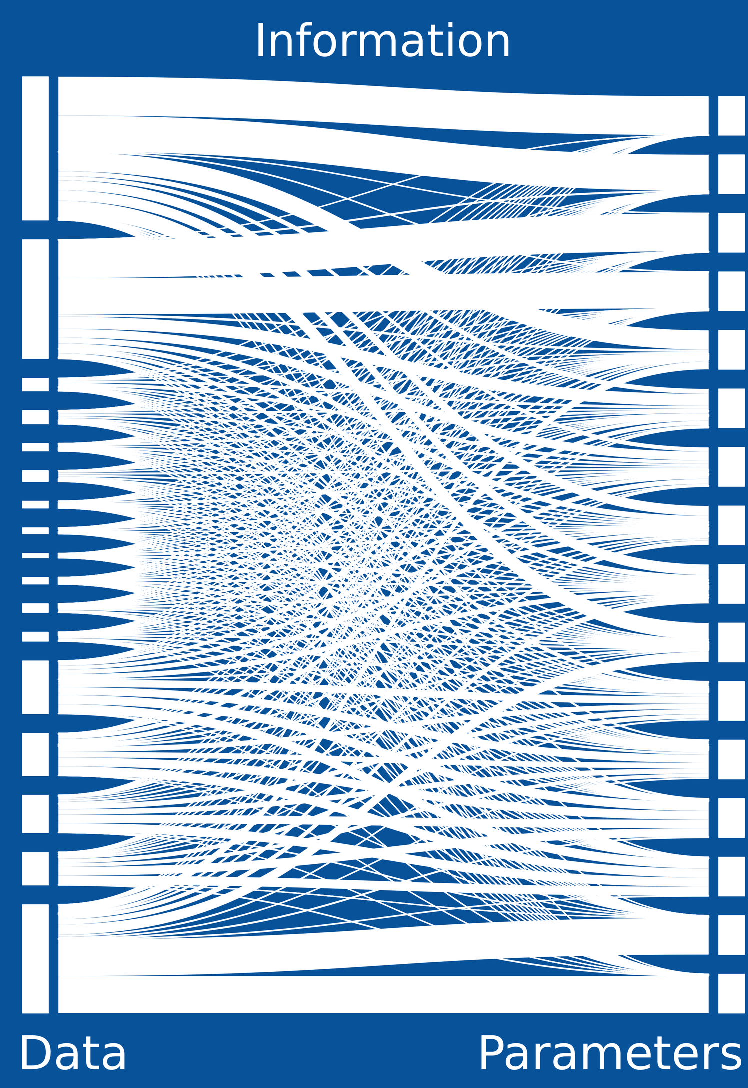
Robustly account for **informative**ness in **ABC** via inverse **ML** models.

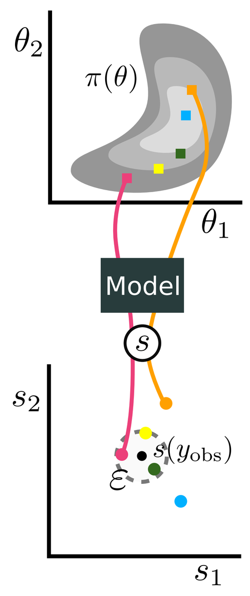


Yannik Schaelte[1,2,3], Jan Hasenauer[1,2,3]

2022-05

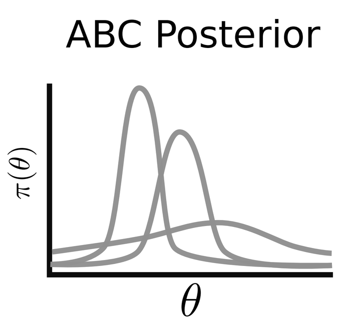
1: Helmholtz Center Munich, Institute of Computational Biology, 2: Technical University Munich, Center for Mathematics, 3: University Bonn, Faculty of Mathematics and Natural Sciences

Mini-Intro: ABC



- Parameter inference method that does't need to evaluate the likelihood function
- until N acceptances:
  1. sample parameters $\theta \sim \pi(\theta)$
  2. simulate data $y \sim \pi(y|\theta)$
  3. accept if $d(s(y), s(y_{\text{obs}})) < \varepsilon$
- often combined with an SMC scheme

The problem



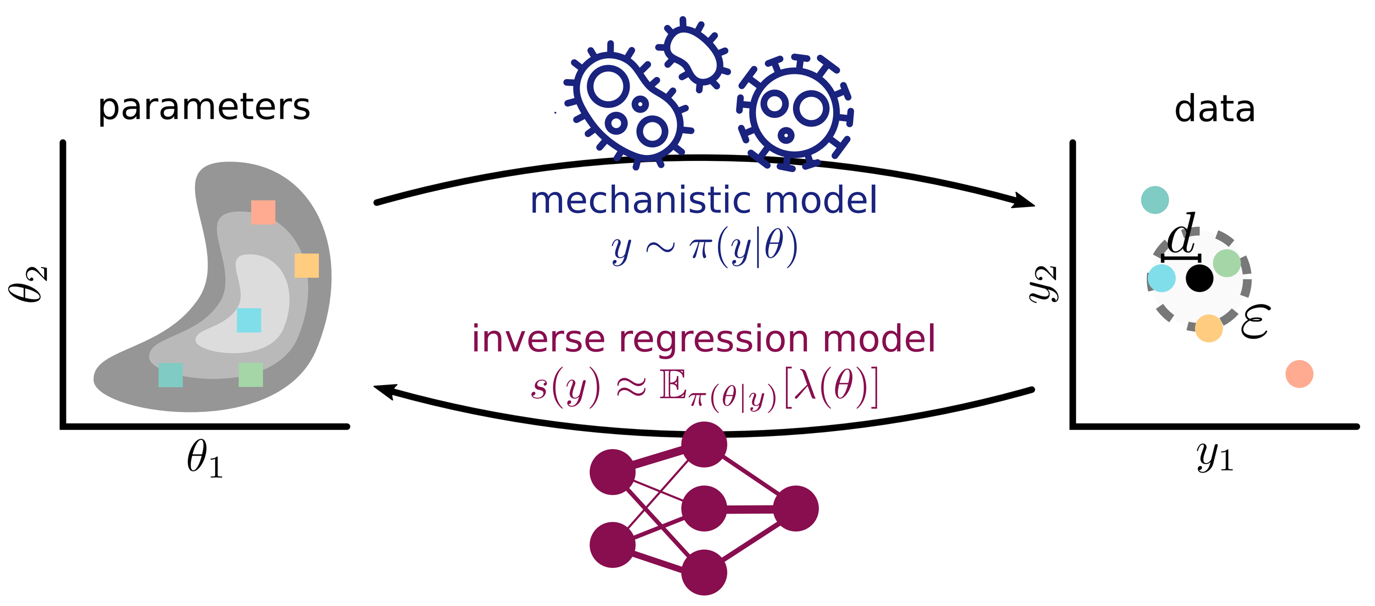
- ABC **easily gives bad results** if summary statistics and distance are not properly calibrated
- ABC posteriors can look very **different,** depending on the **method**

Method: Robust adaptive distances

- Use **robust norms** with **adaptive weights** to normalize for **scale** and down-weight **outliers**

Method: Inverse ML models

- Idea: learn an **inverse regression model** s: y → θ

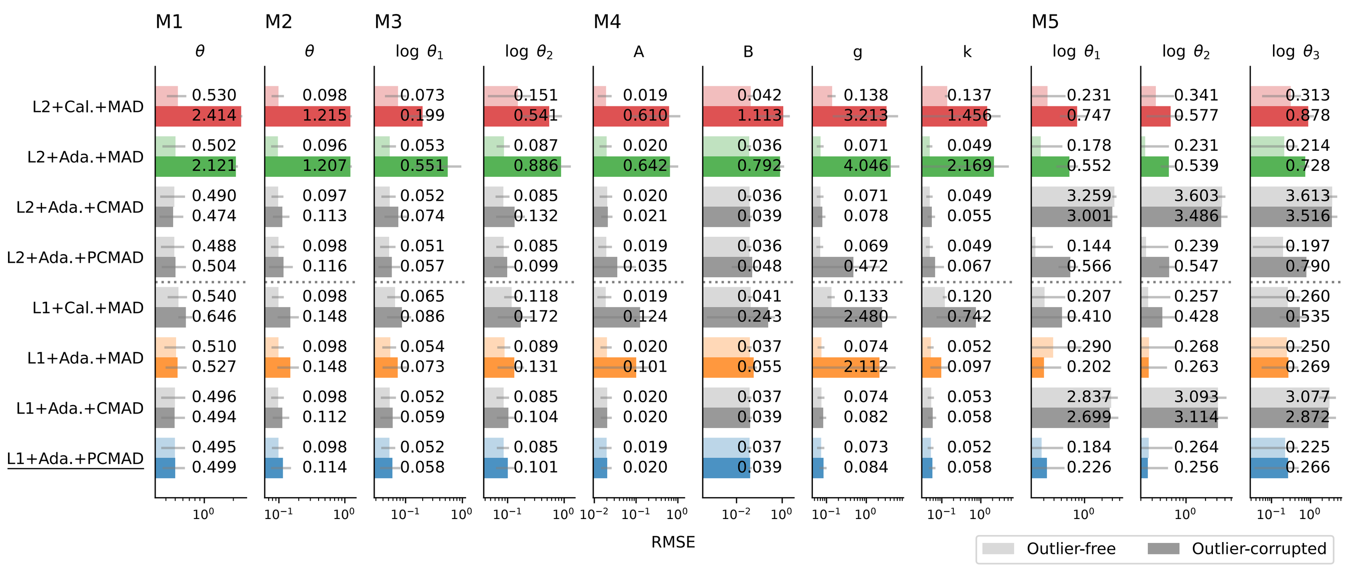


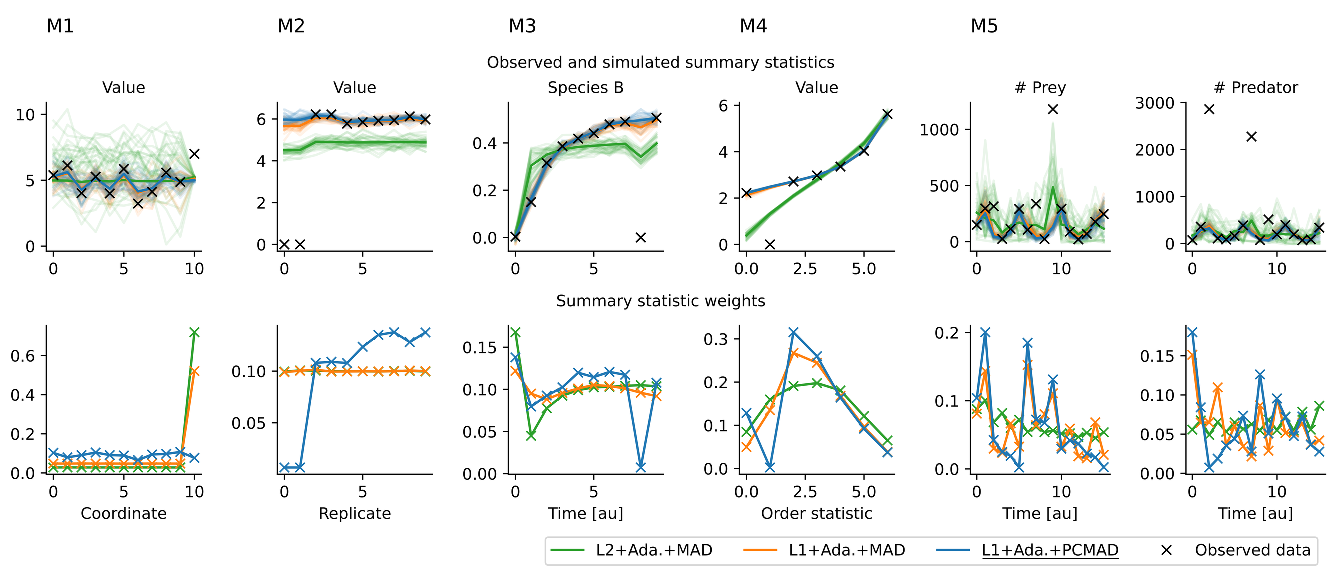
- use that model
  - to construct low-dimensional **summary statistics** ("Summarize all information of data on each parameter in a single value per parameter")
  - to define additional **informativeness weights** given via normalized sensitivities ("If the data were perturbed, how much would the optimal parameters change?")
- learn functions of parameters to capture **higher-order moments**, s: y → λ(θ)

Results

Robustness

the new robust adaptive distances **substantially outperform** established methods, on a variety of test problems with different model and data types

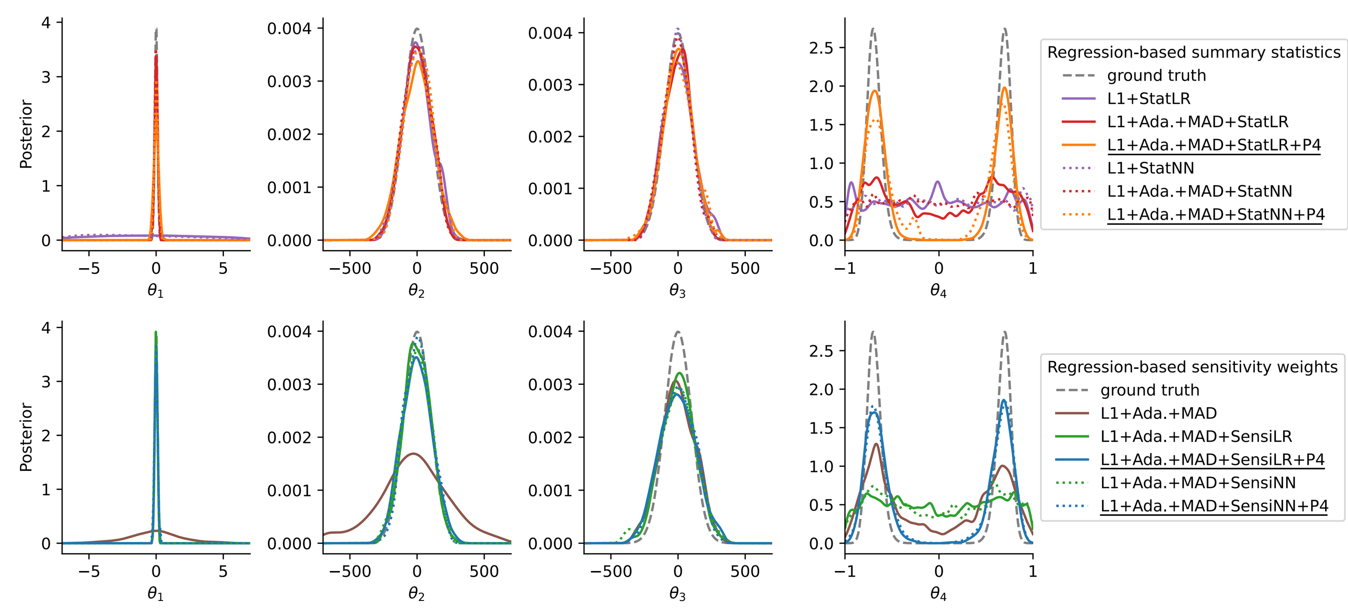


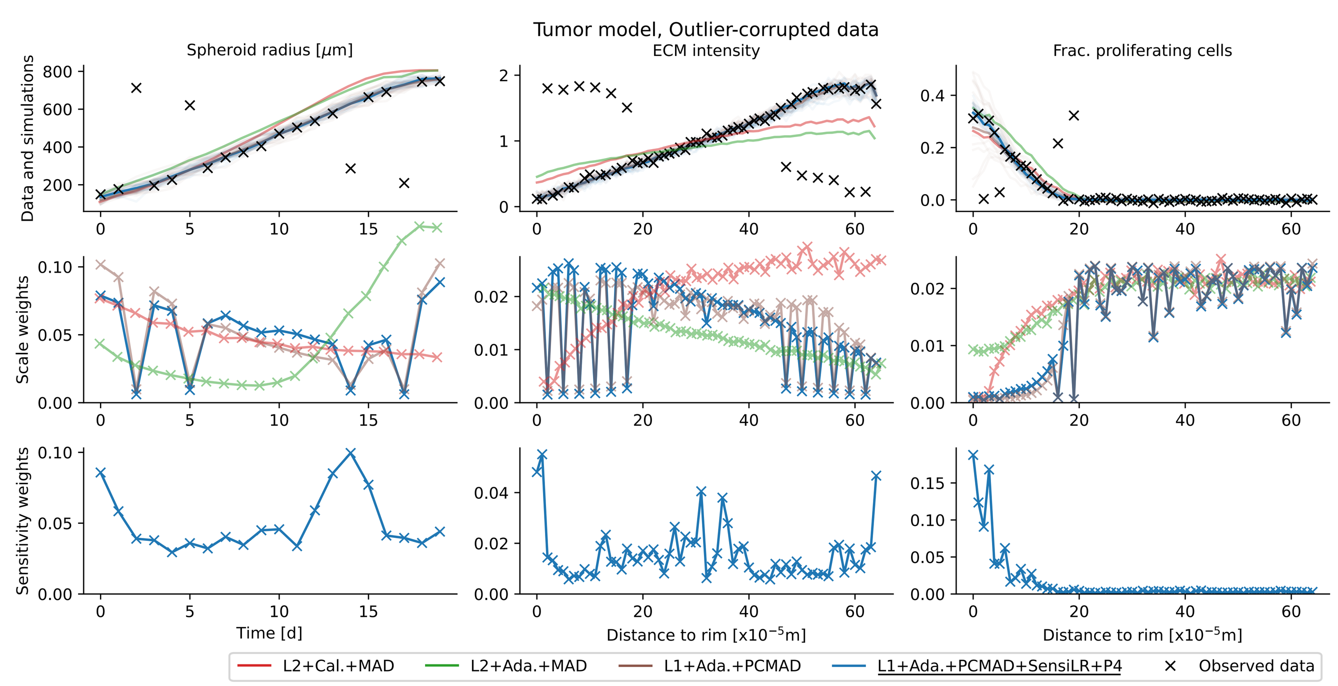


they reliably **identify outliers** and yield accurate fits with **low bias** and **tight confidence intervals**

Informativeness

established methods perform **badly** and can yield **wrong interpretations**; only the **combination** of the new methods (accounting for informativeness; adaptive weighting; assessing higher-order moments) allows to tackle more **challenging problems**





we are able to **simultaneously** robustly detect outliers and account for informativeness also on **complex application problems** (here: an agent-based model of tumor spheroid growth)

Further information

Robust adaptive distance functions for approximate Bayesian inference on outlier-corrupted data. Schälte et al., 2021

Informative and adaptive distances and summary statistics in sequential approximate Bayesian computation. Schälte et al., 2022

pyABC: Efficient and robust easy-to-use approximate Bayesian computation. Schälte et al., 2022

github.com/yannikschaelte
yannik.schaelte@gmail.com
@yannik_schaelte

FitMultiCell

HelmholtzZentrum münchen
German Research Center for Environmental Health

UNIVERSITÄT BONN