

Modélisation Des Précipitations Pour Une Meilleure Conception Et Gestion Des Ouvrages De Protection Côtière A Madagascar

¹HERIPINOANARIMANANA Fenomora Evariste, ²RAMIANDRA Aina Clarc, ³RAMANAMPAMONJY Jean Claude, ⁴VELOMISAINA Dorissia, ⁵VAVIZARA Sylvie, ⁶HARY Jean, ⁷MAXWELL Djaffard

¹Ecole Doctorale du Génie de Vivant et Modélisation
Université de Mahajanga
Mahajanga-MADAGASCAR
herievariste@gmail.com

²Ecole Doctorale du Génie de Vivant et Modélisation
Université de Mahajanga
Mahajanga-MADAGASCAR
clarcaina@gmail.com

³Ecole Doctorale du Génie de Vivant et Modélisation
Université de Mahajanga
Mahajanga-MADAGASCAR
rabotajc@yahoo.fr

⁴Ecole Doctorale du Génie de Vivant et Modélisation
Université de Mahajanga
Mahajanga-MADAGASCAR
dorissia.aina@gmail.com

⁵Ecole Doctorale du Génie de Vivant et Modélisation
Université de Mahajanga
Mahajanga-MADAGASCAR
vavizarasylvie@yahoo.fr

⁶Ecole Doctorale du Génie de Vivant et Modélisation
Université de Mahajanga
Mahajanga-MADAGASCAR
jjeanhary@yahoo.com

⁷Ecole Doctorale du Génie de Vivant et Modélisation
Université de Mahajanga
Mahajanga-MADAGASCAR
djafmax@yahoo.fr

Auteur correspondant: HERIPINOANARIMANANA Fenomora Evariste. E-mail : herievariste@gmail.com



Résumé : Cet article analyse la modélisation des précipitations sur les côtes Est et Ouest de Madagascar afin d'améliorer la conception et la gestion des ouvrages de protection côtière face aux aléas climatiques. Deux modèles sont utilisés : le SARIMA, adapté pour les séries temporelles avec saisonnalité, et le LSTM, efficace pour capturer les dépendances non linéaires et la mémoire à long terme des précipitations mensuelles. Les résultats montrent une forte variabilité interannuelle et des cycles périodiques de précipitations distincts entre les deux zones, avec une tendance positive faible à l'Ouest et une tendance négative à l'Est. Le modèle SARIMA est robuste pour les prévisions à court terme mais présente une incertitude croissante à long terme. Le LSTM reproduit avec précision la saisonnalité annuelle mais sous-estime les événements extrêmes et les anomalies. Ces caractéristiques font que le SARIMA est adapté pour la planification à court terme, tandis que le LSTM convient mieux à la planification saisonnière. L'étude souligne la nécessité d'intégrer les projections climatiques futures pour adapter durablement les protections côtières. Elle appelle aussi à des recherches complémentaires pour intégrer d'autres risques comme l'élévation du niveau de la mer et l'érosion. Ainsi, la modélisation combinée permet une meilleure compréhension et gestion des risques côtiers à Madagascar. **Keywords:** Contextual Teaching and Learning, traditional games, Pancasila values, multicultural education, cooperative learning.

Mots-clés : Modélisation des précipitations, SARIMA, LSTM, protection côtière, Madagascar, variabilité interannuelle, cycles saisonniers, gestion des risques, changements climatiques.

Abstract: This paper analyzes rainfall modeling on the east and west coasts of Madagascar to improve the design and management of coastal protection structures in the face of climate hazards. Two models are used: SARIMA, suitable for time series with seasonality, and LSTM, effective in capturing nonlinear dependencies and long-term memory of monthly rainfall. The results show strong interannual variability and distinct periodic rainfall cycles between the two areas, with a weak positive trend in the west and a negative trend in the east. The SARIMA model is robust for short-term forecasts but exhibits increasing uncertainty in the long term. The LSTM accurately reproduces annual seasonality but underestimates extreme events and anomalies. These characteristics make SARIMA suitable for short-term planning, while LSTM is better suited for seasonal planning. The study highlights the need to integrate future climate projections to sustainably adapt coastal protections. It also calls for further research to integrate other risks such as sea level rise and erosion. Thus, combined modeling allows for a better understanding and management of coastal risks in Madagascar. **Keywords:** Contextual Teaching and Learning, traditional games, Pancasila values, multicultural education, cooperative learning.

Keywords: Rainfall modeling, SARIMA, LSTM, coastal protection, Madagascar, interannual variability, seasonal cycles, risk management, climate change.

I. INTRODUCTION

La protection du littoral malgache face aux aléas climatiques est un enjeu majeur. Cet article présente une étude visant à améliorer la gestion des risques côtiers en s'appuyant sur la modélisation des précipitations. Plus précisément, l'étude s'est focalisée sur les régions côtières Est et Ouest, caractérisées par des régimes pluviométriques contrastés. Pour ce faire, nous avons utilisé des modèles de prévisions avancés, tels que SARIMA et LSTM, afin de simuler les séries temporelles des précipitations et d'identifier les tendances et les événements extrêmes. Les résultats de ces modélisations ont été ensuite utilisés pour évaluer la vulnérabilité des ouvrages de protection côtière existants et pour proposer des améliorations en termes de conception et de gestion.

En effet, Madagascar, île de l'océan Indien, présente une grande diversité climatique et géographique. Sa côte Est est soumise à un climat équatorial caractérisé par l'influence des moussons, caractérisé par des précipitations annuelles supérieures à 2500 mm sans saison sèche distincte. À l'inverse, la côte Ouest présente un climat de savane tropicale avec une pluviométrie bien plus faible, environ 1500 mm, marquée par une alternance nette entre une saison des pluies et une longue saison sèche. Cette exposition directe aux flux océaniques fait également de la côte Est la première zone d'impact des cyclones tropicaux, où ils frappent avec leur énergie maximale. Par conséquent, la côte Ouest est beaucoup moins vulnérable, ne subissant que des systèmes cycloniques généralement affaiblis après leur traversée des terres.

L'objectif de cette étude est d'analyser ces spécificités et d'évaluer les risques associés aux événements extrêmes, notamment les précipitations abondantes, sur les ouvrages de protection côtière.

II. METHODOLOGIE

A. Zones d'étude

Cette étude se concentre sur deux régions côtières aux caractéristiques contrastées : la côte Ouest, représentée par les villes de Mahajanga (longitude 46,19 et latitude -15,43) et de Maroantaly (longitude 44 et latitude -18), et la côte Est, contenant les villes de Fénérive Est (longitude 49,15 et latitude -17,13) et de Toamasina (longitude 49,23 et latitude -18,08). Ces deux zones, telles que la côte Est définie comme zone 1 et la côte Ouest comme zone 2 sont soumises à des régimes pluviométriques distincts, induisant des dynamiques côtières différentes.

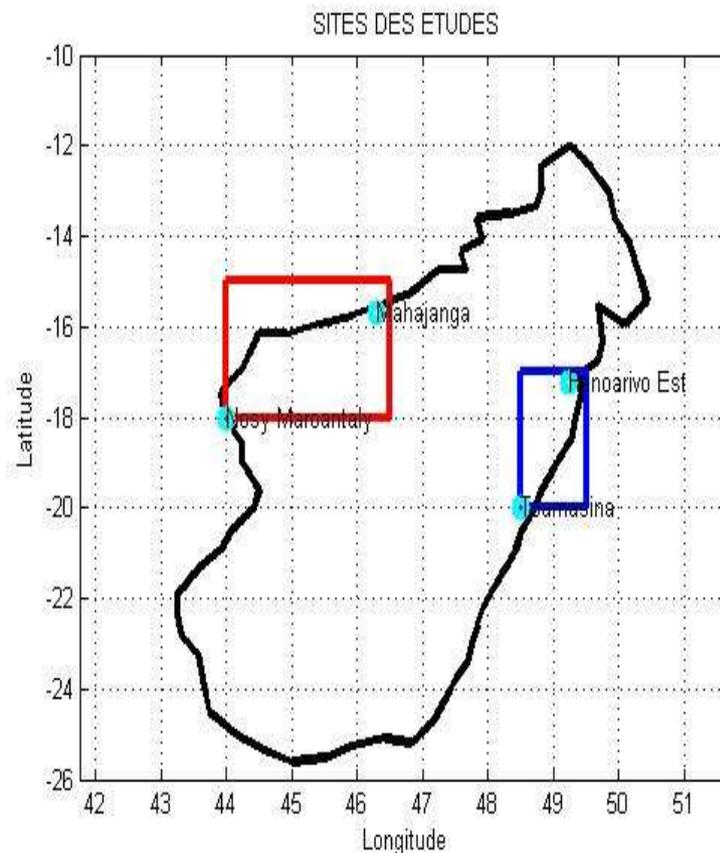


Figure 1: Délimitation de la zone d'études

B. Modèle SARIMA

Afin de mieux comprendre et anticiper les variations temporelles des précipitations, essentielles pour la gestion des risques côtiers, nous avons utilisé le modèle SARIMA. Le choix du modèle SARIMA s'explique par sa capacité à s'adapter à la variabilité naturelle des précipitations, tout en permettant de réaliser des prévisions fiables. Ce modèle statistique est particulièrement adapté aux séries temporelles présentant des tendances et des saisonnalités marquées. Nous avons appliqué ce modèle aux données de précipitations recueillies sur les côtes Est et Ouest de Madagascar dans le but d'identifier les caractéristiques spécifiques de chaque région, notamment les cycles saisonniers et les tendances à long terme.

1. Structure du modèle SARIMA

Le modèle SARIMA est caractérisé par six paramètres : $(p, d, q)(P, D, Q)s$.

- **(p, d, q):** Ces paramètres correspondent à la partie non saisonnière du modèle :
 - **p:** Ordre de l'auto-régression (AR), il indique le nombre de valeurs passées de la série qui influencent la valeur actuelle.
 - **d:** Ordre de différenciation, il correspond au nombre de fois qu'on doit différencier la série pour la rendre stationnaire.
 - **q:** Ordre de la moyenne mobile (MA), il indique le nombre de termes d'erreur passés qui influencent la valeur actuelle.
- **(P, D, Q):** Ces paramètres correspondent à la partie saisonnière du modèle :
 - **P:** Ordre de l'auto-régression saisonnière.
 - **D:** Ordre de différenciation saisonnière.
 - **Q:** Ordre de la moyenne mobile saisonnière.
 - **s:** Période de saisonnalité (par exemple, 12 pour des données mensuelles).

2. *Fonctionnement du modèle*

Le modèle SARIMA exprime la valeur d'une série temporelle à un instant t en fonction :

- Des valeurs passées de la série (composante AR).
- Des erreurs passées (composante MA).
- Des valeurs passées de la série à des intervalles saisonniers (composante AR saisonnière).
- Des erreurs passées à des intervalles saisonniers (composante MA saisonnière).

Mathématiquement, le modèle SARIMA peut s'écrire sous la forme d'une équation aux différences.

3. *Estimation des paramètres*

Les paramètres du modèle SARIMA sont généralement estimés par la méthode du maximum de vraisemblance. Cette méthode consiste à trouver les valeurs des paramètres qui maximisent la probabilité d'observer les données observées.

4. *Application du modèle*

Le modèle SARIMA peut être utilisé pour :

- **Identifier les tendances et les saisonnalités** d'une série temporelle.
- **Effectuer des prévisions** à court et à moyen terme.
- **Analyser l'impact de différents événements** sur la série temporelle.

5. *Choix des paramètres*

Le choix des paramètres du modèle SARIMA est une étape cruciale. Il est généralement réalisé de manière itérative en combinant:

- **L'analyse graphique :** En examinant l'autocorrélogramme (ACF) et la fonction d'autocorrélation partielle (PACF) de la série.
- **Des critères d'information :** Comme le critère d'information d'Akaike (AIC) ou le critère d'information bayésien (BIC).

6. *Équations du modèle SARIMA*

Le modèle SARIMA (Seasonal AutoRegressive Integrated Moving Average) est défini par l'équation suivante :

$$\Phi(B)(1 - B)^d X_t = \Theta(B)(1 - B^s)^D Z_t$$

Où :

- X_t : La valeur de la série à l'instant t.
- B : L'opérateur de retard (backshift operator), tel que $BX_t = X_{t-1}$.
- $\Phi(B)$: Un polynôme d'ordre p en B, représentant la partie auto-régressive non saisonnière.
- $\Theta(B)$: Un polynôme d'ordre q en B, représentant la partie moyenne mobile non saisonnière.
- $(1 - B)^d$: L'opérateur de différenciation d'ordre d, utilisé pour rendre la série stationnaire.
- $(1 - B^s)^D$: L'opérateur de différenciation saisonnière d'ordre D, avec s la période de saisonnalité.
- Z_t : Un bruit blanc (processus aléatoire stationnaire avec moyenne nulle et variance constante).

7. Décomposition de l'équation

En développant les polynômes $\Phi(B)$ et $\Theta(B)$, on obtient une équation plus détaillée :

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d X_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)(1 - B^s)^D Z_t$$

- $\phi_1, \phi_2, \dots, \phi_p$: Les coefficients de la partie auto-régressive non saisonnière.
- $\theta_1, \theta_2, \dots, \theta_q$: Les coefficients de la partie moyenne mobile non saisonnière.

8. Interprétation des termes

Partie auto-régressive (AR) : Les termes $\phi_1 X_{t-1}, \phi_2 X_{t-2}, \dots, \phi_p X_{t-p}$ indiquent que la valeur actuelle de la série est liée linéairement à ses p valeurs passées.

Partie moyenne mobile (MA) : Les termes $\theta_1 Z_{t-1}, \theta_2 Z_{t-2}, \dots, \theta_q Z_{t-q}$ indiquent que la valeur actuelle de la série est liée linéairement aux erreurs passées.

Différenciation (d et D) : Les opérateurs de différenciation $(1 - B)^d$ et $(1 - B^s)^D$ permettent de rendre la série stationnaire en éliminant les tendances et les saisonnalités.

Partie saisonnière: Les termes $(1 - B^s)^D$ et les coefficients saisonniers dans $\Phi(B)$ et $\Theta(B)$ capturent les variations saisonnières de la série.

C. Modèle LSTM

Face à la complexité des séries temporelles pluviométriques, caractérisées par des dépendances non linéaires et des mémoires à long terme, le choix d'un modèle LSTM (Long Short-Term Memory) s'est imposé pour la prévision des précipitations mensuelles sur les côtes Est et Ouest de Madagascar.

L'objectif principal est de développer et de valider un modèle prédictif robuste, capable de générer une prévision continue des cumuls pluviométriques mensuels sur un horizon de 60 mois pour les littoraux Est et Ouest de Madagascar.

1. Structure du modèle LSTM

Chaque cellule LSTM, à un pas de temps t , prend en entrée :

- x_t : L'entrée actuelle de la séquence
- h_{t-1} : L'état caché (ou sortie) de la cellule précédente (mémoire à court terme).
- c_{t-1} : L'état de la cellule précédente (mémoire à long terme).

Elle produit en sortie :

- h_t : Le nouvel état caché
- c_t : Le nouvel état de la cellule

2. Formules détaillées

Chaque porte est en réalité un petit réseau de neurones (une couche dense) avec une fonction d'activation sigmoïde. La sigmoïde (σ) produit une sortie entre 0 et 1, ce qui est parfait pour agir comme un "interrupteur" ou un "filtre" :

- 0 signifie "ne laisse rien passer"
- 1 signifie "laisse tout passer"

2.1 La Porte de l'Oubli (Forget Gate) - f_t

Décider quelle information de la mémoire à long terme (c_{t-1}) doit être oubliée. Elle regarde l'entrée actuelle (x_t) et l'état caché précédent pour déterminer les parties de c_{t-1} à jeter.

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f)$$

- σ : Fonction d'activation sigmoïde.
- w_f : Matrice de poids de la porte de l'oubli.
- b_f : Biais de la porte de l'oubli.
- $[h_{t-1}, x_t]$: Concaténation des vecteurs h_{t-1} et x_t .

2.2 La porte d'Entrée (Input Gate) - i_t et C_t

Elle décide une nouvelle information doit être stockée dans la cellule mémoire. Ce processus se fait en deux étapes.

Étape A : Décider quelles valeurs mettre à jour. La porte d'entrée (i_t) utilise une sigmoïde pour décider quels éléments de la cellule mémoire seront modifiés.

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i)$$

Étape B : Créer un vecteur de nouvelles valeurs candidates. Un second réseau, avec une fonction d'activation \tanh , crée un vecteur de nouvelles informations C_t qui *pourraient* être ajoutées à la mémoire.

$$C_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

- \tanh : fonction d'activation tangente hyperbolique, qui produit des valeurs entre -1 et 1.

2.3 Mise à jour de l'Etat de la Cellule (Cell State Update) - c_t

Mettre à jour la mémoire à long terme en combinant les décisions des portes précédentes.

$$c_t = f_t \otimes c_{t-1} + i_t \otimes C_t$$

Cette formule est le cœur du LSTM :

- $f_t \otimes c_{t-1}$: On multiplie l'ancienne mémoire (c_{t-1}) par le vecteur de la porte de l'oubli (f_t). Si un élément de f_t est proche de 0, l'information correspondante dans c_{t-1} est "oubliée".
- $i_t \otimes C_t$: On multiplie les nouvelles informations candidates (C_t) par le filtre de la porte d'entrée (i_t). Seules les informations jugées pertinentes sont conservées.
- $+$: On additionne les deux résultats pour obtenir le nouvel état de la cellule c_t .
- \otimes : Représente la multiplication élément par élément (produit d'Hadamard).

2.4 La porte de Sortie (Output Gate) - o_t et h_t

Décider quelle partie de l'état de la cellule (c_t) va être utilisée pour générer la sortie (le nouvel état caché h_t).

- **Étape A : Décider quelles parties de la cellule mémoire exporter.** La porte de sortie (o_t) utilise une sigmoïde pour créer un filtre.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

- **Étape B : Générer le nouvel état caché.** L'état de la cellule c_t est passé à travers une fonction \tanh (pour le normaliser entre -1 et 1), puis multiplié par le filtre de la porte de sortie o_t .

$$h_t = o_t \otimes \tanh(c_t)$$

Le h_t résultant est la sortie du pas de temps t et sert également d'entrée h_{t-1} pour le pas de temps suivant $(t-1)$.

3. Résumé des formules

- Porte de l'Oubli : $f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f)$
- Porte d'entrée : $i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i)$
- Candidats Mémoire : $C_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$
- Nouvel État de Cellule : $c_t = f_t \otimes c_{t-1} + i_t \otimes C_t$
- Porte de Sortie : $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$
- Nouvel État Caché : $h_t = o_t \otimes \tanh(c_t)$

Les matrices de poids (w_f, w_i, W_c, W_o) et les biais (b_f, b_i, b_c, b_o) sont les paramètres que le réseau apprend pendant l'entraînement par rétropropagation du gradient.

4. Choix de la fonction non linéaire

La clé du modèle LSTM réside dans le choix de la fonction non linéaire $f(\cdot)$. Les réseaux de neurones artificiels sont particulièrement adaptés à cette tâche, car ils peuvent approximer une large gamme de fonctions non linéaires. Les architectures les plus couramment utilisées sont :

- **Les Fonctions d'Activation** : À l'intérieur de chaque neurone LSTM, les données sont transformées par des fonctions comme la **tangente hyperbolique (tanh)** et la **sigmoïde**. Ces fonctions sont courbes, pas droites. C'est ce qui permet au réseau de "tordre" et de "courber" sa représentation des données pour s'adapter à des schémas complexes.
- **Les Interactions entre les Portes (Gates)** : Le mécanisme interne du LSTM (portes d'oubli, d'entrée, de sortie) implique de multiples multiplications entre les entrées et les états de la cellule. Ces multiplications entre variables créent des relations intrinsèquement non linéaires.

III. RESULTATS

A. Caractéristique de la précipitation moyenne annuelle dans la zone 1

L'évolution des précipitations moyennes annuelles dans la partie Ouest de Madagascar sur la période 1979-2023 est illustrée sur la Figure 2. Ces données révèlent une moyenne de base d'environ 234,7 mm par an, accompagnée d'une légère tendance positive de 0,174 mm/an comme l'indique l'équation de tendance ($y = 234,716 + 0,174 \cdot x$). La caractéristique la plus frappante de cette série temporelle est sa forte variabilité interannuelle, avec des écarts pouvant dépasser 100 mm entre certaines années consécutives. Les valeurs extrêmes oscillent entre un maximum d'environ 305 mm (vers 2002) et un minimum d'environ 175 mm (vers 2005), soit une

amplitude totale de près de 130 mm. On observe clairement une forte variabilité interannuelle avec des phases humides bien définies : un pic majeur en 1982, le maximum absolu en 2003, et une période pluvieuse entre 2014 et 2018. L'hypothèse d'une périodicité est intéressante ; l'intervalle entre les deux pics les plus récents (2003 et 2015) est de 12 ans, ce qui soutient votre suggestion d'un cycle possible de 10-12 ans. La transition rapide entre 2002 et 2005 est particulièrement remarquable, où l'on passe du maximum absolu au minimum absolu en seulement trois ans. La faible pente positive de la tendance linéaire est cohérente avec certains modèles de changement climatique, bien que la forte variabilité interannuelle puisse masquer cette évolution à long terme. Le cumul annuel relativement modeste suggère un climat semi-aride à subhumide sec pour cette région occidentale.

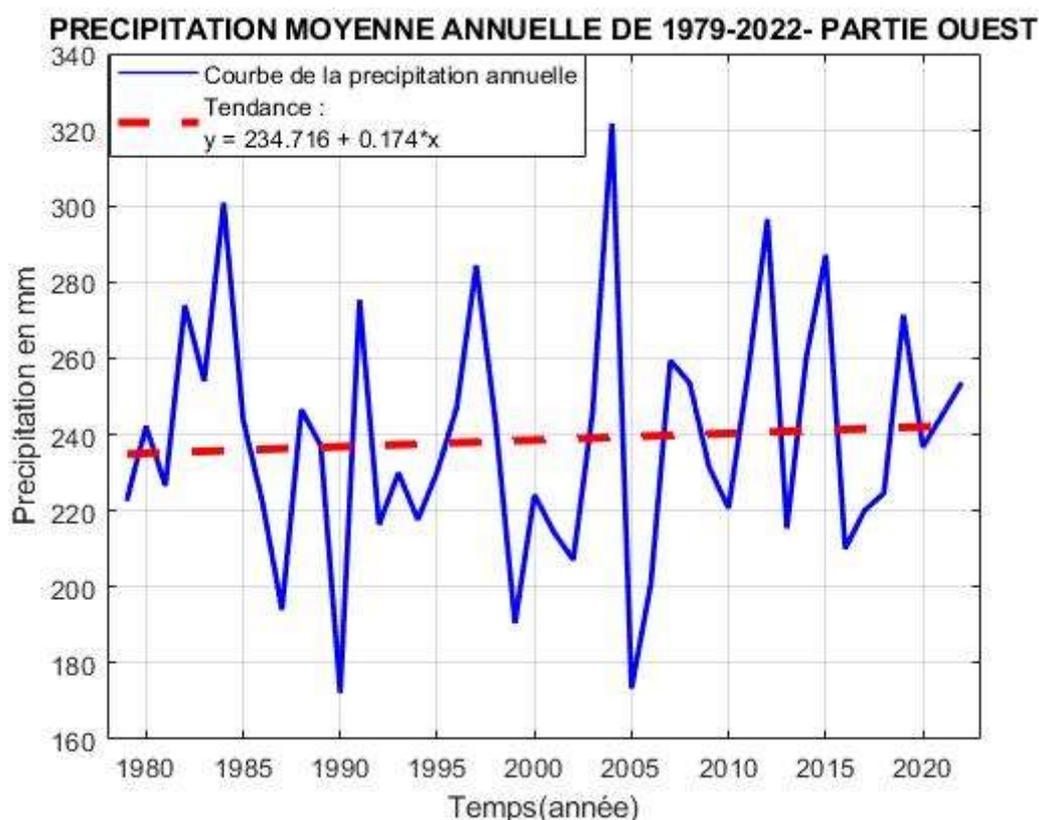


Figure 2 : Précipitations moyennes annuelles de la zone 1 de 1979-2022

B. Caractéristique de la précipitation moyenne annuelle dans la zone 2

La figure 3 illustre l'évolution interannuelle des précipitations moyennes pour la zone Est, couvrant la période d'étude 1979-2022. L'analyse de cette série chronologique révèle une moyenne pluviométrique annuelle de 358,9 mm, une valeur significativement supérieure à celle observée pour la zone Ouest précédemment étudiée avec une faible tendance linéaire négative, estimée à $-0,286 \text{ mm} \cdot \text{an}^{-1}$.

La série temporelle se distingue par une forte variabilité interannuelle, avec des fluctuations pluviométriques d'amplitude considérable, atteignant jusqu'à 150 mm entre certaines années. Les extrêmes enregistrés oscillent entre une valeur maximale avoisinant 450 mm (observée notamment autour de 1981 et 2006) et une valeur minimale proche de 260 mm (vers 1998), ce qui correspond à une amplitude totale de variation d'environ 190 mm sur la période considérée.

L'examen de la chronique suggère une succession de phases de précipitations excédentaires (par exemple, au début des décennies 1980, 1990, 2000 et au milieu de la décennie 2010) alternant avec des périodes de déficit pluviométrique relatif. Cette alternance

évoque une possible pseudo-périodicité d'environ 8 à 10 ans. Un épisode notable est la diminution marquée des précipitations survenue aux environs de 1998, suivie d'une reprise rapide au début des années 2000.

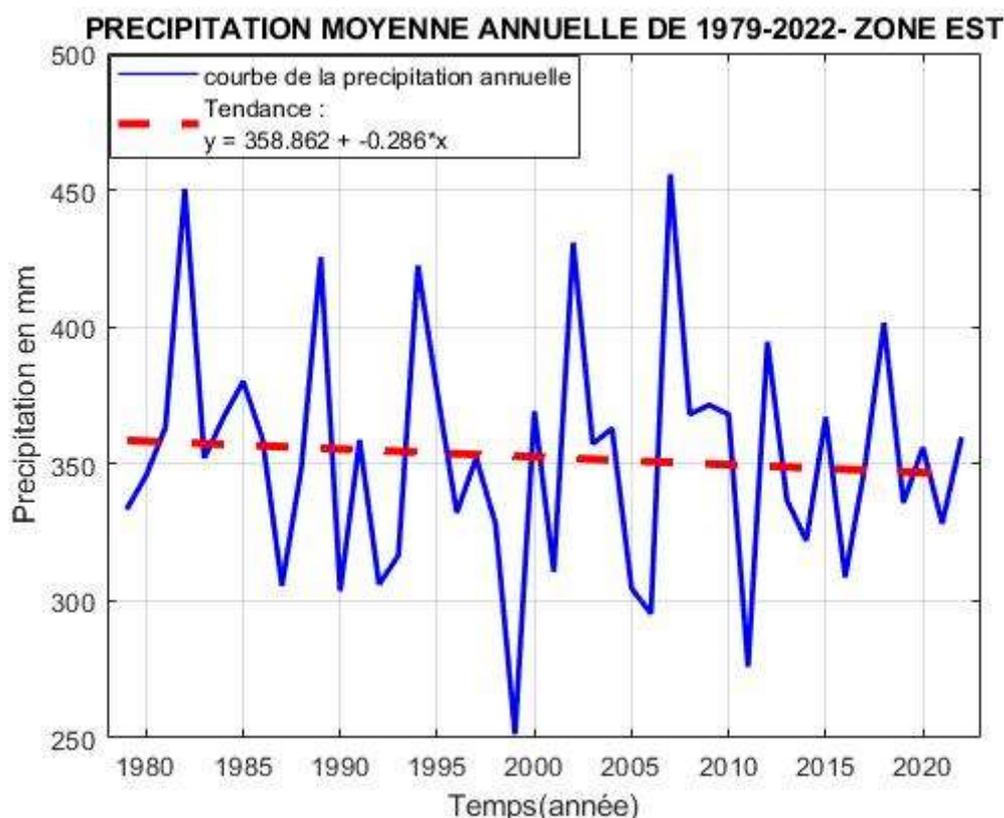


Figure 3: Précipitations moyennes annuelles de la zone 2 de 1979-2022

C. Caractéristique de la précipitation annuelle dans la zone 1 dans une projection future

1. Modèle SARIMA

Le choix du modèle SARIMA est choisi pour la prédiction annuelle des précipitations en raison de sa capacité à bien gérer les séries temporelles présentant à la fois des tendances et des saisonnalités marquées. SARIMA permet d'identifier les tendances à long terme et les cycles saisonniers, ce qui est essentiel pour analyser et prévoir les précipitations annuelles qui ont une forte variabilité interannuelle et des composantes saisonnières claires. Il est adapté pour analyser et prévoir les tendances et saisons sur une base annuelle.

La figure 4 présente une prévision des précipitations annuelles à l'aide d'un modèle SARIMA. Le modèle est entraîné sur des données historiques (bleu) puis validé sur une période de test (vert) avant de prévoir le futur (rouge), le tout encadré par un intervalle de confiance à 95% (tirets).

Faisant suite à l'analyse détaillée par courbe, avec leurs pics respectifs :

Pluie annuelle (ligne bleue) : Représente les données historiques (années 0 à 35 du 1979 au 2014). La courbe montre une forte variabilité sans tendance claire, oscillant autour d'une moyenne de 5.5 m.

Pic notable : Le pic le plus élevé de la série historique est observé autour de l'année 27 (2006), avec une valeur d'environ 5.8 m.

Test (ligne verte) : Représente les données réelles utilisées pour la validation (années 35 à 43 du 2014 au 2022).

Pic notable : Le pic principal sur cette période de test se situe au début, vers l'année 36 (2015), atteignant environ 5.65 m.

Prédiction (ligne rouge) : C'est la prévision du modèle (années 35 à 53 du 2014 au 2032). Sur la période de test, elle suit la tendance de la courbe verte, mais pour la prévision future, elle continue d'osciller.

Pic notable : Le pic le plus haut de la prévision se situe vers l'année 46 (2025), avec une valeur d'environ 5.65 m.

Borne supérieure (tirets noirs) : C'est la limite supérieure de l'intervalle de confiance à 95%. Elle montre la valeur maximale que les précipitations pourraient atteindre selon le modèle.

Pic notable : Elle atteint son point le plus haut vers l'année 48 (2027), approchant 6.2 m, illustrant l'incertitude croissante des prévisions à long terme.

Borne inférieure (tirets cyan) : C'est la limite inférieure de l'intervalle de confiance.

Pic notable : Son point le plus haut se situe vers l'année 36 (2015) à environ 5.25 m. Cette courbe tend à baisser, ce qui, combiné à la hausse de la borne supérieure, montre que l'intervalle de confiance s'élargit avec le temps.

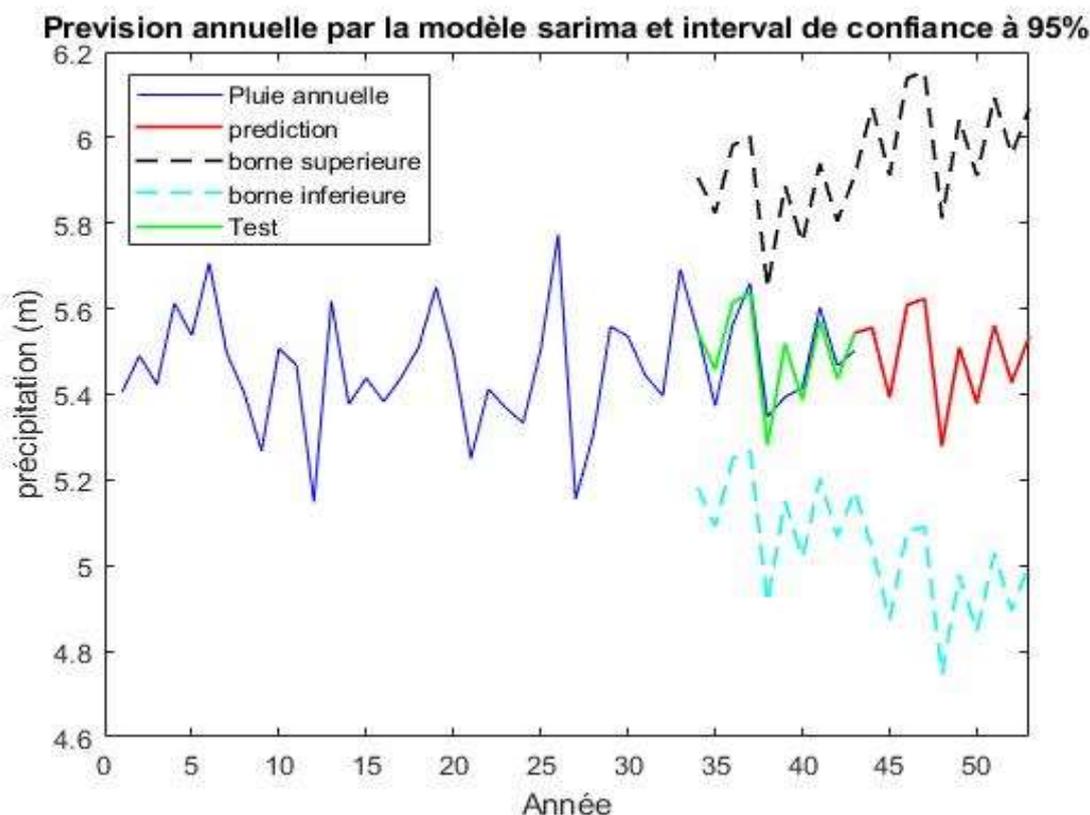


Figure 4: Prédiction de précipitation annuelle de la zone 1

2. Modèle LSTM

Le modèle LSTM est choisi pour la prédiction mensuelle des précipitations car les séries pluviométriques mensuelles présentent des dépendances non linéaires complexes et des mémoires à long terme, que le LSTM est capable de modéliser efficacement. Le LSTM excelle à capturer la saisonnalité annuelle très marquée et les cycles périodiques mensuels, offrant une prévision continue sur un horizon mensuel. Cependant, il reproduit surtout le cycle moyen et peut sous-estimer les anomalies et événements extrêmes.

Il est plus performant pour modéliser les dynamiques complexes et saisonnières sur une résolution mensuelle mais avec des limites dans la prévision d'événements extrêmes.

La figure 5 montre un modèle de prévision pour les précipitations mensuelles (en mm) sur une période de 50 ans (600 mois). Le processus est divisé en trois phases : observation, test et prévision. La caractéristique principale des données est une forte saisonnalité annuelle.

Voici l'analyse détaillée par courbe et leurs pics :

Observation (Ligne Bleue - mois 0 à 480 du 1979 au 2019) :

Rôle : Données historiques utilisées pour entraîner le modèle.

Pics : Les pics de cette courbe sont très réguliers, se produisant environ tous les 12 mois. Chaque pic représente le point culminant de la saison des pluies annuelle. Leur amplitude reste relativement constante, indiquant que le modèle apprend à partir d'un cycle saisonnier stable.

Test (Ligne Verte - mois 480 à 540 du 2019 au 2024) :

Rôle : Données réelles utilisées pour valider la performance du modèle.

Pics : Les pics de cette courbe continuent le même schéma annuel. Ils représentent les précipitations maximales réelles observées pendant la période de validation et servent de référence pour juger de la précision du modèle.

Prévision (Ligne Rouge - mois 540 à 600 du 2024 au 2029) :

Rôle : Prévision future générée par le modèle.

Pics : Les pics de cette prévision sont une extrapolation du cycle saisonnier appris. Ils représentent les futures saisons des pluies prévues par le modèle. Leur régularité et leur amplitude, similaires à celles des données historiques, montrent que le modèle a bien capturé et reproduit la saisonnalité.

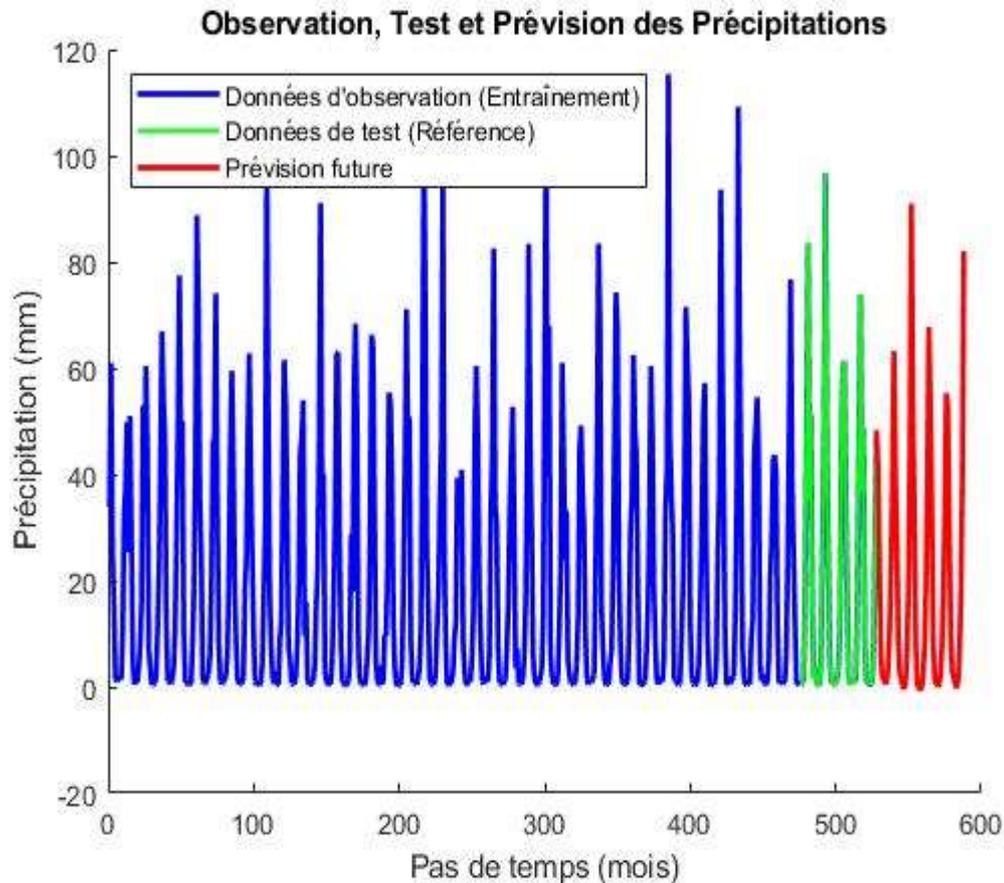


Figure 5:Prévision de précipitation mensuelle de la zone 1

D. Caractéristique de la précipitation annuelle dans la zone 2 dans une projection future

3. Modèle SARIMA

La figure 6 présente une prévision des précipitations annuelles (en mètres) à l'aide d'un modèle SARIMA. Le graphique montre les données historiques, la validation du modèle, la prévision future, ainsi que l'incertitude associée via un intervalle de confiance à 95%.

La figure 6 présente une prévision des précipitations annuelles (en mètres) à l'aide d'un modèle SARIMA. Le graphique montre les données historiques, la validation du modèle, la prévision future, ainsi que l'incertitude associée via un intervalle de confiance à 95%.

Voici l'analyse détaillée par courbe, avec leurs pics respectifs :

Pluie annuelle (ligne bleue) : Représente les données historiques d'entraînement (années 0 à 35 du 1979 au 2014). La série est très variable et ne montre pas de tendance claire.

Pic notable : Le pic le plus élevé de la série historique se situe vers l'année 4 (1983), atteignant environ 6.1 m.

Test (ligne verte) : Représente les données réelles de validation (années 35 à 43 du 2014 au 2022).

Pic notable : Le pic principal sur cette période de test est observé vers l'année 40 (2019), avec une valeur proche de 6.0 m.

Prédiction (ligne rouge) : C'est la prévision générée par le modèle (à partir de l'année 43 ou 2022). Elle continue de simuler les fluctuations observées dans le passé.

Pic notable : Le pic le plus haut de la prévision se produit vers l'année 50 (2029), atteignant environ 5.95 m.

Borne supérieure (tirets noirs) : C'est la limite supérieure de l'intervalle de confiance à 95%. Elle illustre le scénario le plus humide prévu par le modèle.

Pic notable : La courbe monte, reflétant une incertitude croissante. Elle atteint son point le plus haut vers la fin, autour de l'année 52 (2031), à près de 6.6 m.

Borne inférieure (tirets cyan) : C'est la limite inférieure de l'intervalle, représentant le scénario le plus sec.

Pic notable : Son point le plus haut se situe vers l'année 38 ou 2017 à environ 5.5 m, mais la tendance générale est que l'écart avec la borne supérieure s'agrandit.

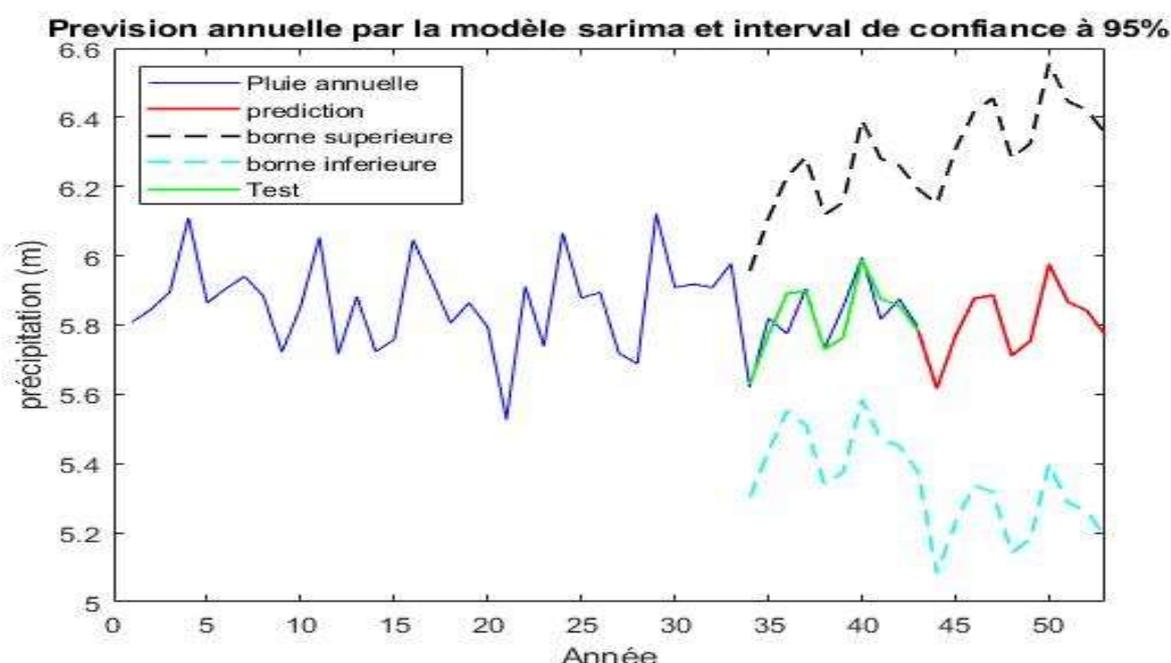


Figure 6: Prédiction de précipitation annuelle de la zone 2

4. Modèle LSTM

La figure 7 présente un modèle de prévision pour les précipitations mensuelles (en mm), divisé en trois phases : observation, test et prévision. La caractéristique visuelle la plus frappante est un cycle saisonnier très fort et régulier, avec des pics de pluie intenses et des périodes de sécheresse bien définies.

L'analyse suivante présente en détail chaque courbe et les pics associés :

Données d'observation (Bleu - Entraînement) :

Rôle : Données historiques (mois 0 à 480 du 1979 au 2019) utilisées pour entraîner le modèle.

Pics : Ces pics se produisent de manière très régulière tous les 12 mois. Ils montrent une grande variabilité, avec les pics les plus élevés de toute la série qui approchent 100 mm.

Données de test (Vert - Référence) :

Rôle : Données réelles (mois 480 à 540 du 1979 au 2024) pour valider la performance du modèle.

Pics : La courbe verte continue le cycle saisonnier. Son pic le plus haut, qui sert de référence pour la validation, atteint environ 95 mm.

Prévision future (Rouge) :

Rôle : Prévision générée par le modèle pour les mois futurs (540 à 600 du 2024 au 2029).

Pics : Le modèle reproduit bien le calendrier saisonnier, mais il a du mal à prévoir l'intensité des pics. Le pic le plus élevé de la prévision est nettement plus bas que ceux observés dans la réalité, plafonnant à environ 90 mm.

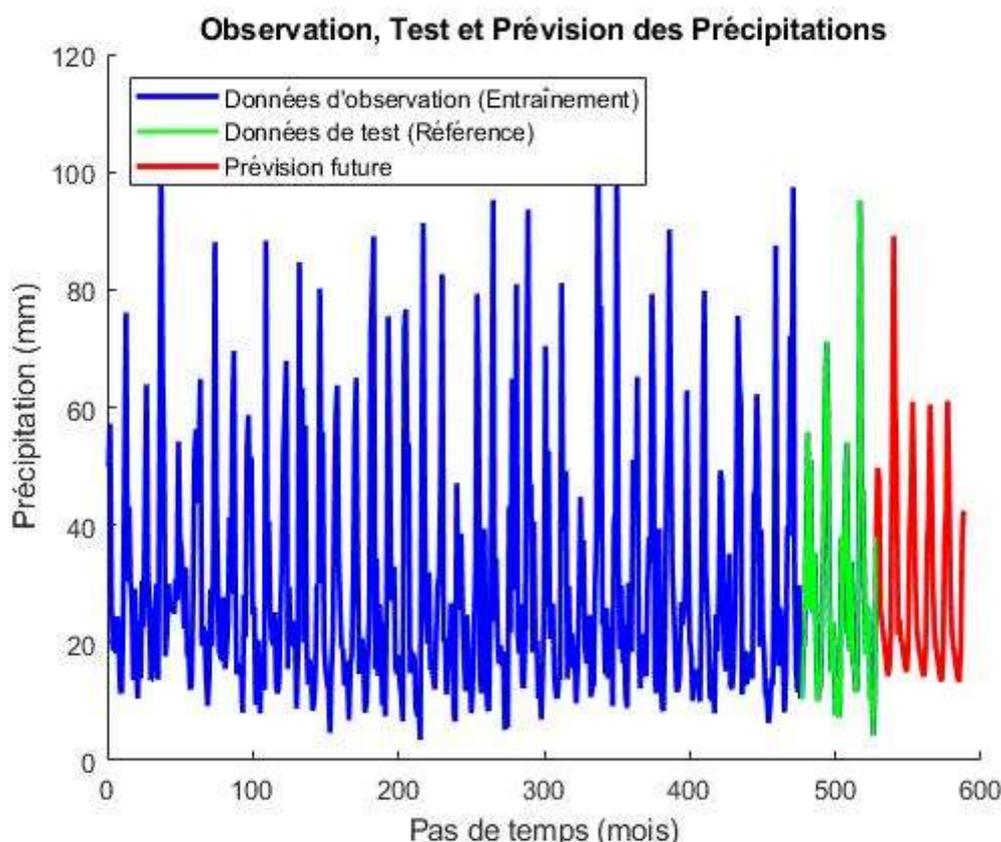


Figure 7: Prévision de précipitation mensuelle dans la zone 2

IV. DISCUSSIONS

A. Caractéristique de la précipitation moyenne annuelle dans la zone 1

L'analyse des précipitations annuelles moyennes dans la région Ouest sur la période 1979-2022 ($N = 44$ ans) révèle un régime moyen de 234,7 mm/an, caractérisant un climat semi-aride à subhumide sec. Une tendance linéaire très faible mais positive (+0,174 mm/an) est détectée, suggérant un accroissement marginal sur la période. Le trait dominant est une forte variabilité interannuelle, avec des précipitations oscillantes entre 175 mm (2005) et 305 mm (2002), soit une amplitude (130 mm) représentant plus de 55% de la moyenne annuelle. Cette variabilité semble s'organiser selon une périodicité potentielle quasi-décennale avec une alternance de phases humides et sèches, et est marquée par une instabilité notable (ex : chute rapide 2002-2005). La prédominance de cette variabilité naturelle sur la faible tendance à long terme rend l'évolution future incertaine et souligne la nécessité d'investigations approfondies sur les forçages climatiques régionaux.

B. Caractéristique de la précipitation moyenne annuelle dans la zone 2

L'analyse des précipitations annuelles moyennes dans la zone Est pour la période 1979-2022 ($N = 44$ ans) indique un régime pluviométrique significativement plus humide que celui de la zone Ouest, avec une moyenne de 358,9 mm/an. Une tendance linéaire positive de +0,286 mm/an est observée, suggérant un léger accroissement sur la période. Le régime est marqué par une forte variabilité interannuelle, les précipitations oscillant entre 260 mm (1998) et 450 mm (1981, 2006), soit une amplitude (190 mm) équivalente à 53% de la moyenne annuelle. Une périodicité apparente de 8-10 ans semble moduler cette variabilité, avec des phases humides et sèches identifiables et une instabilité marquée (ex : chute en 1998, récupération vers 2006). Le gradient pluviométrique Est-Ouest prononcé (≈ 359 vs 235 mm) confirme une hétérogénéité spatiale notable. Les tendances légèrement divergentes entre zones suggèrent des réponses régionales distinctes, nécessitant une analyse approfondie des mécanismes atmosphériques sous-jacents et du contexte de changement climatique. La tendance générale négative observée pour la zone Est, bien que statistiquement faible, pourrait indiquer une évolution progressive vers des conditions plus arides.

C. Caractéristique de la précipitation annuelle dans la zone 1 dans une projection future

1. *Modèle SARIMA*

Le modèle SARIMA parvient à suivre la dynamique générale des précipitations. Cependant, l'intervalle de confiance large et qui s'élargit au fil du temps indique que les prévisions à long terme sont associées à une incertitude considérable.

Ce modèle, calibré sur les données de précipitations de 1979 à 2014 qui oscillent autour d'une moyenne de 5,5 m avec un pic historique à 5,8 m, a démontré sa capacité à suivre la tendance générale durant la phase de validation (2014-2022), bien qu'il sous-estime les extrêmes. Pour le futur, il anticipe un pic vers 2025, mais l'élément le plus critique est l'élargissement progressif de l'intervalle de confiance à 95 %. Cette incertitude croissante rend les prévisions à long terme peu fiables, comme en témoigne la large plage de prévision pour 2027, s'étendant de 5,0 m à 6,2 m. En conclusion, bien que pertinent pour le court terme, le modèle présente des limites pour la planification à long terme, ce qui impose de considérer l'ensemble de l'intervalle de confiance et non la seule prédiction ponctuelle pour toute prise de décision.

2. *Modèle LSTM*

La figure illustre un modèle qui a réussi à identifier et à prévoir le cycle saisonnier annuel des précipitations. Les pics récurrents sur chaque courbe correspondent aux maximums de la saison des pluies, et la prévision (ligne rouge) prolonge ce schéma de manière cohérente.

Le modèle prédictif a été entraîné sur 40 ans (480 mois) de données de précipitations mensuelles, qui révèlent une saisonnalité annuelle très marquée et stable. Ce signal est dominé par un cycle périodique de 12 mois, que le modèle a parfaitement identifié. Lors de la phase de validation (2019-2024), le modèle a démontré sa capacité à généraliser ce schéma saisonnier. La prévision future pour 2024-2029 extrapole avec succès ce cycle annuel, ce qui constitue la principale force du modèle. Cependant, cette approche expose une faiblesse majeure : le modèle ne reproduit qu'un "cycle moyen" et est incapable de prévoir des anomalies comme une année exceptionnellement sèche ou humide. De plus, il ne peut intégrer aucune tendance à long terme, telle que l'impact du changement climatique. En somme, bien qu'efficace pour la planification saisonnière opérationnelle, le modèle est inadapté à la prévision d'événements extrêmes ou à l'analyse de dynamiques non-stationnaires. Sa fiabilité est donc conditionnée à la persistance des cycles passés. Le modèle est ainsi un bon outil de reproduction saisonnière, mais un mauvais prédicteur d'anomalies. Il est donc impératif de comprendre ses limites pour l'utiliser à bon escient.

D. Caractéristique de la précipitation annuelle dans la zone 2 dans une projection future

1. *Modèle SARIMA*

Le modèle SARIMA a été calibré sur les données de précipitations de 1979 à 2014, une période caractérisée par une forte variabilité et sans tendance claire. Cette série historique inclut des pics extrêmes, comme celui de 1983 atteignant environ 6,1 m. En phase de validation (2014-2022), le modèle a démontré sa capacité à reproduire la dynamique générale des données réelles, y

compris le pic significatif de 2019. La prévision future extrapole ces fluctuations, anticipant un pic notable vers 2029 à près de 5,95 m. Cependant, l'information la plus cruciale est l'élargissement progressif de l'intervalle de confiance à 95 %. Ce phénomène traduit une accumulation de l'incertitude qui dégrade la fiabilité des prévisions à long terme. Par exemple, vers 2031, le modèle prévoit une plage de valeurs très large, allant d'environ 5,0 m à un maximum de 6,6 m. En conclusion, le modèle est un outil adéquat pour les prévisions à court terme. Sa principale contribution est de quantifier les limites de la prévisibilité. Pour toute planification à long terme, il est donc impératif de considérer l'ensemble de cet intervalle d'incertitude plutôt que la seule trajectoire médiane.

SARIMA reproduit bien la variabilité des précipitations. Cependant, l'élément le plus important est l'élargissement de l'intervalle de confiance au fil du temps. Cela montre que si les prévisions à court terme sont relativement fiables, l'incertitude pour les prévisions à long terme devient très grande, ce qui est crucial pour la planification.

2. Modèle LSTM

Le modèle prédictif a été entraîné sur 40 ans de données révélant un cycle saisonnier de 12 mois très régulier, mais avec des pics historiques extrêmes approchant 100 mm. Sa principale force est de prédire avec une grande précision le calendrier de ces cycles saisonniers. Cependant, sa faiblesse critique est la sous-estimation systématique de l'intensité des précipitations. En effet, alors que les données de validation montrent des pics à 95 mm, la prévision du modèle plafonne à seulement 90 mm. Ce comportement montre que le modèle lisse la variabilité et prédit une version "moyenne" de la saison des pluies. Par conséquent, il est fiable pour une planification saisonnière générale, mais totalement inadéquat pour la prévision d'événements extrêmes. Son incapacité à capturer les pics intenses le rend dangereux pour la gestion des risques hydrologiques comme les inondations. En conclusion, le modèle est performant pour la structure temporelle mais défaillant pour l'amplitude des pics. Son utilisation doit donc être limitée à des applications non critiques. Pour la gestion des risques, des modèles spécialisés dans les valeurs extrêmes sont indispensables.

LSTM réussit à prévoir le rythme saisonnier des pluies (quand il va pleuvoir). Cependant, sa principale faiblesse est la sous-estimation des événements extrêmes. Les pics de la prévision (rouge) sont systématiquement plus bas que les pics réels les plus élevés (bleu et vert). Le modèle lisse la variabilité et prédit une version "moyenne" de la saison des pluies, le rendant peu fiable pour anticiper les précipitations les plus intenses, qui sont souvent critiques pour la gestion des risques comme les inondations.

V. CONCLUSION

Cette étude a démontré l'importance de la modélisation des précipitations pour une meilleure compréhension des risques côtiers à Madagascar. Les résultats obtenus à partir des modèles SARIMA et LSTM ont mis en évidence de fortes disparités entre les régimes pluviométriques des côtes Est et Ouest, ainsi que la fréquence et l'intensité des événements extrêmes. Ces résultats ont permis d'évaluer la vulnérabilité des ouvrages de protection côtière existants et de proposer des recommandations pour améliorer leur conception et leur gestion. Il apparaît clairement que la prise en compte des projections climatiques futures est essentielle pour adapter les stratégies de protection du littoral aux changements en cours. Des études complémentaires sont nécessaires pour affiner les modèles et intégrer d'autres facteurs de vulnérabilité, tels que l'élévation du niveau de la mer et l'érosion côtière.

REFERENCES

- [1] La Référence Fondamentale (Les Mécanismes de Base), Besancenot, J.-P. (1989). Le climat de Madagascar. Éditions de l'ORSTOM (actuel IRD).
- [2] L'Analyse Statistique et Spatiale (La Quantification), Jury, M. R. (2002). A rainfall climatology of Madagascar. *Journal of Climate*, 15(11), 1753-1766.
- [3] Les Moteurs de la Variabilité Interannuelle (Pourquoi la pluviométrie d'une année diffère-t-elle de celle d'une autre ?)-Fita, L., et al. (2010). A study of the impact of ENSO and of the MJO on the climate of Madagascar. *Theoretical and Applied Climatology*, 102(3-4), 363-382.