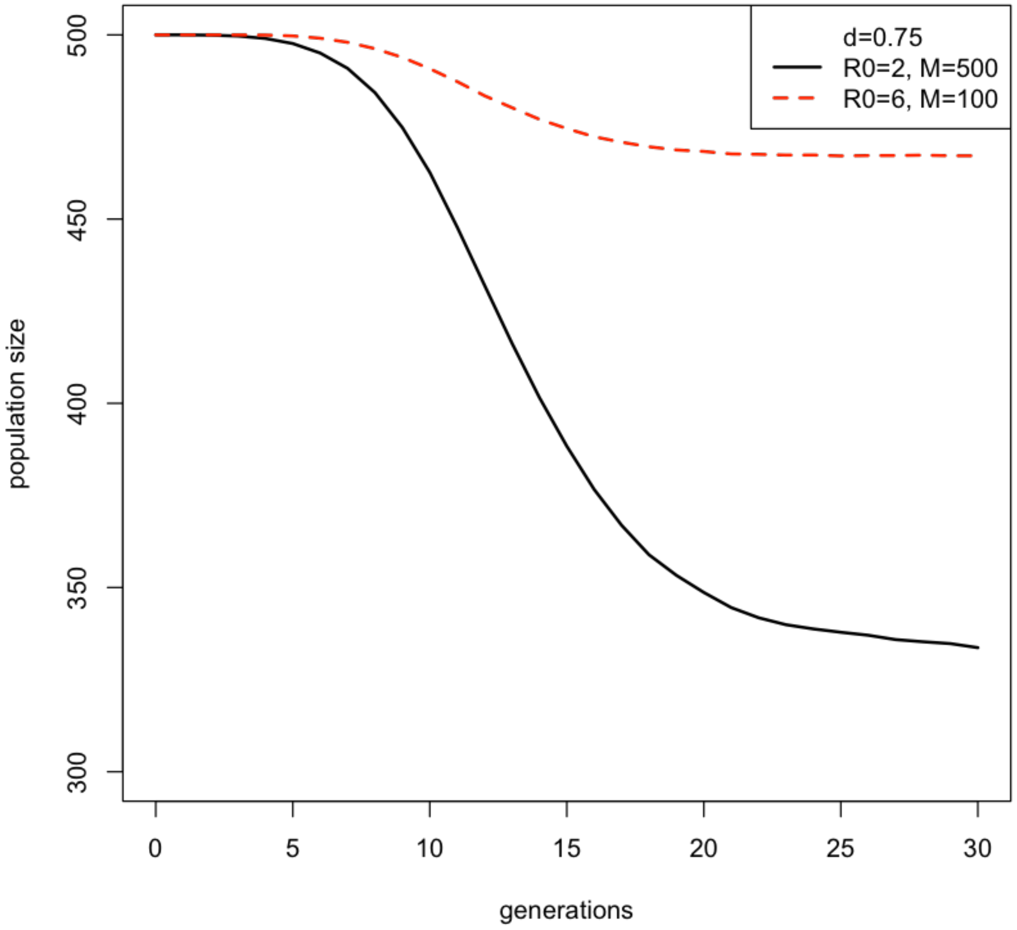
Estimating *Anopheles gambiae Rm* using insecticide and bed net field trial data

Using ecological modelling and MLE inference to inform gene drive models.

Katie Willis

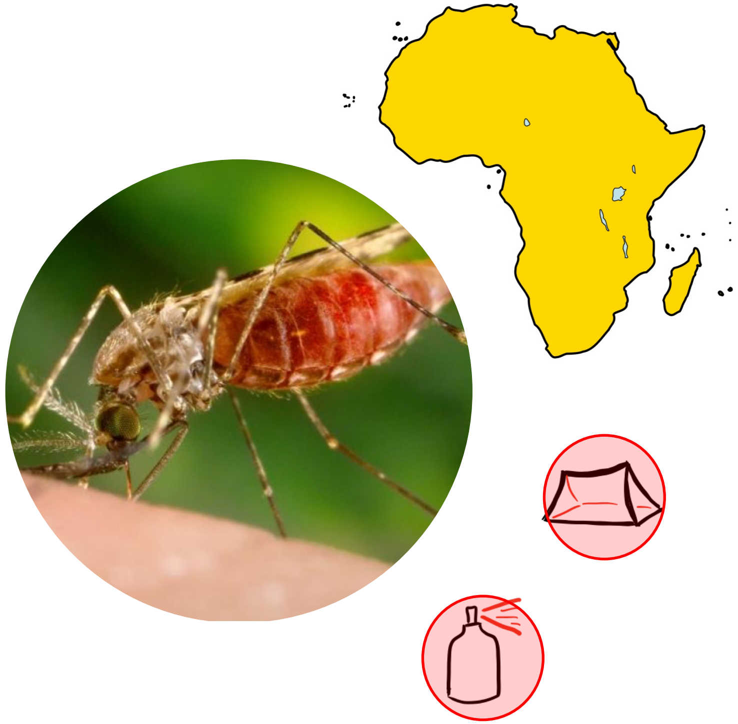
What is the Rm of wild populations of mosquitoes?

- We have seen that populations with different R0 (Rm) respond differently to the same gene drive intervention.
  - To design a HEG/transgene which reduces the population by a desired amount we first need to know the Rm of the target population.



What is the Rm of wild populations of mosquitoes?

- We have seen that populations with different R0 (Rm) respond differently to the same gene drive intervention.
  - To design a HEG which reduces the population by a desired amount we first need to know the Rm of the target population.
- Can we estimate Rm using data from existing field trials evaluating the impact of other interventions on density?
  - Use **Ecological modelling** and **maximum likelihood estimation** to estimate Rm and inform gene drive models.



1. How intervention trial data contains information on Rm
2. Model of An gambiae population dynamics
3. Intervention response field data
4. Maximum likelihood estimation

1. How intervention data contains information on Rm?
2. Model of An gambiae population dynamics
3. Intervention response field data
4. Maximum likelihood estimation

What is Rm?

Beverton-holt model

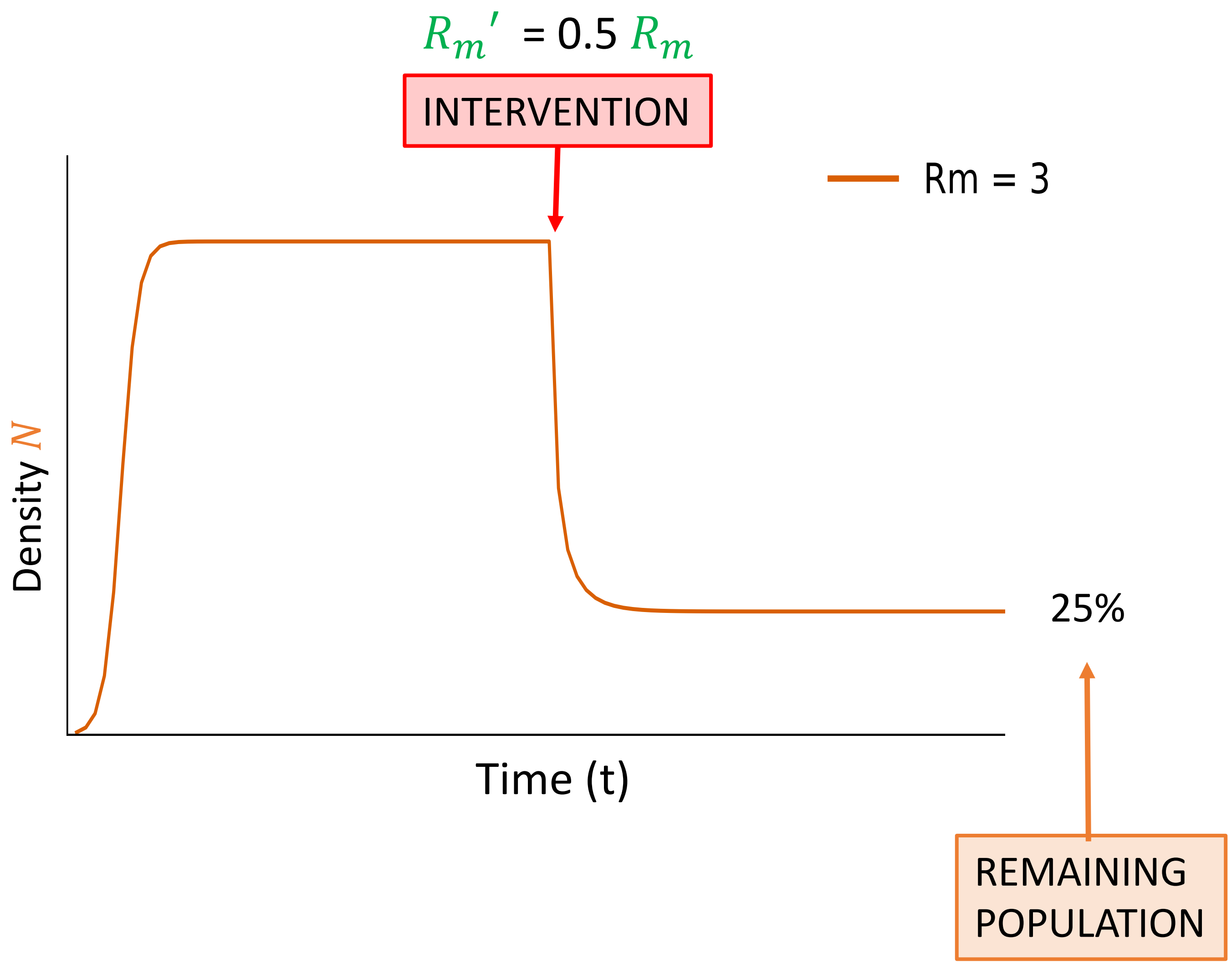
$$N_{t+1} = \frac{N_t R_0}{1+\frac{N_t}{m}} = N_t R_0 \frac{m}{m+N_t}$$

$$N_{t+1} = N_t R_m \frac{\alpha}{\alpha + N_t}$$

R_m = Intrinsic growth rate

Density dependent regulation

Rm estimation using intervention data



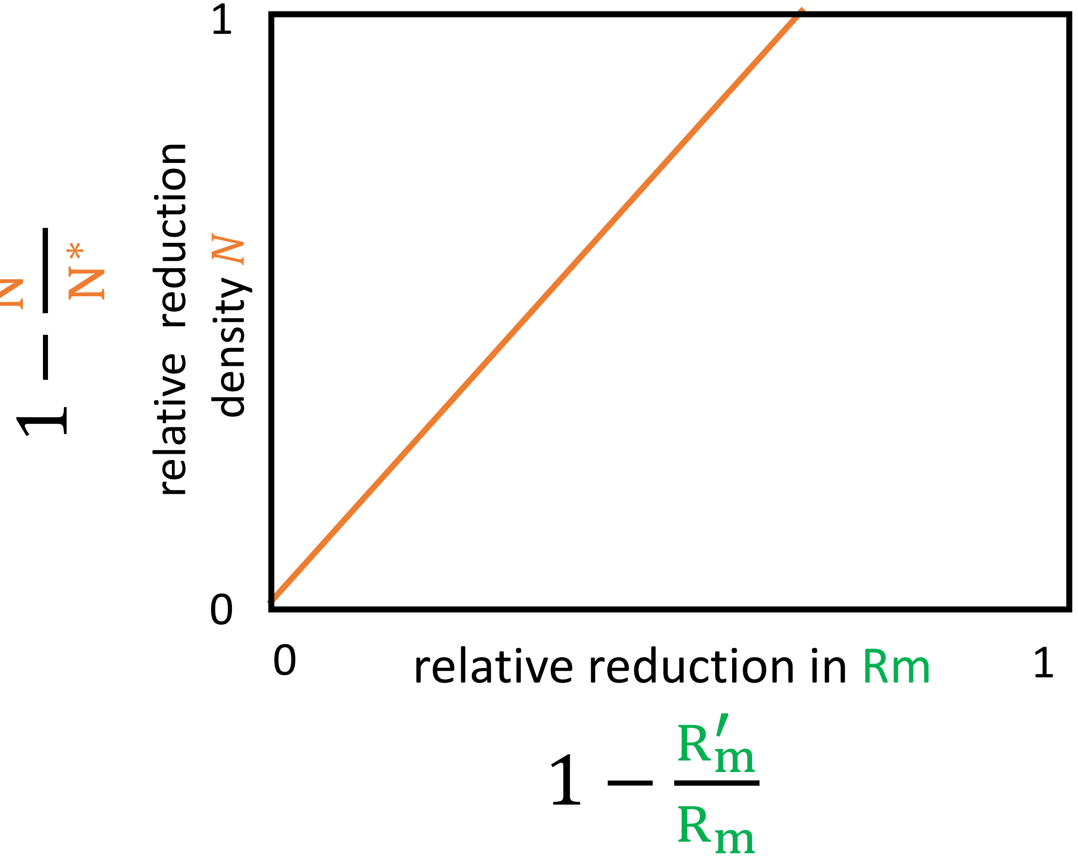
Before intervention

$$N_{t+1} = N_t R_m \frac{\alpha}{\alpha + N_t}$$

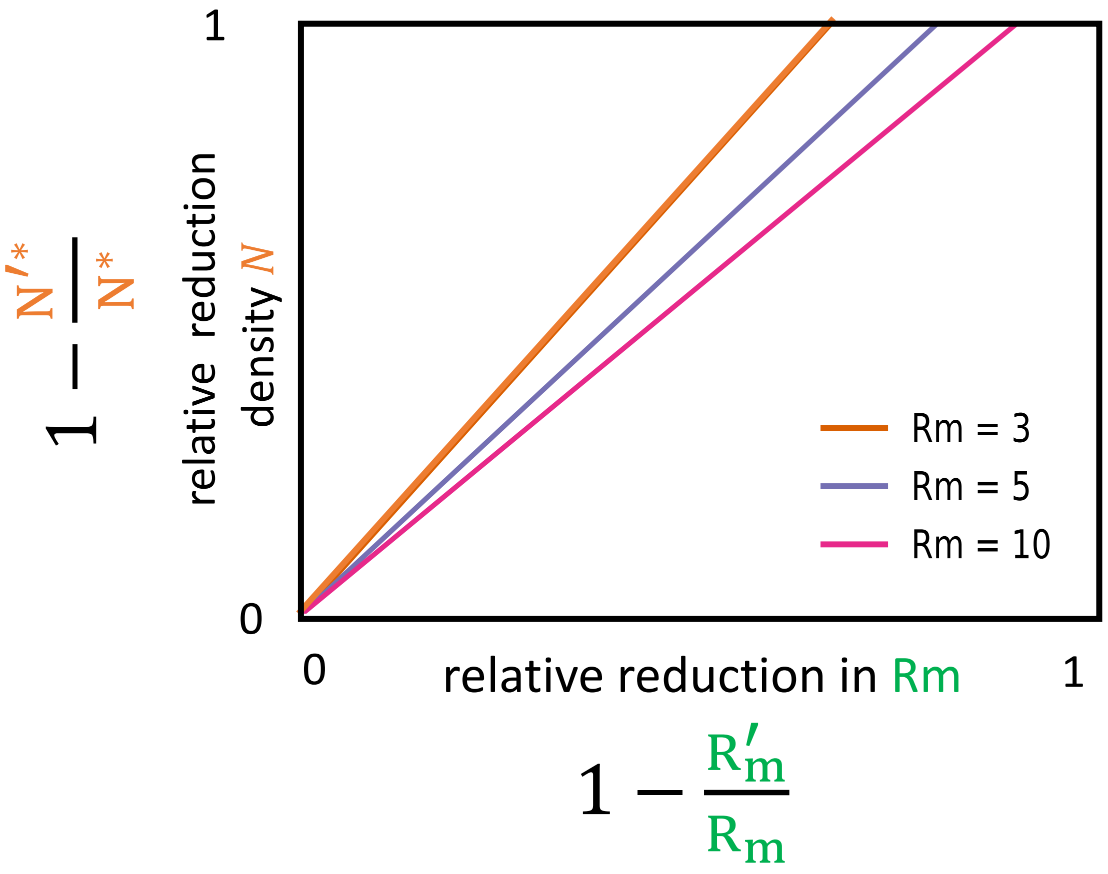
After intervention

$$N_{t+1} = N_t R_m{}' \frac{\alpha}{\alpha + N_t}$$

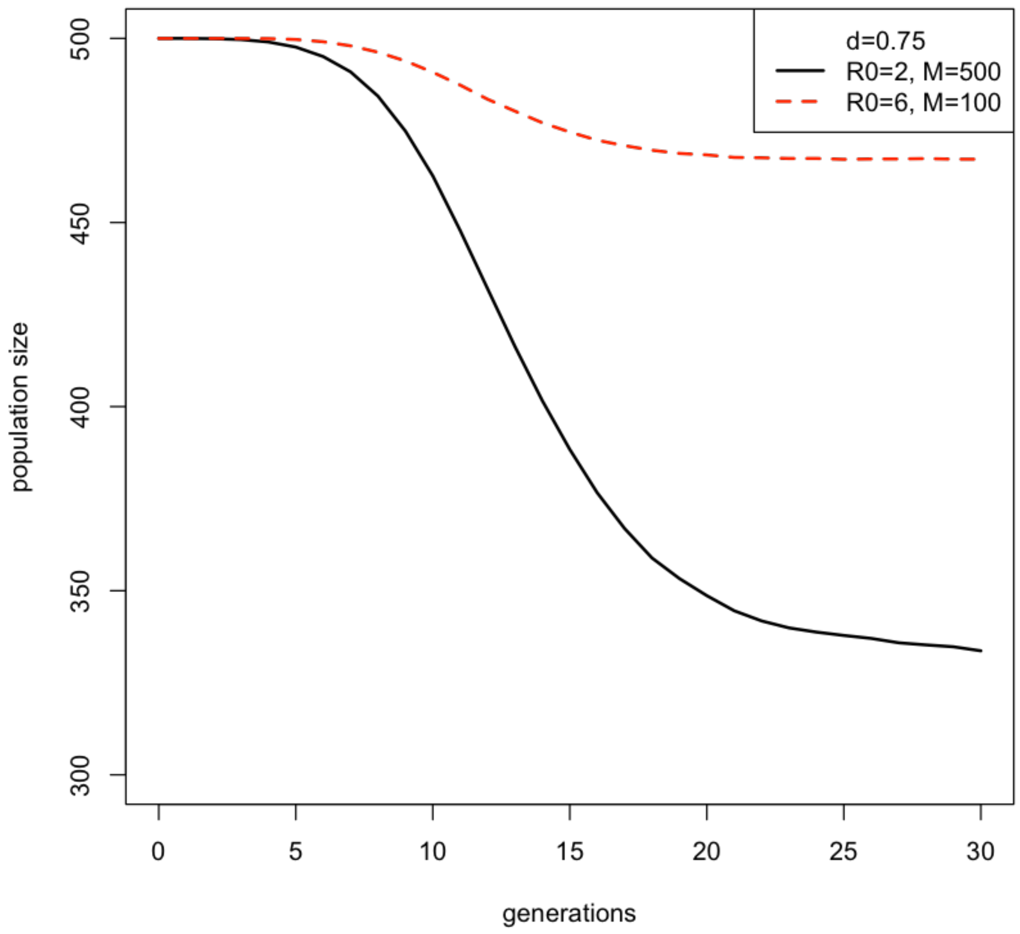
Rm estimation using intervention data



Rm estimation using intervention data

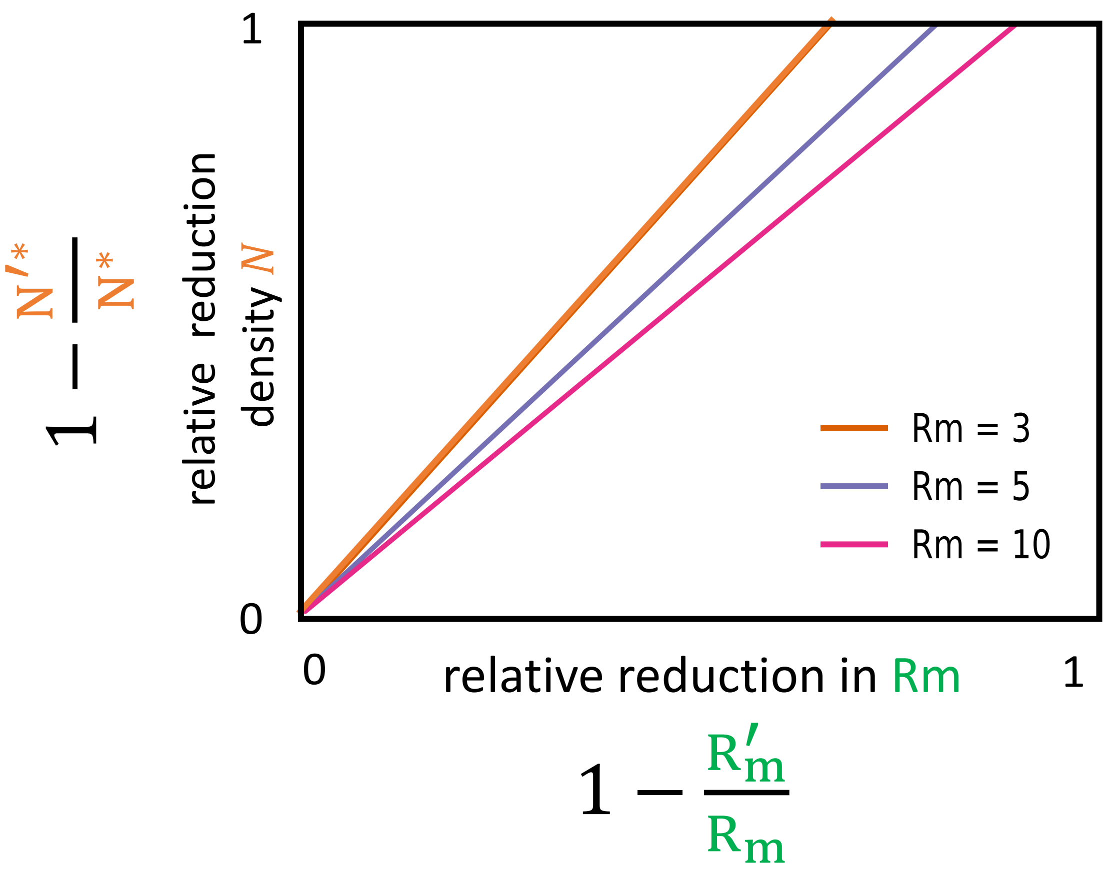


For a given reduction in intrinsic growth rate (R_m), a population with lower R_m will show a greater reduction in density (N).



This is what we saw using our gene drive simulations!

Rm estimation using intervention data



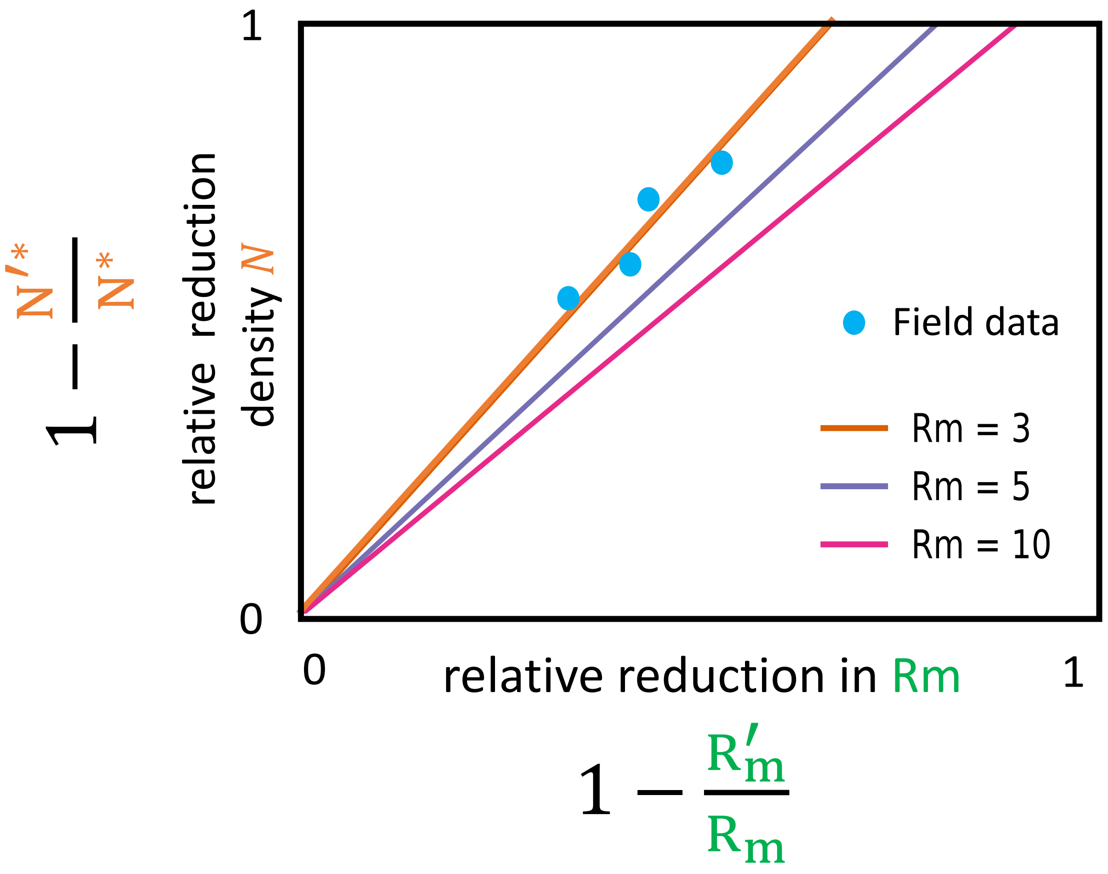
For a given reduction in intrinsic growth rate (R_m), a population with lower R_m will show a greater reduction in density (N).

- **A model of population dynamics**

  Relating a change in R_m to a change in N at equilibrium

  $N^* = (R_m - 1)\alpha \qquad N'^* = (R_m' - 1)\alpha$

Rm estimation using intervention data



For a given reduction in intrinsic growth rate (R_m), a population with lower R_m will show a greater reduction in density (N).

- **A model of population dynamics**

  Relating a change in R_m to a change in N at equilibrium

  $N^* = (R_m - 1)\alpha \qquad N'^* = (R_m' - 1)\alpha$

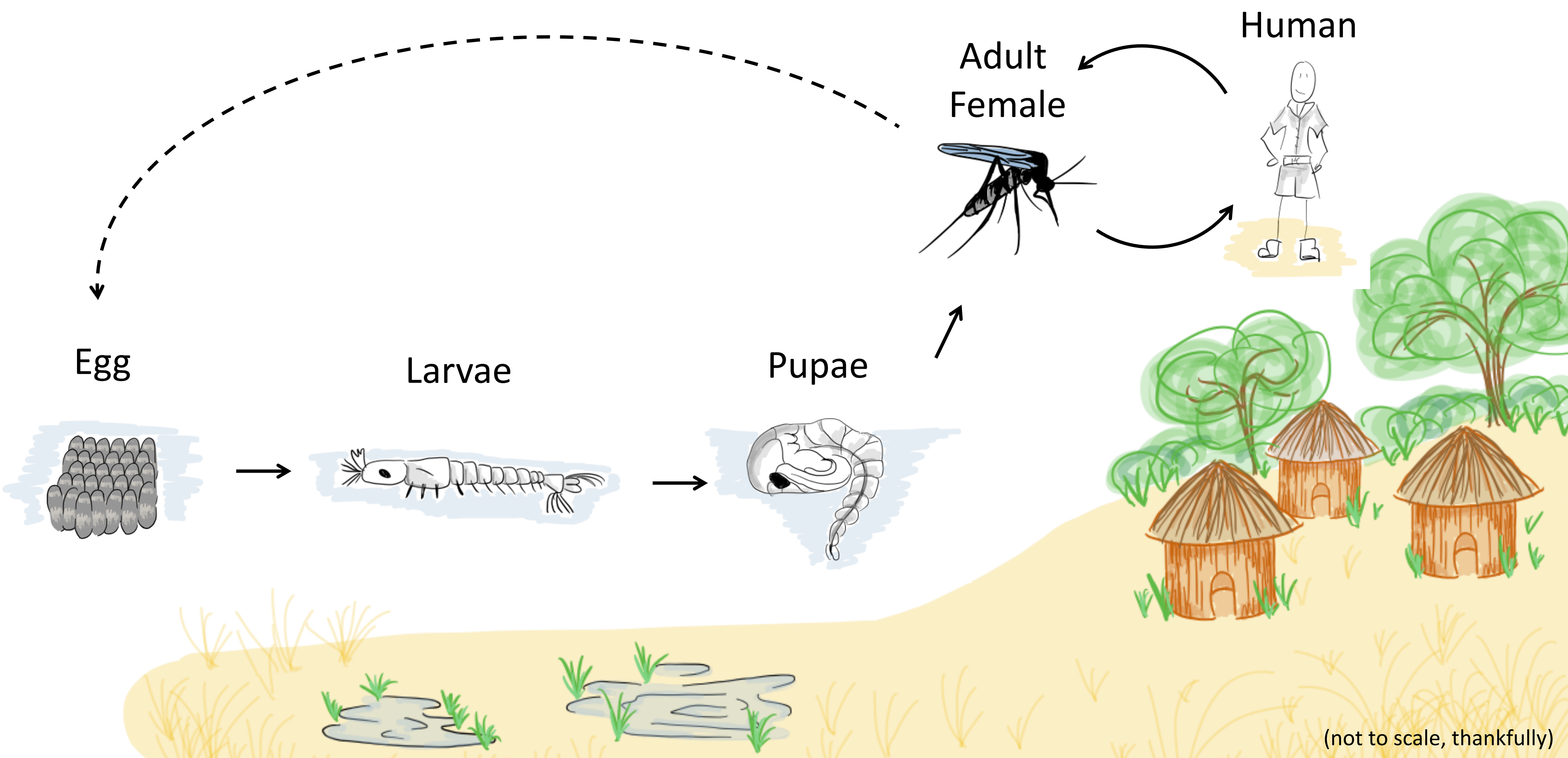
- **Intervention field data**

  x Effect on the intrinsic growth rate R_m

  y Change in equilibrium population density N

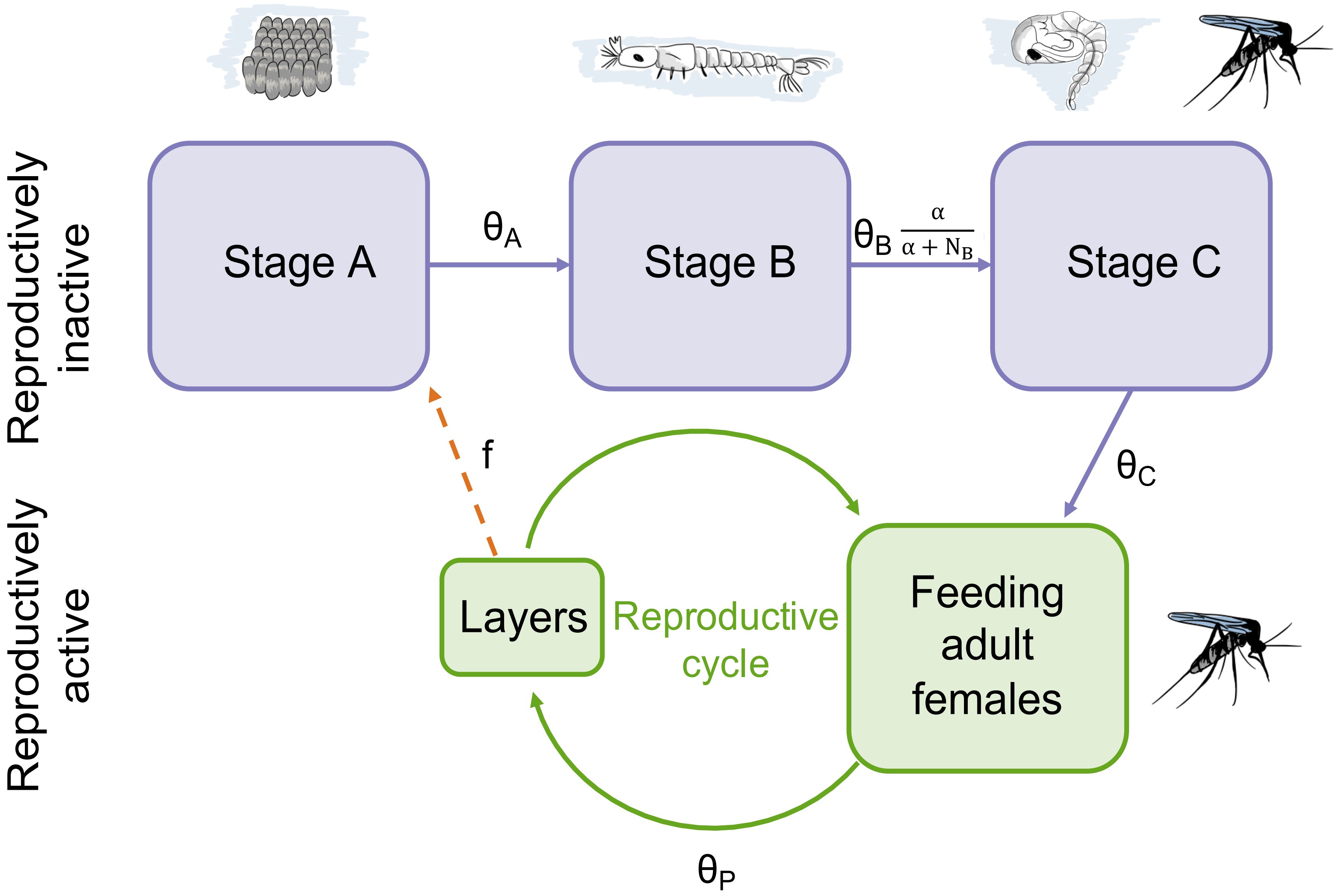
1. How intervention data contains information on Rm?
2. Model of An gambiae population dynamics
3. Intervention response field data
4. Maximum likelihood estimation

Using a model of An gambiae population dynamics

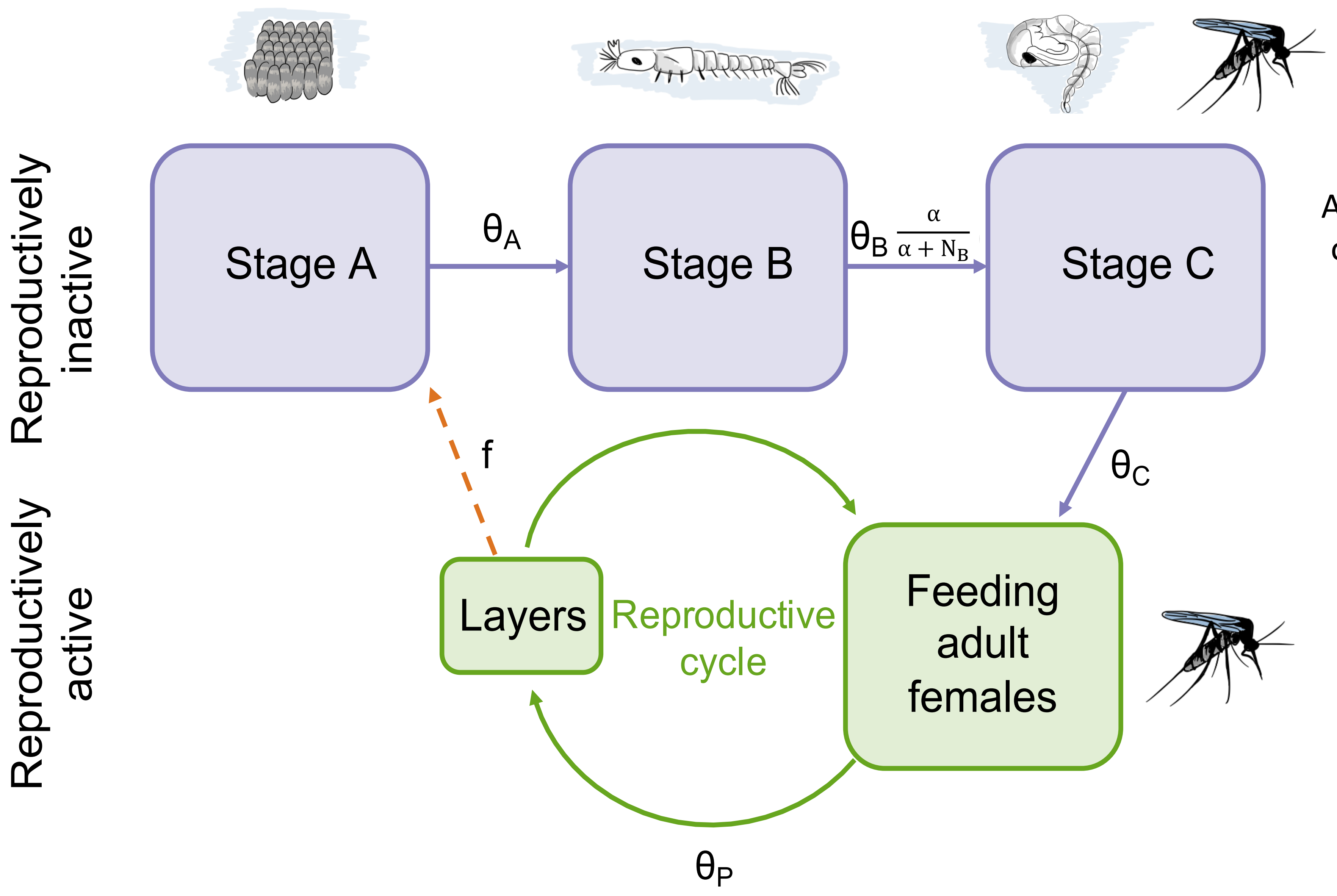


(not to scale, thankfully)

Using a model of An gambiae population dynamics



Using a model of An gambiae population dynamics



Growth rate

$$r = \boxed{\frac{\theta_P}{1-\theta_P} f} \boxed{\theta_A \theta_B \frac{\alpha}{\alpha + N_B} \theta_C}$$

Average number of eggs laid per female in her lifetime

Probability of each surviving the reproductively inactive stage

Intrinsic growth rate

$$R_m = \frac{\theta_P}{1-\theta_P} f\, \theta_A \theta_B \theta_C$$

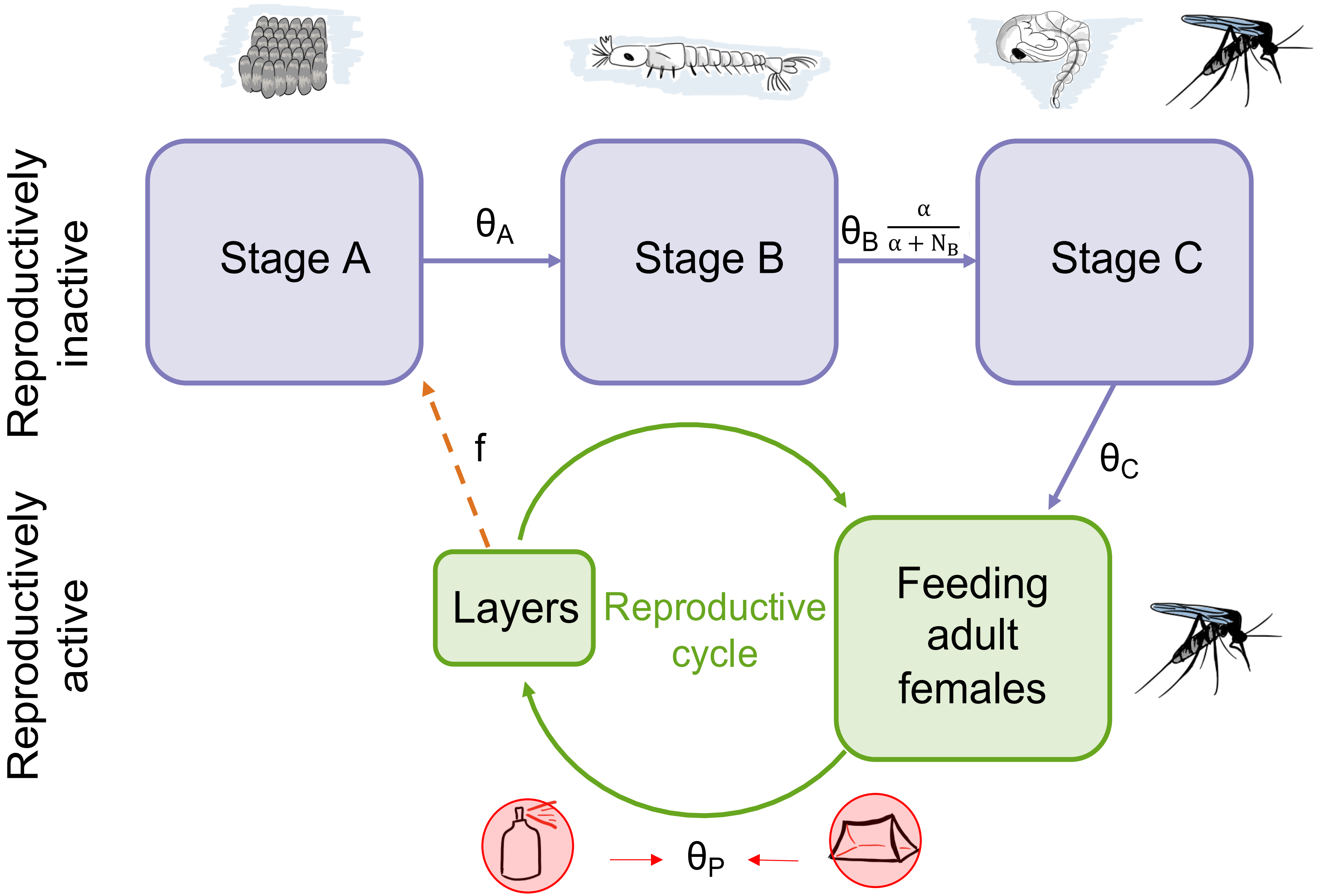
Emerging females

$$N_{EF}^{*} = \frac{\alpha (R_m - 1) \theta_B \theta_C}{2 R_m}$$

Equilibrium adult density

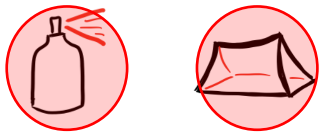
$$N^{*} = N_{EF}^{*} \frac{1}{1-\theta_P}$$

Using a model of An gambiae population dynamics



We are interested in interventions which modify Rm

$$R_m = \frac{\theta_P}{1-\theta_P} f\,\theta_A\theta_B\theta_C$$



Insecticide bed nets and sprays act to reduce Rm by reducing adult survival (θ_P)

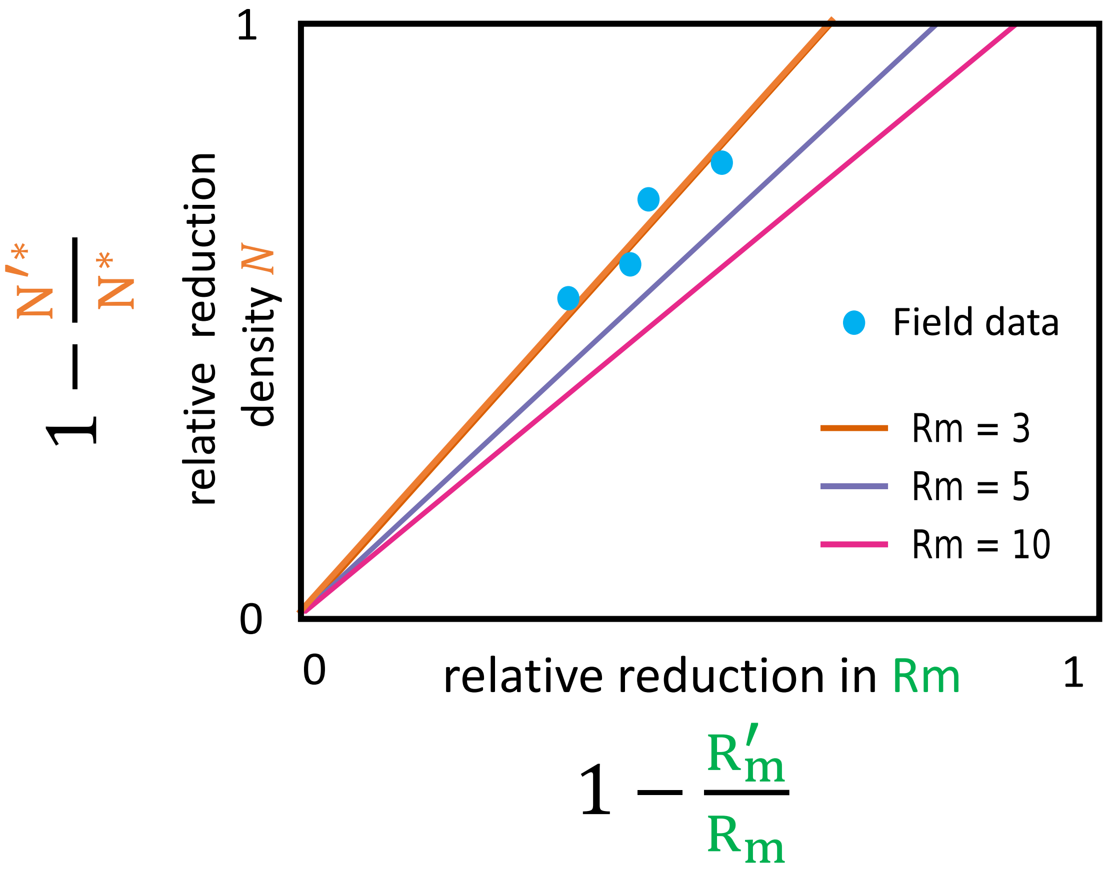
This changes the average number of egg laying cycles by

$$\frac{\theta_P}{1-\theta_P} \text{ to } \frac{\theta_P{}'}{1-\theta_P{}'}$$

And therefore reduces Rm by

$$1-\frac{R_m'}{R_m} = 1-\frac{(1-\theta_P)\theta_{P'}}{(1-\theta_{P'})\theta_P}$$

Rm estimation using intervention data



For a given reduction in intrinsic growth rate (R_m), a population with lower R_m will show a greater reduction in density (N).

A model of population dynamics

Relating a change in R_m to a change in N at equilibrium

- **Intervention field data**

x Effect on the intrinsic growth rate R_m

y Change in equilibrium population density N

1. How intervention data contains information on Rm?
2. Model of An gambiae population dynamics
3. Intervention response field data
4. Maximum likelihood estimation

Intervention response field data

x Effect on the intrinsic growth rate R_m $(\theta_p , \theta_{P'})$

y Change in equilibrium population density N

SAMPLING INTERVENTION SAMPLING

Pre-intervention

Post-intervention

+

Intervention response field data

$$1 - \frac{R'_m}{R_m} = 1 - \frac{(1-\theta_P)\theta_{P'}}{(1-\theta_{P'})\theta_P}$$

x Effect on the intrinsic growth rate R_m $(\theta_p , \theta_{P'})$

y Change in equilibrium population density N

SAMPLING INTERVENTION SAMPLING

Pre-intervention

Post-intervention

+

P = number tested for parity

ρ = number parous

$\theta_p = \frac{\rho}{P}$ = parous rate pre-intervention

P' = number tested for parity

ρ′ = number parous

$\theta'_p = \frac{\rho'}{P\prime}$ = parous rate post-intervention

Intervention response field data

$$1-\frac{R'_m}{R_m}=1-\frac{(1-\theta_P)\theta_{P'}}{(1-\theta_{P'})\theta_P}$$

x Effect on the intrinsic growth rate R_m $(\theta_p\ ,\theta_{P'})$

y Change in equilibrium population density N

SAMPLING

INTERVENTION

SAMPLING

Pre-intervention

Post-intervention

+

P = number tested for parity

ρ = number parous

$\theta_p=\frac{\rho}{P}$ = parous rate pre-intervention

n = number of mosquitoes sampled pre-intervention

P' = number tested for parity

ρ' = number parous

$\theta'_p=\frac{\rho'}{P\prime}$ = parous rate post-intervention

n'= number of mosquitoes sampled pre-intervention

Intervention response field data

x Effect on the intrinsic growth rate R_m $(\theta_p\ , \theta_{P'})$

$$1 - \frac{R'_m}{R_m} = 1 - \frac{(1-\theta_P)\theta_{P'}}{(1-\theta_{P'})\theta_P}$$

y Change in equilibrium population density N

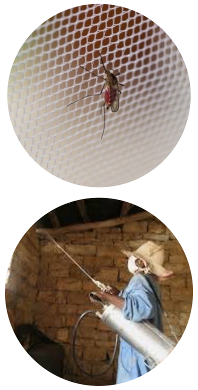
$$1 - \frac{N'^*}{N^*} = 1 - \frac{n'}{n}$$

SAMPLING | INTERVENTION | SAMPLING

Pre-intervention

Post-intervention

+



P = number tested for parity

ρ = number parous

$\theta_p = \frac{\rho}{P}$ = parous rate pre-intervention

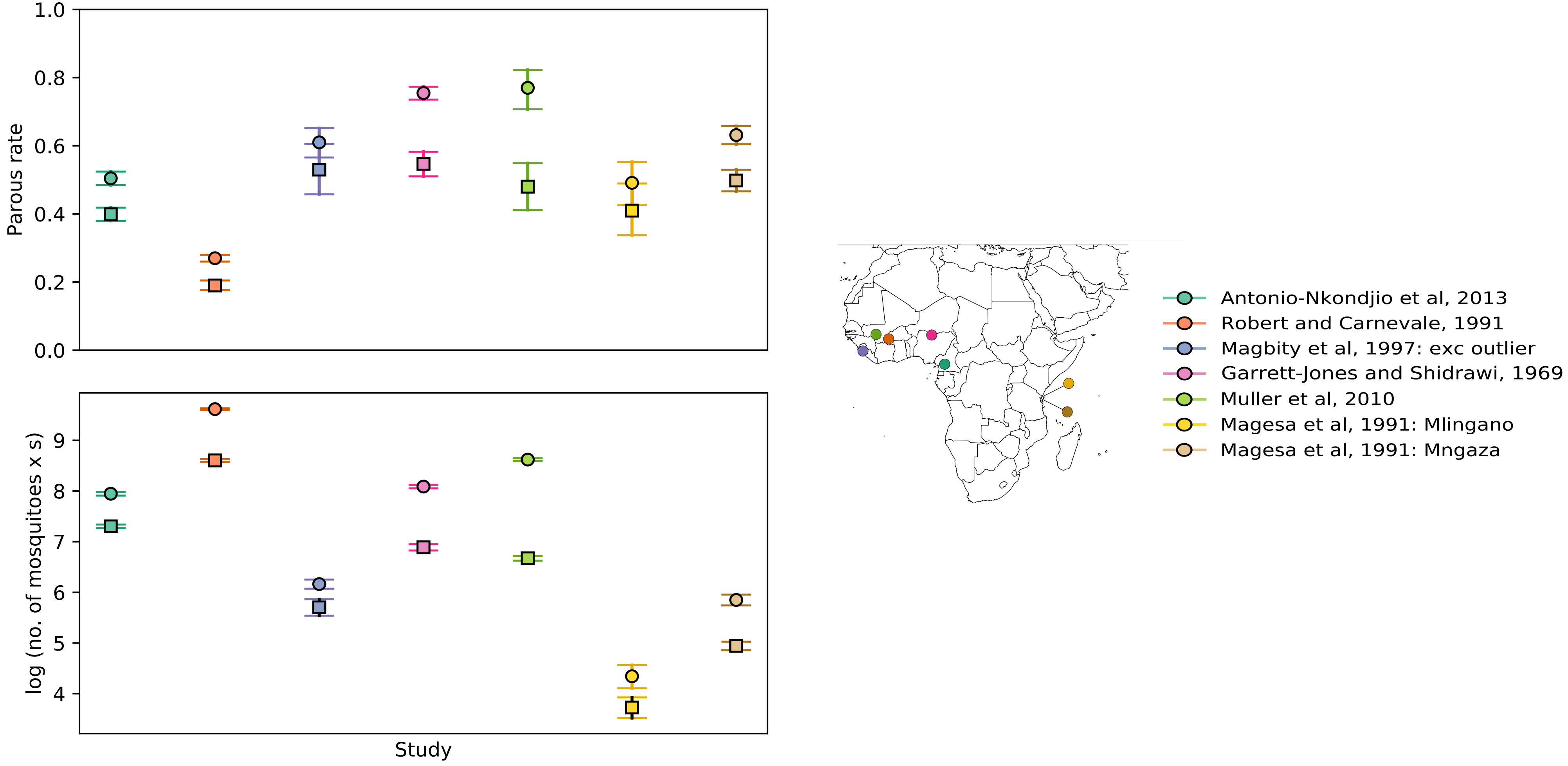
n = number of mosquitoes sampled pre-intervention

P' = number tested for parity

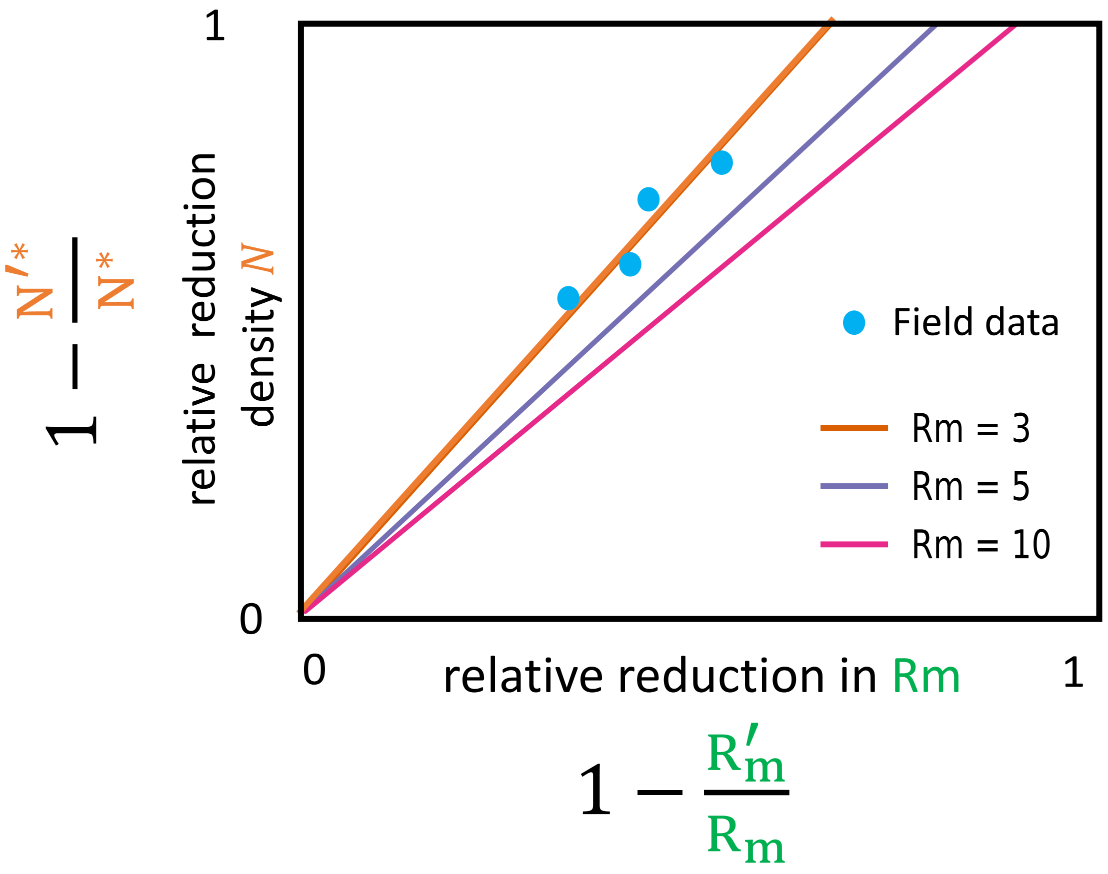
ρ' = number parous

$\theta'_p = \frac{\rho'}{P'}$ = parous rate post-intervention

n'= number of mosquitoes sampled pre-intervention


Parous rate
1.0
0.8
0.6
0.4
0.2
0.0
log (no. of mosquitoes x s)
9
8
7
6
5
4
Study
Antonio-Nkondjio et al, 2013
Robert and Carnevale, 1991
Magbity et al, 1997: exc outlier
Garrett-Jones and Shidrawi, 1969
Muller et al, 2010
Magesa et al, 1991: Mlingano
Magesa et al, 1991: Mngaza

Rm estimation using intervention data



For a given reduction in intrinsic growth rate (R_m), a population with lower R_m will show a greater reduction in density (N).

✓ **A model of population dynamics**

Relating a change in R_m to a change in N at equilibrium

✓ **Intervention field data**

x Effect on the intrinsic growth rate R_m

y Change in equilibrium population density N

1. How intervention data contains information on Rm?
2. Model of An gambiae population dynamics
3. Intervention response field data
4. Maximum likelihood estimation

Maximum likelihood estimation: REMINDER

MLE

- Maximum Likelihood estimation (MLE) is a method to estimate parameters of a statistical model
- When the method is applied to a dataset with a statistical model, MLE provides estimates for the associated parameters.
- "The parameter values that makes the observed dataset most *probable*."

$$L(\theta|x) = p(x|\theta)$$

- The likelihood function $L(\theta|x)$ is used to quantify how likely the parameter values (θ) are given some observation (x).
- Treat the parameters as unknown
- For each sample (x) and a given model, let θ be a parameter value at which $L(\theta|x)$ attains a maximum.
- This is then the maximum likelihood estimate of θ for x

Maximum likelihood estimation:

Our statistical model:

To build our likelihood function we first need to make some assumptions about the sampling probability of our data.

Parous rates

P = number tested for parity

ρ = number parous

$\theta_p = \frac{\rho}{P}$ = parous rate pre-intervention

Binomial distribution

$\theta_P \sim \text{Bin}(\rho, P)$

Sum of P independent and identically distributed Bernoulli random variables, where a Bernoulli r.v. is a binary outcome (success or fail)

Maximum likelihood estimation:

Our statistical model:

To build our likelihood function we first need to make some assumptions about the sampling probability of our data.

Parous rates

P = number tested for parity

ρ = number parous

$\theta_p = \frac{\rho}{P}$ = parous rate pre-intervention

Binomial distribution

$\theta_P \sim \text{Bin}(\rho, P)$

Sum of P independent and identically distributed Bernoulli random variables, where a Bernoulli r.v. is a binary outcome (success or fail)

Mosquito counts

n = number of mosquitoes sampled pre-intervention

Ns = total number of feeding mosquitoes * sampling probability

Poisson distribution

$Ns \sim \text{Pois}(n)$

Number of events occurring in a fixed interval of time.

Maximum likelihood estimation

Our statistical model: $\theta_p \sim \text{Bin}(\rho, P)$ $Ns \sim \text{Pois}(n)$

$$
\begin{aligned}
logL(Parameters|Data) \\
&= \rho log\theta_p + (P-\rho)log(1-\theta_p) \\
&+ \rho' log\theta_p' + (P'-\rho')log(1-\theta_p') \\
&+ (-Ns + nlogNs - log(n!) \\
&+ (-N's + n'logN's - log(n'!)
\end{aligned}
$$

Parous rates → Binomial distribution

Mosquito counts → Poisson distribution

Parameters

θ_p = inferred parous rate

N = equilibrium number of mosquitoes

s = sampling probability

Data

n = number of caught mosquitoes

ρ = number of parous mosquitoes

P = number of dissected mosquitoes

$'$ denotes post intervention

Incorporate the model of mosquito population dynamics:

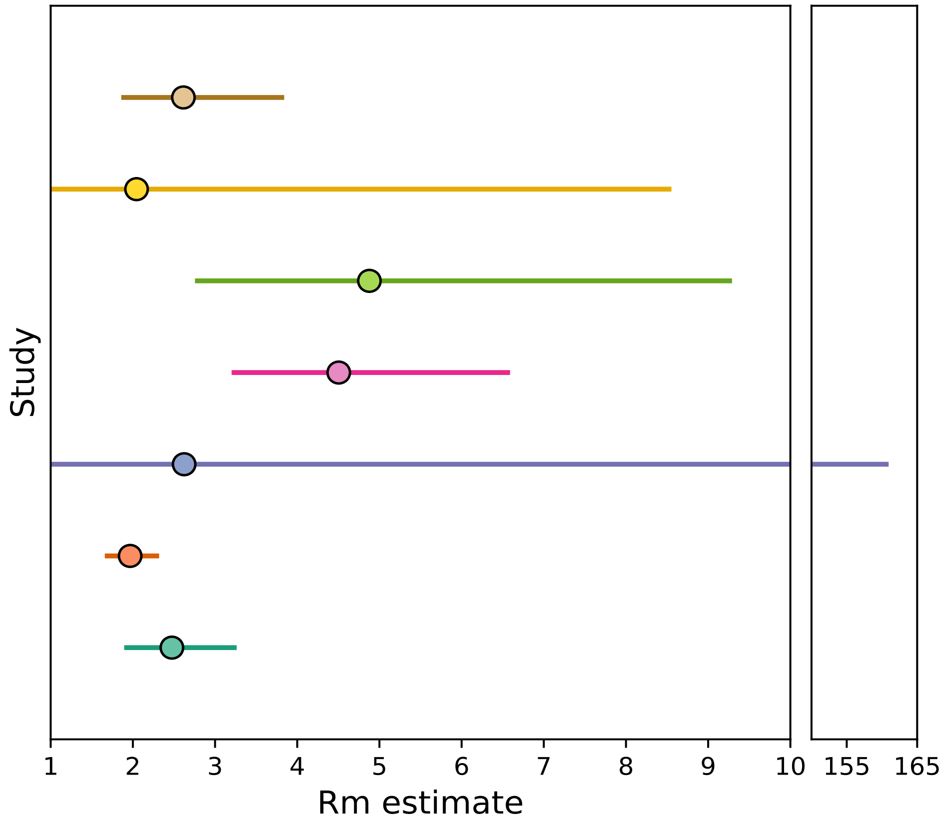
$$N = \frac{\alpha(R_m - 1)\theta_B\theta_C}{2R_m(1-\theta_P)} \qquad N' = \frac{\alpha(R_m' - 1)\theta_B\theta_C}{2R_m'(1-\theta_P')}$$

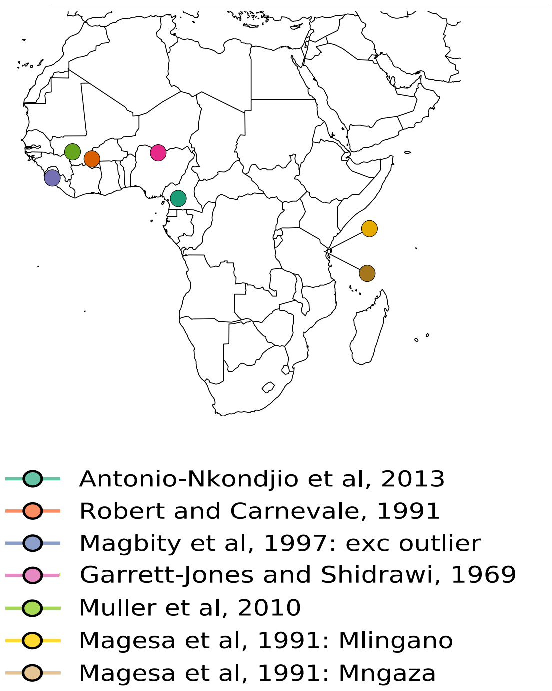
$$\boxed{logL(Rm, \theta_p, \theta_p', k | \rho, P, n, \rho', P', n')}$$

$$k = \alpha\theta_B\theta_C\, s$$

Maximum likelihood estimation: Maximum likelihood estimates

$$logL\left(Rm, \theta_p, \theta_p', k \middle| \rho, P, n, \rho', P', n'\right)$$





Conclusions

- Using an ecological model we can see that observations of changes in density in response to interventions can provide information of the intrinsic growth rate of a population

- MLE can be applied to a **dataset** using a **statistical model** to provide estimates for the associated **parameters**

- Here I apply MLE methods to **intervention field trial data** using a **model of mosquito population dynamics** to estimate the **Rm of the population**

MSc/MRes Project

Modelling the population dynamics of *Anopheles gambiae* mosquitoes.

Anopheles gambiae mosquitoes are the primary vectors of malaria and despite extensive application of population control strategies, the disease remains a substantial health burden. An improved understanding of An. gambiae population dynamics is required to predict, evaluate and optimize the impact of intervention strategies on disease transmission. Of particular interest is the intrinsic population growth rate, a parameter which influences the strength of intervention required to modify population densities. Little is known about An. gambiae growth rate in the field, and current estimations have relied on collecting detailed demographic data, often measured under laboratory conditions. The project involves modifying and extending a statistical inference method developed by our group for estimating the intrinsic growth rate using field measurements of population densities in response to control strategies.

This is a fully theoretical project and is likely to involve:

- Extension of the current population dynamics model to include additional features e.g. **seasonality** and/or **migration**.
- Use of maximum likelihood/Bayesian approaches to infer population biology parameters from the data.
- Use of simulations to assess how well the inference method is able to estimate parameters under different ecological scenarios.
- Application of the method to suitable field data identified from the literature.

The project will require basic computational/programming skills and will provide opportunity to develop skills in mathematical modelling and maximum likelihood/Bayesian methods for parameter inference.