Swiss TPH

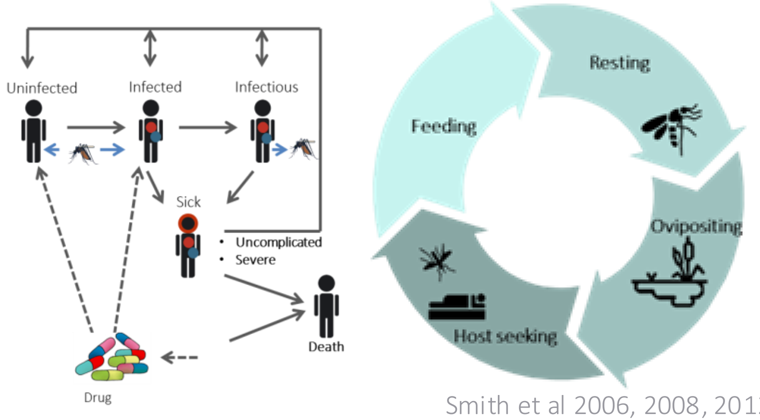
Using Bayesian Optimization to calibrate individual-based models:

Application to OpenMalaria

Aurélien Cavelan <aurelien.cavelan@unibas.ch>
Melissa Penny <melissa.penny@unibas.ch>
Nakul Chitnis <nakul.chitnis@unibas.ch>
Lars Kamber <lars.kamber@unibas.ch>

OpenMalaria

An open-source individual-based stochastic model of malaria epidemiology and control.



Smith et al 2006, 2008, 2012
Reiker 2021

Parasite densities
Infectiousness
Number of infections
Immunity
Drug level

Infectiousness
Mosquito density
Feeding cycle
Parasite development
seasonality

Health System
Drugs & quality
Adherence
Compliance

Drugs
Vector Control (IRS, ITNs)
Vaccines
Mass treatment
Reactive strategies

github.com/SwissTPH/openmalaria

Model parameters

labels	R0000
$1 - -ln(1 - S\infty)$	0.0507
$2 - E^*$	0.0325
$3 - S_{imm}$	0.1382
$4 - X_p^*$	1514.3859
$5 - \gamma_p$	2.0369
$6 - \sigma_i^2$	10.1736
$7 - X_y^*$	35158523.3113
$8 - X_h^*$	97.3347
$9 - ln(1 - \alpha_m)$	2.3303
$10 - a_m^*$	2.5311
$11 - \sigma_0^2$	0.6557
$12 - X_\nu^*$	0.9162
$13 - Y_2^*$	6502.2634
$14 - \alpha$	142601.9125
$15 - DensityBias(nonGarki)$	0.1774
$16 - \sigma_2$	0.05
$17 - log(\phi_1)$	0.7362
$18 - Q_D$	0.0188
$19 - Q_n$	49.539
$20 - DensityBias(Garki)$	4.7961
$21 - Y_{B_1}^*$	784455.6
22 – Immune penalty	1.0
23 – Immune effector decay	0.0
$24 - F_0$	0.0968
$25 - \frac{log2}{\omega}$	0.2754
$26 - Y_1^*$	0.5965
$27 - Asexualimmunitydecay$	0.0
$28 - Y_0^*$	296.3024
$29 - Idetemultiplier$	2.7975
$30 - a_F^*$	0.1174
$31 - Muller_1$	0.0
$32 - Muller_2$	0.0

Infection Incidence

Acquisition of Immunity

Parasite densities

Disease Model:

- Pathogenesis
- Clinical
- Severe
- Mortality

Between 20 and 25 model parameters need to be calibrated.

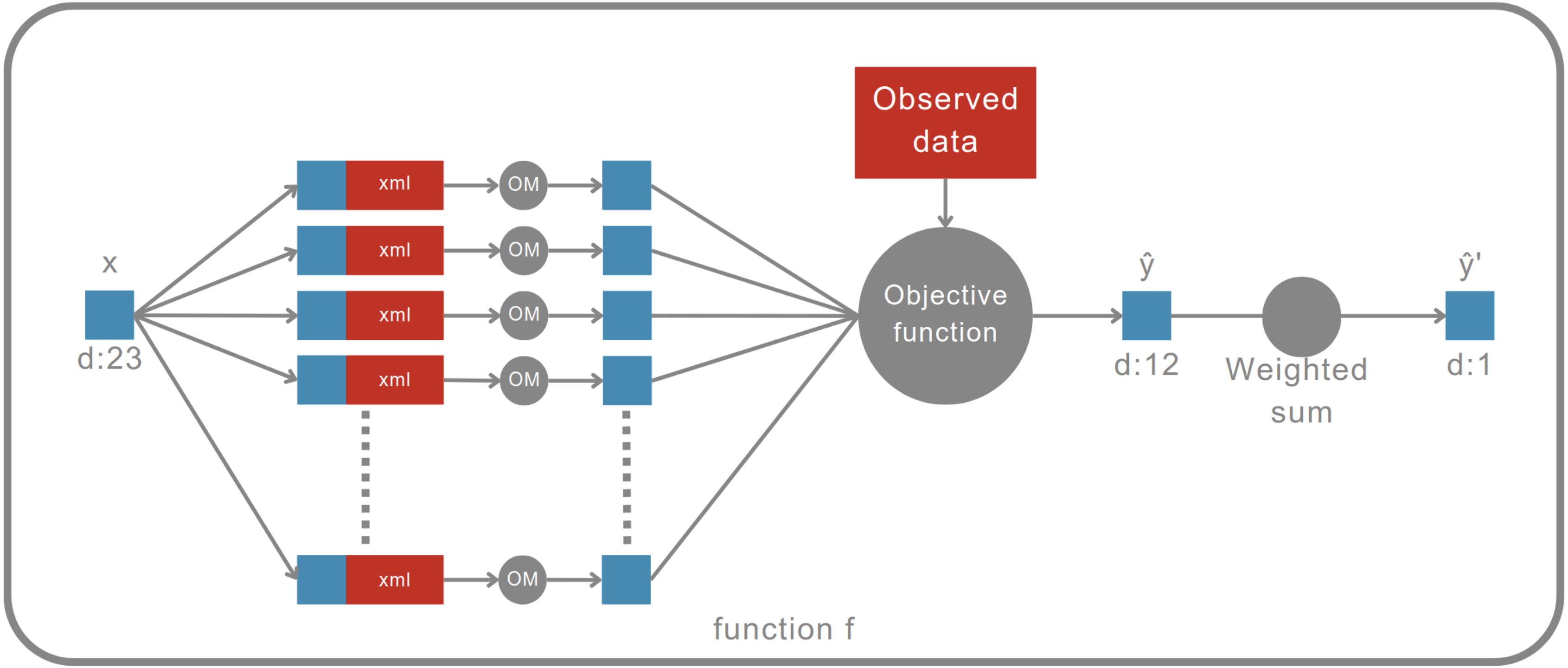
These parameters are not easily 'observable' in practice.

They need to be calibrated using observed data.

Why recalibrate?

- New data
- New model features
- Model variants

Loss function: 23 parameters -> 12 outputs



The curse of dimensionality

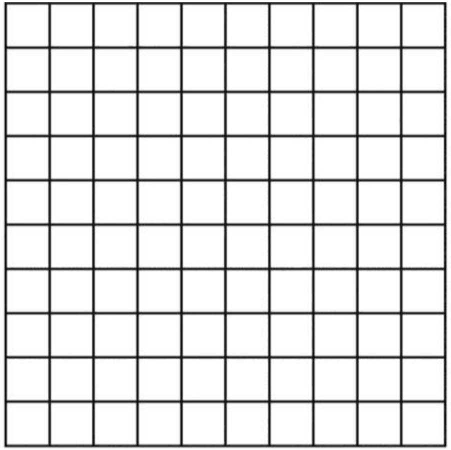
Ideally, we would like 10 samples per dimension.

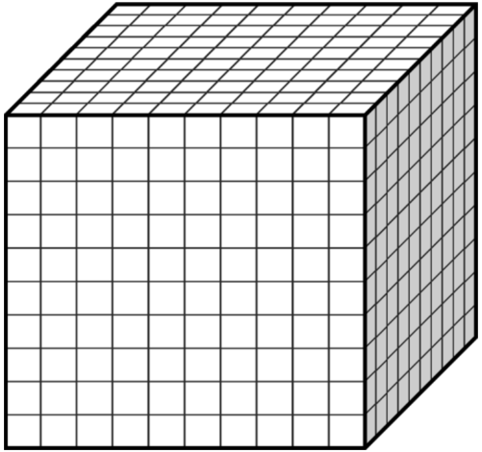
1D: 10 samples
2D: 100 samples
3D: 1000 samples
...
23D: 10^23 samples (more than stars in the observable universe!)

Realistically, we can only do **10,000 samples maximum:**
0.0000000000000001% of 10^23

Bayesian optimization helps in minimizing the number of samples needed to converge.

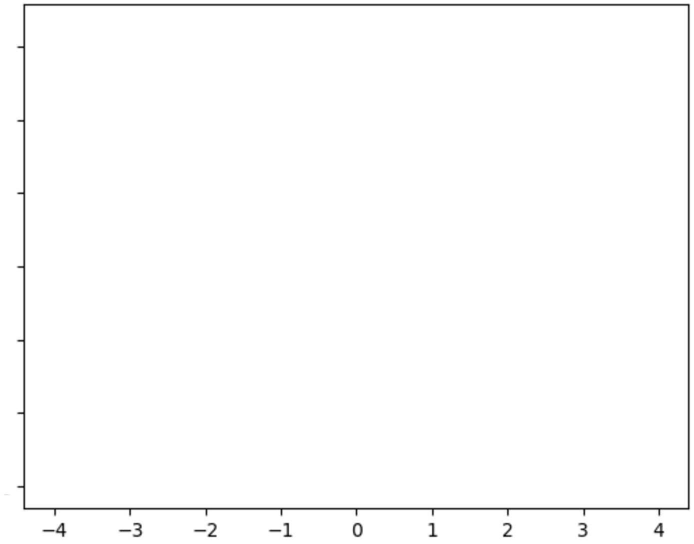
However, exploration is still challenging.





Bayesian Optimization: step-by-step example

Ackley 1D

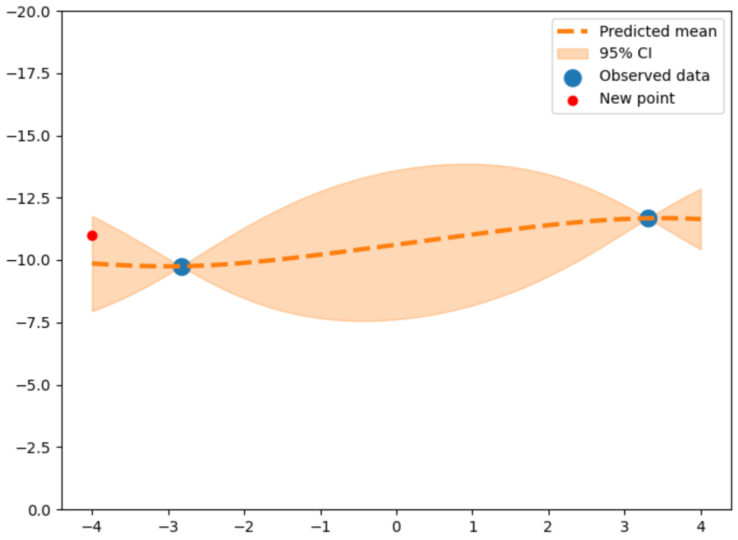


Simple 1D function.

Goal: minimize the function for x

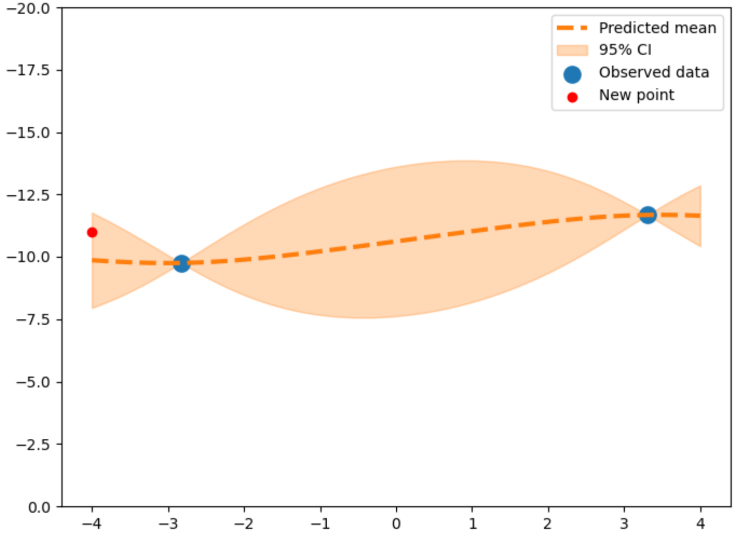
Unfortunately, in real applications we do not know the shape of the function.

Bayesian Optimization: step-by-step example



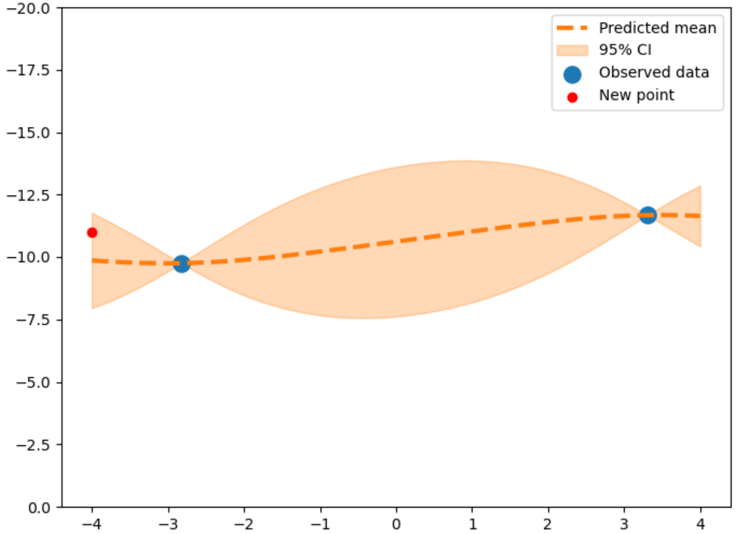
1. **Initial sampling**
   - OpenMalaria: 3000 initial samples

Bayesian Optimization: step-by-step example



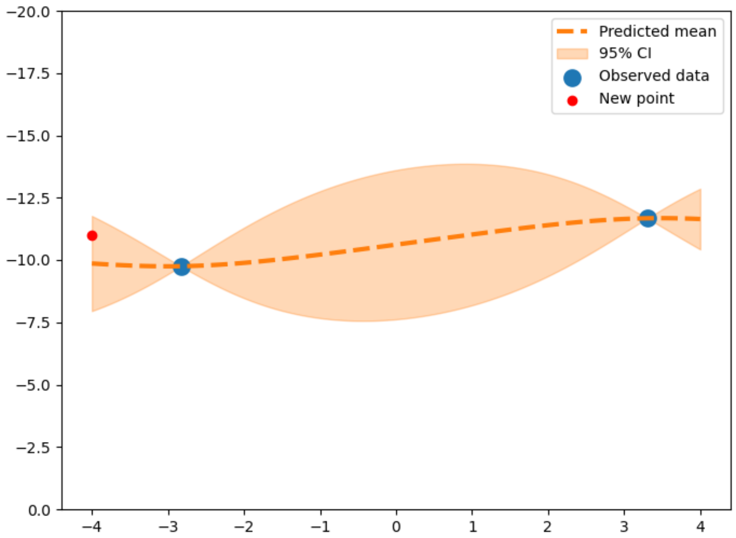
1. **Initial sampling**
   - OpenMalaria: 3000
2. **Fit surrogate model**
   - Single-task Gaussian process
   - Multi-task Gaussian process

Bayesian Optimization: step-by-step example



1. **Initial sampling**
   - OpenMalaria: 3000
2. **Fit surrogate model**
   - Single-task Gaussian
   - Multi-task Gaussian process
3. **Optimize acquisition function**
   - Exploration vs exploitation dilemma ("mutli-armed bandit problem")
   - Optimize for mean vs uncertainty

Bayesian Optimization: step-by-step example

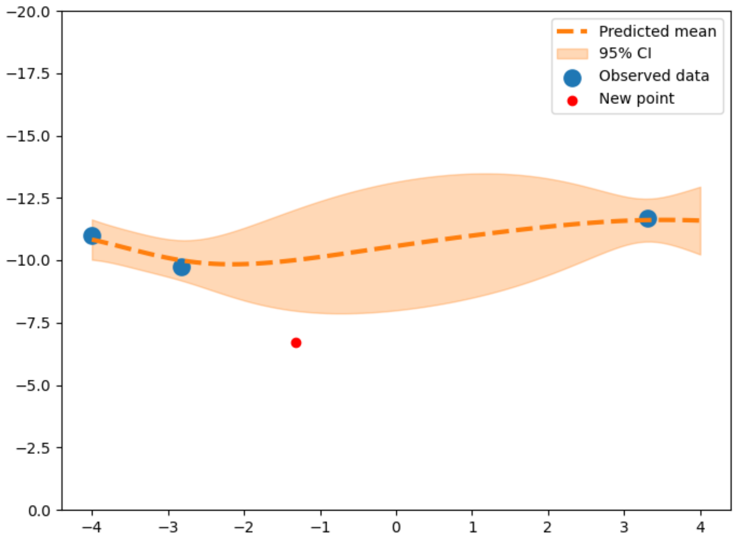


1. **Initial sampling**
   - OpenMalaria: 3000
2. **Fit surrogate model**
   - Single-task Gaussian
   - Multi-task Gaussian process
3. **Optimize acquisition function**
   - Exploration vs exploitation dilemma ("mutli-armed bandit problem")
   - Optimize for mean vs uncertainty

Possible acquisition functions:

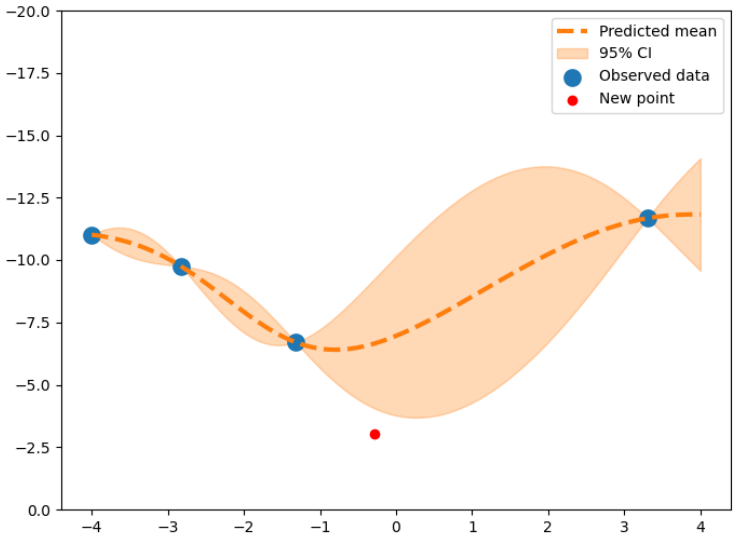
- *Expected improvement* (more exploration)
- *Upper confidence bound* (can be adjusted)
- *Thompson sampling* (more exploitation)
- ...

Bayesian Optimization: step-by-step example

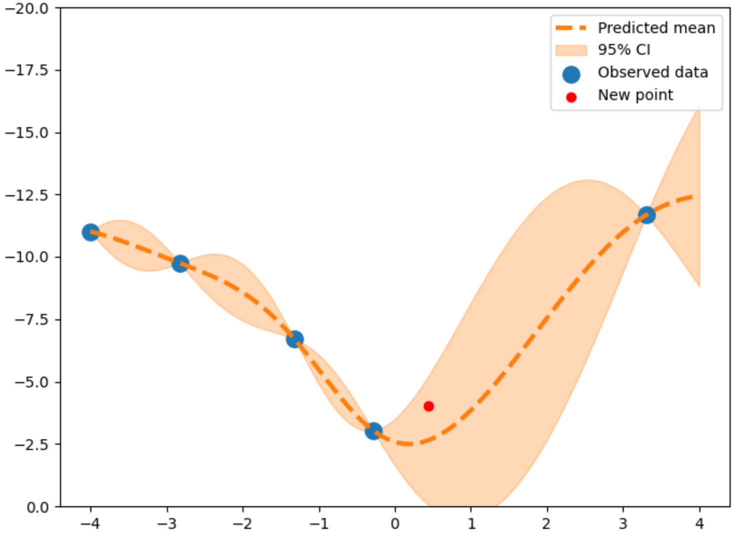


1. **Initial sampling**
   - OpenMalaria: 3000
2. **Fit surrogate model**
   - Single-task Gaussian
   - Multi-task Gaussian process
3. **Optimize acquisition function**
   - Exploration vs exploitation dilemma ("mutli-armed bandit problem")
   - Optimize for mean vs uncertainty
4. **Add point to observed data**
   - OpenMalaria: batches of 64 points
   - Up to 1h on A100 GPU to fit Multi-Task gaussian process

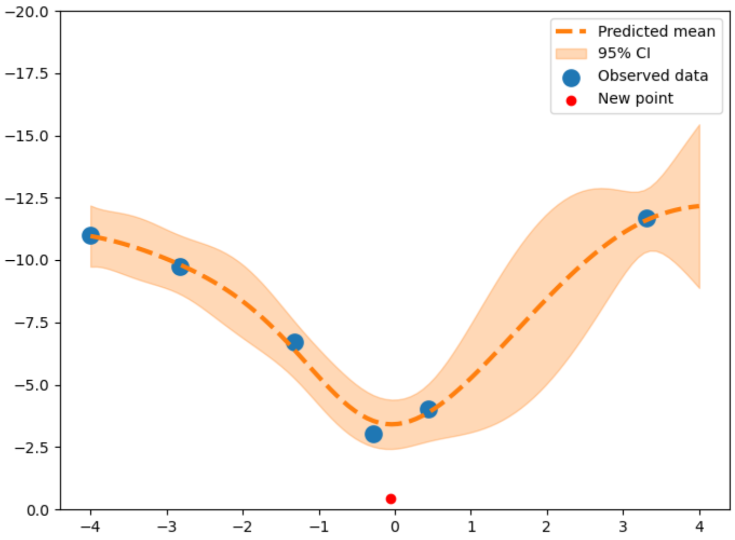
Bayesian Optimization: step-by-step example



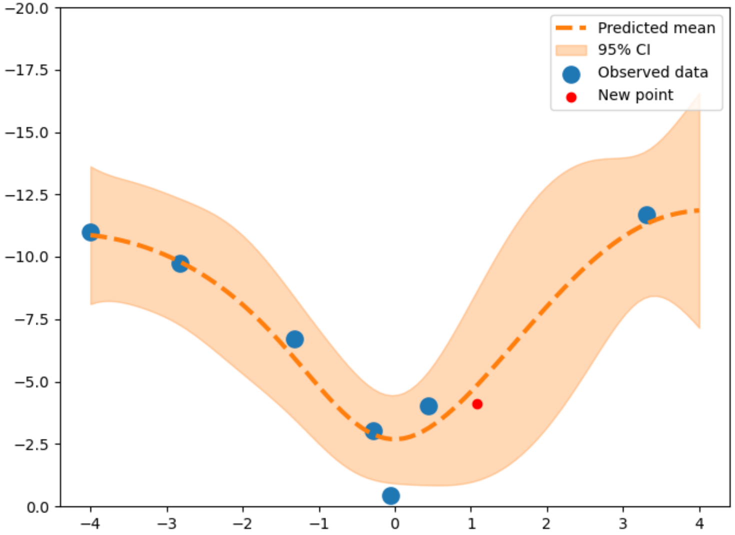
Bayesian Optimization: step-by-step example



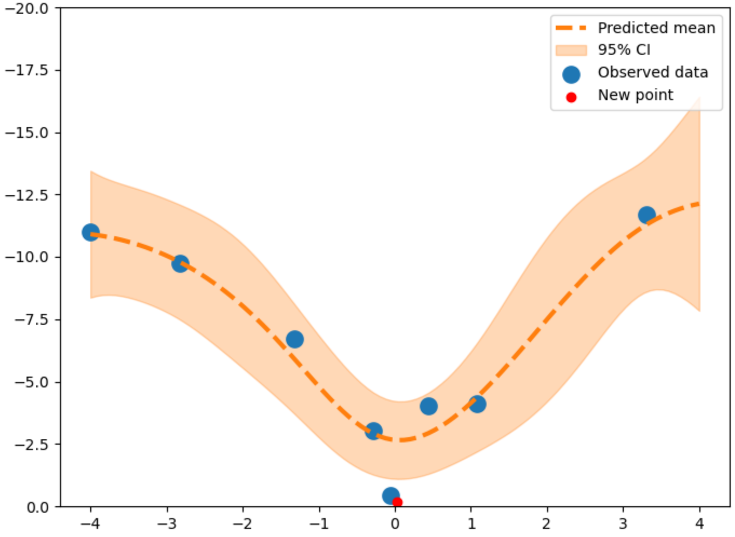
Bayesian Optimization: step-by-step example



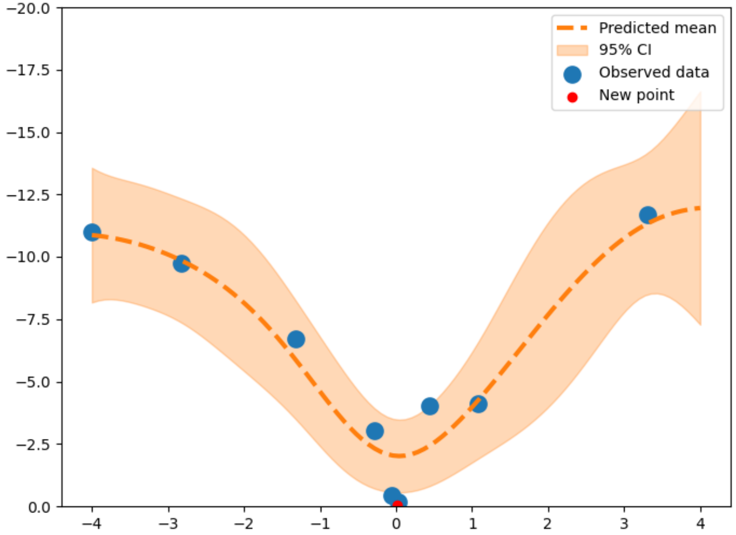
Bayesian Optimization: step-by-step example



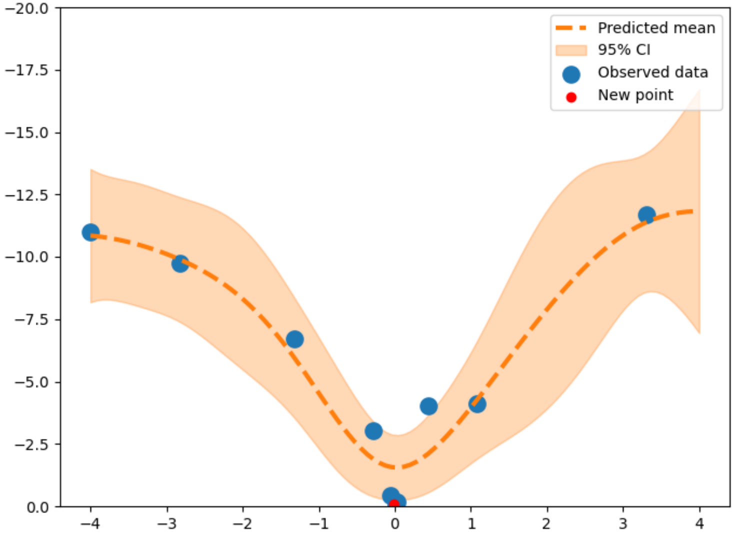
Bayesian Optimization: step-by-step example



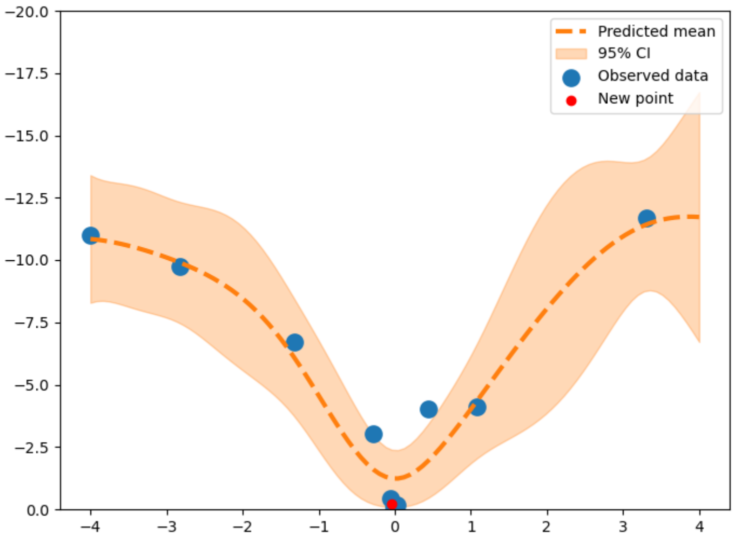
Bayesian Optimization: step-by-step example



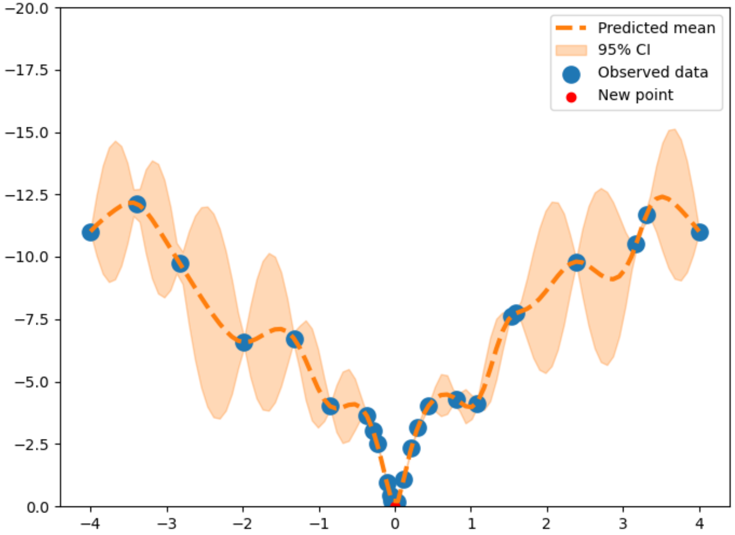
Bayesian Optimization: step-by-step example



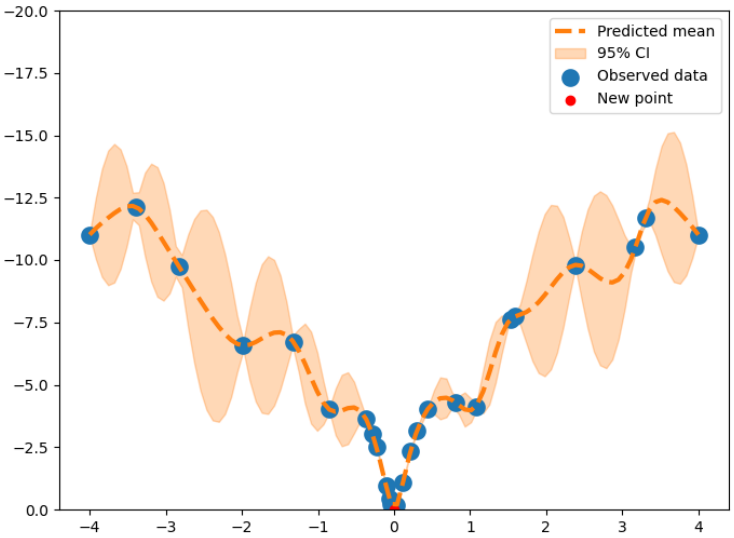
Bayesian Optimization: step-by-step example

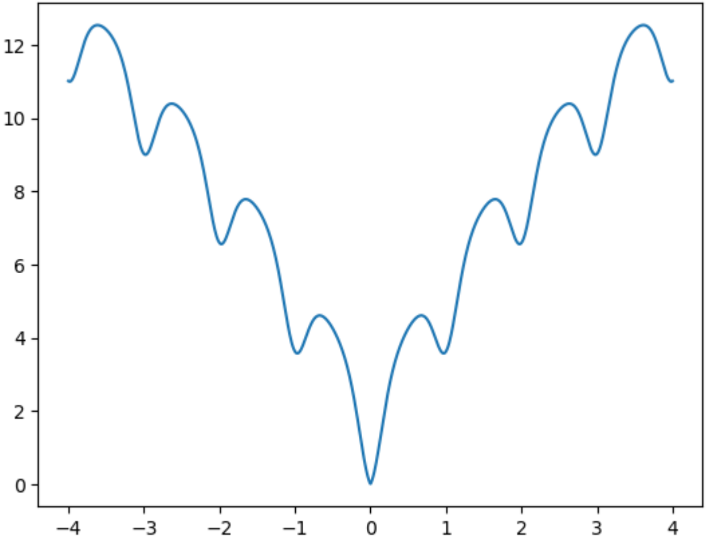


Bayesian Optimization: step-by-step example

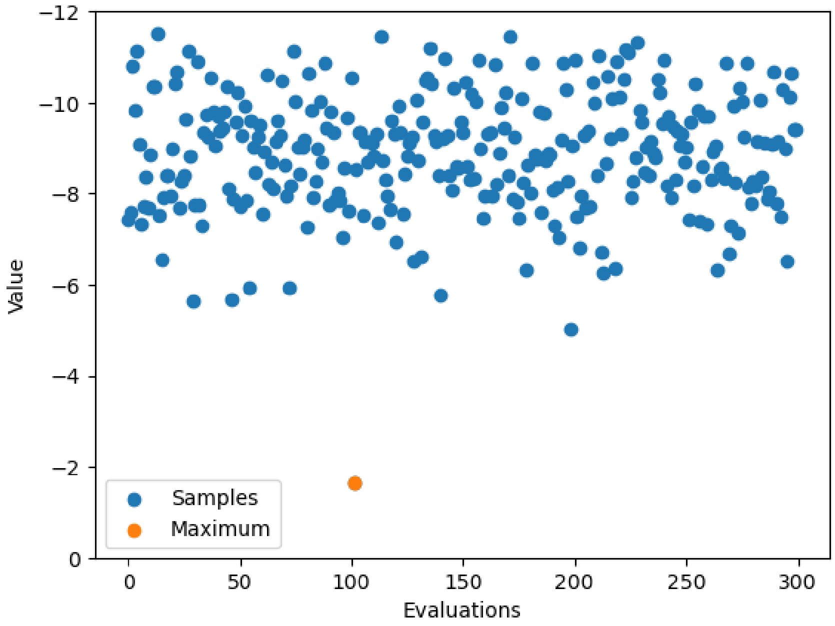


Bayesian Optimization: step-by-step example





Convergence of Ackley 5D



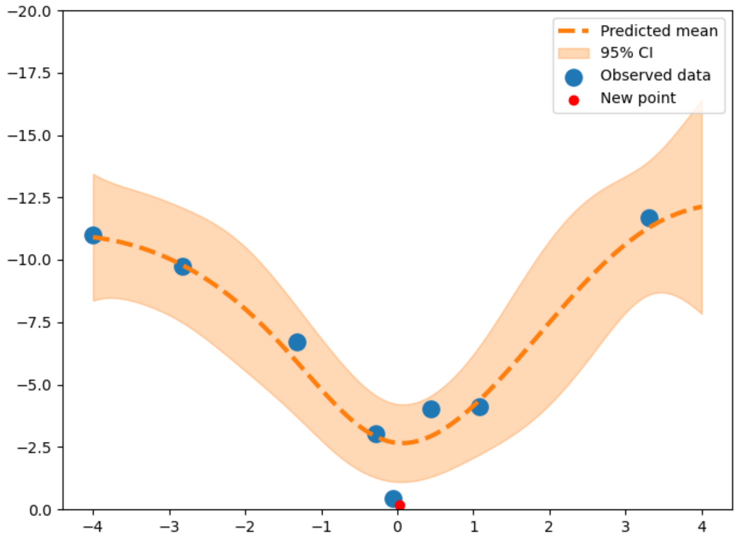
No convergence in 5D.

- The search space is too already large!
- Most of the search space is unknown
- Uncertainty (variance) is high everywhere
- Acquisition function does only exploration (forever)

Trust Region Bayesian Optimization (TuRBO)

Algorithm fails to improve over 3 consecutive iterations

- Cut the search space in half around the best current point

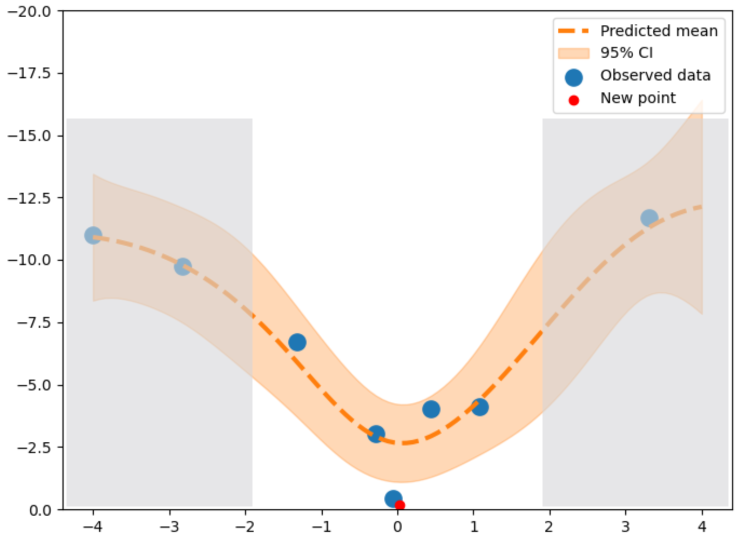


[1]: Eriksson, David, et al. Scalable global optimization via local Bayesian optimization. Advances in Neural Information Processing Systems. 2019

Trust Region Bayesian Optimization (TuRBO)

Algorithm fails to improve over 3 consecutive iterations

- Cut the search space in half around the best current point

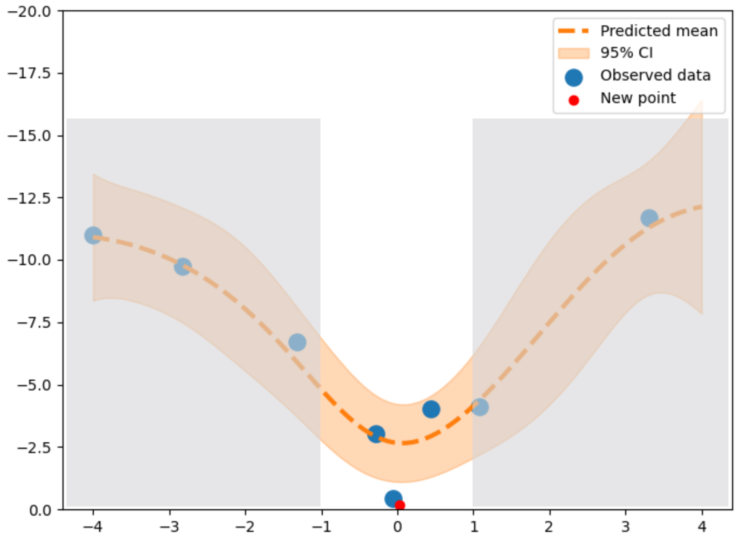


[1]: Eriksson, David, et al. Scalable global optimization via local Bayesian optimization. Advances in Neural Information Processing Systems. 2019

Trust Region Bayesian Optimization (TuRBO)

Algorithm fails to improve over 3 consecutive iterations

- Cut the search space in half around the best current point



[1]: Eriksson, David, et al. Scalable global optimization via local Bayesian optimization. Advances in Neural Information Processing Systems. 2019

Trust Region Bayesian Optimization (TuRBO)

Algorithm fails to improve over 3 consecutive iterations

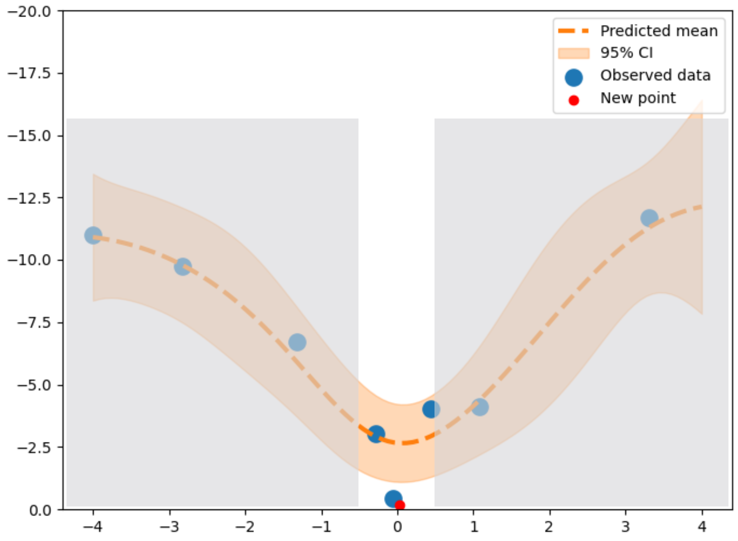
- Cut the search space in half around the best current point

Algorithm succeeds in improving over 3 consecutive iterations

- Double the search space around the best current point

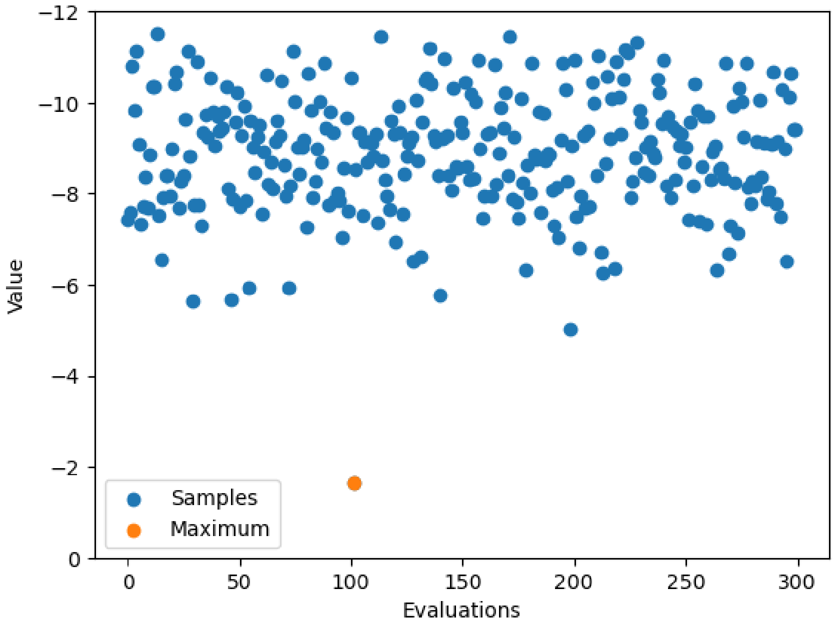
TuRBO:

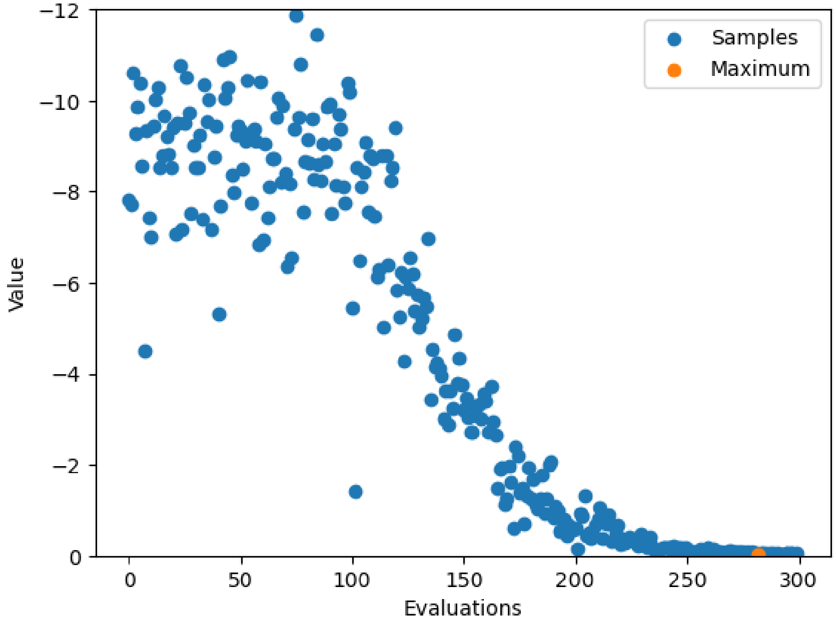
- Guaranteed convergence
- More likely to end up in a local minima
- Requires more initial samples



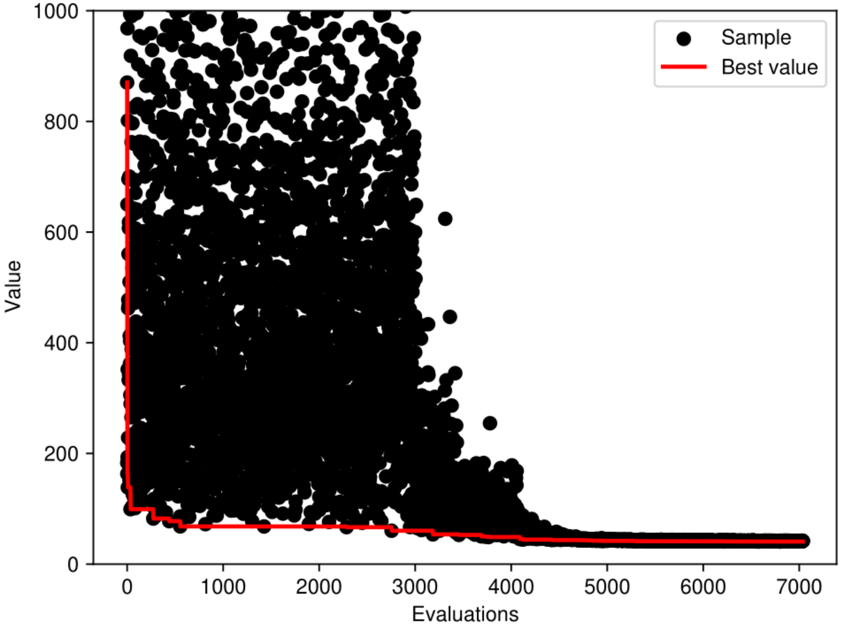
[1]: Eriksson, David, et al. Scalable global optimization via local Bayesian optimization. Advances in Neural Information Processing Systems. 2019

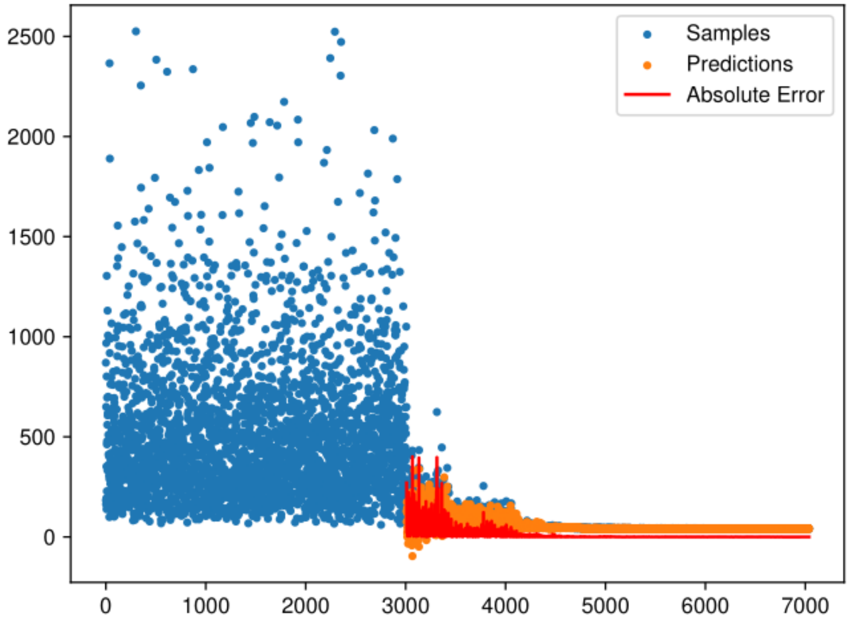
Convergence okf Ackley 5D with TuRBO (right)



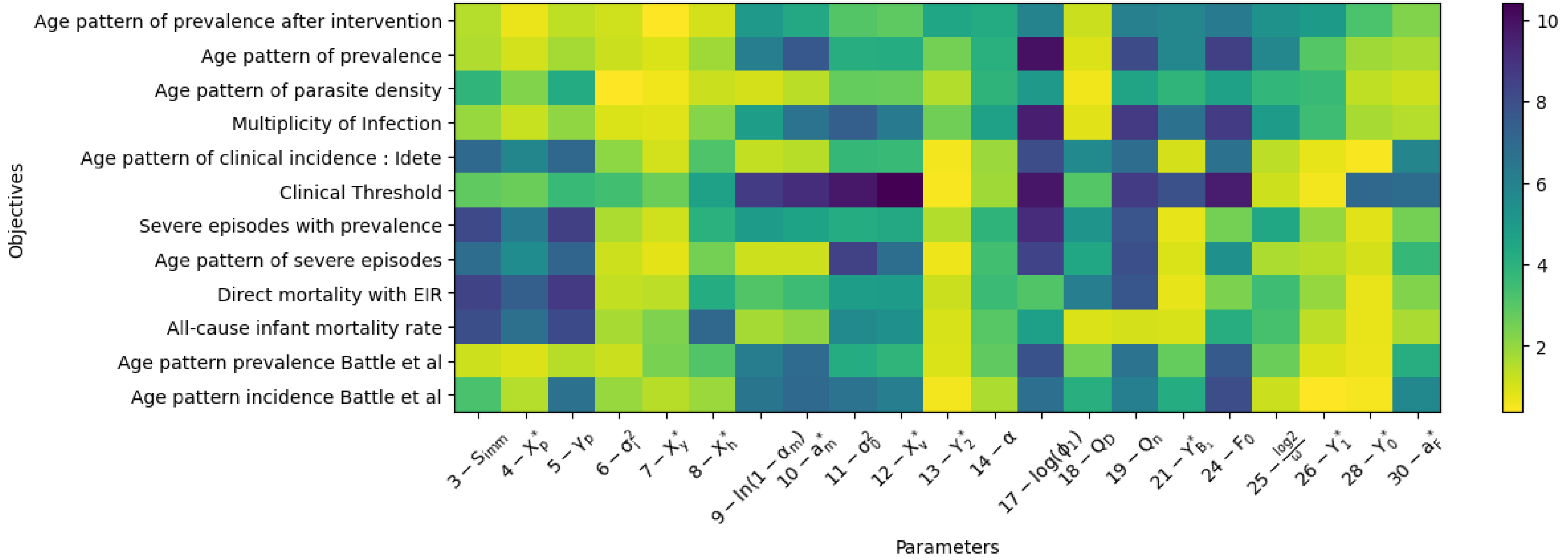


Convergence with OpenMalaria 23 parameters

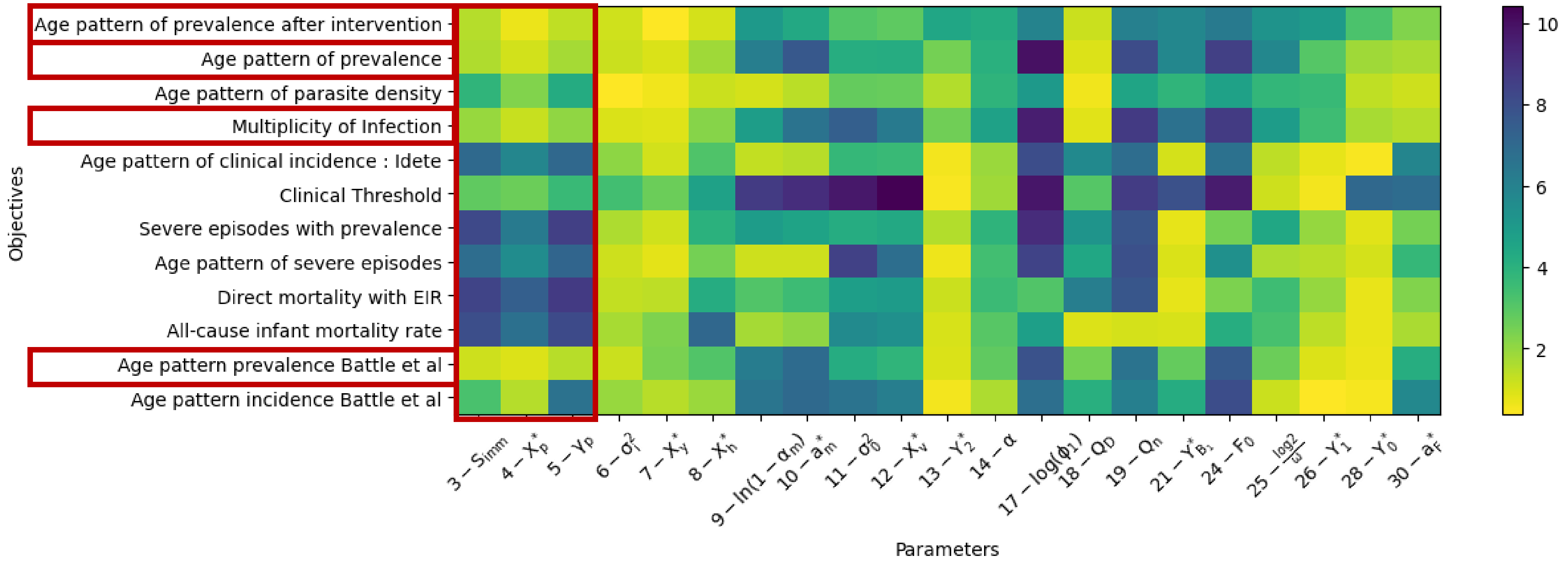




Parameter-objective correlation



Parameter-objective correlation



Takeway messages

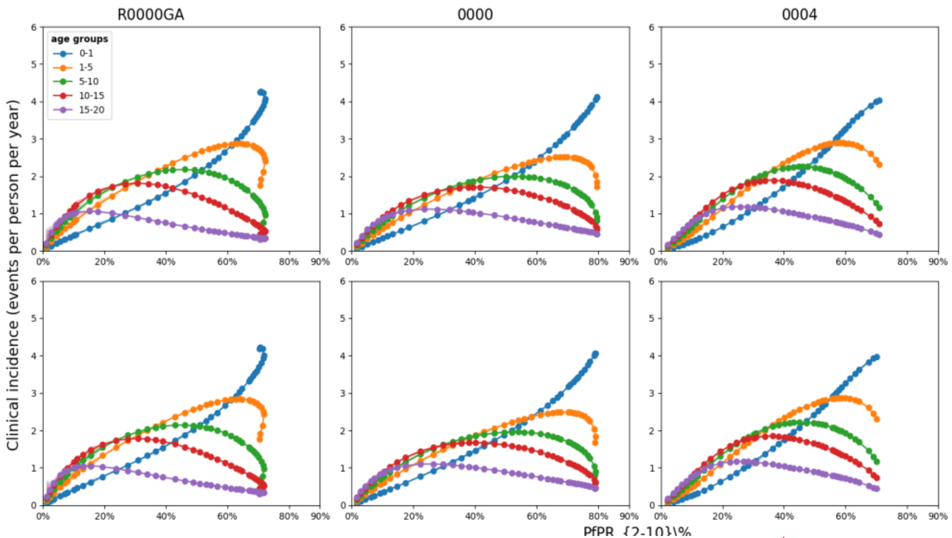
We have developed a framework to calibrate individual-based models (or other functions).

NOT plug and play:
- **It requires high quality, curated data**
- Objectives must be constrained enough
- The search space must be well defined
- Tools to validate the model are needed

Limitations:
- The curse of dimensionality
- Computing cost and time
- Unidentifiability of some parameters
- Model limitations

Technical framework: Python & BoTorch

Also used for:
- EMOD (IDM, Chicago)
- Opisthorchis model (Swiss TPH, Basel)
- Vaccine trials (TKI, Perth)



Identifiability – comparing objectives across multiple fits