

Études de la prévisibilité par l'autocorrélation et l'autoinformation des épisodes sèches dans la zone des hauts plateaux de Madagascar

Théodore TOTOZAFINY^{1,2*}, Roselin RANDRIANANTENAINA²
et Adolphe RATIARISON²

¹ Institut Supérieur des Sciences et Technologies de Mahajanga, Université de Mahajanga,
Bâtiment Ex-LOLO Mahajanga be, Madagascar

² Laboratoire de la Dynamique de l'Atmosphère du Climat et des Océans, Facultés des Sciences,
Université d'Antananarivo, Madagascar

* Correspondance, courriel : theodore.totozafiny@gmail.com

Résumé

Cet article étudie particulièrement la pluviométrie dans la région haute terre de Madagascar. La pluviométrie est l'un des paramètres qui intervient dans le réchauffement climatique. Nous disposons des données de la pluie journalière de 1989 au 2014, fournies par ECMWF (*European Centre for Medium range Weather Forecasts*) à l'échelle synoptique avec une grille de 1° x 1° pour conduire notre étude. La méthode par l'analyse en composantes principales (ACP) a été utilisée afin de caractériser la ressemblance du groupement des données. Les cumules de la pluie journalière, mensuelle et annuelles ont été développés pour observer l'autocorrélation et l'autoinformation. Ces deux informations qui sont combinées avec l'erreur de prédiction régressive, nous ont permis de définir la périodicité pour chaque année de la saison de pluie, du temps d'attente et de l'épisode sèche. Les résultats montrent qu'il existe un épisode d'incertitude ni de pluie ni de sèche nécessitant une nouvelle étude.

Mots-clés : *analyse, pluviométrie, ACP, autocorrélation, autoinformation, épisode sèche.*

Abstract

Studies of the predictability by autocorrelation and autoinformation dry episodes in the area of the highlands of Madagascar

This article specifically studies the pluviometry in the highland area of Madagascar. The pluviometry is a parameter that is involved in global warming. We have collected data of daily rainfall from 1989 to 2014, provided by ECMWF (*European Centre for Medium range Weather Forecasts*) on a synoptic scale with 1° x 1° grid to perform our study. The method by principal component analysis (PCA) was used to characterize the similarity of the data group. The cumulated in daily rainfall, monthly and yearly have been developed to observe the autocorrelation and autoinformation. These information that are combined with the error of regressive prediction, allowed us to define the periodicity for each year of the rainy season, the waiting period and the dry weather episode. The results show that there is an uncertainty episode, neither rain nor dry requiring a new study.

Keywords: *analysis, pluviometry, PCA, autocorrelation, autoinformation.*

1. Introduction

Le changement climatique est un phénomène qui touche le territoire de Madagascar où l'inondation et la sécheresse font chaque année des dégâts considérables selon l'endroit de la Grande Ile [1]. Les pouvoirs publics malgaches, les ONG, les associations et les organismes de volontariats sont très préoccupés par l'impact du réchauffement climatique dans la vie quotidienne de la population malgache. Nombreuses études ont été réalisées par des équipes de recherche se trouvant localement et à l'étranger pour mener des travaux sur le changement climatique à Madagascar [2]. Ce dernier est dû aux phénomènes naturels caractérisés par l'angle d'inclinaison du globe terrestre et aux interventions humaines au niveau de différents écosystèmes tels que l'océan, les forêts, les massifs montagneux, etc. [3]. L'un de facteur climatique déterminant dans la caractérisation climatique des différentes régions de Madagascar est la pluviométrie. Notre étude se focalise sur les phénomènes qui ont à l'origine de pluie, au temps d'attente (période non pluie) et la période sèche. Dans cette optique, la proposition d'un modèle des processus pluvieux dans une zone d'étude et la détermination de l'épisode sèche à partir du temps d'attente sont des étapes primordiales pour conduire cette étude. Les cumules de pluies journalières, mensuelles et annuelles sont à étudier afin d'observer l'autocorrélation et l'autoinformation des données. Nous disposons en effet des données de la pluie journalière de 1989 au 2014, fournies par ECMWF (European Centre for Medium range Weather Forecasts) à l'échelle synoptique avec une grille de $1^\circ \times 1^\circ$ pour mener notre travail de recherche.

2. Méthodologie

La prévisibilité de pluies dans la zone de notre étude nécessite l'analyse successive des données climatologiques de la pluviométrie. C'est un paramètre fondamental que nous sommes intéressés. Les données réelles des quantités journalières que nous disposons ont été obtenues pour les coordonnées spatiales entre la latitude -16° à -22° et longitude et 46° à 49° en utilisant la résolution de $1 \times 1^\circ$. Dans la littérature, il existe plusieurs techniques qui étudient de la pluviométrie [4]. Nous avons sélectionné la méthode par l'analyse en composantes principales pour ses performances de regrouper les données brutes par leurs similarités.

2-1. Analyse en composantes principales (ACP)

L'ACP effectue un traitement sur un tableau rectangulaire de données comportant des valeurs des variables quantitatives p pour n unités. Cette analyse permet la représentation géométrique du couple variable et unité (p, n) [5]. La similarité des données sont représentées à l'aide de ses bornes. Les objectifs de l'utilisation de l'ACP sont de distinguer les groupes dans l'ensemble des unités et de définir la corrélation entre deux variables. Pour ce faire, on définit les paramètres suivants :

La moyenne de la variable x_i notée par \bar{x}_j :

$$\bar{x}_j = \sum \left\{ \frac{m_i}{M} x_{ij} \mid i \in I \right\} \text{ avec } M = \sum \{m_i \mid i \in I\} \quad (1)$$

Où I est l'ensemble de lignes de la matrice X et m_i est le poids affecté à l'individu i :

$$m_i = \frac{1}{\text{card}I} \text{ et } \sum \{m_i \mid i \in I\} = M = 1 \quad (2)$$

La variance de la variable x_i notée par σ_j^2 :

$$\sigma_j^2 = var(x_i) = \sum \left\{ \frac{m_i}{M} (x_{ij} - \bar{x}_j)^2 \mid i \in I \right\} \quad (3)$$

Les composantes sur l'ensemble I de la variable centrée et réduite :

$$X_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j} \quad (4)$$

Où σ_j est l'écart type de x_j :

$$moy(X_j) = \sum \{X_{ij} \mid i \in I\} = 0 \quad (5)$$

$$var(X_j) = \sum \{[X_{ij} - moy(X_j)]^2 \mid i \in I\} = 1 \quad (6)$$

Le coefficient de corrélation linéaire entre deux variables x_j et $x_{j'}$, borné entre -1 et +1 :

$$r_{jj'} = \sum \left\{ \frac{m_i}{M} X_{ij} X_{ij'} \mid i \in I \right\} \quad (7)$$

La distance entre deux points pour une contribution égale à la dispersion totale du nuage :

$$d^2(i, i') = \sum \left\{ \left(\frac{x_{ij} - x_{i'j}}{\sigma_j} \right)^2 \mid j \in J \right\} \quad (8)$$

La distance d'un point de i au centre de gravité G

$$d^2(i, G) = \rho^2(i) = \sum \{X_{ij}^2 \mid j \in J\} \quad (9)$$

2-2. Erreurs de prédiction régressive

On définit également l'erreur de prédiction afin d'étudier la convergence de données. Le modèle de prédiction calcule $x(n)$ à partir de ses valeurs antérieures. En inversant l'axe du temps, on parvient à obtenir l'échantillon $x(n - k)$ à partir de $\{x(n - k + 1), \dots, x(n)\}$ selon l'équation suivante :

$$x(n - k) = -\sum_{i=1}^k A_k(i)x(n - k + i) + e_k^b(n) \quad (10)$$

$$e_k^b(n) = x(n - k) + \sum_{i=1}^k A_k(i)x(n - k + i) \quad (11)$$

où $e_k^b(n)$ représente le bruit blanc à l'ordre k.

3. Résultats et Discussion

3-1. Détermination de la zone d'étude

Notre zone d'étude concerne la latitude comprise entre -22° et -16° avec la longitude comprise entre 46° Est et 49° Est.

C'est une grande portion de région ce qui nécessite une réduction de zone pour obtenir un résultat performant. Pour ce faire, on fait appel à l'analyse en composantes principales (ACP) en prenant la moyenne climatologique mensuelle de la pluviométrie de nos données. Le calcul se fait avec un maillage de $1^\circ \times 1^\circ$. Les individus sont les nœuds dénommés A1 à A7 ; B1 à B7 ; C à C7 et puis D1 à D7. La **Figure 1** représente les résultats de l'opération par l'ACP.

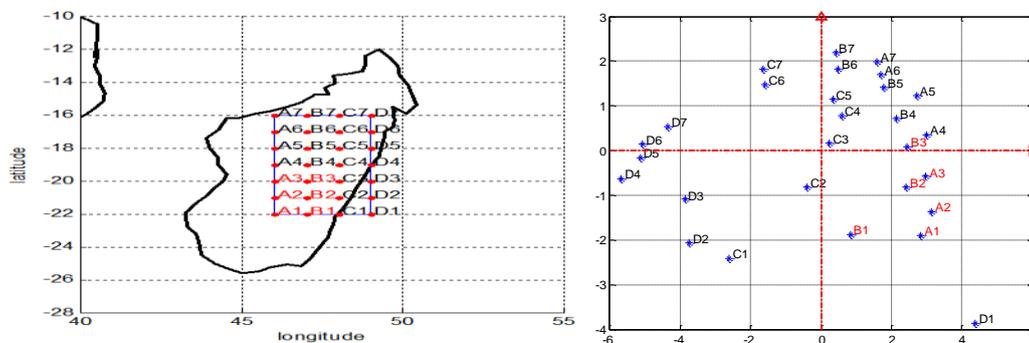


Figure 1 : Zone d'étude, $46^\circ < \text{longitude} < 49^\circ$ et $-22^\circ < \text{latitude} < -16^\circ$ obtenu par l'ACP

On peut diviser notre zone d'étude en 7 sous zones, selon l'ACP :

- ✓ Zone 1 : A1, A2, A3, B1 B2 et B3;
- ✓ Zone 2 : A4, A5 et B4;
- ✓ Zone 3 : A6, A7 et B5;
- ✓ Zone 4 : C3, C4 et C5;
- ✓ Zone 5 : B6, B7, C6 et C7;
- ✓ Zone 6 : C1, C2, D1, D2 et D3;
- ✓ Zone 7 : D4, D5, D6 et D7.

D'après l'ACP, la pluviométrie est similaire dans la *zone 1* ce que nous utilisons pour la suite de notre étude. Cette zone contient les villes ou les villages suivants : Lazarivo, Mangatabohangy, Soavina, Mandrosonoro, Amborompotsy, Itremo, Ambatomainty, Fitampito, Arimbava, Tsitondroina, Fearivo, Fanjakana, Mangidy, Soatanana, Solila, Mahabodo et Besoa.

3-2. Autocorrélation et autoinformation

On s'intéresse à la manifestation des effets mémoire. On décrit donc la structure fractale de ce phénomène par la série temporelle en explorant la corrélation entre les accroissements passés et futurs. Les résultats montrent que la série temporelle présente une mémoire à long terme. Les deux points pertinents suivants sont à soulever :

- les fonctions d'autocorrélation et d'autoinformation se comportent en asymptotique ce qui constituent l'effet de la mémoire à court terme et l'effet de la mémoire à long terme. En effet, l'autocorrélation et l'autoinformation décroît progressivement quand le retard augmente ;
- la décroissance en exponentielle est rapide pour la mémoire à court terme tandis qu'elle est plus lente en hyperbolique pour la mémoire long à terme.

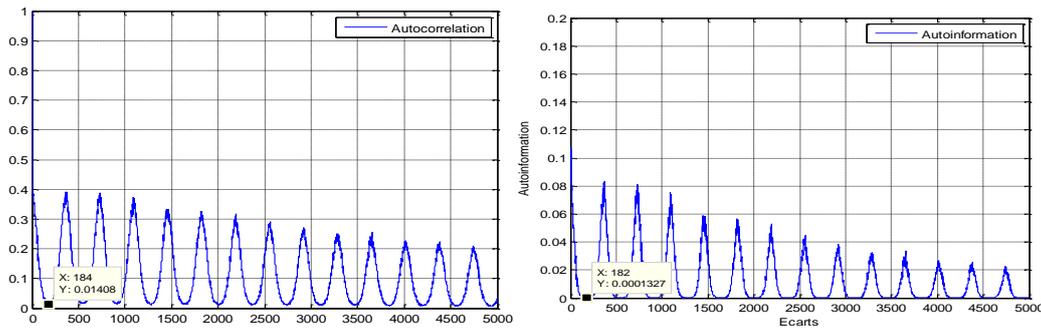


Figure 2: Fonctions autocorrélation et autoinformation

La **Figure 2** représente l'évolution de l'autocorrélation et de l'autoinformation de la chronique des précipitations en représentation linéaire et logarithmique. Les premiers minimums des deux fonctions s'établissent autour de *182 jours* pour la fonction d'autoinformation et de *184 jours* pour la fonction d'autocorrélation.

Le coefficient de corrélation normalisé $\rho(k)$ est donné par l'expression suivante :

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} \tag{12}$$

où $\gamma(0)$ représente les écart-types correspondants à x_t et x_{t+k} et $\gamma(k)$ représente le coefficient de corrélation tels que :

$$\gamma(k) = \frac{1}{T} \sum_{t=1}^T (x_t - \bar{x}_t)(x_{t+k} - \bar{x}_t) \tag{13}$$

$$\gamma(0) = \frac{1}{T} \sum_{t=1}^T (x_t - \bar{x}_t)(x_t - \bar{x}_t) \tag{14}$$

avec :

- ✓ k représente le décalage en jours ;
- ✓ x_t représente la chronique de pluviométrie ;
- ✓ x_{t+k} représente la chronique de débit.

On définit l'erreur sur l'estimation de $\gamma(k)$ afin de déterminer le seuil de corrélation. Ce seuil spécifie l'indépendantes entre les deux séries des données. L'intervalle de confiance calculé sur $\gamma(k)$ permet de connaître le seuil que l'on peut distinguer en deux cas :

- le nombre de mesure grand : $T > 300$;
- le nombre de mesure réduit : $T \leq 300$.

3-3. Cumul des pluies mensuelles

On détermine le point qui présente une anomalie sur les données. Les hauteurs de la pluie cumulée sont représentées en ordonnée pour chaque instant t .

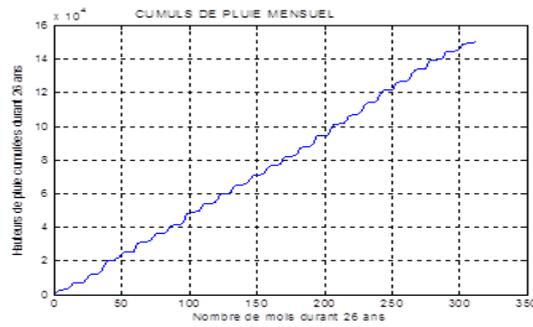


Figure 3 : Hauteurs de pluie cumulées mensuels des précipitations du 1989 au 2014

La variation du cumul des précipitations reste stable et ne présente pas de brusque changement au cours de ces 25 années. Cette variation nous indique que le changement climatique concernant de la pluie n'existe pas dans notre zone d'étude.

3-4. Nombre de jours avec et sans pluie

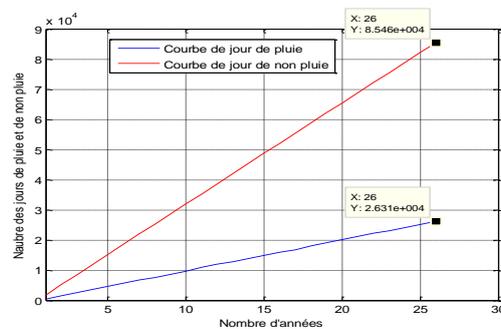


Figure 4 : Courbe de jour avec et sans pluie

Le nombre de jours de la pluie et non pluie restent constant durant la période de 25 années ce qui confirme les résultats des cumuls mensuels discutés dans le paragraphe précédent qu'il n'existe pas de une forte variation de pluie dans notre zone d'étude.

3-5. Moyenne journalière du temps d'attente

La Figure 5 représente les relevés de jours sans pluie consécutive pendant 25 ans.

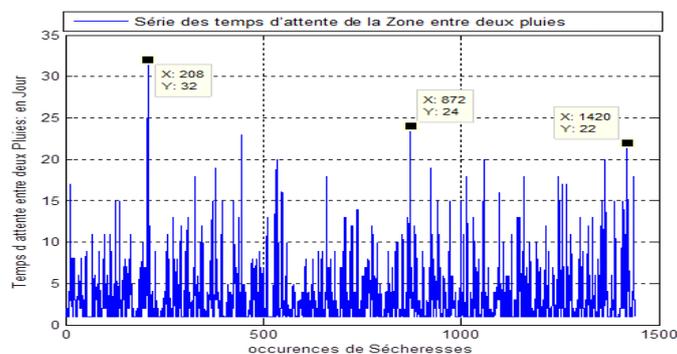


Figure 5 : Temps d'attente entre deux pluies successives

On remarque que la densité de points est la plus basse à partir de 208^{ème} jour. L'intervalle de temps sans pluie est de 32 jours. Elle est cependant de 24 jours à partir 872 jours. Il est impossible de conclure sur les occurrences de sécheresses en se basant sur les temps d'attente.

3-6. Analyse par la MEM

Le filtrage des données par la Méthode d'Entropie Maximale (MEM) consiste à décomposer le signal en une somme de fonction sinusoïdale [6, 7]. Il élimine la fréquence non significative. Dans la **Figure 6**, on observe qu'il y a des pics nets et deux quasi-périodicités. Les deux crêtes se trouvent sur les périodes : $6,63 < \tau_1 < 7,587$ jours et $2,757 < \tau_2 < 5,076$ jours. Les pseudo-périodes du temps d'attente sont donc : 0,95 jour (soit $7,587 - 6,63$) et 2,22 jours (soit $5,076 - 2,757$).

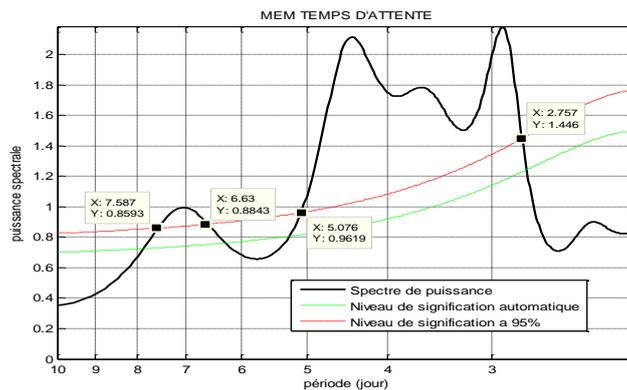


Figure 6 : Temps d'attente par la MEM

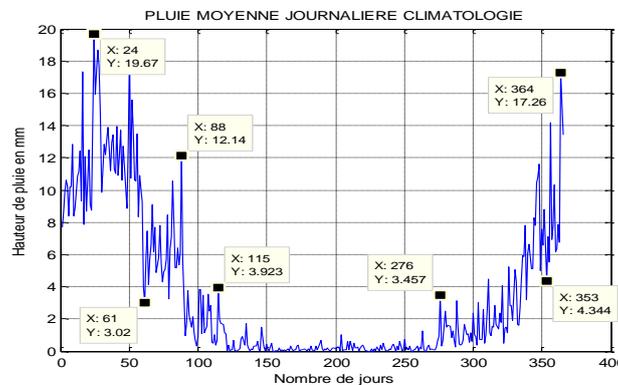


Figure 7 : Moyenne de pluie journalière climatologique

Dans le calcul des hauteurs de pluie que nous réalisons, les hauteurs de pluie qui sont comprises entre 1,68 mm et 35,96 mm durant la saison de pluie de l'année 1995 (de mois de Mars à Mai) représentent l'année où il y a une forte pluie. Tandis que les hauteurs de pluie qui sont comprises entre 1,68 mm et 35,96 mm durant la saison de pluie de l'année 2002 représente l'année où il y a moins de pluie. Dans la **Figure 7**, on constate qu'entre 1^{er} et 88^{ème} jour, la quantité de pluie journalière varie de 3,02 à 19,67 mm. Tandis que du 89^{ème} au 275^{ème} jour, le taux moyen de pluie est environ compris de 0 à 3,923 mm. Enfin du 276^{ème} au 365^{ème} jour, on a une variation entre 0,356 et 17,26 mm. En conclusion, la saison sèche est en moyenne entre 89^{ème} jours et 275^{ème} jours durant 25 ans. Les valeurs moyennes mensuelles climatologiques de la pluie de nos données sont représentées dans le **Tableau 1**. Elles permettent de suivre l'évolution du régime pluviométrique et du régime de sécheresse dans notre zone d'étude.

Tableau 1 : Moyenne mensuelle climatologique de la pluie, exprimée en mm

Mois	Jan	Fév	Mar	Avr	Mai	Juin	Juil	Août	Sep	Oct	Nov	Déc
Pluie (mm)	11,670	12,170	6,080	1,789	0,372	0,108	0,204	0,176	0,209	1,225	2,148	7,776

On observe une augmentation des quantités de pluie du mois de Décembre jusqu'au mois de Mars. *Ce sont donc les mois pluvieux.* Du mois d'Avril au mois de Novembre, on rencontre une diminution des quantités des pluies. *Ce sont des mois secs.* En conclusion, les mois les plus secs sont ainsi : Mai, Juin, Juillet, Août et Septembre.

4. Conclusion

La méthode par l'analyse en composantes principales a été utilisée pour regrouper les données de la pluie journalière que nous avons à la disposition pour mener cette étude. Les informations discriminées sont par la suite utilisées pour l'étude de l'autocorrélation et de l'autoinformation. L'analyse quantitative de la pluie journalière de 1989 à 2014 dans la région haute terre de Madagascar qui se situe entre la longitude 46° Est et 49° Est, la latitude -22° et -16° nous permet de prédire la période sèche dans cette zone. Par conséquent, on rencontre la saison pluvieuse aux mois de : Décembre, Janvier, Février et Mars ce qui reflète de la réalité. On retrouve la saison sèche aux mois de Juillet et Août et Septembre. L'incertitude concerne cependant les mois d'Avril, Mai et Juin. Les études que nous avons réalisées ne parviennent pas sortir une idée précise quant à la prévisibilité des paramètres pour ces trois mois. Une nouvelle étude s'impose en introduisant une ou d'autres méthodes d'analyse pour mieux comprendre la variation de la pluie journalière pendant cette période de l'année. L'autocorrélation et l'autoinformation de quantité de pluie ont montré qu'il y ait un événement de pluie espacé de 78 jours. Le temps d'attente de jour de pluie est de 8 jours en moyenne. La perspective de travail consiste à étendre la zone d'études sur la partie sud de Madagascar où la sécheresse est très fréquente depuis ces dix dernières années.

Références

- [1] - IRD, Climatologie, <http://horizon.documentation.ird.fr>, (Février 2016).
- [2] - GILLES MOLINIE, Mesures des Précipitations, Université de Grenoble, Rapport, (2015).
- [3] - DUPONT and al., Verification of ensemble-based uncertainty circles around tropical cyclone track forecasts, *Wea. Forecasting*, 26 (5) (2011).
- [4] - RICCO RAKOTOMALALA, Analyse de corrélation et étude des dépendances des variables quantitatives, <http://eric.univ-lyon2.fr/~ricco>, (Janvier 2016).
- [5] - I. JOLLIFFE, Principal Component Analysis, 2nd edition, Ed. Springer-Verlag, (2002).
- [6] - B. PICINBONO et BARRET, Nouvelle présentation de la méthode du maximum d'entropie, *Traitement du Signal*, Vol.7, N°2 (1990) 153 - 158.
- [7] - FRANCIS COTTET, Traitement du signal, Ed. DUNOD, (2005).