Climate drivers of malaria transmission seasonality and their relative importance in Sub-Saharan Africa

Edmund I. Yamba[1], Andreas H. Fink[2], Kingsley Badu[3], Ernest O. Asare[4]Adrian M. Tompkins[5]and Leonard K. Amekudzi[1]

[1]Department of Meteorology and Climate Science, Kwame Nkrumah University of Science and Technology (KNUST), Kumasi-Ghana
[2]Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany
[3]Department of Theoretical and Applied Biology, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
[4]Department of Epidemiology of Microbial Diseases, Yale School of Public Health, Yale University, New Haven, CT, USA
[5]International Centre for Theoretical Physics, Earth System Physics, Trieste, Italy

Key Points:

- Seasonal malaria transmission in Sub-Saharan Africa is sustained at temperatures well above 15°C or below 40°C.
- Monthly maximum rainfall for seasonal malaria transmission should not exceed 600 mm.
- Rainfall and temperature are significant drivers of malaria seasonality in all parts of Sub-Saharan Africa except in west Central Africa.
- Topography has significant influence on which climate variable is an important driver of malaria seasonality in East Africa.
- Malaria transmission onset lags behind rainfall only at markedly seasonal rainfall areas, otherwise, malaria transmission is year-round.

Corresponding author: Edmund I. Yamba, `eyilimoan48@gmail.com`

Abstract

A new database of the Entomological Inoculation Rate (EIR) is used to directly link the risk of infectious mosquito bites to climate in Sub-Saharan Africa. Applying a statistical mixed model framework to high-quality monthly EIR measurements collected from field campaigns in Sub-Saharan Africa, we analyzed the impact of rainfall and temperature seasonality on EIR seasonality and determined important climate drivers of malaria seasonality across varied climate settings in the region. We observed that seasonal malaria transmission was within a temperature window of 15°C – 40°C and was sustained if average temperature was well above 15°C or below 40°C. Monthly maximum rainfall for seasonal malaria transmission did not exceed 600 mm in west Central Africa, and 400 mm in the Sahel, Guinea Savannah and East Africa. Based on a multi-regression model approach, rainfall and temperature seasonality were found to be significantly associated with malaria seasonality in all parts of Sub-Saharan Africa except in west Central Africa. Topography was found to have significant influence on which climate variable is an important determinant of malaria seasonality in East Africa. Seasonal malaria transmission onset lags behind rainfall only at markedly seasonal rainfall areas such as Sahel and East Africa; elsewhere, malaria transmission is year-round. High-quality EIR measurements can usefully supplement established metrics for seasonal malaria. The study's outcome is important for the improvement and validation of weather-driven dynamical mathematical malaria models that directly simulate EIR. Our results can contribute to the development of malaria models fit-for-purpose to support health decision-making in the fight to control or eliminate malaria in Sub-Saharan Africa.

1 Introduction

Sub-Saharan Africa remains the world's region with the greatest malaria burden despite massive efforts over the past decades to lower or eliminate malaria (WHO, 2020). Though poor health care systems and low socio-economic status (Degarege et al., 2019; Yadav et al., 2014) are contributing factors, the climate suitability of the region for malaria transmission has a major influence (Caminade et al., 2014). Generally, climate variables such as temperature, rainfall and relative humidity are known to have significant influence on the development and survival of both the malaria parasites and their vectors. Malaria parasite development is not possible at temperatures below 16°C and temperatures above 40°C have adverse effect on mosquito population turnover (Parham and Michael, 2010; Mordecai et al., 2013; Blanford et al., 2013; Shapiro et al., 2017). Rainfall provides the environment for vector breeding (Ermert et al., 2011; Tompkins and Ermert, 2013; Kar et al., 2014) and relative humidity of at least 60% appears necessary for vector survival (Thompson et al., 2005). Rainfall therefore affects the availability, persistence and dimensions of Anopheles vectors and their larval habitats (Fournet et al., 2010; Afrane et al., 2012a; Boyce et al., 2016; Asare et al., 2016a). Previous work studying the relationship between sporozoite development and the survival of infectious mosquitoes found optimal temperatures for efficient malaria transmission between 25°C and 27°C (Bayoh, 2001; Lunde et al., 2013a,b). In Sub-Saharan Africa, most countries have annual mean temperatures between 20°C and 28°C (Lunde et al., 2013a). Given Sub-Saharan Africa's warm tropical climate, a plethora of efficient and effective malaria parasite and vectors thrives in this setting (Sinka et al., 2010; Murray et al., 2012). Understanding the relative importance of climate drivers of malaria seasonality is crucial for describing the geographic patterns of the heterogeneous risk and burden of malaria across the sub-region (Gething et al., 2011; Reiner et al., 2015). This could translate to substantial public health gains, taking into account the seasonality in malaria control and prevention interventions, by helping to determine when, where and how to apply vector and parasite control measures.

To our knowledge, there are insufficient field studies using Entomological Inoculation Rate (EIR, defined as the number of infectious mosquito bites person receives per time) data to relate climate to malaria seasonality in Sub-Saharan Africa. Mabaso et al. (2007) assessed the relationship between EIR seasonality and environmental variables in Africa using a rainfall seasonality index (Markham, 1970). The index fails to capture seasonality at areas with bimodal rainfall regimes, however. Furthermore, the study did not take into consideration the impact of diverse climatic conditions on seasonality outcomes but aggregated data from sites of different climate and environmental settings into a single study, which has the potential to skew the results. Other research has examined the link between malaria and climate variables but primarily relied on clinical data or malaria suitability indices (Lowe et al., 2013; Midekisa et al., 2015; Komen et al., 2015). Both malaria indices and case data have drawbacks for studying malaria seasonality.

Malaria indices are derived using statistical relationships between weather and malaria measures and their out-of-sample generalization over space and time for seasonality studies is subject to significant uncertainties. Clinical case data are also subject to significant uncertainties due to the inaccurate diagnostics (often counts of suspected cases, with temporal inconsistency in the use of Rapid Diagnostic Test, RDT or slide analysis) and under-counting due to varying health-seeking behaviour and health policies (Afrane et al., 2012b). Given that the biology of the malaria parasite and its vector mosquito are temperature and rainfall dependent (Ermert et al., 2011), and that EIR can directly quantify parasite-infected mosquitoes and their propensity to transmit the parasites to humans (MARA, 1998; Shaukat et al., 2010) or estimate the seasonality of the exposure of a population to malaria parasite inoculations (Beier et al., 1999; Takken and Lindsay, 2003), then EIR should be able to usefully relate climate to malaria seasonality better than malaria cases.

In this study, therefore, we investigated the impact of climate variables on EIR seasonality in diverse climate settings across Sub-Saharan Africa with the goal of identifying significant climate determinants of malaria seasonality, their relative importance and variability across the region. To our knowledge, this is the first study to use $\mathrm{EIR_m}$ to explore the impact of climatic variables on malaria seasonality in Sub-Saharan Africa on this wider scale. We applied a mixed model statistiacal framework to a high-quality malaria EIR data (Yamba et al., 2018; Yamba et al., 2020) gathered from publicly available field campaigns of sufficient duration and determined the climate effect that explained significant variations in EIR seasonality. Our findings are intended to provide an understanding of geographical heterogeneous malaria risk from climate effect and support future malaria modeling and forecasting efforts. It will contribute to the development of malaria models especially weather-driven dynamical malaria models fit-for-purpose to support health decision-making in the fight to control or eliminate malaria in Sub-Saharan Africa.

2 Data and Methods

2.1 Study Area

The study area includes locations in Sub-Saharan Africa (as shown in Figure 1), where mosquitoes have previously been collected for malariometrics such as Human Biting Rate (HBR), CircumSporozoite Protein Rates (CSPR), and EIR. The geographical coordinates and elevation of each location are detailed in Tables S1 to S4. The study locations are grouped into four distinct climate zones namely Sahel, Guinea, WCA, and EA (see Figure 1). Each zone has a unique climate conditions from others (see Figure S1) and therefore have different climate implications on malaria seasonality (Yamba, 2016). The division into zones is, therefore, to ensure that malaria transmission patterns are consistent across geographical areas with similar climate characteristics. The seasonal distribution of rainfall

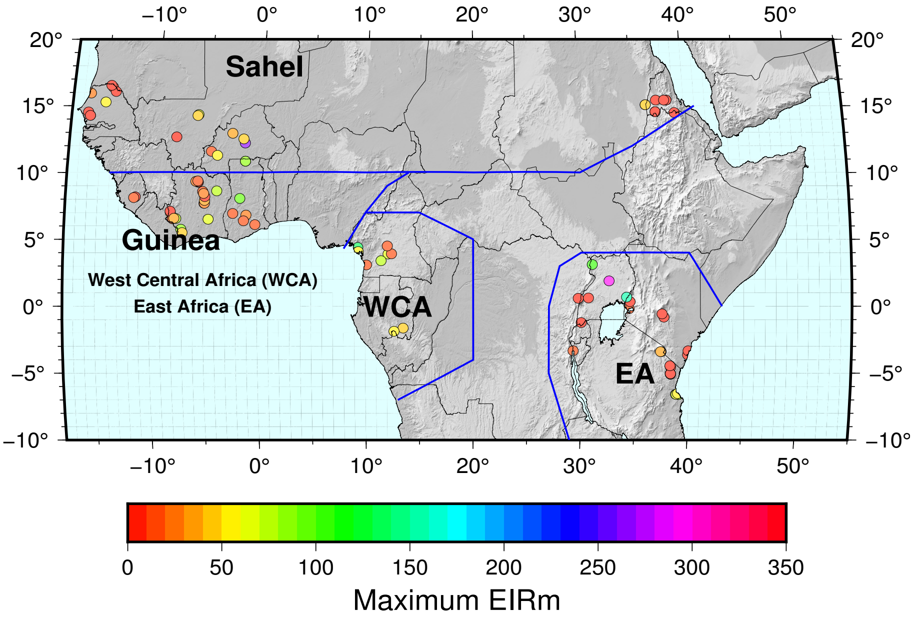


Figure 1: The map of the Sub-Saharan Africa showing field survey sites for EIR. The colour gradient of each site show the maximum EIR available. The blue lines delineate the region into climate zones of Sahel, Guinea, WCA and EA.

and temperature for each zone is shown in Figure S1. In the Sahel, rainfall is markedly seasonal, with a single wet season (usually June to October) and a protracted dry season (November to May). Seasonal temperature ranges between a minimum value of 20 ^{o}C during the harmattan season and to a maximum of about 40 ^{o}C during the pre-monsoon season. In general, temperatures are higher in the Sahel and colder in EA due to the fact that most areas are characterized by higher altitudes.

2.2 Data

2.2.1 Monthly EIR data

Monthly malaria EIR data (hereafter referred to as $\mathrm{EIR_m}$) were obtained from a newly compiled and published monthly malaria EIR database (Yamba et al., 2018; Yamba et al., 2020) for each study location shown in Figure 1. The years and months for which the $\mathrm{EIR_m}$ data were available for each study location is shown in Table S1-S4. Generally, most locations had 12 months of data while other locations had data varying between 24 and 36 months. The data also spanned the period 1983-2013 for all locations. The temporal duration of the data is mostly limited to one year because sampling mosquitoes for EIR is extremely capital and labour intensive (Kilama et al., 2014; Tusting et al., 2014; Badu et al., 2013). The EIR database from which data were extracted for use in this work is a comprehensive one. It was constructed through an all-inclusive literature review using google scholar and PubMed search facilities. All data in that database was generated from publicly available field campaigns of adequate duration and is freely available for public usage in the PANGAEA repository (Yamba et al., 2018). Details of how this database was constructed including compilation, sources, recording, spatial coverage and temporal resolutions are clearly described in Yamba et al. (2020).

2.2.2 Meteorological data

Monthly rainfall (RR) and temperature (minimum (T_{min}), mean (T_{mean}) and maximum (T_{max})) data for each study location were gathered. Rainfall data was obtained from the Global Precipitation Climatology Centre (GPCC) product, version 2018 (Schneider et al., 2018). The GPCC data is a gridded gauge-analysis products and available globally from 1891-2016 at a spatial resolution of 0.25°. GPCC was chosen because it is a rain gauge-analysis product built from quality-controlled rainfall data from ground-based weather stations. Previous validation studies (Manzanas et al., 2014; Atiah et al., 2020) have also found it to be reliable and consistent with ground-based weather observations. The temperature data was obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis, 5th generation (ERA5) (Hersbach et al., 2020). ERA5 is also a gridded re-analysis product and available globally on an hourly time scale from 1979 to present at a high spatial resolution of 0.25° by 0.25°. ERA5 was chosen because previous evaluation studies of the product (Tarek et al., 2020; Gleixner et al., 2020; Oses et al., 2020) have widely recommended it for meteorological research. RR, T_{min} and T_{max} were extracted from the respective database for each study location using the nearest grid point of the location's geographical coordinates. T_{mean} values were estimated by averaging the T_{min} and T_{max} values for the location. The extracted temperature and rainfall data had to also conform with the exact years and months at which EIR_m data were available for each location. The study relied on GPCC and ERA5 because, ground-based local weather stations from which these data could be gathered were mostly not available at the EIR sites or, if present, often have sparse data.

2.3 Data analysis

The analysis was conducted for each classified zone as shown in Fig 1. EIR data from locations characterized with the presence of permanent water bodies and/or irrigation activities were exempted. Irrigation and permanent water bodies (such as damps, rivers, streams, swamps etc) have significant influence on the intensity and length of seasonal malaria transmission (Ermert et al., 2011; Tompkins and Ermert, 2013; Asare et al., 2016b; Asare and Amekudzi, 2017). Their exclusion was, therefore, a means to dissociate the influence of these hydrological parameters on malaria seasonality and reducing the impact to climate factors alone.

2.3.1 pair-wise comparison

The study examined the ranges of RR, $\mathrm{T_{min}}$, $\mathrm{T_{mean}}$ and $\mathrm{T_{max}}$ at which $\mathrm{EIR_m}$ occurred using a simple pair-wise comparison approach. This was done by first aggregating the $\mathrm{EIR_m}$ data from all locations within each zone into a single time series of 12 months irrespective of the year of availability. Similarly, the corresponding RR, $\mathrm{T_{min}}$, $\mathrm{T_{mean}}$ and $\mathrm{T_{max}}$ data were also aggregated. The aggregated monthly timeseries of RR, $\mathrm{T_{min}}$, $\mathrm{T_{mean}}$, $\mathrm{T_{max}}$ and $\mathrm{EIR_m}$ were then matched head-to-head as shown in Figure 2. The ranges of RR, $\mathrm{T_{min}}$, $\mathrm{T_{mean}}$ and $\mathrm{T_{max}}$ at which EIR occurred were then determined for each zone .

2.3.2 Relative importance of climate predictors

The relative importance of RR, $\mathrm{T_{min}}$, $\mathrm{T_{mean}}$ and $\mathrm{T_{max}}$ in predicting $\mathrm{EIR_m}$ for each climate zone was analysed using a multiple regression model of the form:

$$EIR_{\mathrm{m}} \sim RR + T_{\mathrm{max}} + T_{\mathrm{min}} + T_{\mathrm{mean}} \tag{1}$$

where $\mathrm{EIR_m}$ is the response variable and RR, $\mathrm{T_{min}}$, $\mathrm{T_{mean}}$ and $\mathrm{T_{max}}$ are the predictors. The contribution of each individual predictor to $\mathrm{EIR_m}$ outcome was then quantified (see Table 1 and 2). Each regressor's contribution was considered as the R^2 from univariate regression, and all univariate R^2 values add up to the full model R^2 (Grömping, 2007). The R package "relaimpo" (Grömping, 2007) was utilized for the calculation of the contribution of the reqressors in the model. It implements six different metrics for assessing relative importance of regressors namely:first, last, pratt, betasg, lmg and pmvd. Among these, lmg and pmvd are computer intensive and has advantage over others in the sense that they decompose R^2 into non-negative contributions that automatically sum to the total R^2

(Grömping, 2007). In this study, lmg was invoked since pmvd is patent protected. The lmg calculates the relative contribution of each predictor to the R^2 with the consideration of the sequence of predictors appearing in the model. It intuitively decomposes the total R^2 by adding the predictors to the regression model sequentially. Then, the increased R^2 is considered as the contribution by the predictor just added. The following are mathematical descriptions of lmg metric referenced from Grömping (2007):
For a model with regressors in set S, the R^2 is given as:

$$R^2(S) = \frac{ModelSS(modelwithregressorsinS)}{TotalSS} \tag{2}$$

To add regressors in set M to a model with the regressors in set S, the additional R^2 is given as:

$$seqR^2(M|S) = R^2(MUS) - R^2(S) \tag{3}$$

where the order of the regressors is a permutation of the available regressors $x_1, ..., x_{\mathrm{p}}$ denoted by the tuple of indices $r = (r_1, ..., r_{\mathrm{p}})$. Let $S_{\mathrm{k}}(r)$ denote the set of regressors entered into the model before regressor x_{k} in the order r. Then the portion of R^2 allocated to regressor x_{k} in the order r can be written as:

$$seqR^2(\{x_{\mathrm{k}}\}|S_{\mathrm{k}}(r)) = R^2(\{x_{\mathrm{k}}\}US_{\mathrm{k}}(r)) - R^2(S_{\mathrm{k}}(r)) \tag{4}$$

With eq. 4, the metric lmg (in formulae denoted as LMG) can be written as:

$$LMG(x_{\mathrm{k}}) = \frac{1}{P!} \sum_{rpermutation} seqR^2(\{x_{\mathrm{k}}\}|r) \tag{5}$$

Orders with the same $S_{\mathrm{k}}(r) = S$ can be summarized into one summand, which simplifies the formula into:

$$LMG(x_{\mathrm{k}}) = \frac{1}{p!} \sum_{S \subseteq \{x_1, ..., x_{\mathrm{p}}\} \backslash \{x_{\mathrm{k}}\}} seqR^2(\{x_{\mathrm{k}}\}|S) \tag{6}$$

The analysis also assessed the relative importance of each regressor (in eqn 1) by looking at what each regressor alone is able to explain (i.e., comparing the R^2 value of regression model with one regressor only without considering the dependence of others as is the case of the metric *lmg*). The metric *first* in the "relaimpo" package was invoked for this purpose because, unlike *lmg*, it is completely ignorant of the other regressors in the model and so no adjustment takes place (Grömping, 2007). Since *first* does not decompose R^2 into contributions like *lmg*), the contribution of the individual regressors alone do not naturally add up to the overall R^2. The sum of these individual contributions is often far higher than the overall R^2 of the model with all regressors together.

Whether *lmg* or *first*, each metric's outcome were bootstrapped to ensure that the relative importance of each regressor was clearly defined (i.e. those different and those that are similar in terms of relative importance). Bootstrapping in "relaimpo" was done using the function boot in the package. Prior to calculating the *lmg* and *first* metrics, all data series (i.e. EIR_m, RR, T_{min}, T_{mean} and T_{max} timeseries) were log transformed. The essence of the log transformation was to decrease the variabilities in the data pairs and make them conform more closely to normal distribution with similar variance and standard deviation (Curran-Everett, 2018).

2.3.3 EIR lag behind rainfall

Seasonal malaria transmission onset lags behind rainfall season onset because of the time taken for mosquito breeding and vector population growth after rainfall season onset (Tompkins and Ermert, 2013; Badu et al., 2013; Asare and Amekudzi, 2017). This lag time as influenced by climate and whether it varies from one climate zone to another is not known. In this analysis, we quantified this lag time for each climate zone using a cross-correlation statistics performed between RR and EIR_m data pairs. In this statistics, RR was treated as the predictor variable and the corresponding EIR_m as the response variable. The pairs were then cross-correlated at lags of -5 to 0 months and the correlation co-efficient at each lag was calculated. The lag with the strongest positive correlation coefficients was identified as the optimum period of delay between rainfall onset and the EIR season for the zone.

3 Results

3.1 pair-wise comparison

Figure 2 shows the EIR_m response ranges of pairs of rainfall (RR) and temperature (T_{min}, T_{mean} and T_{max}). In the Sahel, maximum rainfall (RR) ranges were about 400 mm per month. Temperature ranges generally varied between 20°C – 40°C in this zone. T_{max} ranges were clustered between 25°C – 40°C, T_{min} within 20°C – 30°C and T_{mean} observed within 25°C – 35°C. In Guinea, RR ranges were also centered around 400 mm per month. Temperature response ranges were mostly observed within 25°C – 35°C for T_{max}, 20°C – 25°C for T_{min} and 24°C – 30°C for T_{mean}. In WCA, maximum RR ranges were centered at about 600 mm per month, which is higher compared to ranges observed in the Sahel, Guinea and EA. Temperature response ranges in this zone were slightly lower than observed in the

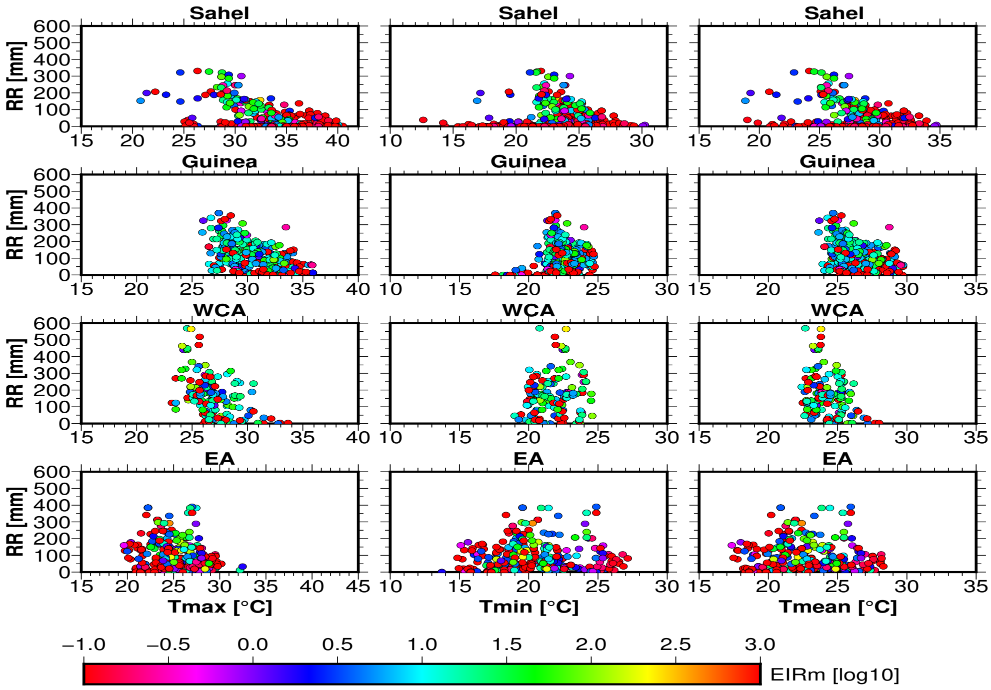


Figure 2: A pair-wise comparison showing the ranges of RR, T_{min}, T_{mean} and T_{max} at which EIR_m occurs. The coloured circles shows log transformed EIR_m values.

Sahel and Guinea. These include 24°C – 32°C for T_{max}, 20°C – 25°C for T_{min} and 22°C – 27°C for T_{mean}. The EA maximum RR ranges were also about 400mm. Temperature ranges of 20°C – 30°C for T_{max}, 15°C – 27°C for T_{min} and 18°C – 29°C for T_{mean} were observed.

3.2 Relative importance of climate predictors

In Table 1 and 2, the relative importance of climate variables in predicting EIR_m is presented for locations with elevations $\leq$ 500 m and $>$ 1000 m respectively. The predictors with p-value ≤ 0.05 were considered significant and interpreted that the respective climate variable significantly predicted the EIR seasonality in that zone. At lower elevations ($\leq$ 500 m) in Sahel, rainfall and temperature were all significant drivers of EIR seasonality cumulatively contributing about 30.72% of the variations in EIR seasonality. At these lower elevation areas, important predictors of EIR_m seasonality were RR and T_{max}. At higher elevations ($>$ 1000 m), rainfall and temperature are together responsible for about 40% of the variations in EIR_m with insignificant contribution from T_{min}. Like the Sahel, temperature and rainfall

Table 1: **The relative contribution of RR, T_{min}, T_{mean} and T_{max} in predicting EIR_m** bootstrapped at confidence interval of 95% for locations with elevations $\leq$ 500 m. Variables with significant p-values contributions are boldfaced. R^2 represents the total proportion of variance in EIR explained by all the climate predictors. lmg values show the individual contribution of each predictor to R^2 relative to others. *First* is the contribution of each predictor alone to R^2 with complete ignorance of the others.

Zone	R^2 [%]	Variable	lmg [%]	First [%]	Coefficient (R)	P-value
Sahel	**30.72**	**RR**	**7.73**	**15.76**	0.3497	**0.0000**
		$\mathbf{T_{max}}$	**12.03**	**17.54**	-15.7276	**0.0000**
		T_{min}	4.89	1.79	3.6380	0.0138
		$\mathbf{T_{mean}}$	**6.07**	**3.20**	-6.8674	**0.0033**
Guinea	**13.59**	**RR**	**5.85**	**10.22**	0.4848	**0.0000**
		$\mathbf{T_{max}}$	**4.09**	**9.65**	-13.3808	**0.0000**
		T_{min}	0.64	0.19	2.7410	0.3760
		$\mathbf{T_{mean}}$	**3.01**	**6.38**	-15.7753	**0.0000**
WCA	**1.69**	RR	0.34	0.23	0.0974	0.5550
		T_{max}	0.60	0.95	5.4640	0.3770
		T_{min}	0.42	0.49	-4.5700	0.5810
		T_{mean}	0.33	0.35	6.8450	0.5280
EA	**31.83**	RR	0.62	0.00	-0.0141	0.9360
		$\mathbf{T_{max}}$	**10.23**	**26.50**	-23.2210	**0.0000**
		$\mathbf{T_{min}}$	**8.84**	**20.10**	-12.1120	**0.0000**
		$\mathbf{T_{mean}}$	**12.14**	**26.04**	-18.2160	**0.0000**

were also significant determinants of EIR_m at lower elevations ($\leq$ 500 m) in Guinea just that their contribution to EIR_m variations is small (about 13.59%) compared to that of Sahel (about 30.72%). In Guinea, also, EIR_m data were unavailable for locations > 1000 m for further analysis in this regard. In WCA, rainfall and temperature were insignificantly associated with EIR_m seasonality whether a lower or higher elevations. Their percentage explanation of the variation in EIR_m were also low (extremely low at lower elevation areas and slighly higher for higher elevation areas) compared to other climate zones. In EA,

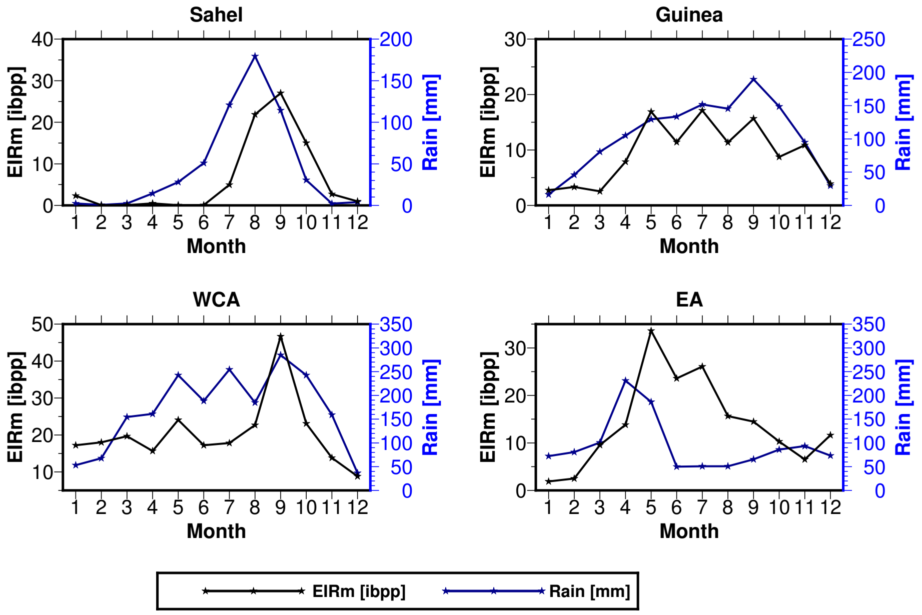
Table 2: **Same as Table 1 but for locations with elevations > 1000 m**. In Guinea, EIR data were unavailable for locations at this elevation hence represented as dashed lines.

Zone	R^2 [%]	Variable	lmg [%]	First [%]	Coefficient (R)	P-value
Sahel	40.47	**RR**	7.43	**14.83**	0.3745	**0.0780**
		T_{max}	**17.66**	**35.91**	-10.8070	**0.0036**
		T_{min}	3.69	4.17	-1.4543	0.5950
		T_{mean}	**11.69**	**23.40**	-7.7660	**0.0513**
Guinea	-	RR	-	-	-	-
		T_{max}	-	-	-	-
		T_{min}	-	-	-	-
		T_{mean}	-	-	-	-
WCA	16.55	RR	1.41	1.25	-0.0844	0.7653
		T_{max}	6.70	0.23	6.1700	0.6620
		T_{min}	1.53	0.39	14.1000	0.5970
		T_{mean}	6.91	1.32	12.2800	0.5300
EA	18.22	**RR**	**10.44**	**13.37**	0.5510	**0.0000**
		T_{max}	**1.82**	**3.94**	8.1570	**0.0289**
		T_{min}	**2.88**	**7.77**	10.2080	**0.0011**
		T_{mean}	**3.08**	**6.95**	11.0460	**0.0026**

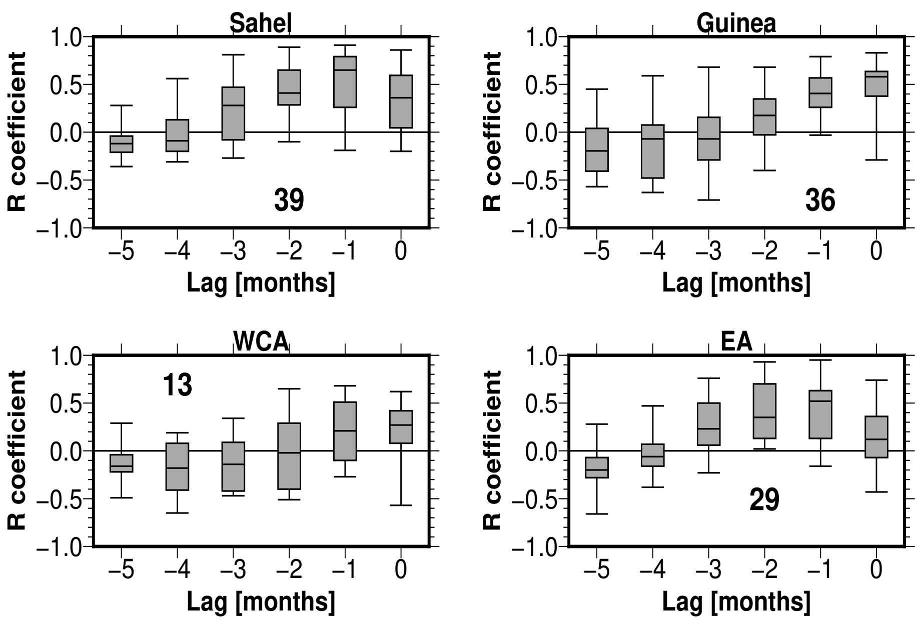
temperature variables (T_{min}, T_{mean} and T_{max}) were the significant drivers of EIR seasonality at locations $\leq$ 500 m. It explained about 31% of the seasonality in EIR_m in these areas with extremely insignificant contribution from rainfall. But at areas > 1000 m, all the climate variables were significant contributors with rainfall showing higher contribution to EIR_m variation than temperature.

3.3 EIR lag behind rainfall

Figure 3 shows the seasonal relationship between rainfall and EIR_m (see Figure 3a) and the lag between rainfall onset and EIR onset (see Figure 3b). It is observed in Figure 3a that EIR_m is positively correlated with rainfall. A lag period of 1 month is observed in Sahel



(a) Average monthly time series of EIR_m and rainfall



(b) The box-and-whisker plot of cross-correlation coefficient between Rainfall and EIR_m at different lag period

Figure 3: The correlation between rainfall and EIR_m. The numbers 39, 36, 13 and 29 shows the number of location observations contributing to the average timeseries (a) and the box-and-whisker plots (b).

and EA but zero month in Guinea and WCA. Similarly, the cross-correlation statistics determining the lag between the onset of rainy season and the start of the EIR season are shown in Figure 3b. Again, it is observed that the lag at which $\mathrm{EIR_m}$ seasonality strongly and positively correlated with rainfall was 1 month in the Sahel and EA but zero month in Guinea and WCA.

4 Discussion

Our study first examined the seasonal ranges of rainfall and temperature at which $\mathrm{EIR_m}$ occurred in a pair-wise comparison study. In general, temperature ranges of $\mathrm{EIR_m}$ response were mostly clustered between a minimum of 15°C and a maximum of 40°C. This outcome suggests that seasonal malaria transmission is barely impossible below 15°C or above 40°C. Previous studies (Shapiro et al., 2017; Parham and Michael, 2010; Lunde et al., 2013a; Mordecai et al., 2013) have indicated that malaria parasite development is not possible at temperatures below 16°C and that temperatures above 40°C have adverse effect on mosquito population turnover. The outcome of our study using $\mathrm{EIR_m}$ corroborates these previous findings. It provides an additional justification that the number of infectious mosquito bites a person receives per time are associated with changes in temperature. While $\mathrm{T_{min}}$ may be below 16°C as observed in the Sahel and EA (see Figure 2), the daily $\mathrm{T_{mean}}$ must be greater than 16°C particularly for the anopheles mosquitoes for transmission to occur. It should also be significantly less than 40°C for anopheles mosquitoes to survive thermal stress and possible death if seasonal transmission has to take place. Similarly, maximum monthly rainfall values for $\mathrm{EIR_m}$ occurrence was 600 mm in WCA but 400 m in the Sahel, Guinea and EA. The higher monthly maximum rainfall in WCA is due to the fact that annual total rainfall is mostly higher in this region than others (Nicholson, 2013; Froidurot and Diedhiou, 2017). Previous works (Craig et al., 1999; Ermert et al., 2011) have demonstrated that the least monthly amount of rainfall required for malaria transmission is about 80 mm. Our findings suggest that the monthly maximum limit required for seasonal malaria transmission should be about 600 mm in WCA but 400 mm in Sahel, Guinea and EA. Excess of these thresholds could result in flooding of breeding grounds and flushing out and killing the water-bound stage vectors (Paaijmans et al., 2010; Ermert et al., 2011).

The evaluation of the relative importance of RR, $\mathrm{T_{min}}$, $\mathrm{T_{mean}}$ and $\mathrm{T_{max}}$ in predicting EIR seasonality (see details in Table 1 and 2) revealed climate variables that were significantly associated with EIR seasonality in Sub-Saharan Africa. These climate variables are observed as the drivers of malaria seasonality in those zones of the sub-region. The climate variables with highest contribution to EIR variance in each zone are attributed as the most significant drivers. This means that any changes in these significant drivers can result in a substantial changes in malaria seasonality in those areas. Elevation or topography was also observed to play a significant role in determining the important climate drivers of seasonal malaria transmission. In EA for instance, temperature was the important determinant of EIR seasonality at lower elevated areas ($\leq$ 500 m). On the contrary, both rainfall and temperature significantly influenced $\mathrm{EIR_m}$ seasonality at higher elevated areas (>1000 m). Though temperature and rainfall are important factors in malaria transmission, our study does not find them to have any significant association with EIR seasonality in WCA. This suggest that malaria seasonality in this zone is importantly driven by other factors other than climate. This require additional studies to unravel these factors driving malaria seasonality in this zone. Mabaso et al. (2007) predicted EIR seasonality from environmental data and found that seasonality in rainfall, minimum temperature, and irrigation were important determinants of seasonality in EIR in Sub-Saharan Africa. Though this study outcome is important, it is not climate specific as it does not justify the implications of diverse climate conditions on EIR seasonality as demonstrated in this study. Other studies (Mabaso et al., 2006; Simple et al., 2018) have used malaria case records from hospitals and found significant correlation between rainfall and temperature. As stated in the introduction, malaria case records have drawbacks for studying malaria seasonality as they are subject to significant uncertainties due to the inaccurate diagnostics and under counting due to varying health-seeking behaviour and health policies (Afrane et al., 2012b).

The cross-correlation statistics showed the lag(s) at which rainfall strongly correlated with $\mathrm{EIR_m}$ in each zone. The lag period suggest the time taken for malaria season to start after rainfall season has started. The lag of 1 month in Sahel and EA signifies that malaria transmission season delays 1 month after the start of rainfall season at these zones. In Guinea and WCA, this lag period was zero month suggesting that there is no delay between rainfall season onset and the start of malaria season. Hence malaria transmission in these zones is year-round. In markedly seasonal rainfall zones such as the Sahel and EA, the delay between rainfall onset and the start of the malaria season is expected. Rainfall in the Sahel is markedly seasonal, lasting from June to October, followed by about six to eight months of dry period (Nicholson, 2013; Froidurot and Diedhiou, 2017). Hence, mosquitoes are barely present during the dry and long hot season. Even if present, they are inactive due to low humidity and high temperature and only recover within the rainy season when rainfall and temperature requirements are suitable. The absence of delay between rainfall season onset and the start of malaria season at Guinea and WCA is also expected. These zones are highly humid with shorter dry seasons (Nicholson, 2013; Froidurot and Diedhiou, 2017). For this reason, vectors are able to persist all year round at these zones resulting in year-round transmission at these areas. Previous studies (Simple et al., 2018; Tompkins and Di Giuseppe, 2015; Reiner et al., 2015; Ikeda et al., 2017) have reported malaria lagging behind rainfall at about 1 to 2 months but our study has further demonstrated that malaria season onset may lag behind rainfall only at markedly seasonal rainfall areas in Sub-Saharan Africa.

5 Conclusion

Clinical malaria case data is commonly utilized as a malariometric in examining the relationship between climate and seasonal malaria transmission in Sub-Saharan Africa. This data, on the other hand, is fraught with uncertainty due to out-of-sample generalization over geography and time, erroneous diagnosis, and under-counting due to varying health-seeking behavior and policy. As a result, in this work, we explored the applicability of high-quality EIR measurements to link rainfall and temperature seasonality to seasonal malaria outcomes in Sub-Saharan Africa. The main goal was to determine the climate variables that significantly drives malaria seasonality and their relative importance in the sub-region. Sub-Saharan Africa was first divided into four distinct climate zones namely Sahel, Guinea, WCA, and EA. The division was necessary because each zone has a unique climate conditions and therefore will have different climate implications on malaria seasonality. Applying a multi-regression statistics, pair-wise comparison and cross-correlation approaches to a $\mathrm{EIR_m}$ database gathered from publicly available field campaigns for each zone, the climate variables that explained significant variations in EIR seasonality were determined.

Findings in this study affirmed previous understanding that seasonal malaria transmission is barely impossible below 16°C or above 40°C temperature threshold (Shapiro et al., 2017; Mordecai et al., 2013). Hence, for seasonal malaria transmission to be sustained, average temperature should be well above the minimum or well below maximum threshold. While previous studies (Craig et al., 1999; Ermert et al., 2011) suggest that the monthly minimum rainfall requirement for seasonal transmission is about 80 mm, our study observed monthly maximum rainfall limit should be about 600 mm in WCA, and 400 mm in the Sahel, Guinea and EA. While rainfall and temperature were found to be significantly associated with EIR_m seasonality in the Sahel, Guinea and EA, they were not important drivers of malaria seasonality in WCA. Important drivers of malaria seasonality in WCA may be due to other factors other than climate variables. In zones characterized by elevations such as EA, topography has a significant influence on which variable is an important determinant of malaria seasonality. At markedly seasonal rainfall areas such as Sahel and EA, malaria seasonal starts one month later after the rainfall season has started. However, for zones where rainfall season is bimodal such as Guinea and WCA, there is no delay between rainfall season onset and malaria season onset.

In this study, therefore, we showed that high quality EIR_m measurements can usefully supplement established metrics for seasonal malaria by demonstrating evidence for the use of EIR to directly link the risk of humans to infectious mosquito bites to climate. The study informs our understanding of the connection between climate variables and both the malaria vector and parasite biology and how that translates into malaria seasonality in Sub-Saharan Africa. This information is key for the improvement and validation of weather-driven dynamical mathematical malaria models that directly simulate EIR. Our findings provide an understanding of geographical heterogeneous malaria risk from climate effect and support future malaria modeling and forecasting efforts. The study also supplements previous works describing clinical patterns of malaria infection and morbidity. Taking into account the seasonality of malaria management, findings in this study could lead to significant public health advantages by assisting in determining when, where, and how to use vector and parasite control strategies. It can, therefore, help stakeholders establish a robust framework for monitoring, forecasting and control of malaria.

This study does not claim to have identified all the EIR_m data available across sub-Saharan Africa. It relied on EIR_m data available in repository (Yamba et al., 2018) with details explained in (Yamba et al., 2020). The study also acknowledges that the observed EIR_m data were both spatially and temporally limited and thus unavailable for many settings (as shown if Figure 1). This limitation was unavoidable because sampling mosquitoes for the determination of EIR is both labour and cost intensive. Hence, it is very difficult to have EIR_m data available for many locations and for a long period of time. Future mosquito sampling should, therefore, focus on areas of unavailable data in order to consolidate the spatial homogeneity of available EIR_m data distribution. However, an important strength of this study is its restricted geographic and climate relevance. To our knowledge, this study is the first of its kind and also that $\mathrm{EIR_m}$ data has not been explored on such a wider scale in Sub-Saharan Africa. With the amount of $\mathrm{EIR_m}$ utilized for each climate zone, it is not anticipated that the inherent limitations may have any major adverse influence on the outcome of the study.

Acknowledgments

Profound gratitude to Katholischer Akademischer Ausländer-Dienst (KAAD) and the University of Cologne, Germany for the financial assistance.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

The work presented here was carried out collaboratively among all the authors. Edmund I. Yamba compiled the database, conducted the analysis for the figures and tables and drafted the manuscript. Leonark K. Amekudzi, Andreas H. Fink and Adrian M. Tompkins co-designed the project, supervised the analysis and co-authored the paper. Enerst O. Asare and Kingsley Badu contributed to result interpretation and co-authored the paper. All the authors proofred the paper.

References

Afrane, Y. A., Githeko, A. K., and Yan, G. (2012a). The ecology of anopheles mosquitoes under climate change: Case studies from the effects of environmental changes in east africa highlands. *Annals of the New York Academy of Sciences*, 1249:204.

Afrane, Y. A., Githeko, A. K., and Yan, G. (2012b). The ecology of anopheles mosquitoes under climate change: case studies from the effects of environmental changes in east africa highlands. *Annals of the New York Academy of Sciences*, 1249:204–210.

Asare, E. O. and Amekudzi, L. K. (2017). Assessing climate driven malaria variability in ghana using a regional scale dynamical model. *Climate*, 5(1):20.

Asare, E. O., Tompkins, A. M., Amekudzi, L. K., and Ermert, V. (2016a). A breeding site model for regional, dynamical malaria simulations evaluated using in situ temporary ponds observations. *Geospatial Health*, 11(s1):390.

Asare, E. O., Tompkins, A. M., and Bomblies, A. (2016b). A regional model for malaria vector developmental habitats evaluated using explicit, pond-resolving surface hydrology simulations. *PLoS One*, 11(3):e0150626.

Atiah, W. A., Tsidu, G. M., and Amekudzi, L. K. (2020). Investigating the merits of gauge and satellite rainfall data at local scales in ghana, west africa. *Weather and Climate Extremes*, 30:100292.

Badu, K., Brenya, R. C., Timmann, C., Garms, R., and Kruppa, T. F. (2013). Malaria transmission intensity and dynamics of clinical malaria incidence in a mountainous forest region of ghana. *MWJ*, 4:14.

Bayoh, M. N. (2001). *Studies on the development and survival of Anopheles gambiae sensu stricto at various temperatures and relative humidities*. PhD thesis, Durham University.

Beier, J. C., Killen, G. F., and Githure, J. I. (1999). A short report: Entomologic inoculation rate and plasmodium falciparum malaria prevalence in africa. *Am. J. Trop. Med. Hyg*, 61(1):109–113.

Blanford, J. I., Blanford, S., Crane, R. G., Mann, M. E., Paaijmans, K. P., Schreiber, K. V., and Thomas, M. B. (2013). Implications of temperature variation for malaria parasite development across africa. *Scientific reports*, 3(1):1–11.

Boyce, R., Reyes, R., Matte, M., Ntaro, M., Mulogo, E., Metlay, J. P., Band, L., and Siedner, M. J. (2016). Severe flooding and malaria transmission in the western ugandan highlands: implications for disease control in an era of global climate change. *The Journal of infectious diseases*, 214(9):1403–1410.

Caminade, C., Kovats, S., Rocklov, J., Tompkins, A. M., Morse, A. P., Colon-Gonzalez, F. J., Stenlund, H., Martens, P., and Lloyd, S. J. (2014). Impact of climate change on global malaria distribution. *PNAS*, 111(9):3286–3291.

Craig, M. H., Snow, R. W., and le Seur, D. (1999). A climate-based distribution model of malaria transmission in sub-saharan africa. *Parasitology Today*, 15(3):105–111.

Curran-Everett, D. (2018). Explorations in statistics: the log transformation. *Advances in physiology education*, 42(2):343–347.

Degarege, A., Fennie, K., Degarege, D., Chennupati, S., and Madhivanan, P. (2019). Improving socioeconomic status may reduce the burden of malaria in sub saharan africa: A systematic review and meta-analysis. *PloS one*, 14(1):e0211205.

Ermert, V., Fink, A. H., Jones, A. E., and Morse, A. P. (2011). Development of a new version of the liverpool malaria model. i. refining the parameter settings and mathematical formulation of basic processes based on a literature review. *Malaria journal*, 10(1):35.

Fournet, F., Cussac, M., Ouari, A., Meyer, P.-E., Toé, H. K., Gouagna, L.-C., and Dabiré, R. K. (2010). Diversity in anopheline larval habitats and adult composition during the dry and wet seasons in ouagadougou (burkina faso). *Malaria journal*, 9(1):78.

Froidurot, S. and Diedhiou, A. (2017). Characteristics of wet and dry spells in the west african monsoon system. *Atmospheric Science Letters*, 18(3):125–131.

Gething, P. W., Patil, A. P., Smith, D. L., Guerra, C. A., Elyazar, I. R. F., Johnston, G. L., Tatem, A. J., and Hay, S. I. (2011). A new world malaria map: Plasmodium falciparum endemicity in 2010. *Malaria journal*, 10(1):378.

Gleixner, S., Demissie, T., and Diro, G. T. (2020). Did era5 improve temperature and precipitation reanalysis over east africa? *Atmosphere*, 11(9):996.

Grömping, U. (2007). Relative importance for linear regression in r: the package relaimpo. *Journal of statistical software*, 17:1–27.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.

Ikeda, T., Behera, S. K., Morioka, Y., Minakawa, N., Hashizume, M., Tsuzuki, A., Maharaj, R., and Kruger, P. (2017). Seasonally lagged effects of climatic factors on malaria incidence in south africa. *Scientific reports*, 7(1):2458.

Kar, N. P., Kumar, A., Singh, O. P., Carlton, J. M., and Nanda, N. (2014). A a review of malaria transmission dynamics in forest ecosystems. *Parasites and Vectors*, 7(265).

Kilama, M., Smith, D. L., Hutchinson, R., Kigozi, R., Yeka, A., Lavoy, G., Kamya, M. R., Staedke, S. G., Donnelly, M. J., Drakeley, C., Greenhouse, B., Dorsey, G., and Lindsay, S. W. (2014). Estimating the annual entomological inoculation rate for plasmodium falciparum transmitted by anopheles gambiae s.l. using three sampling methods in three sites in uganda. *Malaria Journal*, 13(111).

Komen, K., Olwoch, J., Rautenbach, H., Botai, J., and Adebayo, A. (2015). Long-run relative importance of temperature as the main driver to malaria transmission in limpopo province, south africa: A simple econometric approach. *EcoHealth*, 12(1):131–143.

Lowe, R., Chirombo, J., and Tompkins, A. M. (2013). Relative importance of climatic, geographic and socio-economic determinants of malaria in malawi. *Malaria journal*, 12(1):416.

Lunde, T. M., Bayoh, M. N., and Lindtjørn, B. (2013a). How malaria models relate temperature to malaria transmission. *Parasites & vectors*, 6(1):20.

Lunde, T. M., Korecha, D., Loha, E., Sorteberg, A., and Lindtjørn, B. (2013b). A dynamic model of some malaria-transmitting anopheline mosquitoes of the afrotropical region. i. model description and sensitivity analysis. *Malaria Journal*, 12(1):1–29.

Mabaso, M. L., Vounatsou, P., Midzi, S., Da Silva, J., and Smith, T. (2006). Spatio-temporal analysis of the role of climate in inter-annual variation of malaria incidence in zimbabwe. *International Journal of Health Geographics*, 5(1):20.

Mabaso, M. L. H., Craig, M., Ross, A., and Smith, T. (2007). Environmental predictors of the seasonality of malaria transmission in africa: the challenge. *Am. J. Trop. Med. Hyg.*, pages 33–38.

Manzanas, R., Amekudzi, L., Preko, K., Herrera, S., and Gutiérrez, J. (2014). Precipitation variability and trends in ghana: An intercomparison of observational and reanalysis products. *Climatic change*, 124(4):805–819.

MARA (1998). Towards an atlas of malaria risk in africa. towards an atlas of malaria transmission in africa: First technical report of the mara/arma. Technical Report 30, South African Medical Research Council, IDRC, Canada and Wellcome Trust, UK.

Markham, C. G. (1970). Seasonality of precipitation in the united states. *Ann. Ass. Am. Geogr.*, 60:593–597.

Midekisa, A., Beyene, B., Mihretie, A., Bayabil, E., and Wimberly, M. C. (2015). Seasonal associations of climatic drivers and malaria in the highlands of ethiopia. *Parasites & vectors*, 8(1):1–11.

Mordecai, E. A., Paaijmans, K. P., Johnson, L. R., Balzer, C., Ben-Horin, T., de Moor, E., McNally, A., Pawar, S., Ryan, S. J., Smith, T. C., et al. (2013). Optimal temperature for malaria transmission is dramatically lower than previously predicted. *Ecology letters*, 16(1):22–30.

Murray, C. J., Rosenfeld, L. C., Lim, S. S., Andrews, K. G., Foreman, K. J., Haring, D., Fullman, N., Naghavi, M., Lozano, R., and Lopez, A. D. (2012). Global malaria mortality between 1980 and 2010: a systematic analysis. *Lancet*, 379:413–31.

Nicholson, S. E. (2013). The west african sahel: a review of recent studies on the rainfall regime and its interannual variability. *ISRN Meteorol*, 2013:1–32.

Oses, N., Azpiroz, I., Marchi, S., Guidotti, D., Quartulli, M., and G Olaizola, I. (2020). Analysis of copernicus' era5 climate reanalysis data as a replacement for weather station temperature measurements in machine learning models for olive phenology phase prediction. *Sensors*, 20(21):6381.

Paaijmans, K. P., Blanford, S., Bell, A. S., Blanford, J. I., Read, A. F., and Thomas, M. B. (2010). Influence of climate on malaria transmission depends on daily temperature variation. *Proc Natl Acad Sci USA*, 107(33):15135–15139.

Parham, P. E. and Michael, E. (2010). Modeling the effects of weather and climate change on malaria transmission. *Environmental Health Perspectives*, 118(5):620–626.

Reiner, R. C., Geary, M., Atkinson, P. M., Smith, D. L., and Gething, P. W. (2015). Seasonality of plasmodium falciparum transmission: a systematic review. *Malaria journal*, 14(1):343.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., and Ziese, M. (2018). Gpcc full data monthly product version 2018 at 0.25○: Monthly land-surface precipitation from rain-gauges built on gts-based and historical data. *GPCC: Offenbach, Germany.*

Shapiro, L. L., Whitehead, S. A., and Thomas, M. B. (2017). Quantifying the effects of temperature on mosquito and parasite traits that determine the transmission potential of human malaria. *PLoS biology*, 15(10):e2003489.

Shaukat, A. M., Bremen, J. G., and McKenzie, F. E. (2010). Using the entomological inoculation rate to assess the impact of vector control on malaria parasite transmission and elimination. *Malaria Journal*, 9(112).

Simple, O., Mindra, A., Obai, G., Ovuga, E., and Odongo-Aginya, E. I. (2018). Influence of climatic factors on malaria epidemic in gulu district, northern uganda: A 10-year retrospective study. *Malaria research and treatment*, 2018.

Sinka, M. E., Bangs, M. J., Manguin, S., Coetzee, M., Mbogo, C. M., Hemingway, J., Patil, A. P., Temperley, W. H., Gething, P. W., Kabaria, C. W., Okara, R. M., Van Boeckel, T., Godfray, H. C. J., Harbach, R. E., and Hay, S. I. (2010). The dominant anopheles vectors of human malaria in africa, europe and middle east: Occurrence data, distribution maps and bionomic precis. *Parasites and Vectors*, 3(117).

Takken, W. and Lindsay, S. W. (2003). *Ecological aspect for application of genetically modified mosquitoes*. Kluwer academic publishers, Netherlands.

Tarek, M., Brissette, F. P., and Arsenault, R. (2020). Evaluation of the era5 reanalysis as a potential reference dataset for hydrological modelling over north america. *Hydrology and Earth System Sciences*, 24(5):2527–2544.

Thompson, M. C., Mason, S. J., Phindela, T., and Connor, S. J. (2005). Use of rainfall and sea surface temperature monitoring for malaria early warning in bostwana. *Am J Trop Med Hyg*, 73(1):214–221.

Tompkins, A. M. and Di Giuseppe, F. (2015). Potential predictability of malaria in africa using ecmwf monthly and seasonal climate forecast. *Journal of applied meteorology and climatology*, 54:521–540.

Tompkins, A. M. and Ermert, V. (2013). A regional-scale, high resolution dynamical malaria model that accounts for population density, climate and surface hydrology. *Malaria Journal*, 12(65).

Tusting, L. S., Bousema, T., Smith, D. L., and Drakeley, C. (2014). Measuring changes in plasmodium falciparum transmission: precission, accuracy and cost of metrics. *Advance Parasitol*, 84:151–208.

WHO (2020). World malaria report 2020: 20 years of global progress and challenges. In *World malaria report 2020: 20 years of global progress and challenges.*

Yadav, K., Dhiman, S., Rabha, B., Saikia, P., and Veer, V. (2014). Socio-economic determinants for malaria transmission risk in an endemic primary health centre in assam, india. *Infectious diseases of poverty*, 3(1):19.

Yamba, E. I. (2016). *Improvement and validation of dynamical malaria models in Africa.* PhD thesis, University of Cologne.

Yamba, E. I., Tompkins, A. M., Fink, A. H., Ermert, V., Amelie, M. D., Amekudzi, L. K., and Briët, O. J. (2020). Monthly entomological inoculation rate data for studying the seasoanality of malaria transmission in africa. *Data*, 5(2):31.

Yamba, E. I., Tompkins, A. M., Fink, A. H., Ermert, V., Djouda, A., Amekudzi, L. K., and Briët, O. J. T. (2018). Monthly entomological inoculation rates for studying malaria transmission seasonality in Africa.