Auto corrélation et Auto information des données pluviométriques

4.1 Introduction

Dans ce chapitre, nous nous intéressons en particulier à la manifestation des effets de mémoire. La structure fractale du phénomène décrit par la série temporelle est explorée par la dimension fractale et la corrélation entre accroissements passés et futurs. Les résultats obtenus dans cette étude peuvent représenter la série temporelle d'une mémoire à long terme ou d'une mémoire à court terme.

4.2 Fonctions d'auto corrélation et d'auto information

Deux résultats pertinents sont retenus pour la présente étude : le comportement asymptotique de la fonction d'auto corrélation et d'auto-information ; ainsi le distinguo entre mémoire courte et longue porte sur la décroissance de la fonction d'auto corrélation et d'auto-information au fur et à mesure que le retard augmente.

La décroissance est rapide, exponentielle pour une mémoire à court terme et plus lente, hyperbolique pour la mémoire à long terme. (Voir Annexe C)

En utilisant les données pluviométriques que nous disposons, on obtient les courbes d'auto corrélation (Figure 16) et auto information (Figure 17) suivantes :

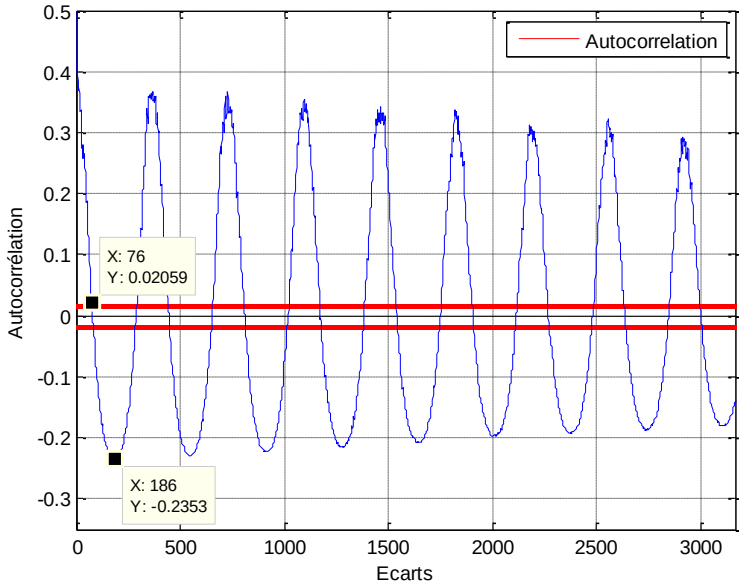

Figure 16 : Courbe d'auto corrélation.

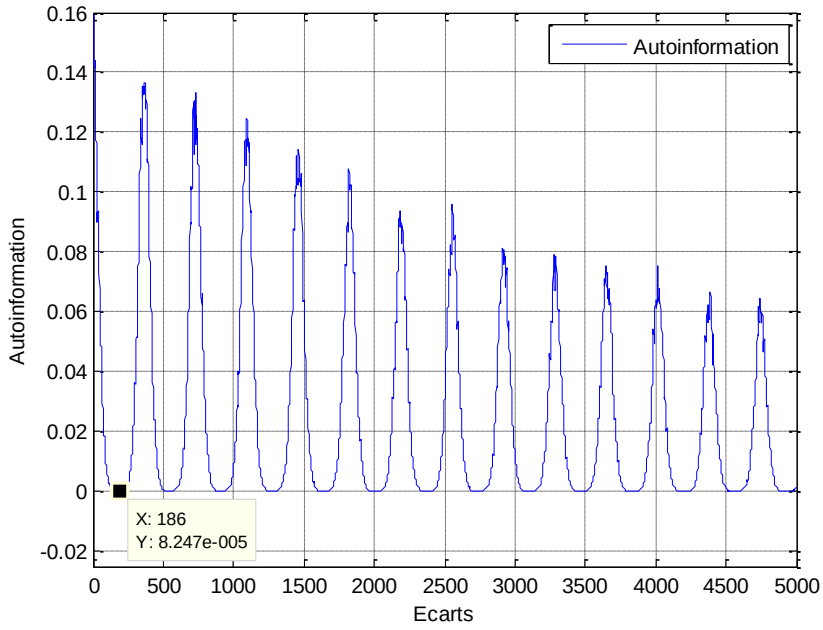

Figure 17 : Courbe d'auto information.

Les figures 16 et figure 17 montrent respectivement les évolutions des fonctions d'auto corrélation et d'auto information de la chronique des précipitations en représentation linéaire et logarithmique.

Les premiers minima des deux fonctions s'établissent autour de 186 jours. On a aussi une courbe linéaire de 0 jusqu'au 120ème jour.

Auto corrélation : $\rho(k) = \frac{\gamma(k)}{\gamma(0)}$, avec $\rho(k)$ le coefficient de corrélation normalisé, $\gamma(k)$ le coefficient de corrélation défini ci-dessous, k le décalage en jours, x_t la chronique de pluviométrie et x_{t+k} la chronique de débit, et $\gamma(0)$ les écart-types correspondants à x_t et x_{t+k} respectivement.

$$\gamma(k) = \frac{1}{T}\sum_{t=1}^{T}(x_t - \overline{x_t})(x_{t+k} - \overline{x_t}) \text{ et}$$

$$\gamma(0) = \frac{1}{T}\sum_{t=1}^{T}(x_t - \overline{x_t})(x_t - \overline{x_t}) \qquad (4.1)$$

Pour déterminer le seuil de corrélation à partir duquel les deux séries sont indépendantes, il faut déterminer l'erreur sur l'estimation de $\gamma(k)$. L'intervalle de confiance calculé sur $\gamma(k)$ permet de connaître le seuil à partir duquel $\rho(k)$ peut être considéré comme égal à 0.

Deux cas sont à distinguer :

- le nombre de mesure grand : (T > 300) ;

- le plus réduit (T ≤ 300).

4.2.1 Grand nombre de mesures (n > 300)

Dans ce cas, on considère que le coefficient de corrélation suit une loi de distribution normale. On peut alors déterminer l'intervalle de confiance de $\gamma(k)$ de la façon suivante:

L'intervalle de confiance au seuil de 95% est de $\gamma(k) = \pm 1{,}96 * \sigma_r$.

Et l'intervalle de confiance au seuil de 99% est de $\gamma(k) = \pm 1{,}96 * \sigma_r$.

Où $\sigma_r = \frac{1}{\sqrt{T}}$. (Voir Annexe E) (4.2)

Auto information : $I(X(t), X(t+\tau)) = \sum_{t,t+\tau} P(X(t), X(t+\tau)) \log \frac{P(X(t),X(t+\tau))}{P(X(t))P(X(t+\tau))}$ (4.3)

Les auto-corrélations notées $\rho(k)$ sont les quotients des covariances par la variance et l'auto information d'une variable désignant l'information mutuelle d'une seule variable ; c'est-à-dire la mesure de dépendance entre une variable et lui-même.

D'après la formule de la corrélation qu'on a utilisée, on trouve l'intersection de l'intervalle de confiance au seuil 95% avec la courbe d'auto corrélation au $76^{ème}$ jour.

Donc les événements de pluie espacés pendant 76 jours (2 mois et 16 jours) sont statistiquement corrélés.

Ces constatations semblent indiquer que l'auto corrélation et l'auto information mesurent les mêmes phénomènes à long terme.

La figure 18 montre une décroissance quasi linéaire (sous une représentation log-log) ce qui induit, pour la courbe en elle-même, une décroissance de type puissance. On peut interpréter ceci par le fait que la persistance d'un temps pluvieux ou d'une période de sècheresse n'influe pas sur le comportement conjugué des deux phénomènes.

Autrement dit, les successions pluie-sècheresse sont relativement indépendantes. Cela décrit des alternances de périodes assez longues de pluie et de sècheresse.

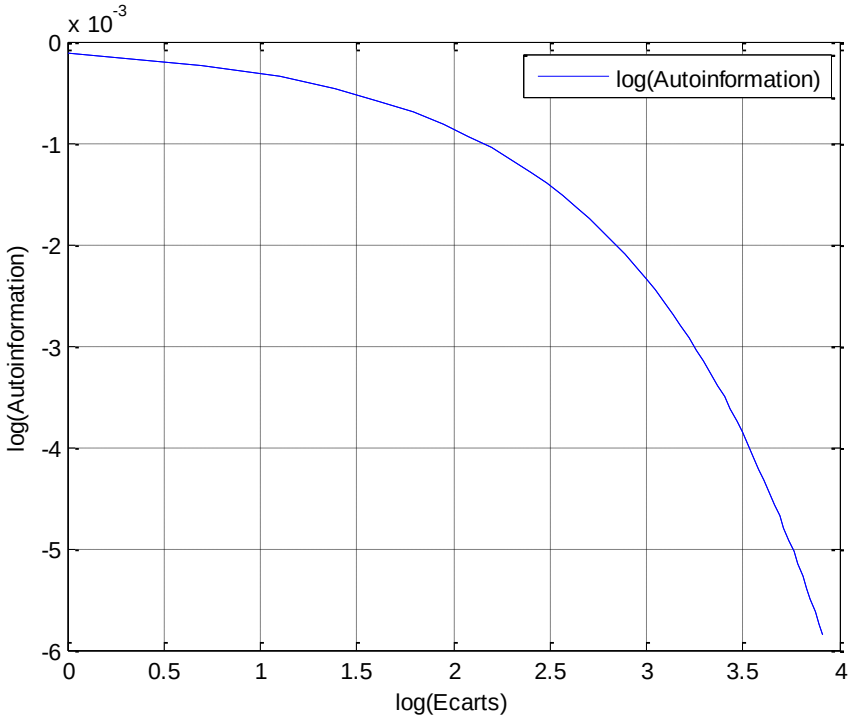

Figure 18 : Courbe de log d'auto information

4.3 Cumul des pluies journalières

Dans cette étude, nous déterminons le point qui a une anomalie sur la courbe pour les données pluviométriques de 1979 à 2015. On interprète éventuellement les observations douteuses durant ces 37 ans passés.

La courbe des hauteurs de pluie cumulées représente en ordonnée, pour chaque instant t, l'intégrale de la hauteur de pluie journalière tombée depuis le début de l'averse durant le nombre des années à étudier.

La courbe représentée à la figure 19 nous donne le cumul de pluie journalière des années 1979 à 2015.

D'après cette figure, on constate que, il n'y a pas de changement climatique du point de vue pluie.

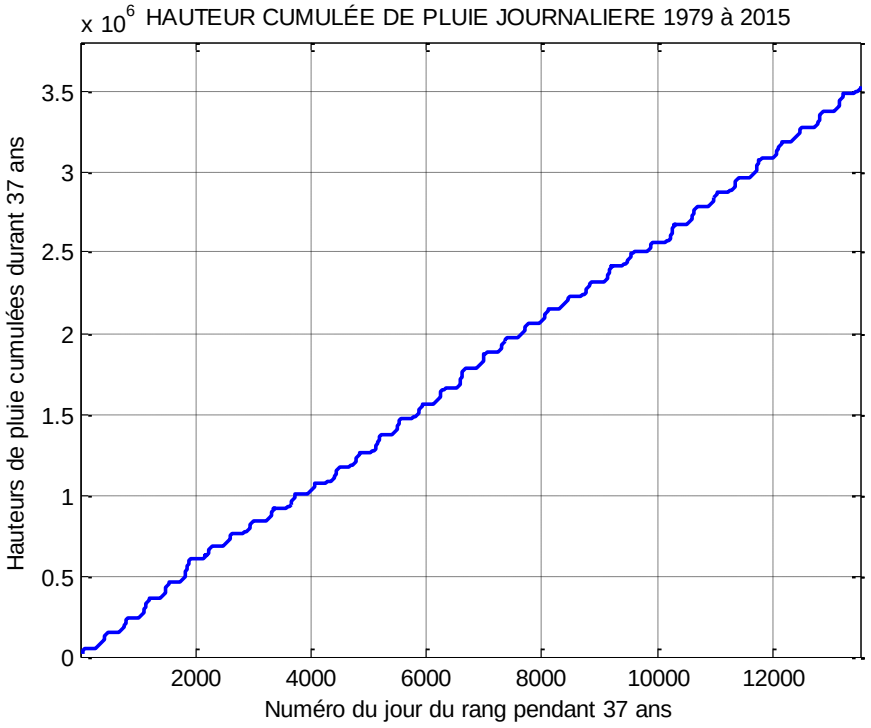

Figure 19 : Courbe des hauteurs de pluie cumulées des précipitations.

4.4 Ajustement des valeurs des cumuls de pluie par une loi de probabilité

Nous allons définir une loi de probabilité à partir de la fonction de répartition des valeurs de cumuls de pluies.

En théorie de probabilité et en statistique, une densité de probabilité est une fonction qui permet de représenter une loi de probabilité sous forme d'intégrales. C'est un outil d'analyse de données utilisé dans des domaines aussi variés que l'étude de la climatologie.

Comme hypothèse, le phénomène pluvieux local est régi par un processus stochastique, qui reste à définir. L'intensité de précipitation est la variable aléatoire, dont on s'attachera à déterminer les caractéristiques.

Ainsi, quelques propriétés mathématiques méritent d'être soulignées pour pouvoir mener bien à terme ces recherches.

Nous avons déjà vu que la fonction de répartition des données pluviométriques (figure 19) est une fonction croissante et dont les valeurs sont positives.

Cette allure de courbe nous incite à choisir des lois mathématiques semblant avoir la même variation pour leur fonction de répartition pour faire la comparaison, à savoir :

- La loi **Log normale,**
- La loi **Beta,**

- La loi **Gamma,**
- Et la loi **Généralisée de Pareto.**

4.4.1 Comparaison de la fonction des valeurs cumulées de pluie avec les différentes lois de probabilité [82] [83] [84] [85] [86]

Loi Log Normale

La loi log-normale qui est aussi appelée loi de Galton s'énonce comme suit :

Une variable aléatoire X suit une loi log-normale quand son logarithme suit une loi normale. C'est à dire que $Y = \mathrm{Ln}(X)$ suit une loi $N(\alpha\,;\,\beta)$ avec α et β les paramètres de la loi normale.

$$g(x) = \frac{1}{x} * \frac{1}{\beta\sqrt{2\pi}} * e^{-\frac{1}{2}(\frac{Ln(x)-\alpha}{\beta})^2} \qquad (4.4)$$

α et β sont les paramètres :

$\alpha = \mu_y$ (espérance de $Ln(x)$) et $\beta = \sigma_y$ (écart type de $Ln(x)$).

Pour la figure 20, la loi log normale a pour paramètres $\mu = -0{,}9826$ et $\sigma = 0{,}9465$. D'où la distribution log normale n'arrive pas suivre la variation de la courbe des cumuls de hauteurs de pluie. Finalement, la courbe de cumul de pluie ne suit pas une loi log normale.

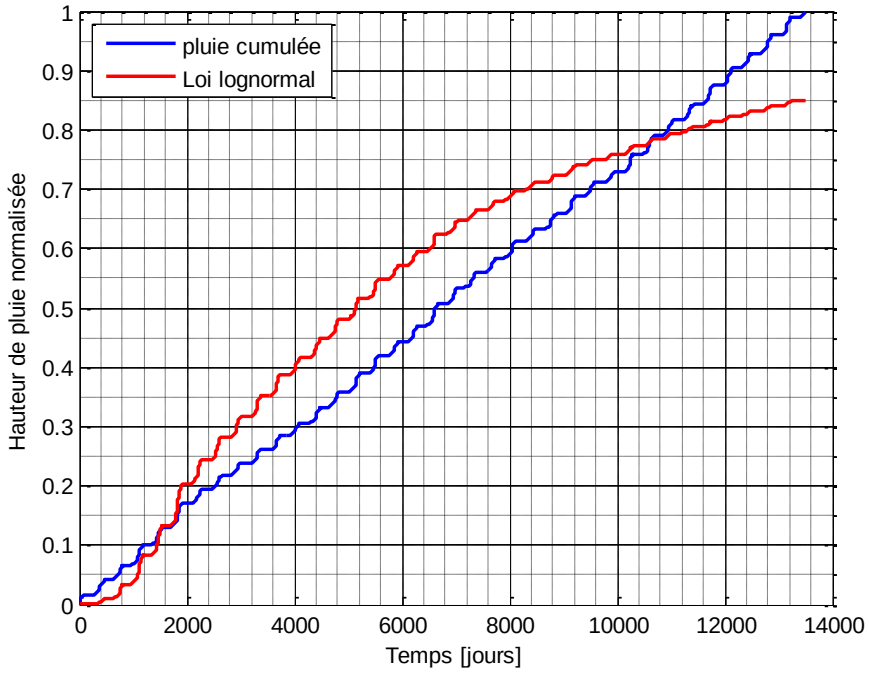

Figure 20 : Ajustement de la loi log normale sur la hauteur de pluie cumulée

Loi Beta

La loi bêta est définie pour $0 \leq x \leq 1$,

$$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+1)} x^{\alpha}(1-x)^{\beta} & si\ x \epsilon]0;1[\\ 0 & si\ x \notin]0;1[\end{cases} \quad (4.5)$$

avec $\alpha = -1$ et $\beta = -1$.

La moyenne vaut $\frac{\alpha}{\alpha+b}$ et la variance $\frac{\alpha b}{(\alpha+b)^2(\alpha+b+1)}$. Les deux paramètres de forme α et b doivent être positifs.

D'après la figure 21, les deux courbes sont quasiment confondues et l'écart entre eux est très petit. Pour notre cas, la loi Beta a pour paramètre n= 1,0396 et p=1.0059.

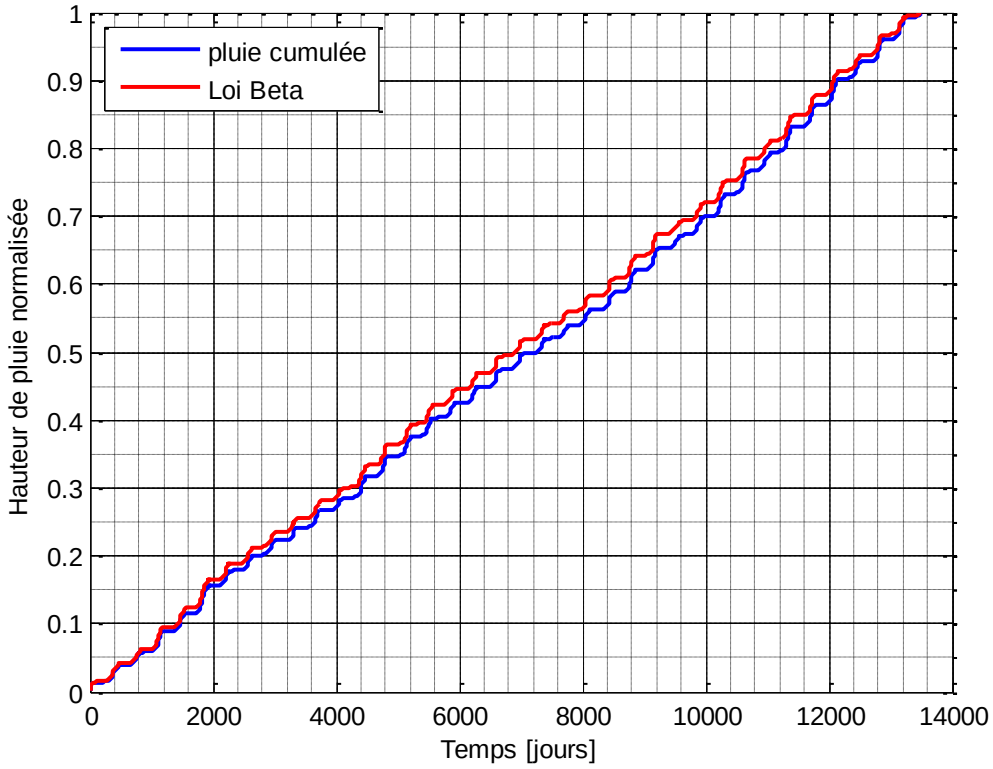

Figure 21 : Ajustement de la loi Beta sur la hauteur de pluie cumulée

Loi Gamma

La loi gamma est définie pour $x \geq 0$ et telle que,

$$\frac{\lambda^k}{\Gamma(k)} x^{k-1} \exp(-\lambda x) \quad (4.6)$$

où λ et k sont deux paramètres positifs représentant respectivement l'échelle et la forme de la distribution. Le nom de la loi provient de son dénominateur $\Gamma(k)$ qui assure la normalisation de la distribution. La moyenne vaut $\frac{k}{\lambda}$ et la variance $\frac{k}{\lambda^2}$. Lorsque $k = 1$, la distribution se réduit à l'exponentielle. La loi Gamma est proche de la loi de Poisson mais avec une queue moins importante, ce qui donne des probabilités d'occurrence plus petites aux valeurs extrêmes.

La courbe de la fonction cumulée de la loi gamma (Figure 22) n'arrive pas à suivre la courbe de la hauteur de pluie cumulée. Cela signifie que la courbe cumulée de pluie ne suit pas la loi de probabilité Gamma. Dans notre cas, les paramètres de la loi Gamma sont k=1.8813 et teta= 0,2656.

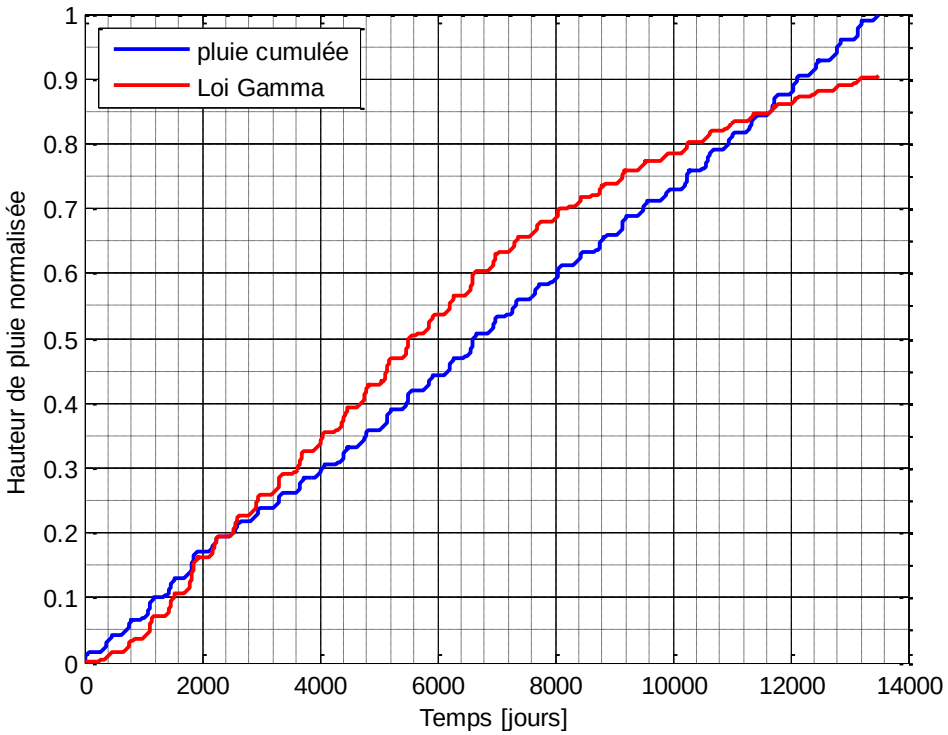

Figure 22 : Ajustement de la loi Gamma sur la hauteur de pluie cumulée

Loi généralisée de PARETO

Cette loi, aussi connue sous le nom de loi hyperbolique ou loi de puissance, doit son nom à l'économiste italien Vilfredo Pareto qui l'a utilisé à la fin du $XIX^{ième}$ siècle pour représenter les enrichissements personnels dans certaines sociétés. Elle est définie pour $x \geq a$ et pour des paramètres de forme a et b positifs,

$$\frac{b\alpha^b}{x^{1+b}} \tag{4.7}$$

La moyenne est égale à $\frac{\alpha b}{(b-1)}$ pour $b > 1$.

Pour $b > 2$, la variance vaut,

$$\frac{\alpha^2 b}{[(b-1)^2(b-2)]} \tag{4.8}$$

et est infinie pour $b \leq 2$.

La loi de Pareto est très souvent utilisée pour représenter les lois d'échelle rencontrées dans la nature. Avec cette loi, la probabilité que la variable $x > u > a$ vaut $(\frac{\alpha}{u})^b$. Une application particulière de la loi de Pareto est la modélisation des pics de crues.

Les paramètres de la loi généralisée de Pareto sont déterminés par les valeurs k = -0,9905 ; x_m = 0,9905 et $\theta = 0$. Pour $k < 0$, la distribution de Pareto généralisée se comporte comme celle de Beta pour les séries de taille finie. Les deux courbes nous montrent que le cumul de hauteur de pluie suit une loi généralisée de Pareto.

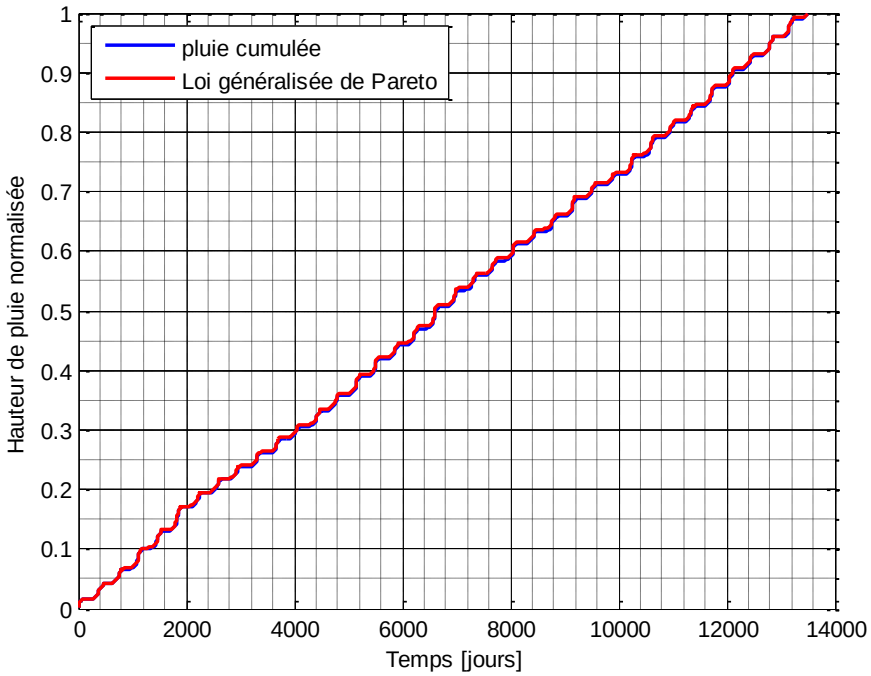

Figure 23 : Ajustement de la loi généralisée de Pareto sur la hauteur de pluie cumulée

Les figures ci dessous (Figure 24 et 25) représentent les répartitions pluviométriques dans la région de Boeny et celle de Madagascar.

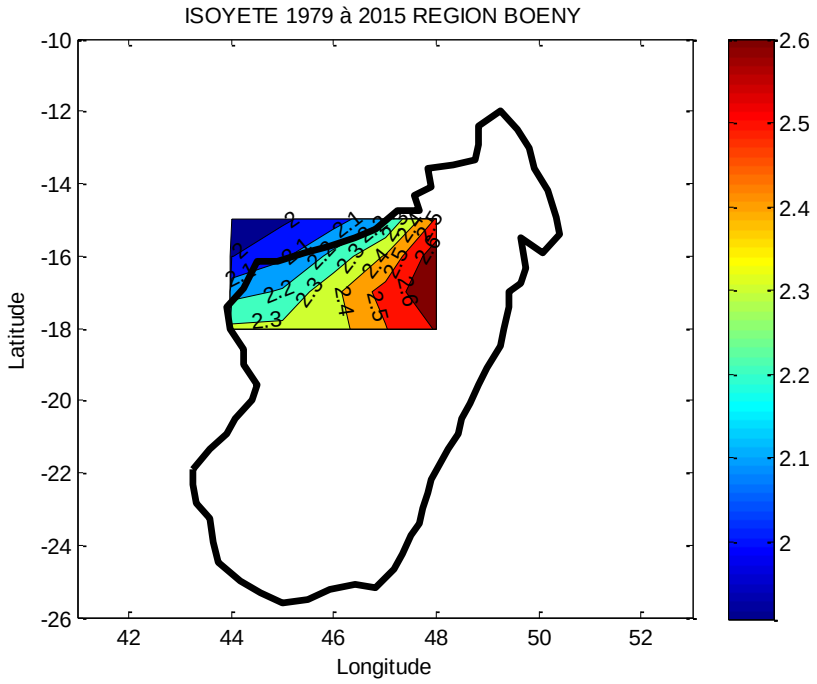

Figure 24 : Répartitions pluviométriques dans la région Boeny

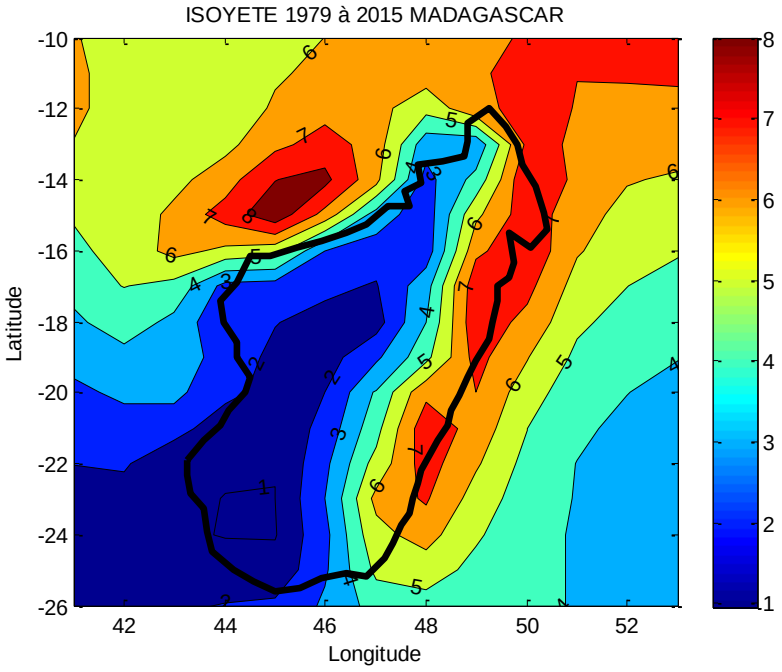

Figure 25 : Répartitions pluviométriques de Madagascar

Dans la région de Boeny, en analysant l'évolution pluviométrique durant ces 37 années, les hauteurs pluviométriques varient de 2 mm à 2,6 mm.

Si on fait la comparaison entre les résultats de la figure 25 et figure 26, les hauteurs de pluie varient de 2 mm à 3 mm dans la région de Boeny. Les valeurs trouvées sont identiques.

4.5 Conclusion

La courbe de précipitation présente l'effet mémoire à long terme durant 37 années. En constatant la courbe de cumul de pluie journalière, on remarque qu'il n'y a pas de changement climatique du point de vue pluie. Pour l'ajustement du cumul de pluie avec la loi de probabilité, il suit une loi généralisée de Pareto.