

Changement De Distribution Statistique Des Cumuls Annuels Des Précipitations Convectives De La Basse Vallée De La Betsiboka, Madagascar

[Change In Statistical Distribution Of Annual Accumulations Of Convective Precipitation In The Lower Betsiboka Valley]

RAZAFINDRALAMBO Hasina Tahiananirina¹, RAZAFIMANDIMBY Honoré², RABEHARISOA Jean Marc³, RATIARISON Adolphe⁴

^{1,2,3} Université d'Antananarivo, Faculté des Sciences, Laboratoire de la Dynamique de l'Atmosphère, du Climat et des Océans (DyACO), B.P. 906, Antananarivo, Madagascar,

hasinatahiananirina@gmail.com¹



Résumé – Dans cet article, nous examinons les effets du réchauffement global de l'atmosphère sur les cumuls annuels des précipitations convectives de la basse vallée de la Betsiboka. Nous explorons particulièrement le changement de distribution statistique, c'est une facette moins discutée mais tout aussi cruciale du dérèglement climatique. Les analyses statistiques ont montré sans équivoques des ruptures significatives de distributions statistiques. Les années de ruptures sont situées entre les années 1985 et 1987. En comprenant ces nuances du dérèglement climatique, nous pourrions mieux appréhender les défis à venir et élaborer des stratégies pour atténuer les effets négatifs et exploiter les opportunités qui se présentent à nous.

Mots-clés – Précipitation convective, changement climatique, rupture en distribution statistique, basse Betsiboka

Abstract – In this article, we examine how the global warming of the atmosphere affects the annual convective precipitation accumulations in the lower Betsiboka valley. We specifically delve into a lesser-discussed yet equally crucial aspect of climate change, the change in statistical distribution. The statistical analyses unequivocally revealed significant breaks in statistical distributions, occurring between 1985 and 1987. This led us to adjust the empirical distributions both before and after the observed break using Gaussian models. By grasping these nuances of climate change, we can gain a deeper understanding of the challenges that lie ahead and develop strategies to mitigate the negative effects while capitalizing on the opportunities that come our way.

Keywords – Convective precipitation, climate change, break in statistical distribution, lower-Betsiboka.

I. INTRODUCTION

A l'échelle mondiale, le dérèglement climatique est une réalité indéniable [1]. Le réchauffement global de l'atmosphère, se traduisant par des épisodes caniculaires de plus en plus fréquents ou des fontes accélérée des étendues glacées, sont manifestes [2]. Cependant, certains aspects du dérèglement climatique sont moins évidents. Des variables climatiques peuvent subir des altérations moins visibles, mais tout aussi importantes pour comprendre les nuances du dérèglement climatique.

Dans cette étude, nous nous intéresserons aux cumuls annuels des précipitations convectives de la basse vallée de la Betsiboka, Madagascar. D'une part, cette zone abrite de vastes plaines, entre autres, la plaine d'Ambatoboeny et celle de Marovoay, qui

jouent un rôle crucial en tant que deuxième grenier à riz du pays [3]. D'autre part, la convection est l'un des processus les plus importants de la circulation atmosphérique [4] mais que le réchauffement global de l'atmosphère a modifiée de manière significative et complexe [5]. A l'instar des plaines rizicoles de la basse Betsiboka, l'agriculture pluviale dépend entièrement des précipitations pour son approvisionnement en eau [6]. Par conséquent, l'analyse des impacts du réchauffement climatique sur les précipitations convectives dans cette zone se révèle impérative.

La question qui se pose est : « dans quelle mesure le réchauffement global de l'atmosphère affecte les cumuls annuels des précipitations convectives de la basse vallée de la Betsiboka ? »

Notre objectif étant de montrer que le dérèglement climatique se manifeste par un changement significatif de distribution statistique.

Dans la section deux, nous présentons les approches statistiques qui nous ont permis d'analyser les données climatiques ainsi que les matériels qui nous ont aidés à réaliser les calculs numériques complexes. Il s'agit des bibliothèques, des classes et des fonctions implémentés dans le langage de programmation PYTHON. Dans les sections 3 et 4, nous présentons successivement les résultats et les discussions sur ce travail.

II. MATÉRIELS ET MÉTHODES

Les données pluviométriques utilisées dans ce travail sont les hauteurs de pluies convectives journalières du 01 janvier 1959 au 31 décembre 2021 aux points de coordonnées géographiques situés au 16,25°, respectivement 16,5°, de latitude Sud et 46,50°, respectivement 46,75°, de longitude Est, correspondant respectivement à la plaine de Marovoay et à celle d'Ambatoboeny (*Figure 2*), localisées dans la partie Nord-Ouest de Madagascar. Ce sont des données satellites, issues de l'agence européenne « European Centre for Medium range Weather Forecasts (ECMWF) ».

Notre démarche méthodologique se décline en cinq étapes : (a) calcul des cumuls annuels (b) localisation des points potentiels de ruptures ; (c) identification des distributions théoriques des observations avant ruptures et des observations après ruptures ; (d) estimations des paramètres de la distribution théorique ; (e) détermination de la significativité des ruptures en distributions.

2.1 Cumuls annuels

La saison de pluie dans cette zone s'étale entre le mois d'octobre au mois d'avril. De ce fait, il convient de prendre comme début et fin des années météorologiques [7] le 1er septembre et le 31 août de l'année suivante. Ainsi, le cumul de la précipitation convective d'une année correspond à la somme des précipitations journalières convectives tombées entre cette période.

2.2 Points de ruptures en distribution

Nous utiliserons l'algorithme de Programmation Dynamique, munie d'un modèle de coût basé sur les distributions normales [8], pour localiser les points potentiels de ruptures en distribution.

La Programmation Dynamique est un algorithme qui consiste à diviser un problème d'optimisation complexe en sous-problèmes plus simples et résoudre chacun de ces sous-problèmes de manière itérative. La Programmation Dynamique muni d'un modèle de coût basé sur les lois normales est une approche algorithmique de la Programmation Dynamique dans laquelle les valeurs à optimiser sont modélisées à l'aide de lois gaussiennes.

Cet algorithme est implémenté dans la classe « *Dynp* » de la bibliothèque « *rupture* » du langage PYTHON [8].

2.3 Ajustement avec les distributions théoriques

A priori, les cumuls annuels de pluie suivent une distribution normale. Par conséquent, nous appliquerons les tests de normalité de Shapiro-Wilk [9] et de Jarque-Bera [10] pour vérifier si les observations avant et après ruptures sont normalement distribuées. Ces tests permettent d'évaluer l'hypothèse nulle H_0 selon laquelle : « *la série des données est issue d'une population normalement distribuée* ».

Le test de Shapiro-Wilk consiste à calculer une statistique de test basée sur les écarts entre les valeurs observées et celles attendues dans une distribution normale. Si les écarts sont trop importants, cela indique que les données ne suivent pas une loi normale.

Le test de Jarque-Bera, quant à lui, calcule une statistique de test en se basant sur deux coefficients : le coefficient d'asymétrie et le coefficient d'aplatissement. Cette statistique suit approximativement une distribution chi-deux. En comparant cette statistique à une valeur critique de la distribution chi-deux, le test permet de décider si les données sont suffisamment proches d'une distribution normale.

La conclusion des tests de normalité de Shapiro-Wilk ou de Jarque-Bera est exprimée en termes de *p-value*. Si la *p-value* est inférieure au seuil de 1%, alors on peut rejeter l'hypothèse nulle et conclure que les données ne suivent pas une distribution normale. Nous disons que l'échantillon n'est pas normalement distribué si l'hypothèse nulle est rejetée par, au moins, l'un des deux tests.

Nous utiliserons les classes « *shapiro* » et « *jarque_bera* » de la bibliothèque « *scipy.stats* » du langage PYTHON [11] pour implémenter ces tests.

2.4 Estimations des paramètres de la distribution théorique

Une fois la distribution théorique fixée, nous utiliserons l'estimateur du maximum de vraisemblance avec la classe « *GaussianMixture* » de la bibliothèque « *sklearn.mixture* » de PYTHON pour l'estimation des paramètres du modèle.

2.5 Significativité des ruptures en distribution

L'algorithme de Programmation Dynamique permet de localiser les points potentiels de ruptures en distribution mais n'évalue pas leur significativité statistique. Pour évaluer la significativité des ruptures en distribution, nous utiliserons le test statistique de Kolmogorov-Smirnov à deux échantillons [12]. Ce test permet d'évaluer l'hypothèse nulle H_0 selon laquelle « *il n'y a pas de différence significative entre les distributions des séries avant et après la rupture* ».

L'idée générale du test de Kolmogorov-Smirnov à deux échantillons est de calculer une statistique basée sur la différence entre les fonctions de répartition empiriques des deux échantillons. La conclusion du test de Kolmogorov-Smirnov est exprimée en termes de *p-value*. Si la *p-value* est inférieure au seuil de 1%, alors on peut rejeter l'hypothèse nulle et conclure que la rupture en distribution est significative.

Nous utiliserons la classe « *ks_2samp* » de la bibliothèque « *scipy.stats* » du langage PYTHON [11] pour implémenter ce test.

III. RÉSULTATS

3.1 Courbes des cumuls annuels des précipitations convectives

Les courbes des cumuls annuels des précipitations convectives sont présentées dans la *Figure 1*. Sur les droites de régression linéaire (lignes bleues discontinues), nous observons de légère tendance à la baisse. En outre, nous remarquons aussi que les hauteurs des pluies convectives de la plaine d'Ambatoboeny sont nettement supérieures à celles de Marovoay. La différence étant, en moyenne, de l'ordre de 200 millimètres par an.

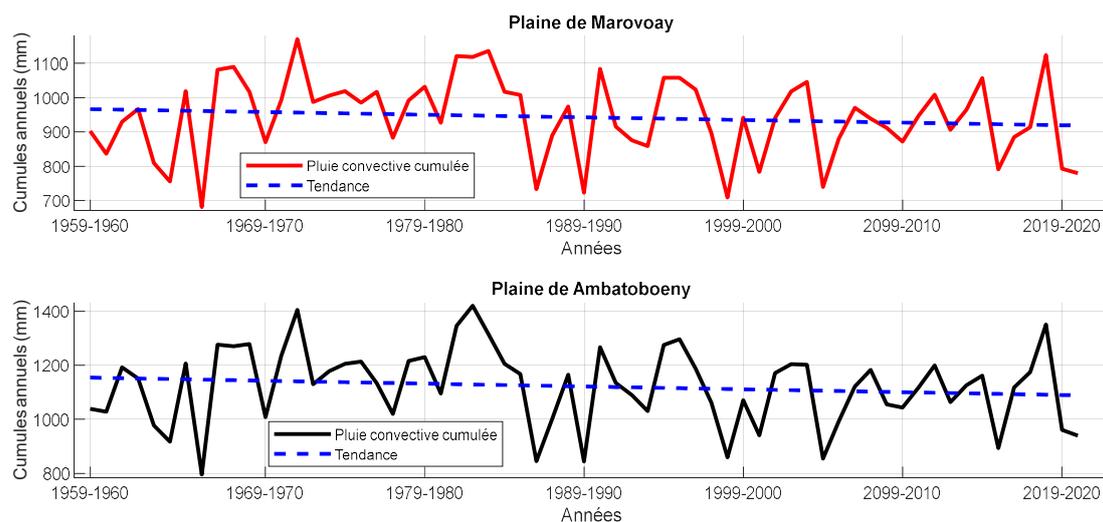


Figure 1 : Cumuls annuels des précipitations convectives

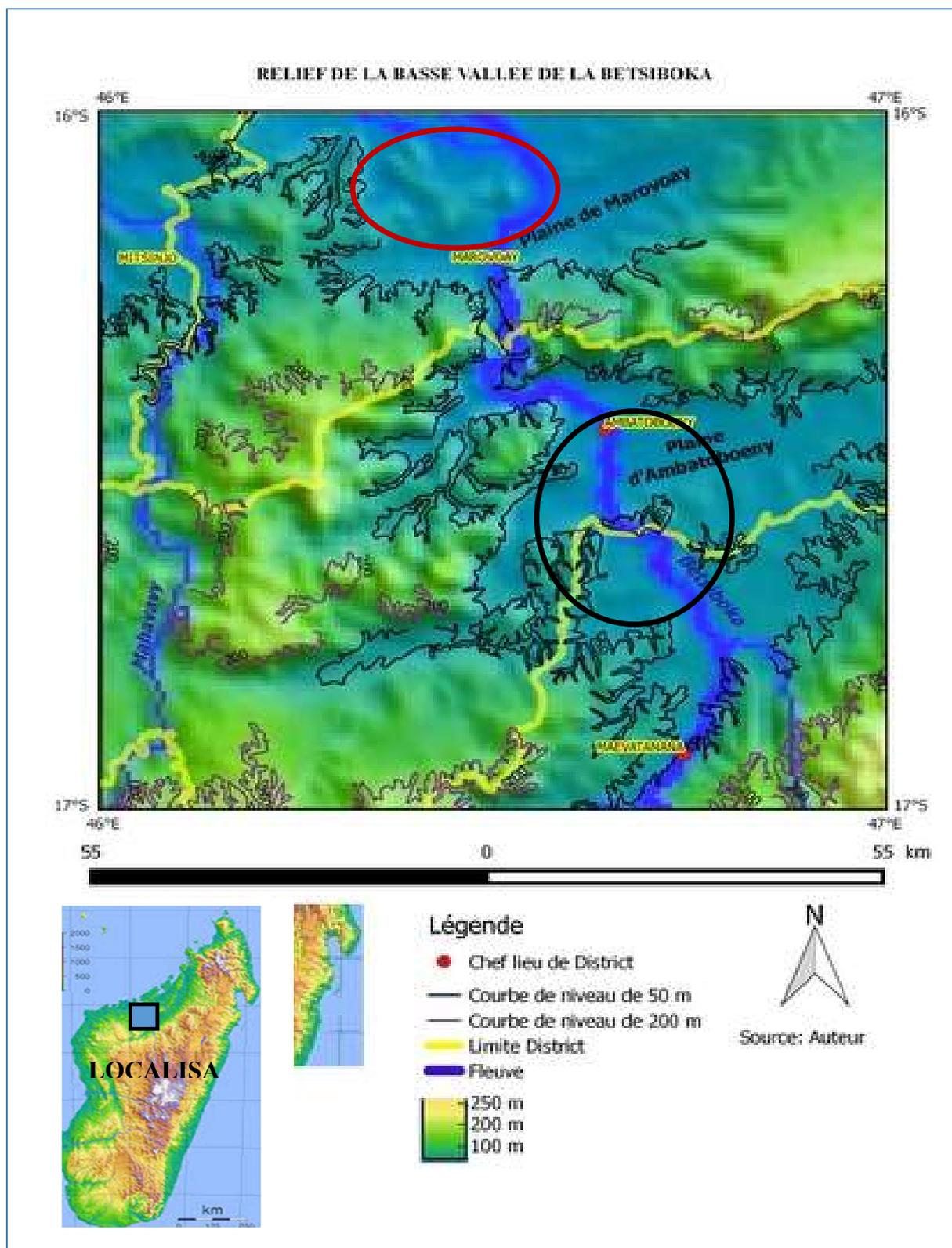


Figure 2 : Localisation de la zone d'étude.

3.2 Points potentiels de ruptures en distribution

La figure 3 présente les résultats des calculs des points potentiels de ruptures. Les lignes bleues verticales discontinues présentent les années de ruptures :

- Plaine de Marovoay : 1986-1987
- Plaine d'Ambatoboeny : 1985-1986.

Les lignes horizontales pleines en vertes indiquent la moyenne de chaque sous-série, et les lignes vertes en pointillées désignent ces mêmes moyennes plus ou moins un écart-type associé à chaque sous-période.

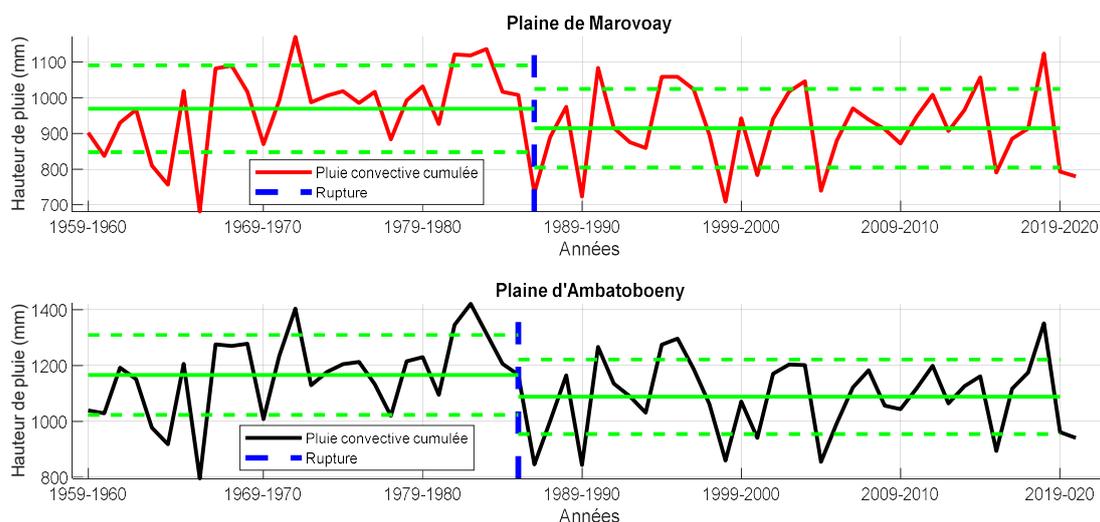


Figure 3 : Points potentiels de ruptures en distribution

3.3 Identification des distributions théoriques

Les tableaux 1 et 2 présentent les résultats des tests de Shapiro-Wilk et ceux de Jarque-Bera. Nous observons que, pour les deux tests, l'hypothèse nulle ne peut pas être rejetée au seuil de 10%. Par conséquent, nous pouvons ajuster les sous-séries avant et après ruptures avec des lois gaussiennes.

Tableau 1 : Test de Shapiro-Wilk

	Plaine de Marovoay		Plaine d'Ambatoboeny	
	Avant	Après	Avant	Après
p-value	0.3574	0.4811	0.4588	0.1142
H0	Non rejetée	Non rejetée	Non rejetée	Non rejetée

Tableau 2 : Test de Jarque-Bera

	Plaine de Marovoay		Plaine d'Ambatoboeny	
	Avant	Après	Avant	Après
p-value	0.3129	0.5380	0.5516	0.1187
H0	Non rejetée	Non rejetée	Non rejetée	Non rejetée

3.4 Estimations des paramètres des distributions théoriques

Les paramètres des lois gaussiennes ajustant les sous-séries avant et après ruptures sont donnés dans les tableaux 3 et 4. Ils montrent une nette diminution des moyennes pour les deux zones : 6,46% pour la plaine de Marovoay et 5,69% pour la plaine d'Ambatoboeny. En outre, nous pouvons aussi observer une diminution des paramètres de dispersions : plus accentuée pour la plaine d'Ambatoboeny mais dans une moindre mesure pour celle de Marovoay.

Tableau 3 : Paramètres des distributions gaussiennes, plaine de Marovoay

	Avant rupture	Après rupture
Moyenne	1023.94	957.73
Ecart-type	120.70	117.43

Tableau 4 : Paramètres des distributions gaussiennes, plaine d'Ambatoboeny

	Avant rupture	Après rupture
Moyenne	1034.31	975.45
Ecart-type	142.41	120.44

3.5 Significativité des ruptures

Le tableau 5 présente les résultats des tests de Kolmogorov-Smirnov. Nous constatons qu'au seuil de $\alpha=5\%$, l'hypothèse nulle est rejetée pour les deux plaines. C'est-à-dire, les ruptures en distributions sont significatives.

Tableau 5 : Test de rupture en distribution de Kolmogorov-Smirnov

	Plaine Marovoay	Plaine d'Ambatoboeny
Année de rupture	1986-1987	1985-1986
p-value	0.0174	0.0090
Rupture	Significative	Significative

La figure 4 présente les distributions gaussiennes des sous-séries avant et après ruptures. Nous remarquons que les histogrammes de fréquences avant ruptures se présentent sous forme de cloche avec un unique sommet alors que ceux d'après rupture a tendance à exhiber des sommets aux extrémités. Ce qui confirme la significativité des ruptures en distribution.

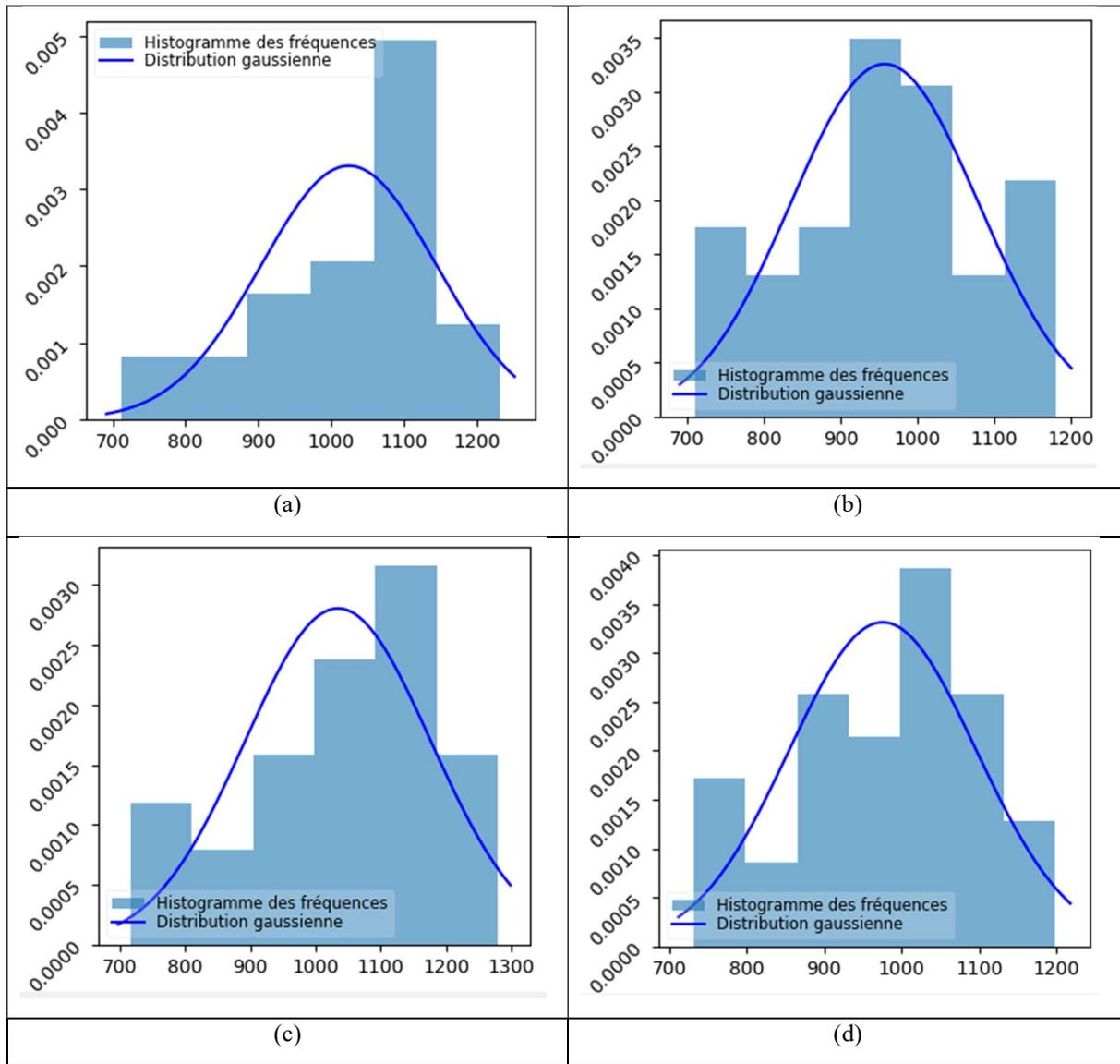


Figure 4 : Histogrammes de fréquences et distribution théorique : (a) Avant rupture, Marovoay, (b) Après rupture, Marovoay (c) Avant rupture, Ambatoboeny (d) Après rupture, Ambatoboeny

IV. DISCUSSION

L’approche couramment appliquée pour évaluer l’impact du changement climatique sur les précipitations consiste à la détection d’une tendance significative avec le test de Mann-Kendall [13] ou d’une rupture significative au niveau de la moyenne avec le test de Pettitt [14].

Appliqués à nos séries de données, le test de Pettitt et le test de Mann-Kendall indiquent que ni les tendances, ni les ruptures en moyenne ne sont significatives.

En somme, l'analyse des cumuls annuels de précipitations convectives dans la basse vallée de la Betsiboka révèle l'absence de tendances significatives et de ruptures notables au niveau de la moyenne. Cependant, une exploration plus approfondie des données dévoile des changements significatifs en distributions statistiques. En se limitant à l'application des tests classiques de détection de tendance et de rupture en moyenne pour évaluer les impacts du changement climatique, on pourrait hâtivement conclure que les variations observées dans cette variable climatique ne sont pas significatives.

V. CONCLUSION

Cette étude met en évidence que les méthodes conventionnelles pour évaluer le changement climatique, axées principalement sur les tendances et les moyennes, peuvent masquer des changements importants qui surviennent au niveau des distributions statistiques. Ces changements dans la distribution des précipitations convectives annuelles suggèrent des altérations profondes dans la répartition, l'intensité ou la fréquence des événements de précipitations convectives, même en l'absence de tendances ou de rupture au niveau de la moyenne.

Il est donc essentiel d'élargir notre compréhension du changement climatique en explorant des aspects subtils tels que les variations de distribution. En intégrant ces informations, nous pouvons nuancer les changements climatiques en cours dans la basse vallée de la Betsiboka. Cette approche plus holistique permettra une meilleure prise de décision et une adaptation plus efficace aux évolutions climatiques locales.

REFERENCES

- [1] Sciamia, Y. (2005), "Le changement climatique : Une nouvelle ère sur la terre", Petite Encyclopédie Larousse
- [2] GIEC. (2021), Bilan 2021 des changements climatiques : Rapport de synthèse
- [3] Dubourdiou, L. (1986), "Gens de riz dans la basse vallée de la Betsiboka : société sans espace ou espace sans cadre ? ", Serie SHS, n° 2, pp. 143-194
- [4] Raymond, David J. (2013) "Atmospheric Convection", Oxford University Press
- [5] Beniston, M. (2004), "Climatic change and its impacts", Kluwer Academic Publishers, USA
- [6] Alizany, N., Rakotondravelo, J.C., Rabarijohn, R., Raharinjanahary, H., Rabeharisoa, L., Ranaivonasy, J. et Tiani, A.M. (2010), "Madagascar : Pour une meilleure adaptation de la riziculture « atriatry » aux changements climatiques ; Le cas de Marovoay", IDRC
- [7] B. Raymond. "Une définition de l'année climatologique en montagne", Revue de géographie alpine, tome 43, n°2, 1955. pp. 449-454;
- [8] Truong, C., Vayatis, N., & Fadili, J. M. (2021). "ruptures : Change Point Detection and Segmentation for Python", <https://github.com/deepcharles/ruptures>
- [9] Shapiro, S. S., & Wilk, M. B. (1965). "An analysis of variance test for normality", *Biometrika*, 52(3/4), 591-611
- [10] Jarque, C. M., & Bera, A. K. (1987). "A test for normality of observations and regression residuals", *International Statistical Review*, 55(2), 163-172.
- [11] Jones, E., Oliphant, T., & Peterson, P. (2001). "SciPy: Open source scientific tools for Python", <https://www.scipy.org/>
- [12] Smirnov, N. (1939). "Estimate of deviation between empirical distribution functions in two independent samples", *Bulletin of the Moscow University*, 2(2), 3-16
- [13] Mann, H. B. (1945). "Nonparametric tests against trend", *Econometrica : Journal of the Econometric Society*, 245-259.
- [14] Pettitt, A. N. (1979). "A non-parametric approach to the change-point problem", *Applied Statistics*, 28(2), 126-135