Mathematical Modeling of Malaria

- Methods for Simulation of Epidemics

Patrik Johansson

    patrijoh@student.chalmers.se

Jacob Leander

    jaclea@student.chalmers.se

Abstract

In this report we investigate the mathematical model for malaria spread introduced by Chitnis [1]. We state and analyze some important mathematical properties of the system. The reproductive number, R_0, of the system is introduced and shown to be important for the qualitative behavior of the system. We use some basic bifurcation and sensitivity analysis to understand how the model depends on important parameters. The most influential parameter in the model is concluded to be the mosquito biting rate. Simulations of the model are presented by solving a system of differential equations. We also perform a stochastic simulation of the model using the Gillespie method. One conclusion is that the stochastic approach is more realistic and can be used to make a probabilistic statement about disease prevalence. A geographical extension of the model is proposed and we simulate the spread of disease on the African continent.

Contents

1 Introduction

Malaria is a life-threatening disease widely spread in tropical and subtropical regions, including Africa, Asia, Latin America, the middle East and some parts of Europe. The most cases and deaths occur in sub-Saharan Africa. In 2006 there were almost 250 million cases of malaria, causing nearly one million deaths [11]. However, malaria is preventable and curable. By making appropriate models for the spread of malaria one can understand the underlying processes and develop effective prevention strategies.

1.1 Malaria

Malaria is caused by parasites of the species *Plasmodium*. The parasites are transmitted to humans through the bites of infectious female mosquitos (vectors). The malaria parasite enters a human when an infectious mosquito bites a person. After entering a human the parasite transforms through a complicated life-cycle. The parasites multiply in the human liver and bloodstream. Finally, when it has developed into an infectious form, it spreads the disease to a new mosquito that bites the infectious human. After approximately 10 to 15 days the mosquito takes her next blood meal and can infect a new person. After a human gets bitten the symptoms appear in about 9 to 14 days [11]. The most common symptoms are headache, fever and vomiting. If the infected human does not get drugs the infection can progress and become life-threatening.

1.2 Modeling Epidemiology

One of the most basic epidemiological models is the so called SIR model from 1927 [7]. This model is widely used to model the spread of a disease, not only the spread of Malaria. The model describes the different states which a human can be in. The three states are susceptible, infectious and recovered. A human moves through the different states at different rates. Humans enter the system in the susceptible state when born at rate μ_1. A susceptible human enters the infectious state at rate σ_1 when receiving the disease. From the infected state the human can either move to the recovered state at rate σ_2 or the human can leave the system by death at rate d. Humans can also leave the system by immigration and natural death at rate μ_2. The total population is denoted as N. The interaction between the states of this model is illustrated in Figure 1.
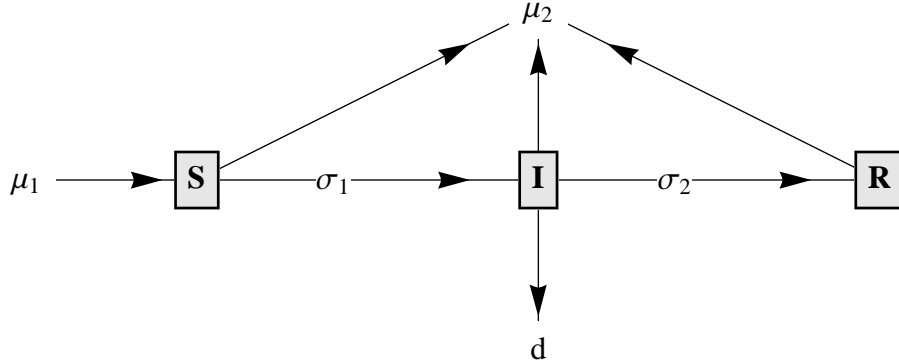
Figure 1: The basic SIR model with rates σ_1, σ_2, μ_1, μ_2 and d

This model can be described by a set of differential equations by using the mass action law. The resulting equation system is:

$$\frac{dS}{dt} = N\mu_1 - S(\mu_2 + \sigma_1)$$
$$\frac{dI}{dt} = S\sigma_1 - I(\sigma_2 + d + \mu_2)$$
$$\frac{dR}{dt} = I\sigma_2 - R\mu_2$$

1.3   Previous work

The mathematical modeling of Malaria began in 1911 with Ross [12], who was awarded with the Nobel price for his work. His model was very simple and has been greatly extended during the years. In 1927 Kermack and McKendrick [7] came up with the improved SIR model of epidemics. In 1957 MacDonald [8] improved the model to a two dimensional model with one variable representing humans and one variable representing mosquitos. An important extension of the model was proposed by Dietz, Molineaux and Thomas [5] who added the inclusion of immunity. Other extensions that have been made is for example environmental dependence and drug resistance. Ngwa and Shu [9] proposed an ordinary differential equation of the model which includes four different states for humans (suceptible-exposed-infectious-recovered) and three different states for mosquitos (suceptible-exposed-infectious). These groups of states interact through different transmission rates. The model by Ngwa and Shu has been improved and studied further by Chitnis [1][2][3].

1.4 Purpose

The purpose of this report is to convey a broad understanding of the current methods of malaria modeling. This will be done by building on the recent work by Chitnis [1][2][3] and Ngwa Shu [9]. We intend to:

- summarize the important features investigated in previous work,

- investigate interesting properties by illustrative examples,

- broaden this understanding with stochastic analysis,

- expand the model to include a geographical dimension and

- make an informed statement about malaria models in general.

2 Malaria Model

In this report we consider the model first proposed Ngwa Shu [9] which was further investigated by Chitnis [1][2][3]. In this section we give a short introduction to the model.

The malaria model that concerns us in this report can be understood as a number of states wherein humans and mosquitos exist depending on their relation to the disease. The different states are explained in Table 1. Parameters describing the behavior of the movement of individuals between these states are introduced in Table 2. The relation between the states and rates of movement between states is illustrated in Figure 2.
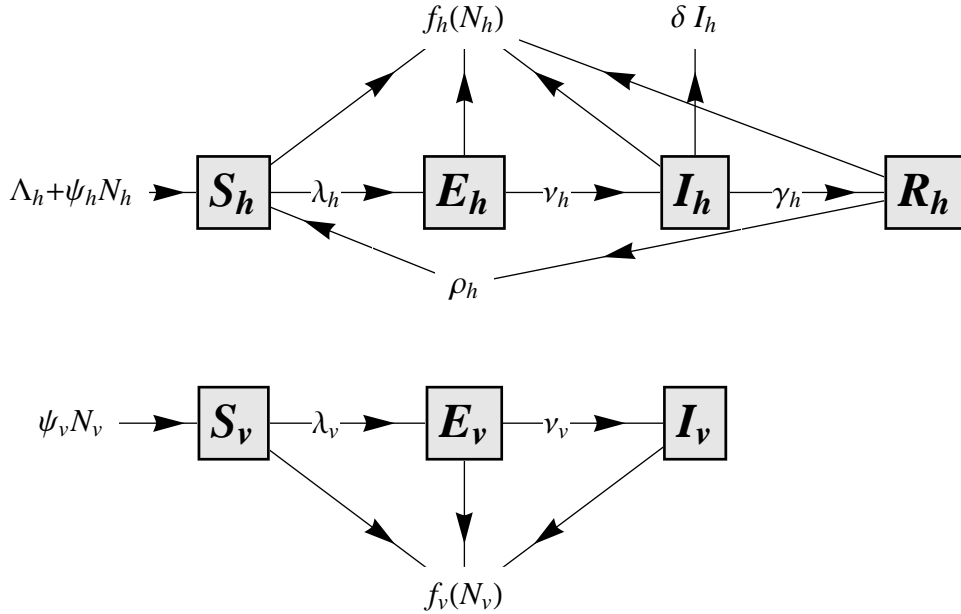
Figure 2: The malaria model with states described in Table 1 and paramters described in Table 2.

Here $f_h(N_h)$ is the per capita density-dependent death and emigration rate for humans and $f_v(N_h)$ is the per capita density-dependent death rate for mosquitos. λ_h and λ_v are the corresponding infection rates. The infection rate for humans is given by the product of the number of mosquito bites that one human can have per time unit, b_h, the probability of transmission from the mosquito to human, β_{hv} and the probability that the mosquito is

Table 1: The different states of the model in Figure 2.

S_h:	Number of susceptible humans at time t
E_h:	Number of exposed humans at time t
I_h:	Number of infectious humans at time t
R_h:	Number of recovered humans at time t
N_h:	Total human population at time t
S_v:	Number of susceptible mosquitoes at time t
E_v:	Number of exposed mosquitoes at time t
I_v:	Number of infectious mosquitoes at time t
N_v:	Total mosquito population at time t

infectious, $\frac{I_v}{N_v}$. In the same fashion the infection rate for mosquitos λ_v is the sum of the force of infection from infectious and recovered humans.

This is written as:

$$f_h(N_h) = \mu_{1h} + \mu_{2h} N_h,$$

$$f_v(N_v) = \mu_{1v} + \mu_{2v} N_v,$$

$$\lambda_h = b_h(N_h, N_v)\beta_{hv}\frac{I_v}{N_v},$$

$$\lambda_v = b_v(N_h, N_v)(\beta_{vh}\frac{I_h}{N_h} + \tilde{\beta}_{vh}\frac{R_h}{N_h}).$$

where b_h and b_v are expressed as:

$$b_h(N_h, N_v) = \frac{\sigma_v N_v \sigma_h}{\sigma_v N_v + \sigma_h N_h}$$

$$b_v(N_h, N_v) = \frac{\sigma_v N_h \sigma_h}{\sigma_v N_v + \sigma_h N_h}.$$

Since ν_h is the rate at which exposed humans move to the infectious state $1/\nu_h$ is the average duration of the latent period for humans. From the infectious state humans move to the recovered state with a rate γ_h. This means that $1/\gamma_h$ is the average duration of the infectious period for humans. In the same fashion we see that $1/\rho_h$ is the average duration of the immune period for humans. By choosing μ_{1h} and μ_{2h} in a way that stabilizes the human population one can investigate the transmission of the disease for an area with a certain population. The same interpretation is valid for corresponding mosquito parameters.

Table 2: The parameters of the model described in Figure 2.

Λ_h:	Immigration rate of humans. $Humans \times Time^{-1}$
ψ_h:	Per capita birth rate of humans. $Time^{-1}$
ψ_v:	Per capita birth rate of mosquitos. $Time^{-1}$
σ_v:	Number of times one mosquito would want to bite humans per unit of time, if humans were freely available. $Time^{-1}$
σ_h:	The maximum number if mosquito bites a human can have per unit of time. $Time^{-1}$
β_{hv}:	Probability of transmission of infection from an infectious mosquito to a susceptible human given that contact between the two occurs. *Dimensionless*
β_{vh}:	Probability of transmission of infection from an infectious human to a susceptible mosquito given that contact between the two occurs. *Dimensionless*
$\tilde{\beta}_{vh}$:	Probability of transmission of infection from a recovered human to a susceptible mosquito given that contact between the two occurs. *Dimensionless*
ν_h:	Per capita rate of progression of humans from the exposed state to the infectious state. $Time^{-1}$
ν_v:	Per capita rate of progression of mosquitos from the exposed state to the infectious state. $Time^{-1}$
γ_h:	Per capita recovery rate of humans from the infectious state to the recovered state. $Time^{-1}$
δ_h:	Per capita disease-induced death rate for humans. $Time^{-1}$
ρ_h:	Per capita rate of loss of immunity for humans. $Time^{-1}$
μ_{1h}:	Density independent part of the death (and emigration) rate for humans. $Time^{-1}$
μ_{2h}:	Density dependent part of the death (and emigration) rate for humans. $Humans^{-1} \times Time^{-1}$
μ_{1v}:	Density independent part of the death rate for mosquitos. $Time^{-1}$
μ_{2v}:	Density dependent part of the death rate for mosquitos. $Mosquitos^{-1} \times Time^{-1}$

Together with the state variables in Table 1 and the parameters in Table 2 the model in Figure 2 satisfies the equation system:

$$\frac{dS_h}{dt} = \Lambda_h + \psi_h N_h + \rho_h R_h - \lambda_h(t)S_H - f_h(N_h)S_H$$

$$\frac{dE_h}{dt} = \lambda_h(t)S_h - \nu_h E_h - f_h(N_h)E_h$$

$$\frac{dI_h}{dt} = \nu_h E_h - \gamma_h I_H - f_h(N_h)I_h - \delta_h I_h$$

$$\frac{dR_h}{dt} = \gamma_h I_h - \rho_h R_h - f_h(N_h)R_h \tag{1}$$

$$\frac{dS_v}{dt} = \psi_v N_v - \lambda_v(t)S_v - f_v(N_v)S_v$$

$$\frac{dE_v}{dt} = \lambda_v(t)S_v - \nu_v E_v - f_v(N_v)E_v$$

$$\frac{dI_v}{dt} = \nu_v E_v - f_v(N_v)I_v$$

where $N_h = S_h + E_h + I_h + R_h$, $N_v = S_v + E_v + I_v$.

3 Analysis of the model

In this section we perform some basic analysis to investigate the model described in section 2. We also state two important theorems which describe the behavior of the system. Basic sensitivity and bifurcation analysis of the system is also presented in this section.

Chitnis [3] shows that by scaling the population sizes in each state by the total population size one gets that;

$$\frac{de_h}{dt} = \frac{\sigma_v \sigma_h N_v \beta_{hv} i_v}{\sigma_v N_v + \sigma_h N_h}(1 - e_h - i_h - r_h) - (\nu_h + \psi_h + \frac{\Lambda_h}{N_h})e_h + \delta_h i_h e_h$$

$$\frac{di_h}{dt} = \nu_h e_h - (\gamma_h + \delta_h + \psi_h + \frac{\Lambda_h}{N_h})i_h + \delta_h i_h^2$$

$$\frac{dr_h}{dt} = \gamma_h i_h - (\rho_h + \psi_h + \frac{\Lambda_h}{N_h})r_h + \delta_h i_h r_h$$

$$\frac{dN_h}{dt} = \Lambda_h + \psi_h N_h - (\mu_{1h} + \mu_{2h} N_h)N_h - \delta_h i_h N_h \qquad (2)$$

$$\frac{de_v}{dt} = \frac{\sigma_v \sigma_h N_h}{\sigma_v N_v + \sigma_h N_h}(\beta_{vh} i_h + \tilde{\beta}_{vh} r_h)(1 - e_v - i_v) - (\nu_v + \psi_v)e_v$$

$$\frac{di_v}{dt} = \nu_v e_v - \psi_v i_v$$

$$\frac{dN_v}{dt} = \psi_v N_v - (\mu_{1v} + \mu_{2v} N_v)N_v$$

where the parameters are described in Table 2 and the state variables in Table 3.

Table 3: The different states of the model scaled by the total population as in (2).

e_h:	Proportion of exposed humans at time t
i_h:	Proportion of infected humans at time t
r_h:	Proportion of recovered humans at time t
N_h:	Total human population at time t
e_v:	Proportion of exposed mosquitoes at time t
i_v:	Proportion of infected mosquitoes at time t
N_v:	Total mosquito population at time t

In [1] it is proved that the model in equation system (2) is epidemiological and mathematically valid in the domain,

$$D = \left\{ \begin{pmatrix} e_h \\ i_h \\ r_h \\ N_h \\ e_v \\ i_v \\ N_v \end{pmatrix} \in \mathbb{R}^7; \begin{array}{l} e_h \geq 0, \\ i_h \geq 0, \\ r_h \geq 0, \\ e_h + i_h + r_h \leq 1, \\ N_h > 0, \\ e_v \geq 0, \\ i_v \geq 0, \\ e_v + i_v \leq 1, \\ N_v > 0 \end{array} \right\}.$$

Using the same notation as Chitnis [3] we denote points in D by $x = (e_h, i_h, r_h, N_h, e_v, i_v, N_v)$. We also define the "diseased" classes as the human or mosquito classes that are either exposed, infectious or recovered.

In [2] a few useful theorems are stated and proved, here they will be stated for reference.

Theorem 1. *Assuming that the initial conditions lie in D, the system of equations for the malaria model (2) has a unique solution that exists and remains in D for all time $t \geq 0$.*

Theorem 2. *The malaria model (2) has exactly one equilibrium point, $x_{dfe} = (0, 0, 0, N_h^*, 0, 0, N_v^*)$, with no disease in the population.*

The positive equilibrium human and mosquito population values, where there is no disease, for (2) are

$$N_h^* = \frac{(\psi_h - \mu_{1h}) + \sqrt{(\psi_h - \mu_{1h})^2 + 4\mu_{2h}\Lambda_h}}{2\mu_{2h}} \quad \text{and } N_v^* = \frac{\psi_v - \mu_{1v}}{\mu_{2v}}.$$

This is obtained by setting the left hand side of (2) to zero, substituting $e_h = i_h = r_h = e_v = i_v = 0$, and then solving for N_h and N_v.

3.1   Reproductive number

A common parameter in epidemiological models is the reproductive number R_0. This number can be understood as the number infections that would result from one infectious individual (human or mosquito) over the infectious period given that all other individuals are susceptible. This number can be defined as

$$R_0 = \sqrt{K_{vh}K_{vh}},$$

where K_{hv} is the number of humans that one mosquito infects through its infectious lifetime if all humans are susceptible and K_{vh} is the number of mosquitos that one human infects through the duration of the infectious period if all mosquitos are susceptible. Mathematically this is written as;

$$K_{hv} = \left(\frac{\nu_v}{\nu_v + \mu_{1v} + \mu_{2v}N_v^*}\right) \cdot \frac{\sigma_v \sigma_h N_h^*}{\sigma_v N_v^* + \sigma_h N_h^*} \cdot \beta_{hv} \cdot \left(\frac{1}{\mu_{1v} + \mu_{2v}N_v^*}\right)$$
$$K_{vh} = \left(\frac{\nu_h}{\nu_h + \mu_{1h} + \mu_{2h}N_h^*}\right) \cdot \frac{\sigma_v N_v^* \sigma_h}{\sigma_v N_v^* + \sigma_h N_h^*} \cdot \left(\frac{1}{\gamma_h + \delta_h + \mu_{1h} + \mu_{2h}N_h^*}\right) \quad (3)$$
$$\cdot \left[\beta_{vh} + \tilde{\beta}_{vh}\left(\frac{\gamma_h}{\rho_h + \mu_{1h} + \mu_{2h}N_h^*}\right)\right]$$

The motivation for this expression is discussed in section 3.2 of [2].

The number R_0 is interesting since it gives us an idea of wether the infection will spread through the population or not. To illustrate this we have the following theorem proved in [2].

Theorem 3. *The disease free equilibrium point, x_{dfe}, is locally asymptotically stable if $R_0 < 1$ and unstable if $R_0 > 1$.*

We can also say a few things about equilibrium points where there exists infection in the population. Chitnis [2] proves that for values of $R_0 > 1$ there exists at least one endemic equilibrium point x_{ee} for the model in (2). That is a steady-state solution where all state variables are positive. In order to find such a point for certain values of the parameters in Table 2 we rely on numerical methods. By setting the left hand side of (2) equal to 0 and then solving the system using *NSolve* in *Mathematica* we easily obtain the possible equilibrium points; $x = (e_h^*, i_h^*, r_h^*, N_h^*, e_v^*, i_v^*, N_v^*)$.

3.2 Bifurcation analysis

Bifurcation analysis is the mathematical study of changes in the solutions when changing the parameters of for example an ODE system. These qualitative changes in the dynamics of the system are called bifurcations. The parameter values where they occur are called bifurcation points. By analyzing the existence of and behavior of the model in such points one can derive much about the systems properties. To understand the following section a brief introduction to bifurcation theory might be appreciated. For space conservation reasons this will not be presented here, instead we recommend reading the basic introduction given by Crawford [4].

By using bifurcation theory Chitnis [2] shows that a endemic equilibrium point exists for all $R_0 > 1$ with a transcritical bifurcation at $R_0 = 1$. He further shows by numerical simulation that for $\delta_h = 0$, and for some small

values of δ_h, there is a supercritical transcritical bifurcation at $R_0 = 1$. This bifurcation is shown to have an exchange of stability between the disease free equilibrium and the endemic equilibrium. Furthermore, there exists a subcritical transcritical bifurcation at $R_0 = 1$ for larger values of δ_h with exchange of stability between the endemic equilibrium and the disease-free equilibrium. There is also a saddle-node bifurcation at $R_0 = R_0^*$ for some $R_0^* < 1$. This means that for some values of $R_0 < 1$ there exists two endemic equilibrium points. One of these is shown to be unstable and the the other to be locally asymptotically stable.

There is no general proof that the endemic equilibrium is unique and stable for $R_0 > 1$. However, Chitnis concludes that numerical results for some parameter sets suggest that this is indeed the case.

The existence of a locally asymptotically stable endemic equilibrium point for $R_0 < 1$, for some parameter values, is interesting from a epidemiological point of view. This means that the threshold for surely eradicating the disease in this case is not $R_0 = 1$ but rather R_0^* which is less than R_0. The saddle-node at $R_0 = R_0^*$ implies that a small change in R_0 can have large impact on malaria prevalence. In an area with malaria a small reduction in R_0 to a value below R_0^* might have great impact on the spread of the disease, since a stable endemic equilibrium vanishes. On the other hand in an area without malaria a small increase in R_0 to a value above R_0^* might cause spread of disease through the population. In this case it is also sufficient to move the system into the basin of attraction of the endemic equilibrium in order for the disease to spread.

3.3   Sensitivity analysis

Sensitivity analysis is a method to measure the relative change in a state variable when a parameter is changed. In [3] Chitnis performs a sensitivity analysis of the model to determine the relative importance of the parameters to disease transmission and prevalence.

We define the normalized forward sensitivity index of a variable (u) to a parameter (p) as the ratio of relative change in the variable to the relative change in the parameter. Mathematically, this is written as:

$$\gamma_p^u = \frac{\partial u}{\partial p} \times \frac{p}{u}$$

By computing the the sensitivity indices of the reproductive number R_0 and the endemic equilibrium point one can conclude which of the parameters are the most important for these variables.

Since we have a explicit formula for R_0 one can easily derive the forward sensitivity index of R_0 for each of the seventeen parameters in Table 2. Chitnis shows that the highest sensitivity indices of the variable R_0 is given by the parameters σ_v with a sensitivity index of 0.76, β_{hv} (0.50) and ψ_v (-0.46). This means that if we increase the parameter value of σ_v by 10 % the value of R_0 increases with 7.6 %.

To investigate the sensitivity indices of the endemic equilibrium point x_{ee} Chitnis relies on numerical results. In the same way as for R_0 the parameter with the highest sensitivity index is σ_v. This result shows that by reducing the mosquito biting rate one can reduce the number of infected humans. This can be done by for example mosquito nets and indoor mosquito sprays.

4   Simulation

In this section we simulate some standard scenarios to explore the behavior of the model in (1). This can be done in an infinite number of ways but we have chosen two main scenarios. One where the transmission rate is relatively low and one where the transmission rate is relatively high. This corresponds in some sense to two real life scenarios were the environmental conditions provide different possibilities for the spread of disease. We also simulate a situation where the existence of a stable endemic equilibrium for values of $R_0 < 1$ is shown. The simulations were computed numerically using *NDsolve* with default settings in *Mathematica*.

4.1   Choice of parameters

The choice of parameters is complicated since most are rather tricky to attain from measurements in real life. In Chitnis [3] a thorough job is done of compiling the interesting factors from reliable sources. These can be stated as two baseline scenarios, one for high and one for low transmission areas, which are shown in Table 4. Many of these parameters are based on studies conducted by various sources. Some values, such as the ones concerning human populations, are based on assumptions about the most common disease situations. That is to simulate spread of the disease in rural areas and small towns.

4.2   Simulation of low transmission area

In areas of low transmission the model in (2) has only one endemic equilibrium point in the domain D. This is shown as previously discussed by numerically solving the system (2) when the left hand side is set to 0 and the parameters are those for low transmission in Table 4. The endemic equilibrium point is

$$x_{ee} = (0.0029, 0.080, 0.10, 578, 0.024, 0.016, 2425).$$

By linearization and calculation of the Jacobian matrix of the system (2) we find the eigenvalues for the system in the point x_{ee}. We conclude that this is a locally asymptotically stable equilibrium since all eigenvalues have strictly negative real part. We can also compute the value of R_0 from (3) to see that for areas of low transmission $R_0 = 1.1$.

If we choose some endemic initial values and simulate the system for a sufficient period of time we see that the solution approaches the endemic equilibrium point, this is shown in Figure 3. Note that this is done for

13

Table 4: The parameter values for the two baseline scenarios for areas of high transmission and low transmission respectively. See Table 2 for definitions and dimensions.

	High	Low
Λ_h	0.033	0.041
ψ_h	$1.1 * 10^{-4}$	$5.5 * 10^{-5}$
ψ_v	0.13	0.13
σ_v	0.50	0.33
σ_h	19	4.3
β_{hv}	0.022	0.022
β_{vh}	0.48	0.24
$\tilde{\beta}_{vh}$	0.048	0.024
ν_h	0.10	0.10
ν_v	0.091	0.083
γ_h	0.0035	0.0035
δ_h	$9.0 * 10^{-5}$	$1.8 * 10^{-5}$
ρ_h	$5.5 * 10^{-4}$	$2.7 * 10^{-3}$
μ_{1h}	$1.6 * 10^{-5}$	$8.8 * 10^{-6}$
μ_{2h}	$3.0 * 10^{-7}$	$2.0 * 10^{-7}$
μ_{1v}	0.033	0.033
μ_{2v}	$2.0 * 10^{-5}$	$4.0 * 10^{-5}$

the original, unscaled system (1). Remember that an unstable disease free equilibrium point exists but since it is unstable it has no effect on the system in this case.

4.3 Simulation of high transmission area

In the same manner as for areas of low transmission, we find that the system (2) only has one endemic equilibrium point in D, for parameters corresponding to a high transmission area. For areas of high transmission $R_0 = 4.5$. The endemic equilibrium point is

$$x_{ee} = (0.0059, 0.16, 0.77, 490, 0.15, 0.11, 4850).$$

By choosing some endemic initial values we can simulate the original system (1) over time. The result is seen in Figure 4. We see that the solution approaches the endemic equilibrium point as expected.

4.4 Simulation of an area on the edge of endemic malaria

In section 3.2 results of bifurcation analysis shows that for some parameter values there exists a stable endemic equilibrium for $R_0 < 1$. One such
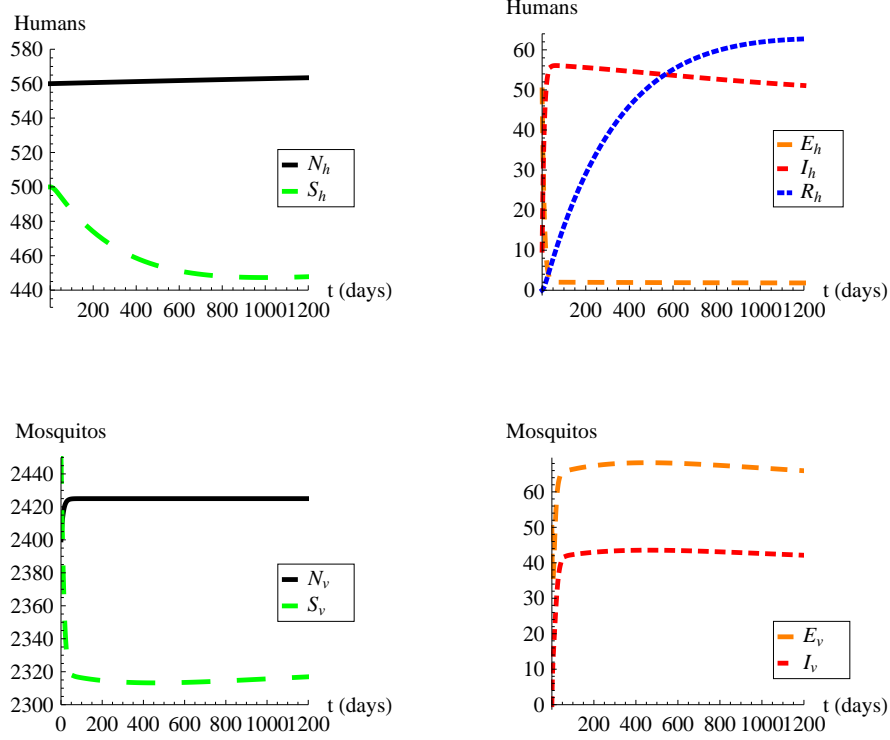
Figure 3: The different state variables over time when solving the system (1) with parameter values for an area of low transmission, as seen in Table 4. The initial conditions are; $N_h = 560, S_h = 500, E_h = 50, I_h = 10, R_h = 0, N_v = 2400, S_v = 2450, E_v = 50$ and $I_v = 0$.

configuration of parameters is shown in Table 5. The value of R_0 for these parameters is 0.9898.

For the values in Table 5 we can numerically find the equilibrium points of the system (2). The interesting equilibrium points are the locally asymptotically stable endemic equilibrium and the disease free equilibrium. Remember that there is also an unstable endemic equilibrium but this is not of any particular interest, instead we concern ourselves with the one that is locally asymptotically stable. The locally asymptotically stable disease free equilibrium is

$$x_{df} = (0, 0, 0, 771, 0, 0, 1129)$$

and the locally asymptotically stable endemic equilibrium is

$$x_{ee} = (0.01622, 0.3297, 0.08279, 301.7, 0.2254, 005635, 1129).$$
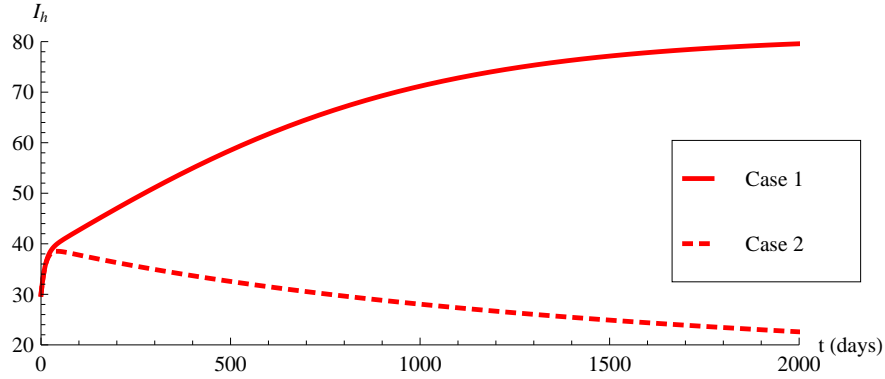
15
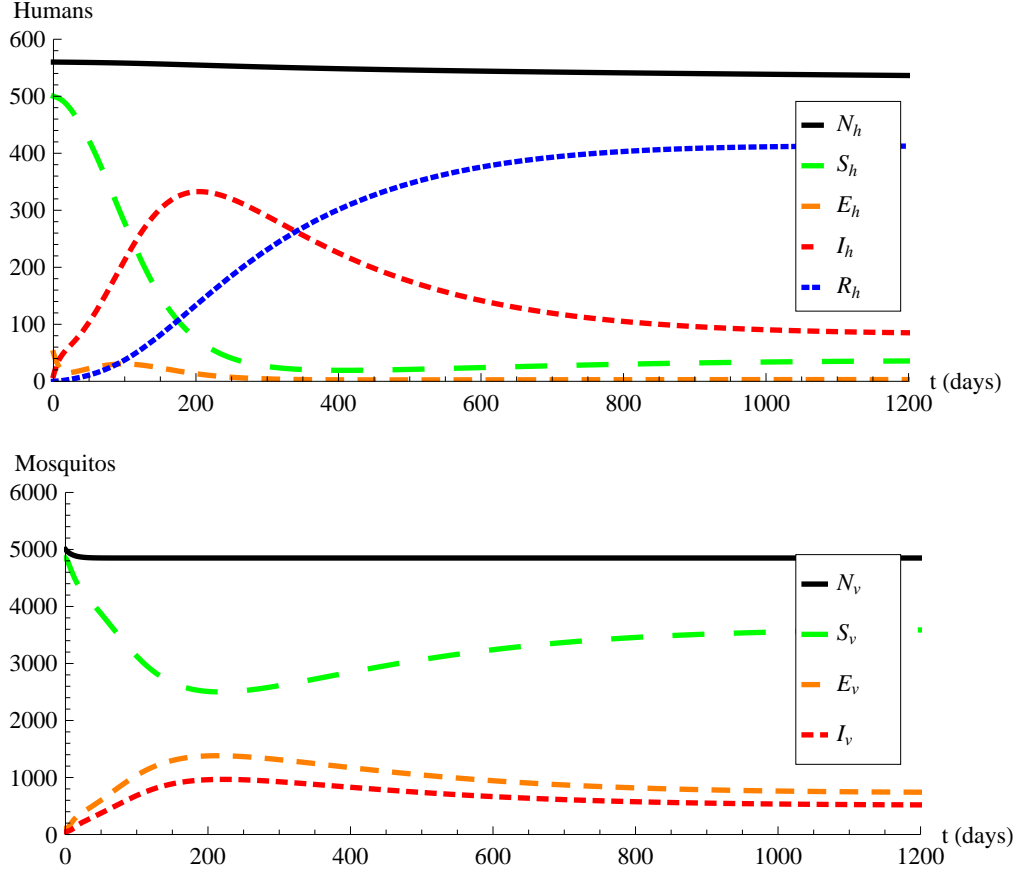
Figure 4: The different state variables over time when solving the system (1) with parameter values for an area of high transmission, as seen in Table 4. The initial conditions are; $N_h = 560, S_h = 500, E_h = 50, I_h = 10, R_h = 0, N_v = 5000, S_v = 4850, E_v = 100$ and $I_v = 50$.

The stability of theses can be shown in the same way as in section 4.2. By choosing two different initial values we show in Figure 5 that one solution approaches the locally asymptotically stable endemic equilibrium and one solution approaches the disease free equilibrium.

Table 5: Parameter values where a stable endemic equilibrium exists for $R_0 < 1$. See Table 2 for definitions and dimensions.

Λ_h	$3.285 * 10^{-2}$
ψ_h	$7.666 * 10^{-5}$
ψ_v	0.4000
σ_v	0.6000
σ_h	18
β_{hv}	$2.000 * 10^{-2}$
β_{vh}	0.8333
$\tilde{\beta}_{vh}$	$8.333 * 10^{-2}$
ν_h	$8.333 * 10^{-2}$
ν_v	0.1000
γ_h	$3.704 * 10^{-3}$
δ_h	$3.454 * 10^{-4}$
ρ_h	$1.460 * 10^{-2}$
μ_{1h}	$4.212 * 10^{-5}$
μ_{2h}	$1.000 * 10^{-7}$
μ_{1v}	0.1429
μ_{2v}	$2.279 * 10^{-4}$

Figure 5: The different state variables over time when solving the system (1) with parameter values from Table 5. The initial conditions for the two cases are as follows. Case 1; $N_h = 740, S_h = 700, E_h = 30, I_h = 10, R_h = 0, N_v = 1150, S_v = 1000, E_v = 100$ and $I_v = 50$. Case 2; $N_h = 440, S_h = 400, E_h = 30, I_h = 10, R_h = 0, N_v = 1150, S_v = 1000, E_v = 100$ and $I_v = 50$.

5 Stochastic approach

Up to this point we have only considered the malaria model deterministically. The deterministic approach has several drawbacks that a stochastic model handles in a more realistic way. Deterministic modeling for instance allows fractional state values, which is not realistic considering the improbability of half a human. This means that the deterministic model "smoothes" the behavior of the system, making it impossible to detect jumps in the state variables as occur in real life when one person gets infected. The stochastic approach remedies this by only considering integer state values. Furthermore the deterministic approach gives the same result every time we run a simulation with the same initial values. This might be mathematically correct but we easily understand why this is not the case in a real epidemic situation. There simply exists many parameters which we can not model entirely realistically, by modeling them deterministically we loose some of the complexity of the system. It might in many cases be more appropriate to assume a stochastic behavior. Consider for example the probability of two people coming in contact with each other. This hardly follows such a strict rule as a the deterministic model assumes but rather a more sporadic behavior, such as in the stochastic model.

In this section we perform a stochastic simulation of the malaria model (1) using the Gillespie algorithm and compare the results with the deterministic approach. The method was implemented in *MATLAB*.

5.1 Gillespie Algorithm

The Gillespie algorithm was introduced by Daniel Gillespie in an article from 1977 [6]. The article describes a way to simulate the behavior of a chemical system by modeling reactions within the system stochastically. This method can be adapted to the malaria model in order to give the system a stochastic behavior. We will here state the algorithm in a general formulation since the particulars are rather tedious and not very informative, for details of why the method is valid we refer to [6].

The right hand side of the system (1) can be understood as a number of probabilities for a certain reaction, namely the increase or decrease of the number of individuals in a certain state. We denote these probabilities $\{p_i\}_{i=1}^n$. In our case $n = 17$ since we can collect 17 terms in the right hand side of the system corresponding to unique events, such as the birth of a human or the infection of a mosquito. We also define $p_s = \sum_{i=1}^N p_i$. With this notation the Gillespie algorithm is executed as follows.

1. Generate two uniformly distributed random numbers, r_1 and r_2, on the interval $[0, 1]$.

2. Calculate the reaction time $\Delta t = \frac{1}{p_s} \ln \frac{1}{r_1}$.

3. Find the smallest m such that $r_2 p_s < \sum_{i=1}^{m} p_i$.

4. Perform the m:th reaction by changing affected state variables.

5. Update the probabilities p_i.

6. Set $p_s = \sum_{i=1}^{N} p_i$.

This is repeated either a set number of steps or until the cumulative sum of the reaction times reaches a specified time limit. Note that before taking the first step we need to set the initial state variable values and calculate the probabilities p_i and p_s.
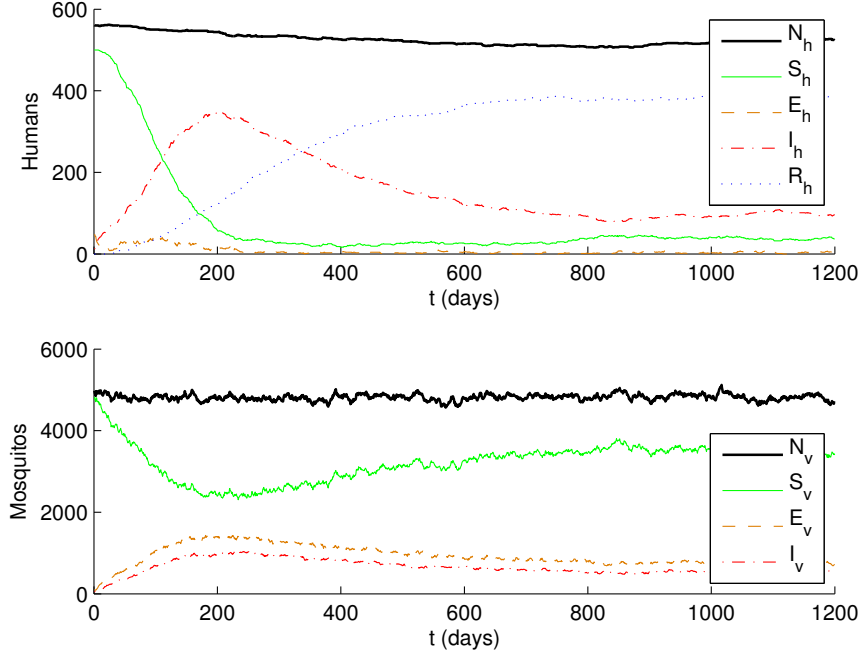
5.2 Stochastic simulation of high transmission area

By running the Gillespie method for the equation system (1) using parameter values for an area of high transmission the result in Figure 6 was obtained.

In Figure 6 we see that the system has an apparent stochastic behavior and that the overall trend of the trajectories follow the same path as the trajectories illustrated in Figure 4. This means that even though fluctuation occurs it in principle has the same behavior as the deterministic model.

Running the simulation once gives us one set of trajectories, i.e. one realization of the stochastic process. This can for example corresponds to one outbreak of the disease in the real world. If we want to study general features of the model it is wise to run the simulation multiple times and look at the average of all simulations. In Figure 7 the average result of 100 realizations of the process is shown. We can see that the stochastic features cancel in some sense, resulting in trajectories that closer resemble the deterministic model when compared to just one realization of the process.

To get a better understanding of the fluctuation around the equilibrium observed in the stochastic simulation in Figure 6 we can approach it another way. By simulating 1000 realizations of the stochastic process and making a histogram of the fraction of infected humans at the end of the process, we illustrate the distribution of values at that time. This is a snapshot of the process when the system has had time to reach the equilibrium points in the deterministic sense. The histogram is illustrated in Figure 8 together with a normal distribution curve fitted to the values, as well as a quantile-quantile plot of the data. By this measure the proportion of infected humans
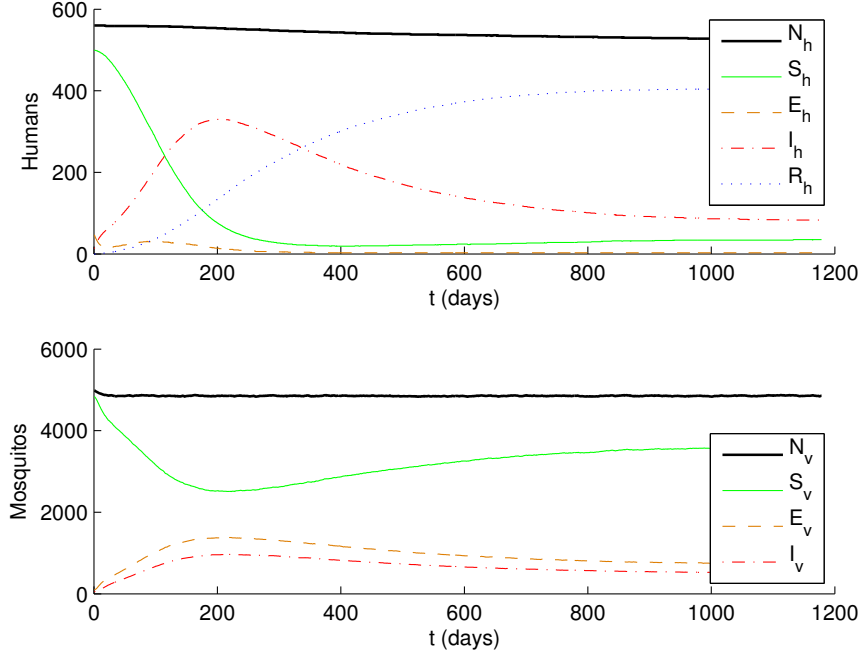
Figure 6: One stochastic trajectory of the malaria model (1) simulated by the Gillespie method with parameter values for an area of high transmission, as seen in Table 4. The initial conditions are; $N_h = 560, S_h = 500, E_h = 50, I_h = 10, R_h = 0, N_v = 5000, S_v = 4850, E_v = 100$ and $I_v = 50$.

appears to be roughly normally distributed around the equilibrium value. The normal distribution appearance is understood as a consequence of the Gillespie method.

5.3 Further investigation of stochastic behavior

In the previous section the effect of the stochastic approach has been limited to random fluctuations around the trajectories in the deterministic model. This is natural since there only exists one stable equilibrium point for the system when looking at parameter values for a high transmission area. Thus there is only so much "damage" the stochastic approach can do compared to the deterministic one.

If we consider a case where there exists multiple stable equilibria the stochastic approach behaves rather differently. Such a case is shown in section 4.4 for parameter values in Table 5. An interesting result is shown if we

Figure 7: The average of 100 stochastic trajectories of the malaria model (1) simulated by the Gillespie method with parameter values for an area of high transmission, as seen in Table 4. The initial conditions are; $N_h = 560, S_h = 500, E_h = 50, I_h = 10, R_h = 0, N_v = 5000, S_v = 4850, E_v = 100$ and $I_v = 50$.

select initial values for the system between the two trajectories in Figure 5. This result is shown in Figure 9. Note that the initial values cause the deterministic trajectory to seemingly fall between the two stable equilibria. (This only looks to be the case however, in fact it tends to the stable disease free equilibrium.) The stochastic trajectory on the other hand sometimes heads for the endemic equilibrium and sometimes for the disease free equilibrium, other times it reaches a value in between. The two stable equilibria can in some sense be said to cause divergence in the stochastic trajectories. This behavior is made apparent in Figure 9. We can also illustrate this behavior in the same way as in Figure 8 to make the difference more clear. This is shown in Figure 10 together with a quantile plot which illuminates the difference in behavior compared to the result in Figure 8.

21
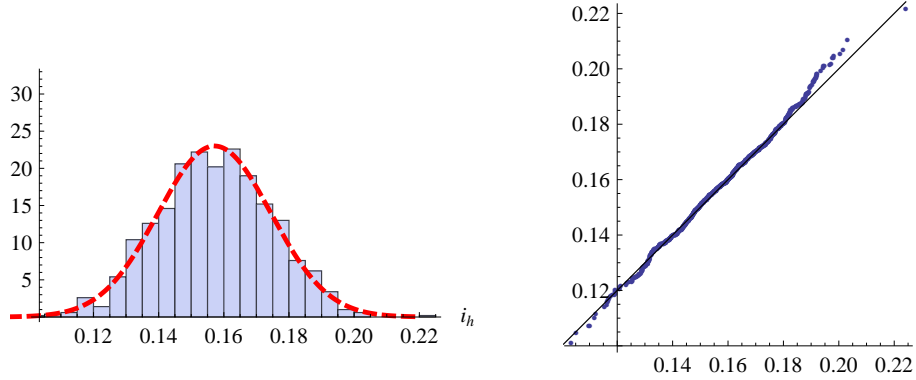
Figure 8: The histogram (LEFT) showing the proportion of infected humans at the end of the simulation time, simulated by the Gillespie method with parameter values for an area of high transmission, as seen in Table 4. The initial conditions are; $N_h = 560, S_h = 500, E_h = 50, I_h = 10, R_h = 0, N_v = 5000, S_v = 4850, E_v = 100$ and $I_v = 50$. Also shown is the quantile-quantile plot comparing the proportion of infected humans to the normal distribution assumption. (RIGHT)
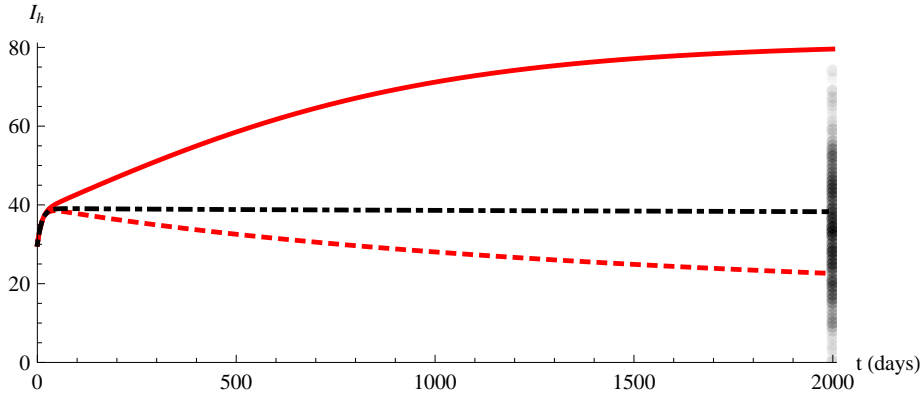
Figure 9: Shown are the two deterministic trajectories introduced in Figure 5, one approaching the endemic equilibrium point and one approaching the disease free equilibrium. Also shown (black, -.-.) is a deterministic trajectory with initial values; $N_h = 640, S_h = 600, E_h = 40, I_h = 10, R_h = 0, N_v = 1150, S_v = 1000, E_v = 100$ and $I_v = 50$ (*) for parameter values in Table 5. At $t = 2000$, 1000 values of I_h simulated by the Gillespie method are shown. The initial values (*) are used. These are plotted with opacity to show the density of the values.

22
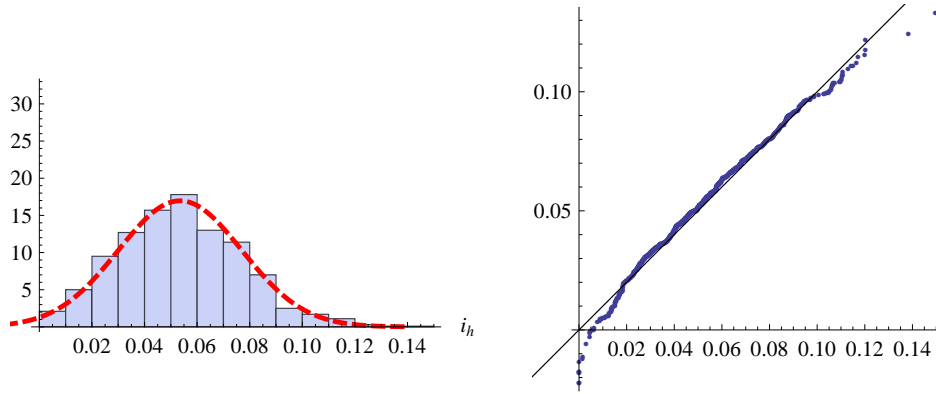
Figure 10: The histogram (LEFT) showing the proportion of infected humans at the end of the simulation time, simulated by the Gillespie method with parameter values from Table 5. The initial conditions are; $N_h = 640, S_h = 600, E_h = 40, I_h = 10, R_h = 0, N_v = 1150, S_v = 1000, E_v = 100$ and $I_v = 50$. The quantile-quantile plot comparing the proportion of infected humans to the normal distribution assumption. (RIGHT)

6 Geographical extension

In this section we introduce a way to extend the model introduced in section 2 to include a geographical dimension. To illustrate the model we implement it for malaria spread in Africa.

6.1 General idea

Consider a set of k regions; $\{R_1, R_2, ..., R_k\}$, that each consist of a system of states as in Figure 2. We define $S_h^{(i)}$ to be the number of susceptible humans in region i. We also define $\Lambda_{ij}^{(S)}$ to be a rate of which susceptible humans emigrate from region i to region j. In the same way we define the number of exposed, infected and recovered individuals and their emigration rates. Note that we only consider emigration of humans. This allows us to express k equation systems, linked by emigration and immigration, that together constitute the model with a geographical dimension.

6.2 Three regions

As an example we choose a system of three regions (e.g. countries, cities, villages). Susceptible humans move between these regions at rates $\Lambda_{12}^{(S)}$, $\Lambda_{13}^{(S)}$, $\Lambda_{21}^{(S)}$, $\Lambda_{23}^{(S)}$, $\Lambda_{31}^{(S)}$ and $\Lambda_{32}^{(S)}$. The geographical dimension of this system is illustrated in Figure 11.
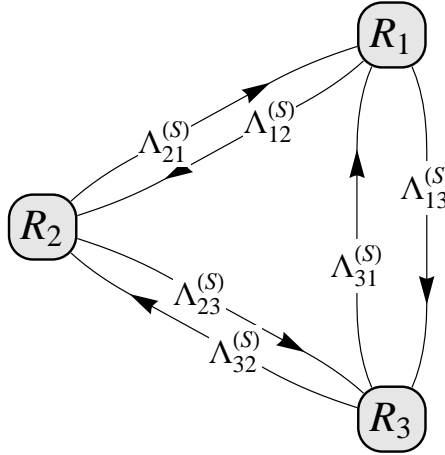
Figure 11: Three regions with rates of emigration of susceptible humans. The same figure is valid for all other human states.

In the same way as before we use the mass action law to set up a system of differential equations. If we just consider the change of susceptible individuals (the same principle applies to other states) caused by emigration of humans we get the equation system:

$$\frac{dS_h^{(1)}}{dt} = \Lambda_{21}^{(S)} S_h^{(2)} + \Lambda_{31}^{(S)} S_h^{(3)} - (\Lambda_{12}^{(S)} + \Lambda_{13}^{(S)}) S_h^1$$

$$\frac{dS_h^{(2)}}{dt} = \Lambda_{12}^{(S)} S_h^{(1)} + \Lambda_{32}^{(S)} S_h^{(3)} - (\Lambda_{21}^{(S)} + \Lambda_{23}^{(S)}) S_h^2 \tag{4}$$

$$\frac{dS_h^{(3)}}{dt} = \Lambda_{13}^{(S)} S_h^{(1)} + \Lambda_{23}^{(S)} S_h^{(2)} - (\Lambda_{31}^{(S)} + \Lambda_{32}^{(S)}) S_h^{(3)}$$

To find the behavior of the entire system, i.e. the systems for all three regions, we combine three equation systems on the form of (1) and solve them simultaneously. The only difference being that we substitute the parameters concerning immigration and emigration of humans with the new definitions introduced in this section.

6.3 Arbitrarily many regions

The notions in the previous section can without any problems be extended to arbitrarily many geographical regions. Together with the other parameters in the malaria model from Table 2 this leads to the system (5). Note that this is the equation system for one region R_i. The important feature of this system is the sums, that for each state express the emigration from an area to all other areas and the immigration from all other areas to that area.

$$\frac{dS_h^{(i)}}{dt} = \sum_{i \neq j} \Lambda_{ji}^{(S)} S_h^{(j)} - S_h^{(j)} \sum_{i \neq j} \Lambda_{ij}^{(S)} + \psi_h N_h^{(i)} + \rho_h R_h^{(i)} - \lambda_h(t) S_H^{(i)} - f_h(N_h^{(i)}) S_H^{(i)}$$

$$\frac{dE_h^{(i)}}{dt} = \sum_{i \neq j} \Lambda_{ji}^{(E)} E_h^{(j)} - E_h^{(j)} \sum_{i \neq j} \Lambda_{ij}^{(E)} + \lambda_h(t) S_h^{(i)} - \nu_h E_h^{(i)} - f_h(N_h^{(i)}) E_h^{(i)}$$

$$\frac{dI_h^{(i)}}{dt} = \sum_{i \neq j} \Lambda_{ji}^{(I)} I_h^{(j)} - I_h^{(j)} \sum_{i \neq j} \Lambda_{ij}^{(I)} + \nu_h E_h^{(i)} - \gamma_h I_H^{(i)} - f_h(N_h^{(i)}) I_h^{(i)} - \delta_h I_h^{(i)}$$

$$\frac{dR_h^{(i)}}{dt} = \sum_{i \neq j} \Lambda_{ji}^{(R)} R_h^{(j)} - R_h^{(j)} \sum_{i \neq j} \Lambda_{ij}^{(R)} + \gamma_h I_h^{(i)} - \rho_h R_h^{(i)} - f_h(N_h^{(i)}) R_h^{(i)} \qquad (5)$$

$$\frac{dS_v^{(i)}}{dt} = \psi_v N_v - \lambda_v(t) S_v^{(i)} - f_v(N_v^{(i)}) S_v^{(i)}$$

$$\frac{dE_v^{(i)}}{dt} = \lambda_v(t) S_v^{(i)} - \nu_v E_v^{(i)} - f_v(N_v^{(i)}) E_v^{(i)}$$

$$\frac{dI_v^{(i)}}{dt} = \nu_v E_v^{(i)} - f_v(N_v^{(i)}) I_v^{(i)}$$

where $N_h^{(i)} = S_h^{(i)} + E_h^{(i)} + I_h^{(i)} + R_h^{(i)}, N_v^{(i)} = S_v^{(i)} + E_v^{(i)} + I_v^{(i)}$, and

$$f_h(N_h^{(i)}) = \mu_{1h} + \mu_{2h} N_h^{(i)},$$
$$f_v(N_v^{(i)}) = \mu_{1v} + \mu_{2v} N_v^{(i)},$$
$$\lambda_h = b_h(N_h^{(i)}, N_v^{(i)}) \beta_{hv} \frac{I_v^{(i)}}{N_v^{(i)}},$$
$$\lambda_v = b_v(N_h^{(i)}, N_v^{(i)})(\beta_{vh} \frac{I_h^{(i)}}{N_h^{(i)}} + \tilde{\beta}_{vh} \frac{R_h^{(i)}}{N_h^{(i)}}).$$

## 6.4   Simulation of geographical system

To present a proper simulation of the geographical malaria system (5) a few parameters are needed. Beyond the original parameters in Table 2 for each region, one also needs the rates at which humans move between the regions. These numbers are hard to obtain so in order to present a simulated scenario we make some assumptions. A reasonable assumption is that the emigration rate $\Lambda_{ij}^{(S)}$ can be written as

$$\Lambda_{ij}^{(S)} = \frac{N_h^{(j)}}{\sum_{k_i} N_h^{(j)}} \Lambda_i^{(S)} \qquad (6)$$

26

Here $k_i = \{$All regions j; j is a neighbor to region $i\}$. $\Lambda_i^{(S)}$ is the emigration rate of susceptible humans from country i. In the same way we define the parameters $\Lambda_{ij}^{(E)}$, $\Lambda_{ij}^{(I)}$, $\Lambda_{ij}^{(R)}$. This expresses that the emigration from one country, which is known, is divided as immigration to the neighboring countries proportional to the relations of the population sizes.

Using the above assumption we have simulated the malaria spread for the African continent when all countries except Angola start as being disease free. Angolas initial value is a 10% infectious population. The population sizes and emigration rates of each county were obtained from a statistical database [10]. The result of this simulation is shown in Figure 12. In this figure a very natural behavior is observed. The disease starts of where it is expected and gradually spreads to neighboring countries.
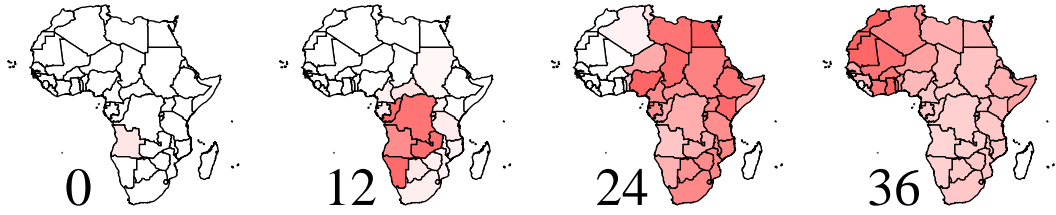
Figure 12: The spread of malaria in Africa. The initial values are disease free in all countries except Angola where 10% of the population is infected. Shading represents the proportion of infected humans in the total population at a certain time. The time scale is months.

7 Discussion and conclusion

In this report we have investigated the malaria model introduced by Ngwa and Shu [9] that was extended by Chitnis [1][2][3]. We gave a brief introduction to the model and the basic properties that describe its behavior. A proof that the system has a unique solution remaining in the epidemiological valid domain when initial conditions are properly defined, was introduced. We also stated a proof that there always exists a disease free equilibrium point with positive population sizes. That is to say that the disease can be eradicated without exterminating the entire population.

A brief discussion about the reproductive number R_0 was conducted and its influence on the behavior of the model was analyzed. We concluded that the disease free equilibrium point of the system is stable if $R_0 < 1$ and unstable if $R_0 > 1$. This is an important property of the malaria model. At least one endemic equilibrium point exists for the model when $R_0 > 1$. Finding such a point, and other equilibrium points, was concluded to best be done by numerical methods.

Bifurcation analysis was used to show an interesting property of the model. Namely that for some parameter values there exists a stable endemic equilibrium point when $R_0 < 1$. This is important when controlling the spread of the malaria since there now exists a number $R_0^* < R_0$ that is the threshold for ensuring that no disease persists in the population.

By introducing some results from sensitivity analysis performed by Chitnis [3] we conclude that the most influential parameter on the persistence of disease is the mosquito biting rate. This is important to consider when choosing a strategy for disease reduction.

To illustrate the behavior of the system a series of simulations were presented in section 4. These illuminate the properties of the model and gives an idea of how aggressive the disease is in areas with different parameter values. This is understood as a consequence of the qualitative change in the model depending on the parameters.

In order to get a more complete picture of epidemiological modeling we also simulated the model using a stochastic approach. We conclude that in many ways a stochastic simulation is better suited to give a correct understanding of the disease. The regular approach has some drawbacks, for example the deterministic property and the fact that it smoothes the behavior of the model. Stochastic modeling on the other hand allows us to make a probabilistic statement about the progression of the disease. Therefor our opinion is that a stochastic approach should be considered when constructing a proper model over disease spread. At the same time the regular approach has some important features, such as the possibility of mathematical anal-

ysis, that also ought to be considered. We therefore conclude that both approaches should be considered for a better understanding.

In section 6 we proposed an extension to the model in which a geographical dimension is added. This extension is important if one wants to investigate the spread of disease in a scenario where the population is not homogeneously mixed. In the basic model an assumption is made that any person is equally likely to infect any other person. This might be true for large and dense population but it is not a good approximation in reality. The geographical extension allows us to split the total population into smaller groups where the population is more homogeneously mixed, such as splitting a larger region into small towns. This gives a more accurate model. To illustrate the capability of the geographical extension a simulation of disease spread over the entire African continent was introduced and results were presented.

The geographical extension proposed in this report is in some sense a crude one and is based on simple assumptions about the emigration of humans. In a more realistic model it might be feasible to consider smaller regions, almost approaching population density expressions. Consider for example a fine mesh where the population states are described in each node. This improvement is something further work might consider.

Many other improvements and extensions of the model are possible. One example is the addition of seasonal dependence. This could easily be done by expressing the dependent parameters as periodic function of time. For example the number of born mosquitos will vary significantly during the year, with different periods having different rainfall, temperature and humidity. Another extension is to add more states to the model. For example a more accurate model might divide the human states by age and gender.

Our final conclusion is that malaria models today are much more advanced than only a brief time ago. The applications of mathematical malaria modeling are many. Analyzing a proper model can for example lead to added understanding of the disease. Modeling the spread of malaria can also be a helpful tool in choosing a strategy to curb the spread of the disease. To fully exploit the benefits of mathematical modeling a broad approach is needed since each approach has its own drawbacks and advantages.

References

[1] N. Chitnis. *Using Mathematical Models in Controlling the Spread of Malaria*. PhD thesis, University of Arizona, 2005.

[2] N. Chitnis, J.M Cushing, and M. Hyman. Bifurcation analysis of a mathematical model for malaria transmission. *Siam J. Appl. Math.*, 67(1):24–45, 2006.

[3] N. Chitnis, J.M. Cushing, and M. Hyman. Determining important parameters in the spread of malaria through the sensitivity analysis of a mathematical model. *Bulletin of Mathematical Biology*, 70:1272–1296, 2008.

[4] J.D. Crawford. Introduction to bifurcation theory. *Reviews of Modern Physics*, 64(4):991–1037, 1991.

[5] K. Dietz, L. Molineaux, and A. Thomas. A malaria model tested in the african savannah. *Bulletin World Health Organ*, 50:347–357, 1990.

[6] D. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22:403–434, 1976.

[7] W.O. Kermack and A.G McKendric. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115:700–721, 1927.

[8] G. MacDonald. *The Epidemiology and Control of Malaria*. Oxford University Press, London, 1957.

[9] A. Ngwa and W.S Shu. A mathematical model for endemic malaria with variable human and mosquito populations. *Math. Comput. Modelling*, 32:747–763, 2000.

[10] Global Healt Observatory. `http://apps.who.int/ghodata/`. (2010-05-10).

[11] RollbackMalaria. What is malaria? `http://www.rollbackmalaria.org/cmc_upload/0/000/015/372/RBMInfosheet_1.pdf`. (2010-05-10).

[12] R. Ross. *The Prevention of Malaria*. John Murray, London, 1911.