

Article

Using the Hurst Exponent and Entropy Measures to Predict Effective Transmissibility in Empirical Series of Malaria Incidence

João Sequeira ^{1,2} , Jorge Louçã ¹, António M. Mendes ³  and Pedro G. Lind ^{4,5,6,*}

¹ ISTAR-IUL, Instituto Universitário de Lisboa (ISCTE-IUL), Av. das Forças Armadas, 1649-026 Lisboa, Portugal; jsequeira@chlo.min-saude.pt (J.S.); jorge.l@iscte-iul.pt (J.L.)

² Hospital Santa Cruz, Av. Prof. Dr. Reinaldo dos Santos, 2790-134 Carnaxide, Portugal

³ Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa, Av. Prof. Egas Moniz, 1649-028 Lisboa, Portugal; antoniomendes@medicina.ulisboa.pt

⁴ Department of Computer Science, OsloMet—Oslo Metropolitan University, N-0130 Oslo, Norway

⁵ Artificial Intelligence Lab, Oslo Metropolitan University, N-0166 Oslo, Norway

⁶ NordSTAR—Nordic Center for Sustainable and Trustworthy AI Research, Pilestredet 52, N-0166 Oslo, Norway

* Correspondence: pedro.lind@oslomet.no

Abstract: We analyze the empirical series of malaria incidence, using the concepts of autocorrelation, Hurst exponent and Shannon entropy with the aim of uncovering hidden variables in those series. From the simulations of an agent model for malaria spreading, we first derive models of the malaria incidence, the Hurst exponent and the entropy as functions of gametocytemia, measuring the infectious power of a mosquito to a human host. Second, upon estimating the values of three observables—incidence, Hurst exponent and entropy—from the data set of different malaria empirical series we predict a value of the gametocytemia for each observable. Finally, we show that the independent predictions show considerable consistency with only a few exceptions which are discussed in further detail.

Keywords: malaria; Hurst exponent; Shannon entropy; long range dependence; autocorrelation function; stochastic long memory; gametocytemia



Citation: Sequeira, J.; Louçã, J.; Mendes, A.M.; Lind, P.G. Using the Hurst Exponent and Entropy Measures to Predict Effective Transmissibility in Empirical Series of Malaria Incidence. *Appl. Sci.* **2022**, *12*, 496. <https://doi.org/10.3390/app12010496>

Academic Editors: Jianbo Gao and José Machado

Received: 17 October 2021

Accepted: 29 December 2021

Published: 5 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In Babylon, during the year 323 BC, Alexander the Great fell ill after returning from Persia, and died shortly after at the age of 32. The reason of his death was most probably a disease now known to be caused by an infection with the parasite *Plasmodium falciparum* [1]. Malaria cases have declined globally from 238 million cases, in 2000, to 229 million, in 2019 [2]. Its mortality also decreased worldwide from around 736 thousand in 2000 to 409 thousand deaths in 2019 [2]. However, this reduction has stabilized since 2016, and several challenges and risks remain, namely, as concerns children under 5 years old. Due to their limited immune protection, they are a highly vulnerable group, accounting for 67% of global fatal malaria cases in 2019 [2].

Of particular importance in malaria transmission is the amount of gametocytes in the blood circulation. Gametocytes are the parasite forms which mediate transmission from humans to mosquitoes and, therefore, are also obvious targets in implementing preventive actions such as vaccine immunization or anti-malarial drugs. Submicroscopic gametocytemia detection is difficult and often not as precise as statistical data collection from e.g., malaria incidence varying heterogeneously from one community to another [3,4].

In this paper we propose an approach to estimate effective *Plasmodium* transmissibility levels, which reflect submicroscopic gametocytemia in individuals of a specific community, based on collected time series of malaria incidence in that same community.

Time series models have been used as important tools and metrics, not only in epidemiology, but also in economics, geophysics, biology and ecology [5]. One of them, the so-called Hurst exponent, was introduced by Hurst in 1951 [6] in the context of hydrology planning, specifically, during the study of flooding levels in the river Nile. The Hurst exponent is a measure of memory in a series of values. Applying it to the series of Nile flooding levels, based upon Egyptian ancient hydrology data collected over 847 years, Hurst was able to estimate the rate at which the autocorrelation of that series decreased as the time interval between measurements increased. In this way, Hurst used the collected data to model the flooding levels of the Nile, and, with that, assessed optimum dam sizes to contain extreme rain events. Soon the applicability of the Hurst exponent extended to many other fields, particularly in the study of financial theory [7] and complex phenomena with evidence of fractal features [8–10]. However, while different time series models have been implemented in epidemic infections in general [11–18], and in malaria incidence forecasting in particular [19,20]—mostly auto-regressive and linear models—the application of the Hurst exponent in malarial spreading is still lacking.

To relate gametocytemia levels with malaria incidence time series we introduce a methodology based upon the use of Hurst exponents. To the best of our knowledge, this is the first analysis of empirical data of malaria incidence series using Hurst exponent to derive an estimate of a hidden variable, namely, the level of transmissibility or easiness of spreading within the community from which the series of malaria incidence is collected.

We show that the Hurst exponent is able to grasp long-range dependencies close to the phase transition between disease elimination and stable prevalence scenarios of malaria. In a more general scope, complexity in time series may also be related to the level of information entropy, which is commonly used to address emergence phenomena and self-organization. In this paper, we analyze the effects of memory in time series of malaria incidence. Two types of memory have been defined in random stochastic processes, long and short memory, where the transition between these regimes may be represented as a phase transition in the context of a stochastic random process. We focus on Hurst exponent estimation and Shannon entropy at different levels of disease transmission intensity, applied to malaria time series derived from simulations with a previously introduced agent-based model [3] with different (parameterized) levels of gametocytemia, as well as to different empirical malaria time series.

Our objective is to properly identify malaria transmission patterns, as well as to link long-range dependence processes in malaria incidence time series to the occurrence of phase transitions in close proximity to disease elimination. We also test the importance of Hurst exponent estimation and Shannon entropy as sound predictors of the presence of long-range dependence and long memory processes in malaria transmission. In particular, we show that from simple models connecting gametocytemia levels and measures easily extractable from empirical series of malaria incidence, such as incidence levels, a Hurst exponent and entropy, we are able to predict an indicator of “effective” gametocytemia for regions where malaria incidence is regularly monitored. Our study uses eight different empirical series, shown in Figure 1.

We start in Section 2 by describing the different empirical data sets analyzed in this study and the agent-model used to produce simulated scenarios of different transmissibility levels, parameterized by the gametocytemia level. Moreover, we briefly describe the basic tools—the Hurst exponent and Shannon entropy—explaining how they are computed from series of malaria incidence. In Sections 3 and 4 we describe, qualitatively, the behavior of Hurst exponent and entropy, and the form of the autocorrelation function decay, in eight different malaria empirical time series examples. In Section 5 we derive models to fit the values of the three observables: the malaria incidence, the Hurst exponent and the entropy for different simulated scenarios, as a function of the gametocytemia level. Using this fitted expressions we then measure the average malaria incidence, Hurst exponent and entropy of the empirical data sets. Introducing these values in the fitted expressions we retrieve an estimate of the associated “effective” gametocytemia level, i.e., the effective transmissibility

level in each empirical case. Finally, in Section 6 we discuss the limitations of our approach, as well as possible extensions to it.

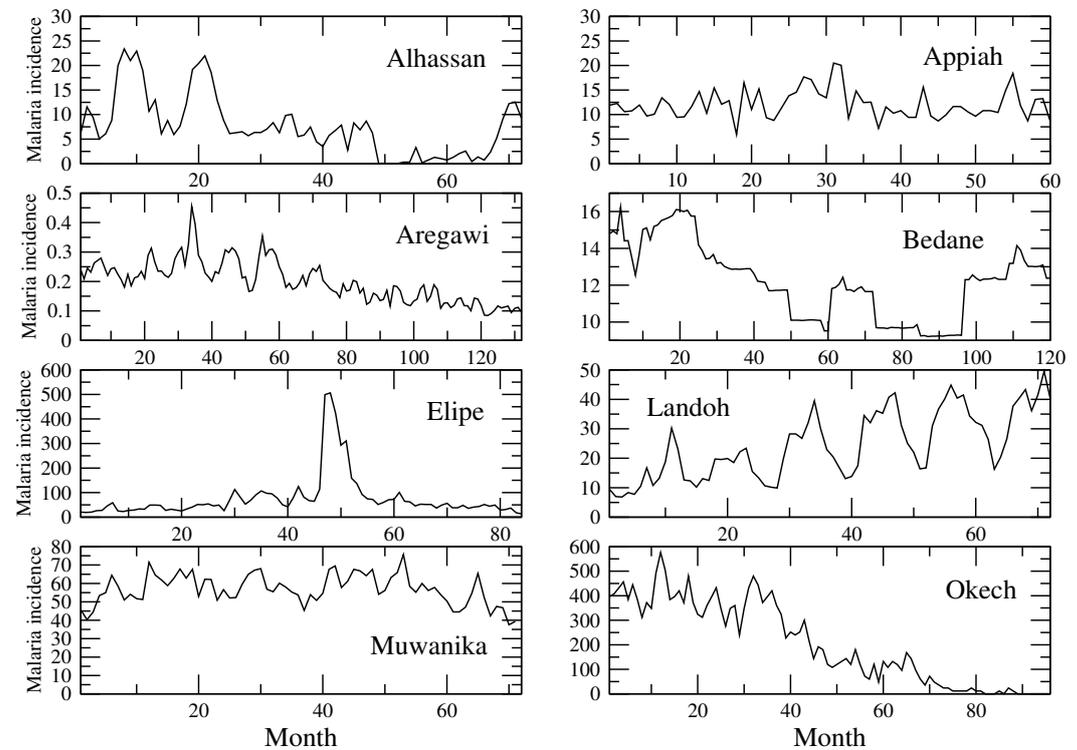


Figure 1. Eight different series of malaria incidence from different studies in the literature: Alhassan et al. (Kasena Nankana municipality in Ghana, 2017) [21], Appiah et al. (Ejisu-Juaben municipality in Ghana, 2015) [16], Aregawi et al. (Ethiopia, 2014) [22], Bedane et al. (Kucha district in Ethiopia, 2016) [23], Gomez-Elipe et al. (Karuzi in Burundi, 2007) [24], Landoh et al. (Est Mono district in Togo, 2012) [25], Muwanika et al. (Uganda, 2017) [26], Okech et al. (Kenya, 2008) [27]. All data series are available from the respective original references from the authors.

2. Data, Modelling Methods and Analysis Tools

2.1. Empirical Series of Malaria Incidence

The empirical part of our investigation comprehends eight series of malaria incidence as presented in Figure 1. These series are available from previous studies namely Alhassan et al. (Kasena Nankana municipality in Ghana, 2017) [21], Appiah et al. (Ejisu-Juaben municipality in Ghana, 2017) [16], Aregawi et al. (Ethiopia, 2014) [22], Bedane et al. (Kucha district in Ethiopia, 2016) [23], Gomez-Elipe et al. (Karuzi in Burundi, 2007) [24], Landoh et al. (Est Mono district in Togo, 2012) [25], Muwanika et al. (Uganda, 2017) [26], Okech et al. (Kenya, 2008) [27]. The data sets were chosen as representative of regions with malaria incidence diversity in Africa. They have a time duration ranging from 60 [16] to 132 months [22], while showing different trends and periodicity—see Figure 1 and Table 1. Malaria incidence was measured at each month as the expected number of new malaria cases per 100 inhabitants in a full year, if malaria incidence were to be kept constant. In this way, to obtain the precise incidence at each month, the yearly presented graphical value must be divided by 12 months.

Table 1. The empirical values of average malaria incidence I (malaria cases per 100 inhabitants, per year), the Hurst exponent H and Shannon entropy S , concerning a 36-month time frame. Error indicate standard deviations from the series by using moving windows of 36 months.

Empirical Series	Year	I 36 Months (I_{36})	H 36 Months (H_{36})	S 36 Months (S_{36})	Duration Months
Aregawi [22]	2014	0.208 ± 0.0495	0.7766 ± 0.1428	5.138 ± 0.0064	132
Alhassan [21]	2017	7.091 ± 2.808	1.014 ± 0.1425	4.792 ± 0.1717	72
Bedane [23]	2016	11.63 ± 1.355	1.077 ± 0.0925	5.161 ± 0.0036	120
Appiah [16]	2015	12.44 ± 0.1318	0.9229 ± 0.0570	5.121 ± 0.0026	60
Landoh [25]	2012	24.54 ± 4.312	0.7326 ± 0.0850	5.065 ± 0.0206	72
Muwanika [26]	2017	58.76 ± 1.112	0.8254 ± 0.1236	5.160 ± 0.0027	71
Elipe [24]	2007	101.0 ± 28.20	0.9367 ± 0.0993	4.710 ± 0.1733	84
Okech [27]	2008	203.4 ± 120.2	1.018 ± 0.1568	4.881 ± 0.2774	96

As can be seen from Figure 1, the eight cases differ in their level of malaria transmission and epidemic behavior. All empirical series show some form of irregular periodicity as a consequence of climate seasonality. Declining malaria incidence is clear in series from Aregawi and Okech, as well as in Alhassan and Bedane, although with a final disease outbreak in these last two cases. High levels of malaria transmission occur in Gomez-Elipe (identified solely as Elipe time series in some of the figures, for clarity) for a brief period of time. Malaria incidence remains quite stable in the Muwanika, Appiah and Landoh empirical series, with a consistent upward trend in the last case.

2.2. Agent Model for Malarial Spreading

To combine results from empirical data with simulations from the agent model introduced and developed in previous works [3,4], we conduct a series of simulations at different levels of disease transmission efficiency. Human-to-mosquito (H-to-M) transmission was defined in terms of the fraction of human disease days with the presence of gametocytemia in blood circulation, henceforth represented as w_h .

Six different scenarios were considered, corresponding to a wide range of different levels of positive gametocytemia duration and disease transmission efficiency, specifically, 110 days of positive gametocytemia during 150 days of expected disease duration (i.e., $w_h = 110/150 = 0.733$), 90 days ($w_h = 0.600$), 75 days ($w_h = 0.500$), 70 days ($w_h = 0.467$), 68 days ($w_h = 0.453$), and 63 days ($w_h = 0.420$). For the simulations, we consider a system of $N_m = 4000$ mosquitoes and $N_h = 2000$ human individuals, both including healthy and infected individuals. We have modeled the number of mosquitoes as a small but effective fraction of the overall mosquito mass that randomly feeds on a human individual, twice daily on average. The simulation time lasts 30 years while evaluating each human individual in terms of disease duration and human-to-mosquito transmission.

The algorithm keeps track of several attributes for each agent, whether human or mosquito, at a particular age, such as the time spent since the first day of infection, and the individual immunity status. Beyond 5 years of persistent human reinfections, the human host will develop partial protective immunity at the maximum possible level, while losing it after 2 years without infection. The computational cycle includes a realistic mosquito daily mortality routine. Dead mosquitoes are to be replaced by uninfected mosquitoes. Human disease duration reflects realistic human recovery from malaria.

Human-to-mosquito transmission efficiency (w_h) is also stochastically defined and directly dependent on the number of days with positive gametocytemia. Upon updating the number of healthy humans individuals and mosquitoes, the algorithm generates one episode of mosquito feeding in a human individual, with the possibility of protection from long lasting insecticide-impregnated nets (LLIN), insecticide-impregnated nets (ITN) or

indoor residual spraying (IRS). Our model is inspired by Mozambique’s seasonality [18,28], considering 150 days for the duration of its high-transmission season—see Figure 2.

Relevant details as well as the flowchart describing the computer implementation of the agent-based model are given in Ref. [3]. That model was also used as support for an additional study concerning the presence of heterogeneity in malaria transmission along with the use of ivermectin [4].

Using the present model we analyzed the behavior of the human–mosquito coupled system, resulting from a complex interaction between the two compartments. Human-to-mosquito transmission efficiency (w_h) was used to define the probability of a sustained presence of gametocytemia in human blood circulation, as well as the survival probability of infected mosquitoes beyond latency. These aspects are considered critical in disease transmission. Our model simulations use gametocytemia as an independent variable affecting human-to-mosquito transmission. Different levels of gametocytemia define different stages of disease transmission efficiency. Theoretical gametocytemia reduction is considered equivalent to an effective treatment with gametocidal agents, such as primaquine or methylene blue, in a fraction of the human population.

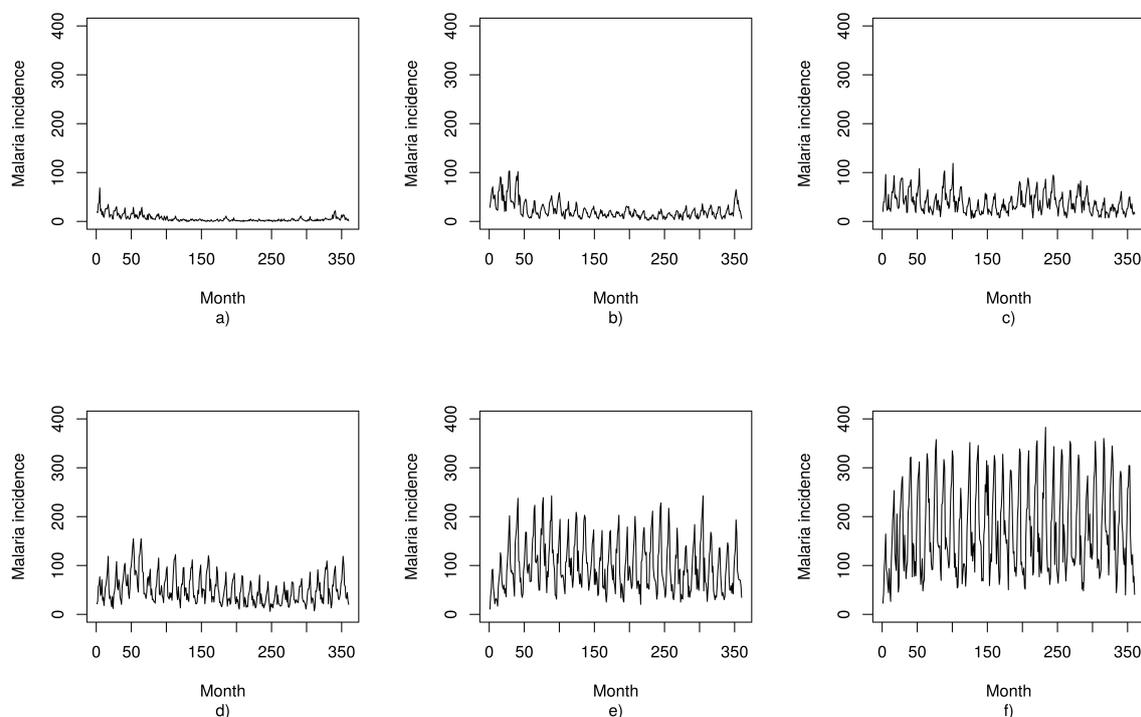


Figure 2. Six different simulations of malaria incidence using different scenarios of gametocytemia, specifically, (a) 63 days ($w_h = 0.420$), (b) 68 days ($w_h = 0.453$), (c) 70 days ($w_h = 0.467$), (d) 75 days ($w_h = 0.500$), (e) 90 days ($w_h = 0.600$), and (f) 110 days of positive gametocytemia over 150 days of expected disease duration (i.e., $w_h = 110/150 = 0.733$).

2.3. Hurst Exponent and Entropy to Assess Memory Effects in Stochastic Series

We use two different metrics to assess memory in series of malaria incidence, empirical and simulated, investigating whether long-range dependencies could occur close to phase transition near disease elimination when compared to more stable epidemic scenarios. The Hurst exponent is defined as

$$\frac{R}{D} = kT^H \quad (1)$$

where the first member represents the rescaled range as a dimensionless ratio between R (represents the maximal range of all observations) and D (represents the standard deviation of all observations). An explicit mathematical definition is given below in Equation (4). T

stands for the time index (number of observations in the time series), k is some constant to be determined and H represents the Hurst exponent.

With the Hurst exponent estimation it became possible to distinguish among three different regimes characterizing the time series: (i) the anti-persistent regime, characterized by $0.0 < H < 0.5$, where, if the series increases (resp. decreases) in one period, it is very likely that it will decrease (resp. increase) in the next period; (ii) the persistent regime, characterized by $0.5 < H < 1$, where, if the series decreases (resp. increases) in one period it is very likely that it will decrease (resp. increase) in the next period; and (iii) the memoryless regime, characterized by $H = 0.5$ when the process is uncorrelated in time.

A Hurst exponent estimation close to 0.5 (random walk process) is found in empirical time series with heavier disease transmission. In malaria time series from our model simulations, transitions from prevalence to disease elimination are characterized by values of the Hurst exponent larger than 0.5 and close to 1, a footprint of a persistence time series.

As for entropy, it is related with the complexity of the time series [29,30]. The complexity of stochastic processes may be calculated with the use of entropy-based measures. For that purpose, several functions may be employed. The significance of complexity, emergence phenomena and self-organization may provide us with useful information concerning continuous as well as discrete systems, in the form of time series results.

Information entropy is supported by the equation:

$$S = - \sum_{i=1}^{N_b} P(x_i) \log P(x_i), \tag{2}$$

where $P(x)$ is the probability density function of the observable x , which, in its discretized form, is estimated by a histogram with a finite set of values $P(x_i)$ for the set of bin-points x_i . The log function is base two. Information theory defines entropy in terms of information uncertainty in the evolution of time series results. Other entropy-related terms, such as mutual information, may be used as alternative methods for time series analysis in malaria.

2.4. Estimating the Hurst Exponent and Entropy in Series of Malaria Incidence

Both the Hurst exponent and Shannon entropy are influenced by the length of the time series sample. Therefore, we introduced a standardization procedure, which is independent of the length of the time series.

The procedure is illustrated in Figure 3 and is as follows. We define a 36-month moving average malaria incidence, symbolized as $I_{36}(m)$, at month m , as the average malaria incidence in the previous 36 months:

$$I_{36}(m) = \frac{1}{36} \sum_{n=m-35}^m I(n), \tag{3}$$

where $I(n)$ is the monthly malaria incidence measured or simulated, composing the series of values.

In the Alhassan malaria time series (Kasena Nankana municipality in Ghana, 2017) [21] one can see a declining trend in malaria incidence in the form of a decreasing moving average of 36 months (I_{36}), from ~ 11 to ~ 4 cases per 100 inhabitants, per year [*phy*], despite the small final outbreak in the last 6 months. Consistent this declining trend in I_{36} , we may also witness a decreasing trend in S_{36} , as well as a rising trend in H_{36} after month 55. In this case, the behavior of both the Hurst exponent and entropy in the preceding 36 months is well correlated with the malaria incidence trend.

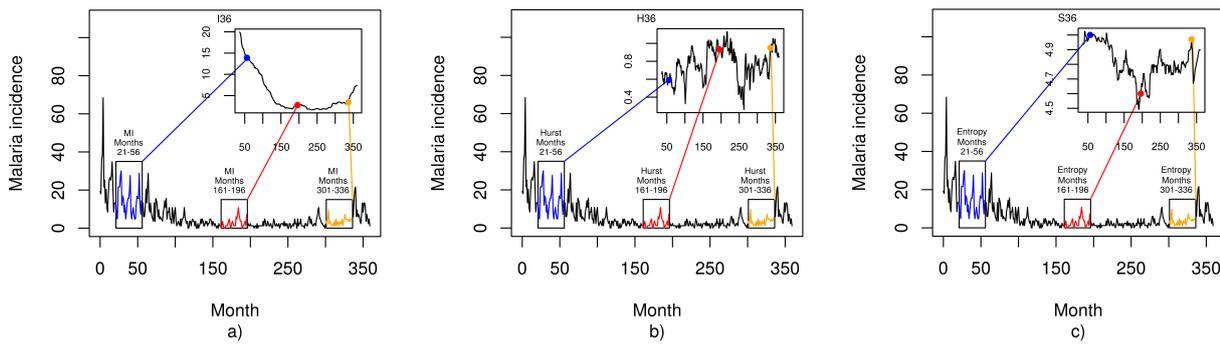


Figure 3. (a) Illustration of how a 36-month moving average of malaria incidence, symbolized as I_{36} , is obtained from an original series of monthly malaria incidence values, either empirical or simulated. Similar procedures are used to compute (b) H_{36} and (c) S_{36} . See text for details. In this case, we used an agent-based simulation with a low human-to-mosquito transmission efficiency of $w_h = 0.420$.

The Gomez-Elipe time series (Karuzi in Burundi, 2007) [24] is quite different from the remaining empirical examples. It shows a consistent stable pattern of malaria incidence below 110 cases *phy*, until month 45, peaking at ~ 500 cases *phy*, around month 48, with a fast downward trend to ~ 60 cases *phy* at month 56, and a slower decline thereafter to a final value of ~ 20 cases *phy*—see Figure 1. The 36-months malaria incidence (I_{36}) reveals a consistent upward trend peaking at month 64, with a downward pattern thereafter. In this case, we find a persistent oscillatory behavior in the value of H_{36} during the entire time series, usually above 0.9. Yet, in parallel with the outbreak in malaria incidence, we see a sudden fall in H_{36} to values close to 0.6 (closer to random noise) at around month 48. Entropy in the form of a 36-months moving average (S_{36}) reveals a similar behavior in relation to I_{36} from month 48 onwards.

The Landoh malaria time series (Est Mono district in Togo, 2012) [25] reveals a consistent upward trend in malaria incidence in the form of a mild increase in the 36-months moving average (I_{36}) during the whole time series, ranging from the initial ~ 18 cases *phy*, to a peak at ~ 30 cases *phy* in month 72.

Consistent with the steady upward trend in I_{36} during the entire time series, we find a declining trend in the 36-months moving average of the Hurst exponent (H_{36}), from a peak of ~ 0.85 at month 47, to an all-time-low value of ~ 0.55 . Along with this downward trend in H_{36} there is a consistent wave-like increase in entropy in the form of a 36-months moving average (S_{36}), during the entire time series. In the Landoh time series, the behavior of H_{36} and S_{36} is reasonably well correlated with the behavior of I_{36} in time.

Of a similar form, but with an opposite trend to the Landoh time series, the Okech time series (Kenya, 2008) [27] also reveals a steady decreasing trend. In the Okech time series, malaria incidence consistently decreases from an initial peak of ~ 400 cases *phy*, to a final value of close to 50 cases *phy*. Consistent with the steady downward trend in I_{36} during the entire time series, we find an upward trend in the 36-months moving average of the Hurst exponent (H_{36}) from an all-time-low of ~ 0.6 at month 42, to a peak of ~ 1.2 at month 82. Along with this upward trend in H_{36} , there is also a consistent decrease in entropy in the form of a 36-months moving average (S_{36}), during the whole time series. In the Okech time series, the behavior of H_{36} and S_{36} is also quite correlated with the behavior of I_{36} in time.

As the Hurst exponent (H) and Shannon entropy (S) are influenced by the length of the time-series sample, we also considered a 36-month moving window estimate. In the case of the Hurst exponent, we compute the quantities R and D as

$$R(T) = \max_{t=1, \dots, T} \left(\sum_{k=1}^t (I(k) - I_{36}(T)) \right) - \min_{t=1, \dots, T} \left(\sum_{k=1}^t (I(k) - I_{36}(T)) \right), \quad (4a)$$

$$D(T) = \left(\frac{1}{36} \sum_{k=T-35}^n (I(k) - I_{36}(T))^2 \right)^{1/2}. \quad (4b)$$

Having the series of 36 values $R(T)/D(T)$ for each window, we then applied Equation (1) to find fitting values for k and for the Hurst exponent H .

As for Shannon entropy, an estimate of the probability density was first computed given by the set of values $P(I_i)$ for an assumed set of bin values I_i , with $i = 1, \dots, N_b$, of malaria incidence (I) and then Equation (2) was applied, running the sum only over the 36 incidence values observed within each time window.

In this way, we have computed the average of malaria incidence I , H and S for 36-month windows, independently of the size of the empirical series. In order to quantify I , H and S for each empirical case, we then obtained the average of the corresponding time series and computed its standard deviations. Table 1 shows the empirical values of the three observables. Notice that, in these cases, higher moments seems to be not particularly relevant, which can be seen comparing the average in Table 1 with the median and quartiles in Table 2.

Table 2. Table with the empirical values of median, quartile 25 and 75 from the series of malaria incidence, entropy and the Hurst exponent, using moving windows of 36 months.

Empirical Series	(I_{36})	(I_{36})	(I_{36})	(H_{36})	(H_{36})	(H_{36})	(S_{36})	(S_{36})	(S_{36})
	Median	Q ₂₅	Q ₇₅	Median	Q ₂₅	Q ₇₅	Median	Q ₂₅	Q ₇₅
Aregawi, 2014	0.22	0.16	0.26	0.80	0.66	0.90	5.14	5.13	5.14
Alhassan, 2017	7.43	4.36	9.68	0.94	0.90	1.15	4.77	4.68	4.96
Bedane, 2016	11.2	10.4	12.4	1.10	1.00	1.15	5.16	5.16	5.16
Appiah, 2015	12.5	12.4	12.5	0.93	0.87	0.96	5.12	5.12	5.12
Landoh, 2012	24.0	21.0	28.4	0.76	0.70	0.80	5.06	5.05	5.08
Muwanika, 2017	59.0	58.2	59.7	0.81	0.77	0.90	5.16	5.16	5.16
Elipe, 2007	115	83.5	122	0.95	0.89	1.00	4.67	4.59	4.72
Okech, 2008	211	89.1	310	1.06	0.98	1.13	4.96	4.85	5.08

3. Qualitative Analysis and Robustness Assessment of the Hurst Exponent and Entropy in Empirical Time Series Behavior

While the eight empirical cases show a broad range of values for malaria incidence and the Hurst exponent, the entropy seems much more resistant to changes. The cases of Alhassan, Elipe and Okech form a separated group from the other series which form a second group. We will address these cases in more detail when building the model for effective human-to-mosquito transmissibility (gametocytemia).

We have also looked at the way those indices behaved in a typical low-transmission empirical time series, such as that obtained from Okech, 2008 [27]. In this empirical time series, when the Hurst exponent was evaluated in 36-month partial intervals (H_{36}), it revealed a clear inverse correlation with malaria annual incidence. Furthermore, the Shannon entropy consistently decreased with progressive lower values of malaria incidence. This correlation pattern was similar to the one found in model simulations when comparing high- and low-transmission scenarios.

By consistently searching for evidence of the presence of long-range dependence in malaria time series we looked into the time evolution of information (Shannon) entropy in scenarios with stronger disease transmission, and compared the results with those from other time series of lower disease transmission. In our model simulations, at low levels of disease burden near an eradication-prevalence transition, we consistently found malaria time series with lower information entropy. In Appendix A we describe, in more detail, the eight particular cases.

Table 3 shows the correlation between the three properties, I , H and S , for the eight empirical data sets. One finds evidence of linear correlation between malaria incidence and

the Hurst exponent or Shannon entropy in seven of the eight presented empirical malaria time series. The case of Bedane is an exception.

Table 3. Pearson correlation coefficient r^2 , between pairs among the three metrics malaria incidence I , the Hurst exponent H and Shannon entropy S . The three metrics were computed for three different time-windows of 24, 36 and 48 months. See also Appendix B. Results with $r^2 > 0.600$ are highlighted in bold.

Average	24 Months		36 Months		48 Months	
Pearson coeff.	$r^2(I, H)$	$r^2(I, S)$	$r^2(I, H)$	$r^2(I, S)$	$r^2(I, H)$	$r^2(I, S)$
Aregawi, 2014 [22]	0.048	0.216	0.323	0.223	0.727	0.012
Alhassan, 2017 [21]	0.528	0.657	0.638	0.879	0.530	0.746
Bedane, 2016 [23]	0.006	0.107	0.094	0.027	0.000	0.181
Appiah, 2015 [16]	0.336	0.620	0.222	0.353	0.299	0.534
Landoh, 2012 [25]	0.310	0.339	0.436	0.633	0.722	0.743
Muwanika, 2017 [26]	0.043	0.547	0.175	0.748	0.449	0.903
Elipe, 2007 [24]	0.074	0.886	0.012	0.655	0.590	0.683
Okech, 2008 [27]	0.068	0.603	0.414	0.069	0.013	0.949

The estimation of the Hurst exponent by R/S -analysis may be biased due to the shorter length of 36 months [31–33]. To ascertain how robust the results with 36-month windows are, we repeated our estimates for 24- and 48-month averages. The correlation is also reasonably evident in the 48-months time frame. Entropy is clearly more linearly correlated to malaria incidence than the Hurst exponent.

In summary, the data shown in Table 3 suggests the presence of significant linear correlation between malaria incidence, and the Hurst exponent or Shannon entropy. In Appendix B we present a more detailed comparison between 36-month averages with 24- and 48-month averages.

4. Autocorrelation Function and Stochastic Memory in Malaria Empirical Series

The autocorrelation function (ρ_k) behavior has been used with reasonable success in different research fields, from finance to hydrology, and climate data time series. It measures the linear relationship between two sequential values of a time series with a specific time lag k . The autocorrelation function (ρ_k) expresses the magnitude of that correlation between k lagged values:

$$\rho_k = \frac{\sum_{t=1}^{N-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^N (x_t - \bar{x})^2} \tag{5}$$

Figure 4 shows the autocorrelation function for each empirical case.

The algebraic decay of the empirical autocorrelation function ρ_k is strongly connected to the memory of stochastic processes, such as long memory in the form of *Long-Range Dependence* (LRD). The existence of LRD assumes the presence of stationarity in the time series. A memory parameter d is defined in relation to the slope of the autocorrelation function (ρ) decay. When $d > 0$ the term *persistent* defines the time series, with progressively larger values in time. In the opposite case, we have $d < 0$ with the presence of *anti-persistence*, where positive values will tend to alternate with negative values and vice versa. Here, we borrow the concept as defined in Ref. [34]: *For $d < 0$ we have anti-persistence; i.e., positive values tend to be followed by negative values and vice versa.* In the special case of $d = 0$, the process will correspond to the presence of white noise, without evidence of autocorrelation and corresponding to a pure Markovian process [34].

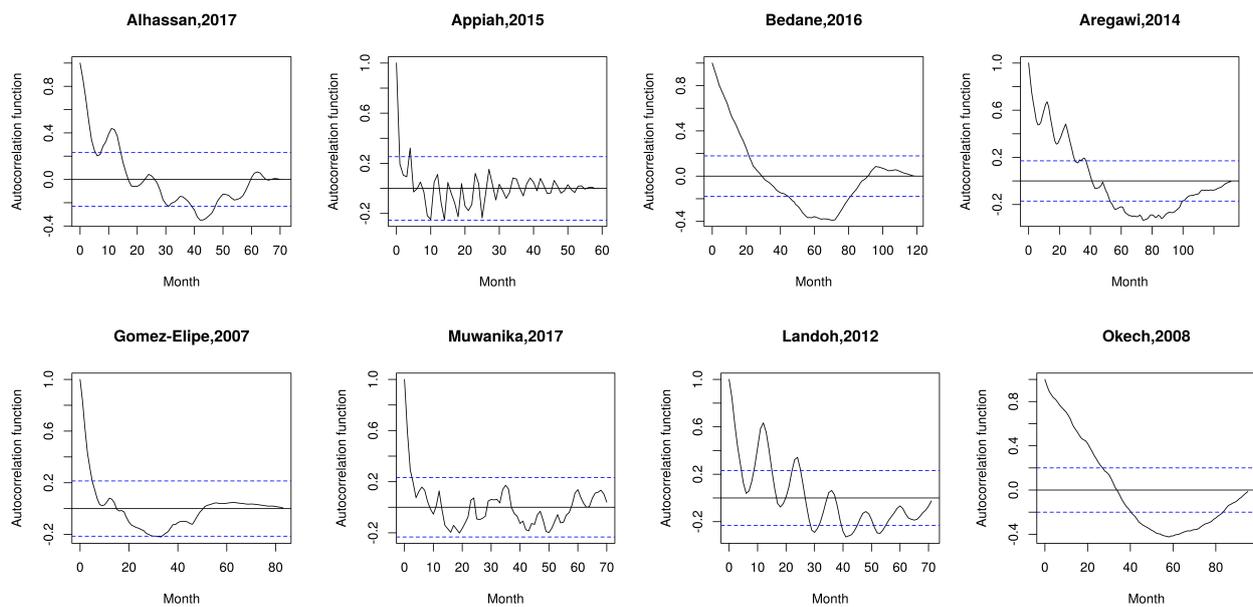


Figure 4. Malaria empirical series: Autocorrelation function from Alhassan et al. (Kasena Nankana municipality in Ghana, 2017) [21]; Appiah et al. (Ejisu-Juaben municipality in Ghana, 2015) [16]; Bedane et al. (Kucha district in Ethiopia, 2016) [23]; Aregawi et al. (Ethiopia, 2014) [22]; Gomez-Elipe et al. (Karuzi in Burundi, 2007) [24]; Muwanika et al. (Uganda, 2017) [26]; Landoh et al. (Est Mono district in Togo, 2012) [25]; Okech et al. (Kenya, 2008) [27].

The decay in time of the autocorrelation function ρ_k in a stochastic process correlates with the presence of memory persistence of past events in the present state of the system. In the case of fast ρ_k the exponential decay the system memory will be short. With slower ρ_k decays (corresponding to power law processes) memory will be longer in relation to the presence of LRD.

For the present eight empirical time series the autocorrelation function ρ_k revealed similar decay patterns. The ρ_k decay seems to have deviated from exponential decay in most of the cases, which would have been expected in the case of a pure-white-noise Markovian process, thus suggesting the presence of memory persistence in all the empirical series shown. The slow ρ_k decay is usually related to the presence of time series' non-stationarity. In more than half of the presented examples, a persistent and ondulatory expression of ρ_k values as a result of seasonality and periodicity in disease transmission overlaps with the background decaying trend.

5. Towards a More Quantitative Malaria Model for Predicting Effective Gametocytemia

Figure 5 shows the result obtained for I , H and S in the six different scenarios of gametocytemia levels. From Figure 5a we observe that the six simulations cover all different incidence regimes, ranging from low incidence ($I \sim 0$) to high incidence ($I \sim 1$). For the same simulations, the Hurst exponent shown in Figure 5b, shows a clear decrease of memory patterns with the increase of the gametocytemia level w_h .

As for the dependence of the entropy S , shown in Figure 5c, we also observe a transition to large entropy values as w_h increases, though the transition seems much more abrupt. At the phase transition near disease extinction we have consistently found lower values of information entropy, clearly defining a stochastic process with long memory. At higher disease transmission rates (higher w_h) entropy becomes higher and more stable, evolving towards a short memory stochastic process. This dichotomy defines the nature of transmission stability and may be useful in defining how distant a malaria time series is from a situation of disease extinction.

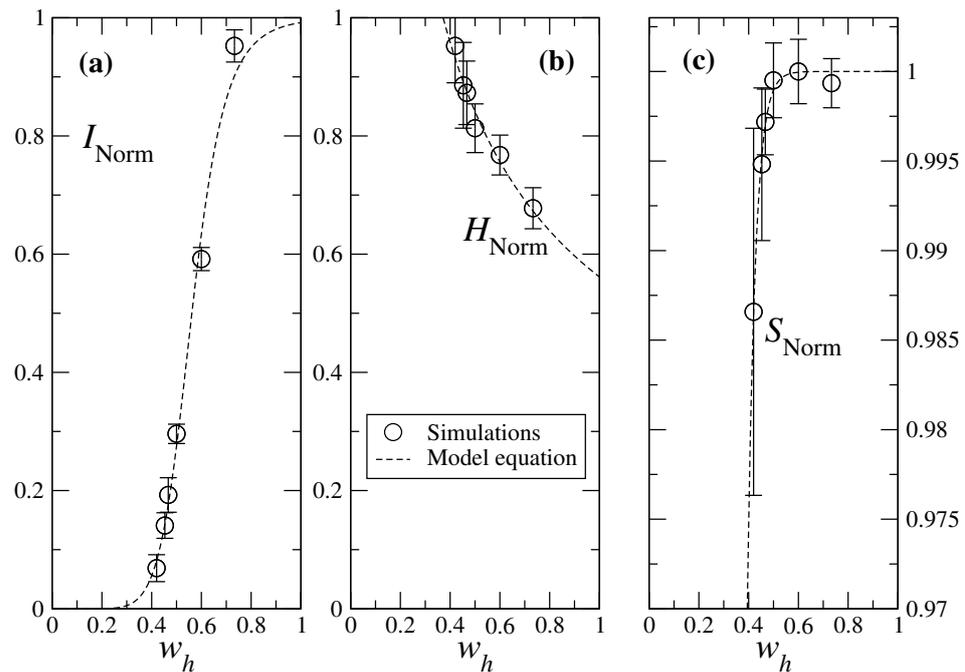


Figure 5. (a) Evaluation of the Malaria incidence in time series from model simulations at different settings of human-to-mosquito transmission efficiency ($0.420 < w_h < 0.733$). The Hurst exponent is consistently close to 1.0 at the phase transition (when $w_h \sim 0.420$), and decreasing to values of ~ 0.5 (close to a random walk stochastic state) with higher values of $w_h \sim 0.700$ in higher disease transmission epidemic stage. For the same series (b) evaluation of the Hurst exponent and (c) evaluation of Shannon entropy. In each case we simulated the agent model 10 times.

5.1. Models for the Three Observables as Function of Parameter Gametocytemia

Having described the values of the incidence, the Hurst exponent and entropy obtained in six simulations with an agent model for malaria spreading, we now derive models for each one of these observables as a function of the central parameter in our approach, the gametocytemia level w_h .

Notice that, being a parameter that cannot be measured directly from empirical series of malaria incidence, a model from simulations relating the observables with this parameter will enable predicting the “effective” gametocytemia level (i.e., the transmissibility) in empirical cases.

To model the malaria incidence I we consider a continuous step function, varying from $I = 0$ for w_h to $I = 1$ for $w_h = 1$:

$$I = \frac{1}{1 + \left(\frac{A}{w_h}\right)^\alpha} \tag{6}$$

The dashed line in Figure 5a shows a function given by Equation (6) for $A = 0.562$ and $\alpha = 8.3$. Here, parameter A gives the gametocytemia level, which brings the malaria incidence to the level of 50%, while the value of parameter α indicates how abrupt the transition from eradication to prevalence occurs when increasing the gametocytemia level.

Through inspection of Figure 5b, we choose to model the Hurst exponent H by a power law

$$H = Bw_h^\beta, \tag{7}$$

for which the best fit yields $B = 0.56$ and $\beta = 0.58$. Here, parameters have no direct interpretation.

Finally, to address the abrupt transition observed for the entropy when the gametocytemia level varies, we choose a step function tuned by an exponential of w_h :

$$S = \frac{1}{1 + C \exp(-\gamma w_h)}, \tag{8}$$

with the best fit yielding $C = 10,136$ and $\gamma = 32$. The parameter C tunes how low the entropy is for the extreme case of $w_h = 0$, while, similarly to parameter α , the γ controls how abrupt the transition from that minimum level to $S = 1$ occurs.

5.2. Prediction of Effective Gametocytemia for Empirical Cases

In relation to the prediction of “effective” gametocytemia in the empirical cases we first invert the functions defined in Equations (6)–(8) with respect to the gametocytemia level. We call to the values obtained estimates of gametocytemia, which, in general, do not coincide:

$$w_h^{(I)} = A \left(\frac{1}{I} - 1 \right)^{-1/\alpha}, \tag{9a}$$

$$w_h^{(H)} = \left(\frac{H}{B} \right)^{-1/\beta}, \tag{9b}$$

$$w_h^{(S)} = -\frac{1}{\gamma} \log \left(\frac{1}{C} \left(\frac{1}{S} - 1 \right) \right). \tag{9c}$$

In the case that all three models, in predicting w_h , retrieve the same value, we can assume almost zero error (maximum consistency of the models). In general, there will be deviations between the three predictions for gametocytemia level. Therefore we take, as an estimate \hat{w}_h for the gametocytemia level, the average of the three independent predictions, and, as the corresponding error $\sigma_{\hat{w}_h}$, the largest deviation of the independent predictions from that estimate:

$$\hat{w}_h = \frac{1}{3} \left(w_h^{(I)} + w_h^{(H)} + w_h^{(S)} \right), \tag{10a}$$

$$\sigma_{\hat{w}_h} = \max \left(\left| \hat{w}_h - w_h^{(I)} \right|, \left| \hat{w}_h - w_h^{(H)} \right|, \left| \hat{w}_h - w_h^{(S)} \right| \right). \tag{10b}$$

Figure 6 repeats the models drawn for the simulations of the agent-based model, together with the estimate for the eight empirical series. While malaria incidence and the Hurst exponent retrieve reasonably acceptable predictions, the entropy seems to be very sensitive. The reason for this may be related with the fact that the values of the entropy are all very similar, making difficult to derive a numerical model that distinguishes between the different values. Moreover, the errors are typically large, showing a broad range of different predictions depending on which models are used (cf. Equation (9)).

In Figure 6 we plot the curves in Figure 5 together with estimates for the eight empirical cases. To predict the gametocytemia in empirical datasets, we assume that the range of values of I , H and S observed for the collection of empirical series covers the range of admissible values between a minimum and a maximum. The same occurs for the collection of our simulations, but, since there is no guarantee of proper calibration, the minimum and maximum values may be different. Still, assuming that those values obtained from simulations should cover the same range of possibilities as those in empirical cases, we normalize the range of observed values in the empirical cases to the range observed for the simulations. This is a necessary step to predict effective gametocytemia, as explained below.

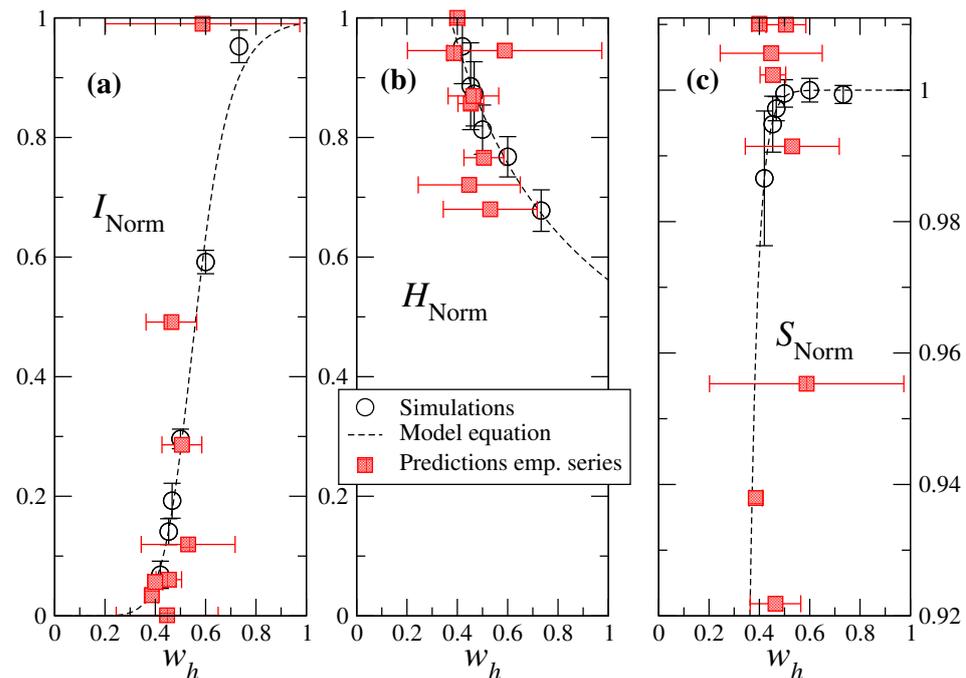


Figure 6. Repeating Figure 5 with the effective gametocytemia for each empirical series (see text).

6. Discussion and Conclusions

The utility of time series models is still a long way from becoming standard practice in malaria prevention. Differences in climate and geographic factors between world regions act as confounding factors in the strictly mathematical time series approach, lowering malaria forecast precision. In recent years the Box–Jenkins theory has become a consistent development in malaria forecasting [5,18,35,36]. However, little attention has been devoted to the Hurst theory, information entropy, short- and long-memory stochastic processes or long-range dependence. It is remarkable that the Hurst theory was initially implemented in the field of hydrology, as malaria surges are clearly correlated with rainfall, temperature and climate seasonality [6].

Malaria epidemic time series consistently present different memory patterns, depending on disease transmission intensity. By comparing time series from our model simulations to real-data malaria time series from different parts of the world, it was possible to obtain a better definition of epidemic stability, according to disease transmission efficiency, from field data time series results. In stationary time series, long-memory processes have been related to the presence of the long-range dependence (LRD) between present and past results [37].

At low H-to-M disease transmission intensity, time series patterns were consistent with the presence of LRD. However, at high disease transmission intensity, this pattern reverted to a low-memory process. By looking at the present model time series with changing w_h , one could witness significant differences in stochastic memory patterns. Additionally, in the presented empirical time series, the Hurst exponent and entropy correlated reasonably well with different epidemic growth rates. A similar pattern was evident when looking at malaria incidence correlation with the Hurst exponent and Shannon entropy. As these parameters may be affected by the time series length, their use in a normalized setting should be considered as a reliable option.

By using the standardized forms of the Hurst exponent (H_{36}) and entropy measurement (S_{36}) it was possible to define the type of memory of stochastic malaria incidence time series with greater precision. This fact may be of significant relevance as both parameters may become additional and useful tools in malaria forecasting.

In this paper we used the 36-month standard time length for a specific analysis. We considered it a compromise between a shorter time length (24 months) with less information available to the Hurst exponent estimation, and a longer time length (48 months) with less data available for analysis in shorter empirical time series, such as the series of Appiah et al. [16] (60 months). Other alternative methods, such as the generalized Hurst exponent and adaptive fractal analysis [38,39], could be applied, but, in our case, the results did not show significant improvements. See Appendix B.

Being a standard approach in time series analysis, we also performed SARIMA models for the simulation and empirical cases. Figures 7 and 8 show two illustrative cases of each. Details of the SARIMA model are given in Appendix C. In the case of the seasonal component, the regressive component of the SARIMA term vanishes in the Okech model such that we have $(P, D, Q) = (0, 1, 1)$. On the contrary, the seasonal SARIMA model of Landoh is purely a regressive model with one-time differencing, where we have $(P, D, Q) = (1, 1, 0)$. The SARIMA forecast of both empirical series is presented in Figure 8, while its equations and coefficients are available in Table A4.

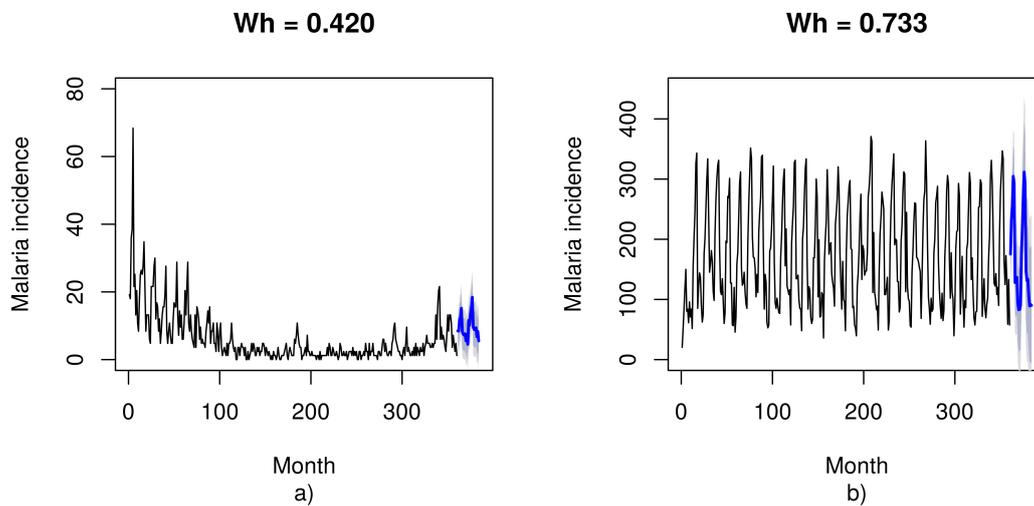


Figure 7. Model simulation (black) and SARIMA model forecasting (blue) of the model simulation at the phase transition: (a) low human-to-mosquito transmission efficiency ($w_h = 0.420$); (b) high human-to-mosquito transmission efficiency ($w_h = 0.733$).

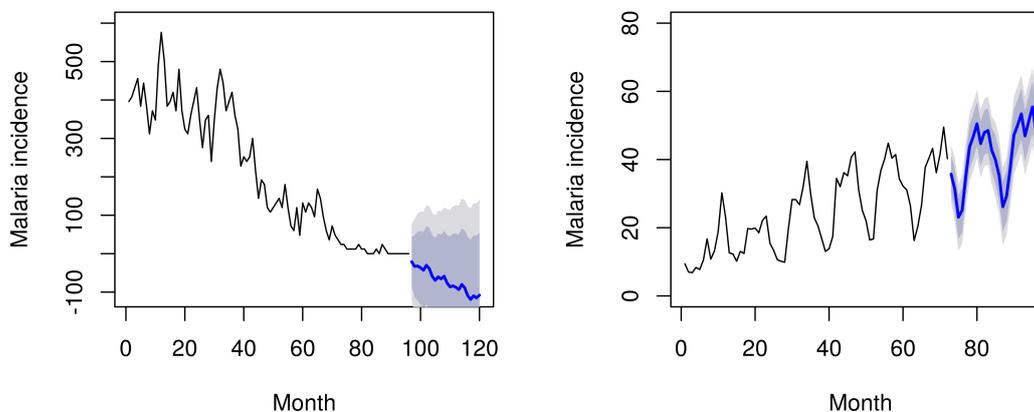


Figure 8. (Left) Empirical time series (black) and SARIMA model forecasting (blue) from Okech, 2008 [27] empirical time series, along with declining malaria incidence. The SARIMA forecast predicts rapid disease elimination evolving towards negative values. (Right) Empirical time series (black) and SARIMA model forecasting (blue) from the Landoh, 2012 [25] empirical time series, along with a steady increase in malaria incidence. The SARIMA forecast predicts a trend of progressive diseasespreading.

All together, our results seem to indicate that, in malaria incidence time series, long-range dependence may occur close to the phase transition between epidemic stability and disease elimination. The presence of these long-memory stochastic processes in malaria incidence time series could become an additional and useful tool in the early detection of epidemic resurgence, as well as a potential improvement in malarial prevention strategies.

Author Contributions: J.S. implemented the model and performed the simulations, J.S. and P.G.L. structured the text and produced the figures. J.L. and A.M.M. assisted with background ideas and the overall research was coordinated by J.S. and P.G.L. All authors revised and contributed to the final version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The empirical datasets in this article are provided in the respective cited references. See Refs. [16,21–27].

Acknowledgments: We thank Miguel Prudêncio (IMM-Lisbon) for his help in defining model parameters, and for his crucial remarks in discussing the model design.

Conflicts of Interest: No competing interests apply to this manuscript.

Appendix A. Qualitative Analysis of Hurst Exponent and Entropy: Case-by-Case Description

Appendix A.1. Alhassan (2017)

In the Alhassan malaria time series (Kasena Nankana municipality in Ghana, 2017) [21] one can see a declining trend in malaria incidence in the form of a decreasing moving average of 36 months (I_{36}), from ~ 11 to ~ 4 cases per 100 inhabitants, per year [phy], despite the small final outbreak in the last 6 months. Consistent with the declining trend in I_{36} , we may also witness a decreasing trend in S_{36} , as well as a rising trend in H_{36} after month 55. In this case, the behavior of both the Hurst exponent and entropy in the preceding 36 months is well correlated with the malaria incidence trend—see Figures A1–A3.

Appendix A.2. Appiah (2015)

By looking at the Appiah malaria time series (Ejisu-Juaben municipality in Ghana, 2015) [16] an irregular oscillation of malaria incidence is detectable, superimposed on a stable trend in malaria incidence ~ 12.4 cases phy , with a range between a peak incidence of ~ 20 cases phy at 32 months and an all-time-low of ~ 5 cases phy at 18 months—see Figure 1. Consistent with the initial declining trend in I_{36} until month 53, we may also witness a declining trend in S_{36} as well as a rising trend in H_{36} . In the present case, the behaviors of both the Hurst exponent and entropy in the initial 53 months are well correlated with the malaria incidence trend—see Figures A1–A3.

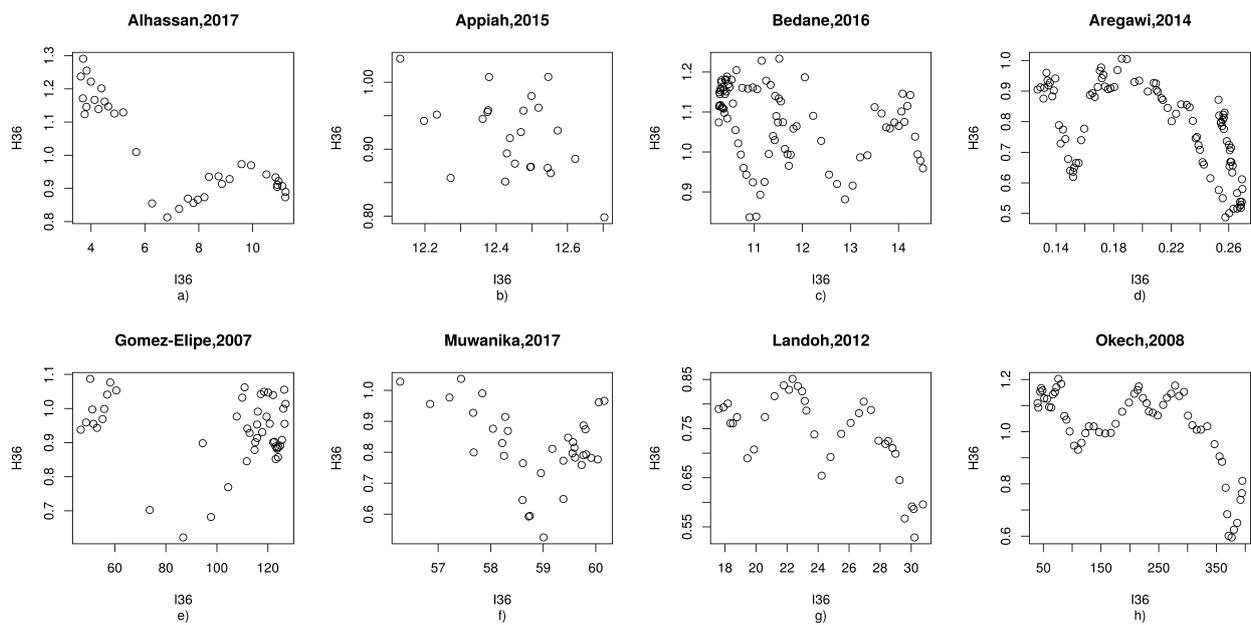


Figure A1. Illustration of how a 36-month moving average malaria incidence in the x -axis, symbolized as I_{36} , relates to the 36-month Hurst exponent H_{36} in the y -axis in malaria empirical time series: (a) Alhassan et al. (Kasena Nankana municipality in Ghana, 2017) [21]. (b) Appiah et al. (Ejisu-Juaben municipality in Ghana, 2015) [16], (c) Bedane et al. (Kucha district in Ethiopia, 2016) [23], (d) Aregawi et al. (Ethiopia, 2014) [22], (e) Gomez-Elipe et al. (Karuzi in Burundi, 2007) [24], (f) Muwanika et al. (Uganda, 2017) [26], (g) Landoh et al. (Est Mono district in Togo, 2012) [25], (h) Okech et al. (Kenya, 2008) [27].

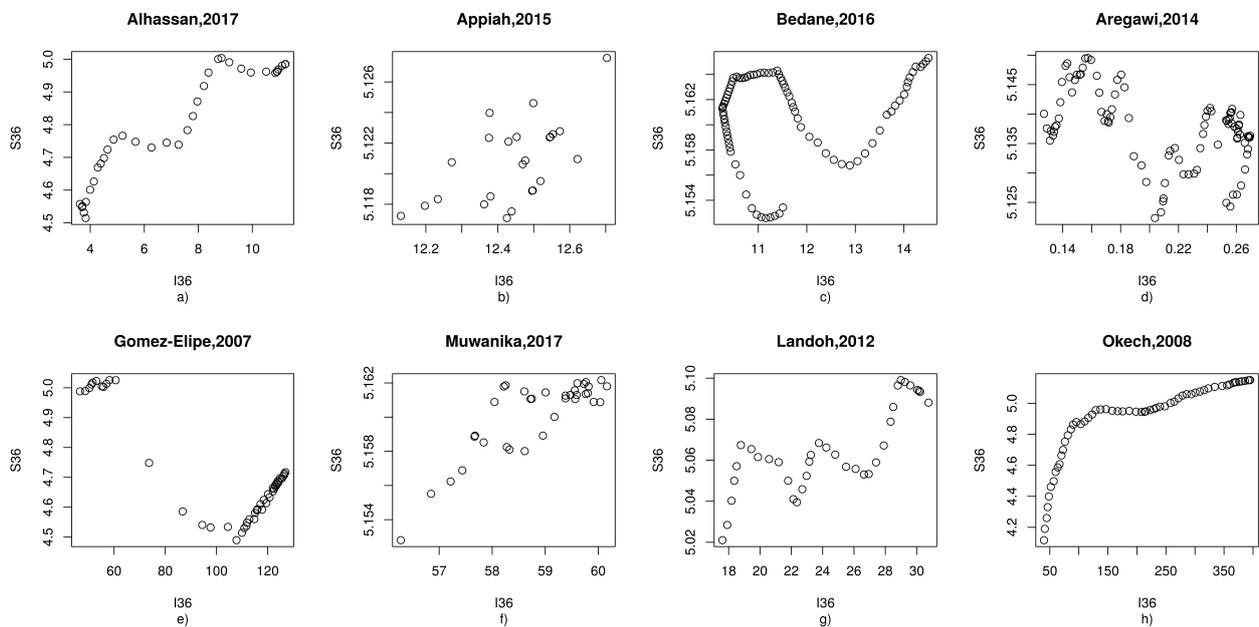


Figure A2. Illustration of how a 36-month moving average malaria incidence in the x -axis, symbolized as I_{36} , relates to the 36-month entropy S_{36} in the y -axis in malaria empirical time series: (a) Alhassan et al. (Kasena Nankana municipality in Ghana, 2017) [21]. (b) Appiah et al. (Ejisu-Juaben municipality in Ghana, 2015) [16], (c) Bedane et al. (Kucha district in Ethiopia, 2016) [23], (d) Aregawi et al. (Ethiopia, 2014) [22], (e) Gomez-Elipe et al. (Karuzi in Burundi, 2007) [24], (f) Muwanika et al. (Uganda, 2017) [26], (g) Landoh et al. (Est Mono district in Togo, 2012) [25], (h) Okech et al. (Kenya, 2008) [27].

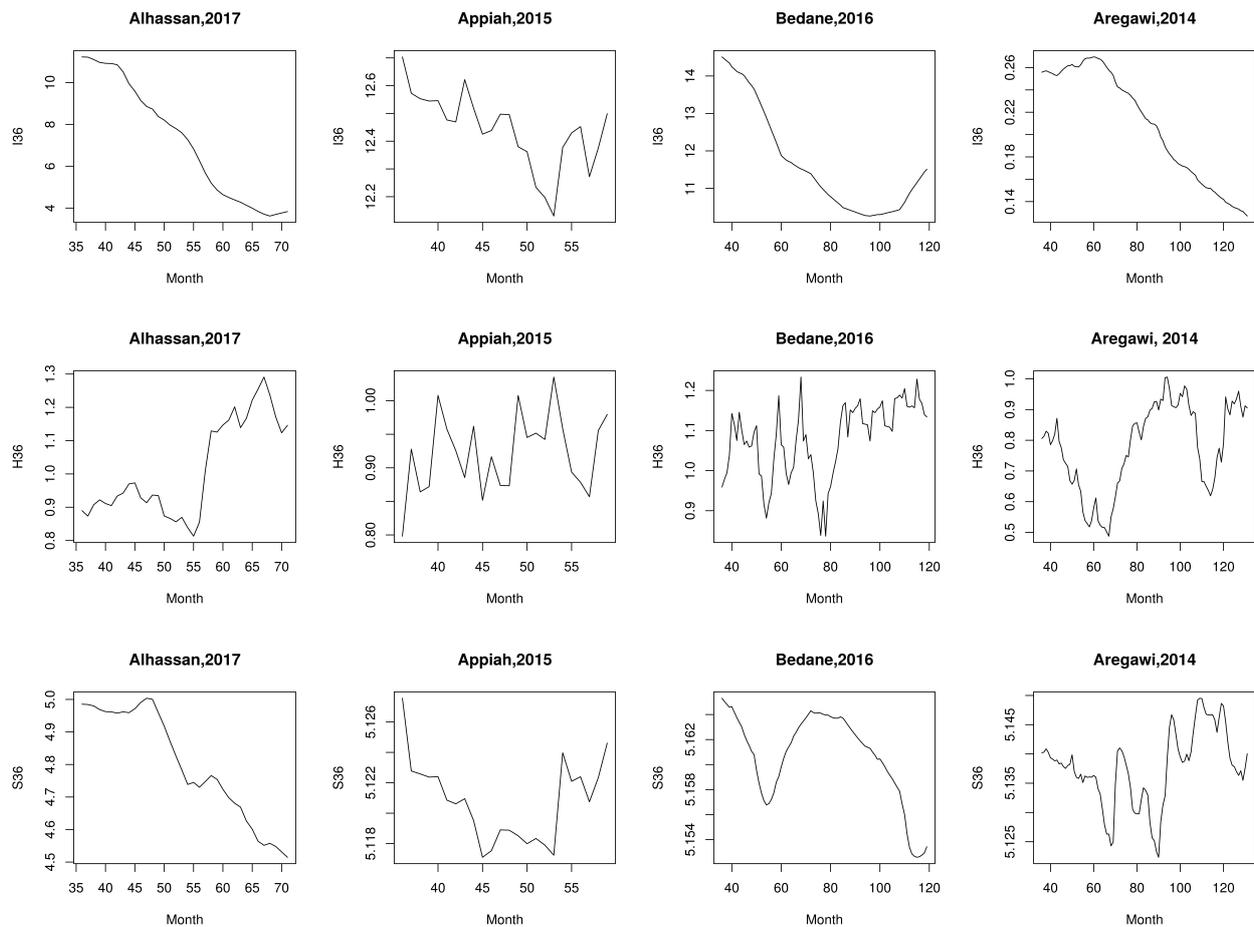


Figure A3. Comparative depiction of I_{36} (top), H_{36} (middle) and S_{36} (bottom) for (from left to right) Alhassan et al. (Kasena Nankana municipality in Ghana, 2017) [21], Appiah et al. (Ejisu-Juaben municipality in Ghana, 2015) [16], Bedane et al. (Kucha district in Ethiopia, 2016) [23], and Aregawi et al. (Ethiopia, 2014) [22].

Appendix A.3. Bedane (2016)

In relation to the Bedane malaria time series (Kucha district in Ethiopia, 2016) [23] we can witness an initial declining trend in malaria incidence in the form of a decreasing moving average of 36 months (I_{36}), lasting until month 98, from ~ 15 cases *phy*, to ~ 10 cases *phy*, followed by a small final upsurge in I_{36} to ~ 11.5 cases *phy*. Consistent with the initial declining trend in I_{36} until month 98, we may also witness a delayed declining trend in S_{36} from months ~ 70 to ~ 115 , as well as a rising trend in H_{36} from months ~ 60 to ~ 80 , despite the presence of a superimposed irregular oscillatory noise pattern. In this case, the behavior of both the Hurst exponent and entropy in the initial ~ 98 months partially shows some degree of correlation with the global malaria incidence trend—see Figures A1–A3.

Appendix A.4. Aregawi (2014)

In the Aregawi malaria time series (Ethiopia, 2014) [22] we can witness an initial small upper trend in malaria incidence in the form of an increasing moving average of 36 months (I_{36}) in the initial ~ 62 months, with ~ 0.26 cases per 100 inhabitants, per year, declining thereafter to less than ~ 0.12 cases per 100 inhabitants, per year.

Consistent with the initial small upper trend in I_{36} , we also witness a declining trend in H_{36} after month 62. In this case, entropy revealed an atypical behavior, peaking a littler later, at month ~ 72 . Along with a declining trend in I_{36} after month 62, one can witness a consistent rise in H_{36} , lasting to the end of the time series despite a transitory fall at month ~ 105 , with a rapid recovery at month ~ 114 . In the Aregawi time series, the behavior of S_{36}

was more unpredictable, with a more delayed response. This fact may be somehow related to the low malaria incidence in the time series.

However, in the present case, the behavior of the Hurst exponent in the form of a 36-month moving average (H_{36}) in the initial 65 months still reveals some degree of correlation with the global malaria incidence trend—see Figure A3. At such low levels of malaria incidence (~ 0.2 *phy*) this behavior could be interpreted as a possible outlier result—see Figures A1–A3.

Appendix A.5. Gomez-Elipe (2007)

The Gomez-Elipe time series (Karuzi in Burundi, 2007) [24] is quite different from the remaining empirical examples. It shows a consistent stable pattern of malaria incidence below 110 cases *phy*, until month 45, peaking to ~ 500 cases *phy*, around month 48, with a rapid fall to ~ 60 cases *phy* at month 56, and with a slower decline thereafter to a final value of ~ 20 cases *phy*—see Figure 1. The 36-months malaria incidence (I_{36}) reveals a consistent upward trend peaking at month 64, with a downward pattern thereafter.

In this case, we find a persistent oscillatory behavior in the value of H_{36} during the entire time series, usually ranging above 0.9. Yet, coinciding with the outbreak in malaria incidence at month 48, it is possible to see a sudden fall in H_{36} to values close to 0.6 (closer to random noise). Entropy in the form of a 36-month moving average (S_{36}) reveals a parallel correlated behavior to I_{36} from month 48 onwards. Despite the presence of a superimposed oscillatory noise pattern, the behaviors of the Hurst exponent and entropy also show some degree of correlation with the malaria incidence trend, globally—see Figures A1, A2 and A4.

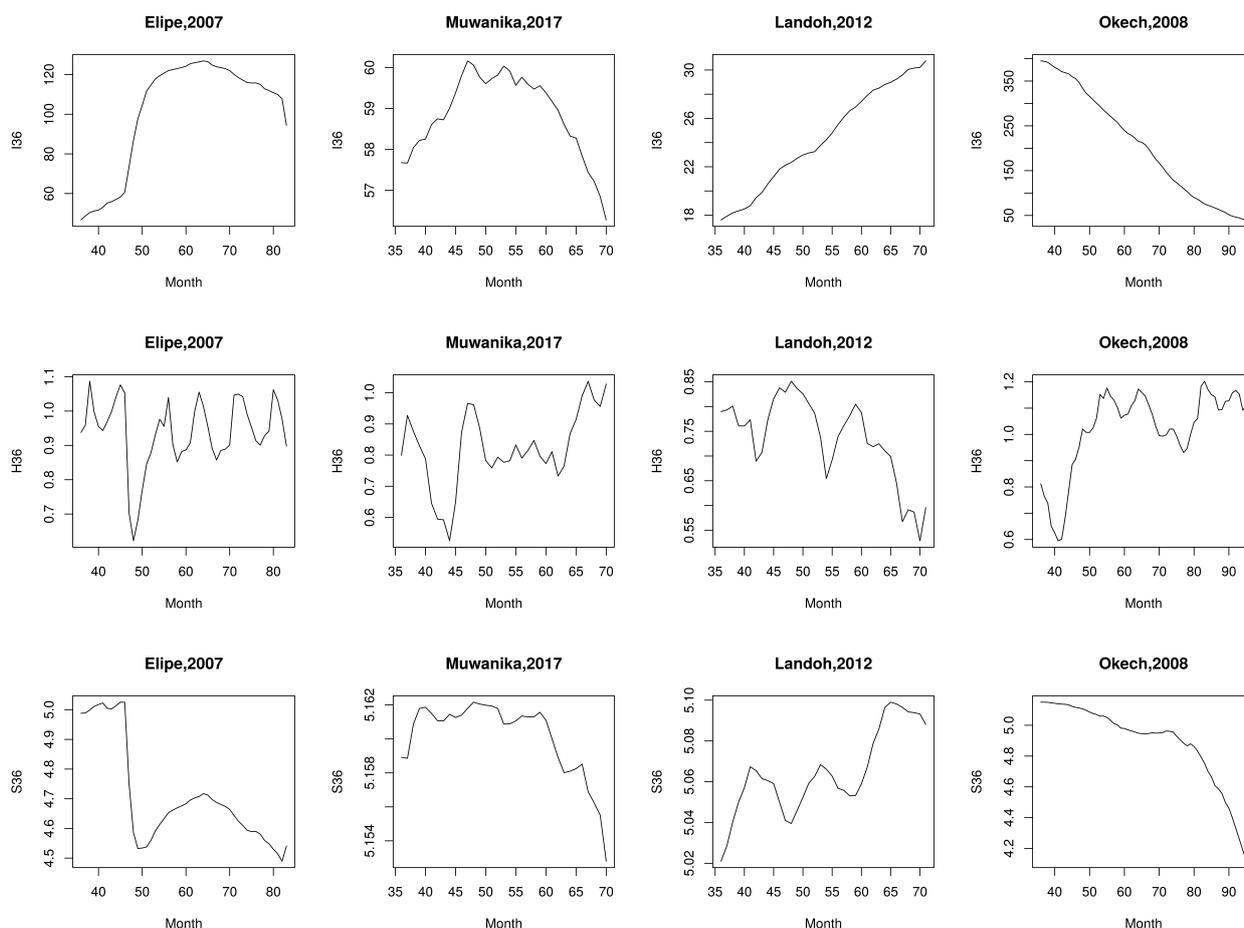


Figure A4. (Cont. Figure A3): Gomez-Elipe et al. (Karuzi in Burundi, 2007) [24], Muwanika et al. (Uganda, 2017) [26], Landoh et al. (Est Mono district in Togo, 2012) [25], Okech et al. (Kenya, 2008) [27].

Appendix A.6. Muwanika (2017)

In the Muwanika malaria time series (Uganda, 2017) [26] we may witness an initial small upper trend in malaria incidence in the form of a mild increase in the 36-month moving average (I_{36}) during the initial ~ 47 months, with a peak at ~ 60 cases *phy*, declining thereafter to ~ 56 cases *phy*.

Consistent with the initial small upper trend in I_{36} until month 47, we may also witness a declining trend in H_{36} from an initial value ~ 0.9 to ~ 0.5 close to month 45. From month 47 onwards, H_{36} consistently increases to values close to 1.0 (long memory process) in parallel with the steady decline in I_{36} . In the present case, entropy, in the form of a 36-month moving average (S_{36}), reveals a consistent decreasing trend, shadowing the decline in I_{36} beginning at \sim month 47. In the Muwanika time series, the behaviors of H_{36} and S_{36} are globally correlated with the behavior of I_{36} in time—see Figures A1, A2 and A4.

Appendix A.7. Landoh (2012)

The Landoh malaria time series (Est Mono district in Togo, 2012) [25] reveals a consistent upper trend in malaria incidence in the form of a mild increase in the 36-month moving average (I_{36}) over the whole time series, ranging from the initial ~ 18 cases *phy*, to a peak at ~ 30 cases *phy* in month 72.

Consistent with the steady upper trend in I_{36} during the entire time series, we find a declining trend in the 36-month moving average of the Hurst exponent (H_{36}) from a peak of ~ 0.85 at month 47, to an all-time-low value of ~ 0.55 . Along with this downward trend in H_{36} , there is a consistent increase in entropy in the form of a 36-month moving average (S_{36}), during the entire time series, despite its undulatory behavior. In the Landoh time series, the behaviors of H_{36} and S_{36} are reasonably well correlated with the behavior of I_{36} in time—see Figure A4.

Appendix A.8. Okech (2008)

In a similar form, but with an opposite trend to the Landoh time series, the Okech time series (Kenya, 2008) [27] also reveals a steady decreasing trend. In the Okech time series, malaria incidence consistently decreases from an initial peak of ~ 400 cases *phy*, to a final value close to 50 cases *phy*.

Consistent with the steady downward trend in I_{36} during the entire time series, we find an upper trend in the 36-month moving average of Hurst exponent (H_{36}), from an all-time-low of ~ 0.6 at month 42, to a peak of ~ 1.2 at month 82. Along with this upper trend in H_{36} , there is a consistent decrease in entropy in the form of a 36-month moving average (S_{36}) over the whole time series. In the Okech time series, the behaviors of H_{36} and S_{36} are also quite correlated with the behavior of I_{36} in time—see Figures A1, A2 and A4.

Appendix B. Inspecting the Robustness of 36-Month Averages

The results in the previous appendix have been derived using average values in windows of 36 months, i.e., three years. While the number of points is small, it covers two annual cycles. As it is known—and as shown in Appendix C—malaria spreading shows periodic behavior following annual seasonality.

To evaluate the robustness of the estimated values for malaria incidence, the Hurst exponent and entropy, shown in the previous appendix and discussed in the main text, we present, next, the results for estimates done from two and four annual cycles, i.e., 24 and 48 months, respectively. The results are shown in Figures A5–A8.

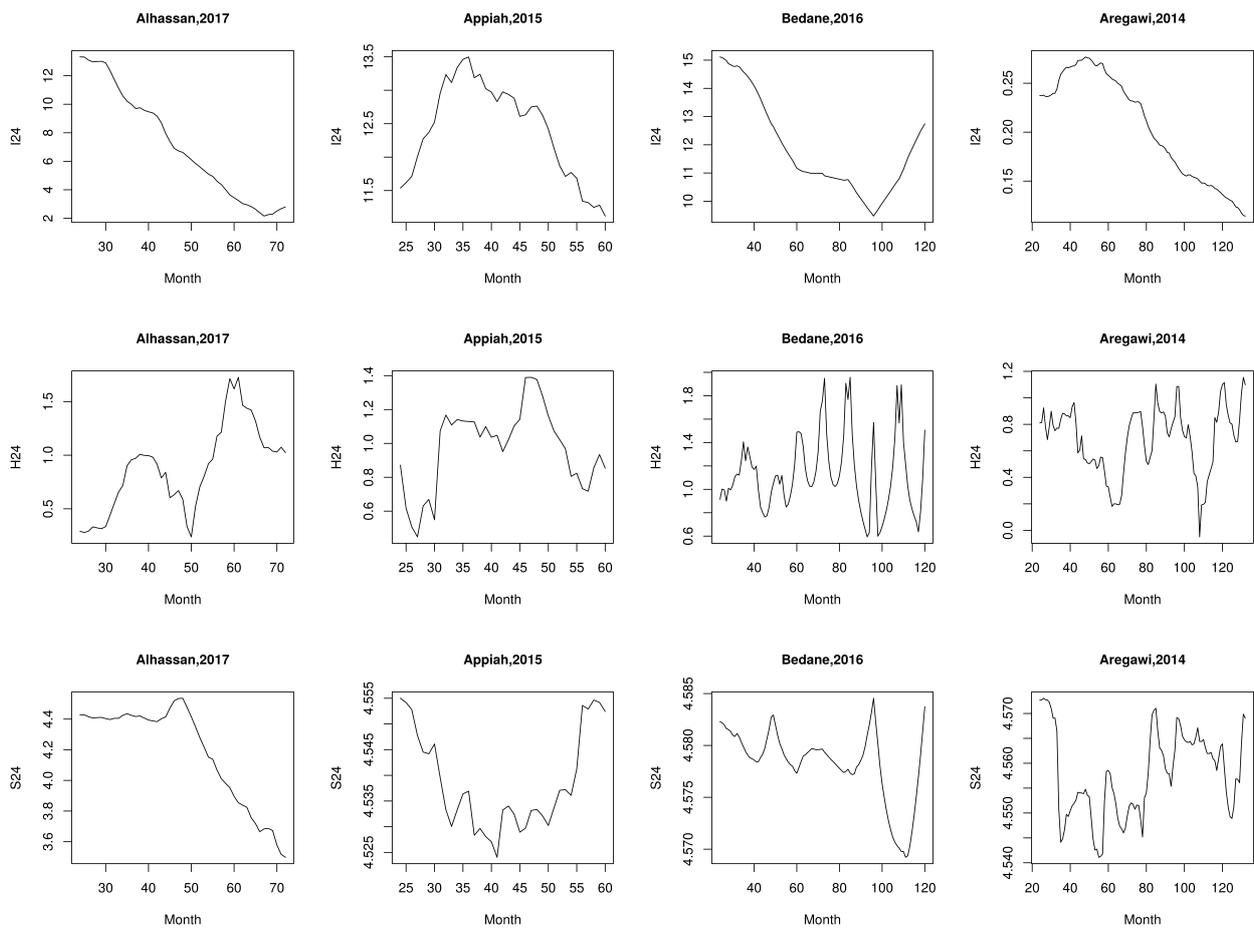


Figure A5. Comparative depiction of I_{24} (top), H_{24} (middle), S_{24} (bottom) for (from left to right): Alhassan et al. (Kasena Nankana municipality in Ghana, 2017) [21], Appiah et al. (Ejisu-Juaben municipality in Ghana, 2015) [16], Bedane et al. (Kucha district in Ethiopia, 2016) [23], and Aregawi et al. (Ethiopia, 2014) [22].

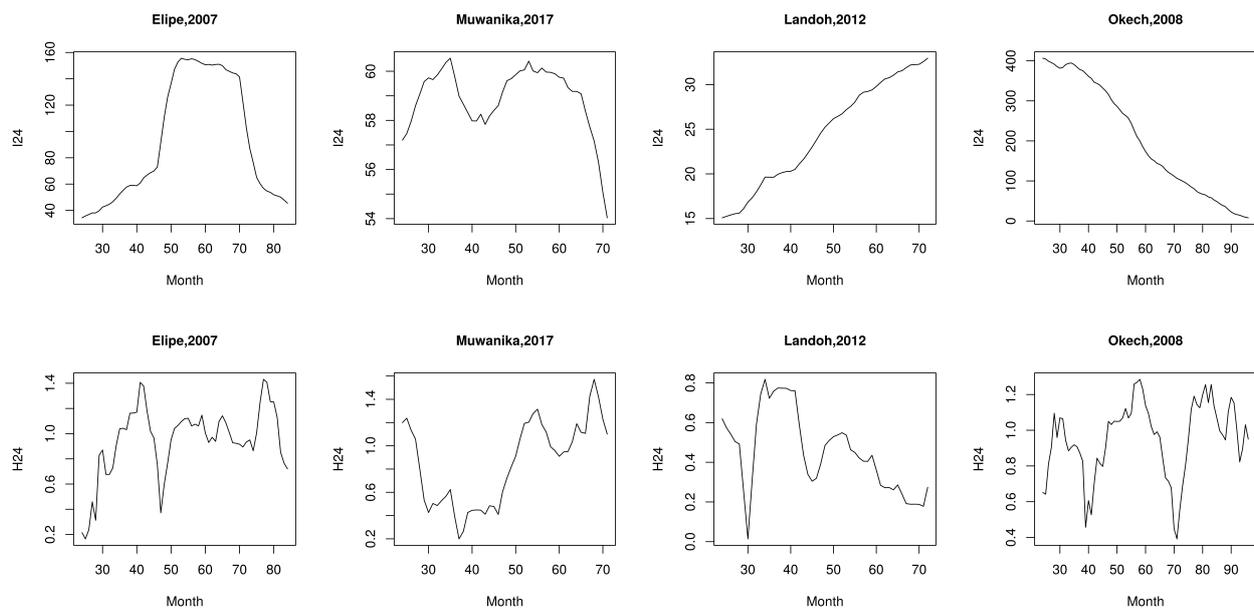


Figure A6. Cont.

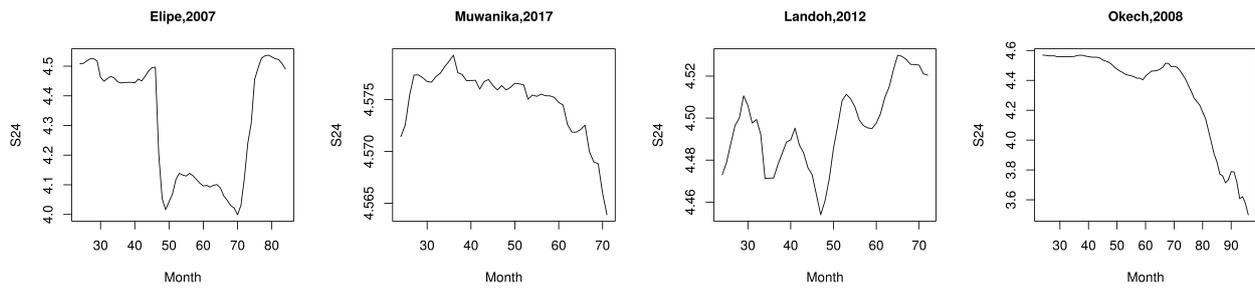


Figure A6. (Cont. Figure A5): Gomez-Elipe et al. (Karuzi in Burundi, 2007) [24], Muwanika et al. (Uganda, 2017) [26], Landoh et al. (Est Mono district in Togo, 2012) [25], and Okech et al. (Kenya, 2008) [27].

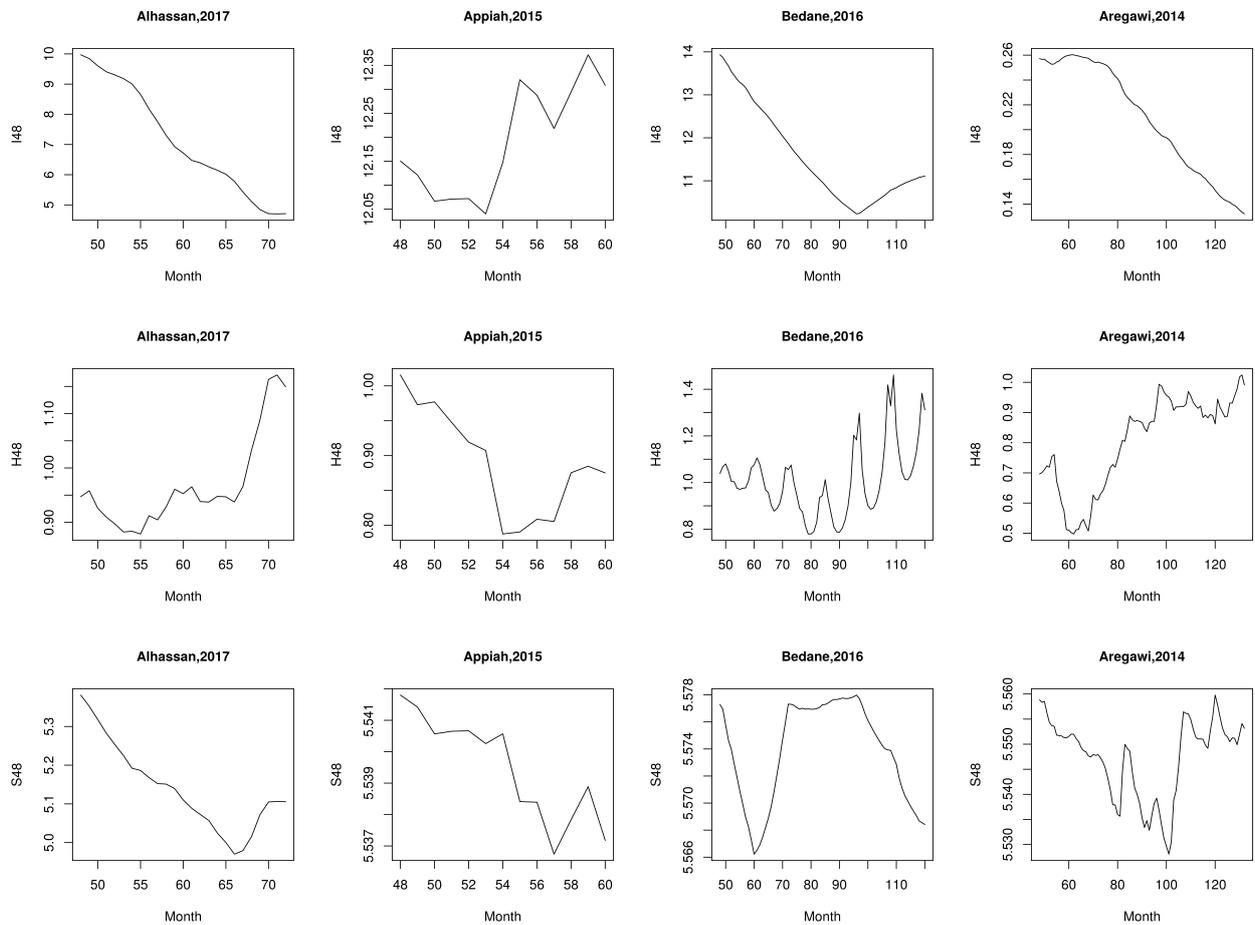


Figure A7. Comparative depiction of I_{48} (top), H_{48} (middle), S_{48} (bottom) for (from left to right): Alhassan et al. (Kasena Nankana municipality in Ghana, 2017) [21], Appiah et al. (Ejisu-Juaben municipality in Ghana, 2015) [16], Bedane et al. (Kucha district in Ethiopia, 2016) [23], and Aregawi et al. (Ethiopia, 2014) [22].

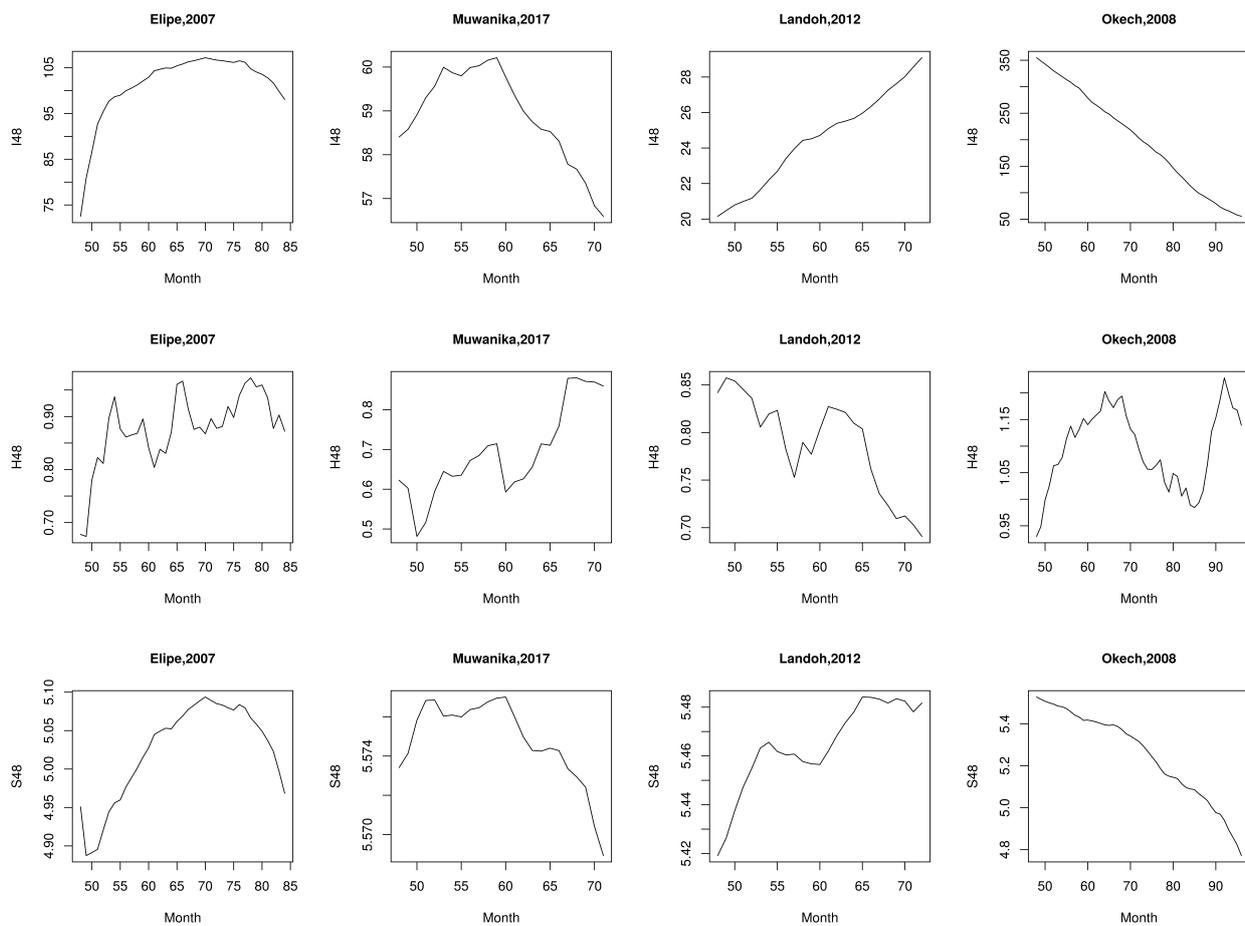


Figure A8. (Cont. Figure A7): Gomez-Elipe et al. (Karuzi in Burundi, 2007) [24], Muwanika et al. (Uganda, 2017) [26], Landoh et al. (Est Mono district in Togo, 2012) [25], and Okech et al. (Kenya, 2008) [27].

We choose to use the standard method to estimate the Hurst exponent, based on the quotient R/S [34,37,40–43]. An alternative method could be the so-called generalized Hurst exponent (GHE), defined from the function

$$K_q(\tau) = \frac{\langle |I(t + \tau) - I(t)|^q \rangle}{\langle I(t)^q \rangle}, \tag{A1}$$

where $\langle \cdot \rangle$ is the average over time t . The GHE H_G depends in general on q as [44]

$$K_q(\tau) \sim \tau^{qH_G(q)}. \tag{A2}$$

A general Hurst exponent varies with the moment q flags the presence of multifractality in the series.

GHE is a modern approach to Hurst analysis with some specific advantages applied in the analysis of complex and inhomogeneous time series in electrocardiography (ECG) signals, and it has been shown to be a promising tool for the study of atrial fibrillation (AF) organization from surface ECG [45]. It is usually recommended in the presence of short time series, where it has been shown to be slightly more efficient. It has been used mainly in the assessment of the stability of financial firms as applied to the stock market. However, GHE does not solve the problem of small time series, such as the ones analysed in this paper. Moreover, deriving the exponent from the scaling relation in Equation (A2) imposes a linear relation between $\log K_q$ and $\log \tau$ that usually occurs only for the smallest values of τ . Consequently, it may overestimate more recent events, reducing the significance of

past events. This is indeed the case for the GHE derived for Okech (see Figure 1), where, as malaria is eradicated towards the end of the data record, it can wrongly show an anti-persistence value ($H_g < 0.5$). See Table A1. For this reason we choose the R/S analysis to estimate the Hurst exponent.

Table A1. Table with the general Hurst exponent (GHE) computed, using Equation (A2), for the entire time series and for $q = 1$ and $q = 2$ of the eight empirical time series.

Empirical Series	GHE for $q = 1$	GHE for $q = 2$
Aregawi [22]	0.48	0.49
Alhassan [21]	0.48	0.56
Bedane [23]	0.73	0.48
Appiah [16]	0.21	0.09
Landoh [25]	0.65	0.62
Muwanika [26]	0.41	0.37
Elipe [24]	0.73	0.60
Okech [27]	0.26	0.42

Appendix C. Arima Models of the Incidence of Malaria

Models for single time series in epidemiology and infectious disease research have been widely used [11,15–18,35,46–53]. In this section, malaria time series will be analysed from a different perspective, involving simulations in two different human-to-mosquito transmission efficiency w_h settings ($w_h = 0.420$ and $w_h = 0.733$), based upon our malaria transmission agent-based stochastic model [3].

Those results will be compared with two empirical time series with different transmission patterns from separate world regions where malaria is still endemic, “Okech” and “Landoh” [25,27]. The Okech time series reveals a pattern consistent with an unstable epidemic state, with a decreasing trend from high levels of malaria incidence, evolving to imminent disease elimination. The other empirical time series reveal an unstable epidemic behavior, but, with a non-stationary increasing trend and an average level of malaria incidence by African standards of between 10 and 50 malaria cases per 100 inhabitants per year. See Table A2.

Table A2. (Top) Human-to-mosquito disease transmission efficiency (low, with $w_h = 0.420$, and high, with $w_h = 0.733$), and average malaria incidence (\pm SE) of ten simulations in all settings. **(Bottom)** Empirical time series with average malaria incidence (\pm SD) in different settings of malaria transmission.

Level	w_h	Malaria Transmission Intensity	Malaria Incidence (\pm SE)	
Low	0.420	Phase transition	7.7 (\pm 0.89)	
High	0.733	Epidemic stability	169.7 (\pm 0.64)	
Series	Follow-Up (months)	Region	Trend	Malaria incidence (\pm SD)
Okech, 2008	96	Kenya	decreasing	205.1 (\pm 169.8)
Landoh, 2012	72	Togo	increasing	24.4 (\pm 11.5)

As our model shows annual seasonality approximately, we choose a SARIMA model with a period of 12 months:

$$\Phi^P(B^s)\phi^p(B)\nabla_s^D\nabla^dX_t = \Theta^Q(B^s)\theta^q(B)Z_t, \tag{A3}$$

where Z_t represents a Gaussian white noise time series (i.e., $Z_t = N(0, \sigma^2)$) and the ordinary ∇ operator represents the one-difference operator, $\nabla X_t = X_t - X_{t-1}$, $\nabla^d X_t = (1 - B)^d X_t$ is the ordinary differencing level ∇^d and $\nabla_s^D X_t = (1 - B^s)^D X_t$ is the differencing level of the seasonal ∇^D with periodicity s . Operator $\phi^p(B)$ is the ordinary autoregressive operator $\phi^p(B) = 1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \dots - \phi_p B^p$, $\theta^q(B)$ the ordinary moving average operator $\theta^q(B) = 1 + \theta_1 B + \theta_2 B^2 + \theta_3 B^3 + \dots + \theta_q B^q$, $\Phi^P(B^s)$ is the seasonal autoregressive operator $\Phi^P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \Phi_3 B^{3s} - \dots - \Phi_P B^{Ps}$, $\Theta^Q(B^s)$ is the seasonal moving average operator $\Theta^Q(B^s) = 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \Theta_3 B^{3s} + \dots + \Theta_Q B^{Qs}$. Here $B^s(X_t) = X_{t-s}$ is the seasonal - or s -order - backshift operator. In particular $B(X_t) = X_{t-1}$.

Table A3 shows all parameters of the SARIMA model for the two simulations, while Table A4 shows all parameters of the SARIMA model for the two simulations.

Table A3. (Top) H-to-M transmission efficiency (w_h) levels from model simulations and in relation to the SARIMA models (α = drift). **(Bottom)** Different H-to-M transmission efficiency levels ($w_h = 0.420$ and $w_h = 0.733$), and coefficients of the SARIMA models.

w_h	SARIMA $(p, d, q) \times (P, D, Q)^s$	Model $\Phi^P(B^s)\phi^p(B)\nabla_s^D\nabla^d X_t = \alpha + \Theta^Q(B^s)\theta^q(B)Z_t$
0.420	$(2, 1, 2) \times (2, 1, 2)^{12}$	$\Phi^2(B^{12})\phi^2(B)\nabla_{12}^1\nabla^1 X_t = \alpha + \Theta^2(B^{12})\theta^2(B)Z_t$
0.733	$(2, 1, 1) \times (2, 1, 1)^{12}$	$\Phi^2(B^{12})\phi^2(B)\nabla_{12}^1\nabla^1 X_t = \alpha + \Theta^1(B^{12})\theta^1(B)Z_t$
w_h	0.420	0.733
$(p, d, q) \times (P, D, Q)^s$	$(2, 1, 2) \times (2, 1, 2)^{12}$	$(2, 1, 1) \times (2, 1, 1)^{12}$
ϕ_1	-0.3635	-0.3495
ϕ_2	0.3415	-0.2344
θ_1	-0.3455	-0.2997
θ_2	-0.5422	-0.2221
Φ_1	0.6564	-0.2273
Φ_2	-0.1574	-0.7408
Θ_1	-1.4685	-
Θ_2	0.7626	-
α (drift)	-	-
σ^2	10.89	1037
AICc	1849.32	3426.29

Table A4. (Top) SARIMA models of empirical time series from Okech [27] and Landoh [25]. **(Bottom)** Coefficients of SARIMA models of empirical time series from both examples.

Empirical Series	SARIMA $(p, d, q) \times (P, D, Q)^s$	$\Phi^P(B^s)\phi^p(B)\nabla_s^D\nabla^d X_t = \Theta^Q(B^s)\theta^q(B)Z_t$
Okech, 2008	$(1, 1, 1) \times (0, 1, 1)^{12}$	$\Phi^0(B^{12})\phi^1(B)\nabla_{12}^1\nabla^1 X_t = \Theta^1(B^{12})\theta^1(B)Z_t$
Landoh, 2012	$(1, 1, 1) \times (1, 1, 0)^{12}$	$\Phi^1(B^{12})\phi^1(B)\nabla_{12}^1\nabla^1 X_t = \Theta^0(B^{12})\theta^1(B)Z_t$
	Okech, 2008	Landoh, 2012
$(p, d, q) \times (P, D, Q)^s$	$(1, 1, 1) \times (0, 1, 1)^{12}$	$(1, 1, 1) \times (1, 1, 0)^{12}$
ϕ_1	0.6357	0.5306
θ_1	-0.8778	-1.000
Φ_1	-	-0.4894
Θ_1	-0.6522	-
σ^2	2518	17.49
AICc	900.8	351.15

From the present model simulations, the SARIMA linear models have been defined according to the level of gametocytemia and disease transmission efficiency. There are small differences in the type of the best SARIMA model between stochastic simulations at the same level of disease transmission intensity. The presented results have been obtained from the first-tested simulation at every transmission level.

Close to the phase transition at the lower transmission level (63 days gametocytemia duration; $w_h = 0.420$) the most reliable SARIMA model has been defined as $(2, 1, 2) \times (2, 1, 2)^{12}$, corresponding to a combination of auto-regressive ($p = 2$ and $P = 2$), and moving average ($q = 2$ and $Q = 2$) procedures, with the need of time series differencing (once, $d = 1$ and $D = 1$), both in the non-seasonal and the 12-month seasonal model fractions.

At the higher transmission level (110 days gametocytemia duration; $w_h = 0.733$) the best SARIMA model has been defined as $(2, 1, 1) \times (2, 1, 1)^{12}$, with a predominantly auto-regressive component ($p = 2$ and $P = 2$) and a lower-order moving average ($q = 1$ and $Q = 1$), with the need for time series differencing (once, $d = 1$ and $D = 1$), both in the non-seasonal and in the 12-month seasonal model fractions. In both cases, the best model has been defined according to Akaike information criterion optimization.

The two presented model scenario simulations of extreme human-to-mosquito transmission conditions ($w_h = 0.420$ and $w_h = 0.733$) have been compared to two typical empirical time series of malaria incidence in different geographical regions, from West and East Africa, with opposite disease transmission trends. They represent western and eastern Africa regions according to their apparent epidemic stability, seasonality and trends: West Africa (Togo) and East Africa (Kenya).

References

- Retief, F.; Cilliers, L. The death of Alexander the Great. *Acta Theol. Suppl.* **2005**, *7*, 14–28. [[CrossRef](#)]
- World Health Organization. *World Malaria Report*; World Health Organization: Geneva, Switzerland, 2020. Available online: <https://www.who.int> (accessed on 16 October 2021).
- Sequeira, J.; Louçã, J.; Mendes, A.M.; Lind, P.G. Transition from endemic behavior to eradication of malaria due to combined drug therapies: An agent-model approach. *J. Theor. Biol.* **2020**, *484*, 110030. [[CrossRef](#)] [[PubMed](#)]
- Sequeira, J.; Louçã, J.; Mendes, A.M.; Lind, P.G. A Model for Assessing the Quantitative Effects of Heterogeneous Affinity in Malaria Transmission along with Ivermectin Mass Administration. *Appl. Sci.* **2020**, *10*, 8696. [[CrossRef](#)]
- Box, G.E.P.; Jenkins, G.M. *Time Series Analysis: Forecasting and Control*; Holden Day: San Francisco, CA, USA, 1970.
- Hurst, H.E. The Long-Term Storage Capacity of Reservoir. *Trans. Am. Soc. Civ. Eng.* **1951**, *116*, 770–799. [[CrossRef](#)]
- Carbone, A.; Castelli, G.; Stanley, H.E. Time-dependent Hurst exponent in financial time series. *Physica A* **2004**, *344*, 267–271. [[CrossRef](#)]
- Mandelbrot, B.B.; Van Ness, J.W. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **1968**, *10*, 422–437. [[CrossRef](#)]
- Mandelbrot, B.B.; Wallis, J.R. Noah, Joseph and operational hydrology. *Water Resour. Res.* **1968**, *4*, 909–918. [[CrossRef](#)]
- Mandelbrot, B.B. *The Fractal Geometry of Nature*; W.H. Freeman Editors: San Francisco, CA, USA, 1982.
- Allard, R. Use of time-series analysis in infectious disease surveillance. *Bull. World Health Organ.* **1998**, *76*, 327–333.
- Permanasari, A.E.; Hidayah, I.; Bustoni, I.A. SARIMA (Seasonal ARIMA) implementation on time series to forecast the number of Malaria incidence. In Proceedings of the International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, Indonesia, 7–8 October 2013; Volume 2, pp. 203–207.
- Golestani, A.; Gras, R. Can we predict the unpredictable? *Sci. Rep.* **2014**, *4*, 6834. [[CrossRef](#)]
- Kumar, V.; Mangal, A.; Panesar, S.; Yadav, G.; Talwar, R.; Raut, D.; Singh, S. Forecasting malaria cases using climatic factors in delhi, India: A time series analysis. *Malar. Res. Treat.* **2014**, *2014*, 482851. [[CrossRef](#)]
- Musa, M.I. Malaria Disease Distribution in Sudan Using Time Series ARIMA Model. *Int. J. Public Health Sci.* **2015**, *4*, 7–16.
- Appiah, S.T.; Otoo, H.; Nabubie, I.B. Times Series Analysis Of Malaria Cases In Ejisu- Juaben Municipality. *Int. J. Sci. Technol. Res.* **2015**, *4*, 220–226.
- Sarkar, R.R.; Chatterjee, C. Application of Different Time Series Models on Epidemiological Data—Comparison and Predictions for Malaria Prevalence. *SM J. Biom. Biostat.* **2017**, *2*, 1022. [[CrossRef](#)]
- Ferrão, J.L.; Mendes, J.M.; Painho, M. Modelling the influence of climate on malaria occurrence in Chimoio Municipality, Mozambique. *Parasites Vectors* **2017**, *10*, 1–12. [[CrossRef](#)] [[PubMed](#)]
- Chatfield, C. *The Analysis of Time Series*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2006.
- Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and Its Applications*; Springer: Berlin/Heidelberg, Germany, 2006.
- Alhassan, E.A.; Isaac, A.M.; Emmanuel, A. Time Series Analysis of Malaria Cases in Kasena Nankana Municipality. *Int. J. Stat. Appl.* **2017**, *7*, 43–56.
- Aregawi, M.; Lynch, M.; Bekele, W.; Kebede, H.; Jima, D.; Taffese, H.S.; Yenehun, M.A.; Lilay, A.; Williams, R.; Thomson, M.; et al. Time series analysis of trends in malaria cases and deaths at hospitals and the effect of antimalarial interventions, 2001–2011. Ethiopia. *PLoS ONE* **2014**, *9*, e106359. [[CrossRef](#)]
- Bedane, A.S.; Tanto, T.K.; Asena, T.F. Malaria Distribution in Kucha District of Gamo Gofa Zone, Ethiopia: A Time Series Approach. *Am. J. Theor. Appl. Stat.* **2016**, *5*, 70–79. [[CrossRef](#)]

24. Gomez-Elipse, A.; Otero, A.; van Herp, M.; Aguirre-Jaime, A. Forecasting malaria incidence based on monthly case reports and environmental factors in Karuzi, Burundi, 1997–2003. *Malar. J.* **2007**, *6*, 129. [[CrossRef](#)]
25. Landoh, E.D.; Tchamdja, P.; Saka, B.; Tint, K.S.; Gitta, S.N.; Wasswa, P.; Christiaan, J. Morbidity and mortality due to malaria in Est Mono district, Togo, from 2005 to 2010: A times series analysis. *Malar. J.* **2012**, *11*, 389. [[CrossRef](#)]
26. Muwanika, F.R.; Atuhaire, L.K.; Ocaya, B. Prediction of Monthly Malaria Incidence in Uganda and its Implications for Preventive Interventions. *J. Med. Diagn. Meth.* **2017**, *6*, 2.
27. Okech, B.A.; Mwobobia, I.K.; Kamau, A.; Muiruri, S.; Mutiso, N.; Nyambura, J.; Mwatele, C.; Amano, T.; Mwandawiro, C.S. Use of integrated malaria management reduces malaria in Kenya. *PLoS ONE* **2008**, *3*, e4050. [[CrossRef](#)]
28. Ferrão, J.L.; Mendes, J.M.; Painho, M.; Zacarias, S. Malaria mortality characterization and the relationship between malaria mortality and climate in Chimoio, Mozambique *Malar. J.* **2017**, *16*, 212. [[CrossRef](#)]
29. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
30. Santamaria-Bonfil, G.; Gershenson, C.; Fernéz, N. A Package for Measuring emergence, Self-organization, and Complexity Based on Shannon entropy. *Front. Robot. AI* **2017**, *4*, 1–12. [[CrossRef](#)]
31. Anis, A.; Lloyd, E.H. The expected value of the adjusted rescaled Hurst range of independent normal summands. *Biometrika* **1976**, *63*, 283–298. [[CrossRef](#)]
32. Sánchez-Granero, M.A.; Trinidad-Segovia, J.E.; García-Pérez, J. Some comments on Hurst exponent and the long memory processes on capital markets. *Physica A* **2008**, *387*, 5543–5551. [[CrossRef](#)]
33. Barunik, J.; Kristoufek, L. On Hurst exponent estimation under heavy tailed distributions. *Physica A* **2010**, *389*, 3844–3855. [[CrossRef](#)]
34. Franzke, C.L.E.; Osprey, S.M.; Davini, P.; Watkins, N.W. A Dynamical Systems Explanation of the Hurst Effect and Atmospheric Low-Frequency Variability *Sci. Rep.* **2015**, *5*, 9068. [[CrossRef](#)]
35. Abeku, T.A.; De Vlas, S.J.; Borsboom, G.; Teklehaimanot, A.; Kebede, A.; Olana, D.; Van Oortmarsen, G.J.; Habbema, J.D.F. Forecasting malaria incidence from historical morbidity patterns in epidemic-prone areas of Ethiopia: A simple seasonal adjustment method performs best. *Trop. Med. Int. Health* **2002**, *7*, 851–857. [[CrossRef](#)] [[PubMed](#)]
36. Lima, M.V.M.; Laporta, G.Z. Evaluation of prediction models for the occurrence of malaria in the state of Amapá, Brazil, 1997–2016: An ecological study. *Epidemiol. Serv. Saúde Brasília* **2021**, *30*, e2020080. [[CrossRef](#)]
37. Graves, T.; Gramacy, R.; Watkins, N.; Franzke, C. A brief history of long memory: Hurst, Mandelbrot and the road to ARFIMA, 1951–1980. *Entropy* **2017**, *19*, 437. [[CrossRef](#)]
38. Gao, J.B.; Hu, J.; Tung, W.W. Tung, Facilitating joint chaos and fractal analysis of biosignals through nonlinear adaptive filtering. *PLoS ONE* **2011**, *6*, e24331. [[CrossRef](#)]
39. Gao, J.B.; Hu, J.; Mao, X.; Perc, M. Culturomics meets random fractal theory: Insights into long-range correlations of social and natural phenomena over the past two centuries. *J. R. Soc. Interface* **2012**, *9*, 1956–1964. [[CrossRef](#)] [[PubMed](#)]
40. Flynn, M.N.; Pereira, W.L.R.S. Ecological diagnosis from biotic data by Hurst exponent and the R/S analysis adaptation to short time series. *Biomatemática* **2013**, *23*, 1–14.
41. Corrêa, C.S.; Schuch, D.A.; Queiroz, A.P.; Fisch, G.; Corrêa, F.N.; Coutinho, M.M. The Long-Range Memory and the Fractal Dimension: A Case Study for Alcântara. *J. Aerosp. Technol. Manag.* **2017**, *9*, 461–468. [[CrossRef](#)]
42. Garcia, M.N.L.; Requena, J.P.R. Different methodologies and uses of the Hurst exponent in econophysics. *Estud. Econ. Appl.* **2019**, *37*, 96–108. [[CrossRef](#)]
43. Kirichenko, L.; Radivilova, T.; Bulakh, V. Generalized approach to Hurst exponent estimating by time series. *Inf. Control Meas. Econ. Environ. Prot.* **2018**, *8*, 28–31. [[CrossRef](#)]
44. Sensoy, A. Generalized Hurst exponent approach to efficiency in MENA markets. *Physica A* **2013**, *392*, 5019–5026. [[CrossRef](#)]
45. Julián, M.; Alcaraz, R.; Rieta, J.J. Generalized Hurst Exponents as a Tool to Estimate Atrial Fibrillation Organization from the Surface ECG. *Comput. Cardiol.* **2013**, *40*, 1199–1202.
46. Ganguly, K.S.; Modak, S.; Chattopadhyay, A.K.; Ganguly, K.S.; Mukherjee, T.K.; Dutta, A.; Biswas, D. Forecasting Based On a SARIMA Model of Urban Malaria for Kolkata. *Am. J. Epidemiol. Infect. Dis.* **2016**, *4*, 22–33.
47. Ostovar, A.; Haghdoost, A.A.; Rahimiforushani, A.; Raeisi, A.; Majdzadeh, R. Time series analysis of meteorological factors influencing Malaria in south eastern Iran. *J. Arthropod. Borne Dis.* **2016**, *10*, 222–237.
48. Getnet, B.; Ayalew, S. Developing Stochastic Model For Forecasting Malaria Cases In Addis Zemen, South Gondar, Ethiopia: A Time Series Analysis. *Nat. Sci.* **2017**, *15*, 64–77.
49. Anokye, R.; Acheampong, E.; Owusu, I.; Obeng, E.I. Time series analysis of malaria in kumasi : Using arima models to forecast Time series analysis of malaria in kumasi: Using arima models to forecast. *Cogent Soc. Sci.* **2018**, *3*, 1–13.
50. Helfenstein, U. The use of transfer function models, intervention analysis and related time series methods in epidemiology. *Int. J. Epidemiol.* **1991**, *20*, 808–815. [[CrossRef](#)] [[PubMed](#)]
51. Briët, O.J.T.; Vounatsou, P.; Gunawardena, D.M.; Galappaththy, G.N.L.; Amerasinghe, P.H. Models for short term malaria prediction in Sri Lanka. *Malar. J.* **2008**, *7*, 76. [[CrossRef](#)]

-
52. Wangdi, K.; Singhasivanon, P.; Silawan, T.; Lawpoolsri, S.; White, N.J.; Kaewkungwal, J. Development of temporal modelling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: A case study in endemic districts of Bhutan. *Malar. J.* **2010**, *9*, 251. [[CrossRef](#)] [[PubMed](#)]
 53. Hay, S.I.; Rogers, D.J.; Shanks, G.D.; Snow, R.W. Malaria early warning in Kenya Simon. *Trends Parasitol.* **2001**, *17*, 95–99. [[CrossRef](#)]