Don’t Simulate Twice!

One-shot Sensitivity Analyses via Automatic Differentiation

Arnau Quera-Bofarull, Ayush Chopra, Anisoara Calinescu, Ramesh Raskar, Michael Wooldridge

AAMAS 2023 — London

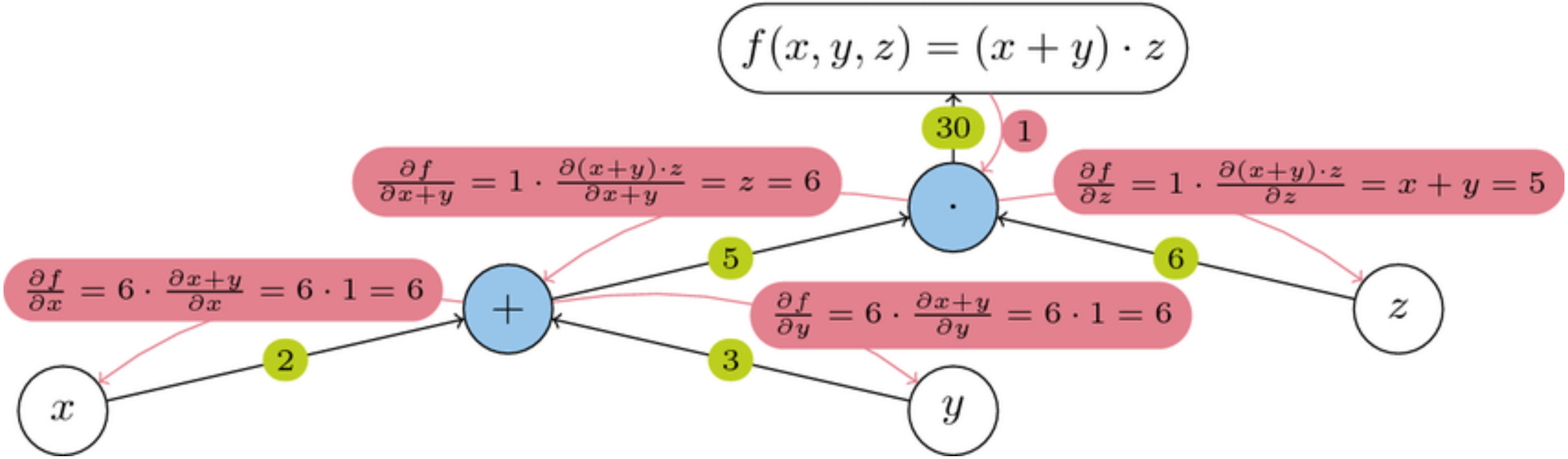
UNIVERSITY OF OXFORD

Agent-Based Models

- ABMs promising tool to model complex systems “bottom-up”
- Wide adoption hindered by (not exhaustive):
  - Scalability
  - Robustness
  - Data availability

Differentiable Agent-Based Models

- Idea: Use Automatic Differentiation in ABMs



Case study: the JUNE epidemiological model

Case study: the JUNE epidemiological model

- JUNE is a 1:1 epi model of England (56 million agents)

Case study: the JUNE epidemiological model

- JUNE is a 1:1 epi model of England (56 million agents)

- GradABM-JUNE is its differentiable implementation (PyTorch).

Case study: the JUNE epidemiological model

- JUNE is a 1:1 epi model of England (56 million agents)

- GradABM-JUNE is its differentiable implementation (PyTorch).

	Simulation
JUNE	50 hours
GradABM-June (GPU)	5 seconds

Tensorisation enables scalability to millions (billions?) of agents

Ref: Ayush Chopra previous talk

Case study: the JUNE epidemiological model

- We can use gradient descent / variational inference for calibration

	Simulation	Calibration (No UQ)	Bayesian Calibration
JUNE	50 hours	-	100k hours
GradABM-June (GPU)	5 seconds	20 minutes	8 hours

Paper coming soon…

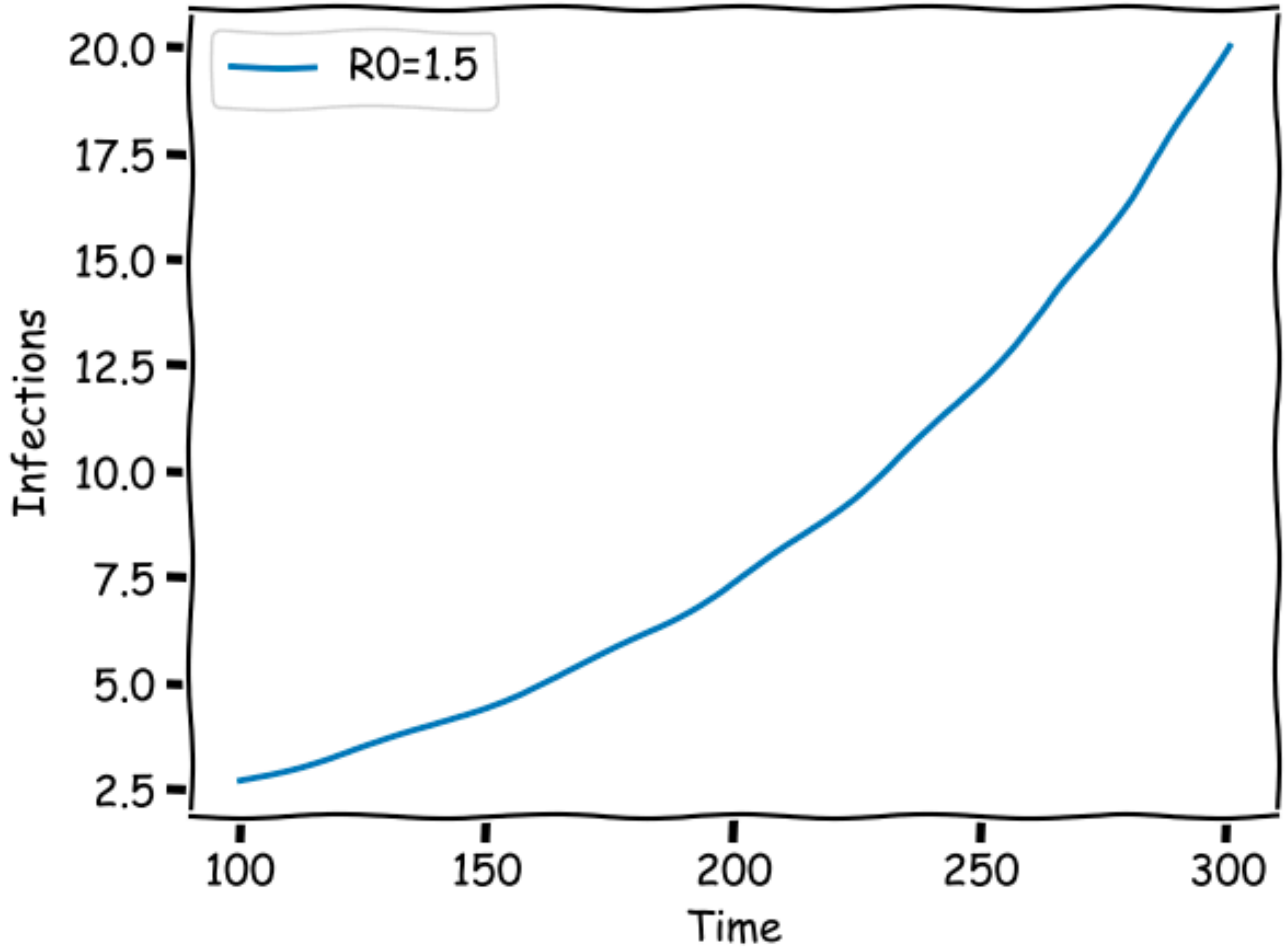
Differentiability enables fast and accurate model calibration

Sensitivity Analyses

Why is it necessary?

1. Robustness

- Example: Epidemiological ABM
- One parameter: R0, an expert measures to be R0 = 1.5 ± 0.3



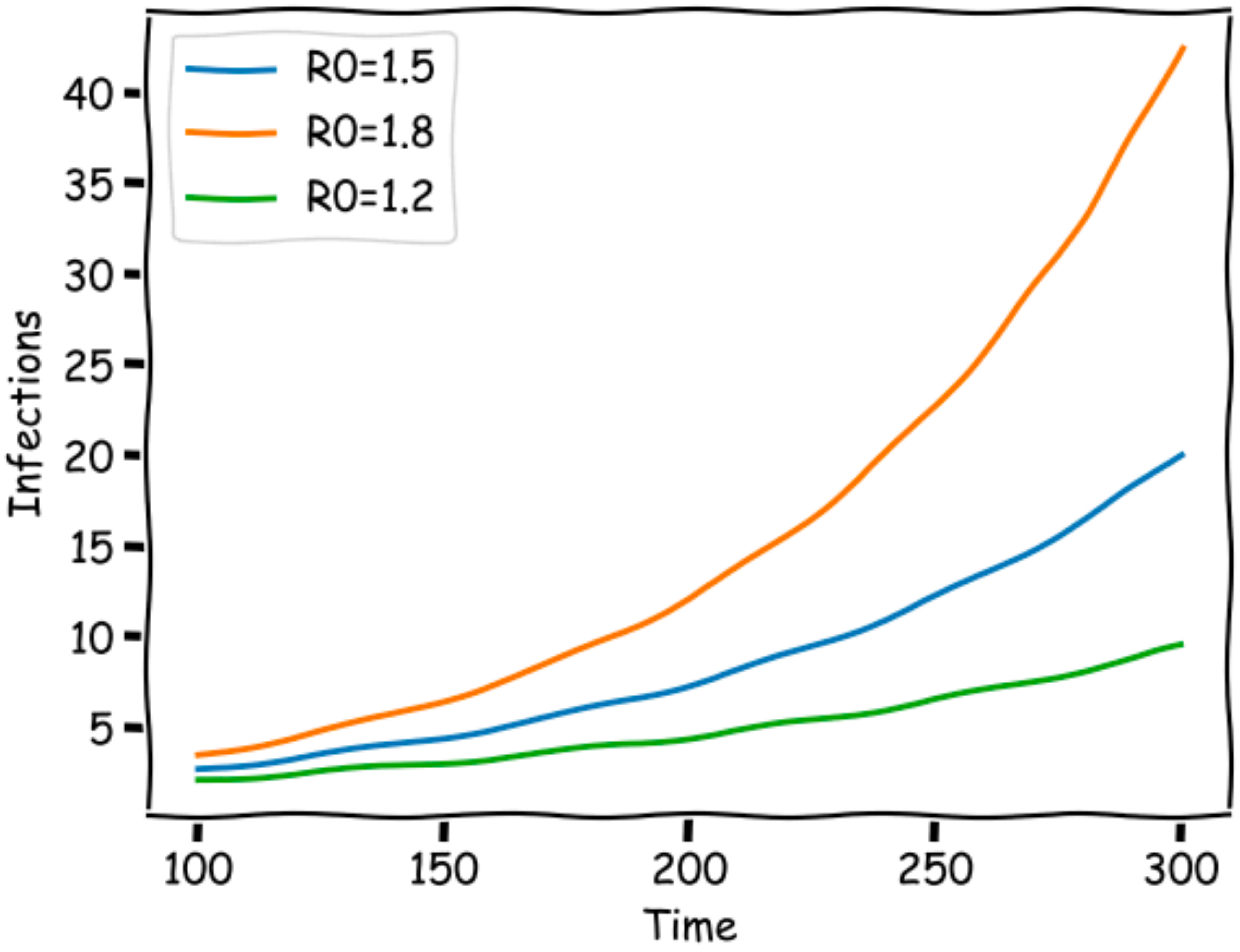
Sensitivity Analyses crucial for policy evaluation

Sensitivity Analyses

Why is it necessary?

1. Robustness

- Example: Epidemiological ABM
- One parameter: R0, an expert measures to be R0 = 1.5 ± 0.3



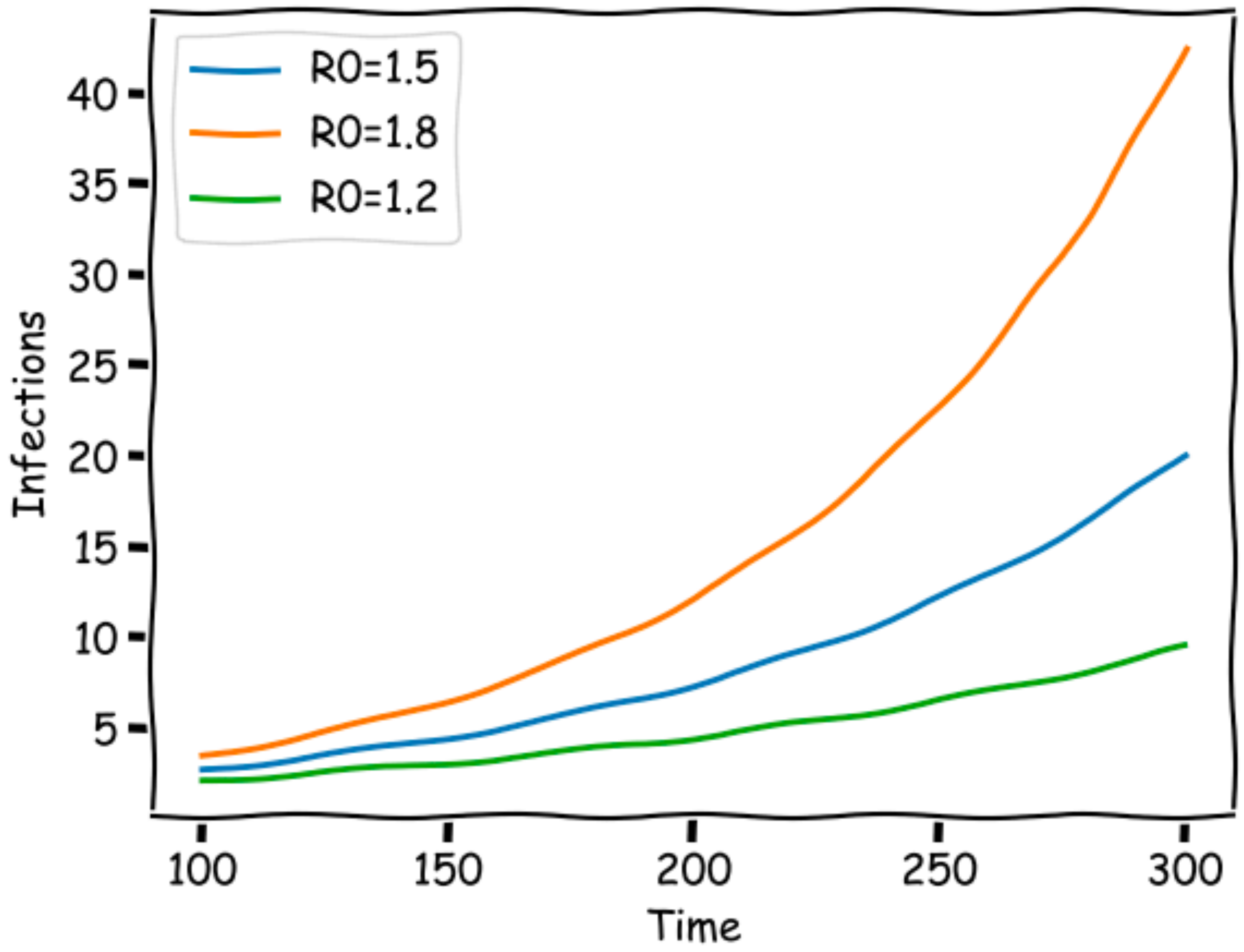
Sensitivity Analyses crucial for policy evaluation

Sensitivity Analyses

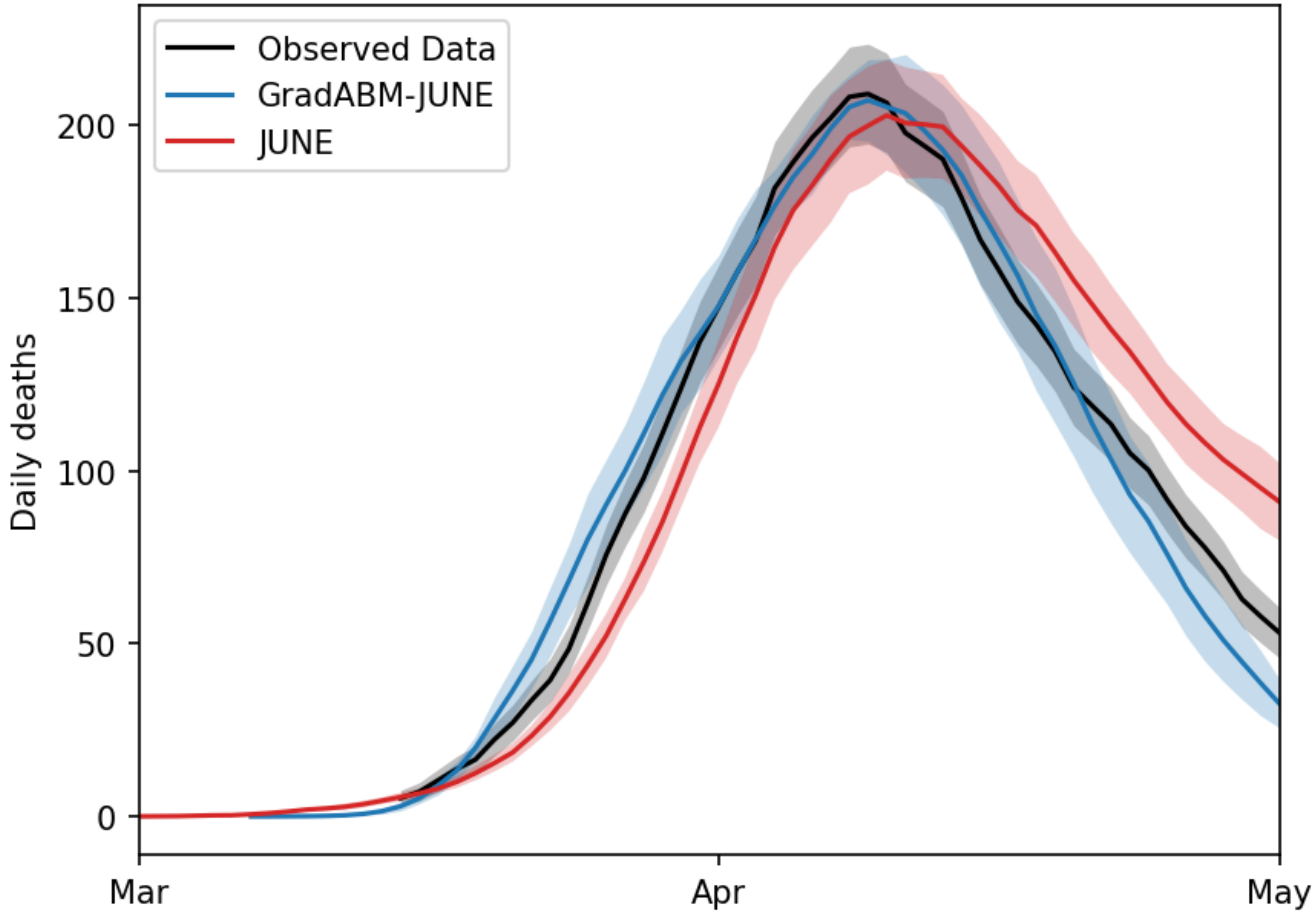
Why is it necessary?

2. Interpretability

- Sensitive parameters tells us what's important in the model.

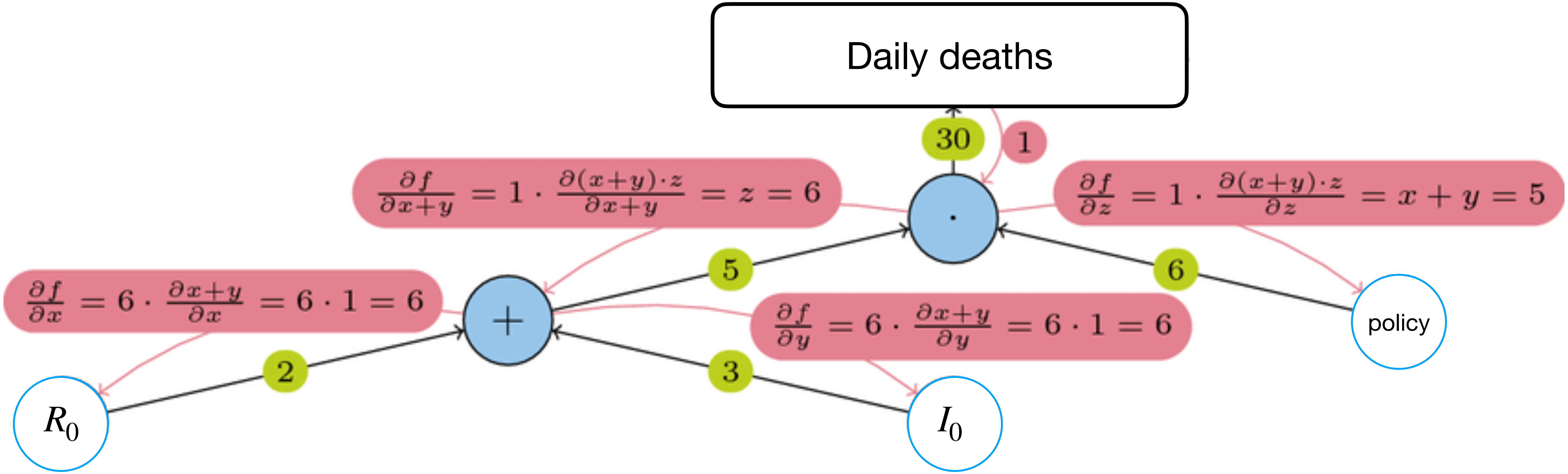


JUNE: London case study

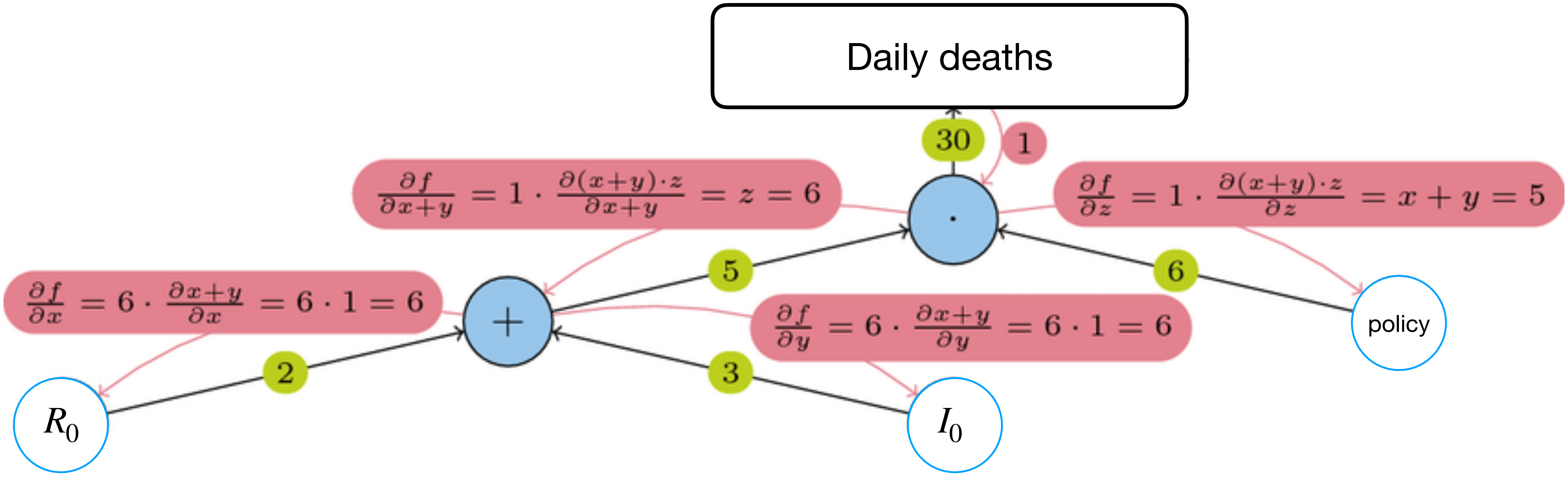


11 Free parameters:
10 contact intensity locations
(schools, companies, etc.)
+
Initial number of cases

Sensitivity Analysis via Automatic Differentiation

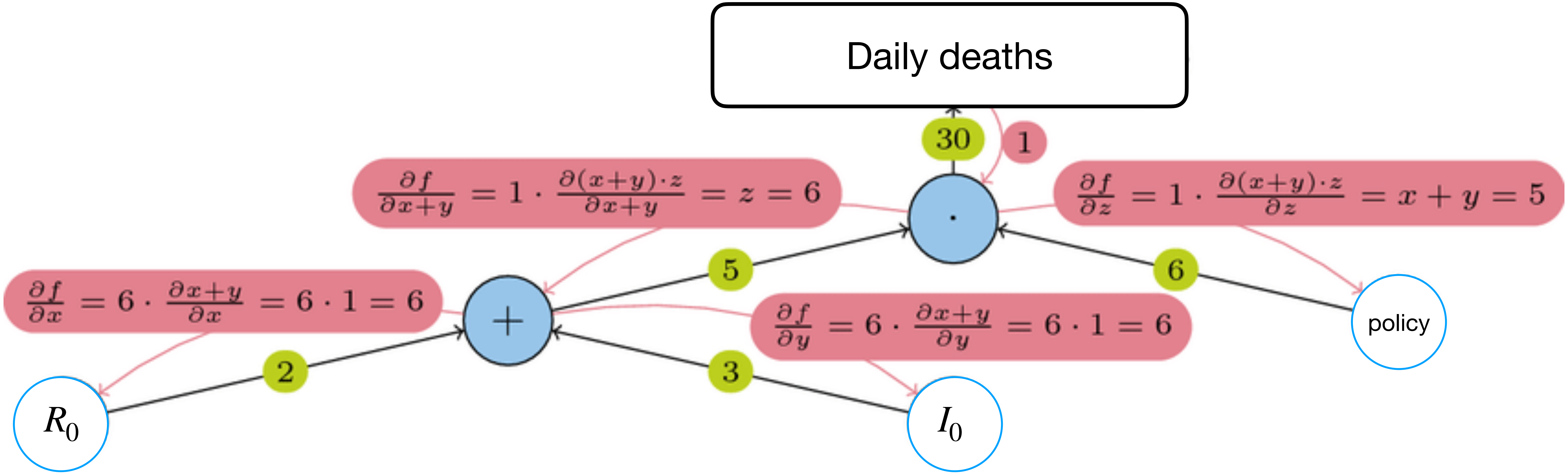


Sensitivity Analysis via Automatic Differentiation



	Simulation	Calibration (No UQ)	Bayesian Calibration	Sensitivity Analysis
JUNE	50 hours	-	100k hours	5k hours
GradABM-June (GPU)	5 seconds	20 minutes	8 hours	10 seconds

Sensitivity Analysis via Automatic Differentiation



				Sensitivity Analysis
JUNE				5k hours
GRADABM-JUNE (GPU)				10 seconds

Reverse-mode AD enables almost instant SA

The impact of uncertainty on predictions of the CovidSim epidemiological code

Wouter Edeling[1], Hamid Arabnejad[2], Robbie Sinclair[3], Diana Suleimenova[2], Krishnakumar Gopalakrishnan[3], Bartosz Bosak[4], Derek Groen[2], Imran Mahmood[2], Daan Crommelin[1,5] and Peter V. Coveney[3,6]

Epidemiological modelling has assisted in identifying interventions that reduce the impact of COVID-19. The UK government relied, in part, on the CovidSim model to guide its policy to contain the rapid spread of the COVID-19 pandemic during March and April 2020; however, CovidSim contains several sources of uncertainty that affect the quality of its predictions: parametric uncertainty, model structure uncertainty and scenario uncertainty. Here we report on parametric sensitivity analysis and uncertainty quantification of the code. From the 940 parameters used as input into CovidSim, we find a subset of 19 to which the code output is most sensitive—imperfect knowledge of these inputs is magnified in the outputs by up to 300%. The model displays substantial bias with respect to observed data, failing to describe validation data well. Quantifying parametric input uncertainty is therefore not sufficient: the effect of model structure and scenario uncertainty must also be properly understood.

Ensemble execution. Consequently, through the use of adaptive methods we make the uncertainty analysis of CovidSim tractable, but our analysis nevertheless required us to perform thousands of runs, each with its own unique set of input parameters. Specifically, we used the Eagle supercomputer at the Posnan

The impact of uncertainty on predictions of the CovidSim epidemiological code

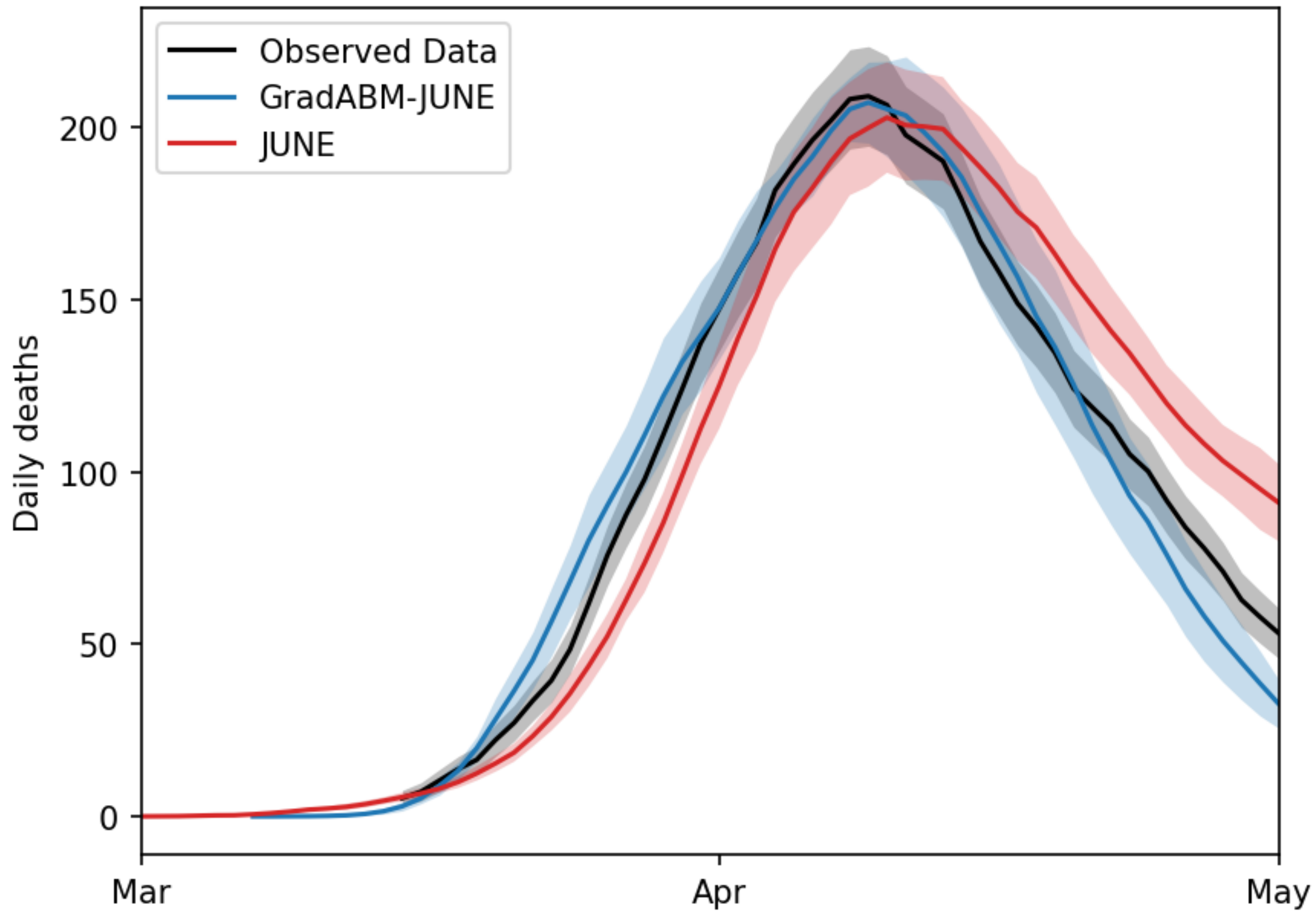
Wouter Edeling[1], Hamid Arabnejad[2], Robbie Sinclair[3], Diana Suleimenova[2], Krishnakumar Gopalakrishnan[3], Bartosz Bosak[4], Derek Groen[2], Imran Mahmood[2], Daan Crommelin[1,5] and Peter V. Coveney[3,6]

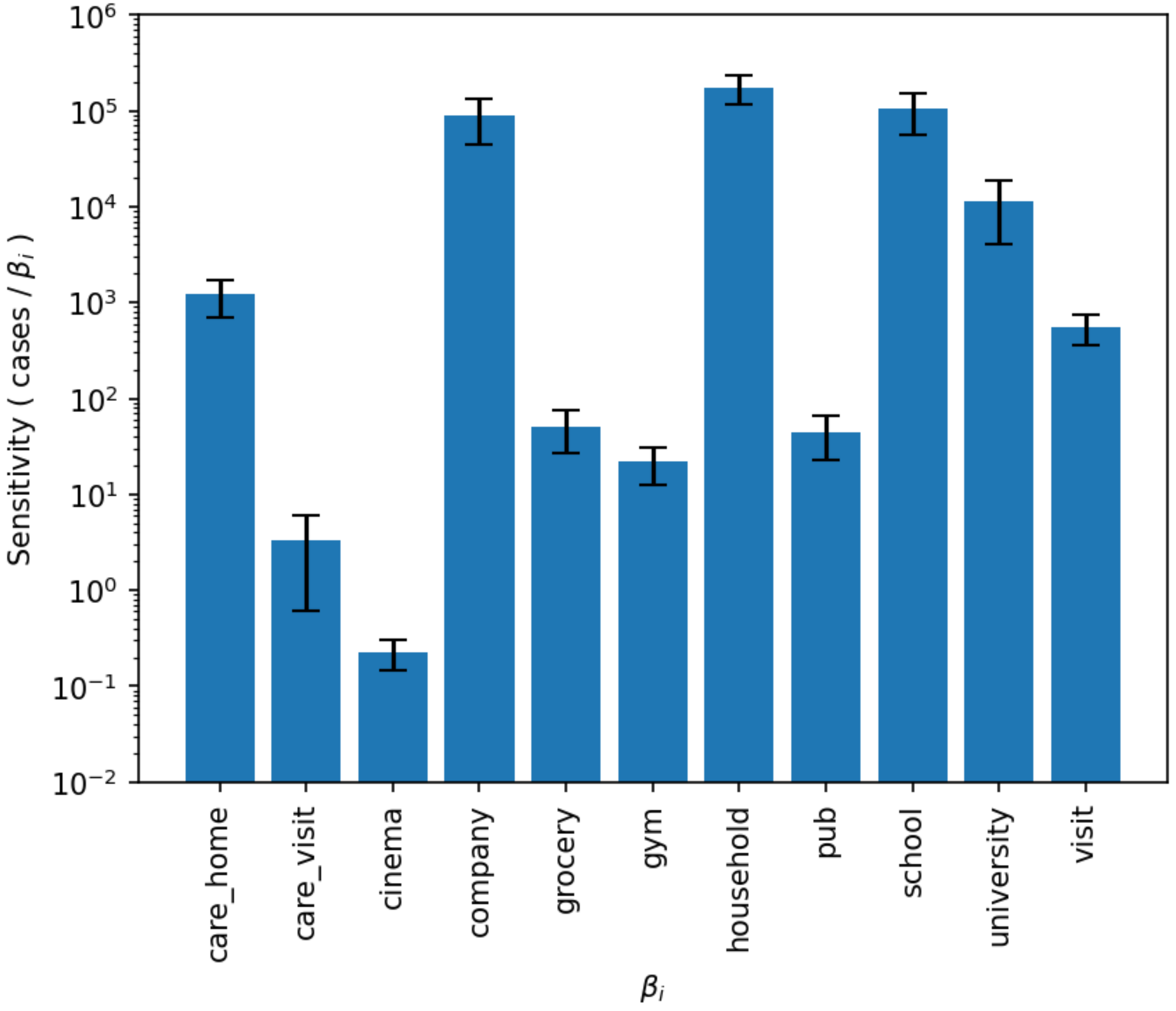
Epidemiological modelling has assisted in identifying interventions that reduce the impact of COVID-19. The UK government relied, in part, on the CovidSim model to guide its policy to contain the rapid spread of the COVID-19 pandemic during March and April 2020; however, CovidSim contains several sources of uncertainty that affect the quality of its predictions: parametric uncertainty, model structure uncertainty and scenario uncertainty. Here we report on parametric sensitivity analysis and uncertainty quantification of the code. From the 940 parameters used as input into CovidSim, we find a subset of 19 to which the code output is most sensitive—imperfect knowledge of these inputs is magnified in the outputs by up to 300%. The model displays substantial bias with respect to observed data, failing to describe validation data well. Quantifying parametric input uncertainty is therefore not sufficient: the effect of model structure and scenario uncertainty must also be properly understood.

Ensemble execution. Consequently, through the use of adaptive methods we make the uncertainty analysis of CovidSim tractable, but our analysis nevertheless required us to perform thousands of runs, each with its own unique set of input parameters. Specifically, we used the Eagle supercomputer at the Posnan

Reverse-mode AD independent of number of parameters!

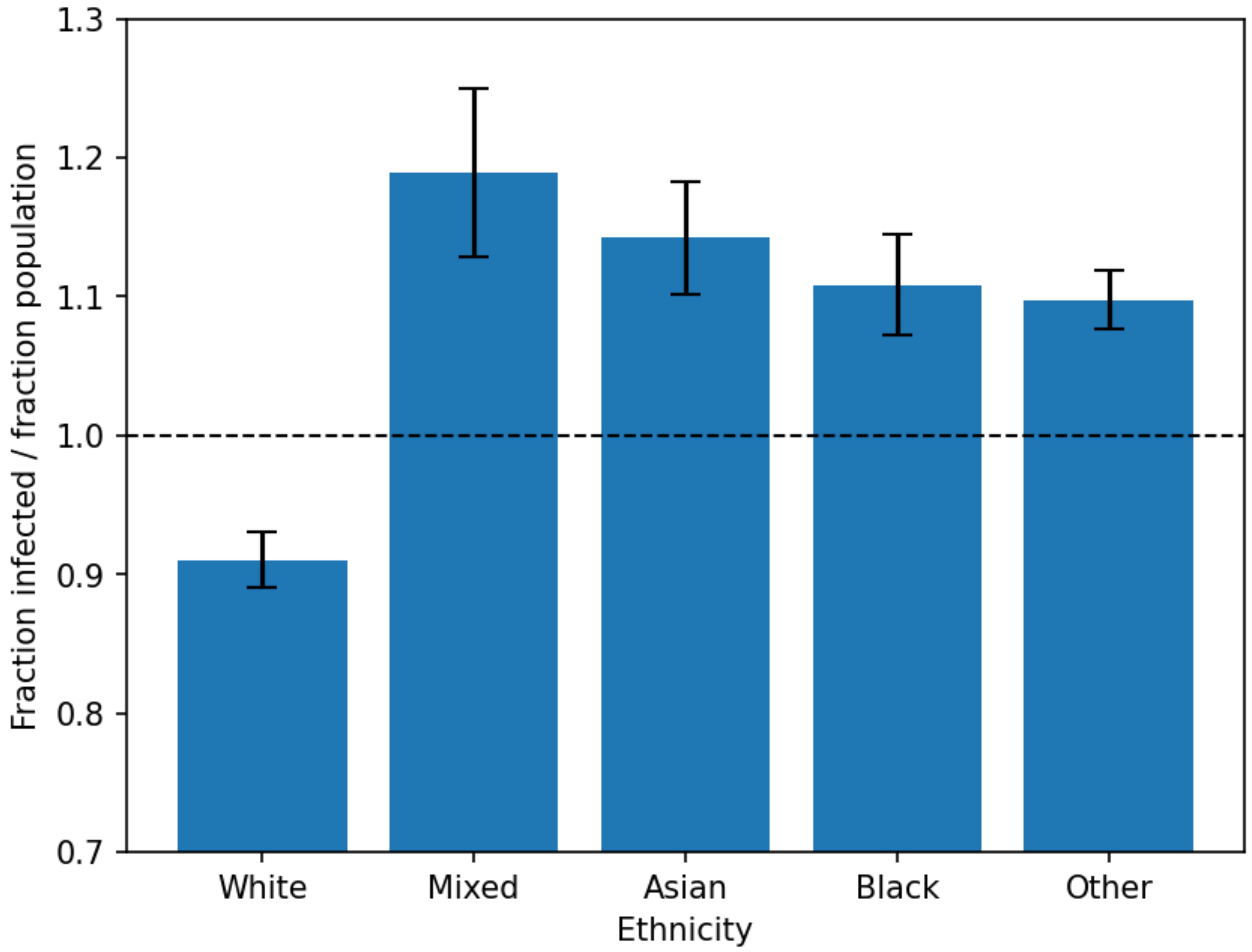
Sensitivity Analysis





Interrogating the model

JUNE recovers infection inequalities across demographic groups **without** explicit calibration



What causes this imbalance?

Analysing the sensitivity of each demographic group

Infected population in group d

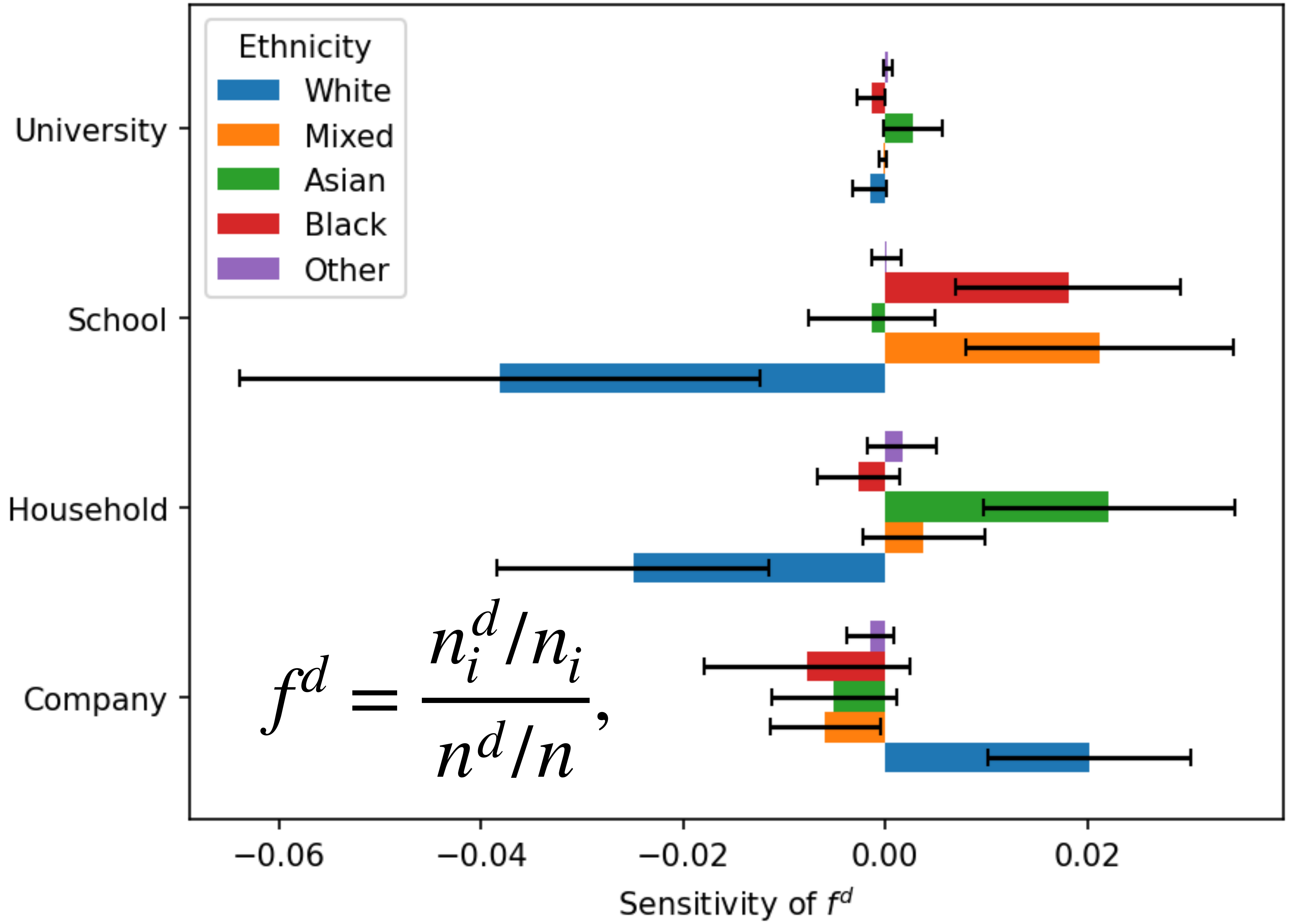
Infected population

$$f^d = \frac{n_i^d / n_i}{n^d / n},$$

Population in group d

Total population

Analysing the sensitivity of each demographic group

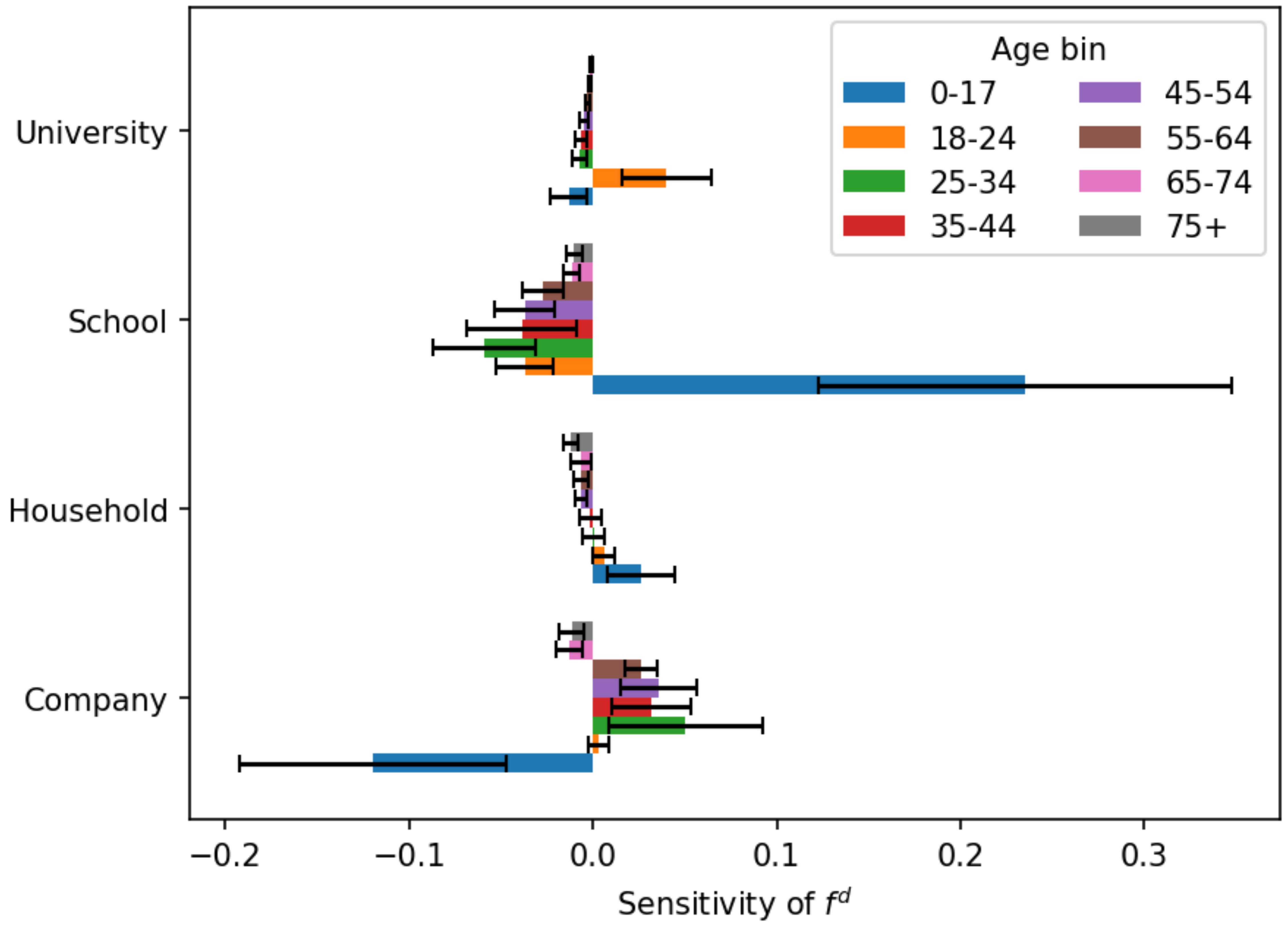


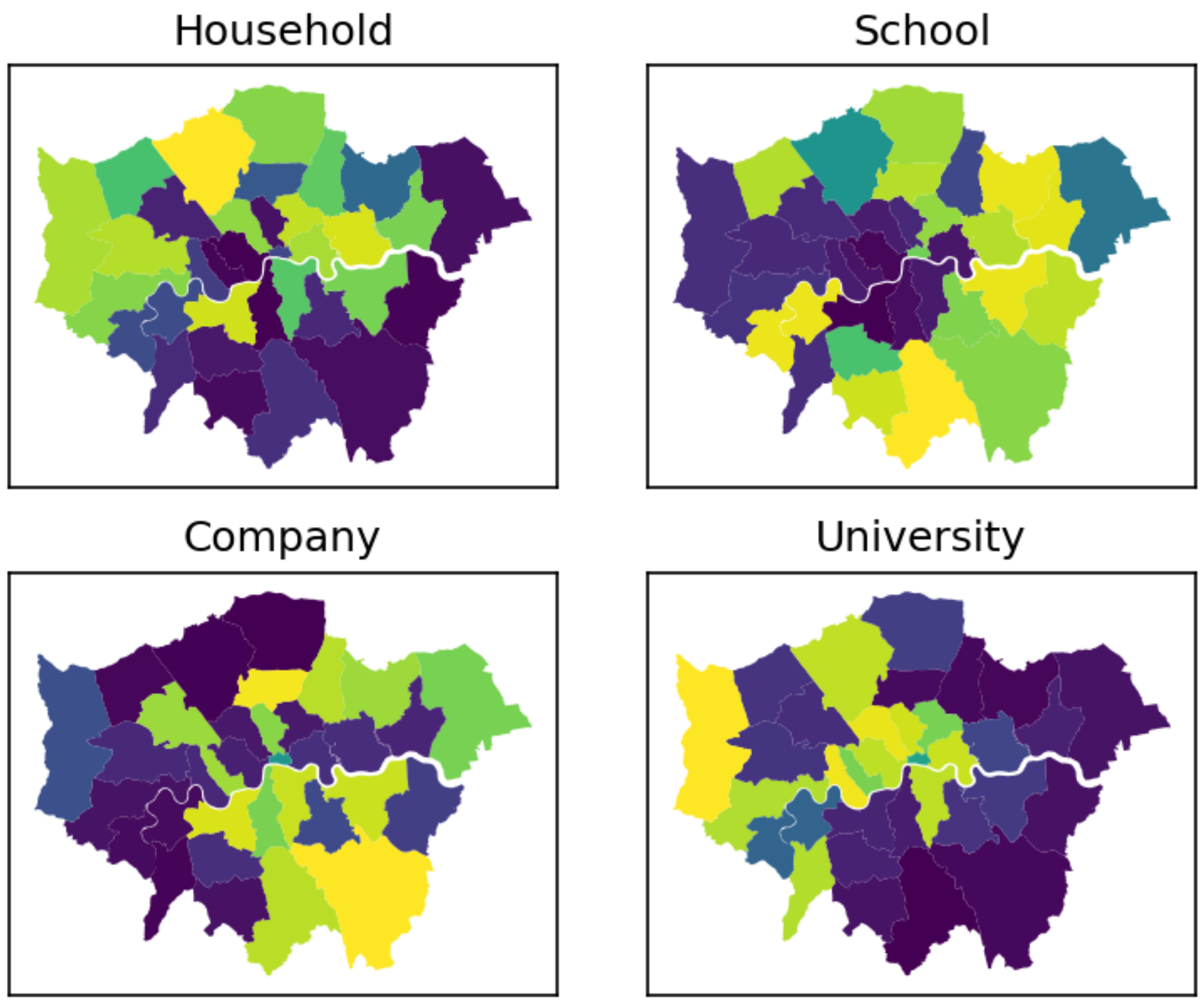
$$f^d = \frac{n_i^d / n_i}{n^d / n},$$

Some Ethnic groups are more vulnerable to infections in certain locations.

This is due to household size, work sector, family structure, etc.

Analysing the sensitivity of each demographic group





Conclusions

Differentiable agent-based models enable:

1. Fast **simulation** via tensorisation.
2. Fast and accurate Bayesian **calibration** via gradients.
3. Fast and accurate **sensitivity analyses** via gradients.

Paper + slides: www.arnau.ai/aamas23