sin\left(\frac{2\pi t_{2j}}{365} + \theta_2\right) - \sum_{j=1}^{\frac{n}{2}} \exp(\beta_2) \int_{t_{2j-1}}^{t_{2j}} \sin\left(\frac{2\pi t}{365} + \theta_2\right) \exp\left(K_2 \sin\left(\frac{2\pi t}{365} + \theta_2\right)\right) dt$$

(5.15)

$$\frac{\partial \ln L^2}{\partial \theta_2} = K_2 \sum_{j=1}^{\frac{n}{2}} \cos\left(\frac{2\pi t_{2j}}{365} + \theta_2\right) - \frac{365 \exp(\beta_2)}{2\pi} \left[\sum_{j=1}^{\frac{n}{2}} \exp\left(K_2 \sin\left(\frac{2\pi t_{2j}}{365} + \theta_2\right)\right) - \exp\left(K_2 \sin\left(\frac{2\pi t_{2j-1}}{365} + \theta_2\right)\right) \right]$$

Les valeurs des paramètres β_1 , K_1 , θ_1 , et β_2 , K_2 , θ_2 , qui maximisent $\ln L^1$ et $\ln L^2$ respectivement sont obtenues en fixant les dérivées partielles à 0 et en résolvant les systèmes d'équations 5.14 et 5.15. Toutefois, la résolution des systèmes d'équations ne peut se faire de manière analytique. On doit alors utiliser une méthode numérique d'optimisation non linéaire. La méthode utilisée dans ce travail est celle de Powell (Press *et al.*, 1992). C'est une méthode d'optimisation non linéaire qui ne nécessite pas de calcul de gradients et qui transforme un problème multidimensionnel en une suite de problèmes unidimensionnels. Les calculs d'optimisation sont effectués à l'aide du langage de programmation C.

5.3.2 Le processus de renouvellement non homogène (RNH)

En estimant les paramètres de $\lambda^k(t)$ à la section 5.3.1, nous estimons en fait un taux de rupture moyen sur les n années observées puisque nous supposons que $\lambda^k(t)$ est une fonction périodique et que $\lambda^k(t+365*p) = \lambda^k(t)$, $p=1..n$.

En appliquant cette méthode d'estimation (MV) aux taux de rupture du RNH nous pouvons donc estimer $E(\varphi^k(t-u, t))$, où:

$$E(\varphi^k(t-u, t)) = E(\varphi^k(x^k)) \lambda^k(t) \quad (5.16)$$

Or $E(\varphi^k(x^k)) = \frac{1}{\mu} = 1$ car $\varphi^k(x^k)$ est le taux de rupture d'un processus de renouvellement homogénéisé dont la loi des durée est une Gamma de moyenne $\mu=1$.

Le taux de rupture moyen $E(\varphi^k(t-u, t))$ est donc égal au paramètre de changement d'échelle, $\lambda^k(t)$. L'estimation de $E(\varphi^k(t-u, t))$ correspond alors à l'estimation $\lambda^k(t)$ telle que décrite à la section 5.3.1 (North, 1981).

5.3.3 Le processus de Poisson non homogène conditionnel (PNHC)

Les intensités du PNHC proposé dans ce travail sont elles-mêmes des processus aléatoires puisqu'elles sont des fonctions de variables aléatoires, à savoir les précipitations et les températures. Toutefois, nous supposons que les précipitations ainsi que les températures sont des observations connues. La méthode d'estimation de la section 5.3.1 peut alors être appliquée à l'estimation des intensités du PNHC. Il s'agit donc de trouver les valeurs de β_1 , K_1 , θ_1 , et K_2 , θ_2 , qui maximisent $\ln L^1$ (5.18) et $\ln L^2$ (5.19) respectivement.

$$\ln L^1 = \sum_{j=1}^{n/2} \ln(\rho^1(t_{2j-1}, H_t)) - \sum_{j=1}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} \rho^1(t(u), H_t) du - \int_{t_n}^{t_r} \rho^1(t(u), H_t) du \quad (5.17)$$

où $\rho^1(t, H_t) = \exp\left[\beta_1 + K_1 \text{nbjoursec}(t) + \theta_1 \sin\left(\frac{2\pi t}{365}\right)\right]$. Nous avons donc:

$$\begin{aligned} \ln L^1 = & \sum_{j=1}^{n/2} \left[\beta_1 + K_1 \frac{\text{nbjoursec}(t_{2j-1})}{365} + \frac{\theta_1 \sin(2\pi t_{2j-1})}{365} \right] \\ & - \sum_{j=1}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} \exp\left[\beta_1 + K_1 \frac{\text{nbjoursec}(u)}{365} + \frac{\theta_1 \sin(2\pi u)}{365}\right] du - \int_{t_n}^{t_r} \exp\left[\beta_1 + K_1 \frac{\text{nbjoursec}(u)}{365} + \frac{\theta_1 \sin(2\pi u)}{365}\right] du \end{aligned} \quad (5.18)$$

$$\ln L^2 = \sum_{j=1}^{n/2} \ln(\rho^2(t_{2j}, H_t)) - \sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \rho^2(t(u), H_t) du$$

où: $\rho^2(t, H_t) = \exp(K_2 \text{pnorm}(t) + \theta_2)$. Nous avons donc:

$$\ln L^2 = \sum_{j=1}^{n/2} K_2 \text{pnorm}(t_{2j}) + \theta_2 - \sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \exp(K_2 \text{pnorm}(u) + \theta_2) du \quad (5.19)$$

Les estimateurs de β_1 , K_1 , θ_1 , et de K_2 , θ_2 sont les solutions des systèmes d'équations (5.20) et (5.21) respectivement. La méthode d'optimisation numérique utilisée est la même qu'à la section 5.3.

$$\begin{aligned} \frac{\partial \ln L^1}{\partial \beta_1} &= \frac{n}{2} - \exp(\beta_1) \sum_{j=2}^{\frac{n}{2}} \int_{t_{2j-2}}^{t_{2j-1}} \exp\left(K_1 n \text{bjour sec}(t) + \theta_1 \sin\left(\frac{2\pi t}{365}\right)\right) dt \\ &\quad - \exp(\beta_1) \int_{t_n}^{t_r} \exp\left(K_1 n \text{bjour sec}(t) + \theta_1 \sin\left(\frac{2\pi t}{365}\right)\right) dt \\ \frac{\partial \ln L^1}{\partial K_1} &= \sum_{j=2}^{\frac{n}{2}} n \text{bjour sec}(t(2 * j - 1)) - \sum_{j=2}^{\frac{n}{2}} \exp(\beta_1) \int_{t_{2j-2}}^{t_{2j-1}} n \text{bjour sec}(t) \exp\left(K_1 n \text{bjour sec}(t) + \theta_1 \sin\left(\frac{2\pi t}{365}\right)\right) dt \\ &\quad - \exp(\beta_1) \int_{t_n}^{t_r} n \text{bjour sec}(t) \exp\left(K_1 n \text{bjour sec}(t) + \theta_1 \sin\left(\frac{2\pi t}{365}\right)\right) dt \end{aligned} \quad (5.20)$$

$$\begin{aligned} \frac{\partial \ln L^1}{\partial \theta_1} &= \sum_{j=2}^{\frac{n}{2}} \sin\left(\frac{2\pi t_{2j-1}}{365}\right) - \exp(\beta_1) \left[\sum_{j=2}^{\frac{n}{2}} \int_{t_{2j-2}}^{t_{2j-1}} \sin\left(\frac{2\pi t}{365}\right) \exp\left(K_1 n \text{bjour sec}(t) + \theta_1 \sin\left(\frac{2\pi t}{365}\right)\right) dt \right] \\ &\quad - \exp(\beta_1) \left[\sum_{j=2}^{\frac{n}{2}} \int_{t_n}^{t_r} \sin\left(\frac{2\pi t}{365}\right) \exp\left(K_1 n \text{bjour sec}(t) + \theta_1 \sin\left(\frac{2\pi t}{365}\right)\right) dt \right] \end{aligned}$$

$$\begin{aligned} \frac{\partial \ln L^2}{\partial K_2} &= \sum_{j=1}^{\frac{n}{2}} p \text{norm}(t_{2j}) - \sum_{j=1}^{\frac{n}{2}} \int_{t_{2j-1}}^{t_{2j}} p \text{norm}(t) \exp(K_2 p \text{norm}(t) + \theta_2) dt \end{aligned} \quad (5.21)$$

$$\frac{\partial \ln L^2}{\partial \theta_2} = \frac{n}{2} - \sum_{j=1}^{\frac{n}{2}} \int_{t_{2j-1}}^{t_{2j}} \exp(K_2 p \text{norm}(t) + \theta_2) dt$$

5.3.4 Les variances des paramètres des fonctions d'intensité des PNH et RNH

Les variances des estimateurs par la méthode du MV, $\hat{\beta}_i$, \hat{K}_i , $\hat{\theta}_i$, et $\hat{\lambda}^i(t)$, $i=1,2$, du PNH et du RNH sont présentés dans l'annexe 3.

5.3.5 Les variances des paramètres des fonctions d'intensité du PNHC

Les variances des estimateurs par la méthode du MV, $\hat{\beta}_i$, \hat{K}_i , $\hat{\rho}^i(t, H_i)$, $i=1,2$, et $\hat{\theta}_1$ du PNHC sont présentés dans l'annexe 6.

5.4 Vérification de l'adéquation des modèles

La vérification de l'adéquation des modèles construits est une étape très importante dans toute démarche de modélisation. Nous vérifierons par des tests graphiques et statistiques si les processus PNH, RNH, et PNHC sont acceptables comme modèles pour PA^1 et PA^2 . Ces tests sont appliqués aux durées homogénéisées et non homogénéisées, et non aux temps d'occurrence. Le modèle alterné sera composé d'une combinaison de deux processus choisis parmi les processus PNH, RNH, et PNHC. Tous les calculs de cette étape sont effectués dans MATLAB.

Les tests statistiques étant applicables surtout sur des processus ponctuels homogènes, nous devons en un premier lieu procéder à l'homogénéisation du processus observé. La vérification de l'adéquation des modèles est effectuée en deux étapes comme cela est présenté au tableau 5.6.

Étapes de la vérification de l'adéquation des modèles

1. *L'homogénéisation du processus observé à l'aide des fonctions d'intensité estimées;*
2. *L'application de tests graphiques et statistiques aux durées homogénéisées et non homogénéisées.*

Tableau 5.6: Étapes de la vérification de l'adéquation des modèles

5.4.1 L'homogénéisation du processus observé

Suite à l'estimation des fonctions d'intensité des étages et des surplus $\rho^k(t, H_t)$ pour chacun des processus PNH et PNHC, le processus alterné observé est homogénéisé en transformant les durées dans l'état $k=1,2$ observées x_j^k , en durées homogénéisées $x_j'^k$ comme suit:

$$x_j'^2 = \int_{x_{j-1}^2}^{x_j^2} \rho^2(u, H_u) du, \quad x_j'^1 = \int_{x_{j-1}^1}^{x_j^1} \rho^1(u, H_u) du \quad (5.22)$$

L'homogénéisation, en pondérant les durées en fonction de l'époque de l'année où on se situe, permet d'enlever la non-stationnarité. Lorsque le processus ponctuel proposé et sa fonction d'intensité sont acceptables, le processus homogénéisé par l'équation (5.22) est alors

un processus de Poisson homogène, d'intensité 1. C'est principalement cette propriété que nous testerons afin de vérifier l'adéquation des modèles PNH et PNHC.

Pour ce qui est du processus RNH, l'homogénéisation par les fonctions $\lambda^k(t)$ implique seulement un changement d'horloge. Les durées homogénéisées x_j^k sont obtenues de la façon suivante:

$$x_j^2 = \int_{x_{j-1}^2}^{x_j^2} \lambda^2(u) du, \quad x_j^1 = \int_{x_{j-1}^1}^{x_j^1} \lambda^1(u) du \quad (5.23)$$

Lorsque l'hypothèse d'un modèle RNH décrite à la section 5.2.2 est acceptable, le processus homogénéisé est alors un processus de renouvellement dont les durées suivent une loi Gamma de moyenne unitaire, de paramètres $\alpha_k = \rho_k = \frac{1}{\hat{\sigma}_k^2}$, $\hat{\sigma}_k^2$ étant la variance estimée des durées homogénéisées x_j^k . C'est cette hypothèse que nous chercherons à vérifier par les tests statistiques.

Il est à noter que les durées x_j^k seront différentes pour chacun des trois processus. Des tests statistiques seront appliqués à ces durées afin de déterminer l'adéquation du processus ponctuel correspondant.

5.4.2 L'application de tests graphiques

Une façon de vérifier l'adéquation des modèles PNH est de comparer la fonction de survie logarithmique (3.4.1.2) des durées empiriques homogénéisées x_j^k à une droite de pente 1 (ce qui correspond à la fonction de survie logarithmique d'un processus de Poisson homogène d'intensité 1). On peut aussi comparer la fonction de survie logarithmique des durées empiriques non homogénéisées x_j^k et $\ln(1 - F(x_j^k))$ où $1 - F(x)$ est définie par l'équation 5.25.

5.4.3 L'application de tests statistiques

Les tests statistiques présentés au chapitre 3 sont tout à fait adéquats pour une série d'événements d'un seul type, c'est-à-dire une série d'étiage ou une série de déficit. Toutefois, ils ne sauraient tous être appliqués à une série alternée de deux types d'événements. Les tests s'appliquant à des intervalles ne sont pas adéquats car l'intervalle entre deux événements d'un

même type k , correspond dans le processus alterné à la somme de deux variables aléatoires: la durée dans l'état i et la durée dans l'état k , $i \neq k$. Les tests basés sur les temps d'arrivée d'un événement d'un type ne peuvent donc pas être utilisés pour les processus alternés. On peut toutefois appliquer des tests basés sur les durées.

Dans ce travail, nous appliquons le test du périodogramme (3.4.2.2) pour tester l'indépendance des durées homogénéisées x_j^k , et les statistiques de Kolmogorov-Smirnov (3.4.2.2.1) et d'Anderson (3.4.2.2.2) afin de tester l'adéquation des loi exponentielle et Gamma à x_j^k , ainsi que l'adéquation de la loi (5.25) aux durées non homogénéisées x_j^k . Ces tests sont appliqués aux durées x_j^1 et x_j^2 séparément. Il est possible de constater par exemple que les durées de surplus d'une rivière donnée sont modélisables par PNH et que les durées d'étiage de la même rivière sont modélisables par un RNH. Dans ce cas on aurait un processus alterné hybride constitué de deux processus ponctuels de nature différente qui alternent.

Le test du périodogramme appliqué aux durées x_j^1 et x_j^2 (PNH, RNH, PNHC)

Ce test (3.4.2.2) permet de tester l'indépendance des durées homogénéisées x_j^1 et x_j^2 . Cette hypothèse doit être vérifiée lorsque le processus alterné appliqué est acceptable. Ce test est utilisé pour les trois modèles PNH, RNH, PNHC.

Hypothèses:

H_0 : les durées homogénéisées x_j^k sont indépendantes et le processus en question n'est pas rejeté automatiquement

H_1 : H_0 est rejetée

Le test d'adéquation de la loi exponentielle aux durées x_j^1 et x_j^2 (PNH, PNHC)

Ce test est utilisé pour les modèles PNH et PNHC. Si les durées homogénéisées suivent une loi exponentielle de moyenne 1, alors ces modèles ne sont pas rejetés.

Hypothèses:

H_0 : x_j^k suivent une loi exponentielle de moyenne 1

H_1 : H_0 est rejetée

Le test d'adéquation de la loi Gamma aux durées (RNH)

Lorsque l'hypothèse d'indépendance des durées homogénéisées x_j^{ik} n'est pas rejetée, un RNH est acceptable. Ce test permet de vérifier si le processus de RNH particulier défini dans notre étude est acceptable en vérifiant si les durées homogénéisées x_j^{ik} de variance $\hat{\sigma}_k^2$ suivent une loi Gamma de moyenne unitaire et de paramètres:

$$\alpha_k = \rho_k = \frac{1}{\hat{\sigma}_k^2} \quad (5.24)$$

Hypothèses:

H_0 : x_j^{ik} suivent une loi Gamma de moyenne unitaire- RNH (section 5.2.2)

H_1 : le processus RNH est rejeté.

Le test d'adéquation de la loi (5.25) aux durées non homogénéisées (PNH)

Ce test permet de vérifier l'adéquation du modèle PNH. Il est appliqué aux durées observées non homogénéisées qui devraient suivre la loi (5.25) (North 1981) lorsque le modèle PNH n'est pas rejeté. On devrait toutefois préciser que North (1981) a obtenu cette expression en faisant l'hypothèse (forte) que le processus des étiages d'un seul type est PNH, ce qui ne s'est pas vérifié dans notre cas.

$$1-F(x) = P[X^k > x] = \frac{\int_0^{t_0} \lambda_*^i(u) \exp\left[-\int_t^{t+x} \lambda^k(u) du\right] dt}{\int_0^{t_0} \lambda_*^i(t) dt} \quad (5.25)$$

Hypothèses:

H_0 : x_j^k suivent la loi 5.25 - PNH

H_1 : PNH est rejeté.

L'application des tests aux trois modèles est résumée au tableau 5.7. La démarche pour vérifier l'adéquation des modèles est résumée à la figure 5.6.

Tests	Modèles
Fonction de survie logarithmique	PNH
Indépendance de x_j^k	PNH, RNH, PNHC
Adéquation de la loi exponentielle à x_j^k	PNH, PNHC
Adéquation de la loi Gamma à x_j^k	RNH
Adéquation de la loi 5.25 à x_j^k	PNH

Tableau 5.7: L'application des tests aux divers modèles

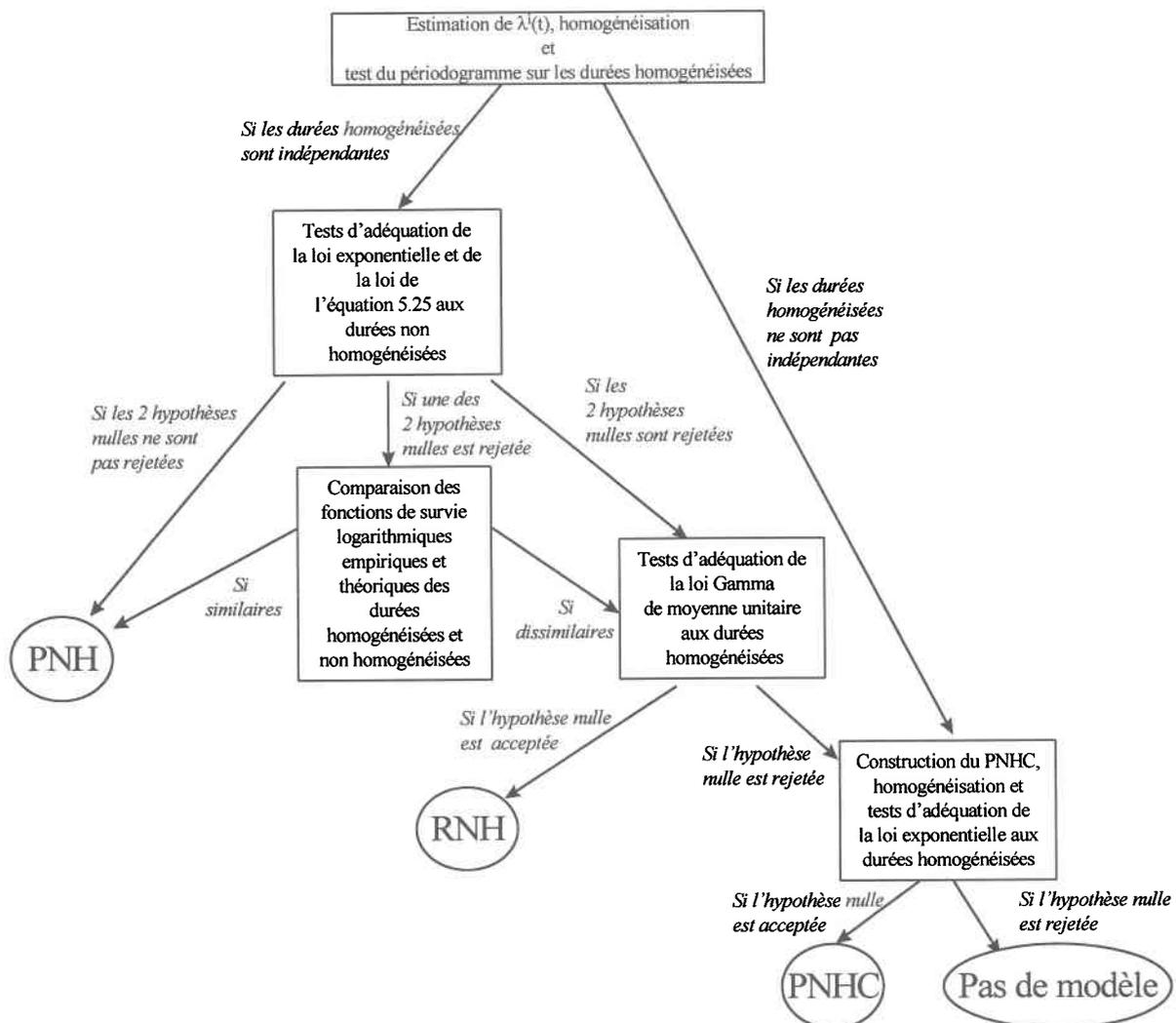


Figure 5.6: La vérification de l'adéquation des modèles

5.4.4 Simulation

Le modèle alterné construit pour les séries d'étiage et de surplus d'une rivière donnée, et composé d'une combinaison de PNH, RNH, ou PNHC, peut être utilisé afin de simuler des séries alternées de surplus et d'étiage. Cette simulation s'avère pratique pour l'analyse de sensibilité ainsi que pour la quantification du risque par des indices tels que décrits au chapitre 7.

En supposant que les paramètres estimés de la fonction d'intensité, $\hat{\beta}_i, \hat{K}_i$, et $\hat{\theta}_i$, $i=1,2$, suivent asymptotiquement une loi normale multivariée conditionnelle aux données, nous pouvons générer des séries de paramètres $\beta_i^n, K_i^n, \theta_i^n$ et des séries d'événements correspondantes.

Pour chacun des états 1 ou 2 tel que le PNH ou le RNH est un modèle acceptable, nous générons une série de $\hat{\beta}_i, \hat{K}_i$, et $\hat{\theta}_i$, $i=1,2$, en tirant dans une loi normale multivariée de moyenne $\hat{\beta}_i, \hat{K}_i, \hat{\theta}_i$ et de matrice variance-covariance \hat{I}^{i-1} telle que définie à l'annexe 3. On fait de même lorsque le PNHC s'applique à l'état 1, mais en utilisant la matrice variance-covariance \hat{I}^{1-1} telle que définie à l'annexe 6. Pour ce qui est de l'application du PNHC à l'état 2, on tire dans une loi normale bivariée de moyennes $\hat{\beta}_2, \hat{K}_2$, et de variances et covariance telle que définies à l'annexe 6. La combinaison de modèles PNH, RNH ou PNHC correspondante est ensuite utilisée pour simuler la série d'événements.

L'algorithme de génération de n variables provenant d'une loi normale multivariée ainsi que les procédures de génération d'un PNH, d'un RNH et d'un PNHC sont décrits à l'annexe 4. Il est important de noter que la simulation du processus alterné doit se faire en simultané, c'est-à-dire qu'on ne peut pas simuler une série d'événements pour chaque état séparément.

5.4.5 Analyse de sensibilité

Un test statistique sur les séries d'événements observées fournit une valeur de test que l'on compare à un seuil afin d'accepter ou de rejeter le modèle proposé pour ces séries. Toutefois, la série observée n'est qu'une réalisation possible du processus sous-jacent. Nous pouvons générer, par simulation, d'autres réalisations possibles à l'aide des modèles construits. Sous l'hypothèse de la validité de notre modèle, notre série initiale aurait donc pu être n'importe laquelle de ces séries générées.

Afin de tenir compte de l'incertitude sur les paramètres et de l'aléa propre au modèle (incertitude hydrologique) nous effectuons une analyse de sensibilité en appliquant les tests statistiques sur 100 séries générées et en calculant avec quelle fréquence le test en question aurait rejeté l'hypothèse nulle. Cette analyse de sensibilité peut nous indiquer d'une façon qualitative l'effet de l'incertitude sur les modèles construits. La série générée est par la suite homogénéisée à l'aide des paramètres $\hat{\beta}_i$, \hat{K}_i , et $\hat{\theta}_i$ ($i=1,2$ PNH, RNH, et $i=1$ PNHC) et $\hat{\beta}_2$, \hat{K}_2 (PNHC) estimés à partir des données observées, et les tests statistiques pour le PNH, le RNH ou le PNHC sont appliqués aux durées homogénéisées. Nous comptons ensuite le nombre de fois que les tests rejettent l'hypothèse du PNH, RNH, ou PNHC. Cette procédure est résumée au tableau 5.8.

Analyse de sensibilité
<i>Pour chaque série d'événements observée</i>
<i>Pour $l=1:100$</i>
<i>Pour chaque état $i: i=1,2$</i>
<i>- générer $\beta_i^l, K_i^l, \theta_i^l$ à partir d'une loi normale multivariée ou β_2, K_2 (PNHC)</i>
<i>- générer la série d'événements correspondante PNH, RNH, ou PNHC</i>
<i>- homogénéiser la série générée par $\hat{\beta}_i, \hat{K}_i$, et $\hat{\theta}_i$ ou $\hat{\beta}_2, \hat{K}_2$ (PNHC)</i>
<i>- si le processus observé est de type PNH ou PNHC effectuer les tests suivants:</i>
<i>le test du périodogramme appliqué aux durées homogénéisées</i>
<i>le test d'adéquation de la loi exponentielle aux durées homogénéisées</i>
<i>- si le processus observé est de type RNH effectuer les tests suivants:</i>
<i>le test du périodogramme appliqué aux durées homogénéisées</i>
<i>le test d'adéquation de la loi Gamma aux durées homogénéisées</i>
<i>- Compter le nombre de fois que chaque test rejette l'hypothèse nulle à 5%</i>

Tableau 5.8: Analyse de sensibilité

6. APPLICATIONS ET RÉSULTATS

Dans ce chapitre nous appliquons et évaluons les modèles PNH, RNH, et PNHC présentés au chapitre 5 en considérant les données de 10 rivières canadiennes et d'une rivière française, et ce à 2 seuils: les 25^{ème} et 10^{ème} quantiles (Q_{25} et Q_{10}). Nous commençons par construire le PNH. Si ce dernier n'est pas adéquat, nous tentons de modéliser la série par le RNH. Nous ne construisons le PNHC que pour les rivières pour lesquelles ni le PNH ni le RNH ne sont acceptables, et ce dans un souci de parcimonie. Nous passons ainsi du modèle le plus simple au plus complexe.

Nous rappelons que la construction et l'évaluation des modèles comprennent (chapitre 5):

- le choix du type de processus et le choix de la fonction d'intensité;
- l'estimation des paramètres de l'intensité, à partir des données qu'on veut modéliser;
- la vérification par des tests statistiques de l'adéquation du modèle aux données; et
- l'analyse de sensibilité.

Nous présentons à la section 6.1 les données qui sont utilisées pour les applications. La section 6.2 contient les résultats de l'estimation des intensités et la section 6.3 les résultats de l'évaluation des PNH et RNH. La section 6.4 est dédiée à l'évaluation du PNHC. La section 6.5 contient l'analyse de sensibilité et les simulations. Finalement, nous concluons le chapitre par la section 6.6.

6.1 Présentation des données

Les données canadiennes étudiées proviennent d'un réseau de stations de mesure identifiées par Anderson *et al.* (1991) comme étant d'excellente qualité (figure 6.1). Ces bassins sont naturels, et plusieurs années de mesure sont disponibles. La rivière française analysée est la Seine à la station de Bar-sur-Seine (figure 6.2).

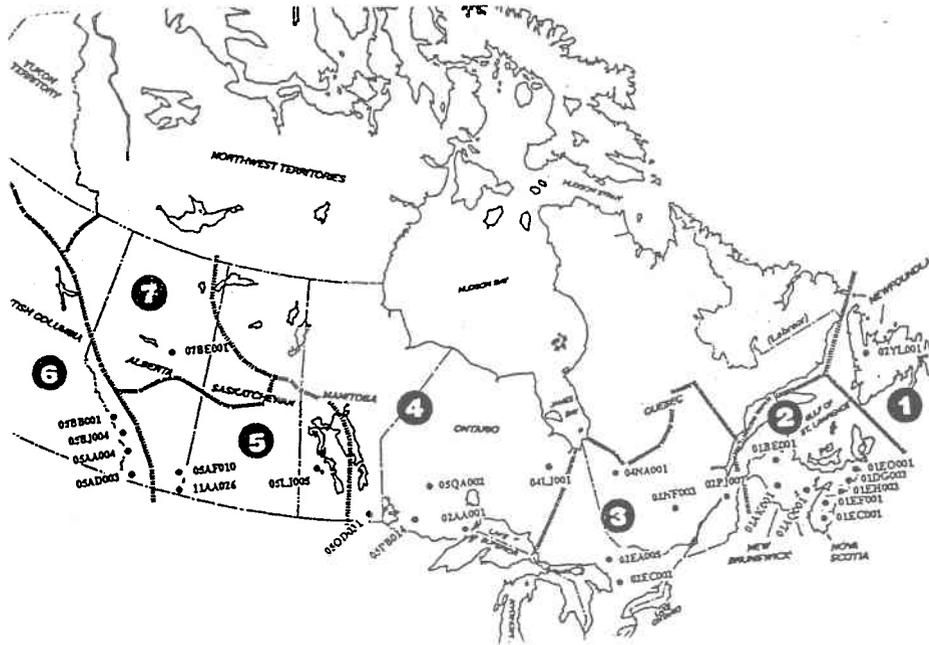


Figure 6.1: Les stations canadiennes (Anderson *et al.*, 1991)

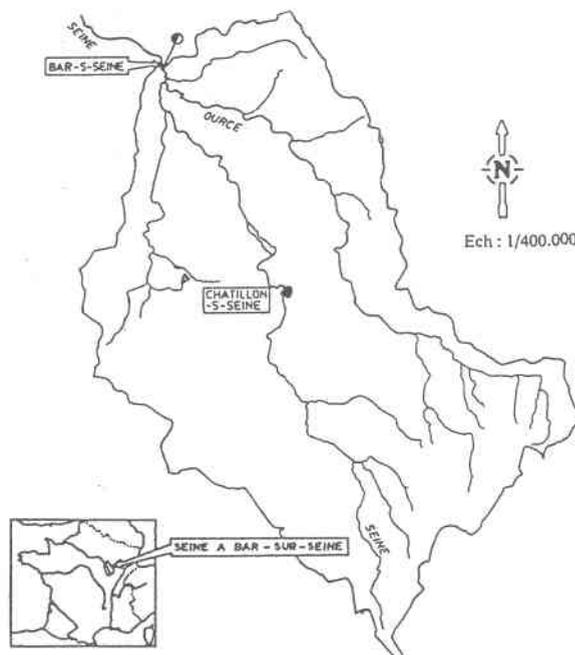


Figure 6.2: Le bassin versant de la Seine à Bar-sur-Seine (Bernier, 1964)

Les données manquantes sont remplacées par les moyennes journalières correspondantes. Le tableau 6.1 présente les stations étudiées, leurs débits moyens, les superficies des bassins, ainsi que le nombre d'années de données journalières disponibles. Les courbes de débit moyen mensuel des rivières du tableau 6.1 sont présentées à l'annexe 7. Les séries d'événements d'étiage et de surplus caractérisées par leurs durées respectives $\{x_j^2\}$ et $\{x_j^1\}$ sont déterminées à partir des données journalières de débit et des seuils Q_{25} et Q_{10} comme cela est décrit à la section 5.1.

Nom de la rivière	Nom de la station	Q moyen (m ³ /s)	Bassin (km ²)	Province/Pays	Nombre d'années
Pigeon	St06 (02AA001)	14.1	1550	Ontario	67
Magnetawan N.	St07 (02AA005)	5.6	321	Ontario	75
Black	St08 (02EC002)	22	1520	Ontario	75
Upper Humber	St14 (02YL001)	79.5	2110	T. Neuve	38
Missinaibi	St15 (04LJ001)	104.2	8940	Ontario	70
Harricana	St16 (04NA001)	59.1	3860	Québec	57
Bow	St18 (05BB001)	39.3	2210	Alberta	80
Sprague	St19 (05OD031)	1.6	455	Manitoba	41
Turtle	St20 (05PB014)	36.7	4870	Ontario	58
Athabasca	St22 (07BE001)	442	74600	Alberta	39
Seine	Stse (Bar-sur-Seine)	78.6	2300	France	28

Tableau 6.1: Description des stations de mesure de débit

Le tableau 6.2 résume le nombre d'événements total (des 2 types) obtenu pour chaque rivière et chaque seuil.

Nom de la station (Anderson <i>et al.</i> ,1991)	Seuil Q_{25}		Seuil Q_{10}	
	Débit m^3/s	Nombre d'événements Q_{25}	Débit m^3/s	Nombre d'événements Q_{10}
St06 (02AA001)	3.4	236	2.38	118
St07 (02AA005)	1.78	234	1.02	152
St08 (02EC002)	4.81	278	2.28	206
St14 (02YL001)	21.5	320	12.1	150
St15 (04LJ001)	20.7	198	13.1	98
St16 (04NA001)	25.5	178	18.1	110
St18 (05BB001)	8.9	300	7.5	212
St19 (05OD031)	.05	188	.03	130
St20 (05PB014)	15.7	92	10.4	26
St22 (07BE001)	113	108	80.3	52
Stse (Bar-sur-Seine)	21	120	10	90

Tableau 6.2: Description des événements aux stations correspondant aux seuils Q_{25} et Q_{10}

6.2 Estimation de $\lambda^i(t)$ (PNH, RNH)

Nous commençons par estimer les paramètres des fonctions $\lambda^i(t)$, $i=1,2$ correspondant aux intensités du PNH de la section 5.3.1 et aux fonctions de changement d'échelle du RNH de la section 5.3.2. Nous rappelons que l'événement de type i correspond à la rupture de l'état i (sortie de l'état i).

Les paramètres $\hat{\beta}_i$, \hat{K}_i , et $\hat{\theta}_i$ de la fonction périodique $\lambda^i(t) = \exp(\beta_i + K_i \sin(\omega t + \theta_i))$, $\omega = \frac{2\pi}{365}$ sont estimés par la méthode du maximum de vraisemblance pour chacune des rivières. Les résultats pour le seuil Q_{25} sont présentés aux tableaux 6.3 (étiages $i=2$) et 6.4 (surplus, $i=1$) et les résultats pour Q_{10} aux tableaux 6.5 (étiages $i=2$) et 6.6 (surplus, $i=1$). $St^{**}(25)$ correspond aux données de la station St^{**} obtenues en considérant le seuil Q_{25} , et

St**(10) à celles de la station St** obtenues en considérant le seuil Q_{10} . Les graphiques des fonctions empiriques $\lambda^i(t)$ et estimées par le MV $\hat{\lambda}_i(t)$ sont présentés à l'annexe 8.

Station Q_{25}	$\hat{\beta}_2$	$\text{var}(\hat{\beta}_2)$	\hat{K}_2	$\text{var}(\hat{K}_2)$	$\hat{\theta}_2$	$\text{var}(\hat{\theta}_2)$
St06(25)	-3.5	.008	.96	.025	6.36	.017
St07(25)	-3.9	.01	-.53	.028	4.97	.06
St08(25)	-3.9	.03	-.9	.028	5.1	.04
St14(25)	-2.9	.006	.42	.016	3.36	.11
St15(25)	-3.9	.01	2.5	.09	4.93	.002
St16(25)	-3.9	.01	2.21	.09	4.86	.003
St18(25)	-1.02	.03	4.33	.084	4.05	.0006
St19(25)	-3.5	.01	1.21	.03	6.35	.015
St20(25)	-4.7	.06	-3.69	.26	3.56	.004
St22(25)	-.09	.1	6.54	.42	5.91	.0006
Stse(25)	-3.5	.03	-.71	.04	4.07	.08

Tableau 6.3: Les paramètres estimés de $\hat{\lambda}^2(t)$, seuil Q_{25} (Étiage)

Station Q_{25}	$\hat{\beta}_1$	$\text{var}(\hat{\beta}_1)$	\hat{K}_1	$\text{var}(\hat{K}_1)$	$\hat{\theta}_1$	$\text{var}(\hat{\theta}_1)$
St06(25)	-5.29	.011	-.91	.02	1.06	.02
St07(25)	-5.29	.014	-1.64	.03	.91	.007
St08(25)	-6.21	.04	-3.23	.07	.72	.002
St14(25)	-4.2	.006	-.14	.01	-.46	.75
St15(25)	-5.29	.011	.96	.03	1.08	.03
St16(25)	-5.29	.014	.95	.03	1.02	.03
St18(25)	-6.2	.05	2.9	.06	1.1	.003
St19(25)	-4.82	.012	-.52	.02	.82	.1
St20(25)	-6.21	.04	-.93	.07	.99	.07
St22(25)	-6.9	.17	-3.27	.22	.16	.007
Stse(25)	-5.11	.03	-1.55	.06	.74	.02

Tableau 6.4: Les paramètres estimés de $\hat{\lambda}^1(t)$, seuil Q_{25} (Surplus)

Station Q_{10}	$\hat{\beta}_2$	$\text{var}(\hat{\beta}_2)$	\hat{K}_2	$\text{var}(\hat{K}_2)$	$\hat{\theta}_2$	$\text{var}(\hat{\theta}_2)$
St06(10)	-3.21	.02	.89	.06	6.32	.04
St07(10)	-3.5	.04	-1.17	.08	5.17	.03
St08(10)	-.82	.10	3.22	.06	3.9	.14
St14(10)	-2.81	.014	.29	.06	4.85	.48
St15(10)	-3.5	.02	1.9	.18	4.9	.006
St16(10)	-3.21	.02	1.84	.15	4.67	.01
St18(10)	-.82	.07	3.22	.15	3.91	.002
St19(10)	-2.99	.019	.65	.02	4.04	.16
St20(10)	-4.82	.43	-4.7	1.09	3.49	.018
St22(10)	0.63	.31	6.5	.97	5.84	.001
Stse(10)	-2.52	.09	1.1	.08	5.48	.10

Tableau 6.5: Les paramètres estimés de $\hat{\lambda}^2(t)$, seuil Q_{10} (Étiage)

Station Q_{10}	$\hat{\beta}_1$	$\text{var}(\hat{\beta}_1)$	\hat{K}_1	$\text{var}(\hat{K}_1)$	$\hat{\theta}_1$	$\text{var}(\hat{\theta}_1)$
St06(10)	-6.21	.02	-.89	.04	1.22	.05
St07(10)	-7.26	.10	-3.04	.15	.81	.004
St08(10)	-6.9	.16	2.89	.21	.97	.002
St14(10)	-5.1	.01	-2.20	.02	2.43	.82
St15(10)	-6.2	.02	.55	.04	.84	.14
St16(10)	-5.8	.02	.87	.04	.74	.04
St18(10)	-6.90	.07	2.9	.1	.98	.004
St19(10)	-5.52	.02	-.79	.04	-.24	.05
St20(10)	-8.11	.35	-1.9	.54	.63	.06
St22(10)	-7.82	.37	-3.1	.5	.12	.01
Stse(10)	-7.26	.30	-3.6	.40	.63	.006

Tableau 6.6: Les paramètres estimés de $\hat{\lambda}^1(t)$, seuil Q_{10} (Surplus)

6.3 Vérification de l'adéquation des modèles (PNH, RNH)

L'étape qui suit l'estimation des paramètres de la fonction périodique $\lambda^i(t)$, $i=1,2$ est l'homogénéisation des durées observées telle que décrite à la section 5.4.1. Suite à l'homogénéisation, des tests d'adéquation sont effectués sur les durées homogénéisées. Afin de pouvoir considérer le PNH ou le RNH, l'hypothèse d'indépendance des durées homogénéisées ne doit pas être rejetée. Pour vérifier cette hypothèse, nous appliquons le test du périodogramme (5.4.3.1).

Dans le cas du PNH, on vérifie si les durées homogénéisées indépendantes suivent une loi exponentielle de paramètre 1. De plus, un test d'adéquation de l'équation 5.25 est appliqué aux durées non homogénéisées. Ces deux tests permettent de déterminer le rejet ou l'acceptation de l'hypothèse du PNH.

Lorsque l'hypothèse d'un PNH est rejetée sans pour autant rejeter l'hypothèse de l'indépendance des durées homogénéisées, un test d'adéquation aux durées homogénéisées de la loi Gamma de moyenne unitaire et de paramètres $\alpha_i = \rho_i = \frac{1}{\hat{\sigma}_i^2}$, $\hat{\sigma}_i^2$ étant la variance estimée des durées homogénéisées x_j^i , permet de rejeter ou non le RNH. Si le RNH est aussi rejeté, nous construisons alors et vérifions le PNHC. Tous les tests sont effectués à un niveau de confiance $\alpha=.05$.

Les étapes décrites précédemment s'appliquent aux événements de type 1 et 2. Nous présentons aux colonnes 2 et 3 des tableaux 6.7 (étiage, $i=2$) et 6.8 (surplus, $i=1$) les résultats du test du périodogramme qui permet de vérifier l'hypothèse de l'indépendance des durées homogénéisées. Les colonnes 4 et 5 contiennent les résultats des tests de Kolmogorov-Smirnov (K-S) et d'Anderson afin de vérifier si les durées homogénéisées suivent une loi exponentielle de paramètre 1 (PNH). La colonne 6 présente les résultats du test d'adéquation de la loi des durées non homogénéisées d'un PNH (équation 5.25) aux durées empiriques observées non homogénéisées. L'hypothèse nulle H_0 est rejetée au niveau de confiance $\alpha=.05$ si le coefficient K-S calculé est supérieur à 1.358, et si le coefficient d'Anderson w^2 est supérieur à 2.492. Les cellules en gris foncé indiquent le rejet de l'hypothèse H_0 au niveau $\alpha=.05$.

6.3.1 Les étiages: résultats du PNH

Test du périodogramme			Test d'adéquation de la loi exponentielle		Test d'adéquation de la loi 5.25	
Nom	K-S	w ²	K-S	w ²	K-S	w ²
St06(25)	.58	.48	.98	.85	.52	.49
St07(25)	.96	1.16	.82	1.00	.92	1.06
St08(25)	.99	1.09	1.09	1.79	1.46	3.3
St14(25)	1.48	1.82	.85	1.08	.54	1.19
St15(25)	.75	.67	1.69	3.86	1.93	8.77
St16(25)	.73	.81	1.52	3.49	1.79	6.87
St18(25)	.85	.99	1.35	2.33	1.57	3.35
St19(25)	.54	.17	1.13	1.18	1.46	1.41
St20(25)	1.88	6.93	.82	.46	1.05	2.16
St22(25)	.63	.27	1.4	3.5	1.16	3.26
Stse(25)	1.14	1.84	1.89	5.68	2.10	6.76
St06(10)	.84	1.69	.93	.88	.91	1.27
St07(10)	.67	.65	.60	.32	.43	.38
St08(10)	1.35	2.07	.94	1.57	1.55	3.23
St14(10)	1.48	3.1	.79	.71	.62	.87
St15(10)	.89	1.83	.74	.73	1.49	3.77
St16(10)	.75	.97	.47	.30	1.19	1.62
St18(10)	.68	.73	1.46	2.53	1.37	2.5
St19(10)	.84	.60	1.11	.73	.70	.51
St20(10)	.99	1.33	.45	.24	.47	.30
St22(10)	.67	.37	.73	.48	.69	.61
Stse(10)	2.39	6.3	2.65	11.55	2.28	9.09

Tableau 6.7: Tests du PNH - i=2 (Étiage)

D'après le tableau 6.7, l'hypothèse d'indépendance des durées homogénéisées est rejetée pour les stations 14(25), 20(25), 14(10), et se(10) par le test de K-S et /ou le test d'Anderson

(w^2). L'hypothèse de l'adéquation de la loi exponentielle de moyenne unitaire aux durées homogénéisées est rejetée par les tests de K-S et/ou d'Anderson pour les stations 15(25), 16(25), 22(25), se(25), 18(10), et se(10). L'hypothèse de l'adéquation de l'équation 5.25 aux durées empiriques non homogénéisées est rejetée par les tests de K-S et/ou d'Anderson pour les stations 08(25), 15(25), 16(25), 18(25), 19(25), et se(25), 08(10), 15(10), 18(10), et se(10). D'après les résultats des tests, on peut conclure que l'hypothèse d'un PNH est tout à fait acceptable pour modéliser les étiages aux stations 06(25), 07(25), 06(10), 07(10), 16(10), 19(10), 20(10), et 22(10).

Le PNH et le RNH sont rejetés pour les stations 20(25), 14(10), et se(10). Cette décision est motivée par le fait que l'hypothèse d'indépendance des durées homogénéisées est rejetée. L'hypothèse d'un PNH est rejetée pour les stations 15(25), 16(25), se(25), et 18(10).

Puisque les tests d'adéquation de loi exponentielle et de l'équation 5.25 présentent des résultats contradictoires quant à l'adéquation d'un PNH aux étiages des stations 08(25), 18(25), 19(25), 08(10), et 15(10), il est difficile de conclure à l'acceptation ou au rejet du modèle. Afin d'avoir plus d'information, nous comparons sur les figures 6.3 à 6.8 les fonctions de survie logarithmiques (3.4.1.2) des durées empiriques non homogénéisées et des durées théoriques non conditionnelles d'un PNH (équation 5.25). De plus, sur les figures 6.9 à 6.13, nous présentons les fonctions de survie logarithmiques des durées empiriques homogénéisées par rapport à la droite représentant la fonction de survie logarithmique d'un PH. Ces graphiques sont aussi présentés sur les figures 6.14 à 6.15 pour la station 14(25) pour laquelle l'hypothèse d'indépendance des durées homogénéisées est rejetée par le test de K-S.

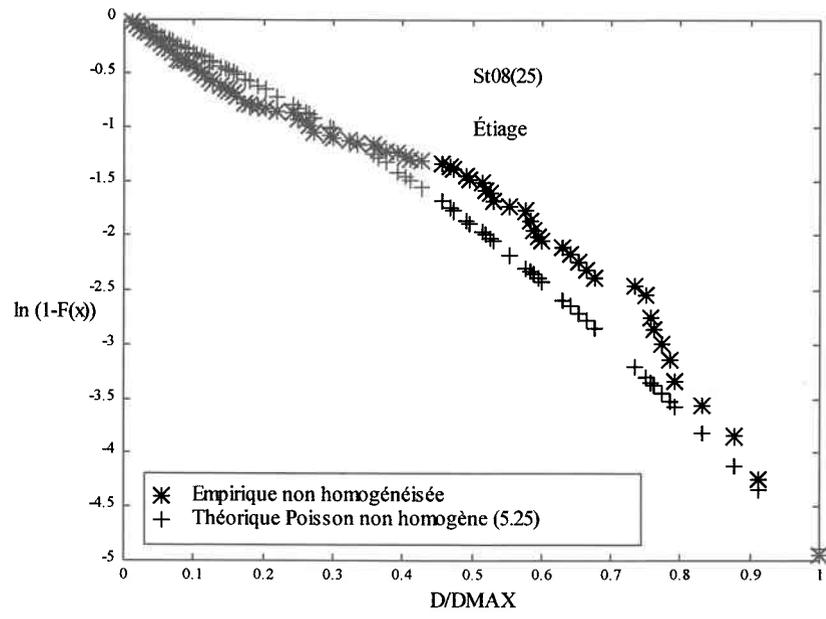


Figure 6.3: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St08(25)

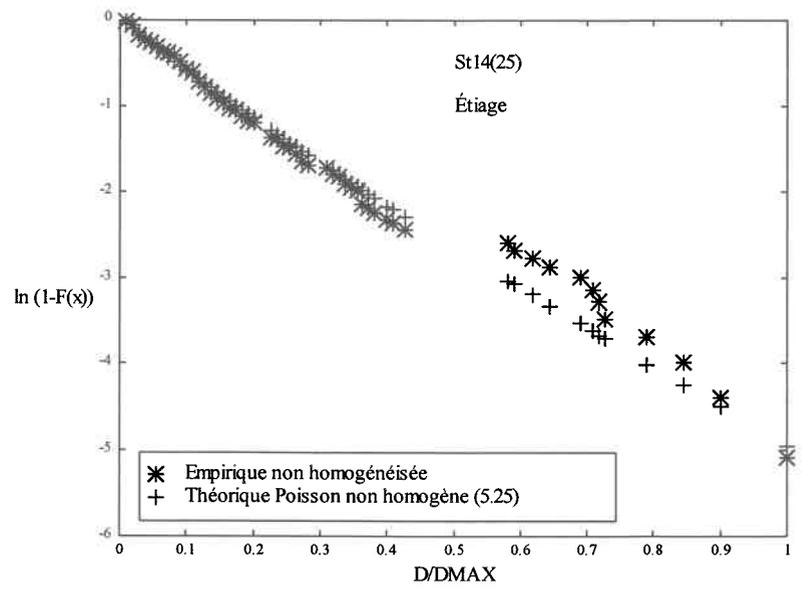


Figure 6.4: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St14(25)

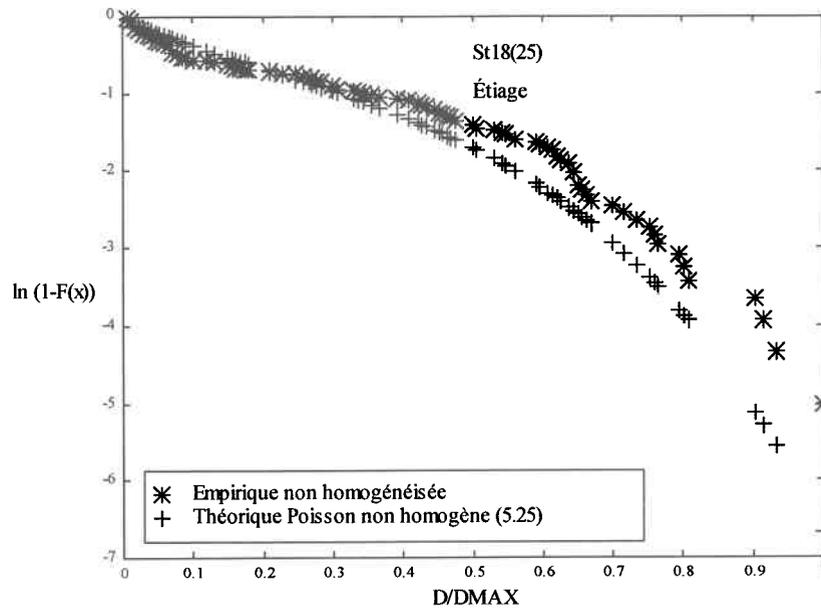


Figure 6.5: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St18(25)

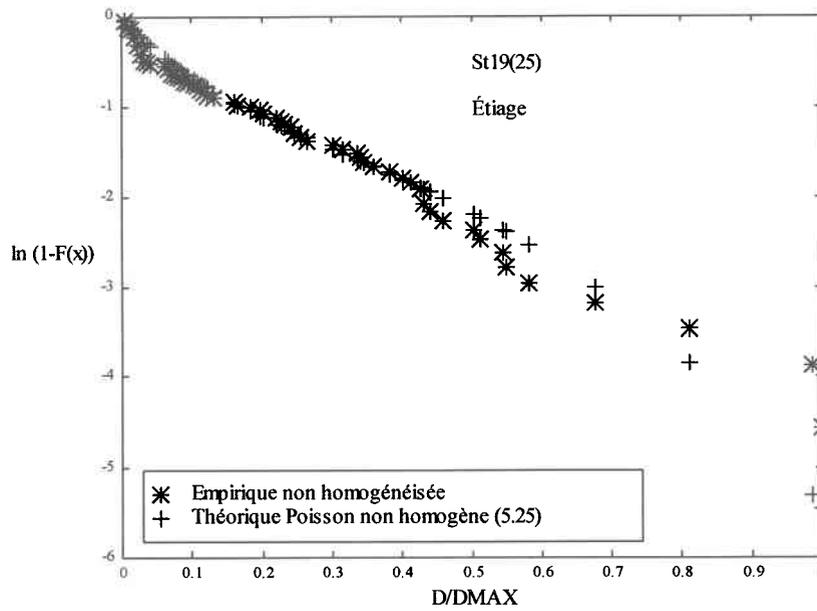


Figure 6.6: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St19(25)

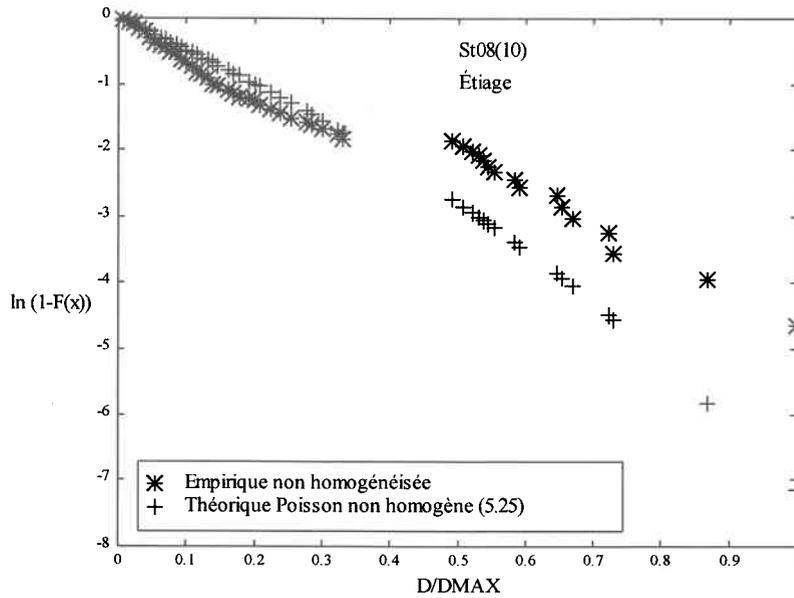


Figure 6.7: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St08(10)

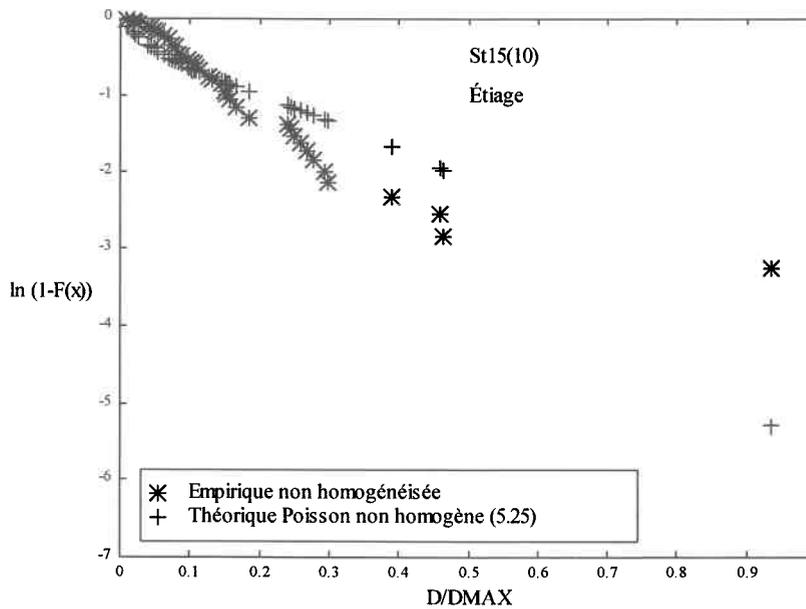


Figure 6.8: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St15(10)

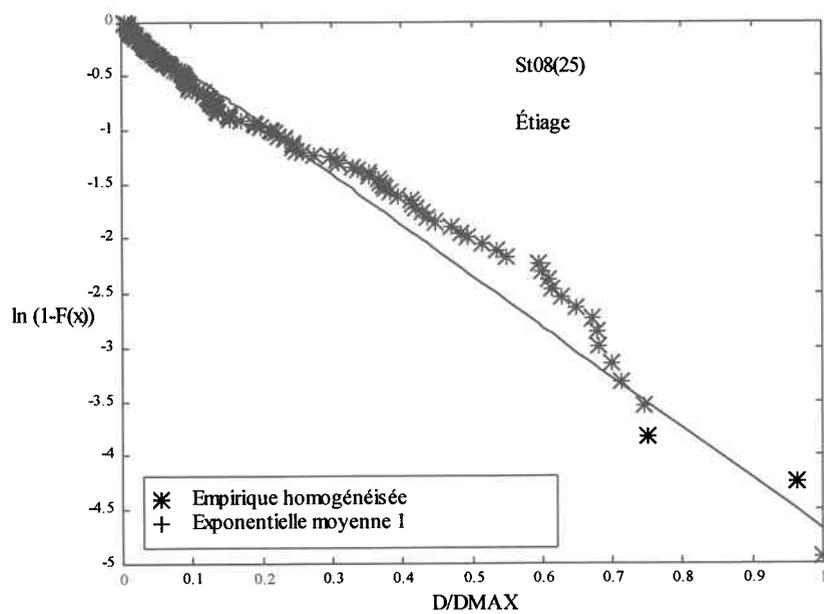


Figure 6.9: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St08(25)

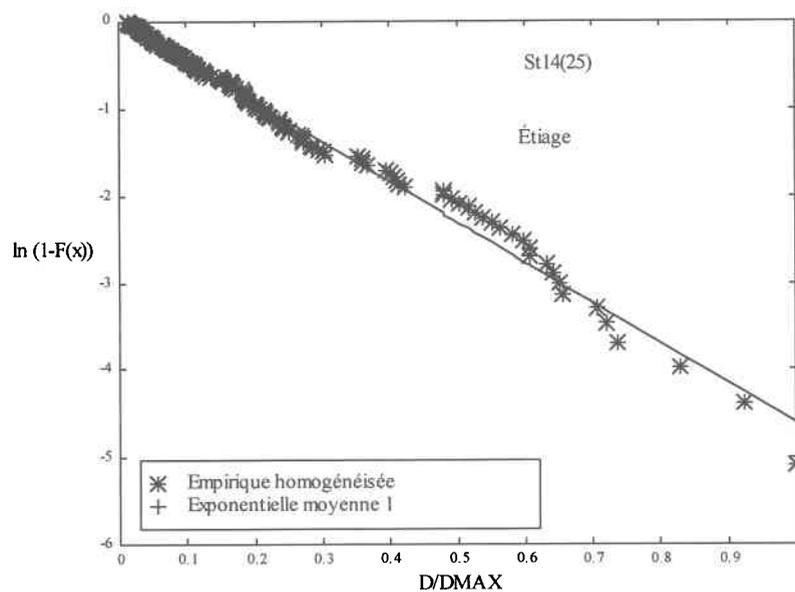


Figure 6.10: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St14(25)

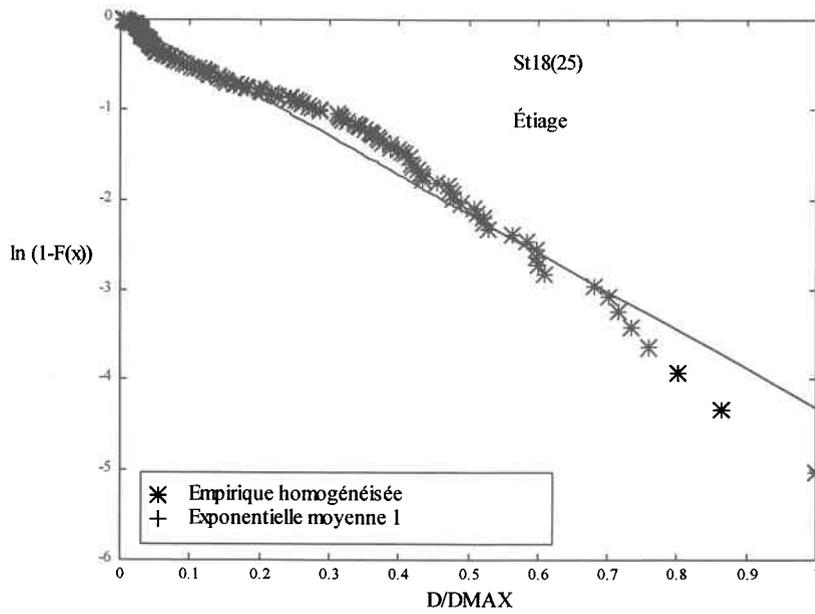


Figure 6.11: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St18(25)

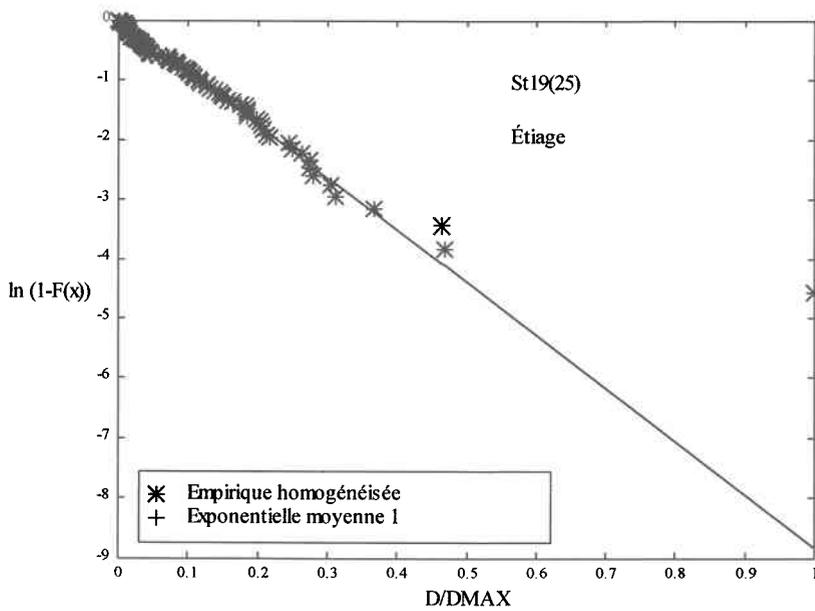


Figure 6.12: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St19(25)

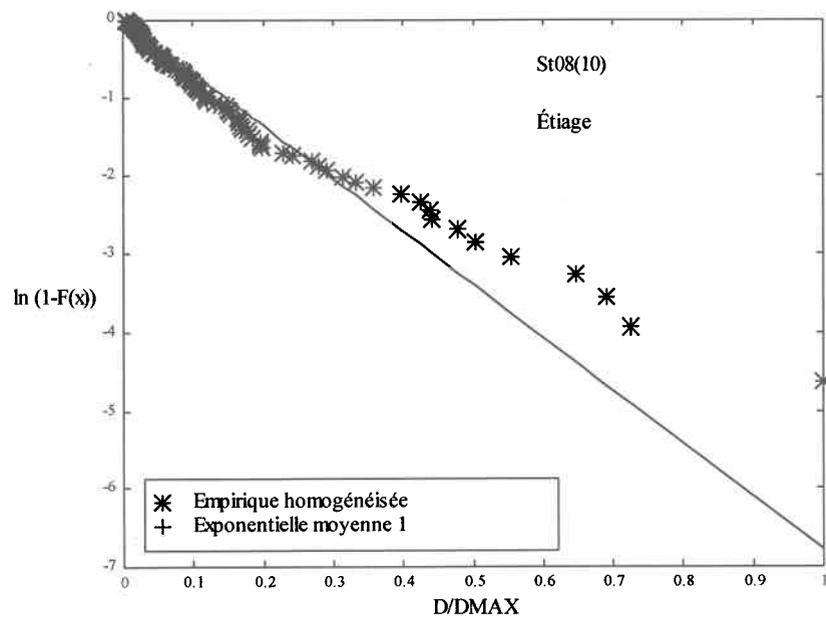


Figure 6.13: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St08(10)

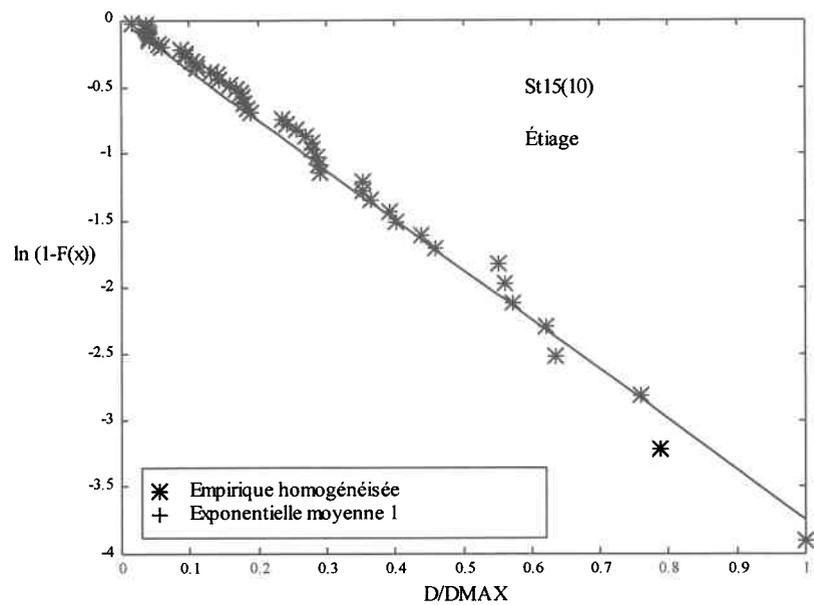


Figure 6.14: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St15(10)

On voit sur les figures 6.3, 6.5, 6.7 et 6.8 que les fonctions de survie logarithmiques des durées empiriques non homogénéisées des étiages des stations 08(25), 18(25), 08(10), et 15(10) ne correspondent pas à $\ln(1-F(x))$ où $1-F(x)$ est donnée par l'équation 5.25. De plus, les fonctions de survie logarithmiques des durées empiriques homogénéisées des étiages de ces stations (figures 6.9, 6.11, 6.13, et 6.14) ne correspondent pas à la fonction de survie logarithmique d'un processus de Poisson homogène d'intensité 1 (droite de pente 1). On conclut donc que le PNH n'est un modèle adéquat pour représenter les étiages des stations 08(25), 18(25), 08(10), et 15(10).

En examinant les figures 6.4, 6.6, 6.10, et 6.12, nous concluons que le PNH est adéquat pour modéliser les étiages des stations 14(25) et 19(25).

6.3.2 Les surplus: résultats du PNH

Test du périodogramme			Test d'adéquation de la loi exponentielle		Test d'adéquation de la loi 5.25	
Nom	K-S	w ²	K-S	w ²	K-S	w ²
St06(25)	.89	1.7	1.66	6.96	2.19	9.65
St07(25)	.72	.82	.95	1.35	1.04	1.63
St08(25)	1.35	2.04	1.22	1.95	1.31	1.78
St14(25)	.82	.72	1.27	2.39	1.41	2.76
St15(25)	1.13	1.21	2.01	5.37	2.0	5.15
St16(25)	1.13	2.52	1.70	3.80	1.68	3.76
St18(25)	1.60	4.92	1.39	2.16	.79	1.09
St19(25)	1.09	1.92	2.89	19.93	3.09	21.43
St20(25)	.97	1.04	1.27	1.41	1.28	1.67
St22(25)	.71	.70	.58	.57	1.13	.88
Stse(25)	1.17	2.59	2.16	9.33	1.85	7.54
St06(10)	1.16	.97	1.89	6.57	1.87	7.27
St07(10)	.63	.42	.61	.48	.76	.49
St08(10)	.33	.12	2.11	9.88	2.12	8.47
St14(10)	1.48	3.80	.82	1.29	.82	1.28
St15(10)	.42	.46	.52	.29	.69	.36
St16(10)	.47	.42	.70	1.14	.89	.90
St18(10)	1.05	1.27	1.55	3.15	1.59	3.25
St19(10)	.46	.29	2.76	21.09	2.74	21.43
St20(10)	.71	.33	1.55	4.23	1.55	4.76
St22(10)	.85	1.1	1.18	1.89	1.47	2.92
Stse(10)	.45	.19	1.96	9.33	2.05	8.57

Tableau 6.8: Tests du PNH - i=1 (Surplus)

D'après le tableau 6.8, l'hypothèse d'indépendance des durées homogénéisées est rejetée pour les stations 16(25), 18(25), se(25), et 14(10) par le test de K-S et /ou le test d'Anderson.

L'hypothèse de l'adéquation de la loi exponentielle de moyenne unitaire aux durées homogénéisées est rejetée par les tests de K-S et/ou d'Anderson pour les stations 06(25), 15(25), 16(25), 19(25), se(25), 06(10), 08(10), 18(10), 19(10), 20(10), et se(10). L'hypothèse d'adéquation de l'équation 5.25 aux durées empiriques non homogénéisées est rejetée par les tests de K-S et/ou d'Anderson pour les stations 06(25), 14(25), 15(25), 16(25), 19(25), se(25), 06(10), 08(10), 18(10), 19(10), 20(10), 22(10), et se(10).

D'après les résultats des tests, on peut conclure que l'hypothèse d'un PNH est tout à fait acceptable pour modéliser les surplus aux stations 07(25), 08(25), 20(25), 22(25), 07(10), 15(10), et 16(10).

Le PNH et le RNH sont rejetés pour les stations 16(25), 18(25), se(25), et 14(10). Cette décision est motivée par le fait que l'hypothèse d'indépendance des durées homogénéisées est rejetée. L'hypothèse d'un PNH est rejetée pour les stations 06(25), 15(25), 19(25), 06(10), 08(10), 18(10), 19(10), 20(10), et se(10).

Puisque les tests d'adéquation de loi exponentielle et de l'équation 5.25 présentent des résultats contradictoires quant à l'adéquation d'un PNH aux surplus des stations 14(25), et 22(10), il est difficile de conclure à l'adéquation ou au rejet du modèle. Afin d'avoir plus d'information, nous comparons sur les figures 6.15 et 6.16 les fonctions de survie logarithmiques des durées empiriques non homogénéisées et des durées théoriques non conditionnelles d'un PNH (équation 5.25). De plus, sur les figures 6.17 et 6.18, nous présentons les fonctions de survie logarithmiques des durées empiriques homogénéisées par rapport à la droite représentant la fonction de survie logarithmique de variables exponentielles. Ces figures nous indiquent que le PNH est inadéquat pour modéliser le surplus de la station 14(25), et qu'il est acceptable pour les surplus de 22(10).

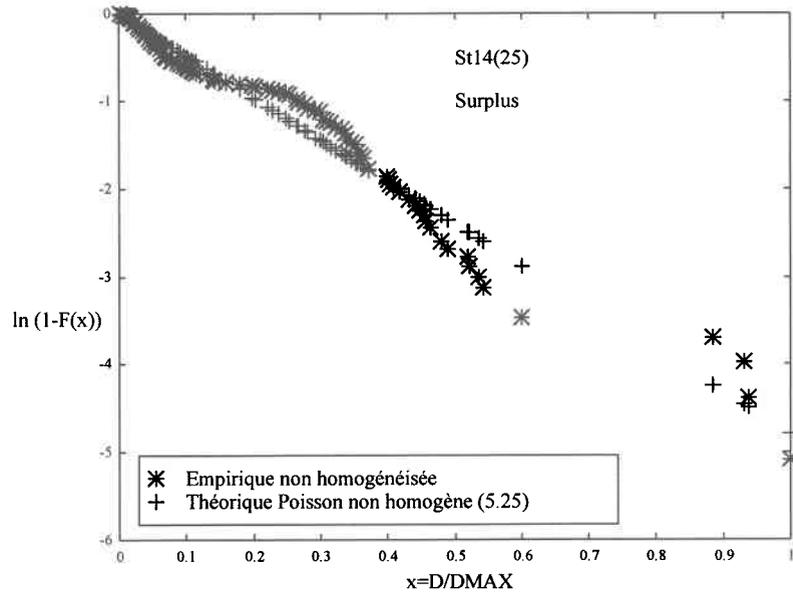


Figure 6.15: Les fonctions de survie logarithmiques des durées de surplus empiriques non homogénéisées et des durées théoriques d'un PNH, St14(25)

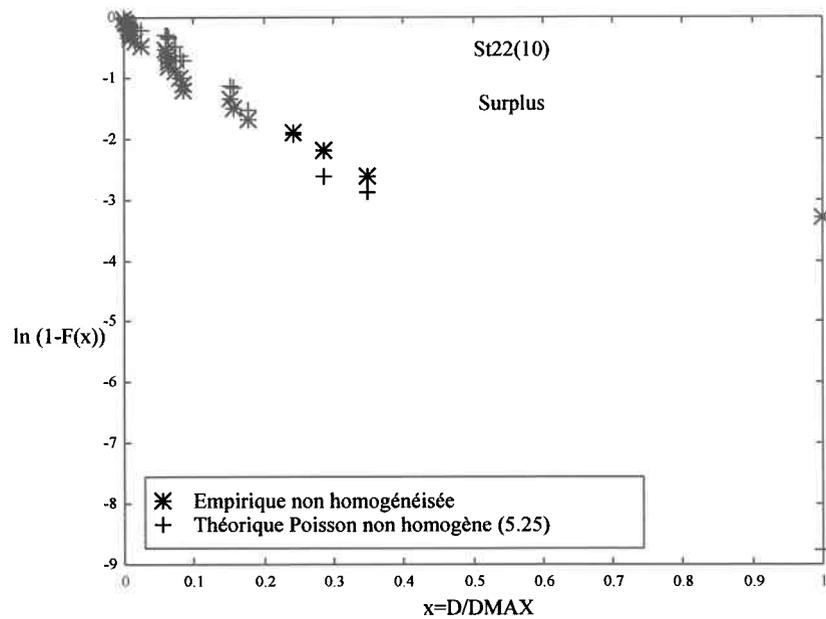


Figure 6.16: Les fonctions de survie logarithmiques des durées de surplus empiriques non homogénéisées et des durées théoriques d'un PNH, St22(10)

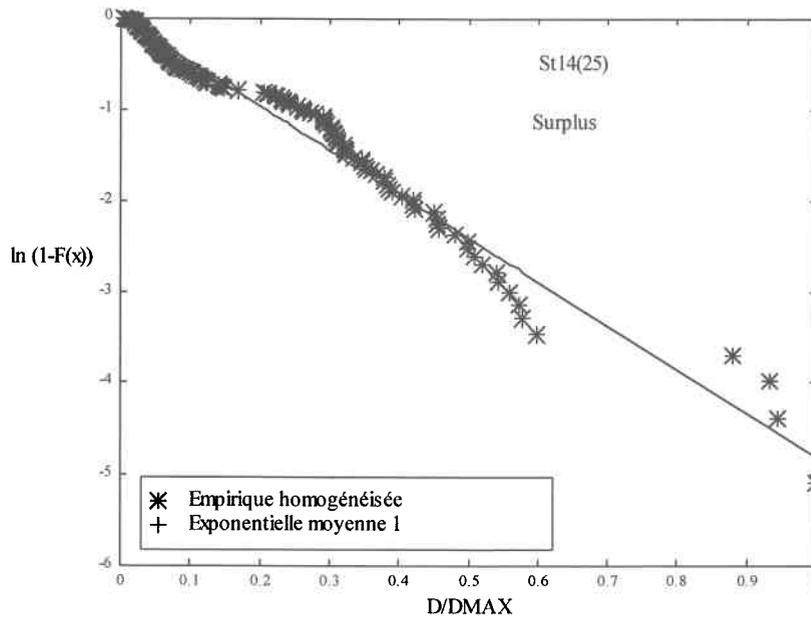


Figure 6.17: Les fonctions de survie logarithmiques des durées de surplus empiriques homogénéisées et des durées théoriques d'un PH, St14(25)

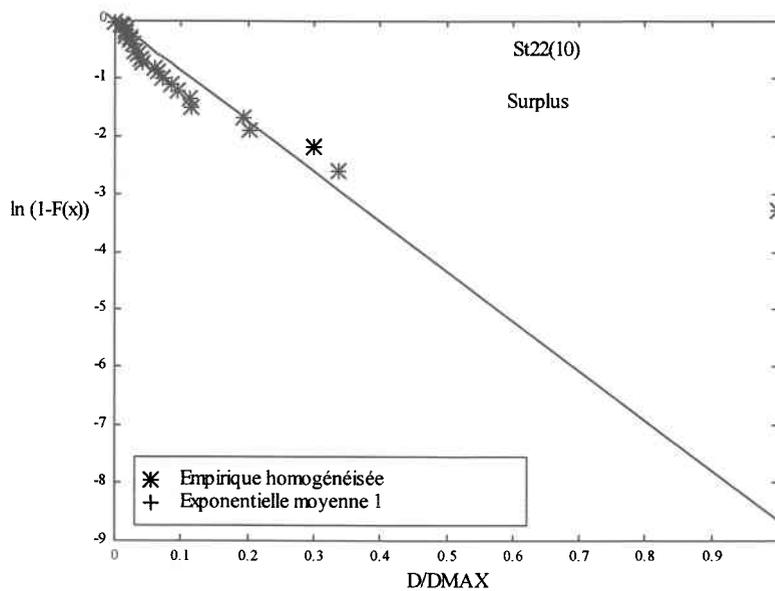


Figure 6.18: Les fonctions de survie logarithmiques des durées empiriques de surplus homogénéisées et des durées théoriques d'un PH, St22(10)

6.3.3 Vérification de l'adéquation du RNH

Lorsque l'hypothèse du PNH est rejetée sans pour autant rejeter l'hypothèse d'indépendance des durées homogénéisées, un test d'adéquation aux durées homogénéisées de la loi Gamma, de moyenne $\frac{\alpha_i}{\rho_i} = 1$ et de variance $\frac{\alpha_i}{\rho_i^2}$ est effectué. Ceci permet de vérifier l'adéquation du RNH. Nous rappelons que α_i est l'inverse de $\hat{\sigma}_i^2$, la variance des durées homogénéisées dans l'état $i = 1, 2$.

6.3.4 Les étiages: résultats du RNH

Les résultats de l'adéquation d'un RNH aux étiages des stations 08(25), 15(25), 16(25), 18(25), 22(25), se(25), 08(10), 15(10), et 18(10) sont présentés au tableau 6.9. On en conclut que le RNH pour lequel les durées d'étiage suivent une loi Gamma de moyenne unitaire et de paramètres $\alpha_2 = \rho_2 = \frac{1}{\hat{\sigma}_2^2}$, $\hat{\sigma}_2^2$ étant la variance estimée des durées homogénéisées x_j^2 , est adéquat pour les étiages des stations 08(25), 15(25), 16(25), se(25), 08(10), 15(10), et 18(10) et qu'il est rejeté pour les étiages de 18(25) et 22(25).

Test de la loi Gamma moyenne unitaire - Étiage			
Station	$\alpha_2 = \rho_2$	K-S	w^2
St08(25)	.9106	.89	1.0
St15(25)	1.86	.48	.36
St16(25)	1.36	1.11	2.03
St18(25)	-	1.7	3.7
St22(25)	-	1.47	4.68
Stse(25)	.50	.84	1.07
St08(10)	.69	1.0	1.43
St15(10)	1.41	.40	.22
St18(10)	.75	1.03	1.23

Tableau 6.9: Test d'adéquation de la loi Gamma aux durées homogénéisées - $i=2$ (Étiage)

En résumé, d'autres modèles doivent être développés pour les étiages des stations 20(25), 14(10), 18(25), 22(25), et se(10) puisque ni le RNH ni le PNH ne sont adéquats pour celles-ci.

6.3.5 Les surplus: résultats du RNH

Les résultats de l'adéquation d'un RNH aux surplus des stations 06(25), 14(25), 15(25), 19(25), 06(10), 08(10), 18(10), 19(10), 20(10), et se(10) sont présentés au tableau 6.10. On en conclut que le RNH pour lequel les durées de surplus suivent une loi Gamma de moyenne unitaire et de paramètres $\alpha_1 = \rho_1 = \frac{1}{\hat{\sigma}_1^2}$, $\hat{\sigma}_1^2$ étant la variance estimée des durées homogénéisées x_j^1 , est adéquat pour les stations 06(25), 15(25), 06(10), 19(10), et se(10).

Test de la loi Gamma moyenne unitaire - Surplus			
Station	$\alpha_1 = \rho_1$	K-S	w^2
St06(25)	.56	.95	1.06
St14(25)		1.50	2.89
St15(25)	2.28	.48	.36
St19(25)		1.7	4.61
St06(10)	.59	.95	1.02
St08(10)		2.94	10.8
St18(10)		2.3	7.54
St19(10)	.37	1.3	2.3
St20(10)		2.14	17.3
Stse(10)	.41	.83	.93

Tableau 6.10: Test d'adéquation de la loi Gamma aux durées homogénéisées - $i=1$ (Surplus)

En résumé, d'autres modèles doivent être développés pour représenter les surplus des stations 14(25), 16(25), 18(25), 19(25), se(25), 08(10), 14(10), 18(10), et 20(10) puisque ni le RNH ni le PNH ne sont adéquats.

6.4 Résumé des résultats du PNH et du RNH

Les résultats nous indiquent que le PNH et le RNH ne sont pas adéquats pour toutes les rivières étudiées. Sur les 22 rivières/seuils étudiés, le PNH est adéquat pour modéliser les étiages de 10 stations et les surplus de 8 stations. Le RNH est adéquat pour modéliser les étiages de 7 stations et les surplus de 5 stations. D'autres modèles doivent donc être développés pour les étiages des stations 18(25), 20(25), 22(25), 14(10), et se(10) et les surplus des stations 14(25), 16(25), 18(25), 19(25), se(25), 08(10), 14(10), 18(10), et 20(10). Nous proposons donc d'évaluer le PNHC ($i=1,2$) présenté à la section 5.2.3.

6.5 Vérification de l'adéquation du PNHC

Nous rappelons que nous appliquons le PNHC uniquement aux étiages/surplus des stations pour lesquelles ni le PNH ni le RNH ne sont adéquats. Afin de construire le PNHC décrit dans la section 5.2.3, il est nécessaire d'obtenir des données journalières de température et de précipitations. Nous avons pu le faire pour les rivières canadiennes du tableau 6.11. Le PNHC ne sera donc pas appliqué aux surplus de St16 ni aux étiages de St22 vu la non-disponibilité des données météorologiques. Bien que les températures ne soient pas disponibles pour la Seine, nous avons quand même tenté d'y ajuster le PNHC en tenant compte seulement des précipitations.

Données météorologiques disponibles (températures et précipitations)	
Station	Nombre d'années
St08 (02EC002)	52 (1939-1990)
St14 (02YL001)	37 (1954-1990)
St15 (04LJ001)	24 (1967-1990)
St18 (05BB001)	61 (1930-1990)
St19 (05OD031)	41 (1941-1981)
St20 (07BE001)	54 (1921-1974)

Tableau 6.11: Stations canadiennes utilisées pour l'évaluation des PNHC

6.5.1 Les étiages: résultats du PNHC

Le tableau 6.12 présente pour $i=2$, les paramètres de l'intensité du PNHC estimés selon 5.3.2, ainsi que leurs variances.

Station	\hat{K}^2	$\text{var}(\hat{K}^2)$	$\hat{\theta}^2$	$\text{var}(\hat{\theta}^2)$
St14(10)	1.05	.264	-3.07	.033
St18(25)	3.98	3.82	-3.98	.012
St20(25)	-.64	1.21	-4.68	.04
Stse(10)	9.96	2.59	-6.45	.40

Tableau 6.12: Les paramètres estimés de $\hat{\rho}^2(t, H_t)$ - PNHC

Le tableau 6.13 présente les résultats du test du périodogramme ainsi que le test d'adéquation de la loi exponentielle de paramètre 1, appliqués aux durées homogénéisées du PNHC pour $i=2$. On constate que ce modèle est adéquat pour modéliser les étiages des stations 18(25), 14(10) et 20(10). Il ne l'est pas pour modéliser les étiages de se(10). Ceci peut s'expliquer par le fait que nous n'avons pas les données de température, et que nous n'en avons donc pas pu en tenir compte dans la formulation de $\hat{\rho}^2(t, H_t)$.

Test du périodogramme			Test d'adéquation de la loi exponentielle	
Nom	K-S	w^2	K-S	w^2
St14(10)	1.23	2.17	1.32	1.67
St18(25)	.85	.52	0.83	1.26
St20(25)	1.03	1.94	1.13	1.51
Stse(10)	1.70	2.90	1.56	3.48

Tableau 6.13: Test du PNHC - $i=2$ (Étiage)

6.5.2 Les surplus: résultats du PNHC

Le tableau 6.14 présente pour $i=1$, les paramètres de l'intensité du PNHC estimés selon 5.3.2, ainsi que leurs variances. On note que $\text{var}(\hat{K}_1)$ est élevée pour les stations se(25), 08(10) et 20(10).

Station	$\hat{\beta}_1$	$\text{var}(\hat{\beta}_1)$	\hat{K}_1	$\text{var}(\hat{K}_1)$	$\hat{\theta}_1$	$\text{var}(\hat{\theta}_1)$
St14(25)	-4.46	.01	6.96	1.2	-.44	.017
St18(25)	-6.57	.07	5.50	1.08	-2.29	.08
St19(25)	-4.79	.01	.225	.67	-.42	.02
Stse(25)	-5.04	.030	21.06	83.5	-1.03	.037
St08(10)	-5.76	.052	-5.8	21.43	-1.8	.07
St14(10)	-5.51	.025	7.86	1.97	-.32	.042
St18(10)	-6.9	.067	6.60	.75	-1.44	.057
St20(10)	-7.49	1.23	-3.2	10.09	-1.95	.39

Tableau 6.14: Les paramètres estimés de $\hat{\rho}^1(t, H_i)$ - PNHC

Le tableau 6.15 présente les résultats du test du périodogramme ainsi que le test d'adéquation de la loi exponentielle de paramètre 1, appliqués aux durées homogénéisées du PNHC pour $i=1$.

Test du périodogramme			Test d'adéquation de la loi exponentielle	
Nom	K-S	w^2	K-S	w^2
St14(25)	.67	.43	.89	1.14
St18(25)	1.09	2.13	.93	.75
St19(25)	.96	1.59	2.9	17.0
Stse(25)	1.15	3.03	1.87	7.12
St08(10)	.72	.63	2.3	7.9
St14(10)	1.2	2.4	.67	.54
St18(10)	.97	1.23	1.21	2.05
St20(10)	.73	.34	1.5	4.38

Tableau 6.15: Tests du PNHC - $i=1$ (Surplus)

On constate que ce modèle est adéquat pour les surplus des stations 14(25), 18(25), 14(10) et 18(10). Il ne l'est pas pour les surplus des stations 19(25), se(25), 08(10), et de 20(10).

Dans le cas de St19(25) on remarque sur la figure 6.19 que les temps d'arrivée des événements de type 1 sont distribués partout sur l'année contrairement à, par exemple, St18(25) et St14(10) (figures 6.22 et 6.23) où ces occurrences sont confinées à des périodes spécifiques. Ceci est aussi vrai pour la station se(25) (figure 6.20). De plus, nous ne disposons pas de données de température pour cette station et nous n'avons donc pas pu en tenir compte. Pour ce qui est de St08(10), on remarque la présence d'une donnée singulière (*outlier*) (figure 6.21) ce qui peut expliquer la forte variance de \hat{K}_1 . Dans le cas de St20(10), le mauvais ajustement est probablement dû au faible nombre d'événements observés (13) (figure 6.22) ce qui explique la forte variance de \hat{K}_1 .

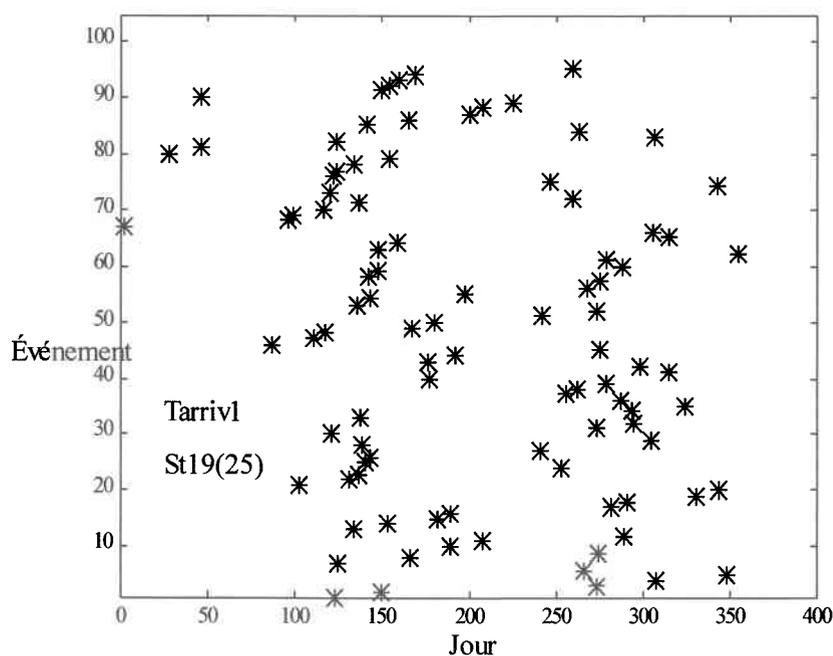


Figure 6.19: Temps d'arrivée dans l'année des événements de rupture de surplus - St19(25)

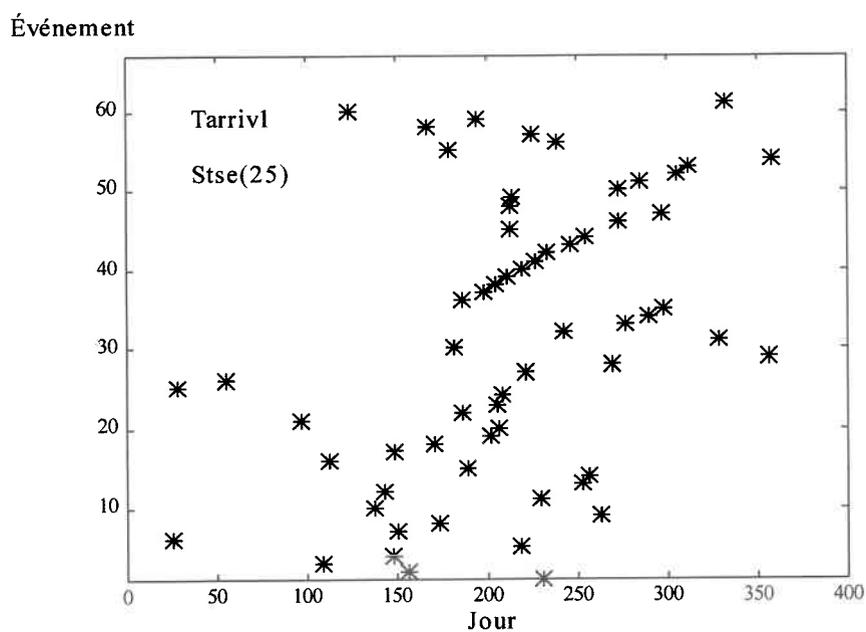


Figure 6.20: Temps d'arrivée dans l'année des événements de rupture de surplus - Stse(25)

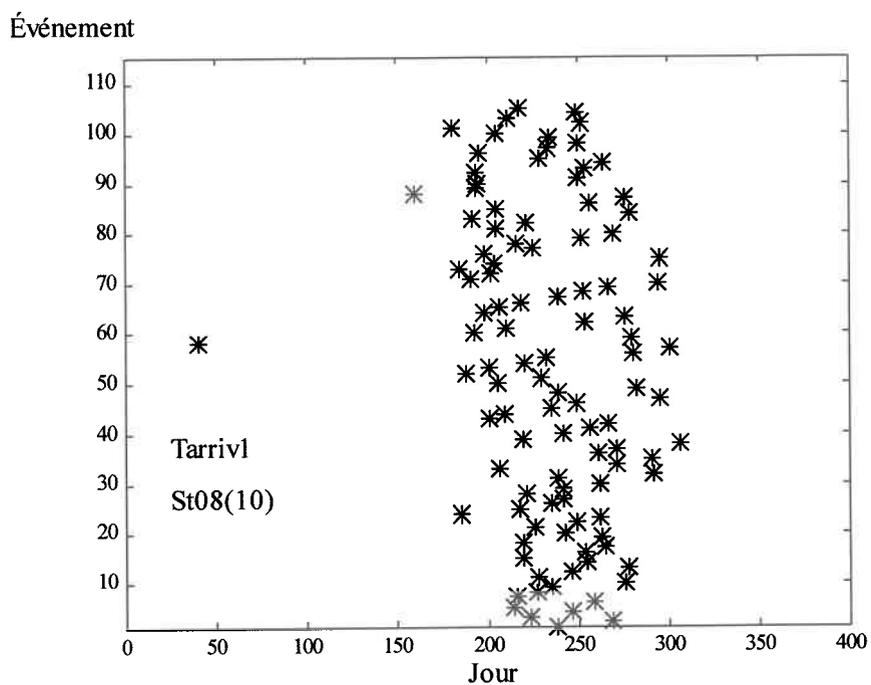


Figure 6.21: Temps d'arrivée dans l'année des événements de rupture de surplus - St08(10)

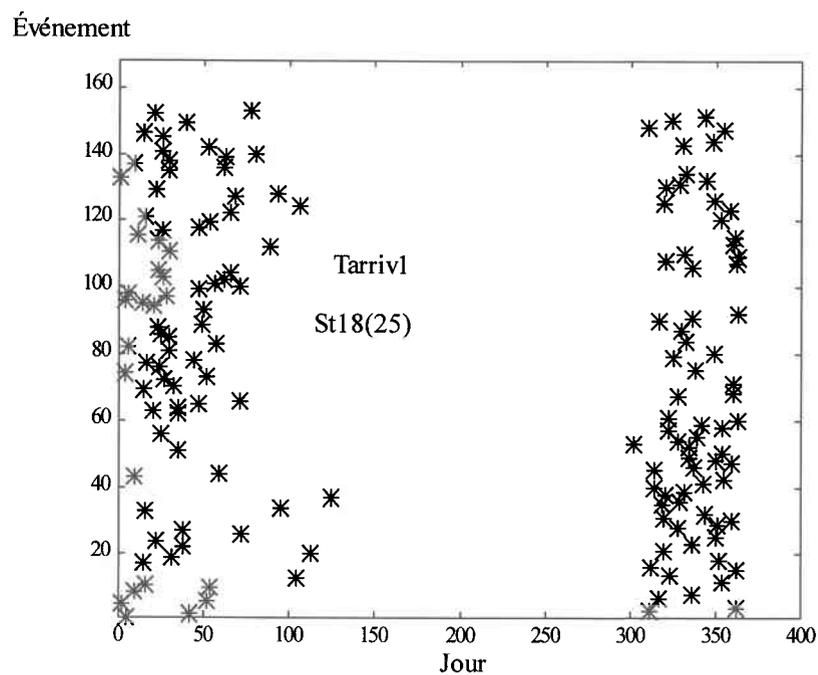


Figure 6.22: Temps d'arrivée dans l'année des événements de rupture de surplus - St18(25)

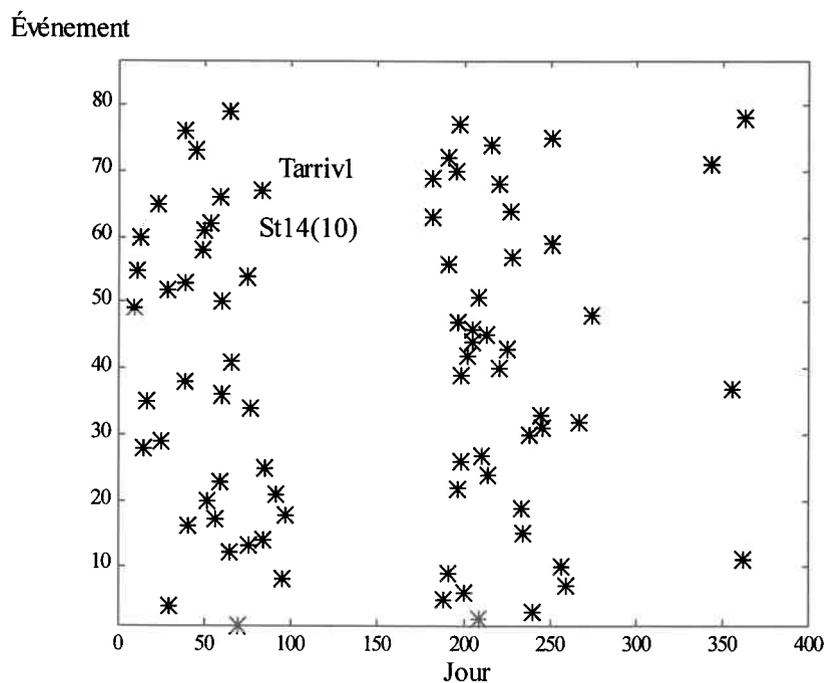


Figure 6.23: Temps d'arrivée dans l'année des événements de rupture de surplus - St14(10)

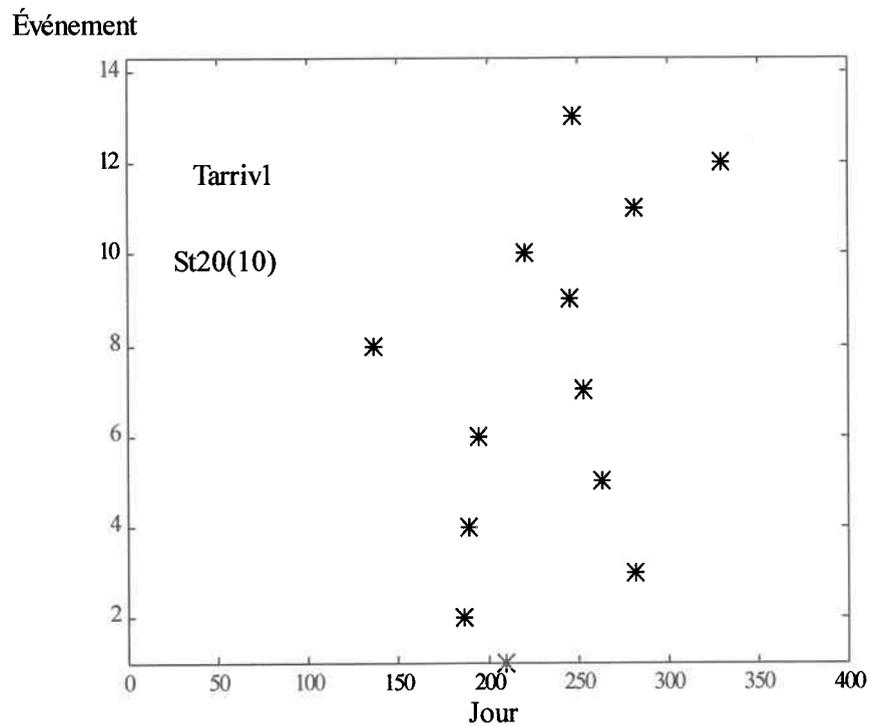


Figure 6.24: Temps d'arrivée dans l'année des événements de rupture de surplus - St20(10)

Le tableau 6.16 présente une synthèse des résultats obtenus. Nous notons que le PNH s'applique aux étiages de 10 et aux surplus de 8 stations. Le RNH s'applique aux étiages de 7 et aux surplus de 5 stations, et le PNHC s'applique aux étiages de 3 et aux surplus de 4 stations. Nous n'avons pas pu construire un modèle adéquat pour les étiages de Stse(10) ni pour les surplus de St19(25), Stse(25), St08(10) et St20(10) pour les raisons évoquées ci-haut. La non-disponibilité (N/D) de données météorologiques ne nous a pas permis de développer un modèle pour les étiages de St22(25) ni pour les surplus de St16(25).

Les modèles adéquats		
Station	Étiages	Surplus
St06(25)	PNH	RNH
St07(25)	PNH	PNH
St08(25)	RNH	PNH
St14(25)	PNH	PNHC
St15(25)	RNH	RNH
St16(25)	RNH	N/D
St18(25)	PNHC	PNHC
St19(25)	PNH	Autre
St20(25)	PNHC	PNH
St22(25)	N/D	PNH
Stse(25)	RNH	Autre
St06(10)	PNH	RNH
St07(10)	PNH	PNH
St08(10)	RNH	Autre
St14(10)	PNHC	PNHC
St15(10)	RNH	PNH
St16(10)	PNH	PNH
St18(10)	RNH	PNHC
St19(10)	PNH	RNH
St20(10)	PNH	Autre
St22(10)	PNH	PNH
Stse(10)	Autre	RNH

Tableau 6.16: Synthèse des résultats

6.6 Analyse de sensibilité et simulation

Nous avons effectué l'analyse de sensibilité pour les étiages et les surplus de chaque rivière/seuil pour lesquels un modèle s'est avéré adéquat (tableaux 6.17 et 6.18). Normalement, le processus alterné devrait être généré en simultané. Pour les rivières où nous

n'avions un modèle adéquat que pour un seul des deux états, nous avons supposé un PNH pour l'état non modélisable pour les fins de simulation et d'analyse de sensibilité.

Nombre de fois où le test rejette l'hypothèse nulle, $\alpha=.05$						
Test du périodogramme			Test d'adéquation de la loi exponentielle		Test d'adéquation de la loi Gamma	
Nom	K-S	w ²	K-S	w ²	K-S	w ²
St06(25) (PNH)	2	4	4	9	-	-
St07(25) (PNH)	2	4	27	31	-	-
St08(25) (RNH)	3	4	-	-	38	40
St14(25) (PNH)	4	2	5	14	-	-
St15(25) (RNH)	8	14	-	-	51	59
St16(25) (RNH)	8	10	-	-	33	39
St18(25)	7	7	54	73	-	-
St19(25) (PNH)	5	3	15	18	-	-
St20(25)	3	4	6	7	-	-
Stse(25) (RNH)	1	1	-	-	10	15
St06(10) (PNH)	4	5	10	15	-	-
St07(10) (PNH)	6	4	27	32	-	-
St08(10) (RNH)	3	3	-	-	46	58
St14(10)	3	3	38	59	-	-
St15(10) (RNH)	6	4	-	-	25	22
St16(10) (PNH)	3	5	32	32	-	-
St18(10) (RNH)	6	7	-	-	57	65
St19(10) (PNH)	2	4	14	14	-	-
St22(10) (PNH)	2	3	49	74	-	-

Tableau 6.17: Analyse de sensibilité - $i=2$ (Étiage)

Nombre de fois où le test rejette l'hypothèse nulle, $\alpha=.05$						
Test du périodogramme			Test d'adéquation de la loi exponentielle		Test d'adéquation de la loi Gamma	
Nom	K-S	w ²	K-S	w ²	K-S	w ²
St06(25) (RNH)	4	4	-	-	11	16
St07(25) (PNH)	9	6	23	24	-	-
St08(25) (PNH)	4	3	46	52	-	-
St14(25) (PNHC)	4	3	17	27		
St15(25) (RNH)	4	8	-	-	35	33
St18(25) (PNHC)	2	6	66	66	-	-
St20(25) (PNH)	0	2	6	6	-	-
St22(25) (PNH)	3	5	52	59	-	-
St06(10) (RNH)	3	4	-	-	7	12
St07(10) (PNH)	4	4	51	54	-	-
St14(10) (PNHC)	1	0	81	84		
St15(10) (PNH)	3	3	16	17	-	-
St16(10) (PNH)	7	7	17	20	-	-
St18(10) (PNHC)	4	3	11	13	-	-
St19(10) (RNH)	3	4	-	-	4	8
St22(10) (PNH)	1	1	38	55	-	-
Stse(10) (RNH)	2	6	-	-	48	54

Tableau 6.18: Analyse de sensibilité - $i=1$ (Surplus)

Nous n'avons pas réussi à effectuer l'analyse de sensibilité pour St20(10) à cause du faible nombre d'événements générés par simulation. Les résultats de l'analyse de sensibilité basée sur 100 séries de paramètres générés nous indiquent que lorsque le PNH est adéquat, il semble l'être pour un grand nombre de réalisations de processus. Ceci est moins vrai du PNHC qui semble sensible à l'incertitude.

Les modèles construits peuvent servir pour générer des processus alternés tels que décrits à l'annexe 4. À titre d'illustration, nous présentons sur les figures 6.25 à 6.28 des exemples de série alternée de surplus/étiage simulée et observée sur 10 ans: St07(25) (PNH/PNH), St08(25) (PNH/RNH), St14(25)(PNHC/PNH), St15(25)(RNH/RNH).

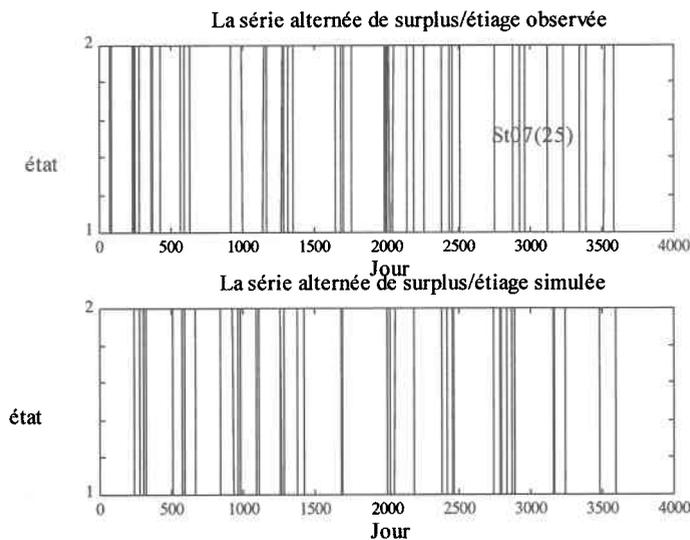


Figure 6.25: Des séries de surplus/étiage simulées et observées à St07(25) (PNH/PNH)

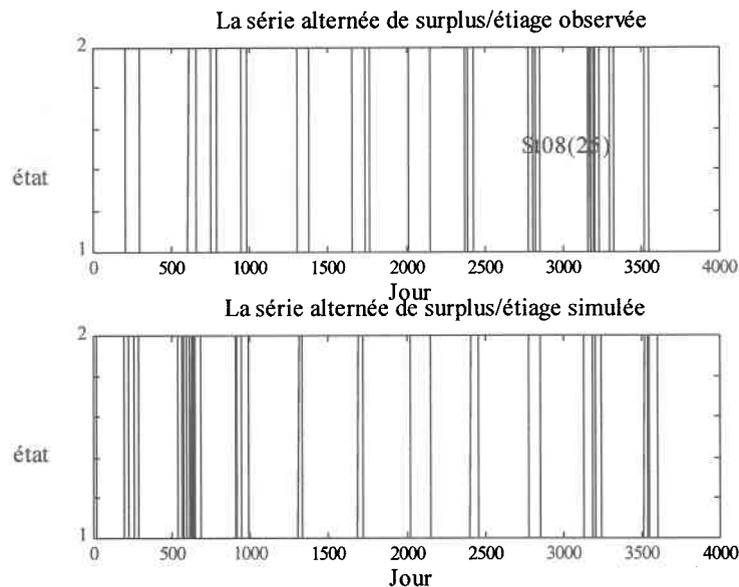


Figure 6.26: Des séries de surplus/étiage simulées et observées à St08(25) (PNH/RNH)

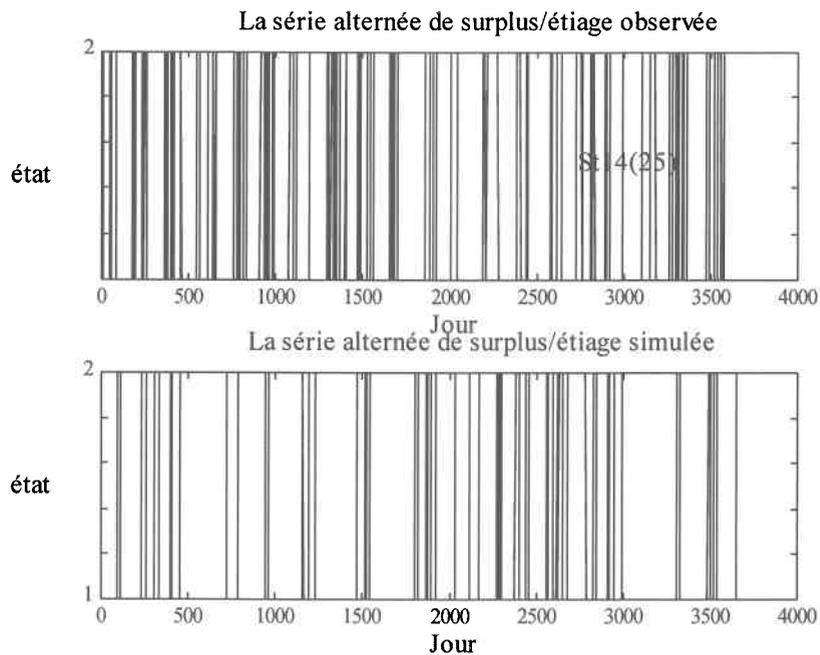


Figure 6.27: Des séries de surplus/étiage simulées et observées à St14(25) (PNHC/PNH)

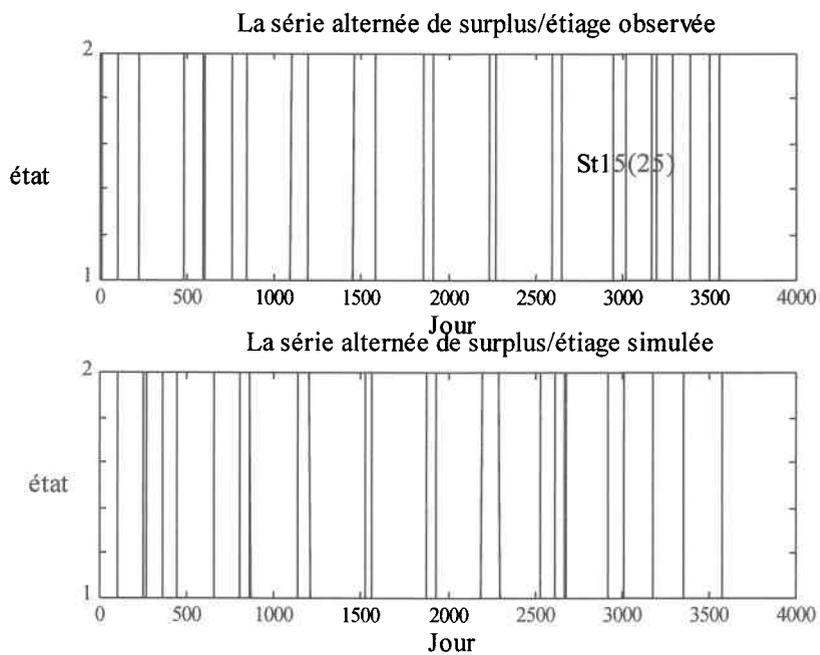


Figure 6.28: Des séries de surplus/étiage simulées et observées à St15(25) (RNH/RNH)

De plus, nous présentons sur les figures 6.29 à 6.32 les histogrammes de durées d'étiage observées et simulées pour les stations St06(25)(PNH), St08(25)(RNH), et St14(10)(PNHC), et les histogrammes de durées de surplus observées et simulées pour St14(10) (PNHC). Les comparaisons entre les données simulées et les données observées sont à titre qualitatif seulement.

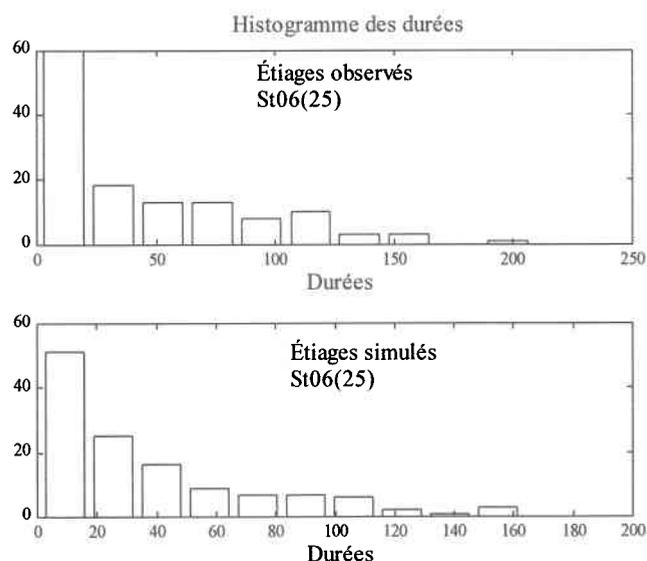


Figure 6.29: Histogramme des durées d'étiage observées et simulées à St06(25)- PNH

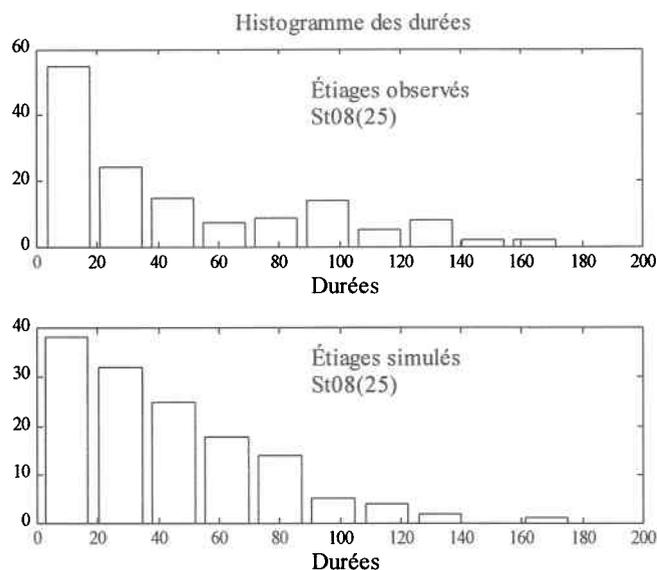


Figure 6.30: Histogramme des durées d'étiage observées et simulées à St08(25)- RNH

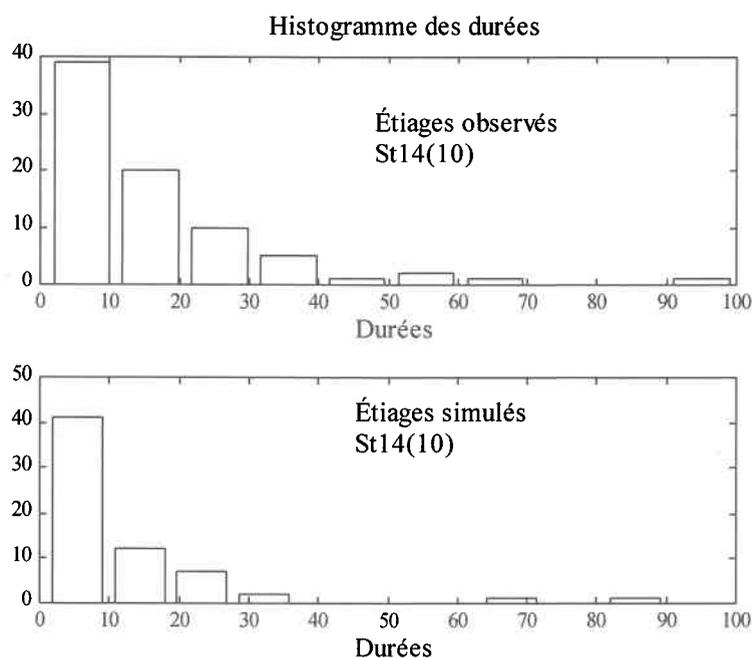


Figure 6.31: Histogramme des durées d'étiage observées et simulées à St14(10)- PNHC

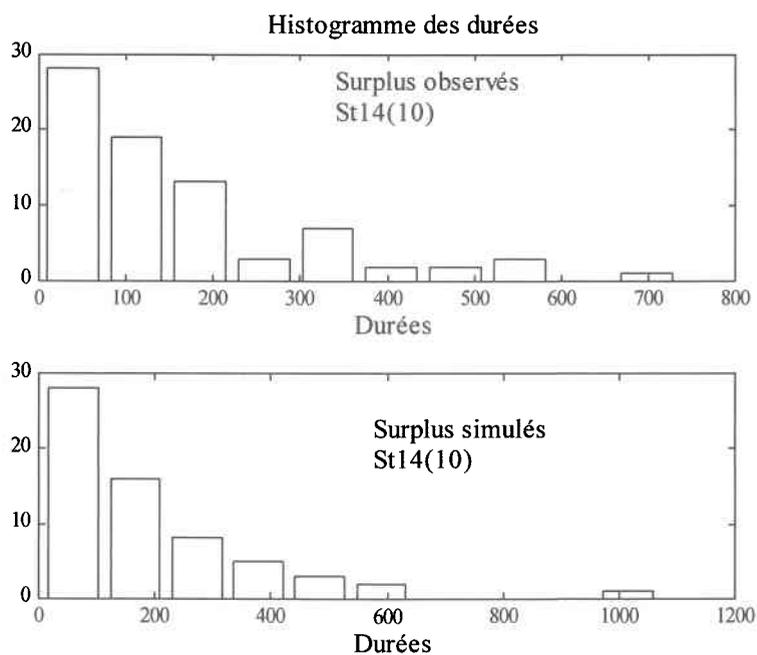


Figure 6.32: Histogramme des durées de surplus observées et simulées à St14(10)- PNHC

6.7 Conclusions

Un de nos sous-objectifs était d'évaluer le PNH et le RNH sur les rivières canadiennes afin de voir si ces modèles sont acceptables en la présence d'étiages d'hiver et de comparer nos résultats avec ceux de North (1981). Nous avons utilisé les données journalières des rivières suivantes: Pigeon, Magnetawan Nord, Black, Missinaibi, et Turtle en Ontario, Upper Humber à Terre Neuve, Harricana au Québec, Bow et Athabasca en Alberta, Sprague au Manitoba, et Bar-sur-Seine en France. Le débit moyen journalier variait entre $1.6 \text{ m}^3/\text{s}$ et $442 \text{ m}^3/\text{s}$, et la superficie du bassin variait entre 321 km^2 et 74600 km^2 . Nous avons défini les étiages à partir des seuils correspondant aux 10^{ème} et 25^{ème} quantiles, c'est-à-dire les débits dépassés 90 et 75% du temps respectivement.

Nous avons constaté l'adéquation du PNH comme modèle d'occurrence de surplus ou d'étiage pour 18 séries (rivières/seuils), et celle du RNH pour 12 séries (rivières/seuils). Ces résultats sont assez surprenants compte tenu de la simplicité relative de ces modèles. Ils appuient toutefois les résultats de North (1981) qui a appliqué le PNH et le RNH à 17 rivières, à un seuil correspondant au 95^{ème} quantile. Il a constaté que le PNH est acceptable pour les petites rivières de plaines n'ayant pas d'étiage d'hiver (11 rivières sur 17). Cette conclusion ne semble pas s'appliquer aux rivières que nous avons étudiées puisque toutes nos rivières canadiennes présentaient des étiages d'hiver. Dans le cas de North (1981), le RNH semblait adéquat pour les 5 autres rivières, et aucun modèle n'était acceptable pour la 17^{ème} rivière. Les surfaces des bassins étudiés variaient entre 0.5 et 65 km^2 , et un bassin de 165 km^2 est considéré comme grand. Le débit moyen variait entre $.01$ et $3.16 \text{ m}^3/\text{s}$.

Afin de tenir compte des deux étiages (été et hiver) nous avons tenté d'ajuster un modèle que nous appellerons PNH2, qui correspond au modèle PNH avec la différence que la fonction d'intensité a la forme suivante: (2 harmoniques). Ce modèle s'avéra adéquat pour une des 7 séries que nous n'avons pas réussi à modéliser, à savoir les surplus de St16(25). Les difficultés de modéliser les 6 autres séries proviennent surtout du fait que les temps d'occurrence des événements sont répartis un peu partout sur l'année.

Nous avons développé et appliqué le PNH2 uniquement pour les rivières pour lesquelles ni le RNH ni le PNH n'était acceptable, et ce, dans un souci de parcimonie et de simplicité. Pour ce faire, nous avons dû essayer plusieurs types de fonctions d'intensité présentées à

l'annexe 5. Nous avons trouvé que le PNHC était adéquat pour les étiages de 3 des 4 séries (rivières/seuils) considérées et pour les surplus de 4 des 8 séries (rivières/seuils) que nous avons tenté de modéliser.

Suite à nos travaux et à ceux de North (1981), il s'avère que des processus ponctuels assez simples tels que le PNH et le RNH peuvent être utilisés pour modéliser en pratique les occurrences des étiages et des surplus. Pour les cas où ces processus ne sont pas adéquats, on peut appliquer le PNHC que nous avons proposé. Ce modèle est parcimonieux (deux paramètres pour les étiages et trois paramètres pour les surplus). Il suppose toutefois la disponibilité de données météorologiques journalières. L'approche alternée est nécessaire si on veut tenir compte des durées des étiages, ce qui est une variable importante. Tous les modèles présentés peuvent être utilisés pour générer par simulation des séries de surplus et d'étiage.

Nous avons remarqué que l'adéquation des modèles d'étiage et de surplus variaient en fonction du seuil pour la plupart des rivières (6 sur 11 pour les étiages et 6 sur 11 pour les surplus, mais pas les mêmes six rivières). Ceci illustre la grande influence du seuil sur le type de modèle qui sera adéquat, d'où l'impossibilité de conclure à un modèle "universel". Le même modèle d'occurrence s'appliquait et aux occurrences de surplus et aux occurrences des étiages pour 7 des 22 rivières/seuils. Nous n'avons pas réussi à dégager une classification associant un type de rivière à un type de modèle en fonction de la situation géographique, ou du débit moyen, ou de la surface du bassin, ou du nombre d'événements.

Les résultats que nous avons présentés ont été obtenus en utilisant un seuil de confiance $\alpha=0.05$. Afin de vérifier la robustesse de nos résultats, nous présentons à l'annexe 9 les résultats des tests statistiques pour $\alpha=0.01$.

En résumé, nous constatons que des modèles basés sur des processus ponctuels relativement simples, sont tout à fait adéquats pour modéliser les occurrences des étiages et des surplus tout en tenant compte de la non-stationnarité du phénomène. Retenons cependant l'influence du choix du seuil sur les résultats.

7. QUANTIFICATION DU RISQUE

Un des objectifs de ce travail est de faire le lien entre l'hydrologie statistique et l'hydrologie opérationnelle. Nous atteindrons ce but en calculant des indices de risque basés sur les durées des étiages à partir des modèles d'occurrence développés. En pratique, les indices de risque proposés peuvent être utilisés comme critères à optimiser lors de l'élaboration de politiques de gestion des ressources en eau. Il est important de noter que les résultats présentés dans ce chapitre sont plutôt de nature exploratoire et s'inscrivent dans un cadre d'évaluation et de quantification du risque.

Dans ce travail, nous considérons une rivière comme un système hydrique qui peut être dans un état non satisfaisant (étiage, débit en dessous d'un seuil) ou dans un état satisfaisant (surplus, débit au-dessus du seuil). X est la durée d'étiage, Y est la durée de surplus, et Z est l'intervalle entre deux débuts d'étiage. Notre approche consiste à définir pour le PNH, le RNH, et le PNHC, cinq indices de risque basés sur les questions suivantes:

- $I_1(x,u)$: quelle est la probabilité qu'un étiage qui a débuté à l'instant u dure plus de x jours, $P(X > x | T = u)$?
- $I_2(x,u)$: quelle est la probabilité qu'un étiage qui a débuté à l'instant u se termine étant donné qu'il a déjà duré x jours, $P(X = x | X \geq x, T = u)$?
- $I_3(y,u)$: quelle est la probabilité d'entrer en étiage sur l'intervalle $(u, u+y)$ étant donné qu'on est entré en surplus à l'instant u , $P(Y < y | T = u)$?
- $I_4(z,u)$: quelle est la probabilité d'entrer à nouveau en étiage sur $(u, u+z)$ étant donné qu'on est entré en étiage à l'instant u , $P(Z < z | T = u) = D(z | T = u)$?
- $I_5(x,u)$: quelle est la probabilité d'être en étiage au temps t étant donné que le dernier étiage a débuté à l'instant u , $P_2(u+x | T = u)$?

Nous présentons dans les paragraphes qui suivent les expressions analytiques de ces indices. Afin d'illustrer leur applicabilité, nous les évaluons pour les stations décrites au tableau 7.1, et ce pour des durées d'étiage (X) et de surplus (Y), ainsi que des intervalles entre deux étiages (Z), entre 1 et 365 jours. Le PNH et le RNH sont des processus périodiques de période 365 jours (ce qui n'est pas le cas du PNHC). Les dates d'occurrence du dernier événement u , (2 janvier pour le PNH, 30 avril pour le RNH) ont été choisies au hasard. Une hypothèse du PNHC développé est que les précipitations et les températures qui conditionnent l'intensité sont connues. C'est pour cette raison que nous avons dû spécifier une date exacte pour u , le dernier événement du PNHC. Ceci explique la notation différente pour u (PNHC) de la notation pour u (PNH et RNH). Nous avons choisi arbitrairement d'utiliser les données observées au 7/12/43 pour le PNHC.

Station	Modèle étiage	Modèle surplus	Paramètres étiages				Paramètres surplus				Date du dernier événement
			β_2	K_2	θ_2	ρ_2	β_1	K_1	θ_1	ρ_1	
St07(25)	PNH	PNH	-3.9	-.53	4.97	-	-5.29	-1.64	.91		2/01
St15(25)	RNH	RNH	-3.9	2.5	4.93	1.86	-5.29	.96	1.08	2.28	30/04
St18(25)	PNHC	PNHC	-	3.98	-3.98	-	-6.57	5.50	-2.29	-	7/12/43

Tableau 7.1: Stations pour lesquelles les indices de risque sont calculés

Nous rappelons que pour le PNHC, les précipitations sont fixées à 0 lorsque la température est inférieure à -2. La variable $nbjoursec(t)$ est le nombre de jours sans précipitation précédant t divisé par 365. La variable $pnorm(t)$ est la hauteur de pluie tombée pendant les 90 jours précédant t normée par cette quantité maximale observée:

$$pnorm(t) = \frac{\sum_{kk=t-90-1}^{kk=t} P(kk)}{\max_t \left(\sum_{kk=t-90-1}^{kk=t} P(kk) \right)}$$

Les résultats que nous présentons sont obtenus à l'aide de MATLAB et MAPLE. Dans le cas des PNH et RNH, les intégrales sont estimées par des algorithmes d'intégration numérique. Dans le cas du PNHC, les intégrales des équations 7.5 et 7.14 sont obtenues par approximation en sommant sur la variable d'intégration a , et ce entre les bornes d'intégration.

7.1 La loi des durées d'étiage, $I_1(x,u)$

Le premier indice proposé $I_1(x,u)$, est la probabilité qu'un étiage qui a débuté à l'instant u , dure plus de x jours: $P(X > x | T = u)$. Il est calculé étant donné l'historique du processus et l'instant d'occurrence u du dernier étiage. Cette quantité est (équations 5.1, 5.3):

$$I_1(x,u) = P(X > x | T = u, H_u) = \exp\left(-\int_u^{u+x} \rho^2(a, H_u) da\right) \quad (7.1)$$

où $\rho^2(t, H_t)$ est l'intensité du processus de rupture des étiages dans le processus alterné, à l'instant t .

PNH

Dans le cas du PNH, $\rho^2(a, H_a) = \exp(\beta_2 + K_2 \sin(\omega a + \theta_2))$. $I_1(x,u)$ est donc:

$$P(X > x | T = u) = \exp\left(-\int_u^{u+x} \exp(\beta_2 + K_2 \sin(\omega a + \theta_2)) da\right) \quad (7.2)$$

où $\omega = \frac{2\pi}{365}$. $I_1(x, 2 \text{ janvier})$ est présenté sur la figure 7.1 en fonction de la durée x pour la station St07(25). On y voit que la probabilité d'étiage de longue durée diminue lorsque la longueur de la durée x augmente.

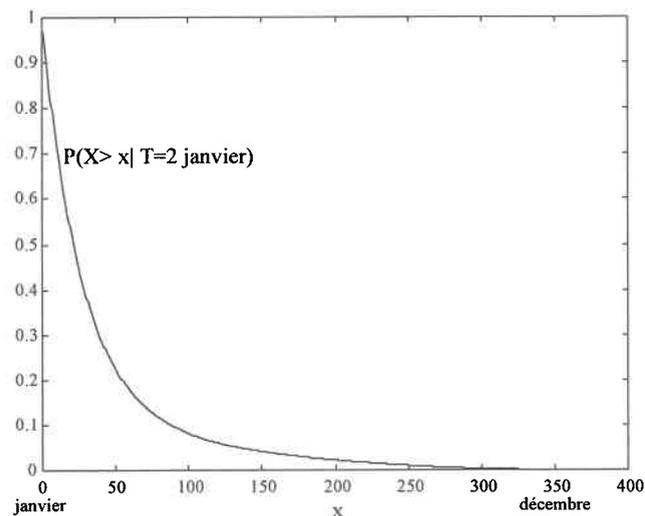


Figure 7.1: $I_1(x, 2 \text{ janvier}) = P(X > x | T = 2 \text{ janvier})$ en fonction de x - PNH (St07(25))

RNH

Dans le cas du RNH, $\rho^2(a, H_u) = \varphi^2(a - u, a) = \varphi(x')\lambda^2(a)$, où $x' = \int_u^a \lambda^2(v)dv$, $\lambda^2(v) = \exp(\beta_2 + K_2 \sin(\omega v + \theta_2))$, et où $\varphi(x')$ est le taux de rupture d'un RH dont les durées x' suivent une loi Gamma de paramètres $\alpha = \rho$. Or $\varphi(x')$ est (Ross, 1985):

$$\varphi(x') = \frac{1}{\int_0^\infty \exp^{-\rho s} \left(1 + \frac{s}{x'}\right)^{\rho-1} ds} \tag{7.3}$$

$I_1(x,u)$ est donc:

$$P(X > x | T = u) = \exp \left(- \int_u^{u+x} \frac{\exp(\beta_2 + K_2 \sin(\omega a + \theta_2))}{\int_0^\infty \exp^{-\rho s} \left(1 + \frac{s}{\int_u^a \exp(\beta_2 + K_2 \sin(\omega v + \theta_2)) dv}\right)^{\rho-1} ds} da \right) \tag{7.4}$$

Nous présentons, sur la figure 7.2, $I_1(x, 30 \text{ avril})$ pour la station St15(25). On y voit que la courbe de probabilité des durées x s'estompe beaucoup plus rapidement que $I_1(x, 2 \text{ janvier})$ pour St07(25). Ceci s'explique par le fait qu'un étiage qui a débuté le 30 avril (figure 7.2) a moins de chance de se poursuivre qu'un étiage qui a débuté le 2 janvier (figure 7.1) puisque le printemps est une saison qui présente beaucoup de précipitations pluvieuses.

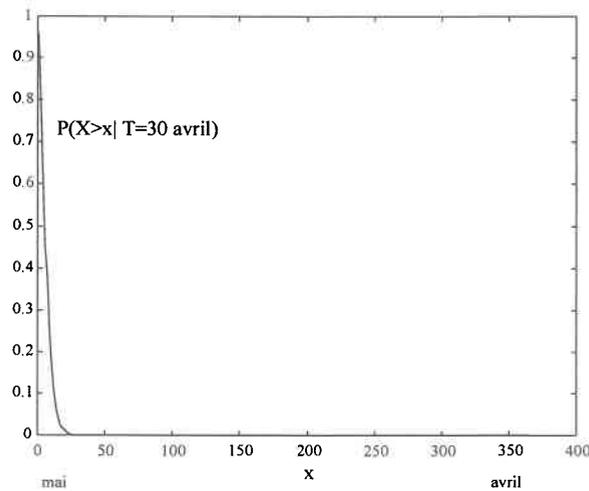


Figure 7.2: $I_1(x, 30 \text{ avril}) = P(X > x | T = 30 \text{ avril})$ en fonction de x - RNH (St15(25))

PNHC

Dans le cas du PNHC, $\rho^2(t, H_t)$ dépend uniquement de la quantité de pluie tombée les 90 jours précédant t mesurée par $\text{pnorm}(t)$. La figure 7.3 présente $\text{pnorm}(t)$ pour t entre le 7 décembre 1943 et le début décembre 1944 (u et $u+365$).

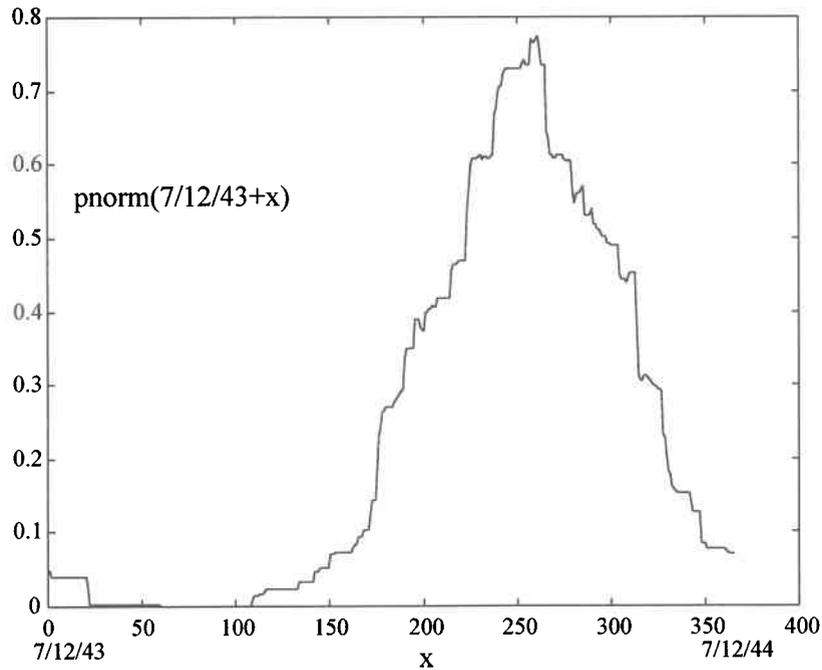


Figure 7.3: $\text{pnorm}(t, t=7/12/43+x)$ St18

Pour le PNHC $\rho^2(a, H_a) = \exp(K_2 \text{pnorm}(a) + \theta_2)$. $I_1(x, u)$ est donc:

$$P(X > x | T = u, H_u) = \exp\left(- \int_u^{u+x} \exp(K_2 \text{pnorm}(a) + \theta_2) da\right) \quad (7.5)$$

Nous avons calculé $I_1(x, 7/12/43)$ pour St18(25). Les résultats sont présentés sur la figure 7.4. On y voit que la probabilité que l'étiage dépasse une durée x d'autour 200 jours est nulle, ce qui s'explique par le fait que la valeur de $\text{pnorm}(7/12/43+x)$ s'accroît rapidement autour de $x=200$ (figure 7.3).

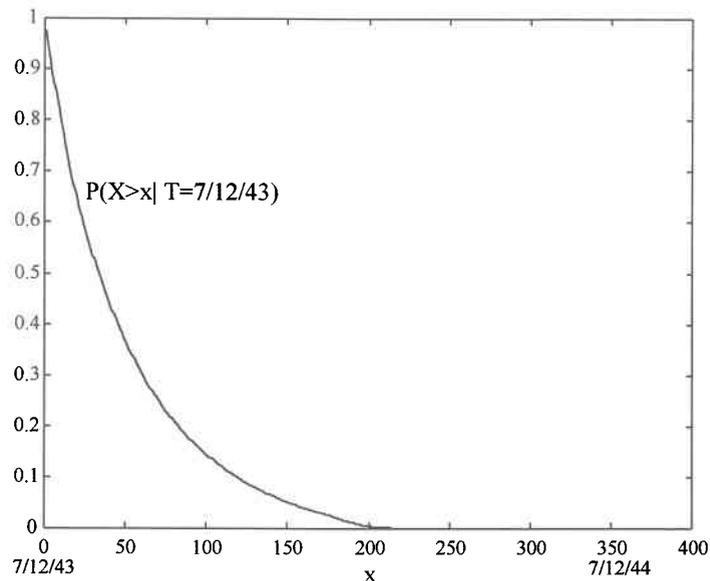


Figure 7.4: $I_1(x, 7/12/43) = P(X > x | T = 7/12/43)$, en fonction de x - PNHC (St18(25))

7.2 La probabilité qu'un étiage se termine, $I_2(x, u)$

Le deuxième indice de risque que nous proposons $I_2(x, u)$, est la probabilité de sortir d'un étiage qui a déjà duré x unités de temps, étant donné qu'il a commencé à l'instant u , $P(X=x | X \geq x, T=u)$. C'est la probabilité de rupture d'un étiage donnée par:

$$I_2(x, u) = P(X=x | X \geq x, T=u) = \rho^2(x+u, H_{x+u}) \Delta t \quad (7.6)$$

Puisque nous utilisons des données journalières dont le pas de temps Δt est 1 jour, $I_2(x, u)$ correspond au taux de rupture des étiages:

$$I_2(x, u) = \rho^2(x+u, H_{x+u}) \quad (7.7)$$

PNH

Dans le cas du PNH, $I_2(x, u)$ est:

$$P(X=x|X \geq x, T=u) = \lambda^2(u+x) = \exp(\beta_2 + K_2 \sin(\omega(u+x) + \theta_2)) \quad (7.8)$$

Nous présentons $I_2(x, 2 \text{ janvier})$ pour la station St07(25), sur la figure 7.5. Cette fonction correspond à la probabilité qu'un étiage qui a débuté un 2 janvier et qui a duré x jours se termine immédiatement. On voit que la probabilité de rupture des étiages décroît rapidement durant l'hiver (étiage d'hiver). Cette fonction peut être utilisée pour obtenir $I_2(x, u')$, pour tout u' entre 1/01 et 31/12, car le taux de rupture est périodique de période 365 jours. De plus, il dépend uniquement de l'instant où l'on se trouve, et non du temps écoulé depuis le dernier événement. Il suffit donc, pour obtenir $I_2(x, u')$, pour tout u' entre 1/01 et 31/12, de décaler la fonction de la figure 7.5 par $u-u'$ jours vers la droite.

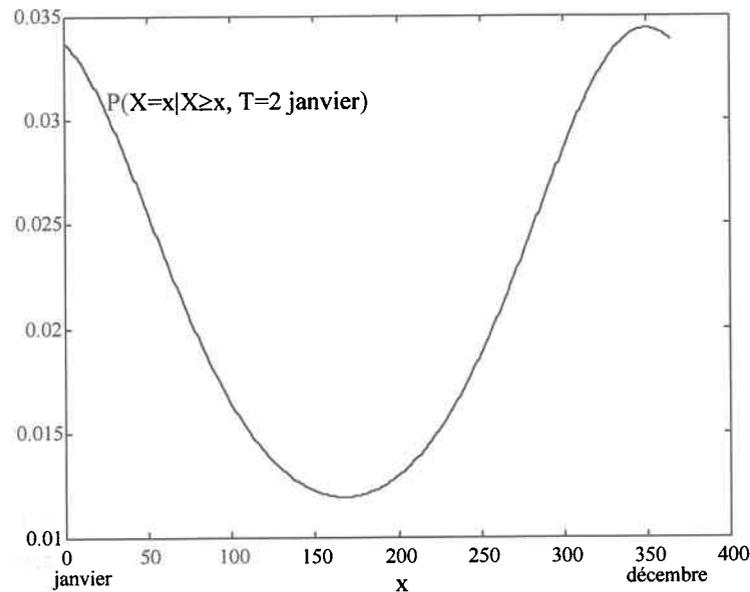


Figure 7.5: $I_2(x, 2 \text{ janvier}) = P(X=x | X \geq x, T=2 \text{ janvier})$ en fonction de x - PNH (St07(25))

RNH

Dans le cas du RNH, $I_2(x, u)$ est:

$$P(X=x | X \geq x, T=u) = \varphi^2(x, u+x) = \frac{\exp(\beta_2 + K_2 \sin(\omega(u+x) + \theta_2))}{\int_0^\infty \exp^{-ps} \left(1 + \frac{s}{\int_u^{u+x} \lambda^2(a) da} \right)^{\rho-1} ds} \quad (7.9)$$

où $\lambda^2(u+x) = \exp(\beta_2 + K_2 \sin(\omega(u+x) + \theta_2))$. Nous présentons $I_2(x, 30 \text{ avril})$ pour la station St15(25) sur la figure 7.6. On y voit que la probabilité de rupture augmente assez rapidement au printemps. Cette fonction est périodique, mais dépend du temps écoulé depuis le dernier événement. Nous ne pouvons donc pas utiliser la même fonction pour calculer $I_2(x, u')$ pour tout u' entre 1/01 et 31/12, comme c'est le cas du PNH.

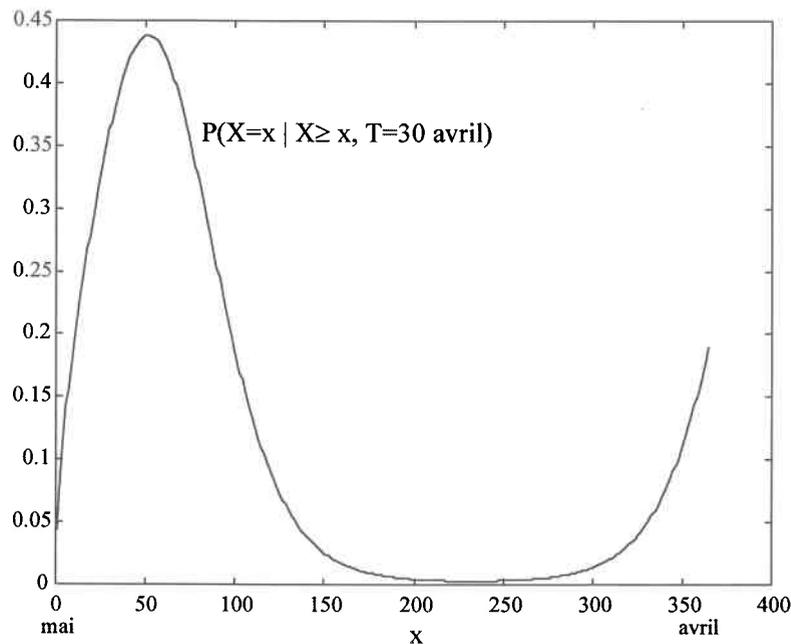


Figure 7.6: $I_2(x, 30 \text{ avril}) = P(X=x | X \geq x, T=30 \text{ avril})$ en fonction de x - RNH (St15(25))

PNHC

Pour le PNHC, $I_2(x, u)$ dépend uniquement de $\text{pnorm}(u+x)$. Cet indice est égal à :

$$P(X=x | X \geq x, T=u) = \rho^2(u+x, H_{u+x}) = \exp(K_2 \text{pnorm}(u+x) + \theta_2) \quad (7.10)$$

Nous présentons sur la figure 7.7, $I_2(x, 7/12/43)$ pour la station St18(25). Cette courbe est proportionnelle à celle de la figure 7.3. Il est clair que $I_2(x, u)$ est grand lorsque $\text{pnorm}(7/12/43+x)$ est grand (autour de $x=250$).

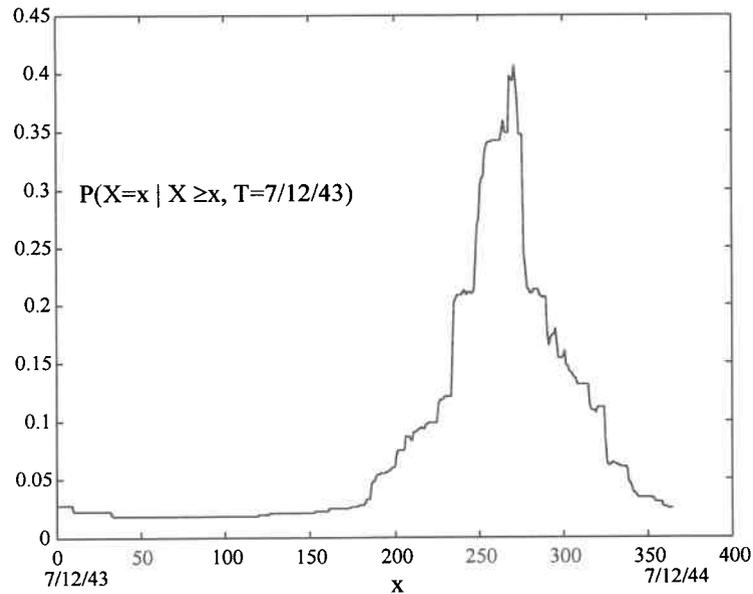


Figure 7.7: $I_2 = P(X=x | X \geq x, T=7/12/43)$ en fonction de x - PNHC (St18(25))

7.3 La probabilité d'entrer en étiage, $I_3(y,u)$

Le troisième indice proposé est $I_3(y,u)$, la probabilité d'entrer en étiage sur $(u, u+y)$, étant donné qu'on est rentré en surplus à l'instant u (dernier événement). Il correspond à la probabilité que la période de surplus dure au plus y jours:

$$I_3(y,u) = P(Y \leq y | T=u) = 1 - P(Y > y | T=u, H_u) = 1 - \exp\left(-\int_u^{u+y} \rho^1(a, H_u) da\right) \quad (7.11)$$

PNH

Dans le cas du PNH, $I_3(y,u)$ est:

$$P(Y \leq y | T=u) = 1 - \exp\left(-\int_u^{u+y} \exp(\beta_1 + K_1 \sin(\omega a + \theta_1)) da\right) \quad (7.12)$$

Nous présentons $I_3(y, 2 \text{ janvier})$ sur la figure 7.8 pour la station St07(25). Il est clair que cet indice augmente avec la durée de l'état surplus.

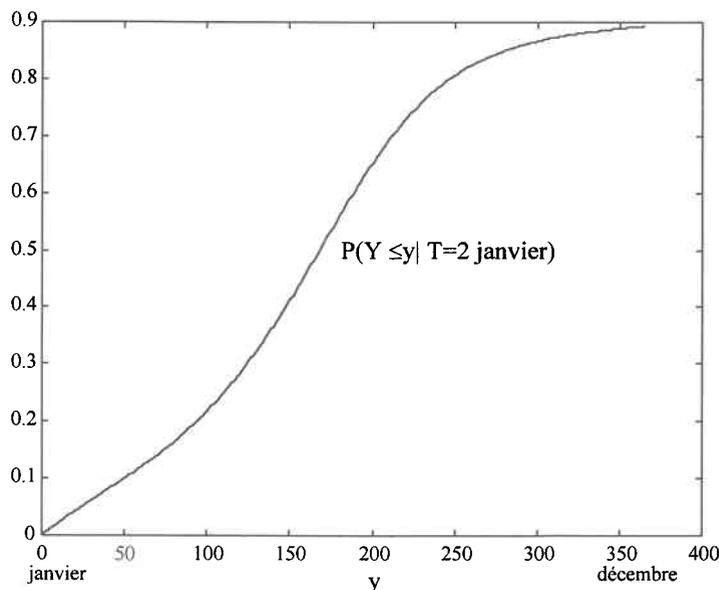


Figure 7.8: $I_3(y,u)=P(Y \leq y | T=2 \text{ janvier})$ en fonction de y - PNH (St07(25))

RNH

Dans le cas du RNH, $I_3(y,u)$ est:

$$P(Y \leq y | T=u) = 1 - \exp \left(- \int_u^{u+y} \frac{\exp(\beta_1 + K_1 \sin(\omega a + \theta_1))}{\int_0^\infty \exp^{-ps} \left(1 + \frac{s}{u+a} \right)^{\rho-1} ds} da \right) \quad (7.13)$$

où $\lambda^1(a) = \exp(\beta_1 + K_1 \sin(\omega(a) + \theta_1))$. Nous présentons sur la figure 7.9 $I_3(y,30 \text{ avril})$ pour St15(25). On voit que la courbe s'accroît moins rapidement que sur la figure 7.8. La probabilité de rentrer en étiage est plus faible 200 jours après le 30 avril (figure 7.9, fin novembre) que 200 jours après le 2 janvier (figure 7.8, début juillet). Ceci s'explique par le fait que les pluies d'automne viennent alimenter les débits de rivière.

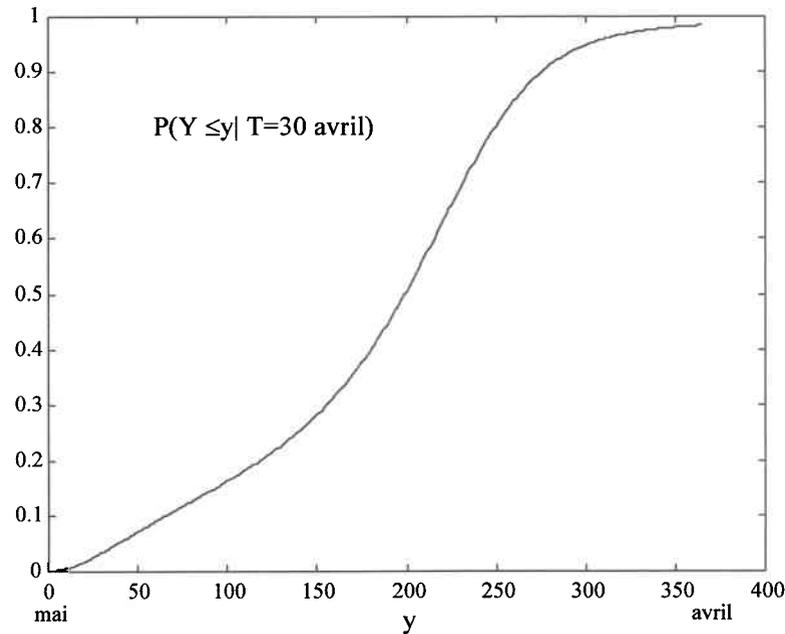


Figure 7.9: $I_3(y, 30 \text{ avril}) = P(Y \leq y | T=30 \text{ avril})$ en fonction de y - RNH (St15(25))

PNHC

Dans le cas du PNHC, $I_3(y,u)$ est fonction du nombre de jour sans précipitations, $\text{nbjoursec}(u+y)$ présenté sur la figure 7.11. $I_3(y,u)$ est:

$$P(Y \leq y | T=u) = \int_u^{u+y} \exp\left(\beta_1 + K_1 \text{nbjoursec}(a) + \theta_1 \sin\left(\frac{2\pi a}{365}\right)\right) da \quad (7.14)$$

Nous présentons sur la figure 7.10, $I_3(y, 7/12/43)$ pour St18(25). On remarque que la probabilité d'entrer en étiage augmente rapidement sur l'intervalle $y=(0,50)$ et devient stable autour de $y=110$, ce qui correspond au temps $t=u+y$, où la variable $\text{nbjoursec}(t)$ qui conditionne $\rho^1(u+y, H_u)$, fait un bond vers le haut.

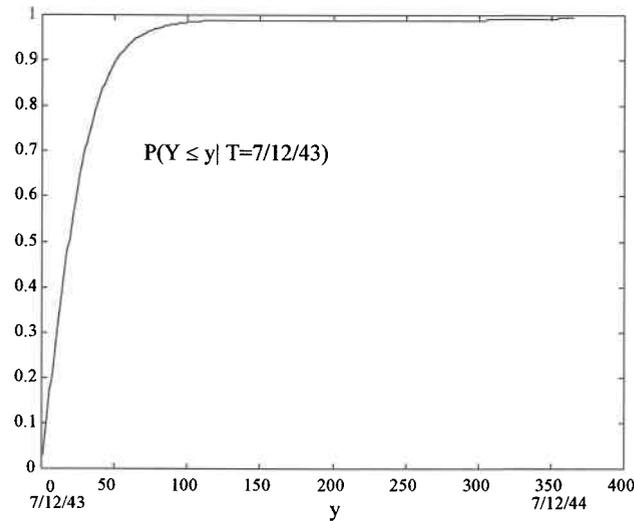


Figure 7.10: $I_3(y, 7/12/43)=P(Y \leq y | T=7/12/43)$ en fonction de y - PNHC (St18(25))

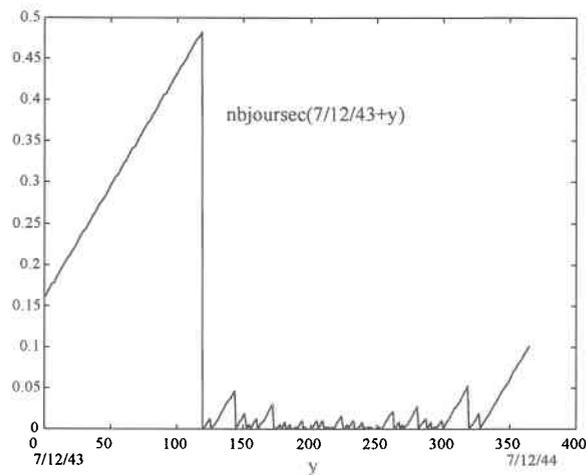


Figure 7.11: $nbjoursec(t), t=7/12/43+y$, St18

7.4 La probabilité d'un nouvel étiage, $I_4(z,u)$

Les trois premiers indices de risque ont été calculés à partir des taux de rupture des étiages et des surplus, $\rho^k(t, H_t)$ $k=1,2$, du processus alterné. Nous présentons maintenant des indices

de risque reliés aux processus d'un seul type. Nous référons le lecteur à la section 5.2 pour un rappel de ces concepts.

Le quatrième indice de risque proposé $I_4(z,u)$, correspond à la probabilité de rentrer à nouveau en étiage sur $(u, u+z)$ sachant qu'on est y entré à l'instant u : $P(Z < z | T = u)$. C'est la probabilité que Z , l'intervalle entre 2 débuts d'étiages, soit plus faible que z . Z est la somme de deux variables aléatoires: la durée de l'étiage X et la durée du surplus Y telle qu'illustrée sur la figure 7.12. Avant de calculer $I_4(z,u)$, il faut obtenir la densité de z , $d(z,u)$.

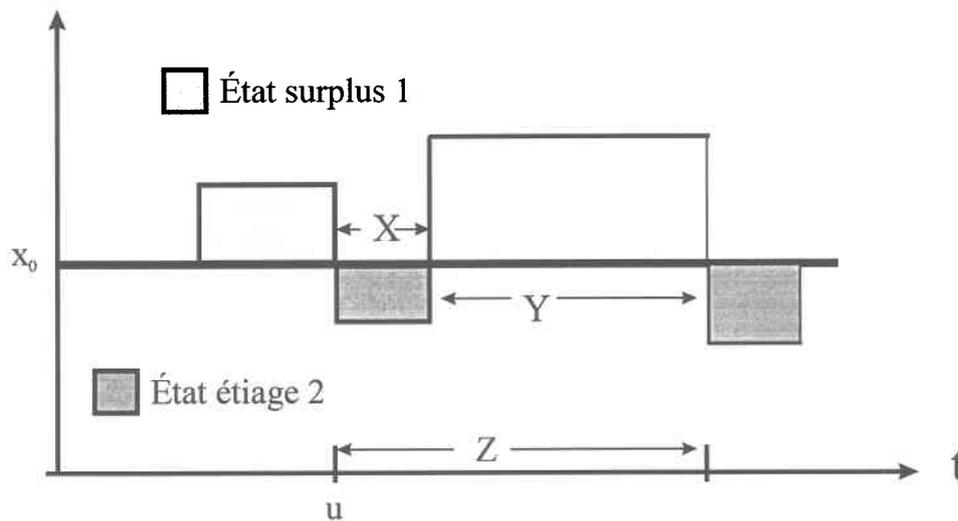


Figure 7.12: Z , l'intervalle entre 2 débuts d'étiage

7.4.1 La densité des intervalles entre les débuts d'étiages

Nous commençons par calculer la densité de z , $d(z)$. Supposons que les durées d'étiage x aient une densité $f(x)$, et les durées de surplus y aient une densité $g(y)$. Alors, la densité de $z=x+y$ est f^*g , la convolution de f et de g :

$$d(z) = f^*g(z) = \int_{-\infty}^{\infty} f(z-y)g(y)dy \quad (7.15)$$

Toutefois, dans notre cas x , y , et z sont non négatifs. On peut donc réécrire la convolution de la façon suivante:

$$d(z) = f^*g(z) = \int_0^z f(z-y)g(y)dy \quad (7.16)$$

Puisque les processus sont non homogènes, les densités des intervalles $d(z)$ doivent dépendre du dernier instant d'occurrence. Nous calculerons donc $d(z,u)$, la densité des intervalles entre deux étiages, étant donné que le processus est rentré en étiage à l'instant u .

En pratique, on calcule numériquement la convolution discrète, pour z entre 1 et 365 de la manière suivante:

$$d(z,u) = \sum_{k=0}^z f(z-k, u+k)g(k,u) \quad (7.17)$$

Les densités des durées d'étiage et des durées de surplus sont des fonctions des intensités comme suit (tableau 3.1):

$$f(x,u) = \rho^2(u+x) \exp\left(-\int_u^{u+x} \rho^2(a) da\right) \quad (7.18)$$

$$g(y,u) = \rho^1(u+y) \exp\left(-\int_u^{u+y} \rho^1(a) da\right) \quad (7.19)$$

PNH

Dans le cas du PNH, $f(x,u)$ la densité des durées des étiages, et $g(y,u)$ la densité des durées de surplus étant donné que l'étiage (surplus) a débuté au temps u sont:

$$f(x,u) = \exp(\beta_2 + K_2 \sin(\omega(u+x) + \theta_2)) \exp\left(-\int_u^{u+x} \exp(\beta_2 + K_2 \sin(\omega a + \theta_2)) da\right) \quad (7.20)$$

$$g(y,u) = \exp(\beta_1 + K_1 \sin(\omega(u+y) + \theta_1)) \exp\left(-\int_u^{u+y} \exp(\beta_1 + K_1 \sin(\omega a + \theta_1)) da\right) \quad (7.21)$$

RNH

Dans le cas du RNH:

$$f(x,u) = \varphi^2(x, u+x) \exp\left(-\int_u^{u+x} \varphi^2(a, u+a) da\right) \quad (7.22)$$

$$g(y,u) = \varphi^1(y, u+y) \exp\left(-\int_u^{u+y} \varphi^1(a, u+a) da\right) \quad (7.23)$$

$$\text{où } \varphi^i(x, u+x) = \frac{\exp(\beta_i + K_i \sin(\omega(u+x) + \theta_i))}{\int_0^\infty \exp^{-ps} \left(1 + \frac{s}{\int_u^{u+x} \lambda^i(a) da} \right)^{p-1} ds}, \lambda^i(u+x) = \exp(\beta_i + K_i \sin(\omega(u+x) + \theta_i))$$

$i=1,2.$

PNHC

Dans le cas du PNHC:

$$f(x,u) = \exp(K_2 \text{pnorm}(u+x) + \theta_2) \exp\left(-\int_u^{u+x} \exp(K_2 \text{pnorm}(a) + \theta_2) da\right) \quad (7.23)$$

$$g(y,u) = \quad (7.24)$$

$$\exp\left[\beta_1 + K_1 \text{nbjoursec}(u+y) + \theta_1 \sin\left(\frac{2\pi(u+y)}{365}\right)\right] \exp\left(-\int_u^{u+y} \exp\left[\beta_1 + K_1 \text{nbjoursec}(a) + \theta_1 \sin\left(\frac{2\pi a}{365}\right)\right] da\right)$$

Calcul de $I_4(z,u)$

$I_4(z,u)$, la probabilité de rentrer à nouveau en étiage sur $(u, u+z)$ sachant qu'on est y entré à l'instant u , correspond à:

$$I_4(z,u) = P(Z < z | T=u) = D(z | T=u) = \sum_{i=0}^z d(i, u) \quad (7.25)$$

Les figures 7.13, 7.14, et 7.15 présentent $I_4(z, 2 \text{ janvier})$, $I_4(z, 30 \text{ avril})$, et $I_4(z, 7/12/43)$ pour St07(25) (PNH), St15(25) (RNH), et St18(25) (PNHC) respectivement. On y voit que cet indice augmente plus rapidement pour St07(25) à partir de $x=250$. La probabilité de rentrer à nouveau en étiage 300 jours après le 2 janvier (fin octobre) est plus grande que 300 jours après le 30 avril (fin février) car dans le deuxième cas, on est probablement déjà en étiage. Pour ce qui est de PNHC, cet indice est grand autour de $z=100$ jours, ce qui correspond à un grand $\text{nbjoursec}(7/12/43+z)$ (figure 7.11).

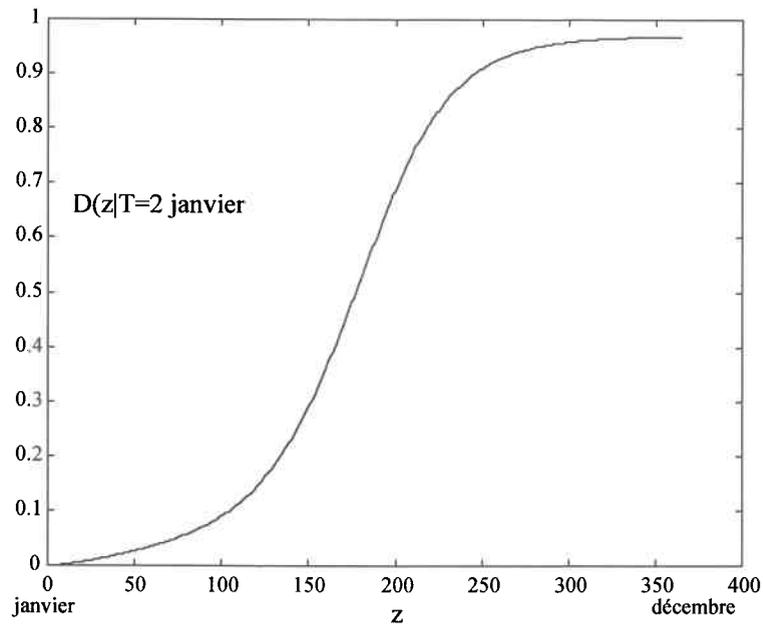


Figure 7.13: $I_4(z, 2 \text{ janvier})=D(z | T=2 \text{ janvier})$ en fonction de z - PNH (St07(25))

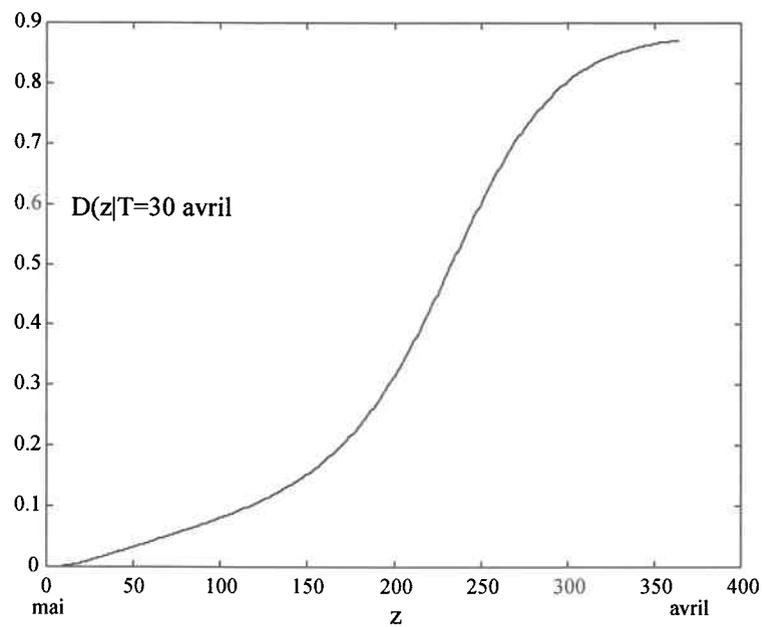


Figure 7.14: $I_4(z, 30 \text{ avril})=D(z | T=30 \text{ avril})$ en fonction de z - RNH (St15(25))

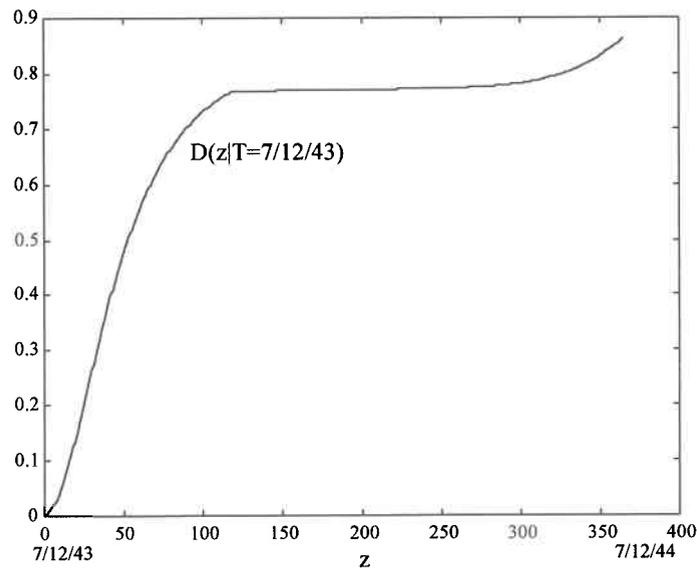


Figure 7.15: $I_4(z, 7/12/43)=D(z | T=7/12/43)$ en fonction de z - PNHC St18(25))

7.5 La probabilité d'être en étiage, $I_5(x,u)$

Le cinquième indice de risque que nous proposons $I_5(x, u)$ est la probabilité d'être en étiage au temps $u+x$ étant donné qu'on y est rentré au temps u , $P_2(u+x|T=u)$. Cet indice correspond à (North, 1981):

$$I_5(x, u) = P_2(u+x|T=u) = \frac{\rho_*^2(u+x|T=u)}{\rho^2(u+x|T=u)} \quad (7.26)$$

$\rho^2(t|T=u)$ est l'intensité de rupture des étiages à l'instant t dans le processus alterné, étant donné qu'on est entré en étiage à l'instant u . Cette intensité est conditionnelle à ce qu'on soit en étiage au temps t . $\rho_*^2(t|T=u)$ est l'intensité de rupture des étiages au temps t dans le processus d'un seul type, étant donné qu'on est entré en étiage à l'instant u . Cette quantité n'est pas conditionnelle à ce que l'on soit en étiage au temps t . Il est possible d'obtenir la probabilité d'être en surplus au temps $u+x$ étant donné qu'on est entré en étiage au temps u ,

$P_1(u+x|T=u)=1- P_2(u+x|T=u)$. Nous commençons par calculer $\rho_*^2(u+x|T=u)$. Cette quantité est (tableau 3.1):

$$\rho_*^2(u+x|T=u)=\frac{d(x,u)}{1-D(x,u)} \quad (7.27)$$

Calcul de $I_5(x,u)$

Les figures 7.16, 7.17 et 7.18 présentent $I_5(x,u)= P_2(u+x|T=u)$ pour St07(25) (PNH), St15(25) (RNH), et St18(25) (PNHC) respectivement. Dans le cas du PNH, la probabilité d'être en étiage est la plus forte au mois de janvier (étiage d'hiver qui se poursuit).

En examinant la figure 7.17 on peut être surpris par la probabilité relativement faible d'être en étiage tout de suite après le début de l'étiage. Or la courbe de probabilité présentée correspond à un étiage qui débute un 30 avril, date à laquelle il y a normalement beaucoup de précipitations pluvieuses, ce qui veut dire qu'un étiage débutant à cette date-là, ne dure pas longtemps. La probabilité d'être en étiage est la plus forte autour de 300 jours après le 30 avril, ce qui correspond à la fin février (étiage d'hiver).

Dans le cas du PNHC (figure 7.18), la probabilité d'être en étiage est la plus forte autour de 50 jours après le 7/12/43 ce qui s'explique par les faibles valeurs de $p_{norm}(t)$ autour de cette date, (figure 7.7). On voit aussi que la probabilité d'être en étiage devient très faible autour de 120 jours après le 7/12/43, date à laquelle la variable $nbjoursec$ (figure 7.11) devient elle aussi très faible.

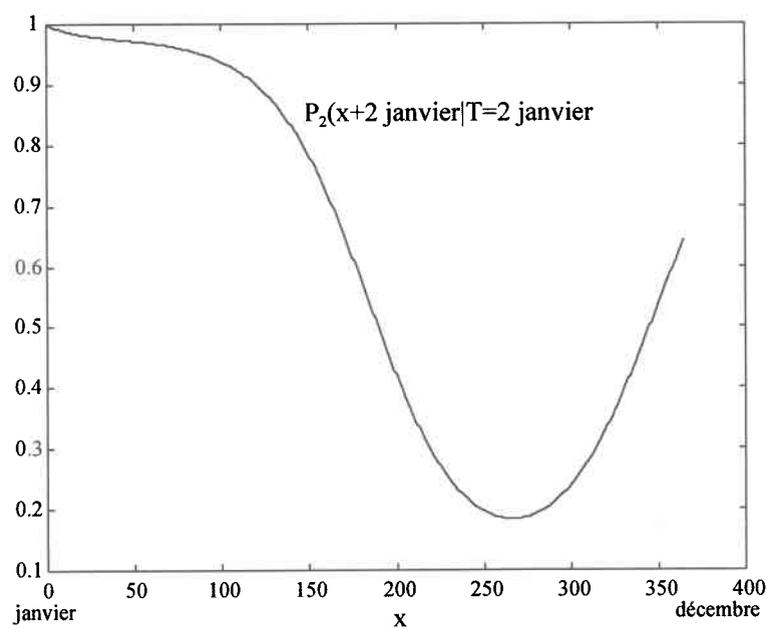


Figure 7.16: $I_5(x, 2 \text{ janvier}) = P_2(x+2 \text{ janvier} | T=2 \text{ janvier})$ en fonction de x - PNH (St07(25))

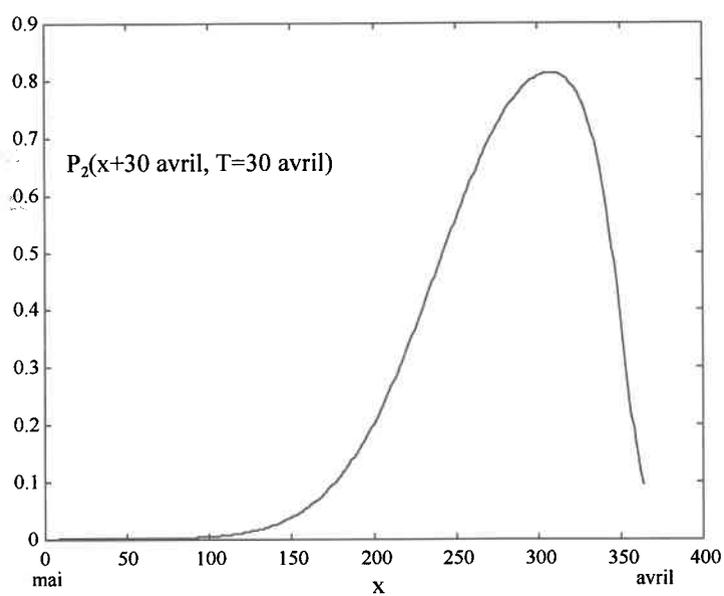


Figure 7.17: $I_5(x, 30 \text{ avril}) = P_2(x+30 \text{ avril}, T=30 \text{ avril})$ en fonction de x - RNH (St15(25))

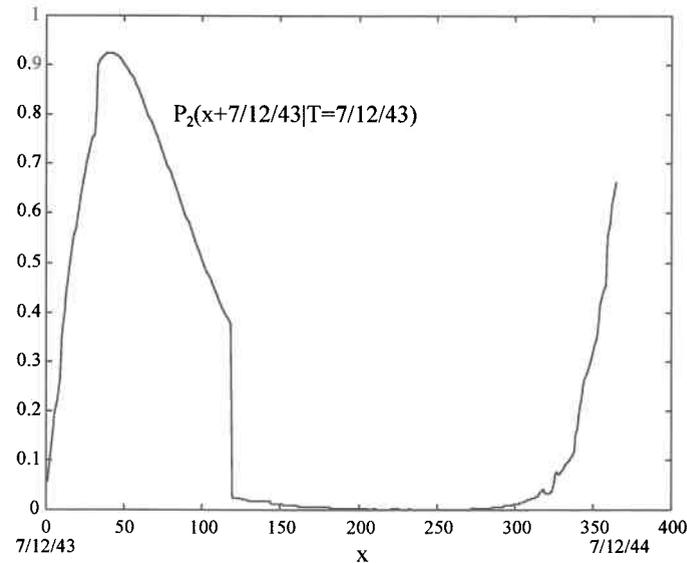


Figure 7.18: $I_5(x, 7/12/43) = P_2(x+7/12/43, T=7/12/43)$ en fonction de x - PNHC (St18(25))

7.6 Conclusions

Nous avons présenté dans ce chapitre cinq indices de risque reliés aux durées des étiages: $I_1(x,u)$, la probabilité qu'un étiage qui a débuté à l'instant u dure plus que x jours; $I_2(x,u)$, la probabilité qu'un étiage qui a débuté à l'instant u se termine étant donné qu'il a déjà duré x jours; $I_3(y,u)$, la probabilité d'entrer en étiage sur l'intervalle $(u, u+y)$ étant donné qu'on est entré en surplus à l'instant u ; $I_4(z,u)$, la probabilité d'entrer à nouveau en étiage sur $(u, u+z)$ sachant qu'on est entré en étiage à l'instant u ; et $I_5(x,u)$, la probabilité d'être en étiage au temps $t=u+x$ étant donné que le dernier étiage a débuté à l'instant u . En référence aux indices de performance présentés à la section 2.7.2, $I_1(x,u)$ correspond à la fonction de survie, $I_2(x,u)$ correspond à une mesure de résilience, $I_3(x,u)$, $I_4(x,u)$, et $I_5(x,u)$ se rapportent à la définition classique du risque hydrologique (probabilité d'une défaillance). En pratique, ces indices peuvent servir d'objectifs à minimiser ou à maximiser lors de l'élaboration de politiques de gestion des ressources en eau.

North (1981) a obtenu les expressions, non conditionnelles à l'instant d'occurrence du dernier événement, de la loi des durées dans un état, la loi conjointe de la durée et de

l'instant d'occurrence, ainsi que la loi de la durée sur un intervalle donné. Pour ce faire, il a estimé $h_i(u)$, la densité d'occurrence d'un événement dans le processus d'un type P^i , par

$$h_i(u) = \frac{\lambda_*^i(u)}{\int_0^u \lambda_*^i(a) da}, \text{ où } \lambda_*^i(t) \text{ est la fonction d'intensité du processus d'un seul type. Il a dû}$$

faire l'hypothèse que le processus d'un seul type est poissonien. Or cette hypothèse est forte, et ne s'est pas avérée acceptable pour les rivières que nous avons étudiées.

Nous avons obtenu les expressions analytiques des cinq indices de risque proposés pour les trois types de processus développés dans ce travail: PNH, RNH, et PNHC. Afin d'illustrer leur applicabilité, nous les avons calculés sur des données réelles. Bien que les occurrences des étiages et des surplus soient modélisées par le même type de processus pour les trois rivières étudiées, ces indices pourraient aussi bien être calculés pour des rivières où ces modèles d'occurrence sont différents.

Les résultats présentés dans ce chapitre sont à titre d'illustration afin de montrer comment les modèles d'occurrence d'étiage et de surplus que nous avons développés dans ce travail peuvent servir à l'ingénieur praticien. D'autres indices devraient être développés selon les besoins des applications. Dans le cas où il serait difficile de calculer ces nouveaux indices analytiquement, des séries synthétiques pourraient être générées par simulation à l'aide des modèles développés et les indices de risque calculés empiriquement. De plus, il serait intéressant d'affecter à ces indices des termes d'erreur reflétant l'incertitude reliée au choix du modèle d'occurrence des étiages.

8. CONCLUSION

8.1 Motivation

La connaissance des caractéristiques statistiques extrêmes des rivières, plus particulièrement celles des étiages, est essentielle pour la planification et la conception des politiques de gestion et de contrôle des ressources hydriques. À l'inverse d'une crue, qui est un phénomène spectaculaire de durée souvent négligeable et dont l'impact est dû surtout à son intensité, un étiage constitue un phénomène qui perdure et dont l'impact est dû surtout à la durée des faibles débits. Il s'avère alors impératif de développer des modèles d'étiages dont la variable principale est la durée du déficit.

La majorité des études statistiques concernant les sécheresses hydrologiques et les étiages ont porté sur la distribution des valeurs minimales annuelles des débits moyens journaliers, ou sur le minimum annuel du débit moyen sur m jours, Q_m . De plus, la plupart des approches de modélisation des étiages telles que l'analyse des débits ne tiennent pas compte des durées et émettent des hypothèses simplificatrices de stationnarité. Ce travail a été motivé par le besoin d'obtenir des modèles d'occurrence des étiages plus réalistes qui tiennent compte de l'aspect non stationnaire du phénomène.

Par ce travail, nous avons cherché à faire le lien entre l'hydrologie statistique et la gestion des ressources hydriques, en calculant des indices de risque basés sur les durées des étiages. Ceci est motivé par le fait que la définition classique du risque hydrologique, la probabilité de dépassement d'un seuil critique n'est pas suffisante pour bien juger et comparer des politiques de gestion dans un cadre d'analyse de risque.

8.2 Objectifs

L'objectif principal de cette thèse était de construire des modèles pour les occurrences des surplus et des étiages en tenant compte de l'aspect non stationnaire du phénomène. L'approche privilégiée était celle des processus ponctuels. Le deuxième objectif était de construire à partir de ces modèles des indices de risque reliés aux durées des étiages. Ces

indices pourraient servir lors de l'évaluation de politiques de gestion des ressources hydriques. Nos objectifs ont été atteints tels que résumés dans les sections 8.3 et 8.4.

8.3 Démarche

Pour atteindre nos objectifs, nous avons effectué au chapitre 2 une revue bibliographique exhaustive des approches reliées à la modélisation des étiages. Nous avons constaté que peu de travaux ont été effectués pour les séries non stationnaires, et que ces travaux émettaient des hypothèses trop simplificatrices. Par exemple, afin de tenir compte des saisonnalités, on procède souvent par découpage de l'année, ce qui contient beaucoup d'arbitraire et augmente le nombre des paramètres. De plus, lorsque les processus ponctuels sont utilisés pour la modélisation des étiages à partir de données journalières, un événement est défini de façon à assurer l'indépendance des occurrences et permettre l'utilisation d'un processus simple de Poisson (Madsen et Rosbjerg (1995), Zelenhasic et Salvai (1987)). Or nous croyons que ce sont les modèles statistiques qui doivent se conformer à la réalité physique et non le contraire.

Au chapitre 3, nous avons rassemblé et résumé à partir de plusieurs documents, les éléments des processus ponctuels nécessaires à la compréhension de notre travail. Nous avons voulu ainsi rendre ces notions complexes accessibles à l'ingénieur praticien.

Le processus de Poisson non homogène étant utilisé souvent en hydrologie, nous avons développé au chapitre 4 une méthode d'estimation bayésienne semi-paramétrique basée sur les travaux de Bernier (1981) et Leonard (1973, 1978). Nous avons comparé cette méthode avec la méthode du maximum de vraisemblance qui présuppose une forme paramétrique de l'intensité. Nous avons aussi développé les variances des estimateurs en utilisant les résultats asymptotiques (Berger, 1985). Nous avons appliqué les deux méthodes d'estimation sur un processus simulé et sur un processus observé de précipitations sur la station française de Bar-Sur-Seine.

Au chapitre 5, nous avons développé deux modèles existants: le PNH et le RNH. Le PNH suppose l'indépendance des événements et une fonction d'intensité périodique qui dépend du temps. Le RNH a une mémoire limitée et tient compte de l'instant où se trouve le processus et du temps d'occurrence du dernier événement. De plus, nous avons proposé et évalué un

nouveau modèle: le PNHC. Ce dernier est un modèle semi-physique où les fonctions d'intensité sont conditionnées par des variables exogènes telles que la précipitations et les températures. Les étapes que nous avons suivies afin de développer les modèles proposés sont les suivantes:

- construire des séries d'étiage et de surplus à partir des débits journaliers et des seuils définis par les 10^{ème} et 25^{ème} quantiles, *i.e.* les débits dépassés 90 et 75% du temps respectivement;
- pour chacun des modèles PNH, RNH et PNHC, estimer les paramètres de la fonction d'intensité par la méthode du maximum de vraisemblance;
- homogénéiser la série observée à l'aide de la fonction d'intensité estimée;
- vérifier l'adéquation du modèle à l'aide des tests statistiques; et
- effectuer une analyse de sensibilité basée sur la simulation de processus à l'aide du modèle développé.

Nous avons évalué au chapitre 6 les trois modèles (PNH, RNH, PNHC) sur 10 rivières canadiennes et une rivière française, et ce à deux seuils: Q_{25} et Q_{10} . Les rivières étudiées ont des bassins variant de 300 à 75000 km², et des débits entre 1.6 et 442 m³/s. Notre approche permet de considérer un modèle différent pour les étiages et les surplus dans une même série alternée.

La deuxième partie de ce travail consiste à obtenir des indices de risque qui peuvent servir comme base d'évaluation et de comparaison de différentes politiques de gestion. Nous avons développé pour les PNH, RNH et PNHC les indices suivants: $I_1(x,u)$, la probabilité qu'un étiage qui a débuté à l'instant u dure plus de x jours; $I_2(x,u)$, la probabilité qu'un étiage qui a débuté à l'instant u se termine étant donné qu'il a déjà duré x jours; $I_3(y,u)$, la probabilité d'entrer en étiage sur l'intervalle $(u,u+y)$ étant donné qu'on est entré en surplus à l'instant u ; $I_4(z,u)$, la probabilité d'entrer à nouveau en étiage sur $(u, u+z)$ étant donné qu'on est entré en étiage à l'instant u ; et $I_5(x,u)$, la probabilité d'être en étiage au temps $t = u+x$ étant donné que le dernier étiage a débuté à l'instant u . Nous avons quantifié ces indices pour des données

réelles afin d'illustrer leur applicabilité. Les résultats présentés au chapitre 7 sont plutôt de nature exploratoire.

8.4 Résumé des résultats

Nous avons constaté l'adéquation du PNH comme modèle d'occurrence de surplus ou d'étiage pour 18 séries (rivières/seuils), et celle du RNH pour 12 séries (rivières/seuils) sur 44 séries considérées (22 étiages et 22 surplus). Ces résultats sont assez surprenants compte tenu de la simplicité relative de ces modèles.

Le PNHC, un modèle que nous avons proposé comme alternative au PNH et RNH, a été appliqué sur les rivières où ni le RNH ni le PNH n'était acceptable. Pour ce faire, nous avons dû essayer plusieurs types de fonction d'intensité présentées à l'annexe 5. Nous avons trouvé que le PNHC était adéquat pour les étiages de 3 des 4 séries (rivières/seuils) et pour les surplus de 4 des 8 séries (rivières/seuils) que nous avons tenté de modéliser.

Suite à nos travaux et à ceux de North (1981), il s'avère que des processus ponctuels assez simples tels que le PNH et le RNH peuvent être utilisés pour modéliser en pratique les occurrences des étiages et des surplus. Pour les cas où ces processus ne sont pas adéquats, on peut envisager le PNHC que nous avons proposé. Ce modèle est parcimonieux (2 paramètres pour les étiages et 3 paramètres pour les surplus). Il suppose toutefois la disponibilité de données météorologiques journalières. L'approche alternée est nécessaire si on veut tenir compte des durées des étiages, qui est une variable importante. Tous nos modèles peuvent être utilisés pour générer, par simulation, des séries de surplus et d'étiage.

Nous avons remarqué que l'adéquation des modèles d'étiage et de surplus varient en fonction du seuil pour 6 des 11 rivières. Ceci illustre la grande influence du seuil sur le type de modèle approprié, d'où l'impossibilité de conclure à un modèle "universel".

Quant à la méthode d'estimation bayésienne de l'intensité d'un PNH qui a été développée dans ce travail, son avantage principal par rapport à la méthode du maximum de vraisemblance, est qu'il n'est pas nécessaire de supposer, a priori, une forme complètement paramétrée de la fonction d'intensité. Par contre, cette approche, moins efficace, présente le désavantage d'avoir à découper le temps en intervalles, n'estimant ainsi que des valeurs

moyennes d'intensité. Son avantage par rapport à la méthode conventionnelle non paramétrique où $\hat{\lambda}_i = \frac{\exp(1/n_i)}{p}$ et $\text{var}(\hat{\lambda}_i) = \frac{\hat{\lambda}_i}{p}$ est sa robustesse aux données manquantes, *i.e.* lorsque le nombre d'événements est 0 pour un intervalle donné.

8.5 Contributions

Tels que résumés aux sections 8.3 et 8.4, nous avons atteint les objectifs que nous nous sommes fixés pour ce travail. Notre contribution à la modélisation des étiages se situe à différents niveaux:

- une revue de littérature exhaustive résumant les approches reliées à la modélisation des étiages;
- une synthèse des éléments des processus ponctuels nécessaires pour modéliser des séries d'événements, surtout au niveau des processus non homogènes et des processus auto excités;
- le développement d'une méthode d'estimation bayésienne semi-paramétrique pour l'estimation de la fonction d'intensité d'un PNH et la quantification des variances des estimateurs;
- le développement des variances des estimateurs du maximum de vraisemblance de la fonction d'intensité paramétrée pour un processus d'un type et pour une série alternée;
- le développement d'une approche méthodologique pour la modélisation des séries de surplus et d'étiage qui insiste sur la différence entre les processus ponctuels alternés de deux types (étiage et surplus), et les processus ponctuels d'un type (étiage ou surplus);
- l'application, l'évaluation, et la simulation des processus PNH et RNH pour 22 séries (rivières/seuils);

- la construction de modèles pour les surplus et les étiages basés sur les processus doublements stochastique (PNHC), dont les fonctions d'intensité dépendent de variables exogènes;
- l'application, l'évaluation, et la simulation du PNHC pour 12 séries (rivières/seuils); et
- le développement analytique et numérique d'indices de risque basés sur les durées pour les processus de type PNH, RNH, et PNHC. Ce travail est toutefois à titre exploratoire.

8.6 Travaux futurs

Plusieurs voies de recherche se présentent comme suite à ce travail. Premièrement, il serait intéressant de tester l'applicabilité des modèles d'occurrence de surplus et d'étiage basés sur le PNH, le RNH, et le PNHC sur plusieurs rivières ayant des caractéristiques différentes et ce, à plusieurs seuils. Ceci permettrait d'associer des types de modèles à des classes de rivières caractérisées par la superficie du bassin, la présence d'un étiage d'hiver, le débit moyen, le seuil, etc.

Une des difficultés que nous avons rencontrées est la définition d'un événement. Afin de tenir compte des erreurs de mesure et de conserver une certaine flexibilité, nous avons défini un changement d'état lorsque l'écart du débit au seuil était de plus de 10%. Il n'en demeure pas moins que cette définition d'événement est rigide. Il serait intéressant d'étudier des seuils flous qui permettent une transition graduelle entre les états et une appartenance partielle à un état. De plus, il est important de bien comprendre l'impact du choix du seuil sur le modèle retenu, car le seuil a une grande influence sur les résultats.

Nous avons développé le PNHC en supposant connues les variables exogènes: la température et les précipitations. Afin d'étendre l'applicabilité du PNHC, il serait intéressant de développer ces modèles à partir de modèles stochastiques des précipitations et des températures et non des séries observées. Cette approche nous donnerait une plus grande liberté prédictive, et le PNHC deviendrait en fait un processus doublement stochastique (PDS).

Pour ce qui est de la méthode d'estimation bayésienne que nous avons développée, l'application à un grand nombre de situations réelles permettrait de déterminer si un gain considérable est obtenu par rapport à d'autres méthodes non paramétriques. Pour ce qui a trait à l'évaluation par simulation de cette méthode, la répétition de l'expérience (réplicats) permettrait d'avoir une idée de l'erreur d'échantillonnage, et un critère de proximité entre la fonction d'intensité estimée et la courbe théorique serait utile pour étudier avec plus de rigueur l'impact du choix des valeurs initiales et du nombre d'intervalles. De plus, il existe un pendant continu à cette méthode, présenté par Leonard (1978), qu'il serait intéressant d'approfondir. L'approche continue fournit une estimation deux fois dérivable de la fonction d'intensité. Cette estimation est obtenue en minimisant le log de la fonctionnelle de vraisemblance postérieure à l'aide de techniques provenant du calcul de variation.

Les indices de risque obtenus l'ont été difficilement compte tenu des processus ponctuels rendus complexes par la non-stationnarité. North (1981) a obtenu des résultats reliés à l'analyse fréquentielle en supposant que le processus d'un type est poissonien. Cette hypothèse, forte d'ailleurs, ne s'est pas vérifiée dans notre cas, et nous avons dû avoir recours à des approximations numériques pour le calcul de convolutions. Il serait utile d'approfondir ces calculs et d'affecter à ces indices des termes d'erreur reflétant l'incertitude reliée au choix du modèle d'occurrence des étiages. D'un point de vue opérationnel, il faudra développer d'autres indices de risque selon les besoins du praticien. De plus, on pourrait examiner la possibilité d'utiliser ces indices pour définir des régions homogènes dans un contexte de régionalisation.

Finalement, nous nous sommes concentrés dans ce travail sur les durées des étiages sans aborder l'aspect des volumes de déficit. Il serait intéressant de combiner nos modèles avec des modèles de volumes. Ceci permettrait de quantifier la sévérité des étiages. De plus, un travail plus hydrologique basé sur les méthodes développées dans ce travail serait intéressant et utile surtout, pour les ingénieurs praticiens.

9. BIBLIOGRAPHIE

Abi-Zeid, I. et B. Bobée (1994), Some reliability measures for nonstationary Markov chains, dans: Engineering Risk and Natural Resources Management, L. Duckstein et E. Parent (eds.), NATO ASI Series, vol. 275.

Adamczyk, M. (1992), Nonparametric low flow frequency analysis, report for M.Sc. Civil Engineering, University of Ottawa.

Anderson, T. W., et D. A. Darling (1952), Asymptotic theory of certain goodness of fit criteria based on stochastic processes, Ann. Math. Stat., vol. 23, pp. 193-212.

Anderson, J.E., S.Y. Shiau, et D. Harvey (1991), Preliminary investigation of trend/patterns in surface water characteristics and climate variations, Proceedings of NHRI Symposium, no. 8, pp. 189-201.

Bellman R.E. and L.A. Zadeh (1970), Decision making in a fuzzy environment, Management Science, Vol. 17, No. 4, pp. B141- B164.

Beran, M.A. (1979), The effect of dependence on the assessment of hydrologic drought severity, Proceedings of the International Symposium on Hydrological Aspect of Drought, New Delhi, vol. 2, pp. 103-111.

Beran, M.A., et J.A. Rodier (1985), Hydrological aspects of drought, Rapport UNESCO-WMO.

Berger, J. O. (1985), Statistical Decision Theory and Bayesian Analysis, Springer- Verlag, 617 pages.

Bernier, J. (1964), La prévision statistique des bas débits, Association Internationale d'Hydrologie, no. 63.

Bernier, J. (1965), Sur les probabilités d'occurrence des sécheresses et des étiages, Bulletin du Centre de Recherche et d'essais de Chatou, no. 11, pp. 3-12.

Bernier, J. (1981a), Le modèle de renouvellement non stationnaire, Électricité de France, Département Laboratoire National d'Hydraulique, no. HE 40 81-11.

Bernier, J. (1981b), Éléments de Statistique Bayésienne, Électricité de France, Département Laboratoire National d'Hydraulique, no. HE 40 81-09.

Bidwell, V. J., (1972), Agricultural response to hydrologic drought, Colorado State University Hydrology Paper, no. 53,.

Bobée B., Ashkar F., Roy R. et L. Perreault (1991), Risk analysis of hydrological data using Gamma family and derived distributions, Water Resources Engineering Risk Assessment, ed. J. Ganoulis, Springer-Verlag, pp. 21-42.

- Boisvert, R (1969), Étude des débits d'étiage en hiver et du comportement de la courbe de tarissement, Mémoire de Maîtrise, Université Laval.
- Bonacci, O. (1993), Hydrological identification of drought, Hydrological Processes, vol. 7, pp. 249-262.
- Chander, S., N.S. Kambo, S.K. Spolia, et A. Kumar (1981), Analysis of surplus and deficit -using runs, Journal of Hydrology, vol. 49, pp. 193-208.
- Chang, T.J (1989), Characteristics of streamflow drought, New Directions for Surface Water Modeling, Proceedings of the Baltimore Symposium, IAHS Publ. no. 181, pp. 333-341.
- Chang, T.J., et J.R. Stenson (1990), Is it realistic to define a 100-year drought for water management?, Water Resources Bulletin, vol. 26, no. 5.
- Chow, V.T. (1979), Risk and reliability analysis applied to water resources in practice, Reliability in Water Resources Management, eds E.A. McBean, K.W. Hipel, and T.E. Unny, Water Resources Publications, Littleton, Colorado, pp. 243-252.
- Chow V.T., D. R. Maidment, et L. W. Mays (1988), Applied Hydrology, McGraw-Hill Publishing Company, Toronto.
- Clarke, R. T.(1994), Statistical Modeling in Hydrology, John Wiley & Sons, 412 pages.
- Çınlar, E. (1973), Superposition of point processes, Stochastic Point Processes: Statistical Analysis, Theory, and Applications, ed. P.A. Lewis , pp. 549:606.
- Colson G. (1985), Theories of risk and MCDM, Multiple Criteria Decision Methods and Applications, eds. G. Fandel and J. Spronk, Springer-Verlag, pp. 171-196.
- Condie, R., et G.A. Nix (1975), Modelling of low flow frequency distributions and parameter estimation, Proceedings of the International Water Resources Symposium, Water for Arid Lands, Iran.
- Correia, F.N., M.A. Santos, et R.R. Rodrigues (1991), Reliability in regional drought studies, Water Resources Engineering Risk Assessment, ed. J. Ganoulis, Springer-Verlag, pp. 43-62.
- Cox, D.R., et V. Isham (1980), Point Processes, Chapman and Hall, 188 pages.
- Cox, D. R., et P. A. W. Lewis (1966), Statistical Analysis of Series of Events, Methuen and Co. Ltd.
- Cramér, H. (1946), Mathematical Methods of Statistics, Princeton University Press.
- C.T.G.R.E.F. (1979), Étude fréquentielle des étiages, Informations Techniques, cahier no. 33.
- Dempster, A.P. (1968), A generalization of Bayesian inference, Journal of the Royal Statistical Society, ser. B. Vol. 30, pp. 207-247.

Diewald, W. (1989), Risk analysis in public works facilities planning, design and construction, Risk-Analysis and Management of Natural and Man-Made Hazards, eds. Y.Y. Haines and E. Z. Stakhiv, ASCE, New York, pp. 331- 340.

Dezman, L.E., B.A. Shafer, H.D. Simpson, et J.A. Danielson (1982), Development of a surface water supply index - A drought severity indicator for Colorado, Proceedings of the International Symposium on Hydrometeorology, AWRA, pp.337-341.

Downer, R.N., M.M. Siddiqi et V. Yevjevich (1967), Applications of runs to hydrologic droughts, Proceedings of the International Hydrology Symposium, Fort Collins, Colorado, pp. 496-505

Dracup, J.A., K.S. Lee, et E.G. Paulson Jr. (1980a), On the definitions of droughts, Water Resources Research, vol. 16, no. 2, pp. 297-302.

Dracup, J.A., K.S. Lee, et E.G. Paulson Jr. (1980b), On the statistical characteristics of drought events , Water Resources Research, vol. 16, no. 2, pp. 289-296.

Dracup, J.A. (1986), Causes and occurrence of drought, Drought Management and its Impact on Public Water Systems, Report on a Colloquium sponsored by the Water Science and Technology Board, pp. 24-48.

Dracup, J.A. (1991), Drought monitoring, Stochastic Hydrology and Hydraulics, vol. 5, pp. 261-266.

Dub et Dzuback (1960), La définition des débits d'étiage et l'illustration de la superficie de leur extension, Bulletin A.I.H.S., no. 51, pp. 151-156.

Duckstein, L. (1979), Imbedding uncertainties into multi-objective decision models in water resources, Reliability in Water Resources Management, eds E.A. McBean, K.W. Hipel, and T.E. Unny, Water Resources Publications, Littleton, Colorado, pp. 3-29.

Duckstein, L., B.P. Shrestha, et E.Z. Stakhiv (1991), Multicriterion risk and reliability in hydrologic system design, Water Resources Engineering Risk Assessment, ed. J. Ganoulis, pp. 363-392.

French, S. (1988), Decision Theory: an Introduction to the Mathematics of Rationality, Ellis Horwood Ltd. 448 pages.

Garen, D.C. (1993), Revised surface-water supply index for western United States, Journal of Water Resources Planning and Management, vol. 119, no. 4, pp. 437-454.

Graybill, F. A. (1983), Matrices with Applications in Statistics, Wadsworth International Group, 2^{ème} édition, 461 pages.

Grigelionis, B. (1963), Convergence of sums of random step processes, Theory of Probability and Applications, Vol. 8.

Guerrero-Salazar, P., et V. Yevjevich (1975), Analysis of drought characteristics by the theory of runs, Colorado State University Hydrology Paper, no. 80.

Gumbel, E.J. (1954), Statistical theory of droughts, Proceedings of the American Society of Civil Engineers, vol. 80, separate 439.

Güven, O. (1983), A simplified semiempirical approach to probabilities of extreme hydrologic droughts, Water Resources Research, vol. 19, no. 2, pp. 441-453.

Haimès, Y.Y. (1991), Total risk management, Risk Analysis, Vol. 11, No. 2, pp. 169-171.

Hashimoto, T. (1980), Robustness, reliability, resilience and vulnerability criteria for planning water resources systems, Thèse de Doctorat, Université Cornell, 115 pages.

Hashimoto, T., J. Stedinger, et D.P. Loucks (1982), Reliability, resilience and vulnerability criteria for water resources system performance evaluation, Water Resources Research, vol. 18, no. 1, pp. 14-26.

Hoang, V.D. (1978), Meilleur ajustement statistique aux débits journaliers de crue et d'étiage des rivières du Québec, Ministère des Richesses Naturelles, Direction générale des eaux, H.P.-47.

Jackson, B. B. (1975), Markov mixture models for drought lengths, Water Resources Research, vol. 11, no. 1.

Joseph, E.S. (1970), Probability distribution of annual droughts, Journal of the irrigation and drainage division, Proceedings of the ASCE, vol. 96, IR4, pp. 461-474.

Kaplan S., et B.J. Garrick (1981), On the quantitative definition of risk, Risk Analysis, Vol. 1, No. 1, pp. 11-27.

Kavvas, M.L. (1975), Point stochastic analysis of daily rainfall with applications to Indiana, Thèse de Doctorat, Université Purdue.

Kebaili Bargaoui, Z. (1990), Modélisation stochastique des sécheresses annuelles, Thèse d'État, Faculté des Sciences de Tunis.

Kendall, D.R., et J.A. Dracup (1990), The development of a hydrologic drought index and termination rate probabilities, Hydraulics/hydrology of arid lands, pp. 323-328.

Kendall, M. G., et A. Stuart (1961), The Advanced Theory of Statistics, vol. 2, Griffin.

Konecny F., et H. P. Nachtnebel (1985), Extreme value processes and the evaluation of risk in flood analysis, Applied Mathematical Modelling, vol. 9, pp. 11-15.

Law, A. M., et W. D. Kelton (1991), Simulation Modeling and Analysis, McGraw-Hill Inc., 2nd edition, 759 pages.

- Lee, K. S., J. Sadeghipour, et J.A. Dracup (1986), An approach for frequency analysis of multiyear drought durations, Water Resources Research, vol. 22, no. 5, pp. 655-662.
- Lehman, E. L. (1983), Theory of Point Estimation, John Wiley and Sons.
- Leonard, T. (1973), A Bayesian method for histograms, Biometrika, vol. 60, no. 2, pp. 297-308.
- Leonard, T. (1978), Density estimation, stochastic processes and their prior information, Journal of the Royal Statistical Society, Series B, vol. 40, no.2, pp. 113-146.
- Lewis, P. A. W. (1972), Recent results in the statistical analysis of univariate point processes, Stochastic Point Processes, ed. P. A. W. Lewis, Wiley and Sons.
- Llamas, J., et M.M. Siddiqui (1969), Runs of precipitation series, Colorado State University Hydrology Paper, no. 33.
- Loganathan, G.V., P. Mathejat, C.Y. Kuo, et M.H. Diskin (1986), Frequency analysis of low flows: Hypothetical distribution methods and a physically based approach, Nordic Hydrology, vol. 17, pp. 129-150.
- Madsen, H., et D. Rosbjerg (1995), On the modelling of extreme droughts, IAHS Publ., no. 231, pp. 377-385.
- Matalas, N.C. (1963), Probability distribution of low flows, U.S. Geological survey, professional report, no. 434 A.
- Matalas, N.C. (1991), Drought description, Stochastic Hydrology and Hydraulics, vol. 5, pp. 255-260.
- Mathier L., L. Perreault, B. Bobée et F. Ashkar (1992), The use of Geometric and Gamma-related distributions for frequency analysis of water deficit, Stochastic Hydrology and Hydraulics, vol. 6, pp. 239-254.
- Millan, J., et V. Yevjevich (1971), Probabilities of observed droughts, Colorado State University Hydrology Paper, no. 50.
- Millan, J. (1972), Drought impact on regional economy, Colorado State University Hydrology Paper, no. 55.
- Millan, J. (1973), Statistical properties of runs as applied to hydrologic droughts, Floods and Droughts, Proceedings of the 2nd international symposium on hydrology, pp. 627-636
- Miquel J., et B. Phien Bou Pha (1979), Évaluation des ressources en eau: défintion des étiages et estimation de leurs probabilités d'occurrence, Association Internationale de Recherches Hydrauliques, A.B.5, pp. 287-294.

Mood, A. M. (1940), The distribution theory of runs, Annals of Mathematical Statistics, vol. 11, pp. 367-392.

Moore, P.G. (1983), The Business of Risk, Cambridge University Press, 244 pages.

Moyé, L.A., A.S. Kapadia, I.M. Cech, et R.J. Hardy (1988), The theory of runs with applications to drought prediction, Journal of Hydrology, vol. 103, pp. 127-137.

Moyé, L. A., et A. S. Kapadia (1995), Predictions of drought length extreme order statistics using run theory, Journal of Hydrology, vol. 169, pp. 95-110.

Nachtnebel, H. P. et F. Konecny (1987), Risk analysis and time-dependent flood models, Journal of Hydrology, vol. 91, pp. 295-318.

Nathan, R.J., et T.A. McMahon (1990), Practical aspects of low-flow frequency analysis, Water Resources Research, vol. 26, no. 9, pp. 2135-2141.

North, M. (1981), Processus Hydrologiques Intermittents, Étude de modélisations stochastiques et applications, Thèse de Doctorat, no. 419, École Polytechnique Fédérale de Lausanne.

Ozga-Zielinska, M. (1989), Droughts and floods - their definitions and modeling, IAHS Publ., no. 181, pp. 313-322.

Palmer, W.C. (1965), Meteorological drought, U.S. Department of Commerce, Weather Bureau, Research Paper, no. 45, 58 pages.

Palmer, W.C. (1968), Keeping track of crop moisture conditions nationwide: The new crop moisture index, Weatherwise, pp. 156-161.

Parent, E., F. Lebdi, et P. Hurand (1991), Gestion stratégique d'un système de ressources en eau: l'exemple du système NESTE, Revue des Sciences de l'Eau, vol. 5, no. 4, pp. 544-567

Press, W. H., S. A. Teukolsky, W. T. Vetterling, et B. P. Flannery (1992), Numerical Recipes in C, The Art of Scientific Computing, Cambridge University Press, 2^{ème} édition, 994 pages.

Raynal-Villasenor, J.A., et J.C. Douriet-Cardenas (1987), The GEV distribution in drought frequency analysis, Engineering hydrology, ed. A.D. Feldman, pp. 532-537.

Riggs, H.C. (1972), Low-flow investigations, Techniques of Water Resources Investigations of the United States Geological Survey.

Riggs, H.C. (1979), Charaterizing streamflow droughts , Proceedings of the International Symposium on Hydrological Aspect of Drought, New Delhi, vol. 2, pp. 331-338.

- Ross, S. M. (1985), Introduction to Probability Models, Academic Press, 3^{ème} édition, 502 pages.
- Ross, S. M. (1990), A Course in Simulation, Macmillan Publishing Company, 202 pages.
- Rowe, W.D. (1988), An Anatomy of Risk, Robert E. Krieger Publishing Company, Florida, 488 pages.
- Russell, C. S., Arey D. G., et R. W. Kates (1970), Droughts and Water Supply, John Hopkins Press.
- Saldarriaga, J. et V. Yevjevich (1970), Application of run-lengths to hydrologic series, Colorado State University Hydrology Paper, no. 40.
- Sen, Z. (1977), Run-Sums of annual flow series, Journal of Hydrology, vol. 35, pp. 311-324.
- Sen, Z. (1980a), The numerical calculation of extreme wet and dry Periods in hydrological time series, Hydrological Sciences Bulletin, vol. 25, no. 2, pp. 135-142.
- Sen, Z. (1980b), Critical drought analysis of periodic-stochastic processes, Journal of Hydrology, vol. 46, pp. 251-263.
- Sen, Z. (1980c), Statistical analysis of hydrologic critical droughts, Proceedings of the ASCE, Journal of the Hydraulics Division, vol. 106, no. HY1, pp. 99-115.
- Sen, Z. (1990), Critical drought analysis by second-order Markov chain, Journal of Hydrology, vol. 120, pp. 183-202.
- Snyder, L. (1975), Random Point Processes, John Wiley & Sons, 485 pages.
- Szladow A. and W. Ziarko (1993), Rough Sets: Working with imperfect data, AI Expert, Juillet, pp. 36-41.
- Tasker, G.D. (1987) A comparison of methods for estimating low flow characteristics of streams, Water Resources Publications, vol. 23, no. 6, pp. 1077-1083.
- Todorovic, P. (1970), On some problems involving random number of random variables, The Annals of Mathematical Statistics, vol. 41, no. 3, pp. 1059-1063.
- Tumeo, M.A. (1993), The meaning of stochasticity, randomness and uncertainty in environmental modeling, Proceedings of Stochastic and Statistical Methods in Hydrology and Environmental Engineering, Waterloo, Ontario, June 21-23.
- Ulmo, J. et J. Bernier (1973), Éléments de Décision Statistique, PUF, 1^{ère} édition, 330 pages.
- Vernos, D. (1982), Application des modèles stochastiques dynamiques linéaires à la description et à la simulation des processus des débits des rivières aux échelles saisonnières et journalières, Thèse de Doctorat, Université Pierre et Marie Curie, Paris 6, 119 pages.

Wilhite, D.A. (1982), Measuring drought severity and assessing impact, Proceedings of the International Symposium on Hydrometeorology, AWRA, pp. 333-335.

Wilhite, D.A. et M.H. Glantz (1987), Understanding the drought phenomenon: The role of definitions, chapitre 2, Planning for Drought - Toward a Reduction of Societal Vulnerability.

Yen, B.C. (1989), Engineering approaches to risk and reliability analysis, Risk-Analysis and Management of Natural and Man-Made Hazards, eds Y.Y. Haimes and E. Z. Stakhiv, ASCE, New York, pp. 22- 49.

Yevjevich, V. (1967), An objective approach to definitions and investigations of continental hydrologic droughts, Colorado State University Hydrology Paper, no. 23.

Yevjevich, V., L. Da Cunha et E. Vlachos (1983) (eds.), Coping with Droughts, Water Resources Publications.

Zelenhasic, E., et A. Salvai (1987), A method of streamflow drought analysis, Water Resources Research, vol. 23, no. 1, pp. 156-168.

10. ANNEXE 1: ESTIMATION BAYÉSIENNE DE L'INTENSITÉ D'UN PNH

Cette annexe est reliée à l'estimation de la fonction d'intensité $\lambda(t)$ d'un processus de Poisson non homogène (PNH). Elle contient trois parties: la première présente la matrice de variance-covariance des intensités moyennes qui rentre dans l'estimation bayésienne. La deuxième partie calcule les estimateurs du MV. La troisième partie développe les variances de ces estimateurs. Ces calculs se rapportent au chapitre 4. Afin d'éviter de trop alourdir le texte, nous avons numéroté uniquement les équations importantes.

10.1 La matrice de variance-covariance V

La matrice de variance-covariance V des intensités moyennes discrètes φ_i ainsi que son déterminant, sont nécessaires pour l'estimation bayésienne de la fonction d'intensité du PNH du chapitre 4. La relation circulaire entre les intensités $\varphi_i = a(\varphi_{i-1} + \varphi_{i+1}) + \varepsilon_i$ implique que $\text{Var}(\varphi_i) = V_0$, et $\text{Cov}(\varphi_i, \varphi_{i+j \bmod K}) = V_j \forall i, i = 1..K$. La solution de ce système telle que les critères de variances égales et de covariances du même ordre égales sont respectés s'écrit:

$$\Phi = M * E \quad (10.1)$$

où :

$$M = \begin{bmatrix} k_1 & k_2 & k_3 & \cdot & \cdot & \cdot & \cdot & k_{K-1} & k_K \\ k_K & k_1 & k_2 & \cdot & \cdot & \cdot & \cdot & k_{K-2} & k_{K-1} \\ k_{K-1} & k_K & k_1 & k_2 & \cdot & \cdot & \cdot & \cdot & k_{K-2} \\ k_{K-2} & \cdot & k_{K-3} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ k_2 & k_3 & \cdot & \cdot & \cdot & \cdot & \cdot & k_K & k_1 \end{bmatrix}, E = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_K].$$

Il s'agit d'obtenir M. Or, le système $\varphi_i = a(\varphi_{i-1} + \varphi_{i+1}) + \varepsilon_i$ s'écrit aussi de la manière suivante:

$$\Phi = a * B * E \tag{10.2}$$

où

$$B = \begin{bmatrix} k_K + k_2 + \frac{1}{a} & k_1 + k_3 & k_2 + k_4 & k_{i-1} + k_{i+1} & \cdot & k_{K-1} + k_1 \\ k_{K-1} + k_1 & k_K + k_2 + \frac{1}{a} & k_1 + k_3 & \cdot & k_{i-1} + k_{i+1} & k_{K-2} + k_K \\ k_{K-2} + k_K & k_{K-1} + k_1 & \cdot & k_1 + k_3 & \cdot & k_{i-1} + k_{i+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ k_1 + k_3 & k_2 + k_4 & \cdot & \cdot & k_{K-1} + k_1 & k_K + k_2 + \frac{1}{a} \end{bmatrix}$$

B est une matrice symétrique et circulaire, chaque ligne est une permutation d'ordre 1 de la ligne précédente.

(10.1) et (10.2) ⇒

$$a * B = M \Rightarrow A * K = E_1 \tag{10.3}$$

$$\text{où } A = \begin{bmatrix} 1 & -a & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & -a \\ -a & 1 & -a & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & -a & 1 & -a & 0 & 0 & 0 & \cdot & \cdot & 0 \\ \cdot & 0 & -a & 1 & -a & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & 0 & -a & 1 & -a & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & -a & 1 & -a & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & -a & 1 & -a & \cdot & 0 \\ \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & -a & 1 & -a \\ -a & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & -a & 1 \end{bmatrix}$$

$$K = [k_1 \quad k_2 \quad \cdot \quad \cdot \quad k_K], \text{ et } E_1 = [1 \quad 0 \quad \cdot \quad \cdot \quad 0].$$

En utilisant l'équation $A * K = E_1$, il est clair que $M^{-1} = A$, car $M * A = I$. Nous obtenons ainsi, M la solution de 10.1 en fonction de A.

Nous voulons maintenant calculer V et son inverse C . L'indépendance des $\{\varepsilon_i\}$ implique que V , la matrice de variance-covariance de Φ , est $V = (M^*M^T) \cdot \sigma^2$. C est donc :

$$C = V^{-1} \Rightarrow C = h^*(M^{-1})^2 \Rightarrow C = h^*A^2 \text{ où} \quad (10.4)$$

$$A^2 = \begin{bmatrix} 2a^2 + 1 & -2a & a^2 & 0 & 0 & \cdot & \cdot & 0 & a^2 & -2a \\ -2a & 2a^2 + 1 & -2a & a^2 & 0 & \cdot & \cdot & \cdot & 0 & a^2 \\ a^2 & -2a & 2a^2 + 1 & -2a & a^2 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & a^2 & -2a & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & a^2 & \cdot \\ \cdot & \cdot & 0 & a^2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & -2a & a^2 \\ a^2 & 0 & \cdot & \cdot & \cdot & 0 & a^2 & -2a & 2a^2 + 1 & -2a \\ -2a & a^2 & 0 & \cdot & \cdot & \cdot & 0 & a^2 & -2a & 2a^2 + 1 \end{bmatrix}$$

Le déterminant de V

Nous avons que $\det(V) = \frac{1}{h \det(A)}$. Or, A étant une matrice régulière circulante, son

déterminant est $\det(A) = \prod_{i=1}^K \lambda_i$ (Graybill, 1983). λ_i sont les valeurs propres de A telles que

$$\lambda_i = \sum_{j=0}^{K-1} a_j \cdot \omega_i^j, \text{ où les } \omega_i \text{ sont les } K \text{ racines de } x^K=1.$$

a_0, a_1, \dots, a_{K-1} représentent les éléments de la première ligne de la matrice A . Nous avons alors, compte tenu de la structure de A :

$$\lambda_i = 1 - a\omega_i - a\omega_i^{K-1} \text{ et } \det(A) = \prod_{i=1}^K (1 - a\omega_i - a\omega_i^{K-1}). \quad (10.5)$$

En développant, on peut aussi écrire le déterminant de A sous la forme :

$$\det(A) = \left\{ -2a^k + \prod_{m=1}^{k-1} \left(1 + 2a \cos\left(\frac{m\pi}{K}\right) \right) - 2a^2 \prod_{m=1}^{k-2} \left(1 + 2a \cos\left(\frac{m\pi}{K-1}\right) \right) \right\} \quad (10.6)$$

Car:

$$\det(A) = -a \det(B_{K-1 \times K-1}) + a(-a \det B_{K-2 \times K-2} + (-1)^{K+1} a \det T_{K-2 \times K-2}) + (-1)^{K+1} (-a)(-a \det T_{K-2 \times K-2} + (-1)^{K+1} a \det B_{K-2 \times K-2})$$

$$\text{où } T_{K \times K} = \begin{bmatrix} -a & 1 & -a & 0 & \cdot & 0 \\ 0 & -a & 1 & -a & 0 & \cdot \\ \cdot & 0 & -a & 1 & -a & 0 \\ \cdot & \cdot & 0 & \cdot & 1 & -a \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & 0 & 0 & 0 & 0 & -a \end{bmatrix} \quad B_{K \times K} = \begin{bmatrix} 1 & -a & 0 & 0 & \cdot & 0 \\ -a & 1 & -a & 0 & 0 & 0 \\ 0 & -a & 1 & -a & \cdot & 0 \\ \cdot & \cdot & -a & 1 & -a & 0 \\ \cdot & \cdot & \cdot & -a & 1 & -a \\ 0 & 0 & 0 & 0 & -a & 1 \end{bmatrix}$$

et où $\det(T_{K \times K}) = (-a)^K$ et $\det(B_{K \times K}) = \prod_{m=1}^K \left(1 + 2a \cos\left(\frac{m\pi}{K+1}\right) \right)$ puisque le déterminant d'une

$$\text{matrice Bande } BB_{K \times K} = \begin{bmatrix} a_0 & a_1 & 0 & 0 & \cdot & 0 \\ a_2 & a_0 & a_1 & 0 & 0 & 0 \\ 0 & a_2 & a_0 & a_1 & \cdot & 0 \\ \cdot & \cdot & a_2 & a_0 & a_1 & 0 \\ \cdot & \cdot & \cdot & a_2 & a_0 & a_1 \\ 0 & 0 & 0 & 0 & a_2 & a_0 \end{bmatrix} \text{ est (Graybill, 1983):}$$

$$\det(BB_{K \times K}) = \prod_{m=1}^K \left(a_0 + 2\sqrt{a_1 a_2} \cos\left(\frac{m\pi}{K+1}\right) \right) \quad (10.7)$$

10.2 L'estimation par la méthode MV des paramètres λ , k , et θ

Considérons la fonction d'intensité $\lambda(t) = \lambda \exp(k \sin(\frac{2\pi t}{365} + \theta))$ d'un PNH défini par les temps d'occurrence $\vec{t} = (t_1, t_2, \dots, t_n)$ sur l'horizon $(0, T)$. Nous voulons estimer les

paramètres de cette fonction, λ , k , et θ par la méthode du MV. Cette méthode consiste à optimiser la fonction de vraisemblance suivante (Lewis, 1972):

$$L(t_j, \lambda, k, \theta) = \prod_{j=1}^n \lambda(t_j) \exp\left(-\int_0^T \lambda(u) du\right) \quad (10.8)$$

$$L(t_j, \lambda, k, \theta) = \prod_{j=1}^n \lambda \exp\left(k \sin\left(\frac{2\pi t_j}{365} + \theta\right)\right) \exp\left(-\int_0^T \lambda \exp\left(k \sin\left(\frac{2\pi u}{365} + \theta\right)\right) du\right)$$

Les valeurs qui maximisent $L(t_j, \lambda, k, \theta)$ sont les solutions du système des dérivées partielles du premier ordre. Ceci revient à maximiser $LL(t_j, \lambda, k, \theta) = \log(L(t_j, \lambda, k, \theta))$ où:

$$LL(t_j, \lambda, k, \theta) = n \ln(\lambda) + \sum_{j=1}^n k \sin\left(\frac{2\pi t_j}{365} + \theta\right) - \lambda \int_0^T \exp\left(k \sin\left(\frac{2\pi u}{365} + \theta\right)\right) du \quad (10.9)$$

Les dérivées partielles d'ordre 1 sont les suivantes:

$$\begin{aligned} \frac{\partial LL}{\partial \lambda} &= \frac{n}{\lambda} - \int_0^T \exp\left(k \sin\left(\frac{2\pi u}{365} + \theta\right)\right) du \\ \frac{\partial LL}{\partial \theta} &= \sum_{j=1}^n k \cos\left(\frac{2\pi t_j}{365} + \theta\right) - \lambda \frac{365}{2\pi} \left(\exp\left(k \sin\left(\frac{2\pi T}{365} + \theta\right)\right) - \exp(k \sin(\theta)) \right) \end{aligned} \quad (10.10)$$

$$\frac{\partial LL}{\partial k} = \sum_{j=1}^n \sin\left(\frac{2\pi t_j}{365} + \theta\right) - \lambda \int_0^T \sin\left(\frac{2\pi u}{365} + \theta\right) \exp\left(k \sin\left(\frac{2\pi u}{365} + \theta\right)\right) du$$

Posons:

$$B = \sum_{j=1}^n \sin\left(\frac{2\pi t_j}{365}\right), \quad A = \sum_{j=1}^n \cos\left(\frac{2\pi t_j}{365}\right)$$

Or, T un multiple de 365 et $\sin\left(\frac{2\pi T}{365} + \theta\right) = \sin(\theta)$ impliquent:

$$\frac{\partial LL}{\partial \theta} = k(A \cos(\theta) - B \sin(\theta)) \quad (10.11)$$

$$\frac{\partial LL}{\partial \lambda} = 0 \Rightarrow \lambda = \frac{n}{\int_0^T \exp\left(k \sin\left(\frac{2\pi u}{365} + \theta\right)\right) du}$$

En substituant la valeur de λ dans $\frac{\partial LL}{\partial k}$, nous obtenons:

$$\frac{\partial LL}{\partial k} = B \cos(\theta) + A \sin(\theta) - \frac{n}{\int_0^T \exp\left(k \sin\left(\frac{2\pi u}{365} + \theta\right)\right) du} \int_0^T \sin\left(\frac{2\pi u}{365} + \theta\right) \exp\left(k \sin\left(\frac{2\pi u}{365} + \theta\right)\right) du$$

Posons:

$$I_0(k) = \int_0^T \exp\left(k \sin\left(\frac{2\pi u}{365} + \theta\right)\right) du$$

$$I_1(k) = \frac{\partial I_0(k)}{\partial k} = \int_0^T \exp\left(k \sin\left(\frac{2\pi u}{365} + \theta\right)\right) \sin\left(\frac{2\pi u}{365} + \theta\right) du \quad (10.12)$$

alors, $\frac{\partial LL}{\partial k} = B \cos(\theta) + A \sin(\theta) - n \frac{I_1(k)}{I_0(k)}$.

$$\frac{\partial LL}{\partial \theta} = 0 \Rightarrow k(A \cos(\theta) - B \sin(\theta)) = 0 \Rightarrow \theta = \arctan\left(\frac{A}{B}\right)$$

$$\frac{\partial LL}{\partial k} = 0 \Rightarrow \frac{1}{n}(B \cos(\theta) + A \sin(\theta)) = \frac{I_1(k)}{I_0(k)}$$

$$\theta = \arctan\left(\frac{A}{B}\right) \Rightarrow B \cos(\theta) = \frac{B}{\sqrt{1 + \frac{A^2}{B^2}}} \text{ et } A \sin(\theta) = \frac{A^2}{B \sqrt{1 + \frac{A^2}{B^2}}}$$

$$\Rightarrow B \cos(\theta) + A \sin(\theta) = \sqrt{B^2 + A^2}$$

$$\Rightarrow \frac{1}{n} \sqrt{A^2 + B^2} = \frac{I_1(k)}{I_0(k)}$$

k est donc la solution de $\frac{1}{n} \sqrt{A^2 + B^2} = \frac{I_1(k)}{I_0(k)}$.

Les estimateurs du maximum de vraisemblance sont:

$$\hat{\lambda} = \frac{n}{\int_0^T \exp\left(\hat{k} \sin\left(\frac{2\pi u}{365} + \hat{\theta}\right)\right) du}$$

$$\hat{\theta} = \tan^{-1}\left(\frac{A}{B}\right) \quad (10.13)$$

\hat{k} est la solution de $\frac{1}{n} \sqrt{A^2 + B^2} = \frac{I_1(k)}{I_0(k)}$.

Les variances des paramètres estimés par la méthode MV pour PNH d'un seul type

Posons $\vec{p} = (p_1, p_2, p_3)$ où $p_1 = \lambda, p_2 = k, p_3 = \theta$. La loi asymptotique des estimateurs du MV, $\hat{\vec{p}} = (\hat{\lambda}, \hat{k}, \hat{\theta})$ est approximativement $N(\hat{\vec{p}}, \hat{I}^{-1})$, où les éléments (r,s) de $\hat{I}(\hat{\lambda}, \hat{k}, \hat{\theta})$

sont: $\hat{I}_{rs} = \left[-\frac{\partial^2}{\partial p_r \partial p_s} \log(L(t_j, \vec{p})) \right]_{\vec{p}=\hat{\vec{p}}}$ (Lehman, 1983).

$$\text{var}(\hat{\lambda}) = \hat{I}^{-1}_{11}$$

$$\text{var}(\hat{k}) = \hat{I}^{-1}_{22}$$

$$\text{var}(\hat{\theta}) = \hat{I}^{-1}_{33}$$

La matrice de variance-covariance de $\hat{\vec{p}} = (\hat{\lambda}, \hat{k}, \hat{\theta})$

Nous obtenons la matrice de Fisher $\hat{I}(\hat{\lambda}, \hat{k}, \hat{\theta})$ suivante:

$$\hat{I}_{11} = \frac{n}{\hat{\lambda}^2}$$

$$\hat{I}_{12} = \hat{I}_{21} = \int_0^{\Gamma} \sin\left(\frac{2\pi u}{365} + \hat{\theta}\right) \exp \hat{k} \sin\left(\frac{2\pi u}{365} + \hat{\theta}\right) du$$

$$\hat{I}_{13} = \hat{I}_{31} = \frac{365}{2\pi} \left(\exp\left(\hat{k} \sin\left(\frac{2\pi\Gamma}{365} + \hat{\theta}\right)\right) - \exp(\hat{k} \sin \hat{\theta}) \right) \quad (10.14)$$

$$\hat{I}_{22} = \hat{\lambda} \int_0^{\Gamma} \sin^2\left(\frac{2\pi u}{365} + \hat{\theta}\right) \exp\left(\hat{k} \sin\left(\frac{2\pi u}{365} + \hat{\theta}\right)\right) du$$

$$\hat{I}_{23} = \hat{I}_{32} = B \sin \hat{\theta} - A \cos \hat{\theta} + \hat{\lambda} \frac{365}{2\pi} \left[\sin\left(\frac{2\pi\Gamma}{365} + \hat{\theta}\right) \exp\left(\hat{k} \sin\left(\frac{2\pi\Gamma}{365} + \hat{\theta}\right)\right) - \sin \hat{\theta} \exp(\hat{k} \sin \hat{\theta}) \right]$$

$$\hat{I}_{33} = B \hat{k} \cos \hat{\theta} + A \hat{k} \sin \hat{\theta} + \hat{\lambda} \frac{365}{2\pi} \left(\hat{k} \cos\left(\frac{2\pi\Gamma}{365} + \hat{\theta}\right) \exp\left(\hat{k} \sin\left(\frac{2\pi\Gamma}{365} + \hat{\theta}\right)\right) - \cos \hat{\theta} \exp(\hat{k} \sin \hat{\theta}) \right)$$

$$\text{var}(\hat{\lambda}) = \hat{\Gamma}^{-1}_{11} \text{ est:} \quad (10.15)$$

$$\begin{aligned} & \left(4 \lambda \%3 \pi^2 k \cos(\theta) B + 4 \lambda \%3 \pi^2 k \sin(\theta) A \right. \\ & \quad + 730 \lambda^2 \%3 \pi k \cos\left(\frac{2}{365} \pi t + \theta\right) e^{k \%4} - 730 \lambda^2 \%3 \pi \cos(\theta) e^{k \sin(\theta)} \\ & \quad - 4 B^2 \sin(\theta)^2 \pi^2 + 8 B \sin(\theta) \pi^2 A \cos(\theta) - 1460 B \sin(\theta) \pi \lambda \%4 e^{k \%4} \\ & \quad + 1460 B \sin(\theta)^2 \pi \lambda e^{k \sin(\theta)} - 4 A^2 \cos(\theta)^2 \pi^2 + 1460 A \cos(\theta) \pi \lambda \%4 e^{k \%4} \\ & \quad - 1460 A \cos(\theta) \pi \lambda \sin(\theta) e^{k \sin(\theta)} - 133225 \lambda^2 \%4^2 \left(e^{k \%4} \right)^2 \\ & \quad \left. + 266450 \lambda^2 \%4 e^{k \%4} \sin(\theta) e^{k \sin(\theta)} - 133225 \lambda^2 \sin(\theta)^2 \left(e^{k \sin(\theta)} \right)^2 \right) \lambda^2 \\ & \quad / \left(266450 n \lambda^2 \%4 e^{k \%4} \sin(\theta) e^{k \sin(\theta)} \right. \\ & \quad - 266450 \%2 \lambda^3 e^{k \%4} \sin(\theta) e^{k \sin(\theta)} \\ & \quad - 730 \%2^2 \lambda^3 \pi k \cos\left(\frac{2}{365} \pi t + \theta\right) e^{k \%4} + 4 n \lambda \%3 \pi^2 k \sin(\theta) A \\ & \quad + 730 n \lambda^2 \%3 \pi k \cos\left(\frac{2}{365} \pi t + \theta\right) e^{k \%4} - 730 n \lambda^2 \%3 \pi \cos(\theta) e^{k \sin(\theta)} \\ & \quad + 4 n \lambda \%3 \pi^2 k \cos(\theta) B - 4 n B^2 \sin(\theta)^2 \pi^2 + 1460 n A \cos(\theta) \pi \lambda \%4 e^{k \%4} \\ & \quad - 133225 n \lambda^2 \%4^2 \left(e^{k \%4} \right)^2 + 8 n B \sin(\theta) \pi^2 A \cos(\theta) \\ & \quad - 1460 n B \sin(\theta) \pi \lambda \%4 e^{k \%4} + 1460 n B \sin(\theta)^2 \pi \lambda e^{k \sin(\theta)} \\ & \quad - 133225 \lambda^3 \%3 \left(e^{k \%4} \right)^2 - 266450 \%2 \lambda^3 e^{k \sin(\theta)} \%4 e^{k \%4} \\ & \quad - 4 n A^2 \cos(\theta)^2 \pi^2 - 1460 n A \cos(\theta) \pi \lambda \sin(\theta) e^{k \sin(\theta)} \\ & \quad + 1460 \%2 \lambda^2 e^{k \%4} B \sin(\theta) \pi - 4 \%2^2 \lambda^2 \pi^2 k \cos(\theta) B \\ & \quad + 730 \%2^2 \lambda^3 \pi \cos(\theta) e^{k \sin(\theta)} - 4 \%2^2 \lambda^2 \pi^2 k \sin(\theta) A \\ & \quad - 1460 \%2 \lambda^2 e^{k \%4} A \cos(\theta) \pi + 266450 \%2 \lambda^3 \%4 \left(e^{k \%4} \right)^2 \\ & \quad - 1460 \%2 \lambda^2 e^{k \sin(\theta)} B \sin(\theta) \pi - 133225 \lambda^3 \%3 \left(e^{k \sin(\theta)} \right)^2 \\ & \quad + 266450 \lambda^3 e^{k \%4} \%3 e^{k \sin(\theta)} + 1460 \%2 \lambda^2 e^{k \sin(\theta)} A \cos(\theta) \pi \\ & \quad \left. + 266450 \%2 \lambda^3 \sin(\theta) \left(e^{k \sin(\theta)} \right)^2 - 133225 n \lambda^2 \sin(\theta)^2 \left(e^{k \sin(\theta)} \right)^2 \right) \end{aligned}$$

$$\%1 := \sin\left(\frac{2}{365} \pi x + \theta\right)$$

$$\%2 := \int \%1 e^{k \%1} dx$$

$$\%3 := \int \%1^2 e^{k \%1} dx$$

$$\%4 := \sin\left(\frac{2}{365} \pi t + \theta\right)$$

$$\text{var}(\hat{k}) = \hat{I}^{-1}_{22} \text{ est:} \quad (10.16)$$

$$\begin{aligned} & \left(4 n \pi^2 k \cos(\theta) B + 4 n \pi^2 k \sin(\theta) A + 730 n \pi \lambda k \cos\left(\frac{2}{365} \pi t + \theta\right) e^{k \%4} \right. \\ & \quad - 730 n \pi \lambda \cos(\theta) e^{k \sin(\theta)} - 133225 \lambda^2 \left(e^{k \%4} \right)^2 \\ & \quad \left. + 266450 \lambda^2 e^{k \%4} e^{k \sin(\theta)} - 133225 \lambda^2 \left(e^{k \sin(\theta)} \right)^2 \right) / \left(\right. \\ & \quad 266450 n \lambda^2 \%4 e^{k \%4} \sin(\theta) e^{k \sin(\theta)} - 266450 \%2 \lambda^3 e^{k \%4} \sin(\theta) e^{k \sin(\theta)} \\ & \quad - 730 \%2^2 \lambda^3 \pi k \cos\left(\frac{2}{365} \pi t + \theta\right) e^{k \%4} + 4 n \lambda \%3 \pi^2 k \sin(\theta) A \\ & \quad + 730 n \lambda^2 \%3 \pi k \cos\left(\frac{2}{365} \pi t + \theta\right) e^{k \%4} - 730 n \lambda^2 \%3 \pi \cos(\theta) e^{k \sin(\theta)} \\ & \quad + 4 n \lambda \%3 \pi^2 k \cos(\theta) B - 4 n B^2 \sin(\theta)^2 \pi^2 + 1460 n A \cos(\theta) \pi \lambda \%4 e^{k \%4} \\ & \quad - 133225 n \lambda^2 \%4^2 \left(e^{k \%4} \right)^2 + 8 n B \sin(\theta) \pi^2 A \cos(\theta) \\ & \quad - 1460 n B \sin(\theta) \pi \lambda \%4 e^{k \%4} + 1460 n B \sin(\theta)^2 \pi \lambda e^{k \sin(\theta)} \\ & \quad - 133225 \lambda^3 \%3 \left(e^{k \%4} \right)^2 - 266450 \%2 \lambda^3 e^{k \sin(\theta)} \%4 e^{k \%4} \\ & \quad - 4 n A^2 \cos(\theta)^2 \pi^2 - 1460 n A \cos(\theta) \pi \lambda \sin(\theta) e^{k \sin(\theta)} \\ & \quad + 1460 \%2 \lambda^2 e^{k \%4} B \sin(\theta) \pi - 4 \%2^2 \lambda^2 \pi^2 k \cos(\theta) B \\ & \quad + 730 \%2^2 \lambda^3 \pi \cos(\theta) e^{k \sin(\theta)} - 4 \%2^2 \lambda^2 \pi^2 k \sin(\theta) A \\ & \quad - 1460 \%2 \lambda^2 e^{k \%4} A \cos(\theta) \pi + 266450 \%2 \lambda^3 \%4 \left(e^{k \%4} \right)^2 \\ & \quad - 1460 \%2 \lambda^2 e^{k \sin(\theta)} B \sin(\theta) \pi - 133225 \lambda^3 \%3 \left(e^{k \sin(\theta)} \right)^2 \\ & \quad \left. + 266450 \lambda^3 e^{k \%4} \%3 e^{k \sin(\theta)} + 1460 \%2 \lambda^2 e^{k \sin(\theta)} A \cos(\theta) \pi \right. \\ & \quad \left. + 266450 \%2 \lambda^3 \sin(\theta) \left(e^{k \sin(\theta)} \right)^2 - 133225 n \lambda^2 \sin(\theta)^2 \left(e^{k \sin(\theta)} \right)^2 \right) \end{aligned}$$

$$\%1 := \sin\left(\frac{2}{365} \pi x + \theta\right)$$

$$\%2 := \int \%1 e^{k \%1} dx$$

$$\%3 := \int \%1^2 e^{k \%1} dx$$

$$\%4 := \sin\left(\frac{2}{365} \pi t + \theta\right)$$

$$\text{var}(\hat{\theta}) = \hat{I}^{-1}_{33} \text{ est:} \quad (10.17)$$

$$\begin{aligned}
& 4 \left(n \%3 - \%2^2 \lambda \right) \lambda \pi^2 / \left(266450 n \lambda^2 \%4 e^{k \%4} \sin(\theta) e^{k \sin(\theta)} \right. \\
& \quad - 266450 \%2 \lambda^3 e^{k \%4} \sin(\theta) e^{k \sin(\theta)} \\
& \quad - 730 \%2^2 \lambda^3 \pi k \cos\left(\frac{2}{365} \pi t + \theta\right) e^{k \%4} + 4 n \lambda \%3 \pi^2 k \sin(\theta) A \\
& \quad + 730 n \lambda^2 \%3 \pi k \cos\left(\frac{2}{365} \pi t + \theta\right) e^{k \%4} - 730 n \lambda^2 \%3 \pi \cos(\theta) e^{k \sin(\theta)} \\
& \quad + 4 n \lambda \%3 \pi^2 k \cos(\theta) B - 4 n B^2 \sin(\theta)^2 \pi^2 + 1460 n A \cos(\theta) \pi \lambda \%4 e^{k \%4} \\
& \quad - 133225 n \lambda^2 \%4^2 \left(e^{k \%4} \right)^2 + 8 n B \sin(\theta) \pi^2 A \cos(\theta) \\
& \quad - 1460 n B \sin(\theta) \pi \lambda \%4 e^{k \%4} + 1460 n B \sin(\theta)^2 \pi \lambda e^{k \sin(\theta)} \\
& \quad - 133225 \lambda^3 \%3 \left(e^{k \%4} \right)^2 - 266450 \%2 \lambda^3 e^{k \sin(\theta)} \%4 e^{k \%4} \\
& \quad - 4 n A^2 \cos(\theta)^2 \pi^2 - 1460 n A \cos(\theta) \pi \lambda \sin(\theta) e^{k \sin(\theta)} \\
& \quad + 1460 \%2 \lambda^2 e^{k \%4} B \sin(\theta) \pi - 4 \%2^2 \lambda^2 \pi^2 k \cos(\theta) B \\
& \quad + 730 \%2^2 \lambda^3 \pi \cos(\theta) e^{k \sin(\theta)} - 4 \%2^2 \lambda^2 \pi^2 k \sin(\theta) A \\
& \quad - 1460 \%2 \lambda^2 e^{k \%4} A \cos(\theta) \pi + 266450 \%2 \lambda^3 \%4 \left(e^{k \%4} \right)^2 \\
& \quad - 1460 \%2 \lambda^2 e^{k \sin(\theta)} B \sin(\theta) \pi - 133225 \lambda^3 \%3 \left(e^{k \sin(\theta)} \right)^2 \\
& \quad + 266450 \lambda^3 e^{k \%4} \%3 e^{k \sin(\theta)} + 1460 \%2 \lambda^2 e^{k \sin(\theta)} A \cos(\theta) \pi \\
& \quad \left. + 266450 \%2 \lambda^3 \sin(\theta) \left(e^{k \sin(\theta)} \right)^2 - 133225 n \lambda^2 \sin(\theta)^2 \left(e^{k \sin(\theta)} \right)^2 \right)
\end{aligned}$$

$$\%1 := \sin\left(\frac{2}{365} \pi x + \theta\right)$$

$$\%2 := \int \%1 e^{k \%1} dx$$

$$\%3 := \int \%1^2 e^{k \%1} dx$$

$$\%4 := \sin\left(\frac{2}{365} \pi t + \theta\right)$$

La variance de $\hat{\lambda}(t) = \hat{\lambda} \exp(\hat{k} \sin(\frac{2\pi t}{365} + \hat{\theta}))$

La propriété d'invariance de l'estimateur du MV assure qu'une fonction des estimateurs MV est elle-même un estimateur du MV (Ulmo et Bernier, 1973). $\hat{\lambda}(t)$ peut donc être considérée comme un estimateur du MV. Afin de calculer l'incertitude sur la fonction $\hat{\lambda}(t) = \hat{\lambda} \exp(\hat{k} \sin(\frac{2\pi t}{365} + \hat{\theta}))$ à partir de la matrice d'information de Fisher $\hat{I}(\hat{\lambda}, \hat{k}, \hat{\theta})$, nous définissons la transformation suivante:

$$\xi_1 = \hat{\lambda} \exp(\hat{k} \sin(\frac{2\pi t}{365} + \hat{\theta})), \xi_2 = \hat{k}, \xi_3 = \hat{\theta}$$

Nous avons alors:

$$p_1 = \xi_1 \exp(-\xi_2 \sin(\frac{2\pi t}{365} + \xi_3)), p_2 = \xi_2, p_3 = \xi_3$$

La matrice d'information de $\vec{\xi} = (\xi_1, \xi_2, \xi_3)$, $I^*(\vec{\xi})$ est (Lehman, 1983):

$$I^*(\vec{\xi}) = JI(\vec{p})J' \quad (10.18)$$

où $J = \left\| \frac{\partial p_s}{\partial \xi_r} \right\|$, $J_{sr} = \frac{\partial p_s}{\partial \xi_r}$ et $I(\vec{p}) = \hat{I}(\hat{\lambda}, \hat{k}, \hat{\theta})$ est la matrice d'information de \vec{p} . Nous

obtenons:

$$\begin{aligned} \frac{\partial p_1}{\partial \xi_1} &= \exp\left(-\xi_2 \sin\left(\frac{2\pi t}{365} + \xi_3\right)\right) \\ \frac{\partial p_1}{\partial \xi_2} &= -\xi_1 \sin\left(\frac{2\pi t}{365} + \xi_3\right) \exp\left(-\xi_2 \sin\left(\frac{2\pi t}{365} + \xi_3\right)\right) \\ \frac{\partial p_1}{\partial \xi_3} &= -\xi_1 \xi_2 \cos\left(\frac{2\pi t}{365} + \xi_3\right) \exp\left(-\xi_2 \sin\left(\frac{2\pi t}{365} + \xi_3\right)\right) \quad \frac{\partial p_2}{\partial \xi_1} = 0, \frac{\partial p_2}{\partial \xi_2} = 1, \frac{\partial p_2}{\partial \xi_3} = 0 \\ \frac{\partial p_3}{\partial \xi_1} &= 0, \frac{\partial p_3}{\partial \xi_2} = 0, \frac{\partial p_3}{\partial \xi_3} = 1 \end{aligned} \quad (10.19)$$

$$J = \begin{bmatrix} C & -\xi_1 \sin\left(\frac{2\pi t}{365} + \xi_3\right) \exp(-C) & -\xi_1 \xi_2 \cos\left(\frac{2\pi t}{365} + \xi_3\right) \exp(-C) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$C = \exp\left(-\xi_2 \sin\left(\frac{2\pi t}{365} + \xi_3\right)\right)$$

La matrice de variance-covariance de $\vec{\xi} = (\xi_1, \xi_2, \xi_3)$, est $\Gamma^{-1} * (\vec{\xi}) = J^{-1} \Gamma^{-1}(\bar{p}) J^{-1}$. La variance de $\hat{\lambda}(t) = \hat{\lambda} \exp(k \sin(\frac{2\pi t}{365} + \hat{\theta}))$ est $\text{var}(\hat{\lambda}(t)) = \Gamma^{-1}_{1,1} * (\vec{\xi})$.

$$\text{var}(\hat{\lambda}(t)) = \tag{10.20}$$

$$\left(\begin{aligned} &4 \lambda \%3 \pi^2 k \cos(\theta) B + 4 \lambda \%3 \pi^2 k \sin(\theta) A \\ &+ 730 \lambda^2 \%3 \pi k \cos\left(\frac{2}{365} \pi t + \theta\right) e^{k \%4} - 730 \lambda^2 \%3 \pi \cos(\theta) e^{k \sin(\theta)} \\ &- 4 B^2 \sin(\theta)^2 \pi^2 + 8 B \sin(\theta) \pi^2 A \cos(\theta) - 1460 B \sin(\theta) \pi \lambda \%4 e^{k \%4} \\ &+ 1460 B \sin(\theta)^2 \pi \lambda e^{k \sin(\theta)} - 4 A^2 \cos(\theta)^2 \pi^2 + 1460 A \cos(\theta) \pi \lambda \%4 e^{k \%4} \\ &- 1460 A \cos(\theta) \pi \lambda \sin(\theta) e^{k \sin(\theta)} - 133225 \lambda^2 \%4^2 \left(e^{k \%4} \right)^2 \\ &+ 266450 \lambda^2 \%4 e^{k \%4} \sin(\theta) e^{k \sin(\theta)} - 133225 \lambda^2 \sin(\theta)^2 \left(e^{k \sin(\theta)} \right)^2 \end{aligned} \right) \lambda^2$$

$$\sqrt{\left(\left(e^{-x2 \sin\left(\frac{2}{365} \pi t + x3\right)} \right)^2 \left(\begin{aligned} &266450 n \lambda^2 \%4 e^{k \%4} \sin(\theta) e^{k \sin(\theta)} \\ &- 266450 \%2 \lambda^3 e^{k \%4} \sin(\theta) e^{k \sin(\theta)} \\ &- 730 \%2^2 \lambda^3 \pi k \cos\left(\frac{2}{365} \pi t + \theta\right) e^{k \%4} + 4 n \lambda \%3 \pi^2 k \sin(\theta) A \\ &+ 730 n \lambda^2 \%3 \pi k \cos\left(\frac{2}{365} \pi t + \theta\right) e^{k \%4} - 730 n \lambda^2 \%3 \pi \cos(\theta) e^{k \sin(\theta)} \\ &+ 4 n \lambda \%3 \pi^2 k \cos(\theta) B - 4 n B^2 \sin(\theta)^2 \pi^2 + 1460 n A \cos(\theta) \pi \lambda \%4 e^{k \%4} \\ &- 133225 n \lambda^2 \%4^2 \left(e^{k \%4} \right)^2 + 8 n B \sin(\theta) \pi^2 A \cos(\theta) \\ &- 1460 n B \sin(\theta) \pi \lambda \%4 e^{k \%4} + 1460 n B \sin(\theta)^2 \pi \lambda e^{k \sin(\theta)} \\ &- 133225 \lambda^3 \%3 \left(e^{k \%4} \right)^2 - 266450 \%2 \lambda^3 e^{k \sin(\theta)} \%4 e^{k \%4} \\ &- 4 n A^2 \cos(\theta)^2 \pi^2 - 1460 n A \cos(\theta) \pi \lambda \sin(\theta) e^{k \sin(\theta)} \\ &+ 1460 \%2 \lambda^2 e^{k \%4} B \sin(\theta) \pi - 4 \%2^2 \lambda^2 \pi^2 k \cos(\theta) B \\ &+ 730 \%2^2 \lambda^3 \pi \cos(\theta) e^{k \sin(\theta)} - 4 \%2^2 \lambda^2 \pi^2 k \sin(\theta) A \\ &- 1460 \%2 \lambda^2 e^{k \%4} A \cos(\theta) \pi + 266450 \%2 \lambda^3 \%4 \left(e^{k \%4} \right)^2 \\ &- 1460 \%2 \lambda^2 e^{k \sin(\theta)} B \sin(\theta) \pi - 133225 \lambda^3 \%3 \left(e^{k \sin(\theta)} \right)^2 \\ &+ 266450 \lambda^3 e^{k \%4} \%3 e^{k \sin(\theta)} + 1460 \%2 \lambda^2 e^{k \sin(\theta)} A \cos(\theta) \pi \end{aligned} \right) \right)}$$

$$\%1 := \sin\left(\frac{2}{365} \pi x + \theta\right)$$

$$\%2 := \int \%1 e^{k \%1} dx$$

$$\%3 := \int \%1^2 e^{k \%1} dx$$

$$\%4 := \sin\left(\frac{2}{365} \pi t + \theta\right)$$

11. ANNEXE 2: LES LOIS CITÉES DANS LE TEXTE

Nous présentons dans cette annexe les lois mentionnées dans le texte.

La fonction Gamma (Chow *et al.*, 1988)

$$\Gamma(t) = \int_0^{\infty} u^{t-1} e^{-u} du$$

La loi exponentielle (Chow *et al.*, 1988)

$$f(x) = \lambda e^{-\lambda x}, \text{ de moyenne } 1/\lambda.$$

La loi Gamma à deux paramètres (Clarke, 1994)

$$f(x) = \frac{\alpha^\rho x^{\rho-1} e^{-\alpha x}}{\Gamma(\rho)}, \text{ de moyenne } \rho/\alpha, \text{ et variance } \rho/\alpha^2.$$

La loi GEV (Clarke, 1994)

$$\frac{1}{\alpha} \left\{ 1 - k \frac{(x-u)}{\alpha} \right\}^{\frac{1}{k}-1} \exp \left[- \left\{ 1 - k \frac{(x-u)}{\alpha} \right\}^{\frac{1}{k}} \right], \alpha, u, \text{ et } k \text{ sont les paramètres d'échelle, de}$$

position, et de forme respectivement.

La loi géométrique (Clarke, 1994)

$$f(x) = p(1-p)^x, \text{ de moyenne } \frac{1-p}{p} \text{ et variance } \frac{1-p}{p^2}.$$

La loi Gumbel (Clarke, 1994)

$$f(x) = \alpha \exp[-\alpha(x-u) - \exp(-\alpha(x-u))], \alpha \text{ et } u \text{ sont les paramètres d'échelle et de position respectivement.}$$

La loi lognormale 3 (Condie et Nix, 1975)

$$f(x) = \frac{1}{(x-a)\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2} [\ln(x-a) - m]^2\right\}$$
, m et σ sont la moyenne et l'écart type respectivement, et a est un paramètre de position.

La loi Log Pearson 3 (Chow *et al.*, 1988)

$$f(x) = \frac{1}{xa\Gamma(b)} \left(\frac{y-m}{a}\right)^{b-1} \exp\left(-\frac{y-m}{a}\right)$$
, $y = \log(x)$, m , b , et a sont les paramètres de position, de forme, et d'échelle respectivement.

La loi normale (Chow *et al.*, 1988)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$
, $N(\mu, \sigma^2)$.

La loi Pearson 3 (Condie et Nix, 1975)

$$f(x) = \frac{1}{a\Gamma(b)} \left(\frac{x-m}{a}\right)^{b-1} \exp\left(-\frac{x-m}{a}\right)$$
, m , b , et a sont les paramètres de position, de forme, et d'échelle respectivement.

La loi Pearson 5 (Condie et Nix, 1975)

$$f(x) = \frac{1}{a\Gamma(b)} \left(\frac{x-m}{a}\right)^{-(b+1)} \exp\left(-\frac{a}{x-m}\right)$$
, m , b , et a sont les paramètres de position, de forme, et d'échelle respectivement.

La loi de Poisson (Clarke, 1994)

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$
, de moyenne λ et variance λ .

La loi Weibull (Clarke, 1994)

$f(x) = \frac{k}{\alpha} \left(\frac{x-u}{\alpha}\right)^{k-1} \exp-\left(\frac{x-u}{\alpha}\right)^k$, α , u , et k sont les paramètres d'échelle, de position, et de forme respectivement.

Le coefficient d'asymétrie (Clarke, 1994)

Ce coefficient est $\frac{E(x-\mu)^3}{\sigma^3}$ pour une loi de moyenne μ et de variance σ^2 .

Le coefficient d'aplatissement (Clarke, 1994)

Ce coefficient est $\frac{E(x-\mu)^4}{\sigma^4}$ pour une loi de moyenne μ et de variance σ^2 .

12. ANNEXE 3: LES VARIANCES DES PARAMÈTRES DES PNH ET RNH

Nous calculons dans ce qui suit la variance des estimateurs $\hat{\beta}_i$, \hat{K}_i , et $\hat{\theta}_i$ des fonctions d'intensité des processus PNH et RNH qui font partie d'un processus alterné tels que définis aux sections 5.2.1 et 5.2.2. Afin d'éviter de trop alourdir le texte, nous avons numéroté uniquement les systèmes d'équations importants.

Posons $\bar{p}^i = (p^i_1, p^i_2, p^i_3)$ où $p^i_1 = \beta_i$, $p^i_2 = K_i$, $p^i_3 = \theta_i$, $i=1,2$. La loi asymptotique des estimateurs du MV, \hat{p}^i , est approximativement $N(\hat{p}^i, \hat{I}^{i-1})$, où les éléments (r,s) de

$\hat{I}^i(\hat{\beta}_i, \hat{K}_i, \hat{\theta}_i)$ sont $\hat{I}^i_{rs} = \left[-\frac{\partial^2}{\partial p^i_r \partial p^i_s} \ln L^i \right]_{\bar{p}^i = \hat{p}^i}$ (Berger, 1985). Nous avons donc:

$$\begin{aligned} \text{var}(\hat{\beta}_i) &= \hat{I}^{i-1}_{11} \\ \text{var}(\hat{K}_i) &= \hat{I}^{i-1}_{22} \\ \text{var}(\hat{\theta}_i) &= \hat{I}^{i-1}_{33} \end{aligned} \quad (12.1)$$

En calculant les dérivées partielles du second ordre, nous obtenons la matrice d'information I suivante, pour $i=1$:

$$\begin{aligned} \hat{I}^1_{11} &= -\frac{\partial^2 \ln L^1}{\partial \beta_1^2} = \\ &\exp(\beta_1) \left(\sum_{j=2}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} \exp\left(K_1 \sin\left(\frac{2\pi t}{365} + \theta_1\right)\right) dt + \int_{t_n}^{t_r} \exp\left(K_1 \sin\left(\frac{2\pi t}{365} + \theta_1\right)\right) dt \right) \end{aligned} \quad (12.2)$$

$$\hat{I}^1_{12} = \hat{I}^1_{21} = -\frac{\partial^2 \ln L^1}{\partial K_1 \partial \beta_1} = \exp(\beta_1) \left(\sum_{j=2}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} \sin\left(\frac{2\pi t}{365} + \theta_1\right) \exp\left(K_1 \sin\left(\frac{2\pi t}{365} + \theta_1\right)\right) dt + \int_{t_n}^{t_r} \sin\left(\frac{2\pi t}{365} + \theta_1\right) \exp\left(K_1 \sin\left(\frac{2\pi t}{365} + \theta_1\right)\right) dt \right)$$

$$\hat{I}_{13}^1 = \hat{I}_{31}^1 = -\frac{\partial^2 \ln L^1}{\partial \theta_1 \partial \beta_1} = \exp(\beta_1) \frac{365}{2\pi} \left[\sum_{j=2}^{n/2} \exp\left(K_1 \sin\left(\frac{2\pi t_{2j-1}}{365} + \theta_1\right)\right) - \exp\left(K_1 \sin\left(\frac{2\pi t_{2j-2}}{365} + \theta_1\right)\right) + \exp\left(K_1 \sin\left(\frac{2\pi t_f}{365} + \theta_1\right)\right) - \exp\left(K_1 \sin\left(\frac{2\pi t_n}{365} + \theta_1\right)\right) \right]$$

$$\hat{I}_{22}^1 = -\frac{\partial^2 \ln L^1}{\partial K_1^2} = \sum_{j=2}^{n/2} \exp(\beta_1) \int_{t_{2j-2}}^{t_{2j-1}} \sin\left(\frac{2\pi t}{365} + \theta_1\right)^2 \exp\left(K_1 \sin\left(\frac{2\pi t}{365} + \theta_1\right)\right) dt + \exp(\beta_1) \int_{t_n}^{t_f} \sin\left(\frac{2\pi t}{365} + \theta_1\right)^2 \exp\left(K_1 \sin\left(\frac{2\pi t}{365} + \theta_1\right)\right) dt$$

$$\hat{I}_{23}^1 = \hat{I}_{32}^1 = -\frac{\partial^2 \ln L^1}{\partial K_1 \partial \theta_1} = -\sum_{j=2}^{n/2} \cos\left(\frac{2\pi t_{2j-1}}{365} + \theta_1\right) + \frac{365 \exp(\beta_1)}{2\pi} \sum_{j=2}^{n/2} \sin\left(\frac{2\pi t_{2j-1}}{365} + \theta_1\right) \exp\left(K_1 \sin\left(\frac{2\pi t_{2j-1}}{365} + \theta_1\right)\right) - \sin\left(\frac{2\pi t_{2j-2}}{365} + \theta_1\right) \exp\left(K_1 \sin\left(\frac{2\pi t_{2j-2}}{365} + \theta_1\right)\right) + \frac{365 \exp(\beta_1)}{2\pi} \left(\sin\left(\frac{2\pi t_f}{365} + \theta_1\right) \exp\left(K_1 \sin\left(\frac{2\pi t_f}{365} + \theta_1\right)\right) - \sin\left(\frac{2\pi t_n}{365} + \theta_1\right) \exp\left(K_1 \sin\left(\frac{2\pi t_n}{365} + \theta_1\right)\right) \right)$$

$$\hat{I}_{33}^1 = -\frac{\partial^2 \ln L^1}{\partial \theta_1^2} = K_1 \sum_{j=2}^{n/2} \sin\left(\frac{2\pi t_{2j-1}}{365} + \theta_1\right) + \frac{365 \exp(\beta_1)}{2\pi} K_1 \sum_{j=2}^{n/2} \cos\left(\frac{2\pi t_{2j-1}}{365} + \theta_1\right) \exp\left(K_1 \sin\left(\frac{2\pi t_{2j-1}}{365} + \theta_1\right)\right) - \cos\left(\frac{2\pi t_{2j-2}}{365} + \theta_1\right) \exp\left(K_1 \sin\left(\frac{2\pi t_{2j-2}}{365} + \theta_1\right)\right) + \frac{365 \exp(\beta_1)}{2\pi} K_1 \left(\cos\left(\frac{2\pi t_f}{365} + \theta_1\right) \exp\left(K_1 \cos\left(\frac{2\pi t_f}{365} + \theta_1\right)\right) - \cos\left(\frac{2\pi t_n}{365} + \theta_1\right) \exp\left(K_1 \cos\left(\frac{2\pi t_n}{365} + \theta_1\right)\right) \right)$$

Les calculs qui suivent sont effectués dans MAPLE où $\beta = \hat{\beta}_1$, $k = \hat{k}^1$, $\theta = \hat{\theta}_1$, $\omega = \frac{2\pi}{365}$, $t(j) = t_j$, $t_f = t_{fin}$, $t_n = t_n$. À cause de l'incompatibilité entre MAPLE et le logiciel de traitement de texte, les polices de caractères ne sont pas exactement pareilles pour toutes les équations. Les variances des estimateurs sont:

$$\text{var}(\hat{\beta}_1) = \hat{I}^{-1}_{11} = \quad (12.3)$$

$$\begin{aligned}
& \left(k \omega^2 \beta^8 + k \omega \beta^7 + k \omega \beta e^{k\beta} - k \omega \beta e^{k\beta} \omega^2 + k \omega^2 \beta^8 + k \omega \beta^7 \right. \\
& + k \omega \beta e^{k\beta} - k \omega \beta e^{k\beta} \omega^2 + 2 \omega \beta^2 + 2 \omega \beta e^{k\beta} \\
& - 2 \omega \beta e^{k\beta} - (\beta^2)^2 - 2 (\beta^2)^2 e^{k\beta} + 2 (\beta^2)^2 e^{k\beta} \\
& - (\beta^2)^2 (e^{k\beta})^2 + 2 (\beta^2)^2 e^{k\beta} e^{k\beta} - (\beta^2)^2 (e^{k\beta})^2 \left. \right) / \left(\beta \left(\right. \right. \\
& - 2 \beta^2 k \omega^2 - 2 \beta^2 k \omega \beta - 2 \beta^2 k \omega \beta e^{k\beta} - 2 \beta^2 k \omega \beta^2 \omega \\
& + 2 \beta^2 (\beta^2)^2 (e^{k\beta})^2 - \beta^2 k \omega \beta^2 e^{k\beta} - \beta^2 k \omega \beta^2 \omega^2 - \beta^2 k \omega^2 \beta^2 \\
& - \beta^2 (\beta^2)^2 + 2 \beta^2 e^{k\beta} - 2 \beta^2 e^{k\beta} - 2 \beta^2 e^{k\beta} (\beta^2)^2 \\
& - \beta^2 (\beta^2)^2 + \beta^2 k \omega \beta^2 e^{k\beta} - \beta^2 k \omega \beta^2 \omega^2 + \beta^2 k \omega \beta^2 e^{k\beta} \\
& - \beta^2 k \omega \beta^2 e^{k\beta} - \beta^2 \beta^2 - \beta^2 (e^{k\beta})^2 + \beta^2 k \omega \beta^2 + \beta^2 k \omega \beta e^{k\beta} \\
& - \beta^2 (\beta^2)^2 + 2 \beta^2 e^{k\beta} e^{k\beta} + 2 e^{k\beta} e^{k\beta} (\beta^2)^2 + 2 \beta^2 e^{k\beta} (\beta^2)^2 \\
& - (e^{k\beta})^2 (\beta^2)^2 - \beta^2 (e^{k\beta})^2 + 2 \beta^2 \omega \beta e^{k\beta} - \beta^2 (\beta^2)^2 (e^{k\beta})^2 \\
& - \beta^2 (\beta^2)^2 (e^{k\beta})^2 + \beta^2 k \omega \beta^2 \omega^2 + \beta^2 k \omega \beta^2 e^{k\beta} + \beta^2 k \omega^2 \beta^2 \\
& + \beta^2 k \omega^2 \beta^2 \omega^2 - \beta^2 k \omega \beta^2 e^{k\beta} - \beta^2 k \omega \beta^2 e^{k\beta} + 2 \beta^2 \omega \beta^2 \\
& - \beta^2 k \omega \beta^2 e^{k\beta} + \beta^2 k \omega^2 \beta^2 \omega^2 - (e^{k\beta})^2 (\beta^2)^2 - \beta^2 (\beta^2)^2 (e^{k\beta})^2 \\
& - 2 \beta^2 \omega \beta e^{k\beta} + 2 \beta^2 (\beta^2)^2 e^{k\beta} e^{k\beta} + 2 \beta^2 (\beta^2)^2 e^{k\beta} e^{k\beta} \\
& - \beta^2 (\beta^2)^2 (e^{k\beta})^2 - 2 \beta^2 (\beta^2)^2 \omega^2 - 2 \beta^2 e^{k\beta} (\beta^2)^2 e^{k\beta} + 2 \beta^2 \beta^2 (\beta^2)^2 \\
& + 2 \beta^2 \beta^2 (\beta^2)^2 e^{k\beta} - 2 \beta^2 \beta^2 (\beta^2)^2 e^{k\beta} + 2 \beta^2 \beta^2 k \omega \beta^2 e^{k\beta} \\
& - 2 \beta^2 \beta^2 e^{k\beta} \omega - 2 \beta^2 \beta^2 e^{k\beta} (\beta^2)^2 e^{k\beta} + 2 \beta^2 (\beta^2)^2 (e^{k\beta})^2 \\
& + 2 \beta^2 \beta^2 (\beta^2)^2 - 2 \beta^2 \beta^2 e^{k\beta} (\beta^2)^2 + 2 \beta^2 \beta^2 e^{k\beta} \omega + 2 \beta^2 \beta^2 e^{k\beta} \omega \\
& - 2 \beta^2 \beta^2 \beta^2 \omega - \beta^2 k \omega^2 \beta^2 - 2 \beta^2 \beta^2 e^{k\beta} \omega - 2 \beta^2 \beta^2 (\beta^2)^2 e^{k\beta} \\
& + 2 \beta^2 \beta^2 e^{k\beta} (\beta^2)^2 + 2 \beta^2 (\beta^2)^2 (e^{k\beta})^2 + 2 \beta^2 \omega \beta e^{k\beta} \\
& + 2 \beta^2 (\beta^2)^2 \left. \right)
\end{aligned}$$

$$\begin{aligned}
\%1 &= \sum_{j=1}^{\frac{n}{2}} \cos(\omega t_{2j-1} + \theta) \\
\%2 &= \sin(\omega u + \theta) \\
\%3 &= \int_{t_n}^{t_{fin}} \exp(k \sin(\omega u + \theta)) du \\
\%4 &= \cos(\omega t_{fin} + \theta) \\
\%5 &= \sin(\omega t_{fin} + \theta) \\
\%6 &= \sum_{j=2}^{\frac{n}{2}} e^{\beta} \int_{t_{2j-2}}^{t_{2j-1}} \exp(k \sin(\omega u + \theta)) \sin(\omega u + \theta)^2 du \\
\%7 &= \sum_{j=2}^{\frac{n}{2}} \exp(k \sin(\omega t_{2j-1} + \theta)) \cos(\omega t_{2j-1} + \theta) - \exp(k \sin(\omega t_{2j-2} + \theta)) \cos(\omega t_{2j-2} + \theta) \\
\%8 &= \sum_{j=2}^{\frac{n}{2}} \sin(\omega t_{2j-1} + \theta) \\
\%9 &= \sin(\omega t_n + \theta) \\
\%10 &= \sum_{j=2}^{\frac{n}{2}} \exp(k \sin(\omega t_{2j-1} + \theta)) \sin(\omega t_{2j-1} + \theta) - \exp(k \sin(\omega t_{2j-2} + \theta)) \sin(\omega t_{2j-2} + \theta) \\
\%11 &= \sum_{j=2}^{\frac{n}{2}} \int_{t_{2j-2}}^{t_{2j-1}} \exp(k \sin(\omega u + \theta)) du \\
\%12 &= \int_{t_n}^{t_{fin}} \exp(k \sin(\omega u + \theta)) \sin(\omega u + \theta) du \\
\%13 &= \sum_{j=2}^{\frac{n}{2}} \int_{t_{2j-2}}^{t_{2j-1}} \exp(k \sin(\omega u + \theta)) \sin(\omega u + \theta) du \\
\%14 &= \sum_{j=2}^{\frac{n}{2}} \exp(k \sin(\omega t_{2j-1} + \theta)) - \exp(k \sin(\omega t_{2j-2} + \theta)) \\
\%15 &= \cos(\omega t_n + \theta) \\
\%16 &= \int_{t_n}^{t_{fin}} \exp(k \sin(\omega u + \theta)) \sin(\omega u + \theta)^2 du \\
\text{var}(\hat{k}^1) &= \hat{I}^{-1}_{22} =
\end{aligned}
\tag{12.4}$$

$$\begin{aligned}
& \left(k \omega^2 \%11 \%8 + k \omega \%11 \%7 + k \omega \%11 e^{k \%5 \%4} - k \omega \%11 e^{k \%9 \%15} + k \omega^2 \%3 \%8 + k \omega \%3 \%7 \right. \\
& + k \omega \%3 e^{k \%5 \%4} - k \omega \%3 e^{k \%9 \%15} - e^{\beta \%14^2} - 2 e^{\beta \%14} e^{k \%5} + 2 e^{\beta \%14} e^{k \%9} - e^{\beta} (e^{k \%5})^2 \\
& \left. + 2 e^{\beta} e^{k \%5} e^{k \%9} - e^{\beta} (e^{k \%9})^2 \right) / \left(-2 e^{\beta \%13} k \omega^2 \%12 \%8 - 2 e^{\beta \%13} k \omega \%12 \%7 \right. \\
& - 2 e^{\beta \%13} k \omega \%12 e^{k \%5 \%4} - 2 e^{\beta \%13} \%14 \%1 \omega + 2 \%13 (e^{\beta})^2 (e^{k \%9})^2 \%9 - e^{\beta} k \omega \%13^2 e^{k \%5 \%4} \\
& - e^{\beta} k \omega \%12^2 \%7 - e^{\beta} k \omega^2 \%13^2 \%8 - \%11 (e^{\beta})^2 \%10^2 + 2 e^{\beta \%14} e^{k \%9} \%6 - 2 e^{\beta \%14} e^{k \%5} \%6 \\
& - 2 \%14 e^{k \%5} (e^{\beta})^2 \%16 - \%14^2 (e^{\beta})^2 \%16 + e^{\beta} k \omega \%12^2 e^{k \%9} \%15 - e^{\beta} k \omega \%13^2 \%7 \\
& + e^{\beta} k \omega \%13^2 e^{k \%9} \%15 - e^{\beta} k \omega \%12^2 e^{k \%5} \%4 - e^{\beta \%14^2} \%6 - e^{\beta} (e^{k \%5})^2 \%6 + \%11 k \omega \%6 \%7 \\
& + \%11 k \omega \%6 e^{k \%5 \%4} - \%3 (e^{\beta})^2 \%10^2 + 2 e^{\beta} e^{k \%5} e^{k \%9} \%6 + 2 e^{k \%5} e^{k \%9} (e^{\beta})^2 \%16 \\
& + 2 \%14 e^{k \%9} (e^{\beta})^2 \%16 - (e^{k \%9})^2 (e^{\beta})^2 \%16 - e^{\beta} (e^{k \%9})^2 \%6 + 2 \%11 \%1 \omega e^{\beta} e^{k \%5} \%5 \\
& - \%11 (e^{\beta})^2 (e^{k \%5})^2 \%5^2 - \%3 (e^{\beta})^2 (e^{k \%5})^2 \%5^2 + \%11 k \omega e^{\beta \%16} \%7 + \%11 k \omega e^{\beta \%16} e^{k \%5 \%4} \\
& + \%11 k \omega^2 \%6 \%8 + \%3 k \omega^2 e^{\beta \%16} \%8 - \%11 k \omega e^{\beta \%16} e^{k \%9} \%15 - \%3 k \omega \%6 e^{k \%9} \%15 \\
& + 2 \%11 \%1 \omega e^{\beta \%10} - \%11 k \omega \%6 e^{k \%9} \%15 + \%11 k \omega^2 e^{\beta \%16} \%8 - (e^{k \%5})^2 (e^{\beta})^2 \%16 \\
& - \%11 (e^{\beta})^2 (e^{k \%9})^2 \%9^2 - 2 \%11 \%1 \omega e^{\beta} e^{k \%9} \%9 + 2 \%11 (e^{\beta})^2 e^{k \%5} \%5 e^{k \%9} \%9 \\
& + 2 \%3 (e^{\beta})^2 e^{k \%5} \%5 e^{k \%9} \%9 - \%3 (e^{\beta})^2 (e^{k \%9})^2 \%9^2 - 2 \%11 (e^{\beta})^2 \%10 e^{k \%5} \%5 \\
& + \%3 k \omega e^{\beta \%16} e^{k \%5} \%4 - \%3 k \omega e^{\beta \%16} e^{k \%9} \%15 + 2 \%3 \%1 \omega e^{\beta} e^{k \%5} \%5 + \%3 k \omega e^{\beta \%16} \%7 \\
& + 2 \%12 e^{k \%5} (e^{\beta})^2 \%10 + 2 \%12 (e^{\beta})^2 (e^{k \%5})^2 \%5 - 2 \%3 \%1 \omega e^{\beta} e^{k \%9} \%9 \\
& - 2 \%3 (e^{\beta})^2 \%10 e^{k \%5} \%5 - 2 \%12 e^{k \%9} (e^{\beta})^2 \%10 - 2 \%12 e^{k \%9} (e^{\beta})^2 e^{k \%5} \%5 \\
& + 2 \%12 \%14 (e^{\beta})^2 e^{k \%5} \%5 - 2 \%13 e^{k \%5} (e^{\beta})^2 e^{k \%9} \%9 - \%11 \%1^2 \omega^2 \\
& - 2 \%12 e^{k \%5} (e^{\beta})^2 e^{k \%9} \%9 + 2 \%13 \%14 (e^{\beta})^2 \%10 + 2 \%13 \%14 (e^{\beta})^2 e^{k \%5} \%5 \\
& - 2 \%13 \%14 (e^{\beta})^2 e^{k \%9} \%9 + 2 e^{\beta \%13} k \omega \%12 e^{k \%9} \%15 - 2 e^{\beta \%13} e^{k \%5} \%1 \omega \\
& - 2 \%13 e^{k \%9} (e^{\beta})^2 e^{k \%5} \%5 + 2 \%13 (e^{\beta})^2 (e^{k \%5})^2 \%5 + 2 \%12 \%14 (e^{\beta})^2 \%10 \\
& - 2 \%13 e^{k \%9} (e^{\beta})^2 \%10 + 2 e^{\beta \%12} e^{k \%9} \%1 \omega + 2 e^{\beta \%13} e^{k \%9} \%1 \omega - 2 e^{\beta \%12} \%14 \%1 \omega \\
& - e^{\beta} k \omega^2 \%12^2 \%8 - 2 e^{\beta \%12} e^{k \%5} \%1 \omega - 2 \%12 \%14 (e^{\beta})^2 e^{k \%9} \%9 + 2 \%13 e^{k \%5} (e^{\beta})^2 \%10 \\
& + 2 \%12 (e^{\beta})^2 (e^{k \%9})^2 \%9 + 2 \%3 (e^{\beta})^2 \%10 e^{k \%9} \%9 + 2 \%3 \%1 \omega e^{\beta \%10} \\
& \left. + 2 \%11 (e^{\beta})^2 \%10 e^{k \%9} \%9 + \%3 k \omega^2 \%6 \%8 + \%3 k \omega \%6 \%7 + \%3 k \omega \%6 e^{k \%5} \%4 - \%3 \%1^2 \omega^2 \right)
\end{aligned}$$

$$\text{var}(\hat{\theta}_1) = \hat{\mathbf{I}}^{-1}_{33} = \quad (12.5)$$

$$\begin{aligned}
& \left(\sigma_{11} \sigma_6 + e^{\beta} \sigma_{11} \sigma_{16} + \sigma_3 \sigma_6 + e^{\beta} \sigma_3 \sigma_{16} - e^{\beta} \sigma_{13}^2 - 2 e^{\beta} \sigma_{13} \sigma_{12} - e^{\beta} \sigma_{12}^2 \right) \omega^2 / \left(\right. \\
& - 2 e^{\beta} \sigma_{13} k \omega^2 \sigma_{12} \sigma_8 - 2 e^{\beta} \sigma_{13} k \omega \sigma_{12} \sigma_7 - 2 e^{\beta} \sigma_{13} k \omega \sigma_{12} e^{k \sigma_5} \sigma_4 - 2 e^{\beta} \sigma_{13} \sigma_{14} \sigma_1 \omega \\
& + 2 \sigma_{13} (e^{\beta})^2 (e^{k \sigma_9})^2 \sigma_9 - e^{\beta} k \omega \sigma_{13}^2 e^{k \sigma_5} \sigma_4 - e^{\beta} k \omega \sigma_{12}^2 \sigma_7 - e^{\beta} k \omega^2 \sigma_{13}^2 \sigma_8 \\
& - \sigma_{11} (e^{\beta})^2 \sigma_{10}^2 + 2 e^{\beta} \sigma_{14} e^{k \sigma_9} \sigma_6 - 2 e^{\beta} \sigma_{14} e^{k \sigma_5} \sigma_6 - 2 \sigma_{14} e^{k \sigma_5} (e^{\beta})^2 \sigma_{16} \\
& - \sigma_{14}^2 (e^{\beta})^2 \sigma_{16} + e^{\beta} k \omega \sigma_{12}^2 e^{k \sigma_9} \sigma_{15} - e^{\beta} k \omega \sigma_{13}^2 \sigma_7 + e^{\beta} k \omega \sigma_{13}^2 e^{k \sigma_9} \sigma_{15} \\
& - e^{\beta} k \omega \sigma_{12}^2 e^{k \sigma_5} \sigma_4 - e^{\beta} \sigma_{14}^2 \sigma_6 - e^{\beta} (e^{k \sigma_5})^2 \sigma_6 + \sigma_{11} k \omega \sigma_6 \sigma_7 + \sigma_{11} k \omega \sigma_6 e^{k \sigma_5} \sigma_4 \\
& - \sigma_3 (e^{\beta})^2 \sigma_{10}^2 + 2 e^{\beta} e^{k \sigma_5} e^{k \sigma_9} \sigma_6 + 2 e^{k \sigma_5} e^{k \sigma_9} (e^{\beta})^2 \sigma_{16} + 2 \sigma_{14} e^{k \sigma_9} (e^{\beta})^2 \sigma_{16} \\
& - (e^{k \sigma_9})^2 (e^{\beta})^2 \sigma_{16} - e^{\beta} (e^{k \sigma_9})^2 \sigma_6 + 2 \sigma_{11} \sigma_1 \omega e^{\beta} e^{k \sigma_5} \sigma_5 - \sigma_{11} (e^{\beta})^2 (e^{k \sigma_5})^2 \sigma_5^2 \\
& - \sigma_3 (e^{\beta})^2 (e^{k \sigma_5})^2 \sigma_5^2 + \sigma_{11} k \omega e^{\beta} \sigma_{16} \sigma_7 + \sigma_{11} k \omega e^{\beta} \sigma_{16} e^{k \sigma_5} \sigma_4 + \sigma_{11} k \omega^2 \sigma_6 \sigma_8 \\
& + \sigma_3 k \omega^2 e^{\beta} \sigma_{16} \sigma_8 - \sigma_{11} k \omega e^{\beta} \sigma_{16} e^{k \sigma_9} \sigma_{15} - \sigma_3 k \omega \sigma_6 e^{k \sigma_9} \sigma_{15} + 2 \sigma_{11} \sigma_1 \omega e^{\beta} \sigma_{10} \\
& - \sigma_{11} k \omega \sigma_6 e^{k \sigma_9} \sigma_{15} + \sigma_{11} k \omega^2 e^{\beta} \sigma_{16} \sigma_8 - (e^{k \sigma_5})^2 (e^{\beta})^2 \sigma_{16} - \sigma_{11} (e^{\beta})^2 (e^{k \sigma_9})^2 \sigma_9^2 \\
& - 2 \sigma_{11} \sigma_1 \omega e^{\beta} e^{k \sigma_9} \sigma_9 + 2 \sigma_{11} (e^{\beta})^2 e^{k \sigma_5} \sigma_5 e^{k \sigma_9} \sigma_9 + 2 \sigma_3 (e^{\beta})^2 e^{k \sigma_5} \sigma_5 e^{k \sigma_9} \sigma_9 \\
& - \sigma_3 (e^{\beta})^2 (e^{k \sigma_9})^2 \sigma_9^2 - 2 \sigma_{11} (e^{\beta})^2 \sigma_{10} e^{k \sigma_5} \sigma_5 + \sigma_3 k \omega e^{\beta} \sigma_{16} e^{k \sigma_5} \sigma_4 \\
& - \sigma_3 k \omega e^{\beta} \sigma_{16} e^{k \sigma_9} \sigma_{15} + 2 \sigma_3 \sigma_1 \omega e^{\beta} e^{k \sigma_5} \sigma_5 + \sigma_3 k \omega e^{\beta} \sigma_{16} \sigma_7 + 2 \sigma_{12} e^{k \sigma_5} (e^{\beta})^2 \sigma_{10} \\
& + 2 \sigma_{12} (e^{\beta})^2 (e^{k \sigma_5})^2 \sigma_5 - 2 \sigma_3 \sigma_1 \omega e^{\beta} e^{k \sigma_9} \sigma_9 - 2 \sigma_3 (e^{\beta})^2 \sigma_{10} e^{k \sigma_5} \sigma_5 \\
& - 2 \sigma_{12} e^{k \sigma_9} (e^{\beta})^2 \sigma_{10} - 2 \sigma_{12} e^{k \sigma_9} (e^{\beta})^2 e^{k \sigma_5} \sigma_5 + 2 \sigma_{12} \sigma_{14} (e^{\beta})^2 e^{k \sigma_5} \sigma_5 \\
& - 2 \sigma_{13} e^{k \sigma_5} (e^{\beta})^2 e^{k \sigma_9} \sigma_9 - \sigma_{11} \sigma_1^2 \omega^2 - 2 \sigma_{12} e^{k \sigma_5} (e^{\beta})^2 e^{k \sigma_9} \sigma_9 + 2 \sigma_{13} \sigma_{14} (e^{\beta})^2 \sigma_{10} \\
& + 2 \sigma_{13} \sigma_{14} (e^{\beta})^2 e^{k \sigma_5} \sigma_5 - 2 \sigma_{13} \sigma_{14} (e^{\beta})^2 e^{k \sigma_9} \sigma_9 + 2 e^{\beta} \sigma_{13} k \omega \sigma_{12} e^{k \sigma_9} \sigma_{15} \\
& - 2 e^{\beta} \sigma_{13} e^{k \sigma_5} \sigma_1 \omega - 2 \sigma_{13} e^{k \sigma_9} (e^{\beta})^2 e^{k \sigma_5} \sigma_5 + 2 \sigma_{13} (e^{\beta})^2 (e^{k \sigma_5})^2 \sigma_5 \\
& + 2 \sigma_{12} \sigma_{14} (e^{\beta})^2 \sigma_{10} - 2 \sigma_{13} e^{k \sigma_9} (e^{\beta})^2 \sigma_{10} + 2 e^{\beta} \sigma_{12} e^{k \sigma_9} \sigma_1 \omega + 2 e^{\beta} \sigma_{13} e^{k \sigma_9} \sigma_1 \omega \\
& - 2 e^{\beta} \sigma_{12} \sigma_{14} \sigma_1 \omega - e^{\beta} k \omega^2 \sigma_{12}^2 \sigma_8 - 2 e^{\beta} \sigma_{12} e^{k \sigma_5} \sigma_1 \omega - 2 \sigma_{12} \sigma_{14} (e^{\beta})^2 e^{k \sigma_9} \sigma_9 \\
& + 2 \sigma_{13} e^{k \sigma_5} (e^{\beta})^2 \sigma_{10} + 2 \sigma_{12} (e^{\beta})^2 (e^{k \sigma_9})^2 \sigma_9 + 2 \sigma_3 (e^{\beta})^2 \sigma_{10} e^{k \sigma_9} \sigma_9 + 2 \sigma_3 \sigma_1 \omega e^{\beta} \sigma_{10} \\
& \left. + 2 \sigma_{11} (e^{\beta})^2 \sigma_{10} e^{k \sigma_9} \sigma_9 + \sigma_3 k \omega^2 \sigma_6 \sigma_8 + \sigma_3 k \omega \sigma_6 \sigma_7 + \sigma_3 k \omega \sigma_6 e^{k \sigma_5} \sigma_4 - \sigma_3 \sigma_1^2 \omega^2 \right)
\end{aligned}$$

En calculant les dérivées partielles du second ordre, nous obtenons la matrice d'information I suivante, pour i=2:

$$\hat{I}_{11}^2 = -\frac{\partial^2 \ln L^2}{\partial \beta_2^2} = \exp(\beta_2) \sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \exp\left(K_2 \sin\left(\frac{2\pi t}{365} + \theta_2\right)\right) dt$$

$$\hat{I}_{12}^2 = \hat{I}_{21}^2 = -\frac{\partial^2 \ln L^2}{\partial K_2 \partial \beta_2} = \exp(\beta_2) \sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \sin\left(\frac{2\pi t}{365} + \theta_2\right) \exp\left(K_2 \sin\left(\frac{2\pi t}{365} + \theta_2\right)\right) dt$$

$$\hat{I}_{13}^2 = \hat{I}_{31}^2 = -\frac{\partial^2 \ln L^2}{\partial \theta_2 \partial \beta_2} = \exp(\beta_2) \frac{365}{2\pi} \left[\begin{array}{c} \sum_{j=1}^{n/2} \exp\left(K_2 \sin\left(\frac{2\pi t_{2j}}{365} + \theta_2\right)\right) \\ - \sum_{j=1}^{n/2} \exp\left(K_2 \sin\left(\frac{2\pi t_{2j-1}}{365} + \theta_2\right)\right) \end{array} \right]$$

$$\hat{I}_{22}^2 = -\frac{\partial^2 \ln L^2}{\partial K_2^2} = \sum_{j=1}^{n/2} \exp(\beta_2) \int_{t_{2j-2}}^{t_{2j-1}} \sin\left(\frac{2\pi t}{365} + \theta_2\right)^2 \exp\left(K_2 \sin\left(\frac{2\pi t}{365} + \theta_2\right)\right) dt \quad (12.6)$$

$$\hat{I}_{23}^2 = \hat{I}_{32}^2 = -\frac{\partial^2 \ln L^2}{\partial K_2 \partial \theta_2} = -\sum_{j=1}^{n/2} \cos\left(\frac{2\pi t_{2j}}{365} + \theta_2\right) + \frac{365\beta_2}{2\pi} \sum_{j=1}^{n/2} \left[\begin{array}{c} \sin\left(\frac{2\pi t_{2j}}{365} + \theta_2\right) \exp\left(K_2 \sin\left(\frac{2\pi t_{2j}}{365} + \theta_2\right)\right) \\ - \sin\left(\frac{2\pi t_{2j-1}}{365} + \theta_2\right) \exp\left(K_2 \sin\left(\frac{2\pi t_{2j-1}}{365} + \theta_2\right)\right) \end{array} \right]$$

$$\hat{I}_{33}^2 = -\frac{\partial^2 \ln L^2}{\partial \theta_2^2} = K_2 \sum_{j=1}^{n/2} \sin\left(\frac{2\pi t_{2j}}{365} + \theta_2\right) + \frac{365\beta_2}{2\pi} K_2 \sum_{j=1}^{n/2} \left[\begin{array}{c} \cos\left(\frac{2\pi t_{2j}}{365} + \theta_2\right) \exp\left(K_2 \sin\left(\frac{2\pi t_{2j}}{365} + \theta_2\right)\right) \\ - \cos\left(\frac{2\pi t_{2j-1}}{365} + \theta_2\right) \exp\left(K_2 \sin\left(\frac{2\pi t_{2j-1}}{365} + \theta_2\right)\right) \end{array} \right]$$

Les calculs qui suivent sont effectués dans MAPLE où $\beta = \hat{\beta}_2$, $k = \hat{k}_2$, $\theta = \hat{\theta}_2$, $\omega = \frac{2\pi}{365}$,

$t(j) = t_j$, $tf = t_{fin}$, $tn = t_n$. Les variances des estimateurs sont:

$$\begin{aligned}
\text{var}(\hat{\beta}_2) &= \hat{I}^{-1}_{11} = & (12.7) \\
&\left(\begin{aligned}
&\%1 k \omega^2 \left(\sum_{j=1}^{\frac{1}{2}n} \sin(\omega t(2j) + \theta) \right) \\
&+ \%1 k \omega \left(\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \cos(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \cos(\omega t(2j-1) + \theta) \right) \right) \\
&- \%4^2 \omega^2 + 2 \%4 \omega e^\beta \%2 - (e^\beta)^2 \%2^2 \Big) / \left(e^\beta \left(\%5 \%1 k \omega^2 \left(\sum_{j=1}^{\frac{1}{2}n} \sin(\omega t(2j) + \theta) \right) \right. \right. \\
&+ \%5 \%1 k \omega \left(\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \cos(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \cos(\omega t(2j-1) + \theta) \right) \right) \\
&- \%5 \%4^2 \omega^2 + 2 \%5 \%4 \omega e^\beta \%2 - \%5 (e^\beta)^2 \%2^2 - e^\beta \%3^2 k \omega^2 \left(\sum_{j=1}^{\frac{1}{2}n} \sin(\omega t(2j) + \theta) \right) \\
&- e^\beta \%3^2 k \omega \left(\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \cos(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \cos(\omega t(2j-1) + \theta) \right) \right) \\
&- 2 e^\beta \%3 \left(\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} - e^{k \sin(\omega t(2j-1) + \theta)} \right) \right) \%4 \omega \\
&+ 2 (e^\beta)^2 \%3 \left(\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} - e^{k \sin(\omega t(2j-1) + \theta)} \right) \right) \%2 \\
&\left. \left. - e^\beta \left(\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} - e^{k \sin(\omega t(2j-1) + \theta)} \right) \right)^2 \%1 \right) \right) \\
\%1 &:= \sum_{j=1}^{\frac{1}{2}n} e^\beta \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} \sin(\omega u + \theta)^2 du \\
\%2 &:= \sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \sin(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \sin(\omega t(2j-1) + \theta) \right) \\
\%3 &:= \sum_{j=1}^{\frac{1}{2}n} \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} \sin(\omega u + \theta) du \\
\%4 &:= \sum_{j=1}^{\frac{1}{2}n} \cos(\omega t(2j) + \theta) \\
\%5 &:= \sum_{j=1}^{\frac{1}{2}n} \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} du
\end{aligned}
\right)
\end{aligned}$$

$$\text{var}(\hat{k}_2) = \hat{I}^{2-1}_{22} = \tag{12.8}$$

$$\left(\begin{aligned} & \frac{1}{2} k \omega^2 \left(\sum_{j=1}^{\frac{1}{2}n} \sin(\omega t(2j) + \theta) \right) \\ & + \frac{1}{2} k \omega \left(\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \cos(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \cos(\omega t(2j-1) + \theta) \right) \right) \\ & - e^{\beta} \frac{1}{2} \left(\sum_{j=1}^{\frac{1}{2}n} e^{\beta} \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} \sin(\omega u + \theta)^2 du \right) k \omega^2 \left(\sum_{j=1}^{\frac{1}{2}n} \sin(\omega t(2j) + \theta) \right) + \frac{1}{2} \\ & \left(\sum_{j=1}^{\frac{1}{2}n} e^{\beta} \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} \sin(\omega u + \theta)^2 du \right) k \omega \\ & \left(\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \cos(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \cos(\omega t(2j-1) + \theta) \right) \right) \\ & - \frac{1}{2} \left(\sum_{j=1}^{\frac{1}{2}n} \cos(\omega t(2j) + \theta) \right)^2 \omega^2 + 2 \frac{1}{2} \left(\sum_{j=1}^{\frac{1}{2}n} \cos(\omega t(2j) + \theta) \right) \omega e^{\beta} \\ & \left(\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \sin(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \sin(\omega t(2j-1) + \theta) \right) \right) \\ & - \frac{1}{2} \left(e^{\beta} \right)^2 \left(\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \sin(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \sin(\omega t(2j-1) + \theta) \right) \right)^2 \\ & - e^{\beta} \frac{1}{2} k \omega^2 \left(\sum_{j=1}^{\frac{1}{2}n} \sin(\omega t(2j) + \theta) \right) \\ & - e^{\beta} \frac{1}{2} k \omega \left(\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \cos(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \cos(\omega t(2j-1) + \theta) \right) \right) \\ & - 2 e^{\beta} \frac{1}{2} \left(\sum_{j=1}^{\frac{1}{2}n} \cos(\omega t(2j) + \theta) \right) \omega \\ & + 2 \left(e^{\beta} \right)^2 \frac{1}{2} \left(\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \sin(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \sin(\omega t(2j-1) + \theta) \right) \right) \\ & - e^{\beta} \frac{1}{2} \left(\sum_{j=1}^{\frac{1}{2}n} e^{\beta} \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} \sin(\omega u + \theta)^2 du \right) \end{aligned} \right)$$

$$\%1 := \sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} - e^{k \sin(\omega t(2j-1) + \theta)} \right)$$

$$\%2 := \sum_{j=1}^{\frac{1}{2}n} \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} \sin(\omega u + \theta) du$$

$$\%3 := \sum_{j=1}^{\frac{1}{2}n} \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} du$$

$$\begin{aligned}
\text{var}(\hat{\theta}_2) &= \hat{I}_{33}^{-1} = & (12.9) \\
& \left(\frac{1}{n} - e^{\beta} \frac{1}{n} \right) \omega^2 / \left(\frac{1}{n} k \omega^2 \left[\sum_{j=1}^{\frac{1}{2}n} \sin(\omega t(2j) + \theta) \right] \right. \\
& \left. + \frac{1}{n} k \omega \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \cos(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \cos(\omega t(2j-1) + \theta) \right) \right] \right) \\
& - \frac{1}{n} \left[\sum_{j=1}^{\frac{1}{2}n} \cos(\omega t(2j) + \theta) \right]^2 \omega^2 + 2 \frac{1}{n} \left[\sum_{j=1}^{\frac{1}{2}n} \cos(\omega t(2j) + \theta) \right] \omega e^{\beta} \\
& \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \sin(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \sin(\omega t(2j-1) + \theta) \right) \right] \\
& - \frac{1}{n} (e^{\beta})^2 \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \sin(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \sin(\omega t(2j-1) + \theta) \right) \right]^2 \\
& - e^{\beta} \frac{1}{n} k \omega^2 \left[\sum_{j=1}^{\frac{1}{2}n} \sin(\omega t(2j) + \theta) \right] \\
& - e^{\beta} \frac{1}{n} k \omega \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \cos(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \cos(\omega t(2j-1) + \theta) \right) \right] \\
& - 2 e^{\beta} \frac{1}{n} \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} - e^{k \sin(\omega t(2j-1) + \theta)} \right) \right] \left[\sum_{j=1}^{\frac{1}{2}n} \cos(\omega t(2j) + \theta) \right] \omega + 2 (e^{\beta})^2 \frac{1}{n} \\
& \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} - e^{k \sin(\omega t(2j-1) + \theta)} \right) \right] \\
& \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \sin(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \sin(\omega t(2j-1) + \theta) \right) \right] \\
& - e^{\beta} \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} - e^{k \sin(\omega t(2j-1) + \theta)} \right) \right]^2 \frac{1}{n} \\
\frac{1}{n} & := \sum_{j=1}^{\frac{1}{2}n} e^{\beta} \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} \sin(\omega u + \theta)^2 du \\
\frac{1}{n} & := \sum_{j=1}^{\frac{1}{2}n} \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} \sin(\omega u + \theta) du \\
\frac{1}{n} & := \sum_{j=1}^{\frac{1}{2}n} \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} du
\end{aligned}$$

La variance de $\hat{\lambda}^i(t)$

Nous utilisons les mêmes résultats qu'à l'annexe 1 afin d'obtenir la variance de $\hat{\lambda}^i(t)$. Posons

$\xi^i_1 = \exp(\hat{\beta}_i + \hat{K}_i \sin(\frac{2\pi t}{365} + \hat{\theta}_i))$, $\xi^i_2 = \hat{K}_i$, $\xi^i_3 = \hat{\theta}_i$. Nous avons alors:

$$p^i_1 = \ln(\xi^i_1) - \xi^i_2 \sin\left(\frac{2\pi t}{365} + \xi^i_3\right), p^i_2 = \xi^i_2, p^i_3 = \xi^i_3$$

La variance de $\hat{\lambda}^i(t) = \exp(\hat{\beta}_i + \hat{K}_i \sin(\frac{2\pi t}{365} + \hat{\theta}_i))$ est $\Gamma^{i-1} * (\vec{\xi}^i)$ (Lehman, 1983), où

$\Gamma^{i-1} * (\vec{\xi}^i) = J^{i-1} \Gamma^{i-1}(\vec{p}^i) J^{i-1}$ est la matrice de variance-covariance de $\vec{\xi}^i = (\xi^i_1, \xi^i_2, \xi^i_3)$.

$\Gamma^i(\vec{p}^i) = \hat{\Gamma}^i(\hat{\beta}_i, \hat{K}_i, \hat{\theta}_i)$ est la matrice d'information de \vec{p}^i et $J^i_{sr} = \frac{\partial p^i_s}{\partial \xi^i_r}$ est:

$$J^i = \begin{bmatrix} \frac{1}{\xi^i_1} & -\sin\left(\frac{2\pi t}{365} + \xi^i_3\right) & -\xi^i_2 \cos\left(\frac{2\pi t}{365} + \xi^i_3\right) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (12.10)$$

Pour le calcul de $\text{var}(\hat{\lambda}^1(t))$, $\beta = \hat{\beta}_1$, $k = \hat{k}_1$, $\theta = \hat{\theta}_1$, $\omega = \frac{2\pi}{365}$, $t(j) = t_j$, $t_f = t_{fin}$, $t_n = t_n$, et %1 à

%16 sont telles que définies à la page 210 de cette annexe. Pour le calcul de $\text{var}(\hat{\lambda}_2(t))$,

$\beta = \hat{\beta}_2$, $k = \hat{k}_2$, $\theta = \hat{\theta}_2$, $\omega = \frac{2\pi}{365}$, $t(j) = t_j$, $t_f = t_{fin}$, et $t_n = t_n$.

$$\text{var}(\hat{\lambda}^1(t)) = \quad (12.11)$$

$$\begin{aligned} \text{var}(\hat{\lambda}_2(t)) = & \tag{12.12} \\ & - \left(-\%1 k \omega^2 \left[\sum_{j=1}^{\frac{1}{2}n} \sin(\omega t(2j) + \theta) \right] \right. \\ & - \%1 k \omega \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \cos(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \cos(\omega t(2j-1) + \theta) \right) \right] \\ & + \%4^2 \omega^2 - 2 \%4 \omega e^\beta \%2 + (e^\beta)^2 \%4^2 \left. \right) \beta^2 / \left(e^\beta \left[\%5 \%1 k \omega^2 \left[\sum_{j=1}^{\frac{1}{2}n} \sin(\omega t(2j) + \theta) \right] \right. \right. \\ & + \%5 \%1 k \omega \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \cos(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \cos(\omega t(2j-1) + \theta) \right) \right] \\ & - \%5 \%4^2 \omega^2 + 2 \%5 \%4 \omega e^\beta \%2 - \%5 (e^\beta)^2 \%2^2 - e^\beta \%3^2 k \omega^2 \left[\sum_{j=1}^{\frac{1}{2}n} \sin(\omega t(2j) + \theta) \right] \\ & - e^\beta \%3^2 k \omega \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \cos(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \cos(\omega t(2j-1) + \theta) \right) \right] \\ & - 2 e^\beta \%3 \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} - e^{k \sin(\omega t(2j-1) + \theta)} \right) \right] \%4 \omega \\ & + 2 (e^\beta)^2 \%3 \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} - e^{k \sin(\omega t(2j-1) + \theta)} \right) \right] \%2 \\ & \left. - e^\beta \left[\sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} - e^{k \sin(\omega t(2j-1) + \theta)} \right) \right]^2 \%1 \right) \right) \\ \%1 := & \sum_{j=1}^{\frac{1}{2}n} e^\beta \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} \sin(\omega u + \theta)^2 du \\ \%2 := & \sum_{j=1}^{\frac{1}{2}n} \left(e^{k \sin(\omega t(2j) + \theta)} \sin(\omega t(2j) + \theta) - e^{k \sin(\omega t(2j-1) + \theta)} \sin(\omega t(2j-1) + \theta) \right) \\ \%3 := & \sum_{j=1}^{\frac{1}{2}n} \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} \sin(\omega u + \theta) du \\ \%4 := & \sum_{j=1}^{\frac{1}{2}n} \cos(\omega t(2j) + \theta) \\ \%5 := & \sum_{j=1}^{\frac{1}{2}n} \int_{t(2j-1)}^{t(2j)} e^{k \sin(\omega u + \theta)} du \end{aligned}$$

13. ANNEXE 4: ALGORITHMES DE SIMULATION

Nous présentons dans cette annexe les algorithmes de simulation utilisés au chapitre 6.

Génération d'une loi normale multivariée

L'algorithme utilisé pour la génération de n variables X_i provenant d'une loi normale multivariée de moyenne $\bar{\mu} = (\mu_1, \mu_2, \dots, \mu_n)$ et de matrice de variance covariance Σ est le suivant (Law et Kelton, 1991):

- Générer Z_1, \dots, Z_n , n variables provenant d'une loi $N(0,1)$
- Pour $i = 1, \dots, n$, poser $X_i = \mu_i + c_{ij}Z_j$, tel que $\Sigma = CC^t$.

Les algorithmes pour simuler des variables uniforme, exponentielle, Gamma, et normale sont ceux fournis dans MATLAB.

Génération d'un PNH (horizon spécifié)

L'algorithme de simulation d'un processus de Poisson non homogène d'intensité $\lambda(t)$ sur un horizon de k années, et dont les temps d'arrivée sont $tarriv(s)$ est le suivant:

$tt=0$;

Pour $i=1$ à k

 Pour $j=1$ à 365

$tt:=tt+1$

 Générer u une variable uniforme sur $(0,1)$

 Si $u < \lambda(j)$ alors nouvel événement et $tarriv(\text{nouvel événement})= tt$

 fin

 fin

fin

Génération d'un RNH (nombre d'événements spécifié)

L'algorithme de génération d'un processus de renouvellement non homogène défini par les temps d'arrivée $tarriv(s)$, composé de $nbev$ événements, d'intensité $\rho(t, H_t) = \varphi(x)\lambda(t)$ où $\varphi(x)$ est le taux de rupture d'un processus de renouvellement homogène dont les durées x suivent

une loi Gamma de moyenne unitaire $= \frac{\alpha}{\rho}$ et de variance égale à $\frac{\alpha}{\rho^2}$, et où $\lambda(t)$ est une fonction de changement d'échelle est:

tarriv(1)=1;

Pour i=1:nbev

Générer γ une variable Gamma de moyenne unitaire et de paramètres $\alpha = \rho$

Résoudre pour tarriv(i+1): $\int_{\text{tarriv}(i)}^{\text{tarriv}(i+1)} \lambda(u) du = \gamma$

fin

Génération d'un PNHC (horizon spécifié)

L'algorithme de simulation d'un processus de Poisson non homogène conditionnel défini par les temps d'arrivée tarriv(s) sur un horizon de longueur tfin, d'intensité $\rho(t, H_t) = f(PE(t))$, étant donné le processus externe PE, est le suivant:

tt=0;

Pour i=1:tfin

tt:=tt+1

Générer u une variable uniforme sur (0,1)

Si $u < \rho(i, H_i)$ alors nouvel événement et tarriv(nouvel événement)= tt

fin

fin

Ceci revient à générer un processus de Poisson non homogène dont la fonction d'intensité dépend d'un processus externe PE(t).

Génération d'une série alternée

Nous avons présenté les algorithmes de génération de processus PNH, RNH, et PNHC d'un seul type. Ces algorithmes doivent être utilisés simultanément pour générer une série alternée. Dans cette situation, nous avons deux types de tarriv (tarriv1, et tarriv2) et deux types d'intensité ($\rho_1(t, H_t)$ et $\rho_2(t, H_t)$).

Il est en général aisé de générer un processus alterné composé d'une combinaison de PNH/PNHC en utilisant ces algorithmes. Toutefois, il est plus difficile de générer une combinaison RNH/(PNH,PNHC) car la génération de RNH se fait en fonction d'un nombre spécifié d'événements, tandis que la génération de PNH et PNHC se fait sur un horizon spécifié. Lorsqu'on doit générer un processus alterné RNH avec PNHC ou PNH, la génération du PNHC ou PNH se fait selon l'algorithme basé sur les durées présenté dans ce qui suit.

Génération d'un PNH ou d'un PNHC (durées)

La génération d'un PNH ou RNH basée sur les durée peut être effectuée en générant des variables exponentielles de paramètre 1 et en effectuant l'inverse d'une homogénéisation à l'aide de la fonction d'intensité $\rho(u, H_u)$. Dans le cas d'un PNH, $\rho(u, H_u) = \lambda(t)$. L'algorithme pour générer $dur(s)$, $s = 1..nbev$, est le suivant:

tarriv(1)=1;

Pour i=1:nbev

Générer ε une variable exponentielle de paramètre 1

Résoudre pour tarriv(i+1): $\int_{tarriv(i)}^{tarriv(i+1)} \rho(u, H_u) du = \varepsilon$

dur(i)=tarriv(i)-tarriv(i-1)

fin

14. ANNEXE 5: MODÈLES DE $\rho^1(t;H_t)$ POUR LE PNHC

Lors du développement du PNHC comme modèle d'occurrence de surplus, nous avons essayé sans succès plusieurs modèles pour l'intensité des surplus $\rho^1(t;H_t)$. Ces modèles sont les suivants:

$$\rho^1(t, H_t) = \exp[K_1 \sin(\omega t + \theta_1)]$$

$$\rho^1(t, H_t) = \exp[K_1 \sin(\omega t + \theta_1)] \exp(\alpha_1) / \text{pnorm}(t)$$

$$\rho^1(t, H_t) = \exp[K_1 \text{pnorm}(t) + \sin(\omega t + \theta_1)]$$

$$\rho^1(t, H_t) = \exp[K_1 \text{pnorm}(t) \sin(\omega t + \theta_1)]$$

$$\rho^1(t, H_t) = \exp[K_1 \text{pnorm}(t) + \theta_1]$$

$$\rho^1(t, H_t) = \exp[K_1 \sin(\omega t + \theta_1)] * \text{nbjour sec}(t)$$

$$\rho^1(t, H_t) = \exp[K_1 \sin(\omega t + \theta_1)] \exp(\text{nbjour sec}(t))$$

$$\rho^1(t, H_t) = \exp[K_1 \sin(\omega t + \theta_1)] \text{nbjour sec}(t) \exp(\alpha_1)$$

$$\rho^1(t, H_t) = \exp[\alpha_1 + K_1 \text{nbjour sec}(t)]$$

$$\rho^1(t, H_t) = \left(\exp[\alpha_1 + K_1 \text{nbjour sec}(t)] \right)^{\theta_1}$$

$$\rho^1(t, H_t) = \exp[\alpha_1 + K_1 \text{nbjour sec}(t) + \theta_1]$$

$$\rho^1(t, H_t) = \exp[\alpha_1 + K_1 (\text{nbjour sec}(t) + \theta_1)]$$

où $P(t)$ est la quantité de précipitations au temps t , $P(t)$ étant égale à zéro si la température est inférieure à -2 C, $\text{nbjoursec}(t)$ est le nombre de jours $s < t$ tels que $P(s)=0$ divisé par 365,

$$\text{et } \text{pnorm}(t) = \frac{\sum_{kk=t-nn-1}^{kk=t} P(kk)}{\max_t \left(\sum_{kk=t-nn-1}^{kk=t} P(kk) \right)}, \text{ pour } nn = 7, 30, \text{ et } 90.$$

15. ANNEXE 6: LES VARIANCES DES PARAMÈTRES DU PNHC

Nous calculons dans ce qui suit les variances des estimateurs $\hat{\beta}_1$, \hat{K}_1 , et $\hat{\theta}_1$, estimés par la méthode du MV pour le PNHC des surplus. Posons $\bar{p}^1 = (p^1_1, p^1_2, p^1_3)$ où $p^1_1 = \beta_1, p^1_2 = K_1, p^1_3 = \theta_1$. La loi asymptotique des estimateurs du MV, \hat{p}^1 , est approximativement $N(\hat{p}^1, \hat{I}^1)$, où les éléments (r,s) de $\hat{I}^1(\hat{\beta}_1, \hat{K}_1, \hat{\theta}_1)$ sont

$$\hat{I}^1_{rs} = \left[-\frac{\partial^2}{\partial p^1_r \partial p^1_s} \ln L^1 \right]_{\bar{p}^1 = \hat{p}^1} \quad (\text{Berger, 1985}). \text{ Nous avons donc:}$$

$$\begin{aligned} \text{var}(\hat{\beta}_1) &= \hat{I}^1_{11} \\ \text{var}(\hat{K}_1) &= \hat{I}^1_{22} \\ \text{var}(\hat{\theta}_1) &= \hat{I}^1_{33} \end{aligned} \quad (15.1)$$

Les dérivées partielles du deuxième ordre sont:

$$\begin{aligned} \hat{I}^1_{11} &= -\frac{\partial^2 \ln L^1}{\partial^2 \beta_1} = \sum_{j=2}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} f(u) du + \int_{t_n}^{t_{2f}} f(u) du \\ \hat{I}^1_{21} = \hat{I}^1_{12} &= -\frac{\partial^2 \ln L^1}{\partial^2 \beta_1 K_1} = \sum_{j=2}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} \text{nbjoursec}(u) f(u) du + \int_{t_n}^{t_{2f}} \text{nbjoursec}(u) f(u) du \\ \hat{I}^1_{31} = \hat{I}^1_{13} &= -\frac{\partial^2 \ln L^1}{\partial^2 \beta_1 \theta_1} = \sum_{j=2}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} \sin\left(\frac{2\pi u}{365}\right) f(u) du + \int_{t_n}^{t_{2f}} \sin\left(\frac{2\pi u}{365}\right) f(u) du \\ \hat{I}^1_{22} &= -\frac{\partial^2 \ln L^1}{\partial^2 K_1^2} = \sum_{j=2}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} \text{nbjoursec}(u)^2 f(u) du + \int_{t_n}^{t_{2f}} \text{nbjoursec}(u)^2 f(u) du \\ \hat{I}^1_{32} = \hat{I}^1_{23} &= -\frac{\partial^2 \ln L^1}{\partial^2 K_1 \theta_1} = \sum_{j=2}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} \text{nbjoursec}(u) \sin\left(\frac{2\pi u}{365}\right) f(u) du + \int_{t_n}^{t_{2f}} \text{nbjoursec}(u) \sin\left(\frac{2\pi u}{365}\right) f(u) du \\ \hat{I}^1_{33} &= -\frac{\partial^2 \ln L^1}{\partial^2 \theta_1^2} = \sum_{j=2}^{n/2} \int_{t_{2j-2}}^{t_{2j-1}} \sin^2\left(\frac{2\pi u}{365}\right) f(u) du + \int_{t_n}^{t_{2f}} \sin^2\left(\frac{2\pi u}{365}\right) f(u) du \end{aligned} \quad (15.2)$$

où $f(u) = \exp\left(\beta_1 + K_1 \text{nbjoursec}(u) + \theta_1 \sin\left(\frac{2\pi u}{365}\right)\right)$. Dans ce qui suit, $\beta = \hat{\beta}_1$, $k = \hat{k}_1$, $\theta = \hat{\theta}_1$,

$$\omega = \frac{2\pi}{365}, \quad t(j) = t_j, \quad t_f = t_{fin}, \quad t_n = t_n, \quad \text{et nbjour}_{[u]} \text{ est nbjoursec}(u).$$

$$\text{var}(\hat{\beta}_1) = \hat{I}^{-1}_{11} = \quad (15.3)$$

$$\left[\begin{aligned} & \%11 \%1 + \%11 \int_{tn}^{tf} \%4 du + \%9 \%1 + \%9 \int_{tn}^{tf} \%4 du - \%7^2 - 2 \%7 \int_{tn}^{tf} \%6 du - \left(\int_{tn}^{tf} \%6 du \right)^2 \right] / \left[\%8 \%11 \%1 \right. \\ & + \%8 \%11 \int_{tn}^{tf} \%4 du + \%8 \%9 \%1 + \%8 \%9 \int_{tn}^{tf} \%4 du - 2 \%8 \%7 \int_{tn}^{tf} \%6 du + \int_{tn}^{tf} \%2 du \%11 \%1 \\ & + \int_{tn}^{tf} \%2 du \%11 \int_{tn}^{tf} \%4 du + \int_{tn}^{tf} \%2 du \%9 \%1 + \int_{tn}^{tf} \%2 du \%9 \int_{tn}^{tf} \%4 du - 2 \int_{tn}^{tf} \%2 du \%7 \int_{tn}^{tf} \%6 du \\ & - 2 \%5 \%3 \%1 - 2 \%5 \%3 \int_{tn}^{tf} \%4 du + 2 \%5 \%12 \%7 + 2 \%5 \%12 \int_{tn}^{tf} \%6 du + 2 \%5 \%10 \%7 \\ & + 2 \%5 \%10 \int_{tn}^{tf} \%6 du + 2 \%3 \%12 \%7 + 2 \%3 \%12 \int_{tn}^{tf} \%6 du + 2 \%3 \%10 \%7 + 2 \%3 \%10 \int_{tn}^{tf} \%6 du \\ & - 2 \%12 \%10 \%11 - 2 \%12 \%10 \%9 - \%3^2 \int_{tn}^{tf} \%4 du - \%12^2 \%11 - \%12^2 \%9 - \%10^2 \%11 - \%10^2 \%9 \\ & - \%8 \%7^2 - \%8 \left(\int_{tn}^{tf} \%6 du \right)^2 - \int_{tn}^{tf} \%2 du \%7^2 - \int_{tn}^{tf} \%2 du \left(\int_{tn}^{tf} \%6 du \right)^2 - \%5^2 \%1 - \%5^2 \int_{tn}^{tf} \%4 du \\ & \left. - \%3^2 \%1 \right] \end{aligned} \right)$$

$$\%1 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \sin(\omega u)^2 \%2 du$$

$$\%2 := e^{\alpha + k \text{nbjour}_{[u]} + \theta \sin(\omega u)}$$

$$\%3 := \int_{tn}^{tf} \text{nbjour}_{[u]} \%2 du$$

$$\%4 := \sin(\omega u)^2 \%2$$

$$\%5 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \text{nbjour}_{[u]} \%2 du$$

$$\%6 := \text{nbjour}_{[u]} \sin(\omega u) \%2$$

$$\%7 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \%6 du$$

$$\%8 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \%2 du$$

$$\%9 := \int_{tn}^{tf} \text{nbjour}_{[u]}^2 \%2 du$$

$$\%10 := \int_{tn}^{tf} \sin(\omega u) \%2 du$$

$$\%11 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \text{nbjour}_{[u]}^2 \%2 du$$

$$\%12 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \sin(\omega u) \%2 du$$

$$\text{var}(\hat{K}_1) = \hat{I}^{-1}_{22} = \quad (15.4)$$

$$\left(\begin{aligned} & \%8 \%1 + \%8 \int_{tn}^{tf} \%4 du + \int_{tn}^{tf} \%2 du \%1 + \int_{tn}^{tf} \%2 du \int_{tn}^{tf} \%4 du - \%12^2 - 2 \%12 \%10 - \%10^2 \\ & + \%8 \%11 \int_{tn}^{tf} \%4 du + \%8 \%9 \%1 + \%8 \%9 \int_{tn}^{tf} \%4 du - 2 \%8 \%7 \int_{tn}^{tf} \%6 du + \int_{tn}^{tf} \%2 du \%11 \%1 \\ & + \int_{tn}^{tf} \%2 du \%11 \int_{tn}^{tf} \%4 du + \int_{tn}^{tf} \%2 du \%9 \%1 + \int_{tn}^{tf} \%2 du \%9 \int_{tn}^{tf} \%4 du - 2 \int_{tn}^{tf} \%2 du \%7 \int_{tn}^{tf} \%6 du \\ & - 2 \%5 \%3 \%1 - 2 \%5 \%3 \int_{tn}^{tf} \%4 du + 2 \%5 \%12 \%7 + 2 \%5 \%12 \int_{tn}^{tf} \%6 du + 2 \%5 \%10 \%7 \\ & + 2 \%5 \%10 \int_{tn}^{tf} \%6 du + 2 \%3 \%12 \%7 + 2 \%3 \%12 \int_{tn}^{tf} \%6 du + 2 \%3 \%10 \%7 + 2 \%3 \%10 \int_{tn}^{tf} \%6 du \\ & - 2 \%12 \%10 \%11 - 2 \%12 \%10 \%9 - \%3^2 \int_{tn}^{tf} \%4 du - \%12^2 \%11 - \%12^2 \%9 - \%10^2 \%11 - \%10^2 \%9 \\ & - \%8 \%7^2 - \%8 \left(\int_{tn}^{tf} \%6 du \right)^2 - \int_{tn}^{tf} \%2 du \%7^2 - \int_{tn}^{tf} \%2 du \left(\int_{tn}^{tf} \%6 du \right)^2 - \%5^2 \%1 - \%5^2 \int_{tn}^{tf} \%4 du \\ & - \%3^2 \%1 \end{aligned} \right) / \left(\%8 \%11 \%1 \right.$$

$$\begin{aligned} \%1 &:= \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \sin(\omega u)^2 \%2 du \\ \%2 &:= e^{\alpha + k \text{nbjour}_{[u]} + \theta \sin(\omega u)} \\ \%3 &:= \int_{tn}^{tf} \text{nbjour}_{[u]} \%2 du \\ \%4 &:= \sin(\omega u)^2 \%2 \\ \%5 &:= \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \text{nbjour}_{[u]} \%2 du \\ \%6 &:= \text{nbjour}_{[u]} \sin(\omega u) \%2 \\ \%7 &:= \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \%6 du \\ \%8 &:= \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \%2 du \\ \%9 &:= \int_{tn}^{tf} \text{nbjour}_{[u]}^2 \%2 du \\ \%10 &:= \int_{tn}^{tf} \sin(\omega u) \%2 du \\ \%11 &:= \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \text{nbjour}_{[u]}^2 \%2 du \\ \%12 &:= \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \sin(\omega u) \%2 du \end{aligned}$$

$$\text{var}(\hat{\theta}_1) = \hat{I}^{-1}_{33} = \quad (15.5)$$

$$\left(\begin{aligned} & \%8 \%11 + \%8 \%9 + \int_{tn}^{tf} \%2 du \%11 + \int_{tn}^{tf} \%2 du \%9 - \%5^2 - 2 \%5 \%3 - \%3^2 \end{aligned} \right) / \left(\begin{aligned} & \%8 \%11 \%1 \\ & + \%8 \%11 \int_{tn}^{tf} \%4 du + \%8 \%9 \%1 + \%8 \%9 \int_{tn}^{tf} \%4 du - 2 \%8 \%7 \int_{tn}^{tf} \%6 du + \int_{tn}^{tf} \%2 du \%11 \%1 \\ & + \int_{tn}^{tf} \%2 du \%11 \int_{tn}^{tf} \%4 du + \int_{tn}^{tf} \%2 du \%9 \%1 + \int_{tn}^{tf} \%2 du \%9 \int_{tn}^{tf} \%4 du - 2 \int_{tn}^{tf} \%2 du \%7 \int_{tn}^{tf} \%6 du \\ & - 2 \%5 \%3 \%1 - 2 \%5 \%3 \int_{tn}^{tf} \%4 du + 2 \%5 \%12 \%7 + 2 \%5 \%12 \int_{tn}^{tf} \%6 du + 2 \%5 \%10 \%7 \\ & + 2 \%5 \%10 \int_{tn}^{tf} \%6 du + 2 \%3 \%12 \%7 + 2 \%3 \%12 \int_{tn}^{tf} \%6 du + 2 \%3 \%10 \%7 + 2 \%3 \%10 \int_{tn}^{tf} \%6 du \\ & - 2 \%12 \%10 \%11 - 2 \%12 \%10 \%9 - \%3^2 \int_{tn}^{tf} \%4 du - \%12^2 \%11 - \%12^2 \%9 - \%10^2 \%11 - \%10^2 \%9 \\ & - \%8 \%7^2 - \%8 \left(\int_{tn}^{tf} \%6 du \right)^2 - \int_{tn}^{tf} \%2 du \%7^2 - \int_{tn}^{tf} \%2 du \left(\int_{tn}^{tf} \%6 du \right)^2 - \%5^2 \%1 - \%5^2 \int_{tn}^{tf} \%4 du \\ & - \%3^2 \%1 \end{aligned} \right)$$

$$\%1 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \sin(\omega u)^2 \%2 du$$

$$\%2 := e^{\alpha + k \text{nbjour}_{[u]} + \theta \sin(\omega u)}$$

$$\%3 := \int_{tn}^{tf} \text{nbjour}_{[u]} \%2 du$$

$$\%4 := \sin(\omega u)^2 \%2$$

$$\%5 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \text{nbjour}_{[u]} \%2 du$$

$$\%6 := \text{nbjour}_{[u]} \sin(\omega u) \%2$$

$$\%7 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \%6 du$$

$$\%8 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \%2 du$$

$$\%9 := \int_{tn}^{tf} \text{nbjour}_{[u]}^2 \%2 du$$

$$\%10 := \int_{tn}^{tf} \sin(\omega u) \%2 du$$

$$\%11 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \text{nbjour}_{[u]}^2 \%2 du$$

$$\%12 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \sin(\omega u) \%2 du$$

La variance de $\hat{\rho}^1(t, H_t)$

Nous utilisons les mêmes résultats qu'à l'annexe 1 afin d'obtenir la variance de $\hat{\rho}^1(t, H_t)$.

Posons $\xi^1_1 = \exp\left(\hat{\beta}_1 + \hat{K}_1 \text{nbjour} \sec(t) + \hat{\theta}_1 \sin\left(\frac{2\pi t}{365}\right)\right)$, $\xi^1_2 = \hat{K}_1$, $\xi^1_3 = \hat{\theta}_1$. Nous avons alors:

$$p^1_1 = \ln(\xi^1_1) - \xi^1_2 \text{nbjour}(t) - \xi^1_3 \sin\left(\frac{2\pi t}{365}\right), p^1_2 = \xi^1_2, p^1_3 = \xi^1_3 \quad (15.6)$$

La variance de $\hat{\rho}^1(t, H_t) = \exp\left[\hat{\beta}_1 + \hat{K}_1 \text{nbjour} \sec(t) + \hat{\theta}_1 \sin\left(\frac{2\pi t}{365}\right)\right]$ est (Lehman, 1983):

$$\text{var}(\hat{\rho}^1(t, H_t)) = I^{1-1}_{1,1} * (\vec{\xi}^1) \quad (15.7)$$

où $I^{1-1} * (\vec{\xi}^1) = J^{1-1} I^{-1}(\vec{p}^1) J^{1-1}$ est la matrice de variance-covariance de $\vec{\xi}^1 = (\xi^1_1, \xi^1_2, \xi^1_3)$,

$I^1(\vec{p}^1) = \hat{I}^1(\hat{\beta}_1, \hat{K}_1, \hat{\theta}_1)$ est la matrice d'information de \vec{p}^1 , $J^1 = \left\| \frac{\partial p^1_s}{\partial \xi^1_r} \right\|$, $J^1_{sr} = \frac{\partial p^1_s}{\partial \xi^1_r}$ telle que:

$$J^1 = \begin{bmatrix} \frac{1}{\xi^1_1} & -\text{nbjour} \sec(t) & -\sin\left(\frac{2\pi t}{365}\right) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (15.8)$$

$$\text{var}(\hat{\rho}^1(t, H_t)) = \Gamma_{1,1}^{-1} * (\vec{\xi}^1) = \tag{15.9}$$

$$- \left(\begin{aligned} & \%11 \%1 + \%11 \int_{tn}^{tf} \%4 du + \%9 \%1 + \%9 \int_{tn}^{tf} \%4 du - \%7^2 - 2 \%7 \int_{tn}^{tf} \%6 du - \left(\int_{tn}^{tf} \%6 du \right)^2 \alpha^2 / \left(\right. \\ & - \%8 \%11 \%1 - \%8 \%11 \int_{tn}^{tf} \%4 du - \%8 \%9 \%1 - \%8 \%9 \int_{tn}^{tf} \%4 du + 2 \%8 \%7 \int_{tn}^{tf} \%6 du - \int_{tn}^{tf} \%2 du \%11 \%1 \\ & - \int_{tn}^{tf} \%2 du \%11 \int_{tn}^{tf} \%4 du - \int_{tn}^{tf} \%2 du \%9 \%1 - \int_{tn}^{tf} \%2 du \%9 \int_{tn}^{tf} \%4 du + 2 \int_{tn}^{tf} \%2 du \%7 \int_{tn}^{tf} \%6 du \\ & + 2 \%5 \%3 \%1 + 2 \%5 \%3 \int_{tn}^{tf} \%4 du - 2 \%5 \%12 \%7 - 2 \%5 \%12 \int_{tn}^{tf} \%6 du - 2 \%5 \%10 \%7 \\ & - 2 \%5 \%10 \int_{tn}^{tf} \%6 du - 2 \%3 \%12 \%7 - 2 \%3 \%12 \int_{tn}^{tf} \%6 du - 2 \%3 \%10 \%7 - 2 \%3 \%10 \int_{tn}^{tf} \%6 du \\ & + 2 \%12 \%10 \%11 + 2 \%12 \%10 \%9 + \%3^2 \int_{tn}^{tf} \%4 du + \%12^2 \%11 + \%12^2 \%9 + \%10^2 \%11 + \%10^2 \%9 \\ & + \%8 \%7^2 + \%8 \left(\int_{tn}^{tf} \%6 du \right)^2 + \int_{tn}^{tf} \%2 du \%7^2 + \int_{tn}^{tf} \%2 du \left(\int_{tn}^{tf} \%6 du \right)^2 + \%5^2 \%1 + \%5^2 \int_{tn}^{tf} \%4 du \\ & \left. + \%3^2 \%1 \right) \end{aligned} \right)$$

$$\%1 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \sin(\omega u)^2 \%2 du$$

$$\%2 := e^{\alpha + k \text{nbjour}_{[u]} + \theta \sin(\omega u)}$$

$$\%3 := \int_{tn}^{tf} \text{nbjour}_{[u]} \%2 du$$

$$\%4 := \sin(\omega u)^2 \%2$$

$$\%5 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \text{nbjour}_{[u]} \%2 du$$

$$\%6 := \text{nbjour}_{[u]} \sin(\omega u) \%2$$

$$\%7 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \%6 du$$

$$\%8 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \%2 du$$

$$\%9 := \int_{tn}^{tf} \text{nbjour}_{[u]}^2 \%2 du$$

$$\%10 := \int_{tn}^{tf} \sin(\omega u) \%2 du$$

$$\%11 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \text{nbjour}_{[u]}^2 \%2 du$$

$$\%12 := \sum_{j=2}^{\frac{1}{2}n} \int_{t(2j-2)}^{t(2j-1)} \sin(\omega u) \%2 du$$

Pour ce qui est de la fonction d'intensité des étiages du PNHC, $\rho^2(t, H_t) = \exp(K_2 \text{pnorm}(t) + \theta_2)$, la loi asymptotique des estimateurs du MV $\hat{p}^2 = (\hat{K}^2, \hat{\theta}^2)$, est approximativement $N(\hat{p}^2, \hat{I}^{2-1})$, où $\hat{I}^{2-1} = \left[-\frac{\partial^2}{\partial p_r^2 \partial p_s^2} \ln L^2 \right]_{\hat{p}^2 = \hat{p}^2}$, (Lehman, 1983) où:

$$\begin{aligned}\hat{I}_{11}^2 &= -\frac{\partial^2 \ln L^2}{\partial^2 K_2} = \sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \text{pnorm}^2(u) \exp(K_2 \text{pnorm}(u) + \theta_2) du \\ \hat{I}_{12}^2 &= \hat{I}_{21}^2 = -\frac{\partial^2 \ln L^2}{\partial K_2 \partial \theta_2} = \sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \text{pnorm}(u) \exp(K_2 \text{pnorm}(u) + \theta_2) du \\ \hat{I}_{22}^2 &= -\frac{\partial^2 \ln L^2}{\partial^2 \theta_2} = \sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \exp(K_2 \text{pnorm}(u) + \theta_2) du\end{aligned}\quad (15.10)$$

La matrice de variance-covariance est \hat{I}^{2-1} telle que:

$$\text{var}(\hat{K}^2) = \hat{I}_{11}^{2-1} = \quad (15.11)$$

$$\frac{\sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \exp(K_2 \text{pnorm}(u) + \theta_2) du}{\left(\sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \text{pnorm}^2(u) \exp(K_2 \text{pnorm}(u) + \theta_2) du \sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \exp(K_2 \text{pnorm}(u) + \theta_2) du - \sum_{j=1}^{n/2} \left(\int_{t_{2j-1}}^{t_{2j}} \text{pnorm}(u) \exp(K_2 \text{pnorm}(u) + \theta_2) du \right)^2 \right)}$$

$$\text{cov}(\hat{K}^2, \hat{\theta}^2) = \hat{I}_{21}^{2-1} = \quad (15.12)$$

$$\frac{-\sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \text{pnorm}(u) \exp(K_2 \text{pnorm}(u) + \theta_2) du}{\left(\sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \text{pnorm}^2(u) \exp(K_2 \text{pnorm}(u) + \theta_2) du \sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \exp(K_2 \text{pnorm}(u) + \theta_2) du - \sum_{j=1}^{n/2} \left(\int_{t_{2j-1}}^{t_{2j}} \text{pnorm}(u) \exp(K_2 \text{pnorm}(u) + \theta_2) du \right)^2 \right)}$$

$$\text{var}(\hat{\theta}^2) = \hat{I}_{22}^{2-1} = \quad (15.13)$$

$$\frac{\sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \text{pnorm}^2(u) \exp(K_2 \text{pnorm}(u) + \theta_2) du}{\left(\sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \text{pnorm}^2(u) \exp(K_2 \text{pnorm}(u) + \theta_2) du \sum_{j=1}^{n/2} \int_{t_{2j-1}}^{t_{2j}} \exp(K_2 \text{pnorm}(u) + \theta_2) du - \sum_{j=1}^{n/2} \left(\int_{t_{2j-1}}^{t_{2j}} \text{pnorm}(u) \exp(K_2 \text{pnorm}(u) + \theta_2) du \right)^2 \right)}$$

La variance de $\hat{p}^2(t)$

Calculons $\text{var}(\hat{p}^2(t)) = I^{2-1}_{1,1} * (\vec{\xi})$. Posons: $p_1 = K_2$, $p_2 = \theta_2$, et $\xi_1 = \exp(K_2 p_{\text{norm}}(t) + \theta_2)$, $\xi_2 = \theta_2$. Nous avons alors:

$$p_1 = \frac{\ln \xi_1 - \xi_2}{p_{\text{norm}}(t)}, p_2 = \xi_2 \tag{15.14}$$

La matrice d'information de $\vec{\xi} = (\xi_1, \xi_2)$, $I^2 * (\vec{\xi})$ est $I^2 * (\vec{\xi}) = J^2 I^2(\vec{p}) J^{2'}$ (Lehman, 1983) où $J^2_{sr} = \frac{\partial p_s}{\partial \xi_r}$ et $\hat{I}^2(\vec{p}) = \hat{I}^2(\hat{K}_2, \hat{\theta}_2)$ est la matrice d'information de \vec{p} . Nous avons:

$$J^2 = \begin{bmatrix} \frac{1}{p_{\text{norm}}(t) \exp(K_2 p_{\text{norm}}(t) + \theta_2)} & \frac{-1}{p_{\text{norm}}(t)} \\ 0 & 1 \end{bmatrix} \tag{15.15}$$

et donc:

$$\text{var}(\hat{p}^2(t)) = I^2_{11} * (\vec{\xi}) = \tag{15.16}$$

$$\frac{\sum_{j=1}^{n/2} \int_{t_{j-1}}^{t_j} \exp(K_2 p_{\text{norm}}(u) + \theta_2) du * p_{\text{norm}}^2(t) * (\exp(K_2 p_{\text{norm}}(t) + \theta_2))^2}{\sum_{j=1}^{n/2} \int_{t_{j-1}}^{t_j} p_{\text{norm}}^2(u) \exp(K_2 p_{\text{norm}}(u) + \theta_2) du \sum_{j=1}^{n/2} \int_{t_{j-1}}^{t_j} \exp(K_2 p_{\text{norm}}(u) + \theta_2) du - \left(\sum_{j=1}^{n/2} \int_{t_{j-1}}^{t_j} p_{\text{norm}}(u) \exp(K_2 p_{\text{norm}}(u) + \theta_2) du \right)^2}$$

16. ANNEXE 7: LES DÉBITS MOYENS MENSUELS

Nous présentons dans cette annexe les débits moyens mensuels des rivières étudiées. Le mois 1 correspond à janvier et le mois 12 à décembre.

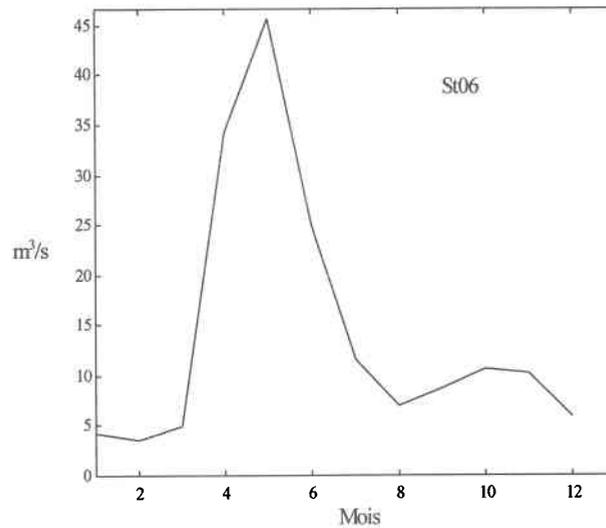


Figure 16.1: Les débits moyens mensuels - St06

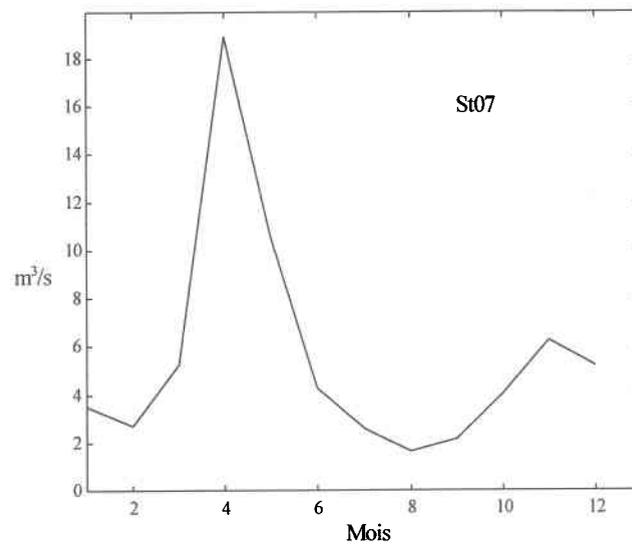


Figure 16.2: Les débits moyens mensuels - St07

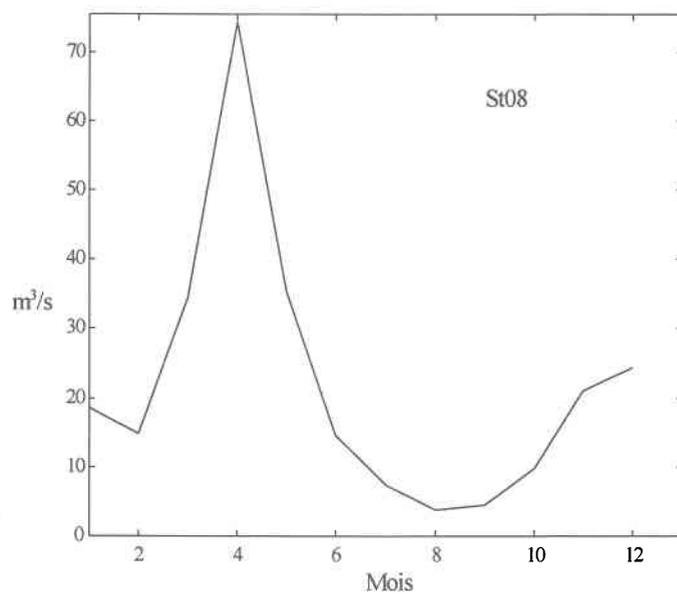


Figure 16.3: Les débits moyens mensuels - St08

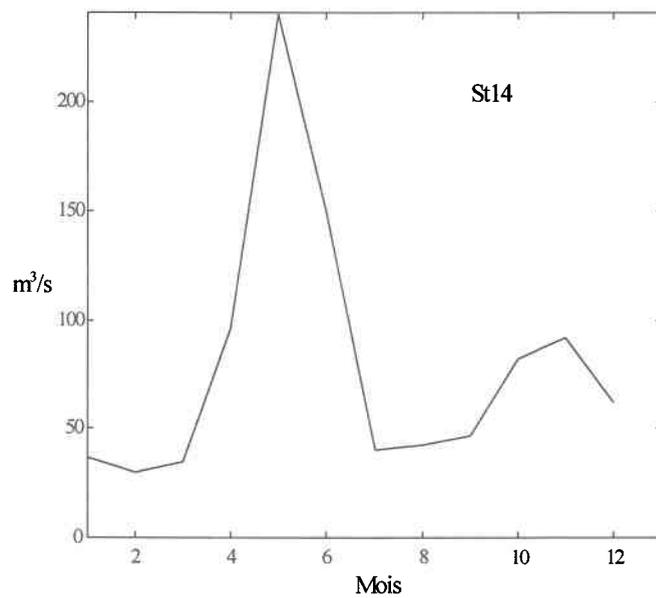


Figure 16.4: Les débits moyens mensuels - St14

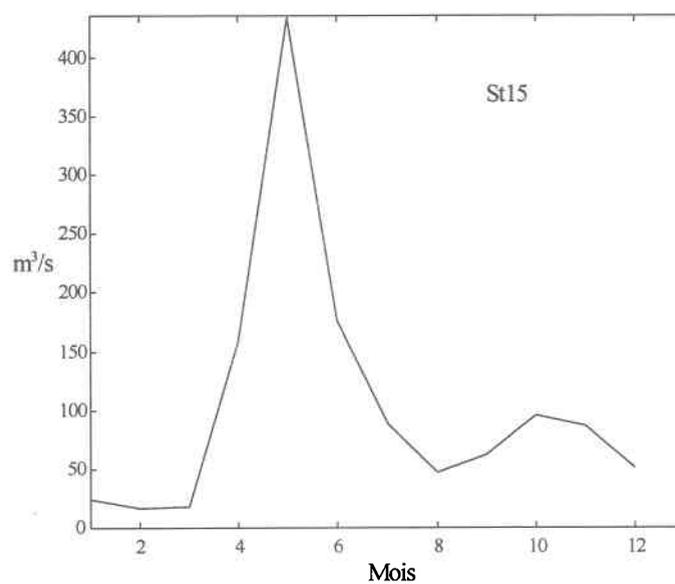


Figure 16.5: Les débits moyens mensuels - St15

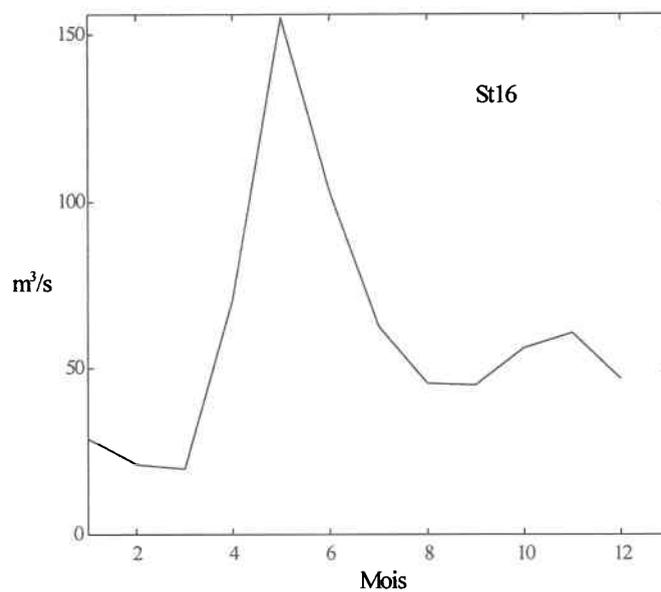


Figure 16.6: Les débits moyens mensuels - St16

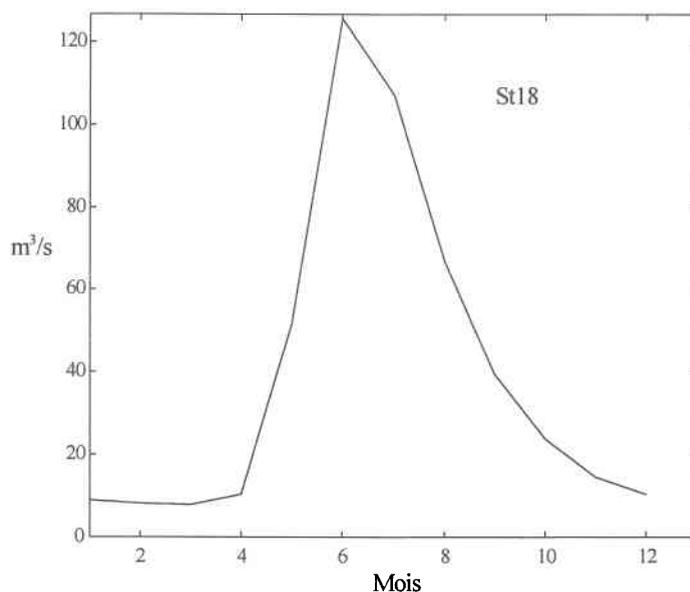


Figure 16.7: Les débits moyens mensuels - St18

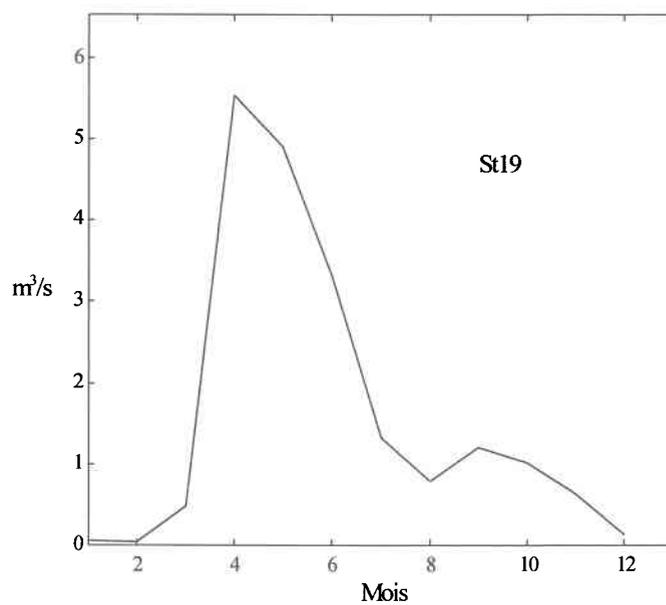


Figure 16.8: Les débits moyens mensuels - St19

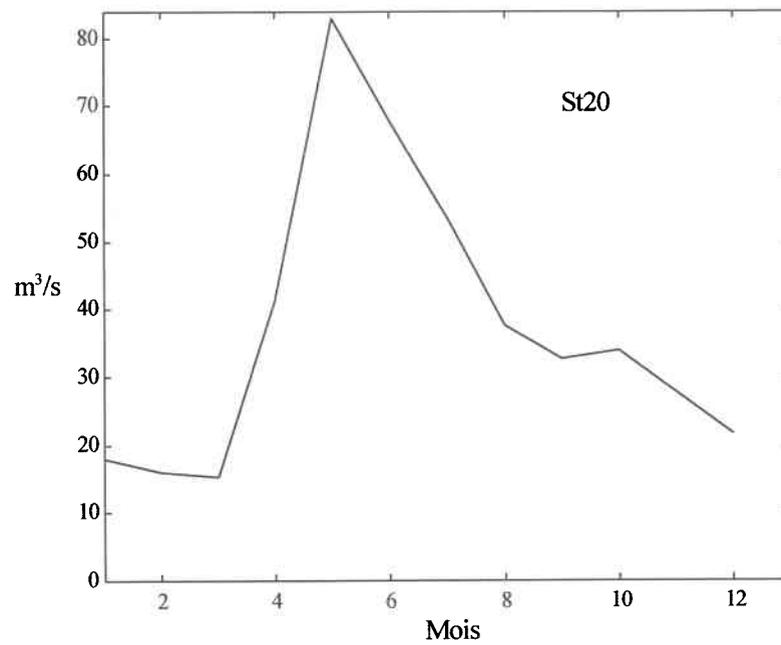


Figure 16.9: Les débits moyens mensuels - St20

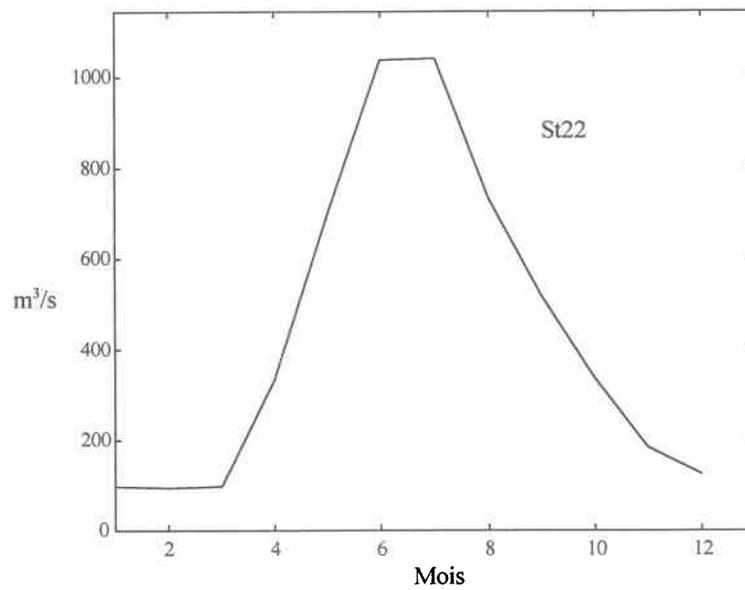


Figure 16.10: Les débits moyens mensuels - St22

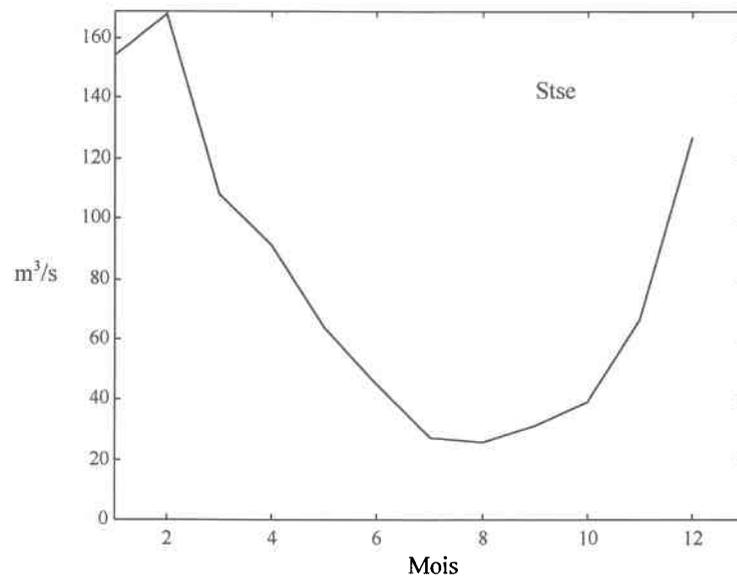


Figure 16.11: Les débits moyens mensuels - Stse

17. ANNEXE 8: LES FONCTIONS ESTIMÉES $\lambda^i(t)$

Nous présentons dans cette annexe les formes graphiques des fonctions de $\lambda^i(t)$, $i=1,2$, empiriques et estimées par la méthode du MV. Dans ce dernier cas, les paramètres sont donnés aux tableaux 6.3 à 6.6. Lorsque le PNH est adéquat, cette fonction correspond à l'intensité du processus, lorsque le RNH est adéquat, elle correspond à la fonction de changement d'échelle, et elle n'a pas de signification lorsque ni le PNH ni le RNH ne sont adéquats.

Afin d'approximer les fonctions empiriques $\lambda^i(t)$, $i=1,2$ nous avons divisé l'année en 12 intervalles et calculé n_j^i le nombre d'événements de type i observés sur l'intervalle j , $j=1..12$. La quantité empirique $\lambda^i(t)$ correspond à l'intensité journalière moyenne sur 30 jours donnée par $\frac{n_{ij}}{30 * p}$ où p est le nombre d'années sur lesquelles le processus a été observé. Cette approximation suppose que l'intensité est constante sur un intervalle de 30 jours.

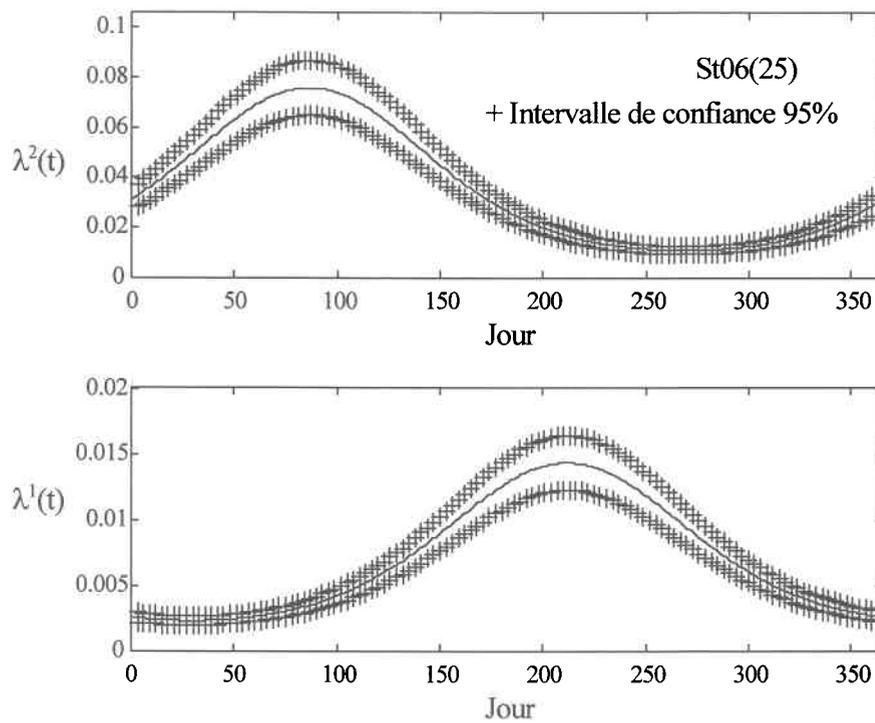


Figure 17.1: Les fonctions estimées par le mv $\lambda^2(t)$ et $\lambda^1(t)$ - St06(25)

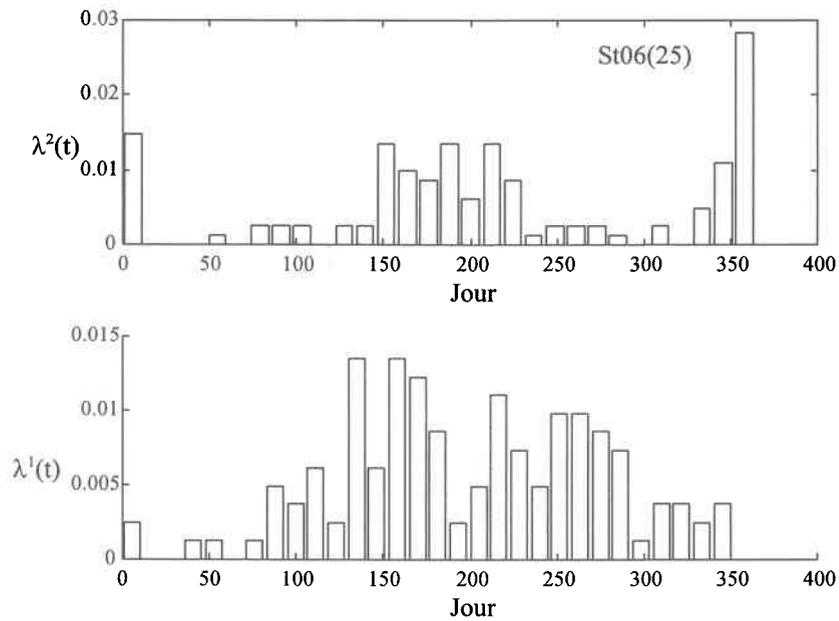
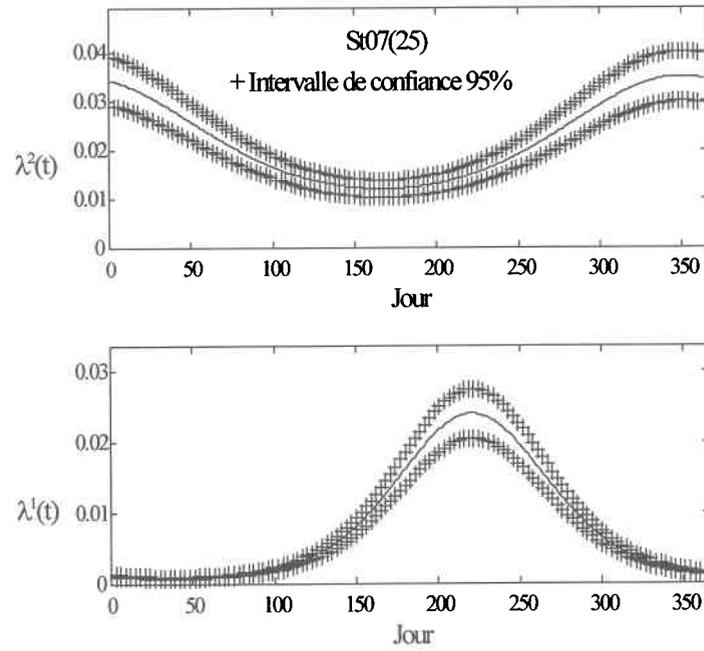
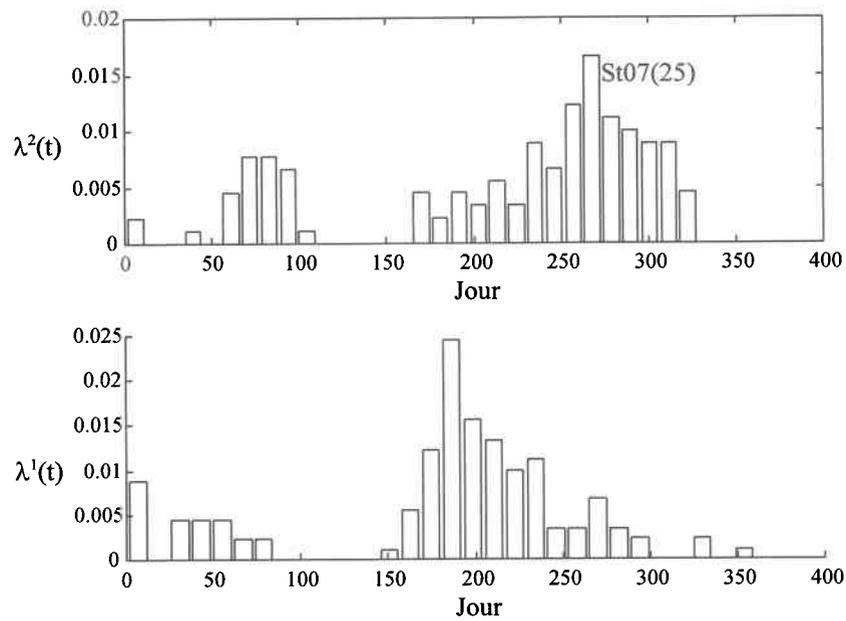
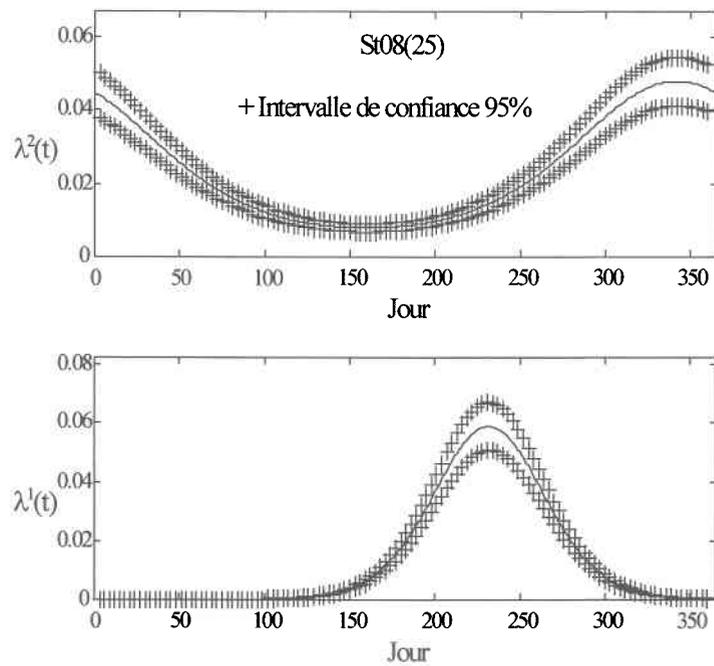
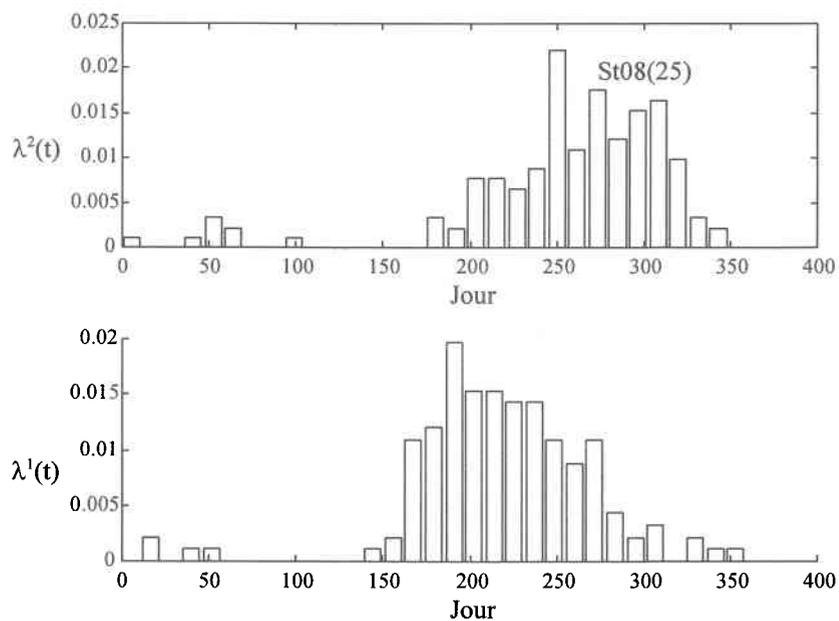
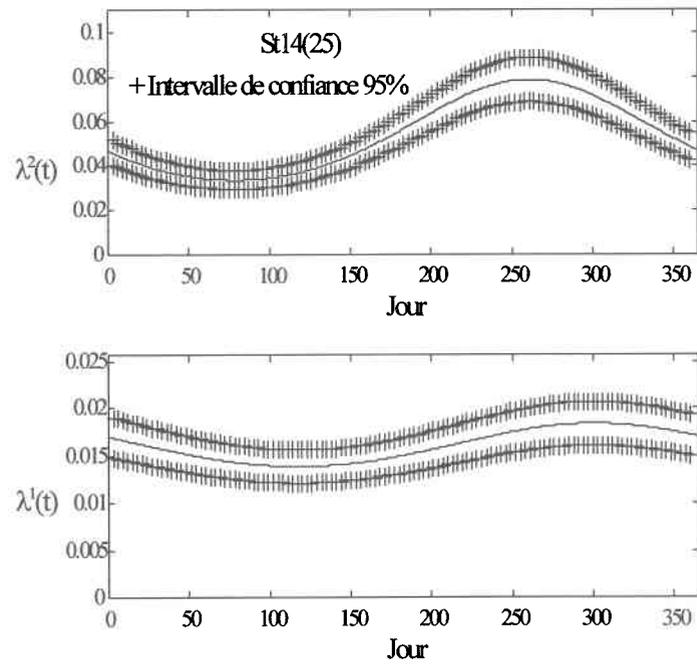
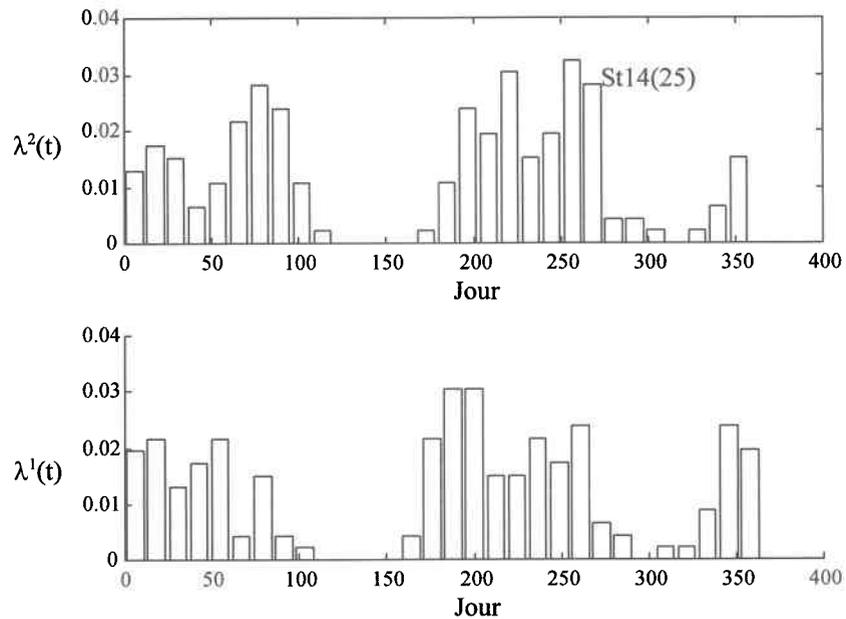
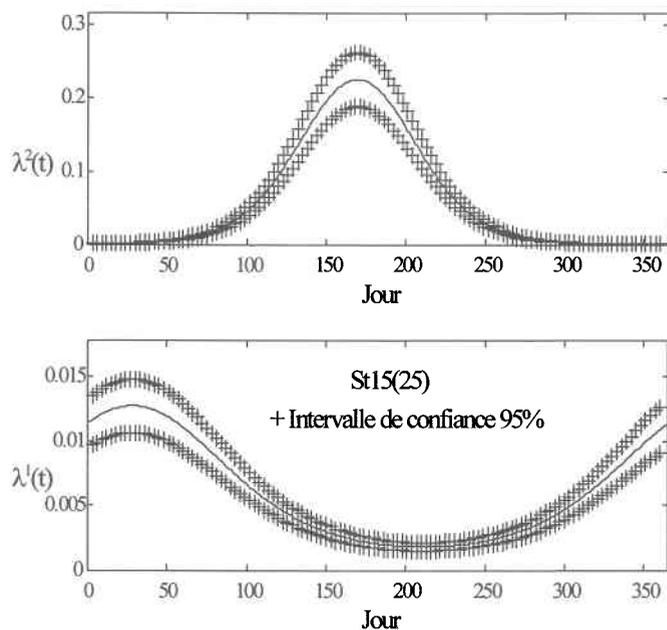
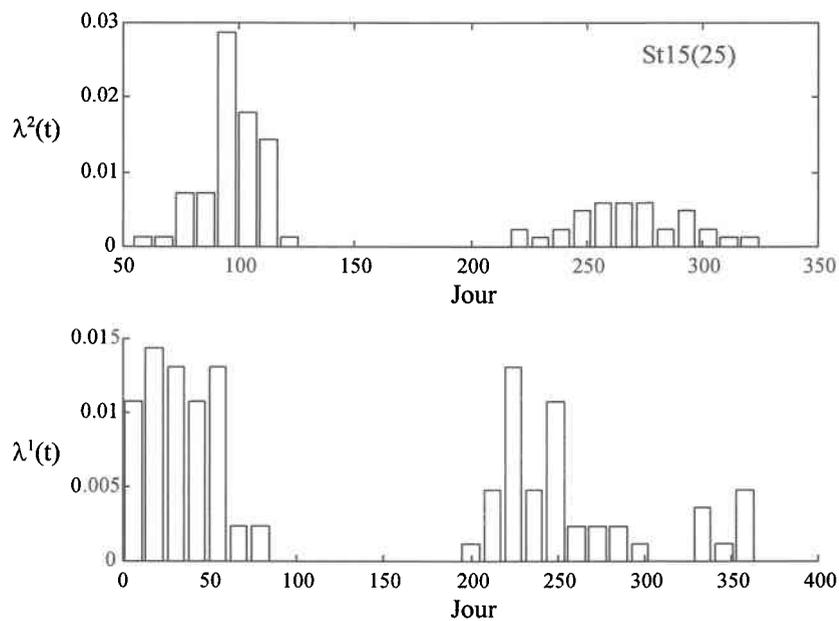


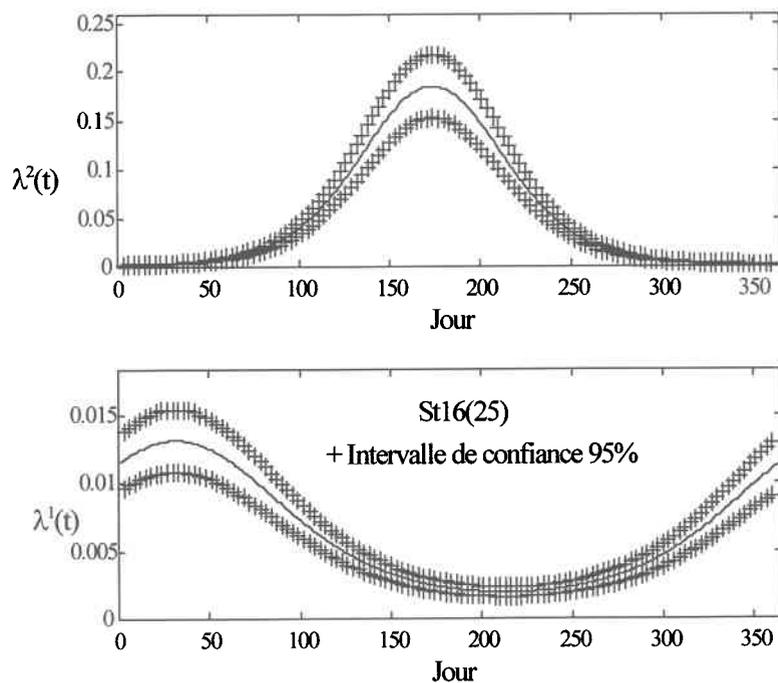
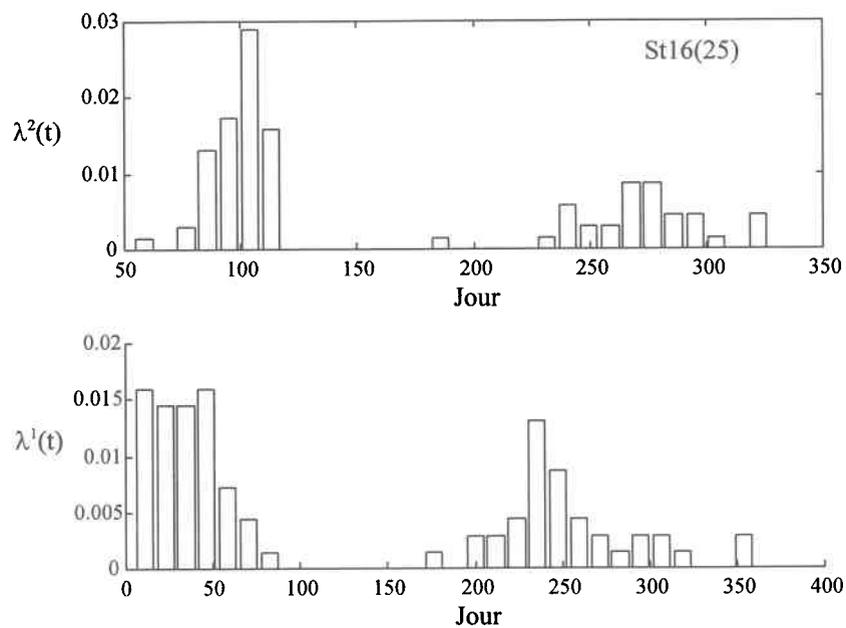
Figure 17.2: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St06(25)

Figure 17.3: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St07(25)Figure 17.4: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St07(25)

Figure 17.5: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St08(25)Figure 17.6: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St08(25)

Figure 17.7: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St14(25)Figure 17.8: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St14(25)

Figure 17.9: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St15(25)Figure 17.10: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St15(25)

Figure 17.11: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St16(25)Figure 17.12: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St16(25)

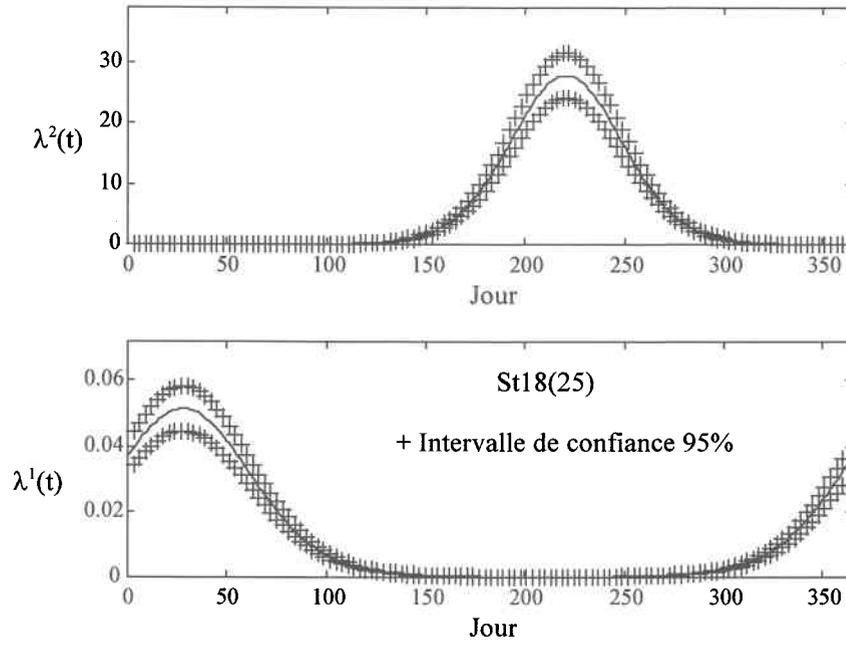


Figure 17.13: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St18(25)

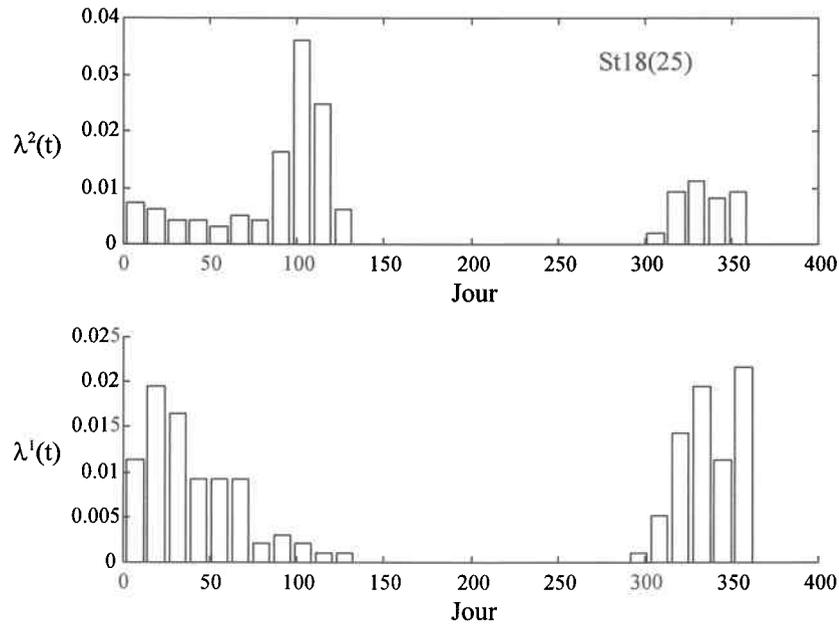
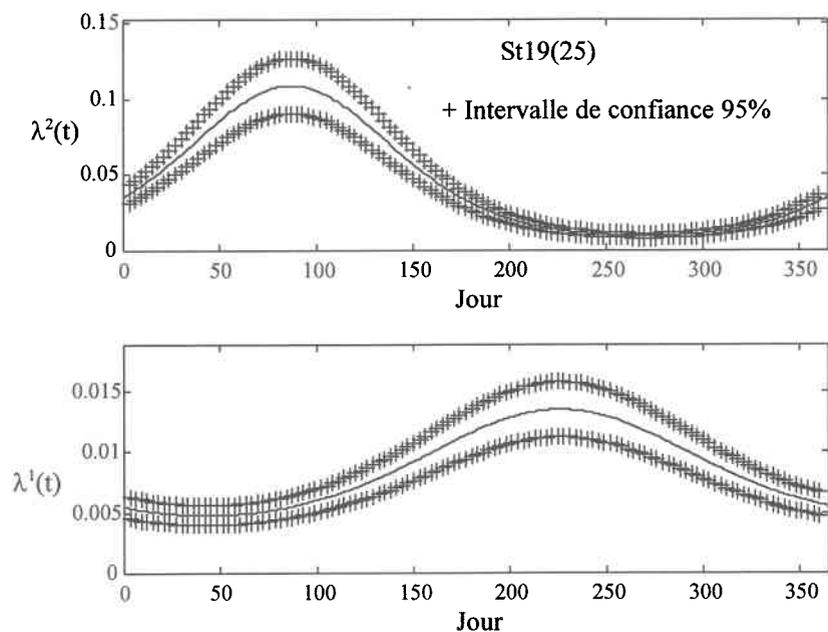
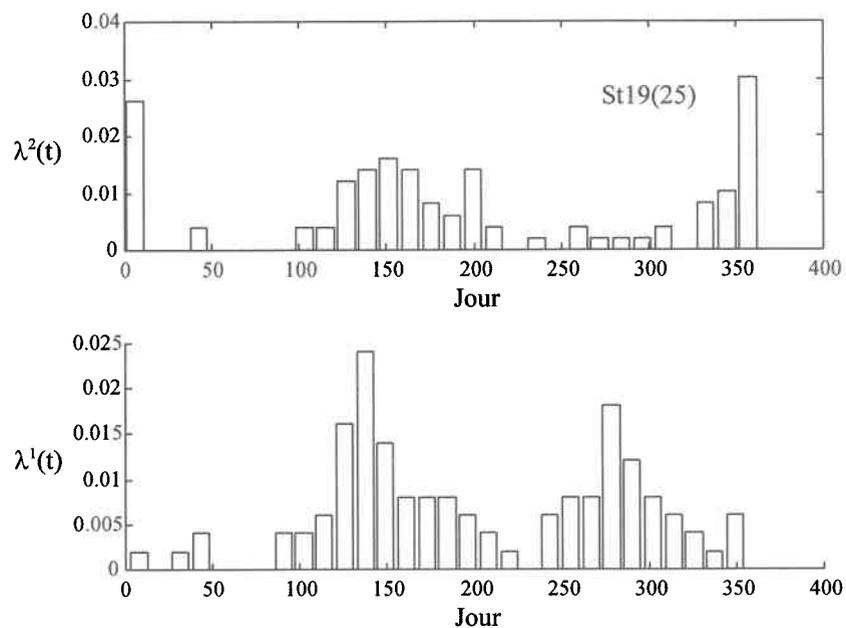


Figure 17.14: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St18(25)

Figure 17.15: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St19(25)Figure 17.16: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St19(25)

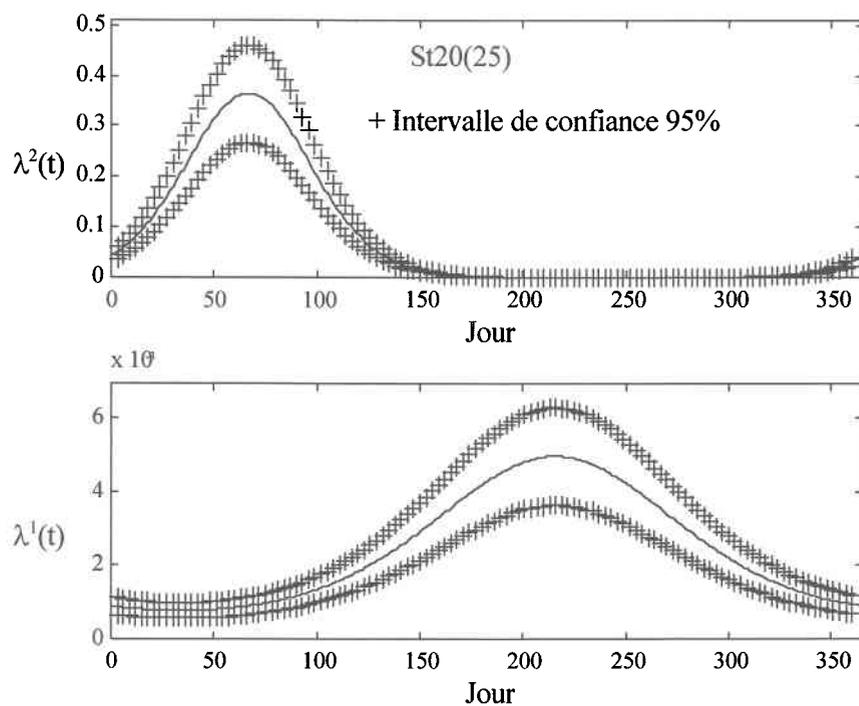


Figure 17.17: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St20(25)

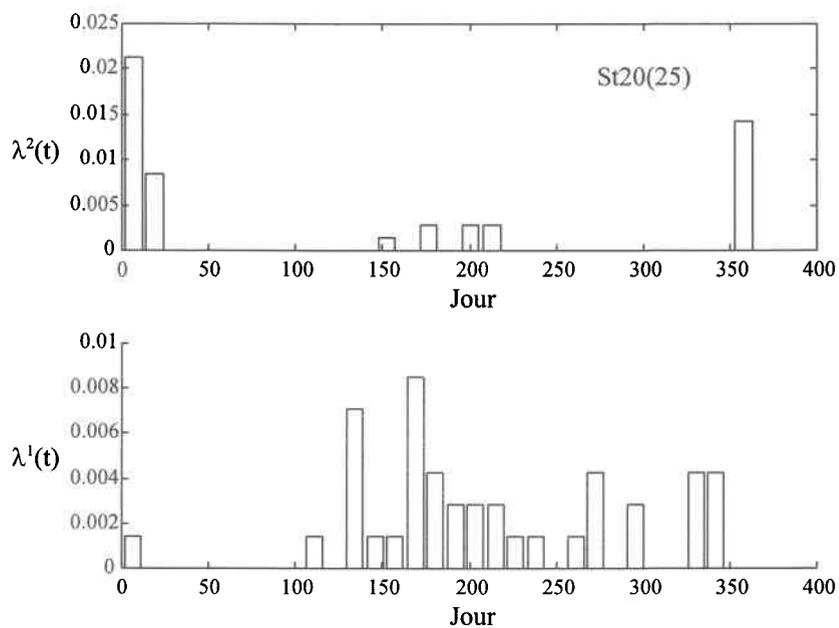
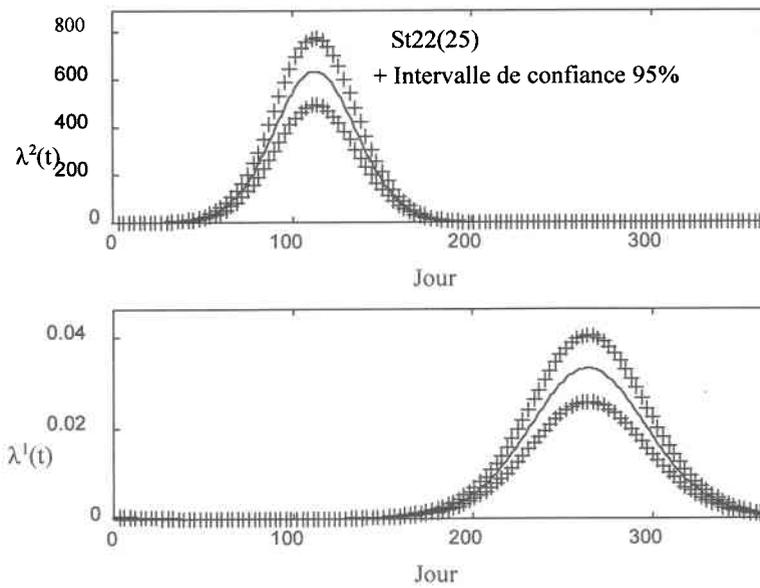
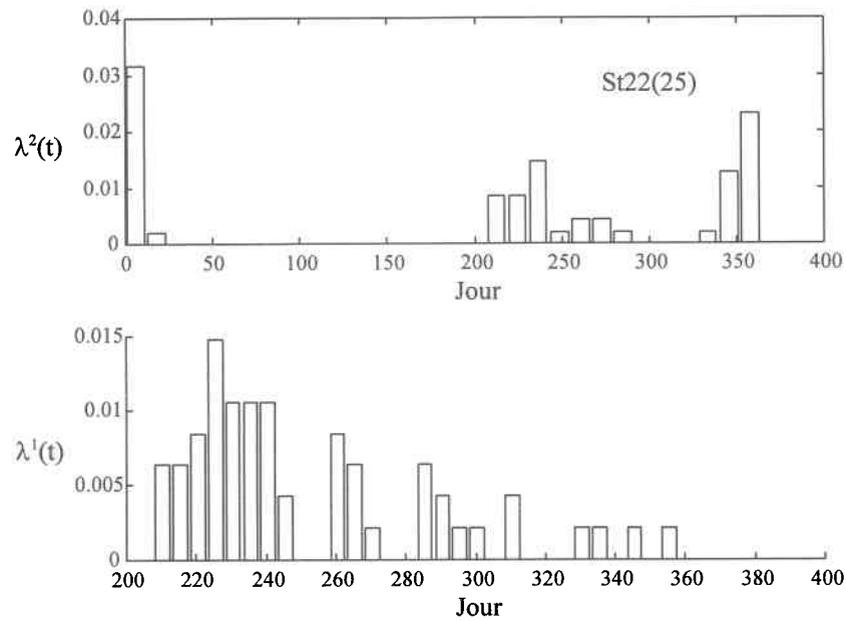


Figure 17.18: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St20(25)

Figure 17.19: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St22(25)Figure 17.20: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St22(25)

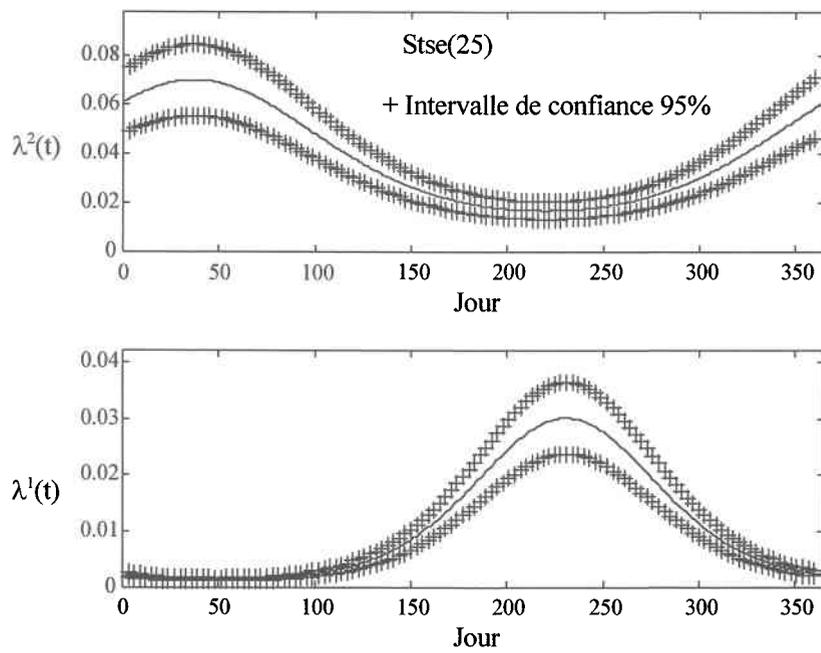


Figure 17.21: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - Stse(25)

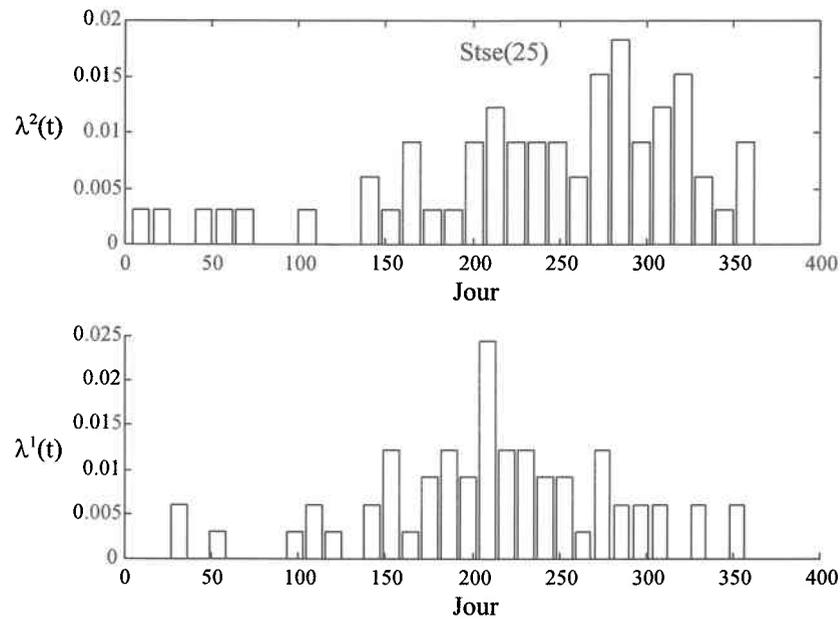
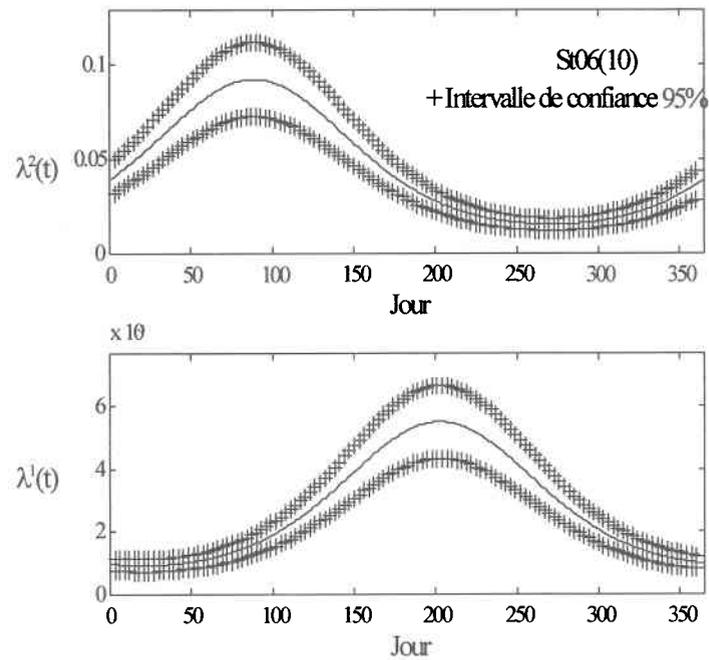
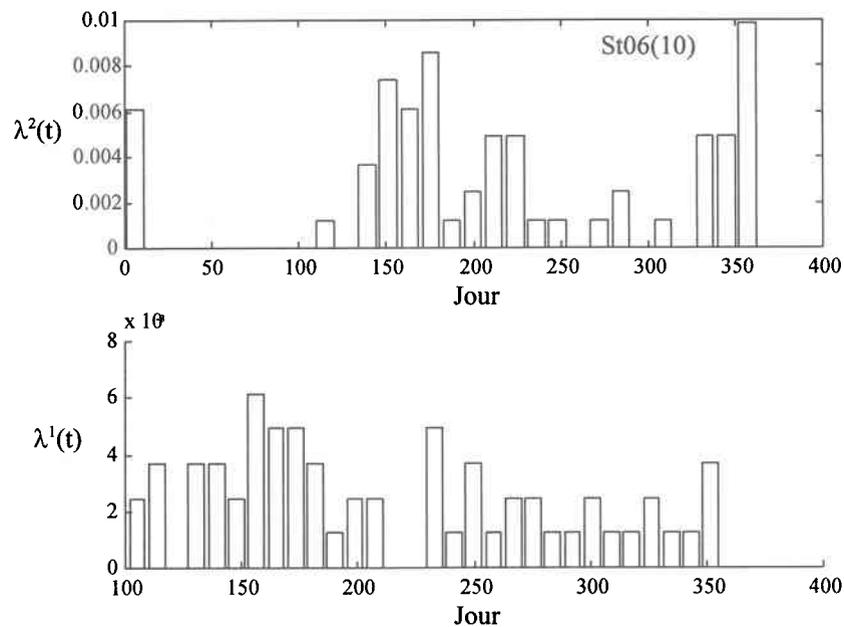


Figure 17.22: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - Stse(25)

Figure 17.23: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St06(10)Figure 17.24: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St06(10)

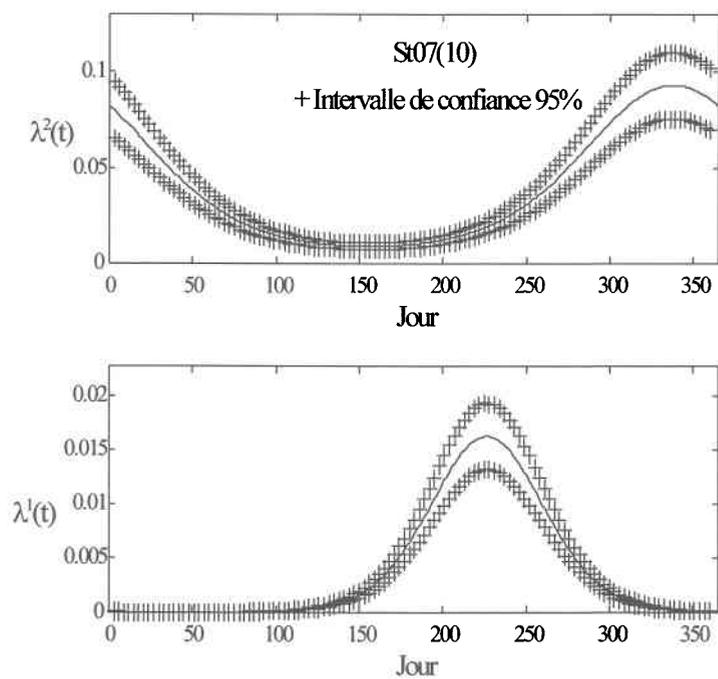


Figure 17.25: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St07(10)

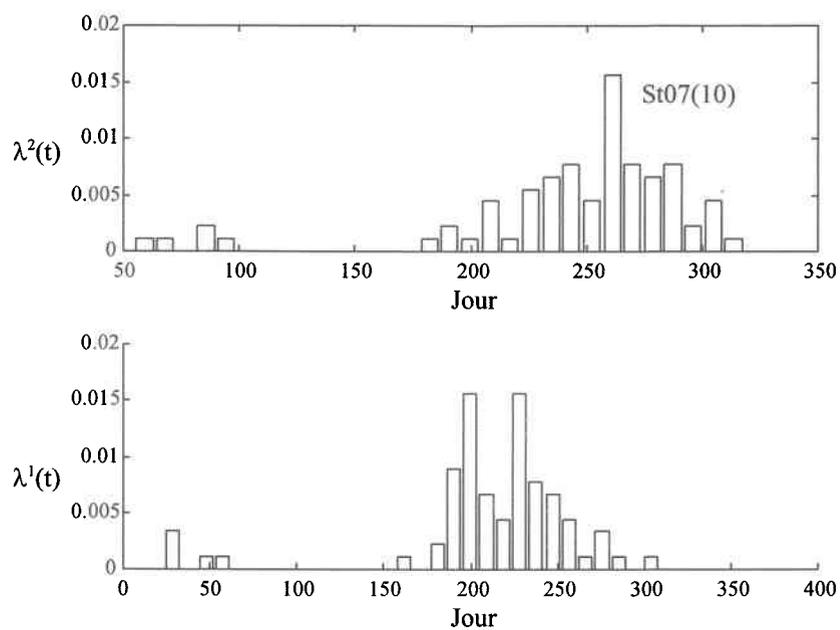
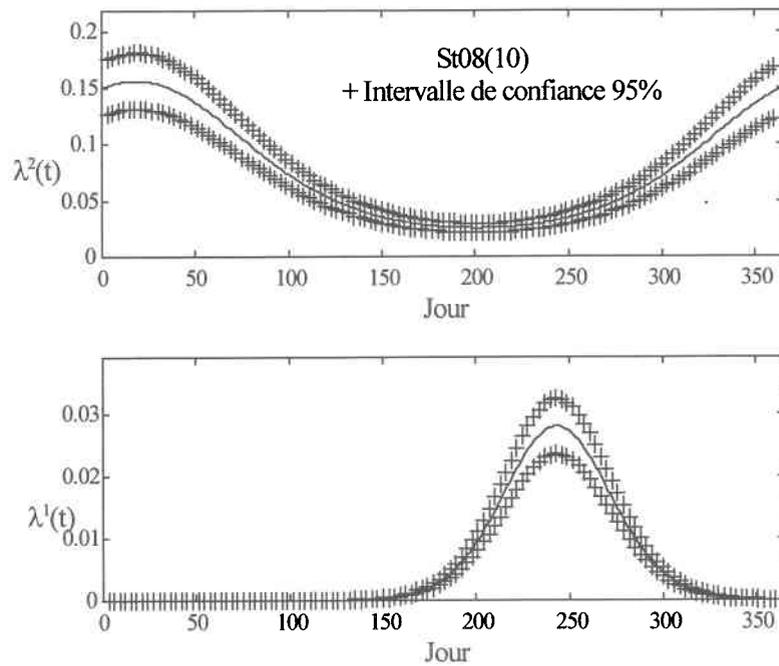
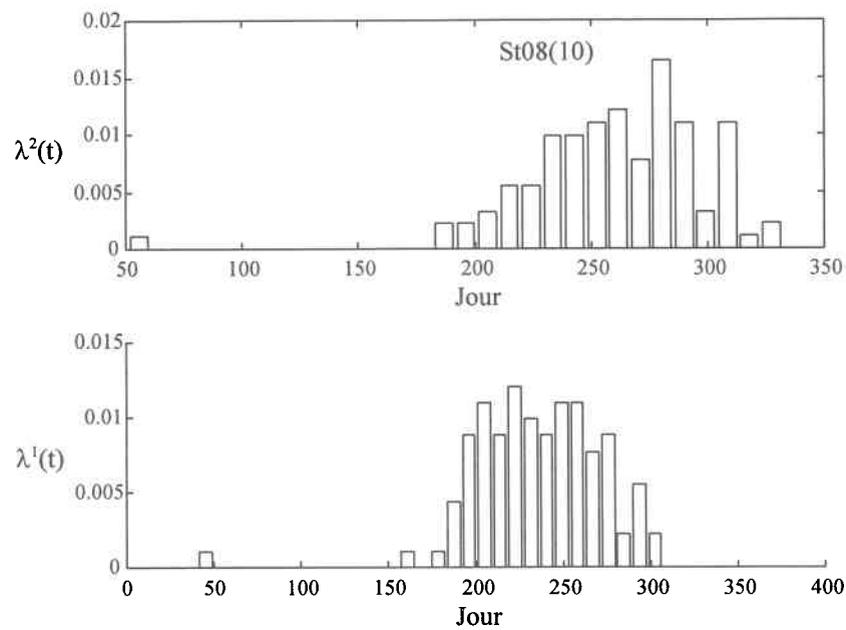
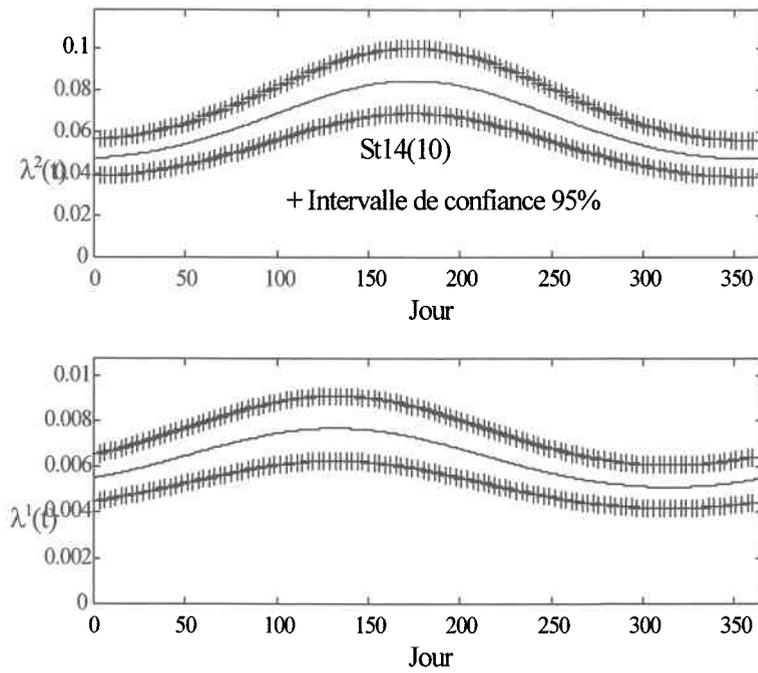
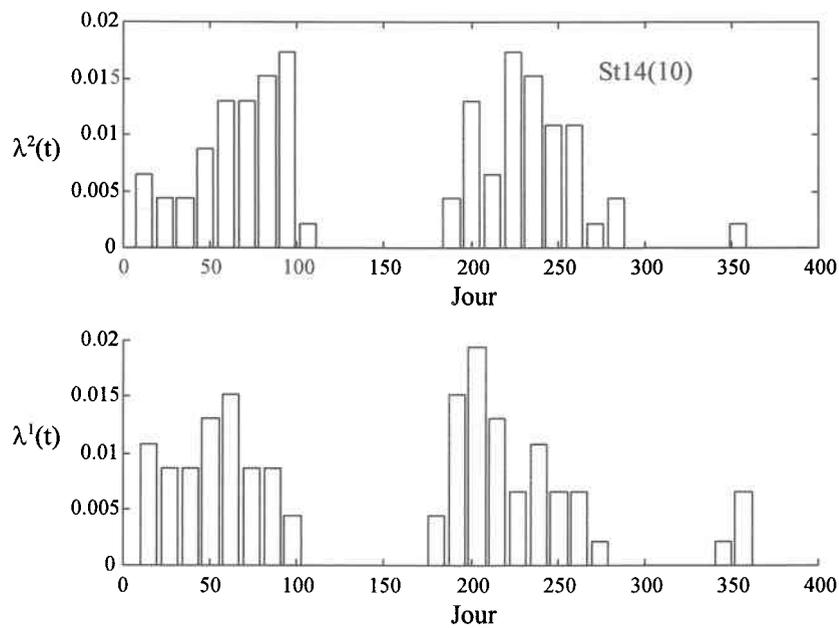
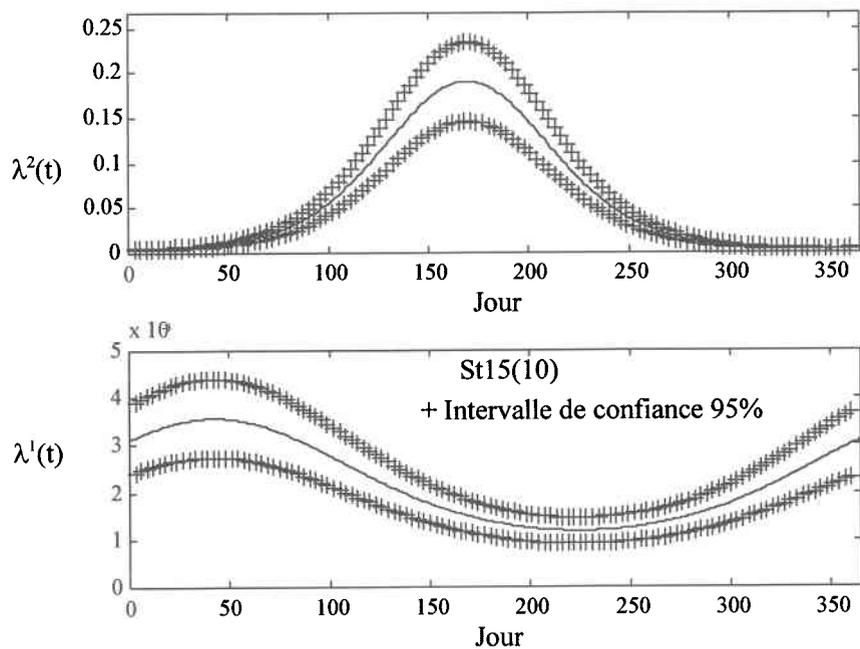
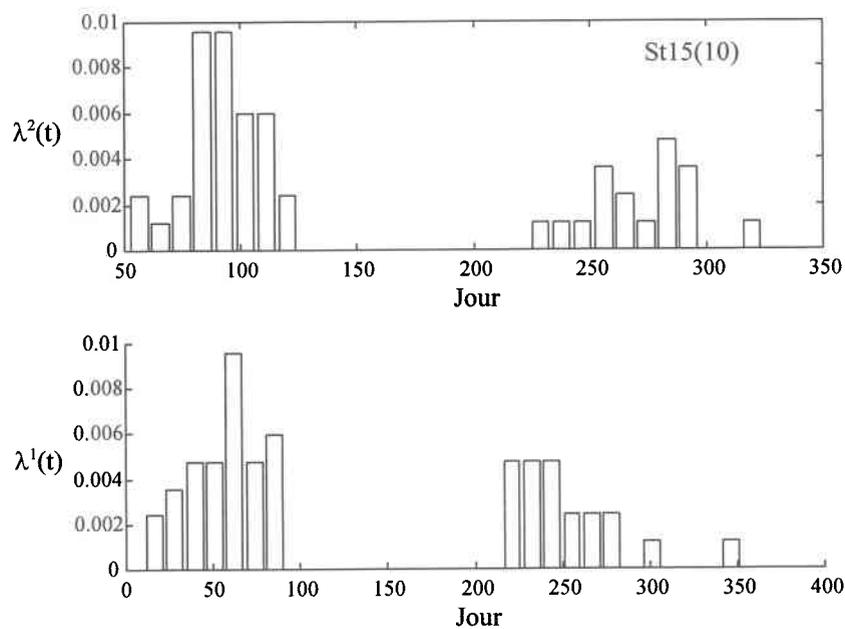
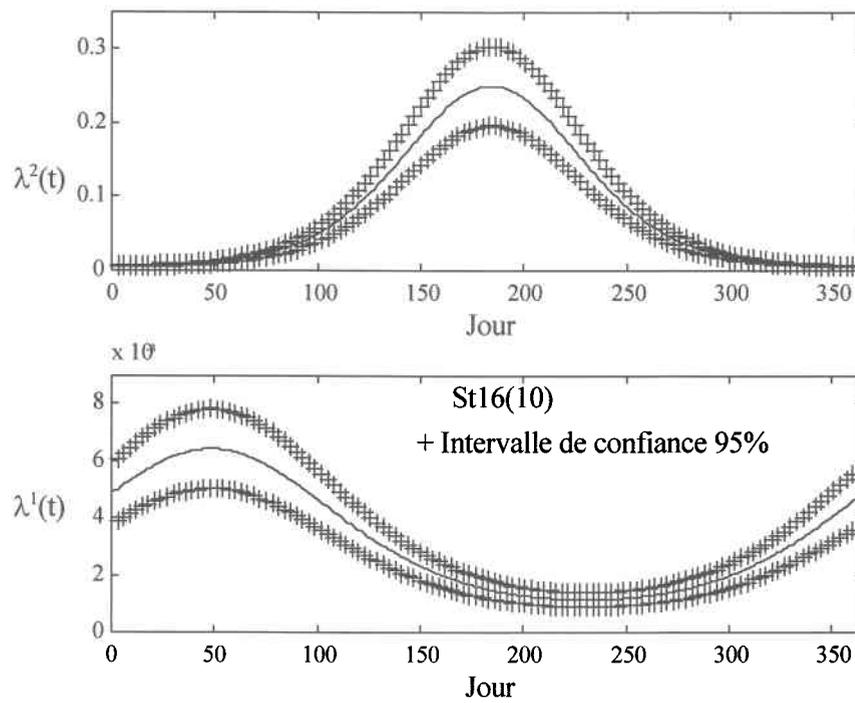
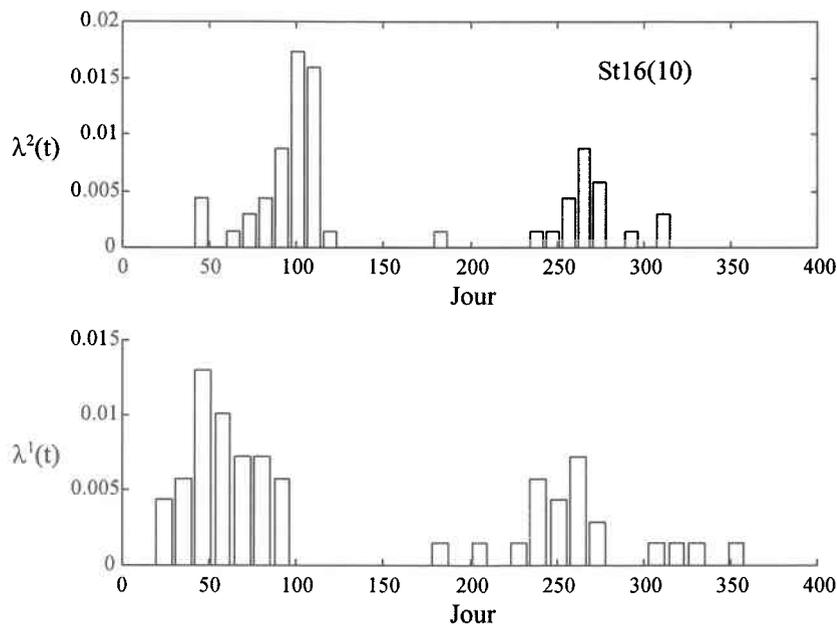


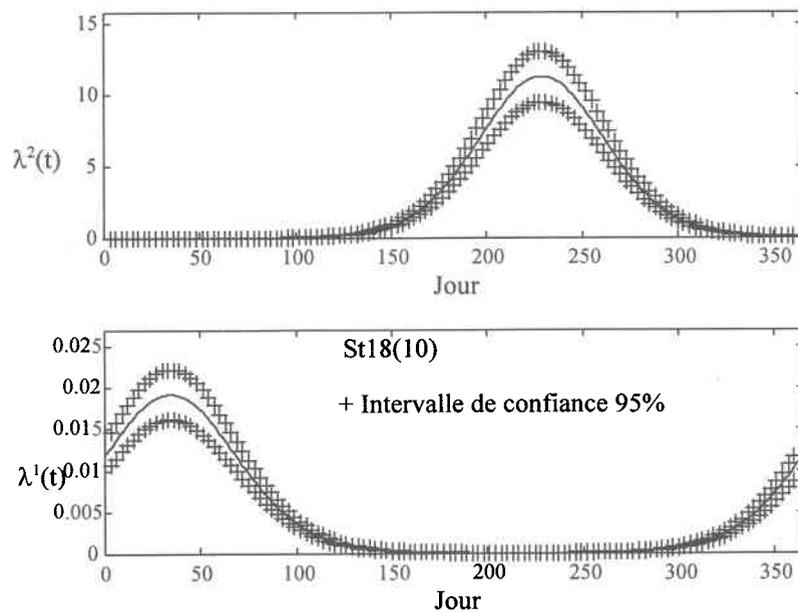
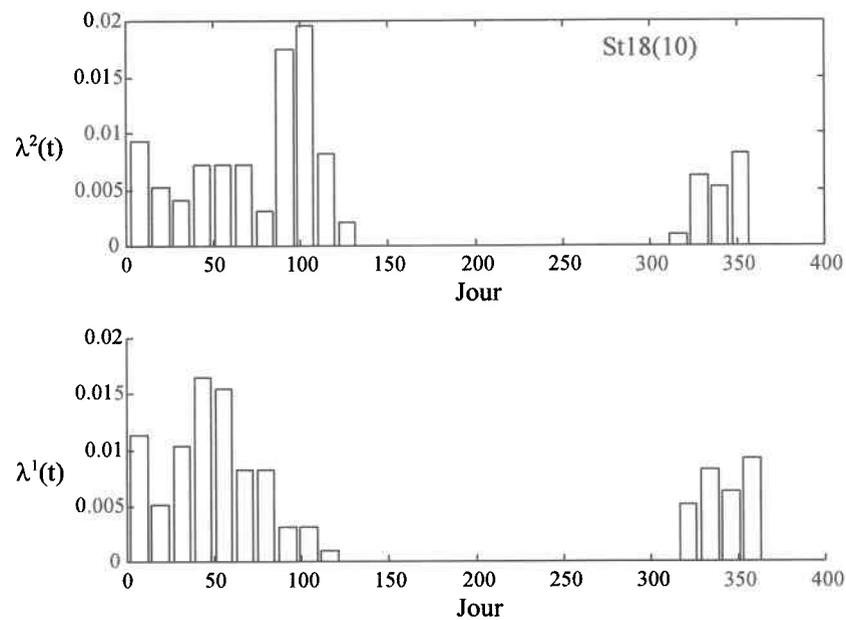
Figure 17.26: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St07(10)

Figure 17.27: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St08(10)Figure 17.28: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St08(10)

Figure 17.29: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St14(10)Figure 17.30: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St14(10)

Figure 17.31: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St15(10)Figure 17.32: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St15(10)

Figure 17.33: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St16(10)Figure 17.34: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St16(10)

Figure 17.35: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St18(10)Figure 17.36: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St18(10)

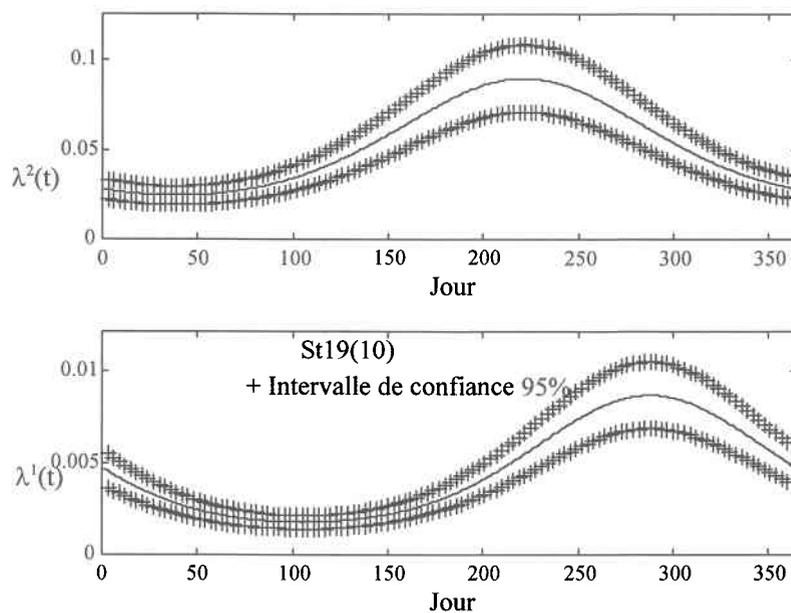


Figure 17.37: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St19(10)

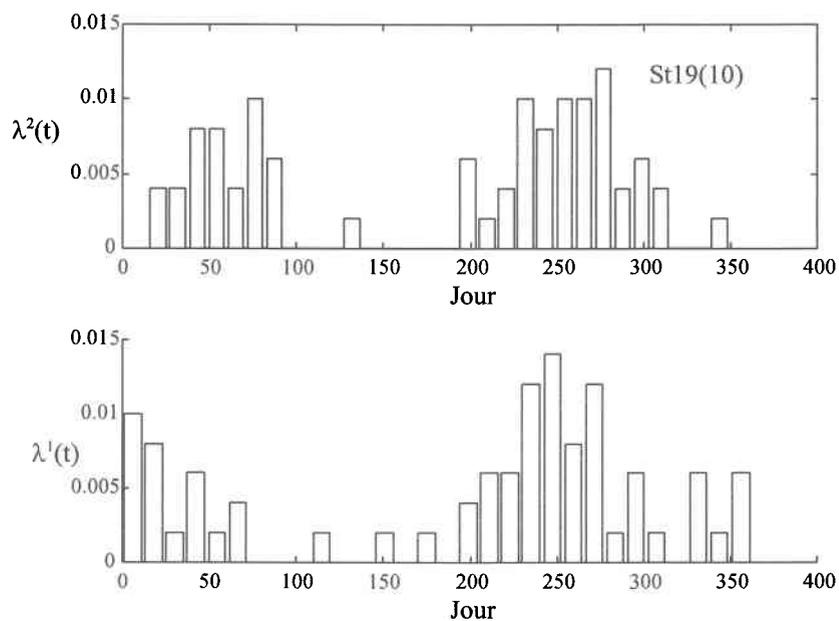
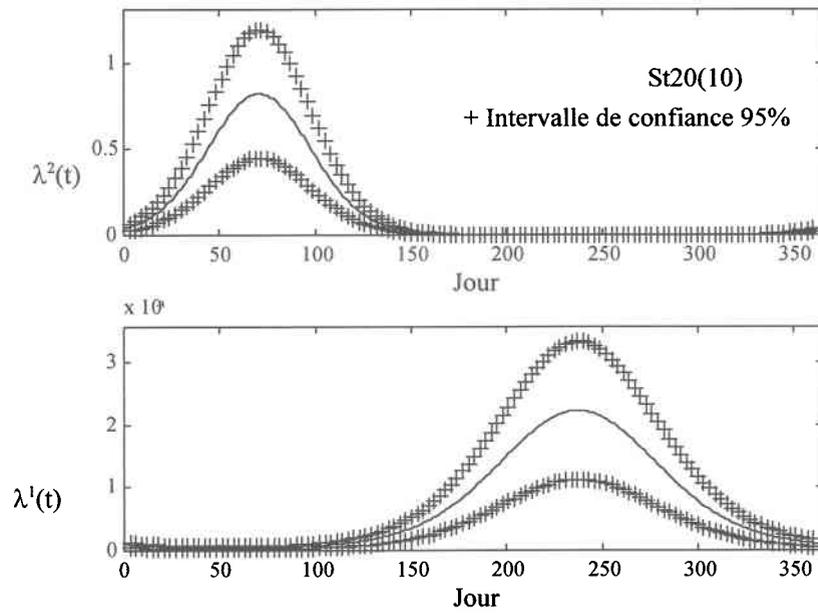
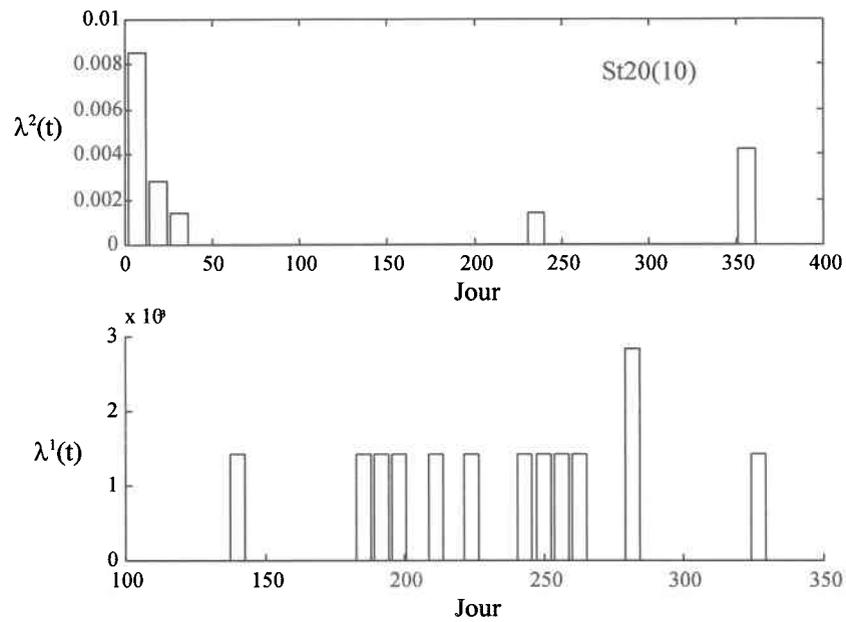
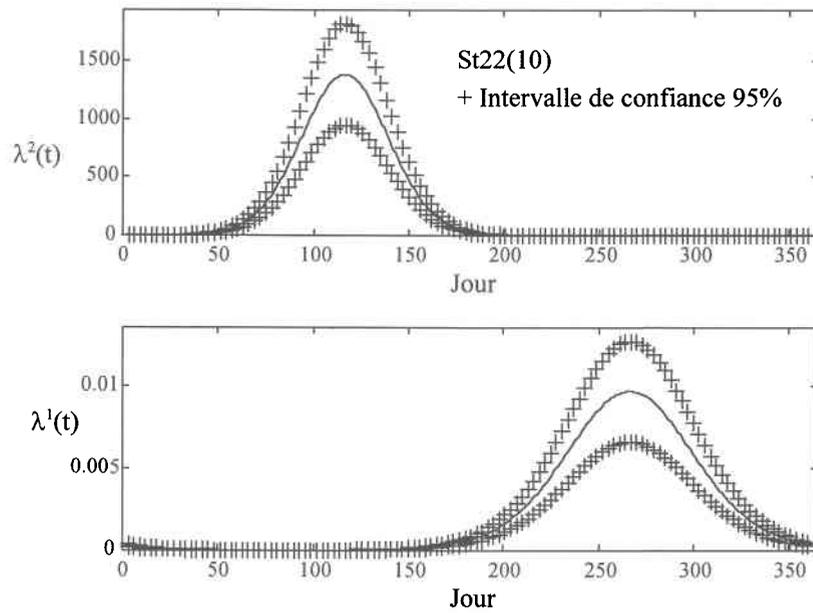
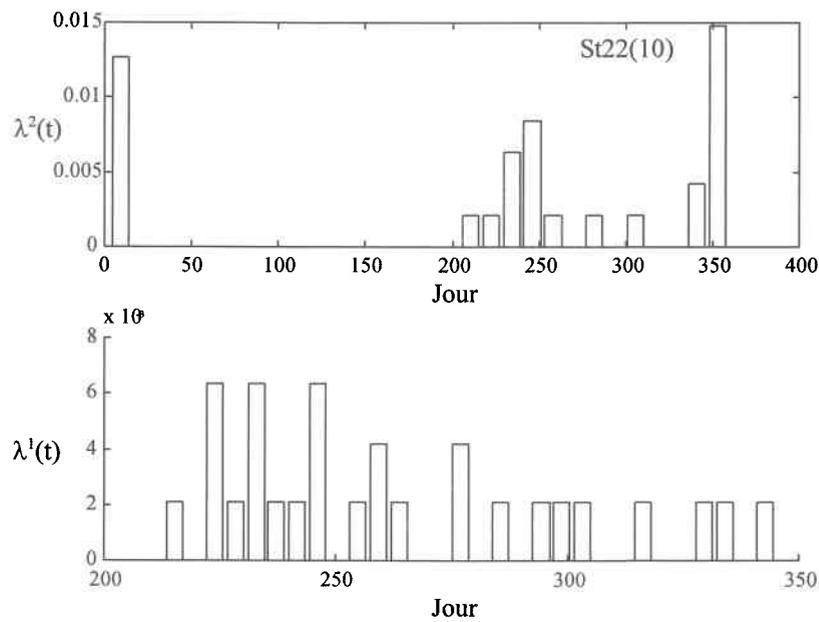
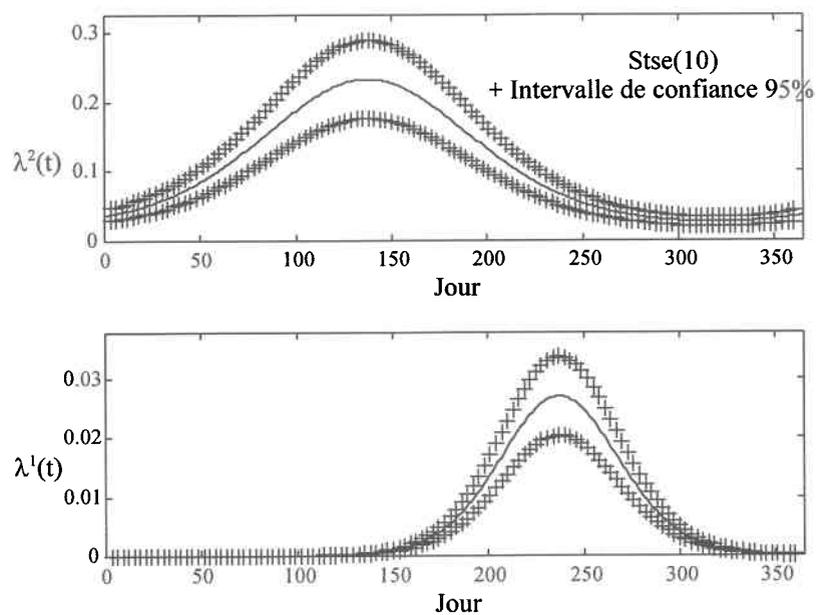
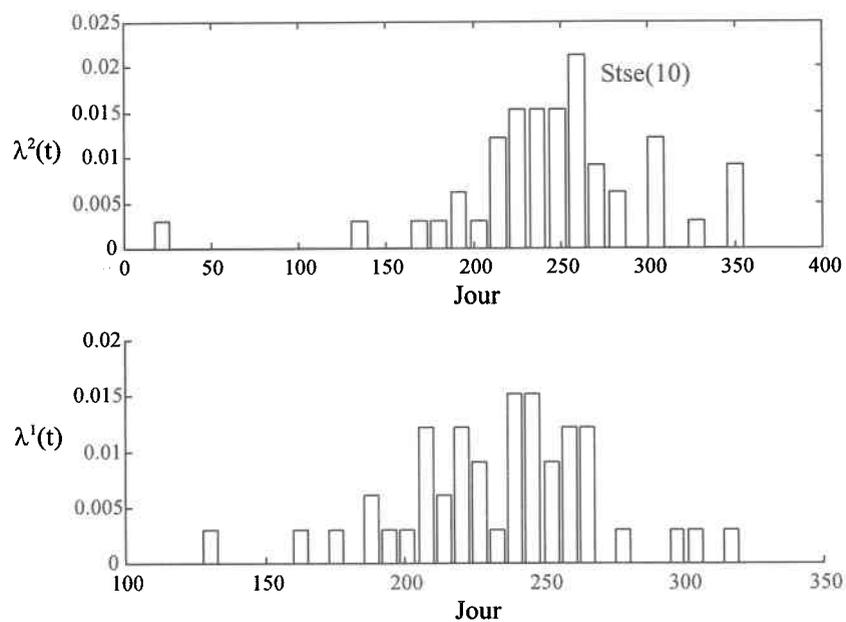


Figure 17.38: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St19(10)

Figure 17.39: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St20(10)Figure 17.40: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St20(10)

Figure 17.41: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - St22(10)Figure 17.42: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - St22(10)

Figure 17.43: Les fonctions estimées $\lambda^2(t)$ et $\lambda^1(t)$ - Stse(10)Figure 17.44: Les fonctions empiriques $\lambda^2(t)$ et $\lambda^1(t)$ - Stse(10)

18. ANNEXE 9: RÉSULTATS À $\alpha=.01$

Nous présentons dans cette annexe les résultats de l'adéquation des modèles PNH, RNH, et PNHC lorsqu'on considère le niveau de confiance $\alpha=.01$. À ce niveau, les valeurs critiques des statistiques de K-S et d'Anderson sont 1.62 et 3.85 respectivement.

En examinant les résultats du tableau 6.7, on constate que le PNH, rejeté à $\alpha=.05$, serait accepté à $\alpha=.01$ comme modèle d'occurrence des étiages des stations suivantes: St08(25), St18(25), St22(25), St08(10), St14(10), St15(10), et St18(10). Or l'examen des figures 6.3, 6.5, 6.7, 6.8, 6.9, 6.11, 6.13, et 6.14 nous révèlent que le PNH n'est pas adéquat pour modéliser les étiages de St08(25), St18(25), St08(10), et St15(10). Les figures 18.1, 18.2, 18.5, et 18.6 nous confirment le rejet du PNH comme modèle d'occurrence des étiages de St18(10) et St22(25). Toutefois, le PNH semble acceptable pour modéliser les étiages de St14(10) (figures 18.3 et 18.4).

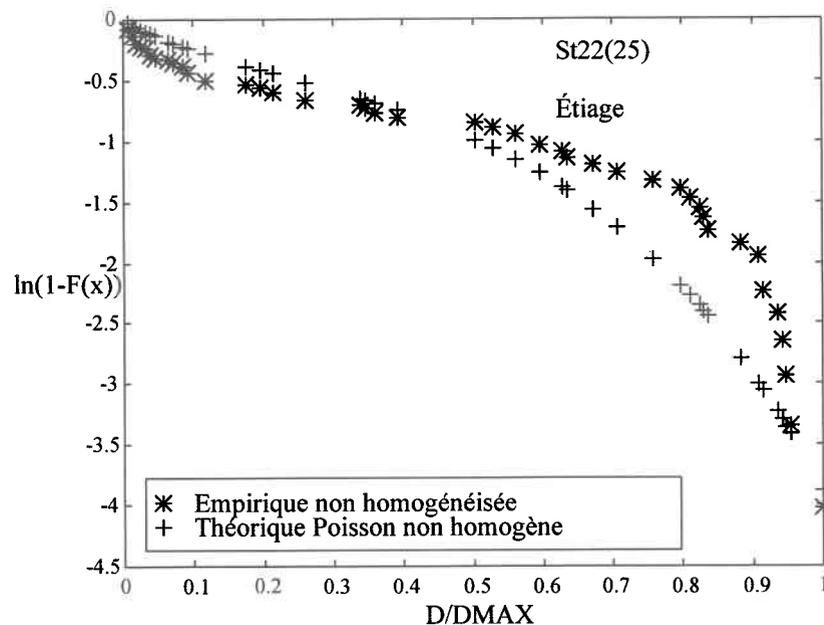


Figure 18.1: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St22(25)

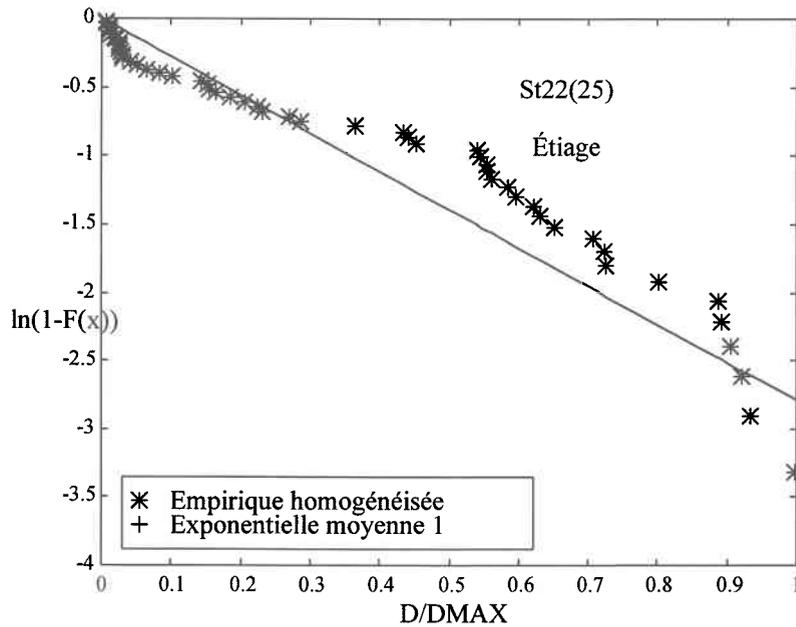


Figure 18.2: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St22(25)

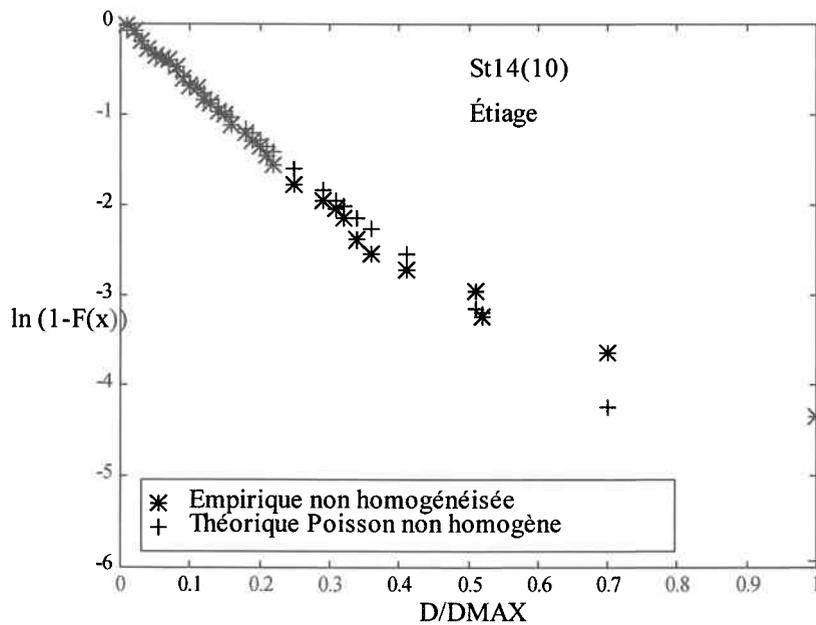


Figure 18.3: Les fonctions de survie logarithmiques des durées d'étiage empiriques non homogénéisées et des durées théoriques d'un PNH, St14(10)

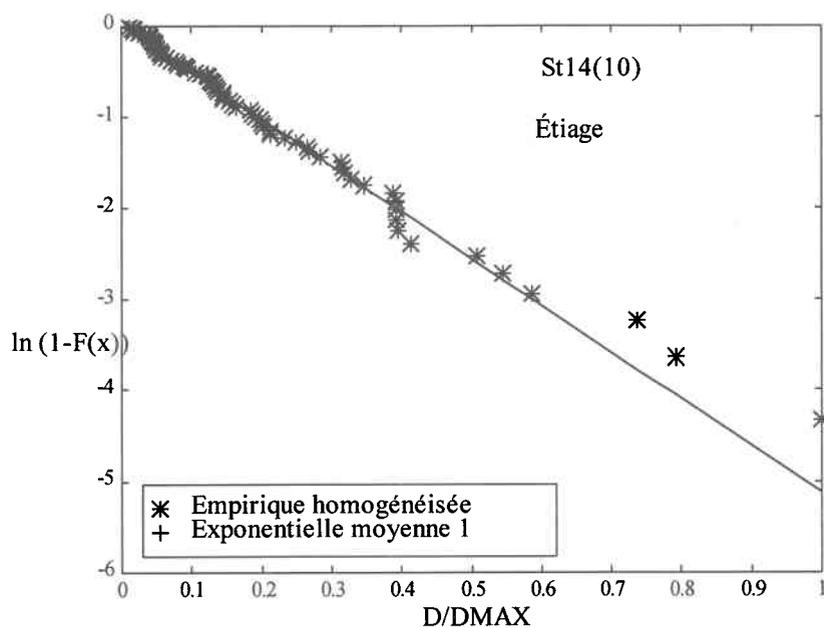


Figure 18.4: Les fonctions de survie logarithmiques des durées d'étéiage empiriques homogénéisées et des durées théoriques d'un PH, St14(10)

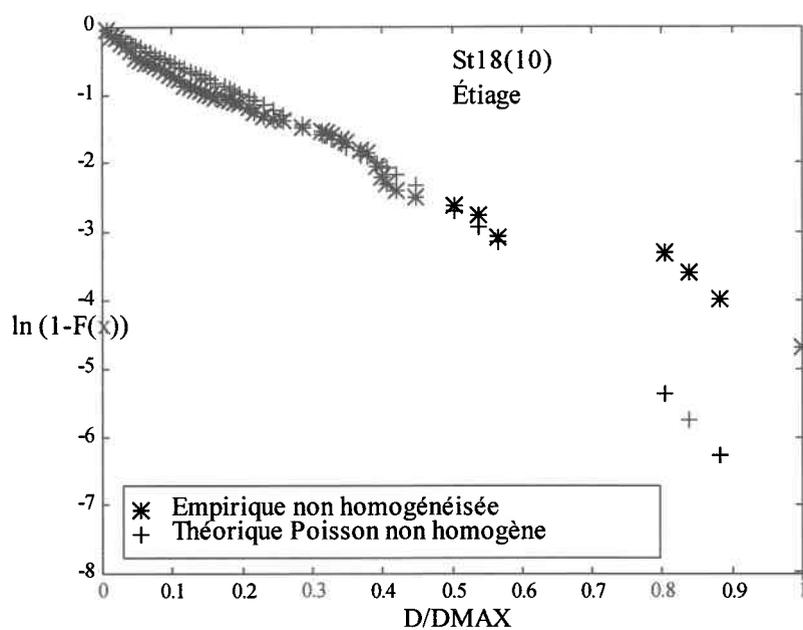


Figure 18.5: Les fonctions de survie logarithmiques des durées d'étéiage empiriques non homogénéisées et des durées théoriques d'un PNH, St18(10)

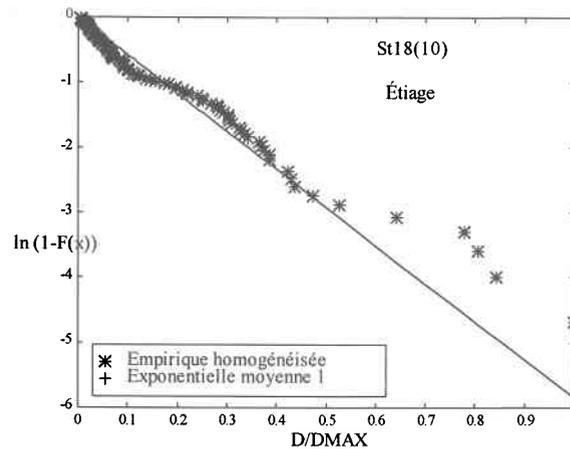


Figure 18.6: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées et des durées théoriques d'un PH, St18(10)

En examinant les résultats du tableau 6.8 sur l'adéquation du PNH comme modèle d'occurrence des surplus, on constate que le PNH, rejeté à $\alpha=.05$, serait accepté à $\alpha=.01$ pour modéliser les surplus des stations suivantes: St14(25) et St18(10). Or l'examen des figures 6.15 et 6.17 nous révèlent que le PNH n'est pas adéquat pour modéliser les surplus de St14(25). Les figures 18.7 et 18.8 nous confirment le rejet du PNH comme modèle d'occurrence des surplus de St18(10).

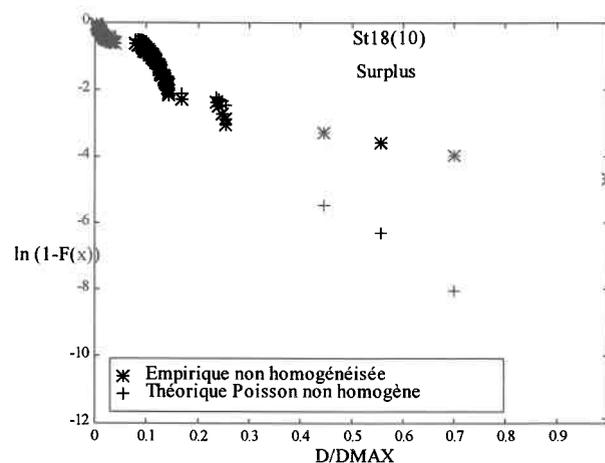


Figure 18.7: Les fonctions de survie logarithmiques des durées de surplus empiriques non homogénéisées et des durées théoriques d'un PNH, St18(10)

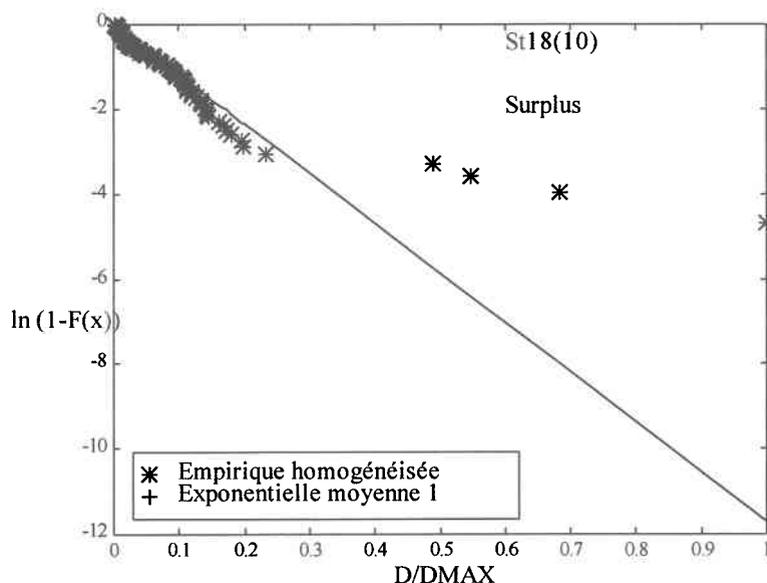


Figure 18.8: Les fonctions de survie logarithmiques des durées de surplus empiriques homogénéisées et des durées théoriques d'un PH, St18(10)

Les résultats de l'adéquation du RNH aux processus d'occurrence des étiages (tableau 6.9) et du PNHC aux processus d'occurrence des surplus (tableau 6.15) demeurent inchangés pour $\alpha=.01$. Toutefois, à ce niveau de confiance, le RNH semble acceptable pour modéliser les surplus de St14(25) (tableau 6.10) et le PNHC pour modéliser les étiages de Stse(10) (tableau 6.13).

Si le RNH est adéquat pour modéliser les surplus de St14(25), alors la fonction de survie logarithmique des durées empiriques homogénéisées par la fonction d'intensité estimée $\hat{\lambda}^1(t)$, devrait correspondre à celle des durées provenant d'une loi Gamma de moyenne unitaire et de paramètres $\hat{\alpha}_1 = \hat{\rho}_1 = \frac{1}{\hat{\sigma}_1^2}$, où $\hat{\sigma}_1^2$ est la variance des durées de surplus empiriques homogénéisées (égale à 1.08 pour St14(25)). La figure 18.9 présente la comparaison entre ces deux courbes. On en déduit qu'elles sont dissimilaires, et que l'hypothèse du RNH comme modèle d'occurrence des surplus de St14(25) est rejetée.

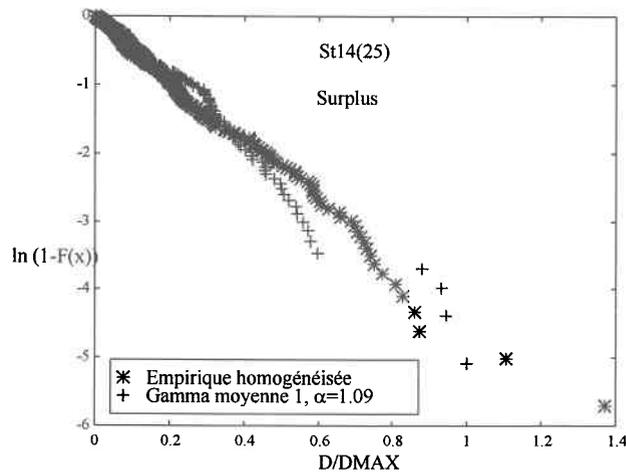


Figure 18.9: Les fonctions de survie logarithmiques des durées de surplus empiriques homogénéisées (RNH) et des durées théoriques d'un RH, St14(25)

Si le PNHC est adéquat pour modéliser les étiages de Stse(10), alors le processus obtenu après homogénéisation par la fonction d'intensité du PNHC devrait être un PH (sections 3.3.1 et 3.3.2). Or la figure 18.10, sur laquelle on compare la fonction de survie empirique des durées d'étiage homogénéisées par la fonction d'intensité du PNHC, et la fonction de survie théorique d'un PH rejette le PNHC comme modèle pour les étiages de Stse(10).

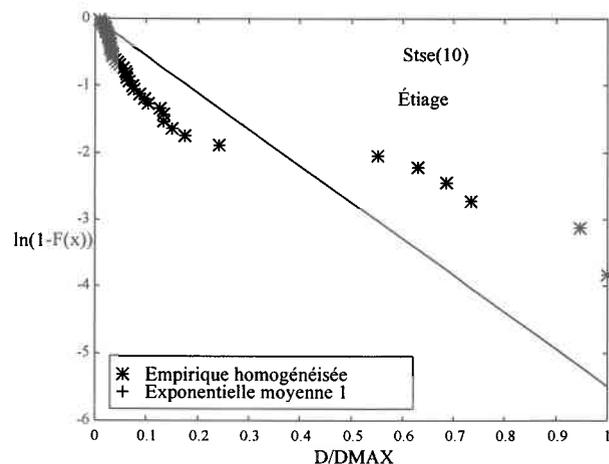


Figure 18.10: Les fonctions de survie logarithmiques des durées d'étiage empiriques homogénéisées (PNHC) et des durées théoriques d'un PH, Stse(10)

En résumé, les résultats de l'adéquation des modèles PNH, RNH et PNHC aux occurrences des surplus et étiages des 22 stations (rivière/seuil) étudiées sont robustes par rapport au niveau de confiance α choisi. Ces résultats sont les mêmes à $\alpha=.01$ qu'à $\alpha=.05$ à l'exception d'une station: les étiages de St14(10) (PNH).

19. ANNEXE 10: LES COURBES DÉBITS-DURÉES

Nous présentons dans cette annexe les courbes de fréquence cumulative débit-durée des rivières étudiées. Ces courbes montrent le pourcentage de temps durant lequel certains débits sont dépassés pendant la période couverte par l'échantillon.

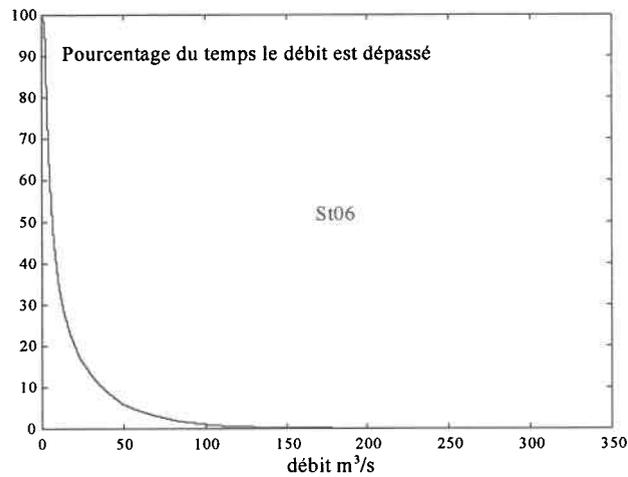


Figure 19.1: La courbe débit-durée - St06

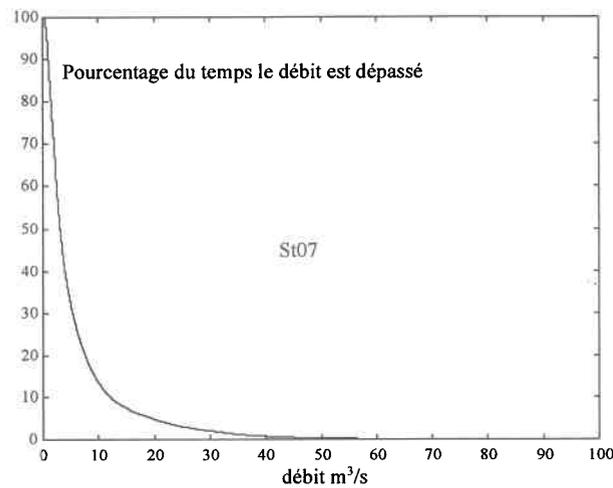


Figure 19.2: La courbe débit-durée - St07

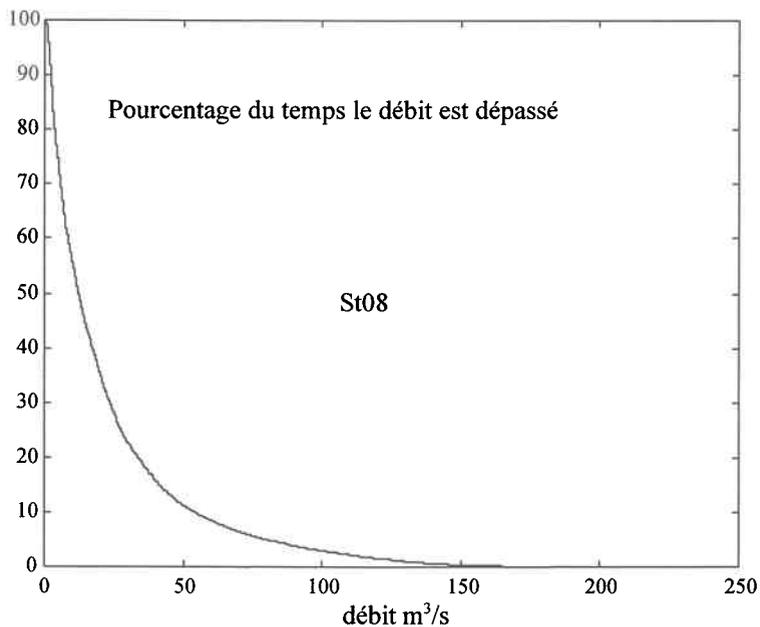


Figure 19.3: La courbe débit-durée - St08

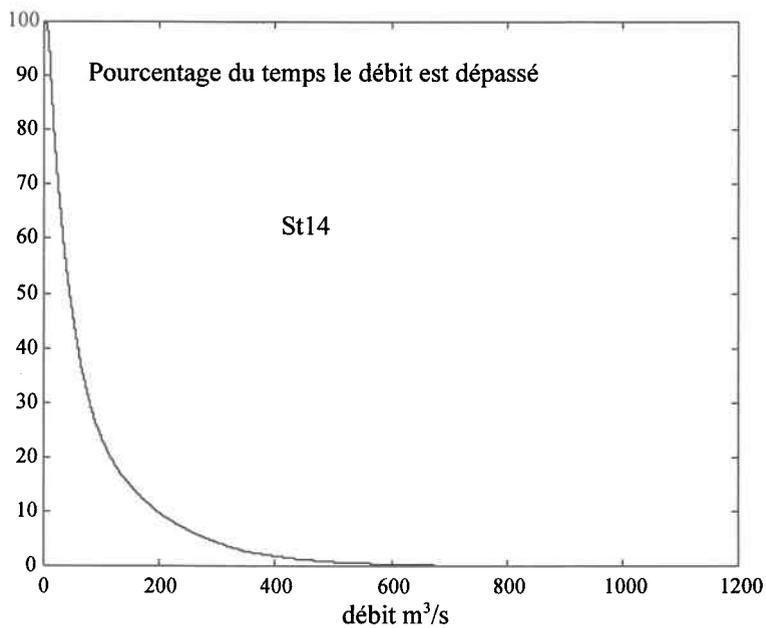


Figure 19.4: La courbe débit-durée - St14

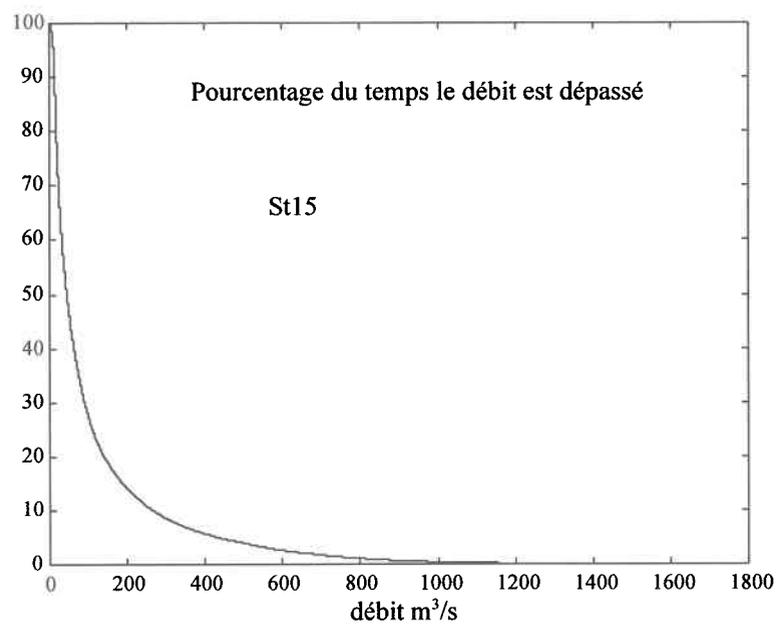


Figure 19.5: La courbe débit-durée - St15

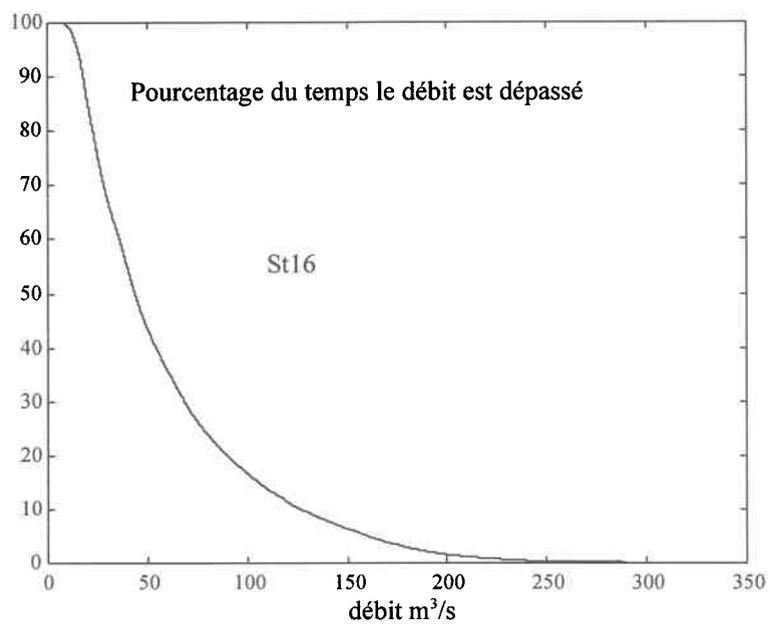


Figure 19.6: La courbe débit-durée - St16

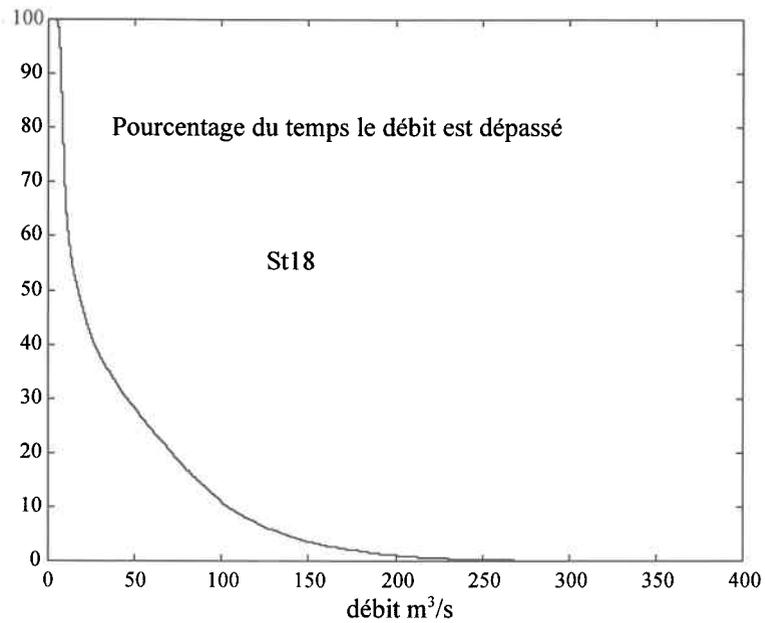


Figure 19.7: La courbe débit-durée - St18

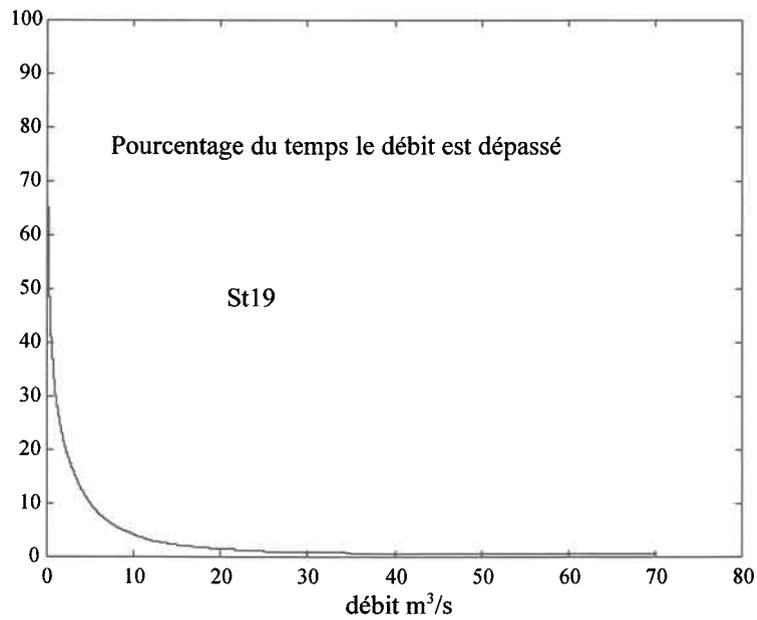


Figure 19.8: La courbe débit-durée - St19

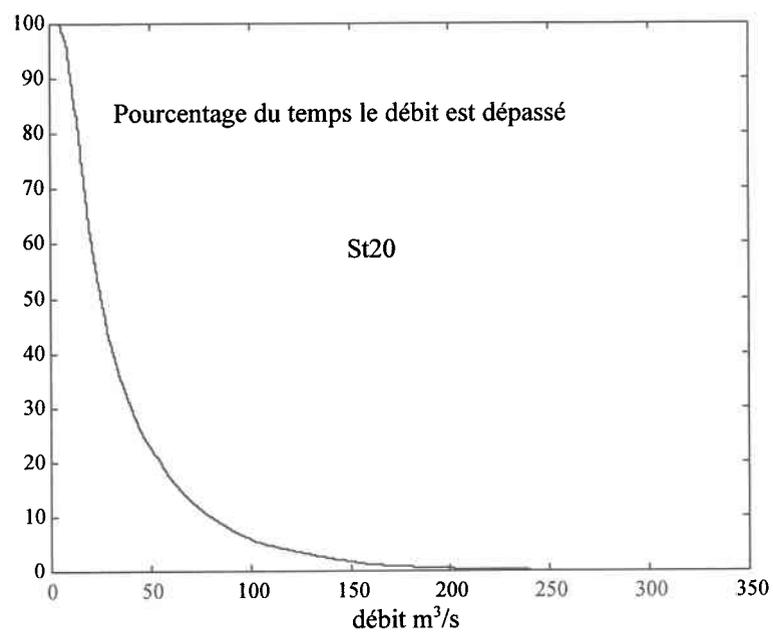


Figure 19.9: La courbe débit-durée - St20

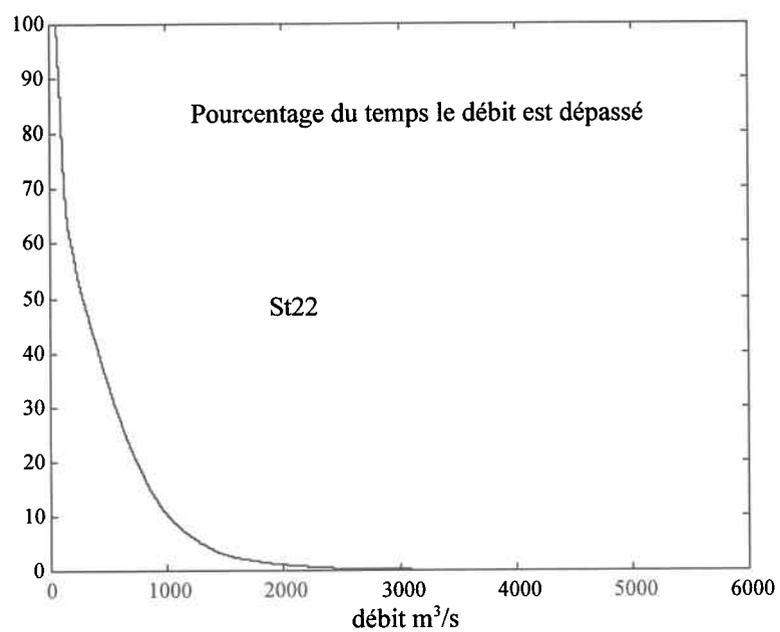


Figure 19.10: La courbe débit-durée - St22

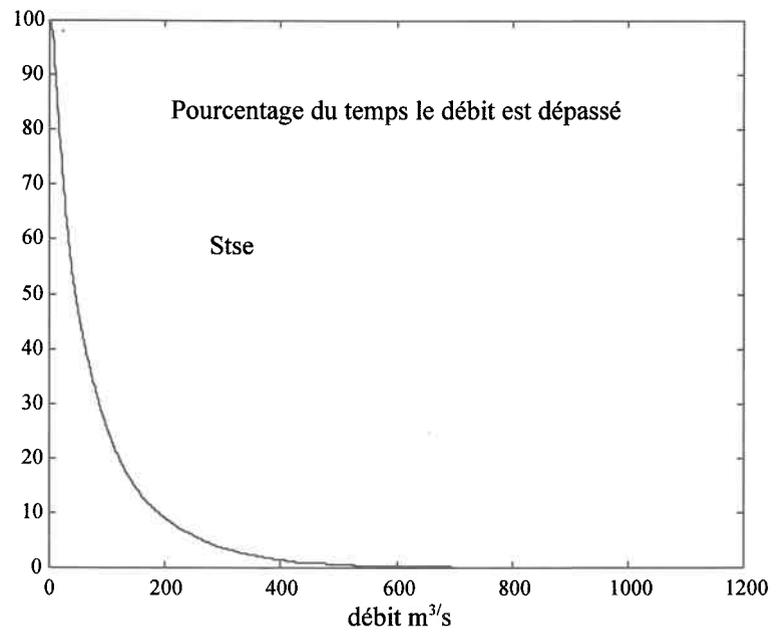


Figure 19.11: La courbe débit-durée - Stse