Université du Québec INRS-Eau

Évaluation de la performance de modèles ponctuels de précipitation à l'aide des courbes IDF.

Par Dominic Grenier

Baccalauréat en mathématiques (statistiques)

Mémoire présenté pour l'obtention du grade de Maître ès science (M.Sc.)

Jury d'évaluation

Examinateur externe

Christian Onof Département de génie civil Imperial College, London

Examinateur interne

Daniel Cluis INRS-Eau

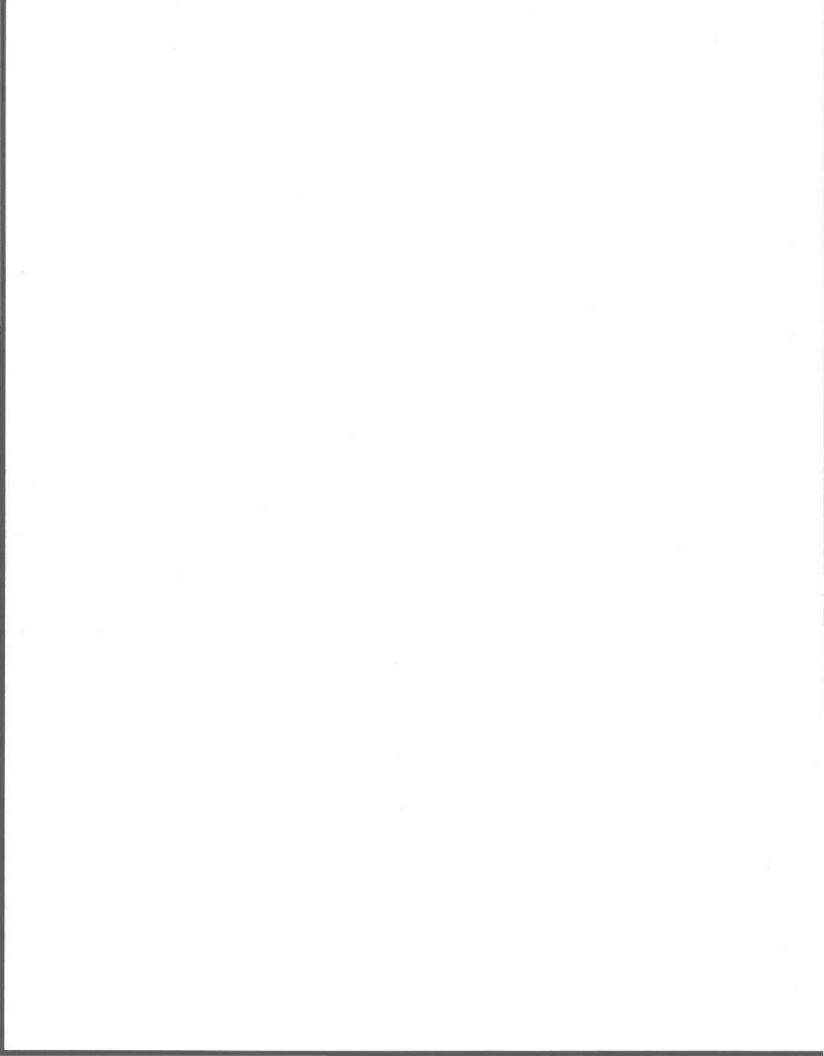
Directeur de recherche

Peter F. Rasmussen INRS-Eau

Codirecteur

Bernard Bobée INRS-Eau

12 décembre 1999



Remerciements

Je tiens à remercier mon directeur, monsieur Peter F. Rasmussen pour son soutien, sa disponibilité, son professionnalisme et toutes ses qualités qui font de lui un excellent directeur de maîtrise, monsieur Bernard Bobée et l'équipe de la Chaire industrielle en hydrologie statistique pour son support scientifique et monétaire, ainsi que monsieur Alin Andrei Carsteanu pour avoir pris la relève après le départ de Peter pour Winnipeg.

Des remerciements très spéciaux vont à ma future épouse, Doris, mes parents et ma soeur. Ils ont été là pour me soutenir, m'aider et me redonner confiance à chaque fois que j'en avais besoin. Je remercie aussi tous mes amis et amies qui m'ont encouragé ou qui m'ont fait partager leur expérience à un moment ou à un autre : Claude, Rossana, Patrick, Daniel, Dany, Marie, Stéphanie, Nicholas et la *e-gang*. Leur présence, physique ou électronique, a fait que j'ai pu passer au travers de ces deux années beaucoup plus agréablement et rapidement.

Résumé

La modélisation des précipitations peut être fort utile dans la planification de la ressource eau, le dimensionnement des ouvrages hydrauliques ou tout autre projet demandant des connaissances sur le phénomène des précipitations. Le but du présent travail est de montrer la diversité des modèles qui existent pour modéliser la pluie et donner une évaluation objective de la performance d'une sélection de modèles typiques, basée sur leur capacité de générer des relations intensité-durée-fréquence réalistes. Cet outil servira à donner un indice de l'adéquation des modèles par rapport aux séries de précipitations extrêmes.

Un inventaire non exhaustif des processus ponctuels de précipitation est fait. Ensuite, la performance de quatre de ces modèles est évaluée. Les modèles sélectionnés sont les suivants :

- modèle de Poisson à impulsions rectangulaires
- modèle de Neyman-Scott à impulsions rectangulaires
- modèle de Bartlett-Lewis avec perturbateur
- modèle de Bartlett-Lewis avec intensité de loi Gamma

Pour chacun de ces modèles, les courbes IDF sont tracées pour les séries de durée partielle. On vérifiera également si les moments sont bien préservés ainsi que les proportions d'intervalles secs. Ces caractéristiques sont importantes dans le cas où, par exemple, on voudrait incorporer un modèle de précipitation dans un modèle "pluie-débit". Les intervalles secs influencent directement les propriétés de ruissellement.

On abordera également l'estimation de paramètres car cela semble devenir un obstacle lorsqu'on veut modifier un modèle et que la modification requiert un ajout de paramètres.

Table des matières

REMERCIEMENTS	iii
RÉSUMÉ	iv
TABLE DES MATIÈRES	v
LISTE DES TABLEAUX	viii
LISTE DES FIGURES	ix
NOTATIONS	xi
1. INTRODUCTION	1
1.1 Problématique	1
1.2 Objectif et méthodologie	2
2. OUTILS STATISTIQUES PERTINENTS	5
2.1 Fonction caractéristique	5
2.2 Fonction génératrice de probabilités (fgp)	6
2.3 Processus de Poisson	9
2.3.1 Définition	9
2.3.2 Processus de Poisson non-homogène	
2.3.3 Variables associées aux processus de Poisson	
2.3.4 Processus ponctuel de Poisson	10
2.3.5 Processus de Poisson filtré	10
2.3.6 Processus de Poisson cumulé	11
3. MÉTHODE DE VALIDATION	13
3.1 L'étude des événements extrêmes	
3.1.1 Intensités de précipitations maximales annuelles (PMA)	13
3.1.2 Pointes d'intensité dépassant un seuil (PDS)	14
2.1.2 Construction de courbes IDE (Intensité Durée Fréquence)	1.5

3.2 Te	sts d'ajustement	_ 15
3.2.1	Test de Kolmogorov-Smirnov	16
3.2.2	2 Test d'ajustement modifié	_16
3.3 Dis	stributions statistiques pour les événements extrêmes	_ 17
3.3.1	Loi de Pareto généralisée	17
3	.3.1.1 Propriétés statistiques	_17
3	3.1.2 Estimation des paramètres	18
3	.3.1.3 Estimation des quantiles X_T de période de retour T	19
3.3.2	2 Loi généralisée des valeurs extrêmes (GEV)	19
3	.3.2.1 Propriétés statistiques	_19
3	.3.2.2 Estimation des paramètres	_21
3	.3.2.3 Estimation des quantiles X_T de période de retour T	22
4. MOD	ÈLES STOCHASTIQUES DE PRÉCIPITATION	_ 25
4.1 Mo	odèles de précipitation de Poisson (PIR)	_ 25
4.1.1	Modèle de Poisson avec impulsions rectangulaires	_25
4.1.2	2 Modèle bruits blancs de Poisson	_32
4.2 Mo	odèles de Neyman-Scott (N-S)	_ 33
4.2.1	Introduction	33
4.2.2	2 Modèle avec bruits blancs de Neyman-Scott	35
4.2.3	8 Modèle de N-S à impulsions rectangulaires (NSIR)	_38
4.2.4	Modèle de Neyman-Scott η-aléatoire	_41
4.3 Mo	odèles de Bartlett-Lewis (B-L)	_ 42
4.3.1	Introduction	_42
4.3.2	2 Modèle de Bartlett-Lewis avec impulsions rectangulaires (BLIRP)	43
4.3.3	3 Modèle de Bartlett-Lewis η-aléatoire	47
4.3.4	Introduction d'une perturbation (jitter) dans le modèle à impulsions rectangulaires	50
4.3.5	5 Intensité de loi gamma pour s'ajuster aux valeurs extrêmes	54
5. APPL	ICATION DE CERTAINS MODÈLES AUX DONNÉES DE L'ALABAMA_	_ 57
5.1 An	alyse qualitative des données	_ 57
5.2 Mc	odèle de Poisson à impulsions rectangulaires (PIR)	_ 59

5.2.1 Les paramètres	59
5.2.2 Analyse du modèle	61
5.3 Neyman-Scott à impulsions rectangulaires (NSIR)	66
5.3.1 Estimation des paramètres	66
5.3.2 Application du modèle	69
5.4 Bartlett-Lewis à impulsions rectangulaires avec perturbateur (BLIRP)	71
5.4.1 Estimation des paramètres	71
5.4.2 Application du modèle	73
5.5 Bartlett-Lewis à impulsions rectangulaires et intensités de loi Gamma (BLIRIG) _	75
5.5.1 Estimation des paramètres	75
5.5.2 Application du modèle	76
5.6 Comparaison des modèles	78
5.6.1 Station 8	78
5.6.2 Station 63	80
5.6.3 Station 369	81
5.6.4 Station 748	82
5.6.5 Synthèse de l'étude des résultats	83
6. CONCLUSION ET DISCUSSION	87
7. BIBLIOGRAPHIE	89
ANNEXE A	03

Liste des tableaux

Tableau 5.1: Valeurs des paramètres du modèle de Poisson à impulsions rectangulaires d
différentes échelles de temps pour la station 859
Tableau 5.2 : Hauteurs moyennes de précipitation à différentes échelles de temps pour les
quatre stations60
Tableau 5.3 : Paramètres pour le modèle de PIR aux quatre stations61
Tableau 5.4: Propriétés statistiques des données historiques (en gras) et simulées à la
station 8. (*: moments utilisé dans l'estimation des paramètres)66
Tableau 5.5: Valeurs des paramètres obtenus avec la méthode de la fonction objectif selon
les valeurs initiales, les échelles utilisées et les valeurs finales pour la fonction objecti
à la station 868
Tableau 5.6 : Caractéristiques statistiques des données historiques (en gras) et simulées de
la station 8 avec E1 et E271
Tableau 5.7: Variances et autocorrélations de délai un, historiques et simulées, pour le
processus non-perturbé (BLIR) et perturbé (BLIRP) (station 8)73
Tableau 5.8: Moments et proportions d'intervalles secs pour le modèle de BLIRP à la
station 8 (les moments historiques sont en gras)75
Tableau 5.9 : Moments et proportions d'intervalles secs pour le modèle de BLIRIG à la
station 8 (les moments historiques sont en gras)78
Tableau A.1 : Moments historiques et simulés et probabilités de périodes sèches pour les
quatre modèles étudiés à la station 894
Tableau A.2 : Moments historiques et simulés et probabilités de périodes sèches pour les
quatre modèles étudiés à la station 6395
Tableau A.3 : Moments historiques et simulés et probabilités de périodes sèches pour les
quatre modèles étudiés à la station 36996
Tableau A.4 : Moments historiques et simulés et probabilités de périodes sèches pour les
quatre modèles étudiés à la station 748

Liste des figures

Figure 2.1: Illustration du temps de séjour et du temps d'attente10
Figure 4.1: Scénario type avec le modèle de Poisson à impulsions rectangulaires. Dans la
partie du haut, les durées et les intensités sont simulées et dans la partie du bas, la
forme du scénario final est obtenue par superposition25
Figure 5.1: Moyennes journalières lissées. Les intervalles stationnaires sont indiqués dans
le graphique de chaque station58
Figure 5.2 : Ajustements des lois de Pareto généralisée (PG) et GEV aux séries de durées
partielles pour les différentes échelles de temps63
Figure 5.3 : Courbes IDF pour les séries de PDS des quatre stations. Les courbes
historiques sont en gras et les simulées correspondent aux traits minces65
Figure 5.4: Courbes IDF avec la méthode des PDS pour la station 8 avec E1 et E2. Les
courbes historiques sont en gras et les simulées correspondent aux traits minces70
Figure 5.5 : Résultats retenus pour les quatre stations parmi les ensembles de paramètres
trouvés avec le modèle de NSIR. Les courbes historiques sont en gras et les simulées
correspondent aux minces72
Figure 5.6: Courbes IDF des quatre stations pour le modèle de BLIRP. Les courbes
historiques sont en gras et les simulées correspondent aux minces74
Figure 5.7 : Courbes IDF des quatre stations pour le modèle de BLIRIG. Les courbes
historiques sont en gras et les simulées correspondent aux minces77
Figure 5.8 : Courbes IDF résultantes des quatre modèles considérés, pour la station 8. Les
courbes historiques sont en gras et les simulées correspondent aux minces79
Figure 5.9 : Courbes IDF résultantes des quatre modèles considérés, pour la station 63
Les courbes historiques sont en gras et les simulées correspondent aux minces80
Figure 5.10 : Courbes IDF résultantes des quatre modèles considérés, pour la station 369
Les courbes historiques sont en gras et les simulées correspondent aux minces.

Figure 5.11 : Courbes IDF résultantes des quatre modèles considérés, pour la station	748.
Les courbes historiques sont en gras et les simulées correspondent aux minces.	_ 82
Figure 5.12 : Courbes IDF obtenues avec le modèle de BLIRP à la station 8	_ 84
Figure 5.13 : Courbes IDF obtenues avec le modèle de BLIRP à la station 63	_ 84
Figure 5.14 : Courbes IDF obtenues avec le modèle de NSIR à la station 369.	_ 85
Figure 5.15 : Courbes IDF obtenues avec le modèle de NSIR à la station 748.	_ 86

Notations

C: nombre de précipitations encore actives à un temps donné

 $f_{()}$: fonction de densité de la variable en indice

G(): fonction génératrice de probabilités de la variable aléatoire en indice

h : échelle de temps utilisée ;

L: variable aléatoire correspondant à la durée d'une précipitation pour les modèles à impulsions rectangulaires

T: vecteur aléatoire donnant la position temporelle des précipitations pour un front

c : variable aléatoire correspondant au nombre de précipitations générées par un front

X: variable aléatoire correspondant à l'intensité d'une précipitation pour les modèles à impulsions rectangulaires ;

X: vecteur d'intensités;

Y(t): intensité totale de précipitation à un instant t

 Y_i^T : processus de précipitation agrégé sur des intervalles de longueur T

lpha : paramètre de forme de la loi gamma qui gère le paramètre η dans les modèles η -aléatoire de N-S et B-L

 β : - paramètre de la loi exponentielle de la variable gérant la position temporelle des précipitations (T) dans le modèle de N-S;

- paramètre du processus de Poisson qui engendre les origines des précipitations dans le modèle de B-L

 $\phi = \gamma/\eta$: paramètre sans dimension dans le modèle de B-L

 γ : paramètre de la loi exponentielle qui donne la durée d'un front dans le modèle de B-L

 η : paramètre de la loi exponentielle gérant la variable de durée (L) dans les modèles à impulsions rectangulaires

 $\kappa = \beta/\eta$: paramètre sans dimension dans le modèle de B-L

 \(\lambda\) : paramètre du processus de Poisson qui engendre l'origine des fronts dans les modèles de Neyman-Scott et de Bartlett-Lewis et l'origine des précipitations dans les modèles de bruits blancs et de Poisson simple

 $\mu_{()}$: premier moment de la variable en indice

 μ_a : durée moyenne qu'un front est actif (B-L)

 $\mu_{\scriptscriptstyle D}$: durée moyenne des périodes sèches

 $\mu_{\scriptscriptstyle W}$: durée moyenne des périodes humides

 ν : paramètre d'échelle de la loi gamma qui gère le paramètre η dans les modèles η -aléatoire de N-S et B-L

 $\rho = \lambda/\eta$: paramètre sans dimension dans les modèles de N-S et B-L

 $\rho_{()}(t)$: fonction d'autocorrélation de délai t de la variable en indice

 $\sigma_{()}^2$: variance de la variable en indice

1. INTRODUCTION

1.1 Problématique

La pluie est un phénomène qui fait partie intégrante de notre climat. L'immensité et la complexité du processus qui engendre les précipitations sont telles que nous ne pouvons pas le contrôler, ni le modéliser exactement de façon à avoir des prévisions précises à long terme. Plusieurs disciplines comme l'hydrologie, la météorologie, la physique atmosphérique et la climatologie sont nécessaires pour expliquer ce phénomène complexe.

Il s'agit donc d'un point important lorsque nous devons résoudre des problèmes d'hydrologie. La fréquence, la durée, l'intensité et toutes les propriétés de la pluie influencent le dimensionnement des ouvrages hydrauliques, la gestion des réservoirs et l'opération des systèmes de drainage urbains. L'agriculture et la prévention des feux de forêts sont également des domaines qui dépendent fortement des précipitations. Cette multiplicité de champs d'applications fait qu'il devient primordial de bien étudier et connaître les propriétés du phénomène.

Même si la pluie est un phénomène tout à fait déterministe, les modèles stochastiques constituent un moyen de reproduire certaines informations, que l'utilisateur du modèle choisira, que les séries de données de précipitations nous procurent. On jugera de l'adéquation d'un modèle par sa capacité de reproduire ces propriétés. Les résultats obtenus avec les modèles stochastiques peuvent être utilisés pour l'étude de scénarios d'épisodes pluvieux à long terme. Le niveau de la performance d'un modèle peut alors devenir critique lorsque celui-ci est utilisé pour la planification de la ressource eau ou comme intrant dans un modèle pluie-débit.

Bien sûr, un modèle relativement simple ne pourra servir à décrire complètement le processus. Le choix des propriétés que nous voulons préserver ou reproduire avec un modèle influencera grandement le choix du modèle lui-même. Dans plusieurs cas, les événements extrêmes et les épisodes pluvieux qui les précèdent et les suivent seront attentivement étudiés étant donné leur impact immédiat sur le débit d'une rivière ou sur le niveau d'un réservoir par exemple.

1.2 Objectif et méthodologie

Devant le grand nombre de modèles disponibles pour la modélisation de la pluie, un utilisateur est en droit de se demander lequel est le meilleur pour reproduire les propriétés qui l'intéressent. L'objectif principal de la présente étude sera donc de donner une liste, non exhaustive, des modèles stochastiques de précipitation de type ponctuel et stationnaires, et d'étudier numériquement et graphiquement leur capacité à reproduire les événements extrêmes.

Numériquement, nous porterons attention aux moments et aux propriétés d'intervalles secs. Comme outil graphique, les courbes IDF (Intensité-Durée-Fréquence) seront exploitées. Ces courbes, par leur construction, donnent de l'information sur l'intensité des précipitations en fonction de la durée pour chacune des périodes de retour préalablement choisies. Un événement X_T de période de retour T (en années ou toute autre unité de temps) est tel que la probabilité de dépassement est égale à 1/T, c'est-à-dire que $P(X > X_T) = T^{-1}$. Cette notion sera exposée plus en détail dans ce travail. Cette quantité qui est très importante dans l'étude des précipitations maximales annuelles ou autre série d'événements extrêmes, sera utilisée dans la construction de courbes IDF.

Les quatre modèles qui seront considérés sont les suivants:

- le modèle de Poisson à impulsions rectangulaires
- le modèle de Neyman-Scott à impulsions rectangulaires
- le modèle de Bartlett-Lewis η -aléatoire avec perturbateur
- le modèle de Bartlett-Lewis η -aléatoire avec intensité de gamma

Chacun d'entre eux sera utilisé pour simuler une série d'événements pluvieux. De chacune de ces simulations sera extraite une série d'événements extrêmes ainsi que les moments usuels. Ces propriétés seront comparées avec les propriétés historiques correspondantes. Des courbes IDF seront tracées pour chaque station.

7.			

2. OUTILS STATISTIQUES PERTINENTS

2.1 Fonction caractéristique

Soit un vecteur aléatoire X. Johnson et al. (1992) expriment sa fonction caractéristique sous la forme suivante:

$$\varphi_{\mathbf{X}}(t) = E[e^{it\mathbf{X}}] \tag{2.1}$$

pour n'importe quel scalaire à valeurs réelles t. Notons que dans le cas présent, l'exponentielle d'un vecteur \mathbf{X} sera considérée comme étant le vecteur contenant l'exponentielle de chacun des éléments du vecteur \mathbf{X} .

Dans la plupart des livres de référence pertinents, on trouve une panoplie de propriétés de la fonction caractéristique. Voici les principales qui seront utilisées dans ce travail. *Snyder* (1975) montre par récurrence la propriété suivante pour un vecteur aléatoire de longueur *n*, dont les composantes sont indépendamment distribuées :

$$E[e^{itX}] = \prod_{j=1}^{n} E[e^{itX_j}]$$
(2.2)

De la fonction caractéristique, on peut tirer les moments de la variable aléatoire. Il existe, entre autres, deux moyens d'utiliser la fonction caractéristique. La première donne les moments successifs en dérivant directement la fonction caractéristique et en l'évaluant en t=0. On a donc :

$$E(X^{k}) = \left(i^{k} \frac{\partial^{k} \varphi_{X}(t)}{\partial t^{k}}\right)_{t=0}$$
(2.3)

On peut aussi considérer le logarithme de la fonction caractéristique. En dérivant et en évaluant toujours en t=0, on a les propriétés suivantes :

$$iE(X(\tau)) = \left(\frac{d}{dt}\log\varphi_{X(\tau)}(t)\right)_{t=0}$$

$$i^{2}Var(X(\tau)) = \left(\frac{d^{2}}{dt^{2}}\log\varphi_{X(\tau)}(t)\right)_{t=0}$$

$$i^{2}Cov(X(\tau_{1}), X(\tau_{2})) = \left(\frac{\partial^{2}}{\partial t_{1}\partial t_{2}}\log\varphi_{X(\tau_{1}), X(\tau_{2})}(t_{1}, t_{2})\right)_{(0,0)}$$

$$(2.4)$$

2.2 Fonction génératrice de probabilités (fgp)

Supposons Y, une variable aléatoire discrète non-négative. Soit sa fonction de masse $P_j = P[X = j]$, j = 0,1,... La fonction génératrice de probabilités d'une variable aléatoire Y à valeurs entières positives a la forme suivante [Johnson et al. (1992)]:

$$G_{Y}(z) = \sum_{j>0} P_{j} z^{j} = E[Z^{Y}]$$
 (2.5)

où nous pouvons supposer $-1 \le z \le 1$. Cette fonction donne une description complète de la distribution de Y. Comme la fgp est définie par les probabilités, l'unicité du développement en série de puissance implique que la fgp donne les probabilités de la façon suivante:

$$P_{j} = \left[\frac{1}{j!} \frac{d^{j} G(z)}{dz^{j}}\right]_{z=0}$$

$$(2.6)$$

Le r^e moment, s'il existe, est obtenu en évaluant la dérivée de G en t = 0.

$$E[Y'] = \sum_{j=0}^{\infty} j' P_j$$

$$= \left[\frac{d' G(e^t)}{dt'} \right]_{t=0}$$
(2.7)

Supposons que N(A) soit un processus qui compte le nombre d'occurrences d'un événement dans un sous-ensemble A de l'espace H (non borné) et qu'on ait une partition de H, A_1 , A_2 , ..., A_n . Soit le vecteur aléatoire $N = [N(A_1), ..., N(A_n)]$, la fonction génératrice de probabilités (forme multivariée) a la forme suivante [Reiss (1993)]:

$$G_{N}(z_{1},...,z_{n}) = E\left[z_{1}^{N(A_{1})} \cdots z_{n}^{N(A_{n})}\right]$$

$$= E\left[\exp\left\{\sum_{i=1}^{n} N(A_{i}) \log z_{i}\right\}\right] \qquad 0 < z_{i} \le 1$$
(2.8)

On peut en déduire que [Waymire et Gupta (1981)] :

$$G_{N}(z_{1},...,z_{n}) = E \left[\exp \left\{ \int_{H} \log \xi(x) dN(x) \right\} \right]$$

$$avec \quad \xi(x) = \begin{cases} z_{i} & \text{si } x \in A_{i} \\ 1 & \text{si } x \notin \bigcup A_{i} \end{cases}$$

$$(2.9)$$

Plus généralement, pour toute fonction ξ telle que

$$0 < \xi(x) \le 1 \qquad \forall x \in \mathbf{H} \\
et \quad \xi(x) = 1 \qquad \forall x \notin \mathbf{H}$$
(2.10)

la fonctionnelle $G_N[\xi]$ peut s'écrire des façons suivantes :

$$G_{N}[\xi] = E\left[\exp\left\{\int_{\mathbb{H}} \log \xi(x) dN(x)\right\}\right]$$

$$= E\left[\prod_{x_{i} \in \mathbb{H}} \xi(x_{i})\right]$$
(2.11)

où

$$dN(x) = \begin{cases} 1 & \text{s'il y a un \'ev\'enement au temps } x \\ 0 & \text{sinon} \end{cases}$$
 (2.12)

et où $\left\{ \int_{H} \log \xi(t) \, dN(t) \right\}$ vaut 0 si $\xi(T_n) = 0$ pour certains n et vaut 1 si $\xi(T_n) = 1$ pour tout n. On supposera que $0 \le \xi(t) \le 1$ pour tout t réel et ξ est identiquement égal à 1 en dehors de certaines limites, ceci assure que (2.11) converge.

Par exemple, la fonctionnelle génératrice de probabilités d'un processus de Poisson de taux $\rho(t)$ est

$$G[\xi] = \exp\left[-\int_{\mathcal{H}} \left\{1 - \xi(t)\right\} \rho(t) dt\right]$$
(2.13)

On peut noter aussi que la fonction génératrice de probabilités d'une somme de variables aléatoires indépendantes est le produit des fonctions génératrice de probabilités. De la même manière, la fonctionnelle génératrice de probabilités d'une superposition de deux processus indépendants est le produit des fgp.

La fonction génératrice de probabilités et la fonction caractéristique sont très intimement liées. Le lien est intéressant à faire et on peut établir la relation suivante (selon la même notation que dans la section précédente) :

$$\varphi_X(t) = E[e^{itX}] = G_X(e^{it}) \tag{2.14}$$

2.3 Processus de Poisson

2.3.1 Définition

Un processus aléatoire $\{N(t), t \ge 0\}$ qui compte le nombre d'événements qui se produisent dans l'intervalle (0,t) (processus de dénombrement) est un processus de Poisson de paramètre $\lambda > 0$ si les conditions suivantes [Taylor et Karlin (1993)] sont satisfaites:

- (i) pour $0 = t_0 < t_1 < ... < t_n < ..., N(t_1) N(t_0), N(t_2) N(t_1),...$ sont des variables aléatoires indépendantes;
- (ii) pour s et t positifs, les variables aléatoires N(t+s) N(s) suivent une loi de Poisson de paramètre λt ;
- (iii) N(0) = 0.

2.3.2 Processus de Poisson non-homogène

Si le taux d'occurrence λ d'un processus de Poisson est fonction du temps $(\lambda(t))$, on dit que le processus est non-homogène. Donc, si N(t) est un processus de Poisson non-homogène, alors la variable N(s) - N(t) suit une loi de Poisson de paramètre ζ avec $\zeta = \int_{s}^{s} \lambda(u) du$.

2.3.3 Variables associées aux processus de Poisson

Sur la Figure 2.1, on peut observer trois variables aléatoires. En ordonnée, il y a la variable N(t) telle que définie plus haut. En abscisse, il y a le temps de séjour S_i qui est le temps pendant lequel le processus reste à l'état i. Il y a aussi le temps d'attente W_i qui est la longueur de l'intervalle de temps entre l'origine et l'arrivée à l'état i. Taylor et Karlin (1993) montrent que ces variables sont distribuées comme suit :

- W_n suit une loi Gamma $(n,1/\lambda)$, donc $f_{W_n}(t) = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}$ avec $t \ge 0$
- S_k suit une loi exponentielle de paramètre $1/\lambda$, donc $f_{S_k}(t) = \lambda e^{-\lambda t}$ avec $t \ge 0$

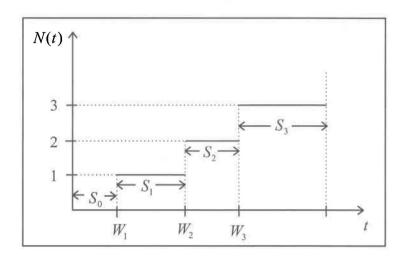


Figure 2.1: Illustration du temps de séjour et du temps d'attente

2.3.4 Processus ponctuel de Poisson

Soit N((s,t]) une variable aléatoire qui compte le nombre d'occurrences dans l'intervalle (s,t]. N((s,t]) est un processus ponctuel de Poisson de paramètre λ si les conditions suivantes [Taylor et Karlin (1993)] sont satisfaites :

- (i) pour tout m=2, 3, ..., et des temps $0=t_0 < t_1 < ... < t_m$ les variables aléatoires $N((t_0,t_1])$, $N((t_1,t_2])$, ..., $N((t_{m-1},t_m])$ sont indépendantes
- (ii) pour tout s < t, N((s,t]) suit une loi de Poisson de paramètre $\lambda(t-s)$

2.3.5 Processus de Poisson filtré

Un processus stochastique $\{X(t), t \ge 0\}$ est dit de Poisson filtré s'il peut être représenté par l'expression suivante:

$$X(t) = \sum_{m=1}^{N(t)} h(t, t_m, Y_m)$$
 (2.15)

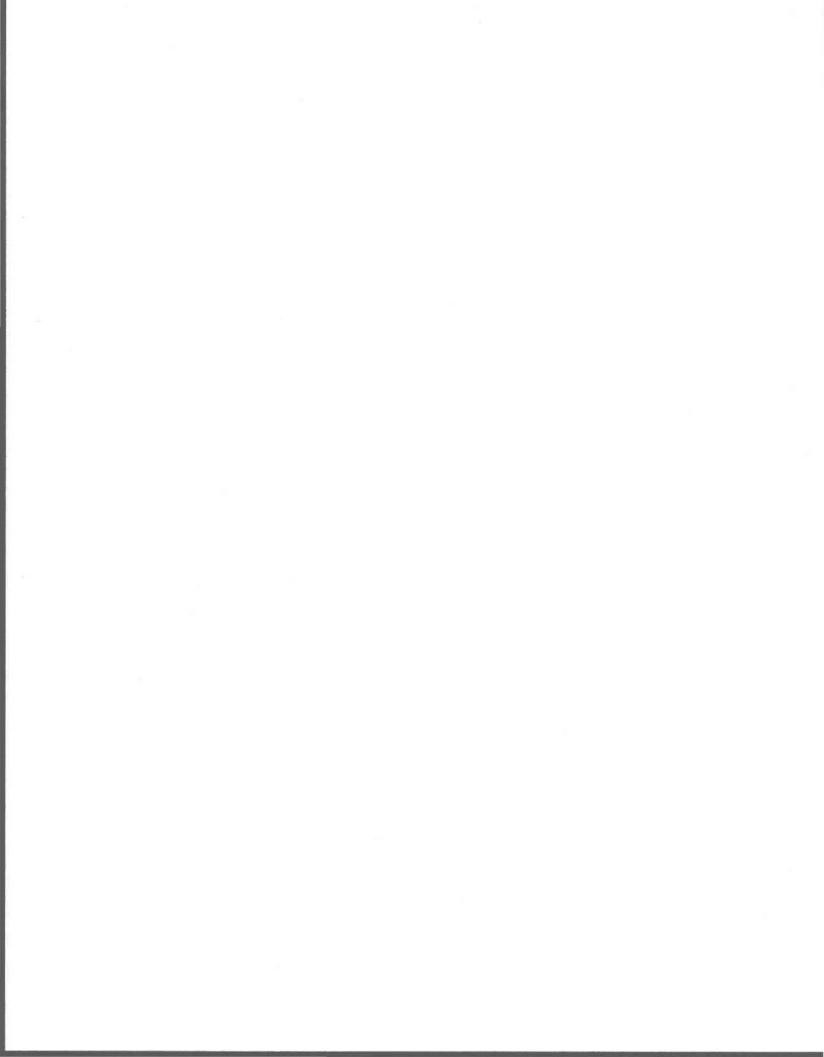
où les conditions suivantes [Parzen (1962)]doivent être satisfaites :

- $\{N(t), t \ge 0\}$ est un processus de Poisson, de paramètre λ ;
- $\{Y_n\}$ est une série de variables aléatoires identiquement distribuées et indépendantes du processus $\{N(t), t \ge 0\}$;
- $h(t, t_m, Y_m)$ est ce qu'on appelle la fonction de réponse ou fonction de transfert et t est le temps auquel on se réfère, t_m est le temps d'occurrence de la variable Y_m .

2.3.6 Processus de Poisson cumulé

Supposons que nous ayons un processus de Poisson N(t) homogène de paramètre λ tel que défini plus haut. Supposons aussi que nous associons à chaque événement une variable aléatoire Y_1, Y_2, \ldots , i.i.d. (indépendantes et identiquement distribuées) de fonction de distribution $G(y) = P(Y_k \le y)$. Un processus de Poisson cumulé Z(t) est défini de la façon

suivante [Taylor et Karlin (1993)]:
$$Z(t) = \sum_{k=1}^{N(t)} Y_k$$
 avec $t \ge 0$



3. MÉTHODE DE VALIDATION

3.1 L'étude des événements extrêmes

La construction des courbes IDF est généralement basée sur des données de précipitations extrêmes. Il existe plusieurs façons de les déterminer. Dans le présent document, deux approches sont exposées. La première est la méthode des précipitations maximales annuelles qui prend la précipitation maximum d'une durée donnée de chaque année. La seconde est la méthode des pointes dépassant un seuil. Les sections suivantes décrivent ces deux méthodes.

3.1.1 Intensités de précipitations maximales annuelles (PMA)

Cette méthode est certainement la plus simple des deux, mais elle comporte quand même quelques désavantages par rapport à la seconde. Il s'agit de faire glisser une fenêtre de longueur h (l'échelle d'agrégation) sur les données de l'année i pour trouver quel intervalle donne l'intensité maximum X_i pour cette année. On obtient ainsi un vecteur d'événements $\left(X_1, \ldots X_n\right)$ pour les n années de données disponibles. En plaçant en ordre décroissant ces événements, $\left(X_{(1)} \geq X_{(2)} \geq \ldots \geq X_{(n)}\right)$, on extrait les probabilités empiriques que l'on fait correspondre à une période de retour T de la façon suivante:

$$P(X > X_{(i)}) = \frac{i}{n+1} = \frac{1}{T}$$
(3.1)

et l'événement $X_{(i)}$ est alors associé à la période de retour $T = \frac{n+1}{i}$. Un tel événement sera noté X_T . Pour déterminer les événements de période de retour plus élevée que la longueur (n) de la série, on ajuste une loi de probabilité pour ensuite extrapoler les

événements X_T avec T > n. Les lois de Pareto généralisée et la GEV (loi généralisée des valeurs extrêmes) seront utilisées pour modéliser ce type de série.

3.1.2 Pointes d'intensité dépassant un seuil (PDS)

Parfois aussi appelées séries de durée partielle [*Ouarda et Ashkar* (1995)], ce second type de séries ne comporte que les événements qui dépassent un certain seuil. La méthode précédente ne conservait que la valeur maximum d'une année, sans tenir compte du fait que, par exemple, le deuxième événement le plus grand de cette même année puisse dépasser le maximum d'une autre année moins pluvieuse. De plus, comme nous le verrons à la section 5.1, lorsqu'une période relativement courte de l'année est considérée, il peut arriver qu'une année plus sèche donne un maximum annuel qui ne devrait pas être considéré comme étant une précipitation extrême.

Une des raisons pour lesquelles cette méthode n'est pas souvent utilisée est que le choix du seuil comporte un élément subjectif. Un moyen de contourner un peu cette subjectivité est de fixer un nombre d'événements par série. Encore une fois, en raison de la faible longueur de l'intervalle de temps considéré, il pourrait être raisonnable de fixer le nombre d'événements extrêmes sous 1 par an. Lorsqu'on ajustera la loi de probabilité sur la série, ce sera uniquement dans le but d'extrapoler l'intensité de période de retour 100 ans. Dans ce but, on voudra que la loi s'ajuste bien aux grandes valeurs plutôt que d'avoir un bon ajustement pour les valeurs moins grandes (près du seuil), qui sont plus nombreuses.

De plus, pour s'assurer de l'indépendance entre deux précipitations considérées extrêmes, on rejette les événements qui ont lieu dans la même semaine. Cela signifie que, de toutes les pointes dépassant le seuil et ayant lieu à l'intérieur d'un intervalle d'une semaine, on prend la plus haute et on rejette les autres. La série qui sera considérée comprendra tous les dépassements, c'est-à-dire les observations dépassant le seuil, diminuées de ce seuil. Ainsi, l'observation la plus petite sera près de 0.

Comme le nombre d'événements que contient une série de durée partielle n'est pas nécessairement égal au nombre d'années de données, un ajustement des périodes de retour doit être effectué. Supposons que la série de durée partielle à notre disposition soit de longueur n_e et que les données historiques soient d'une durée de n_a années. On peut alors dire que chaque événement représente une période de n_a / n_e années [Aron et al. (1987)]. L'ajustement qu'il faut effectuer sur les périodes de retour est la suivante:

$$T_a = T_e \cdot \frac{n_a}{n_e} \tag{3.2}$$

où T_e est la période de retour équivalente dans les unités des n_a / n_e années et T_a est la période de retour correspondante en années standards.

3.1.3 Construction de courbes IDF (Intensité-Durée-Fréquence)

Un graphique de courbes IDF illustre à la fois la relation entre l'intensité, la durée et la fréquence d'une précipitation. En abscisse, on retrouve les échelles d'agrégation, en ordonnée on a l'intensité et une courbe est tracée pour chaque période de retour.

Après avoir obtenu la série de précipitations considérées extrêmes selon l'une ou l'autre des méthodes, on extrapole les intensités de périodes de retour supérieures au nombre d'années disponibles dans les données. Pour ce faire, une loi de probabilité est ajustée à la série. Pour l'une ou l'autre des méthodes, les lois de Pareto généralisée et la GEV (loi généralisée de valeurs extrêmes) s'ajustent généralement assez bien.

3.2 Tests d'ajustement

Afin d'avoir une mesure de l'adéquation d'une loi de probabilité à un ensemble d'observations, deux tests d'ajustement seront utilisés. Le test de *Kolmogorov-Smirnov* dans sa version originale et une version modifiée.

3.2.1 Test de Kolmogorov-Smirnov

Ce test sert à vérifier l'adéquation d'une fonction de distribution à une série d'observations. Dans l'article de *Massey* (1951), on retrouve une description et une table de valeurs critiques pour le test de *Kolmogorov-Smirnov*.

Nous posons comme hypothèse que les observations sont issues d'une population qui est distribuée selon la fonction de distribution cumulative $F_0(x)$. On veut tester si un échantillon de taille n peut être considéré comme étant issu de cette distribution. En posant $\hat{F}(x) = k / n$ où k est le nombre d'observations plus petites ou égales à x (probabilités empiriques), la distribution de la statistique $d = \max \left| F_0(x) - \hat{F}(x) \right|$ est connue et est indépendante de $F_0(x)$ si $F_0(x)$ est continue. Les valeurs critiques qui permettent de rejeter ou non l'hypothèse sont disponibles dans l'article de Massey (1951) pour différents niveaux de confiance et différentes tailles d'échantillons.

3.2.2 Test d'ajustement modifié

Le test présenté par *Finkelstein et Schafer* (1971) est une version modifiée du test de *Kolmogorov-Smirnov*. Les auteurs montrent que, dans les cas traités, cette nouvelle version est plus puissante.

Posons

$$\delta_k = \max[|F_0(x) - (k-1)/n|, |F_0(x) - k/n|] \qquad (k = 1, ..., n)$$
(3.3)

en utilisant la même notation que dans la section précédente. Pour un échantillon de taille n, il y a n valeurs de δ_k et on peut déduire la statistique suivante:

$$S_n = \sum_{k=1}^n \delta_k \tag{3.4}$$

Les auteurs proposent d'utiliser cette statistique pour tester la même hypothèse que celle qu'on veut tester avec le test original. Une table des valeurs critiques est également disponible dans l'article pour différents niveaux de confiance et différentes tailles d'échantillons.

Ce test est sensiblement équivalent à sa version originale mais *Finkelstein et Schafer* montrent que pour le cas des lois normale et exponentielle, cet apport donne un test plus puissant (au sens de la puissance statistique). Dans la présente étude, les deux tests seront utilisés pour tester l'ajustement des lois aux séries d'événements extrêmes.

3.3 Distributions statistiques pour les événements extrêmes

3.3.1 Loi de Pareto généralisée

3.3.1.1 Propriétés statistiques

Supposons que nous avons une variable aléatoire Y qui soit distribuée selon une loi exponentielle. On dit alors de la variable

$$X = \frac{\alpha}{k} \left(1 - e^{-kY} \right) \tag{3.5}$$

qu'elle est distribuée selon la loi Pareto généralisée [Hosking et Wallis (1987)]. La fonction de distribution est définie par

$$F(x) = \begin{cases} 1 - \left(1 - \frac{k}{\alpha}x\right)^{1/k}, & k \neq 0\\ 1 - \exp(-x/\alpha), & k = 0 \end{cases}$$
(3.6)

et la fonction de densité par

$$f(x) = \begin{cases} \frac{1}{\alpha} \left(1 - \frac{k}{\alpha} x \right)^{1/k - 1}, & k \neq 0 \\ \frac{1}{\alpha} \exp(-x / \alpha), & k = 0 \end{cases}$$
 (3.7)

où x varie sur l'intervalle $[0,\infty)$ pour k < 0 et sur $[0,\alpha/k]$ pour k > 0. α est le paramètre d'échelle et k est le paramètre de forme. Les cas k = 0 et k = 1 correspondent respectivement à la loi exponentielle de moyenne α et à la loi uniforme sur l'intervalle $[0,\alpha]$.

3.3.1.2 Estimation des paramètres

Pour estimer les paramètres, la méthode des moments sera retenue car c'est celle-ci qui est utilisée le plus fréquemment. Sachant que la fonction génératrice des moments d'une variable aléatoire de loi exponentielle de paramètre θ est

$$M_{Y}(t) = \frac{1}{1 + \theta t} \tag{3.8}$$

si $t < 1/\theta$, Hosking et Wallis (1987) déduisent que

$$E\left(1 - \frac{k}{\alpha}X\right)' = \frac{1}{1 + rk}\tag{3.9}$$

si rk > -1. C'est donc sous cette condition que le moment d'ordre r existe. En supposant qu'ils existent, les moments d'ordre 1 et 2 donnent des expressions pour la moyenne et la variance:

$$\mu_X = \frac{\alpha}{1+k} \tag{3.10}$$

$$\sigma_X^2 = \frac{\alpha^2}{(1+k)^2(1+2k)} \tag{3.11}$$

En manipulant algébriquement ces expressions, on obtient les estimateurs de α et k,

$$\hat{\alpha} = \frac{\overline{x}}{2} \left(\frac{\overline{x}^2}{s^2} + 1 \right)$$

$$\hat{k} = \frac{1}{2} \left(\frac{\overline{x}^2}{s^2} - 1 \right)$$
(3.12)

où \bar{x} et s^2 sont respectivement la moyenne et la variance échantillonnale.

3.3.1.3 Estimation des quantiles X_T de période de retour T

Les quantiles X_T sont définis d'après la fonction de répartition F. En isolant x dans l'expression (3.6), on obtient

$$X_{T} = \begin{cases} \frac{\alpha}{k} \left[1 - \left(1 - F(X_{T}) \right)^{k} \right], & k \neq 0 \\ -\alpha \log \left(1 - F(X_{T}) \right), & k = 0 \end{cases}$$

$$(3.13)$$

Évidemment, on obtient l'estimateur de X_T , \hat{X}_T , en remplaçant α et k par leurs estimateurs, $\hat{\alpha}$ et \hat{k} respectivement, et en remplaçant $F(X_T)$ par T^{-1} , tel que donné dans l'expression (3.1).

3.3.2 Loi généralisée des valeurs extrêmes (GEV)

3.3.2.1 Propriétés statistiques

On distingue trois types de distributions des valeurs extrêmes [Perreault et Bobée (1992)]. Elles ont été nommées les lois des valeurs extrêmes et sont dénotées EV1, EV2 et EV3. La loi généralisée des valeurs extrêmes est la distribution suivante, qui combine ces trois lois :

$$F(x) = \exp \left[-\left(1 - \frac{k}{\alpha}(x - u)\right)^{1/k} \right], \qquad -\infty < x < \infty$$
 (3.14)

où u, α (>0) et k sont les paramètres. Le signe du paramètre de forme k déterminera de laquelle des trois lois des valeurs extrêmes est applicable.

a) Le cas où $k \to 0$ correspond à la loi EV1 qui est mieux connue sous le nom de distribution de Gumbel. La fonction de distribution de la Gumbel a la forme

$$F(x) = \exp\left[-\exp\left(-\frac{x-u}{\alpha}\right)\right], \quad -\infty < x < \infty$$
 (3.15)

- b) k < 0 correspond à la loi EV2 parfois appelée loi log-Gumbel car si X (X > 0) est une variable aléatoire EV2, alors $\log(X)$ suit une loi de Gumbel. Dans ce cas, X est bornée inférieurement par la valeur $u + \alpha / k$.
- c) k > 0 correspond à la loi EV3 appelée Weibull négative. En effet, si X suit une loi EV3, alors -X suit une loi de Weibull à 3 paramètres. X, dans ce cas, est bornée supérieurement par $u + \alpha / k$.

La densité de probabilité s'écrit, pour les deux cas:

$$f(x) = \frac{1}{\alpha} \exp\left[-\frac{(x-u)}{\alpha} - \exp\left(-\frac{x-u}{\alpha}\right)\right], \quad k = 0$$
 (3.16)

$$f(x) = \frac{1}{\alpha} \left[1 - \frac{k}{\alpha} (x - u) \right]^{1/k - 1} \exp \left[-\left(1 - \frac{k}{\alpha} (x - u) \right)^{1/k} \right], \quad k \neq 0 \quad (3.17)$$

3.3.2.2 Estimation des paramètres

Les paramètres peuvent être estimés par plusieurs méthodes telles que la méthode des moments, la méthode des moments pondérés et la méthode du maximum de vraisemblance. La méthode qui sera utilisée ici est la méthode des moments pondérés.

Pour la loi EV1 (i.e. k=0), on n'a besoin que de deux moments pondérés qui sont les suivants [Perreault et Bobée (1992)]:

$$\beta_0 = u + C\alpha \tag{3.18}$$

$$\beta_1 = \frac{1}{2} \left[u + \alpha \left(\ln(2) + C \right) \right] \tag{3.19}$$

On calcule les estimateurs de β_0 et β_1 , b_0 et b_1 , à partir de l'expression suivante :

$$b_r = \frac{1}{n} \sum_{j=1}^n \frac{(j-1)(j-2)...(j-r)}{(n-1)(n-2)...(n-r)} x_{(j)}$$
(3.20)

où les $x_{(j)}$ sont les statistiques d'ordre. Les estimateurs de α et u sont donnés par :

$$\hat{\alpha} = \frac{2b_1 - b_0}{\ln 2} \tag{3.21}$$

$$\hat{u} = b_0 - \gamma \hat{\alpha} \tag{3.22}$$

où γ est la constante d'Euler ($\approx 0,5772$).

Dans les cas des lois EV2 et EV3, le calcul est effectué à partir des trois expressions suivantes :

$$\beta_0 = u + \frac{\alpha}{k} \left[1 - \Gamma(1+k) \right] \tag{3.23}$$

$$2\beta_{l} - \beta_{0} = \frac{\alpha}{k} \left[\Gamma(l+k)(l-2^{-k}) \right]$$
(3.24)

$$\frac{3\beta_2 - \beta_0}{2\beta_1 - \beta_0} = \frac{1 - 3^{-k}}{1 - 2^{-k}} \tag{3.25}$$

En remplaçant les β_i par leurs estimateurs b_i (3.20), on résout d'abord, par méthode itérative, l'équation (3.25) qui ne dépend que de k. Ensuite, on déduit \hat{u} et $\hat{\alpha}$ des relations

$$\hat{\alpha} = \hat{k} \left[\frac{2b_1 - b_0}{\Gamma(1 + \hat{k})(1 - 2^{-k})} \right]$$
(3.26)

$$\hat{u} = b_o + \frac{\hat{\alpha}}{\hat{k}} \left[\Gamma \left(1 + \hat{k} \right) - 1 \right] \tag{3.27}$$

Il existe des tests pour vérifier que k est significativement différent de 0. Dans ce cas, quand on ne peut rejeter l'hypothèse que k=0, on recalcule les paramètres en supposant que nous sommes dans le cas d'une loi de Gumbel. Un tel test est décrit dans *Perreault et Bobée* (1992).

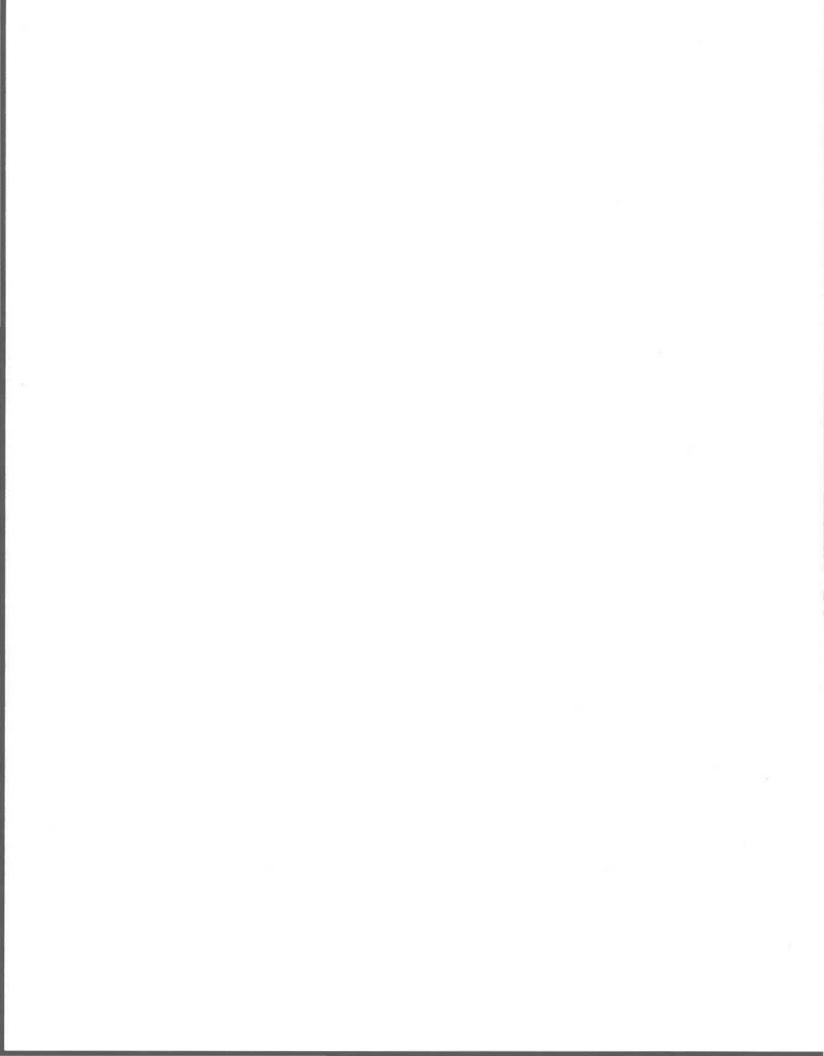
3.3.2.3 Estimation des quantiles X_T de période de retour T

L'estimation des quantiles de la loi GEV peut être effectuée aisément puisque la fonction de distribution est inversible. Ainsi, pour la loi EV1 (Gumbel), les estimateurs des quantiles sont donnés par [Perreault et Bobée (1992)] :

$$\hat{X}_T = \hat{u} - \hat{\alpha} \ln \left[-\ln \left(1 - T^{-1} \right) \right] \tag{3.28}$$

De la même manière pour $k \neq 0$, on obtient

$$\hat{X}_{T} = \hat{u} + \frac{\hat{\alpha}}{\hat{k}} \left\{ I - \left[-\ln\left(I - T^{-l}\right) \right]^{\hat{k}} \right\}$$
(3.29)



4. MODÈLES STOCHASTIQUES DE PRÉCIPITATION

4.1 Modèles de précipitation de Poisson (PIR)

4.1.1 Modèle de Poisson avec impulsions rectangulaires

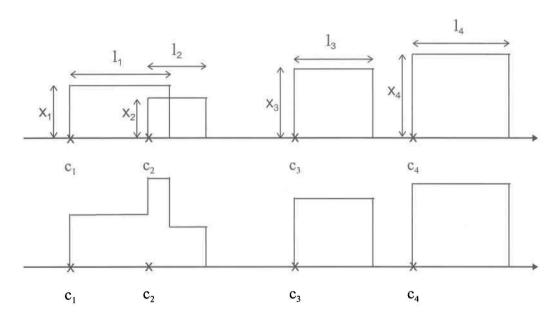


Figure 4.1: Scénario type avec le modèle de Poisson à impulsions rectangulaires. Dans la partie du haut, les durées et les intensités sont simulées et dans la partie du bas, la forme du scénario final est obtenue par superposition.

Bien que peu réaliste pour représenter les précipitations, le modèle de Poisson à impulsions rectangulaires a été la base de nombreux développements dans la modélisation de ce phénomène. Il consiste à associer à chaque point c_k d'un processus de Poisson N(t) de paramètre λ , une impulsion rectangulaire consistant en une hauteur x_k et une largeur l_k

correspondants respectivement à l'intensité et à la durée d'une précipitation. Ces deux valeurs sont aléatoires. On dit de l'impulsion qu'elle est rectangulaire parce qu'on suppose l'intensité constante pendant la durée de la précipitation. C'est ce qu'illustre la partie du haut de la Figure 4.1. On considère le processus comme étant additif, c'est-à-dire que lorsque plusieurs cellules de pluie se chevauchent, on peut additionner les intensités pour la durée où les cellules se chevauchent, comme on le voit dans la partie du bas de la Figure 4.1.

Dans le modèle original [Rodriguez-Iturbe et al. (1984)], l'intensité X et la durée L proviennent de loi exponentielle respectivement de paramètre ξ^{-1} et η^{-1} . Fixons les notations suivantes :

$$f_X(x) = \xi e^{-\xi x}$$
 $E(X) = \xi^{-1}$ $Var(X) = \xi^{-2}$
 $f_L(l) = \eta e^{-\eta l}$ $E(L) = \eta^{-1}$ $Var(L) = \eta^{-2}$

désignant dans l'ordre, la fonction de densité, la moyenne et la variance.

À un instant t, si plusieurs précipitations se chevauchent, l'intensité totale est simplement la somme des intensités.

$$Y(t) = \sum_{k=1}^{N(t)} h(t, t_k, L_k)$$
(4.1)

où $h(t,t_k,L_k)$ est l'intensité au temps t de la précipitation qui a débutée au temps t_k

$$h(t, t_k, L_k) = \begin{cases} X_k & \text{si} \quad t_k \le t \le t_k + L_k \\ 0 & \text{autrement} \end{cases}$$
 (4.2)

Rodriguez-Iturbe et al. (1984) donnent les propriétés suivantes :

$$E(h(t,t_k,L_k)) = E(X)P(L_k > t - t_k)$$

$$= \frac{1}{\xi} \cdot e^{-\eta(t-t_k)}$$
(4.3)

$$E(Y(t)) = E\left[\int_{-\infty}^{t} h(t, t_{k}, L_{k}) dN(t_{k})\right]$$

$$= \int_{-\infty}^{t} E\left[h(t, t_{k}, L_{k})\right] E\left(N(t)\right) dt_{k}$$

$$= \int_{-\infty}^{t} \frac{1}{\xi} e^{-\eta(t-t_{k})} \lambda dt_{k}$$

$$= \frac{\lambda}{\xi \eta}$$
(4.4)

Pour calculer la variance et l'autocovariance de Y(t), il faut faire appel aux fonctions caractéristiques. Rappelons que la fonction caractéristique d'un vecteur aléatoire \mathbf{X} est donnée par $\phi_X(t) = E(e^{it\cdot \mathbf{X}})$. La fonction caractéristique de Y(t) est obtenue de la manière suivante [Parzen (1962)]. On pose

$$u_1 Y(t_1) + u_2 Y(t_2) = \sum_{m=1}^{N(t_2)} g(t_m, L_m)$$
(4.5)

où

$$g(t, L) = u_1 h(t_1, t, L) + u_2 h(t_2, t, L)$$
(4.6)

avec $t_1 < t_2$.

Pour évaluer la fonction caractéristique de Y(t), il suffit d'évaluer

$$\Phi = E[e^{iZ}] \tag{4.7}$$

où

$$Z = \sum_{m=1}^{N(t_2)} g(t_m, L_m)$$
 (4.8)

On définit la variable suivante

$$\varphi(t) = E\left[e^{ig(t,L)} - 1\right] \tag{4.9}$$

pour montrer que

$$\Phi = \exp\left\{\lambda \int_0^{t_2} \varphi(\tau) d\tau\right\} \tag{4.10}$$

Notons d'abord l'expression suivante

$$E[e^{iZ}] = \sum_{n=0}^{\infty} E[e^{iZ}|N(t_2) - N(0) = n]P[N(t_2) - N(0) = n]$$
(4.11)

À l'aide de la relation suivante [Parzen (1962), section 4-5]

$$E\left[\exp\left(i\sum_{m=1}^{n}g(t_{m},L_{m})\right)|N(t_{2})=n\right]=\left\{\frac{1}{t_{2}}\int_{0}^{t_{2}}E\left[\exp(ig(\tau,L)]d\tau\right]^{n}$$
(4.12)

on obtient

$$E\left[e^{iZ}\right] = \sum_{n=0}^{\infty} e^{-\lambda t_2} \frac{\left(\lambda t_2\right)^n}{n!} \left\{ \frac{1}{t_2} \int_0^{t_2} E\left[e^{ig(\tau,L)}\right] d\tau \right\}^n$$

$$= \exp\left\{\lambda \int_0^{t_2} \varphi(\tau) d\tau \right\}$$
(4.13)

donc

$$\varphi_{Y(t)}(u) = \exp\left\{\lambda \int_0^t E\left[e^{iuh(t,\tau,L)} - 1\right]d\tau\right\}$$
(4.14)

Pour trouver l'autocovariance, $(cov(Y(t_1), Y(t_2)))$, il faut la fonction caractéristique conjointe que l'on peut exprimer en transformant g(t, L) dans sa forme longue. En tenant compte que $h(t_1, t, L) = 0$ si $t > t_1$, on obtient

$$\varphi_{Y(t_1),Y(t_2)}(u_1,u_2) = \exp\left\{\lambda \int_0^{t_1} E\left[e^{i(u_1h(t_1,\tau,L)+u_2h(t_2,\tau,L))} - 1\right]d\tau + \lambda \int_{t_1}^{t_2} E\left[e^{iu_2h(t_2,\tau,L)} - 1\right]d\tau\right\}$$
(4.15)

Les différents moments sont trouvés en se servant des propriétés exprimées en (2.4). Voici un exemple de la façon de manipuler ces expressions. Exprimons d'abord le terme en espérance mathématique.

$$E\left[e^{i(u_{1}h(t_{1},\tau,L)+u_{2}h(t_{2},\tau,L))}\right] = e^{i(u_{1}X+u_{2}X)} \cdot P(\tau \leq t_{1}, t_{2} \leq \tau + L)$$

$$+ e^{iu_{1}X} \cdot P(\tau \leq t_{1} \leq \tau + L; t_{2} > \tau + L)$$

$$+ e^{iu_{2}X} \cdot P(\tau > t_{1}; \tau \leq t_{2} \leq \tau + L)$$

$$+ 1 \cdot P(t_{1}, t_{2} \notin [\tau, \tau + L])$$

$$(4.16)$$

Comme pour trouver la covariance, il faut dériver par rapport aux deux variables u_1 et u_2 . Il est clair que seul le premier terme du membre de droite de l'expression (4.16) intervient. Dans les conditions présentes nous pouvons inverser l'ordre de la dérivée et de l'intégrale pour obtenir

$$\frac{\partial^{2}}{\partial u_{1} \partial u_{2}} \log \varphi_{Y(t_{1}),Y(t_{2})}(u_{1},u_{2}) = \lambda \int_{0}^{t_{1}} i^{2} X \cdot X \cdot e^{i(u_{1}+u_{2})X} \cdot P(\tau \leq t_{1},t_{2} \leq \tau + L)$$

$$= i^{2} \lambda \int_{0}^{t_{1}} E[h(t_{1},\tau,L)h(t_{2},\tau,L)] d\tau$$
(4.17)

d'où nous pouvons déduire que

$$Cov(Y(t_1), Y(t_2)) = \lambda \int_0^{t_1} E[h(t_1, \tau, L)h(t_2, \tau, L)] d\tau$$

$$(4.18)$$

Pour avoir l'expression de l'autocovariance, il reste à trouver $E[h(t_1, \tau, L)h(t_2, \tau, L)]$:

$$E[h(t_1, \tau, L_k)h(t_2, \tau, L_k)] = E(X_k^2)P(t_k \ge t_2 - \tau)$$

$$= \frac{2}{\xi^2}e^{-\eta(t_2 - \tau)}$$
(4.19)

En remplaçant (4.19) dans (4.18), on obtient le résultat suivant :

$$Cov(Y(t_1), Y(t_2)) = \frac{2\lambda}{\eta \xi^2} e^{-\eta \tau} \quad (\tau = t_2 - t_1 \ge 0)$$
 (4.20)

En posant $t_1 = t_2$, $(\tau = 0)$, on obtient la variance

$$Var(Y(t)) = \frac{2\lambda}{\eta \xi^2}$$
 (4.21)

Comme les données dont nous disposons sont agrégées en totaux sur des intervalles de temps de longueurs égales h, nous devons connaître les caractéristiques du processus agrégé. Posons

$$Y_i^h = \int_{(i-1)h}^{ih} Y(s)ds$$
 $i = 1, 2, ...$ (4.22)

qui est l'intensité cumulative de précipitations. En fait, les Y_i^h correspondent à la forme des données de précipitations.

L'estimation des paramètres est effectuée par la méthode des moments. Dans le cas exponentiel (cas où f_X et f_L sont de lois exponentielles), il y a trois paramètres (λ, η, ξ) à estimer. En prenant l'espérance de chaque membre de (4.22) et en résolvant l'intégrale, on obtient

$$E(Y_i^h) = \frac{h\lambda}{n\mathcal{E}} \tag{4.23}$$

Une expression pour la covariance peut aussi être obtenue à l'aide de (4.20).

$$Cov(Y_1^h, Y_k^h) = \int_{(k-1)h}^{kh} dt_1 \int_0^h Cov(Y(t_1), Y(t_2)) dt_2 \quad k \ge 1$$
 (4.24)

En faisant les substitutions appropriées, on peut en tirer les expressions pour la variance et la covariance. Pour la variance, on doit tenir compte de la condition d'application de l'équation (4.20), i.e. $t_2 \ge t_1$, et on résout

$$Var(Y_i^h) = \frac{2\lambda}{\eta \xi^2} \int_0^h \left[\int_0^{t_2} \exp(-\eta (t_2 - t_1)) dt_1 \right] dt_2$$
 (4.25)

pour obtenir

$$Var(Y_i^h) = \frac{4\lambda}{\eta^3 \xi^2} (\eta h - 1 + e^{-\eta h})$$
 (k = 1)

Pour la covariance, le calcul se fait directement pour $k \ge 2$ et on obtient l'expression

$$Cov(Y_1^h, Y_k^h) = \frac{2\lambda}{\eta^3 \xi^2} (1 - e^{-\eta h})^2 e^{-\eta (k-2)h} \quad (k \ge 2)$$
 (4.27)

Pour estimer les paramètres, on utilise les expressions (4.23), (4.26) et (4.27) avec k=2. Le modèle préservera ainsi la moyenne, la variance et l'autocorrélation de délai un des données historiques. Par contre, il y a plusieurs propriétés qui ne seront pas préservées, dont entre autres la proportion d'intervalles (de longueur h) secs. Rodriguez-Iturbe et al. (1984) testent aussi la sensibilité aux changements d'échelles. Les paramètres, lorsqu'ils sont estimés pour des échelles de temps différentes, sont significativement différents. Or, une des qualités que nous recherchons dans un modèle est la consistance des paramètres d'une échelle à une autre.

4.1.2 Modèle bruits blancs de Poisson

Ce modèle associe à chaque événement d'un processus de Poisson N(t), de paramètre λ , une précipitation. Dans ce modèle, seule l'intensité instantanée X est déterminée aléatoirement. Ce modèle s'appuie sur le fait que les données historiques sont agrégées sur des intervalles de temps égaux et disjoints de longueur h. On définit le processus cumulé

de Poisson [Rodriguez-Iturbe et al., (1984)] par
$$N_X(t) = \sum_{n=0}^{N(t)} X_n$$
 pour $t > 0$

où $X_0 = 0$. Les expressions des moments pour un tel processus sont connues :

$$E(N_X(t)) = E(X)\lambda t$$

$$Var(N_X(t)) = E(X^2)\lambda t$$

$$Cov(N_X(t_1), N_X(t_2)) = E(X^2)\lambda t_1 \quad (t_1 < t_2)$$

$$(4.28)$$

Les données agrégées peuvent donc être représentées de la façon suivante :

$$Y_i^h = N_X(ih) - N_X((i-1)h) \tag{4.29}$$

et de cette expression on tire

$$E(Y) = E(X)\lambda h$$

$$Var(Y) = E(X^{2})\lambda h$$

$$Cov(Y_{i}, Y_{j}) = 0 \quad (i \neq j)$$

$$(4.30)$$

Si on suppose que X suit une loi exponentielle, on a $E(X^2) = 2E^2(X)$, ce qui réduit le nombre de paramètres à estimer et ainsi, il ne reste que λ et E(X) à déterminer, ce qui sera fait par la méthode des moments. Si la covariance estimée à partir des données est significativement différente de 0, le modèle n'est pas adéquat pour les données traitées. Ce modèle ne réagit pas très bien non plus aux changements d'échelle et la valeur des paramètres en est fortement affectée. Ce modèle est une forme simplifiée du précédent qui

facilite l'estimation des paramètres par le biais d'une représentation plus parcimonieuse du phénomène.

4.2 Modèles de Neyman-Scott (N-S)

4.2.1 Introduction

Kavvas et Delleur (1981) exposent les hypothèses qui leur ont permis d'appliquer le modèle de Neyman-Scott, originalement conçu pour modéliser la distribution des galaxies, à la modélisation des précipitations.

Ce type de modèle fait partie d'une famille de modèles que l'on appelle modèles « cluster ». Il s'agit de modèles stochastiques à deux niveaux. Au premier niveau, il y a le temps d'arrivée du groupe (cluster). Au second niveau, chaque groupe se décompose en plusieurs événements. L'analogie avec les précipitations est expliquée de la façon suivante. Le premier niveau correspond à l'arrivée de fronts de basse pression (groupe de précipitations). Les temps d'occurrences sont générés aléatoirement. Ensuite, au second niveau, à chaque front est associé un nombre aléatoire d'averses c (ou précipitations). Voici plus clairement les 5 hypothèses du modèle telles qu'énoncées par Kavvas et Delleur (1981).

- 1. Les précipitations se produisent sur l'axe du temps sous forme de groupes. On définit le processus de dénombrement N(t) qui compte le nombre de précipitations sur l'intervalle de temps [0, t].
- 2. Les temps d'occurrence des fronts forment un autre processus de dénombrement, R(t), qui compte le nombre de fronts sur l'intervalle de temps [0, t]. Les origines des fronts sont distribuées aléatoirement à partir de l'origine selon une loi de Poisson de paramètre λ . Pour un intervalle de longueur t on a donc

$$P(R(t) = r) = \frac{(\lambda t)^r \exp(-\lambda t)}{r!}$$

et la fonction génératrice de probabilité (FGP) de la variable R(t) pour le même intervalle est

$$G_R(z) = \sum_{r=0}^{\infty} z^{\lambda} P(R(t) = r)$$
$$= e^{-t\lambda(1-z)} \qquad |z| \le 1$$

- 3. À chaque front de basse pression correspond un groupe de précipitations dont la structure est la suivante : si le front a lieu au temps u, la variable c(u) est le nombre aléatoire de précipitations que le front génère et la vecteur T est l'ensemble des positions temporelles de ces précipitations.
- 4. Les variables aléatoires c sont mutuellement indépendantes et sont aussi indépendantes de toutes les autres variables aléatoires du processus.
- 5. Étant donné c(u), les positions temporelles \mathbf{T} des précipitations pour un front sont des variables aléatoires i.i.d. ayant pour fonction de densité $f_T(\tau-u)$, qui dépend seulement de la distance $\tau-u$ où u est la position de l'origine du front et τ est la position de la précipitation $(\tau \in \mathbf{T})$. Les positions temporelles dans \mathbf{T} sont aussi indépendantes mutuellement et indépendantes des autres variables aléatoires du processus.

Les processus de Poissons à taux fixe sont des processus stationnaires. Ces hypothèses impliquent que ce modèle, de type "cluster", ne peut être utilisé pour modéliser que des séries stationnaires.

Kavvas et Delleur (1981) présentent une méthode d'estimation des paramètres qui n'est pas détaillée ici. En effet, d'autres méthodes plus simples ont étés développées et le modèle est maintenant très différent de sa version originale. Ces auteurs montrent aussi que le processus de Poisson généralisé est un cas dégénéré du modèle de N-S.

4.2.2 Modèle avec bruits blancs de Neyman-Scott

Dans ce cas particulier, nous avons, comme dans la section précédente, le processus N(t) qui compte le nombre de précipitations dans l'intervalle de temps [0, t]. Il s'agit d'un processus de Poisson de paramètre λ . Les variables c et T sont utilisées comme dans les sections précédentes. Mais à cela s'ajoute le vecteur X qui est l'ensemble des intensités instantanées associées à chaque précipitation d'un front. On a donc [Waymire et Gupta (1981)]

$$\mathbf{T} = (T_1, T_2, \dots, T_{\nu})$$

$$\mathbf{X} = (X_1, X_2, \dots, X_{\nu})$$

$$(4.31)$$

où les T_j sont i.i.d. $f_T(t)$ et les X_i sont également i.i.d. L'expression pour l'intensité totale est la suivante :

$$Y(t)dt = X(t)dN_1(t) (4.32)$$

où

$$dN_1(t) = \begin{cases} 1 & \text{s'il y a une précipitation dans l'intervalle } (t, t + dt) \\ 0 & \text{autrement} \end{cases}$$

et

$$N_1(t) = \int dN_1(s)$$
 (4.33)

En prenant l'espérance de chaque côté dans (4.32), on obtient

$$E(Y(t)) = E\left[X(t) \cdot \frac{dN_1(t)}{dt}\right] \tag{4.34}$$

où $dN_1(t)/dt$ est le taux de variation du nombre de précipitations sur un intervalle de temps par rapport au temps. Or, les paramètres λ et c sont respectivement, le nombre moyen d'occurrence de fronts par unité de temps et le nombre de précipitations par front. En considérant $E(\lambda c)$, on obtient le nombre moyen de précipitations par unité de temps. Donc, l'expression pour l'intensité moyenne est

$$E(Y(t)) = E(X(t))E(c)\lambda \tag{4.35}$$

Rodriguez-Iturbe et al. (1984) donnent la propriété suivante

$$E[Y(t), Y(t+\tau)] = E(X^{2})E[N'_{1}(t)]\delta(\tau) + \mu_{X}^{2}E[N'_{1}(t)N'_{1}(t+\tau)]$$
(4.36)

où $\delta(\tau)$ est la fonction de Dirac et $N_1(t) = dN_1(t)/dt$. On peut déduire de (4.36) la relation

$$Cov[Y(t), Y(t+\tau)] = E(X^2)E[N_1'(t)]\delta(\tau) + \mu_X^2 Cov[N_1'(t)N_1'(t+\tau)]$$
 (4.37)

À l'aide de la fonction génératrice de probabilités, Waymire et Gupta (1981b) montrent que

$$Cov[Y(t), Y(t+\tau)] = E(X^{2})\lambda E(c)\delta(\tau) +$$

$$\mu_{X}^{2}\lambda E(c^{2}-c)\int_{-\infty}^{\infty} f_{T}(u)f_{T}(u+\tau)du$$
(4.38)

pour $\tau \ge 0$. En supposant que $f_T(t) = \beta e^{-\beta t}$, (exponentielle de paramètre $\beta^{-1} > 0$), que X est distribué selon une loi exponentielle également et que c suit une loi de Poisson, on obtient :

$$E(X^{2}) = 2\mu_{X}^{2}$$

$$E(c^{2}) = E(c) + E^{2}(c)$$
(4.39)

et en substituant dans(4.38), on obtient

$$Cov(Y(t), Y(t+\tau)) = \theta_1 \delta(\tau) + \theta_2 e^{-\beta|\tau|} \qquad -\infty < \tau < \infty$$
(4.40)

où $\theta_1=2\lambda\mu_X^2E(c)$ et $\theta_2=\beta\lambda\mu_X^2E^2(c)$. Pour $\tau=0$ on obtient

$$Var(Y(t)) = \theta_1 \delta(0) + \theta_2 \tag{4.41}$$

Les quatre paramètres μ_X , E(v), λ et β sont inconnus. Mais, tel que mentionné pour les processus de Poisson, les données sont agrégées et il faut connaître les propriétés du processus agrégé. Le processus discret sera noté de la façon suivante :

$$Y_i^h = \int_{(i-1)h}^{ih} Y(s)ds \quad i \ge 1$$
 (4.42)

où T est l'échelle considérée. Les Y_i^T sont identiquement distribués et on peut déduire de (4.35), (4.40) et (4.41) les expressions pour l'espérance, la variance et la covariance :

$$E(Y_i^h) = h\mu_X E(c)\lambda \tag{4.43a}$$

$$Var(Y_i^h) = \theta_1 h + \frac{2\theta_2}{\beta^2} \left[\beta h - 1 + e^{-\beta h}\right]$$
 (4.43b)

$$Cov(Y_1^h, Y_k^h) = \frac{\theta_2}{\beta^2} (1 - e^{-\beta h}) e^{-\beta(k-2)h} \quad (k \ge 2)$$
 (4.43c)

Pour avoir l'autocorrélation de délai un, on prend k = 2 et on peut supposer h = 1,

$$\rho_{\gamma}(1) = \frac{\theta_2 (1 - e^{-\beta})^2}{\theta_1 \beta^2 + 2\theta_2 [\beta - 1 + e^{-\beta}]}$$
(4.44)

Pour l'estimation des paramètres, on procède comme suit :

- on fixe β
- on résout (4.43a), (4.43b)et (4.44) simultanément
- on compare la valeur théorique donnée par (4.43c) et celle que donnent les données historiques
- on ajuste β au besoin et on recommence si nécessaire.

4.2.3 Modèle de N-S à impulsions rectangulaires (NSIR)

Sur le même principe que le modèle de Poisson avec impulsions rectangulaires, on associe à chaque précipitation une intensité et une durée. La différence est que dans ce cas-ci, les précipitations sont issues d'un front de basse pression simulé. En conservant la même notation que celle utilisée dans la section 4.1.1, regardons les propriétés du processus qui décrit l'intensité de précipitation à un instant t [Rodriguez-Iturbe et al. (1987a)].

$$Y(t) = \sum_{n=0}^{R(t)} \sum_{k=0}^{\nu_n} h(t, t_{n,k}, L_{n,k})$$

$$= \sum_{k=0}^{N(t)} h(t, t_k, L_k)$$
(4.45)

En prenant l'espérance de chaque côté, on obtient (comme dans la section 4.1.1)

$$E(Y(t)) = E\left[\int_{-\infty}^{t} h(t, t_k, L_k) dN(t_k)\right]$$

$$= \frac{\lambda}{\eta \xi} E(c)$$
(4.46)

Rodriguez-Iturbe et al. (1987a) obtiennent l'autocovariance de la manière suivante :

$$\operatorname{cov}[Y(t), Y(t+\tau)] = \int_0^\infty \int_0^\infty E[X(u)X(v)] \cdot \operatorname{cov}[dN(t-u), dN(t+\tau-v)]$$
(4.47)

L'expression $cov[dN(t-u), dN(t+\tau-v)]$ dépend de la fonction d'intensité conditionnelle h(u), dont les détails sont donnés dans Cox et Isham (1980), et peut s'écrire de la façon suivante:

$$cov(dN(t), dN(t+u)) = \lambda E(c) \cdot (\delta(u) + h(u) - \lambda E(c)) dt du$$
 (4.48)

où

$$h(u) = \lambda E(c) + \frac{1}{2E(c)} E(c^2 - c) \beta e^{-\beta u}$$

$$(4.49)$$

Sachant que

$$X(u) = \begin{cases} X & avec \ probabilit\'e \ e^{-\eta u} \\ 0 & avec \ probabilit\'e \ 1 - e^{-\eta u} \end{cases}$$
 (4.50)

on peut déduire la relation pour l'autocovariance

$$cov[Y(t), Y(t+h)] = \frac{\lambda e^{-\eta h}}{\eta} \left[E(c)E(X^2) + \frac{\mu_X^2}{2} E(c^2 - c) \frac{\beta^2}{\beta^2 - \eta^2} \right] - \frac{\lambda \mu_X^2}{2} E(c^2 - c) \frac{\beta e^{-\beta h}}{\beta^2 - \eta^2} \tag{4.51}$$

En posant h = 0, on obtient la variance

$$\operatorname{var}(Y(t)) = \frac{\lambda E(X^2)E(c)}{\eta} + \frac{\lambda \mu_X^2}{2\eta}E(c^2 - c)\frac{\beta}{\beta + \eta}$$
(4.52)

Connaissant maintenant les propriétés du processus continu, il faut, comme pour les modèles précédents, déterminer les relations pour le processus agrégé, ou discrétisé.

Si on pose N(t), un processus qui compte le nombre de précipitations dans l'intervalle (0,t), le taux d'occurrence de ce processus est $\lambda E(c)$. Rodriguez-Iturbe et al. (1987b) expriment le processus agrégé sous la forme suivante :

$$Y_i^h = \int_{(i-1)h}^{ih} Y(s)ds \quad i \ge 1$$
 (4.53)

on peut déduire les caractéristiques statistiques suivantes.

$$E(Y_i^h) = \rho E(c) \mu_X h \tag{4.54a}$$

$$Var(Y_{i}^{h}) = \left[\frac{\rho}{\eta^{2}} \left(\eta h - 1 + e^{-\eta h}\right)\right] \cdot \left[2E(c)E(X^{2}) + \frac{E(c^{2} - c)\mu_{X}^{2}\beta^{2}}{\beta^{2} - \eta^{2}}\right] - \lambda \left(\beta h - 1 + e^{-\beta h}\right) \cdot \frac{E(c^{2} - c)\mu_{X}^{2}}{\beta(\beta^{2} - \eta^{2})}$$
(4.54b)

$$Cov(Y_{i}^{h}, Y_{i+k}^{h}) = \left[\frac{\rho}{\eta^{2}} (1 - e^{-\eta h})^{2} e^{-\eta(k-1)h}\right] \cdot \left[E(c)E(X^{2}) + \frac{1}{2} \frac{E(c^{2} - c)\mu_{X}^{2}\beta^{2}}{\beta^{2} - \eta^{2}}\right]$$
$$-\lambda (1 - e^{-\beta h})^{2} e^{-\beta(k-1)h} \left[\frac{1}{2} \frac{E(c^{2} - c)\mu_{X}^{2}}{\beta(\beta^{2} - \eta^{2})}\right] \qquad (k \ge 1)$$

$$(4.54c)$$

où $\rho = \lambda/\eta$. c peut suivre une loi géométrique ou une loi de Poisson sans que cela ait un impact significatif et le choix ne change en rien la forme générale de la fonction d'autocovariance. Si c est de loi géométrique, alors on a

$$P(c=j) = E^{-1}(c)(1 - E^{-1}(c))^{j-1} j = 1,2,... (4.55)$$

$$E(c^{2}-c) = 2E(c)(E(c)-1)$$
(4.56)

Si toutefois c-1 est de loi de Poisson (c-1 car c doit être strictement positive) alors on a

$$E(c^{2} - c) = E(c)(E(c) + 2)$$
(4.57)

4.2.4 Modèle de Neyman-Scott η-aléatoire

La description du modèle η -aléatoire est exactement la même que celle du modèle à impulsions rectangulaires, à l'exception que le paramètre de la loi exponentielle qui gère la durée des précipitations à l'intérieur d'un front devient aléatoire suivant une loi gamma de paramètres α et c. Cette nouvelle variable aléatoire est considérée comme indépendante de toutes les autres du processus. L'article d'*Entekhabi et al.* (1989) donne le développement des expressions pour le processus agrégé.

$$E(Y_i^h) = E(c)\mu_Y \lambda h \cdot I(1,0) \tag{4.58a}$$

$$Var(Y_{i}^{h}) = \left[E(c)\mu_{X}\lambda h \cdot I(1,0)\right]^{2} + \left\{2C_{1}h + C_{2}\beta^{-3}(\beta h + e^{-\beta T} - 1) + \left(E(v)\mu_{X}\lambda h\right)^{2}\right\}I(2,0) - 2C_{1}I(3,0) - C_{2}T \cdot I(4,0) + C_{2}I(5,0) + 2C_{1}I(3,h) - C_{2}I(5,h)$$

$$(4.58b)$$

$$Cov(Y_i^h, Y_{i+k}^h) = C_1 I(3, kh - h) - 2C_1 I(3, kh) + C_1 I(3, kh + h) - \frac{1}{2} C_2 I(5, kh - h)$$

$$+ C_2 I(5, kh) - \frac{1}{2} C_2 I(5, kh + h) + \frac{1}{2} C_2 \beta^{-3} (1 - e^{-\beta h})^2 e^{-\beta (k-1)h} I(2, 0)$$

$$+ \left[E(c) \mu_X \lambda h \cdot I(1, 0) \right]^2 \cdot \left[I(2, 0) - I^2(1, 0) \right]$$
(4.58c)

où on a

$$C_{1} = \lambda E(c)E(X^{2})$$

$$C_{2} = \lambda E(c^{2} - c)\mu_{Y}^{2}\beta^{2}$$
(4.59)

$$I(x,y) = E\left(\eta^{-x}e^{-\eta y}\right) = \frac{\Gamma(\alpha - x)}{\Gamma(\alpha)} v^{\alpha} (\nu + y)^{x - \alpha} \qquad x > 0, y \ge 0$$
 (4.60)

Pour avoir les simplifications habituelles, on peut supposer que X suit une loi exponentielle et que c suit une loi géométrique ou une loi de Poisson.

Pour estimer les paramètres, *Entekhabi et al.* (1989) suggèrent la méthode suivante. Soit \mathbf{u} le vecteur $(n \times 1)$ de paramètres à estimer. On pose la fonction $f(\mathbf{u}) = [f_1(\mathbf{u}), ..., f_n(\mathbf{u})]$ qui contient les expressions du modèle pour les moments correspondant au vecteur $\hat{\mathbf{f}} = (\hat{f}_1, ..., \hat{f}_n)$ de moments estimés à différents niveaux d'agrégation à partir des données historiques. Pour que le modèle s'ajuste parfaitement, il faudrait que $\hat{\mathbf{f}} - f(\mathbf{u}) = \mathbf{0}$. On pose donc $\hat{\mathbf{F}} = diag(\hat{\mathbf{f}})$ et $F(\mathbf{u}) = diag(f(\mathbf{u}))$ et on minimise l'expression

$$trace\left\{ \left[\mathbf{I} - F(\mathbf{u})\hat{\mathbf{F}}^{-1}\right]^{2}\right\}$$
(4.61)

D'après *Entekhabi et al.* (1989), cette méthode est une façon pratique et rapide pour estimer les paramètres.

4.3 Modèles de Bartlett-Lewis (B-L)

4.3.1 Introduction

Le modèle de B-L tel que présenté par *Rodriguez-Iturbe et al* (1987a) est, selon eux, la forme de modèle de type « cluster » d'utilisation la plus directe pour les applications. Ce

modèle est très ressemblant au modèle de N-S et dans bien des cas, donne à peu près les mêmes performances.

Tout comme dans N-S, les fronts se produisent selon un processus de Poisson de paramètre λ . Mais plutôt que de générer des positions temporelles aléatoires pour les précipitations engendrées par le fronts, le modèle de B-L démarre un nouveau processus de Poisson de paramètre β correspondant aux temps d'origine des précipitations. Ce processus arrête au bout d'un certain temps déterminé aléatoirement selon une loi exponentielle de paramètre γ . Contrairement au processus de N-S, on associe une précipitation à l'origine du front.

4.3.2 Modèle de Bartlett-Lewis avec impulsions rectangulaires (BLIRP)

Avec la même notation que pour les modèles précédents avec impulsions rectangulaires, nous avons X pour l'intensité de l'impulsion, la durée de l'impulsion (L) suit une loi exponentielle de paramètre η . Il sera pratique d'utiliser les paramètres sans dimensions suivants : $\kappa = \beta/\eta$ et $\phi = \gamma/\eta$. Le nombre c de précipitations par front suit donc une loi géométrique avec $E(c) = 1 + \kappa/\phi$.

On dira d'un front qu'il est actif tant que la dernière précipitation ne sera pas terminée. Pour sa part, le front se termine quand le processus qui génère les précipitations se termine. Ainsi, un front peut être encore actif, même s'il a terminé de générer des précipitations. Rodriguez-Iturbe et al. (1987a, b) définissent $p_r(t)$ comme étant la probabilité qu'au temps t, il y ait r précipitations actives sans toutefois que le front soit terminé et $q_r(t)$ comme étant la probabilité qu'au temps t, il y ait encore r précipitations actives et que le front soit terminé. Ils définissent aussi les fonctions génératrices suivantes :

$$G_P^*(z,s) = \sum_{r=0}^{\infty} \int_0^{\infty} z^r e^{-st} p_r(t) dt$$
 (4.62a)

$$G_{Q}^{*}(z,s) = \sum_{r=0}^{\infty} \int_{0}^{\infty} z^{r} e^{-st} q_{r}(t) dt$$
 (4.62b)

Avec $p_r(t)$, $q_r(t)$, (4.62a) et (4.62b), ils trouvent une expression pour la durée moyenne pendant laquelle un front est actif

$$\mu_{a} = \frac{\phi}{\eta} \int_{0}^{1} du \int_{0}^{1} dt \, \frac{t^{\phi - 1}}{u} \Big[1 - (1 - ut)e^{-\kappa u(1 - t)} \Big]$$
(4.63)

Il ne faut pas confondre cette expression avec le temps moyen avant que le front termine, qui est simplement γ^{-1} . L'expression (4.63) doit être évaluée numériquement ou encore, lorsque κ et ϕ sont petits, on obtient

$$\mu_{a} \approx \gamma^{-1} \left[1 + \phi(\kappa + \phi) - \frac{1}{4} \phi(\kappa + \phi)(\kappa + 4\phi) + \frac{1}{72} \phi(\kappa + \phi)(4\kappa^{2} + 27\kappa\phi + 72\phi^{2}) \right]$$
(4.64)

Si C est le nombre de précipitations encore actives à un temps donné, alors :

$$P(C = r) = \mu_a^{-1} \int_0^\infty (p_r(t) + q_r(t)) dt \qquad (r \ge 1)$$

$$P(C = 0) = \mu_a^{-1} \int_0^\infty p_r(t) dt \qquad (4.65)$$

Nous avons donc la fonction génératrice de probabilités suivante :

$$G_C(z) = \mu_a^{-1} \lim_{s \to 0} \left[G_P^*(z, s) + G_Q^*(z, s) - q_0^*(s) \right]$$
(4.66)

où $q_0^*(s) = G_{\mathcal{Q}}^*(0,s)$. On en tire donc les moments pour la distribution de C:

$$E(C) = \frac{\kappa + \phi}{\mu_a \eta \phi}$$

$$E(C^2) = \frac{(\kappa + \phi)(\kappa + \phi + 1)}{\mu_a \eta \phi(\phi + 1)}$$

$$E(C^3) = \frac{\kappa^2 (2\phi(\phi + 3) - \kappa(\phi - 3)(\phi + 2)}{\mu_a \eta \phi(\phi + 1)(\phi + 2)(\phi + 3)} + 3E(C^2) - 2E(C)$$
(4.67)

Comme les origines des fronts sont distribuées selon un processus de Poisson et que chaque front a une durée d'activité moyenne de μ_a , alors le nombre de fronts actifs à un instant donné doit suivre une loi de Poisson de moyenne $\lambda\mu_a$. Donc la fonction génératrice des probabilités du nombre total de précipitations actives, issues de tous les fronts actifs, est

$$\exp\left(-\lambda\mu_{\alpha}(1-G_{C}(z))\right) \tag{4.68}$$

On peut en tirer la probabilité qu'il n'y ait pas de précipitation à un instant donné

$$\pi_0 = \exp\left(-\lambda \mu_a + \lambda G_P^*(0,0)\right) \tag{4.69}$$

La valeur exacte de $G_p^*(0,0)$ doit être évaluée numériquement. Mais lorsque κ et ϕ sont petits, on a l'estimation suivante :

$$G_P^*(0,0) \approx \gamma^{-1} \left(1 - \kappa - \phi + \frac{3}{2} \kappa \phi + \phi^2 + \frac{1}{2} \kappa^2 \right)$$

$$\Rightarrow \pi_0 \approx \exp\left\{ -\rho \left(1 + \kappa/\phi \right) \left(1 - \kappa/2 \right) \right\}$$
(4.70)

La connaissance de cette probabilité pour un intervalle de temps de longueur *h* sera utile pour le processus agrégé. *Rodriguez-Iturbe et al.* (1987a) trouvent la façon de l'exprimer :

$$\omega_h = \exp\left\{-\lambda(h+\mu_a) + \lambda G_P^*(0,0) \frac{\gamma + \beta e^{-(\beta+\gamma)h}}{\beta+\gamma}\right\}$$
(4.71)

qui satisfait $\omega_0 = \pi_0$. Ces mêmes auteurs déterminent aussi la durée moyenne des périodes sèches (temps entre la fin d'une période humide et le début de la suivante)

$$\mu_D = \lambda^{-1} \left[1 + \beta G_P^*(0,0) \right]^{-1} \tag{4.72}$$

Pour cette dernière, l'approximation lorsque κ et ϕ sont petits est donnée par :

$$\mu_D \approx \lambda^{-1} \left[1 + \kappa \phi^{-1} \left(1 - \kappa - \phi + \frac{3}{2} \kappa \phi + \phi^2 + \frac{1}{2} \kappa^2 \right) \right]^{-1}$$
(4.73)

On peut ensuite déduire la durée moyenne de période humide μ_{W} de la condition simple suivante :

$$\pi_0 = \frac{\mu_D}{\mu_D + \mu_W}$$

$$\Rightarrow \mu_W = \mu_D \left(\pi_0^{-1} - 1\right)$$
(4.74)

Il ne reste plus qu'à connaître les propriétés du processus agrégé. Rodriguez-Iturbe et al. (1987a) obtiennent :

$$E(Y_i^h) = h\rho E(c)\mu_X \tag{4.75a}$$

$$Var(Y_i^h) = 2\rho E(c) \left\{ E(X^2) + \frac{\beta \mu_X^2}{\gamma} \right\} \frac{h}{\eta} + 2\rho E(c) \mu_X^2 \beta \eta \frac{1 - e^{-\gamma h}}{\gamma^2 (\gamma^2 - \eta^2)}$$

$$-2\rho E(c) \left\{ E(X^2) + \frac{\beta \mu_X^2}{\gamma^2 - \eta^2} \right\} \frac{1 - e^{-\gamma h}}{\eta^2}$$
(4.75b)

$$Cov(Y_{i}^{h}, Y_{i+k}^{h}) = \rho E(c) \left\{ E(X^{2}) + \frac{\beta \mu_{X}^{2}}{\gamma^{2} - \eta^{2}} \right\} (1 - e^{-\gamma h})^{2} \frac{e^{-\eta(k-1)h}}{\eta^{2}}$$

$$-\rho E(c) \mu_{X}^{2} \beta \eta (1 - e^{-\gamma h})^{2} \frac{e^{-\eta(k-1)h}}{\gamma^{2} (\gamma^{2} - \eta^{2})}$$

$$(4.75c)$$

On obtient un modèle à cinq paramètres, à savoir λ , η , μ_X , E(c) et β . L'estimation des paramètres peut se faire de la même façon que pour le modèle de Neyman-Scott [Rodriguez-Iturbe et al. (1987b)], c'est-à-dire en minimisant une fonction objectif qui est une somme des carrés des différences entre les moments historiques et calculés. Cependant, il faut noter que l'estimation des paramètres pour le modèle de Bartlett-Lewis n'est pas une tâche facile, comme nous le verrons dans le chapitre 5 où on en discute plus en détails.

Ce modèle, comme nous l'avons mentionné lors de l'introduction, donne des performances comparables à celles du modèle de N-S, tout en offrant une maniabilité relativement plus facile des expressions analytiques.

4.3.3 Modèle de Bartlett-Lewis η -aléatoire

Le modèle de B-L, tel que décrit dans la section précédente, s'ajuste très bien aux données historiques, sauf dans le cas de la proportion de périodes sèches. Celles-ci sont systématiquement surestimées. Cette faiblesse dans le modèle pourrait devenir grave dans le cas où nous voudrions nous servir des données simulées comme intrants d'un modèle pluie-débit. En effet, le ruissellement en serait fortement affecté à cause des caractéristiques d'infiltration du sol, qui sont évidemment très sensibles aux périodes sèches. Le modèle η -aléatoire change la structure des fronts en faisant en sorte que chaque front génère des précipitations provenant d'une population différente d'un front à l'autre.

[Rodriguez-Iturbe et al. (1988)] ont suggéré de rendre aléatoire le paramètre de la loi exponentielle qui gère la durée des précipitations. Ils ont donc proposé que les η soient i.i.d. suivant une loi $\Gamma(\alpha, \nu)$ avec $\alpha > 1$, le paramètre de forme et ν le paramètre d'échelle et de fonction de densité

$$f_{\eta}(\theta) = \frac{1}{\Gamma(\alpha)\nu^{\alpha}} \theta^{\alpha-1} e^{-\theta/\nu}$$
(4.76)

Comme dans le modèle précédent, on pose $\kappa = \beta/\eta$ et $\phi = \gamma/\eta$ et on fixe leur valeur. Ainsi, β et γ varieront aussi. La transition du modèle η -fixe à η -aléatoire se fait assez directement. Dans l'expression (4.63) de μ_a , η intervient seulement comme un facteur multiplicatif η^{-1} . On le remplace donc simplement par $E(\eta^{-1}) = v/(\alpha - 1)$, la durée moyenne des précipitations. Aussi, dans la fonction génératrice de probabilités de C, le nombre de précipitations actives à un instant donné est un facteur de μ_a η qui est constant, donc la distribution de C reste inchangée. Rodriguez-Iturbe et al. (1988) donnent les expressions pour le processus agrégé en tenant compte du η aléatoire. On a

$$E(Y_i^h) = \frac{\nu}{\alpha - 1} h \lambda (1 + \kappa/\phi) \mu_X \tag{4.77a}$$

$$Var(Y_i^h) = 2A_1 \left\{ (\alpha - 3)h v^{2-\alpha} - v^{3-\alpha} + (\nu + h)^{3-\alpha} \right\}$$

$$-2A_2 \left\{ \phi(\alpha - 3)h v^{2-\alpha} - v^{3-\alpha} + (\nu + \phi h)^{3-\alpha} \right\}$$
(4.77b)

$$Cov(Y_{i}^{h}, Y_{i+k}^{h}) = A_{1} \left\{ \left(\nu + (k+1)h \right)^{3-\alpha} - 2(\nu + kh)^{3-\alpha} + \left(\nu + (k-1)h \right)^{3-\alpha} \right\}$$

$$- A_{2} \left\{ \left(\nu + (k+1)\phi h \right)^{3-\alpha} - 2(\nu + k\phi h)^{3-\alpha} + \left(\nu + (k-1)\phi h \right)^{3-\alpha} \right\}$$

$$(k>0)$$

$$(4.77c)$$

avec

$$A_{1} = \frac{\lambda E(c)v^{\alpha}}{(\alpha - 1)(\alpha - 2)(\alpha - 3)} \left[E(X^{2}) + \frac{\kappa \phi \mu_{X}}{\phi^{2} - 1} \right]$$
$$A_{2} = \frac{\lambda E(c)\kappa \mu_{X}^{2}v^{\alpha}}{\phi^{2}(\phi^{2} - 1)(\alpha - 1)(\alpha - 2)(\alpha - 3)}$$

La probabilité qu'une période de longueur h soit sèche est donnée par l'expression [Onof et Wheater (1993)]

$$P(Y_i^h = 0) \approx \exp\left\{-\lambda h - \frac{\lambda \nu}{\phi(\alpha - 1)} \left[1 + \phi(\kappa + \phi/2) - \frac{1}{4}\phi(5\phi\kappa + \kappa^2 + 2\phi^2)\right]\right\}$$

$$\frac{1}{72}\phi(4\kappa^3 + 31\kappa^2\phi + 99\kappa\phi^2 + 36\phi^3) + \frac{\lambda \nu}{\phi(\alpha - 1)}$$

$$\times \left(1 - \kappa - \phi + \frac{3}{2}\kappa\phi + \phi^2 + \frac{1}{2}\kappa^2\right) \cdot \left(\frac{\phi}{\phi + \kappa} + \frac{\kappa}{\phi + \kappa} \left[\frac{\nu}{\nu + (\kappa + \phi)h}\right]^{\alpha - 1}\right)\right\}$$

$$(4.78)$$

Pour l'estimation des paramètres, *Onof et Wheater* (1993) utilisent deux ensembles d'équations mais un seul a donné des résultats. Il s'agit de l'ensemble suivant : la moyenne, la variance et l'autocovariance de délai un pour l'échelle horaire (h=1), la proportion de périodes sèches pour l'échelle horaire et journalière (h=1) et la variance pour h=6.

Le modèle réussit à corriger ce que le modèle de B-L à impulsions rectangulaire ne faisait pas bien, à savoir s'ajuster aux proportions de périodes sèches. Toutefois *Onof et Wheater* (1993) soulignent que les deux caractéristiques statistiques suivantes ne sont pas préservées :

- quelques autocorrélations pour des délais plus grands que 12 heures
- les valeurs extrêmes des périodes de retour plus grandes que la longueur des données historiques

Onof et Wheater (1993) donnent aussi des expressions pour les variables caractéristiques suivantes :

- la probabilité qu'une période sèche de longueur $\geq h$ succède à une période pluvieuse
- la distribution de probabilité pour la période entre deux événements
- la moyenne et la variance de la durée des périodes entre deux événements
- le nombre moyen de précipitations par mois
- la durée moyenne d'événements pluvieux en données horaires

4.3.4 Introduction d'une perturbation (jitter) dans le modèle à impulsions rectangulaires

Les modèles présentés utilisent souvent des cellules de pluie de forme rectangulaire parce que la paramétrisation est simplifiée. Les modèles à impulsions rectangulaires possèdent jusqu'à cinq paramètres et leur estimation des complexe. Tout ajout de nouveaux paramètres vient considérablement alourdir la phase d'estimation. Un moyen simple de donner des formes moins régulières, donc plus réalistes, aux cellules de pluie, a été proposée par *Rodriguez-Iturbe et al.* (1987a) et reprise par *Onof et Wheater* (1994a).

Rodriguez-Iturbe et al. (1987a) proposent d'introduire un perturbateur dans un modèle à impulsions rectangulaires. Le processus perturbé $\widetilde{Y}(t)$ aura la forme

$$\widetilde{Y}(t) = e^{Z(t)} \cdot Y(t) \tag{4.79}$$

où Z(t) est un processus aléatoire normal indépendant de tous les autres processus du modèle. En supposant que $E[e^{Z(t)}] = 1$, selon les propriétés de la loi log-normale [Casella et Berger (1990)], on aura :

$$\mu_Z + \frac{1}{2}\sigma_Z^2 = 0 \tag{4.80}$$

où μ_Z et σ_Z^2 sont respectivement la moyenne et la variance de Z(t). Sous cette hypothèse, on s'assure que $E\big[\widetilde{Y}(t)\big] = E\big[Y(t)\big]$ et $\widetilde{Y}(t) = 0$, si et seulement si Y(t) = 0. Ceci fait en sorte que la moyenne et les périodes sèches ne sont pas affectées par la perturbation. La forme exponentielle permet aussi de conserver la positivité du processus.

Casella et Berger (1990) donnent la propriété suivante pour la loi log-normale:

$$Var(e^{Z(t)}) = e^{2(\mu_Z + \sigma_Z^2)} - e^{2\mu_Z + \sigma_Z^2}$$
 (4.81a)

La condition (4.80) permet d'affirmer que dans le cas présent on a $\mu_Z = -\frac{1}{2}\sigma_Z^2$. On obtient donc la simplification suivante

$$Var(e^{Z(t)}) = e^{\sigma_Z^2} - 1. \tag{4.81b}$$

En décomposant l'expression de $Var(\widetilde{Y}(t))$ de la façon suivante :

$$Var(\widetilde{Y}(t)) = Var(e^{Z(t)}Y(t))$$

$$= E[(e^{Z(t)})^2 Y^2(t)] - E^2[e^{Z(t)}Y(t)]$$

$$= [Var(e^{Z(t)}) + 1] \cdot [\sigma_Y^2 + \mu_Y^2] - \mu_Y^2$$

$$(4.82)$$

il en découle directement l'expression pour la variance du processus perturbé :

$$Var\left[\widetilde{Y}(t)\right] = \sigma_Y^2 e^{\sigma_Z^2} + \mu_Y^2 \left(e^{\sigma_Z^2} - 1\right) \tag{4.83}$$

Et de la même façon, on obtient :

$$Cov[\widetilde{Y}(t), \widetilde{Y}(t+\tau)] = c_{\gamma}(\tau) \exp\{c_{Z}(\tau)\} + \mu_{\gamma}^{2} \left[\exp\{c_{Z}(\tau)\} - 1\right]$$
(4.84)

où pour simplifier l'expression, nous utilisons la notation $c_Y(\tau) = Cov[Y(t), Y(t+\tau)]$ et $c_Z(\tau) = Cov[Z(t), Z(t+\tau)]$. Dans le cas où σ_Z^2 est petit, on peut réécrire (4.83) et (4.84) sous la forme suivante

$$Var\left[\widetilde{Y}(t)\right] = \sigma_Y^2 + \sigma_Z^2 \left(\mu_Y^2 + \sigma_Y^2\right) + O\left(\sigma_Z^4\right) \tag{4.85}$$

$$Cov[\widetilde{Y}(t), \widetilde{Y}(t+\tau)] = c_{Y}(\tau) + c_{Z}(\tau)[\mu_{Y}^{2} + c_{Y}(\tau)]$$
(4.86)

Ainsi, l'autocorrélation décroîtra si et seulement si la condition suivante est respectée [Onof et Wheater (1994a)]

$$\frac{Cov[Z(t), Z(t+\tau)]}{\sigma_Z^2} < \frac{Cov[Y(t), Y(t+\tau)]}{\sigma_Y^2}$$
(4.87)

Les moments du processus discrétisé dans le cas où l'équation (4.87) est respectée sont :

$$E\left[\widetilde{Y}_{i}^{h}\right] = h \cdot E[Y] \tag{4.88a}$$

$$Var\left[\widetilde{Y}_{i}^{h}\right] = Var\left[Y_{i}^{h}\right] + 2h\left\{\sigma_{Y}^{2} + \mu_{Y}^{2}\right\} \int_{0}^{\infty} c_{z}(u)du$$
(4.88b)

$$Cov\left[\widetilde{Y}_{i}^{h},\widetilde{Y}_{i+k}^{h}\right] = Cov\left[Y_{i}^{h},Y_{i+k}^{h}\right] + h\left\{\sigma_{\gamma}^{2} + \mu_{\gamma}^{2}\right\} \int_{0}^{\infty} c_{Z}(kh + \nu)d\nu \tag{4.88c}$$

où le second terme est négligeable pour k > 1.

Pour pouvoir continuer, il faut définir l'autocovariance $c_z(\tau)$. On pose donc

$$c_Z(\tau) = 10^{-2} \exp\left[-\zeta \tau\right] \tag{4.89}$$

avec ζ suffisamment grand pour avoir $c_Z(\tau) << c_Y(\tau)$. Notons que si $\tau=0$, on a $\sigma_Z^2=10^{-2}$, qui est petit. Selon cette hypothèse, les relations (4.88a-c) deviennent

$$E\left[\widetilde{Y}_{i}^{h}\right] = h \cdot E[Y] \tag{4.90a}$$

$$Var\left[\widetilde{Y}_{i}^{h}\right] = Var\left[Y_{i}^{h}\right] + 2h\left\{\sigma_{Y}^{2} + \mu_{Y}^{2}\right\} \frac{10^{-2}}{\zeta}$$

$$\tag{4.90b}$$

$$Cov[\widetilde{Y}_{i}^{h}, \widetilde{Y}_{i+1}^{h}] = Cov[Y_{i}^{h}, Y_{i+1}^{h}] + h\{\sigma_{Y}^{2} + \mu_{Y}^{2}\} \frac{\exp(-\zeta h)10^{-2}}{\zeta}$$
(4.90c)

et pour k > 1,

$$Cov\left[\widetilde{Y}_{i}^{h}, \widetilde{Y}_{i+k}^{h}\right] = Cov\left[Y_{i}^{h}, Y_{i+k}^{h}\right]$$

$$(4.90d)$$

En adoptant (4.89) pour $c_Z(\tau)$, nous considérerons Z(t) comme étant un processus autorégressif d'ordre un (AR(1)) ayant les propriétés suivantes: le paramètre ϕ_1 , dans un tel modèle, prend la valeur de ρ_1 [Box et Jenkins (1976)], qui vaut ici $\rho_1 = 10^{-2} \exp[-\zeta]$. Le processus Z(t) sera donc un processus AR(1)

$$Z(t) = \phi_1 Z(t-1) + \varepsilon_t \tag{4.91}$$

où ε_{ι} est un bruit blanc avec $\sigma_{\varepsilon}^2 = (1 - \phi_1^2)\sigma_z^2$.

4.3.5 Intensité de loi gamma pour s'ajuster aux valeurs extrêmes

Pour le problème relié aux valeurs extrêmes, mentionné dans la section 4.3.3, *Onof et Wheater* (1994a) suggèrent de représenter l'intensité des précipitations, X, par une loi $\Gamma(p,\delta)$, de densité $f(x) = \frac{1}{\Gamma(p)\delta^p} x^{p-1} e^{-x/\delta}$ et où p est le paramètre de forme et δ est le paramètre d'échelle, plutôt que par une loi exponentielle ou géométrique. Pour le processus agrégé, *Onof et Wheater* (1994a) obtiennent :

$$E(Y_i^h) = \lambda h \mu_X (1 + \kappa/\phi) \frac{v}{\alpha - 1}$$
 (4.92a)

$$Var(Y_i^h) = 2A_1[(\alpha - 3)hv^{2-\alpha} - v^{3-\alpha} + (\nu + h)^{3-\alpha}]$$

$$-2A_2[\phi(\alpha - 3)hv^{2-\alpha} - v^{3-\alpha} + (\nu + \phi h)^{3-\alpha}]$$
(4.92b)

$$Cov(Y_{i}^{h}, Y_{i+k}^{h}) = A_{1} \{ [\nu + (k+1)h]^{3-\alpha} - 2(\nu + kh)^{3-\alpha} + [\nu + (k-1)h]^{3-\alpha} \}$$

$$- A_{2} \{ [\nu + (k+1)\phi h]^{3-\alpha} - 2(\nu + k\phi h)^{3-\alpha} + [\nu + (k-1)\phi h]^{3-\alpha} \}$$

$$(4.92c)$$

où

$$A_{1} = \frac{\lambda (1 + \kappa/\phi) v^{\alpha}}{\delta^{2} (\alpha - 1)(\alpha - 2)(\alpha - 3)} \left[p(p+1) + \frac{\kappa \phi p^{2}}{\phi^{2} - 1} \right]$$
$$A_{2} = \frac{\lambda (1 + \kappa/\phi) \kappa \mu_{X} v^{\alpha}}{\phi^{2} (\phi^{2} - 1)(\alpha - 1)(\alpha - 2)(\alpha - 3)}$$

et
$$\mu_X = \frac{p}{\delta}$$
.

Ce nouvel ajout remplit bien son rôle et améliore la reproduction des valeurs extrêmes. Mais l'ajout d'un paramètre augmente considérablement la difficulté pour estimer les paramètres. Il est suggéré de fixer δ et d'optimiser les autres variables.

5. APPLICATION DE CERTAINS MODÈLES AUX DONNÉES DE L'ALABAMA

5.1 Analyse qualitative des données

Les modèles décrits précédemment sont des modèles stationnaires. Les données de précipitations comportent d'importantes fluctuations dans les moyennes journalières qui font qu'un modèle stationnaire ne peut simuler de scénarios couvrant toute une année. Souvent, dans la littérature, on divise les données en saisons qui peuvent être de longueur variable. Il arrive même qu'on prenne chaque mois séparément [Cowpertwait (1991); Onof et Wheater, (1993) et (1994b)]. Le modèle global sera dans ce cas un ensemble de modèles qui représentent chaque saison.

Comme le but de l'étude n'est pas de trouver un modèle représentatif de toute l'année mais bien de décrire la performance de quelques modèles par rapport au séries d'événements extrêmes, une saison d'une durée d'environ trois mois sera déterminée et c'est sur celle-ci que l'analyse portera. Ceci aura aussi l'avantage de réduire le temps de simulation pour chaque modèle.

L'ensemble de données provient de l'Alabama et elles couvrent une période de 49 ans pour quatre stations. Il s'agit de données horaires. Il est à noter que le climat en Alabama est tel que le mois le plus froid est le mois de janvier. À cette période de l'année, la température descend rarement sous -5° C et les précipitations de neige sont extrêmement rares. Lorsque cela se produit, il n'y a jamais d'accumulation, c'est-à-dire qu'il s'agit toujours de neige fondante. Il importe donc peu de choisir une saison en particulier puisque le phénomène de fonte de neige n'intervient pas. C'est d'ailleurs la raison pour laquelle les données ont été

prises d'une telle région plutôt que d'avoir à se préoccuper des précipitations de neiges, comme il y en a souvent au Québec.

En portant sur un graphique les moyennes journalières et en les lissant à l'aide d'une régression non paramétrique standard, il est possible avec un simple examen visuel de déterminer un tel intervalle. L'intervalle des jours 10 à 100 a été choisi et indiqué sur les graphiques de la Figure 5.1.

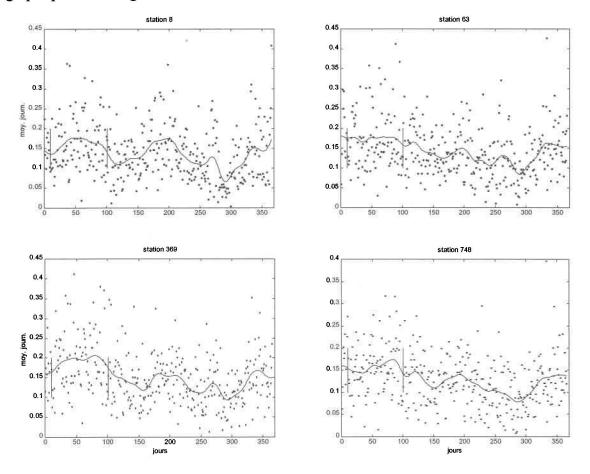


Figure 5.1: Moyennes journalières lissées. Les intervalles stationnaires sont indiqués dans le graphique de chaque station.

Étant donné la courte durée de l'intervalle de temps choisi, il semble que la validation selon la méthode des séries de PMA (précipitation maximum annuelle) ne soit pas très indiquée.

Les courbes IDF (intensité-durée-fréquence) seront donc tracées seulement pour les séries de durées partielles et la performance sera évaluée sur cette base.

5.2 Modèle de Poisson à impulsions rectangulaires (PIR)

Le modèle de PIR a été utilisé avec les données de l'Alabama que l'on vient de décrire. Les résultats sont présentés ici pour venir compléter les analyses déjà disponibles dans la littérature en mettant l'accent sur l'adéquation aux séries d'événements extrêmes.

5.2.1 Les paramètres

La variabilité des paramètres d'une échelle à une autre est évidente lorsqu'on observe le Tableau 5.1. On y voit les solutions aux systèmes de trois équations (4.23, 4.26 et 4.27) à trois inconnues (λ, ξ, η) pour chacune des échelles indiquées. La valeur historique de chacun des moments, obtenue à l'aide de l'agrégation des données au niveau désiré, est égalée à l'expression analytique correspondante et une solution exacte peut être ainsi trouvée. Cependant, à cause de la façon dont les données sont agrégées, la relation $\mu_h = h \cdot \mu_1$, où μ_h est la moyenne à l'échelle h (en heures), est toujours satisfaite et c'est aussi ce qu'indique l'expression (4.23), qui est une fonction linéaire de l'échelle. C'est d'ailleurs ce qu'on peut constater dans le Tableau 5.2 Il est à noter que cette relation n'est pas tout à fait vérifiée pour les échelles dépassant un jour. Cela est dû à la perte d'informations qu'il y a lors de l'agrégation des données dans les cas où il y a des journées manquantes.

	λ (hre ⁻¹)	ξ (hre)	η (po/hre)	
1 hre	0.0315	3.6044	1.3823	
6 hres	0.0189	6.6477	0.4491	
12 hres	0.0179	7.7902	0.3622	
24 hres	0.0181	8.9264	0.3211	

Tableau 5.1: Valeurs des paramètres du modèle de Poisson à impulsions rectangulaires à différentes échelles de temps pour la station 8

À cause de la structure de dépendance du processus ponctuel de Poisson à impulsions rectangulaires, plusieurs auteurs [Waymire and Gupta (1981b); Rodriguez-Iturbe et al. (1984)] s'entendent sur le fait que ce modèle n'est pas approprié pour modéliser les séries temporelles de précipitations. Aussi, puisque les paramètres semblent avoir des tendances monotones, il est clair qu'ils atteindront des valeurs qui seront physiquement irréalistes.

Échelle →	1 hre	6 hres	12 hres	1 jour	2 jours	3 jours	4 jours	5 jours
station 8	0.0063	0.0380	0.0760	0.1520	0.3353	0.5092	0.6759	0.8410
station 63	0.0069	0.0411	0.0823	0.1645	0.3544	0.5331	0.7059	0.8879
station 369	0.0079	0.0476	0.0951	0.1903	0.3983	0.5971	0.7991	0.9817
station 748	0.0060	0.0361	0.0721	0.1443	0.3220	0.4832	0.6451	0.8030

Tableau 5.2 : Hauteurs moyennes de précipitation à différentes échelles de temps pour les quatre stations

Comme le principal outil de validation utilisé dans cette étude est l'ensemble des courbes IDF pour un site donné, il est important que le modèle utilisé pour les simulations soit le même à toutes les échelles de temps. Aussi, une des propriétés que nous voulons vérifier est justement la capacité qu'a un modèle de bien simuler à différentes niveaux d'agrégation. Cependant, les différentes valeurs de paramètres obtenues pour chaque échelle viennent un peu compliquer le choix du modèle. En effet, une solution existe pour le système d'équation que constituent les expressions (4.23), (4.26) et (4.27) mais la solution est différente lorsque l'échelle change. Plutôt que de choisir subjectivement l'ensemble de paramètres à utiliser, la méthode d'estimation choisie est la même que pour le modèle de Neyman-Scott. Elle consiste à minimiser la somme des carrés des différences entre les moments historiques et les moments estimés à l'aide des expressions analytiques pour plusieurs échelles. Les paramètres étaient donc estimés en minimisant la fonction suivante :

$$f_{obj} = 10 \cdot \left[moy_{obs}^{(1)} - moy_{th}^{(1)} \right]^{2} + \left[var_{obs}^{(12)} - var_{th}^{(12)} \right]^{2} + \left[var_{obs}^{(24)} - var_{th}^{(24)} \right]^{2} \right)$$

$$10 \cdot \left(\left[acv(1)_{obs}^{(1)} - acv(1)_{th}^{(1)} \right]^{2} + \left[acv(1)_{obs}^{(12)} - acv(1)_{th}^{(12)} \right]^{2} + \left[acv(1)_{obs}^{(24)} - acv(1)_{th}^{(24)} \right]^{2} \right)$$

$$(5.1)$$

où $\left\{ \right\}_{obs}^{(h)}$ indique le moment observé à l'échelle h et $\left\{ \right\}_{th}^{(h)}$ indique le moment théorique, calculé avec les paramètres. Les moyennes sont multipliées par 10 parce que le modèle doit absolument préserver cette caractéristique. Les termes d'autocovariances sont aussi multipliées par 10 pour leur donner la même "importance", en valeur, que les termes de variances. Ceci a été déterminé en observant l'ordre de grandeur de chacun de ces moments. En minimisant cette fonction objectif, on a obtenu les valeurs présentées dans le Tableau 5.3.

	λ	ξ	η
Station 8	0.0200	3.0891	1.0263
Station 63	0.0185	12.9608	0.2083
Station 369	0.0195	9.9699	0.2473
Station 748	0.0158	11.5380	0.2276

Tableau 5.3 : Paramètres pour le modèle de PIR aux quatre stations.

5.2.2 Analyse du modèle

Les données simulées ont étés obtenues à l'aide du logiciel MATLAB dans lequel le modèle a été programmé. 500 ans de données horaires ont étés simulées avec un seul modèle pour chaque station et ces données étaient agrégées au besoin pour l'analyse des résultats aux échelles plus élevées.

Avant de montrer les résultats obtenus avec le modèle de PIR, nous allons montrer, en exemple, la procédure qui a été suivie pour la construction des courbes IDF. Pour tous les niveaux d'agrégation, il a été convenu que les séries de durées partielles comporteraient environ une trentaine d'événements. C'est de cette façon que les seuils ont été trouvés. Une série de durée partielle était extraite avec un certain seuil, qu'on ajustait ensuite de façon à obtenir une série de taille n=30. La taille de la série a été déterminée de façon à conserver tous les événements extrêmes sans toutefois garder les événements qui sont élevés, mais courants. La loi Pareto généralisée (PG) et la loi généralisée des valeurs extrêmes (GEV) ont été ajustées aux données et un test de Kolmogorov-Smirnov a été appliqué pour déterminer laquelle des deux distributions serait choisie. Dans tous les cas, au moins une des deux distributions ne pouvait être rejetée avec un niveau de signification de 5%. Dans certains cas, la statistique du test de Kolmogorov-Smirnov était à peu près la même pour les deux distributions. Dans ces rares circonstances, le deuxième test, présenté à la section 3.2.2, était utilisé pour trancher.

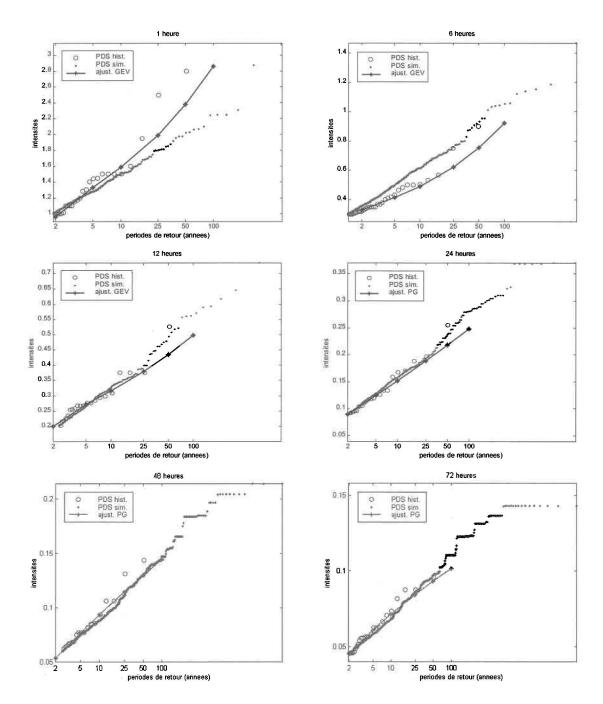


Figure 5.2 : Ajustements des lois de Pareto généralisée (PG) et GEV aux séries de durées partielles pour les différentes échelles de temps.

Les graphiques de la Figure 5.2 montrent l'ajustement de la meilleure des deux lois aux séries de PDS à chaque échelle de temps. Le fait que l'ajustement ne croise pas l'événement

le plus rare n'est pas vraiment un problème. En effet, il est légitime de croire que la pluie la plus intense dans les 49 ans de données à notre disposition n'ait pas nécessairement une période de retour de 50 ans. Cet événement aurait très bien pu être le plus grand, même si 100 ans de données avaient été disponibles. C'est aussi la raison pour laquelle les courbes IDF historiques seront déduites des événements qui correspondent à la courbe ajustée, et non à la série de durée partielle originale.

Les graphiques de la Figure 5.3 montrent les résultats ainsi obtenus pour chaque station. La forme "brisée" des courbes est due au fait que la même loi n'était pas toujours utilisée pour ajuster les séries des différents niveaux d'agrégation.

Malgré une assez bonne performance aux échelles de 12 heures et plus dans le cas de la station 8, la performance générale aux autres stations vient appuyer ce que la plupart des auteurs prétendent, c'est-à-dire que le modèle n'est pas du tout adéquat. En particulier pour la station 748, aucun événement n'a dépassé le seuil fixé pour l'échelle horaire. Il aurait fallu abaisser considérablement le seuil pour réussir à extraire quelques événements de la série de données simulées. Nous pouvons cependant noter que le modèle a une tendance à surestimer les événements pour les échelles plus grandes que l'échelle horaire. Cette dernière est largement sous-estimée de façon systématique. Cette surestimation pour les grandes échelles est plus ou moins prononcée selon la station. Pour améliorer ces résultats, il faudrait que le modèle accentue la pente générale des courbes simulées.

Regardons maintenant les propriétés statistiques des simulations. Dans le Tableau 5.4, on retrouve les différentes valeurs des moments et des proportions d'intervalles secs pour les données historiques et simulées.

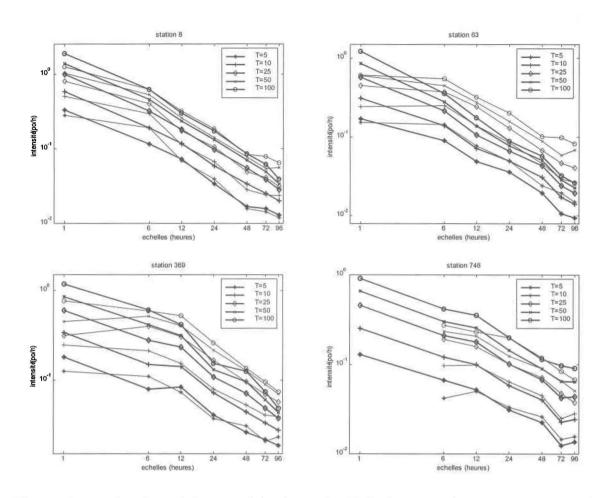


Figure 5.3 : Courbes IDF pour les séries de PDS des quatre stations. Les courbes historiques sont en gras et les simulées correspondent aux traits minces.

La moyenne et la variance aux échelles plus petites ou égales à un jour sont bien reproduites. Mais pour ce qui est de l'autocorrélation, on voit qu'elle est presque toujours sous-estimée et ceci, même pour le délai un, qui était utilisé dans l'estimation des paramètres. Même chose pour les proportions d'intervalles secs. Notons que ces statistiques, sont celles de la station 8, qui a donné les meilleurs résultats pour les courbes IDF. Il est donc clair que la performance du modèle n'est pas satisfaisante pour les cas étudiés ici. Des résultats comparables ont été obtenus pour les trois autres stations. À cette différence près que les trois autres modèles donnaient de meilleurs résultats dans le cas de l'autocorrélation de délai un.

	1 hre	6 hres	12 hres	24 hres	48 hres	72 hres	96 hres
moyenne*	0,0063	0,0380	0,0760	0,1520	0,3353	0,5092	0,6759
	0,0062	0,0375	0,0744	0,1488	0,2979	0,4464	0,5959
variance	0,0023	0,0333	0,0833	0,1846	0,4289	0,6459	0,9078
	0,0029	0,0380	0,0841	0,1756	0,3554	0,5458	0,7195
acr (1)*	0,4429	0,2467	0,1450	0,0745	-0,0147	-0,0678	-0,0957
	0,5298	0,0985	0,0352	0,0077	-0,0229	-0,0423	-0,0482
acr (2)	0,2460	0,0556	0,0092	-0,0198	-0,0198	-0,0393	-0,1043
	0,1860	-0,0048	-0,0075	-0,0242	-0,0152	-0,0409	-0,0625
acr (3)	0,1864	0,0125	-0,0236	-0,0176	-0.0176	0,0231	0,0758
	0,0646	-0,0061	-0,0042	-0,0138	-0.0235	-0,0377	-0,0519
P(int. Sec)	0,9484	0,8870	0,8325	0,7365	0,5330	0,3989	0,2980
	0,9617	0,8712	0,7743	0,6110	0,3795	0,2361	0,1490

Tableau 5.4: Propriétés statistiques des données historiques (en gras) et simulées à la station 8. (*: moments utilisé dans l'estimation des paramètres)

5.3 Neyman-Scott à impulsions rectangulaires (NSIR)

5.3.1 Estimation des paramètres

La méthode d'estimation des paramètres est celle qui a été utilisée dans le cas du modèle de PIR. On minimise la fonction objectif de l'expression (5.1) qui est une somme des carrés des différences entre les moments historiques et les moments théoriques du modèle, exprimés en fonction des paramètres du modèle. L'avantage de cette méthode est que plusieurs échelles sont utilisées. Ceci a pour conséquence que le modèle est unique, c'est-à-dire qu'un seul modèle suffit à chaque station pour simuler à toutes les échelles.

Théoriquement, la quantité d'éléments qu'on introduit dans la fonction objectif n'est pas limitée, mais en pratique les méthodes itératives utilisées pour optimiser cette fonction le sont. La fonction objectif a donc la forme de (5.1) et les termes d'espérance, de variance et d'autocovariance sont donnés par les expressions (4.54a), (4.54b) et (4.54c). Dans la

littérature [Entekhabi et al. (1989); Rodriguez-Iturbe et al. (1987b)] on utilise la plupart du temps l'espérance, la variance et l'autocovariance de délai un à l'échelle horaire, en plus de la variance à une autre échelle (6 heures, 12 heures ou 24 heures) et l'autocovariance de délai un à la même échelle que celle choisie pour la variance.

Puisque dans le cas étudié ici on veut voir comment le modèle se comporte sur une plus large étendue d'échelles, on a voulu introduire des échelles supérieures dans la fonction objectif. Mais à mesure que l'on augmente l'échelle, les valeurs des autocovariances deviennent de plus en plus faibles et, même si la valeur minimale obtenue de la fonction objectif est très près de 0, les moments ne sont pas nécessairement bien préservés. Pour cette raison, les moments pour les échelles plus grandes que journalières seront laissés de côté dans l'estimation des paramètres.

La résolution d'un système de cinq équations à cinq inconnues serait visiblement fastidieuse. Les algorithmes d'optimisation et de résolution de systèmes d'équations qui ont été utilisés n'ont pas réussi à obtenir de solution au bout de 5000 itérations et ce, pour plusieurs combinaisons possibles d'expressions. C'est pourquoi la méthode la plus directe consiste à utiliser la fonction objectif de l'expression (5.1). On réussit à s'approcher de très près d'une solution à l'aide de la fonction objectif mais il faut vérifier la validité et la crédibilité des valeurs obtenues.

Vérifier la crédibilité de la valeur d'un paramètre signifie que l'on vérifie si la valeur est physiquement réaliste. La fonction objectif peut être très près de 0 mais avec des valeurs de paramètres tout à fait invraisemblables pour un modèle de pluie. Si on obtient par exemple une valeur de λ très petite (ex: λ = 0,0005), cela voudrait dire que les fronts de basse pression arrivent à un taux de 1 par 2000 heures. L'intervalle de temps considéré ici est d'environ trois mois, ce qui constitue près de 2000 heures.

Il apparaît donc que la fonction objectif comporte plusieurs minima locaux, que l'on peut découvrir seulement en changeant les valeurs initiales lors de l'optimisation. Le Tableau 5.5 donne les résultats obtenus pour différentes valeurs initiales, aux échelles utilisées, ainsi que la valeur finale de la fonction objectif.

	Valeu	ırs in	itiale	S	λ	β	η	μ_X	E(c)	échelles	fonction
λ	β	η	μ_X	E(c)		# 6 6 7 7 8 8 8 8 8 8				utilisées	objectif
0.01	1	8.0	0.4	3	0.0051	0.0089	0.5015	0.1554	4.0243	6 et 12 hres	8 × 10 ⁻⁷
0.01	1	8.0	0.4	3	0.0071	0.2175	1.6835	0.2257	6.6184	12 et 24 hres	7×10 ⁻⁶
.005	0.4	1	0.5	6	0.0072	0.2214	1.3850	0.2014	6.0340	6 et 12 hres	7×10^{-7}
0.01	0.03	8.0	1	3	0.0091	0.1966	0.7953	0.1493	3.7147	6 et 24 hres	2×10 ⁻⁶
0.01	0.02	0.3	0.4	3	0.0104	0.1862	0.5730	0.1339	2.5436	6 et 24 hres	9×10 ⁻⁶

Tableau 5.5: Valeurs des paramètres obtenus avec la méthode de la fonction objectif selon les valeurs initiales, les échelles utilisées et les valeurs finales pour la fonction objectif à la station 8.

Afin de s'assurer de partir avec des valeurs réalistes physiquement, on a utilisé un arrondissement des valeurs obtenues par *Entekhabi et al.* (1989). Ensuite, un ajustement était fait selon la tendance des valeurs obtenues et on faisait ainsi varier les valeurs initiales.

Pour valider les valeurs de paramètres obtenues, les moments étaient recalculés à l'aide des expressions analytiques pour ensuite être comparés aux valeurs des moments historiques. Aucun test n'était appliqué pour décider si les valeurs seraient retenues car les paramètres obtenus donnaient des résultats, soit très éloignés des valeurs historiques, ou soit près de ces valeurs. Seules les valeurs des paramètres donnant des moments près des moments historiques étaient conservées. On remarquait aussi des ensembles de valeurs qui se ressemblaient. C'est ce qu'on remarque en regardant les valeurs de la 2° et la 3° ligne du Tableau 5.5. Les valeurs en gras dans le Tableau 5.5 indiquent les valeurs qui ont été utilisées pour des simulations. Nous noterons ces ensembles E1 (2° ligne du tableau) et E2

(4e ligne du tableau). On peut remarquer que les ensembles de valeurs de paramètre des première et quatrième lignes n'ont pas été retenus, malgré que la valeur de la fonction objectif soit plus faible. C'est que les échelles utilisées étaient 6 et 12 heures. Or, ce que nous souhaitons c'est d'obtenir un modèle qui sera aussi efficace aux petites qu'aux grandes échelles. C'est pourquoi nous voulons introduire des expressions à des échelles les plus grandes possibles.

Notons que les mêmes difficultés ne sont pas toutes apparues aux mêmes endroits pour les autres stations. Des paramètres pour les stations 63 et 369 ont facilement été obtenus et ce, avec plusieurs combinaisons d'échelles. Par contre, dans le cas de la station 748, très peu d'ensembles de valeurs pour les paramètres ont été trouvés. À chaque station, 2 ou 3 ensembles de paramètres étaient retenus et des simulations étaient faites avec chacun de ces ensembles.

5.3.2 Application du modèle

La Figure 5.4 montre les résultats obtenus avec les ensembles E1 et E2 pour la station 8. Une propriété surprenante du modèle de NSIR est que malgré la présence des expressions à l'échelle horaire dans l'estimation des paramètres, le modèle semble éprouver des difficultés à préserver les propriétés des courbes IDF à cette échelle. Mais il est clair, selon ces graphiques, que le second ensemble de paramètres offre de biens meilleurs résultats.

Dans le Tableau 5.6, on peut voir comment le modèle a réussi à reproduire les moments, selon l'ensemble de paramètres aux différentes échelles de temps. Les moments simulés se ressemblent entre eux et sont raisonnablement proches des valeurs historiques dans le cas des moments utilisés pour l'optimisation des paramètres. Comme nous pouvions nous y attendre toutefois, l'écart croît avec l'échelle de temps dans le cas de la moyenne et de la variance. Cependant, les proportions d'intervalles secs décroissent plus rapidement avec le deuxième ensemble de paramètres qu'avec le premier et sont donc plus proches des valeurs historiques pour E2 que pour E1. En se basant seulement sur ces informations, il serait

encore difficile de discerner quel modèle est le plus adéquat. Mais en les combinant à l'information supplémentaire que fournissent les courbes IDF, il devient clair que E2 est le modèle le plus adéquat.

Pour ce qui est des trois autres stations, les meilleurs résultats obtenus sont montrés sur les graphiques de la Figure 5.5. La justification du choix parmi toutes les simulations à une même station pouvait se faire visuellement. Les courbes que les graphiques illustraient dans chaque cas étaient très différentes et le choix était sans équivoque. La performance par rapport aux moments est comparable à celle obtenue dans le cas de la station 8 qui a été exposé plus en détail. Les valeurs des moments simulés pour les quatre stations sont donnés dans l'annexe A.

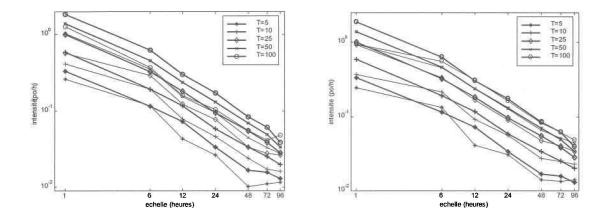


Figure 5.4: Courbes IDF avec la méthode des PDS pour la station 8 avec E1 et E2. Les courbes historiques sont en gras et les simulées correspondent aux traits minces.

	1 hre	6 hres	12 hres	24 hres	48 hres	72 hres	96 hres
moyenne*	0,0063	0,0380	0,0760	0,1520	0,3353	0,5092	0,6759
E1	0,0063	0,0378	0,0756	0,1512	0,3031	0,4537	0,6062
E2	0,0061	0,0369	0,0738	0,1476	0,2956	0,4427	0,5912
variance*	0,0023	0,0333	0,0833	0,1846	0,4289	0,6459	0,9078
E1	0,0023	0,0326	0,0806	0,1851	0,3994	0,6042	0,8306
E2	0,0017	0,0321	0,0803	0,1847	0,4049	0,6094	0,8425
acr (1)*	0,4429	0,2467	0,1450	0,0745	-0,0147	-0,0678	-0,0957
E1	0,5051	0,2388	0,1407	0,0575	0,0099	-0,0162	-0,0412
E2	0,6765	0,2479	0,1432	0,0672	0,0064	-0,0111	-0,0493
acr (2)	0,2460	0,0556	0,0092	-0,0198	-0,0198	-0,0393	-0,1043
E1	0,2216	0,0507	-0,0007	-0,0136	-0,0370	-0,0378	-0,0570
E2	0,3837	0,0468	0,0080	-0,0156	-0,0314	-0,0459	-0,0619
acr (3)	0,1864	0,0125	-0,0236	-0,0176	-0.0176	0,0231	0,0758
E1	0,1458	0,0043	-0,0089	-0,0142	-0,0269	-0,0460	-0,0438
E2	0,2372	0,0129	-0,0101	-0,0116	-0,0321	-0,0509	-0,0551
P(int. Sec)	0,9484	0,8870	0,8325	0,7365	0,5330	0,3989	0,2980
E1	0,9523	0,8939	0,8509	0,7802	0,6593	0,5573	0,4697
E2	0,9459	0,8836	0,8317	0,7481	0,6057	0,4916	0,3978

Tableau 5.6 : Caractéristiques statistiques des données historiques (en gras) et simulées à la station 8 avec E1 et E2.

5.4 Bartlett-Lewis à impulsions rectangulaires avec perturbateur (BLIRP)

5.4.1 Estimation des paramètres

Onof et Wheater (1993) ont développé une méthode pour estimer les paramètres qui dépendaient du temps moyen entre deux événements pluvieux (de la fin d'une précipitation au début de la suivante) et du nombre moyen d'averses par mois. Ils ont développé cette

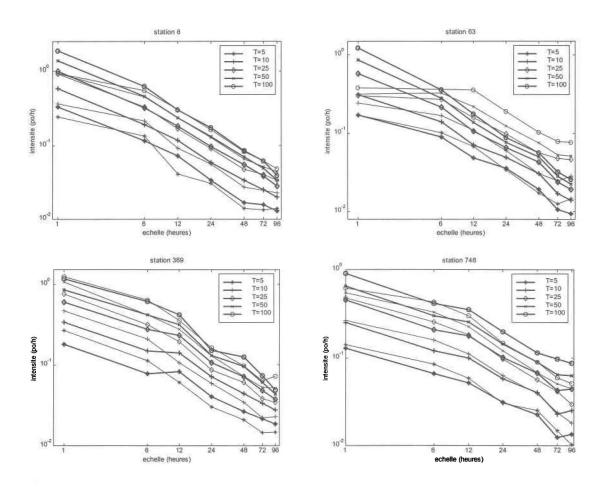


Figure 5.5 : Résultats retenus pour les quatre stations parmi les ensembles de paramètres trouvés avec le modèle de NSIR. Les courbes historiques sont en gras et les simulées correspondent aux minces.

méthode après avoir rencontré des difficultés lors de l'optimisation des paramètres du modèle de BLIRP appliqué aux données britanniques. Le climat du sud des États-Unis étant différent du climat britannique, de tels problèmes n'ont pas été rencontrés dans le cas présent. La même procédure que dans les modèles précédents a donc été conservée. Une fonction objectif, consistant en une somme de carrés de différences entre les moments historiques et les moments calculés, était minimisée. La moyenne à l'échelle horaire, la variance ainsi que l'autocovariance de délai un à trois échelles de temps ont été considérées dans cette fonction objectif.

		1 hre	6 hres	12 hres	24 hres	48 hres	72 hres	96 hres
	hist.	0.0023	0.0333	0.0833	0.1846	0.4289	0.6459	0.9078
var.	BL	0.0023	0.0317	0.0762	0.1811	0.3882	0.5880	0.8296
	perturb.	0.0024	0.0325	0.0778	0.1845	0.3945	0.5969	0.8460
	hist.	0.4429	0.2467	0.1450	0.0745	-0.0147	-0.0678	-0.0957
acr(1)	BL	0.4853	0.2223	0.1663	0.0670	0.0126	-0.0025	-0.0532
	perturb.	0.4733	0.2179	0.1644	0.0656	0.0129	-0.0017	-0.0554

Tableau 5.7: Variances et autocorrélations de délai un, historiques et simulées, pour le processus non-perturbé (BLIR) et perturbé (BLIRP) (station 8).

Cependant, si on observe les expressions (4.90b) et (4.90c), on peut voir que si ζ est suffisamment grand pour satisfaire (4.87), les termes $2h\left\{\sigma_{\gamma}^2 + \mu_{\gamma}^2\right\}\frac{10^{-2}}{\zeta}$ et $h\left\{\sigma_{\gamma}^2 + \mu_{\gamma}^2\right\}\frac{\exp(-\zeta h)10^{-2}}{\zeta}$, qui correspondent respectivement à l'écart entre la variance du processus perturbé et la variance du processus non perturbé et entre l'autocovariance des mêmes processus, peuvent être très petits. Comme le paramètre ζ prenait généralement des valeurs très près de 0,05 (entre 0,043 et 0,065 avec une asymétrie vers les petites valeurs de l'intervalle), il a été convenu de fixer ζ à une valeur moyenne, soit à 0,05 et d'effectuer ensuite l'optimisation pour obtenir les autres paramètres, λ , ϕ , κ , η et μ_{χ} . Le Tableau 5.7 illustre bien la faible différence entre les deux processus, perturbé et non perturbé.

5.4.2 Application du modèle

Tel que mentionné plus haut, l'estimation des paramètres n'a donné lieu à aucun problème majeur et par conséquent, il n'y a pas eu de choix d'ensemble de paramètres à faire. Les valeurs obtenues variaient un peu mais restaient toujours dans un petit intervalle. Les résultats des simulations effectuées pour chaque station sont présentés sur les graphiques de la Figure 5.6..

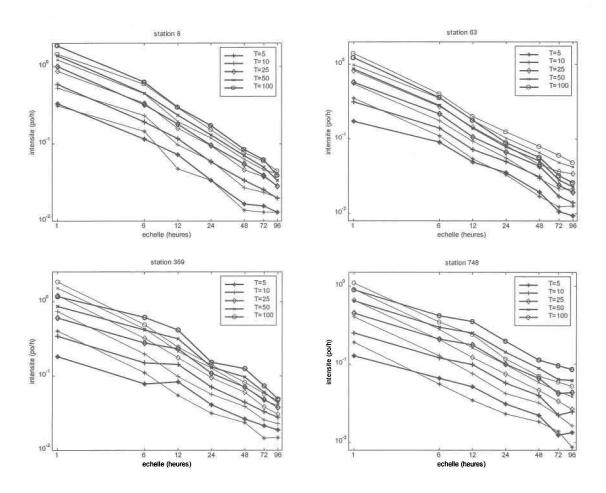


Figure 5.6: Courbes IDF des quatre stations pour le modèle de BLIRP. Les courbes historiques sont en gras et les simulées correspondent aux minces.

On remarque que le modèle de BLIRP a réussi à "remonter" les courbes simulées à l'échelle horaire. Notamment dans le cas de la station 8, où les événements extrêmes à la plus petite échelle sont assez bien reproduits. On observe également que les courbes simulées ont des allures plus linéaires que dans les modèles précédents. C'est d'ailleurs l'allure qu'auraient les courbes historiques si la même distribution avait été utilisée pour ajuster chaque série de durée partielle (Figure 5.2).

Le Tableau 5.8 montre comment le modèle de BLIRP a reproduit les moments et les proportions d'intervalles secs. La performance à ce niveau est sensiblement la même que

celle obtenue avec le modèle de NSIR (voir tableau récapitulatif en annexe A) et l'information que fournissent les courbes IDF pourrait servir de base pour décider du choix du modèle à utiliser.

	1 hre	6 hres	12 hres	24 hres	48 hres	72 hres	96 hres
moyenne*	0,0063	0,0380	0,0760	0,1520	0,3353	0,5092	0,6759
	0.0064	0.0381	0.0762	0.1525	0.3046	0.4574	0.6092
variance*	0,0023	0,0333	0,0833	0,1846	0,4289	0,6459	0,9078
	0.0027	0.0363	0.0873	0.2013	0.4339	0.6737	0.9009
acr (1)*	0,4429	0,2467	0,1450	0,0745	-0,0147	-0,0678	-0,0957
	0.4777	0.2108	0.1436	0.0691	0.0131	-0.0195	-0.0333
acr (2)	0,2460	0,0556	0,0092	-0,0198	-0,0198	-0,0393	-0,1043
	0.1920	0.0525	0.0102	-0.0134	-0.0270	-0.0319	-0.0478
acr (3)	0,1864	0,0125	-0,0236	-0,0176	-0.0176	0,0231	0,0758
	0.1222	0.0192	-0.0068	-0.0177	-0.0265	-0.0434	-0.0642
P(int. Sec)	0,9484	0,8870	0,8325	0,7365	0,5330	0,3989	0,2980
	0.9567	0.8964	0.8410	0.7404	0.5767	0.4488	0.3538

Tableau 5.8: Moments et proportions d'intervalles secs pour le modèle de BLIRP à la station 8 (les moments historiques sont en gras).

5.5 Bartlett-Lewis à impulsions rectangulaires et intensités de loi Gamma (BLIRIG)

5.5.1 Estimation des paramètres

Pour ce modèle à sept paramètres (λ , κ , ϕ , α , ν , p et δ), un de ceux-ci était fixé avant l'estimation comme cela a été suggéré par *Onof et Wheater* (1994a). En prenant les paramètres du modèle précédent (BLIRP) et en égalant δ à la valeur de μ_X^{-1} , on obtenait aisément des valeurs pour chacun des autres paramètres. La raison qui a motivé à fixer la valeur de δ vient de la résolution des relations suivantes:

$$\mu_X = \frac{p}{\delta} \tag{5.2}$$

$$\sigma_X^2 = \frac{p}{\delta^2} \tag{5.3}$$

où la relation (5.3) est basée sur l'hypothèse que la variance de l'intensité est la même que dans le modèle de BLIRP. On obtient ainsi comme première valeur de δ , $\delta = \mu_X^{-1}$. Plusieurs ensembles de paramètres ont pu être trouvés pour la même station en ne changeant que les valeurs de départ. La valeur de λ ne changeant pas beaucoup, la même opération a été répétée mais cette fois-ci, avec λ fixé à la valeur obtenue et en laissant δ libre. Ainsi, après quelques essais, on a pu fixer les valeur de p et δ pour obtenir un ensemble de paramètres optimaux. La même procédure a été employée pour les quatre stations et les paramètres se sont fixés dans le même ordre à chaque fois.

5.5.2 Application du modèle

La simulation de 500 années a été faite à chaque station avec le modèle de BLIRIG et les courbes IDF résultantes sont présentées à la Figure 5.7. Les résultats obtenus ressemblent beaucoup à ceux obtenus pour le modèle précédent. Une légère amélioration à l'échelle horaire est perceptible en général, tandis que les autres échelles sont à peu près semblables. C'est le même constat qu'on peut faire en observant les valeurs des moments et des proportions d'intervalles secs obtenus, qui sont présentés dans le Tableau 5.9 pour la station 8. Comme les tableaux de l'annexe A le montrent, les faibles différences entre les deux modèles de BL ne sont pas assez grandes pour déterminer lequel a le mieux reproduit les caractéristiques statistiques étudiées ici.

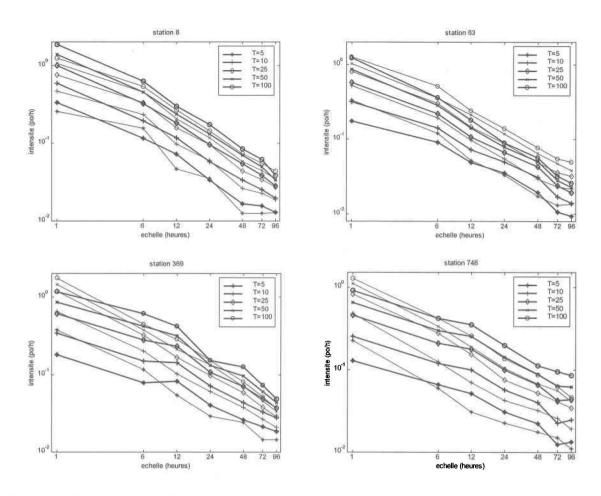


Figure 5.7 : Courbes IDF des quatre stations pour le modèle de BLIRIG. Les courbes historiques sont en gras et les simulées correspondent aux minces.

	1 hre	6 hres	12 hres	24 hres	48 hres	72 hres	96 hres
moyenne*	0,0063	0,0380	0,0760	0,1520	0,3353	0,5092	0,6759
	0.0062	0.0373	0.0746	0.1491	0.2983	0.4474	0.5967
variance	0,0023	0,0333	0,0833	0,1846	0,4289	0,6459	0,9078
	0.0025	0.0340	0.0827	0.1908	0.4160	0.6413	0.8612
acr (1)*	0,4429	0,2467	0,1450	0,0745	-0,0147	-0,0678	-0,0957
	0.4840	0.2200	0.1454	0.0716	0.0083	-0.0137	-0.0380
acr (2)	0,2460	0,0556	0,0092	-0,0198	-0,0198	-0,0393	-0,1043
	0.1940	0.0556	0.0119	-0.0178	-0.0271	-0.0266	-0.0357
acr (3)	0,1864	0,0125	-0,0236	-0,0176	-0.0176	0,0231	0,0758
V. 1000	0.1210	0.0188	-0.0074	-0.0131	-0.0175	-0.0360	-0.0588
P(int. Sec)	0,9484	0,8870	0,8325	0,7365	0,5330	0,3989	0,2980
M	0.9568	0.8964	0.8416	0.7415	0.5779	0.4495	0.3543

Tableau 5.9 : Moments et proportions d'intervalles secs pour le modèle de BLIRIG à la station 8 (les moments historiques sont en gras)

5.6 Comparaison des modèles

On a bien vu précédemment comment chaque modèle se comporte pour les quatre stations. Il est intéressant de montrer maintenant comment le choix du modèle le plus adéquat se fait pour chaque station.

5.6.1 Station 8

On a pu constater au cours de ce travail que la station 8 était généralement bien modélisée par les modèles ponctuels présentés ici. La moyenne et la variance sont bien reproduites par les quatre modèles (annexe A - Tableau A.1). Cependant, la fonction d'autocorrélation de délai un pour le modèle de PIR décroît trop rapidement et les proportions d'intervalles secs du même modèle sont trop faibles aux échelles élevées pour que ce modèle soit considéré adéquat.

Les modèles de NSIR et de BLIRP corrigent cette lacune mais le modèle de BLIRP donne toujours des résultats plus près des valeurs historiques que le modèle de NSIR. Pour décider du modèle final à utiliser, la Figure 5.8 donne pour les quatre modèles, l'information que fournissent les courbes IDF de la station 8. Encore une fois, les modèles de NSIR et de BLIRP donnent des résultats qui sont très semblables. La seule différence notable est à l'échelle horaire où le modèle de BLIRP redresse la courbe. C'est donc ce dernier modèle qui devrait être choisi pour modéliser les précipitations à la station 8, selon les critères de sélection définis dans le présent travail.

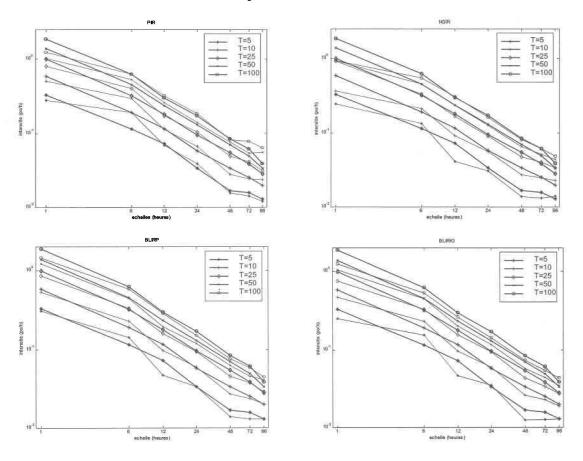


Figure 5.8 : Courbes IDF résultantes des quatre modèles considérés, pour la station 8. Les courbes historiques sont en gras et les simulées correspondent aux minces.

5.6.2 Station 63

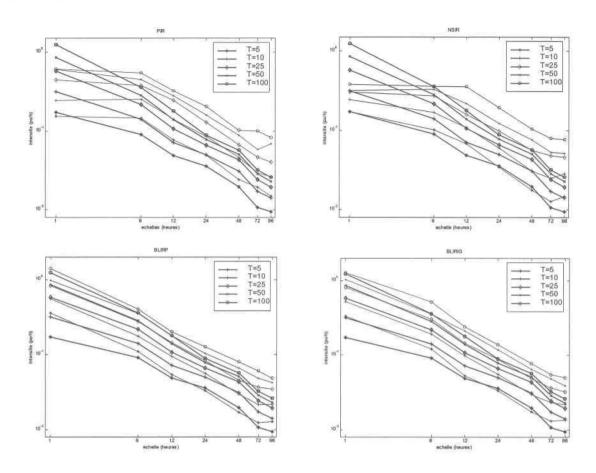


Figure 5.9 : Courbes IDF résultantes des quatre modèles considérés, pour la station 63.

Les courbes historiques sont en gras et les simulées correspondent aux minces.

En ne se basant seulement que sur l'allure des courbes IDF (Figure 5.9) obtenues, le choix du modèle de BLIRP semble évident. Cependant, en s'attardant aux moments (Tableau A.2), le choix n'est plus aussi clair. Tout d'abord, le modèle de PIR peut être éliminé dès le départ, à cause de l'ensemble des résultats obtenus. Comparons donc les modèles de Neyman-Scott et de Bartlett-Lewis. La forme décroissante de la fonction d'autocorrélation de délai un du modèle de NSIR s'approche plus des valeurs historiques que celles du modèle de BLIRP. Les intervalles secs sont surestimés aux grandes échelles par les modèles de NSIR et de BLIRP. Cependant, le modèle de NSIR les surestime moins.

Malgré ces légères différences au niveau de la capacité à reproduire les moments et les proportions d'intervalles secs, il serait difficile de ne pas tenir compte de l'information que donnent les courbes IDF. Le modèle le plus adéquat serait donc le modèle de BLIRP.

5.6.3 Station 369

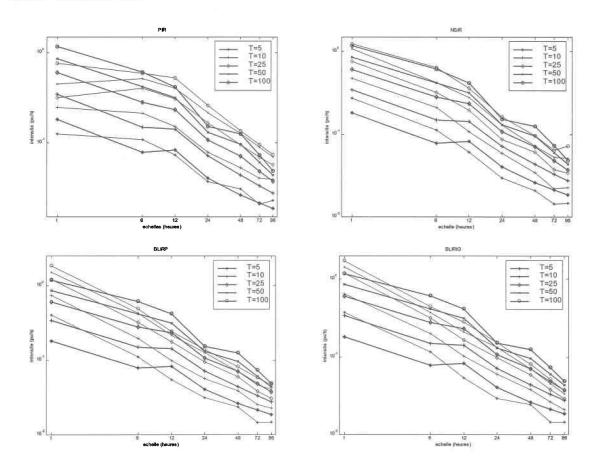


Figure 5.10 : Courbes IDF résultantes des quatre modèles considérés, pour la station 369.

Les courbes historiques sont en gras et les simulées correspondent aux minces.

On peut voir que dans le cas de la station 369, le redressement des courbes simulées à l'échelle horaire que donnent les modèles de Bartlett-Lewis ne vient pas améliorer l'adéquation du modèle aux événements historiques (Figure 5.10), contrairement à ce qui avait été observé à la station 63. On peut noter également que la fonction d'autocorrélation

(Tableau A.3) est généralement mieux reproduite par le modèle de NSIR que par le modèle de BLIRP. Seules les proportions d'intervalles secs sont mieux reproduites par le modèle de BLIRP que par le modèle de NSIR ; cependant, la différence est relativement faible. Ce dernier modèle reproduit généralement mieux les caractéristiques statistiques qui nous intéressent dans le présent travail et le modèle de NSIR devrait être retenu pour modéliser les données de la station 369.

5.6.4 Station 748

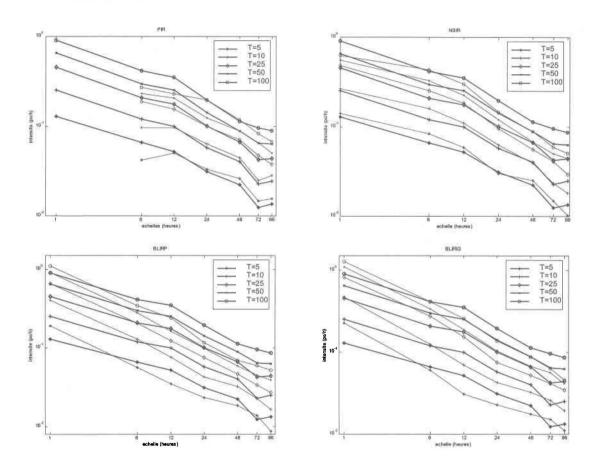


Figure 5.11 : Courbes IDF résultantes des quatre modèles considérés, pour la station 748.

Les courbes historiques sont en gras et les simulées correspondent aux minces.

Les résultats de la station 748 montrent une situation inverse des autres stations. C'est-à-dire que l'analyse des courbes IDF à elle seule serait probablement insuffisante pour déterminer clairement le choix de modèle à utiliser. À part à l'échelle horaire, les événements de périodes de retour 50 et 100 ans sont toujours visiblement sous-estimés par les modèles de Neyman-Scott et de Bartlett-Lewis (Figure 5.11). Aux échelles élevées, on peut noter que le modèle de PIR reproduit très bien les séries d'événements extrêmes, mais aucun événement simulé avec ce modèle à l'échelle horaire ne dépassait le seuil fixé. Il est donc clair que le modèle de PIR ne sera pas retenu. Les fonctions d'autocorrélation donnent des résultats très voisins et ce sont les proportions d'intervalles secs qui tranchent le débat. Le modèle de NSIR reproduit très bien cette caractéristique alors que les modèles de Bartlett-Lewis la surestiment aux grandes échelles. Comme les événements de périodes de retour plus petites que 50 sont bien reproduits par le modèle de NSIR, ce modèle est donc le plus adéquat pour les données de la station 748.

5.6.5 Synthèse de l'étude des résultats

Pour terminer l'étude, la description du modèle final qui devrait être choisit parmi les modèles étudiés pour chaque station, est présentée. Ce choix est basé sur les résultats obtenus et cet ouvrage n'a pas la prétention d'affirmer qu'il s'agit du *meilleur* modèle pour ce type de données.

Le modèle de BLIRP avait donné les meilleurs résultats pour la station 8. En voici les valeurs des paramètres :

λ	μ_X	η	ϕ	K
0,0106	0,2627	1,6025	0,1222	0,3215

C'est à partir de ce modèle que nous avions obtenu les courbes IDF de la Figure 5.12 :

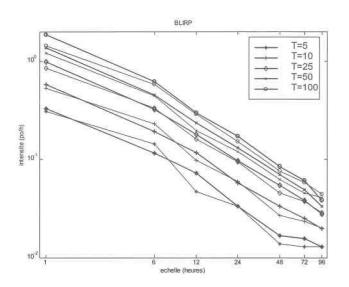


Figure 5.12 : Courbes IDF obtenues avec le modèle de BLIRP à la station 8.

Le modèle de BLIRP avait également été choisi pour représenter la station 63. Les paramètres utilisés sont les suivants :

λ	$\mu_{\scriptscriptstyle X}$	η	ϕ	K
0,0076	0,2892	1,9286	0,0243	0,1208

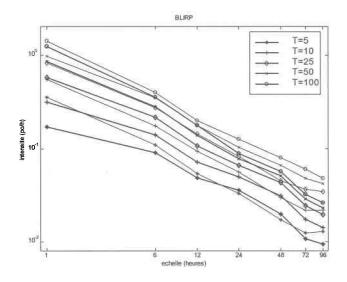


Figure 5.13 : Courbes IDF obtenues avec le modèle de BLIRP à la station 63.

Les courbes IDF de la Figure 5.13 s'en sont suivies.

Le modèle de NSIR avait été désigné pour simuler les séries d'événements pluvieux aux deux autres stations. Les valeurs des paramètres utilisés à chacune de ces stations ainsi que les courbes IDF associées aux modèles sont présentées ici.

• Station 369

λ	β	η	$\mu_{\scriptscriptstyle X}$	μ_{v}
0,0091	0,1705	1,6049	0,2734	5,1697

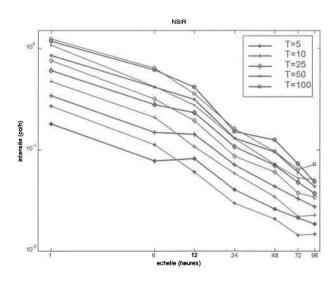


Figure 5.14: Courbes IDF obtenues avec le modèle de NSIR à la station 369.

Station 748

μ_X μ_V
0,1978 2,1727
Ç

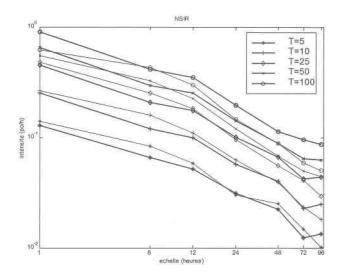


Figure 5.15 : Courbes IDF obtenues avec le modèle de NSIR à la station 748.

6. CONCLUSION ET DISCUSSION

La modélisation de la pluie n'est pas une tâche facile, comme nous avons pu le constater. La complexité du phénomène de précipitation explique pourquoi de nouveaux modèles ont été proposés et pourquoi plusieurs autres sont encore en développement. Les modèles considérés dans ce mémoire ne représentent qu'un échantillon des modèles connus.

Une liste non exhaustive de modèles de type ponctuel a été faite dans une première étape. Par la suite, le principe des courbes IDF a été présenté pour finalement être appliqué sur quatre modèles sélectionnés parmi ceux de la liste de la première étape. À partir de ces résultats, les conclusions qui suivent sont tirées.

Il est important de rappeler que tout ce qui est dit au sujet de la performance des modèles étudiés ici n'est basé que sur les critères de sélection qui avaient été établis au départ, c'est-à-dire la reproduction des moments de premier et deuxième ordre, les proportions d'intervalles secs ainsi que les courbes IDF. L'analyse est à recommencer dès que les critères de sélection du modèle sont modifiés.

En annexe A, chaque tableau (un par station) donne une vue d'ensemble de la performance des quatre modèles mis en application. Comme les deux modèles de B-L ont donné des résultats très semblables, ils seront désignés indistinctement par B-L. On peut voir que tous les modèles donnent des résultats voisins pour ce qui est de la moyenne. Les modèles de N-S et de B-L réussissent toutefois à mieux reproduire les autres moments et les proportions d'intervalles secs que le modèle de PIR. C'est ce à quoi nous devions nous attendre étant donnée la nature plus réaliste des modèles de type *cluster*.

Toute l'information que contiennent les courbes IDF peut servir de base pour le choix du modèle qui sera jugé le plus adéquat. En regardant les graphiques des courbes IDF pour chaque station, la décision devient plus évidente, malgré le fait que le choix n'est pas le même pour toutes les stations. Les stations 8 et 63 sont mieux modélisées par les modèles de BL, tandis que le modèle de NSIR semble plus adéquat pour les deux autres stations, 369 et 748.

Évidemment, plusieurs autres caractéristiques auraient pu être étudiées, comme la longueur moyenne des intervalles secs, la distribution de probabilité pour la période entre deux événements ou encore, le nombre moyen de précipitations par saison. Des modèles exponentiels bivariés ont aussi été développés [Bacchi et al. (1994); Kuroth et al. (1997)] pour tenir compte de la corrélation négative qui existe entre l'intensité et la durée d'une précipitation, qui est une propriété très intéressante.

Les résultats sont obtenus par optimisation en faisant certaines approximations car les méthodes d'estimation existantes ne sont pas toujours satisfaisantes. L'amélioration des méthodes d'estimation pour les modèles est une direction de recherche à privilégier.

On voit donc qu'il y a encore place à l'amélioration et que l'étude des modèles de précipitation continuera sûrement à se développer. Par exemple, les modèles invariants d'échelles (*scaling*) constituent une avenue intéressante pour les problèmes de changements d'échelles.

7. BIBLIOGRAPHIE

- Aron, G., D. J. Wall, E. L. White and C.N. Dunn (1987) Regional rainfall intensity-duration-frequency curves for Pennsylvania. *Water Res. Bulletin*, 23(3), 479-485
- Box, G. E. P. and G. M. Jenkins (1976) Time series analysis, forecasting and control, 2nd ed., Holden-Day, San Francisco
- Bacchi, B., G. Becciu et N. T. Kottegoda (1994) Bivariate exponential model applied to intensities and durations of extreme rainfall. *J. Hydrol.*, 155, 225-236
- Casella, G. and R. L. Berger (1990) Statistical inference, California, Duxbury
- Cox, D.R. and V. Isham (1980) Point processes, London, Chapman and Hall
- Cowpertwait, Paul S. P. (1991) Further developments of the Neyman-Scott clustered point process for modeling rainfall. *Water Resour. Res.*, 27(7), 1431-1438
- Entekhabi, D., I. Rodriguez-Iturbe, and P.S. Eagleson (1989) Probabilistic representation of the temporal rainfall process by the modified Neyman-Scott rectangular pulses model: Parameter estimation and validation. *Water Resour. Res.*, 25(2), 295-302.
- Finkelstein, J. M. and R. E. Schafer (1971) Improved goodness-of-fit tests, *Biometrika*, 58(3), 641-645
- Hosking, J.R.M. and J.R. Wallis (1987) Parameter and quantile estimation for the generalized Pareto distributions, *Technometrics*, 29(3), 339-349
- Johnson, L.N., S. Kotz, A.W. Kemp (1992) Univariate discrete distributions, 2nd edition, New York, Wiley series in probability and mathematical statistics.
- Kavvas, M.L. and J. Delleur (1981) A stochastic cluster model of daily rainfall occurrences. *Water Resour. Res.*, 17(4), 1151-1160.

- Kuroth, R. S., N. K. Goel et B. S. Mathur (1997) Derived flood frequency distribution for negatively correlated rainfall intensity and duration. Water Resour. Res., 33(9), 2103-2107
- Massey, Frank J. (1951) The Kolmogorov-Smirnov test for goodness of fit. *J. Am. Statist.*Ass. 46, 68-78
- Onof, C. and H.S. Wheater (1993) Modeling of British rainfall using a random parameter Bartlett-Lewis rectangular pulse model. *J. Hydrol.*, 149, 67-95.
- Onof, C. and H.S. Wheater (1994) Improvements to the modelling of British rainfall using a modified random parameter Bartlett-Lewis rectangular pulse model. *J. Hydrol.*, 157, 177-195.
- Ouarda, T. M. J. and F. Ashkar (1995) The peaks-over-threshold method for regional flood frequency estimation, 48^e conférence annuelle de l'ACRH: gérer le milieu aquatique.
- Parzen, E. (1962) Stochastic processes, San Francisco, Holden day series in probability and statistics.
- Perreault, L. et B. Bobée (1992) Loi généralisée des valeurs extrêmes: propriétés mathématiques et statistiques et estimation des paramètres et des quantiles X_T de période de retour T, Rapport scientifique no. 350, INRS-Eau
- Reiss, R.-D. (1993) A course on point processes, Springer-Verlag New York inc.
- Rodriguez-Iturbe, I., D.R. Cox, and V. Isham (1987a) Some models for rainfall based on stochastic point processes. *Proc. R. Soc. London Ser. A*, 410, 269-288.
- Rodriguez-Iturbe, I., D.R. Cox, and V. Isham (1988) A point process model for rainfall: further developments. *Proc. R. Soc. London Ser. A*, 283-298.
- Rodriguez-Iturbe, I., B. Febres De Power, and J.B. Valdes (1987b) Rectangular pulses point process models for rainfall: Analysis of empirical data. *J. Geophys. Res.*, 92(D8), 9645-9656.

- Rodriguez-Iturbe, I., V.K. Gupta, and E. Waymire (1984) Scale considerations in the modeling of temporal rainfall. *Water Resour. Res.*, 20(11), 1611-1619.
- Snyder, D. L. (1975) Random point processes, John Wiley and sons inc. (pp. 164-173).
- Taylor, H.M. and Karlin S. (1993) An introduction to stochastic modeling, Academic press.
- Waymire, E. and V.K. Gupta (1981) The mathematical structure of rainfall representations, 3, Some applications of the point process theory to rainfall processes. *Water*

Resourc. Res., 17(5), 1287-1294.

B1		

ANNEXE A

Tableaux des moments historiques et simulés pour les quatre stations

	1 hre	6 hres	12 hres	24 hres	48 hres	72 hres	96 hres
moy. Hist	0,0063	0,0380	0,0760	0,1520	0,3353	0,5092	0,6759
PIR	0,0062	0,0375	0,0744	0,1488	0,2979	0,4464	0,5959
NSIR	0,0061	0,0369	0,0738	0,1476	0,2956	0,4427	0,5912
BLIRP	0.0064	0.0381	0.0762	0.1525	0.3046	0.4574	0.6092
BLIRIG	0.0062	0.0373	0.0746	0.1491	0.2983	0.4474	0.5967
variance hist.	0,0023	0,0333	0,0833	0,1846	0,4289	0,6459	0,9078
PIR	0,0029	0,0380	0,0841	0,1756	0,3554	0,5458	0,7195
NSIR	0,0017	0,0321	0,0803	0,1847	0,4049	0,6094	0,8425
BLIRP	0.0027	0.0363	0.0873	0.2013	0.4339	0.6737	0.9009
BLIRIG	0.0025	0.0340	0.0827	0.1908	0.4160	0.6413	0.8612
acr(1) hist	0,4429	0,2467	0,1450	0,0745	-0,0147	-0,0678	-0,0957
PIR	0,5298	0,0985	0,0352	0,0077	-0,0229	-0,0423	-0,0482
NSIR	0,6765	0,2479	0,1432	0,0672	0,0064	-0,0111	-0,0493
BLIRP	0.4777	0.2108	0.1436	0.0691	0.0131	-0.0195	-0.0333
BLIRIG	0.4840	0.2200	0.1454	0.0716	0.0083	-0.0137	-0.0380
acr(2) hist	0,2460	0,0556	0,0092	-0,0198	-0,0198	-0,0393	-0,1043
PIR	0,1860	-0,0048	-0,0075	-0,0242	-0,0152	-0,0409	-0,0625
NSIR	0,3837	0,0468	0,0080	-0,0156	-0,0314	-0,0459	-0,0619
BLIRP	0.1920	0.0525	0.0102	-0.0134	-0.0270	-0.0319	-0.0478
BLIRIG	0.1940	0.0556	0.0119	-0.0178	-0.0271	-0.0266	-0.0357
acr(3) hist	0,1864	0,0125	-0,0236	-0,0176	-0.0176	0,0231	0,0758
PIR	0,0646	-0,0061	-0,0042	-0,0138	-0,0235	-0,0377	-0,0519
NSIR	0,2372	0,0129	-0,0101	-0,0116	-0,0321	-0,0509	-0,0551
BLIRP	0.1222	0.0192	-0.0068	-0.0177	-0.0265	-0.0434	-0.0642
BLIRIG	0.1210	0.0188	-0.0074	-0.0131	-0.0175	-0.0360	-0.0588
P(int. sec) hist	0,9484	0,8870	0,8325	0,7365	0,5330	0,3989	0,2980
PIR	0,9617	0,8712	0,7743	0,6110	0,3795	0,2361	0,1490
NSIR	0,9459	0,8836	0,8317	0,7481	0,6057	0,4916	0,3978
BLIRP	0.9567	0.8964	0.8410	0.7404	0.5767	0.4488	0.3538
BLIRIG	0.9568	0.8964	0.8416	0.7415	0.5779	0.4495	0.3543

Tableau A.1 : Moments historiques et simulés et probabilités de périodes sèches pour les quatre modèles étudiés à la station 8.

	1 hre	6 hres	12 hres	24 hres	48 hres	72 hres	96 hres
moy. Hist	0.0069	0.0411	0.0823	0.1645	0.3544	0.5331	0.7059
PIR	0.0068	0.0409	0.0819	0.1637	0.3275	0.4912	0.6549
NSIR	0.0066	0.0398	0.0796	0.1591	0.3183	0.4774	0.6366
BLIRP	0.0067	0.0402	0.0804	0.1609	0.3217	0.4826	0.6434
BLIRIG	0.0066	0.0398	0.0796	0.1592	0.3183	0.4775	0.6367
variance hist.	0.0018	0.0289	0.0763	0.1957	0.4843	0.7488	0.9856
PIR	0.0017	0.0354	0.0881	0.1984	0.4145	0.6336	0.8478
NSIR	0.0015	0.0298	0.0789	0.1925	0.4358	0.6808	0.9328
BLIRP	0.0025	0.0296	0.0735	0.1877	0.4738	0.7987	1.1090
BLIRIG	0.0024	0.0292	0.0725	0.1850	0.4678	0.7941	1.1032
acr(1) hist	0.5215	0.3375	0.2639	0.1215	0.0795	-0.0001	0.0926
PIR	0.7592	0.2521	0.1090	0.0368	0.0041	-0.0180	-0.0382
NSIR	0.7350	0.3171	0.2118	0.1163	0.0439	0.0091	-0.0230
BLIRP	0.3806	0.2415	0.2731	0.2496	0.1427	0.0777	0.0409
BLIRIG	0.3930	0.2418	0.2735	0.2522	0.1530	0.0753	0.0337
acr(2) hist	0.3474	0.1314	0.0296	-0.0122	-0.0208	0.0422	-0.1801
PIR	0.4940	0.0086	-0.0086	-0.0136	-0.0307	-0.0398	-0.0510
NSIR	0.4645	0.0991	0.0325	0.0021	-0.0262	-0.0387	-0.0548
BLIRP	0.1198	0.1598	0.1435	0.0592	-0.0113	-0.0393	-0.0485
BLIRIG	0.1225	0.1595	0.1431	0.0652	-0.0197	-0.0403	-0.0501
acr(3) hist	0.2418	0.0652	-0.0061	-0.0272	-0.0191	-0.2067	0.0945
PIR	0.3204	-0.0072	-0.0028	-0.0088	-0.0295	-0.0296	-0.0293
NSIR	0.3064	0.0436	0.0069	-0.0167	-0.0309	-0.0467	-0.0562
BLIRP	0.0828	0.1189	0.0678	0.0025	-0.0231	-0.0370	-0.0484
BLIRIG	0.0831	0.1210	0.0769	0.0030	-0.0246	-0.0417	-0.0456
P(int. sec) hist	0.9349	0.8641	0.8074	0.7120	0.5184	0.3810	0.2873
PIR	0.9348	0.8439	0.7464	0.5846	0.3592	0.2154	0.1323
NSIR	0.9284	0.8539	0.7953	0.7069	0.5721	0.4651	0.3797
BLIRP	0.9447	0.8530	0.8019	0.7296	0.6102	0.5126	0.4283
BLIRIG	0.9450	0.8539	0.8031	0.7322	0.6140	0.5190	0.4334

Tableau A.2 : Moments historiques et simulés et probabilités de périodes sèches pour les quatre modèles étudiés à la station 63.

	1 hre	6 hres	12 hres	24 hres	48 hres	72 hres	96 hres
moy. Hist	0.0079	0.0476	0.0951	0.1903	0.3983	0.5971	0.7991
PIR	0.0077	0.0461	0.0922	0.1845	0.3689	0.5546	0.7395
NSIR	0.0080	0.0479	0.0958	0.1915	0.3830	0.5743	0.7658
BLIRP	0.0083	0.0497	0.0994	0.1988	0.3975	0.5964	0.7952
BLIRIG	0.0080	0.0478	0.0956	0.1912	0.3823	0.5721	0.7628
variance hist.	0.0023	0.0390	0.1056	0.2566	0.6051	0.9089	1.1367
PIR	0.0021	0.0440	0.1129	0.2510	0.5231	0.8028	1.0864
NSIR	0.0033	0.0450	0.1114	0.2581	0.5540	0.8702	1.1630
BLIRP	0.0044	0.0438	0.1080	0.2684	0.6278	1.0082	1.3774
BLIRIG	0.0041	0.0409	0.1014	0.2495	0.5865	0.9341	1.2700
acr(1) hist	0.5374	0.3424	0.2173	0.1037	0.0165	-0.0406	0.0192
PIR	0.7642	0.2639	0.1158	0.0428	0.0003	-0.0264	-0.0350
NSIR	0.4805	0.2259	0.1511	0.0720	0.0221	-0.0134	-0.0187
BLIRP	0.2352	0.2333	0.2288	0.1580	0.0666	0.0164	-0.0057
BLIRIG	0.2427	0.2401	0.2285	0.1634	0.0622	-0.0009	-0.0334
acr(2) hist	0.3490	0.1067	0.0234	-0.0428	-0.0517	0.0159	0.0099
PIR	0.5062	0.0225	-0.0086	-0.0231	-0.0228	-0.0123	-0.0311
NSIR	0.1979	0.0657	0.0100	-0.0109	-0.0185	-0.0271	-0.0541
BLIRP	0.0876	0.1267	0.0719	0.0026	-0.0238	-0.0421	-0.0498
BLIRIG	0.0919	0.1241	0.0734	-0.0009	-0.0456	-0.0488	-0.0417
acr(3) hist	0.2551	0.0454	-0.0272	-0.0284	-0.0002	-0.0026	0.1273
PIR	0.3371	0.0012	-0.0139	-0.0133	-0.0143	-0.0406	-0.0754
NSIR	0.1222	0.0200	-0.0077	-0.0108	-0.0182	-0.0415	-0.0573
BLIRP	0.0751	0.0713	0.0202	-0.0138	-0.0324	-0.0398	-0.0649
BLIRIG	0.0776	0.0709	0.0160	-0.0240	-0.0414	-0.0383	-0.0685
P(int. sec) hist	0.9342	0.8608	0.8016	0.6974	0.5034	0.3757	0.2553
PIR	0.9318	0.8407	0.7438	0.5812	0.3531	0.2169	0.1290
NSIR	0.9443	0.8638	0.8072	0.7207	0.5786	0.4686	0.3733
BLIRP	0.9475	0.8550	0.7978	0.7017	0.5429	0.4249	0.3302
BLIRIG	0.9486	0.8577	0.8009	0.7056	0.5514	0.4316	0.3382

Tableau A.3 : Moments historiques et simulés et probabilités de périodes sèches pour les quatre modèles étudiés à la station 369.

	1 hre	6 hres	12 hres	24 hres	48 hres	72 hres	96 hres
moy. Hist	0.0060	0.0361	0.0721	0.1443	0.3220	0.4832	0.6451
PIR	0.0059	0.0354	0.0708	0.1417	0.2839	0.4258	0.5660
NSIR	0.0059	0.0352	0.0705	0.1409	0.2826	0.4238	0.5683
BLIRP	0.0059	0.0354	0.0709	0.1418	0.2841	0.4261	0.5685
BLIRIG	0.0061	0.0369	0.0738	0.1476	0.2955	0.4433	0.5924
variance hist.	0.0017	0.0271	0.0708	0.1801	0.4556	0.7363	0.9760
PIR	0.0010	0.0246	0.0712	0.1760	0.3925	0.6085	0.8185
NSIR	0.0020	0.0356	0.0866	0.1965	0.4434	0.6946	0.9911
BLIRP	0.0023	0.0263	0.0674	0.1719	0.4175	0.6685	0.9385
BLIRIG	0.0025	0.0282	0.0720	0.1849	0.4544	0.7209	1.0092
acr(1) hist	0.4944	0.3434	0.2449	0.1092	0.0864	-0.0491	-0.0267
PIR	0.8627	0.4393	0.2279	0.0950	0.0164	-0.0135	-0.0429
NSIR	0.6925	0.2078	0.1309	0.1124	0.0737	0.0536	0.0127
BLIRP	0.3294	0.2733	0.2665	0.1934	0.0894	0.0412	0.0091
BLIRIG	0.3441	0.2778	0.2761	0.1979	0.0900	0.0411	-0.0009
acr(2) hist	0.3271	0.1155	0.0280	-0.0045	-0.0279	0.0014	-0.0547
PIR	0.6869	0.1017	0.0057	-0.0181	-0.0236	-0.0309	-0.0345
NSIR	0.3876	0.0458	0.0464	0.0201	-0.0014	-0.0305	-0.0488
BLIRP	0.1221	0.1584	0.0957	0.0223	-0.0190	-0.0278	-0.0725
BLIRIG	0.1196	0.1634	0.0993	0.0230	-0.0223	-0.0402	-0.0531
acr(3) hist	0.2382	0.0392	-0.0013	0.0013	-0.0199	-0.1107	-0.2183
PIR	0.5463	0.0182	-0.0058	-0.0173	-0.0234	-0.0391	-0.0697
NSIR	0.2197	0.0282	0.0226	0.0148	-0.0283	-0.0379	-0.0486
BLIRP	0.0996	0.0947	0.0340	-0.0146	-0.0286	-0.0628	-0.0569
BLIRIG	0.1004	0.0946	0.0317	-0.0110	-0.0364	-0.0283	-0.0500
P(int. sec) hist	0.9464	0.8873	0.8371	0.7513	0.5774	0.4530	0.3661
PIR	0.9196	0.8510	0.7752	0.6442	0.4427	0.3049	0.2078
NSIR	0.9558	0.8858	0.8182	0.7143	0.5712	0.4695	0.3880
BLIRP	0.9529	0.8819	0.8393	0.7671	0.6404	0.5361	0.4503
BLIRIG	0.9513	0.8790	0.8364	0.7639	0.6368	0.5336	0.4403

Tableau A.4 : Moments historiques et simulés et probabilités de périodes sèches pour les quatre modèles étudiés à la station 748.