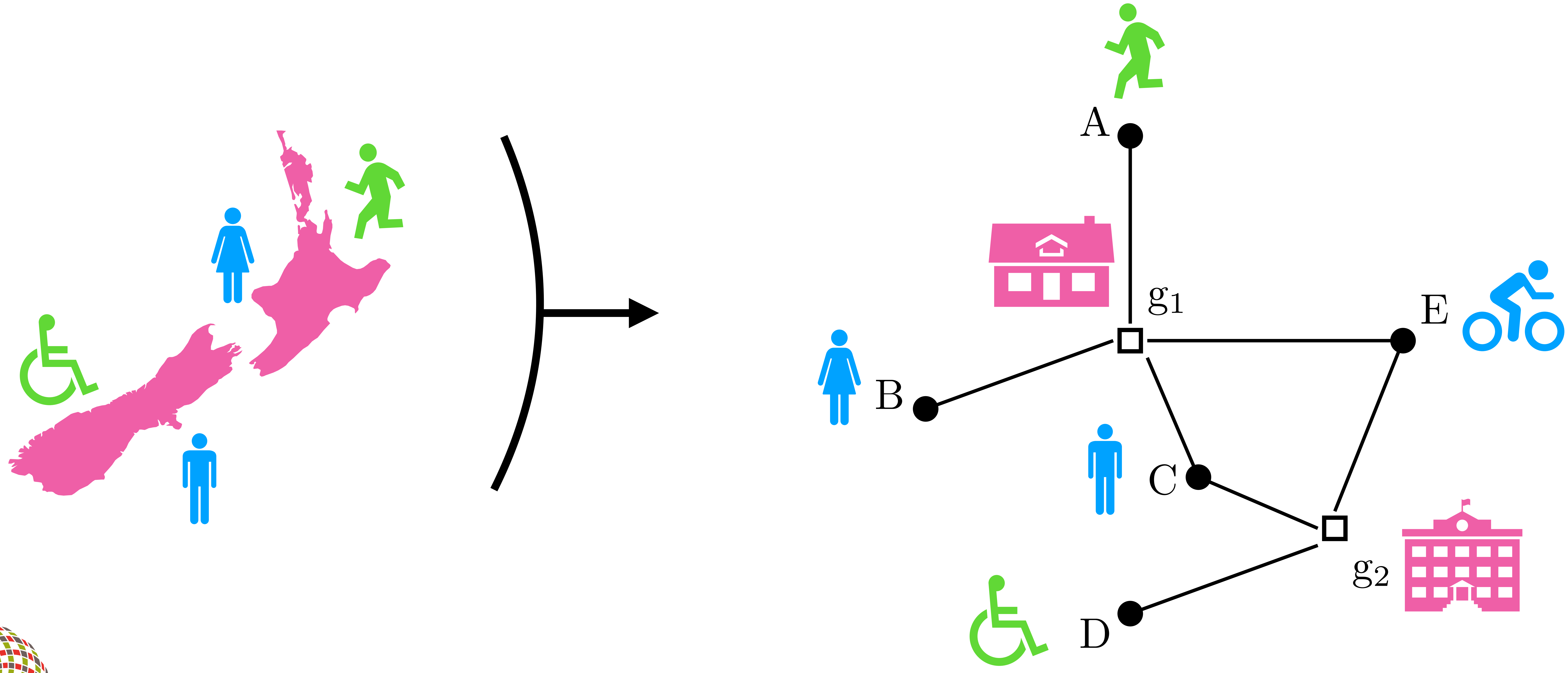
Uncertainty Quantification for Big, Complicated Network Contagion Models

Presenting: ***Oliver Maclaren*** (The University of Auckland)

Team includes: Dion O'Neale, Emily Harvey, ***Frankie Patten-Elliott***, David Wu, Adrian Ortiz Cervantes, Steven Turnbull, Demi Vasques, Tom/James Gilmour, Frank Mackenzie

Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

Context: complex, dynamic network contagion models



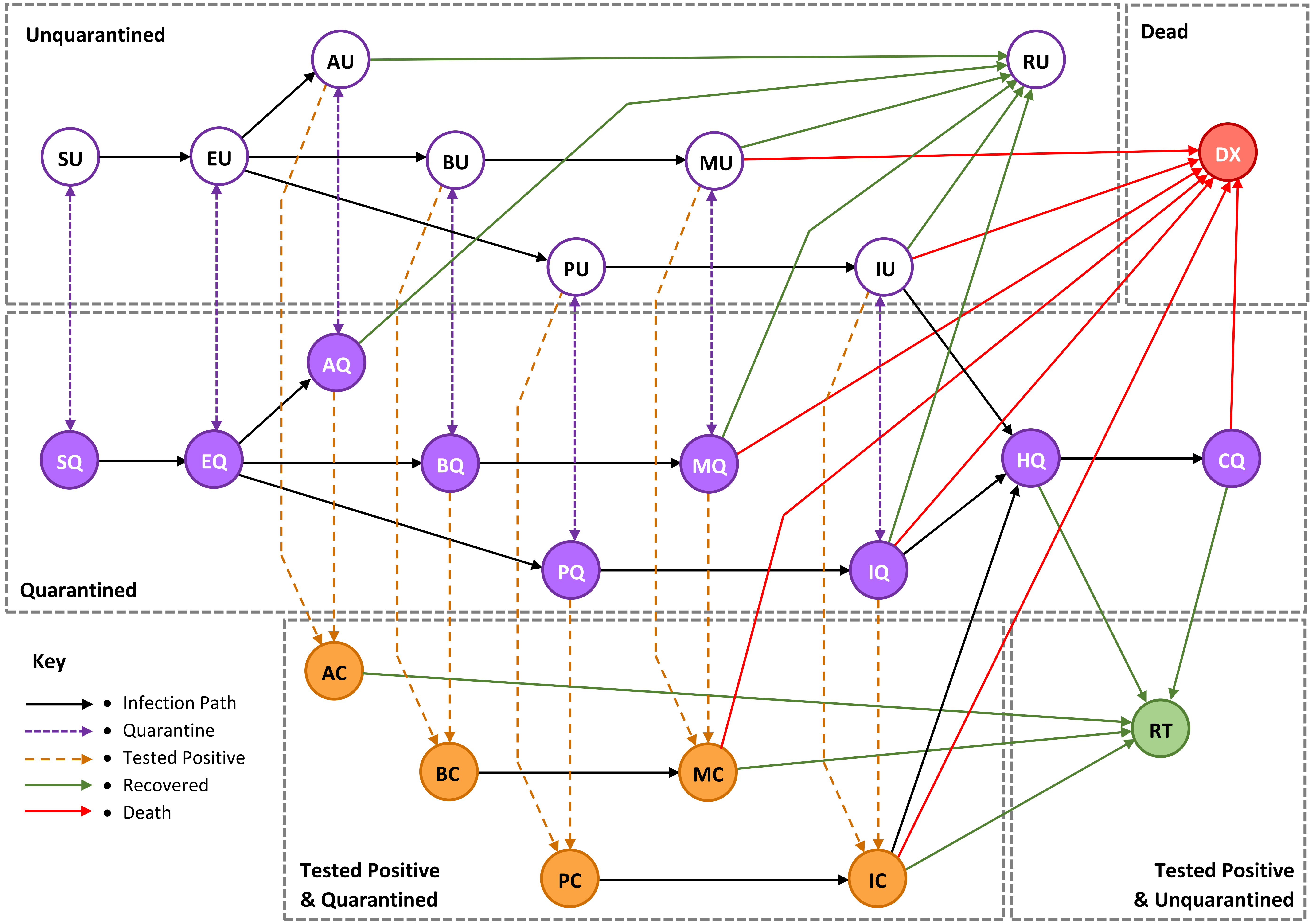
Network models

- **Everybody**! In principle…

- Their ***interaction networks***, ***stochastic*** rates of interactions, additional 'states'

- We also take a relatively novel ***bipartite approach*** to defining network: ***individuals*** and ***groups*** *or* ***interaction contexts*** (work, school, dwelling, etc.)

Under the hood: state transitions



Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

Under the hood: equations and parameters

$$r(SU \to X) = \underbrace{\beta \phi_{age} \frac{\begin{pmatrix} \varepsilon_A|AU| + \varepsilon_B|BU| + \varepsilon_M|MU| + \varepsilon_P|PU| + |IU| + \varepsilon_H|HQ| + \varepsilon_C|CQ| \\ +\varepsilon_A|AC|_{hh} + \varepsilon_B|BC|_{hh} + \varepsilon_M|MC|_{hh} + \varepsilon_P|PC|_{hh} + |IC|_{hh} \\ +\omega(\varepsilon_A|AC|_{\cancel{hh}} + \varepsilon_B|BC|_{\cancel{hh}} + \varepsilon_M|MC|_{\cancel{hh}} + \varepsilon_P|PC|_{\cancel{hh}} + |IC|_{\cancel{hh}}) \end{pmatrix}}{|N|}}_{\to EU}$$

$$r(EU \to X) = \underbrace{p_A\gamma}_{\to AU} + \underbrace{p_M\gamma}_{\to BU} + \underbrace{(1-p_A)(1-p_M)\gamma}_{\to PU}$$

$$r(AU \to X) = \underbrace{\alpha_A}_{\to RU} + \underbrace{\theta_0}_{\to AC}$$

$$r(BU \to X) = \underbrace{\delta_M}_{\to MU} + \underbrace{\theta_0}_{\to BC}$$

$$r(MU \to X) = \underbrace{\mu_M\alpha_{MD}}_{\to DX} + \underbrace{(1-\mu_M)\alpha_{MR}}_{\to RU} + \underbrace{\theta_M}_{\to MC}$$

$$r(PU \to X) = \underbrace{\delta_I}_{\to IU} + \underbrace{\theta_0}_{\to PC}$$

$$r(IU \to X) = \underbrace{\mu_I\alpha_{ID}}_{\to DX} + \underbrace{\eta_H\delta_H}_{\to HQ} + \underbrace{(1-\mu_I)(1-\eta_H)\alpha_{IR}}_{\to RU} + \underbrace{\theta_I}_{\to IC}$$

$$r(AC \to X) = \underbrace{\alpha_A}_{\to RT}$$

$$r(BC \to X) = \underbrace{\delta_M}_{\to MC}$$

$$r(MC \to X) = \underbrace{\mu_M\alpha_{MD}}_{\to DX}\underbrace{(1-\mu_M)\alpha_{MR}}_{\to RT}$$

$$r(PC \to X) = \underbrace{\delta_I}_{\to IC}$$

$$r(IC \to X) = \underbrace{\mu_I\alpha_{ID}}_{\to DX} + \underbrace{\eta_H\delta_H}_{\to HQ} + \underbrace{(1-\mu_I)(1-\eta_H)\alpha_{IR}}_{\to RT}$$

$$r(HQ \to X) = \underbrace{\eta_C\delta_C}_{\to CQ} + \underbrace{(1-\eta_C)\alpha_{HR}}_{\to RT}$$

$$r(CQ \to X) = \underbrace{(1-\mu_c)\alpha_{CR}}_{\to RT} + \underbrace{\mu_c\alpha_{CD}}_{\to DX}$$

Symbol	Value	Units	Description	Source
ε_B	0.5	–	Relative infectivity of presymptomatic mild individuals	Earlier TPM modelling[1,2]. More up to date estimates required.
ε_C	0.0	–	Relative infectivity of critically hospitalised individuals	Model assumption
ε_H	0.0	–	Relative infectivity of hospitalised individuals	Model assumption
ε_M	0.72	–	Relative infectivity of mild symptomatic individuals	Fraser Group parameters[16]. More up to date estimates required.
ε_P	0.5	–	Relative infectivity of presymptomatic moderate/severe individuals	Earlier TPM modelling[1,2]. More up to date estimates required.
η_C	0.05, 0.05, 0.1, 0.5	per day	Age structured rates at which hospitalised cases becoming critically hospitalised for age bands 0–14, 15–29, 30–64 and 65+ respectively (**H** → **C**)	Estimated from Fraser Group parameters[16] from the UK.
η_H	0.003, 0.012, 0.222, 0.4	per day	Age structured rates of hospitalisation for age bands 0–14, 15–29, 30–64 and 65+ respectively (**I** → **H**)	Estimated from Fraser Group parameters[16] from the UK.
γ	1	per day	Latent period rate, rate of becoming infectious from contracting infection (**E** → **P**)	Existing TPM modelling[1,2]
μ_C	0.3, 0.417, 0.571, 0.88	per day	Age structured mortality rates of critically hospitalised individuals for age bands 0–14, 15–29, 30–64 and 65+ respectively	Estimated from Fraser Group parameters[16] from the UK.
μ_I	0.0	–	Mortality probability for moderate/severely infected individuals	Assume deaths only occur from hospital and critical care
μ_M	0.0	–	Mortality probability for mildly infected individuals	Assume deaths only occur from hospital and critical care.
ω	0.01	–	Relative risk of infection from confirmed cases in isolation	Related to compliance and ability to self-isolate. Not well known.
p_A	0.434, 0.391, 0.283, 0.215	–	Age structured proportions of infectious cases that are asymptomatic for age bands 0–14, 15–29, 30–64 and 65+ respectively	Estimated from Fraser Group parameters[16]
p_M	0.551, 0.583, 0.601, 0.518	–	Age structured proportions of infectious cases that are mild for age bands 0–14, 15–29, 30–64 and 65+ respectively	Estimated from Fraser Group parameters[16]

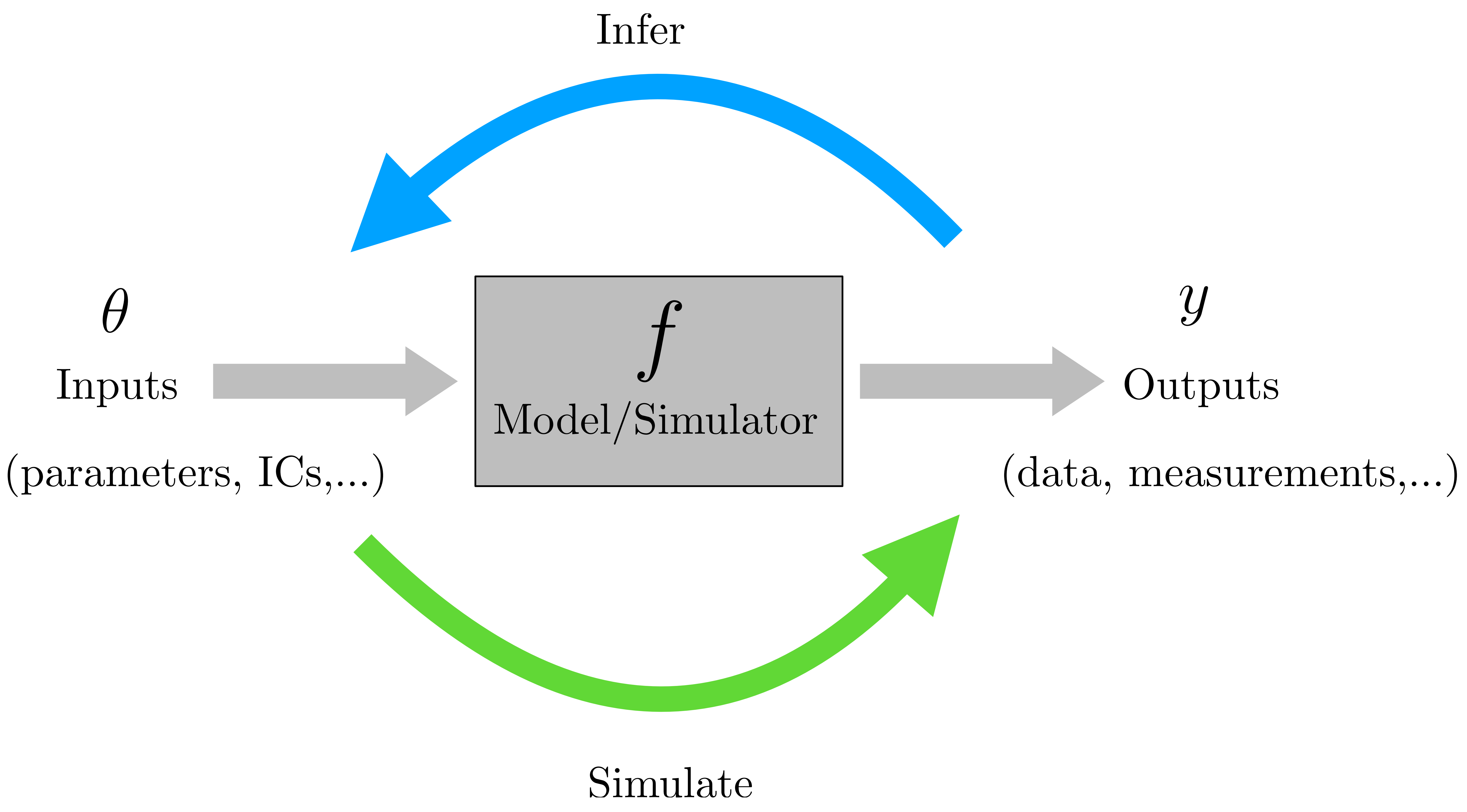
Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

An incomplete selection!

Under the hood: complexity!

- 25+ states for ***each*** of ~5 million individuals
- Rate ***parameters for all transitions*** plus ***modifiers*** depending on demographic properties, group context type, policy settings etc
- Some ***non-Markovian*** processes (delay processes, semi-Markov processes etc)

Using the model as an input-output machine: forward and 'inverse' tasks



Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

Types of uncertainty

- Uncertainty in ***parameters*** (inverse or inferential uncertainty)
- Uncertainty in ***predictions*** (predictive uncertainty)
- Uncertainty in the ***model representation*** (misspecification)
- Uncertainty in the ***data*** compared to (noise, censoring, selection bias etc)

Minimalist goals

- Today I'm just going to talk mostly about ***comparing our high-level model predictions with data***, and ***updating of predictions as we collect more data***

    - **Prior** (before data) **distribution of predictions** compared to subsequent data

    - **Posterior** (after data) **distributions of predictions** as data arrives

- Will also mention a little about updating **parameter distributions** given data

Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

Challenges for complex models

- Besides the usual (sheer complexity, difficulties with identifiability, inherent sensitivity of epidemic predictions etc), our model is ***expensive to run***!

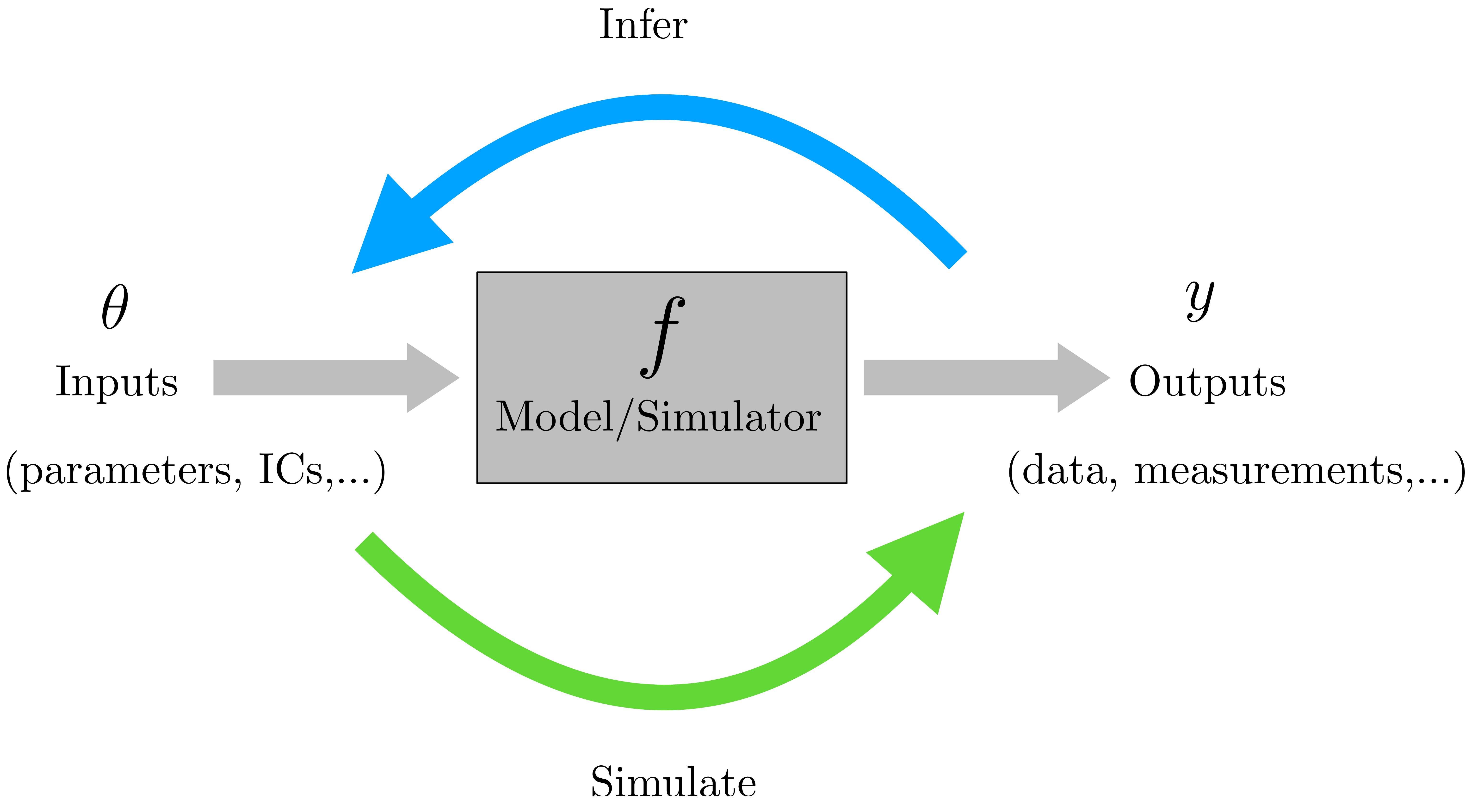
- We hence want to ***make the most of the runs we do have***

Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

Approach

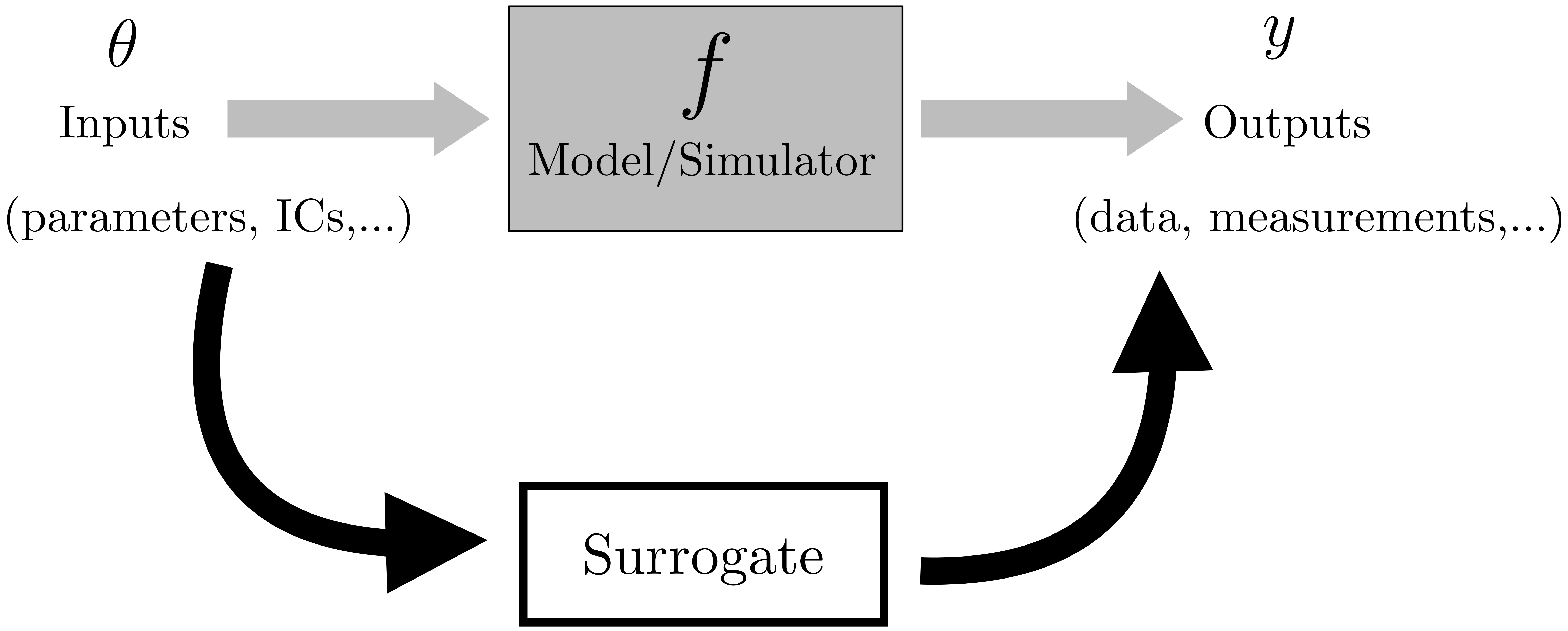
Simulation-based inference using a **surrogate model/ emulator**

Surrogate modelling



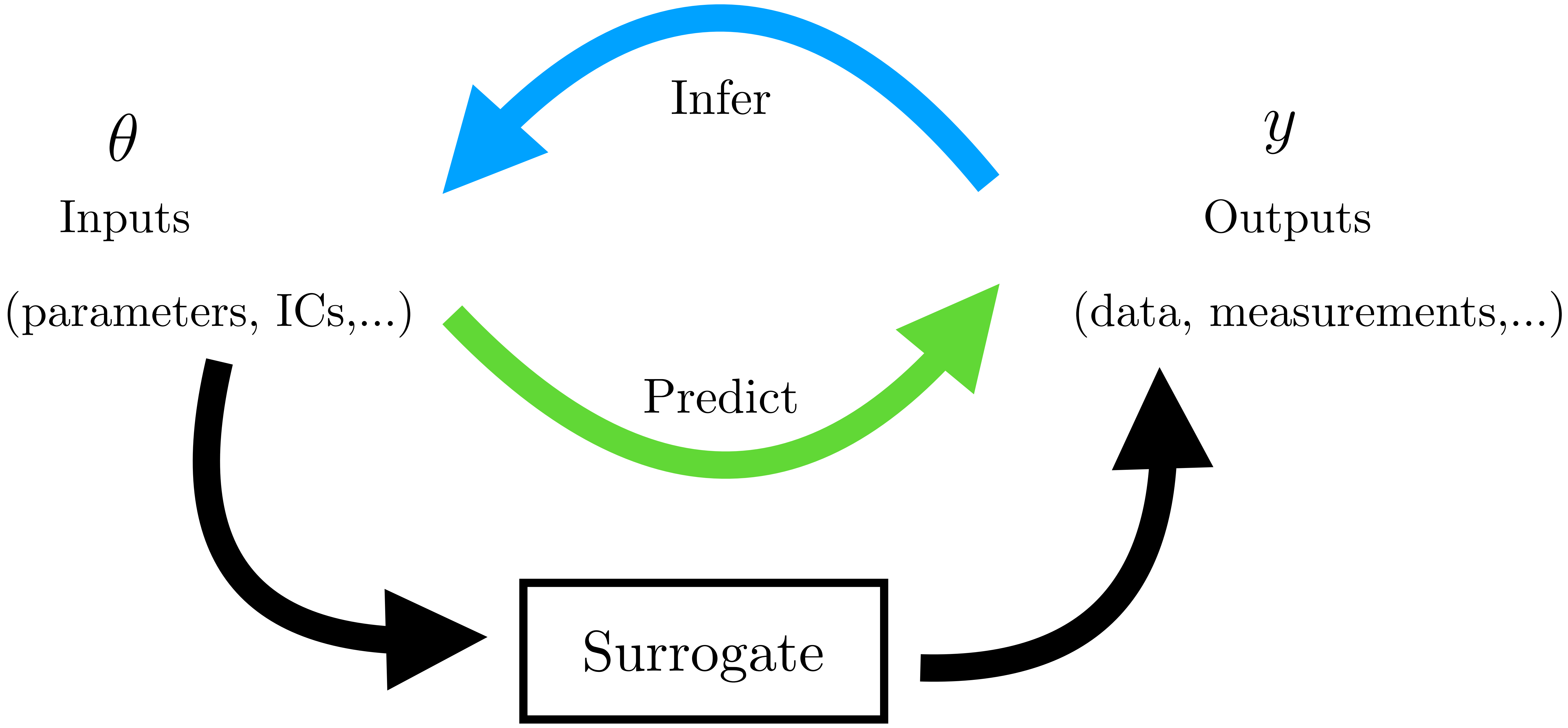
Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

Surrogate modelling



Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

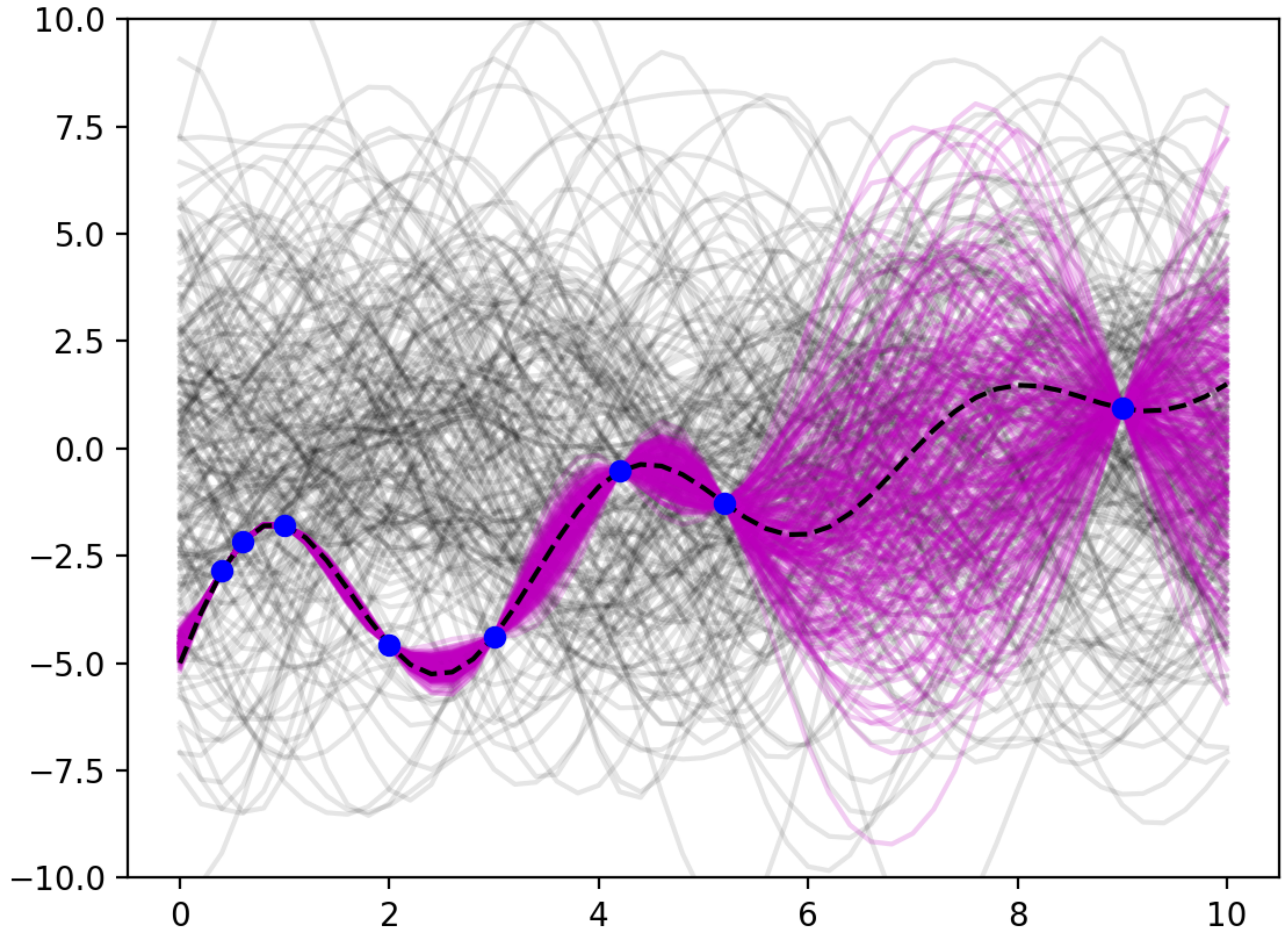
Surrogate modelling



Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

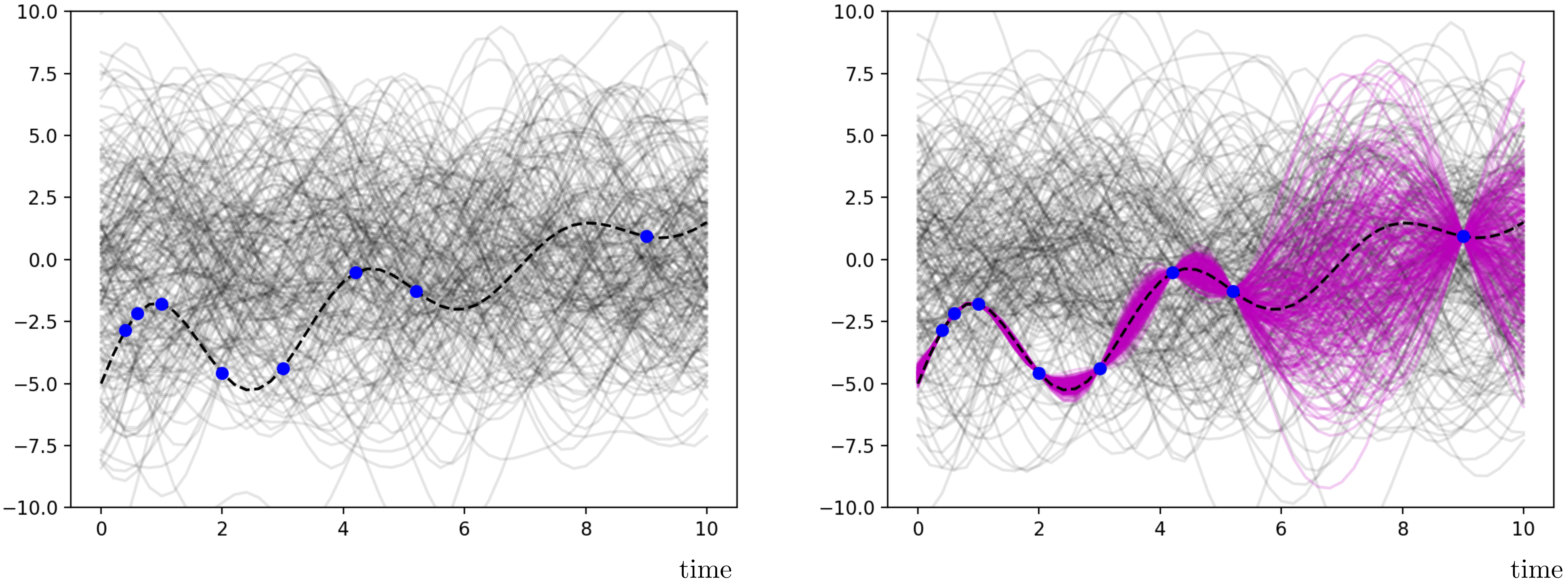
Gaussians and Gaussian processes

- Computationally ***cheap and easy*** to fit,
- Surprisingly ***flexible***,
- Naturally model and incorporate ***uncertainty***
- Easy to ***condition (update)*** given new data



Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

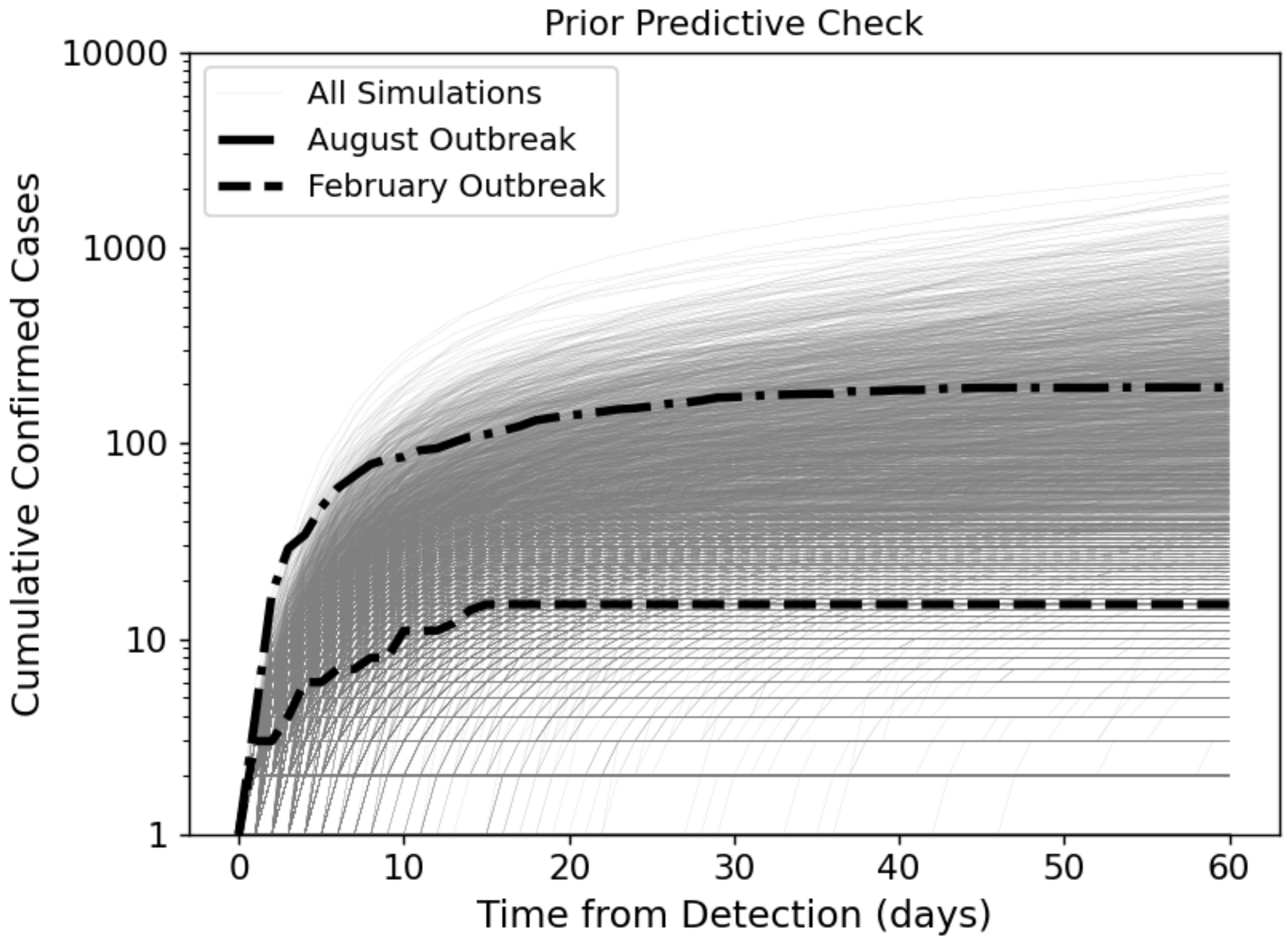
Gaussians and Gaussian processes



Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

Easy to update with (condition on) new data

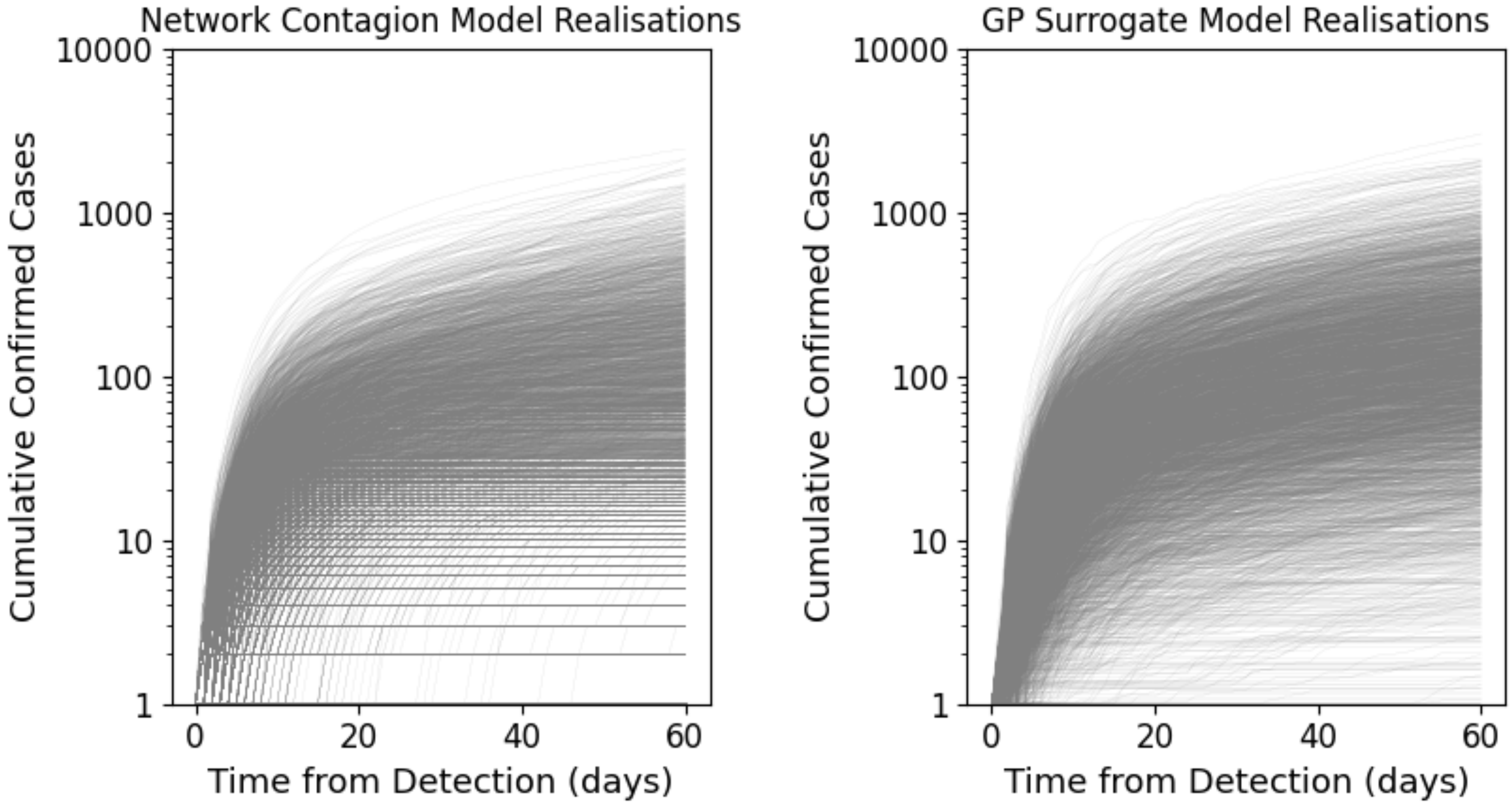
Results: prior predictive checks of original network sims



Te Pūnaha Matatini

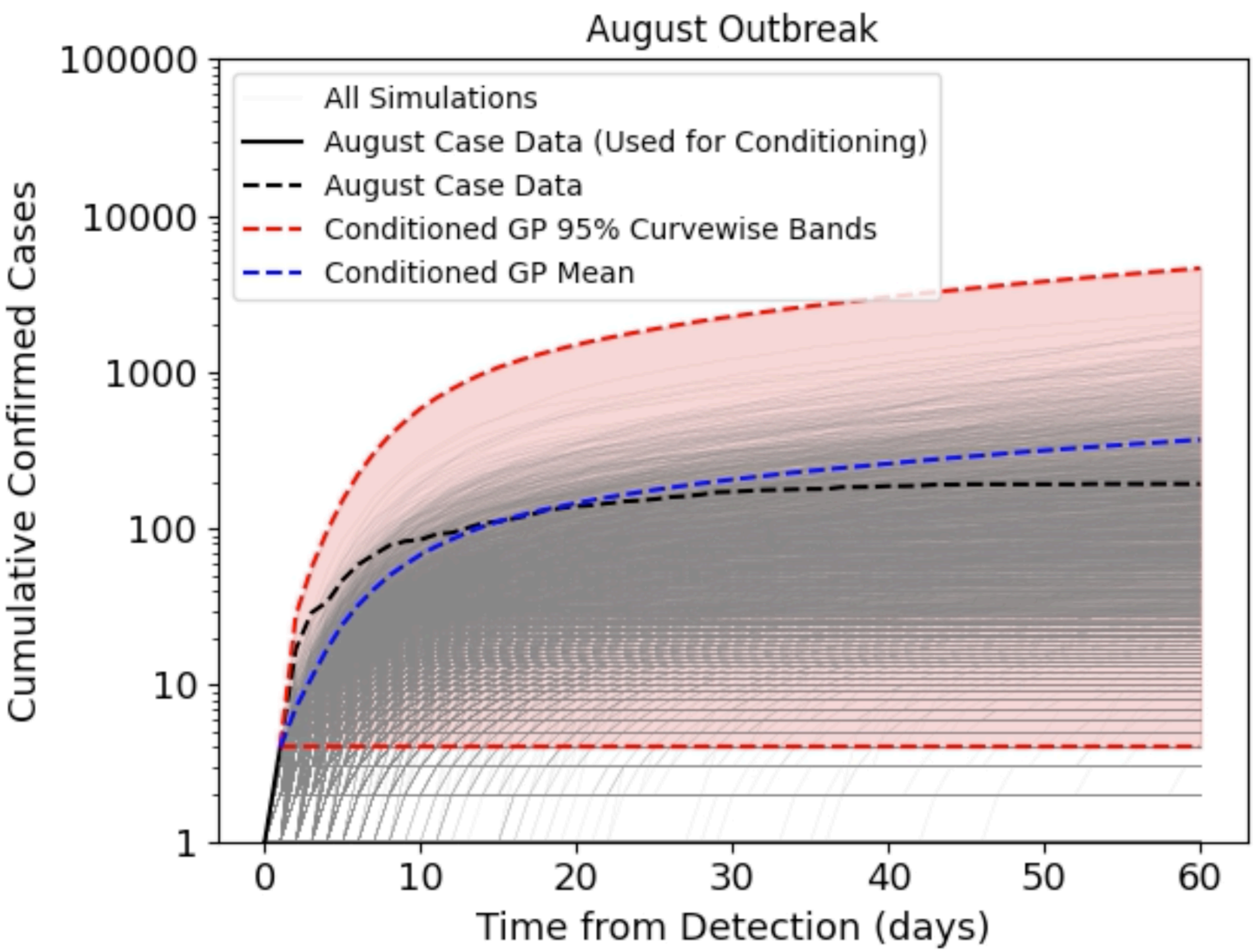
Data ▪ Knowledge ▪ Insight

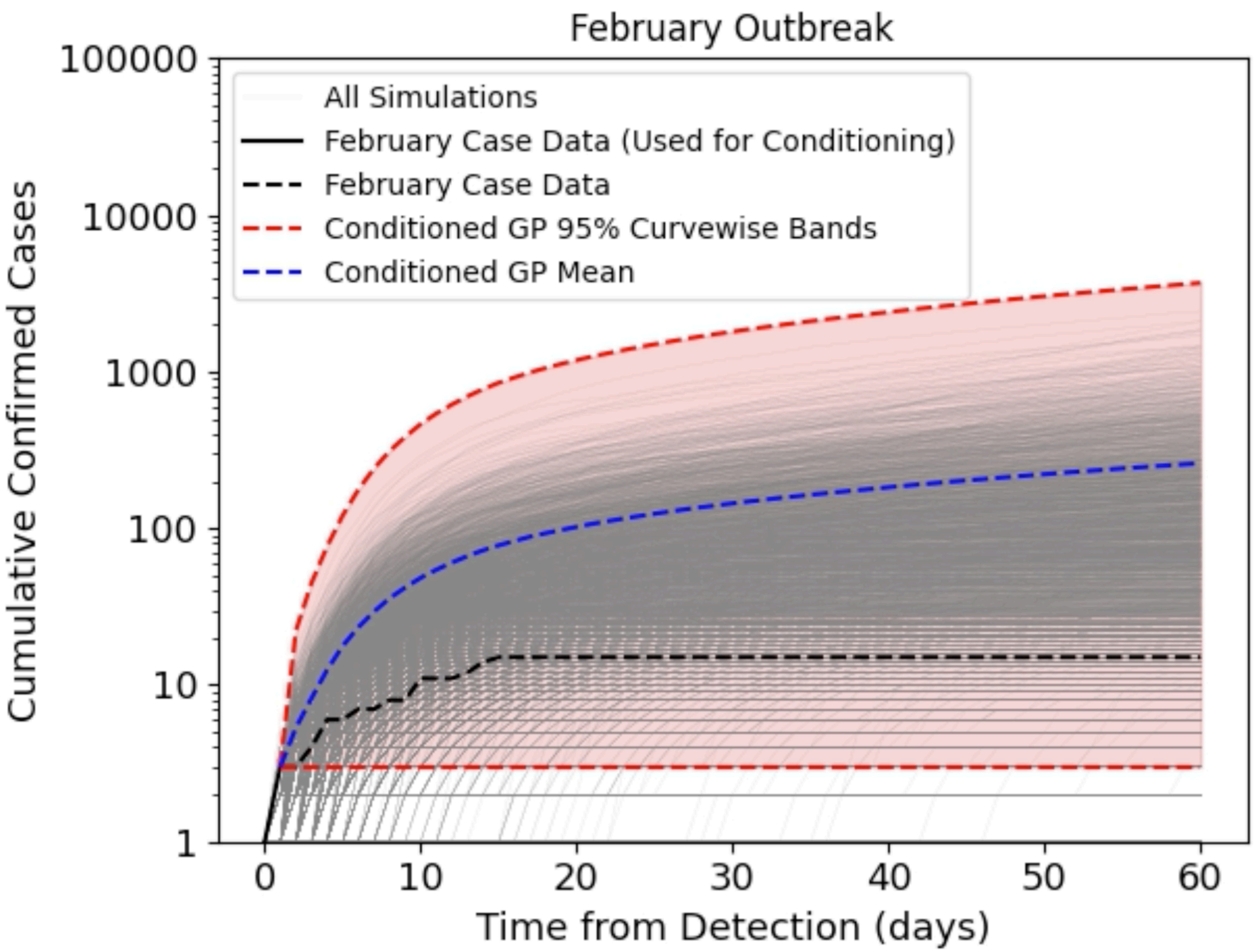
Results: pre-data emulation of network behaviour



Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

Example results: pre-data emulation of network behaviour

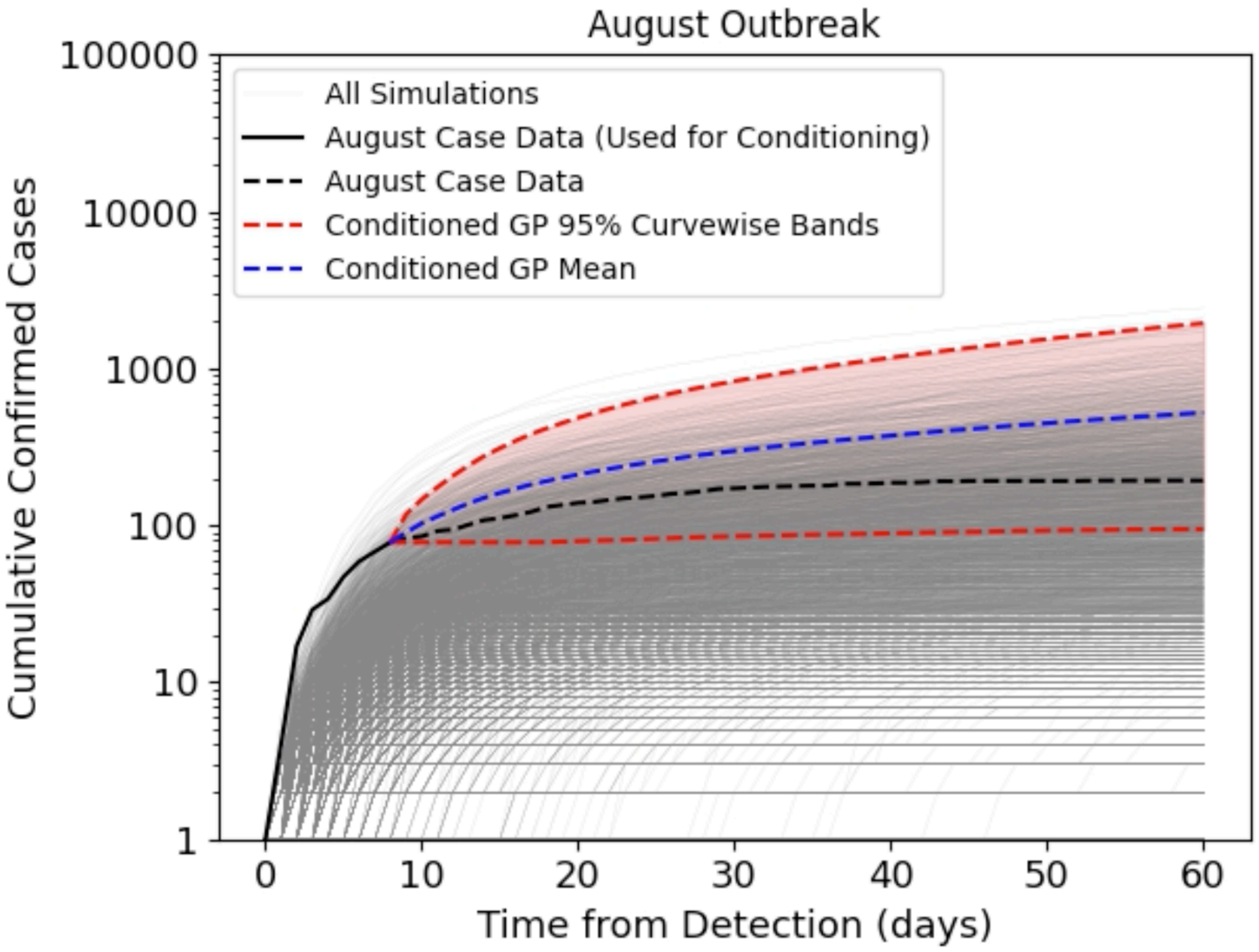


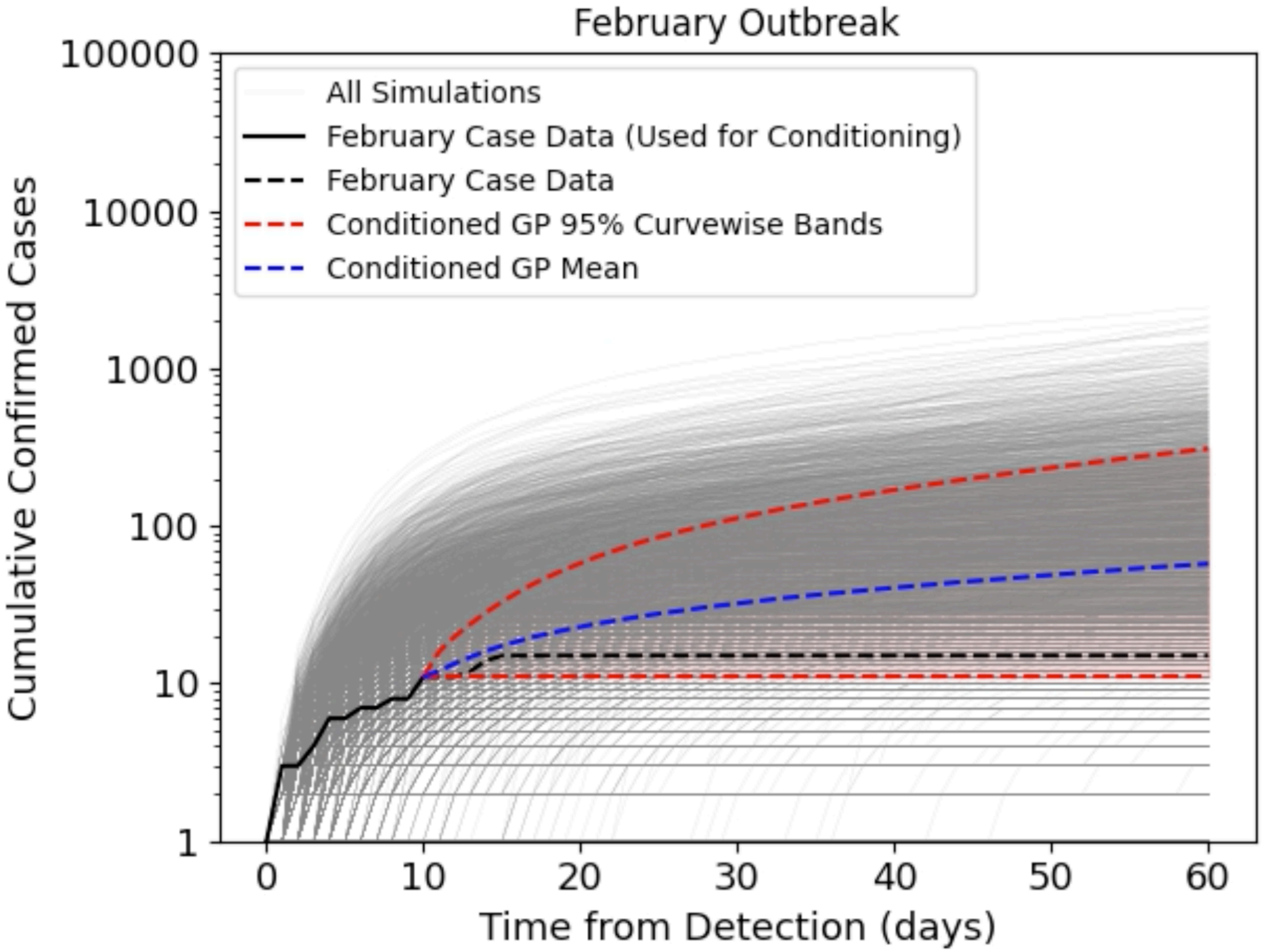


August 2020 **Feb 2021**

Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

Example results: updating prediction bands by conditioning on new data



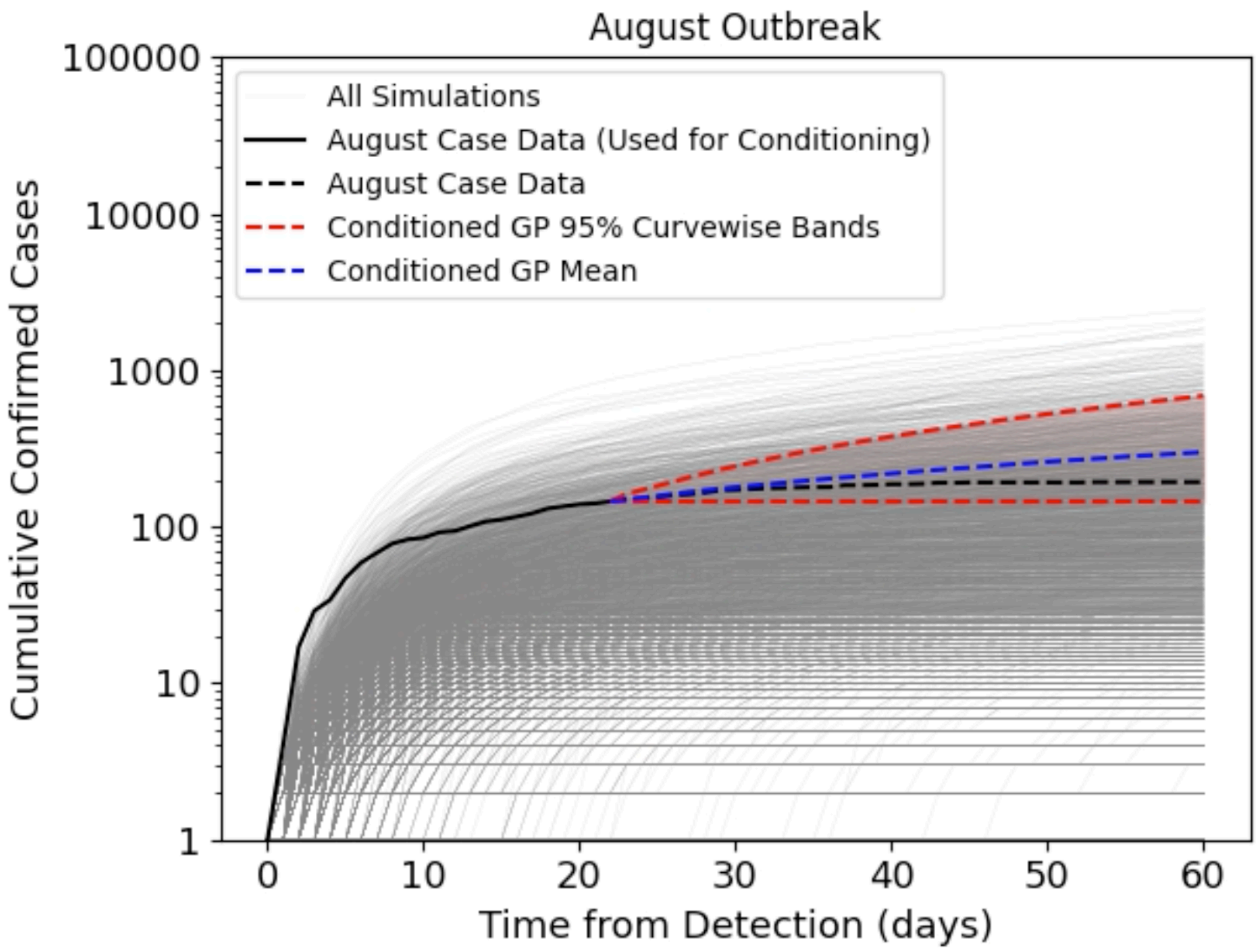


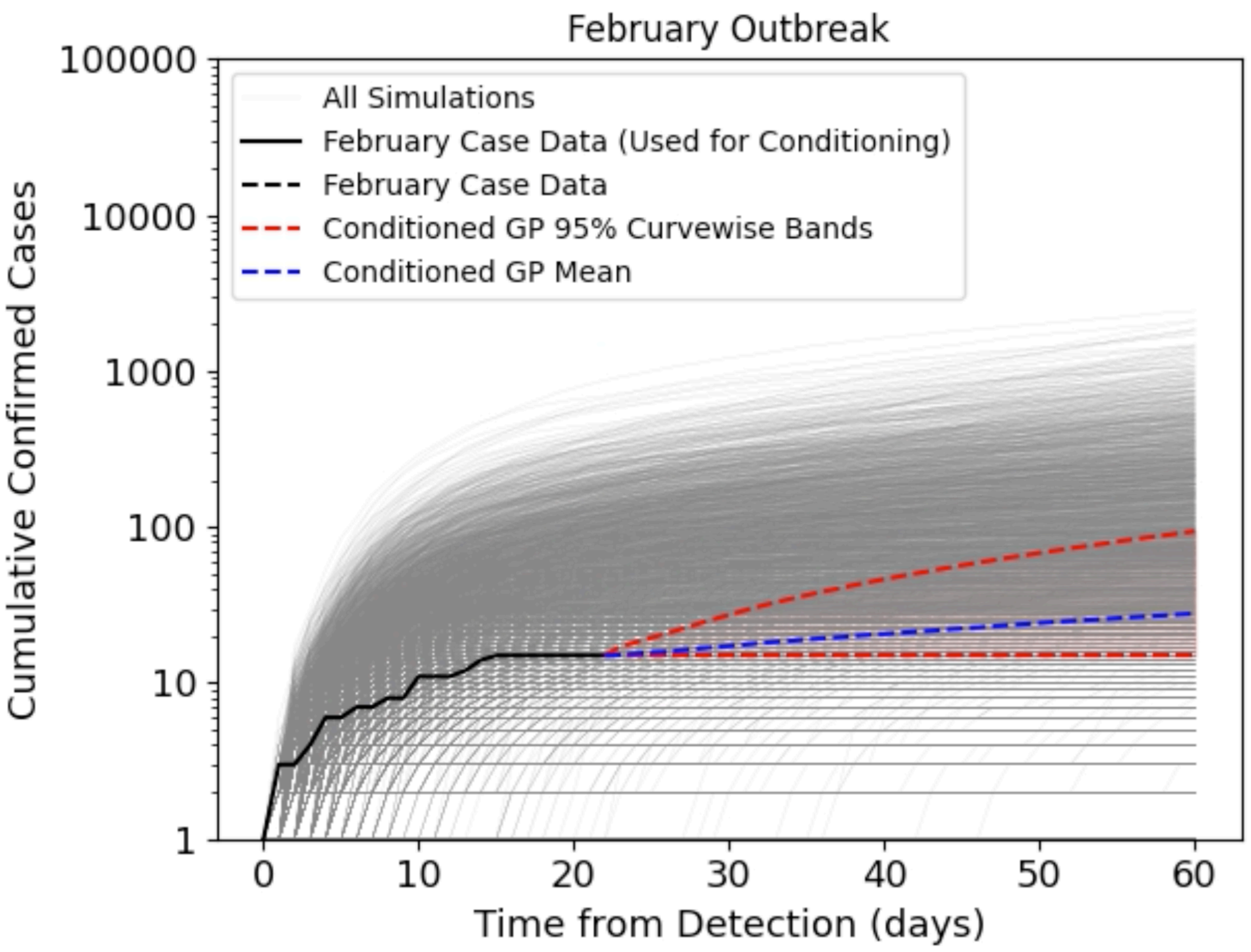
Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

August 2020

Feb 2021

Example results: updating prediction bands by conditioning on new data



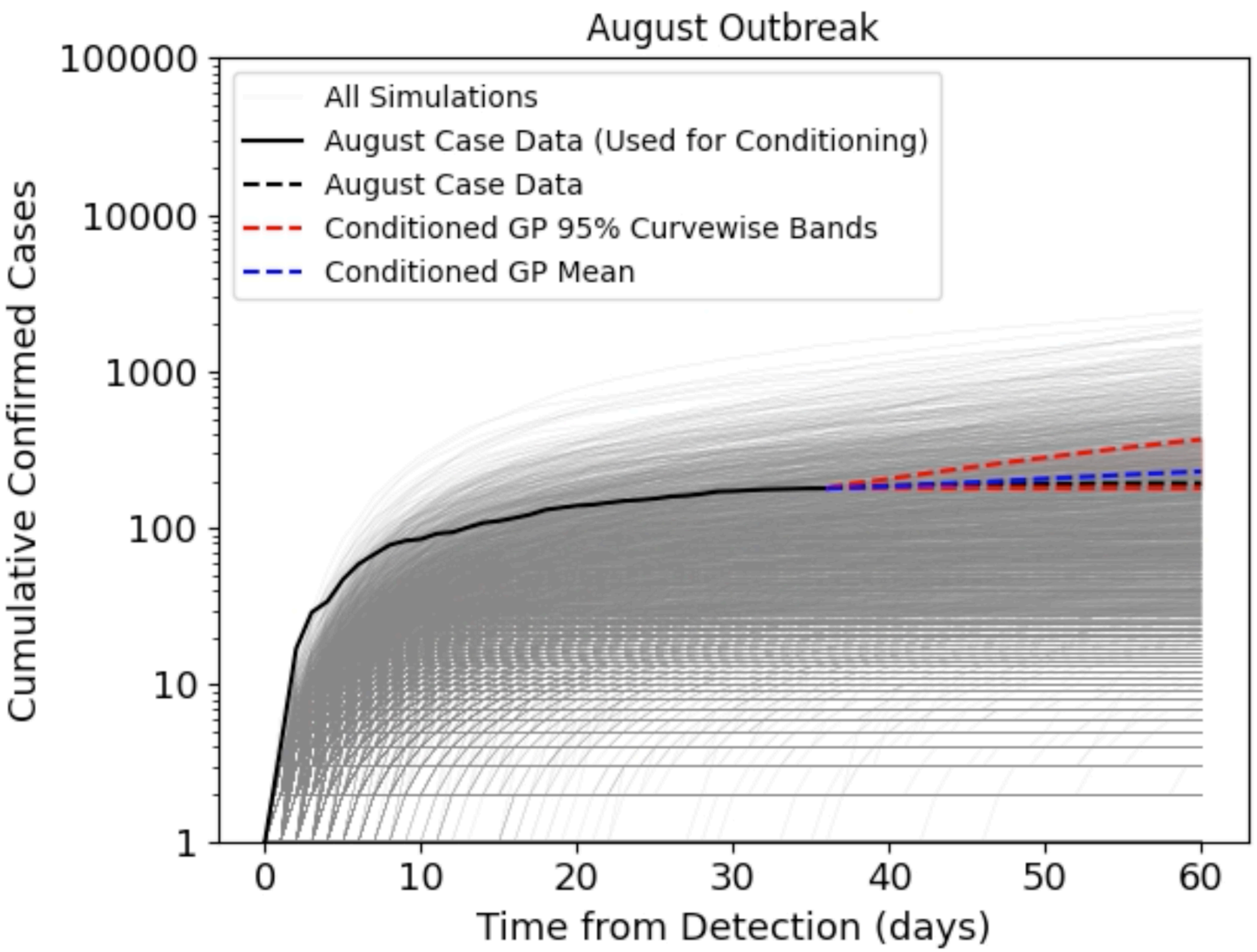


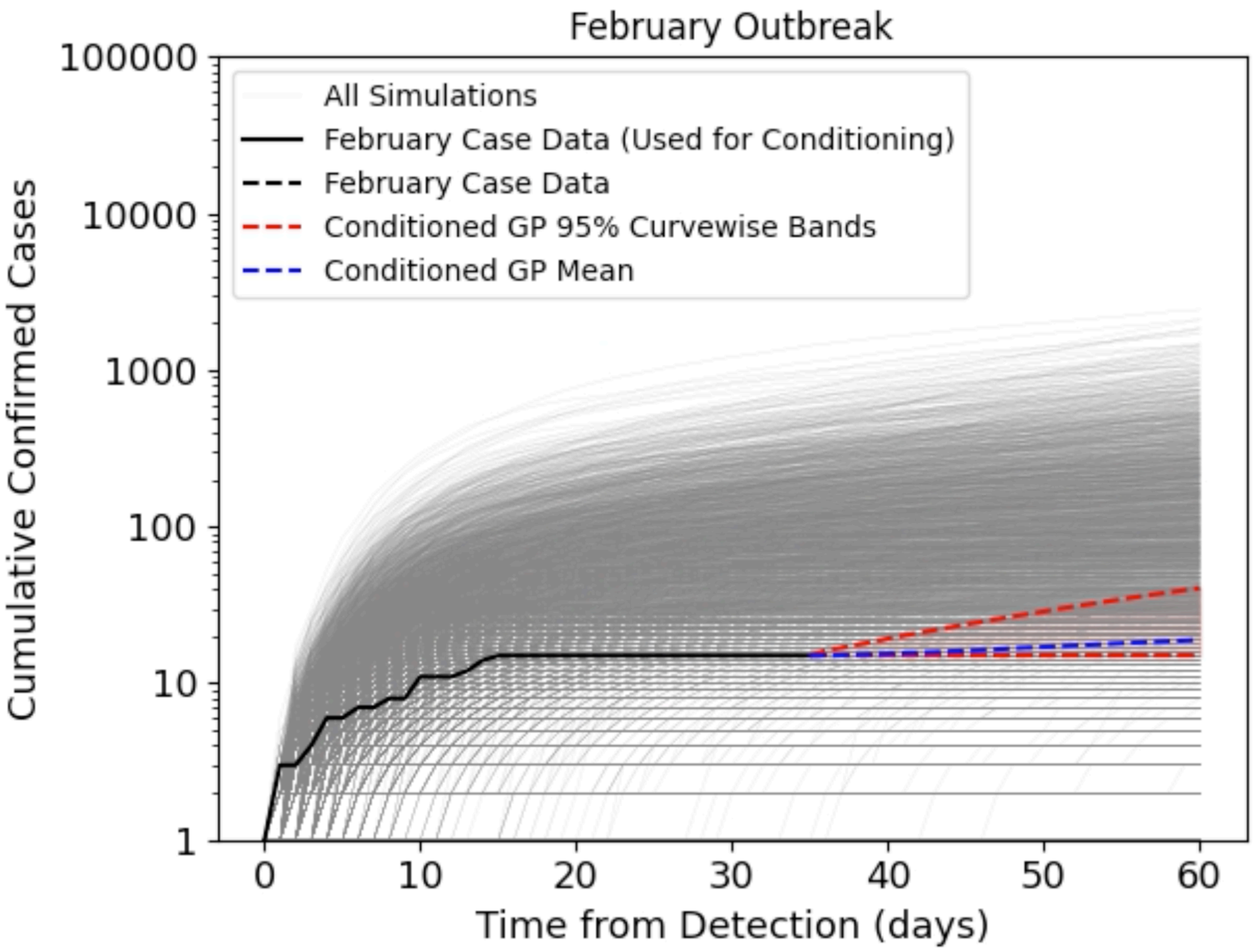
August 2020

Feb 2021

Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

Example results: updating prediction bands by conditioning on new data





Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

August 2020

Feb 2021

Limitations and future work

- Naive use of Gaussians ***doesn't naturally respect shape and other constraints*** that typically arise in the context of epidemics and associated stochastic processes (e.g. counting processes)
  - Can enforce *ad hoc*, or ***blend with differential equation***-based information on epidemic trajectories (see recent work on measles in Samoa)
- Alternatively, could use ***other surrogate stochastic processes*** that share pros of Gaussian processes but more naturally accommodate typical shape constants
- ***A lot*** of work still to do on more ***fine-grained uncertainty analysis, parameter inference and identifiability analysis!***

Thanks!