A predictive parameter of the mosquito population in malaria prediction models

Mentors:
Elizabeth Itskovich
Arnon Houri Yafin (Zzapp CEO)

Authors:
Almog Jakov Maatuf
Itay Rafee
Neta Roth

Department of Computer Science, University of Ariel

July 2022

Abstract

Malaria is a fatal disease that can kill within 24 hours. It is mostly common in Africa, and it is one of the main causes of deaths in the world, normally caused by Plasmodium mosquito.
To deal with the Malaria, mathematicians developed different models to try and predict malaria spread affected by number of factors. Changing the effect of these factors can help medical teams adjust the relevant treatment. The simplest example to illustrate this is so: imagine you leave in a small dense village surrounded with swamps; spraying may be the smartest treatment method since the swamps are nearby the village. This can be learned by the model as it simulates the mosquito's population (by adding a parameter that represents the spraying and decreases population by a fixed factor).
The goal of the project is to implement standard prediction models for malaria, also to evaluate the mosquito population measure more accurately, since it is fixed in such models. We collaborated with the ZzappMalaria company, the developer of a mobile-app and managerial dashboard that helps eliminate malaria and has won IBM Watson AI XPRIZE Competition in 2022.

Malaria

Malaria is a tropical infectious disease caused by Plasmodium parasites that affects humans and other animals. There are at least five different species of the parasite, each of which may show differential sensitivity to drugs, and which have different drug-resistance profiles. It is one of the leading causes of death in the world, especially among children. The disease is widespread in the tropical and subtropical regions that exist in a broad band around the equator. This includes much of sub-Saharan Africa, Asia, and Latin America. In 2020 there were 241 million cases of malaria worldwide resulting in an estimated 627,000 deaths. Approximately 95% of the cases and deaths occurred in sub-Saharan Africa. Rates of disease have decreased from 2010 to 2014 but increased from 2015 to 2020. Malaria is commonly associated with poverty and has a significant negative effect on economic development. In Africa, it is estimated to result in losses of US$12 billion a year due to increased healthcare costs, lost ability to work, and adverse effects on tourism.

Clinical symptoms of malaria

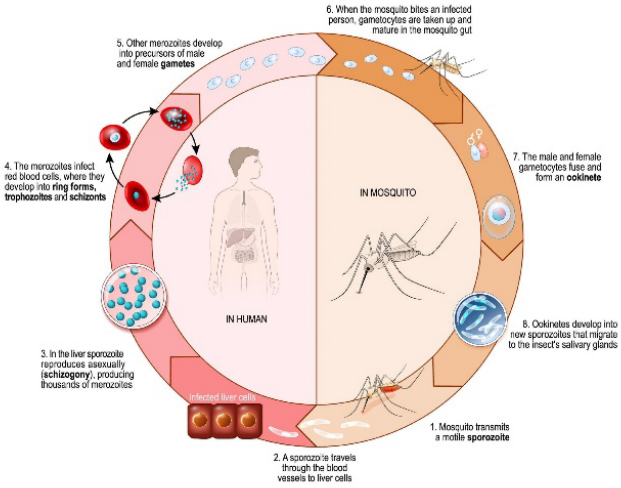
The symptoms of malaria include high fever, tremor, tiredness, joint pain, severe headache, vomiting that may contain blood, excessive sweating, chills, diarrhea, abdominal pain, general malaise, feeling of stinging in the skin, and lack of fabric sugar. Complications of malaria include coma, and in about 20% of untreated malaria cases, death. Young children are particularly vulnerable. Symptoms usually begin ten to fifteen days after being bitten by an infected mosquito. If not properly treated, people may have recurrences of the disease months later. In those who have recently survived an infection, reinfection usually causes milder symptoms. This partial resistance disappears over months to years if the person has no continuing exposure to malaria.

Transmission of malaria

The major mode of transmission of malaria is through the bite of female Anopheles mosquitoes, Five species of Plasmodium can infect and be spread by humans. Most deaths are caused by P. falciparum, whereas P. vivax, P. ovale, and P. malariae generally cause a milder form of malaria. The species P. knowlesi rarely causes disease in humans.
Other modes of transmission include blood product transfusion and contaminated needles. Placental transmission from mother to fetus is known to cause congenital malaria.

The human-mosquito infection cycle

Female Anopheles mosquitoes carry Plasmodium spores in their salivary glands and when it bites a healthy human, it injects Plasmodium into the human's cells which then migrate to the liver and multiply in the liver cells. Afterward, Plasmodium is released into the blood circulation.
In the other direction if a healthy female Anopheles mosquito bites an infected human the parasite arrives with the blood to its stomach, multiplies there, and passes to its salivary glands.



Diagnosis

Malaria is typically diagnosed by the microscopic examination of blood using blood films, or with antigen-based rapid diagnostic tests. Methods that use the polymerase chain reaction to detect the parasite's DNA have been developed, but are not widely used in areas where malaria is common due to their cost and complexity.

Malaria re-occurrence

Malaria re-occurs by three mechanisms:
1. Recrudescence when not treated well the first time.
2. Re-eruption of liver hypnozoites (dormant forms of Plasmodium from when the human got bitten for the first time).
3. Reinfection following mosquito bite after complete resolution of the previous infection.

Fight malaria

The risk of disease can be reduced by preventing mosquito bites through the use of mosquito nets and insect repellents or with mosquito-control measures such as spraying insecticides and draining standing water. Several medications are available to prevent malaria for travellers in areas where the disease is common. Occasional doses of the combination medication sulfadoxine/pyrimethamine are recommended in infants and after the first trimester of pregnancy in areas with high rates of malaria. As of 2020, there is one vaccine which has been shown to reduce the risk of malaria by about 40% in children in AfricaEfforts to develop more effective vaccines are ongoing. The recommended treatment for malaria is a combination of antimalarial medications that includes artemisinin. The second medication may be either mefloquine, lumefantrine, or sulfadoxine/pyrimethamine. Quinine, along with doxycycline, may be used if artemisinin is not available. It is recommended that in areas where the disease is common, malaria is confirmed if possible before treatment is started due to concerns of increasing drug resistance. Resistance among the parasites has developed to several antimalarial medications; for example, chloroquine-resistant P. falciparum has spread to most malarial areas, and resistance to artemisinin has become a problem in some parts of Southeast Asia.

In general, we can divide the common methods to fight malaria to:
1. Drugs
2. Bite prevention
3. Pesticide
4. Vaccine
5. Prediction

Factors that affect the spread of Malaria

1. Climatic factors: As the temperature decreases, the number of days required for a parasite to complete its development inside the mosquito increases. Another important related factor is that a mosquito larva develops more quickly at higher temperatures.
2. Habitats and water collections - The main malaria transmitter can be found in a variety of water collections, mainly closer to human habitations, stagnant water collections in borrow pits, ponds, micro-dams, and pools in small rivers, and streams created immediately after the rainy season.
3. Immunity - Immune people often have a better chance of tolerating the effects of malaria and surviving the disease than non-immune people. Children and pregnant women are the most at risk because they have weak immunity to malaria infection. Immunity to malaria develops slowly after several infections and children need at least five years to develop their immunity. Pregnant women have less immunity to malaria due to their pregnancy.

4. Insecticide resistance in vectors - After repeated application of insect-killing chemicals, the mosquitoes develop insecticide resistance, which means that they are no longer killed by the insecticides.
5. Drug resistance in malaria parasites - Like the insecticide resistance mentioned above, after repeated use of anti-malaria medicine, the parasite can develop resistance to that drug or similar medicines.
6. Anopheles mosquito's growth habits – Females Anopheles can survive up to a month. Adult females lay 50 to 300 eggs per oviposition and can lay between 800 and 1000 eggs during their life. It takes 6-10 days for an egg to become an adult depending on the climate.
7. Migration – migration from an area with malaria to a free-malaria area can also be a crucial factor since people from free-malaria areas are usually more vulnerable.
8. Water development projects - big and small water-related development projects, such as irrigation channels, dams, and ponds, can increase the incidence of malaria in villages that are located near such projects since they create new habitats for the Anopheles.
Not all factors can be measured, but to predict the mosquito's growth correctly, we'll consider some of these factors.

Mathematical models

To know the most effective way to treat malaria in a specific area mathematicians try to predict a set of data equations that each of them gives a piece of the puzzle, such as the ratio of female Anopheles mosquitoes, the ratio of female infected Anopheles mosquitoes, the ratio of infected humans, number of swamps, etc...
A mathematical model is a set of equations – each of them provides new information that will ultimately create a graph of the growth.
The most important and useful thing about mathematical models is that they can help us decide the way to treat malaria – each equation is being fed with initial values, and changing those values will give us a different scenario.
For example, if the model has an equation that computes the number of vivid mosquitoes per day and uses a parameter representing the capacity of a swamp, changing that parameter can mimic the effect of pesticides.

Model Basics - SEIR

In epidemiological compartment models of infectious diseases, transmission of infectious agents in the host population is the fundamental process to be described. When a pathogen appears in a host community, it partitions individuals in the community into categories depending on parasite density inside them and the type of infection. These categories or compartments are represented by standard notation of *S*-*E*-*I*-*R*.

In a simple form they are as follows:

- S- the first group consists of the fraction of host population that is Susceptible to infection.
- E- the Exposed class - the fraction of population whose individuals are affected by the pathogen, but not capable of passing on the infection to others during a latent period.
- I- The Infectious individuals- who give rise to more infected individuals through interaction with the Susceptible.
- R- those individuals who recover from the infection.

There may be variations in the compartment structure depending on the type of disease. For example, the *I* class of individuals may not recover at all and die; *R* can consist of individuals, who recover with temporary or permanent immunity, thereby further subdividing the epidemiological compartments. Using these notations, eight classes of compartmental models are possible - *SI, SIS, SEI, SEIS, SIR, SIRS, SEIR* and *SEIRS*.
For example, in an *SEIRS* model, a fraction of the susceptible (*S*) population gets exposed (*E*) to infection, a part of which then becomes infectious (*I*). Some from the *I* class recover from the disease, and become part of the *R* class with temporary immunity. When immunity is lost, they become susceptible to pathogen attack again, and enter the *S* class. The *Plasmodium* parasite requires both human and mosquito for its life cycle to complete, and the infection is transferred between susceptible human individuals through the bite of infected mosquitoes, which acquire infection through a blood meal from infected humans. In malaria models, therefore, these compartments have been applied to both human (host) and vector (mosquito).

Overview of models

We started of by a simple SEIR model to predict COVID and moved on to Malaria prediction.

COVID prediction

Predicts the rate of infected humans using the SEIR model we mentioned in the last section. The list of parameters:

1. α is the inverse of the incubation period
2. β is the average contact rate in the population
3. γ is the inverse of the mean infectious period

Here are the equations:

1. $\dot{S} = -\beta SI$
2. $\dot{E} = \beta SI - \alpha E$
3. $\dot{I} = \alpha E - \gamma I$
4. $\dot{R} = \gamma I$
5. $N = S + E + I + R$

To indicate the effectiveness of the vaccine on the population we added the parameter δ, which created a new model:

1. $\dot{S} = -\beta SI$
2. $\dot{E} = \beta SI - \alpha E(1 - \delta)$
3. $\dot{I} = \alpha E(1 - \delta) - \gamma I$
4. $\dot{R} = \gamma I$
5. $N = S + E + I + R$

Malaria prediction

Three main models to predict malaria were considered in this paper, in all this is the list of parameters:

1. a- Man biting rate
2. b- Proportion of bites that produce infection in human
3. c- Proportion of bites by which one susceptible mosquito becomes infected
4. m- Ratio of number of female mosquitoes to that of humans
5. r- Average recovery rate of human
6. μ_1- Per capita rate of human mortality
7. μ_2- Per capita rate of mosquito mortality
8. τ_m- Latent period of mosquito
9. τ_h- Latent period of human

Later on, we will replace the m parameter by a set of equations.

The Ross Model

Predicts the rate of infected humans using 2 primary parameters – newly infected humans and recovering. The model also predicts the rate of infected mosquitoes using 2 primary parameters - newly infected mosquitoes and mosquito mortality.

1. $\frac{dI_h}{dt} = a \cdot b \cdot m \cdot I_m(1 - I_h) - r \cdot I_h$
2. $\frac{dI_m}{dt} = a \cdot c \cdot I_h(1 - I_m) - \mu_2 \cdot I_m$

MacDonald model

Extends the Ross model to predicts the proportion of infected humans and the proportion of infected mosquitoes using the mosquito incubation period parameter. It also adds an equation to predict the percentage of exposed mosquitoes.

1. $\frac{dI_h}{dt} = a \cdot b \cdot m \cdot I_m[t](1 - I_h[t]) - r \cdot I_h[t]$
2. $\frac{dE_m}{dt} = a \cdot c \cdot I_h[t] \cdot (1 - E_m[t] - I_m[t]) - a \cdot c \cdot I_h[t - \tau_m] \cdot (1 - E_m[t - \tau_m] - I_m[t - \tau_m]) \cdot e^{-\mu_2 \tau_m} - \mu_2 \cdot E_m[t]$
3. $\frac{dI_m}{dt} = a \cdot c \cdot I_h[t - \tau_m] \cdot (1 - E_m[t - \tau_m] - I_m[t - \tau_m]) \cdot e^{-\mu_2 \tau_m} - \mu_2 \cdot I_m[t]$

Anderson-May model

Extends the McDonald model. It adds an equation to predict the percentage of exposed people given the human incubation period parameter.

1. $\frac{dI_h}{dt} = a \cdot b \cdot m \cdot I_m[t](1 - I_h[t]) - r \cdot I_h[t]$
2. $\frac{dE_m}{dt} = a \cdot c \cdot I_h[t] \cdot (1 - E_m[t] - I_m[t]) - a \cdot c \cdot I_h[t - \tau_m] \cdot (1 - E_m[t - \tau_m] - I_m[t - \tau_m]) \cdot e^{-\mu_2 \tau_m} - \mu_2 \cdot E_m[t]$
3. $\frac{dI_m}{dt} = a \cdot c \cdot I_h[t - \tau_m] \cdot (1 - E_m[t - \tau_m] - I_m[t - \tau_m]) \cdot e^{-\mu_2 \tau_m} - \mu_2 \cdot I_m[t]$

Our algorithm

To predict the changing number of mosquitoes in the swamps (parameter $m \cdot (number\ of\ human\ population)$ in the previous models) a few indices will be taken under consideration.

Initial values

1. d_m- Probability of a mosquito death on a given day.
2. η_m- Period of time taken for an egg to become an adult.
(This parameter depends on the climate and other factors as mentioned in the introduction, for the sake of simplicity it will be considered fixed).
3. η_p- Pregnancy duration.
4. n_m- Number of female mosquitoes born per pregnancy.
5. s- Number of swamps.
6. c_S- Capacity of mosquitoes per swamp.
8. $h(t)$- Number of humans on a given day.
* t- current time.

To predict the requested ratio, it is needed to predict several helpful indicators such as:
1. c_{tot}- Capacity of all mosquitoes at all swamps – This parameter will be calculated in this manner- $c_{tot} = s \cdot c_s$.
2. k- Number of female mosquitoes born on a given day per female mosquito.
3. $g(t)$- Number of female mosquitoes being laid in a swamp on a given day before dilution.
4. tot- Number of female mosquitoes in the swamps before dilution.
5. p- Number of female mosquitoes that needed to be diluted in a given day – This parameter will be calculated in this manner- $p = tot - c_{tot}$.
6. $f(t)$- Number of female mosquitoes being laid in a swamp on a given day after dilution.
8. $m(t)$- Number of mosquitoes on a given day – what the model ultimately tries to predict.

* Dilution is the process in which excess mosquitoes die due to a lack of space in swamps (The parameter p calculates the exact number).

Computation of $k(t)$:

$$k = \begin{cases} k(0) = n_m \\ k(1) = k(2) \dots = k(\eta_p - 1) = 0 \end{cases}$$

As soon as a female mosquito finishes her pregnancy period (after n_m days) she lays n_m new female mosquitoes. Therefore, on days that divide without a reminder the k function will output n_m , on any other day, it will output 0. (The modulo part will be handled later in the g function).

Computation of $g(t)$:

$$g(i) = \begin{cases} k\left(t \bmod (\eta_p)\right) \cdot \left(m_{t-1} \cdot (1 - d_m)\right), & i = t \\ f(i-1), & i \in \{t - \eta_m, \dots t-1\} \\ 0, & i < t - \eta_m \end{cases}$$

For days before any egg had developed to be an adult yet (when $i < t - \eta_m$) occurred g will output 0 because no new eggs laid in any swamp.
For days past days (when $i \in \{t - \eta_m, .. t-1\}$) g will output the last updated value of eggs in all swamps ($f(i)$- after dilution!).
For the current day ($i = t$) g is checking if t can be divided by η_p without residue, g is distinguishing between 2 cases:
1. $t\ (mod\ \eta_m) \equiv 0$: In this case the meaning is that t is a day on which a pregnancy is ended, therefore k will output n_m "new offspring". This value then is being multiply by the number of living mosquitoes (that survived the whole pregnancy) - $\left(m_{t-1} \cdot (1 - d_m)\right)$ this value takes under consideration the survival probability $(1 - d_m)$ and the current number of female mosquitoes.
2. $t(mod\ \eta_m) \not\equiv 0$: In this case the meaning is that all female mosquitoes are in their pregnancy period and no new offspring are being born. Note that in this case k will output 0.
* This function hides inside it the assumption that each female mosquito gets pregnant as soon as it lays the last pregnancy's eggs in the swamp.
Now that the model computed the exact number of eggs in all swamps at any given day it is time to dilute the swamps (the function that responsible for this part is f as mentioned before).

Computation of tot:

$$tot = g(t - \eta_m) + g(t - \eta_m + 1) + \cdots g(t)$$

Sum of all eggs laid that are still in their development period (from egg to adult mosquito).
* Any day before $t - \eta_m$ should not be considered since eggs laid on that day has already grown to adults.

Computation of $f(t)$:

$$f(i) = \begin{cases} g(i), & if\ \ i \in \{t - \eta_m, \dots t\}\ and\ p \le 0 \\ g(i) - \dfrac{g(i)}{tot} \cdot p & if\ \ i \in \{t - \eta_m, \dots t\}\ and\ p > 0 \\ 0, & else \end{cases}$$

For days before the current pregnancy period (when $i \notin \{t - \eta_m, \dots t\}$ there are no new eggs on the swamps therefore f will output 0.

For days in the current pregnancy period ($i \in \{t-\eta_m, \dots t\}$) f is distinguishing between 2 cases:
1. $p \leq 0$: Dilution is not needed because the swamps did not reach their maximum capacity, therefore f will output the value without dilution (the value that g computed for that day). Note that in this case $f = g$.
2. $p > 0$: Dilution is needed. To perform an equal dilution f calculates the percentage of dilution it needs to subtract from each day. The expression $\frac{g(i)}{tot}$ represent the precent of female mosquitoes that day i contributed to the total sum, multiplying it by p (total number of mosquitoes to dilution) will give f the exact number of mosquitoes it needs to subtract from the number g output for day i (before dilution).

Computation of $m(t)$:

$$m_t = \begin{cases} m_{t-1} \cdot (1-d_m) + f(t-\eta_m) & t > 1 \\ f(t-\eta_m) & else \end{cases}$$

$m_{t-1} \cdot (1-d_m)$ represents the number of live mosquitoes from yesterday (this part is not relevant when $t \leq 1$) added $f(t-\eta_m)$ that represents the number of mosquitoes that survived dilution and finished their development period (became adults).
Lastly:
To compute ratio of female mosquitoes to human in day t do: $\frac{m(t)}{h(t)}$.

Tests and results

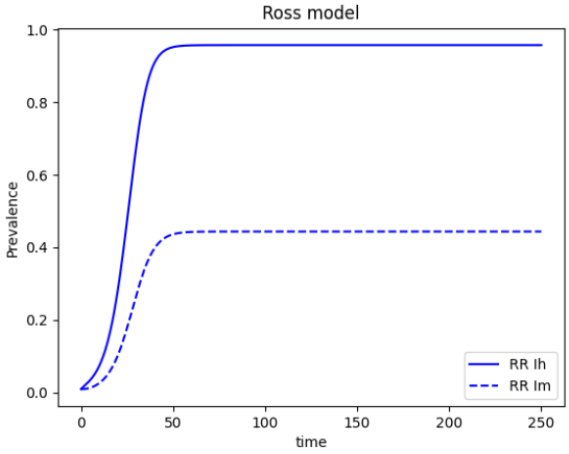
In this section we will present the result of basic models implementation (of both COVID and Malria) and the difference when our m parameter is replaced in the Malaria models.

Corona Model (SEIR)

The result of adding the δ parameter was an increase in the percentage of exposed people (E) and a decrease in the percentage of infected people (I), the larger δ is, the smaller the number of infected people.

Ross model

For that model we got the following output:

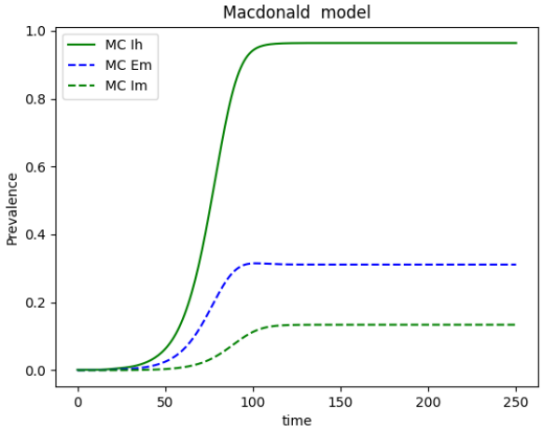


[Parameters used are: $a = 0.2\ day^{-1}, b = 0.5, c = 0.5, m = 20, r = 0.01\ day^{-1}, \mu_2 = 0.12\ day^{-1}$]

In the picture above note the prediction of the coefficients I_h and I_m according to the Ross Model equations.

Macdonald model

For that model we got the following output:



[Parameters used are: $a = 0.2\ day^{-1}, b = 0.5, c = 0.5, m = 20, r = 0.01\ day^{-1}, \mu_2 = 0.12\ day^{-1}, \tau_m = 10\ days$]

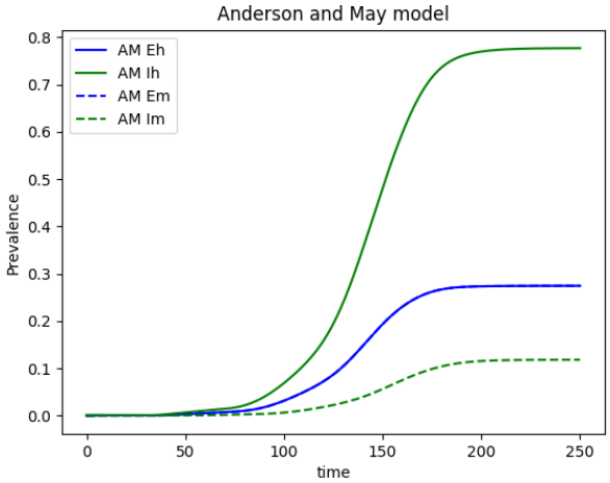
In the picture above note the prediction of the coefficients I_h, I_m and E_m according to the Macdonald Model equations.

The difference from Ross Model:

- In Macdonald Model, E_m is added.
- The growth rate of I_h is more delayed. (Incubation period of the mosquito will only delay the growth)
- I_m is significantly smaller. (During incubation more mosquitoes die)

Anderson and May Model

For that model we got the following output:



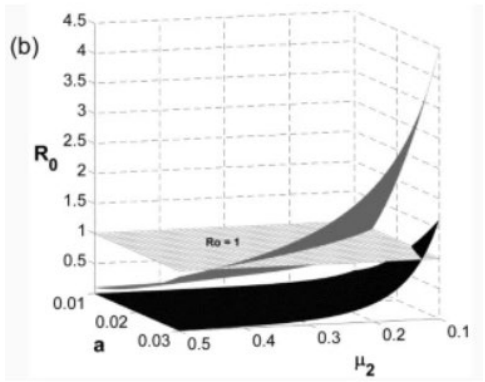
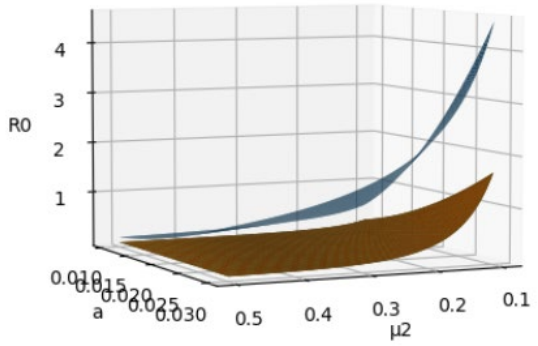
[Parameters used are: $a = 0.2\ day^{-1}, b = 0.5, c = 0.5, m = 20, r = 0.01\ day^{-1}, \mu_1 = 0.017 year^{-1}, \mu_2 = 0.12\ day^{-1}, \tau_m = 10\ days, \tau_h = 21 days$]

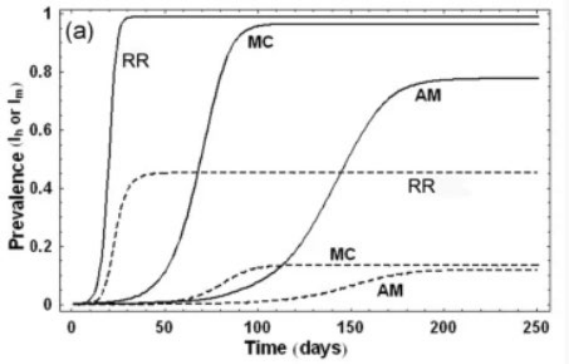
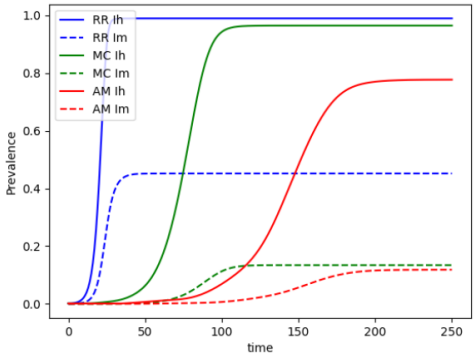
In the picture above note the prediction of the coefficients I_h, I_m, E_m and E_h according to the Anderson and May Model equations.

The difference from Macdonald Model:

- In Anderson and May Model, E_h is added.
- The growth rate of I_m is more delayed. (Incubation period in humans will only delay growth)
- I_m is significantly smaller. (During incubation period, more people succumb)

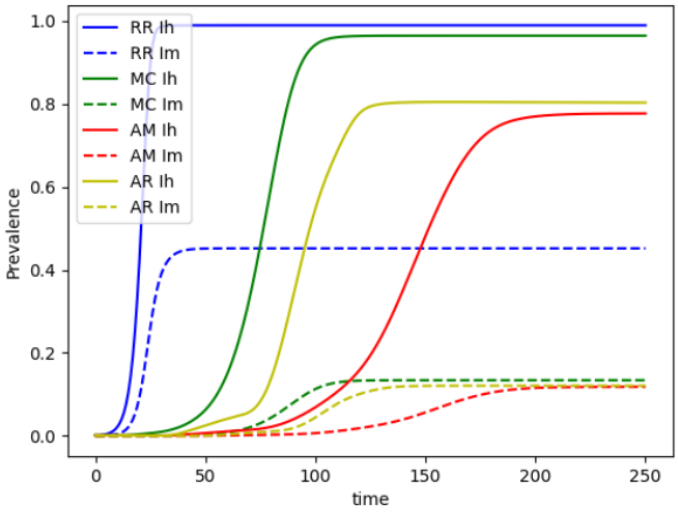
Comparison with the original article

R coefficient of the article	*The R coefficient is calculated using our implementation*
	

Implementation of the article	*Our Implementation*
	

Results of models with our “m” parameter

In the new model the parameter m is recalculated each time based on additional parameters.
The new model (In yellow) in relation to the other models:



m parameters used are: $\eta_m = 12, \eta_p = 4, n_m = 3, s = 100, c_s = 300, mos = 5000, hum = 100$

The difference from Anderson and May Model:

- The growth rate of I_m and I_h is faster. (The parameter m grows faster)

Future work

In the future we would like to develop more precise equations of the parameters we used to predict m and consider several factors that can affect them, for example:

1. Parameter η_m- Period taken for an egg to become an adult as it is now fixed. This parameter can be affected by several factors such as __.

2. Function f in terms of time between pregnancies (Our f function consider consecutive pregnancies).

3. Parameter n_m- Number of female mosquitoes born per pregnancy. This parameter can be affected by mother’s health and diet, climate, etc...

4. Removal of the k function- this function considers egg laying as something done cyclically until the mosquito dies when this is not the case in reality.

Another possible improvement that can be made is to add new parameters that refers to these situations:

1. variable capacity of the swamps (In case of Pesticide)

2. climate parameter (these cannot be controlled but affect a lot so might need to be taken under consideration).

3. Drugs and Immunity factors (such as number of people vaccinated, Immunity resistance).

We will also want to feed our equations into various models besides the ones we had shown here.

In conclusion

In this paper we have presented the importance of the malaria problem, factors that affect it and methods of treatment. As stated above, we focused on methods for predicting mosquito and human population parameters. The innovation presented here is changing the parameter m as a function of several factors such as duration of pregnancy, capacity in swamps and more.

References

[1] A. Pandey, D. Shingadia, Treatment and prevention of malaria in children
[2] T. Yang, S.Ottillie, MalDA, Accelerating Malaria Drug Discovery

[3] Communicable Diseases Module: 6. Factors that Affect Malaria Transmission
[4] NATURE TODAY, Life cycle Anopheles mosquito

[5] Target Malaria, Anopheles gambiae s.l.: morphology, life-cycle, ecology