Brazilian Journal of Biometrics, **42**, 289-306 (2024)
doi:10.28951/bjb.v42i3.703

BRAZILIAN JOURNAL OF BIOMETRICS
ISSN:2764-5290

ARTICLE

A comprehensive statistical analysis of Malaria dynamics in the Adamawa region of Cameroon, from 2018 to 2022[1]

Apollinaire Batoure Bamana[*1], Ezekiel Dangbe[1], Hamadjam Abboubakar[1] and Mahdi Shafiee Kamalabad[2]

[1]Department of Computer Engineering, University Institute of Technology, The University of Ngaoundere, Cameroon
[2]Department of Methodology and Statistics, Faculty of Social and Behavioural Sciences, Utrecht University, The Netherlands
[*]Corresponding author. Email: apollinaire.batoure@univ-ndere.cm

(Received: November 10, 2023; Revised: February 07, 2024; Accepted: February 27, 2024; Published: August 30, 2024)

Abstract
Malaria remains a prominent public health concern in Cameroon, with the potential for epidemic outbreaks, necessitating a robust understanding of its dynamics. This paper uses routinely collected surveillance data from health facilities in the Adamawa Region since January 2018. By applying statistical analysis, this study aims to enhance comprehension, enable data predictions, and facilitate informed decision-making for public health policy implementation. Focusing on weekly health districts data spanning from 2018 to 2022, our analysis employs key statistical metrics for central tendency, data spread, distribution shape, and variable dependence. The study reveals distinctive trends, highlighting peak malaria transmission periods consistently occurring between August and November each year. The highest weekly recorded case count in any health district reached 1,294. The data exhibits leptokurtic distributions, skewed to the left of the median. And in 2022, 11% of the population was reported to have contracted malaria. Despite an overall region-wide average growth rate of -1.21% over the past five years, maintaining vigilant attention to this critical health issue is imperative. Auto dependence analysis indicates that observations are weekly correlated, assuming the time series as stationary. The stationarity has been confirmed by ADF and KPSS tests that we performed. This comprehensive data analysis helps our understanding of the malaria landscape in the Adamawa Region of Cameroon. The paper also recommends the inclusion of additional variables in data collection for a more holistic perspective. These findings provide a basis for the formulation and implementation of targeted interventions by relevant stakeholders, aiding in the prediction of future cases and ultimately contributing to the effective management of malaria in the region.

Keywords: Malaria Time Series, Statistical Analysis, Progression Rate, Prevalence Rate, Adamawa Region of Cameroon.

1. Introduction

In this section we present the context of the study carried out and relevant related works.

1.1 Context

In our study, we turn our attention to a disease of significant global concern: malaria (2024 ICD-10-CM Diagnosis Code B54[2]). This tropical affliction continues to exact a grievous toll, claiming numerous lives within households each year, with a particularly profound impact on the African continent (Wahedi, J. A., et al. 2020; Ibrahim O. R., 2021; Ozulonye, O. S., Okolo,

[2] https://www.icd10data.com/ICD10CM/Codes/A00-B99/B50-B64/B54-/B54

A., Torsen, E. & Tiwah O. J., 2022). According to WHO's World malaria report for 2022[3], there were an estimated 247 million malaria cases worldwide in 2021, an increase from 245 million in 2020. The WHO African Region, with an estimated 234 million cases in 2021, accounted for about 95% of global cases. More than 600,000 people still die of malaria every year, most of them children (Ibrahim O. R. et al. 2021; Danwang et al., 2021; Ramirez, J. H. et al. 2020). Approximately 3.2 billion individuals across the globe face a heightened risk of contracting this ailment. These statistics underscore the persistent and substantial public health concern posed by this disease, necessitating dedicated attention from various stakeholders (Ramirez, J. H. et al. 2020).

Malaria is classified as potentially epidemiological in Cameroon, among thirty others diseases (Sohanang Nodem F. S., Ymele D., Fadimatou M., Fodouop S. C., 2023, Mbouna, A. D., et al, 2019) This classification means that special attention must be paid to them when implementing public health policies. In order to closely monitor these diseases, the Cameroon Ministry of Public Health, through its Health Information Unit, collects data from Health Facilities, to measure and monitor the evolution of the diseases. These data sets, structured as temporal sequences, offer a fertile ground for the application of specialized analytical methodologies. Consequently, this work is primarily motivated by the pronounced potential for robust investigation within this context.

The paper focuses on health data taken as time series. Time series are data collected chronologically over an approved period (Hyndman, R. J. & Athanasopoulos, G., 2021). They have long been used in economics, finance and meteorology (Makridakis, S., Assimakopoulos, V. & Spiliotis, E., 2018); but less so in epidemiology, especially in Africa. We therefore plan to carry out several studies, in connection with Data Science (Zhang Q., 2021). The final aim is to gain a better understanding of how these data behave over time, so that appropriate prevention measures can be taken. The work should lead to recommendations on the use of time series in epidemics monitoring and prediction in Africa. But first, we considered it useful to describe the data statistically. We therefore turned to a comprehensive descriptive statistics and rates.

The data we are working on includes cases of malaria in Health Districts of the Adamawa Region of Cameroon, in Central Africa. This study area encompasses 11 Health Districts, 88 Health Areas and 207 Health Units, for a total population estimated at 1,509,210 people in 2022, covering an area of 63,701 km^2. The 11 districts are, in alphabetical order: Bankim, Banyo, Belel, Dang, Djohong, Meiganga, Ngaoundal, Ngaoundere Rural (Ndere_Rur), Ngaoundere Urbain (Ndere_Urb), Tibati and Tignère. Ngaoundere is the regional capital and is divided into two Health Districts (urban and rural)[4]. The data are collected weekly via the DHIS[5] (District Health Information Software) platform. Concerning Cameroon, data are collected weekly, since January 2018 and saved in the DHIS data base. These datasets will serve as the foundation for our statistical analysis, incorporating key indicators such as measures of concentration, dispersion, shape, dependence, as well as the assessment of progression and prevalence rates.

1.2 Related works

Numerous studies have focused on statistically describing malaria data most of time in view of applying preventive measures, or for analyzing and forecasting.

The work of Landoh et al., 2012, sought to assess the trends of malaria incidence and mortality due to malaria in Est Mono district (Togo) from 2005 to 2010. Data on confirmed and suspected malaria cases reported were obtained from the district health information system. From January 2005 to December 2010, 114,654 malaria cases (annual mean 19,109 ± 6,622) were reported with an increase of all malaria cases from 10,299 in 2005 to 26,678 cases in 2010 ($p<0.001$). Of the 114,654 malaria cases 52,539 (45.8%) were confirmed cases. The prevalence

[3] World malaria report 2022. Geneva: World Health Organization, https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2022
[4] Ministère de la Santé Publique, Carte sanitaire programmatique, Janvier 2023, https://www.minsante.cm/site/?q=fr/content/carte-sanitaire-programmatique-officiel-janv-23-v3-1
[5] https://dhis-minsante-cm.org

of confirmed malaria cases increased from 23.1 per 1,000 in 2005 to 257.5 per 1,000 population in 2010 ($p < 0.001$). The mortality rate decreased from 7.2 per 10,000 in 2005 to 3.6 per 10,000 in 2010 ($p < 0.001$), with a significant reduction of 43.9% of annual number of death due to malaria. This study showed an increase of malaria prevalence despite the implementation of the use of relevant strategies.

Alhassan, E. A., Adjei, M. I., Aidoo, E., 2017 undertaken their research work with the prior motivation to develop an adequate model for forecasting future trends of malaria in the Kasena Nankana Municipality (Ghana). Descriptively, the study revealed that, the average number of patients diagnosed with malaria is 698.7 having slightly flat tail at right side (positively skewed) which implies that the malaria cases are heading towards more positive values with the value of Kurtosis being less than 3 hence making them not normally distributed (platykurtic) which means the variables exhibit broad peaks or high kurtosis. The model was used to forecast monthly cases of malaria for the next two years.

In their works, Rodríguez, S. N. I., Rodríguez, J. A. I., Rodríguez, J. C. P. & Olivera, M. J., 2021 describe the malaria mortality rates from 2009-2018 in Colombia. During the study, 148 malaria-related deaths were registered. The average annual mortality rate was 0.032 deaths/100,000. Two peaks were observed in 2010 and 2016. The unstable downward trend of malaria mortality rates calls for greater emphasis on surveillance and interventions.

Dian et al., 2021 in their study aimed to analyse trends of malaria cases in urban Kuala Lumpur (Malaysia). All suspected cases presented to a university hospital in Kuala Lumpur from January 2005 to December 2020 were examined by microscopy. Infection status was analysed using descriptive statistics and curve estimation analysis. Of 3,105 blood films examined, 92 (3%) were microscopically confirmed malaria cases. Plasmodium vivax infections accounted for the majority (36.9%) of all malaria cases. The curve estimation analysis showed significant decreases in malaria cases due to P. vivax ($R2 = 0.598$; $p < 0.001$) and Plasmodium falciparum ($R2 = 0.298$, $p = 0.029$), but increases for Plasmodium knowlesi ($R2 = 0.325$, $p = 0.021$) during the 16 years. This study highlighted the importance of continued vigilance and improved surveillance.

Finally, the work of Danwang et al. 2021 aims to provide a fine-scale spatiotemporal estimate of malaria incidence among Cameroonian under-5, using routine data on symptomatic malaria collected in health facilities, between 2012 and 2018. In total, 4,052,216 cases of malaria were diagnosed between 2012 and 2018. There was a gradual increase per year, from 369,178 in 2012 to 652,661 in 2018. After adjusting the data for completeness, the national incidence ranged from 489‰ in 2012 to 603‰ in 2018.

In the present research, our primary emphasis is on comprehensive description of the dataset, based on statistics approaches and rates. This will open wide ways for predictions, using statistical, machine learning and Deep learning methods (Twumasi-Ankrah, S, Pels, W. A., Nyantakyi, K. & Addo, D. K., 2019; Yihang, D., 2023). The study takes advantage of the above presented works. In addition, the dataset and the environment under consideration are drawn directly from the local context. To the best of our knowledge, no prior scientific work has been conducted in this specific region, regarding data description.

The paper begins by putting in place statistical framework. This is followed by a comprehensive overview of the methodology employed and the resultant findings. Our presentation culminates in an interpretation of the findings within the context of a thorough discussion, leading to a conclusive summary.

2. Statistical Framework and Derived Rates

In this section, we delve into the statistical concepts applied and the rates that have been derived through our analysis.

2.1 Statistical measures

The purpose of this paper is to provide a comprehensive understanding of dynamics of malaria in the Adamawa Region using statistics and rates. Statistics incorporate a comprehensive set of methodologies, encompassing procedures for the measurement, classification, computation, description, synthesis, analysis, and systematic interpretation of acquired data (Mishra et al., 2019; Yihang, D., 2023). Within the field of statistics, two fundamental branches emerge: Descriptive Statistics and Inferential Statistics (Guetterman, T. C., 2019; Kaur, P., Stoltzfus, J. & Yellapu, V., 2018). Descriptive Statistics offer a toolbox of numerical and graphical techniques designed to succinctly summarize data collections, rendering complex information in a comprehensible format. In contrast, Inferential Statistics equips researchers with methods to draw meaningful inferences about populations based on observations from a sample. For the focus of this discussion, our primary emphasis lies on descriptive statistics. This particular facet of statistical analysis holds significant importance within the realm of biomedical research (Binu, V. S., Mayya, S. S. & Dhar, M., 2014; Satake, E. B., 2015; Guetterman, T. C., 2019). It serves as a vital tool to depict the foundational characteristics of data under study. In essence, it empowers researchers to delve into the intricacies of data, ultimately facilitating the generation of informed conclusions and decisions based on a solid data-driven foundation. Descriptive statistics comprises four pivotal measures: central tendency, dispersion, shape and dependences (Alhassan, E. A., Adjei, M. I., Aidoo, E., 2017; Guetterman, T. C., 2019; Mishra et al., 2019; Twumasi-Ankrah, S, Pels, W. A., Nyantakyi, K. & Addo, D. K., 2019; Yihang, D., 2023).

Measures of the central tendency provide information about the center of the observations. Among them, we have mean, median and quartiles. Let remind that quartiles are the three points that divide the dataset into four equal groups. Each group comprising a quarter of the data, for a set of data values which are arranged in either ascending or descending order. Each dataset has three quartiles, (Q1, Q2, and Q3) representing the first, second, and third quartile's value. Q2 corresponds to the median.

Measures of dispersion indicate how much observations tend to deviate from the average or central values. We focus in this work on some of them, namely: Variance, Standard deviation, Range and Inter Quartile Range (IQR).

Measures of shape are commonly skewness and kurtosis (Demir, S. 2022, Hatem et al. 2022). Skewness is the measure of the horizontal distance between mode and mean. It represents the asymmetry of data or the symmetric distortion. It's the third statistical moment, after mean and variance. Kurtosis provides information about the shape of a distribution, specifically the heaviness or lightness of its tails. It is the fourth statistical moment and serves as a measure of the degree of curvature, indicating whether the distribution is more or less peaked than a normal distribution. Let recall that the statistical moments are quantitative measure that describes the specific characteristics of a probability distribution. It is also the way to measure how spread out or concentrated the number in a dataset is around the central value, such as the mean (Novák, L.; Novák, D. 2020).

As well as the time series we are working on are univariate, the measures of dependence will focus on autocorrelation function (ACF) and partial autocorrelation function (PACF). These functions can help to assume whether or not a time series is stationary. Stationarity has then to be confirmed by tests such as the Augmented Dickey-Fuller (ADF) (Mushtaq, R., 2011; Paparoditis, E. & Politis, D. N., 2018) or the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) (Shin, Y., Schmidt, P., 1992; Hornok, A., & Larsson, R. 2000). A stationarity series reflects the time-invariance of all process characteristics. This property is a prerequisite for the application of a number of statistical and machine learning methods on time series. It also open wide possibilities for time series forecasting. ACF can be used to identify autoregressive (AR) dependencies, and PACF is used for moving average (MA) dependencies (Brockwell, P. J. & Davis R. A., 2016). The autocorrelation function is a measure of the correlation between observations of a time series that are separated by k time units (x_t and x_{t-k}). Meanwhile, the partial autocorrelation function is a measure of the correlation between the observations of a time series that are separated by k time units (x_t and x_{t-k}), after adjusting for the presence of all the other terms of shorter lag (x_{t-1}, x_{t-2}, ..., x_{t-k-1}).

Refer to Table 1, for mathematical representations of these measures.

2.2 Progression and prevalence rates

Time series show the evolution of a statistical variable over time. One of the tools used to define changes in time series is the progression rate. These tools measure variations in an observation between two or more dates, with or without a regular time step. In this work, we use both the annual progression rate and the average progression rate, based on the geometric mean. The progression rate measures the gross change and the direction (positive or negative) of a quantity in relation to a reference situation V_0. The geometric mean of n positive values x_i is defined as the n^{th} root of the product of these values. It is used to establish average rates, particularly annual average rates (Satake, E. B., 2015; Twumasi-Ankrah, S, Pels, W. A., Nyantakyi, K. & Addo, D. K., 2019; Mousa, A. et al. 2020; Danwang et al., 2021; Esum, M. E., Ndip, R. N. & Sumbele I., 2022; Adewole, A. I., Amurawaye, F. F. & Oladipupo J. O., 2023).

According to the dictionary of the Académie de Médecine[6], prevalence in epidemiology is the total number of cases of a given disease existing in a defined population, without distinction between new and old cases, over a defined period of time or at a defined moment in time. It is usually expressed by a ratio where the numerator is the total number of cases and the denominator is the size of the population in question.

To summarize this section, Table 1 represents mathematical formulas used to figure main measures. Source of tables and figures is from authors.

Table 1. Main measures and their formula

Measure	Formula
i-th Quartile	$Q_i = [i * (n+1)/4]th, where\ i = 1, 2\ or\ 3$
Skewness	$Sk = \frac{1}{n} * \frac{\sum_1^n (x_i + \bar{x})^3}{s^3}$
Kurtosis	$K = \frac{1}{n} * \frac{\sum_1^n (x_i + \bar{x})^4}{s^4}$
Auto Correlation Function	$\rho(h) = \frac{Cov(x_t, x_{t+h})}{\sqrt{x_t * x_{t+h}}}$
Annual progression rate	$p_i = \frac{v_{t+1}}{v_t} - 1$
Geometric mean for four rates	$gm = \sqrt[4]{(1+p_1)(1+p_2)(1+p_3)(1+p_4)}$
Prevalence rate	$pr = \frac{cases}{population}$
Main variables	n = number of observations, v_t = number of cases at year t x_i = measure of i-th observation, $i = 1..n$

After this recall, we now move to the presentation of the approach we used.

3. Methodology

The methodology we adopt for the present work encompasses three main steps. They are: Data extraction, checking and plotting; Rates evaluation; and finally statistical description. This last step includes: Concentration, Dispersion, Shape and Dependance analysis.

Before moving to data extraction, we need to describe them.

3.1 Data description

The data come from the DHIS platform[7], set up by the Cameroon Ministry of Public Health. The platform contains data on several diseases collected from Health Facilities and aggregated on a weekly basis to obtain data at upper geographical granularities. The top granularity is the national level, where data are aggregated for the whole country. Data are selected according to

[6] www.academie-medecine.fr/le-dictionnaire
[7] https://dhis-minsante-cm.org

the disease, period and geographical scale required. They are then extracted as .csv or .xls files. These files serve as datasets. In this dataset, we have 4 categorical variables and 11 numerical ones. The categorical variables all refer to the same information, i.e., the week in which data were collected. The 11 numerical variables represent values collected for each of the 11 Health Districts. In addition, there is a further 12th numerical variable, which is the cumulative sum for each week of the 11 Health Districts. This last variable represents the total number of cases for the entire region.

The data on which we are working are discrete quantitative data. For experimentation, we use the scientific programming language Python[8]. It is adapted for statistics, through several specialized libraries. These libraries include Statistics[9] for descriptive statistics; Pandas[10] for numerical computing; Mathplotlib[11] combined with Seaborn[12] for graphics and data visualization. We also used Timeserieslab (Lit, R., Koopman, S.J., and Harvey. A.C. 2023) to corroborate some of the results.

We take into account data between January 2018 and December 2022. The designated region (Adamawa Region), in conjunction with the specified timeframe spanning from 2018 to 2022, encompasses no fewer than 70 distinct time series datasets. We will base the studies on data at the Health District granularity level. The data will afterward be aggregated to the regional level.

3.2 Extraction, checking and plotting data

Data extraction is about getting data from relevant sources, according to criteria and goal we need to achieve. This step helps then extracting and gathering relevant data for the project. The sources are health databases. After that, we check if the extracted values are completed, well-typed, and that there are not outliers. We then select relevant variables to work with. Data plotting is used to produce graphs for a better understanding of behaviours of the datasets. This phase is known as data visualization. Graph types commonly used in univariate time series analysis include line graphs, histograms, density and boxplots. For the present work, we will restrict ourselves to the linear representation of time series.

3.3 Rates evaluation

The rates evaluation helps us determining the changes in the time series at different given dates. We evaluate the annual progression rate and the average progression rate. The progression rate is based on geometric mean. The formulas for annual progression rate, geometric mean and prevalence rate are given in Table 1.

3.4 Data analysis

This stage, which is the most intensive, involves carrying out analyses in order to come out with described and synthesized data. It is broken down into several sub-steps. Concentration analysis will determine the central indices of the time series. These include mean, median and quartiles. Dispersion analysis will enable us to understand the spread of the data in the series, relative to the central values. Shape analysis is based on Kurtosis and Skewness indices. Dependency analysis highlight relationships between the values in a series, at different points in time, and here will be focused of ACF and PACF. We find formulas used in Table 1.

Applying the methodology leads us to results.

4. Results

Results will be presented according to the methodology points.

[8] www.python.org
[9] https://docs.python.org/3/library/statistics.html
[10] https://pandas.pydata.org/
[11] https://matplotlib.org/
[12] https://seaborn.pydata.org/

4.1 Extraction, checking and plotting data

As we previously stated, data are extracted from DHIS platform and stored in .csv files as time-series. These files serve as datasets for the work. What we can notice here is that, the Health Information Unit is doing a great job. When extracting data from the database, we remark that the values are completed and well-typed, for all the Health District. There are also no outlier data. For each Health District, there is a variable called *periodid*, representing weeks. Every value of *periodid* is associated to a number, representing cases occurring each week in that area. These constituted datasets open ways for next steps.

The curves in Figure 1 show the cases occurring in the 11 Health Districts, between 2018 and 2022, on a weekly basis. The Figure is spilt into two, to allow better visibility. Figure 1a represent data for the first 6th Health Districts and Figure 1b for the rest.

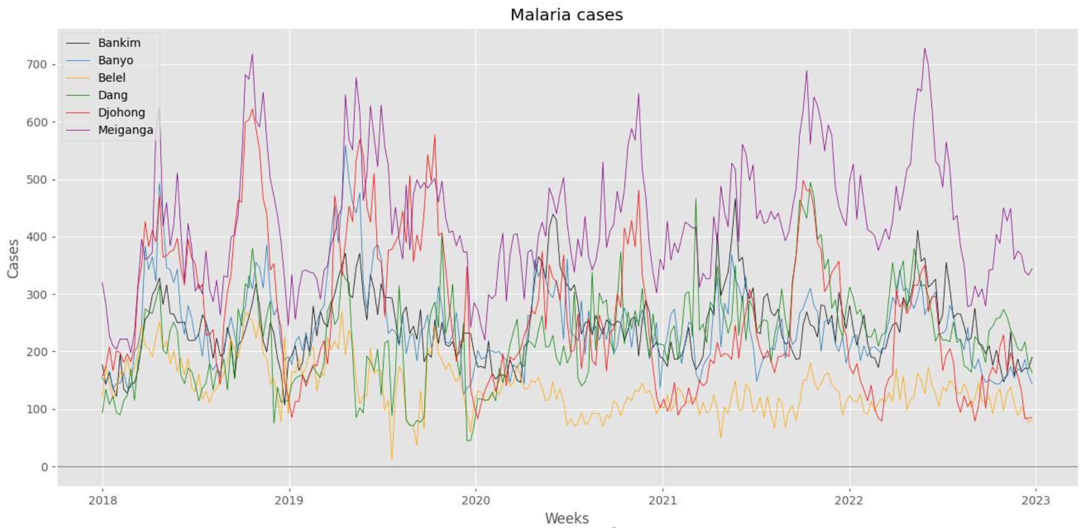

Figure 1a. Data time series for 6th first Health Districts.

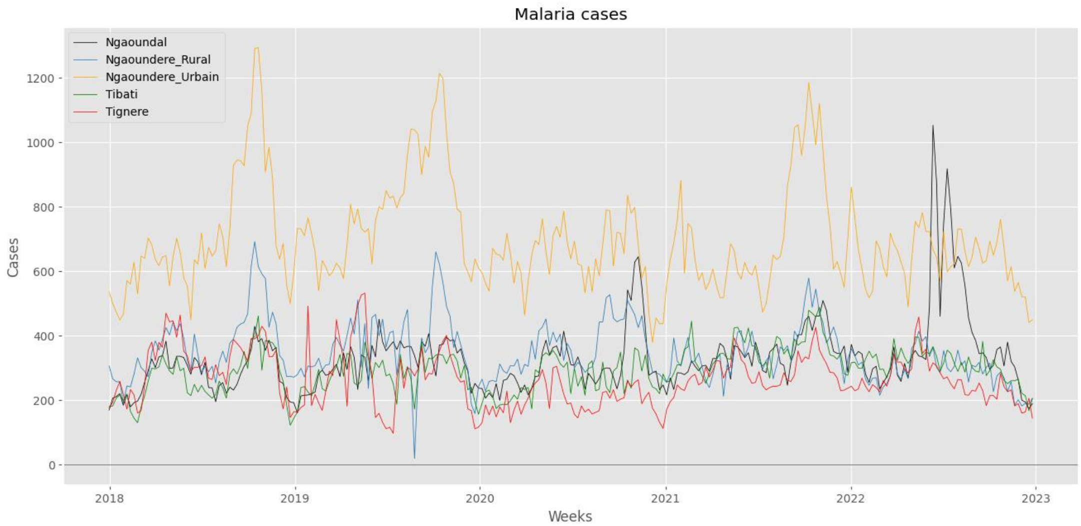

Figure 1b. Data time series for the rest of Health Districts.

Figure 1. Data time series for Health Districts.

Observing figure 2, we note that the process reproduces itself identically over an approximative annual period (52 weeks). The series peaks around week 42 of each year. In other words, the Adamawa Region is much more affected by malaria between the 33rd and 50th week each year. This period runs from mid-August to mid-December, as delimitated on Figure 2 with blue lines. We notice the same observations in Alhassan, E. A., Adjei, M. I., Aidoo, E., 2017 and Eunice, A., Wanjoya, A. & Luboobi, L. 2017, where periods with highest cases occurring are respectively between June to November and August to November. Those are periods characterized by high rainfall, followed by a transition between the rainy season and the dry season. Figure 3, for data of the year 2018 illustrate it more. It is the same case for years 2019, 2020 and 2021 (see Appendices, Figures 6 to 8). However, year 2022 (see Appendices, Figure 9) don't follow the same trend, assuming that a special policy has been deployed in that period of the year.

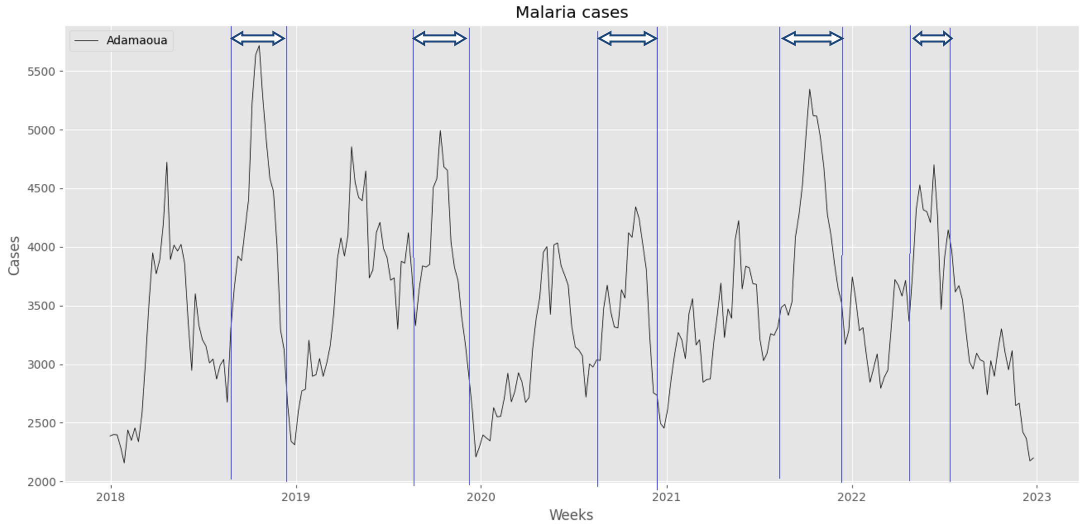

Figure 2. Time series data for the Adamawa Region.

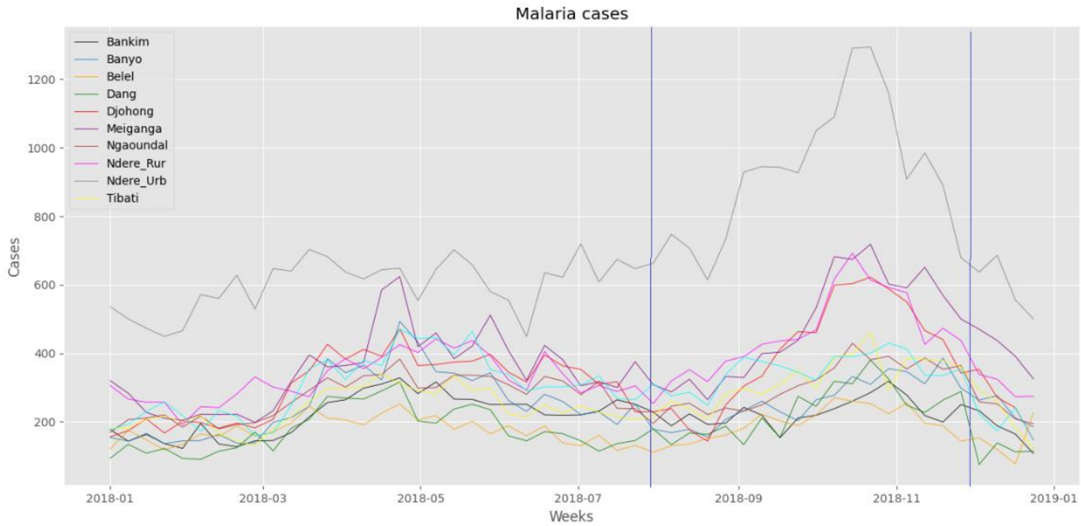

Figure 3. Health Districts data for year 2018.

4.2 Progression and prevalence rates

Table 2 shows (rows 1 to 5) the total number of cases occurring each year in each Health Districts. It also shows the total number of cases over 5 years (row 6). The information is used to determine the annual progression rate of the disease. They are presented in rows 7 to 10. And row 11 is the rate for the whole studied period (2018 to 2022). Columns are the 11 Health Districts and the total for the Region. From the 04 annual progression rates obtained (row 7 to 10), we calculate the average progression rate over the entire period. This rate is based on the geometric mean, and is found in row 11. Refer to Table 1 for different formulas.

Table 2. Progression rate of cases

N°	Year	Bankim	Banyo	Belel	Dang	Djohong	Meiganga	Ngaoundal	Ndere_Rur.	Ndere_Urb.	Tibati	Tignère	Total
	Total cases per year and for the whole period												
1	2018	11,419	13,219	9,299	9,910	17,349	20,608	14,609	19,184	36,913	13,780	16,127	**182,417**
2	2019	13,217	14,617	8,034	9,709	17,450	22,842	16,523	19,209	41,215	14,595	14,378	**191,789**
3	2020	13,175	12,470	6,203	11,135	12,692	21,161	16,492	18,922	33,687	13,850	10,480	**170,267**
4	2021	13,395	12,316	5,973	5,973	12,325	24,019	17,785	18,546	36,912	18,287	14,659	**180,190**
5	2022	12,538	11,422	6,275	13,003	9,353	22,974	20,471	15,029	33,325	16,029	13,328	**173,747**
6	**18-22**	63,744	64,044	35,784	58,907	69,175	111,604	85,880	90,890	182,059	76,541	68,972	**907,600**
	Rates per year and for the whole period												
7	18-19	**15.75%**	10.58 %	-13.60%	-2.03%	0.58%	10.84%	13.10%	0.13%	11.65%	5.91%	***-10.85%***	5.14%
8	19-20	-0.32%	- 14.69 %	-22.79%	**14.69%**	-2.,27%	-7.36%	-0.19%	-1.49%	-18.27%	- 5.10%	**-27.11%**	-11.22%
9	20-21	1.67%	- 1.23%	-3.71%	**-46.36%**	-2.89%	13.51%	7.84%	-1.99%	9.57%	**32.04 %**	**39.88%**	5.83%
10	21-22	-6.40%	- 7.26%	5.06%	**117.7%**	***-24.11%***	-4.35%	15.10%	-18.96%	-9.72%	- 12.35 %	-9.08%	-3.58%
11	**18-22**	2.36%	- 3.59%	**-9.37%**	**7.03%**	**-14.31%**	2.75%	**8.80%**	-5.92%	-2.52%	3.85%	-4.65%	-1.21%

The highest annual growth rate is seen in Dang District, between 2021 and 2022, where there was an increase of more than 100% in the number of cases. The number of cases rose from 5,973 in 2021 to 13,003 in 2022. Next, between 2020 and 2021, we have Tibati and Tignère, which recorded a growth rate of over 30% i.e., 32.04% and 39.88% respectively.

The biggest drop was again in Dang, this time between 2020 and 2021. We had a drop of -46.36% in the number of cases recorded. In the Djohong Health District, we have only seen decreases. Over the 5 years, the average growth rate is -14.31%, dropping from 17,349 cases in 2018 to 9,353 cases in 2022. Djohong is followed by the Bélel Health District, with a growth rate of -9.37%.

On the other hand, the Health Districts of Ngaoundal and Dang show a significant increase in the number of cases, when considering the data for the entire collection period. Between 2018 and 2022, the number of cases rose from 9,910 to 13,003 in Dang, and from 14,609 to 20,471 in Ngaoundal. This gives an average growth rate of 7.03% and 8.80% for Dang and Ngaoundal respectively.

Although the growth rate for the whole Region is slightly below zero (-1.21%), the number of cases remains significant. It represents 907,600 cases over the 5 years, with 173,747 cases occurring in 2022.

In order to determine the most affected Health Districts over the population, we determined the prevalence rate. We based our calculations on the year 2022, for which we have complete population data. The results are shown in Table 3.

Table 3. Prevalence rate

Indicator	Bankim	Banyo	Belel	Dang	Djohong	Meiganga	Ngaoundal	Ndere_Rur.	Ndere_Urb.	Tibati	Tignère	Total
Population	125,043	172,457	533,86	93,038	85,827	179,081	99,087	118,947	388,865	116,674	130,192	1,562,596
Cases	12,538	11,422	6,275	13,003	9,353	22,974	20,471	15,029	33,325	16,029	13,328	173,747
Prevalence	10%	**7%**	12%	14%	11%	13%	**21%**	13%	**9%**	14%	10%	11%

According to the results containing in Table 3, only the Districts of Banyo and Ngaoundéré Urbain have a prevalence rate below 10%. For the rest, at least 10% of the total population was affected by malaria in 2022. The District of Ngaoundal beats the record with a prevalence rate of 21%, followed by the Health Districts of Dang and Tibati, each with 14%. For the Region as a whole, the prevalence rate is 11%. So, out of 1,562,596 people living in the Adamawa Region in 2022, 173,747 were infected.

Let’s explore now statistical characteristics of the dataset.

4.3 Data description

From the statistical description of the data, we obtain 05 tables corresponding to the 05 study years, and 01 table for the entire studied period. The statistical summaries for the annual tables are available in the Appendices (Tables 7 to 11). We present here the statistical summary table for the entire studied period (Table 4).

Table 4. Statistical synthesis of data between 2018 and 2022

Indicator	Bankim	Banyo	Belel	Dang	Djohong	Meiganga	Ngaoundal	Ndere_Rur	Ndere_Urb	Tibati	Tignère
μ (mean)	244.23	245.38	**137.10**	225.70	265.04	427.60	329.04	348.24	**697.54**	293.26	264.26
σ^2	3,870.2	5,535.4	1,991.1	6672.7	15791.13	11,948.2	13,081.53	8,644.56	27,356.05	4,538.0	6,939.18
σ	62.21	74.40	**44.62**	81.69	125.66	109.31	114.37	92.98	**165.40**	67.36	83.30
Min	107	125	**11**	45	79	198	160	**19**	379	122	97
Max	466	559	270	495	622	728	**1,053**	692	**1,294**	488	532
Q_1	197	195	109	171	167	352	265	285	595	242	206
Me or Q_2	240	231	130	226	238	415	311	328	654	297	256
Q_3	279	280	161	273	356	494	365	405	754	340	313
IQR	82	85	52	102	189	142	100	120	159	98	107
Range	359	434	259	450	543	530	893	673	915	366	435
Skewness	0.6444	1.13	0.6217	0.4061	0.6426	0.4332	2.66	0.6926	1.24	0.0520	0.5919
Kurtosis	3.54	4.74	3.38	3.46	2.65	2.88	14.05	4.37	4.56	2.95	3.25

The abbreviation of indicators collected, for the sample, are summarized in Table 5.

Table 5: Abbreviation of indicators

Indicator	Meaning	Indicator	Meaning
μ	Expected or plausible value	Q_1, Q_2, Q_3	First, Second and Third quartile
σ^2 / σ	Variance / Standard deviation	Me	Median also representing second quartile (Q2)
Min/Max	Minimum / Maximum value	IQR	Interquartile range

Table 4 summarizes, for the 05 years of the study (2018 to 2022), the statistical indicators for the observations recorded each week. It encompasses central, spread, and shape indices.

The average number of weekly cases varies between 137 (Bélel District) and 697 (Ngaoundéré Urban District). The deviation from the central value is 44 in Bélel (the lowest value) and 165 in Ngaoundéré Urbain (the highest value). The week with the fewest cases is also in Bélel, with only 11 cases recorded (30/2019, corresponding to the end of June 2019). It is followed by Ngaoundéré Rural, with 19 cases (35/2019, corresponding to the beginning of August 2019). On the other hand, the week with the highest number of cases recorded has 1,294 cases (43/2018, corresponding to the end of September 2018). This record is beaten by Ngaoundéré urban District. It is followed by Ngaoundal, with 1,053 cases recorded in one week (24/2022, corresponding to mid-May).

Concerning shape indicators, all the 11 distributions obtained from the Adamawa Health Districts show positive skewness. This suggests that the distributions are shifted to the left of the median and that the tail of the distribution is therefore skewed to the right. Similarly, the kurtosis coefficients are all positive. This may mean that the curves are all leptokurtic i.e., their ends are thicker than normal. All the coefficients are greater than 3, except for the Districts of Djohong, Meiganga and Tibati. For Districts with a coefficient greater than 3, the distributions should follow Laplace's law (Chen, K., van Laarhoven, T. & Marchiori, E. 2021). And for those below 3, which are all above 2, the distribution should be hyperbolic secant (M. Hilton, R. Alexandru and P. L. Dragotti, 2021).

Figures 4 and 5 show respectively autocorrelation and partial autocorrelation of the time series representing cases for the whole region. We can see from these figures that, as delays or lags increase (on the x-axis), the values of autocorrelations and partial autocorrelations decrease very rapidly and are contained in the insignificance area around zero. This rapid decrease suggests that the observations are weakly correlated with each other, meaning that the series may be stationary. The stationarity has been confirmed by ADF and KPSS tests, as shown in Table 6.

Table 6: ADF and KPSS tests of unit root

Augmented Dickey-Fuller (ADF)	Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
Statistic = -5.71 P-value = 8.9224e-06; Lags = 4 H0: unit root is present, Test result: reject H0	Statistic = 0.0563 P-value > 0.1000; Lags = 9 H0: series is stationary, Test result: do not reject H0

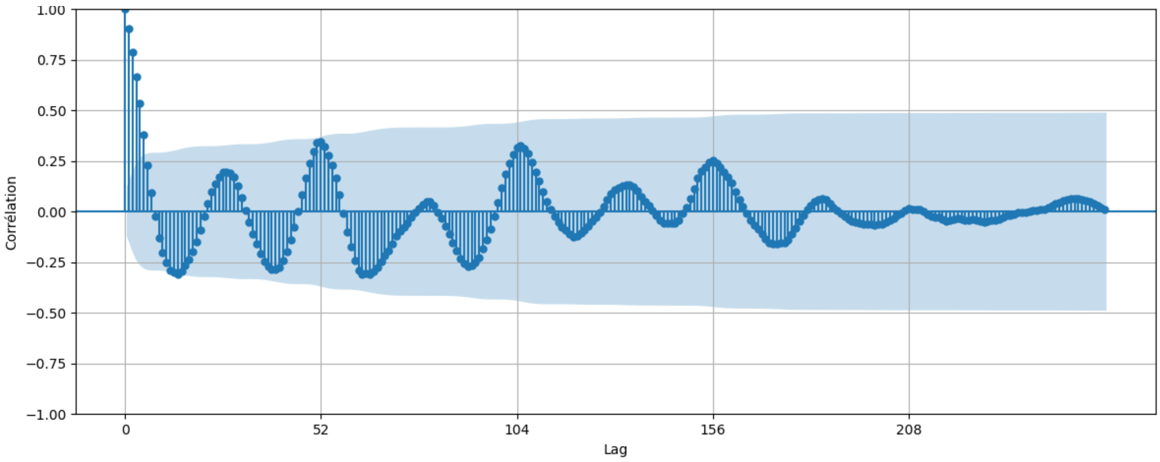

Figure 4. Autocorrelation graph.

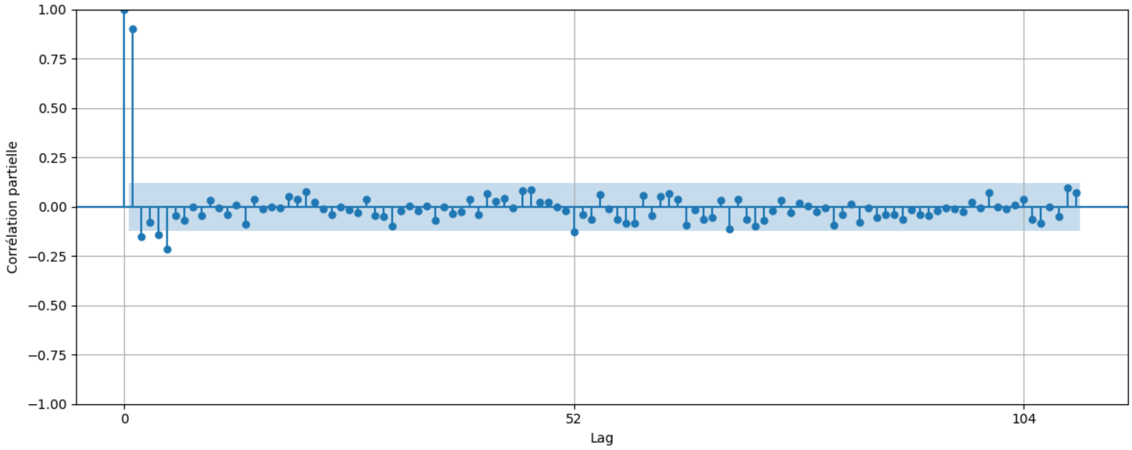

Figure 5. Partial Autocorrelation graph.

At the end of that section, we get data entirely described and rates evaluated. We can now move to discussion and concluded the work.

5. Discussion

The overarching theme of this work revolves around the application of time series analysis in epidemics monitoring and prediction. The ultimate objective is to create models, characterize, analyze, and predict based on time series data. Our goal, in the present preliminary work, is to generate a statistical summary of the collected data and ascertain growth and prevalence rates.

The primary limitation of this study lies in the lack of relevant data for multivariate analysis. We only have access to data on weekly cases, with no information regarding age, sex, environmental conditions, social status, or mortality. Including these variables would have enhanced the relevance of our analysis. It's essential to acknowledge that the statistical description provided here may not be exhaustive. Nevertheless, it offers a necessary and sufficiently comprehensive foundation for understanding and further work on time series analysis and prediction.

The study revealed the following key findings:

- Peak periods of contamination are mostly between August and November, during raining season. Whereas, periods of relaxation are found at the beginning of the year (January to March). The highest level of cases recorded in one week is 1,294, in the Ngaoundéré Urbain Health District, on October 2018. The less one is 11 cases, in Belel Health District, recorded on July 2019. The average number of weekly cases varies between 137 (Bélel District) and 697 (Ngaoundéré Urban District). All the skewness measures are positive, showing that distributions are shifted to the left of the median. Likewise, kurtosis coefficients are all positive and above 2, showing that curves are leptokurtic. Concerning autodependences, we can notice that autocorrelation and partial autocorrelation values decrease rapidly, as lags grow. This points that the observations are weakly correlated with each other, and that the time series can be stationary. The stationarity of the times series has been confirmed by ADF and KPSS tests that we performed.

- The highest annual growth has been recorded between the years 2021 and 2022 in the Dang District, with an increase of 117,70%. And the less one has also been recorded in Dang, between years 2020 and 2021, with an increase rate of -46,36%. On the entire period (2018 to 2022), Djohong Health District recorded the less average rate (-14,31%), whereas Ngaoundal District recorded the highest growth rate (8,80%). For the whole Region, the growth rate is slightly below zero (-1,21%). But the total number of cases (907,600 over the 5 years, with 173,747 cases occurring in 2022) remain really important, and thus need special attention.
- For the year 2022, the Health District of Ngaoundal has the highest prevalence rate (21%), and Banyo Health District the less one (7%). For the whole Region, the prevalence rate is 11%, revealing that on 2022, over 173,747 persons were infected over a population of 1,562,596 persons. This corroborates the previous point.

As for recommendations, the first one pertains to strengthening disease prevention measures, which may include the use of insecticides, insecticide-treated nets, rapid diagnostic tests, artemisinin-based combination therapy, preventive and anti-malarial drugs, and improvements in environmental sanitation, as also suggested in references Talipouo, 2019; Esayas, 2020; Mousa, A. et al. 2020; Li, G. et al. 2022. This is particularly crucial, given that not all patients seek medical care at health facilities. Many individual resort to self-medication, employing conventional or non-conventional treatments.

The second recommendation, which holds significant value for our research, is to gather data exhaustively, including information on age, sex, environmental conditions, social status, and mortality. This comprehensive dataset would enable more relevant and meaningful multivariable data analysis, and afterward focus on data predictions, using statistical, machine leaning and deep learning methods (Moskalaï Ngossaha, J., Ynsufu, A., Batoure Bamana, A., Djeumen, R., Bowong Tsakou, S. & Ayissi Eteme, A., 2024).

In our forthcoming work related to these dataset, spatial frameworks, and statistical analysis, we already furthering works on time series analysis and forecasting using Machine Learning and Deep Learning methods (Batoure Bamana, A., Shafiee Kamalabad, M., and Oberski, D. L., 2024). This will lead to create models and predictions based on time series data, in the field of disease with epidemiological potential, such as Malaria.

6. Conclusion

At the end of this study, which focused on the statistical description of data on malaria cases occurring weekly from 2018 to 2022 in the Adamawa Region of Cameroon, we can state that malaria continues to be highly prevalent, claiming numerous lives within households each year, with a particularly profound impact on the African continent. Appropriate measures should be taken throughout the year, and reinforced during periods of high contamination. This is a matter for Government, Partners and general public alike. The present work is a preliminary to others that will also be related to epidemics time-series studies, and focusing on ways of predictions using Machine Learning and Deep Learning methods, along with Big Data and Artificial Intelligence.

Acknowledgments

We acknowledge the Health Information Unit of the Ministry of Public Health of Cameroun, and the Adamawa Regional Delegation of the same Ministry for providing us with data and relevant information for this work.

Data Availability

The data used to support the findings of this study are available from the corresponding author, on request.

Authors' Contributions

All the authors contributed to the work, have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Adewole, A. I., Amurawaye, F. F. & Oladipupo J. O. Time series analysis of malaria fever prevalence in Ogun State. *TASUED Journal of Pure and Applied Sciences* **2** (1), 201-211 (2023).
2. Alhassan, E. A., Adjei, M. I., Aidoo, E. Time Series Analysis of Malaria Cases in Kasena Nankana Municipality. *International Journal of Statistics and Applications* **7** (2), 43-56 (2017) DOI: 10.5923/j.statistics.20170702.01.
3. Batoure Bamana, A., Shafiee Kamalabad, M., & Oberski, D. L. A systematic literature review of time series methods applied to epidemic prediction. Informatics in Medicine Unlocked (2024), doi: https://doi.org/10.1016/j.imu.2024.101571
4. Binu, V. S., Mayya, S. S. & Dhar, M. Some basic aspects of statistical methods and sample size determination in health science research. *Ayu* **35**, 119-23 (2014).
5. Brockwell, P. J. & Davis R. A. Introduction to time series and forecasting. *Springer*, 3rd Edition 2016.
6. Chen, K., van Laarhoven, T. & Marchiori, E. Gaussian processes with skewed Laplace spectral mixture kernels for long-term forecasting. Mach Learn **110**, 2213–2238 (2021). https://doi.org/10.1007/s10994-021-06031-5.
7. Danwang, C., Khalil, E., Achu, D., Ateba, M., Abomabo, M., Souopgui, J., De Keukeleire, M. & Robert, A. Fine scale analysis of malaria incidence in under-5: hierarchical Bayesian spatio-temporal modelling of routinely collected malaria data between 2012–2018 in Cameroon. *Nature portfolio, Scientific Reports* **11**, (2021) 11408| https://doi.org/10.1038/s41598-021-90997-8.
8. Demir, S. (2022). Comparison of Normality Tests in Terms of Sample Sizes under Different Skewness and Kurtosis Coefficients. International Journal of Assessment Tools in Education, **9** (2), 397-409. https://doi.org/10.21449/ijate.1101295.
9. Dian, N. D., Mohd Salleh, A. F., Rahim, M. A. F. A., Munajat, M. B., Abd Manap, S. N. A., Ghazali, N., Hassan, N. W. & Idris, Z. M. Malaria Cases in a Tertiary Hospital in Kuala Lumpur, Malaysia: A 16-Year (2005–2020) Retrospective Review. *Trop. Med. Infect. Dis.* **6**, 177 (2021) https://doi.org/10.3390/tropicalmed6040177.
10. Esayas, E., Tufa, A., Massebo, F., Ahemed, A., Ibrahim, I., Dillu, D., Bogale, E. A., Yared, S., & Kassaye, K. Malaria epidemiology and stratification of malaria incidence in the malaria elimination setting. *Harari Region, Eastern Ethiopia* **2**. University of Sussex. (2020) https://hdl.handle.net/10779/uos.23478434.v2.
11. Esum, M. E., Ndip, R. N. & Sumbele I. Trends of Malaria in the South West Region of Cameroon: Overview, Challenges and Perspectives. *SDG3, Why It Matters, UN Cameroon*, (2022).
12. Eunice, A., Wanjoya, A. & Luboobi, L. Statistical Modeling of Malaria Incidences in Apac District, Uganda. *Open Journal of Statistics* **7**, 901-919 (2017) https://doi.org/10.4236/ojs.2017.76063.
13. Guetterman, T. C. Basics of statistics for primary care research. *Family medicine and community health* **7** (2), e000067 (2019).

14. Hatem, G., Zeidan, J., Goossens, M., and Moreira, C. Normality testing methods and the importance of skewness and kurtosis in statistical analysis. BAU Journal - Science and Technology **3** (2), Art.7 (2022) DOI: https://doi.org/10.54729/KTPE9512.
15. Hilton, M., Alexandru, R., and Dragotti, P. L. "Time Encoding Using the Hyperbolic Secant Kernel," 2020 28th European Signal Processing Conference (EUSIPCO), Amsterdam, Netherlands, 2304-2308 (2021) doi: 10.23919/Eusipco47968.2020.9287806.
16. Hornok, A., & Larsson, R. The finite sample distribution of the KPSS test. The Econometrics Journal **3** (1), 108–121. (2000) http://www.jstor.org/stable/23115053.
17. Hyndman, R. J. & Athanasopoulos, G. *Forecasting: Principles and Practice*. (Third Print Edition, OTEXTS 2021).
18. Ibrahim OR, Lugga AS, Ibrahim N, Aladesua O, Ibrahim LM, Suleiman BA, Suleiman BM. Impact of climatic variables on childhood severe malaria in a tertiary health facility in northern Nigeria. Sudan J Paediatr. **21** (2), 173-181 (2021). doi: 10.24911/SJP.106-1599226765. PMID: 35221429; PMCID: PMC8879348.
19. Kaur, P., Stoltzfus, J. & Yellapu, V. Descriptive statistics. *Int J Acad Med* **4**, 60-3 (2018).
20. Landoh, E. D., Potougnima, T., Bayaki, S., Khin, S. Tint., Sheba, N. G., Wasswa, P. & de Jager, C. Morbidity and mortality due to malaria in Est Mono district, Togo, from 2005 to 2010: a times series analysis. *Malaria Journal* **11**, 389 (2012).
21. Li, G., Zhang, D., Chen, Z., Feng, D., Cai, X., Chen, X., Tang, S. & Feng, Z. Risk factors for the accuracy of the initial diagnosis of malaria cases in China: a decision-tree modelling approach. *Malaria Journal* **21**, (11) (2022) https://doi.org/10.1186/s12936-021-04006-4.
22. Lit, R., S.J. Koopman, and A.C. Harvey (2023), Time Series Lab: https://timeserieslab.com.
23. Makridakis, S., Assimakopoulos, V. & Spiliotis, E. Objectivity, reproducibility and replicability in forecasting research. *International Journal of Forecasting* **34** (4), 835–838, (2018).
24. Mbouna, A. D., Tompkins, A. M., Lenouo, A., Asare, E. O., Yamba, E. I. & Tchawoua C. Modelled and observed mean and seasonal relationships between climate, population density and malaria indicators in Cameroon. *Malaria Journal* **18**, 359 (2019) https://doi.org/10.1186/s12936-019-2991-8.
25. Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu C. & Keshri A. Descriptive statistics and normality tests for statistical data. *Ann Card Anaesth* **22** 67-72 (2019).
26. Moskalaï Ngossaha, J., Ynsufu, A., Batoure Bamana, A., Djeumen, R., Bowong Tsakou, S. & Ayissi Eteme, A. Towards a Flexible Urbanization Based Approach for Integration and Interoperability in Heterogeneous Health Information Systems: Case of Cameroon. In: Tchakounte, F., Atemkeng, M., Rajagopalan, R.P. (eds) Safe, Secure, Ethical, Responsible Technologies and Emerging Applications. SAFER-TEA 2023. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 566. Springer, (2024), Cham. https://doi.org/10.1007/978-3-031-56396-6_16
27. Mousa, A., Al-Taiar, A. Anstey, N. M., Badaut, C., Barber, B. E. & Bassat Q. The impact of delayed treatment of uncomplicated P. falciparum malaria on progression to severe malaria: A systematic review and a pooled multicentre individual-patient meta-analysis. *PLoS Med* **17**, (10) e1003359 (2020).
28. Mushtaq, R. Augmented Dickey Fuller Test. 2011. http://dx.doi.org/10.2139/ssrn.1911068.
29. Novák, L.; Novák, D. On Taylor Series Expansion for Statistical Moments of Functions of Correlated Random Variables. Symmetry **12**, 1379 (2020). https://doi.org/10.3390/sym12081379.
30. Ozulonye, O. S., Okolo, A., Torsen, E. & Tiwah O. J. Multivariate time series analysis in modelling malaria cases in Jimeta metropolis of Adamawa state, Nigeria. *FUDMA Journal of Sciences* **6** (3), 62-69 (2022).
31. Paparoditis, E. & Politis, D. N. The asymptotic size and power of the augmented Dickey–Fuller test for a unit root, Econometric Reviews **37** (9), 955-973, (2018) DOI:

10.1080/00927872.2016.1178887.

32. Ramirez, J. H., Urtasun, A., Roselló, M., Garrido, M., Peman, J. & Otero, M. C. Estudio descriptivo de los casos de malaria en la población pediátrica en un hospital de referencia de Valencia, Espana, entre 1993 y 2015. *An Pediatr (Barc)*; **92** 21-27 (2020).
33. Rodríguez, S. N. I., Rodríguez, J. A. I., Rodríguez, J. C. P. & Olivera, M. J. Malaria mortality in Colombia from 2009 to 2018: a descriptive study. *Journal of the Brazilian Society of Tropical Medicine* **54**, e0441-2020 (2021) https://doi.org/10.1590/0037-8682-0441-2020.
34. Satake, E. B. Statistical Methods and Reasoning for the Clinical Sciences Evidence-Based Practice. *I^st ed. San Diego: Plural Publishing, Inc*. 1-19 (2015).
35. Shin, Y., Schmidt, P. The KPSS stationarity test as a unit root test, Economics Letters **38** (4), 387-392 (1992) , ISSN 0165-1765, https://doi.org/10.1016/0165-1765(92)90023-R.
36. Sohanang Nodem F. S., Ymele D., Fadimatou M., Fodouop S. C. Malaria and Typhoid Fever Coinfection among Febrile Patients in Ngaoundéré (Adamawa, Cameroon): A Cross-Sectional Study. J Parasitol Res. (2023) doi: 10.1155/2023/5334813. PMID: 37790287; PMCID: PMC10545472.
37. Talipouo, A., Ngadjeu, C. S., Doumbe-Belisse, P., Djamouko-Djonkam, L., Sonhafouo-Chiana, N., Kopya, E., Bamou, R., Awono-Ambene, P., Woromogo, S., Kekeunou, S., Wondji, C. S. & Antonio-Nkondjio, C. Malaria prevention in the city of Yaoundé: knowledge and practices of urban dwellers. *Malaria Journal* **18** (167) (2019) https://doi.org/10.1186/s12936-019-2799-6.
38. Twumasi-Ankrah, S, Pels, W. A., Nyantakyi, K. & Addo, D. K. Comparison of Statistical Techniques for Forecasting Malaria Cases in Ghana. *J Biostat Biometric App* **4**, (1) 102 (2019).
39. Wahedi JA, Ande AT, Oduola AO, Obembe A, Tola M, Oyeniyi TA, Awolola TS. Dynamics of malaria vector indices in two vegetation zones within North Eastern Adamawa State, Nigeria. Trop Biomed. (2020). doi: 10.47665/tb.37.3.637. PMID: 33612778.
40. Yihang, D. Descriptive Statistics and Its Applications, Highlights in Science, Engineering and Technology. *AMMMP* **47**, (2023).
41. Zhang Q. Data science approaches to infectious disease surveillance. *Phil. Trans. R. Soc.* **380** (2021) doi.org/10.1098/rsta.2021.0115.

Appendices:

Table 7. Statistics of 2018

Indicator	Bankim	Banyo	Belel	Dang	Djohong	Meiganga	Ngaoundal	Ndere_Rur.	Ndere_Urb.	Tibati	Tignère
μ	219.60	254.21	178.83	190.58	333.63	396.31	280.94	368.92	709.87	265	310.13
σ^2	2,990.16	7,062.01	1,931.68	5,428.86	15,901.77	18,358.75	4,259.29	11,219.49	40,810.89	5,386.46	7,171.12
σ	54.68	84.04	43.95	73.68	126.10	135.49	65.26	105.92	202.02	73.39	84.68
Median	220	248.50	179.50	170.50	343.50	378	280	344.50	647	259	322.50
Min	107	136	77	75	143	198	169	191	449	122	146
25%	185.25	183.25	146	133	223.75	307.25	220.75	288.75	577.75	214	245.50
75%	257	320.50	206.50	248.75	400.50	461.75	332.50	426	736.50	300	376.75
IQR	71,75	137,25	60,5	115,75	176,75	154,5	111,75	137,25	158,75	86	131,25
Max	328	493	270	380	622	718	429	692	1294	461	470
Range	221	357	193	305	479	520	260	501	845	339	324
Kurtosis	2.30	2.72	2.39	2.32	2.56	2.61	2.01	3.92	4.04	2.74	2.15
Skewness	-0.1062	0.5254	0.0969	0.5295	0.4495	0.6413	0.1584	1.07	1.27	0.3337	-0.1454
Sum Ann.	11,419	13,219	9,299	9,910	17,349	20,608	14,609	19,184	36,913	13,780	16,127

Table 8. Statistics of 2019

Indicator	Bankim	Banyo	Belel	Dang	Djohong	Meiganga	Ngaoundal	Ndere_Rur.	Ndere_Urb.	Tibati	Tignère
μ	254.17	281.10	154.50	186.71	335.58	439.27	317.75	369.40	792.60	280.67	276.50
σ²	2,659.49	9,931.28	3,000.06	7,157.47	18,921.09	10,988.27	4,104.96	12,279.13	28,782.70	2,953.87	12,268.90
σ	51.57	99.66	54.77	84.60	137.55	104.82	64.07	110.81	169.65	54.35	110.77
Median	250.50	240	162.50	177.50	368.50	434.50	333.50	369	774.50	282.50	275.50
Min	167	125	11	45	85	242	160	19	540	144	97
25%	211.75	216.25	123	121.50	213.50	358.25	276.50	301.25	636.75	245	180
75%	294	330.25	195.25	247.50	432.50	499.50	367.25	438.50	880.50	328.50	340
IQR	82,25	114,00	72,25	126,00	219,00	141,25	90,75	137,25	243,75	83,50	160,00
Max	371	559	270	407	577	677	451	660	1214	380	532
Range	204	434	259	362	492	435	291	641	674	236	435
Kurtosis	2.40	3.09	2.85	2.48	1.87	2.48	2.45	4.34	2.73	2.66	2.58
Skewness	0.4373	0.9377	-0.4632	0.3355	-0.1481	0.2557	-0.4915	-0.0545	0.6858	-0.4266	0.4683
Sum Ann.	13,217	14,617	8,034	9,709	17,450	22,842	16,523	19,209	41,215	14,595	14,378

Table 9. Statistics of 2020

Indicator	Bankim	Banyo	Belel	Dang	Djohong	Meiganga	Ngaoundal	Ndere_Rur.	Ndere_Urb.	Tibati	Tignère
μ	248.58	235.28	117.04	210.09	239.47	399.26	311.17	357.02	635.60	261.32	197.74
σ²	5,179.49	2,940.43	687.06	3,857.14	8,248.06	7,778.31	10,130.93	6,382.58	9,948.39	3,156.75	2,192.46
σ	71.97	54.23	26.21	62.11	90.82	88.19	100.65	79.89	99.74	56.18	46.82
Median	245	227	117	210	234	395	277	327	639	241	189
Min	146	137	69	71	83	220	199	228	379	156	112
25%	192	195	93	174	173	342	247	291	586	220	162
75%	285	263	135	244	291	442	341	427	695	311	228
IQR	93,00	68,00	42,00	70,00	118,00	100,00	94,00	136,00	109,00	91,00	66,00
Max	439	361	177	373	481	649	645	527	835	362	305
Range	293	224	108	302	398	429	446	299	456	206	193
Kurtosis	3.16	2.76	2.27	2.88	2.75	3.21	5.79	1.98	2.96	1.87	2.50
Skewness	0.7612	0.5386	-0.097	0.2069	0.5263	0.4932	1.78	0.4254	-0.3880	0.1531	0.3879
Sum Ann.	13,175	12,470	6,203	11,135	12,692	21,161	16,492	18,922	33,687	13,850	10,480

Table 10. Statistics of 2021

Indicator	Bankim	Banyo	Belel	Dang	Djohong	Meiganga	Ngaoundal	Ndere_Rur.	Ndere_Urb.	Tibati	Tignère
μ	257.60	236.85	114.87	291.35	237.02	461.90	342.02	356.65	709.85	351.67	281.90
σ²	4,003.86	2,514.28	670.65	6335.65	1,1785.83	7,808.70	3,984.13	5,725.30	33,796.17	3,405.80	2,561.93
σ	63.28	50.14	25.90	79.60	108.56	88.37	63.12	75.67	183.84	58.36	50.62
Median	247	231	111.50	268.50	197	434	336.50	344.50	635	344	273
Min	168	146	50	182	89	312	216	212	473	252	171
25%	213	200	97	233.50	159.50	413.25	291.75	310	585	311.50	246.50
75%	292.50	270.25	133.25	328.75	310	525.75	370.25	395.75	776.75	386.75	314
IQR	79,50	70,25	36,25	95,25	150,50	112,50	78,50	85,75	191,75	75,25	67,50
Max	466	369	180	495	498	689	509	579	1186	488	426
Range	298	223	130	313	409	377	293	367	713	236	255
Kurtosis	4.10	2.63	2.92	3.15	2.76	2.67	3.09	3.65	2.92	2.62	3.31
Skewness	1.01	0.4104	0.0900	1.03	0.8162	0.5994	0.5907	0.7622	1.06	0.4222	0.5846
Sum Ann.	13,395	12,316	5,973	5,973	12,325	24,019	17,785	18,546	36,912	18,287	14,659

Table 11. Statistics of 2022

Indicator	Bankim	Banyo	Belel	Dang	Djohong	Meiganga	Ngaoundal	Ndere_Rur.	Ndere_Urb.	Tibati	Tignère
μ	241.12	219.65	120.67	250.06	179.87	441.81	393.67	289.02	640.87	308.25	256.31
σ²	3,579.72	3,087.03	473.14	2,733.86	5,879.50	11,574.92	35,872.84	3,127.67	7,381.92	2,180.26	3,453.37
σ	59.83	55.56	21.75	52.29	76.68	107.59	189.40	55.93	85.92	46.69	58.77
Median	233.50	209	120	237.50	161.50	427	337	287.50	644	314	243
Min	147	143	77	163	79	278	173	182	441	167	144
25%	194.50	170	108.50	217.75	114	366.75	291.50	256.75	591.75	288.50	225.75
75%	276	268.25	132	278.25	231.75	509.50	391.25	319	697	336	285
IQR	81,50	98,25	23,50	60,50	117,75	142,75	99,75	62,25	105,25	47,50	59,25
Max	411	342	173	379	350	728	1053	414	861	396	458
Range	264	199	96	216	271	450	880	232	420	229	314
Kurtosis	2.95	1.97	2.81	2.67	2.27	3.09	5.68	2.55	2.94	3.72	4.84
Skewness	0.6860	0.3553	0.2913	0.5573	0.6193	0.7359	1.79	-0.0032	-0.1460	-0.5746	1.01
Sum Ann.	12,538	11,422	6,275	13,003	9,353	22,974	20,471	15,029	33,325	16,029	13,328

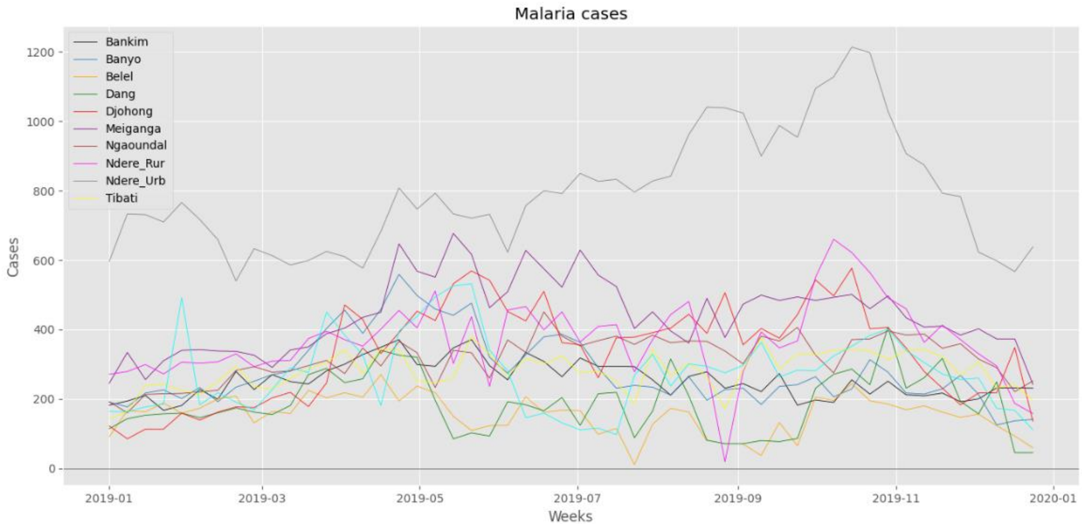

Figure 6. Health Districts data for year 2019.

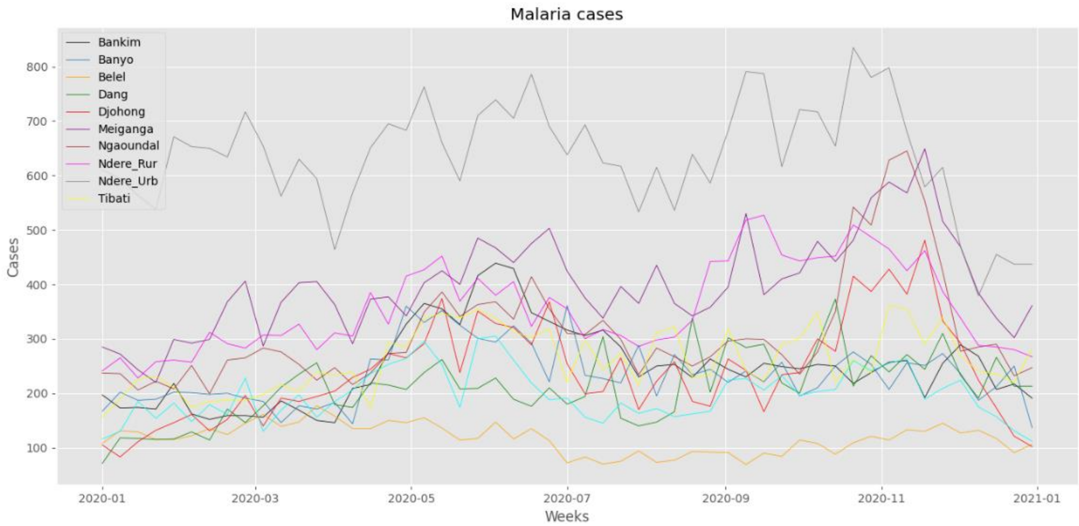

Figure 7. Health Districts data for year 2020.

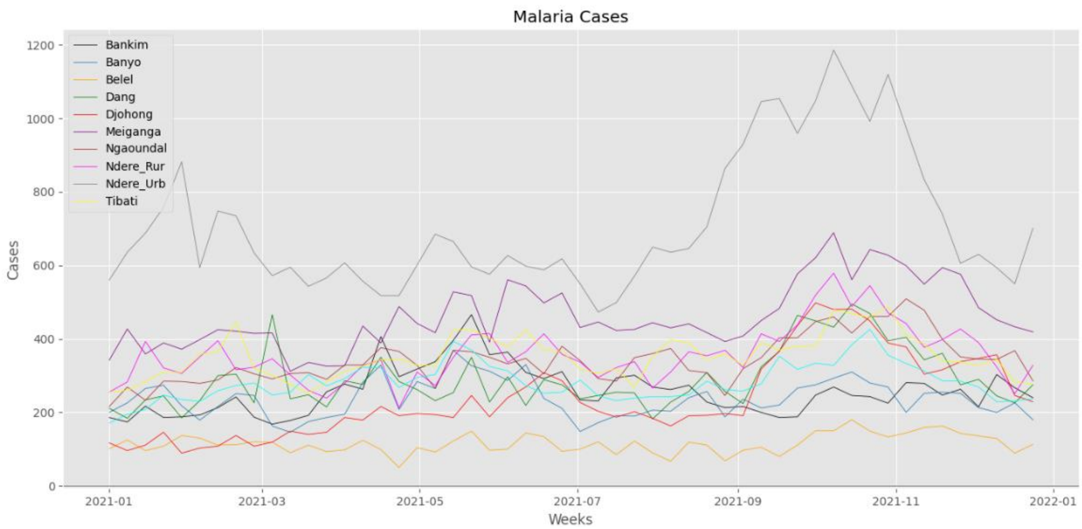

Figure 8. Health Districts data for year 2021.

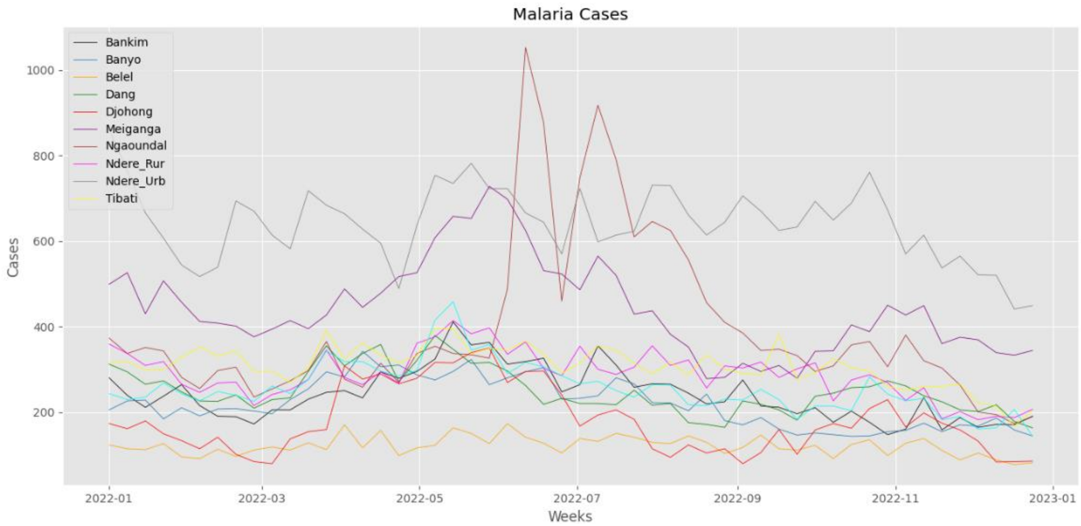

Figure 9. Health Districts data for year 2022.