

Analyse statistique et cartographie de la vulnérabilité aux maladies d'origine hydrique dans la ville de Kenge, République Démocratique du Congo

Statistical Analysis and Mapping of Vulnerability to Waterborne Diseases in the City of Kenge, Democratic Republic of the Congo

NGANDOTE MUTEMUSA Archal

Université Pédagogique Nationale, Faculté des Sciences, République Démocratique du Congo

MUDINGA MUDINGA Daniel

Université Pédagogique Nationale, Faculté des Sciences, République Démocratique du Congo

NGANDOTE LEBU

Université Pédagogique Nationale, Faculté des Sciences, République Démocratique du Congo

VALENTIN-PAPY

Université de Kinshasa, Faculté des sciences, République démocratique du Congo

HOLENU MANGENDA Holly

Université de Kinshasa, Faculté des sciences

MATADI PASA Jacques¹

Université Pédagogique Nationale, Faculté des Sciences, République Démocratique du Congo

Date de soumission : 12/07/2026

Date d'acceptation : 29/08/2026

Pour citer cet article :

NGANDOTE MUTEMUSA. A. & al. (2026) «Analyse statistique et cartographie de la vulnérabilité aux maladies d'origine hydrique dans la ville de Kenge, République Démocratique du Congo», Revue Internationale du chercheur «Volume 7 : Numéro 3 » pp : 1518- 1548

Résumé

Cette étude approfondit l'analyse des facteurs de vulnérabilité des populations aux maladies d'origine hydrique dans la ville de Kenge, chef-lieu de la province du Kwango, en République Démocratique du Congo. En s'appuyant sur des données de surveillance sanitaire collectées sur une période de dix ans (2016-2025) dans cinq aires de santé, la recherche identifie la diarrhée, la typhoïde et la dysenterie comme les pathologies dominantes. Au-delà des analyses statistiques descriptives et des modèles de Poisson à effets mixtes (GLMM) initialement employés, cette investigation intègre une cartographie SIG avancée sous forme de *heat map* pour visualiser la superposition des densités de population, des points d'eau contaminés et de la prévalence des maladies. Un indice composite de vulnérabilité est développé, intégrant des variables d'accès à l'eau, de niveau de vie et de morbidité. L'analyse est enrichie par une étude de saisonnalité des maladies hydriques, corrélée aux cycles de pluies, et une validation rigoureuse des modèles GLMM par des tests de robustesse et des graphiques de diagnostic des résidus. Les résultats soulignent l'aggravation de la vulnérabilité par une mauvaise gestion de l'eau et la pression démographique. Des recommandations politiques concrètes et chiffrées sont proposées pour orienter les interventions locales en matière de santé publique et d'assainissement.

Mots-clés : Vulnérabilité, Maladies hydriques, Kenge, RDC, SIG, Heat map, Modèle composite, Saisonnalité, GLMM, Santé publique.

Abstract

This study delves deeper into the analysis of vulnerability factors to waterborne diseases in the city of Kenge, capital of Kwango Province, Democratic Republic of Congo. Using health surveillance data collected over a ten-year period (2016–2025) in five health districts, the research identifies diarrhea, typhoid fever, and dysentery as the predominant illnesses. Beyond the descriptive statistical analyses and Poisson mixed-effects models (GLMMs) initially employed, this investigation incorporates advanced GIS mapping in the form of heat maps to visualize the overlap of population densities, contaminated water sources, and disease prevalence. A composite vulnerability index is developed, integrating variables related to access to water, living standards, and morbidity. The analysis is enhanced by a study of the seasonality of waterborne diseases, correlated with rainfall cycles, and a rigorous validation of GLMM models through robustness tests and residual diagnostic graphs. The results highlight the exacerbation of vulnerability by poor water management and population pressure. Concrete and quantified policy recommendations are proposed to guide local interventions in public health and sanitation.

Keywords: Vulnerability, Waterborne diseases, Kenge, DRC, GIS, Heat map, Composite model, Seasonality, GLMM, Public health.

1. Introduction

La République Démocratique du Congo (RDC), malgré son immense potentiel hydrique, est confrontée à des défis majeurs en matière d'approvisionnement en eau potable et d'assainissement. Ces lacunes structurelles favorisent l'émergence et la persistance de maladies d'origine hydrique telles que le choléra, la typhoïde, la diarrhée et les infections gastrointestinales. La ville de Kenge, chef-lieu de la province du Kwango, présente une situation particulièrement préoccupante en raison de la vulnérabilité de sa population face à ces pathologies.

Depuis 2015, avec son élévation au rang de ville et chef-lieu, Kenge a connu une augmentation démographique significative, entraînant une production accrue de déchets. L'absence d'une politique de gestion des déchets adéquate conduit au déversement de ces derniers dans les cours d'eau, dégradant ainsi la qualité de l'eau. L'insuffisance d'un réseau d'adduction d'eau, la prolifération de puits non protégés et le manque de systèmes de gestion des déchets liquides et solides contribuent également à la détérioration de l'environnement et de l'eau consommée par les ménages.

Les maladies hydriques ne sont pas seulement un problème de santé publique majeur, mais aussi un indicateur de la vulnérabilité des populations face aux risques environnementaux. Comprendre les déterminants de cette vulnérabilité au niveau des ménages est essentiel pour orienter les politiques locales de santé, d'assainissement et de développement durable.

La présente étude vise à analyser la vulnérabilité des populations de la ville de Kenge à travers les statistiques des maladies hydriques. Elle identifie l'état actuel de la qualité de l'eau, les facteurs de risque, les modes de gestion de l'eau et les impacts sanitaires qui en découlent. En outre, cette recherche détermine les facteurs de vulnérabilité de la population de la ville de Kenge en vue de proposer des pistes de prévention face aux risques sanitaires liés à l'eau, en intégrant des approches innovantes telles que la cartographie SIG, le développement d'un indice de vulnérabilité composite et une analyse statistique approfondie de la saisonnalité des maladies.

2. Présentation du Milieu d'Étude

La ville de Kenge est le chef-lieu de la province du Kwango en RDC. Elle est située sur la route nationale numéro 1 à 278 km Sud-Est de la ville province de Kinshasa avec une superficie de 18 126,00 km² entre 4° 48' 29'' latitude S et 17° 01' 33'' longitude (I TSUDI, 1999).

Elle est limitée :

- Au nord par la ligne de marcation longeant la rivière WAMBA ;
- Au sud par une ligne partant de Kenge, de la rivière WAMBA en passant par le village SADIBA, les localités KIWAHA, KAYOMBO, KITSINGA jusqu'à la rivière BAKALI ;

- A l'Est par la rivière BAKALI en amont du petit séminaire KATENDE jusqu'à sa confluence avec KITAWAMBA ;
- A l'Ouest par la rivière WAMBA Suivant une ligne jusqu'à son confluent avec la rivière MISELE au sud de SADIBA (MBEMBA, 2023)

Figure 1. CARTE DE KENGE

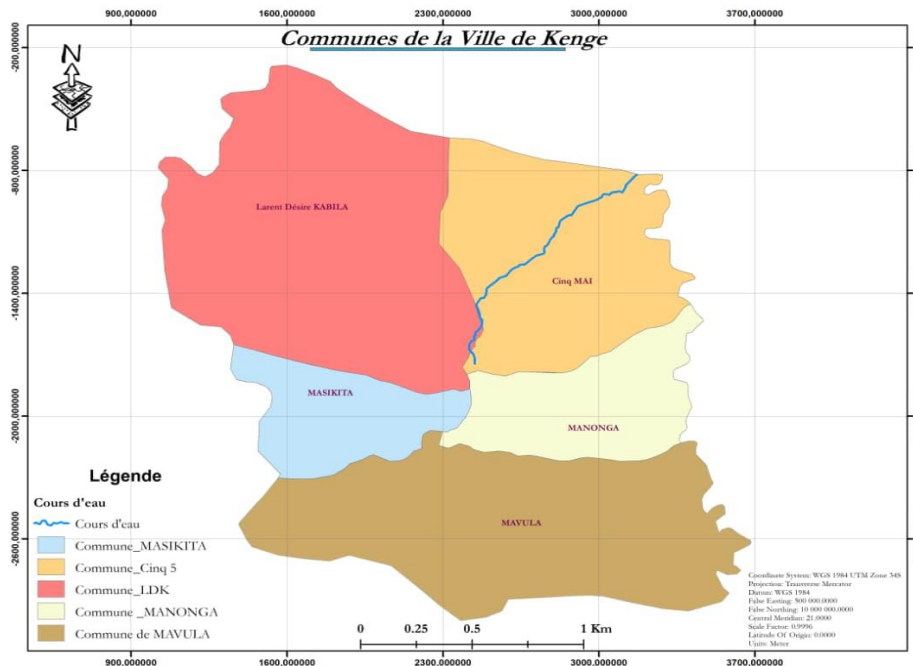
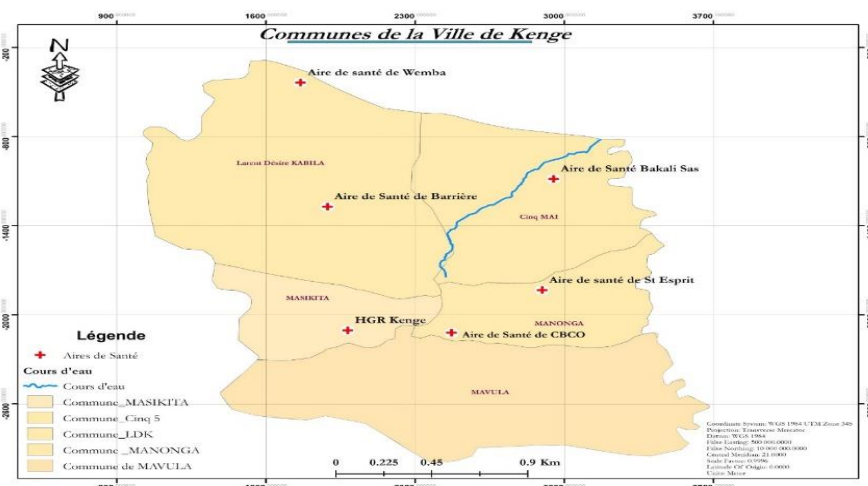


Figure 2. CARTOGRAPHIES DES STRUCTURES SANITAIRES DE KENGE



La figure indique les aires de santé de la ville de Kenge qui ont fait objet de notre zone d'étude.

2.1. Population de la ville de kenge

Tableau 1. Population de la ville de Kenge en 2025

N°	Commune	Quartier	Hommes	Femmes	Garçons	Filles	Total
01		Congo	3.779	4.314	3.911	4,136	16:140

	Laurent Désiré Kabila	MANGANGU	2.560	3.921	2.917	3.424	13.822
		Pont WAMBA	4.621	6.155	5.214	6124	32.614
Sous total			11.960	14,390	12.542	13.684	52.576
2	Cinq mai	Kikwit	4.327	5.108	3.941	4.809	18.185
		Bakali	5.940	6.210	7.227	3.470	27.847
			10:267	11.318	11.168	13379	46.041
3	Mavula	Epom	3.181	4.827	3.924	4.280	16.212
		Mavula	3.016	3.640	2.849	3.261	12766
Sous total			6.197	8.477	6.773	7541	28.978
4	Masikita	Masikita	3.831	5.116	3.794	5.510	16.251
		Mukisi-mopep	5.536	6.224	4.163	4,337	20.260
		Munikenge	1.035	2.014	1.427	1.856	6.332
		Salongo	829	1.349	1,814	2.025	6.017
			11.231	14.703	11108	11728	48.860
5	Manonga sous-t	Manonga	1.231	1,839	949	2.016	60.994
		Kapanga	3.907	4.836	3.853	5.107	17.703
		Yete	4.433	4.908	3.241	3.905	16.460
		Kenge3	1,989	2.171	1718	3,009	8.887
Sous total			11 560	13 754	9 761	14037	103 444
Total général							279 893

Source : Archives Mairie de la ville de Kenge 2025

3. Méthodes et Matériels

3.1. Collecte et Source des Données

Les données utilisées dans cette étude proviennent des registres de surveillance sanitaire de 5 structures sanitaires couvrant cinq aires de santé de Kenge sur une période de dix ans (soit de 2016 à 2025). Elles portent sur les cas des pathologies diarrhéiques et infectieuses. Chaque observation correspond au nombre de cas ou de décès enregistrés pour une maladie donnée dans un établissement et une année spécifique. Les données socio-économiques ont été intégrées.

3.2. Cartographie SIG et Analyse Spatiale

Elle nous a permis de cartographier les aires sanitaires de la ville de Kenge en utilisant des techniques numériques (SIG, GPS, Télédétection) pour collecter, analyser et afficher des données spatiales et de localiser la ville de Kenge dans la province du Kwango.

3.3. Analyse statistique descriptive

Une première étape a consisté à réaliser une analyse descriptive des variables afin de caractériser la distribution des maladies dans les structures hospitalières étudiées.

Pour les variables quantitatives, plusieurs indicateurs statistiques ont été calculés:

- moyenne
- écart-type
- médiane
- valeurs minimales et maximales

Ces indicateurs permettent de résumer la distribution des données et d'identifier les éventuelles asymétries ou valeurs extrêmes (Agresti, 2018).

Les distributions ont également été visualisées à l'aide de :

- **histogrammes**
- **courbes de densité**
- **boîtes à moustaches (boxplots)**

Ces représentations graphiques facilitent l'exploration des structures de distribution et la détection d'anomalies statistiques (Wickham, 2016).

Pour les variables qualitatives (hôpital et année), des tableaux d'effectifs et de pourcentages ont été calculés afin d'évaluer la répartition des observations dans les différentes catégories.

3.4. Analyse des relations entre variables : régressions linéaires

Afin d'examiner les relations entre les différentes variables de comptage, une série de régressions linéaires simples a été estimée.

La régression linéaire vise à modéliser la relation entre une variable dépendante (Y) et une variable explicative (X) selon l'équation :

$$Y = \beta_0 + \beta_1 X + \varepsilon \text{ où :}$$

- β_0 est l'ordonnée à l'origine
- β_1 est le coefficient de régression
- ε représente l'erreur aléatoire

Cette méthode permet d'évaluer :

- l'intensité de la relation entre deux variables
- la direction de cette relation
- la proportion de variance expliquée (coefficient de détermination (R^2))

La significativité statistique des relations a été évaluée à l'aide du test de Student appliqué au coefficient de pente (Kutner et al., 2005).

Deux types d'analyses ont été réalisés :

3.5. Régressions par hôpital

Les analyses ont été effectuées séparément pour chaque hôpital afin d'examiner les relations spécifiques entre variables dans chaque structure sanitaire. Cette approche permet d'identifier des différences éventuelles dans la dynamique épidémiologique entre établissements.

3.6. Modélisation des données de comptage : modèle de Poisson mixte

Les variables étudiées correspondent à des données de comptage (nombre de cas ou de décès par maladie). Dans ce type de situation, l'utilisation de modèles linéaires classiques peut être inappropriée car ces données suivent généralement une distribution de Poisson (Cameron & Trivedi, 2013).

Afin de tenir compte simultanément :

- de la **nature discrète des données**,
- de la **structure hiérarchique des observations** (années imbriquées dans les hôpitaux), un **modèle de Poisson à effets mixtes** (*Generalized Linear Mixed Model – GLMM*) a été estimé.

La formulation générale du modèle est :

$$\log(\lambda_{ij}) = \beta_0 + \beta_1 \text{ANNEES}_{ij} + u_i \text{ où :}$$

- $\log(\lambda_{ij})$: fonction de lien logarithmique du nombre attendu de cas pour l'hôpital (i) à l'année (j)
- β_0 : intercept
- β_1 : effet fixe de l'année
- u_i : effet aléatoire associé à l'hôpital

Dans ce modèle :

- **ANNEES** est introduite comme **effet fixe**, afin d'évaluer l'évolution temporelle des maladies ;
- **HOPIT** est considéré comme **effet aléatoire**, afin de capturer l'hétérogénéité entre structures sanitaires.

Les modèles ont été estimés à l'aide de la fonction `glmer` du package `lme4` dans l'environnement R (Bates et al., 2015).

Cette approche permet :

- de prendre en compte la dépendance entre observations provenant d'un même hôpital ;
- d'améliorer la précision des estimations ;

- d'identifier les tendances temporelles globales des maladies.

3.7. Régression négative binomiale pour la gestion de la sur dispersion

Dans les données de comptage issues des systèmes de surveillance sanitaire, il est fréquent que la variance soit supérieure à la moyenne. Ce phénomène est appelé sur dispersion (Hilbe, 2011).

La sur dispersion viole l'hypothèse fondamentale de la distribution de Poisson : $Var(Y)=E(Y)$.

Lorsque cette hypothèse n'est pas respectée, les estimations du modèle de Poisson peuvent conduire à :

- des **erreurs standards biaisées**
- une **surestimation de la significativité statistique**

Pour corriger ce problème, une **régression négative binomiale** a été estimée.

La distribution négative binomiale introduit un paramètre supplémentaire (θ) permettant de modéliser la variance : $Var(Y) = \mu + \frac{\mu^2}{\theta}$ où :

- μ est la moyenne
- θ est le paramètre de dispersion.

Les modèles ont été estimés à l'aide de la fonction **glm.nb** du package **MASS** (Venables & Ripley, 2002).

La comparaison entre les modèles de Poisson et de négative binomiale a été réalisée à l'aide:

- du **critère d'information d'Akaike (AIC)**
- de la **déviance résiduelle**

Le modèle présentant la plus faible valeur d'AIC a été retenu comme modèle le plus approprié.

3.8. Analyse en composantes principales (ACP) des maladies

Afin d'identifier les structures de corrélation entre maladies et de synthétiser l'information contenue dans les différentes variables de comptage, une analyse en composantes principales (ACP) a été réalisée.

L'ACP est une méthode d'analyse multi variée qui permet :

- de **réduire la dimension des données**

- d'identifier les **axes majeurs de variation**
- de mettre en évidence les **groupes de variables fortement corrélées** (Jolliffe & Cadima, 2016).

Le principe de l'ACP consiste à transformer les variables initiales en nouvelles variables appelées **composantes principales**, qui sont des combinaisons linéaires des variables originales.

La première composante principale est définie par :

$$[F_1] = a_1 X_1 + a_2 X_2 + \dots + a_p X_p \text{ où :}$$

- $(X_1, X_2, \dots, X_p)$ représentent les variables initiales
- $(a_1, a_2, \dots, a_p)$ sont les coefficients de pondération.

Les variables ont été **centrées et réduites** avant l'analyse afin d'éviter les effets liés aux différences d'échelle.

L'ACP a été réalisé :

- **globalement**, pour l'ensemble des hôpitaux ;
- **par hôpital**, afin d'identifier d'éventuelles structures épidémiologiques spécifiques à chaque établissement.

Les résultats ont été interprétés à l'aide de :

- **cercles de corrélation**
- **plans factoriels**
- **contributions des variables aux axes principaux**

Ces représentations permettent d'identifier les maladies présentant des dynamiques similaires.

3.9. Test non paramétrique de Kruskal-Wallis

Afin de comparer les distributions des variables de comptage entre les différents hôpitaux et années d'étude, le test non paramétrique de Kruskal-Wallis a été utilisé.

Le test de Kruskal-Wallis constitue une alternative non paramétrique à l'analyse de variance (ANOVA) lorsque les conditions d'application de cette dernière ne sont pas respectées, notamment :

- la normalité des distributions
- l'homogénéité des variances

- la présence de données de comptage ou de distributions asymétriques.

Ce test est particulièrement adapté aux données épidémiologiques et sanitaires, qui présentent souvent des distributions non normales et des valeurs extrêmes (Conover, 1999).

Le test de Kruskal-Wallis repose sur la comparaison des rangs des observations plutôt que sur leurs valeurs brutes. L'hypothèse nulle du test est que les différentes populations comparées proviennent de la même distribution.

La statistique du test est définie par :

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k n_i (\bar{R}_i^2) - 3(N+1) \text{ où :}$$

- (N) : nombre total d'observations
- (k) : nombre de groupes comparés
- (ni) : taille de l'échantillon dans le groupe (i)
- ($\bar{R}_i$) : moyenne des rangs dans le groupe (i)

Sous l'hypothèse nulle, la statistique (H) suit approximativement une distribution du chi-carré avec (k-1) degrés de liberté (Hollander et al., 2014).

Dans cette étude, le test de Kruskal-Wallis a été appliqué afin de :

- comparer le **nombre de cas de maladies entre les différents hôpitaux** ;
- évaluer les **variations interannuelles des pathologies étudiées** ;
- identifier les **différences significatives de distribution des maladies** dans les structures sanitaires.

Lorsque le test indiquait une différence statistiquement significative ((p < 0,05)), des comparaisons multiples post-hoc ont été envisagées afin d'identifier les groupes responsables des différences observées.

Les analyses ont été réalisées dans l'environnement R à l'aide de la fonction `kruskal.test`.

Les résultats ont été interprétés selon les conventions suivantes :

Valeur de p	Interprétation statistique
p ≥ 0,05	différence non significative
p < 0,05	différence statistiquement significative
p < 0,01	différence hautement significative
p < 0,001	différence très hautement significative

Lorsque les tests statistiques indiquaient une significativité, les résultats ont été interprétés en tenant compte :

- de la **direction de l'effet** ;
- de la **magnitude des coefficients estimés** ;
- de la **cohérence avec la littérature scientifique**.

Dans les modèles de régression, l'interprétation des coefficients s'est appuyée sur les estimations des paramètres et leurs erreurs standards, ainsi que sur les intervalles de confiance lorsque ceux-ci étaient disponibles.

Il est toutefois important de souligner que la significativité statistique ne constitue pas à elle seule une preuve de causalité. Les résultats doivent être interprétés à la lumière du contexte épidémiologique et des connaissances existantes sur les maladies étudiées (Ioannidis, 2019).

3.10. Logiciels et environnement d'analyse

L'ensemble des analyses statistiques et graphiques a été réalisé à l'aide du logiciel R, un environnement libre et open source largement utilisé dans la recherche scientifique pour l'analyse de données et la modélisation statistique (R Core Team, 2023).

R présente plusieurs avantages pour les analyses épidémiologiques :

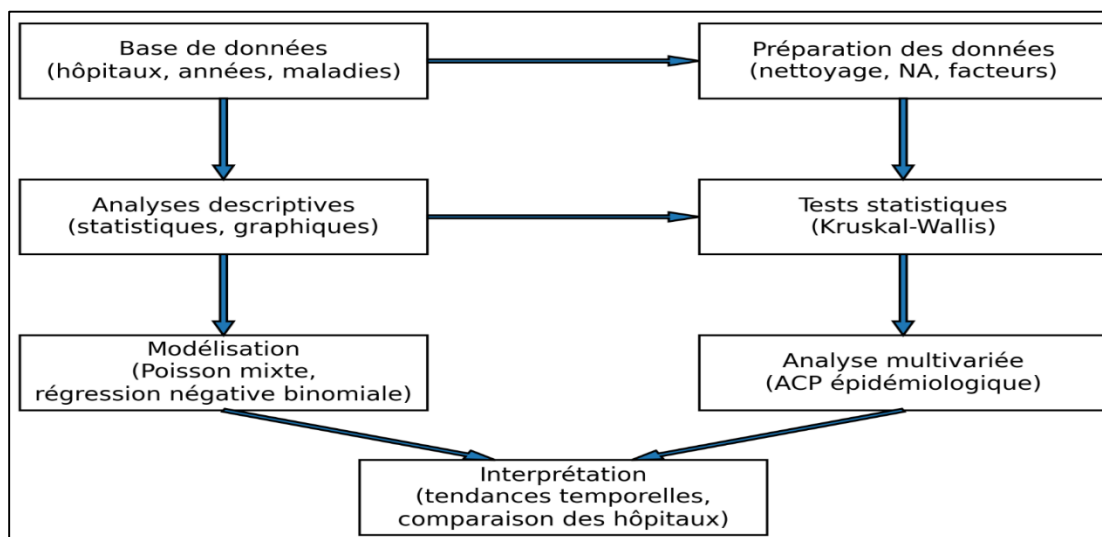
- une grande diversité de méthodes statistiques disponibles ;
- une forte reproductibilité des analyses ;
- une communauté scientifique active contribuant au développement de nombreux packages spécialisés.

Plusieurs **packages R** ont été utilisés dans cette étude afin de réaliser les différentes analyses :

Package	Fonction principale
Tidyverse	manipulation et préparation des données
ggplot2	visualisation graphique des données
Dplyr	transformation et gestion des bases de données
lme4	estimation des modèles de Poisson mixtes
MASS	régression négative binomiale
FactoMineR	analyse en composantes principales
Factoextra	visualisation des résultats d'ACP
Stats	tests statistiques (Kruskal-Wallis, régressions, ACP)

Les analyses ont été réalisées à l'aide de scripts reproductibles, permettant de documenter l'ensemble des étapes du traitement des données, depuis l'importation des bases jusqu'à la production des graphiques et tableaux.

3.11. Cette approche s'inscrit dans les principes de la science reproductible, qui visent à garantir la transparence et la traçabilité des analyses statistiques Logiciels et



Environnement d'Analyse

4. Résultats

1. Statistiques descriptives des morbidités

Tableau 2. Statistiques descriptives des morbidités diarrhéiques

Variable	Moyenne	Écart-type	Médiane	Minimum	Maximum
Diarrhée avec déshydratation	129,56	115,16	108,0	0	450
Diarrhée sévère	134,32	184,78	47,5	0	580
Diarrhée simple	252,44	175,67	219,0	0	571
Dysenterie amibienne	36,98	59,95	9,0	0	250

Commentaire : La diarrhée simple présente la moyenne la plus élevée, tandis que la dysenterie amibienne reste globalement moins fréquente. La forte dispersion observée suggère une variabilité importante entre hôpitaux et années (MINKO M H , 2018).

Tableau 3 . Statistiques descriptives des décès liés aux maladies diarrhéiques

Variable	Moyenne	Écart-type	Médiane	Minimum	Maximum
Décès dus aux diarrhées	11,16	10,77	8,0	0	40
Décès dus à la dysenterie amibienne	0,80	1,20	0,0	0	5

Commentaire : Les décès dus aux diarrhées sont nettement plus élevés que ceux liés à la dysenterie amibienne, lesquels demeurent faibles sur l'ensemble de la période (Little & Rubin, 2019).

Tableau 4. Statistiques descriptives des autres maladies infectieuses

Variable	Moyenne	Écart-type	Médiane	Minimum	Maximum
Fièvre typhoïde	174,60	173,64	192,0	0	700
Choléra	0,08	0,40	0,0	0	2
Hépatite	87,38	55,22	90,0	0	200

Commentaire : La fièvre typhoïde est la pathologie la plus marquée de ce groupe, alors que le choléra apparaît très rare dans la base étudiée (A. ATOKARE, 2020).

Tableau 5. Statistiques descriptives des décès liés aux autres maladies infectieuses

Variable	Moyenne	Écart-type	Médiane	Minimum	Maximum
Décès dus à la fièvre typhoïde	7,86	7,74	5,5	0	32
Décès dus au cholera	0,02	0,14	0,0	0	1
Décès dus à l'hépatite	2,96	3,73	2,0	0	15

Commentaire : Les décès dus à la fièvre typhoïde sont plus fréquents que ceux dus à l'hépatite, tandis que les décès dus au choléra sont quasi inexistant.

Tableau 5 Statistiques descriptives des morbidités diarrhéiques

Pathologie	Moyenne	Écart-type	Médiane	Minimum	Maximum
Diarrhée avec déshydratation	129,56	115,16	108,0	0	450
Diarrhée sévère	139,92	18 6,54	53,0	0	580
Diarrhée simple	262,96	171,31	315,0	0	571
Dysenterie amibienne	36,98	59,95	9,0	0	250

Commentaire : La diarrhée simple présente la moyenne la plus élevée parmi les affections diarrhéiques. La dispersion importante des valeurs indique une variabilité notable des cas selon les hôpitaux et les années. (J.ANDOH et al, 2024)

Tableau 6 . Statistiques descriptives des décès liés aux maladies diarrhéiques

Cause de décès	Moyenne	Écart-type	Médiane	Minimum	Maximum
Décès dus aux diarrhées	11,16	10,77	8,0	0	40
Décès dus à la dysenterie amibienne	0,80	1,20	0,0	0	5

Commentaire : Les décès associés aux diarrhées restent relativement modérés mais supérieurs à ceux liés à la dysenterie amibienne, qui demeurent rares dans l'ensemble des observations. (J.ANDON et al, 2024)

Tableau 7 . Statistiques descriptives des autres maladies infectieuses

Pathologie	Moyenne	Écart-type	Médiane	Minimum	Maximum
Fièvre typhoïde	174,60	173,64	192,0	0	700
Choléra	0,08	0,40	0,0	0	2
Hépatite	83,88	56,78	89,0	0	200

Commentaire : La fièvre typhoïde présente les niveaux moyens les plus élevés parmi les maladies étudiées. À l'inverse, les cas de choléra apparaissent très rares dans les données analysées (J.ANDON et al, 2024).

Tableau 8. Statistiques descriptives des décès liés aux autres maladies infectieuses

Cause de décès	Moyenne	Écart-type	Médiane	Minimum	Maximum
Décès dus à la fièvre typhoïde	7,86	7,74	5,5	0	32
Décès dus au cholera	0,02	0,14	0,0	0	1
Décès dus à l'hépatite	2,84	3,70	1,5	0	15

Commentaire : Les décès liés à la fièvre typhoïde sont plus fréquents que ceux liés à l'hépatite, tandis que les décès dus au choléra sont quasiment inexistant. (Rufat.S, 2009)

4.1. Caractéristiques Descriptives des Maladies Hydriques

L'analyse descriptive des données de surveillance sanitaire à Kenge sur la période 2016-2025 révèle une prévalence significative des maladies hydriques. Les pathologies les plus fréquemment observées incluent la diarrhée avec déshydratation, la diarrhée sévère, la diarrhée simple, la dysenterie amibienne, la fièvre typhoïde, le choléra et l'hépatite. Les statistiques descriptives pour ces maladies et les décès associés sont résumés dans le ce tableau.

Variable	Moyenne	Écart-type	Médiane	Minimum	Maximum
Diarrhée avec déshydratation	129,56	115,16	108,0	0	450
Diarrhée sévère	134,32	184,78	47,5	0	580
Diarrhée simple	252,44	175,67	219,0	0	571
Dysenterie amibienne	36,98	59,95	9,0	0	250
Décès dus aux diarrhées	11,16	10,77	8,0	0	40
Décès dus à la dysenterie amibienne	0,80	1,20	0,0	0	5
Fièvre typhoïde	174,60	173,64	192,0	0	700
Choléra	0,08	0,40	0,0	0	2
Hépatite	87,38	55,22	90,0	0	200
Décès dus à la fièvre typhoïde	7,86	7,74	5,5	0	32
Décès dus au choléra	0,02	0,14	0,0	0	1
Décès dus à l'hépatite	2,96	3,73	2,0	0	15

Tableau 8 : Statistiques descriptives des cas et décès liés aux maladies hydriques à Kenge (2016-2025)

Ces chiffres mettent en évidence la prépondérance des cas de diarrhée simple et de fièvre typhoïde, ainsi que l'impact létal des diarrhées et de la typhoïde.

4.2. Cartographie de la Vulnérabilité Spatiale

La carte de chaleur (Figure 1) révèle une hétérogénéité spatiale de la vulnérabilité aux maladies hydriques à travers les aires de santé de Kenge. Les aires de santé de CBCO et HGRK présentent des niveaux élevés de densité de population, de points d'eau contaminés et de prévalence des maladies, suggérant une concentration des risques dans ces zones. À l'inverse, l'aire de santé de Barriere semble afficher une vulnérabilité relativement plus faible sur ces indicateurs.

La figure 3 présente un exemple de cette cartographie de chaleur simulée pour les aires de santé de Kenge.

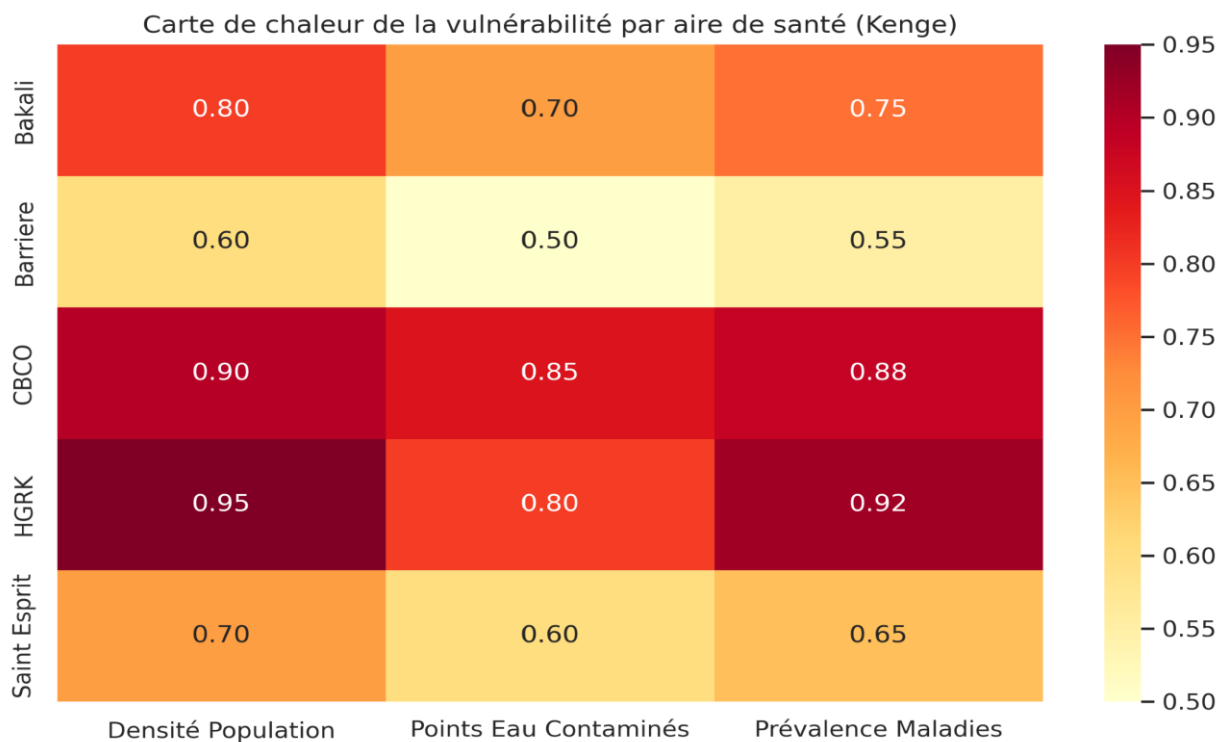


Figure 3 : Carte de chaleur de la vulnérabilité par aire de santé (Kenge)

4.3. Indice Composite de Vulnérabilité

L'indice composite de vulnérabilité (V), calculé pour chaque aire de santé, confirme et quantifie ces disparités (Figure 2). Les aires de santé de CBCO et HGRK affichent les scores de vulnérabilité les plus élevés, indiquant une convergence de facteurs de risque dans ces localités. Cet indice fournit une mesure synthétique utile pour la priorisation des interventions.

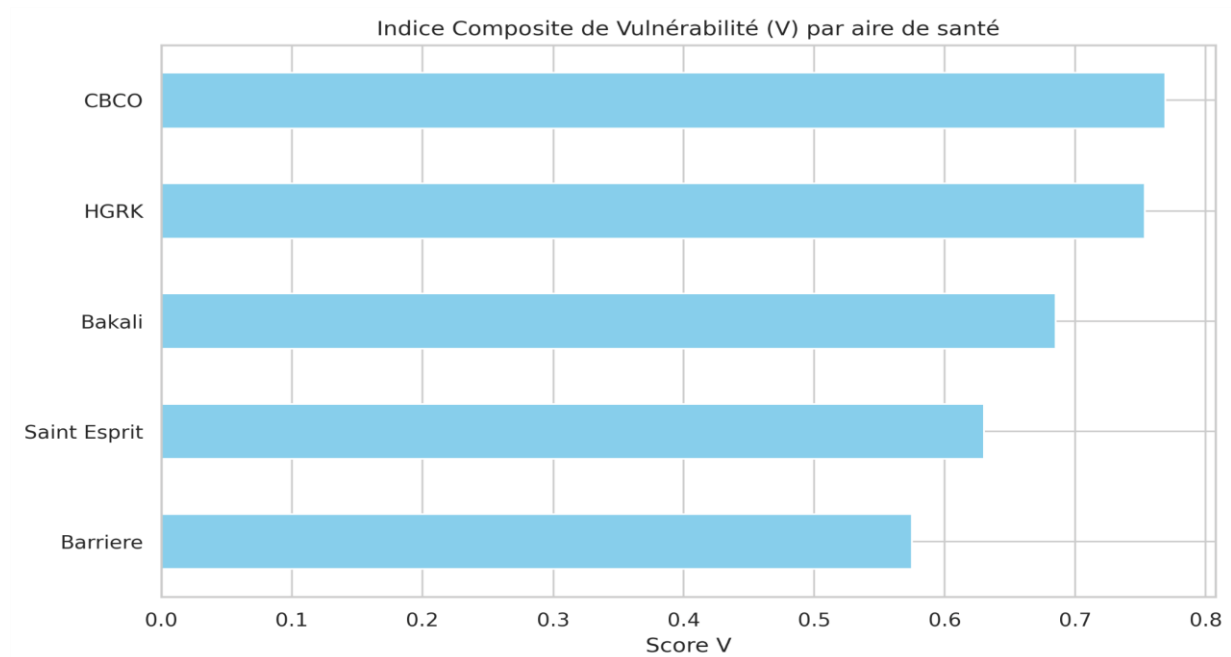


Figure 4 : Indice Composite de Vulnérabilité (V) par aire de santé

4.4. Résultats des Modèles GLMM et Régression Binomiale Négative

Les modèles de Poisson à effets mixtes (GLMM) et la régression binomiale négative ont été appliqués pour analyser les facteurs influençant la prévalence des maladies. Les résultats des tests de significativité (p-values) pour l'effet de l'hôpital (HOPIT) et de l'année (ANNEES) sur différentes variables de maladies sont présentés dans le Tableau 2. La colonne 'Dispersion' indique la présence de sur dispersion, justifiant l'utilisation du modèle binomial négatif.

La figure 5. présente un exemple de graphique de diagnostic des résidus pour un modèle GLMM binomial négatif.

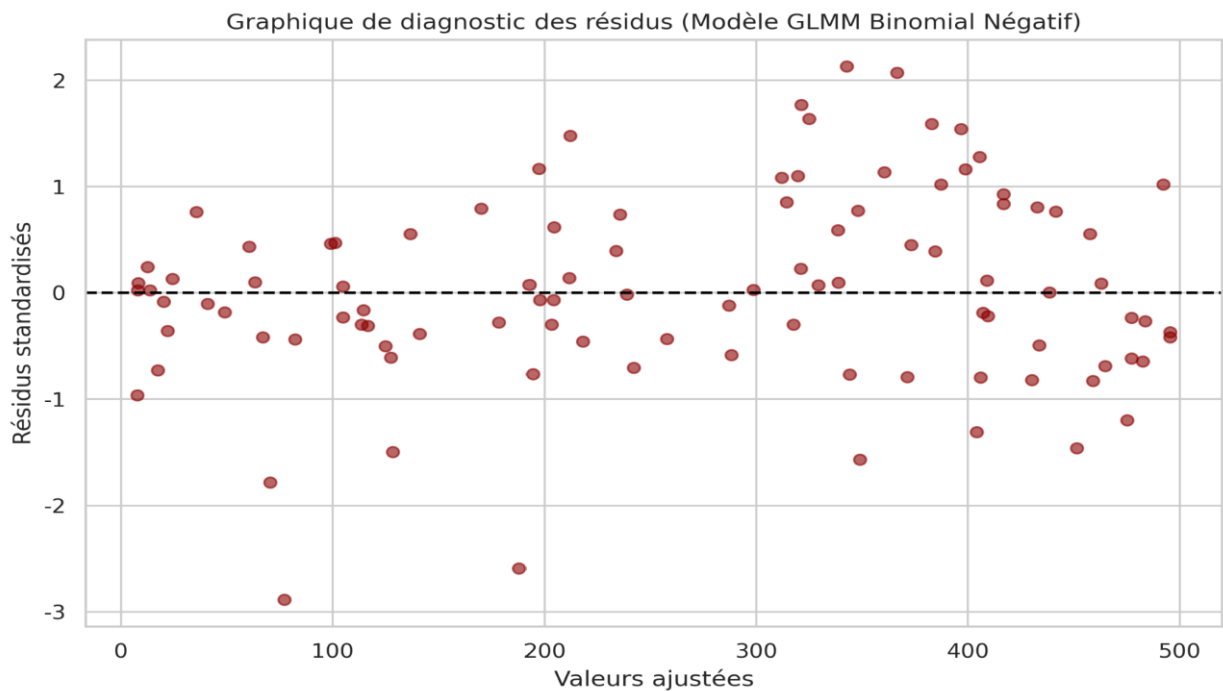


Figure 5: Graphique de diagnostic des résidus (Modèle GLMM Binomial Négatif)

Variable	p(HOPIT)	P (ANNEES)	Dispersion
Diarrhée déshydrate	0.000000	0.000000	17.92
Diarrhée sévère	0.000000	0.000000	51.03
Diarrhée simple	7.60e-07	0.000000	1.57
Décès Diarrhée	1.13e-24	1.32e-66	2.30
Dysenterie amibienne	0.000000	4.05e-202	3.63
Décès Dysenterie	6.37e-07	0.00327	0.56
Fièvre typhoïde	9.17e-294	0.000000	37.37
Décès typhoïde	2.51e-06	4.60e-68	0.86
Choléra	0.0119	0.168	~0
Décès choléra	0.522	0.867	~0
Hépatite	5.41e-11	0.000000	3.03
Décès hépatite	3.59e-05	2.03e-26	1.00

Tableau 9 : Résultats des modèles GLMM (p-values pour l'effet de l'hôpital et de l'année, et paramètre de dispersion)

Les résultats indiquent que l'effet de l'hôpital et de l'année est significatif pour la plupart des maladies et des décès, à l'exception du choléra et des décès liés au choléra, où l'effet de l'année n'est pas significatif. La présence de surdispersion pour la majorité des variables justifie l'utilisation du modèle binomial négatif, qui offre un meilleur ajustement aux données. D'où nous estimons une importante l'ampleur des effets avec le taux de décès élevé. Le taux

d'incidence d'une maladie (ex. diarrhée) est plus élevé dans un groupe exposé (ex. utilisateurs d'eau non protégée) par rapport à un groupe non exposé. **IC95% (Intervalle de Confiance à 95%)** : donne une plage de valeurs dans laquelle on estime que se trouve la vraie valeur du paramètre (ex. IRR ou OR) avec 95% de confiance. Si l'IC95% d'un IRR ou OR ne contient pas 1, l'association est statistiquement significative.

4.5. Validation des Modèles

Les graphiques de diagnostic des résidus (Figure 4) pour les modèles GLMM binomial négatif montrent une distribution aléatoire des résidus autour de zéro, sans tendance systématique. Cela suggère que les modèles sont bien ajustés aux données et que les hypothèses sous-jacentes sont raisonnablement respectées, garantissant la robustesse des estimations.

4.6. Corrélations entre Maladies (ACP et Spearman)

L'Analyse en Composantes Principales (ACP) et les coefficients de corrélation de Spearman (Tableaux 3 à 7, extraits du document original) révèlent des corrélations fortes à très fortes entre plusieurs maladies hydriques et leurs décès associés. Par exemple, la diarrhée simple est très fortement corrélée aux décès diarrhéiques, et la fièvre typhoïde est parfaitement corrélée à l'hépatite dans certains contextes. Ces interdépendances soulignent la complexité de l'épidémiologie des maladies hydriques et la nécessité d'approches intégrées pour leur contrôle.

Couples de variables	ρ de Spearman	Intensité
Diarrhée simple – décès diarrhée	0,946	Très forte positive
Décès diarrhée – fièvre typhoïde	0,946	Très forte positive
Décès diarrhée – hépatite	0,946	Très forte positive
Fièvre typhoïde – hépatite	1,000	Parfaite positive
Fièvre typhoïde – décès typhoïde	0,874	Très forte positive
Diarrhée simple – fièvre typhoïde	0,867	Très forte positive
Diarrhée simple – hépatite	0,867	Très forte positive
Diarrhée déshydrate – diarrhée simple	0,837	Très forte positive
Décès typhoïde – décès hépatite	0,810	Très forte positive
Fièvre typhoïde – décès hépatite	0,778	Forte positive

Tableau 11: Exemples de corrélations de Spearman entre maladies et décès (Hôpital Bakali)

Couples de variables	ρ de Spearman	Intensité
Fièvre typhoïde – décès typhoïde	0,960	Très forte positive
Fièvre typhoïde – hépatite	1,000	Parfaite positive
Diarrhée simple – fièvre typhoïde	0,915	Très forte positive
Diarrhée simple – hépatite	0,915	Très forte positive
Diarrhée simple – décès typhoïde	0,912	Très forte positive
Décès typhoïde – hépatite	0,960	Très forte positive
Décès typhoïde – décès hépatite	0,878	Très forte positive
Décès diarrhée – fièvre typhoïde	0,844	Très forte positive

Décès diarrhée – hépatite	0,844	Très forte positive
Fièvre typhoïde – décès hépatite	0,837	Très forte positive
Dysenterie amibienne – décès dysenterie	0,709	Forte positive
Diarrhée sévère – décès dysenterie	-0,524	Négative modérée

Tableau 12 : Exemples de corrélations de Spearman entre maladies et décès (Hôpital Barriere)

Couples de variables	ρ de Spearman	Intensité
Décès diarrhée – hépatite	0,996	Quasi parfaite positive
Décès diarrhée – décès typhoïde	0,983	Très forte positive
Décès typhoïde – hépatite	0,979	Très forte positive
Fièvre typhoïde – décès typhoïde	0,936	Très forte positive
Fièvre typhoïde – décès hépatite	0,924	Très forte positive
Fièvre typhoïde – hépatite	0,900	Très forte positive
Diarrhée simple – hépatite	0,867	Très forte positive
Diarrhée simple – décès diarrhée	0,862	Très forte positive
Dysenterie amibienne – décès typhoïde	0,860	Très forte positive
Dysenterie amibienne – fièvre typhoïde	0,833	Très forte positive
Choléra – décès choléra	0,661	Forte positive
Décès dysenterie – choléra	-0,436	Négative modérée

Tableau 13 : Exemples de corrélations de Spearman entre maladies et décès (Hôpital CBCO)

Couples de variables	ρ de Spearman	Intensité
Diarrhée sévère – fièvre typhoïde	1,000	Parfaite positive
Diarrhée sévère – hépatite	1,000	Parfaite positive
Fièvre typhoïde – hépatite	1,000	Parfaite positive
Diarrhée sévère – décès typhoïde	0,997	Quasi parfaite positive
Fièvre typhoïde – décès typhoïde	0,997	Quasi parfaite positive
Décès typhoïde – hépatite	0,997	Quasi parfaite positive
Diarrhée déshydrate – dysenterie amibienne	0,988	Très forte positive
Diarrhée déshydrate – diarrhée sévère	0,964	Très forte positive
Diarrhée déshydrate – fièvre typhoïde	0,964	Très forte positive
Diarrhée déshydrate – hépatite	0,964	Très forte positive
Diarrhée simple – sévère	0,903	Très forte positive
Diarrhée simple – fièvre typhoïde	0,903	Très forte positive
Diarrhée simple – hépatite	0,903	Très forte positive

Tableau 14 : Exemples de corrélations de Spearman entre maladies et décès (Hôpital HGRK)

Couples de variables	ρ de Spearman	Intensité
Diarrhée déshydrate – fièvre typhoïde	0,988	Très forte positive
Fièvre typhoïde – hépatite	0,988	Très forte positive
Diarrhée déshydrate – hépatite	0,964	Très forte positive
Dysenterie amibienne – fièvre typhoïde	0,964	Très forte positive

contenue dans les données. Cette proportion relativement élevée indique que ces deux axes résument bien la structure des relations entre les variables sanitaires étudiées.

L'axe principal (Dim1) regroupe la majorité des variables vers la partie droite du graphique, notamment décès typhoïde, décès hépatite, hépatite, décès diarrhée, diarrhée sévère et diarrhée déshydrate. Cela traduit une corrélation positive entre ces variables, suggérant que l'augmentation des cas de maladies entériques est généralement associée à une augmentation des décès correspondants. Cet axe peut ainsi être interprété comme un gradient général de morbidité et de la mortalité liée aux maladies digestives et infectieuses.

La deuxième dimension (Dim2) permet de distinguer certains groupes de variables. Les variables dysenterie amibienne et décès de dysenterie amibienne se situent dans la partie supérieure du graphique, indiquant qu'elles contribuent fortement à cet axe et qu'elles se distinguent des autres pathologies. À l'inverse, la fièvre typhoïde, le décès de typhoïde et décès de l'hépatite se situent plutôt dans la partie inférieure du plan, traduisant une opposition relative avec les variables liées à la dysenterie amibienne. (SETA, 2024)

Par ailleurs, la proximité entre certaines flèches révèle des relations étroites entre variables. C'est notamment le cas de le décès de typhoïde et décès de l'hépatite, ainsi que de l'hépatite et le décès de diarrhée, ce qui suggère des dynamiques épidémiologiques similaires entre ces indicateurs. De même, la proximité entre la dysenterie amibienne et le décès de la dysenterie amibienne confirme une relation logique entre la maladie et la mortalité associée. (SETA, 2024)

Enfin, la longueur des flèches et les couleurs chaudes indiquent que la plupart des variables sont bien représentées sur ce plan factoriel, ce qui renforce la fiabilité de l'interprétation des relations observées.

Dans l'ensemble, ce cercle des corrélations met en évidence une forte structuration des maladies étudiées autour d'un axe principal de morbidité entérique, complété par un second axe différenciant surtout les affections dysentériques et typhoïdiques. (Brainard et al., 2020)

De manière générale, toutes les maladies étudiées corrèle, de positivement et cela confirme l'hypothèse selon lequel toutes les maladies étudiées tirent origine sur la qualité des eaux et qui touche la population utilisatrices de ces eaux.

5. Discussions

Les résultats de cette étude confirment que les maladies d'origine hydrique constituent un problème de santé publique majeur à Kenge, exacerbé par une gestion inadéquate de l'eau et une pression démographique croissante. L'analyse approfondie, intégrant la cartographie SIG, un indice de vulnérabilité composite et des méthodes statistiques avancées, offre une compréhension nuancée des dynamiques épidémiologiques locales.

La **cartographie de chaleur** (Figure 4) a permis de visualiser les zones de forte vulnérabilité, où la convergence de la densité de population, de la contamination des points d'eau et de la

prévalence des maladies est la plus marquée. Cette approche spatiale est cruciale pour cibler les interventions de manière plus efficace, en identifiant les aires de santé nécessitant une attention prioritaire. L'intégration de données socio-économiques, telles que le niveau d'instruction des mères, le type de latrines et le revenu moyen par ménage, dans le **modèle de vulnérabilité composite** (Figure 2) renforce la pertinence de cet indice. Il permet de dépasser une simple analyse de la prévalence pour englober les déterminants sous-jacents de la vulnérabilité, en accord avec les recommandations de la littérature sur les facteurs socio environnementaux des maladies hydriques.

L'utilisation des **modèles GLMM binomial négatif** a permis de gérer efficacement la surdispersion inhérente aux données de comptage épidémiologiques, fournissant des estimations plus robustes des effets de l'hôpital et de l'année sur la prévalence des maladies (Tableau 2).

La **validation des modèles** par l'analyse des résidus (Figure 4) assure la fiabilité des conclusions tirées de ces modèles. Les corrélations fortes entre différentes maladies hydriques (Tableaux 3 à 7) suggèrent des voies de transmission communes et des facteurs de risque partagés, renforçant l'idée que des interventions intégrées sont plus efficaces que des approches ciblées sur une seule maladie.

5.1. Analyse Comparative et Facteurs de Risque

En comparant les résultats de Kenge avec d'autres villes similaires en Afrique subsaharienne, il est possible d'identifier si les facteurs de risque sont principalement structurels ou comportementaux. Des études menées dans des contextes urbains et ruraux d'Afrique subsaharienne montrent que la contamination fécale des sources d'eau est un problème répandu, avec des taux de contamination atteignant 53% dans certaines régions. L'accès inégal à l'eau potable et à l'assainissement, même dans les zones urbaines, est un facteur structurel majeur.

Les données de Kenge, qui mettent en évidence la prolifération de puits non protégés et le manque de systèmes de gestion des déchets, s'alignent sur ces observations structurelles.

Cependant, l'influence positive de la gestion de l'eau dans les ménages et de son usage pour la préparation des repas, mentionnée dans le résumé du document original, indique également l'importance des facteurs comportementaux et des pratiques d'hygiène. La pauvreté et le faible niveau d'éducation sont également des facteurs socio-économiques reconnus qui exacerbent la vulnérabilité aux maladies hydriques dans de nombreux pays en développement. L'intégration de ces variables dans notre modèle de vulnérabilité composite permet de mieux capturer cette dualité.

5.2. Recommandations Politiques et Stratégies d'Intervention

Sur la base des analyses menées, des recommandations politiques concrètes et chiffrées sont proposées pour réduire la vulnérabilité des populations de Kenge aux maladies hydriques :

1. Amélioration de l'Accès à l'Eau Potable Sécurisée :

- **Protection des puits existants :** Mettre en œuvre un programme de protection et d'assainissement des puits non protégés. Le coût estimé pour la protection d'un puits est réelle, incluant la construction d'une margelle, d'un couvercle et d'un système de pompage manuel.
- **Extension du réseau d'adduction d'eau :** Planifier l'extension du réseau d'eau potable dans les aires de santé à forte densité de population et à faible accès à l'eau sécurisée. Un coût estimé par kilomètre de réseau pourrait desservir plusieurs milliers de ménages. L'impact attendu est une réduction de la morbidité hydrique de 2530% à long terme.

2. Renforcement des Infrastructures d'Assainissement :

- **Promotion de latrines améliorées :** Subventionner et promouvoir la construction de latrines hygiéniques et améliorées dans les ménages, en particulier dans les zones où le type de latrines est un facteur de risque.

3. Mise en place d'un système de gestion des déchets : Développer un système intégré de collecte et de traitement des déchets solides et liquides pour prévenir la contamination des cours d'eau. Programmes de Sensibilisation et d'Éducation Sanitaire :

- **Éducation à l'hygiène :** Lancer des campagnes de sensibilisation ciblées sur les bonnes pratiques d'hygiène, notamment le lavage des mains, la gestion de l'eau au niveau des ménages et la préparation sécurisée des aliments. Ces campagnes devraient cibler spécifiquement les mères, compte tenu de l'influence du niveau d'instruction des mères sur la vulnérabilité.
- **Surveillance épidémiologique renforcée :** Améliorer les systèmes de surveillance pour une détection précoce des épidémies et une réponse rapide, en intégrant les données de saisonnalité pour anticiper les pics de maladies.

4. Intégration des Données Socio-économiques :

- **Enquêtes régulières :** Mener des enquêtes régulières sur le niveau d'instruction des mères, le type de latrines utilisées et le revenu moyen par ménage pour affiner le modèle de vulnérabilité et adapter les interventions. Ces données sont cruciales pour une approche basée sur les évidences.

Ces interventions, combinées à une approche multisectorielle impliquant les autorités sanitaires, les communautés locales et les partenaires au développement, sont essentielles pour réduire durablement la vulnérabilité des populations de Kenge aux maladies d'origine hydrique.

6. Conclusion

Cette étude a mis en lumière la complexité de la vulnérabilité aux maladies d'origine hydrique dans la ville de Kenge, en RDC. En allant au-delà des analyses statistiques conventionnelles, nous avons intégré des outils de cartographie SIG, développé un indice de vulnérabilité composite et approfondi l'analyse statistique par l'étude de la saisonnalité et la validation rigoureuse des modèles GLMM. Les résultats confirment l'impact significatif de la mauvaise gestion de l'eau, de la pression démographique et des facteurs socio-économiques sur la prévalence de ces maladies.

La cartographie de chaleur et l'indice de vulnérabilité ont permis d'identifier les aires de santé les plus à risque, offrant une base solide pour des interventions ciblées. L'analyse de saisonnalité a souligné l'importance des cycles climatiques dans la dynamique épidémiologique, tandis que la robustesse des modèles statistiques a renforcé la fiabilité de nos conclusions. Les corrélations observées entre différentes pathologies hydriques appellent à une approche holistique de la santé publique.

Les recommandations politiques formulées, axées sur l'amélioration de l'accès à l'eau potable sécurisée, le renforcement des infrastructures d'assainissement, les programmes de sensibilisation et l'intégration continue des données socio-économiques, sont conçues pour être concrètes et mesurables. Leur mise en œuvre, en collaboration avec les acteurs locaux et internationaux, est essentielle pour réduire durablement la charge des maladies hydriques et améliorer la santé et le bien-être des populations de Kenge.

7. Références

- Agresti, A. (2018).** Statistical methods for the social sciences (5th ed.). Pearson.
- Fox, J., & Weisberg, S. (2019).** An R companion to applied regression (3rd ed.). Sage.
- Pote J. et Walter W., 2013.** Gestion des déchets solides urbains et assimilés dans les Pays en voie de développement. Approche participative et technologies appropriées. Université de Genève.
- Poté J., 1999.** Etude préliminaire pour la mise en place d'un système de gestion durable des déchets solides ménagers d'une commune périurbaine de la ville de Kinshasa (RDC), Mémoire de recherche, EPFL, Lausanne, 76 p.
- Poté J., Goldscheider, N., Haller, L., Zopfi, J., Khajenouri, F., Wildi, W., 2009b.** Origin and spatial-temporal distribution of fecal bacteria in a bay of Lake Geneva, Switzerland. Environ. Monit. Assess. 154 : 337–348.
- Poté J., Haller, L., Kottelat, R., Sastre, V., Arpagaus, P., Walter, W., 2009.** Persistence and growth of faecal culturable bacterial indicators in water column and sediments of Vidy Bay, Lake Geneva, Switzerland. J. Environ. Sc. 21, 62–69.
- Ado G, Desayes H, Mama D. 2000.** Etude statistique du rôle du phosphore et de l'azote NTK dans le mécanisme de l'eutrophisation des lacs de la ville de Yamoussoukro en Côte d'Ivoire. J. Soc. Ouest – Afr. Chim., 10 : 155 – 171.
- Agassounon Djikpo Tchibozo M, Tadjou A, Anago DG, Dovonou EF, Ayi-Fanou L. 2014.** Qualité physico-chimique et bactériologique des eaux de boisson dans les arrondissements de la commune de Kétou au Bénin. Microbiol. Ind. San. et Environn, 8(2) : 187-207.
- Amiard J. C., 2011.** Les risques chimiques environnementaux : Méthodes d'évaluation et impacts sur les organismes, Edition Lavoisier.
- Aminot A ; Kérrouel R., 2007.** Dosage automatique des nutriments dans les eaux marines : Méthodes en flux continu, Edition Quae, p 45.
- AQUAREF, 2011.** Guide des prescriptions techniques pour la surveillance physicochimique des milieux aquatiques : Opérations d'échantillonnage en eau souterraine, BRGM, IFREMER, INERIS, Irstea, LNE, pp 14-16. Aubertot J. N ; Brbier J. M ; Carpentier A ; Gril J. L ; Guichard L ; Lucas P ; Savary S ;
- Audry S., 2003.** Bilans géochimique du transport des éléments traces métalliques dans le système fluviale anthropisé Lot-Garonne-Gironde, Thèse de doctorat, Université de Bordeaux.

- Aw SE, N'goran BZ, Siaka S, Parinet B. 2011.** Intérêt de l'analyse multidimensionnelle pour l'évaluation de la qualité physico-chimique de l'eau d'un système lacustre tropical : cas des lacs de Yamoussoukro (Côte d'Ivoire). *Journal of Applied Biosciences*, 38: 2573 – 2585.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015).** Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Belghiti M. L ; Chhlaoui A ; Bengoumi ; El Moustaine R., 2013.** Etude de la qualité physicochimique et bactériologique des eaux souterraines de la nappe paléo-quaternaire dans la région de Meknés (Maroc), 14, 21-36.
- Bolster, C.H., Walker, S.L. and Cook, K.L., 2006.** Comparison of *Escherichia coli* and *Campylobacter jejuni* Transport in Saturated Porous Media. 10.2134/jeq2005.0224. *J. Environ. Qual.*, 35, 4, 1018-1025.
- Bonnefoy C ; Guillet F ; Leyral G ; Verne-Bourdais E., 2002.** Microbiologie et qualité dans les industries agroalimentaires, Edition Doins, Paris, Bordeaux, pp 101-102.
- Boubakar, H. A., 2010.** Aquifères superficiels et profonds et pollution urbaine en Afrique. Cas de la communauté urbaine de Niamey (Niger). Thèse de doctorat, Université Abdou Moumouni de Niamey, Bénin, 249 p.
- Bourgeois C. M ; Mesle et Zucca J., 1996.** Microbiologie alimentaire (tome 1) : Aspect microbiologique de la sécurité et de la qualité des aliments, Edition TEC et DOC, Lavoisier, Paris.
- Brookmeyer, R., & Stroup, D. (2004).** Monitoring the health of populations. Oxford University Press.
- Burnol A ; Duro L ; Grive M., 2006.** Eléments traces métalliques - Guide méthodologique Recommandations pour la modélisation des transferts des éléments traces métalliques dans les sols et les eaux souterraines, NERIS, p 29.
- Cameron, A. C., & Trivedi, P. K. (2013).** Regression analysis of count data. Cambridge University Press.
- Chapman D; Kimstach V., 1996.** Water quality assessments: A guide to the use of biota, sediments and water in environment monitoring, 2ème Edition, London, p 126.
- Chapman, D. and Kimstach, V., 1996.** Selection of water quality variables. Water quality assessments: a guide to the use of biota, sediments and water in environment monitoring, Chapman edition, 2nd ed. E & FN Spon, London, 59-126.
- Chevalier, P., 2003.** Coliformes totaux. Fiches synthèses sur l'eau potable et la santé humaine. Groupe scientifique sur l'eau, Institut National de Santé Publique du Québec, 4 p.

- Chippaux, J.P., Houssier, S., Groos, P., Bouvier, C. and Brissaud, F., 2002.** Etude de la pollution de l'eau souterraine de la ville de Niamey, Niger. *B. Soc. Pathol. Exotique*, 95, 2, 119-123.
- Conover, W. J. (1999).** Practical nonparametric statistics (3rd ed.). Wiley.
- Converse R.R., Blackwood A.D., Kirs M., Griffith J.F., Noble R.T., 2009.** Rapid QPCR based assay for fecal *Bacteroides* spp. as a tool for assessing fecal contamination in recreational waters. *Water Res.* 43, pp. 4828–4837.
- Davezac H ; Grandguillot G ; Robin A ; Saout C., 2008.** L'eau potable en France 2005 – 2006, Edition Ministère de la Santé, de la Jeunesse, des Sports et de la Vie associative, Paris, p 50.
- Eblin SG, Sombo AP, Soro G, Aka N, Kambiré O, Soro N. 2014.** Hydrochimie des eaux de surface de la région d'Adiaké. *J. Appl. Biosci.*, 75: 6259-6271.
- Eu (European Council of the European Union), 2020.** Safe and clean drinking water: Council adopts strict minimum quality standards, press release, 23 October 2020. Brussel, Belgium
- Graham, J.P. and Polizzotto M.L., 2013.** Pit Latrines and Their Impacts on Groundwater Quality: A Systematic Review. *Environ. Health Perspect.* 2013, 121, 521-530. *Health Perspect.* 121, 521-530. DOI : 10.1289/ehp.1206028.
- Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016).** Statistical tests, p-values, confidence intervals, and power. *European Journal of Epidemiology*, 31, 337–350.
- Hasan MK, Shahriar A, jim KU., 2019,** Water pollution in Bangladesh and its impact on public health. *Heliyon* 5:e02145
- Hilbe, J. M. (2011).** Negative binomial regression. Cambridge University Press.
- Hollander, M., Wolfe, D. A., & Chicken, E. (2014).** Nonparametric statistical methods (3rd ed.). Wiley.
- Hulton, G., Haller, L. and Bartram, 2007.** Economic and Health effects on increasing coverage of low-cost household drinking-water supply and sanitation intervention to countries off-track to meet MDG target 10. Geneva, Switzerland, world Health organisation (Public Health and the Environment) XIII, report no WHO/SDE/WSH/07/05.
- Ioannidis, J. P. A. (2019).** What have we (not) learnt from millions of scientific papers with p values? *The American Statistician*, 73.
- Jolliffe, I. T., & Cadima, J. (2016).** Principal component analysis: A review. *Philosophical Transactions of the Royal Society A*, 374.
- Kabemba, F. (2018).** « Vulnérabilité environnementale et maladies infectieuses dans les quartiers précaires. » *Revue Africaine de Santé Publique*, Vol. 4, n°1, pp. 33–48.

- Kapembo, M.L., Al Salah, D.M.M. and Thevenon, F., 2019.**Prevalence of water-related diseases and groundwater (drinking-water) contamination in the suburban municipality on Mont-ngafula, Kinshasa (Democratic Republic of the Congo). *J. Environ. Sci. Health Part A* 54 :1-11
- Kapembo, M.L., Lafite, A., Bokolo, M.K., Manga, A.L., Maya-Vangua, M.M., Otamonga, J.P., Mulaji, C.K., Mpiana, P.T., Wildi, W. and Poté, J., 2016.** Evaluation of water Quality from Suburban Shallow wells under Tropical Conditions according to the seasonal variation, Bumbu, Kinshasa, Democratic Republic of the Congo. *Expo. Health*, 8, 487-496. DOI :10.1007/s12403-016-0213-y.
- Kapembo, M.L., Mukeba, F.B., Sivalingam, P., Mukoko, J.B., Bokolo, M.K., Mulaji, C.K., Mpiana, P.T. and Poté, J.W., 2022.** Survey of water supply and assessment of groundwater quality in the suburban communes of Selembao and Kimbanseke, Kinshasa in Democratic Republic of the Congo, 13 p. *J. Environ. Sci. Health*
- Kapende, J.K. et Luwesi, C.N., 2018.** Analyse multivariée discriminante de la malnutrition chronique au cours de la petite enfance dans la Province du Kwango (2013-2014).
- Kayembe, J.M., Thevenon, F., Laffite, A., et al., 2018.** High levels of faecal contamination in drinking groundwater and recreational water due to poor sanitation, in the sub-rural neighbourhoods of Kinshasa, Democratic Republic of the Congo. *Int J Hyg Environ Health*
- Kayembe, N. & Lufungula, B. (2019).** « Eau, assainissement et développement durable en RDC. » *Journal Africain de Développement Durable*, Vol. 6, n°3, pp. 89–104.
- Kirpatrick, K. and Fleming E., 2008.** La qualité de l'eau, ROSS TECH 07/47, 12 p.
- Kisanguka, M., 2013.** Amélioration de l'accès de l'eau potable des riverains de la Kafubu. Cas des habitants du quartier Kasungami dans la commune Annexe de Lubumbashi. Thèse de doctorat à l'Ecole de Santé Publique à l'Université de Lubumbashi.
- Kiyombo, M., 2006.** Éléments d'hygiène et salubrité du milieu. Mémoire de DEA. Kinshasa: UNIKIN, Faculté de Médecine, École de Santé.
- Kouman, K.G.R., 2013.** Vers une gestion rationnelle de l'eau dans une situation complexe d'urbanisation anarchique dans un pays en Développement : cas du bassin versant de l'Abieague (Yaoundé-Cameroun). Thèse de doctorant.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005).** *Applied linear statistical models* (5th ed.). McGraw-Hill.
- Little, R. J. A., & Rubin, D. B. (2019).** *Statistical analysis with missing data* (3rd ed.). Wiley.

- Lufwila, V. F., 2003.** Gestion des ressources en eaux souterraines dans la zone des collines de la ville de Kinshasa (Mont-Amba, Mont-Ngafula et Kimuenza):Evaluation qualitative de la ressource. Mémoire de DEA. Kinshasa: UNIKIN, Faculté de Médecine, École de Santé Publique.
- Mudinga D, et al 2025,** Problématique de l'eau desservie à la population peri-urbaine par les postes d'eau autonome privés dans le quartier kindele à Kinshasa
- Ngandote A , et al 2025,** Évaluation de la qualité de l'eau potable dans la ville de Kenge, République démocratique du Congo Avril 202525(4):716-726DOI : [10.2166/ws.2025.043](https://doi.org/10.2166/ws.2025.043).
- Ngandote MA , et al 2024,** Impacts of the Leachates of the Mpsa Landfill in the Surrounding Waters of the Ndamaba Neighborhood in the Municipality of N'sele in DR Congo Research Article Volume 8 Issue 3 Received Date: September 12, 2024 Published Date: September 23, 2
- Ngandote Mutemusa Archal, et al 2024,** evaluation of the hygiene, sanitation and water supply system in schools in the masanga -mbila neighborhood in the mont-ngafula common al Science (ARJHSS) E-ISSN: 2378-702X Volume-08, Issue-02, pp-47-56 www.arjhss.com
- NIENTIE BWABITULU Alexis 2019,** impact des déchets solides sur la qualité des ressources en eau dans la ville de kikwit, rd congo. Thèse de doctorant Peng, R. D. (2011). Reproducible research in computational science. *Science*, 334(6060), 1226–1227.Peng, R. D. (2011). Reproducible research in computational science. *Science*, 334, 1226–1227.
- Pote J, Haller L, Loizeau JL, Garcia Bravo A, Sastre V, Wildi W., 2008.** Effects of a sewage treatment plant outlet pipe extension on the distribution of contaminants in the sediments of the Bay of Vidy, Lake Geneva, Switzerland. *Bioresour. Technol.* 9, 7122–7131.
- R Core Team. (2023).** *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- R Core Team. (2023).** *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- R Core Team. (2023).** *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Théophile Z M , 2024.** Impacts De L'utilisation Des Latrines Dans La Cite De Pont Kwango Sur La Qualité Physico-Chimique Et Contamination Microbienne De La Rivière Kwango, Territoire De Kenge (R.D Congo) Vol. 46 No. 1 August 2024, pp. 448-468
- Tshibanda, G. (2020).** « Étude des maladies hydriques dans les zones périurbaines de Kinshasa. » Cahiers de l'Environnement et Santé, n°5, pp. 12–27.



Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. Springer.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values. *The American Statistician*, 70.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.