Bayesian optimization framework for recalibration of EMOD's within-host malaria model

Tobias M. Holden

IDM Symposium

October 2, 2024

Northwestern University Malaria Modeling

What is ‘calibration’?

- Fitting the *model* to *data* to infer values of unknown model parameters

What is ‘calibration’?

- Fitting the *model* to *data* to infer values of unknown model parameters
  - Parameters may be unknown because they…
    - Are impossible to measure
    - Require elaborate, costly, or unethical experiments

What is ‘calibration’?

- Fitting the *model* to *data* to infer values of unknown model parameters
  - Parameters may be unknown because they…
    - Are impossible to measure
    - Require elaborate, costly, or unethical experiments
- Using reference sites with data on transmission and malaria epidemiology

What is ‘calibration’?

- Fitting the *model* to *data* to infer values of unknown model parameters
  - Parameters may be unknown because they…
    - Are impossible to measure
    - Require elaborate, costly, or unethical experiments
- Using reference sites with data on transmission and malaria epidemiology
  - Older datasets: capture natural history of disease without interventions
  - Newer datasets: capture transmission dynamics in the context of interventions, and may include higher quality measurements

EMOD within-host dynamics could be improved

Motivation: Mathematical models need to be able to reproduce real-life disease dynamics to make meaningful predictions

EMOD within-host dynamics could be improved

Motivation: Mathematical models need to be able to reproduce real-life disease dynamics to make meaningful predictions

1. EMOD over-attributes transmission to infections that were once symptomatic

EMOD within-host dynamics could be improved

Motivation: Mathematical models need to be able to reproduce real-life disease dynamics to make meaningful predictions

1. EMOD over-attributes transmission to infections that were once symptomatic

2. EMOD simulations demonstrate a sudden and extreme rebound in clinical incidence following mass drug administration

What makes model calibration difficult?

Complexity of models like EMOD often comes with **long simulation times** and involves **large numbers of unknown input parameters**.

- At a certain point, calibration requires high-performance computing infrastructure

"**Curse of dimensionality**" – the number of evaluations required increases exponentially with the number of parameters under calibration.

Highly-irregular and multi-dimensional goodness-of-fit space with many local optima.

Methods

Overall goal is to find input parameters $\mathbf{X_i}$ that best fit reference data and maximize $\mathbf{Y(X_i)}$

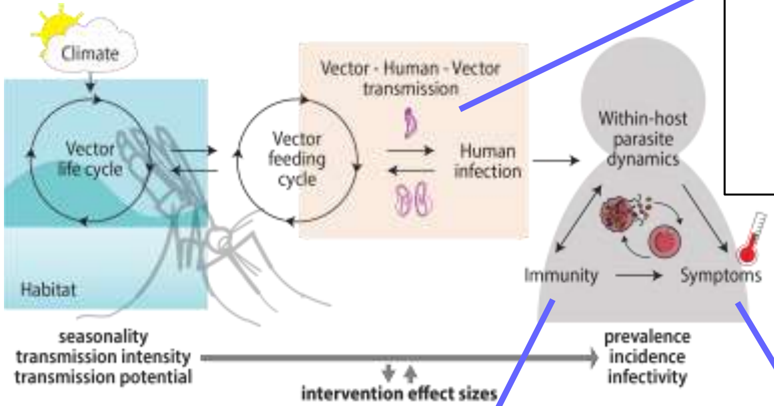
Input Parameters

17 EMOD config parameters related to:

- Immunity
- Symptoms
- Human-Mosquito Transmission

+ 5 different innate immune variation models

$$\mathbf{X_i} = \begin{bmatrix} x_{1,i} \\ \cdots \\ x_{17,i} \end{bmatrix}$$



Gametocyte…
Production Rate
Fraction Male
Stage Survival
Mosquito Stage Survival

Antigen Switch Rate
MSP Variants
MSP Merozoite Kill Fraction
Nonspecific Types
Nonspecific Antibody Growth Rate
Nonspecific Antigenicity Factor
PfEMP1 Variants
Cytokine Gametocyte Inactivation

Max Fever Kill Rate of iRBCs
Pyrogenic Threshold
Erythropoiesis Anemia Effect
RBC Destruction Multiplier
Max Individual Infections

Overall goal is to find input parameters $\mathbf{X_i}$ that best fit reference data and maximize $\mathbf{Y}(\mathbf{X_i})$

Input Parameters

17 EMOD config parameters related to:

- Immunity
- Symptoms
- Human-Mosquito Transmission

$$\mathbf{X_i} = \begin{bmatrix} x_{1,i} \\ \cdots \\ x_{17,i} \end{bmatrix}$$

Types of Data Objectives

y_1 = Incidence vs. Age
y_2 = Prevalence vs. Age
y_3 = Parasite Density Distribution (by age)
y_4 = Gametocyte Density Distribution (by age)
y_5 = Infectiousness vs. Gametocyte Density (by age)

Corresponding to trial data from **8 sites** across 4 countries in Sub-Saharan Africa

Overall goal is to find input parameters $\mathbf{X_i}$ that best fit reference data and maximize $\mathbf{Y}(\mathbf{X_i})$

Input Parameters

17 EMOD config parameters related to:

- Immunity
- Symptoms
- Human-Mosquito Transmission

$$\mathbf{X_i} = \begin{bmatrix} x_{1,i} \\ \cdots \\ x_{17,i} \end{bmatrix}$$

Types of Data Objectives

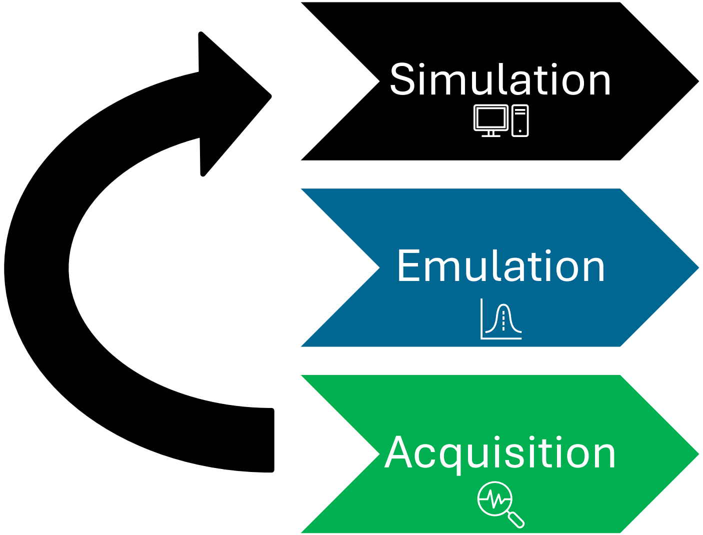
y_1 = Incidence vs. Age
y_2 = Prevalence vs. Age
y_3 = Parasite Density Distribution (by age)
y_4 = Gametocyte Density Distribution (by age)
y_5 = Infectiousness vs. Gametocyte Density (by age)

$across\ sites$:

$$\boxed{\mathbf{Y}(\mathbf{X_i}) = \max\left(\frac{y_{n,\,site,\,i}}{y_{n,\,site,\,default}}\right)}$$

Bayesian Optimization with Gaussian Processes

Based on *Reiker et al. (Nat comms 2021)* with support from Melissa Penny and Aurélien Cavelan*



Bayesian Optimization with Gaussian Processes

Based on *Reiker et al. (Nat comms 2021)* with support from Melissa Penny and Aurélien Cavelan*

Simulations are tuned to match conditions at each reference site:

- Seasonality and transmission intensity
- Antimalarial treatment
- Diagnostics

Bayesian Optimization with Gaussian Processes

Based on *Reiker et al. (Nat comms 2021)* with support from Melissa Penny and Aurélien Cavelan*

Simulations are tuned to match conditions at each reference site:

- Seasonality and transmission intensity
- Antimalarial treatment
- Diagnostics

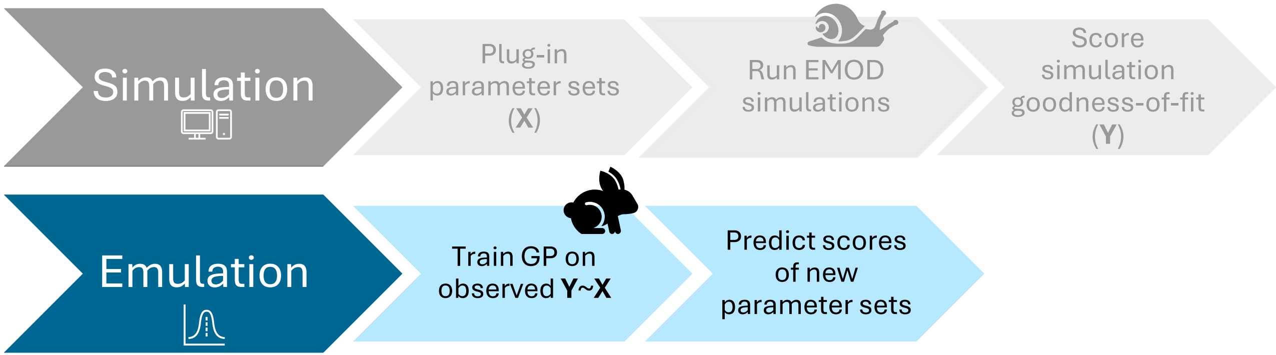
Scoring: **likelihood of observing simulation outputs given the reference data**

- Per site-specific objective, relative to baseline parameterization
- **Y** for simulated parameter set is the *maximum* site-specific objective score

(better < 1 < worse)

Bayesian Optimization with Gaussian Processes

Based on *Reiker et al. (Nat comms 2021)* with support from Melissa Penny and Aurélien Cavelan*

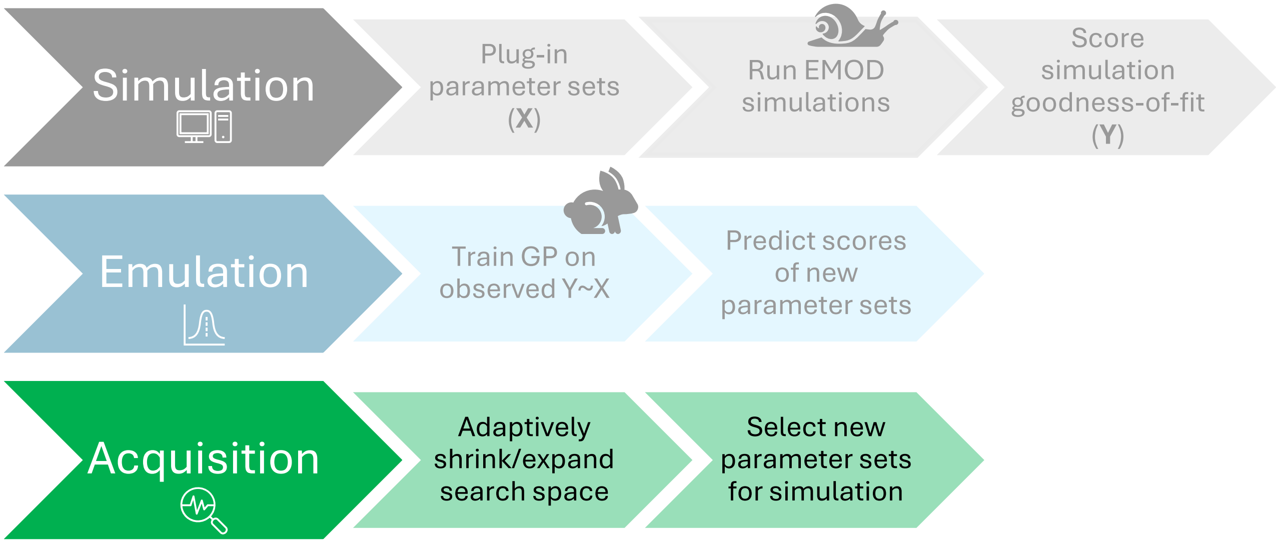


The emulator is faster than running EMOD (minutes vs. hours)

- 5,000 candidates >>> 100 simulations

Bayesian Optimization with Gaussian Processes

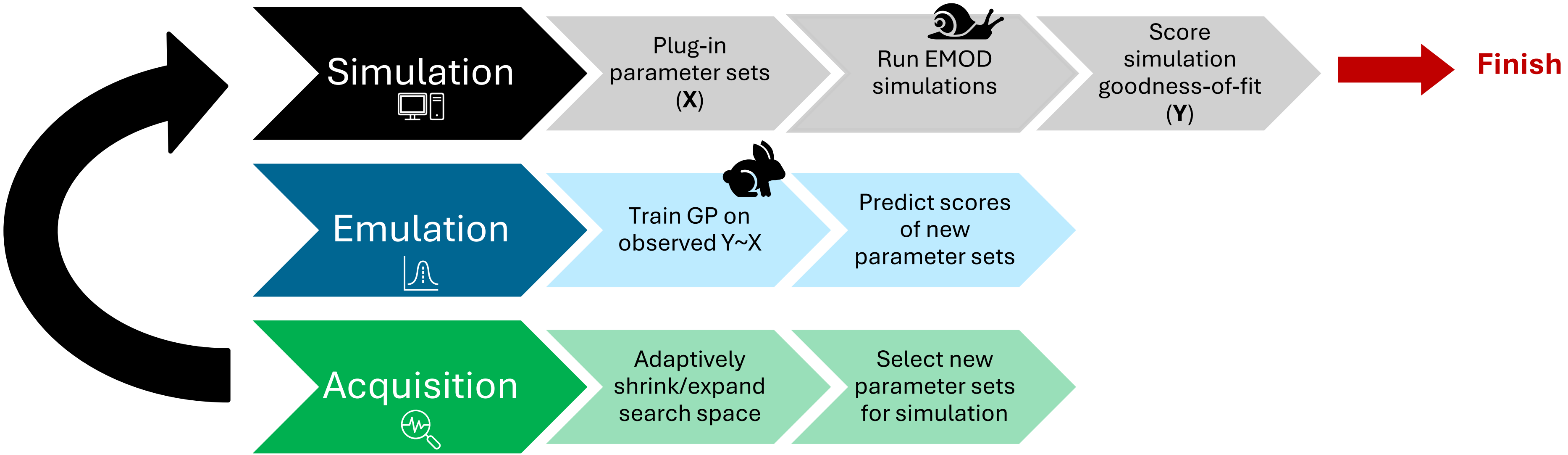
Based on *Reiker et al. (Nat comms 2021)* with support from Melissa Penny and Aurélien Cavelan*



Trust Region-Based Thompson Sampling balances exploration against exploitation to select new samples.

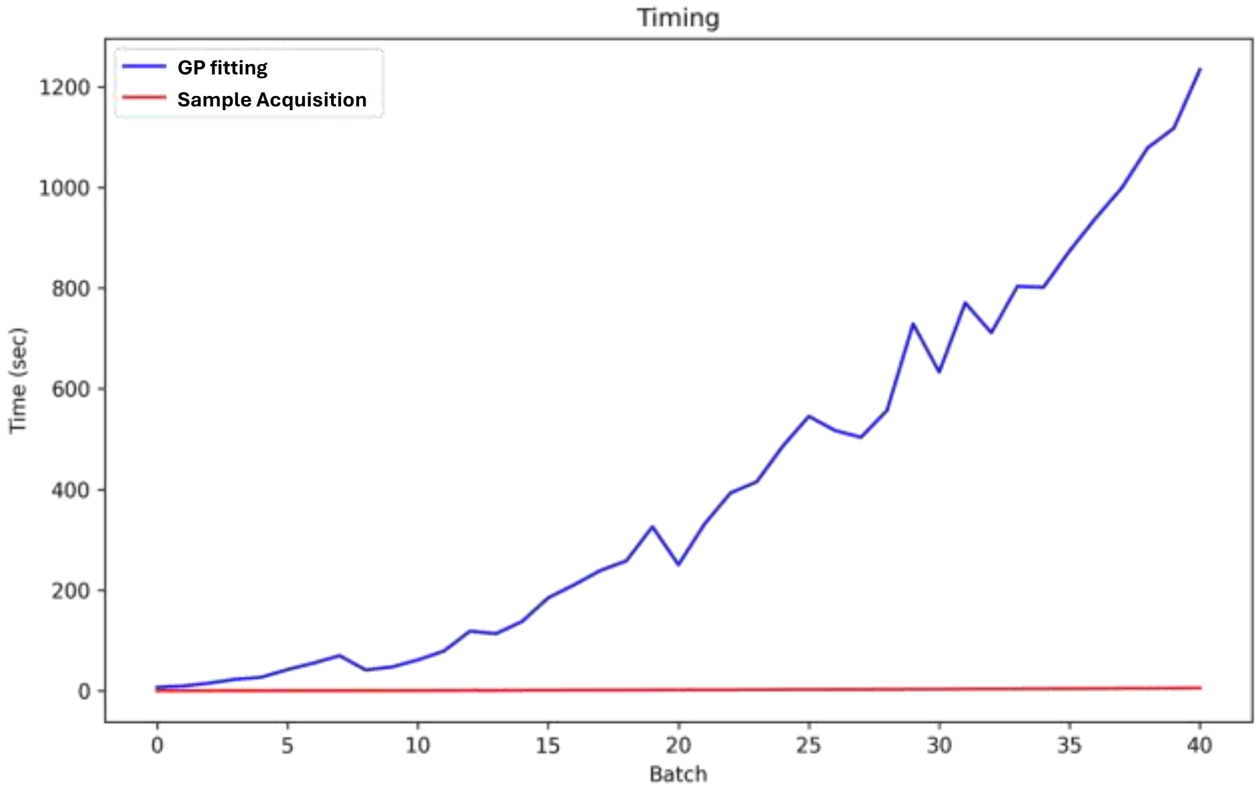
Bayesian Optimization with Gaussian Processes

Based on *Reiker et al. (Nat comms 2021)* with support from Melissa Penny and Aurélien Cavelan*



The cycle of simulation, emulation, and acquisition repeats for 40 cycles, until 5,000 locations in **X** are simulated and scored

Optimization steps are much faster than EMOD simulations, and contribute little to overall runtime

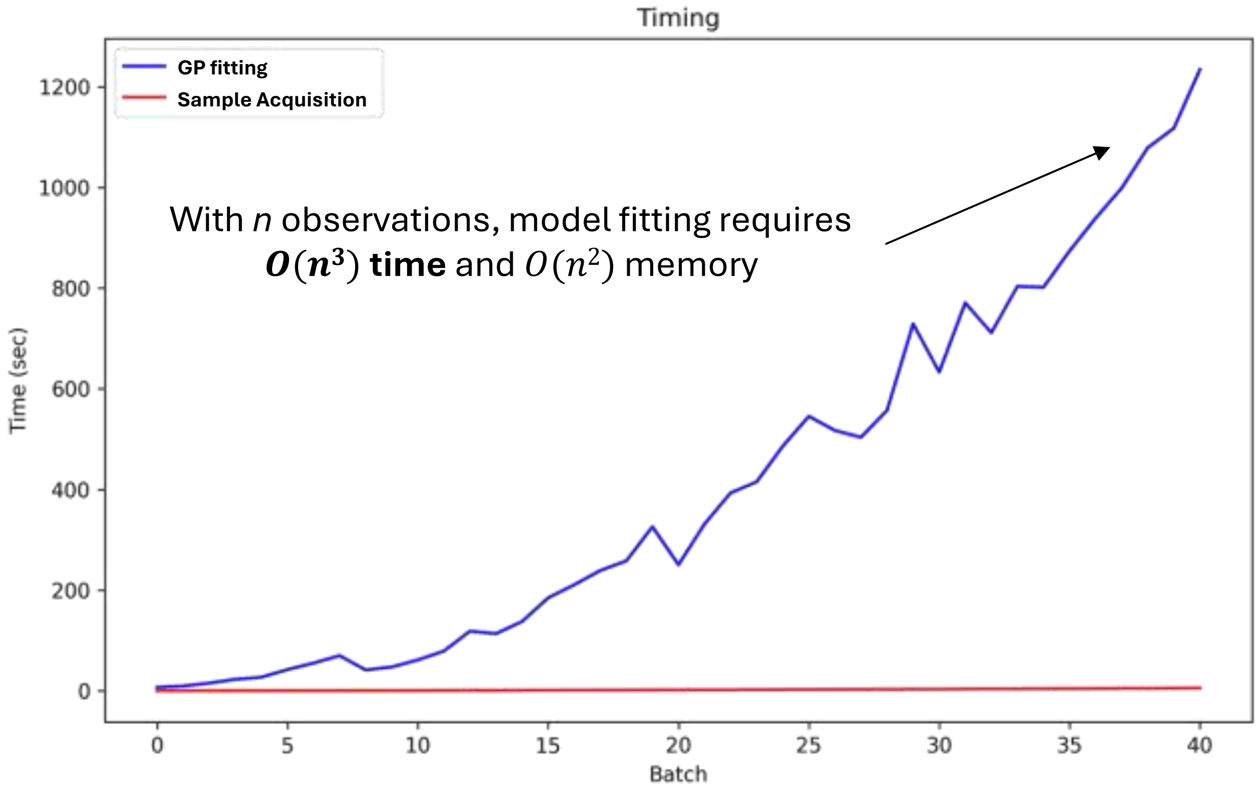


Initial EMOD Simulations
~ 5-10 hours per 100 samples

Later EMOD Simulations
~ 2 hours per 100 samples

Emulator **< *20 minutes***
per 5,000 candidates

Optimization steps are much faster than EMOD simulations, and contribute little to overall runtime

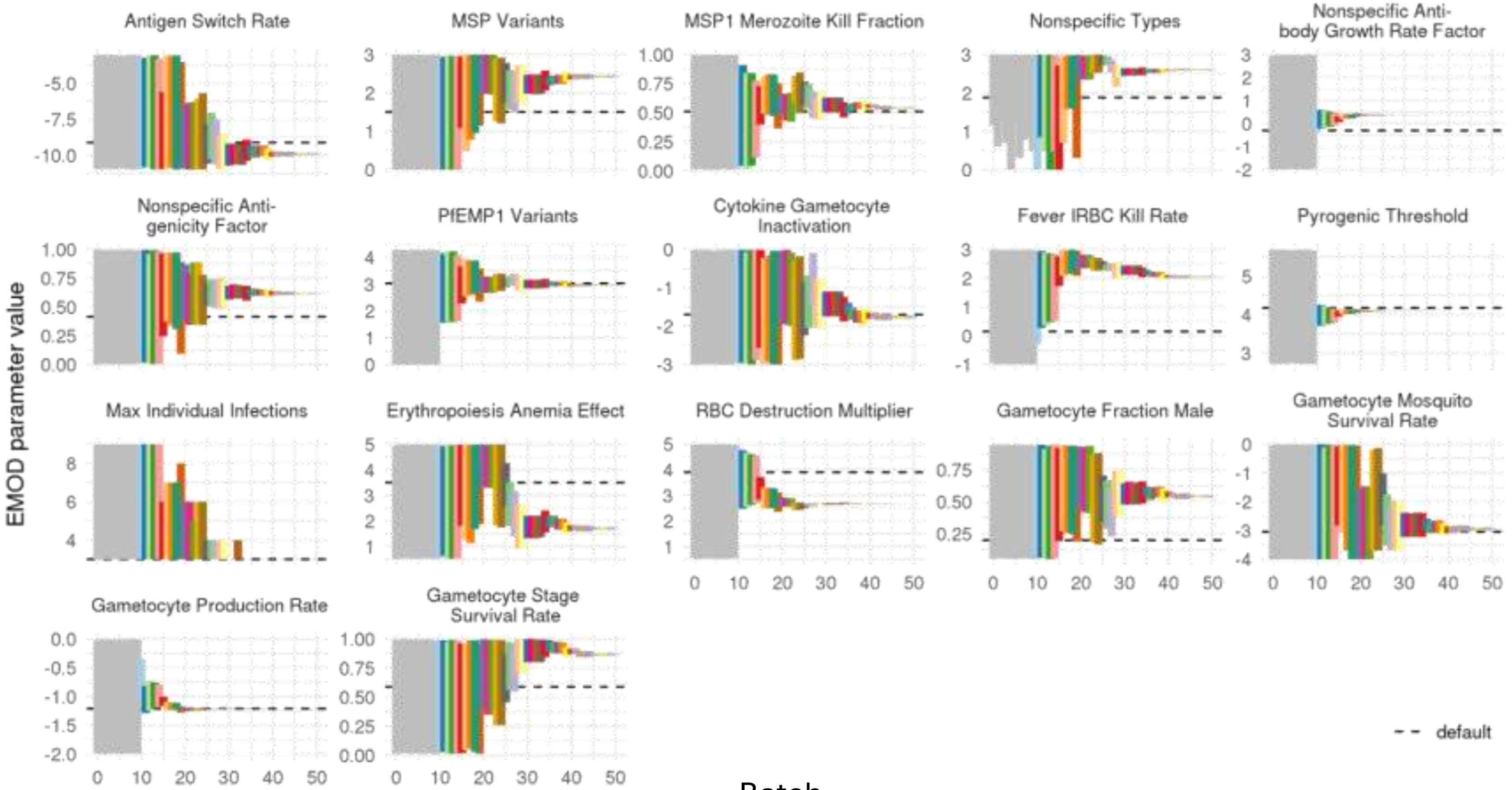


Initial EMOD Simulations
~ 5-10 hours per 100 samples

Later EMOD Simulations
~ 2 hours per 100 samples

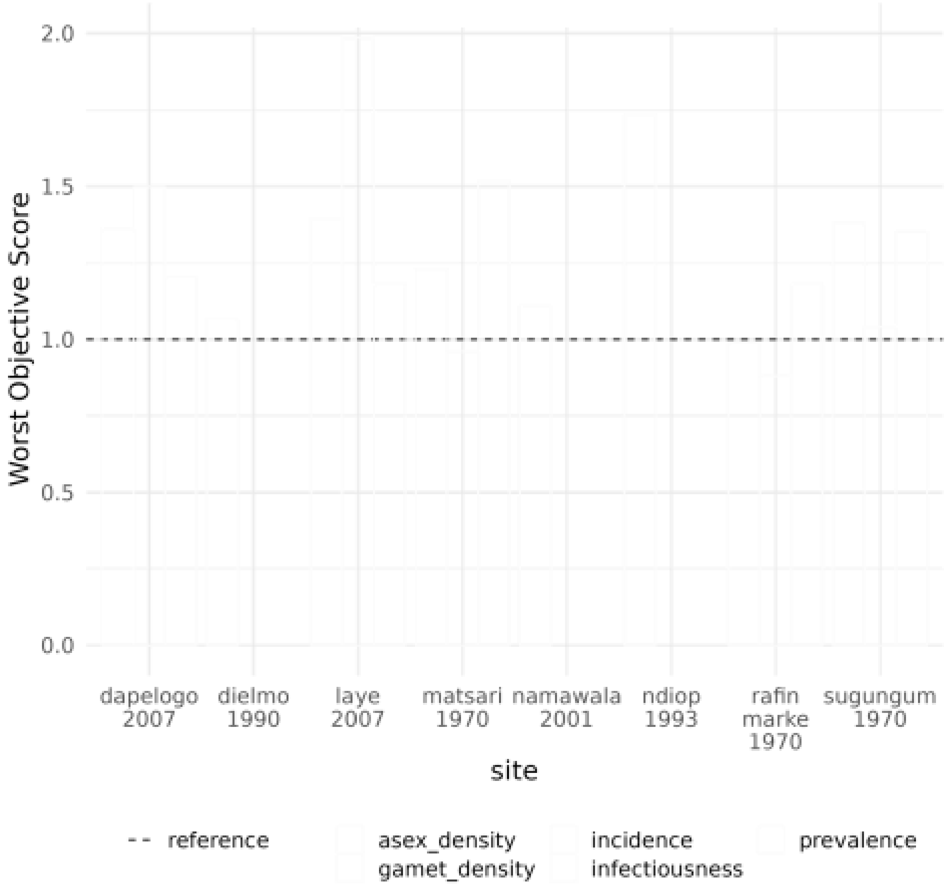
Emulator **< 20 *minutes***
per 5,000 candidates

Parameter search converges on region of best fit



Batch

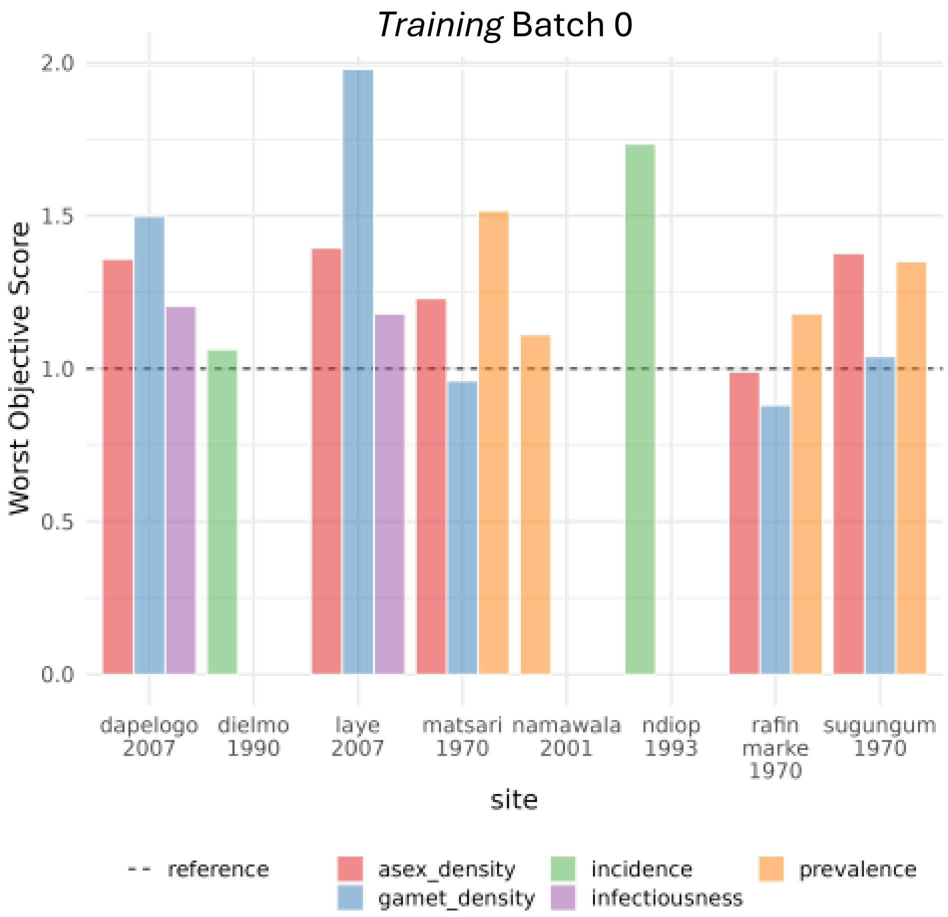
Default model parameterization fit to data



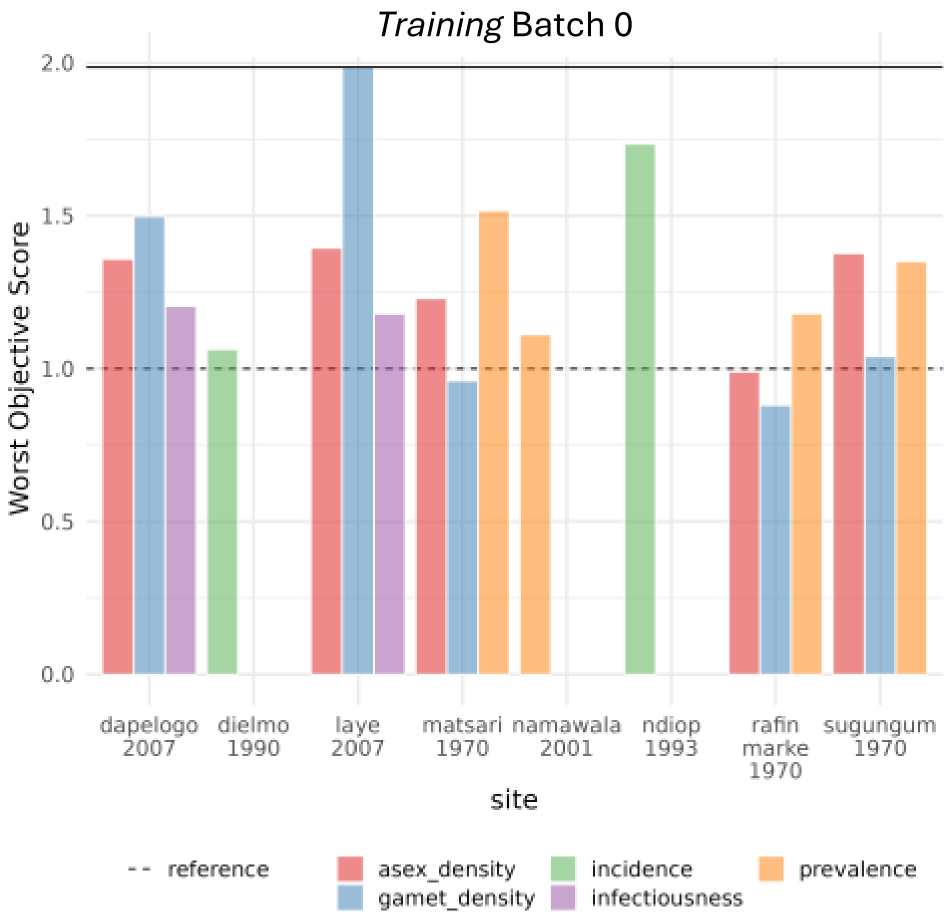
$Y_{baseline}$

Default model parameterization fit to data

Improvement in goodness-of-fit over batches

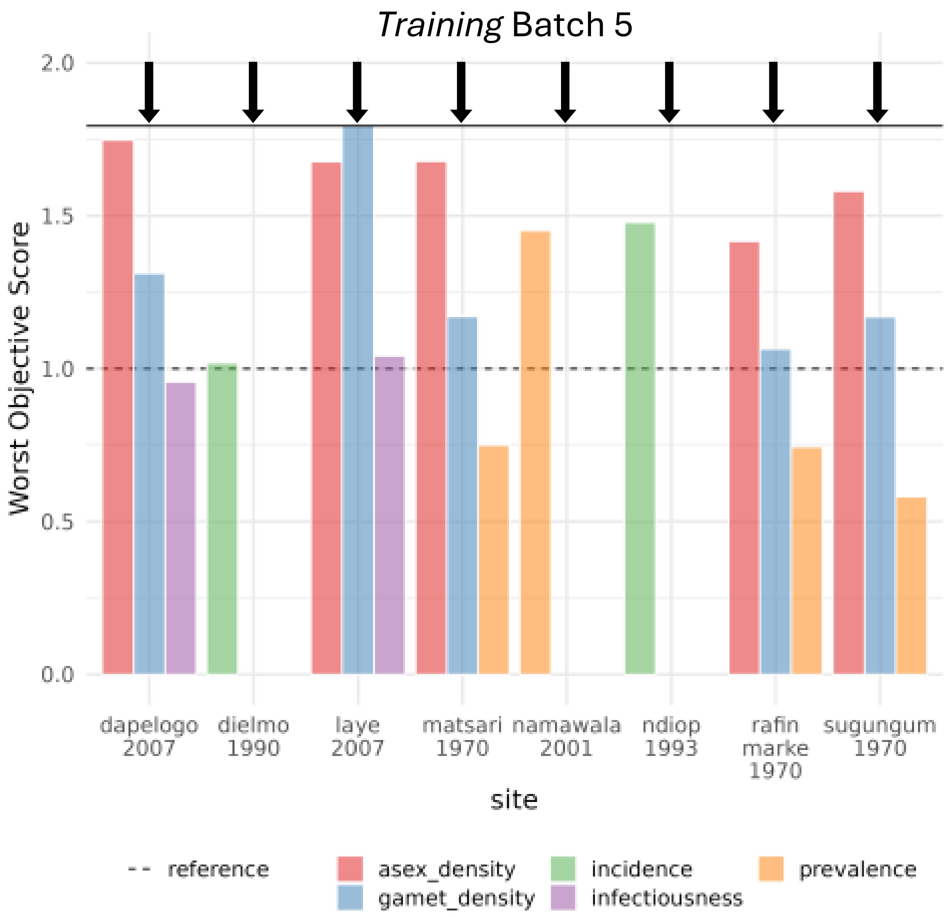


Improvement in goodness-of-fit over batches

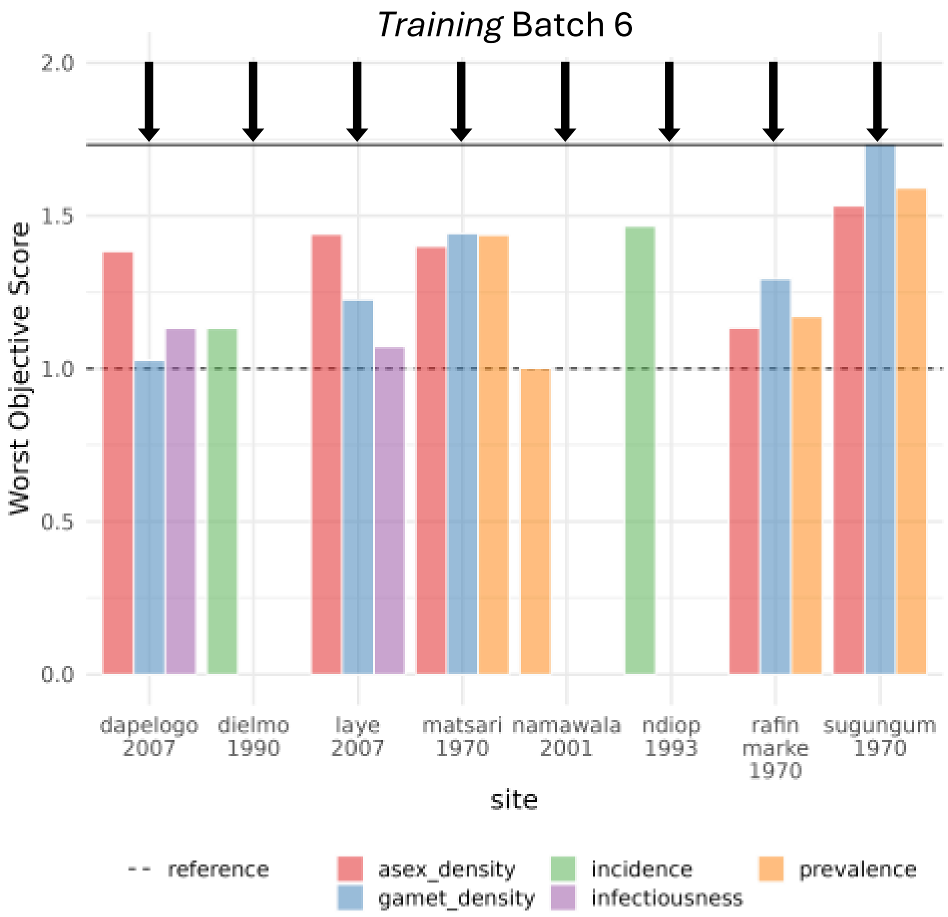


Score, Y

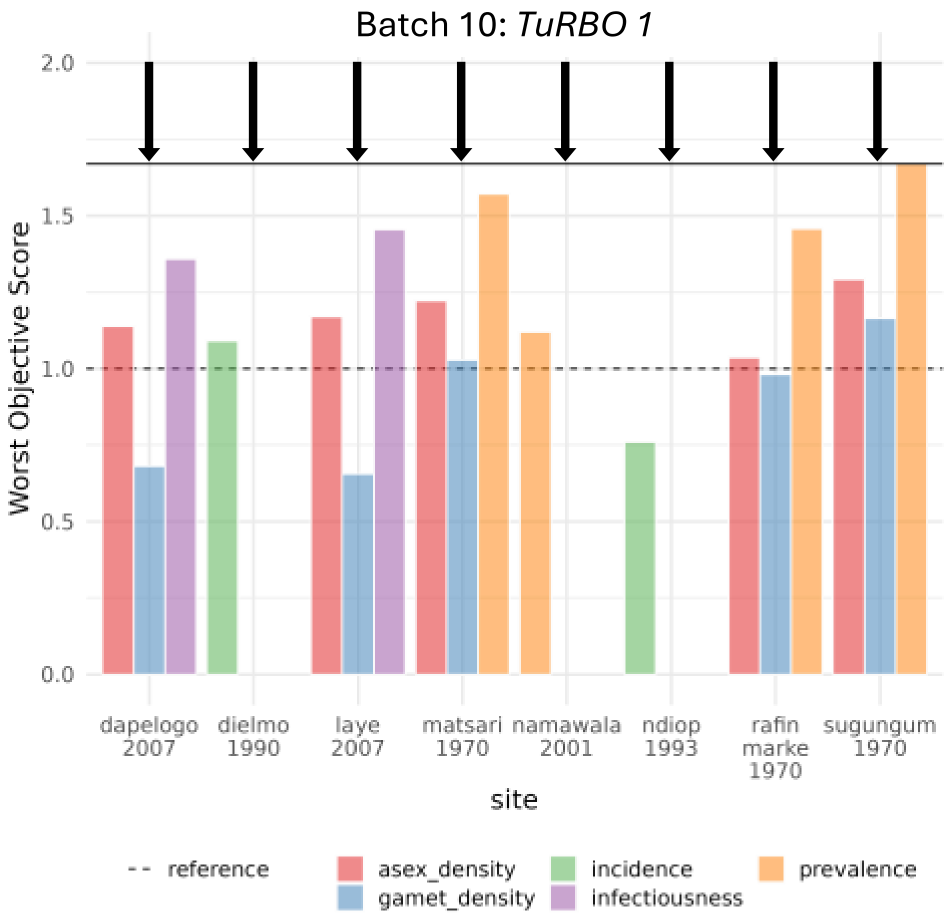
Improvement in goodness-of-fit over batches



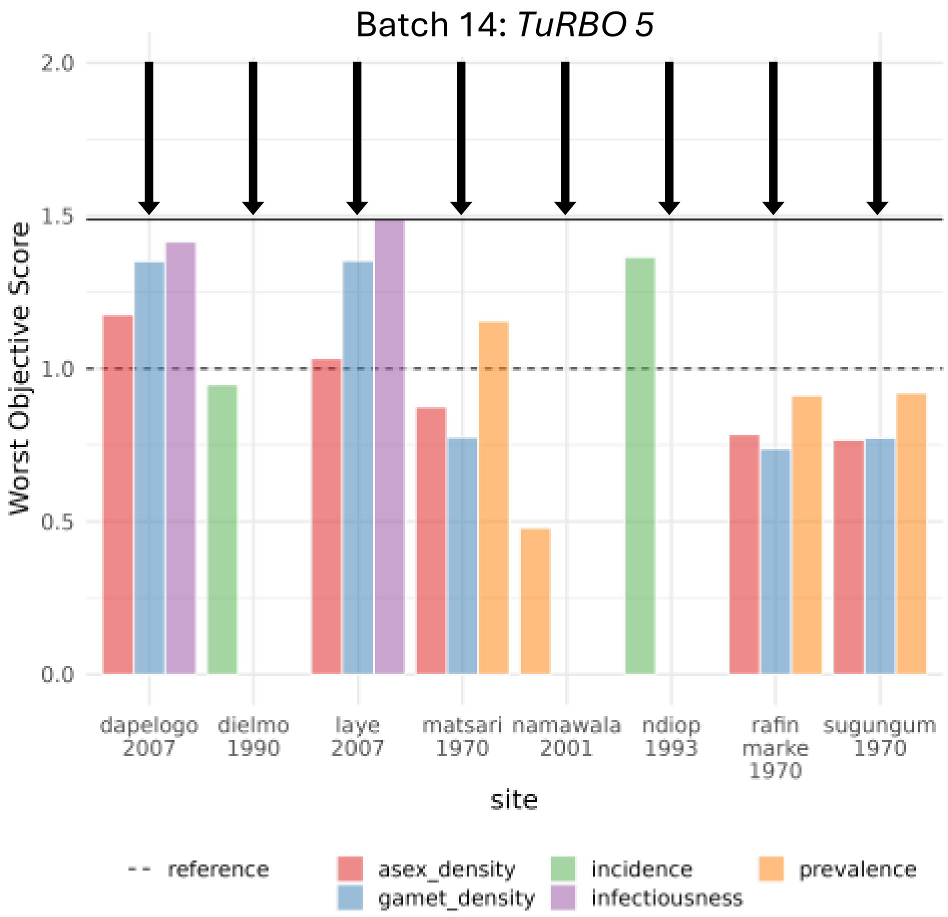
Improvement in goodness-of-fit over batches



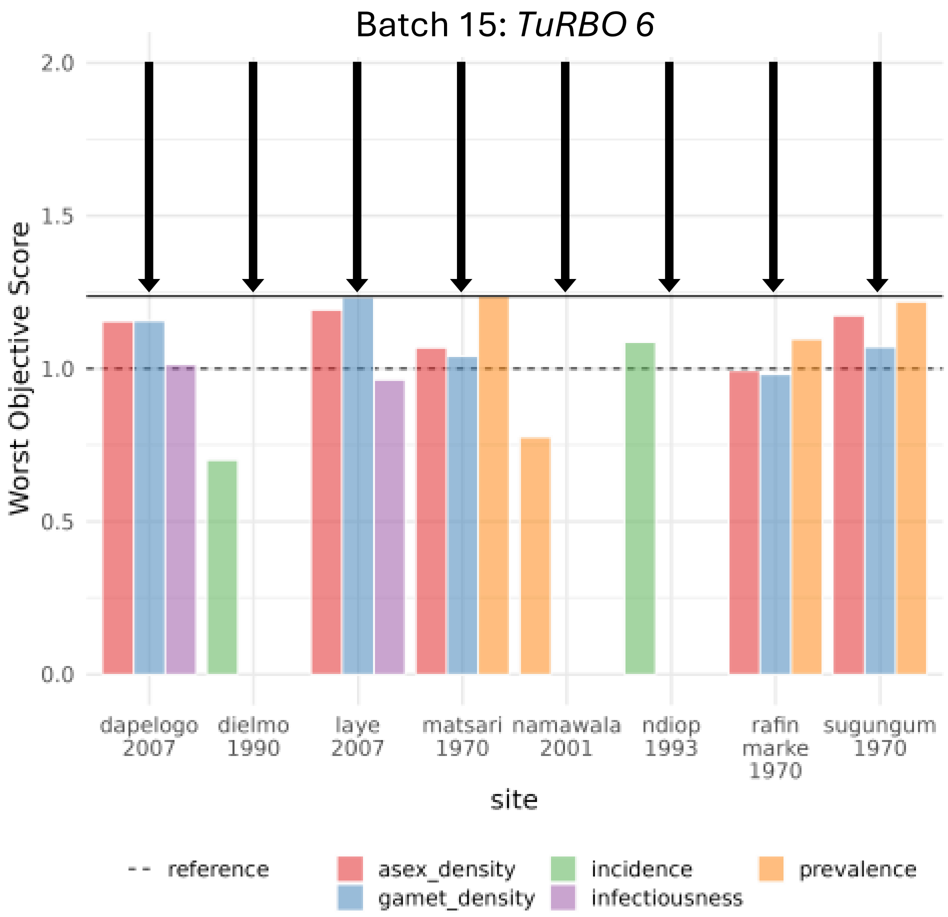
Improvement in goodness-of-fit over batches



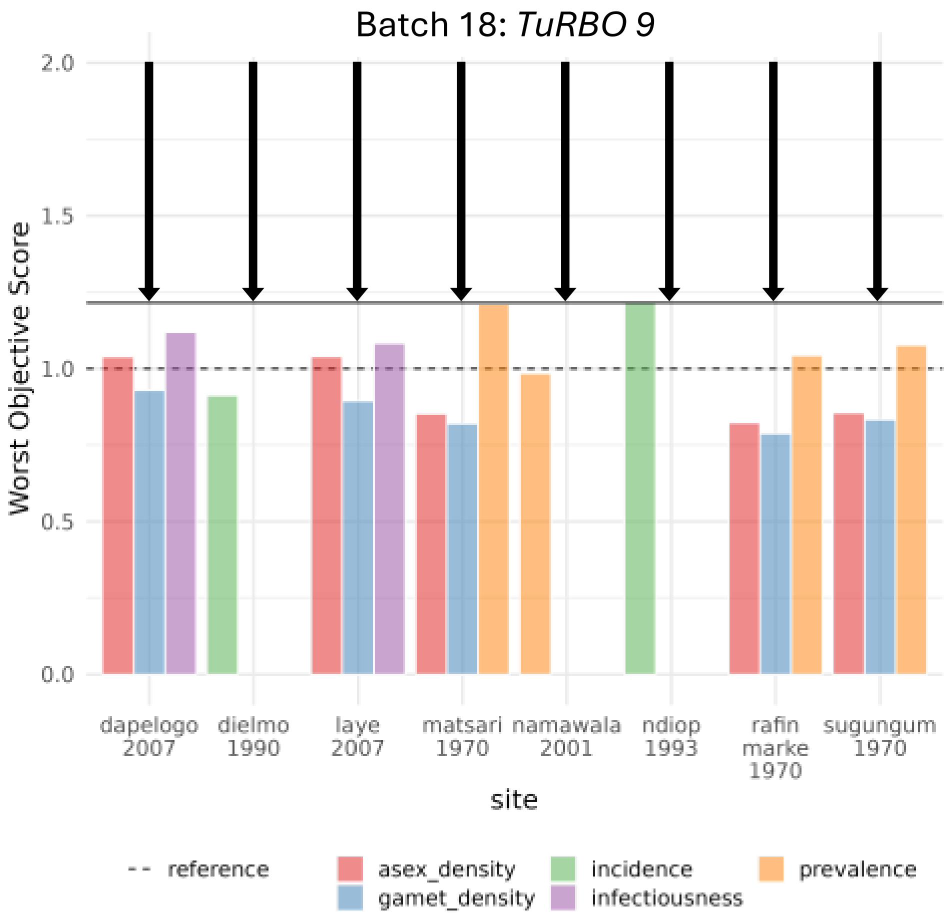
Improvement in goodness-of-fit over batches



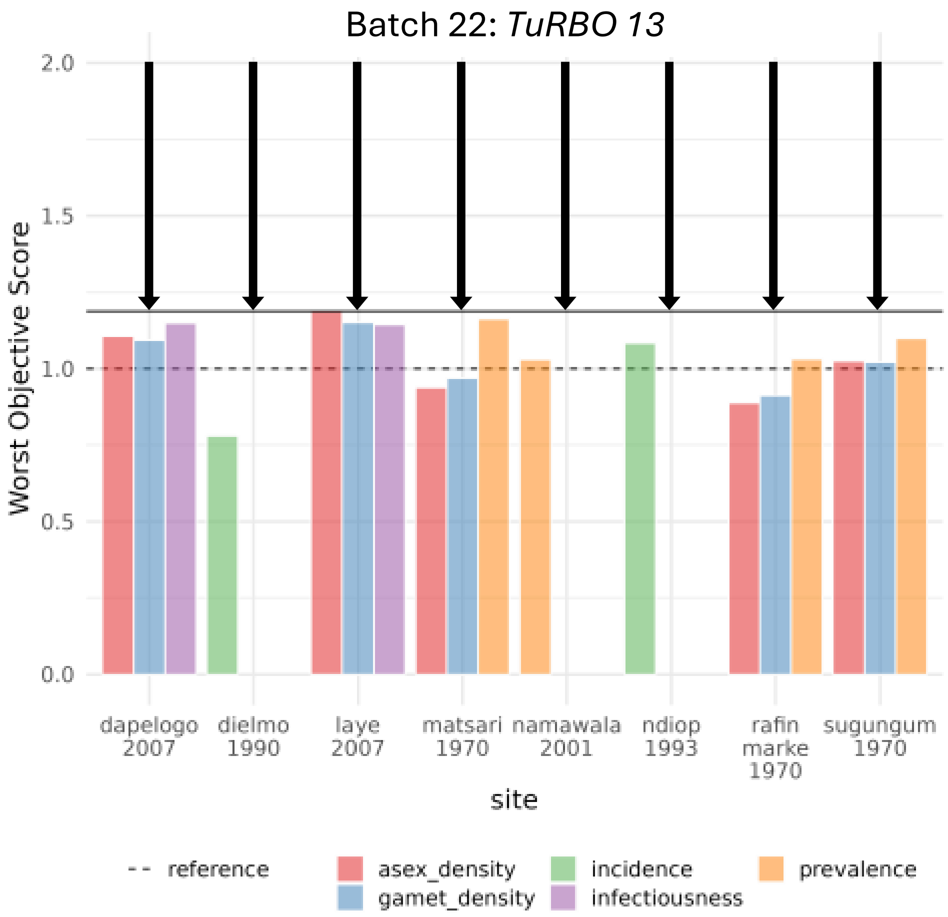
Improvement in goodness-of-fit over batches



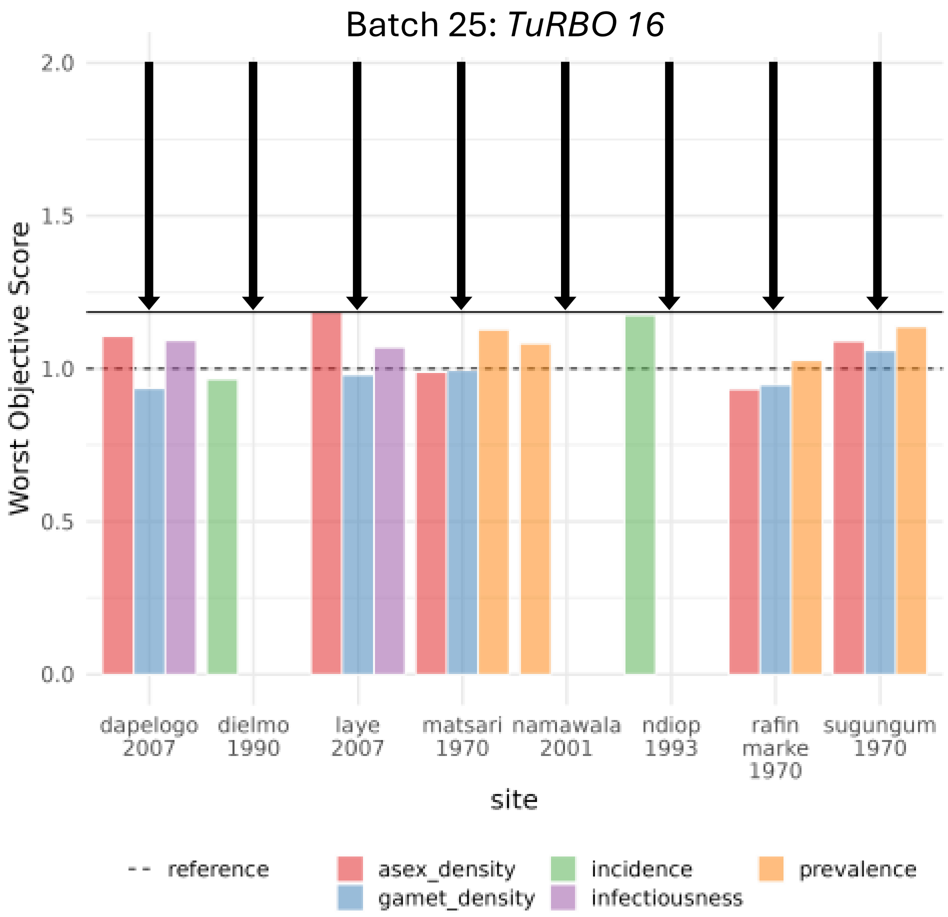
Improvement in goodness-of-fit over batches



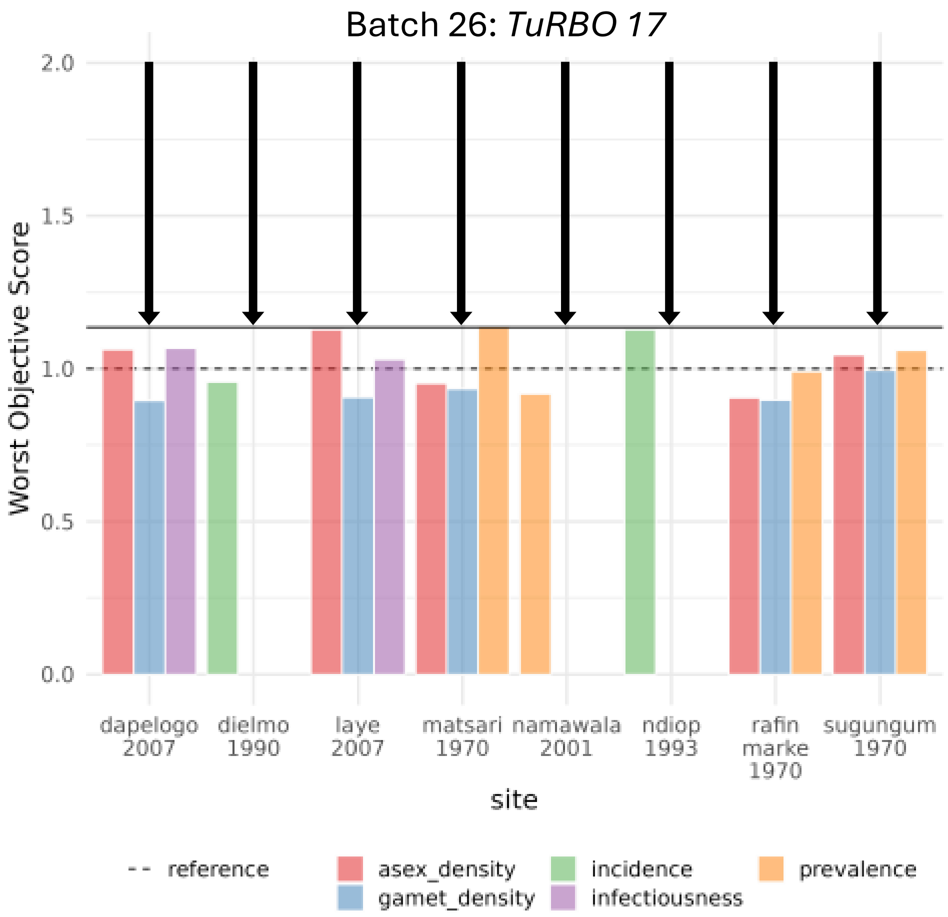
Improvement in goodness-of-fit over batches



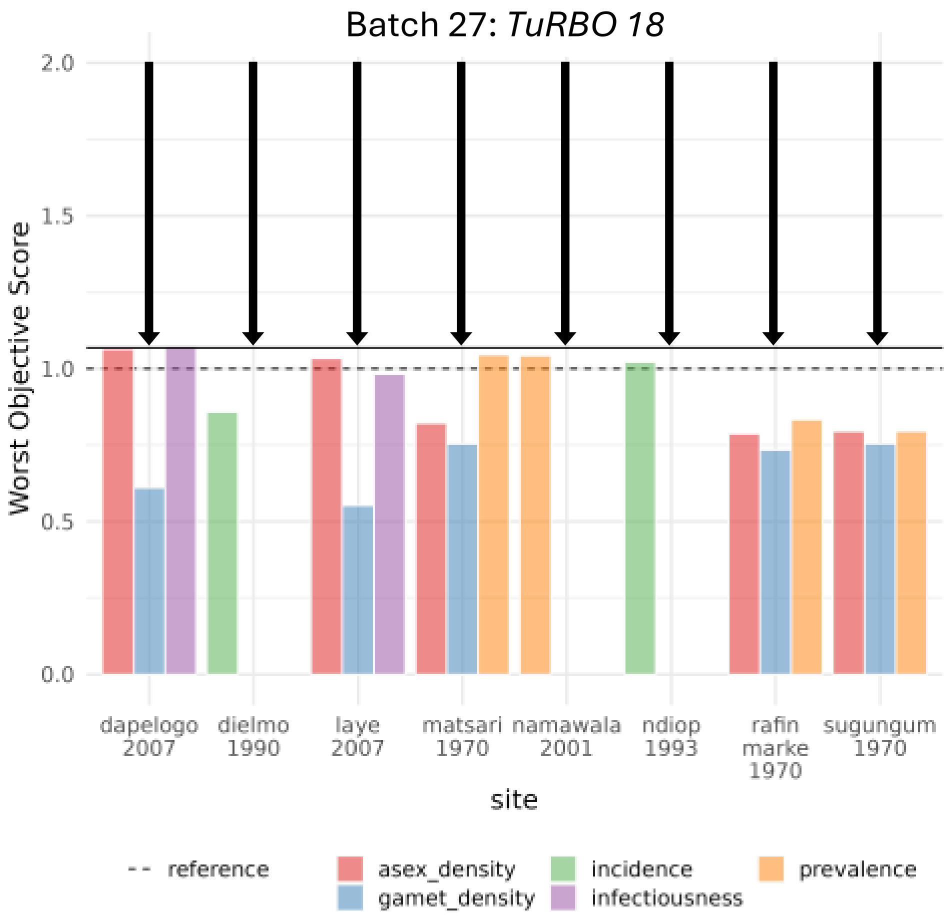
Improvement in goodness-of-fit over batches



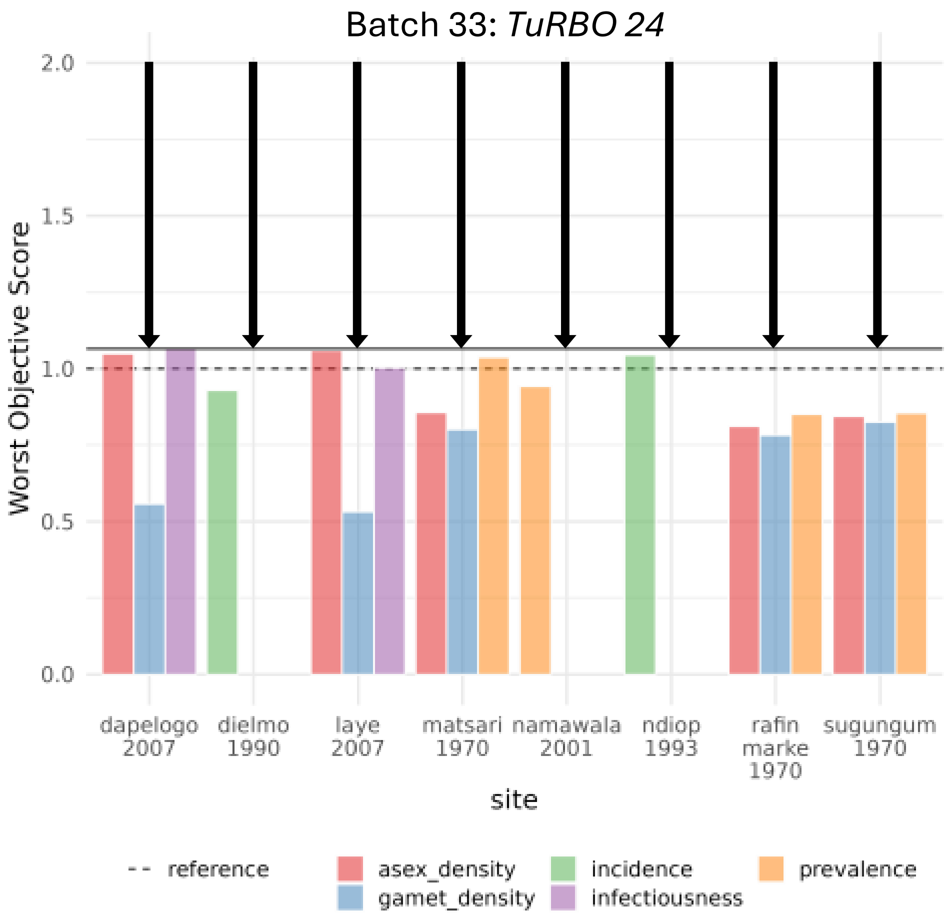
Improvement in goodness-of-fit over batches



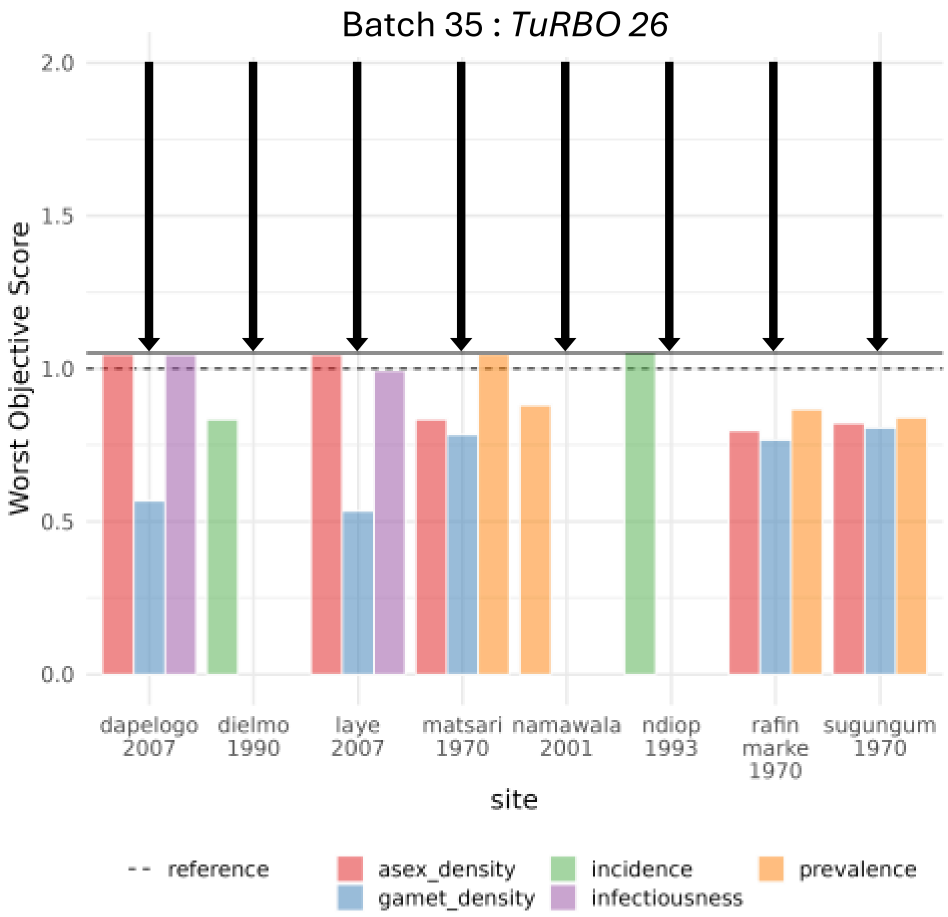
Improvement in goodness-of-fit over batches



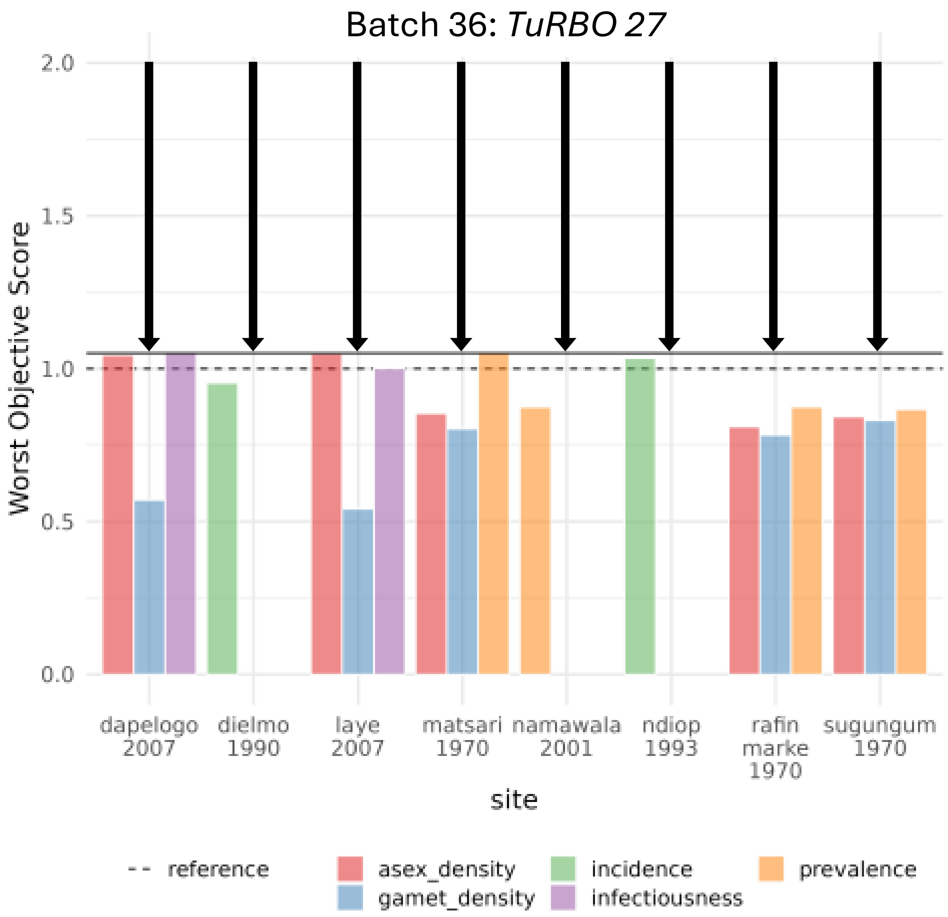
Improvement in goodness-of-fit over batches



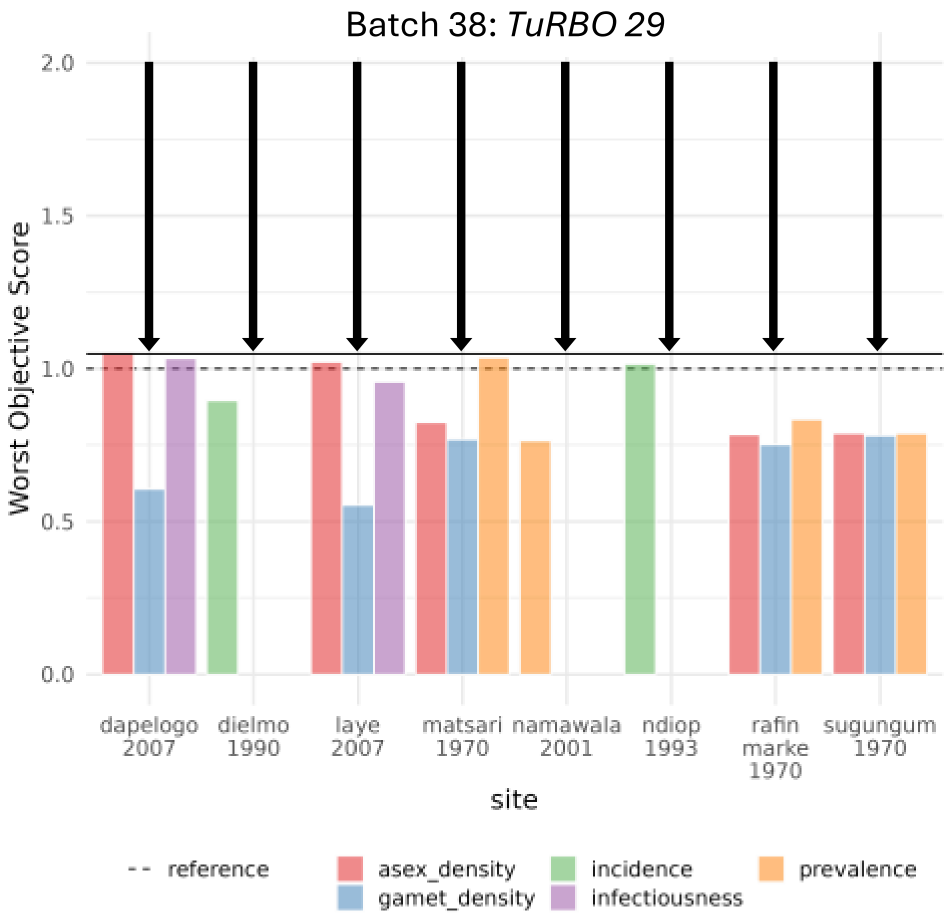
Improvement in goodness-of-fit over batches



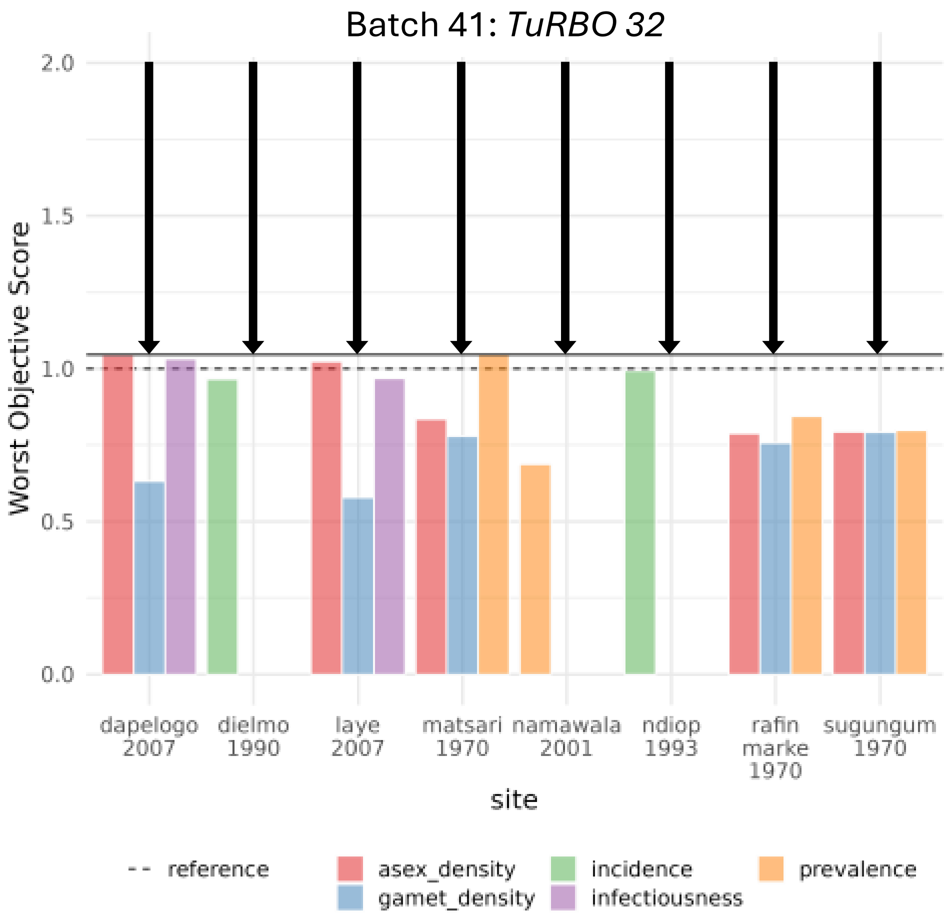
Improvement in goodness-of-fit over batches



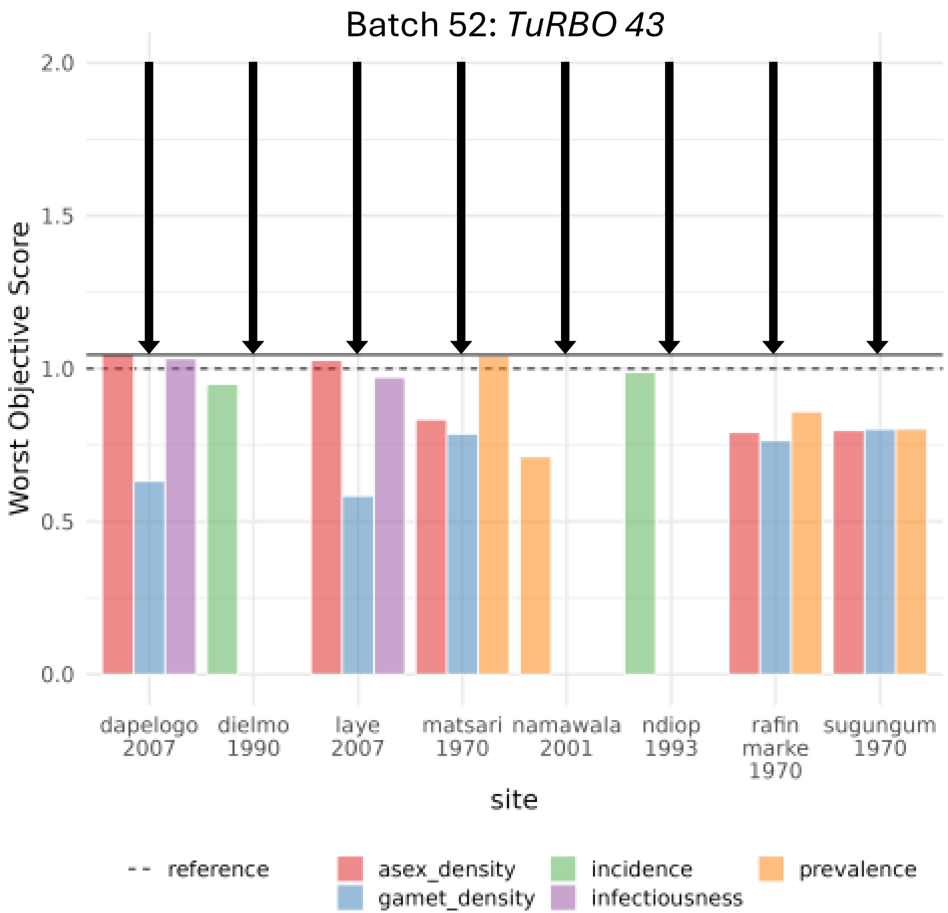
Improvement in goodness-of-fit over batches



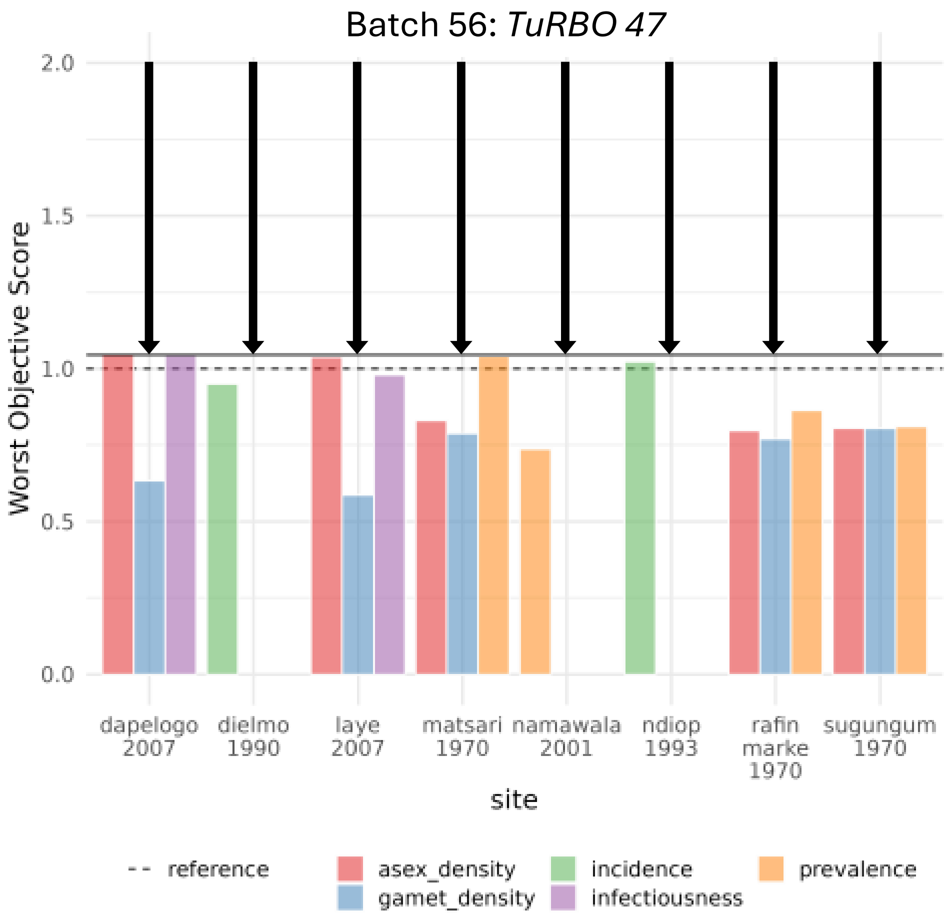
Improvement in goodness-of-fit over batches



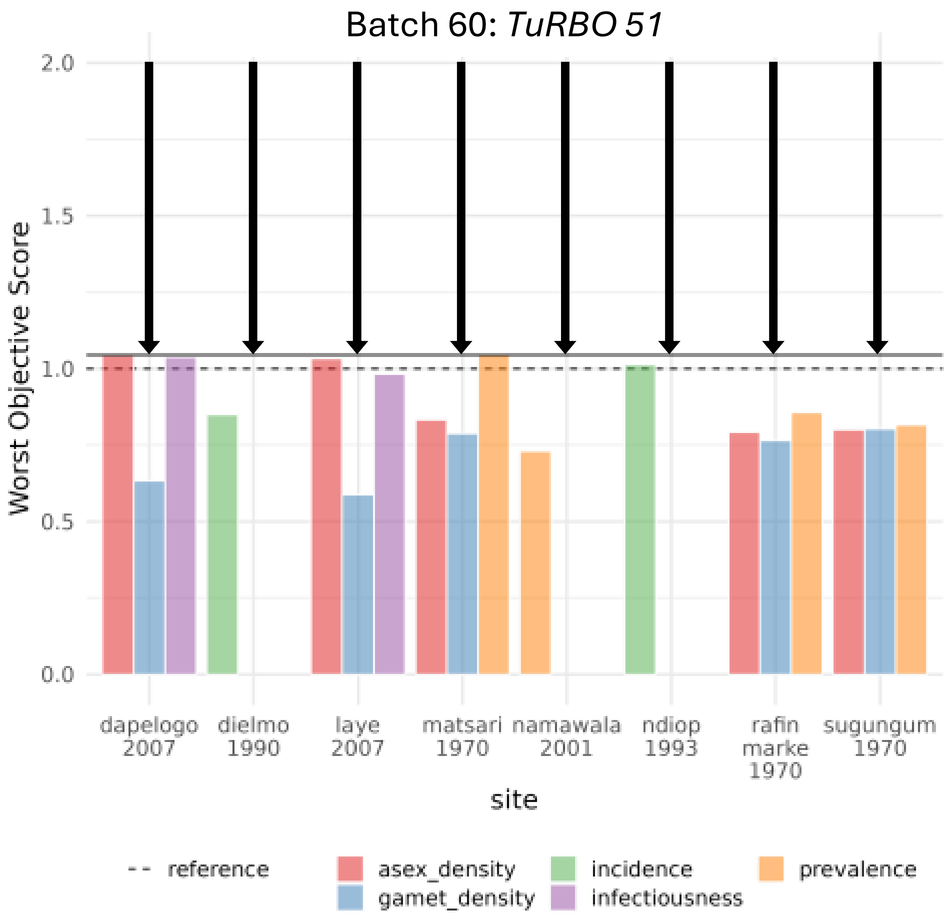
Improvement in goodness-of-fit over batches



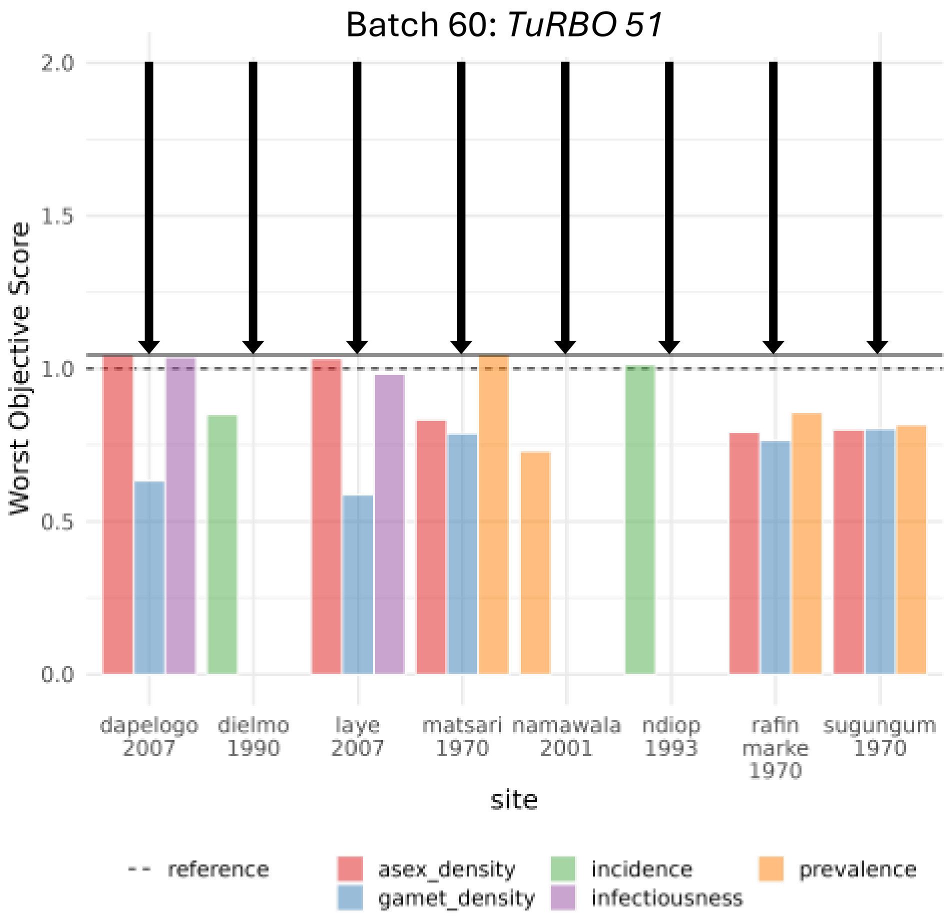
Improvement in goodness-of-fit over batches



Improvement in goodness-of-fit over batches



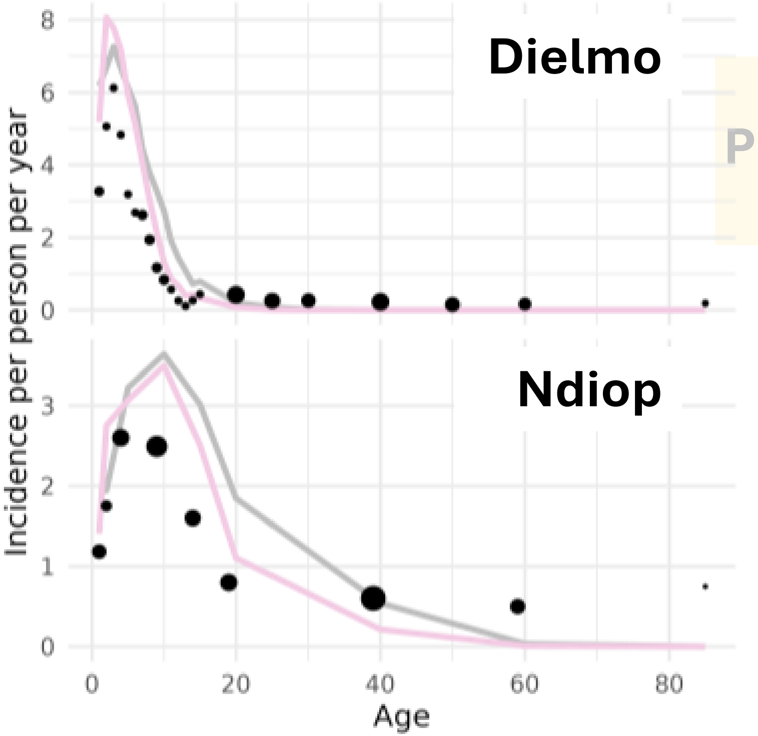
No new parameter set has improved on all objectives



Maybe due to:

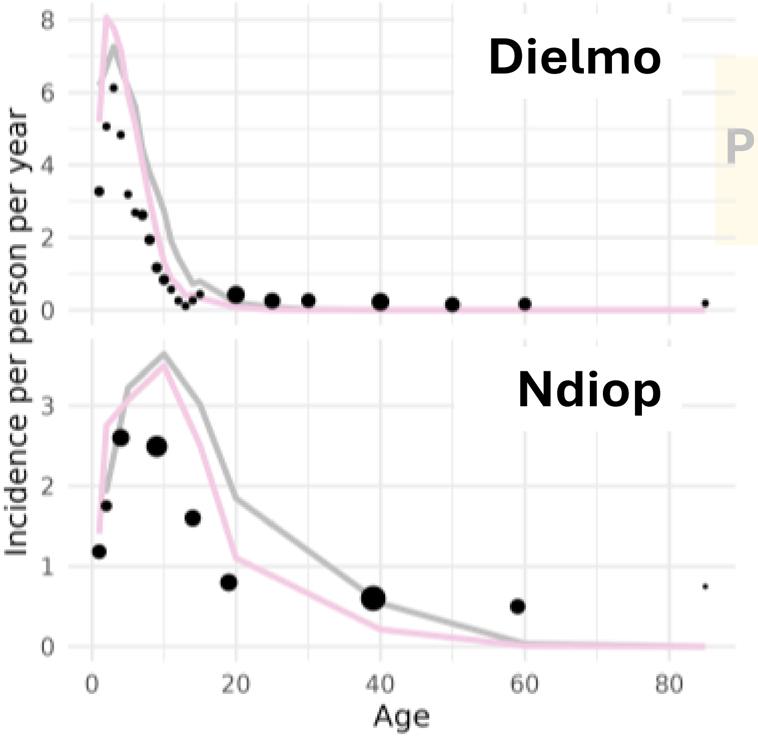
- TuRBO search
- Scoring method
- Baseline performance

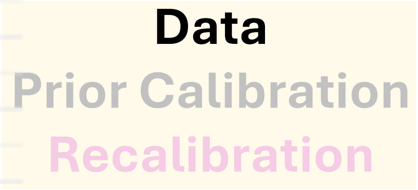
Recalibrated vs. Prior Calibrated Parameter Set

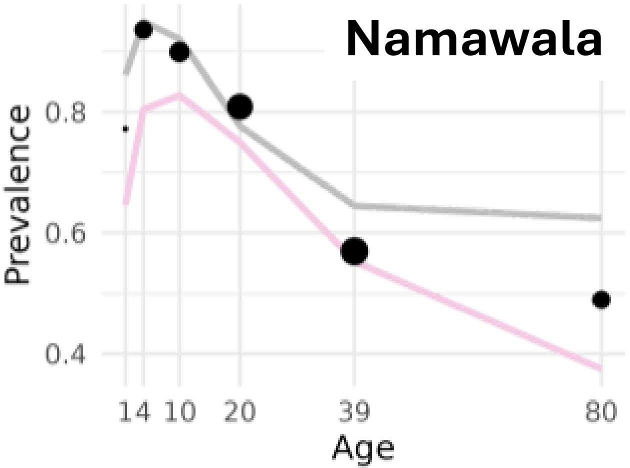


Data
Prior Calibration
Recalibration

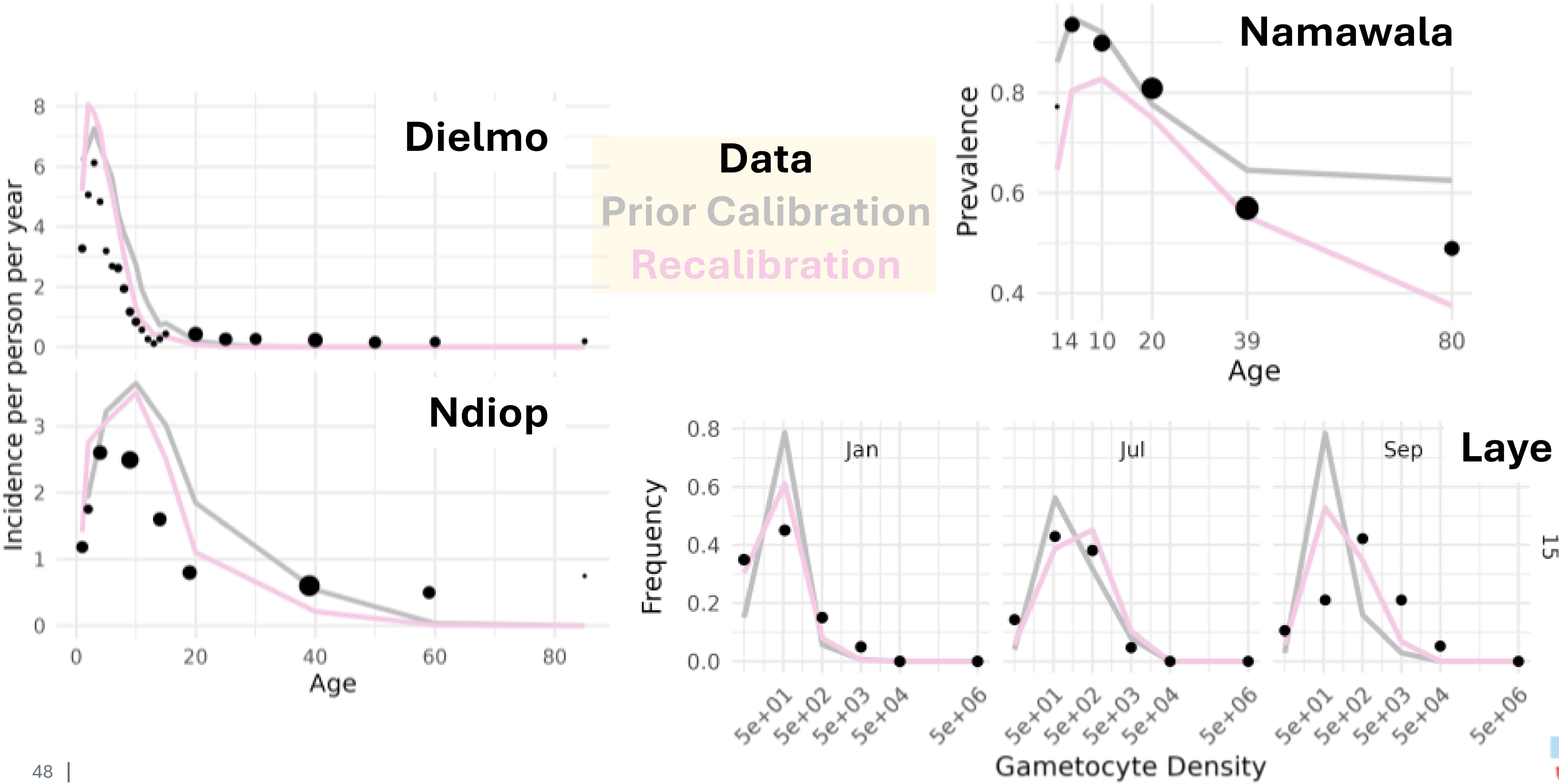
Recalibrated vs. Prior Calibrated Parameter Set





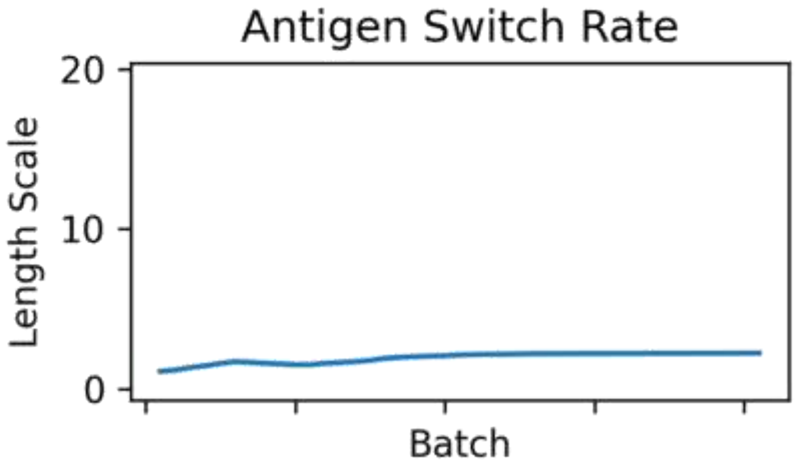


Recalibrated vs. Prior Calibrated Parameter Set

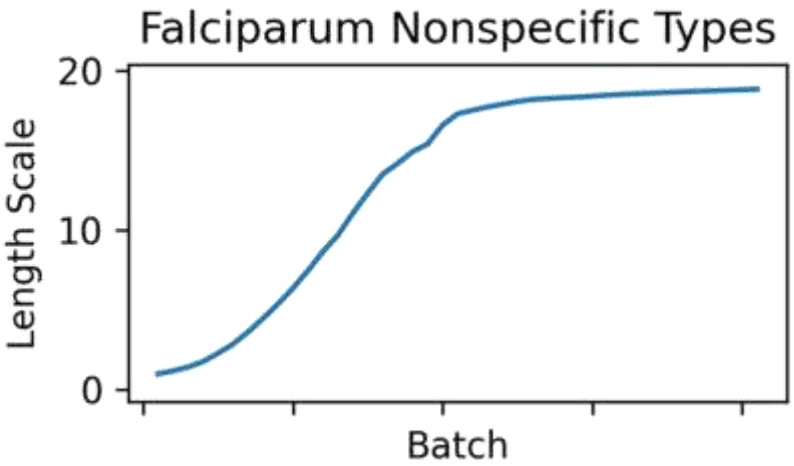


15

The length scale GP hyperparameter describes the correlation between scores over stretches of parameter space

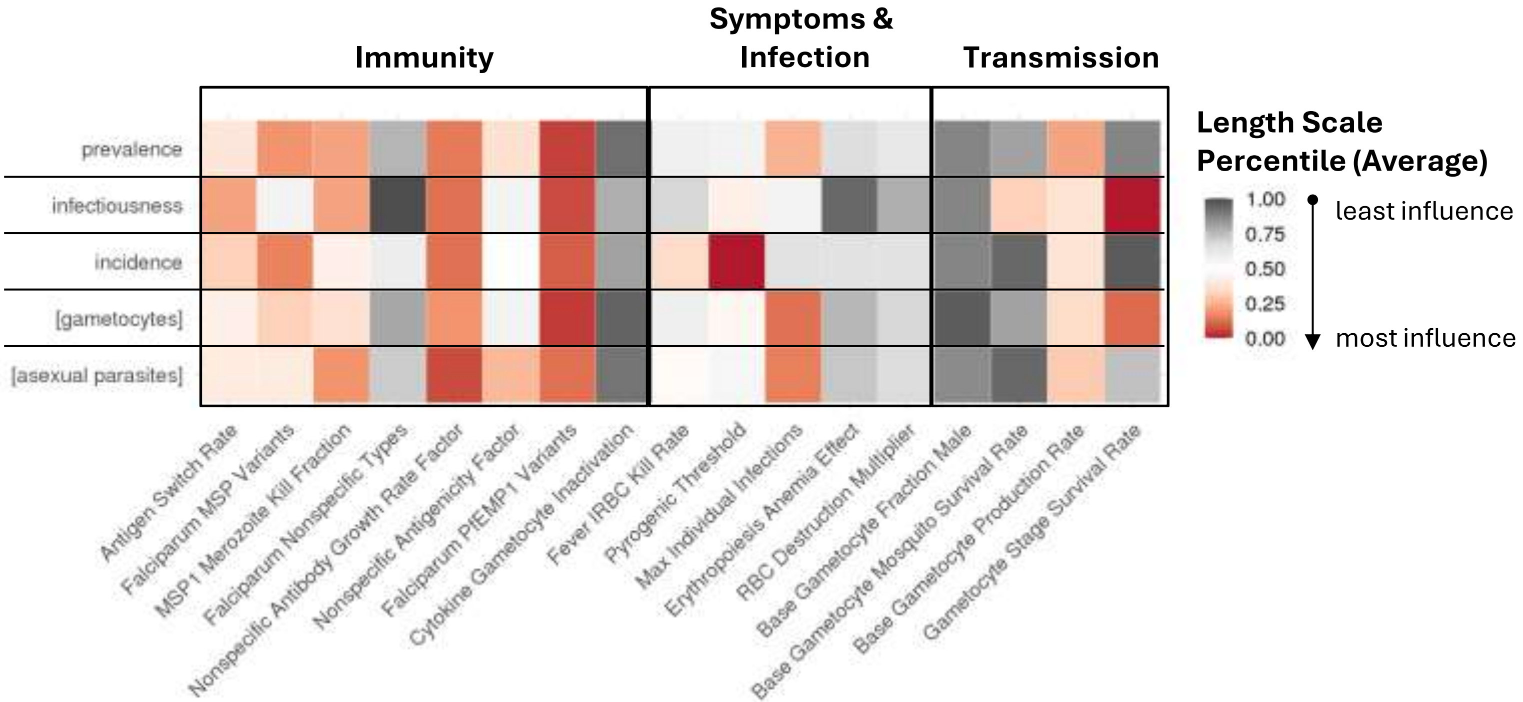


Short length scale = Strong influence
traveling a short "distance" in parameter space results in drastically different scores



Long length scale = Weak influence
capable of extrapolating scores across long "distances" in parameter space

Length scales per-objective show specific parameter influence



Next Steps

Validating and Extending Framework

- Assess recalibrated model performance against datasets not used for fitting
  - Peak parasitemia, gametocytemia, and duration of naive infections (*malariatherapy, 1940-1963*)
  - Severe disease (*The Gambia and Kenya, 1990-1996*)
  - Prevalence, densities, and infectiousness by DMFA (*Sapone, Burkina Faso, 2018-2020*)

Validating and Extending Framework

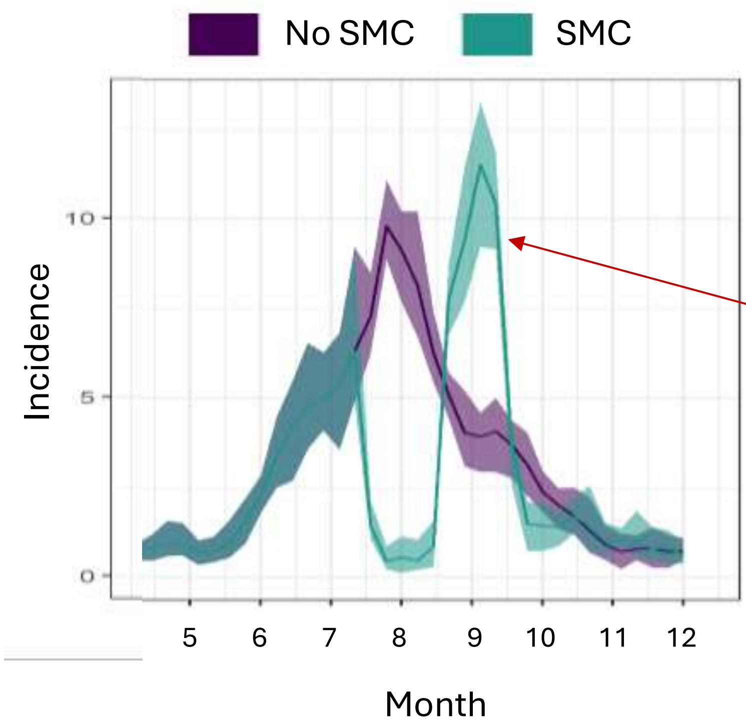
- Assess recalibrated model performance against datasets not used for fitting
  - Peak parasitemia, gametocytemia, and duration of naive infections (*malariatherapy, 1940-1963*)
  - Severe disease (*The Gambia and Kenya*, *1990-1996*)
  - Prevalence, densities, and infectiousness by DMFA (*Sapone, Burkina Faso, 2018-2020*)

- Repeat calibration for alternative innate immune models

Synchronized innate immune responses may drive unexpected model behavior

SMC = Seasonal Malaria Chemoprevention

Mass drug campaign for children



When modeled without variation in innate immune parameters, paradoxical rebound is observed

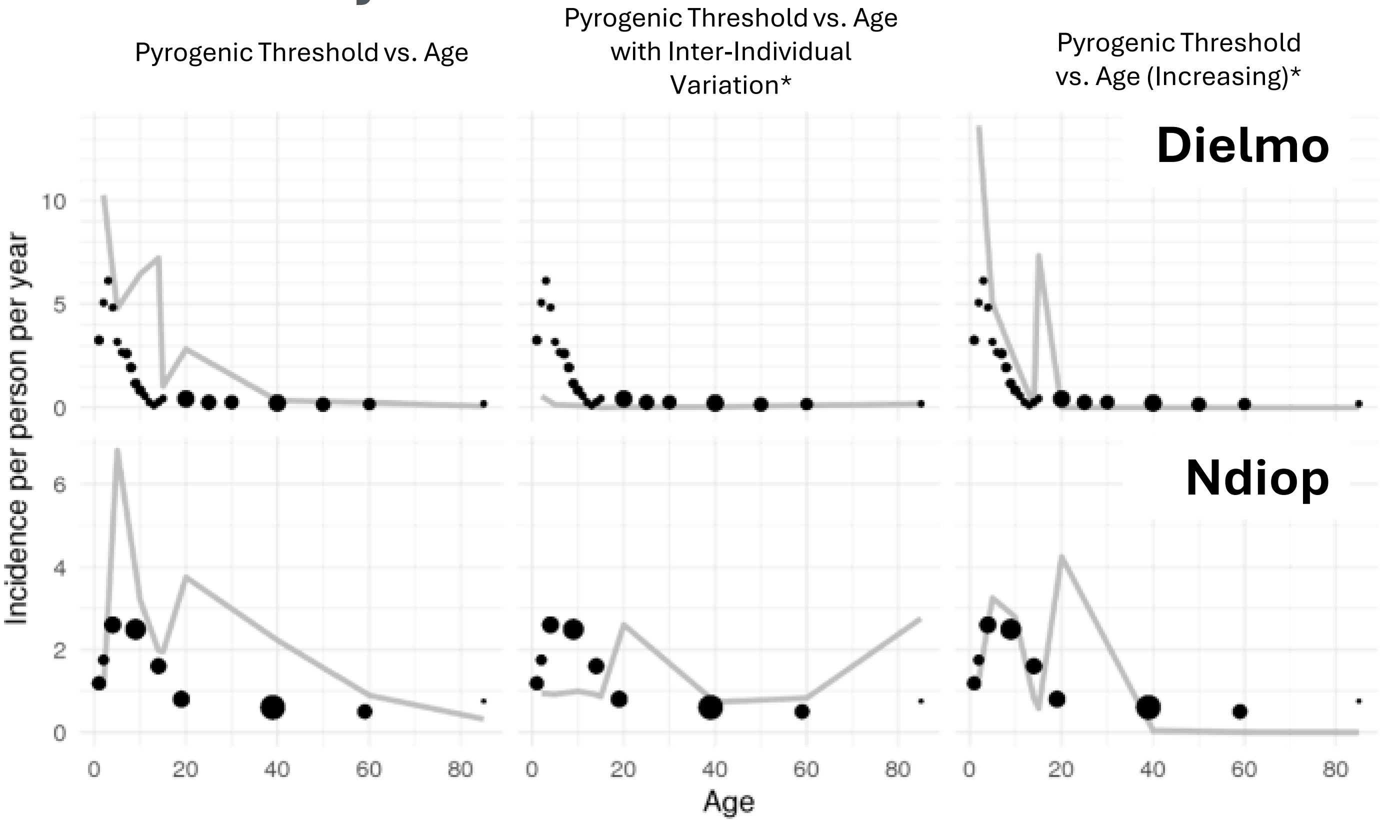
Calibration with heterogeneous innate immunity

Other innate immune models exist in EMOD, adding *age-based* or *inter-individual* heterogeneity to:

Pyrogenic threshold – the concentration [iRBC/μL] at which stimulation of the innate inflammatory immune response is half its maximum value

The **maximum kill rate for iRBCs due to the inflammatory innate immune response**, which increases along a sigmoidal curve as fever increases above 38.5 degrees Celsius

Default parameter set performance varies across innate immunity models with constant distribution



* Rodriguez-Barraquer et al. "Quantification of anti-parasite and anti-disease immunity to malaria as a function of age and exposure" eLife (2018)
* Building on work by Annie Stahlfeld

In Conclusion

- Model calibration is important, but challenging
- Bayesian optimization with gaussian process emulation accelerates calibration without sacrificing goodness-of-fit
- Fitting reveals key parameter-output relationships in EMOD malaria
- Changes to model structure (i.e. innate immune variation) warrant separate recalibrations

Thank You

Caitlin Bever

Aurélien Cavelan

Jaline Gerardin

Melissa Penny

Ricky Richter

Prashanth Selvaraj

Anne Stahlfeld

Josh Suresh

Contact: tobias.holden@northwestern.edu

NIAID 1F31AI172387-01