Inferences Under Varying Stochasticity in a Simple Malaria Model

Sam Dumas, Miles Abssy, Dylan Bradizza, Ofek Gross

Abstract*:* The Ross-MacDonald model is a deterministic seven-parameter susceptible-infectious model for malaria host-vector dynamics. While researchers have extended the base model to incorporate additional parameters, there are fewer reports of stochastic modifications to the base Ross-MacDonald model. Here, we apply the Gillespie Algorithm to investigate the degree to which stochasticity influences our ability to model malaria dynamics using the Ross-MacDonald model. Estimations of *b* fall within one parameter in 97% of simulations and 33% of the time for *c*. This research provides insight into novel models for modeling malaria dynamics, demonstrating alternative means of introducing or increasing stochasticity should be explored.

Introduction: Malaria is a deadly mosquito-borne disease caused by *Plasmodium spp* [1,2]. Malaria is primarily transmitted by mosquitoes (*Anopheles*), whose bite allows the entry of *Plasmodium* into the bloodstream. Once in a human host, infected individuals exhibit non-specific symptoms, such as diarrhea, fever, vomiting, and pulmonary complications [1]. To better understand vector-human relationships, researchers have developed mathematical models, like the Ross-Macdonald model. The Ross-Macdonald model is a simple SI model used to predict disease dynamics over time [2]. The model was first proposed by Sir Ronald Ross in the 1910s, and later extended by George MacDonald in the late 1950s [3,4]. Since then, others have made contributions to the model, aiming to improve its accuracy [3]. Frequently, this comes in the form of considering fixed variables to be dynamic, such as population size, or by introducing new variables such as quarantine [3]. While stochastic incorporation has been loosely explored, the relationship between population size and stochasticity in the fixed Ross-MacDonald Model remains to be characterized [3]. Here, we describe an application of the Gillespie algorithm to incorporate stochasticity into a deterministic model. Through this, we investigate how incorporating stochasticity through changing population sizes influences our ability to model malaria dynamics. The Ross-Macdonald model is a deterministic model consisting of two differential equations, one describes the rate of change in the number of infected humans, whereas the other describes the rate of change in the number of infected mosquitoes (see Figure 1) [1,2,3]. In this project, we specifically focused on estimating the probability of transmission from a mosquito vector to a human host (variable *b*), and the probability of transmission from a human host to a vector (variable *c*).

Although accurate in highly scaled models, the lack of stochasticity in deterministic models can result in dynamics that do not accurately represent slower processes, especially in small-scale finite populations. To account for the deterministic shortcomings of the Ross Macdonald function, the team implemented the Gillespie algorithm. Gillespie assumes that populations of finite individuals are distributed over a finite number of states, where changes in such states occur via reactions from other states. Implementing the Gillespie in terms of the Ross MacDonald, our team identified the following two states and four reactions occurring: the infected human population [illegible], the infected vector population transmission from infected vector to susceptible human [$abI_v/H(H-I_h)$], transmission from infected human to susceptible vector $acI_h/H(V-I_v)$, recovery of infected human γI_h, and mortality rate of vector μI_v. The state change matrix represents the stochastic model, where the top row represents the infected human state and the bottom is the infected vector state (see Figure 2).

Methods: Generally, the approach to this study entailed simulated data using the Gillespie Algorithm and thereafter estimating which parameters would most likely give rise to the data observed from the Gillespie Algorithm.

Data Description: The parameters selected reflect a human-vector population where the

biting rate is low, both probabilities of transmission are moderate, the recovery rate of humans is low, and the mosquito death rate is also low (See Figure 3). The ratio of vectors to mosquitoes was assumed to be 1 throughout this study, and populations of size 10, 50, 100, 500, and 1000 were considered for simulations. Using the Gillespie Algorithm, 100 simulations were conducted for each population size on an interval of 1000 days. The resulting data was used to estimate which parameters were most likely to give rise to the data. Notably, the parameters *b* and *c* were the parameters that were estimated within this study. The probability of observing the data was compared to a normal distribution with a mean equal to the number of infected humans observed that day and a standard deviation of 1. As mosquito population data can be difficult to reliably determine in real life, estimations will be centered on human infection data. Cleaned data will consist of iterations where at least one parameter was estimated, and mixed, where both parameters were estimated.

Data Analysis: While no formal hypothesis testing will be conducted within this study, the data will be analyzed by considering how many estimations fall within one parameter of the actual value. For *b*, this includes the parameters 0.168 and 0.247, and for *c*, the parameters 0.2424 and 0.356. This approach is motivated by the fact that the actual parameter is not contained in the set of values that can be estimated, therefore, the best estimations would fall within one of the two plausible parameters mentioned above for *b* and *c*, respectively.

Results: Across all the estimations (n = 1 500), the mean distance was 0.015 units with a standard deviation of 0.024 units. Estimations of *b* were one parameter value away from the actual value in 97% of simulations, whereas this occurred only 33% of the time when estimating *c*. When grouping by population sizes, this yields 92%, 98%, 94%, 100%, and 100% for estimations of *b* within one parameter (smallest to largest populations, respectively), and 66%, 24%, 38%, 42%, and 50% for estimations of *c* within one parameter (smallest to largest populations, respectively). Furthermore, grouping by population sizes where *c* and *b* were both estimated, yields 84%, 96%, 88%, 100%, and 100% for estimations of *b* within one parameter, and 8%, 16%, 16%, 16%, and 28% for estimations of *c* within one parameter.

Discussion: There are two clear trends within the data; estimations of *b* are more accurate than estimations of *c*, and estimations are more accurate for larger populations. The former is likely a result of the range of parameters used for estimations, whereas the latter is a result of stochastic principles relating to population sizes. While both *b* and *c* had an equal number of parameters that could be estimated (11 possible options), the parameters used for *c* were more condensed. Empirically, *b* has been bounded between 0.01 and 0.8, whereas *c* has been bounded between 0.072 and 0.64. To mitigate this bias, it would be preferable to select possible parameters that increment equally, rather than having an equal number of options. This may also be caused by the formulation of the Ross-Macdonald model, as *c* does not explicitly appear in the differential equation for human infections, however, *b* does. It would be worthwhile to make inferences using vector data to see if the converse occurs (accurate estimations of *c* but not *b*), to properly consider this possibility. Furthermore, it is well understood that stochastic processes pose a greater effect on smaller populations than larger ones [6]. Therefore, it is unsurprising that greater population sizes yielded more accurate estimations, however, this study intended to observe how much stochasticity is required for estimations to deviate from expected values. Based on the data, larger populations had an overwhelming accuracy for estimations of *b*. To account for such, there is a need for a greater magnitude of stochasticity, as under current conditions large population data can be accurately explained.

Figures

Figure 1: The Ross-Macdonald model equations and their parameters.

Ross-MacDonald Model

$$\frac{dI_h}{dt} = ab\frac{I_v}{H}(H - I_h) - \gamma I_h \quad (1)$$

$$\frac{dI_v}{dt} = ac\frac{I_h}{H}(V - I_v) - \mu I_v \quad (2)$$

Figure 1: The variables in the Ross-Macdonald model are: Human Recovery Rate, Vector Death Rate (μ), Infectious Humans (IH), Human Population (H), Infectious Vectors (IV), Vector Population (V), Mosquito Biting Rate (a), Probability of Transmission from Vector to Human (b), and Probability of Transmission from Human to Vector (c)

Figure 2: The state change matrix of the stochastic Gillespie and Ross-Macdonald model

$$\begin{bmatrix} +1 & 0 & -1 & 0 \\ 0 & +1 & 0 & -1 \end{bmatrix}$$

Figure 3: Parameters used within the Ross-MacDonald model to simulate the data

a	b	c	γ	μ
0.1	0.2	0.3	0.01	0.05

Figure 4: Distribution of estimated *b* parameters (CLEANED)

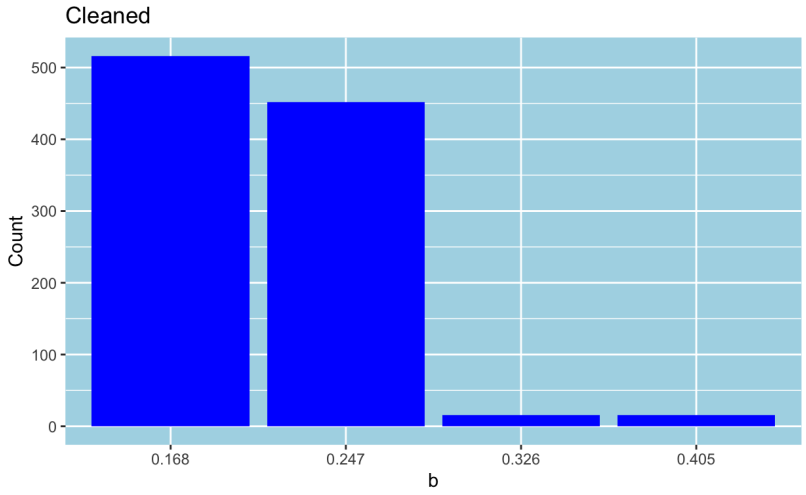


Figure 5: Distribution of estimated *b* parameters (MIXED)

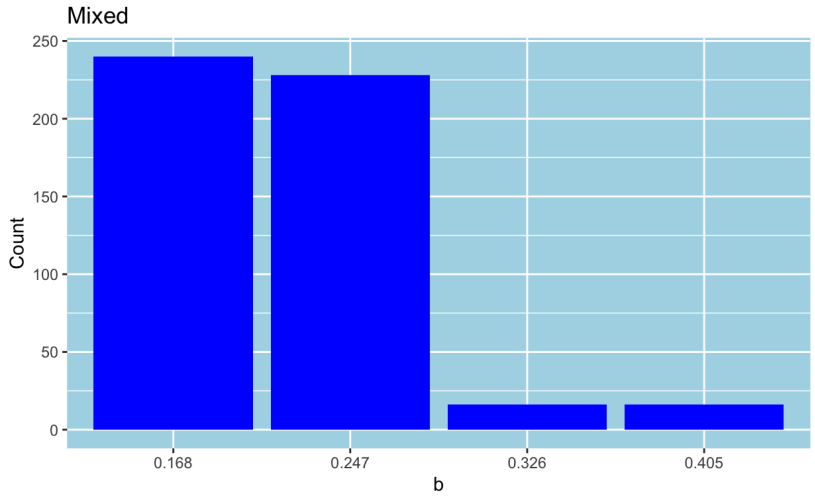


Figure 6: Distribution of estimated *c* parameters (CLEANED)

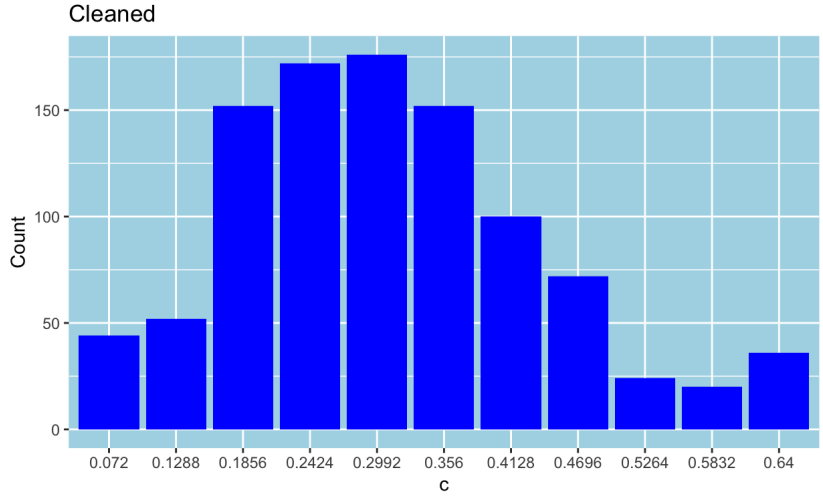


Figure 7: Distribution of estimated *c* parameters (MIXED)

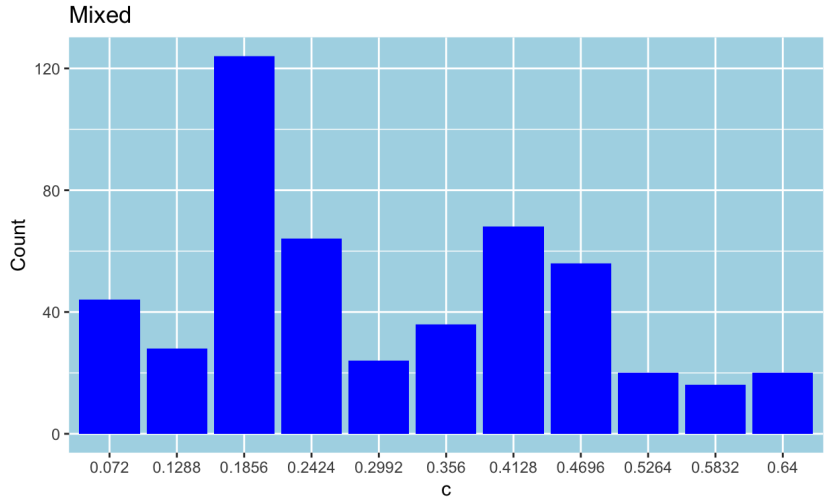


Figure 8: Observed parameter deviations by population size (CLEANED)

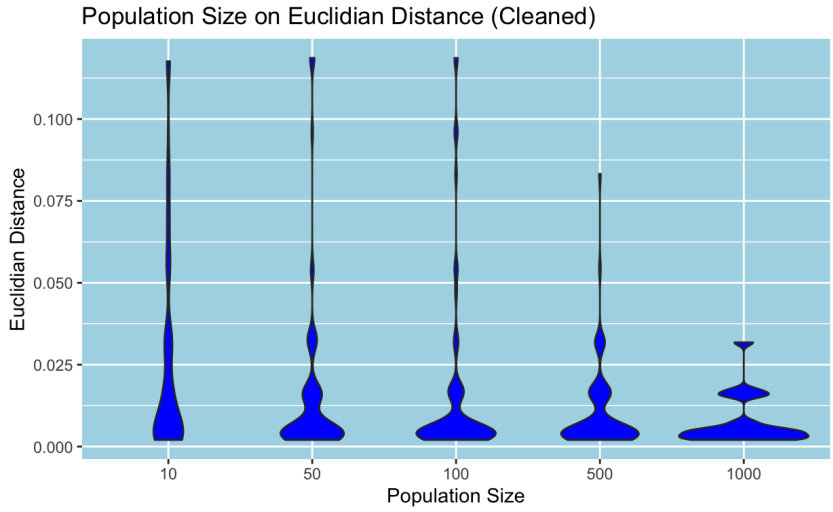
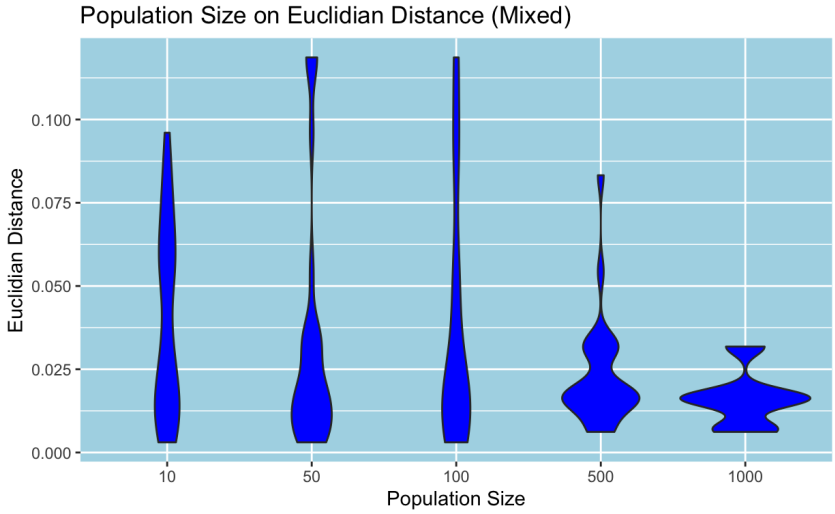


Figure 9: Observed parameter deviations by population size (MIXED)



Works Cited

1. M. A. Phillips, J. N. Burrows, C. Manyando, R. H. van Huijsduijnen, W. C. Van Voorhis, and T. N. Wells, “Malaria,” Nature Reviews Disease Primers, vol. 3, no. 1, 2017, doi: 10.1038/nrdp.2017.50.
2. J. C. Tavares, Malaria. Morgan & Claypool, 2013.
3. S. Mandal, R. R. Sarkar, and S. Sinha, “Mathematical models of malaria - a review,” Malaria Journal, vol. 10, no. 1, 2011, doi: 10.1186/1475-2875-10-202.
4. D. L. Smith, K. E. Battle, S. I. Hay, C. M. Barker, T. W. Scott, and F. E. McKenzie, “Ross, MacDonald, and a theory for the dynamics and control of mosquito-transmitted pathogens,” PLoS Pathogens, vol. 8, no. 4, 2012, doi: 10.1371/journal.ppat.1002588.
5. K. Magori and J. M. Drake, “The population dynamics of vector-borne diseases,” Nature Education Knowledge, vol. 4, Jan. 2013.
6. D. H. Reed. “Conservation Biology: Evolution in Action”, 2008.