# Epidemic Forecasting Software: Lessons from the COVID-19 Pandemic

Canadian Institute for Health Information – 2024-10-10

Steve Walker (McMaster University)

# Acknowledgements


- Canadian Institute for Health Information for the invitation
- CANMOD 😢 and PHAC for funding
- Original developers were real modellers (Ben Bolker, Jonathan Dushoff, David Earn, Mike Li, Irena Papst)
- C++ programming (Weiguang Guan, Sharcnet), model library (Jen Freeman), and product models (Darren Flynn-Primrose)
- Amy Hurford, Lisa Kanary, Caroline Colijn, Zahra Cook, Claude Nadeau, Philippe Berthiaume, Evan Mitchell, Brian Gaas, Kevin Zhao, Maya Earn
- Irena Papst and Mike Li

# Story

- Why Models?
- Background
- Motivation
- Software
- Model Library
- Calibration
- Illustrations
- Open-Source Tools

# Story

- Why Models? : Why even bother at all with modelling?
- Background
- Motivation
- Software
- Model Library
- Calibration
- Illustrations
- Open-Source Tools

# Story

- Why Models?
- Background : McMasterPandemic software was developed for COVID-19 forecasting.
- Motivation
- Software
- Model Library
- Calibration
- Illustrations
- Open-Source Tools

# Story

- Why Models?
- Background
- Motivation : The urgency of producing regular forecasts during the pandemic impeded software development.
- Software
- Model Library
- Calibration
- Illustrations
- Open-Source Tools

# Story

- Why Models?
- Background
- Motivation
- Software : The `macpan2` project is a more versatile public health modelling tool being actively developed to address these challenges.
- Model Library
- Calibration
- Illustrations
- Open-Source Tools

# Story

- Why Models?
- Background
- Motivation
- Software
- Model Library : Instead of building models from the ground up, applied modellers can leverage `macpan2`'s library of predefined models as starting points.
- Calibration
- Illustrations
- Open-Source Tools

# Story

- Why Models?
- Background
- Motivation
- Software
- Model Library
- Calibration : `macpan2` employs formal mathematical optimization for the efficient calibration of model parameters to data.
- Illustrations
- Open-Source Tools

# Story

- Why Models?
- Background
- Motivation
- Software
- Model Library
- Calibration
- Illustrations : I'll dig into a few examples of work using `macpan2`.
- Open-Source Tools

# Story

- Why Models?
- Background
- Motivation
- Software
- Model Library
- Calibration
- Illustrations
- Open-Source Tools : The set of open-source tools available for epidemiologists to build forecasting models is improving and expanding.

# Why Models?

Why even bother at all with modelling?

# What do you get out of mechanistic modelling?

(in applied, non-academic, data-informed, infectious disease, public health work)

- Amplifies utility of data
- Forecasting (accessing ‘data’ of the future)
- Understanding mechanism (what caused the data we see?)
- Counter-factuals (what-if scenarios)

# Impact on the Public Health Agency of Canada (examples)

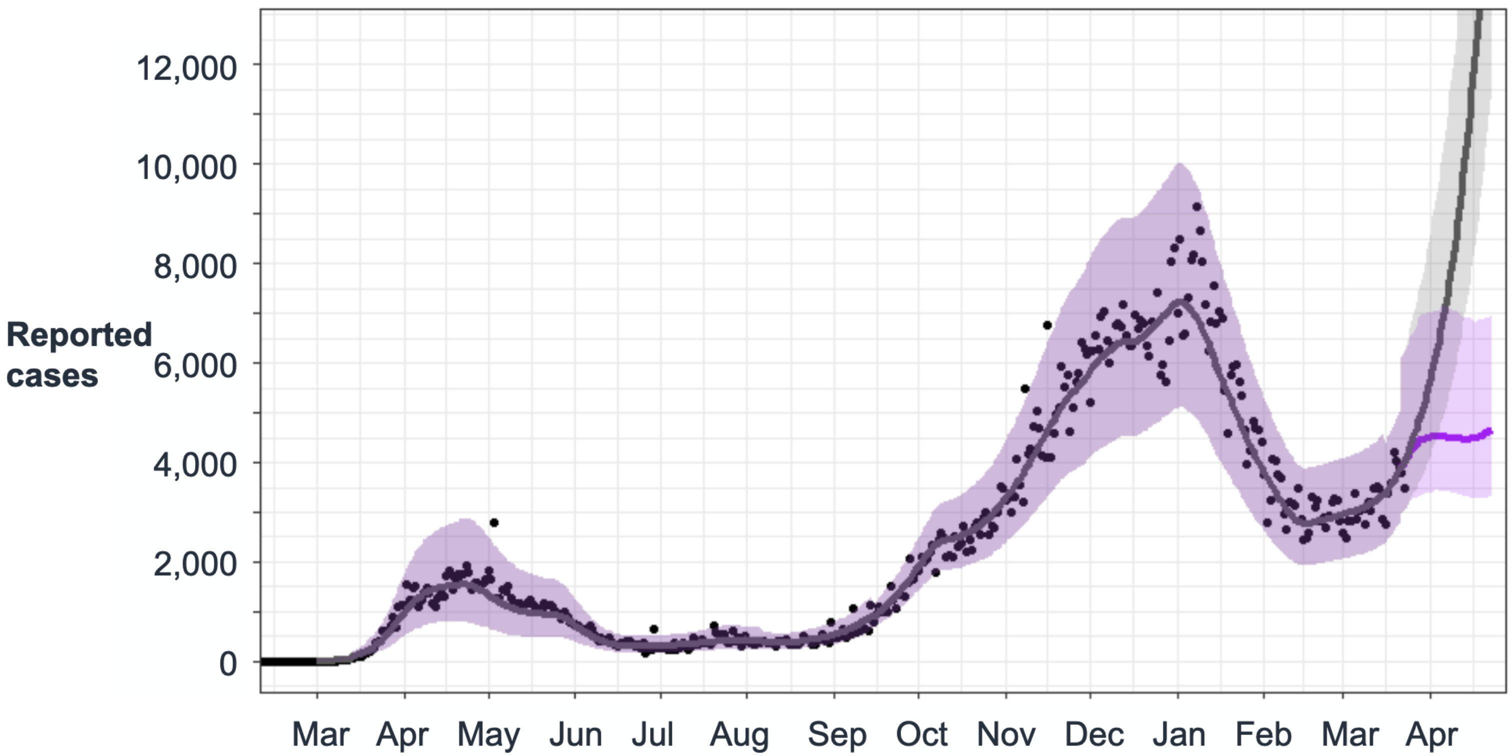
- Measles 2024: https://www.canada.ca/content/dam/phac-aspc/documents/services/emergency-preparedness-response/rapid-risk-assessments-public-health-professionals/rapid-risk-assessment-measles-public-health-implications-2024/rapid-risk-assessment-measles-public-health-implications-2024.pdf
- MPox 2022: https://onlinelibrary.wiley.com/doi/10.1002/jmv.29256
- COVID-19 2021: https://www.canada.ca/content/dam/phac-aspc/documents/services/diseases-maladies/coronavirus-disease-covid-19/epidemiological-economic-research-data/update-covid-19-canada-epidemiology-modelling-20210326-en.pdf

# Background

McMasterPandemic was developed for COVID-19 modelling and forecasting.

# PHAC report involving `McMasterPandemic`

## Longer-range forecast shows stronger public health measures will be required to counter more transmissible variants of concern



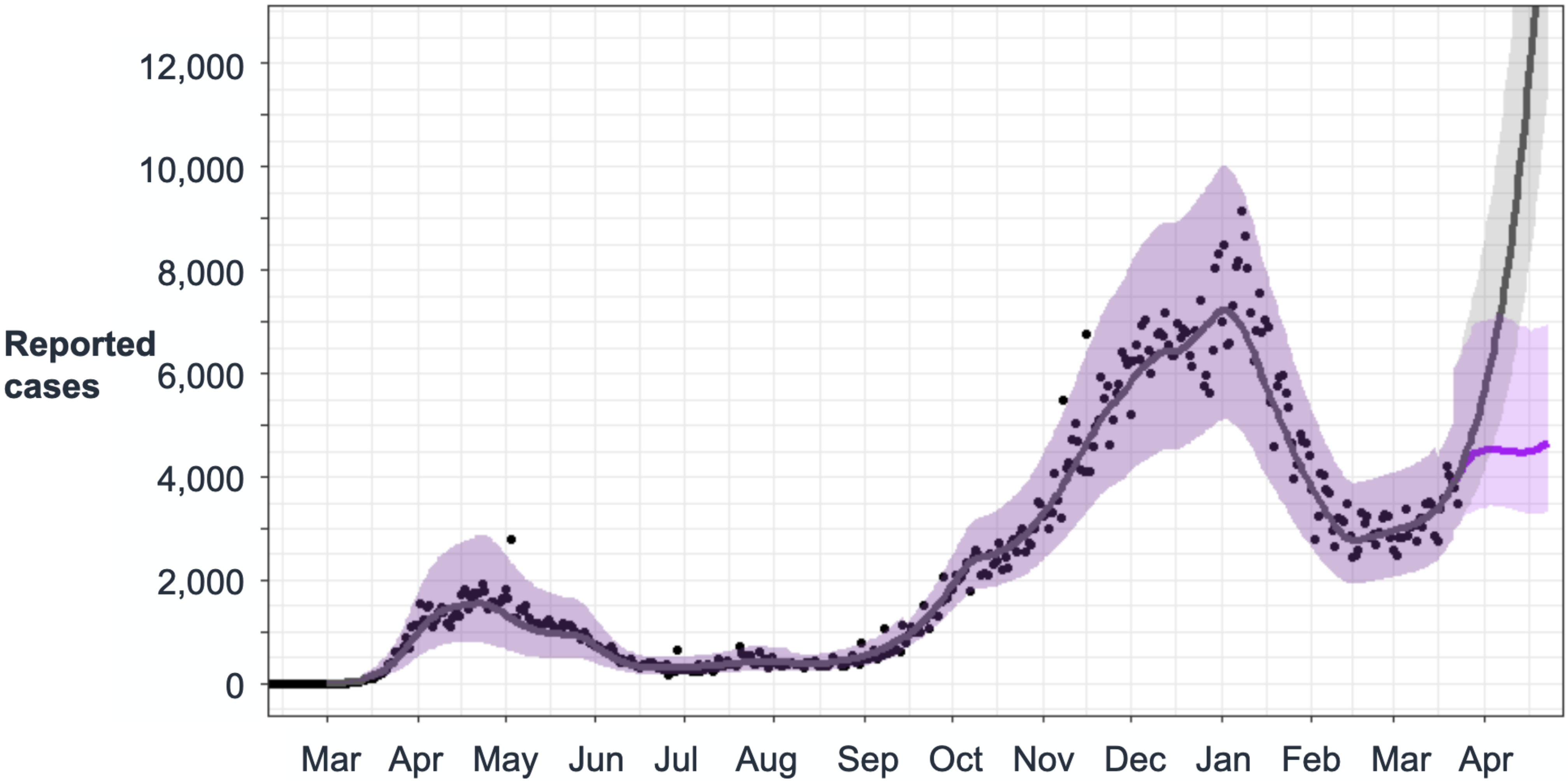

With **spread** of VOCs **and** we **maintain or increase** the current number of people we contact each day

If VOCs are controlled by **reducing** the current number of people we contact each day

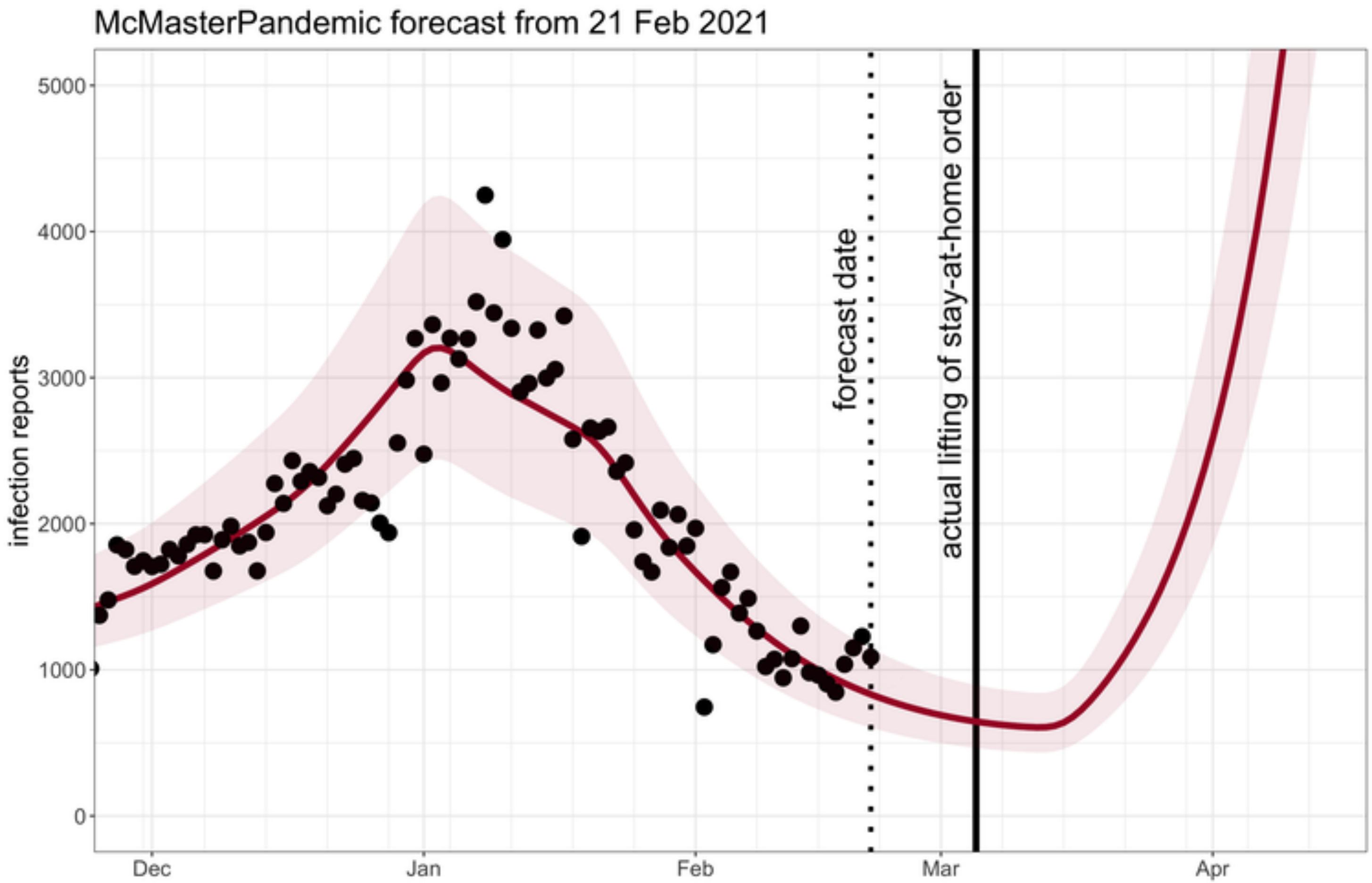
*Data as of March 24, 2021*
***Note:*** *Ensemble of output from PHAC-McMaster and Simon Fraser University models*

The PHAC McMaster forecast is based on current estimates transmission rates fitted to reported cases. It assumes VOCs are introduced in mid-Dec (~1 week prior to first detected case in Canada) at very low prevalence; VOCs (all VOCs known to date) are 50% more transmissible than wild-type; growth rate AND replacement rate are negatively correlated with the strength of public health measures. Proportion of VOC is obtained by a combination of calibrating to surveillance data as well as information on proportions of cases that are VOC. Recent changes in testing rates are not taken into account in this forecast. SFU methods are at https://www.sfu.ca/magpie/blog/variant-simple-proactive.html

# PHAC report involving `McMasterPandemic`

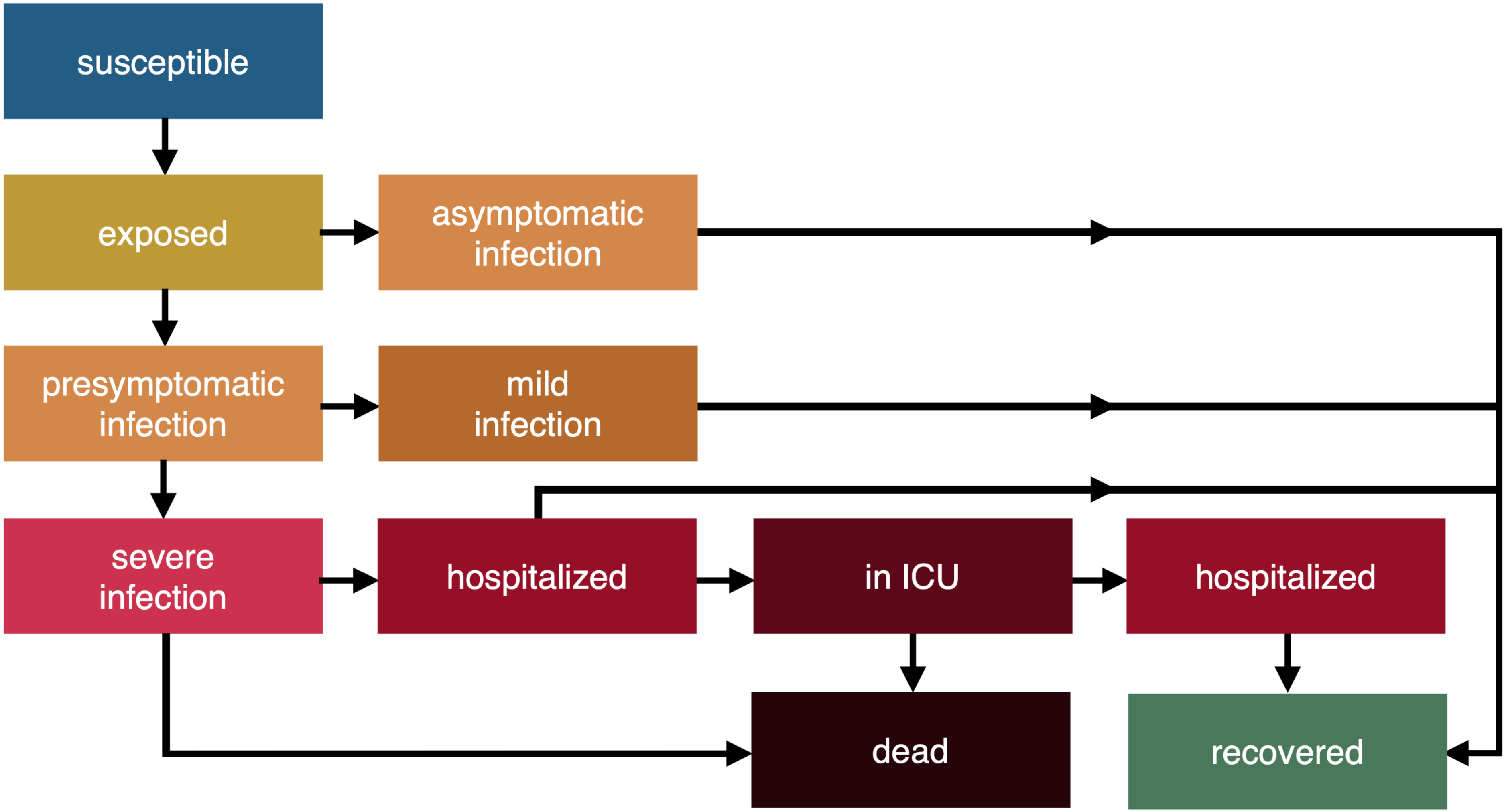


# Forecasting



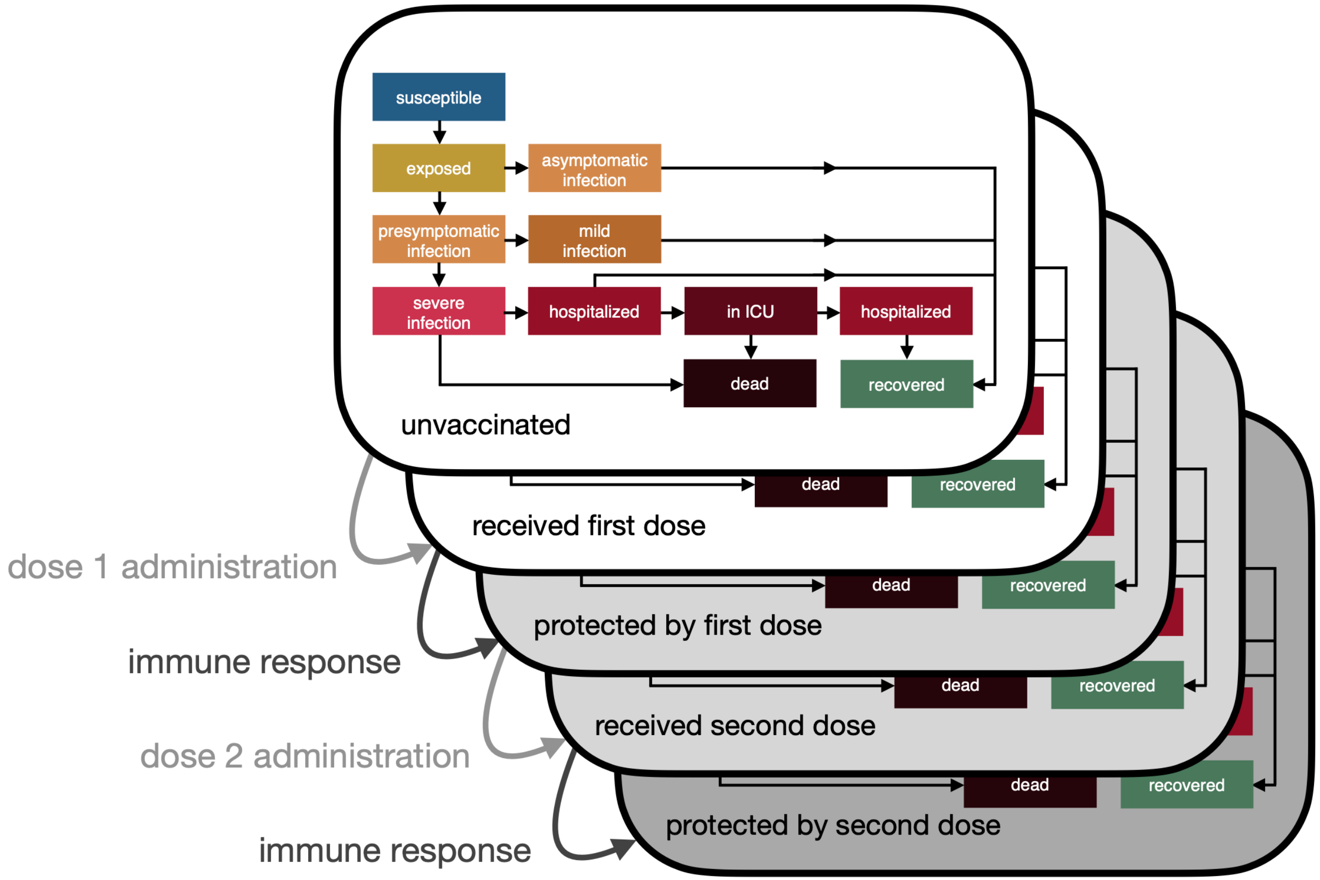

mac-theobio.github.io/forecasts/outputs/ON_accuracy.html

# McMasterPandemic COVID-19 (Mechanistic) Model




mac-theobio.github.io/covid-19

# COVID-19 Vaccination (Mechanistic) Model

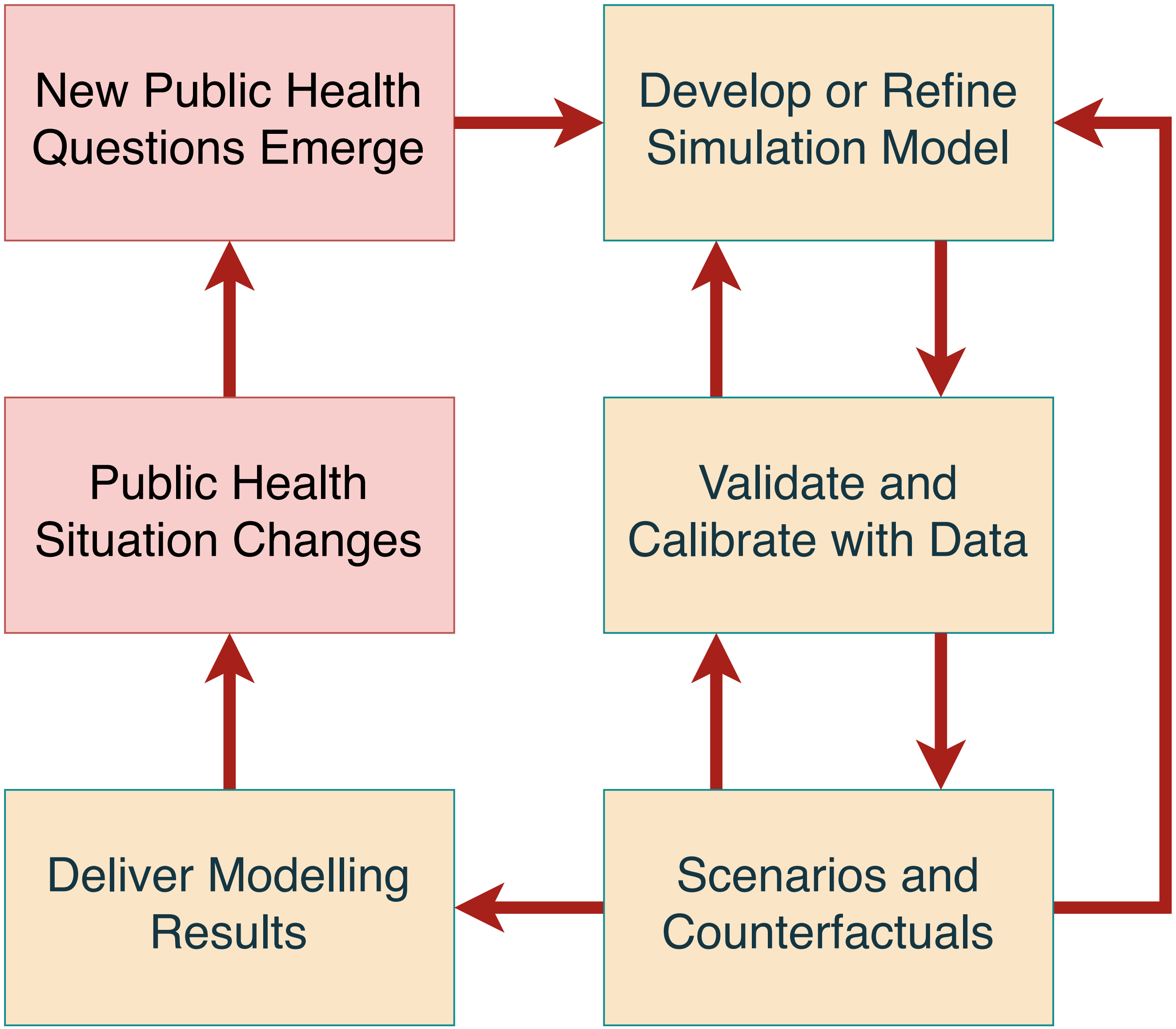



mac-theobio.github.io/covid-19

# Motivation

The urgency of producing regular forecasts during the pandemic impeded software development.

# Model Development Cycle



# Lessons – modeller needs and wants

Mechanistic modelling

Fast simulations and calibrations

Statistically principled

Functionality-rich

Modular model building

Easy to use

canmod.github.io/macpan-book/index.html#vision-and-direction

# Lessons – modeller needs and wants

Mechanistic modelling : Prediction is not enough. Must explain why.

Fast simulations and calibrations

Statistically principled

Functionality-rich

Modular model building

Easy to use

canmod.github.io/macpan-book/index.html#vision-and-direction

# Lessons – modeller needs and wants

Mechanistic modelling

Fast simulations and calibrations : Computation cannot hold back thinking.

Statistically principled

Functionality-rich

Modular model building

Easy to use

canmod.github.io/macpan-book/index.html#vision-and-direction

# Lessons – modeller needs and wants

Mechanistic modelling

Fast simulations and calibrations

Statistically principled : Uncertainty estimation (confidence intervals, priors)

Functionality-rich

Modular model building

Easy to use

canmod.github.io/macpan-book/index.html#vision-and-direction

# Lessons – modeller needs and wants

Mechanistic modelling

Fast simulations and calibrations

Statistically principled

Functionality-rich : "But I need to use fancy modelling idea X."

Modular model building

Easy to use

canmod.github.io/macpan-book/index.html#vision-and-direction

# Lessons – modeller needs and wants

Mechanistic modelling

Fast simulations and calibrations

Statistically principled

Functionality-rich

Modular model building : “Need to stratify my in-use MPox model by age.”

Easy to use

canmod.github.io/macpan-book/index.html#vision-and-direction

# Lessons – modeller needs and wants

Mechanistic modelling

Fast simulations and calibrations

Statistically principled

Functionality-rich

Modular model building

Easy to use : “Why would I switch from the tool I already know how to use?”

canmod.github.io/macpan-book/index.html#vision-and-direction

# Lessons – focus here on these

Mechanistic modelling

Fast simulations and calibrations

Statistically principled

Functionality-rich

Modular model building

Easy to use

canmod.github.io/macpan-book/index.html#vision-and-direction

# Software

To address these challenges, the `macpan2` project was launched to create a more versatile public health modelling tool, integrating lessons learned from the pandemic.

https://canmod.github.io/macpan2

# Timeline

| nickname | year | description |
|---|---|---|
| macpan1 | 2020 | original covid models in R engine |
| macpan1.5 | 2021 | fast C++ engine and flexibility for VOC modelling |
| macpan2 | 2023 | general architecture to implement the lessons learned |
| macpan2 | 2024 | simple interface to general architecture |

macpan1 = github.com/mac-theobio/McMasterPandemic/releases/tag/v0.0.20.1

macpan1.5 = github.com/mac-theobio/McMasterPandemic

macpan2 = github.com/canmod/macpan2

# Lessons

Mechanistic modelling

Fast simulations and calibrations

Statistically principled

Functionality-rich

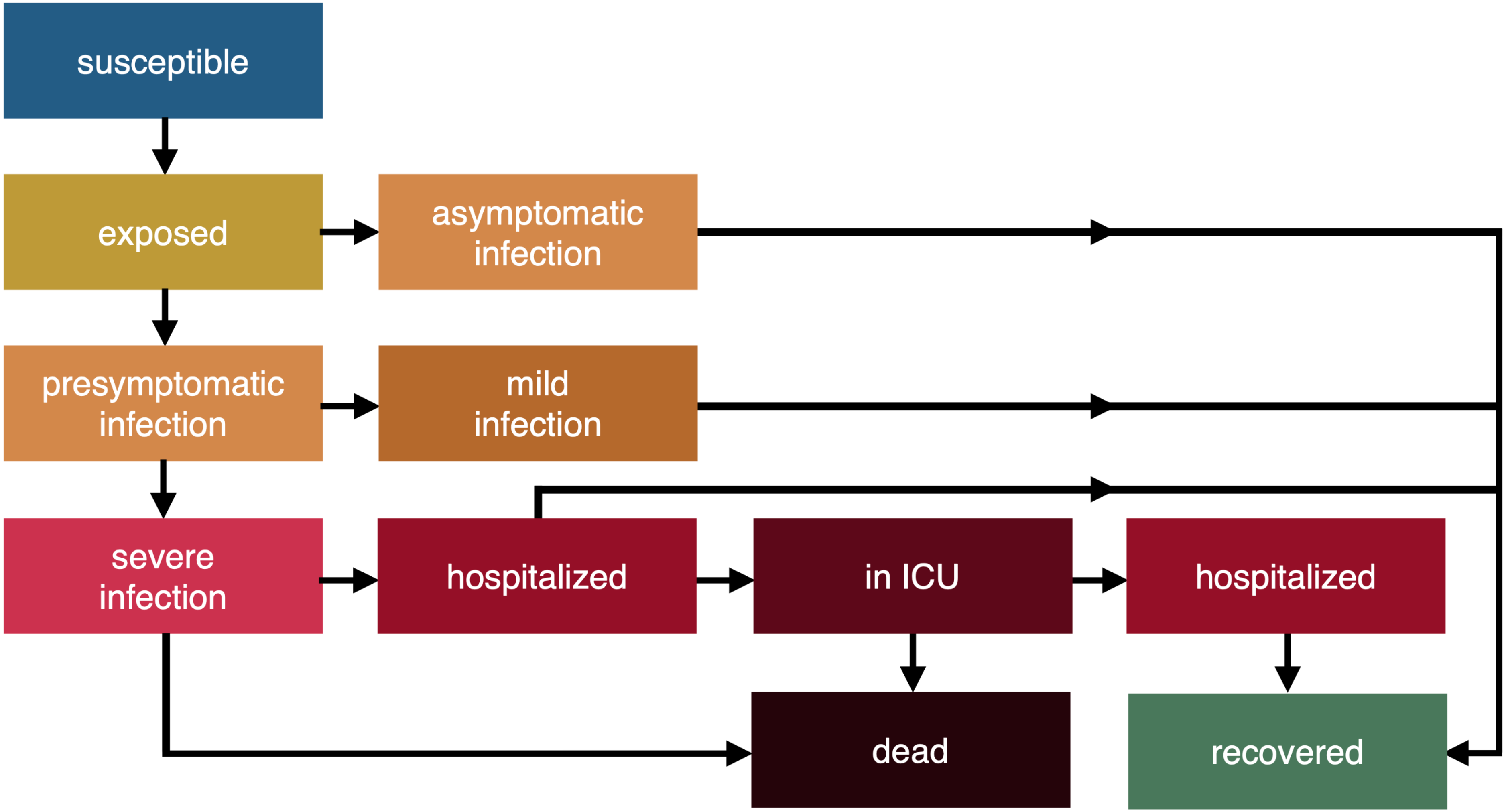
Modular model building

Easy to use

canmod.github.io/macpan-book/index.html#vision-and-direction

# Mechanistic Modelling

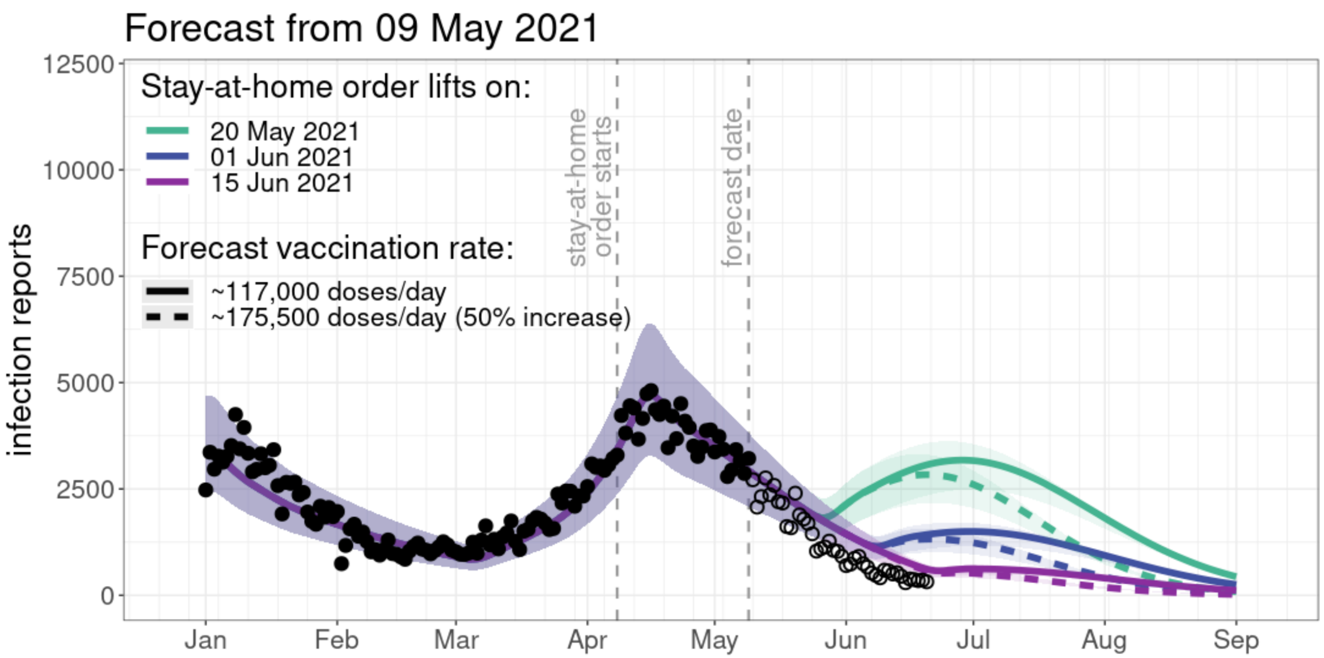
## Why Not Just Use AI for Everything?




mac-theobio.github.io/covid-19

# Mechanistic Modelling

- Useful for exploring “what if” scenarios.
- Causal, epidemiological principles build confidence in predictions beyond just fitting data.



# Mechanistic Modelling

- Limitations of AI
- Weather Forecasting as a Gold Standard
- Responsibility to Decision-Makers

# Mechanistic Modelling

- Limitations of AI : AI predicts well but doesn’t generate hypotheses or uncover causal relationships like mechanistic models.
- Weather Forecasting as a Gold Standard
- Responsibility to Decision-Makers

# Mechanistic Modelling

- Limitations of AI
- Weather Forecasting as a Gold Standard : Mechanistic models drive accuracy and uncertainty quantification, unlikely surpassed by AI.
- Responsibility to Decision-Makers

# Mechanistic Modelling

- Limitations of AI
- Weather Forecasting as a Gold Standard
- Responsibility to Decision-Makers : Mechanistic models provide clear, explainable rationale.
  - AI often can’t explain beyond data fitting
  - Maybe large language models are challenging this
  - Even ChatGPT should use mechanisms when explaining

# Lessons

Mechanistic modelling

Fast simulations and calibrations

Statistically principled

Functionality-rich

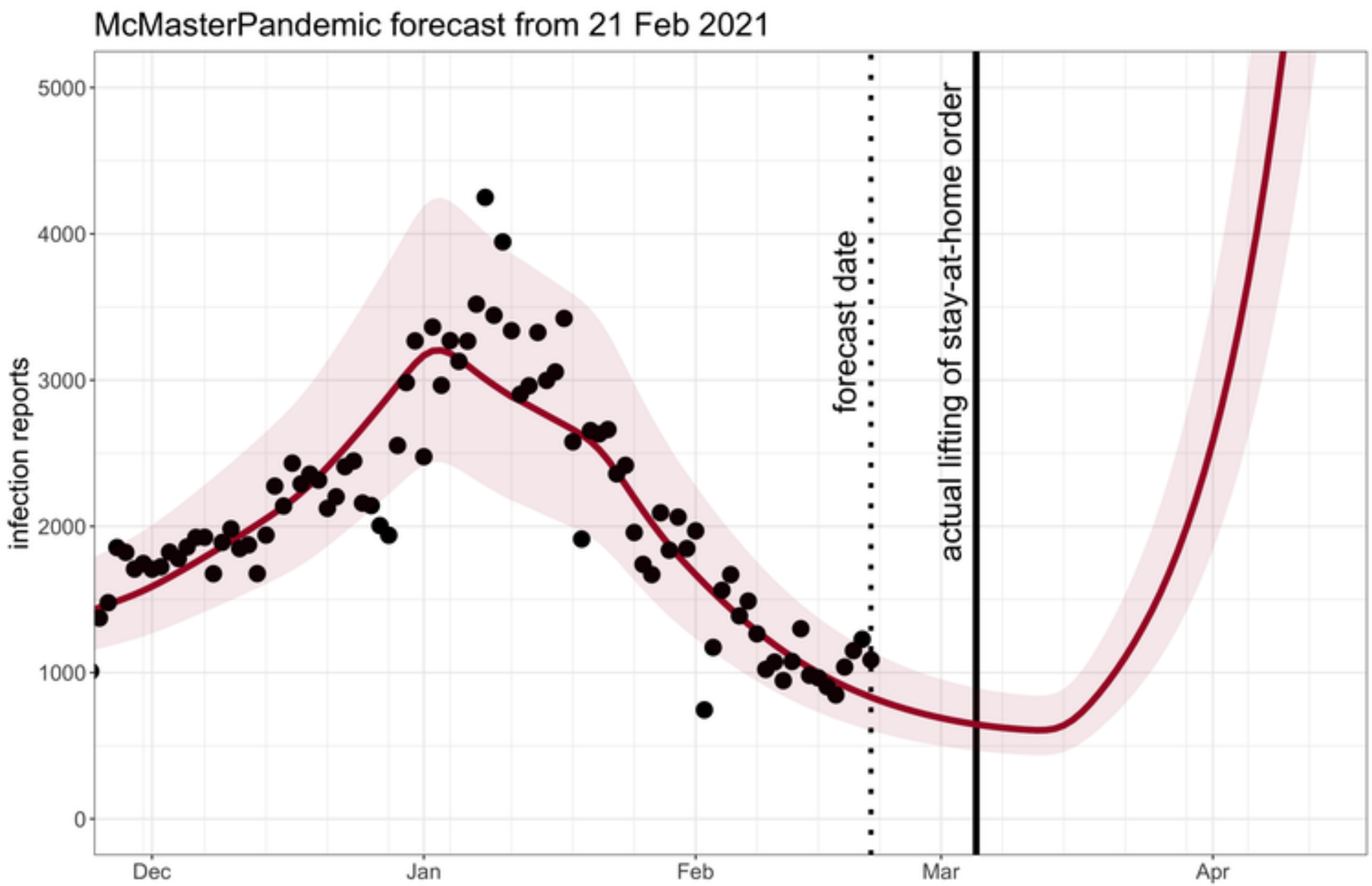
Modular model building

Easy to use

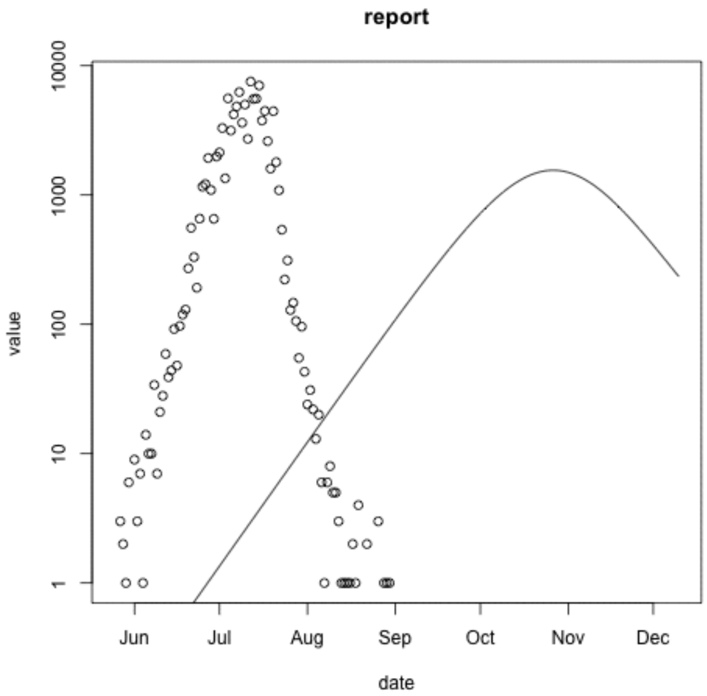
canmod.github.io/macpan-book/index.html#vision-and-direction

# Computational Efficiency

Why is it important? What goes into a picture like this?



# Calibration – Trajectory Matching




- Toy simulated data example
- “Observed” case reports – dots
- Model-predicted case reports – line
- Optimize transmission rate using maximum likelihood

# Lessons

Mechanistic modelling

Fast simulations and calibrations

Statistically principled

Functionality-rich

Modular model building

Easy to use

canmod.github.io/macpan-book/index.html#vision-and-direction

# Interface

```
spec = mp_tmb_model_spec(
    before = S ~ N - I
  , during = mp_per_capita_flow("S", "I", "beta * I / N", "infection")
  , default = list(N = 100, I = 1, beta = 0.2)
)
```

- Will slowly walk through what code means
- Users interact by supplying R code
- Will talk later about similar tools in Python
- Goal: easy to read for modellers

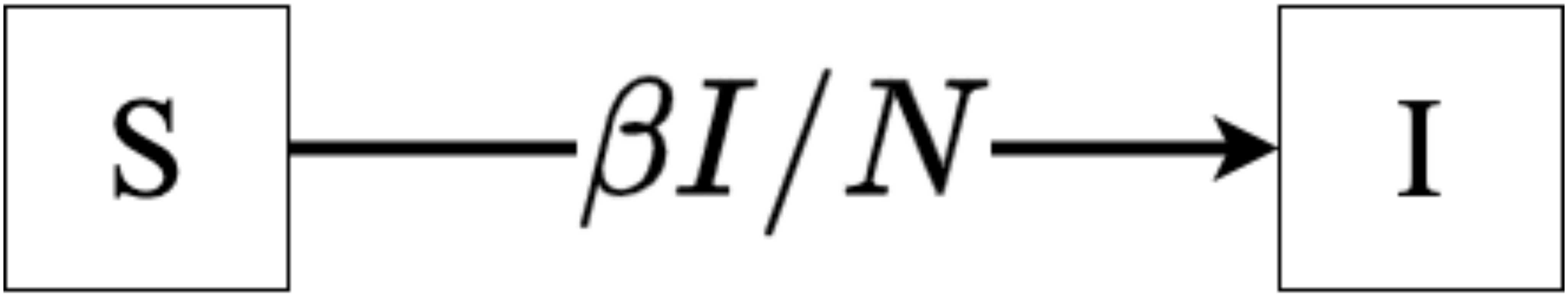
https://canmod.github.io/macpan2#hello-world

# Specifying Per-Capita Flows

```
spec = mp_tmb_model_spec(
  ...
  mp_per_capita_flow("S", "I", "beta * I / N", "infection")
  ...
)
```

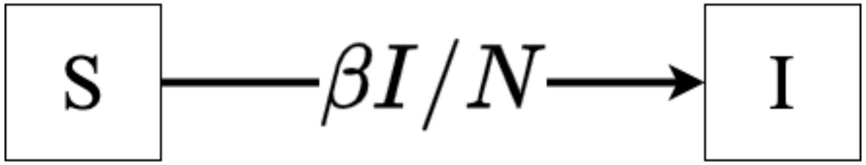
https://canmod.github.io/macpan2#hello-world

# Specifying Per-Capita Flows




- S : number of susceptible individuals
- I : number of infectious individuals
- $\beta$ : transmission rate
- N : total population size

# Specifying Per-Capita Flows



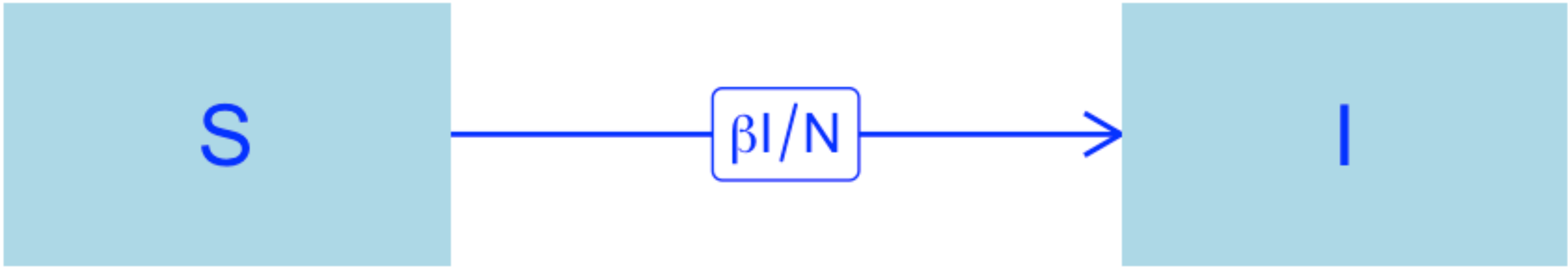

```
spec = mp_tmb_model_spec(
  ...
  mp_per_capita_flow("S", "I", "beta * I / N", "infection")
  ...
)
```

https://canmod.github.io/macpan2#hello-world

# Drawing Flow Diagrams

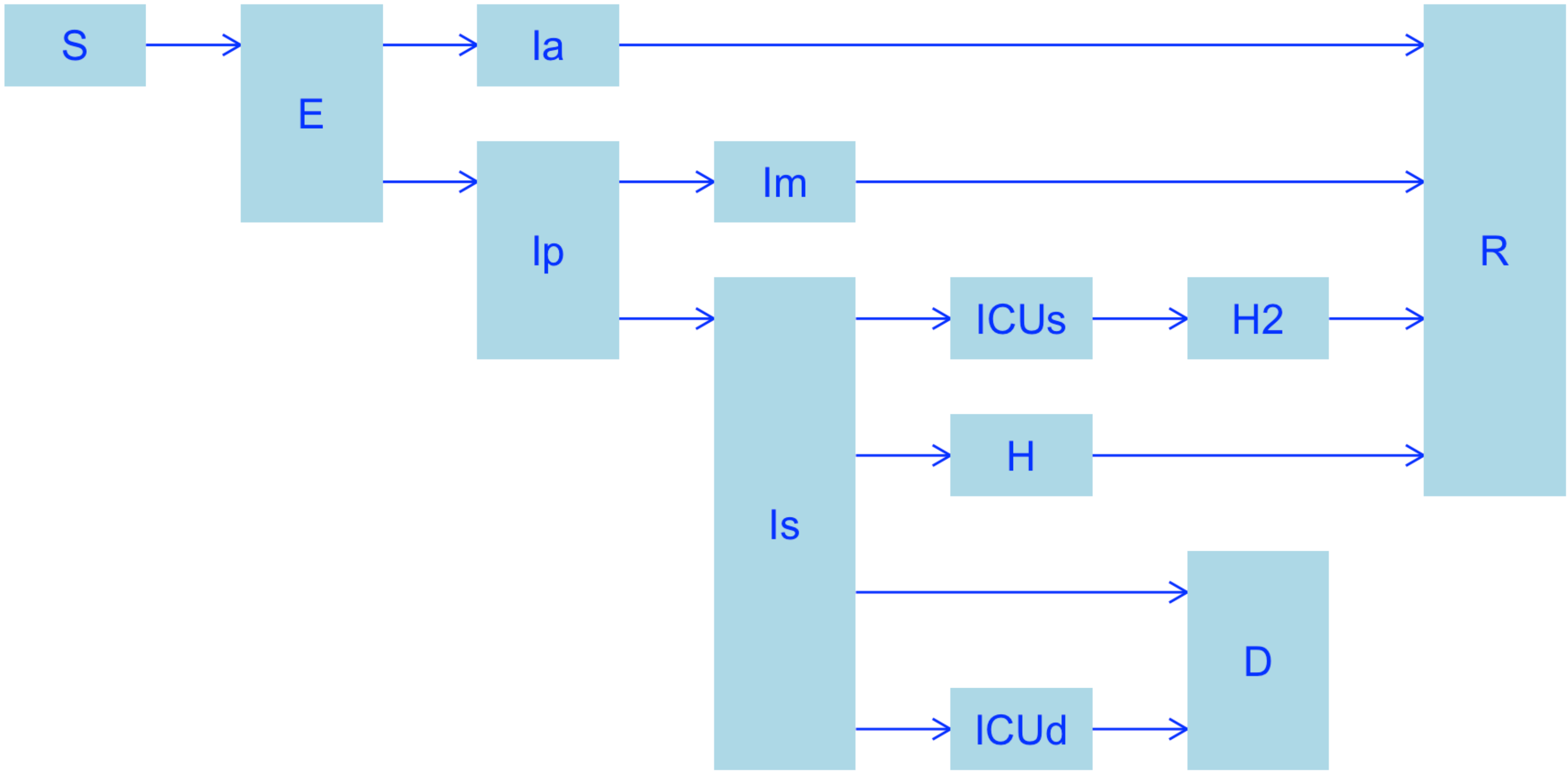
Specifications allow us to draw flow diagrams.

```
(spec                      ## model specification
  |> mp_layout_paths()     ## find where to place boxes
  |> plot_flow_diagram() ## draw diagram
)
```



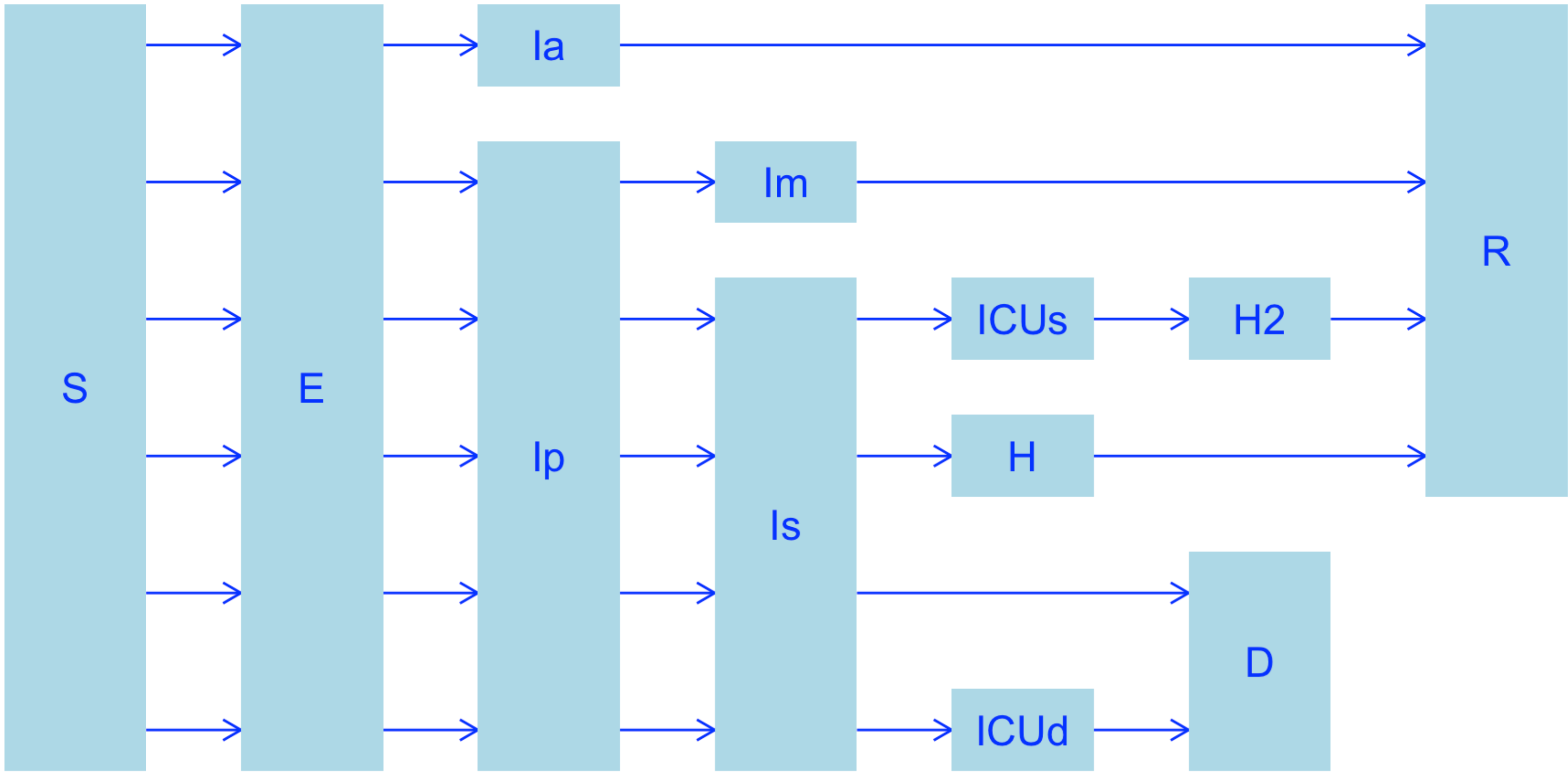
# Drawing Flow Diagrams

- Automatic layout often looks nice for more realistic models.
- Based on finding all paths through the system.



# Drawing Flow Diagrams

- Automatic layout often looks nice for more realistic models.
- Based on finding all paths through the system.



# Flows to Differential Equations

```
spec = mp_tmb_model_spec(
  ...
  mp_per_capita_flow("S", "I", "beta * I / N", "infection")
  ...
)
```

$$\text{infection} = S \times \beta I/N$$

# Flows to Differential Equations

```
spec = mp_tmb_model_spec(
  ...
  mp_per_capita_flow("S", "I", "beta * I / N", "infection")
  ...
)
```

$$
\begin{aligned}
\text{infection} &= S \times \beta I/N \\
\frac{dS}{dt} &= -\text{infection} \\
\frac{dI}{dt} &= +\text{infection}
\end{aligned}
$$

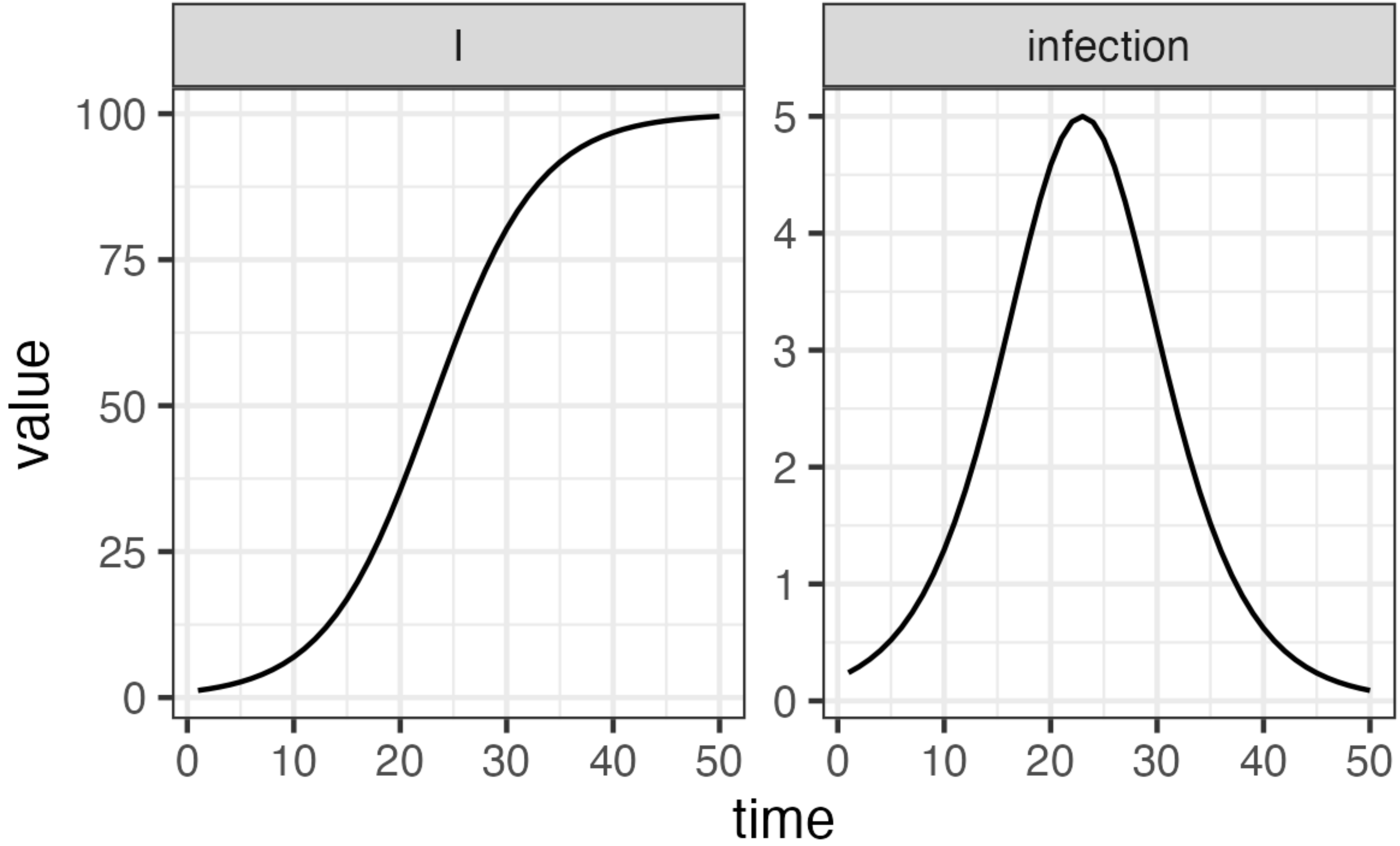
Sorry for the math. This is how our users think.

# Simulation Syntax

- You do not need to read this code
- Philosophy: Do not reinvent the wheel
- Simulation output standardized for all models
- Easily plugs into existing popular tools (e.g., `ggplot2`)

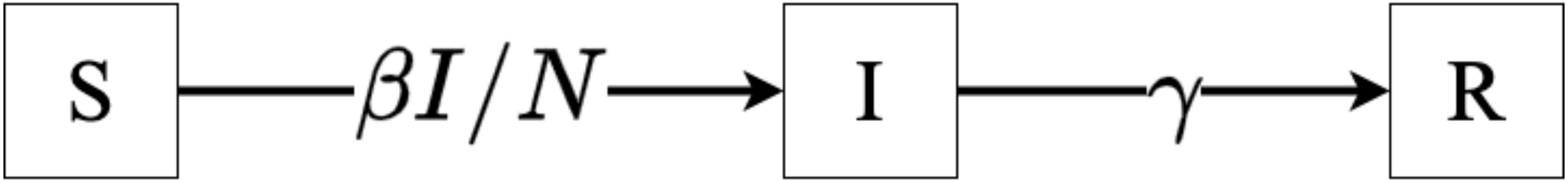
```
(spec
 ## macpan2 part ----------------
  |> mp_simulator(time_steps = 50, outputs = c("I", "infection"))
  |> mp_trajectory()

 ## not macpan2 ----------------
  |> ggplot()
 + geom_line(aes(time, value))
 + facet_wrap(~ matrix, scales = "free")
 + theme_bw()
)
```

# Simulation Syntax



# Specifying Models

```
spec = mp_tmb_model_spec(
  ...
  , mp_per_capita_flow("S", "I", "beta * I / N", "infection")
  , mp_per_capita_flow("I", "R", "gamma"       , "recovery" )
  ...
)
```



# Specifying Models

```
spec = mp_tmb_model_spec(
  ...
  , mp_per_capita_flow("S", "I", "beta * I / N", "infection")
  , mp_per_capita_flow("I", "R", "gamma"       , "recovery" )
  ...
)
```

$$\frac{dS}{dt} = -\text{infection}$$
$$\frac{dI}{dt} = -\text{recovery} + \text{infection}$$
$$\frac{dR}{dt} = +\text{recovery}$$

# Rendering Model Specifications

- Showed how specification can be interpreted as an ODE.
- Specifications with per-capita flows allow us to easily switch simulation strategies.

```
mp_euler(spec)              ## difference equation (default)
mp_rk4(spec)                ## ODE solver
mp_euler_multinomial(spec) ## stochasticity
mp_hazard(spec)             ## McMaster group hack during COVID emergency
```

https://canmod.github.io/macpan2/articles/state_updaters

# Using the Model Library

macpan2 1.8.1 Reference Articles Changelog

## Example Models

Source: `vignettes/example_models.Rmd`

status mature draft

### Finding Example Models

The `macpan2` comes with a set of example model definitions, which can be listed with the `show_models` function.

```
show_models()
```

| dir | title | index_entry |
|---|---|---|
| awareness | awareness models | behaviour modifications in response to death |
| hiv | HIV | A simple HIV model |
| lotka_volterra_competition | Lotka-Volterra | simple two-species competition model |
| lotka_volterra_preda- | Lotka- | simple predator-prey model |

https://canmod.github.io/macpan2/articles/example_models

# Illustrations

We illustrate `macpan2`'s application with some examples.

# Examples of Usage

- COVID-19 forecasts for agencies
  - Public Health Agency of Canada
  - Ontario Science Table
  - World Health Organization
    - South Africa – provincial and national level
    - Afro-Regional – 49 countries
  - Public Health Ontario

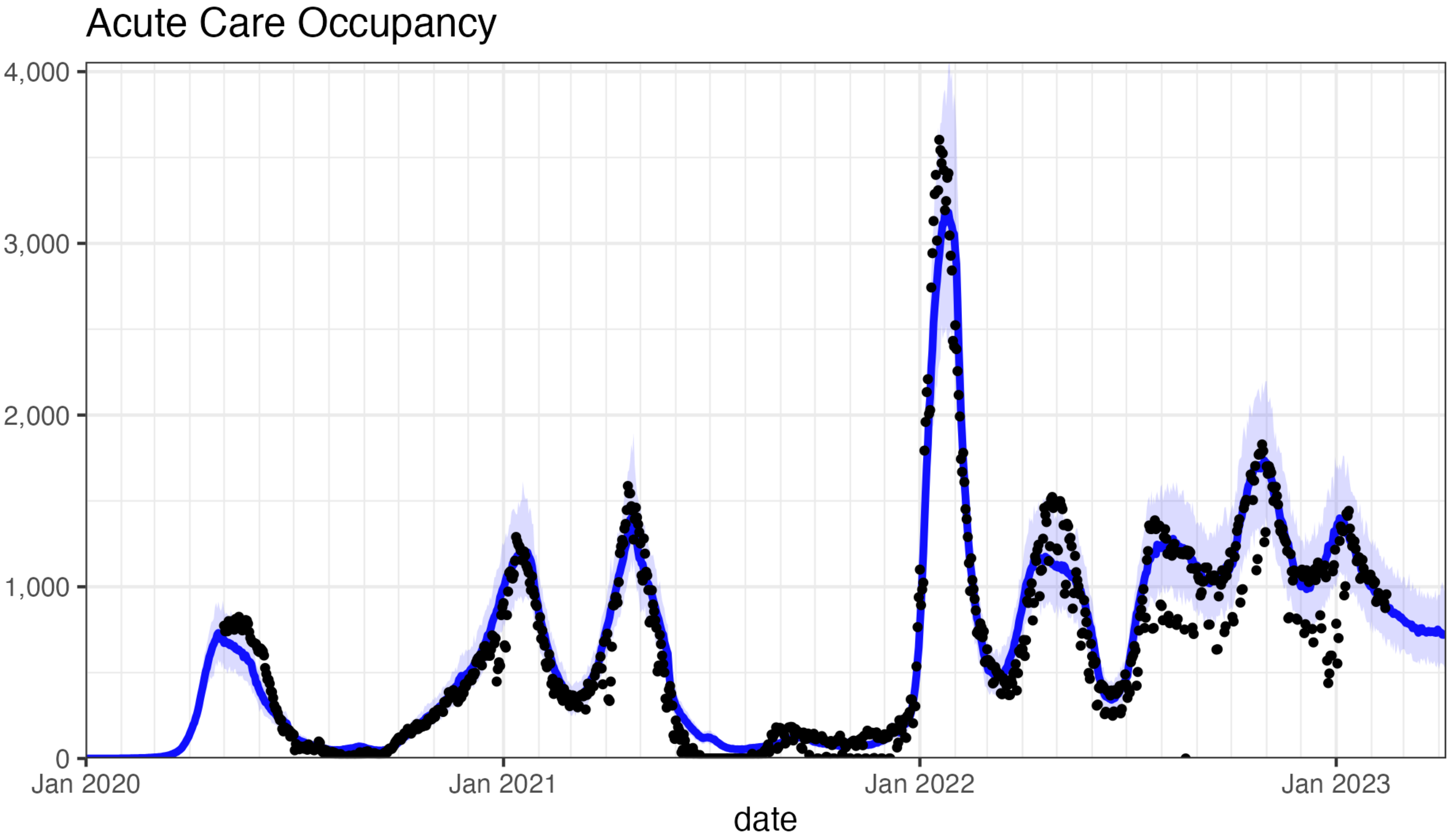
# Examples of Usage

- **Large populations**
  - Base COVID-19 model
  - More COVID-19 models (vaccination, VOC, waste-water, endemicity)
  - Mpox (Mildwid et al 2023)[1]
  - Scarlet Fever in UK
- **Small populations**
  - Base COVID-19 model in Yukon & NFLD
  - Metapopulation model of NFLD (1000s of compartments)
  - Measles (1000-person communities)

1: https://doi.org/10.1002/jmv.29256

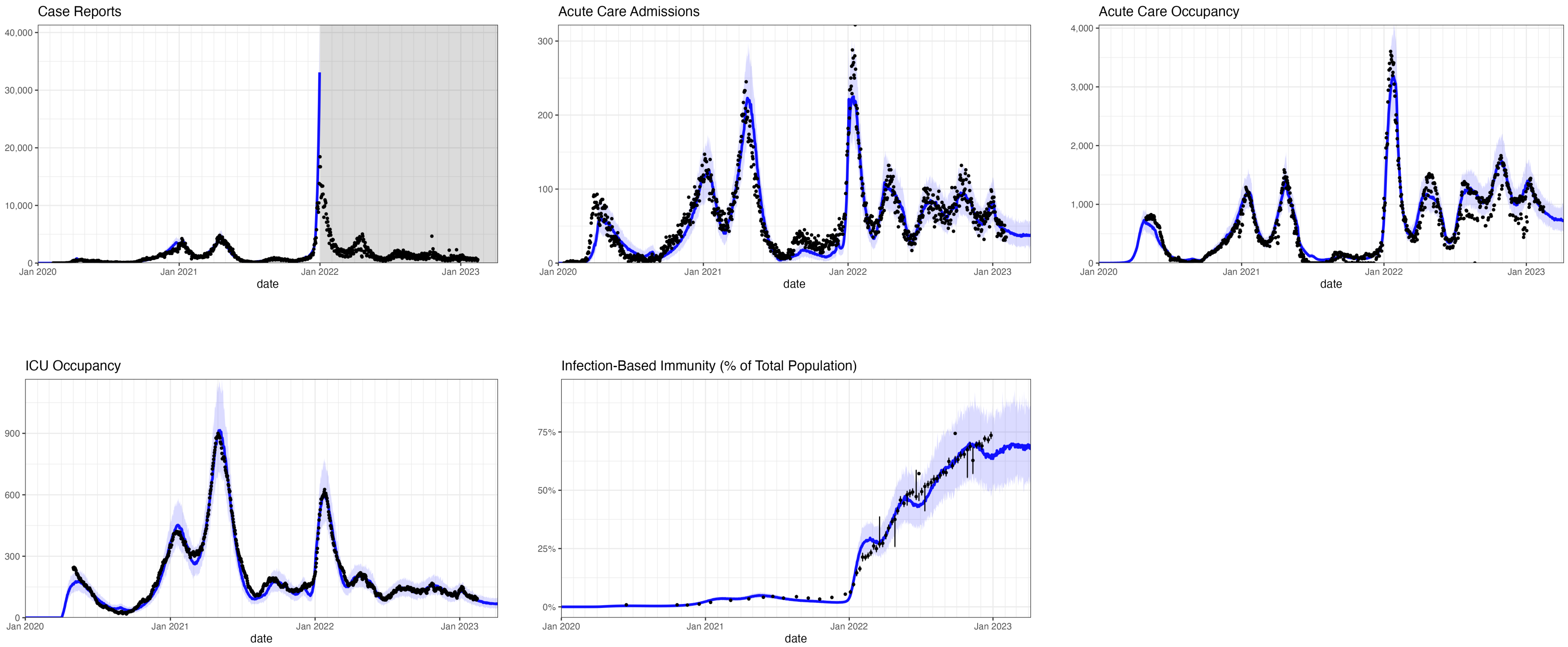
# COVID-19 in Ontario

- e.g., Public Health Ontario Forecasts

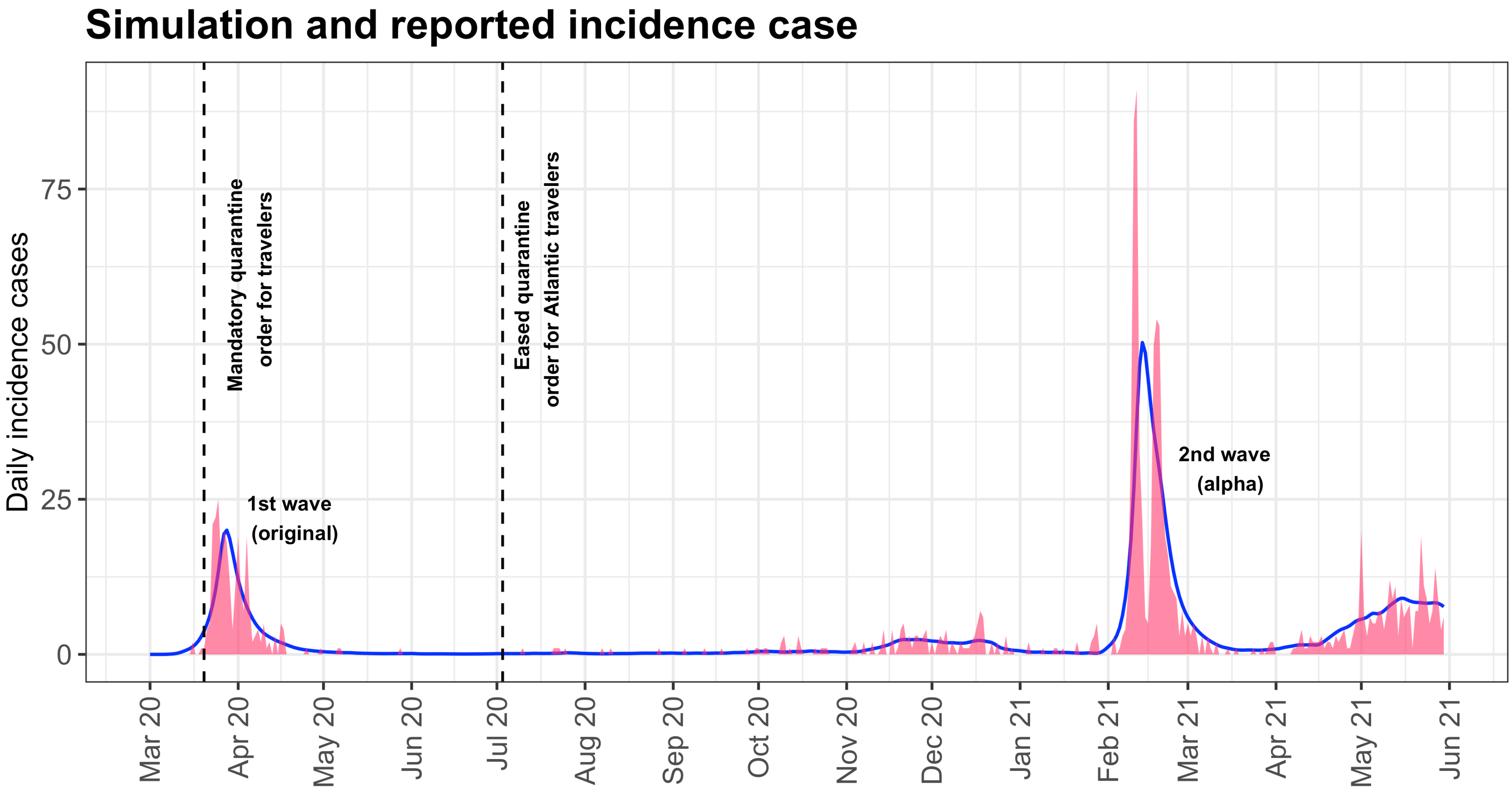


# COVID-19 in Ontario

- e.g., Public Health Ontario Forecasts
- ~100 parameters / five data-streams

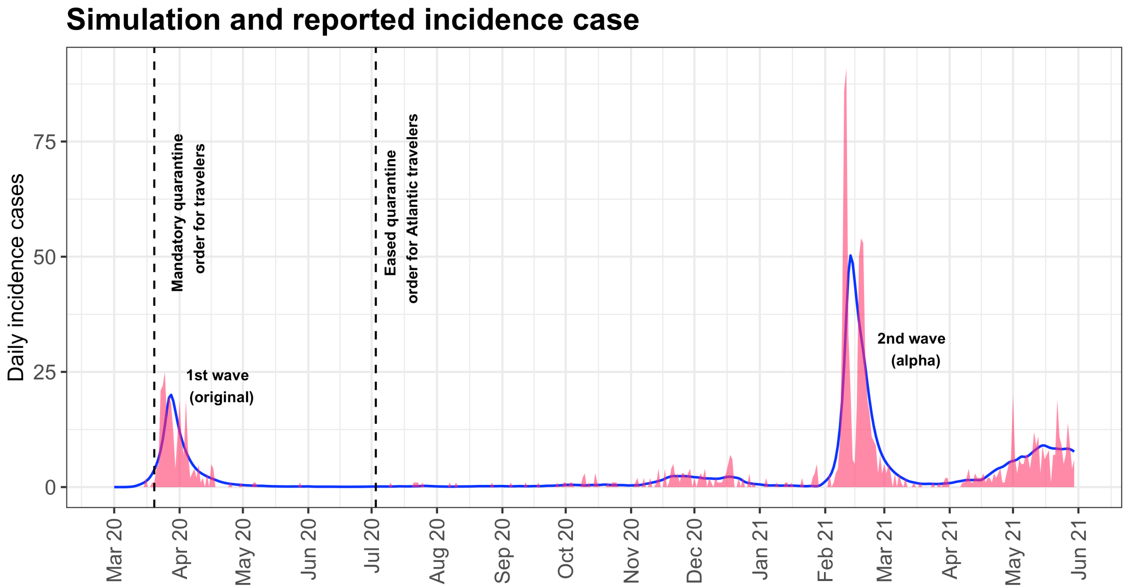


# COVID-19 in Newfoundland




Zahra Cook et al., ongoing project

# COVID-19 in Newfoundland




- Randomly spaced outbreaks caused by importation, not community spread
- Different from large provinces, driven by community spread
- Fitted blue line based on model dominated by community spread, hacked to look like importation
- Need to do better to finish this project (mechanism matters!)

Zahra Cook et al., ongoing project

# Small-N Awareness Models

- **Awareness models**: Population-level awareness of death leads to less risky behaviour (Weitz et al 2020)[1].
- **Random importation models**: Small populations can transition between zero cases and outbreaks caused by imported cases and stochasticity (Hurford et al 2023)[2].
- **Awareness-importation models**: Implemented in `macpan2`[3].

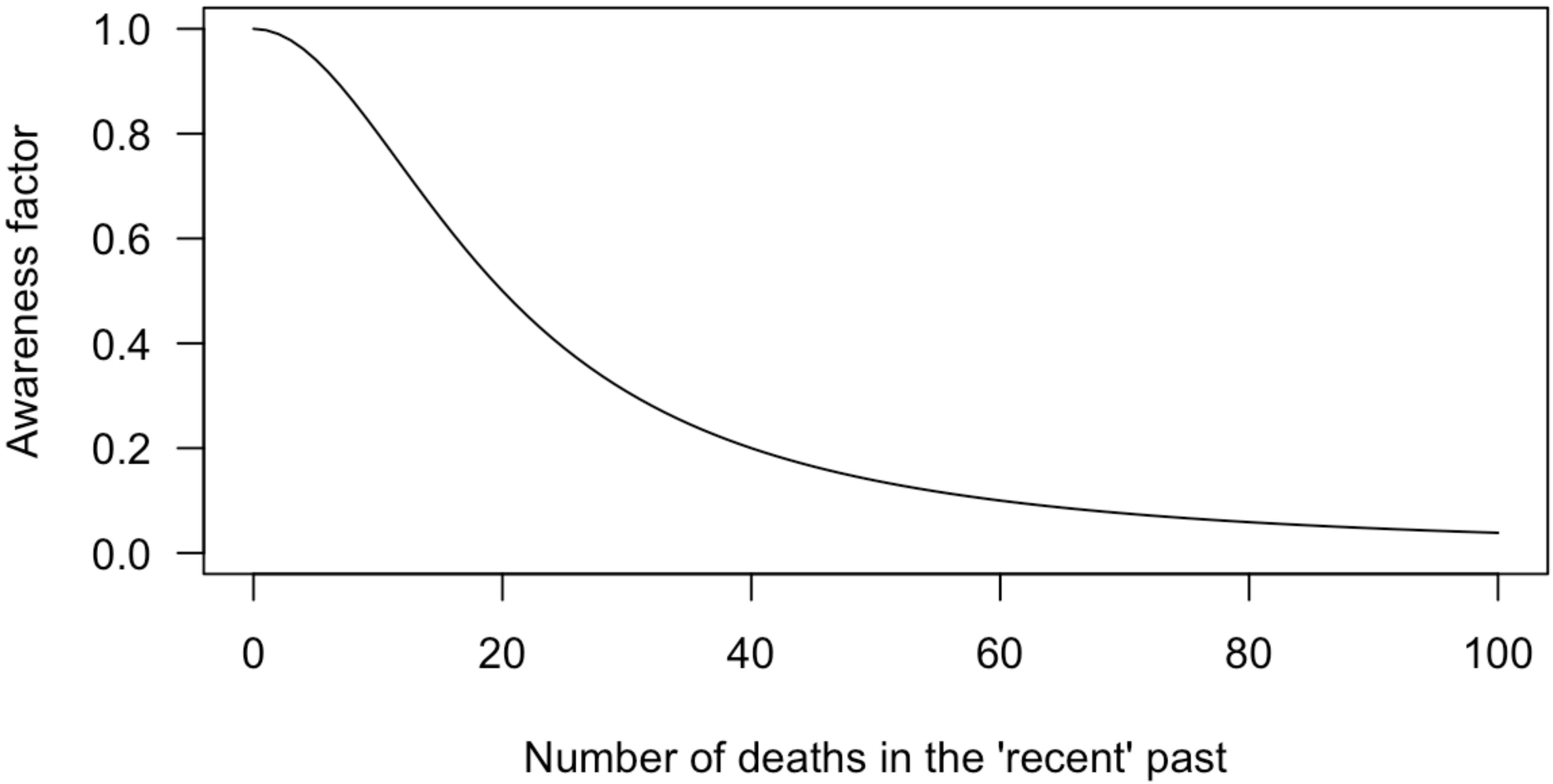
1: https://doi.org/10.1073/pnas.2009911117

2: https://doi.org/10.1016/j.jtbi.2022.111378

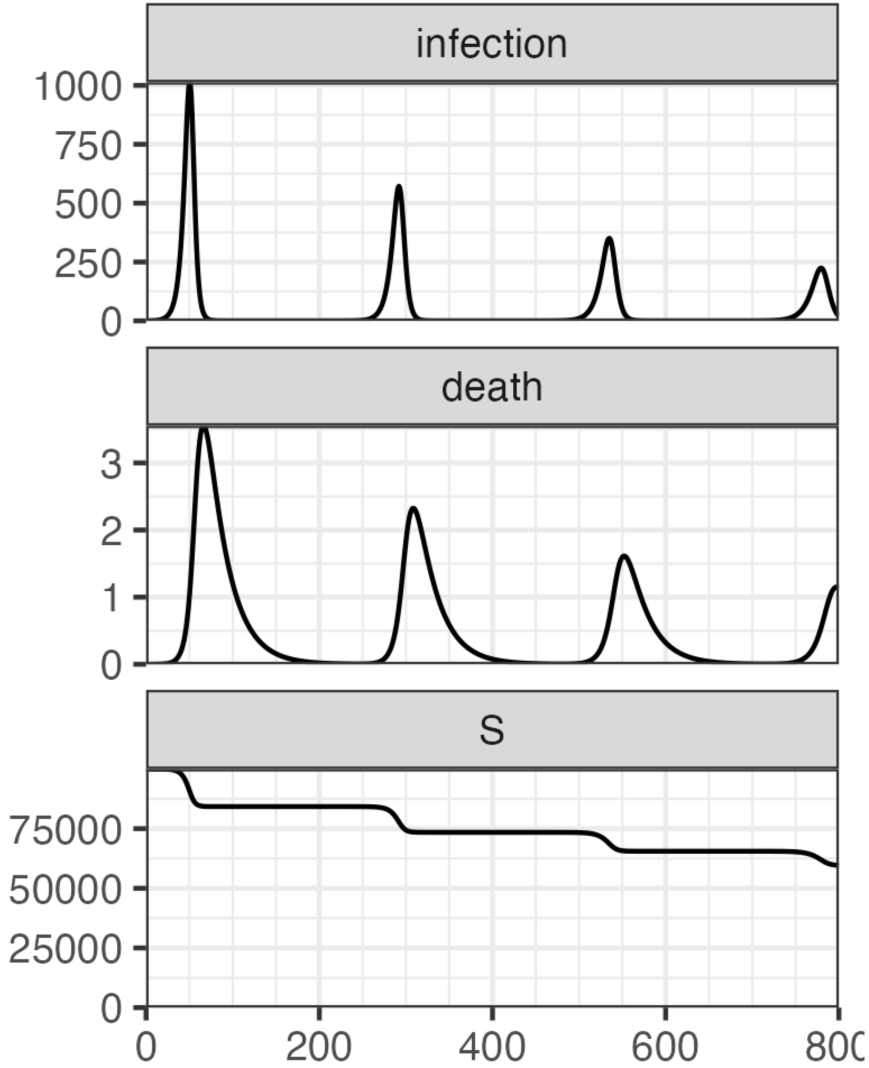
3: https://github.com/canmod/macpan2/tree/main/inst/starter_models/awareness

# Small-N Awareness Models

$$\text{force of infection} = \text{awareness factor} \times \frac{\beta I}{N}$$

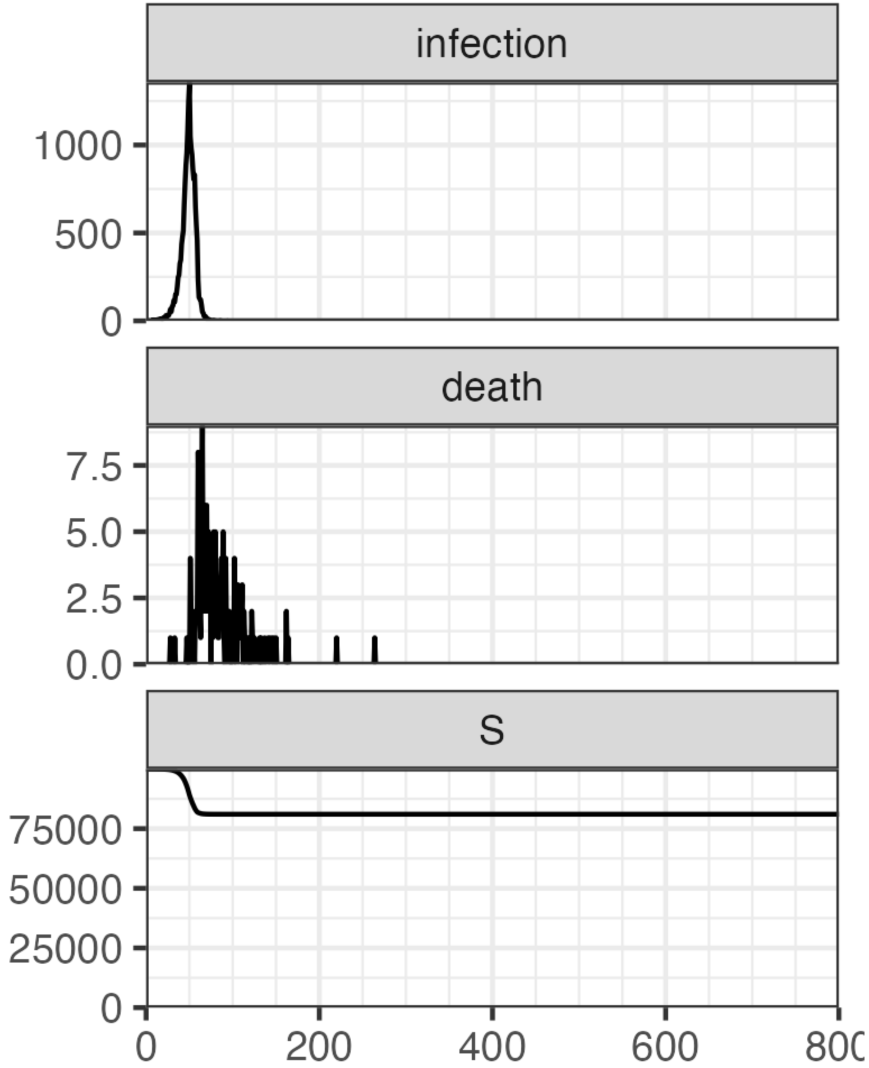


# Small-N Awareness Models



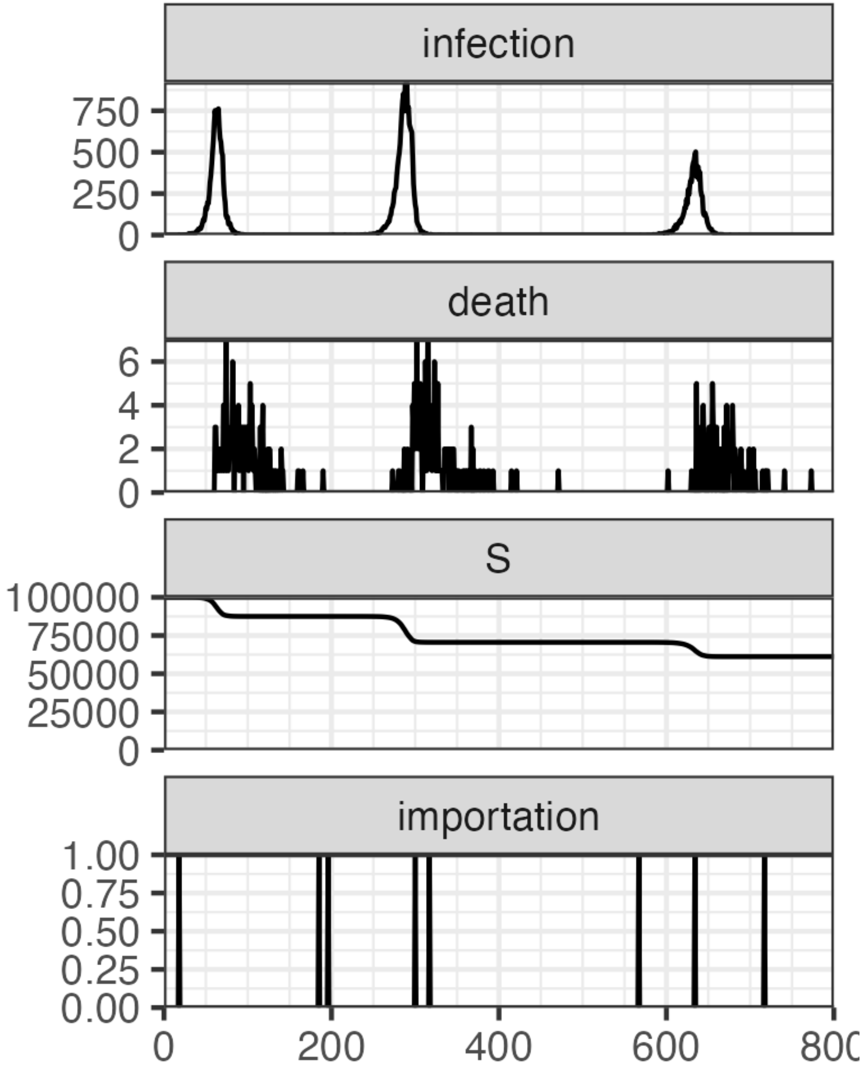

- No stochasticity, no importation
- Regularly-spaced cycles that slowly decrease in amplitude.
- Suceptible depletion is limited. Why?
- Turn-over behavioural (not due to herd immunity).

# Small-N Awareness Models




- With stochasticity we can have stochastic eradiction.
- Why do we not have a second wave?
- Because infectious population goes to exactly zero before behaviour gets risky again.

# Small-N Awareness Models




- Adding importation can yield randomly located outbreaks.
- Importation keeps the infectious population from staying at zero.
- Looks much more like Newfoundland.
- I'll be in touch Zahra

# Open-Source Tools

The set of open-source tools available for epidemiologists to build forecasting models is improving and expanding.

# List of Tools

- General purpose flexible model fitting/simulation tool-kits
  - https://mc-stan.org/
  - https://mcmc-jags.sourceforge.io/
  - https://kingaa.github.io/pomp/
  - https://github.com/kaskr/adcomp (aka TMB)
  - https://github.com/kaskr/RTMB (new!)
- Epidemiology focused tool-kits
  - https://www.epimodel.org/
  - https://epiverse-trace.github.io/
  - https://github.com/benmaier/epipack (Python)
  - https://github.com/ukhsa-collaboration/pygom (Python)
  - https://canmod.github.io/macpan2/

# Resources on `macpan2`

- canmod.github.io/macpan2
- canmod.github.io/macpan2#installation
- canmod.github.io/macpan2/articles/quickstart
- canmod.github.io/macpan2/articles/example_models
- canmod.github.io/macpan2/articles/calibration
- https://canmod.github.io/macpan-workshop/syllabus

# Conclusions

- Modelling adds value to data
- Translating pandemic modeling experience into open-source software for future preparedness.
- The `macpan2` project is part of this effort.
- Sustaining incremental improvement is difficult but critical for epi-focused tools.
- Focus on model specification, simulation, and calibration.
- Training modellers on `macpan2` (Interested?)

# Acknowledgements

- Canadian Institute for Health Information for the invitation
- CANMOD 😢 and PHAC for funding
- Original developers were real modellers (Ben Bolker, Jonathan Dushoff, David Earn, Mike Li, Irena Papst)
- C++ programming (Weiguang Guan, Sharcnet), model library (Jen Freeman), and product models (Darren Flynn-Primrose)
- Amy Hurford, Lisa Kanary, Caroline Colijn, Zahra Cook, Claude Nadeau, Philippe Berthiaume, Evan Mitchell, Brian Gaas, Kevin Zhao, Maya Earn
- Irena Papst and Mike Li