International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/

Vol. 58 No. 1 July 2026, pp. 544-559

Régionalisation Des Précipitations Convectives Dans La Région d'Analamanga (Madagascar) Par Approche D'apprentissage Non Supervisé

Hasinjanahary Rasamimanana[1], Adolphe Andriamanga RATIARISON[2], Chrysologue RATSIMAVO[3]

[1]Laboratoire de Dynamique de l'Atmosphère, du Climat et des Océans (DyACO), Université d'Antananarivo, Madagascar

[2]Laboratoire de Dynamique de l'Atmosphère, du Climat et des Océans (DyACO), Université d'Antananarivo, Madagascar

[3]Laboratoire de Dynamique de l'Atmosphère, du Climat et des Océans (DyACO), Université de Tulear, Madagascar

Corresponding Author : Hasinjanahary Rasamimanana, rahalahyhasina@gmail.com

Résumé – Cet article présente une étude de la régionalisation de la région d'Analamanga, située sur les Hautes Terres Centrales de Madagascar, en zones homogènes de précipitations convectives. L'objectif est de structurer la variabilité spatio-temporelle de ces pluies, essentielles au cycle hydrologique mais souvent associées à des risques d'inondations et d'érosion. En utilisant des données de réanalyse ERA5 sur la période 1979-2023, trois méthodes de clustering ont été appliquées aux séries temporelles de précipitations mensuelles: l'Analyse en Composantes Principales (ACP) couplée aux K-Means, le t-SNE couplé à K-Means, et les cartes auto-organisatrices de Kohonen (SOM) couplées à la Classification Ascendante Hiérarchique (CAH). Les résultats démontrent la supériorité de l'approche t-SNE + K-Means, qui identifie deux zones climatiques distinctes et obtient la meilleure qualité de partition selon l'indice de Dunn (0,332), contre 0,135 pour ACP + K-Means et 0,063 pour SOM + CAH. Cette structuration spatiale permet d'optimiser les futurs modèles de prévision par réseaux de neurones récurrents.

Mots clés : Précipitations convectives, régionalisation, ACP, t-SNE, K-Means, SOM, Analamanga

Abstract – This article presents a regionalization study of the Analamanga region, located in the Central Highlands of Madagascar, into homogeneous zones of convective rainfall. The objective is to characterize the spatio-temporal variability of these rainfalls, which are essential to the hydrological cycle but are often associated with flood and erosion risks. Using ERA5 reanalysis data for the period 1979–2023, three clustering methods were applied to monthly rainfall time series: Principal Component Analysis (PCA) coupled with K-Means, t-distributed Stochastic Neighbor Embedding (t-SNE) coupled with K-Means, and Kohonen Self-Organizing Maps (SOM) coupled with Hierarchical Ascending Classification (HAC). The results demonstrate the superiority of the t-SNE + K-Means approach, which identifies two distinct climatic zones and achieves the highest clustering quality according to the Dunn Index (0.332), compared with 0.135 for PCA + K-Means and 0.063 for SOM + HAC. This spatial structuring provides a valuable basis for optimizing future forecasting models based on recurrent neural networks.

Keywords: Convective rainfall, regionalization, PCA, t-SNE, K-Means, SOM, Analamanga

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 1 July 2026, pp. 544-559

1. INTRODUCTION

La région d'Analamanga, cœur économique de Madagascar, est caractérisée par une topographie complexe influençant fortement la convection atmosphérique[1], ce qui entraîne une érosion sur les hauts plateaux et inondations sur les bas plateaux[2],[3]. Les précipitations y sont majoritairement de nature convective, se manifestant par des événements brefs mais intenses. Face au changement climatique, la variabilité de ces pluies s'accentue, rendant les prévisions globales moins précises à l'échelle locale. La régionalisation devient alors une étape cruciale : elle consiste à regrouper des points de grille géographiques partageant des comportements pluviométriques similaires afin de réduire la dimensionnalité du problème et d'améliorer la robustesse des modèles de prédiction.

2. DONNÉES ET ZONE D'ÉTUDE

2.1 Zone d'étude

L'étude se concentre sur Analamanga, représentée en bleu sur la carte de Madagascar à la figure 1, située entre 17,5° et 19,5° de latitude Sud et entre 46,5° et 48° de longitude Est. Cette zone de transition entre les versants est et les plaines de l'ouest subit des influences météorologiques variées.

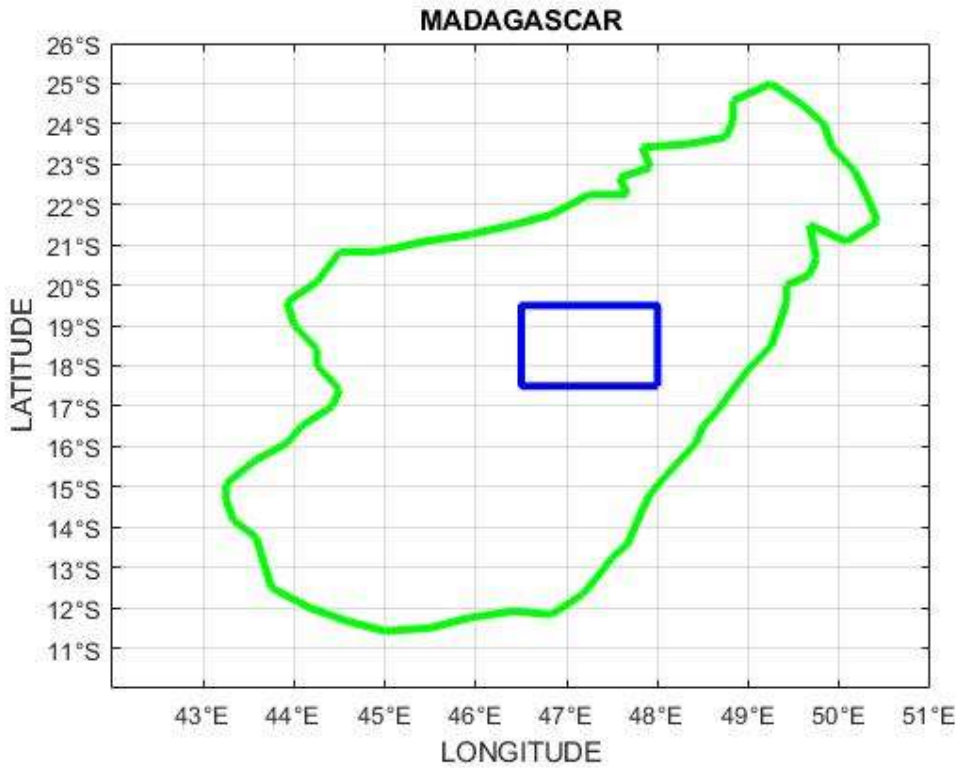

Figure 1: Représentation d'Analamanga sur la carte de Madagascar

2.2 Source des données

Nous utilisons les données de "Convective Precipitation" issues de la réanalyse ERA5 du CEPMMT (Centre Européen pour les Prévisions Météorologiques à Moyen Terme) (https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=download). Les valeurs initialement exprimées en mètres ont été converties en millimètres (mm) afin de faciliter l'interprétation climatologique. Le jeu de données comprend 63 points de grille avec une résolution spatiale de 0,25° x 0,25°, couvrant la période de janvier 1979 à décembre 2023 (soit 540 mois).

3. MÉTHODOLOGIE

3.1. Prétraitements des données

Les données téléchargées ont été traitées afin d'obtenir la matrice M correspondant à la moyenne mensuelle climatologique des précipitations.

- Lecture et fusion des données afin d'obtenir une série temporelle continue couvrant la période d'étude allant de 1979 à 2023.
- Calcul de la climatologie mensuelle en moyennant, pour chaque mois, les valeurs correspondantes sur l'ensemble des années de la période 1979–2023.
- Organisation des données sous forme matricielle en conservant la structure spatiale (latitude × longitude × mois), conduisant à une matrice M de climatologie mensuelle.

Les précipitations convectives ont été agrégées à l'échelle mensuelle en calculant, pour chaque année, la moyenne des valeurs journalières correspondant à chaque mois. Une climatologie mensuelle a ensuite été établie en moyennant ces valeurs sur l'ensemble de la période 1979–2023. La variable obtenue représente ainsi la précipitation convective moyenne journalière climatologique pour chaque mois, exprimée en millimètres par jour (mm/jour).

L'ensemble des points de grille est formé 63 points. La nomenclature des points de grille est comme l'illustre la Figure 2.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 1 July 2026, pp. 544-559

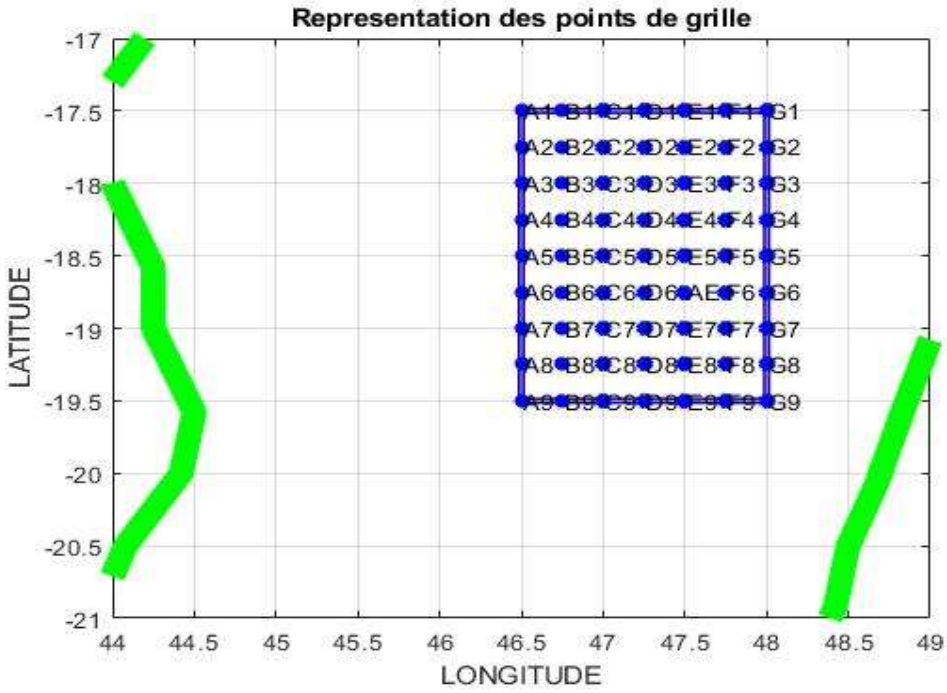

Figure 2 : Points de grille

3.2 Méthodes

L'approche repose sur la comparaison de trois protocoles de clustering :

Réduction de dimension linéaire (ACP + K-Means) : L'Analyse en Composantes Principales est utilisée pour extraire les facteurs expliquant la plus grande variance, suivie d'un partitionnement par K-Means.

Réduction de dimension non-linéaire (t-SNE + K-Means) : Le t-Distributed Stochastic Neighbor Embedding (t-SNE) est appliqué pour capturer les structures locales complexes non-linéaires avant le clustering.

Topologie Neuronale (SOM + CAH) : Les cartes de Kohonen projettent les données sur une grille 2D préservant la topologie, laquelle est ensuite segmentée par une classification hiérarchique pour déterminer le nombre optimal de clusters.

Indice de validation : L'indice de Dunn a été utilisé pour évaluer la qualité du partitionnement (séparation des classes et cohésion interne).

a) Réduction de dimension linéaire (ACP + K-Means)

L'Analyse en Composantes Principales (ACP) constitue une méthode de réduction de dimension linéaire largement utilisée pour synthétiser l'information contenue dans des données multivariées[4]. Elle consiste à projeter les données initiales dans un nouvel espace orthogonal formé par des combinaisons linéaires des variables d'origine, appelées composantes principales, maximisant la variance expliquée [4], [5].

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 1 July 2026, pp. 544-559

La projection des données sur les composantes principales peut s'écrire sous la forme : Z=XW

où X représente la matrice des données initiales, W la matrice des vecteurs propres associés aux plus grandes valeurs propres de la matrice de covariance, et Z les données projetées dans l'espace réduit.

Cette transformation permet de réduire la redondance et de filtrer le bruit, tout en conservant l'essentiel de la structure globale des données.

Dans cette étude, les premières composantes principales expliquant la majorité de la variance cumulée ont été retenues comme variables d'entrée pour le processus de classification[6]. Par la suite, l'algorithme de K-means clustering a été appliqué afin de partitionner les observations en groupes homogènes. K-Means repose sur la minimisation de l'inertie intra-classe en assignant chaque observation au centroïde le plus proche [7]. La fonction objectif minimisée par l'algorithme est donnée par:

$$J = \sum_{i=1}^{k} \sum_{x_j \in C_i} \left\| x_j - \mu_i \right\|^2$$

où k désigne le nombre de clusters, C_i le $i^{ème}$ cluster et μ_i le centroïde associé.

Cette combinaison ACP–K-Means permet ainsi d'améliorer la performance du clustering en réduisant la dimension tout en conservant les structures dominantes.

b) Réduction de dimension non-linéaire (t-SNE + K-Means)

Afin de capturer les structures non linéaires des précipitations convectives, la méthode t-SNE a été appliquée. Contrairement à l'ACP, elle préserve les voisinages locaux en optimisant les similarités probabilistes. Une attention particulière a été portée au réglage de la perplexité, paramètre crucial qui définit l'équilibre entre la sensibilité aux structures locales et globales des données. [8],[9]

Dans t-SNE, la similarité entre deux observations x_i et x_j dans l'espace original est définie par une probabilité conditionnelle donnée par : $p_{j/i} = \dfrac{\exp\left(-\dfrac{\left\|x_i - x_j\right\|^2}{2\sigma_i^2}\right)}{\sum_{k \neq i} \exp\left(-\dfrac{\left\|x_i - x_j\right\|^2}{2\sigma_i^2}\right)}$

Où σ_i est un paramètre lié à la perplexité contrôlant la taille effective du voisinage local.

Cette approche est particulièrement adaptée à l'exploration de structures cachées dans des données fortement non linéaires, comme celles issues de phénomènes atmosphériques complexes (ex : pluie convective). Après projection par t-SNE, l'algorithme de K-means clustering a été appliqué pour identifier les groupes. Cette combinaison permet de mieux capturer des motifs locaux qui pourraient être ignorés par des méthodes linéaires, bien que la sensibilité de t-SNE aux paramètres (notamment la perplexité) nécessite une attention particulière.[7]

c) *Topologie neuronale (SOM + CAH)*

Une approche alternative basée sur les réseaux de neurones non supervisés a été adoptée à travers les cartes auto-organisatrices (Self-Organizing Maps, SOM)également appelées cartes de Kohonen. Cette méthode projette les données multidimensionnelles sur une grille bidimensionnelle tout en préservant les relations topologiques, c'est-à-dire que les observations similaires sont projetées dans des régions proches de la carte.[10]

Le principe d'apprentissage du SOM repose sur l'adaptation itérative des vecteurs de poids des neurones selon la relation :

$$\omega_i(t+1) = \omega_i(t) + \alpha(t) h_{ci}(t)\left[x(t) - \omega_i(t)\right]$$

Où $\omega_i(t)$ représente le vecteur de poids du neurone i à l'itération t, $\alpha(t)$ le taux d'apprentissage, $h_{ci}(t)$ la fonction de voisinage centrée sur le neurone gagnant, et x(t) le vecteur d'entrée.

Après l'apprentissage du SOM, une Classification ascendante hiérarchique (CAH) a été appliquée sur les vecteurs prototypes (neurones) afin de regrouper les unités en classes homogènes. Cette combinaison SOM–CAH présente l'avantage de structurer les données tout en facilitant la visualisation et l'interprétation des clusters. De plus, la CAH permet de déterminer le nombre optimal de classes à travers l'analyse du dendrogramme.[11]

d) *Indices de validation des clusters*

La qualité des partitions obtenues a été évaluée à l'aide de l'indice de Dunn. Cet indice mesure le rapport entre la distance minimale inter-clusters et la distance maximale intra-cluster. Une valeur élevée de l'indice de Dunn indique des clusters bien séparés et compacts. [12]

L'indice de Dunn est défini par :

$$D = \frac{\min_{1\le i\le j\le k} d\left(C_i, C_j\right)}{\max_{1\le l\le k} \delta\left(C_l\right)}$$

Où $d\left(C_i, C_j\right)$ représente la distance entre les clusters C_i et C_j, et $\delta\left(C_l\right)$ le diamètre du cluster C_l.

Ainsi, cet indice permet d'évaluer la performance du clustering en tenant compte à la fois de la compacité interne des groupes et de leur séparation externe.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 1 July 2026, pp. 544-559

4. RÉSULTATS ET ANALYSE

a) Résultat en ACP couplé avec k-means

Visualisons le regroupement des individus par k-means sur le plan ACP :

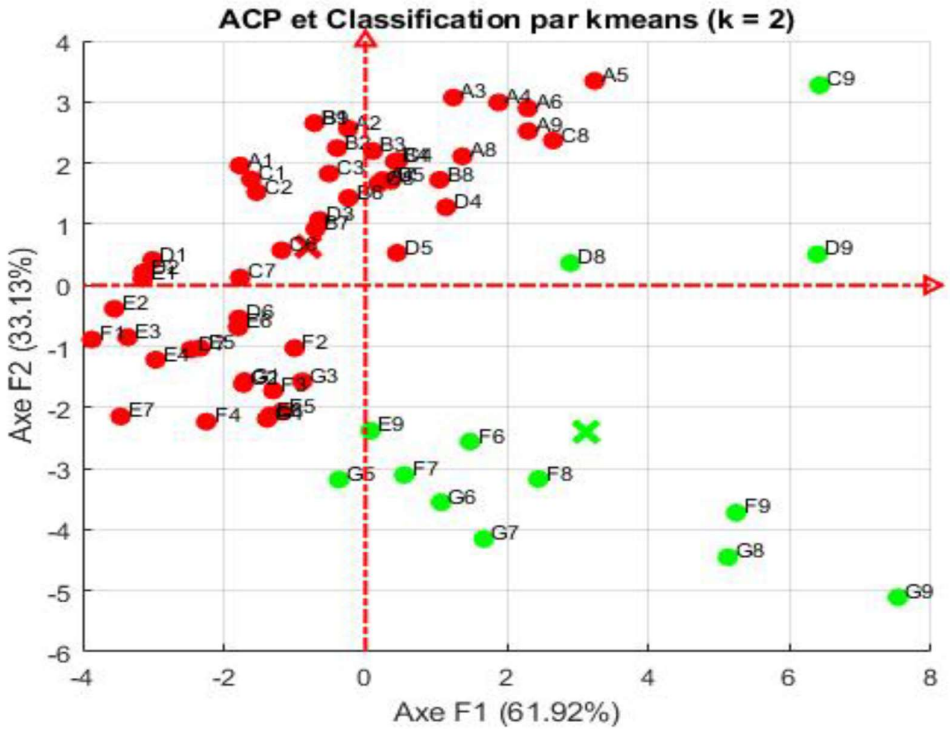

Figure 3 : Regroupement des individus par la méthode de l'ACP couplé avec K-Means

Ce résultat montre que notre étude a permis d'identifier deux zones distinctes :
- Zone 1 (rouge) : comprend les points A1, A2, A3, A4, A5, A6, A7, A8, A9, B1, B2, B3, B4, B5, B6, B7, B8, B9, C1, C2, C3, C4, C5, C6, C7, C8, D1, D2, D3, D4, D5, D6, D7, E1, E2, E3, E4, E5, E6, E7, E8, F1, F2, F3, F4, F5, G1, G2, G3, G4. Elle présente des précipitations convectives modérément élevées, dominantes entre mai et septembre.
- Zone 2 (verte) : regroupe les points situés principalement à l'extrémité sud et sud-est de la région (C9, D8, D9, E9, F6, F7, F8, F9, G5, G6, G7, G8, G9,). Cette zone se distingue par des précipitations convectives élevées, généralement concentrées entre octobre et avril.

La figure ci-dessous montre le résultat des différentes sous-zones obtenue de notre étude :

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 1 July 2026, pp. 544-559

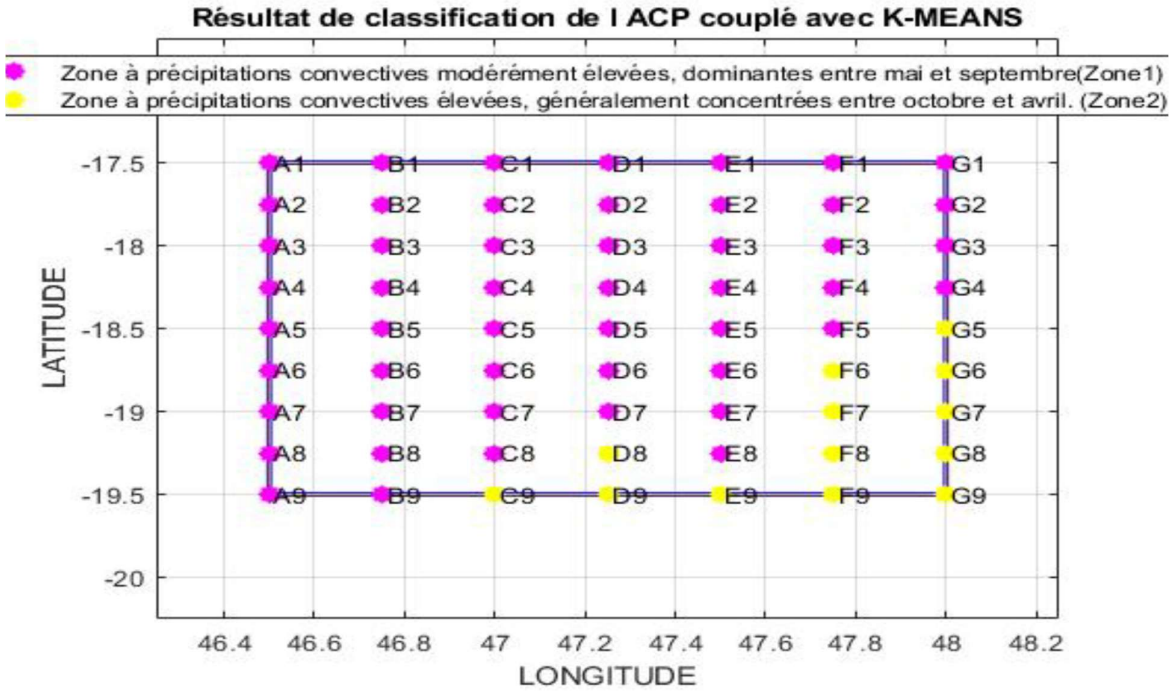

Figure 4 : Résultat de la classification par ACP couplé avec K-Means

b) ***Résultat en TSNE couplé avec k-means***

Après avoir exploré soixante visualisations correspondant à différentes valeurs de perplexité, la valeur 15 a été retenue, car elle permet d'obtenir une séparation plus nette et une structuration plus cohérente des données. Cette configuration met en évidence l'organisation des individus ainsi que les regroupements dans l'espace de représentation. La figure suivante présente la projection des individus sur le plan t-SNE avec une perplexité fixée à 15.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 1 July 2026, pp. 544-559

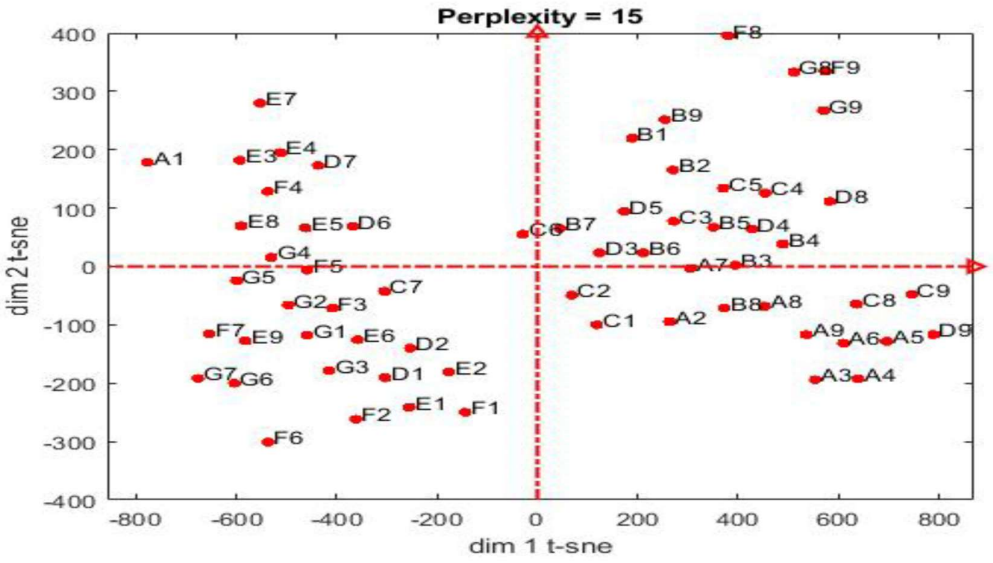

Figure 5 : Visualisation des individus sur le plan T-SNE

En faisant un regroupement de deux par k-means, nous avons le résultat :

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 1 July 2026, pp. 544-559

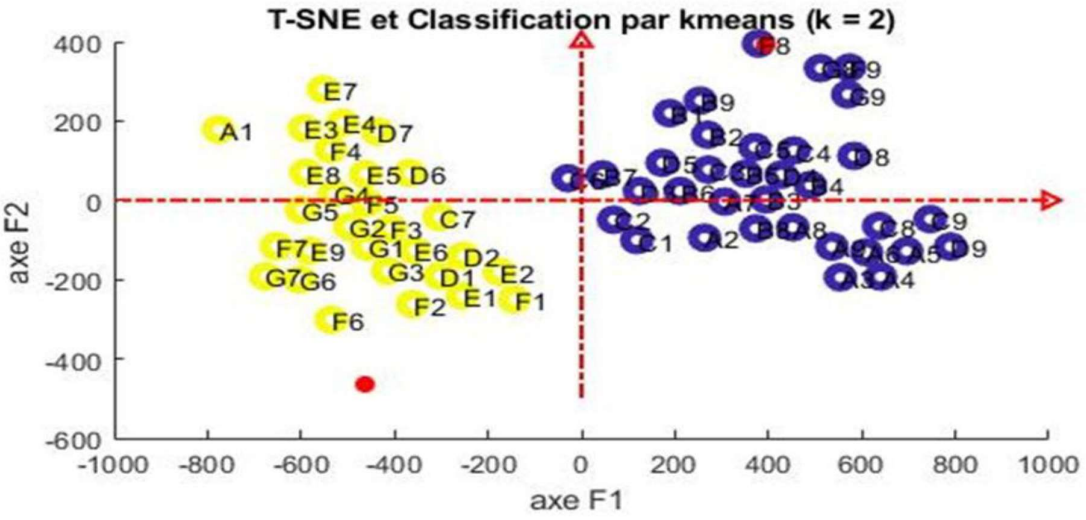

Figure 6 : Regroupement des individus par la méthode de T-SNE couplé avec K-Means

Après le regroupement, les résultats obtenus sont les suivants :

• Zone 1 (en violet) : cette zone regroupe les points A2, A3, A4, A5, A6, A7, A8, A9, B1, B2, B3, B4, B5, B6, B7, B8, B9, C1, C2, C3, C4, C5, C6, C8, C9, D3, D4, D5, D8, D9, F8, F9, G8 et G9. Elle correspond à des précipitations convectives élevées, principalement observées entre octobre et avril c'est-à-dire le cœur de la saison de pluie.
• Zone 2 (en jaune) : Cette zone comprend les points A1, C7, D1, D2, D6, D7, E1, E2, E3, E4, E5, E6, E7, E8, E9, F1, F2, F3, F4, F5, F6, F7, G1, G2, G3, G4, G5, G6 et G7. Elle correspond à des précipitations convectives modérées moins élevées, concentrées sur le mois de mai jusqu en aout.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 1 July 2026, pp. 544-559

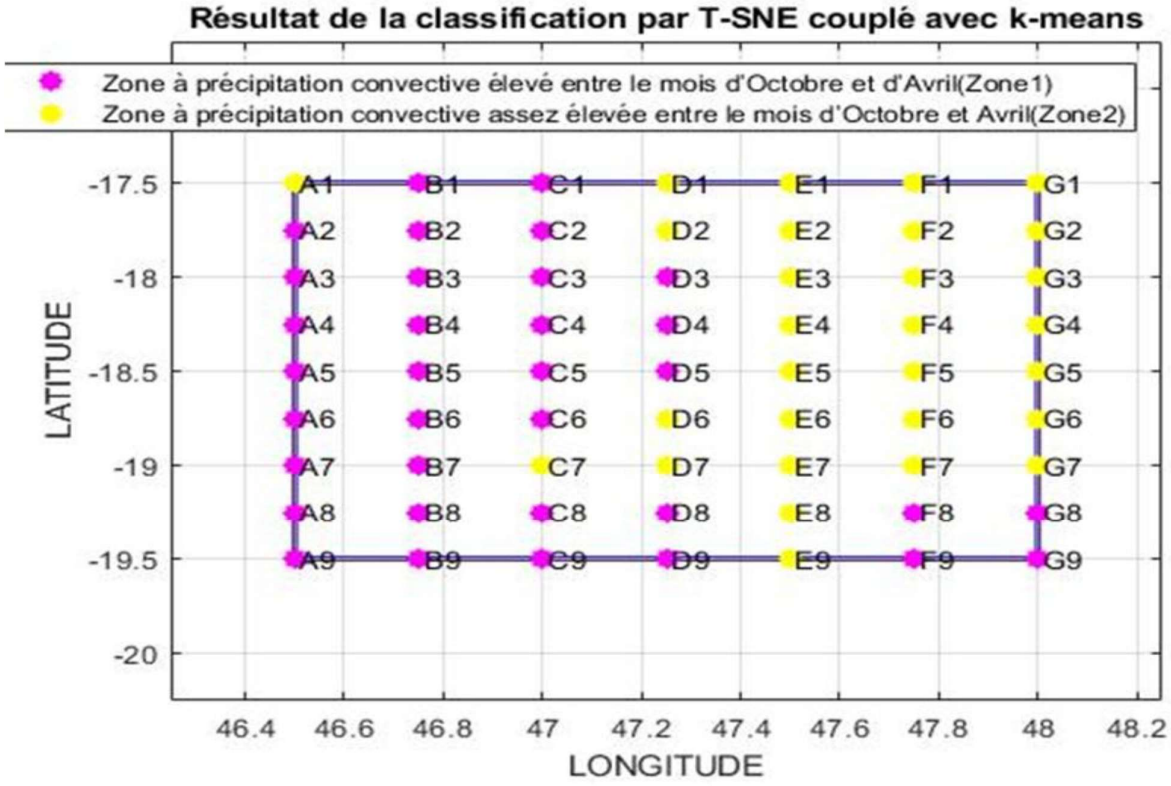

Figure 7 : Résultat de la classification par T-SNE couplé avec K-Means

c) Resultat avec le réseau de Kohonen couplé avec la Classification Hiérarchique Ascendante

Les individus ont été affectés aux différents neurones, et le résultat du classement est présenté en suivant un ordre de lecture de gauche à droite et de bas en haut :

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 1 July 2026, pp. 544-559

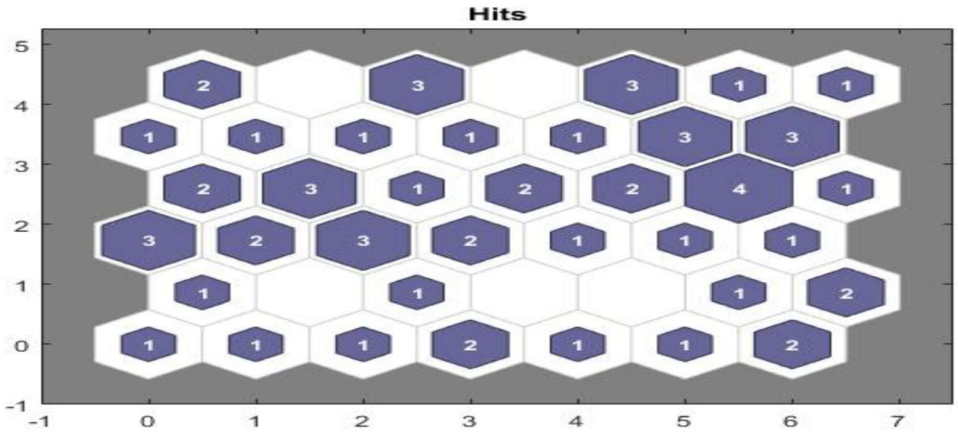

Figure 8 : Classement d'individus dans chaque neurone

Le tableau suivant montre le regroupement de chaque individu dans chaque neurone :

Indiv	Neur	Indiv	Neur	Indiv	Neur	Indiv	Neur	Indiv	Neur	Indiv	Neur	Indiv	Neur
A1	N38	B1	N36	C1	N38	D1	N40	E1	N40	F1	N42	G1	N34
A2	N29	B2	N30	C2	N38	D2	N40	E2	N41	F2	N33	G2	N34
A3	N16	B3	N23	C3	N31	D3	N32	E3	N35	F3	N34	G3	N26
A4	N15	B4	N17	C4	N23	D4	N17	E4	N35	F4	N35	G4	N27
A5	N8	B5	N17	C5	N23	D5	N18	E5	N27	F5	N27	G5	N27
A6	N15	B6	N18	C6	N25	D6	N20	E6	N26	F6	N5	G6	N7
A7	N24	B7	N25	C7	N19	D7	N21	E7	N28	F7	N14	G7	N7
A8	N22	B8	N22	C8	N15	D8	N10	E8	N13	F8	N6	G8	N4
A9	N16	B9	N36	C9	N1	D9	N2	E9	N14	F9	N42	G9	N3

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 1 July 2026, pp. 544-559

Afin de regrouper les neurones ayant des profils d'individus similaires, la méthode de classification hiérarchique ascendante (CAH) a été utilisée. Le résultat obtenu est présenté ci-dessous :

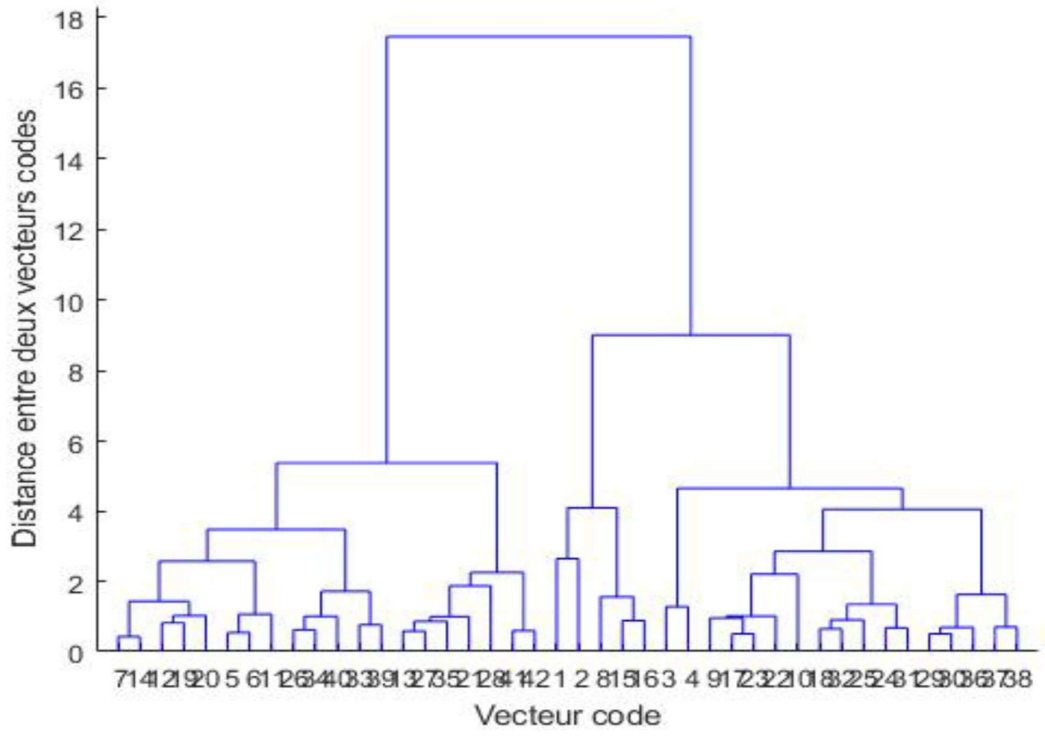

Figure 9 : Regroupements des neurones

Nous pouvons donc conclure que :

- **la première zone** regroupe les points A1, A2, A3, A4, A5, A6, A7, A8, A9, B1, B2, B3, B4, B5, B6, B7, B8, B9, C1, C2, C3, C4, C5, C6, C8, C9, D3, D4, D5, D8, D9, G8 et G9.
- **la deuxième zone** regroupe les points C7, D1, D2, D6, D7, E1, E2, E3, E4, E5, E6, E7, E8, E9, F1, F2, F3, F4, F5, F6, F7, F8, F9, G1, G2, G3, G4, G5, G6 et G7.

Il est alors possible de représenter sur la carte géographique de Madagascar les individus classés en fonction de la quantité de précipitations convectives des différentes sous-zones, comme le montre la figure suivante :

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 1 July 2026, pp. 544-559

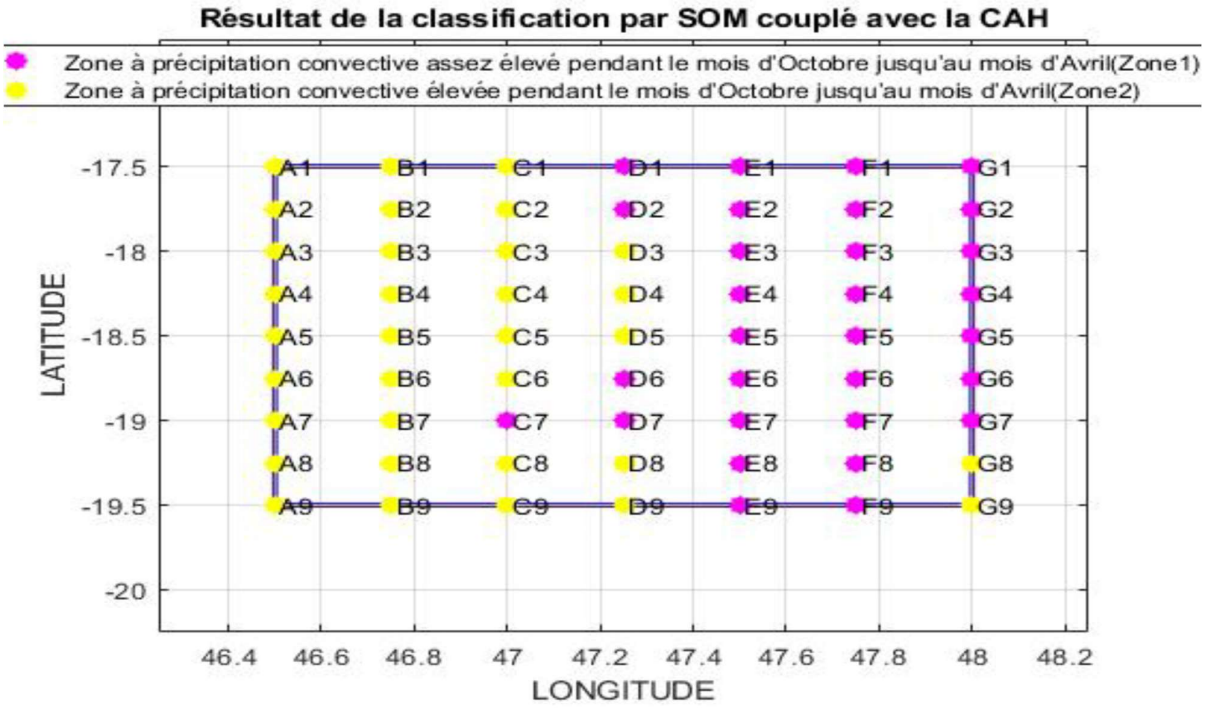

Figure 10 : Résultat avec la méthode de Kohonen couplé avec la CAH

d) Performance de chaque résultat par l'indice de Dunn :

La qualité de chaque regroupement est résumée dans le tableau II.6, lequel indique les valeurs de l'indice de Dunn obtenues pour chaque méthode de classification.

MÉTHODE	VALEUR DE L'INDICE
T-SNE avec K-MEANS	0,3320905
ACP avec K-MEANS	0,1346411
SOM ou Réseau de Kohonen avec CAH	0,06297147

L'analyse comparative montre que le couplage t-SNE + K-Means améliore la séparation visuelle des clusters mais peut introduire une variabilité dans la structure globale. L'indice de Dunn obtenu (0.332) indique une séparation nette entre les groupes.

e) Caractérisation des Zones Identifiées

La régionalisation a permis de diviser Analamanga en deux entités majeures :

Zone 1 (Zone à forte intensité) : Située principalement sur les reliefs, elle enregistre les cumuls de pluie convective les plus élevés. La saison humide y est plus précoce (débutant dès octobre).

Zone 2 (Zone modérée) : Correspond aux zones de vallées et de plateaux inférieurs, où l'activité convective est moins intense en moyenne annuelle.

5. DISCUSSION

Plusieurs recherches scientifiques sur la régionalisation climatique utilisent des paramètres tels que les précipitations, la température, l'humidité ou encore le vent afin d'identifier des zones homogènes. Les méthodes les plus utilisées reposent généralement sur l'ACP associée à des techniques de clustering comme K-Means [13], [14]. D'autres études appliquent également des méthodes de réduction de dimension non linéaires telles que le t-SNE combiné à des algorithmes de classification afin d'améliorer la séparation des groupes climatiques [15],[8].

Notre démarche est similaire à ces travaux, mais elle est appliquée spécifiquement à la pluie convective dans la région d'Analamanga. De plus, cette étude compare trois approches différentes de régionalisation : ACP + K-Means, t-SNE + K-Means et SOM + CAH, tout en utilisant l'indice de Dunn pour évaluer la qualité des regroupements obtenus. Les résultats montrent que la représentation obtenue avec le t-SNE met davantage en évidence la séparation des sous-zones climatiques par rapport à l'ACP classique. Bien que t-SNE ne conserve pas les distances globales, il est utilisé ici comme outil de projection pour améliorer la séparation visuelle avant clustering.

Globalement, ces résultats sont en accord avec la littérature scientifique récente qui montre que les méthodes non linéaires sont mieux adaptées à l'analyse des phénomènes climatiques complexes. La régionalisation obtenue constitue ainsi une base importante pour améliorer les modèles de prévision des précipitations à l'échelle locale.

6. CONCLUSION ET PERSPECTIVES

Cette étude a permis de cartographier la pluie convective à Analamanga et d'identifier des zones homogènes. Les méthodes non linéaires, notamment le t-SNE, ont montré une meilleure capacité de séparation des structures climatiques.

Ces zones homogènes constituent une base pour l'entraînement de modèles de deep learning de type CNN et GRU, afin d'améliorer la précision des prévisions de pluie convective à court terme. Cette approche peut contribuer au développement d'un système d'alerte précoce face aux risques d'inondation.

References

[1] R. B. Smith, « The Influence of Mountains on the Atmosphere », *Adv. Geophys.*, vol. 21, p. 87-230, 1979, doi: 10.1016/S0065-2687(08)60262-9.

[2] Bureau National de Gestion des Risques et des Catastrophes (BNGRC), « Point de situation n°3 – Fortes pluies », Antananarivo, janv. 2024.

[3] Anatra R., « Fortes pluies – Plus de 2 000 sinistrés à Analamanga », La Vérité, 16 février 2025. Consulté le: 11 avril 2026. [En ligne]. Disponible sur: https://laverite.mg/societe/item/23620-fortes-pluies-%20plus-de-2-000-sinistr%C3%A9s-%C3%A0-analamanga.html

[4] P.-L. Gonzalez, « L'Analyse en Composantes Principales (ACP) ». Conservatoire National des Arts et Métiers, Paris, 2017. [En ligne]. Disponible sur: https://maths.cnam.fr/IMG/pdf/A-C-P-.pdf

[5] M. Denis, « L'analyse en composantes principales ». École Polytechnique de Montréal, Montréal, Canada, 2021. Consulté le: 21 avril 2026. [En ligne]. Disponible sur: https://cours.polymtl.ca/geo/marcotte/glq3402/chapitre3.pdf

[6] S. Mestiri, « Chapitre 2 : Analyse en composante principales (ACP) ».

[7] Z. Ansari, M. F. Azeem, W. Ahmed, et A. V. Babu, « Quantitative Evaluation of Performance and Validity Indices for Clustering the Web Navigational Sessions », 2011.

[8] L. van der Maaten et G. E. Hinton, « Visualizing Data using t-SNE », vol. 9, nº November (Nov), p. 2579-2605, 2008, doi: 10.5555/1390681.1390693.

[9] G. C. Linderman et S. Steinerberger, « Clustering with t-SNE, Provably », *SIAM Journal on Mathematics of Data Science*, vol. 1, nº 2, p. 313-332, janv. 2019, doi: 10.1137/18M1216134.

[10] T. Kohonen, « The Self-Organizing Map. », vol. 78, nº 9, Institute of Electrical and Electronics Engineers (IEEE), p. 1464-1480, 1990. doi: DOI: 10.1109/5.58325.

[11] J. Vesanto et E. Alhoniemi, « Clustering of the Self-Organizing Map », vol. 11, nº 3, IEEE Transactions on Neural Networks, p. : 586-600. doi: 10.1109/72.846731.

[12] M. Gagolewski, M. Bartoszuk, et A. Cena, « Are Cluster Validity Measures (In)valid? », *Information Sciences*, vol. 581, p. 620-636, déc. 2021, doi: 10.1016/j.ins.2021.10.004.

[13] V. V. Srinivas, « Regionalization of Precipitation in India—A Review ».

[14] D. S. Wilks, *Statistical methods in the atmospheric sciences*, 2nd ed. in International geophysics series, no. volume 91. Amsterdam Paris: Elsevier, 2006.

[15] D. Kobak et P. Berens, « The art of using t-SNE for single-cell transcriptomics », *Nat Commun*, vol. 10, nº 1, p. 5416, nov. 2019, doi: 10.1038/s41467-019-13056-x.