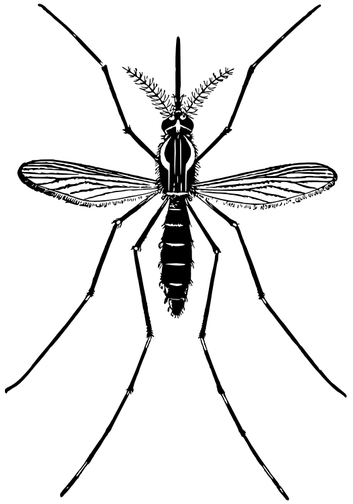
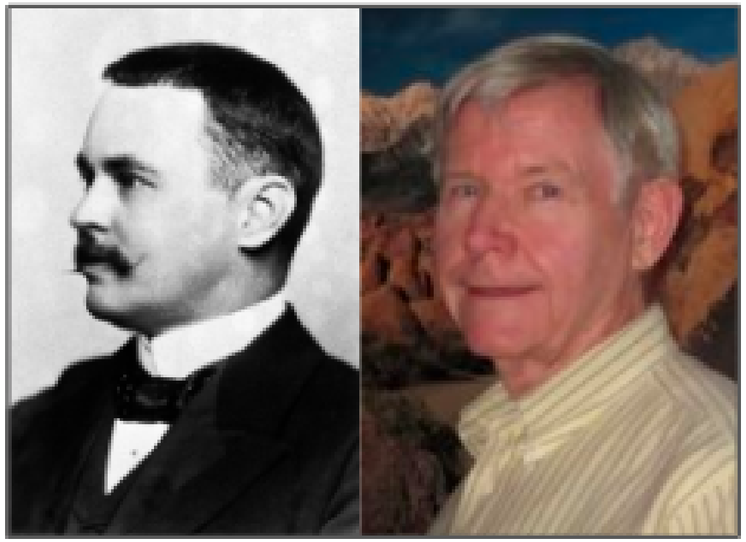
Inferences Under Varying Stochasticity in a Simple Malaria Model

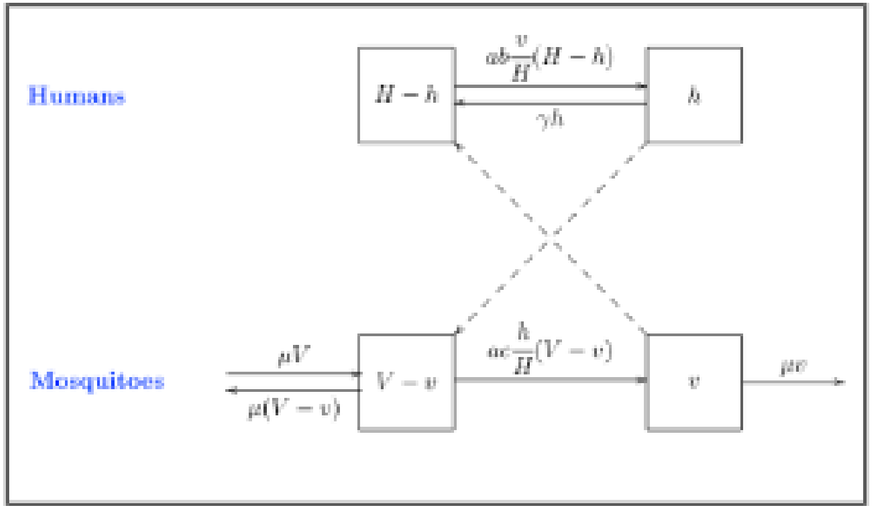
EEB313 - Final Project Presentation, Dec 5, 2023

Sam Dumas, Miles Abssy, Dylan Bradizza, Ofek Gross

Presentation Overview







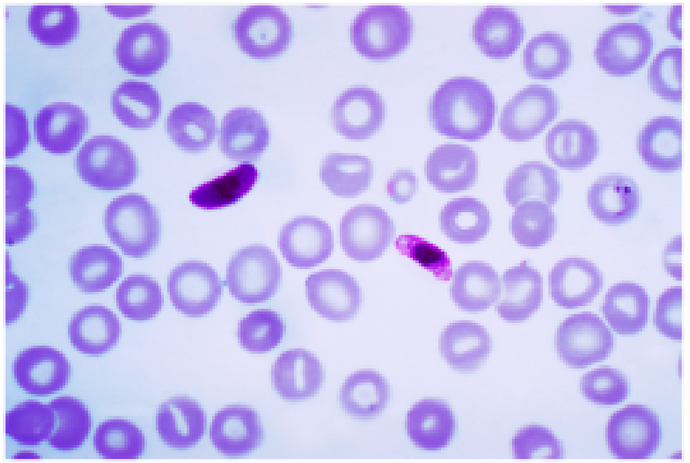
Background

Methods

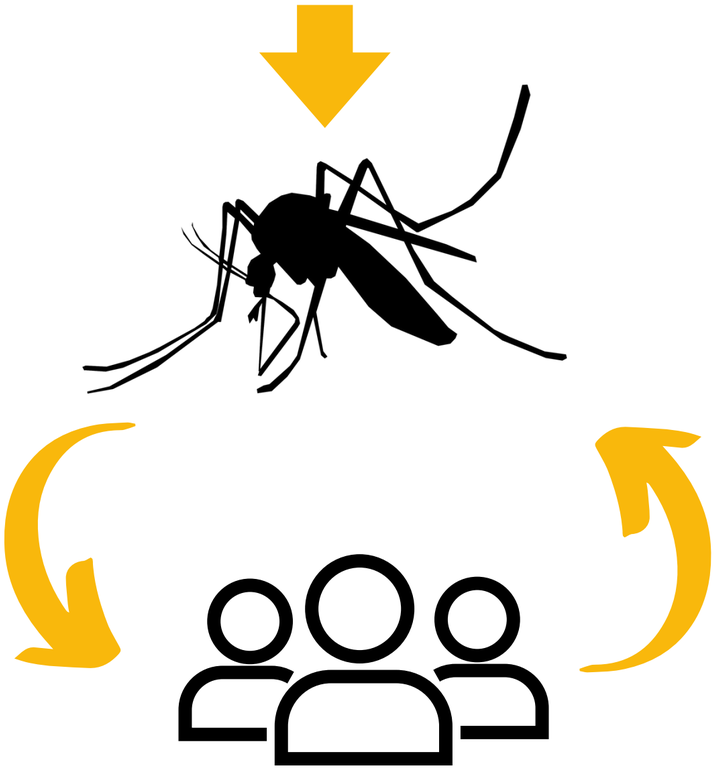
Results

Malaria

- Deadly mosquito-borne disease caused by Plasmodium spp.
- Carried by Anopheles (mosquitoes)
- Fever, death, pulmonary complications, and death
- Sub-saharan Africa is disproportionately burdened by this disease. 600,000 malaria-related deaths in 2021 alone.



P. falciparum



Mathematical Modelling in Biology

Mathematical abstraction of reality

Use empirical data and sets of assumptions to:

01 → 02 → 03 → 04

Step 1

Make predictions and build a hypothesis

Step 2

Determine the appropriate methods to test assumptions

Step 3

Test hypothesis based on analysis of data/model fit

Step 4

Interpret the results and make connections to the data/real world

Mathematical Modelling & Malaria

Many different ways to investigate malaria dynamic using mathematical models:

Ross-MacDonald Model

- Simple *fixed* seven-parameter

- model parameters:
  - SI (Susceptible, Infectious)
  - mosquito biting rate
  - transmission probability (mosquito to human)
  - transmission probability (human to mosquito)
  - ratio of mosquitoes to humans
  - recovery rate of humans
  - recovery rate of mosquitoes

Mathematical Modelling & Malaria

Some Modifications to This Model:

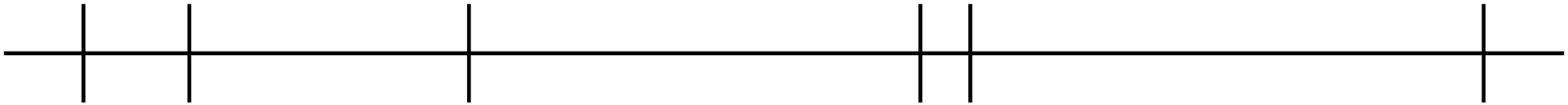
- **Torrez-Sorrando & Rodriguez,** space fragmentation
- **Parham & Michael**, rainfall and temperature
- **Yang**, global warming and social/economic factors
- **Aron & May**, age specific immunity
- Stochasticity in spatial contact structure, temporal forcing, unlimited superinfection
- Population size/stochasticity are intimately related, **haven’t been fully characterized**

Deterministic Model: Ross MacDonald
Incorporating Stochasticity: Gillespie Algorithm

“How does varying stochasticity (population size) affect the ability to make inferences when fitting a mathematical model to data?”

What is a Gillespie Algorithm?

Stochastic simulation where each state is driven by a **discrete event** and its associated **propensity rate**. The time from event-to-event is determined stochastically.



Time

$\mathbf{a}_i$ = Reactions rates, $\boldsymbol{\alpha}_i$ = Propensities, $\boldsymbol{\alpha}_0$ = Total Propensity, $\boldsymbol{\tau}$ =Time increment

Application of the Ross Macdonald Function

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

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

Human Recovery Rate = **γ**, Vector Death Rate = **μ**, Infectious Humans = **IH**, Human Population = **H**, Infectious Vectors = **IV**, Vector Population = **V**, Mosquito Biting Rate = **a**, Prob. Transmission from V to H = **b**, Prob. Transmission from H to V = **c**

Incorporating Stochasticity: Gillespie Algorithm

Propensities

$$ab\frac{I_v}{H}(H - I_h) \quad \gamma I_h \quad ac\frac{I_h}{H}(V - I_v) \quad \mu I_v$$

In each step of the Gillespie Algorithm, one must occur:

Human: +1 human infected population, **-1** human infected population

Vector: +1 infected vector population, **-1** vector infected population

Research Question

“How does varying stochasticity affects the ability to make inferences when fitting a mathematical model to data?”

In the context of malaria:

“How does incorporating stochasticity through changing population sizes influences our ability to model malaria dynamics with the Ross MacDonald model?”

Methodology

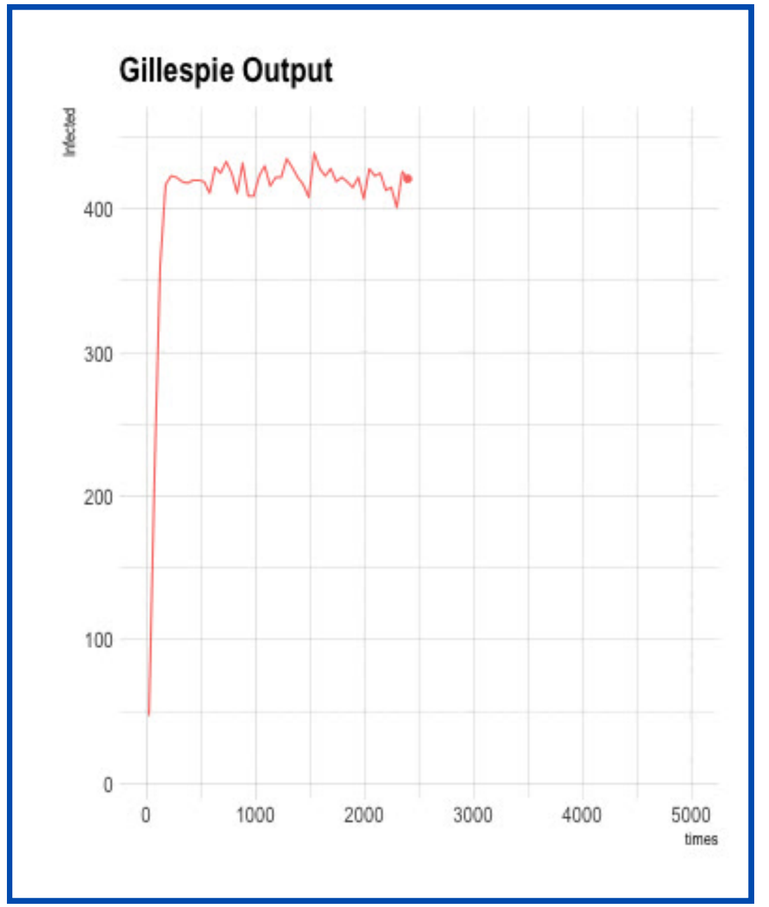
1. Establish Ross-MacDonald Model in R

2. Incorporate Gillespie Algorithm (assign propensities)

3. Vary stochasticity (population sizes)

4. Fit to model (likelihood) and plot output

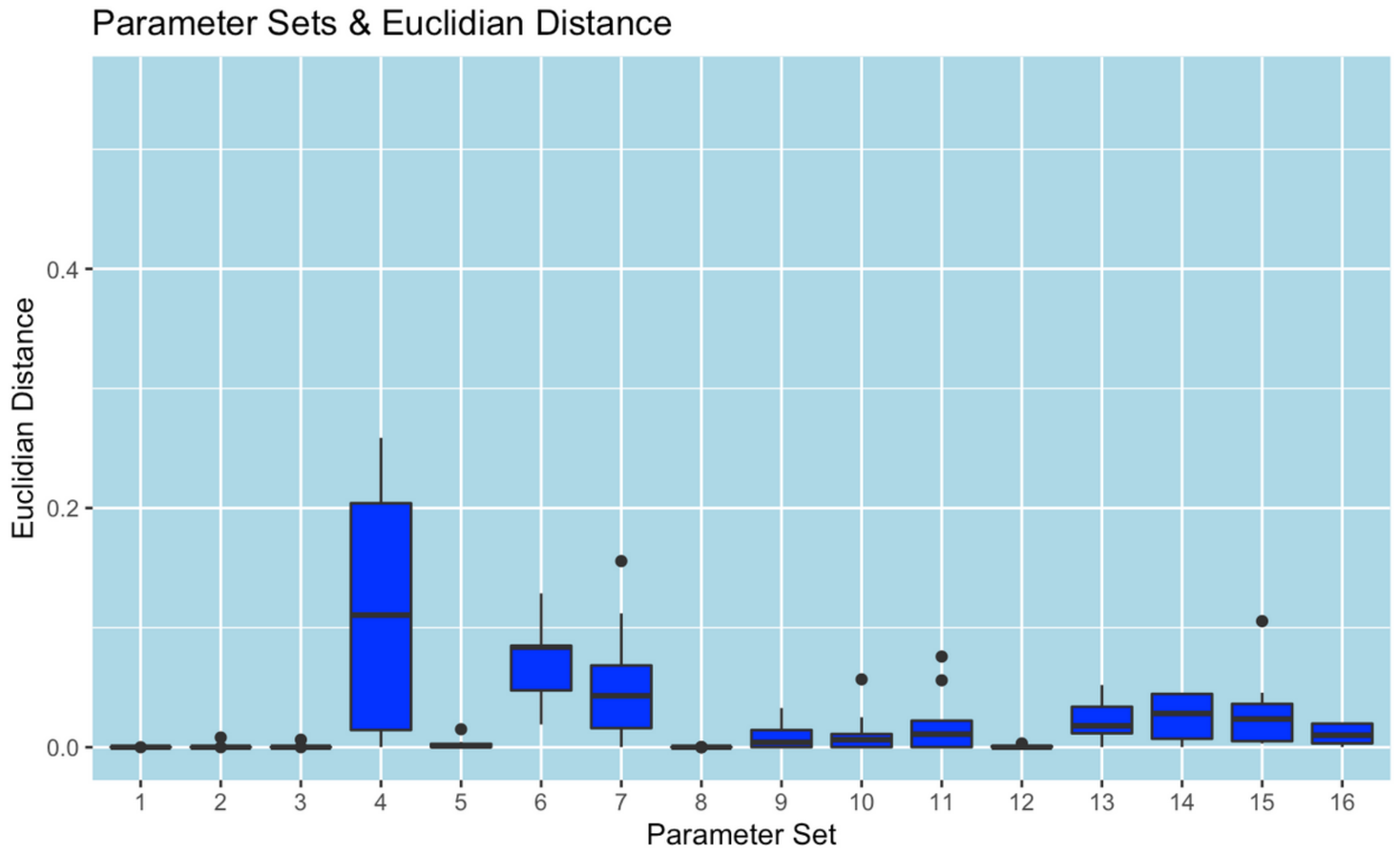
5. Assess accuracy of results



Summary Statistics

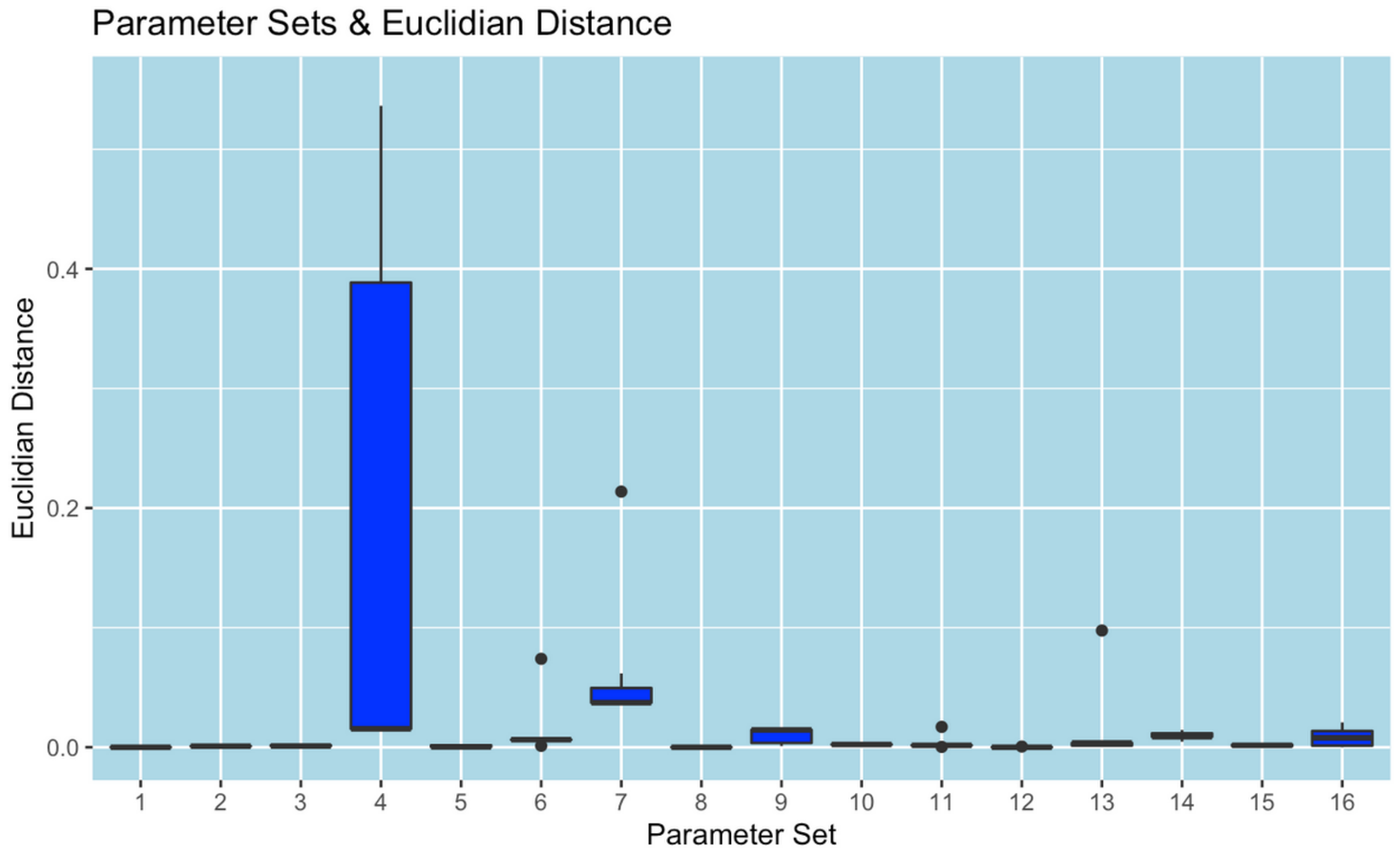
- Human infection data was used to derive inferences.
- 3 sets of parameters were analyzed throughout this study
- Parameter sets had sample sizes of 8, 7, and 6, for the first set, second set, and third set respectively.
- By Population
  - For small populations (n = 9), mean distance of 0.03
  - For medium populations (n = 9), mean distance of 0.02
  - For large populations (n = 3), mean distance of 0.01

Results: Parameter Set 1



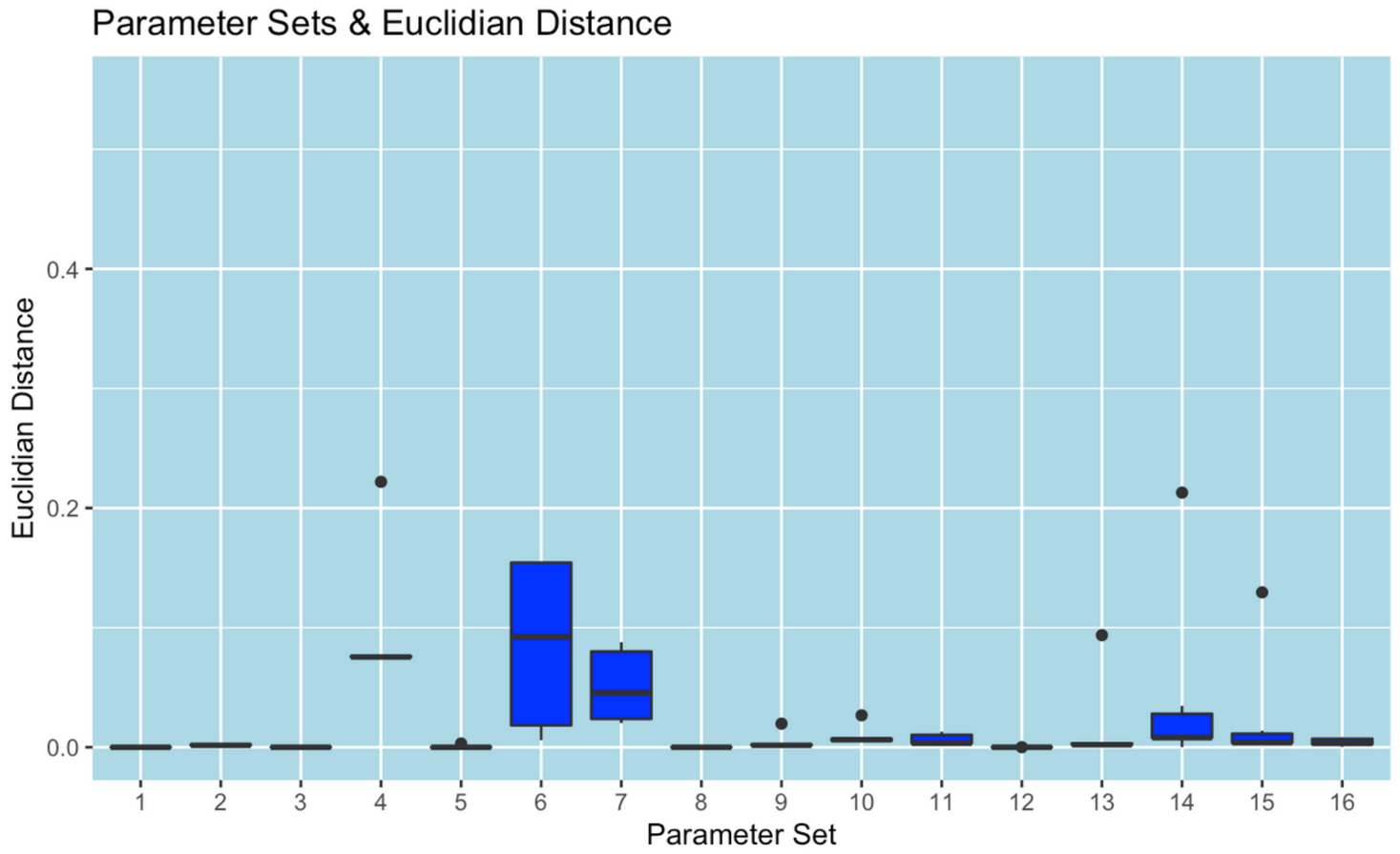
a=0.55, b=0.405, c=0.365, gamma=0.0275, mu = 0.19

Results: Parameter Set 2



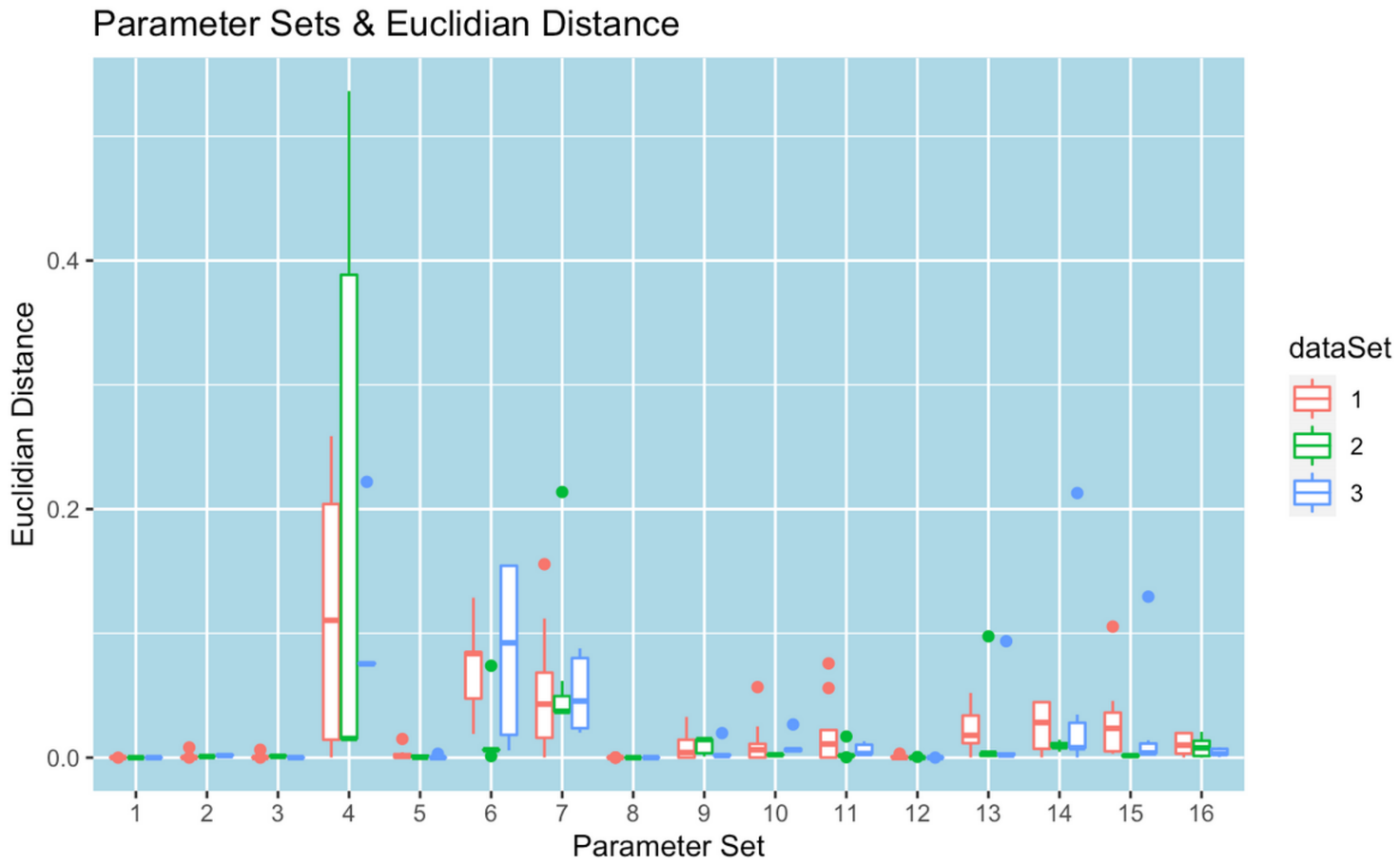
a=0.7, b=0.2, c=0.6, gamma=0.04, mu = 0.25

Results: Parameter Set 3



a=0.5, b=0.4, c=0.3, gamma=0.025, mu = 0.1

Results: Comparisons



Optimization and Scaling-Up

- **Changing how data was saved.**
  - **Overall process had a high run time, thus a step-by-step approach may be preferable**
- **Minimizing the number of parameters estimated.**
  - **Estimating up to two parameters entails 1266 iterations!**
  - **Doing up to five would yield 248832.**
- **Improved stochastic approach.**
- **Utilizing clusters.**
  - **More data leads to stronger predictions.**

Conclusion

- ***Generally, the distance between the real parameters and our estimates decreases with population size.***
  - ***This should be taken lightly without formal analysis and larger datasets.***

- ***Briefly, when parameters were estimated standalone, there was reasonable success.***

- ***Estimating a and b, a and c, b and c, b and mu, yielded the largest distances***
  - ***This suggests that b is more difficult to estimate with this model***
  - ***Gamma was frequently estimated correctly, including instances where another parameter was also being estimated.***

References

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.

J. C. Tavares, Malaria. Morgan & Claypool, 2013.

UNICEF, UNICEF. May 2023. Available: https://data.unicef.org/topic/child-health/malaria/#:~:text=Nearly%20every%20minute%2C%20a%20child%20under%20five%20dies%20of%20malaria,to%20619%2C000%20deaths%20in%20total

THANK YOU!