JOURNAL AFRICAIN DES SCIENCES
VOL. 1, **1** (2024) 156-169

https://jafrisci.org

ISSN (online) : 3107-0949 ; ISSN (print) : 3107-0930

OPEN ACCESS

ARTICLE DE RECHERCHE

Article Info.:
Reçu : le 29/06/2024
Accepté : le 15/08/2024
Publié : le 22/08/2024

ANALYSE DE L'INFLUENCE DES PARAMETRES METEOROLOGIQUES SUR LA PLUVIOMETRIE A KINSHASA

Paulo Rostha Lohalo [1,*], Elie Kianga [1], Roger Mvumilia Kizungu [2]

https://doi.org/10.70237/jafrisci.2024.v1.i1.15

Résumé

La compréhension des paramètres météorologiques influençant la pluviométrie est essentiellepour répondre efficacement aux défis environnementaux et anticiper les conséquences des phénomènes météorologiques extrêmes. Dans cette étude, nous analysons les données météorologiques fournies par la Régie des Voies Aériennes (RVA) de Kinshasa, couvrant la période de1992 à 2019. Cette analyse statistique approfondie vise à mieux cerner les facteurs qui conditionnent les précipitations dans la région et à identifier les tendances climatiques pertinentes pour anticiperles futurs changements environnementaux. Nous avons utilisé la régression logistique pour évaluer l'impact de divers facteurs climatiques tels que la température (les extrema journaliers),l'humidité relative et la pression atmosphérique sur les précipitations. Les résultats indiquent que la température minimale, l'humidité relative et la pression atmosphérique présentent une forte corrélation avec la pluviométrie à Kinshasa, soulignant leur influence significative sur la probabilité de précipitations. En revanche, la température maximalen'a qu'un impact limité, ce qui en fait un indicateur moins fiable de la pluie dans la ville. Les conclusions de cette étude offrent des outils précieux pour les autorités locales dans le développement de stratégies efficaces pour la gestion des variations climatiques et la prévisiondes phénomènes météorologiques. Ces insights permettent de modéliser avec précision les impacts des fluctuations météorologiques sur les écosystèmes et les infrastructures urbaines. De plus, les résultats de cette recherche peuvent orienter les agriculteurs locaux dans l'optimisation de leurs pratiques agricoles, en ajustant les calendriers de plantation et de récolte en fonction des prévisions météorologiques détaillées, permettant ainsi de minimiser les pertes économiques et d'améliorer la résilience des cultures face aux variations climatiques. Enfin, cette étude offre aux climatologues des données robustes pour une analyse approfondie des effets des changements climatiques sur les phénomènes météorologiques à Kinshasa. Ces connaissances sont capitales pour affiner les modèles climatiques régionaux et développer desstratégies d'adaptation spécifiques à d'autres zones urbaines de l'Afrique subsaharienne.

Mots clés : Influence, température, humidité relative, pression atmosphérique, pluviométrie, météorologie, régression logistique

Abstract

The understanding of meteorological parameters influencing rainfall is essential to effectivelyaddress environmental challenges and anticipate the consequences of extreme weather events.In this study, we analyze meteorological data provided by the Airways Authority of Kinshasa(RVA), covering the period from 1992 to 2019. This in-depth statistical analysis aims to better understand the factors that influence precipitation in the region and to identify relevant climate trends for anticipating future environmentalchanges. We used logistic regression to assess the impact of various climatic factors such as temperature (daily extremes), relative humidity, and atmospheric pressure on precipitation. The results indicate that minimum temperature, relative humidity, and atmospheric pressure show a strong correlation with rainfall in Kinshasa, highlighting their significant influence onthe likelihood of precipitation. In contrast, maximum temperature has a limited impact, makingit a less reliable indicator of rainfall in the city. The conclusions of this study provide valuable tools for local authorities in developing effective strategies for managing climate variations and predicting weather phenomena. These insights enable precise modeling of the impacts of meteorological fluctuations on ecosystems and urban infrastructure. Furthermore, the results of this research can guide local farmers in optimizing their agricultural practices by adjusting planting and harvesting schedules based ondetailed weather forecasts, thereby minimizing economic losses and enhancing crop resilienceto climate variations. Finally, this study provides climatologists with robust data for in-depth analysis of the effectsof climate change on meteorological phenomena in Kinshasa. This knowledge is crucial for refining regional climate models and developing specific adaptation strategies for other urbanareas in sub-Saharan Africa.

Key words : Influence, temperature, relative humidity, atmospheric pressure, precipitation, meteorology, logistic regression.

Correspondance : *paulo.rostha@unikin.ac.cd* ***(P. R. Lohalo)***

The Author(s) Published under a Creative Commons Attribution 4.0 International Licence (CC BY 4.0)

[1] Département de Physique Faculté de Sciences Université de Kinshasa (UNIKIN), BP190 Kinshasa 11
[2] Département de Chimie et Industrie Agro-alimentaire, Faculté des Sciences Agronomiques, Université de Kinshasa, BP. 117 Kinshasa 11, République Démocratique du Congo

1. INTRODUCTION

1.1. Problématique

Les précipitations sont un élément clé du système climatique, impactant directement l'agriculture, la gestion des ressources en eau, la production d'énergie hydroélectrique et la prévention des catastrophes naturelles telles que les inondations et les sécheresses. Les variations de la pluviométrie peuvent entraîner des conséquences socioéconomiques importantes. Pour mieux comprendre ces enjeux, de nombreuses études se sont penchées sur les mécanismes de formation des précipitations et leurs relations avec les différents paramètres atmosphériques (Trenberth, 2011). Ces recherches ont permis de développer des modèles numériques de prévision des précipitations de plus en plus sophistiqués (Kain et al., 2004).

Plusieurs études ont exploré l'utilisation de modèles linéaires, tels que la régression linéaire multiple, pour prédire les précipitations (Ahmed et Karmakar, 1993 ; Salinger et Griffiths, 2001). Ces modèles ont permis d'établir des corrélations entre les précipitations et des variables météorologiques comme la température, l'humidité ou la pression atmosphérique. Cependant, leur capacité à capturer la complexité des phénomènes météorologiques est limitée par leur hypothèse de linéarité.

La régression logistique, en tant que modèle linéaire généralisé, s'est révélée particulièrement adaptée à la prévision des précipitations. En modélisant la probabilité qu'un événement binaire se produise (ici, la présence ou l'absence de précipitations), elle permet de s'affranchir de l'hypothèse de linéarité et de capturer des relations plus complexes entre les variables prédictives, comme l'ont souligné de nombreux auteurs (Wilks, 2011).

La pertinence de la régression logistique pour l'étude des précipitations réside dans plusieurs aspects. Tout d'abord, elle permet de modéliser des relations non linéaires entre les variables, ce qui est essentiel pour rendre compte de la complexité des phénomènes météorologiques (Wilks, 2011). Ensuite, les coefficients de la régression logistique sont facilement interprétables en termes de rapports de cotes, facilitant l'analyse de l'impact de chaque variable prédictive sur la probabilité de précipitations (Hosmer & Lemeshow, 2000). Ces rapports de cotes permettent d'évaluer l'augmentation ou la diminution du risque de précipitations associée à une variation d'une unité d'une variable explicative donnée.

Les avancées récentes en matière d'apprentissage automatique offrent de nouvelles perspectives. Les réseaux de neurones artificiels, notamment les réseaux de neurones convolutifs, ont démontré leur capacité à capturer des patterns complexes dans les données météorologiques haute résolution (Rasp & Leroux, 2018). Bien que ces méthodes offrent de nouvelles perspectives, la régression logistique conserve une place importante dans la boîte à outils du climatologue en raison de sa simplicité, de son interprétabilité et de sa capacité à fournir des résultats puissants dans de nombreux contextes.

Dans cette étude, nous appliquons la régression logistique aux données météorologiques fournies par la Régie des Voies Aériennes (RVA) de Kinshasa pour évaluer l'impact des paramètres météorologiques locaux sur la pluviométrie. Nous visons à déterminer comment les extrema journaliers de température, l'humidité relative et la pression atmosphérique influencent la présence ou l'absence de pluie à Kinshasa, et à quantifier l'effet de chacun de ces facteurs sur les précipitations dans cette région.

1.2. Source des données

Les données météorologiques analysées dans la présente étude proviennent de la régie des voies aériennes de l'aéroport international de N'djili, situé à Kinshasa en République Démocratique du Congo. Cet aéroport, principal desservant la capitale, est un site de référence pour les mesures météorologiques dans la région. Localisé dans la commune de N'sele, à environ 28 km du centre-ville de Kinshasa, l'aéroport de N'djili est bordé au nord par le marécage et le fleuve Congo. Il est délimité au sud par les quartiers Mangongele et Bel Air, à l'ouest par la commune de Masina et à l'est par le quartier Mikonga.

En tant qu'entité spécialisée dans le domaine aéronautique, la régie des voies aériennes dispose d'équipements de pointe et d'un personnel qualifié pour effectuer des relevés météorologiques précis et fiables sur le site même de l'aéroport. Cette source d'informations météorologiques de qualité garantit la rigueur et la fiabilité des données exploitées dans nos analyses.

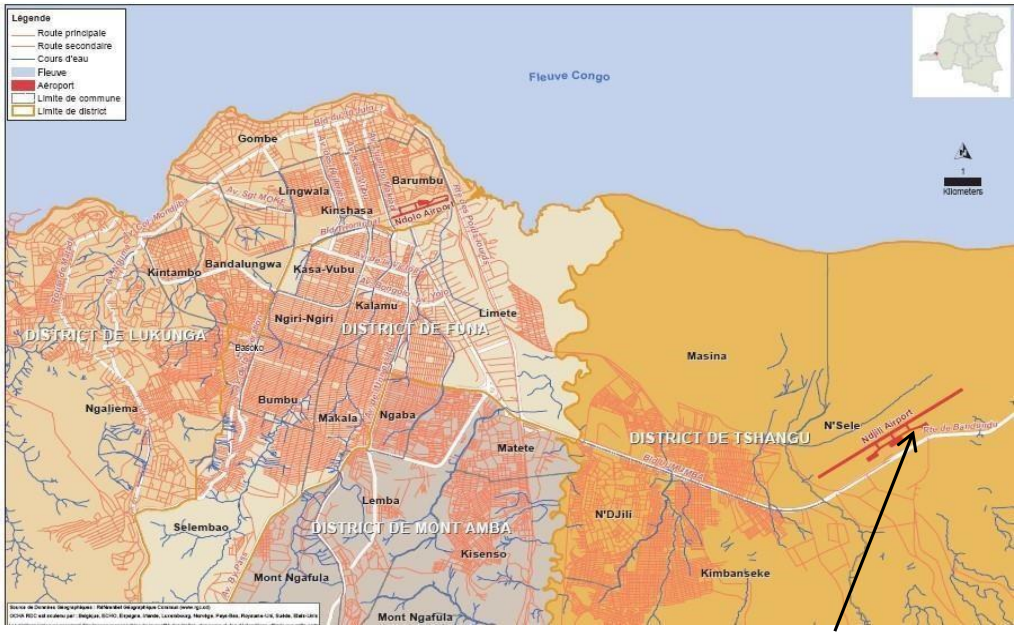


Figure 1 : *Carte administrative de la ville de Kinshasa*

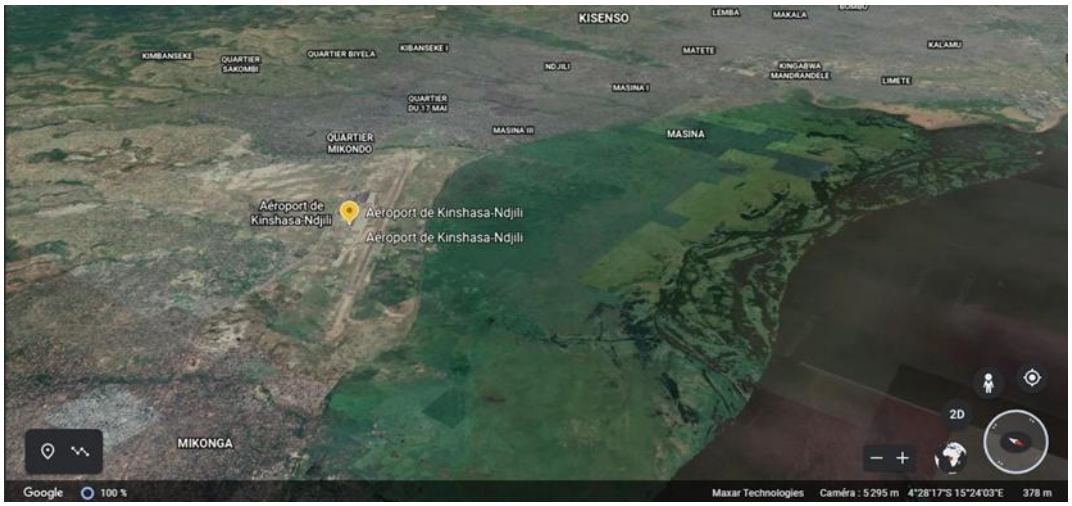


Figure 2 : *Zone de prélèvement des données à l'aéroport de N'djili (Source : Google Earth)*

2. MATERIELS ET METHODES

2.1. Matériels et logiciels utilisés

Pour la collecte des données météorologiques, les appareils suivants ont été employés :

- Baromètre à mercure (produit par la Sté Précis-Mécanique 95870 BEZONS FRANCE N°502)
- Baromètre enregistreur ou barographe (Barographe Lambrecht TYP 0002901-010011 N° 66T829-0001, Diagramme 82B106,7hpa, 0-930 Meter)
- Pluviomètre SDM6A, MC86, 00000139, 0.05-10mm
- Thermomètre, Psychromètre, N° 200107



Figure 3 : *Station météorologique de l'aéroport de N'djili*

Pour le traitement et l'analyse des données collectées, les logiciels suivants ont été utilisés :

- Microsoft Excel : Ce tableur a permis la saisie, l'organisation et le prétraitement des données météorologiques.
- R Studio : Cet environnement de développement intégré (IDE) pour le langage de programmation R a été employé pour effectuer des analyses statistiques approfondies et la visualisation des données.

L'utilisation conjointe de ces appareils de mesure précis et de ces logiciels d'analyse puissants a permis une collecte et un traitement rigoureux des données météorologiques, conformément aux normes scientifiques en vigueur.

2.2. Méthodologie

L'analyse de l'influence des paramètres météorologiques sur la pluviométrie implique une démarche rigoureuse en plusieurs étapes clés. Dans un premier temps, les données relatives à la pluviométrie ainsi qu'aux différents paramètres météorologiques tels que les extrema journaliers de température, l'humidité relative et la pression atmosphérique ont été collectées.
Ensuite, une analyse statistique approfondie est réalisée sur ces données à l'aide de la méthode de régression logistique. Cette approche permet de déterminer de manière précise l'impact des différents paramètres météorologiques sur la variable dépendante binaire, en l'occurrence la présence ou l'absence de pluie.

En suivant cette méthodologie, alliant la collecte de données fiables et l'application d'une technique statistique éprouvée, notre étude vise à fournir des résultats sur l'impact des conditions météorologiques sur la pluviométrie à Kinshasa. Cette approche systématique permettra d'identifier les paramètres clés et de proposer un modèle de prévision des précipitations adapté au contexte local.

2.2.1. La régression logistique

La régression logistique est une méthode statistique qui permet d'établir la relation entre une variable binaire (variable dépendante) et une ou plusieurs variables explicatives (variables indépendantes). Elle est utilisée pour analyser les relations entre des variables quantitatives et des réponses binaires. (Lohalo, 2015).
Le problème spécifique de la régression logistique consiste à prédire une variable binaire y dont les deux issues possibles représentent un succès (Y = 1) et un échec (Y = 0).
Dans la régression linéaire multiple, on modélise Y comme une combinaison linéaire des variables explicatives :

$Y = a + b_1x_1 + b_2x_2 + \cdots + b_nx_n + \varepsilon$ (1) Où a est l'ordonnée à l'origine (intercepte), $b_1, b_2, \ldots, b_n$ sont les coefficients associés aux variables explicatives, et ε est un terme d'erreuraléatoire.
La fonction logistique, notée f(x), est une fonction sigmoïde strictement croissante, variant de 0 à 1 lorsque x varie de -∞ à +∞. Elle est définie par :

$f(x) = \frac{1}{1+e^{-(ax+b)}}$ (2a) ou $f(x) = \frac{e^Y}{1+e^Y}$ (2b)

Cette fonction possède un point d'inflexion à x = 0, où $f(x) = \frac{1}{2}$ et est symétrique autour de ce point.
Dans le modèle de régression logistique, on remplace la combinaison linéaire de (1) par une fonction logistique :

$P(Y = 1 \mid X_1, X_2, \ldots, X_n) = f(a + b_1X_1 + b_2X_2 + \ldots + b_nX_n)$ (3)
Où $P(Y = 1 \mid X_1, X_2, \ldots, X_n)$ représente la probabilité que Y = 1 (succès) étant donné les valeurs des variables explicatives.
Les paramètres $a, b_1, b_2, \ldots, b_n$ sosont estimés par la méthode du maximum de vraisemblance, qui consiste à maximiser la fonction de vraisemblance des observations par rapport à ces paramètres.
Le logit, note $logit(p) = \log(\frac{p}{1-p})$ est la fonction inverse de la fonction logistique.
Il permet de linéariser le modèle et d'exprimer le logarithme du rapport de chances (odds ratio) comme une combinaison linéairedes variables explicatives.
Dans notre étude, la régression logistique s'avère particulièrement pertinente pour modéliser la probabilité d'absence ou de présence de pluie en fonction des paramètres

météorologiques tels que les extrema journaliers de température, d'humidité relative et depression atmosphérique.
En effet, cette méthode nous permettra d'exprimer la variable dépendante binaire(pluie/pas de pluie) comme une fonction logistique des variables explicatives météorologiques. Nous pourrons ainsi estimer la probabilité de précipitations pour différentes combinaisons de ces paramètres. En plus,l'examen des coefficients estimés pour chaque variable explicative nous aidera à identifier les facteurs météorologiques clés influençant la probabilité de pluie. Par exemple, nous pourrons déterminer si destempératures élevées, une forte humidité ou une baisse de pression augmentent significativement les risques de précipitations.

Enfin, l'interprétation des rapports de cotes (odds ratios) facilitera la compréhension de l'ampleur et de la direction des effets de ces variables météorologiques sur la probabilité de pluie. Cela nous fournira des informations précieuses pour mieux appréhender les mécanismes sous-jacents et les interactions complexes en jeu.

3. RÉSULTATS ET INTERPRÉSENTATIONS

3.1. Température mavimale vs pluviomètre

Dans cette analyse, la variable à expliquer Y représente la survenue ou non d'un événement pluvieux. Il s'agit d'une variable binaire codée comme suit : présence de pluie ($Rnfl^{+}$) ou absence de la pluie $Rnfl^{-}$. Nousconsidérons un modèle de régression logistique simple impliquant une seule variable explicative, à savoir la température maximale journalière (T_{max}).
Le modèle s'exprime selon l'équation suivante :

$$P(\text{Rnfl}|T_{max}) = f(T_{max}) = \frac{e^{(aT_{max}+b)}}{1+e^{-(aT_{max}+b)}} \quad (4)$$

Où

- $P(Rnfl^{+}|T_{max})$: la probabilité d'occurrence de pluie si la variable maximum journalier est prise en compte.
- $f(T_{max})$: la fonction logistique.
- a et b sont les coefficients du modèle à estimer.

Estimation des coefficients

Les coefficients a et b sont déterminés par une régression logistique implémentée avec la fonction $glm(Rnfl \sim Tmax, family = "binomial")$ du logiciel R. Il s'agit d'un modèle linéaire généralisé dans lequel on spécifie la fonction de lien, la fonction binomiale,adaptée à la nature binaire de la variable dépendante.

Les résultats de cette régression sont résumésdans le tableau suivant :

Tableau 1: Coefficients de la régression logistique $Rnfl(T_{max})$

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2,478785	0,278681	-8,895	< 2e-16
$Tmax$	0,049289	0,008985	5,486	4,12e-08

Les coefficients estimés permettent d'exprimer le modèle de la façon suivante :

$$P(\text{Rnfl}|T_{max}) = f(T_{max}) = \frac{e^{(0,049289T_{max}-2,478785)}}{1+e^{-(0,049289T_{max}+2,478785)}} \quad (5)$$

Le coefficient associé à T_{max} (0,049289) est positif, indiquant qu'une augmentation de la température maximale accroît la probabilité de pluie. Cet effet est statistiquement significatif avec une p-value inférieure à 0,001. Plusprécisément, l'erreur standard de 0,008985 permet de calculer un intervalle de confiance pour le coefficient. Par exemple, un intervalle de confiance à 95% serait [0,031679 ; 0,066899], confirmant l'effet positif de T_{max} sur la probabilité de pluie. Le coefficient de la constante (-2,478785) n'a pas d'interprétation directe mais sert à ajuster la fonction logistique au point d'origine. La statistique z permet de tester l'hypothèse nulle selon laquelle le coefficient est nul.Avec une valeur de 5,486 pour T_{max}, l'hypothèse nulle est rejetée au seuil de 0,001, renforçant la significativité de cette variable.

a) Représentation graphique du modèle des observations

La représentation graphique de (5) en fonction de T_{max} donne la fonction logistique ci-dessous :

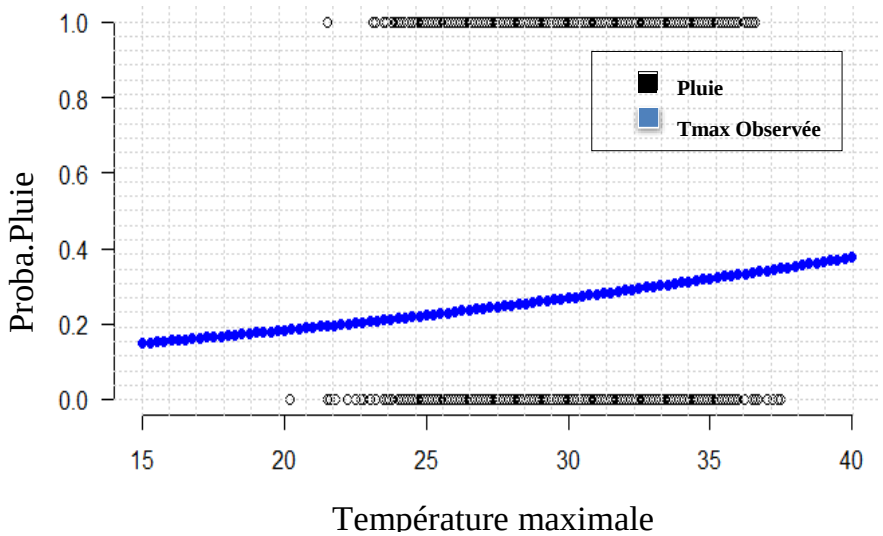


Figure 4: *Influence de la Température Maximale sur la pluviométrie*

b) Estimation du modèle

Utilisons la fonction $glm(rbinom(n, 1, p) \sim Tmax, family = "binomial")$ du logiciel R pour estimer les paramètres du modèle à l'aide de l'échantillon. L'output de cette régressionest résumé dans le tableau 2, qui présente les coefficients de la régression logistique pour lemodèle estimé.

Tableau 2 : Coefficients de la régression logistique $\boldsymbol{Rnfl(Tmax)}$ du modèle estimé

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2,752430	0,279819	-9,836	< 2e-16
$Tmax$	0,058220	0,009014	6,459	1.06e-10

L'intercept (constante) du modèle est estiméà -2,752430 avec une erreur standard de 0,279819, ce qui correspond à un z-value de
-9,836 et une p-valeur très faible, inférieure à 2e-16. Cela indique que l'intercept est statistiquement significatif. Le coefficient associé à la variable $Tmax$ est estimé à 0,058220 avec une erreur standard de 0,009014, un z-value de 6,459 et une p-valeur de 1,06e-10. Cela suggère que la variable $Tmax$ a également une signification statistique dans le modèle.

Le modèle estimé peut s'écrire comme :

$$\log\left(\frac{p}{1-p}\right) = 0{,}058220\, T_{max} - 2{,}752430 \quad (6)$$

$$\log\left(\frac{Rnfl}{1-Rnfl}\right) = 0{,}058220\, T_{max} - 2{,}752430 \quad (7)$$

Ces résultats indiquent que la température maximale (T_{max}) a un effet significatif sur la probabilité de survenue de précipitations, comme le suggèrent les estimations des coefficients et leurs significativités statistiques.

c) Validation par la méthode du maximumde vraisemblance

En régression logistique, l'estimation des coefficients des variables indépendantes repose sur la méthode du maximum de vraisemblance (Maximum Likelihood Estimation, MLE), car la méthode des moindres carrés, utilisée pour la regression linéaire, n'est pas appropriée pour les données binaires. La MLE consiste à identifier les valeurs des coefficients qui maximisent la probabilité d'observer les données telles qu'elles se présentent. Autrement dit, elle trouve le meilleur ajustement entre le modèle statistique et les données observées. (James etal., 2013)
Démarche de la MLE :

- **Construire la fonction de vraisemblance :**

La fonction de vraisemblance est basée sur laprobabilité de chaque observation donnée parle modèle. Pour un modèle de régression logistique, cela implique de considérer les probabilités prédites pour la présence ou l'absence d'un événement.

- **Maximiser le logarithme de la fonction de vraisemblance :**

Étant donné que la fonction de vraisemblancepeut devenir très petite, il est plus pratique de travailler avec son logarithme. La maximisation du logarithme simplifie également les calculs mathématiques. En maximisant ce logarithme, on obtient les valeurs des paramètres qui rendentl'observation des données la plus probable. (James et al., 2013).

Dans notre étude, nous avons utilisé le logiciel R pour simuler et valider notre modèle. La fonction $rbinom(n, 1, p)$ en R simule n essais de Bernoulli, où chaque essaia une probabilité p de succès. Cette simulation permet de générer des résultats théoriques pour les comparer aux données observées, fournissant ainsi un moyen de valider la précision du modèle.

Voici les calculs basés sur nos données :

- Nombre total d'observations : n=9846
- Présences observées : $\sum_{i=1}^{n} \text{rbinom}(\text{n}, 1, \text{p}) = 2734$
- Absences observées : $n - \sum_{i=1}^{n} \text{rbinom}(\text{n}, 1, \text{p}) = 7112$

La probabilité de présence est :

$$p = \frac{\sum_{i=1}^{n} \text{rbinom}(\text{n}, 1, \text{p})}{n} = 0{,}2776762$$

Et par conséquent, la probabilité d'absence est :
$1 - p = 0{,}72232381$

Validation du modèle par la vraisemblance :

Sous l'hypothèse que notre modèle estcorrect, la vraisemblance des observations estdonnée par :
$P(Observations) = L(p)$

$p(Observations) = L(p) =$
$\prod_{i=1}^{n} p(y_i) = 0{,}2776762^{2734}\ x\ 0{,}7223238^{7112}$

Calcul du logarithme de la vraisemblance:

$\log\big(p(Observations)\big) = \log\big(L(p)\big)$
$= 2734log(0{,}2776762)\ x\ 7112log(0{,}2776762)$

Ce calcul résume l'ajustement global du modèle aux données

observées, prenant en compte la fréquence des présences et des absences.

Déviance du modèle :

La déviance mesure la qualité de l'ajustementd'un modèle. La déviance est calculéecomme suit pour le modèle testé :

$-2\log(L(P)) = -2 \times (-5816{,}47686) = 11633$

Nous comparons cette déviance à la déviancedu modèle nul, obtenue dans R avec 9845 degrés de liberté : "11633 on 9845 degrees of freedom". La valeur minimale obtenue valide notre modèle car elle montre que la probabilité estimée p = 0.2776762 réduit la déviance, signifiant un ajustement statistiquement valable.
La méthode du maximum de vraisemblance, en utilisant la fonction de vraisemblance et la déviance, démontre que notre estimation de p = 0.2776762 est conforme aux observations, indiquant que notre modèle de régression logistique est statistiquement robuste et bien ajusté aux données.

d) Représentation graphique du modèle des observations et de l'estimation

En représentant sur un même graphique le modèle : $Rnfl^{+}(T_{max})$ et $\boldsymbol{Rnfl(Tmax)}$, nous obtenons :

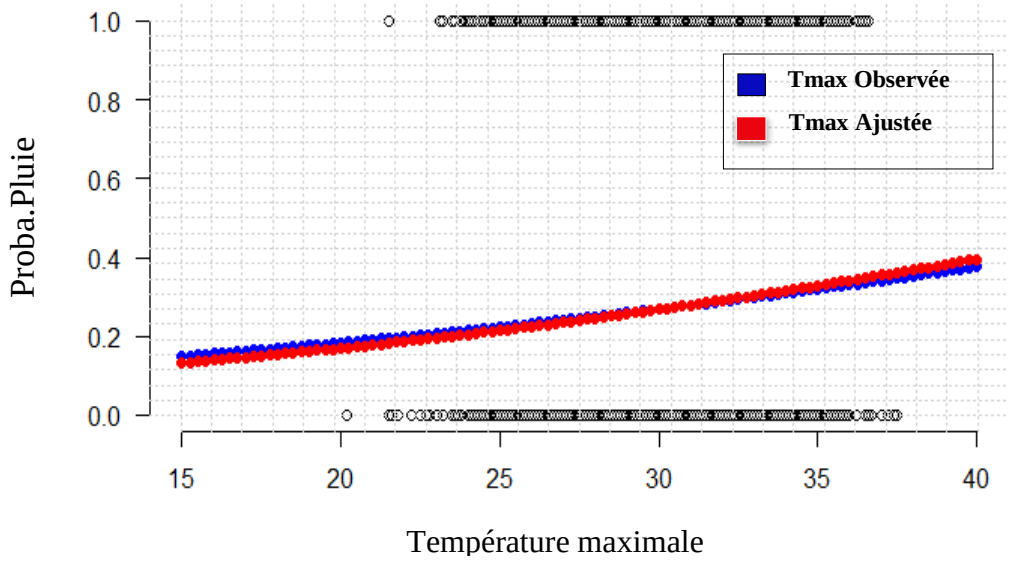


Figure 5 : *Influence de la Température Maximale (observéeetajustée) sur la pluviométrie*

La courbe bleue (les valeurs observées) représente la probabilité de pluie pour chaque valeur de température. Elle montre commentles données réelles se distribuent avec la tem pérature maximale. La courbe rouge (le modèle estimé) montre les probabilités de pluie prédites par le modèle de régression logistique ajusté sur les données. Cette courbe est une estimation statistique de la relation entre la température maximale et la probabilité de pluie.

Pour toutes les valeurs de maxima journaliercomprises entre 15 et 20°C, la probabilité qu'il pleuve varie entre 0,15 et 0,18. Pour lestempératures maximales journalières comprises entre (20 et 25) °C, la probabilité d'unepluie augmente de de 0,18 à 0,22. De 25 à 30°C, nous constatons un accroissement de 0,22 à 0,27. De 30 à 35°C, la probabilité monte de 0,27 à 0,32. De 35 à 40°C, nous remarquons une variation est de 0,32 à 0,38.
Le model observé et estimé indiquent que le modèle logistique est un bon ajustement pourles données observées. Le model estimé est légèrement plus lisse que le model observé, ce qui est typique pour les modèles ajustés qui tentent de capturer la tendance centrale tout en réduisant la variabilité aléatoire des données observées.
Les points noirs montrent la distribution des observations réelles, avec des données regroupées principalement autour des valeurs de probabilité proches de 0 et 1. Cela indiqueque la plupart des jours sont soit secs soit pluvieux, avec moins de jours ayant une probabilité intermédiaire de pluie. Le graphique 5 illustre donc que le modèle de régression logistique ajuste adéquatement la relation entre la température maximale et la probabilité de pluie.

3.2. Température minimale vs pluviomètre

En considérant $Rnfl$ comme la variable à expliquer (indiquant la survenue de la pluie) et T_{min} comme la variable explicative (température minimale), nous pouvons exprimer notre modèle de régression logistique à l'aidede l'équation suivante :

$$p(Rnfl^{+}|T_{min}) = f(T_{min}) = \frac{e^{(aT_{min}+b)}}{1+e^{-(aT_{min}+b)}} \quad (9)$$

- $P(Rnfl^{+}|T_{min})$: Probabilité qu'il pleuve, si la température minimale quotidienne est prise en compte.
- $f(T_{min})$: la fonction logistique.
- a et b : Coefficients de la régression logistique déterminés empiriquement par la procédure de maximumde vraisemblance via le logiciel R.

Les résultats de la régression logistique, obtenus en utilisant glm ($Rnfl$~$Tmin$, $family$ = "$binomial$") dans le logiciel R, sont présentés dans le tableau suivant :

Tableau 3 : Coefficients de la régressionlogistique $\boldsymbol{Rnfl(Tmin)}$

	Estimate	Std. Error	z value	Pr(>|z|)
Intercept)	- 8,66374	0,33009	- 26,25	<2e-16
T_{min}	0,35435	0,01499	23,65	<2e-16

De ces coefficients estimés, nous pouvonsécrire le modèle :

$$p(Rnfl^{+}|T_{min}) = f(T_{min}) = \frac{e^{(0,35435T_{min}-8,66374)}}{1+e^{-(0,35435T_{min}+8,66374)}} \quad (10)$$

La représentation graphique de $P(Rnfl^{+}|T_{min})$ en fonction de $Tmin$ donne la fonction logistique ci-dessous :

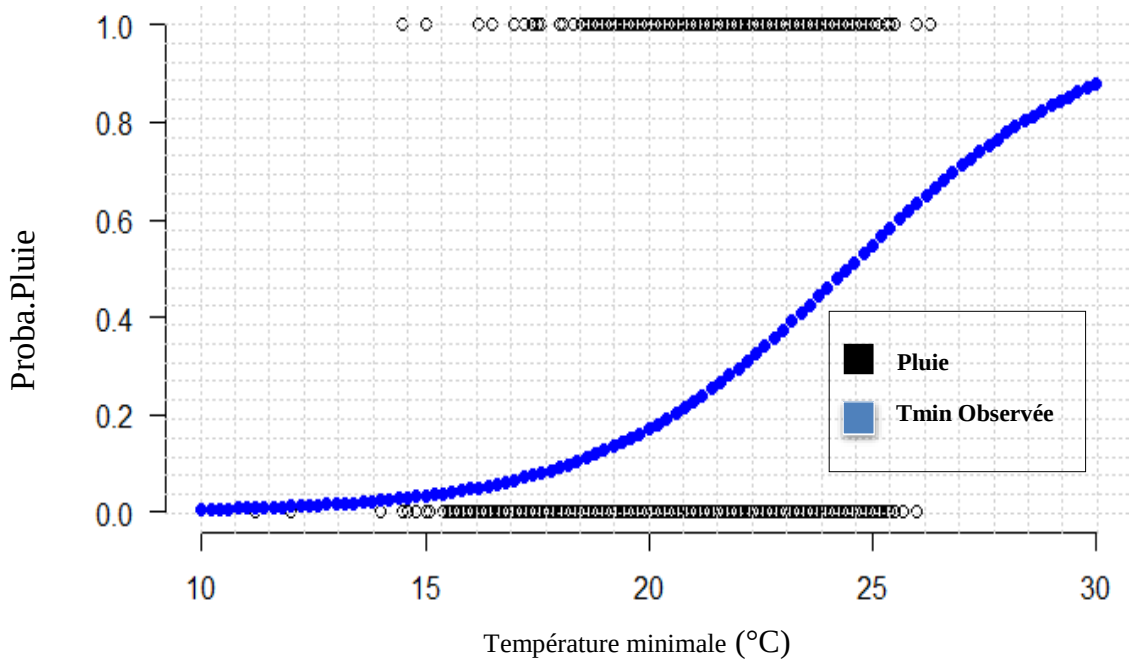


Figure 6 : *Influence de la Température Minimale (observée) sur la pluviométrie*

Estimation du modèle

Pour estimer les paramètres du model avec l'échantillon, nous avons utilisé : $glm(Rnfl/T_{min},\ family = binomial)$ du logiciel R. Les résultats issus de l'estimation dumodèle sont présentés dans le tableau suivant :

Tableau 4 : Coefficients de la régression logistique $\boldsymbol{Rnfl(Tmin)}$ du modèle estimé

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-8,44678	0,32856	-25,71	<2e-16
T_{min}	0,34393	0,01492	23,05	<2e-16

Le modèle estimé peut alors s'écrire comme : $\log\left(\frac{p}{1-p}\right) = 0,34393T_{min} - 8,44678$ (11)

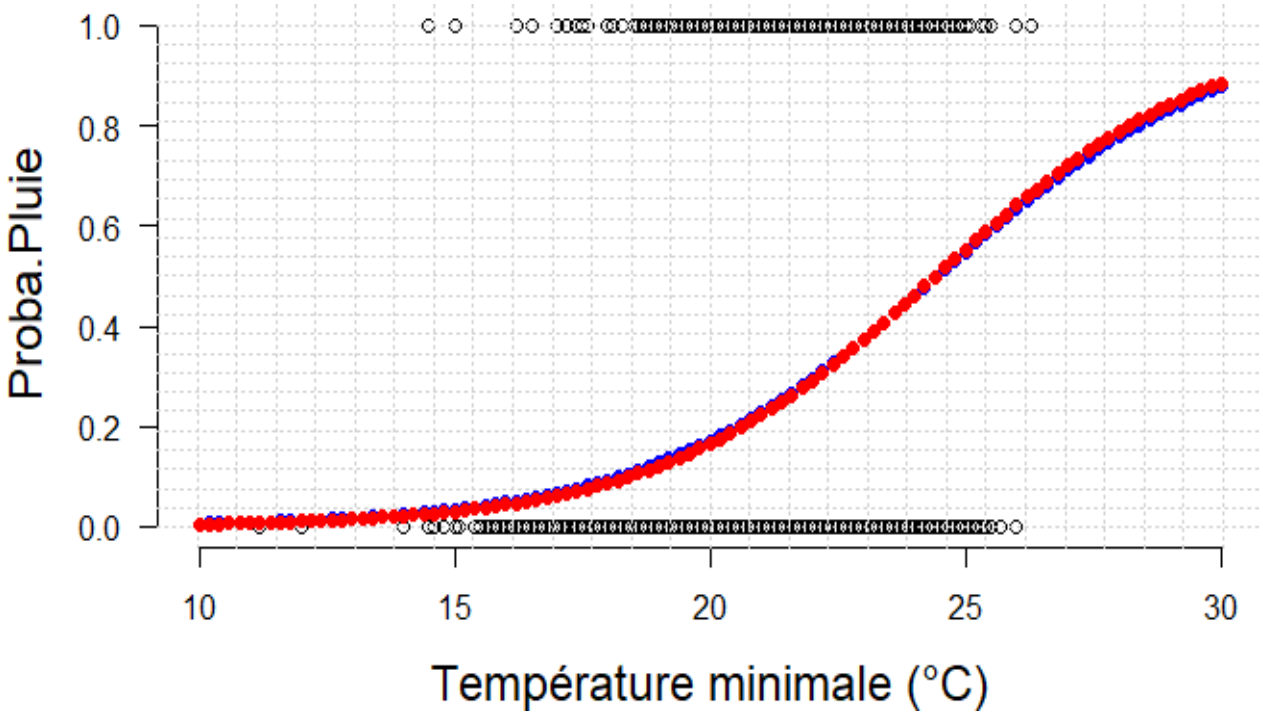


Figure 7 : *Influence de la Température Minimale (observée et ajustée) sur la pluviométrie comme en (3.1. c).*

Ainsi, en représentant graphiquement le modèle, les observations et l'estimation, nous avons la figure 7.
Le graphique démontre que le modèle de régression logistique capture efficacement la relationentre la température minimale et la probabilité de pluie. La concordance entre les observations réelles et données ajustées atteste que le modèle reproduit fidèlement les tendances présentes dans les données, tout en fournissant une estimation lissée et fiable de la probabilité de pluieen fonction de la température minimale.

D'après cette figure, la probabilité qu'il pleuve et la température minimale sont positivement corrélées. Pour toute valeur de $Tmin \leq 10°C$, la probabilité qu'il pleuve est nulle ; quand le minimum nocturne augmente de 10°C à 15°C, la probabilité qu'il pleuve reste encore très faible (elle varie de 0,01 à 0,04) ; lorsque le minimum nocturne passe de 15°C à 20°C, la probabilité qu'il pleuve augmente de 0,04 à 0,17 (un accroissement de 0,13) ; lorsque le minimum nocturne augmente de 20°C à 25°C, la probabilité d'une pluie accroit de 0,17 à 0,54 (soit une augmentation de 0,37) ; lorsque le minimum nocturne augmente de 25°C à 30°C, la probabilité d'une pluie varie de 0,54 à 0,87 (soit une augmentation de 0,33). Le minimum estdonc un bon prédicteur de la pluie pour la ville de Kinshasa. Ces observations indiquent que la température minimale est un indicateur pertinent pour prédire les précipitations à Kinshasa, avec des probabilités de pluie qui augmentent de manière notable à mesure que les températures minimales s'élèvent.

3.3. Humidité relative vs pluviomètre

Soit une variable de milieu $HrMoy$ représentant l'humidité relative, variant de 1 à 100%. Nous souhaitons modéliser la probabilité d'observer une pluie, notée $P(Rnfl)$, en fonction de HrMoy. La relation entre HrMoy et la probabilité de pluie est modélisée par une fonctionlogistique.

La fonction logistique est donnée par :

$$p(Rnfl|HrMoy) = f(HrMoy) = \frac{e^{(\mathrm{a}HrMoy+\mathrm{b})}}{1+e^{-(\mathrm{a}HrMoy+\mathrm{b})}} \quad (12)$$

Les coefficients a et b sont déterminés par la régression logistique. C'est un modèle linéairegénéralisé dans lequel on a spécifié que la fonction de lien est une fonction binomial $g(glm(Rnfl\ HrMoyi, family$ = binomial)). Le tableau 5 donne l'output de cette régression.

Tableau 5 : Coefficients de la régression logistique $Rnfl^{+}$(HrMoyi)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-9,852396	0,298999	-32,95	<2e-16
$HrMoy$	0,119826	0,003975	30,14	<2e-16

En substituant ces valeurs dans la fonction logistique, nous obtenons :

$$f(HrMoy) = \frac{1}{1+e^{-(0,119826HrMoy-9,852396)}} \quad (13)$$

La représentation graphique donne la fonction logistique (figure 8) ci-dessous :

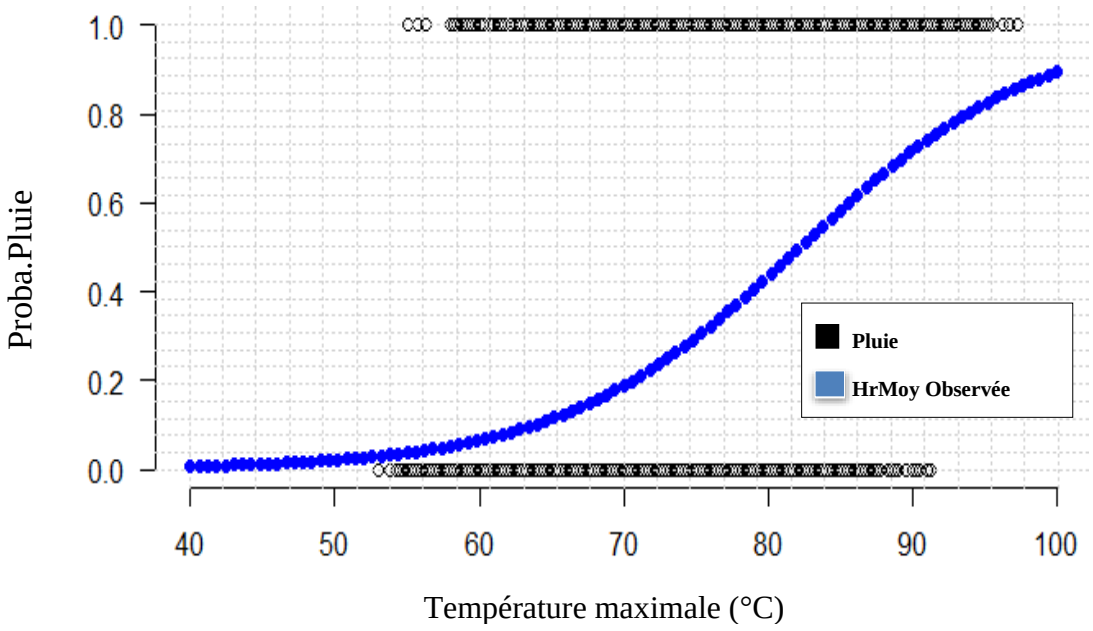


Figure 8 : *Fonction logistique Rnfl (HrMoy) observée*

Estimation du modèle

Pour trouver le modèle estimé, supposons qu'on échantillonne la présence-absence de la pluie et que chaque échantillon soit indépendant des autres. Pour une valeur $HrMoy$, la mesure suit la loi de Bernoulli de paramètre $p_{hr}(HrMoy)$. On peut simuler un tel résultat avec $rbinom(n, 1, p_{hr})$ du logiciel R.

n : nombre d'observations. L'estimation directe de la probabilité à partir de $HrMoy$ n'a pas de sens. Pour estimer les paramètres du modèle avec l'échantillon, nous avons utilisé : $glm(rbinom(n, 1, p_{hr})\ HrMoyi, family$ = binomial) et son output est résumé dans le tableau 6 ci-dessous :

Tableau 6 : Coefficients de la régression logistique $\boldsymbol{Rnfl(HrMoy)}$ du modèle estimé

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-9,928839	0,298602	-33,25	<2e-16
$HrMoy$	0,1211360	0,003973	30,55	<2e-16

Nous pouvons écrire le modèle estimé comme :

$$\log\left(\frac{p}{1-p}\right) = 0{,}1211360\text{HrMoy} - 9{,}928839 \quad (14a)$$

$$\log\left(\frac{Rnfl}{1-Rnfl}\right) = 0{,}1211360\text{HrMoy} - 9{,}928839 \quad (14b)$$

La validation du modèle estimé est réalisée à l'aide de la méthode du maximum de vraisemblance, conformément à la procédure décrite en (3.1. c). La figure 9 illustre graphiquement lemodèle, les observations, et les estimations correspondantes.

Représentation graphique du modèle, des observations et de l'estimation

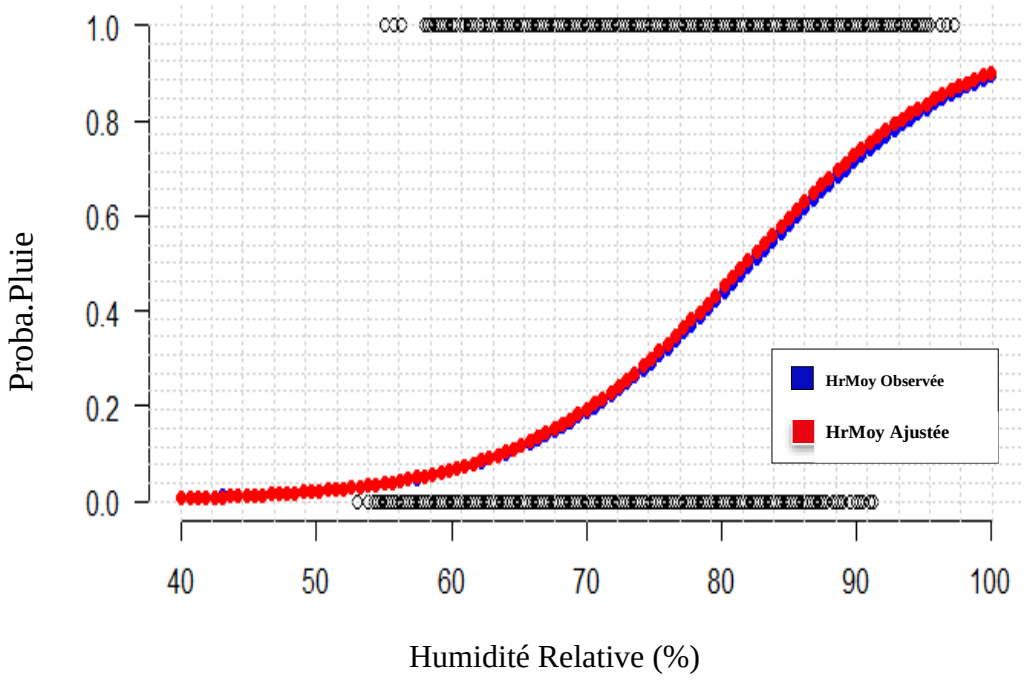


Figure 9 : *Influence de l'humidité relative (observée et ajustée) sur la pluviométrie*

La figure 9 nous renseigne que quand le taux d'humidité relative est inférieur à 50%, la probabilité qu'il pleuve est nulle. Pour une valeur d'humidité relative comprise entre (50 et 60) %, la probabilité d'une pluie est comprise entre (0,02 et 0,07), soit une augmentation de 0,05.Entre (60 et 70) %, la probabilité d'une pluie augmente de 0,12 (soit de 0,07 à 0,19). Lorsque l'humidité relative augmente de (70 à 80) %, la probabilité croit de (0,19 à 0,44), soit 0,25 deplus. Entre 80 et 90, la probabilité d'une pluie va de 0,44 à 0,73, une augmentation de 0,29. Entre (90 et 100) %, la probabilité de la pluie est comprise entre 0,73 et 0,90, une augmentationde 0,17. Par conséquent, l'humidité relative peut être considéré comme un excellent prédicteurde pluie pour le cas de la ville de Kinshasa.

L'étroite superposition entre les valeurs observée et ajustée témoigne de la capacité du modèlede régression logistique à reproduire fidèlement la relation observée entre l'humidité relative et la probabilité de pluie. Cela atteste de l'efficacité du modèle pour prédire la probabilité de précipitations en fonction de l'humidité relative.

Ainsi, l'humidité relative se révèle être un excellent indicateur de la probabilité de pluie, notamment dans le contexte de la ville de Kinshasa

3.4. Influence de la Pression Atmosphérique sur la Pluviométrie

Soit la variable à expliquer $Rnfl$, qui indique la survenue ou non de la pluie. En prenant une seule variable explicative, $PrMoy$(pression atmosphérique moyenne), nous utilisons un modèle de régression logistique simple représenté par l'équation suivante :

$$p(Rnfl|PrMoy) = f(PrMoy) = \frac{e^{(\mathrm{a}PrMoy+\mathrm{b})}}{1+e^{-(\mathrm{a}PrMoy+\mathrm{b})}} \quad (15)$$

Avec :

$P(Rnfl^{+}|PrMoy)$: est la probabilité de pluie en fonction de la pression atmosphérique.

$f(PrMoy)$: La fonction logistique.

Les coefficients (a et b) sont déterminés en faisant une régression logistique avec
$glm(rbinom(n, 1, p_{Prm})\ PrMoy, family$ = binomial) du logiciel R et son output est résumé dans le tableau 8 ci-dessous :

Tableau 8 : les coefficients de la régression logistique Rnfl(PrMoy)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	377,08177	14,26946	26,43	<2e-16
$PrMoy$	-0,38751	0,01463	-26,49	<2e-16

L'équation (15), peut s'écrire :

$$p(Rnfl|PrMoy) = \frac{1}{1+e^{-(-0{,}38751 \times PrMoy + 377{,}08177)}} \quad (16)$$

La représentation graphique de fonction de $f(PrMoy)$ donne la fonction logistique ci-dessous :

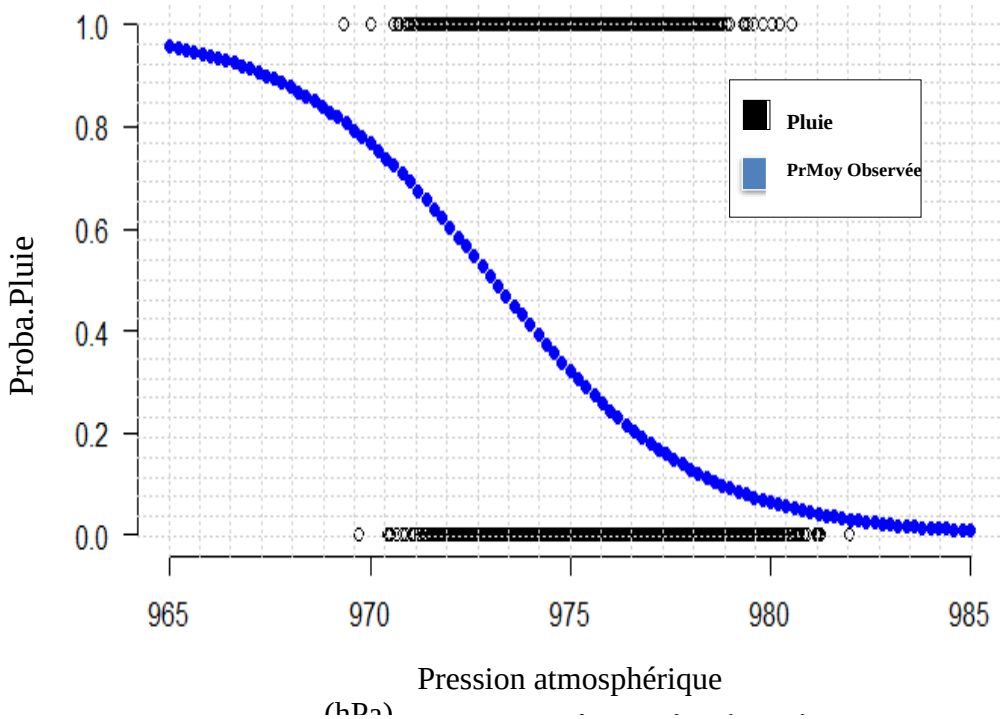


Figure 10 : *Influence de la Pression atmosphérique (observée) sur la pluviométrie*

Le graphique illustre une relation inverse bien définie entre la pression atmosphérique et laprobabilité de pluie. En d'autres termes, à mesure que la pression atmosphérique augmente,la

Estimation du modèle

Les résultats de l'estimation du modèle sont synthétisés dans le tableau ci-dessous :

Tableau 9 : les coefficients de la régression logistique Rnfl(PrMoy) du modèle estimé

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	363,48996	14,19649	25,60	<2e-16
$PrMoy$	-0,37359	0,01456	-25,67	<2e-16

Le modèle estimé peut alors s'écrire comme :

$$\log\left(\frac{p}{1-p}\right) = -0{,}37359\text{PrMoy} + 363{,}48996 \quad (17a)$$

$$\log\left(\frac{Rnfl}{1-Rnfl}\right) = -0{,}37359\text{PrMoy} - 363{,}48996 \quad (17b)$$

Pour valider le modèle estimé, nous utilisons la méthode du maximum de vraisemblance, comme détaillé dans la section (3.1. c). La figure 11 est une représentation graphique du modèle ainsi que des données observées et des prévisions, permettant de visualiser leur concordance.

Représentation graphique du modèle, des observations et de l'estimation

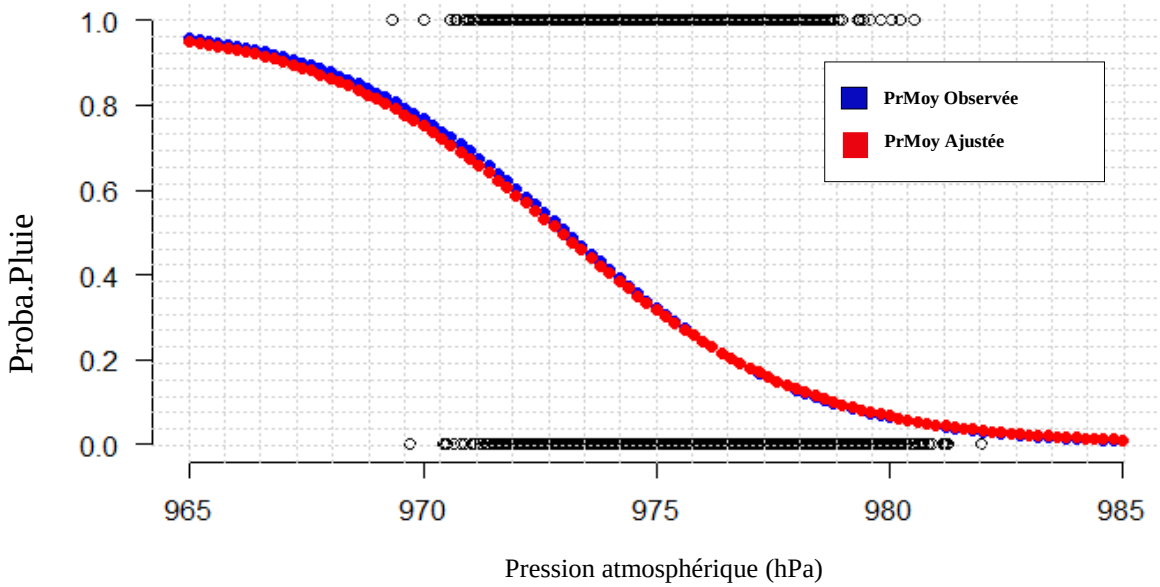


Figure 11 : Influence de la Pression Atmosphérique (observée et ajustée) sur la pluviométrie

La pression atmosphérique et la probabilité qu'il pleuve sont négativement corrélées. Pour toute valeur de $PrMoy \leq 965hPa$, la probabilité qu'il pleuve est $\geq$ 0,95 ; quand la pression atmosphérique augmente de 965 à 970 hPa, la probabilité qu'il pleuve diminue de 0,95 à 0,75 ; lorsque la pression atmosphérique augmente de 970 à 975 hPa, la probabilité qu'il pleuve diminue de 0,75 à 0,32 ; de 975 à 980 hPa, la probabilité qu'il pleuve décroit de 0,32 à 0,07 ; pour toute valeur de $PrMoy \geq 980hPa$, la probabilité qu'il pleuve est presque nulle. La pression atmosphérique influence de façon inverse la probabilité d'une pluie. Quand la pression atmosphérique diminue, la probabilité qu'il pleuve augmente.

La concordance étroite entre les courbes observée et ajustée illustre l'exactitude du modèle de régression logistique dans la reproduction de la relation entre la pression atmosphérique et la probabilité de pluie. Cela confirme l'efficacité du modèle pour prédire les précipitations en fonction de la pression atmosphérique. En un mot, la pression atmosphérique se révèle être un prédicteur fiable de la probabilité de pluie, en particulier pour la ville de Kinshasa

4. CONCLUSION

Les résultats de cette étude suggèrent que l'humidité relative, le minimum nocturne, et la pression atmosphérique jouent un rôle important dans la formation de la pluie à Kinshasa. Grâce à la régression logistique, une méthode d'analyse multivariée puissante permettant d'obtenir une quantification de l'occasion entre une pluie étudiée et chacun des facteurs l'influençant tout en tenant compte de l'effet simultané des autres facteurs, nous avons pu examiner et illustrer de combien les différents paramètres climatiques influencent la pluie à Kinshasa.

Le minima journalier et la probabilité d'une pluie sont positivement corrélés. La probabilité qu'il pleuve est grande lorsque $T_{min} \geq 25\ °C$. Il en est de même L'humidité relative et de la probabilité qu'il pleuve sont positivement corrélées aussi. Quand l'humidité relative augmente, la probabilité d'une pluie augmente, surtout quand $Hr \geq 80\%$
Tandis que la pression atmosphérique et la probabilité de pluie sont négativement corrélées. Quand la pression diminue, la probabilité d'une pluie augmente à Kinshasa, surtout quand la pression est inférieure à 965 hPa.

En revanche, nos résultats indiquent que le maximum journalier n'émerge pas comme un prédicteur efficace des précipitations, contrairement aux attentes initiales. Cette observation est illustrée par les figures 4 et 5, qui démontrent que l'effet du maximum journalier sur les précipitations est relativement marginal.

Pour de futures recherches, il serait bénéfique d'intégrer des données supplémentaires, telles que la couverture nuageuse et la vitesse de vent. Ces paramètres pourraient non seulement affiner les prévisions mais aussi offrir des perspectives nouvelles sur les mécanismes climatiques locaux.

5. REMERCIEMENTS

Les auteurs expriment leur gratitude envers la RVA (N'djili) pour la mise à disposition des données, ainsi que pour avoir accordé à l'un d'entre eux l'opportunité d'effectuer un stage en accédant à leur site.

6. RÉFÉRENCES

[1] Ahmed, S., & Karmakar, S. (1993). A linear regression model for predicting rainfall in the Ganges basin. Journal of Hydrology, 150(1-4), 215-228.
[2] Alpert, P., & Meitner, M. (2003). "Evaluation des tendances climatiques dans les régions semi-arides". Journal of Climate, 16(18).
[3] Hosmer, D. W., & Lemeshow, S. (2000). Applied logistic regression. Wiley.
[4] Deshpande, V. V. & Patil, P. V. (2012). "Analyse de la variabilité spatio-temporelle de la pluie dans le bassin de la rivière Narmada en Inde centrale". Journal of Water Resource and Protection, 4(12).
[5] Goula, B. T., & Gizaw, A. A. (2014). "Impact des facteurs météorologiques sur la pluviométrie dans la région du Bassin Omo Région du Sud de l'Éthiopie". Ethiopian Journal of Environmental Studies & Management, 7(2).
[6] Huang, C., Yang, Y., & You, H. (2018). "Influence of meteorological factors on precipitation in the North China Plain using multivariate analysis". Scientific Reports, 8(1).
[7] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning with Applications in R. Springer.
[8] Ishia, S. Z., Nyagol, K. C., & Omulo M. A. (2017). "Analyse de la variabilité spatiale de la pluviométrie dans le bassin versant de la rivière Nzoia, ouest du Kenya". Journal of Hydrology and Environment Research, 5(1).
[9] Kain, J. S., et al. (2004). The KainFritsch convective parameterization: An update. Journal of the Atmospheric Sciences, 61(13), 2726-2741.
[10] Khedr, M., & Al-Sharif, A. A. (2015). "Analyse des facteurs météorologiques influençant la pluviométrie en Jordanie". Journal of Hydrology and Hydromechanics, 63(1). Kinshasa,Kisangani and Goma
[11] (D.R.Congo)", Ann.Fac.Sci, Unikin Vol.1 11-22.
[12] Rostha Lohalo Paulo, Mushigo Nnaka Bruce, Lufuluabo Mwabila Moise, Kizungu Vumilia Roger (2022). Etude de l'évolution des paramètres Météorologiques par le Modèle Linéaire Généralisé (cas de la ville de Kinshasa), Annales Fac.Sciences et Technologies, Unikin, Vol. 1(2022) 71 – 86.
[13] Rostha Lohalo Paulo, Kizungu Vumilia Roger,

Lukanda Mwamba Vincent (2019). Variabilité temporelle des quelques paramètres météorologiques dans la ville de Goma (R. D. Congo), Annales Fac. Sciences, Unikin, Vol. 1(2019)63 – 78.

[14] Mauvilly, B. (2012). "Impact des changements climatiques sur la pluviométrie dans les pays tropicaux : cas de la ville de Kinshasa". International Journal of Hydrology Science and Technology, 2(1).

[15] Ministry of the Environment, Conservation of Nature and Tourism, (2014). Climat et Météorologie. Paramètres météorologiques dans la ville de Goma(R.D. du Congo)"Ann. Fac. Sci, Unikin, Vol.1 pp. 63-78

[16] Paulo Lohalo Rostha, Vincent Lukanda, Zalewski Maceij, Antoine F. Mulaba, (2014). Contribution to the evaluation of Relative Humidity in Boma, Matadi, Lukala, Mbanza-Ngungu, Kinshasa, Kisangani and Goma (D. R. Congo), Annales de la Faculté Des Sciences, Vol. 1, pp. 11-22

[17] Pregibon, D. (1981). Logistic regression diagnostics. Annals of Statistics,9(4), 705-724

[18] Taillet et All (2009), Dictionnaire dePhysique, de boeck, Paris.

[19] Trenberth, K. E. (2011). Changes in precipitation with climate change. Climatic Change, 105(1-2), 15-38.

[20] United Nations Development Programme, (2016). Rapport National surle Développement Humain en République Démocratique du Congo.

[21] United Nations Framework Convention on Climate Change, (2021). National