Epidemic Forecasting Software: Lessons from the COVID-19 Pandemic

Canadian Institute for Health Information (CIHI) – 2024-10-10

Steve Walker (McMaster University)

Acknowledgements

- 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 Mohammadi, Claude Nadeau, Philippe Berthiaume, Evan Mitchell, Brian Gaas, Kevin Zhao, Maya Earn
- Irena Papst and Mike Li

Story

- Background
- Motivation
- Software
- Model Library
- Calibration
- Illustrations
- Open-Source Tools

Story

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

Story

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

Story

- 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

- 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

- 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

- Background
- Motivation
- Software
- Model Library
- Calibration
- Illustrations : I'll dig into a few examples of `macpan2` (COVID-19, Mpox, and Measles).
- Open-Source Tools

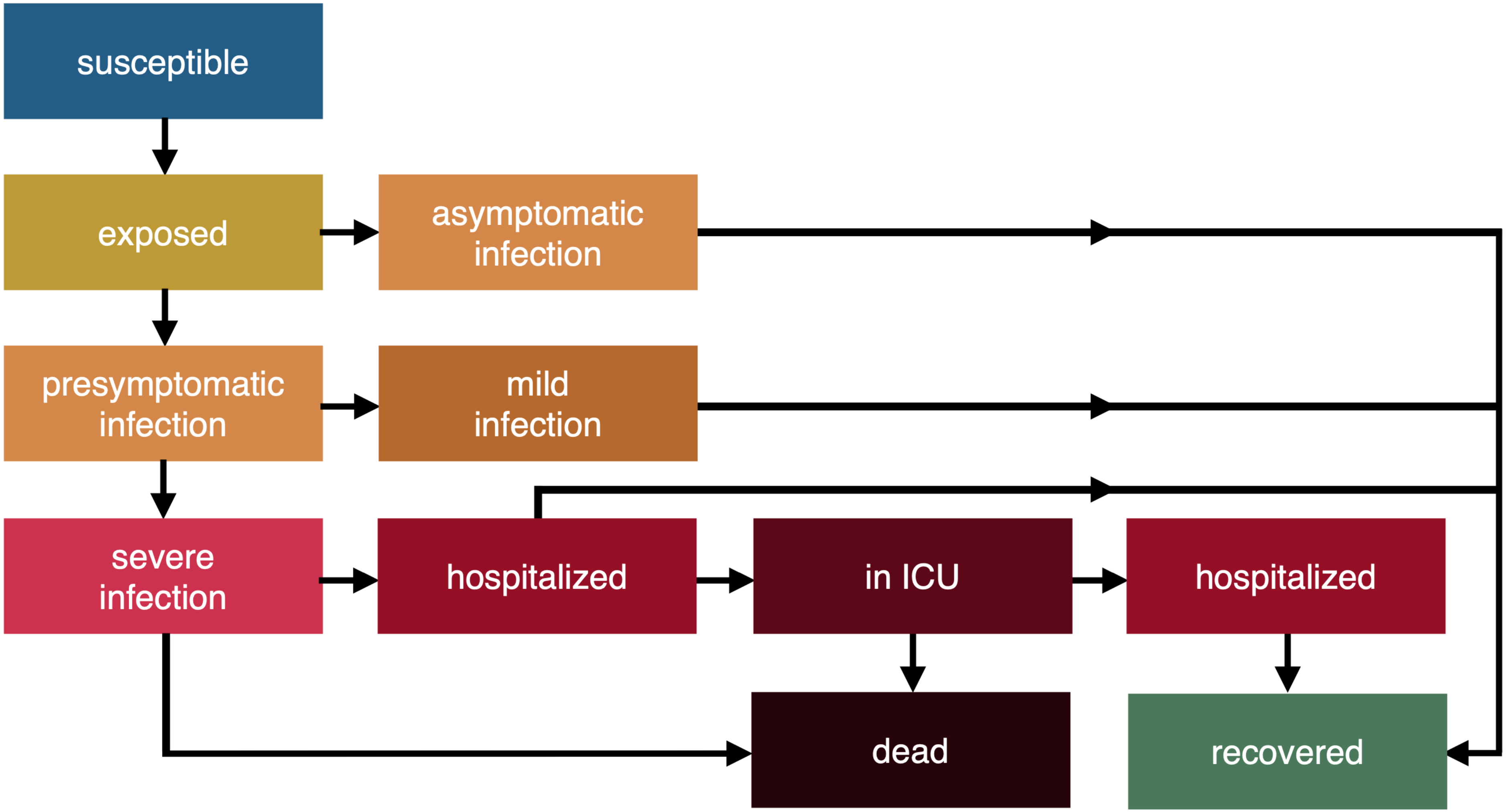
Story

- 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.

Background

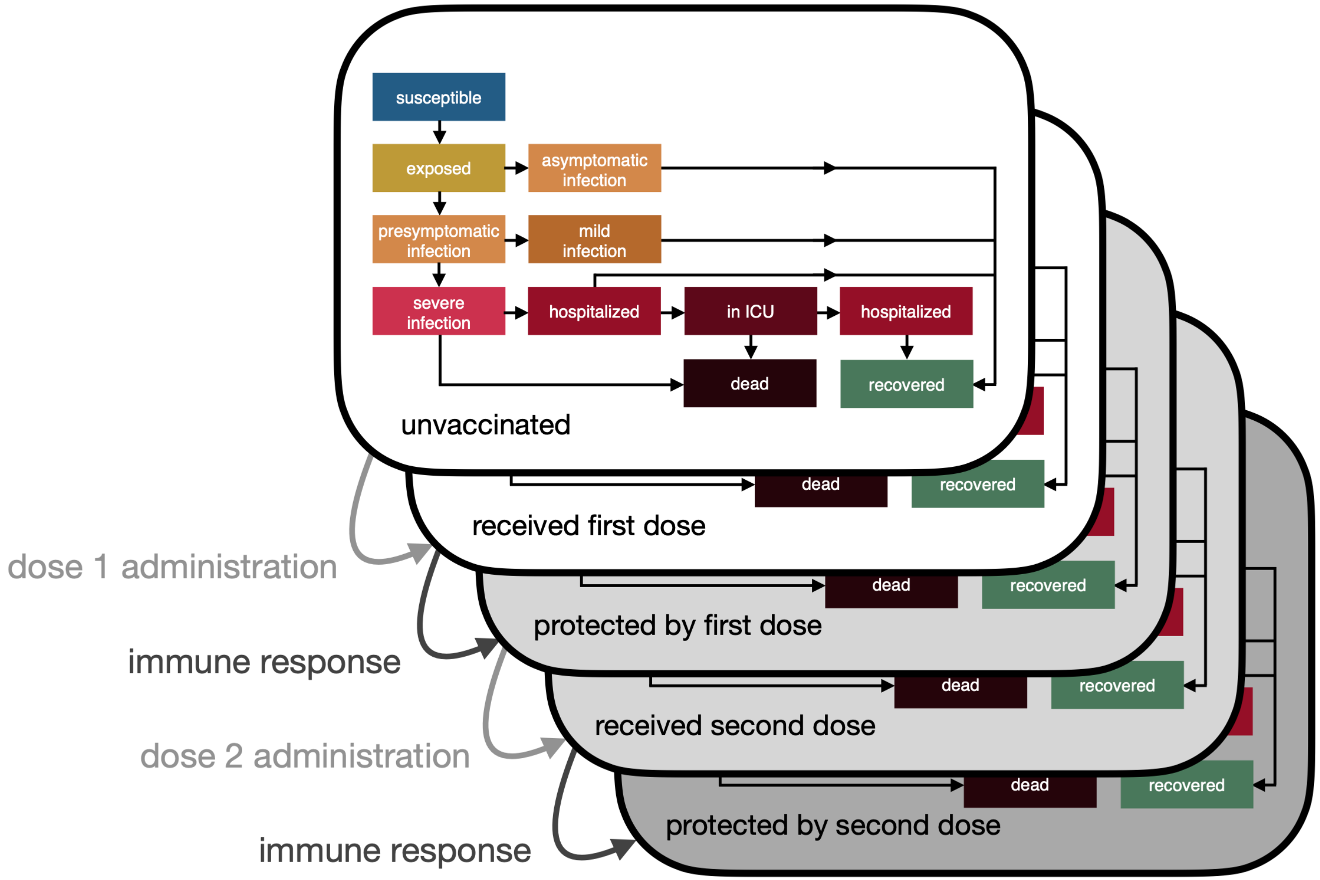
McMasterPandemic was developed for COVID-19 forecasting.

McMasterPandemic COVID-19 Model



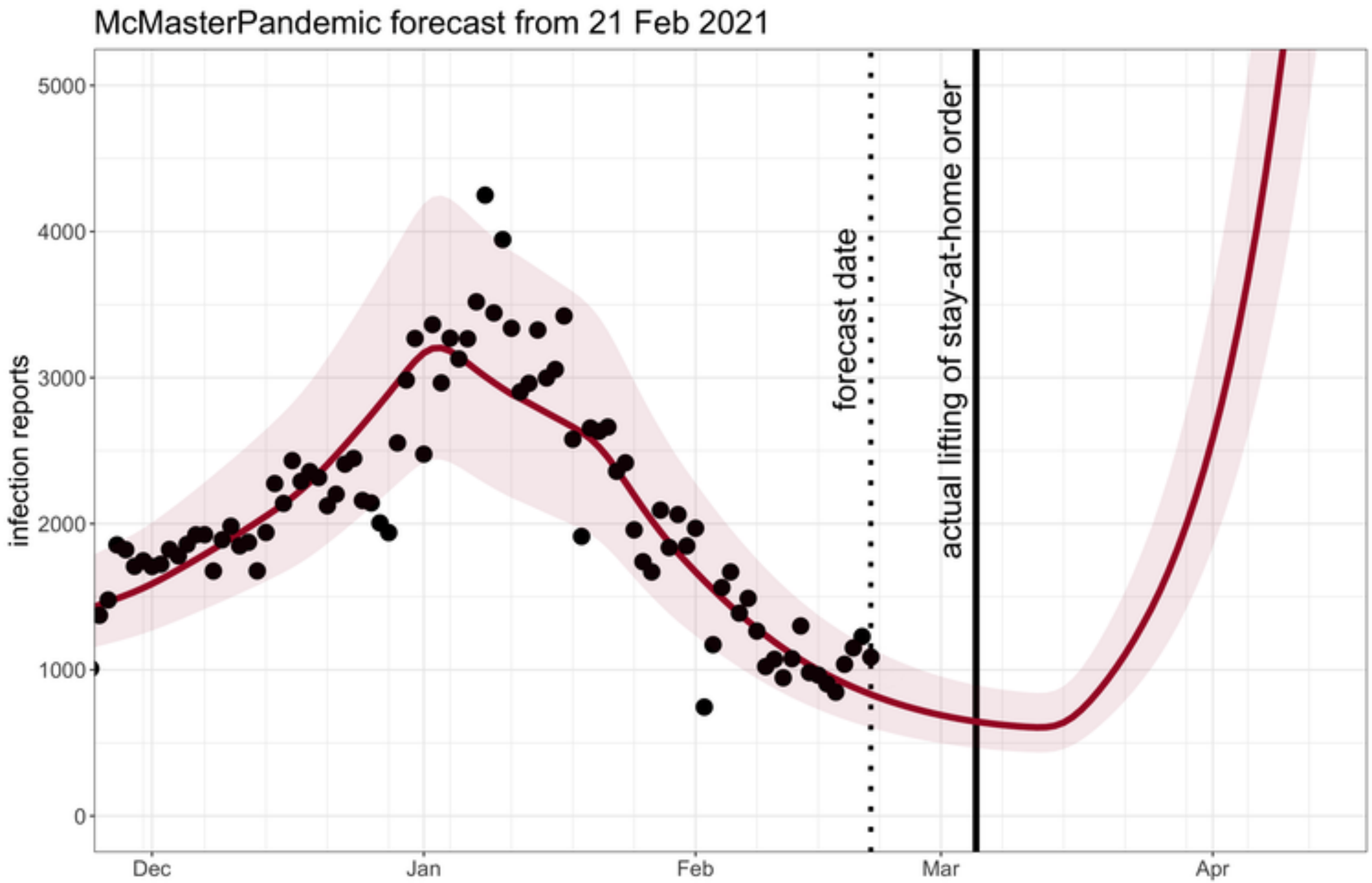
mac-theobio.github.io/covid-19

COVID-19 Vaccination Model



mac-theobio.github.io/covid-19

Forecasting



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

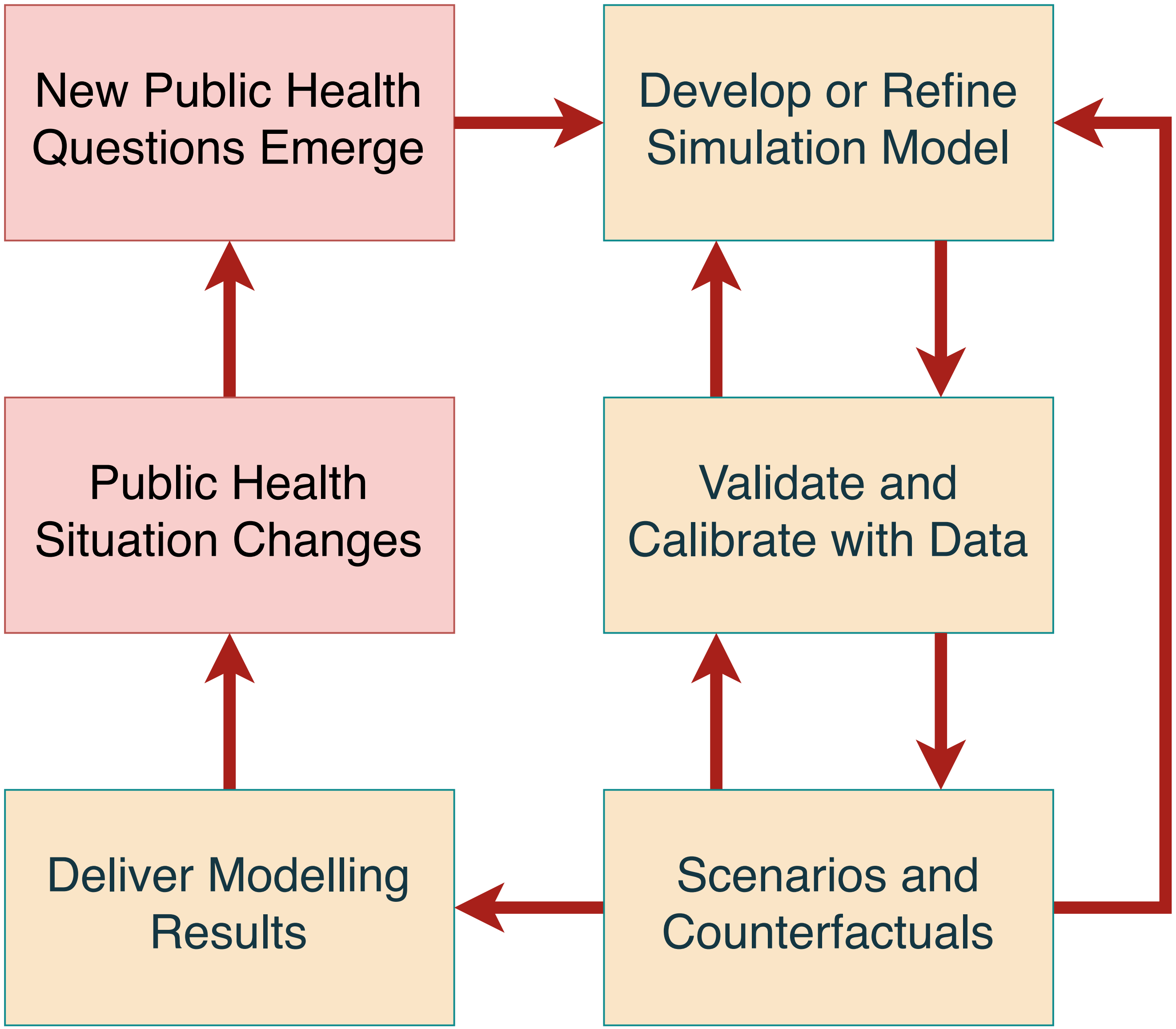
Impact

TODO: illustrate how this package was used for regular reports

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              |
+-----------+------+-------------------------------------------------------+
```

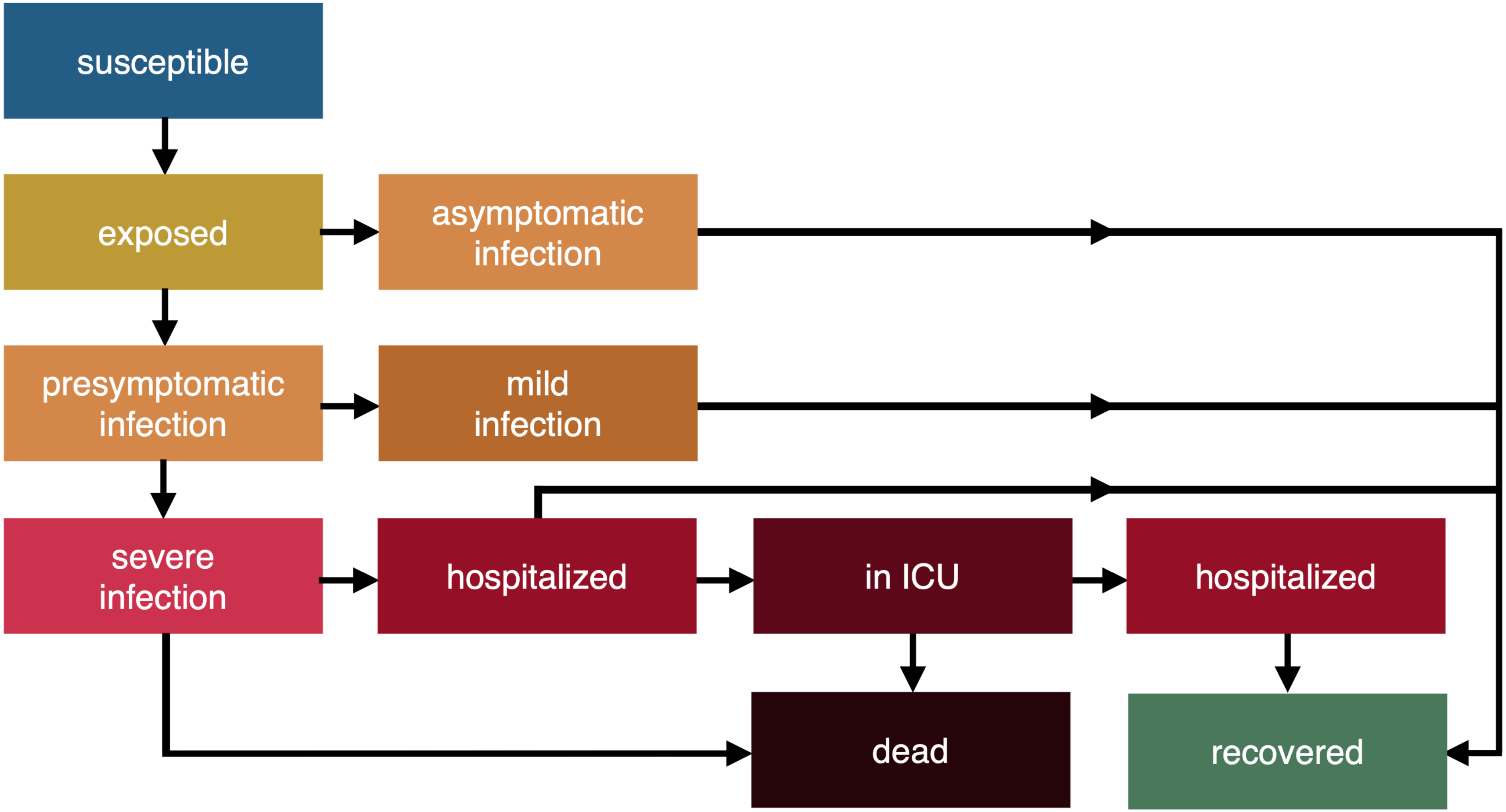
macpan1 = github.com/mac-theobio/McMasterPandemic/releases/tag/v0.0.20.1

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

macpan2 = github.com/canmod/macpan2

Mechanistic Modelling

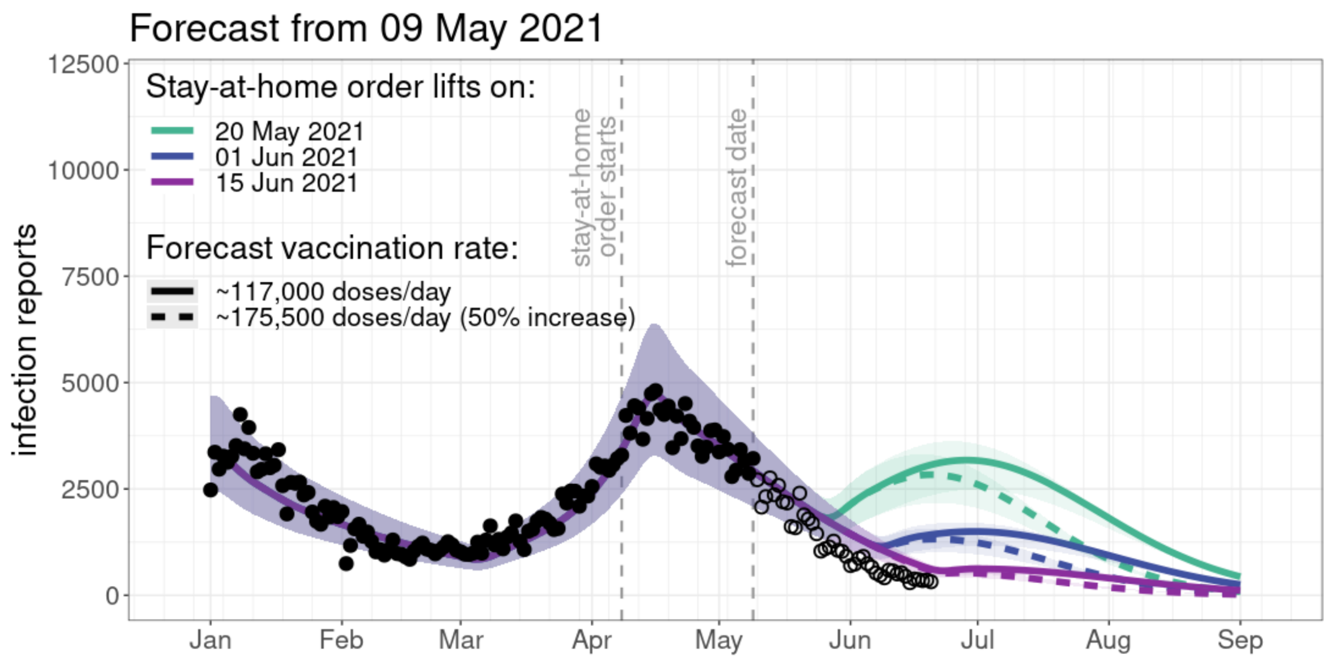
Why Not Just Use AI for Everything?



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

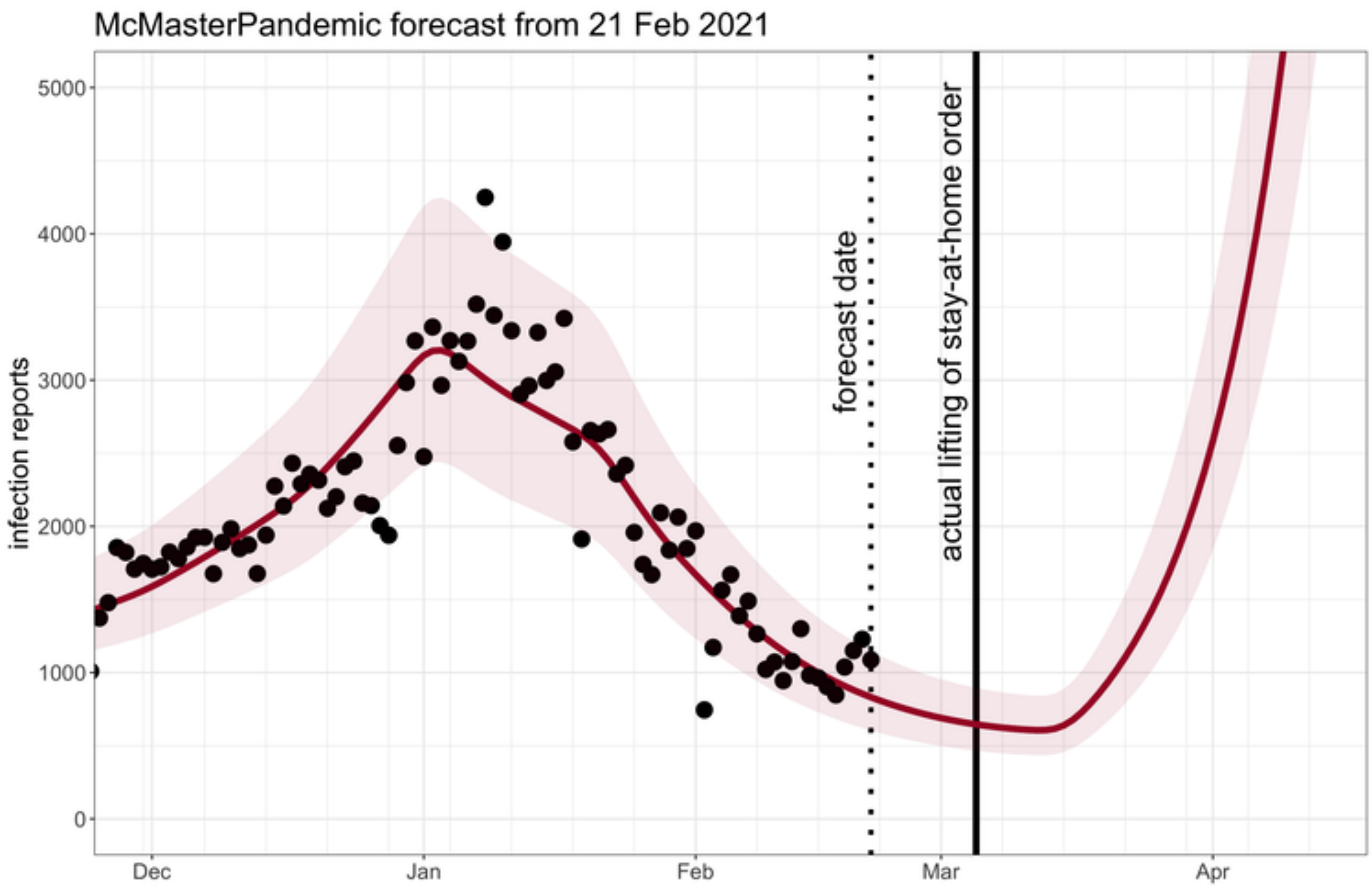
- Limitations of AI
- Weather Forecasting as a Gold Standard : Mechanistic models drive accuracy and uncertainty quantification, unlikely surpassed by AI.
- Responsibility to Decision-Makers

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

Computational Efficiency

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



Computational Efficiency

- Many simulations required.
- Technology choice: Template Model Builder – TMB[1].
- Simulation code translated to C++ (fast).
- Optimization uses automatic differentiation (fast).
- Options for statistical approaches.

1: https://github.com/kaskr/adcomp

Thinking about 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)
)
```

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

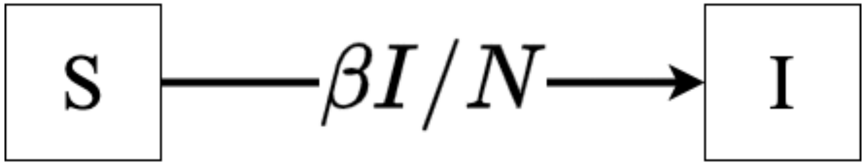
Specifying Models

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

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

Specifying Models

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



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

Specifying Models

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

$$\text{infection} = S(\beta I/N) = \beta SI/N$$

$$\frac{dS}{dt} = -\text{infection}$$

$$\frac{dI}{dt} = +\text{infection}$$

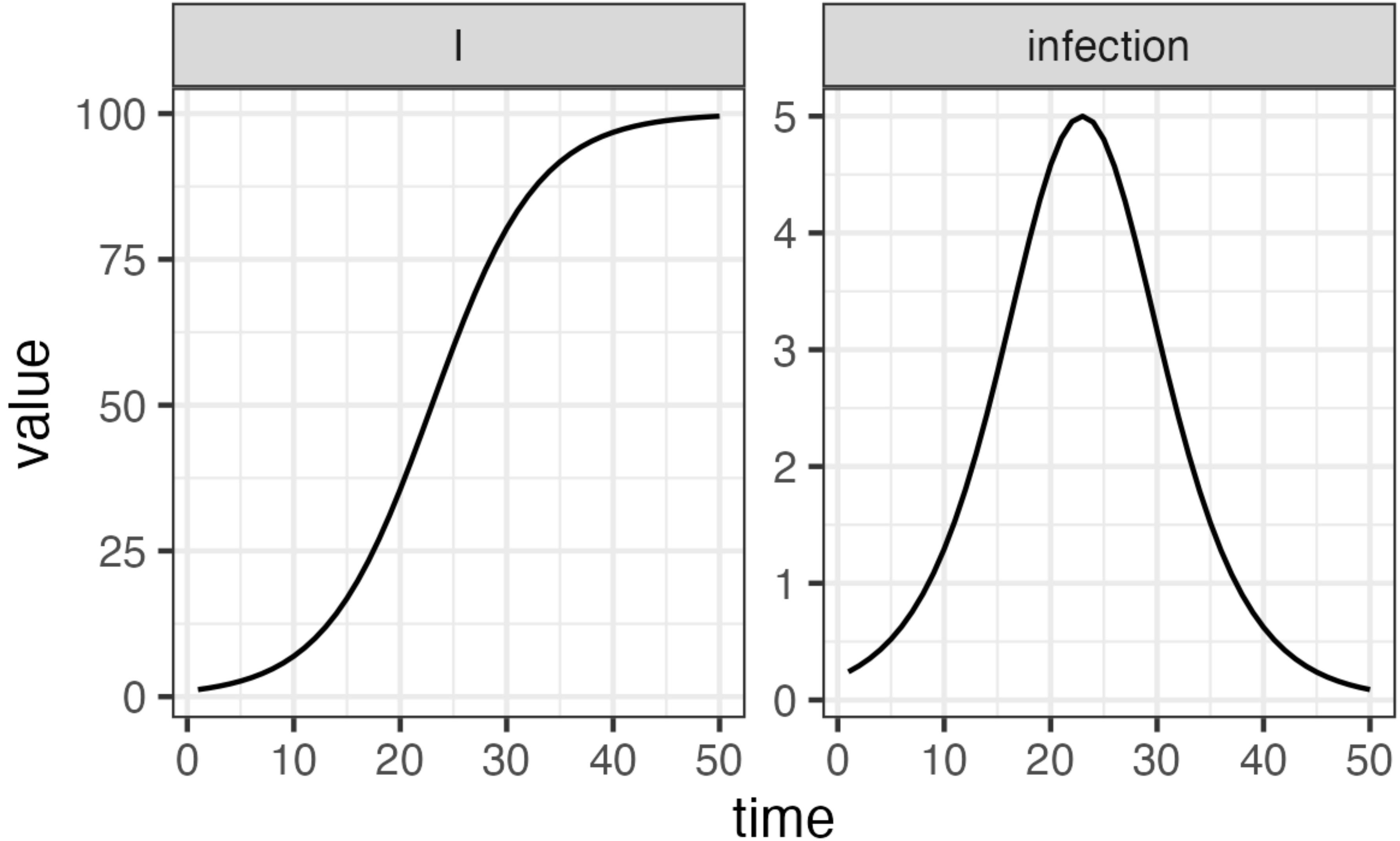
Sorry for the math. This is how our users think.

Simulation

- Philosophy: Do not reinvent the wheel
- Simulation output standardized for all models
- Easily plugs into existing popular tools
- If this code looks strange … you should check out `ggplot2` and `dplyr`
- *Many* people have found them to be very useful

```
library(macpan2); library(ggplot2); library(dplyr)
(spec
 |> mp_simulator(time_steps = 50, outputs = c("I", "infection"))
 |> mp_trajectory()

 ## macpan2 ends here and standard
 ## tools are used for plotting
 |> ggplot()
 + geom_line(aes(time, value))
 + facet_wrap(~ matrix, scales = "free")
 + theme_bw()
)
```

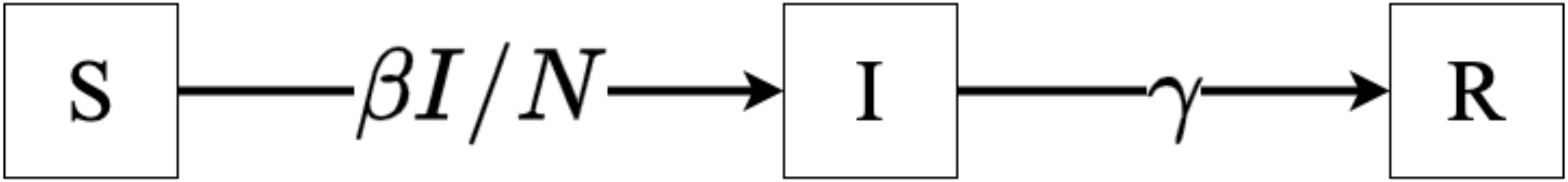
Simulation



I
infection
value
100
75
50
25
0
5
4
3
2
1
0
0 10 20 30 40 50
0 10 20 30 40 50
time

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" )
  ...
)
```

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

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" )
  ...
)
```

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

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

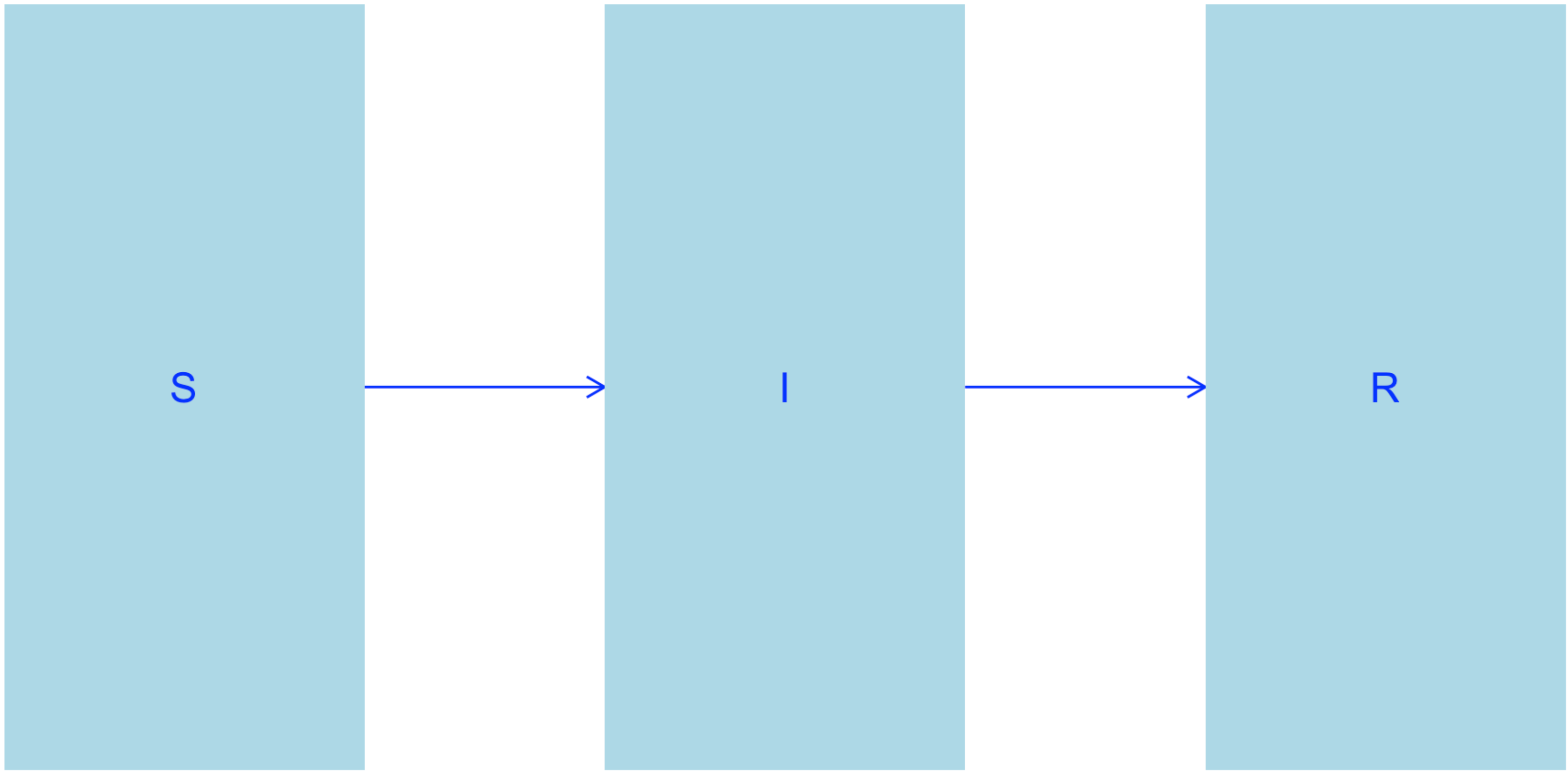
Drawing Flow Diagrams

- Showed how specification can be used to simulate epidemic trajectories.
- Specifications allow us to draw flow diagrams (experimental feature).

```
(spec                   ## model specification
  |> mp_flow_frame()   ## data frame describing flows
  |> mp_layout_path() ## find where to place boxes
  |> mp_plot_layout() ## draw diagram
)
```

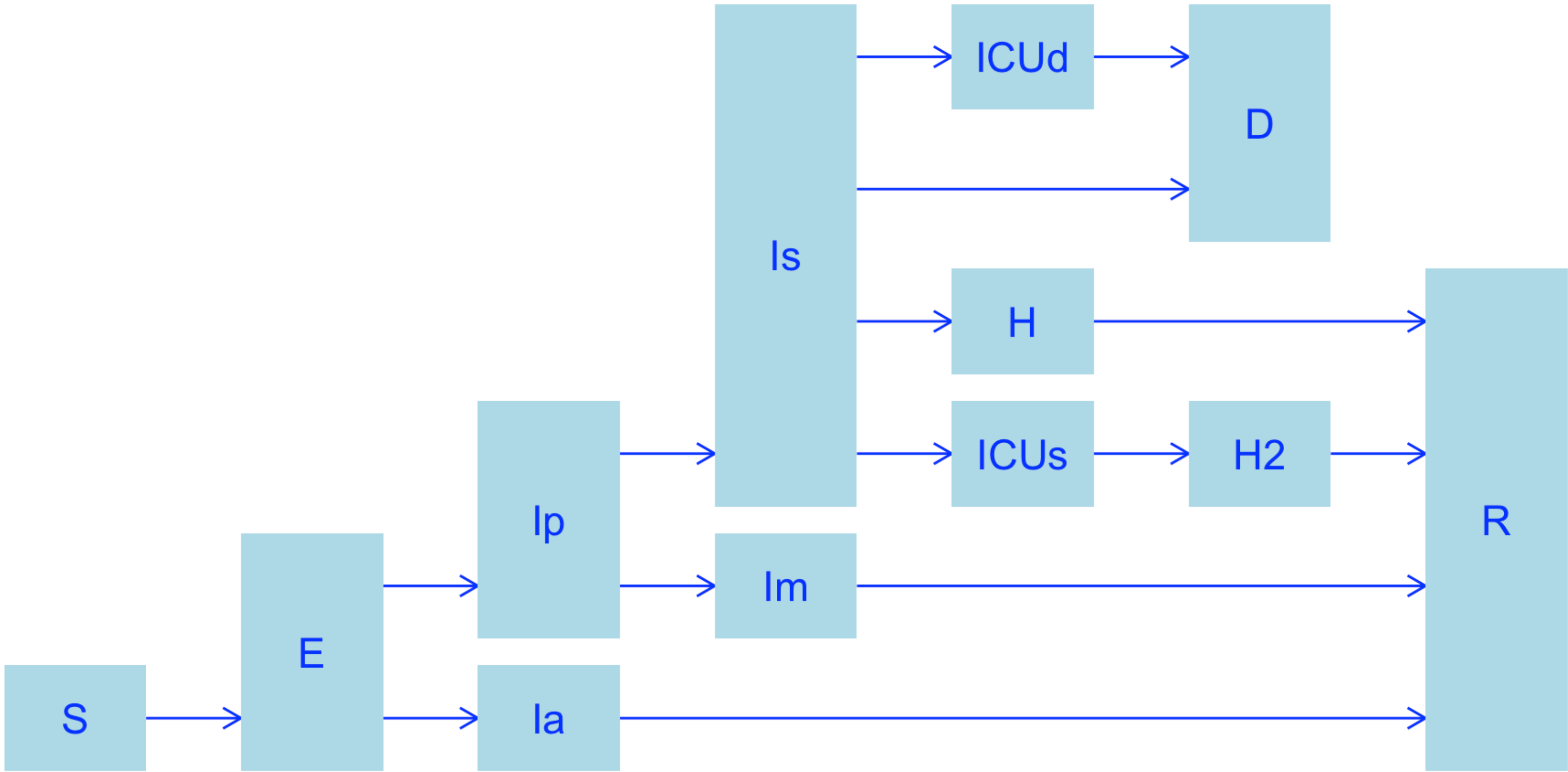
Drawing Flow Diagrams

- Showed how specification can be used to simulate epidemic trajectories.
- Specifications allow us to draw flow diagrams (experimental feature).



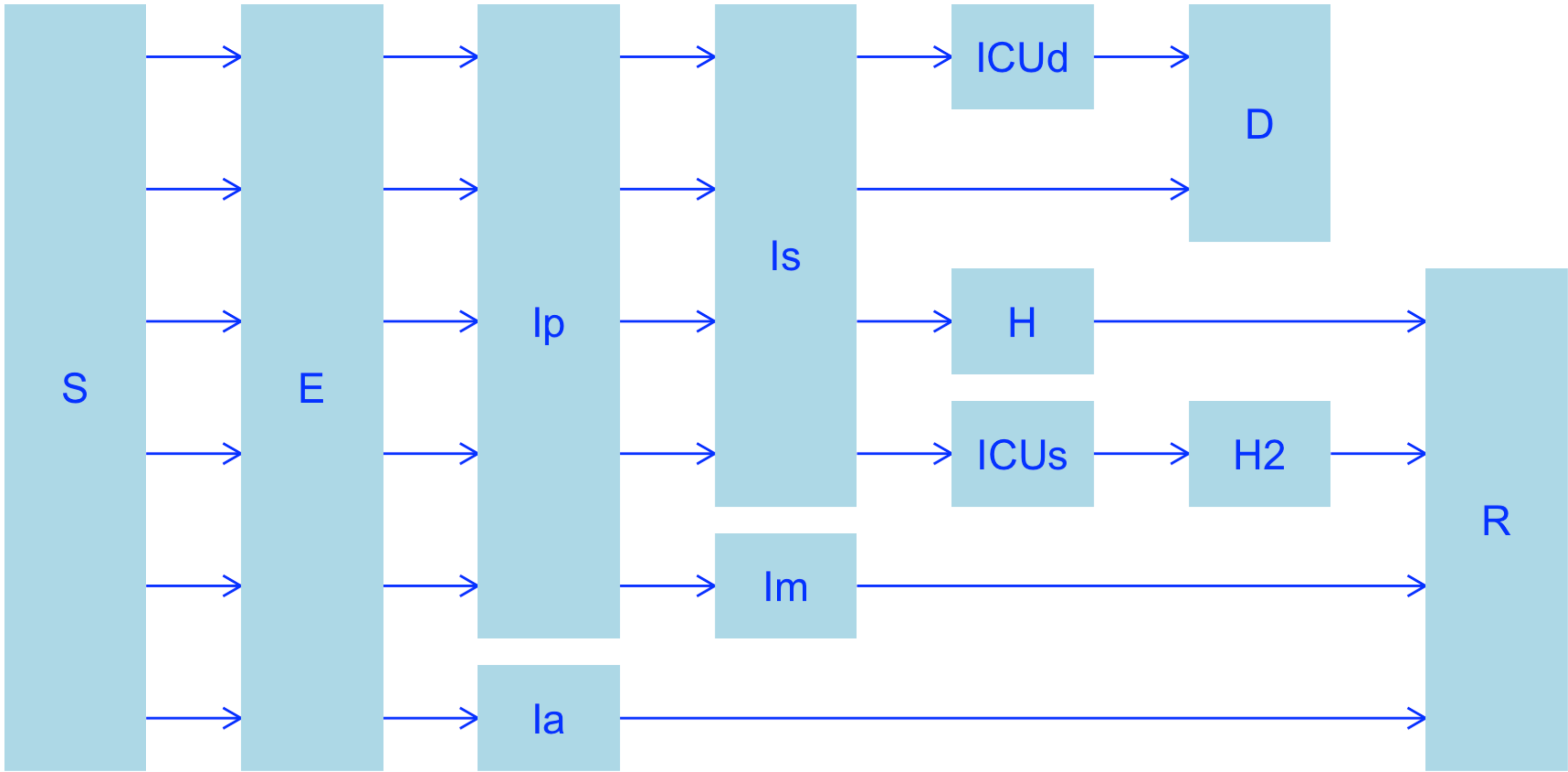
Drawing Flow Diagrams

- 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.



Workshops on `macpan2`

- **Exploration** : Compare simulations with each other & data
- **Parameterization** : Calibrate/refine parameters for a specific question.
- **Inference** : Use parameterized models to make inferences.
- **Stratification** : Separate each compartment into many (e.g. by age, space, vaccination status).

https://canmod.github.io/macpan-workshop/syllabus

Resources

- 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

Model Library

Instead of building models from the ground up, applied modellers can leverage macpan2's library of predefined models as starting points.

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

Using the Model Library

Read in example models using code like this.

```
mp_tmb_library("starter_models", "hiv"        , package = "macpan2")
mp_tmb_library("starter_models", "macpan_base", package = "macpan2")
mp_tmb_library("starter_models", "ww"         , package = "macpan2")
```

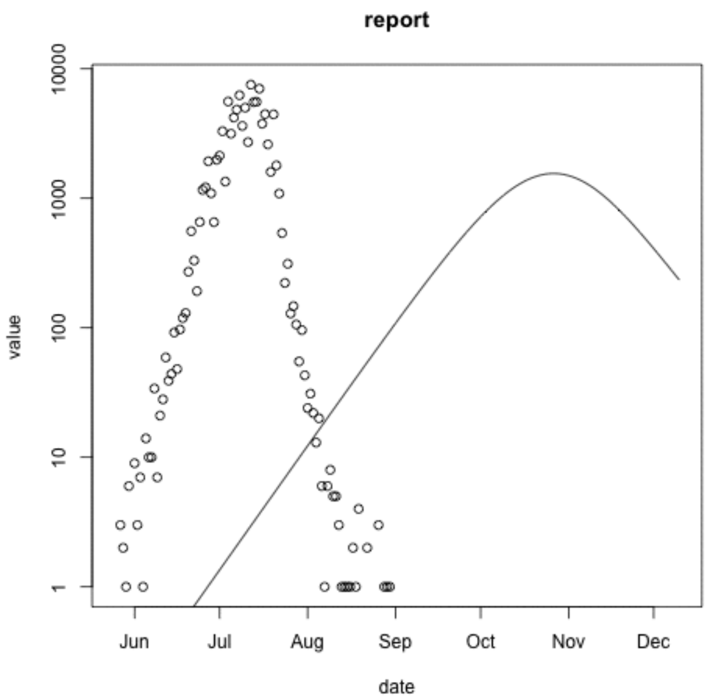
Using the Model Library

```
mp_tmb_library("starter_models", "si", package = "macpan2")
---------------------
Default values:
---------------------
 matrix row col value
      N          100.0
   beta            0.2
      I            1.0

---------------------
Before the simulation loop (t = 0):
---------------------
1: S ~ N - 1

---------------------
At every iteration of the simulation loop (t = 1 to T):
---------------------
1: mp_per_capita_flow(from = "S", to = "I", rate = "beta * I / N",
     abs_rate = "infection")
```

Calibration

`macpan2` employs formal mathematical optimization for the efficient calibration of model parameters to data, even when these parameters evolve over time.

Calibration – Trajectory Matching



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

Calibration – Interface

TODO Maybe

Illustrations

We can illustrate macpan2's application with a variety of examples including COVID-19, Mpox, and Measles.

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

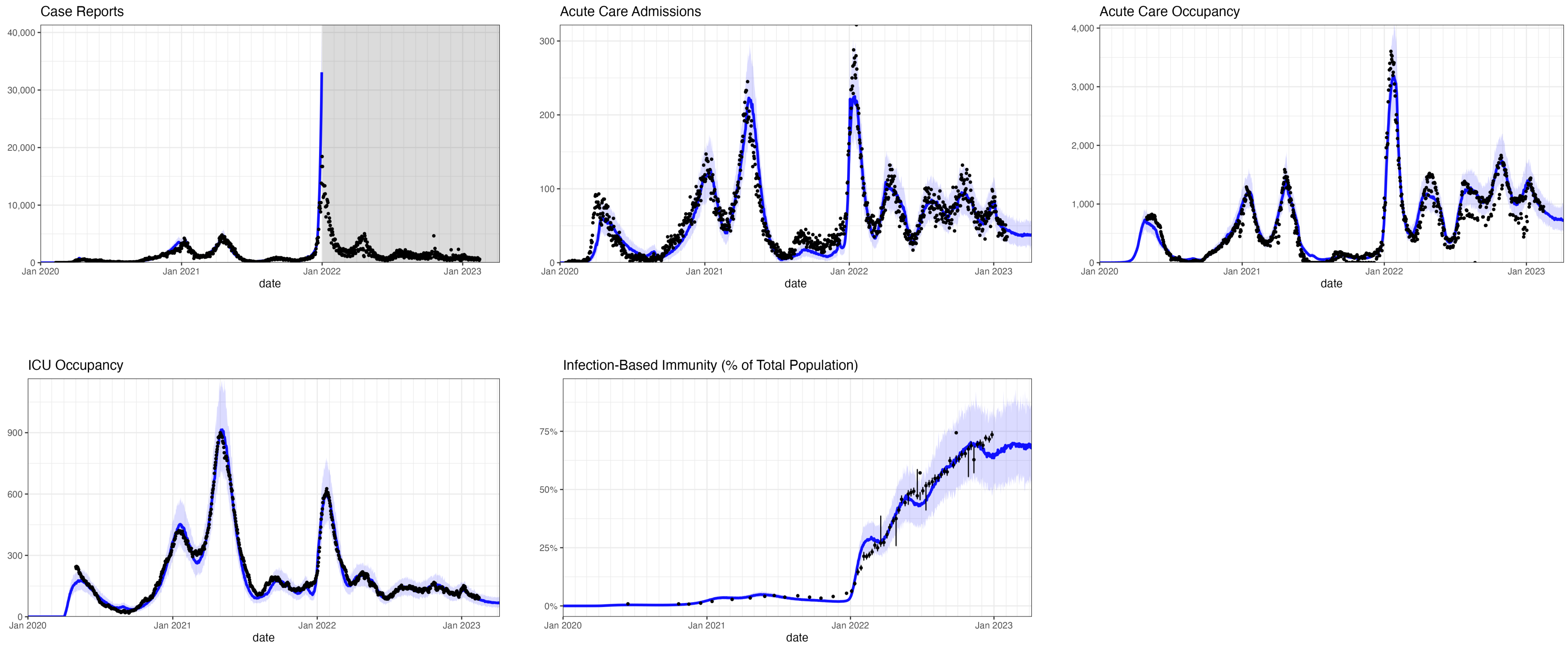
Examples of Usage

- **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

Works on 'Big' Problems

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

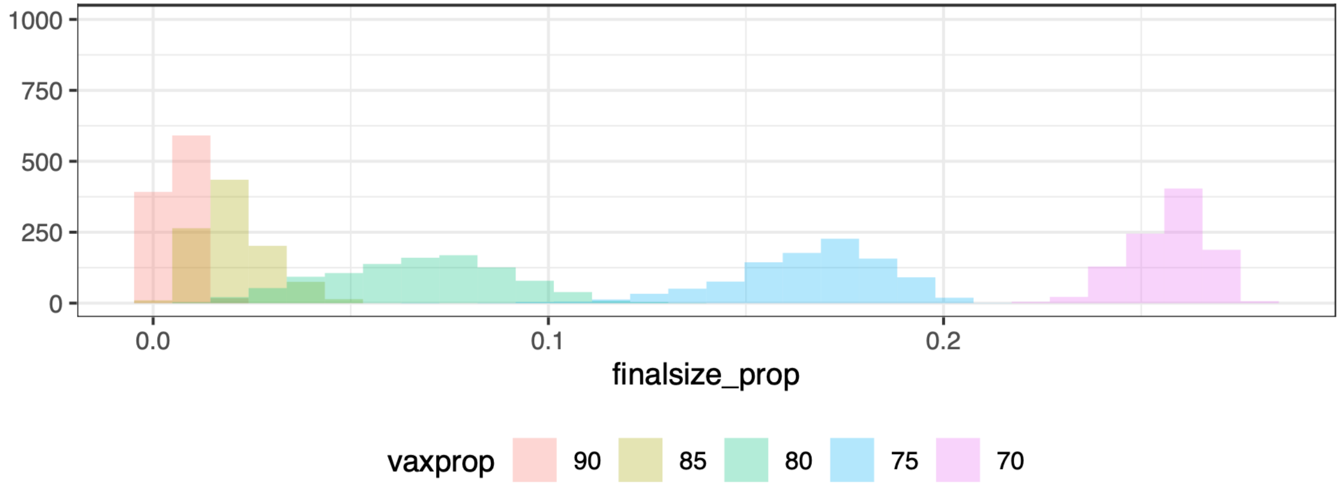


Effect of Vaccination

TODO: add this work maybe?

Measles

- Vaccination in small (~1000-person) communities
- Stochasticity generates final size distributions



Collaboration with Mike Li (PHAC) to support public health response.

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

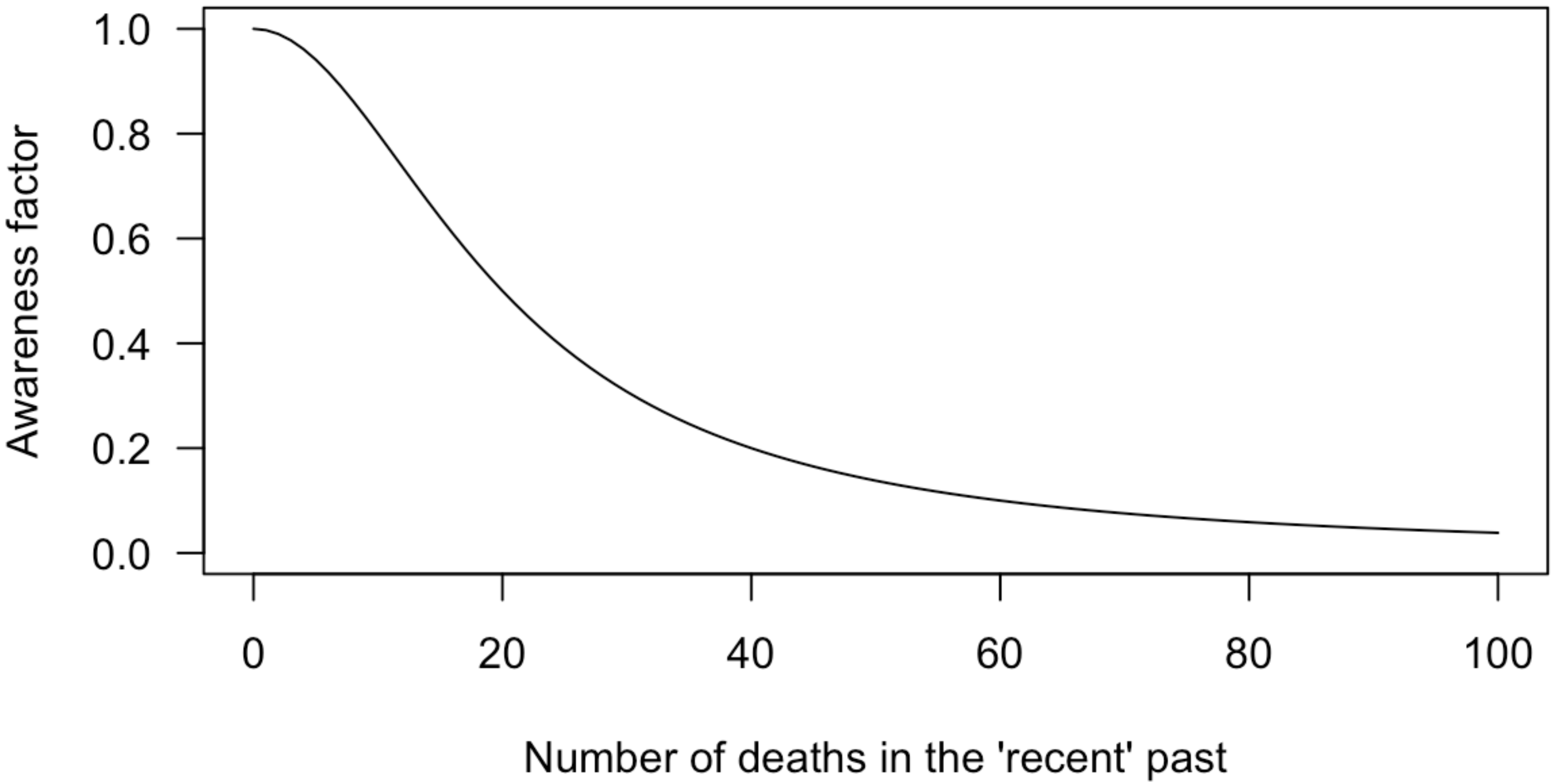
3: https://github.com/canmod/macpan2/tree/main/inst/starter_models/awareness

Small-N Awareness Models

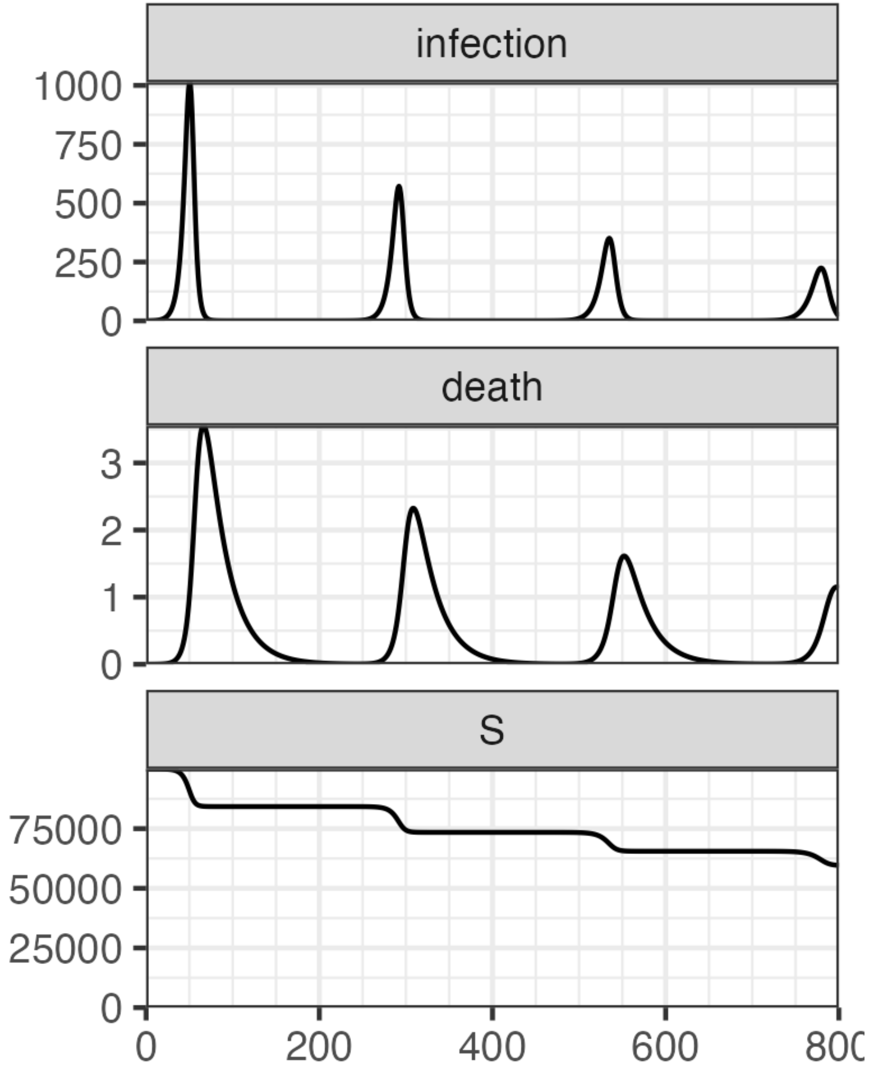
- Why am I showing you small-N awareness models?
- Not because I think they are novel.
- But because they illustrate the software.
- Provides a simple but realistic example (good model of COVID in NFLD).
- Vehicle for digging into epidemiological details of a particular example.
- A colleague was interested in these models, and I wanted to try it out.
- I made this model in ‘no time’.
- It is now in the library.

Small-N Awareness Models

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

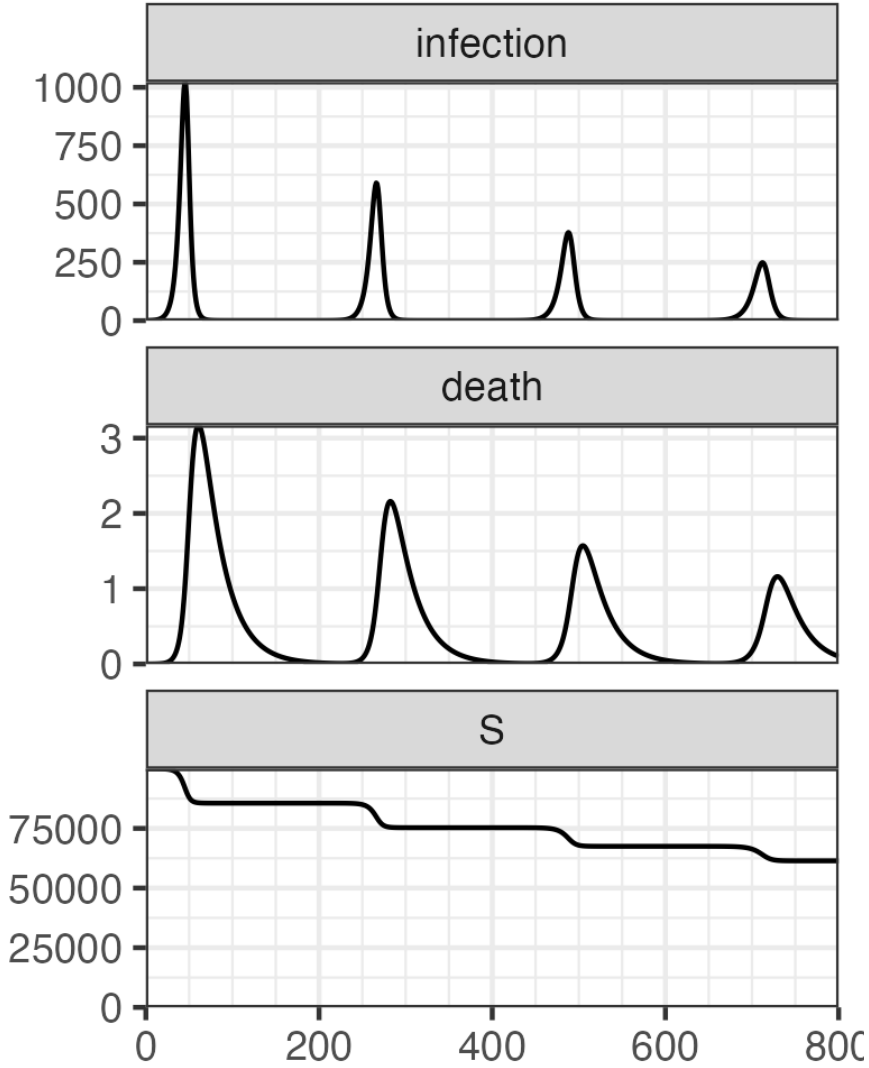


Small-N Awareness Models



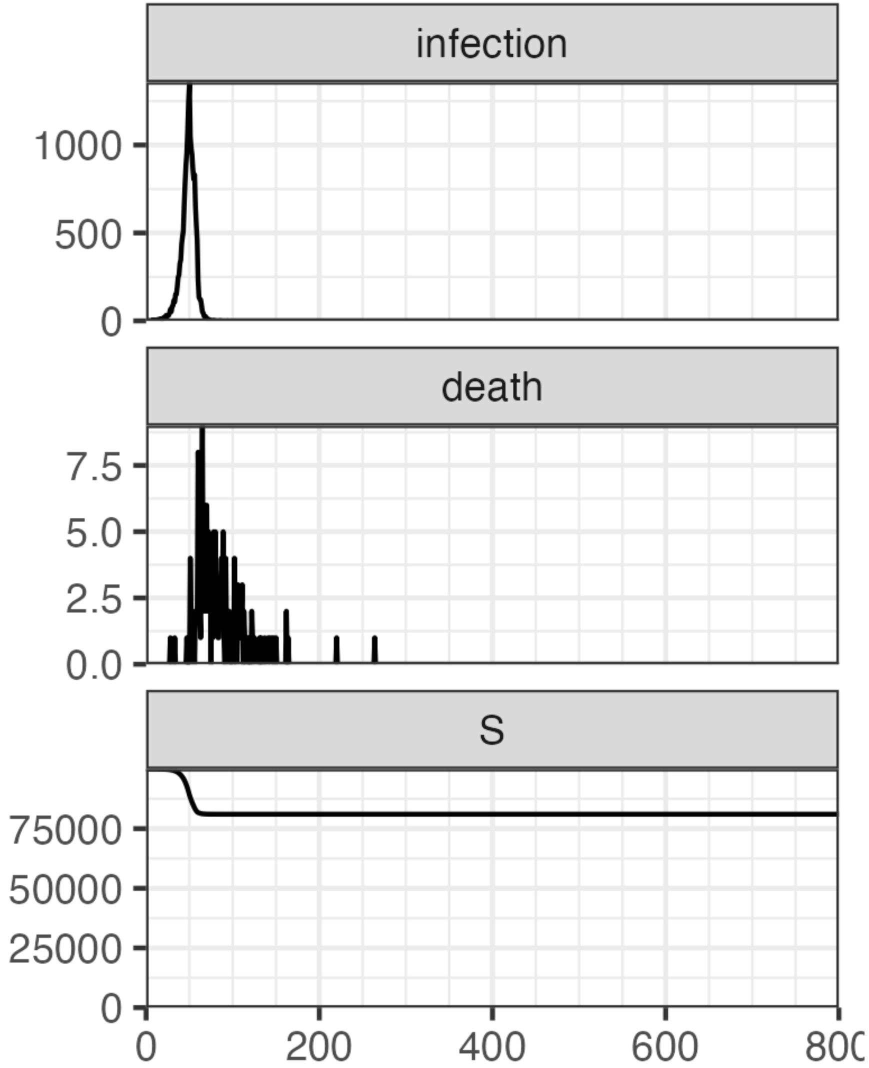
- No stochasticity, no importation
- Regularly-spaced cycles that slowly decrease in amplitude.
- Suceptible depletion is limited. Why?
- Turn-over behavioural (not due to herd immunity).

Small-N Awareness Models



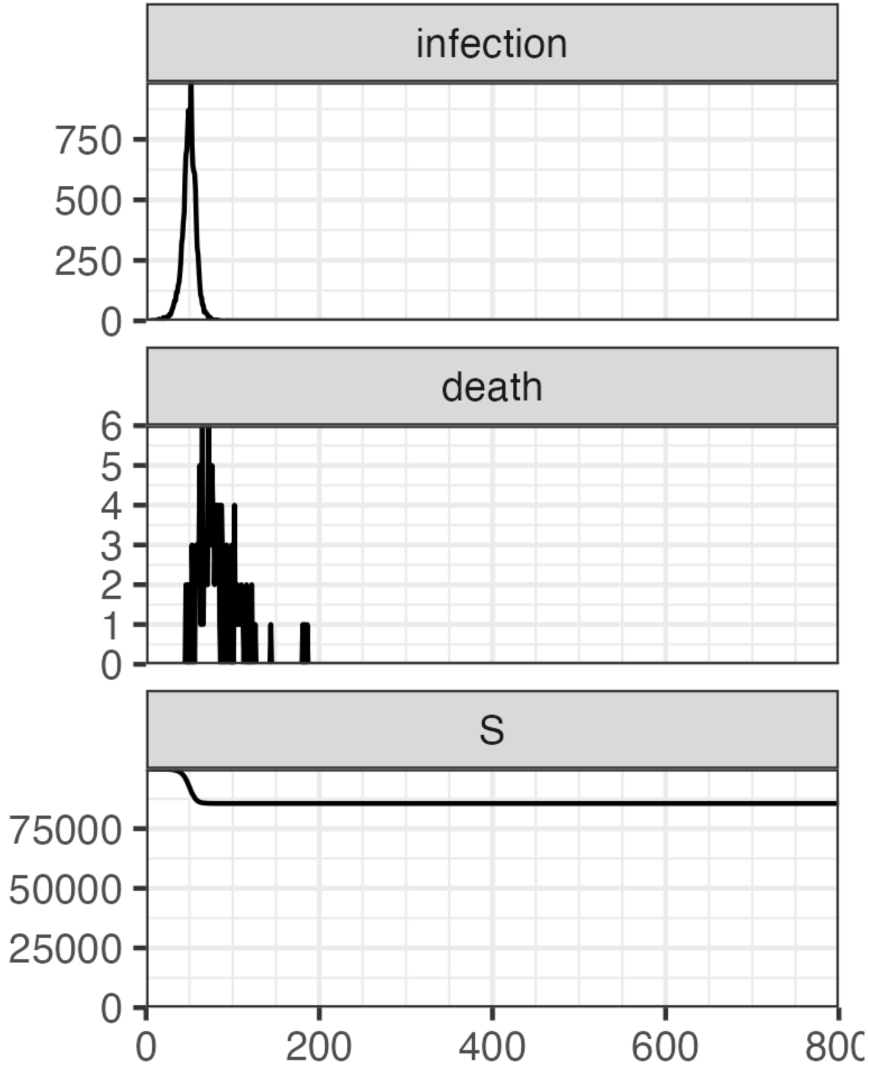
- Doesn’t matter what ODE solver is used.
- This is Runge-Kutta-4, the previous slide was Euler.

Small-N Awareness Models



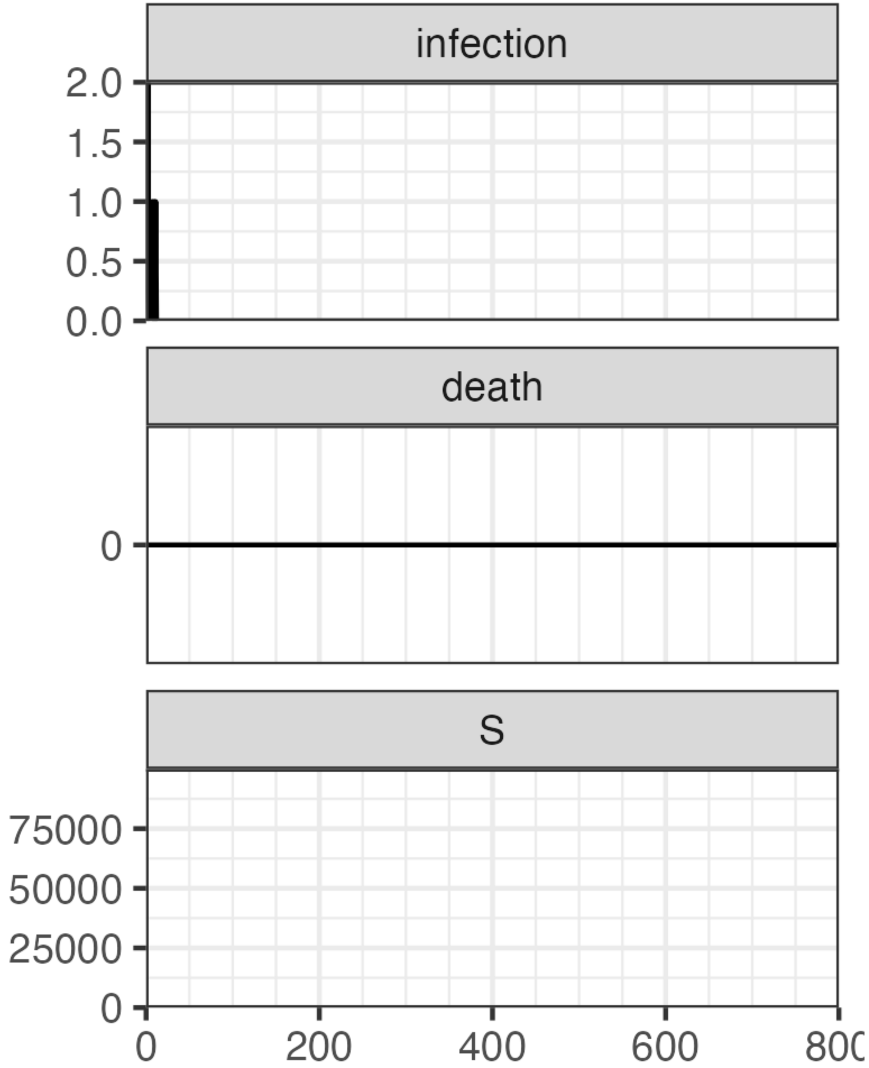
- With stochasticity we can have stochastic eradiction.
- Why do we not have a second wave?
- Because infectious population goes to exactly zero before behaviour gets risky again.

Small-N Awareness Models



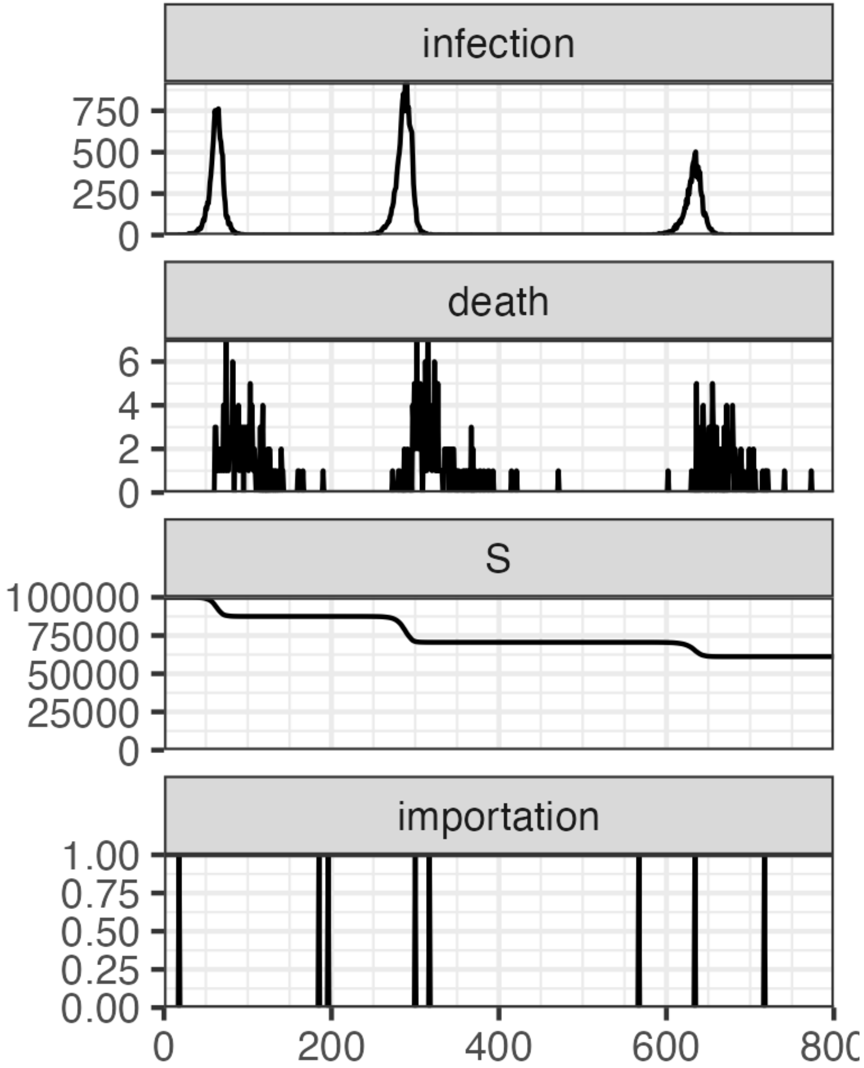
- The exact shape of the outbreak depends on the random seed.

Small-N Awareness Models



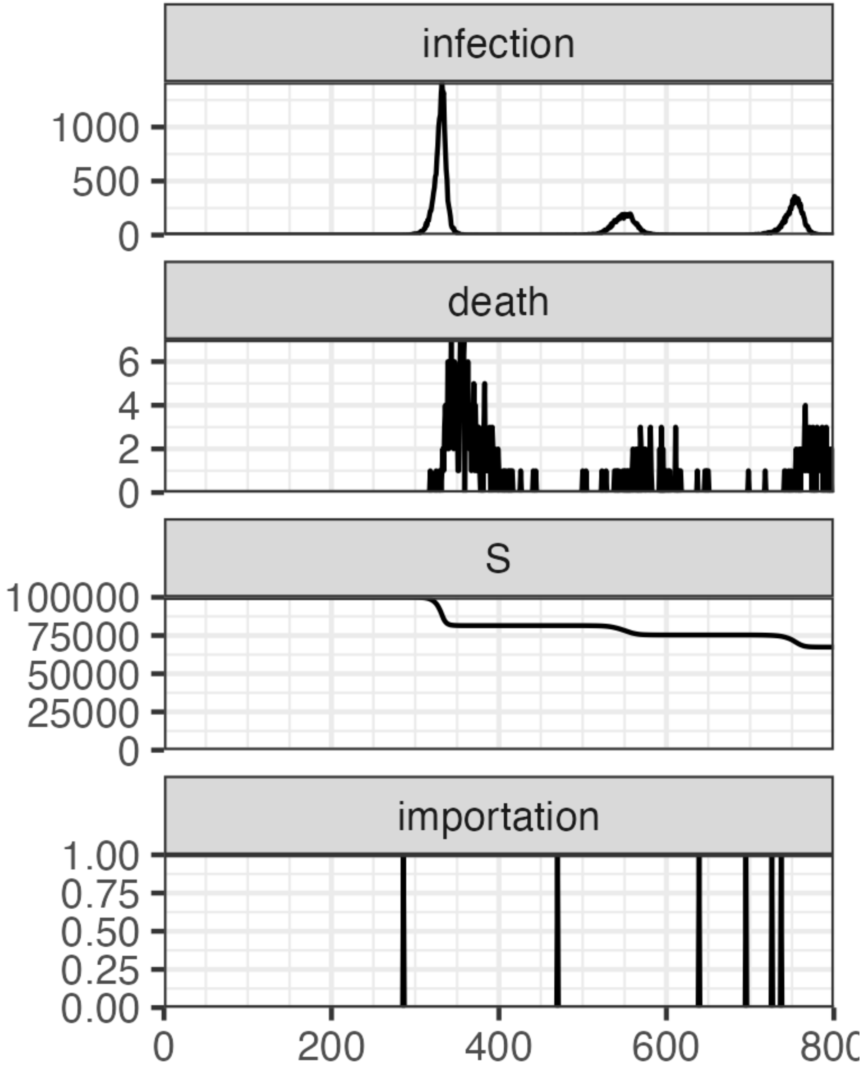
- Sometimes the random seed yields no outbreak at all.

Small-N Awareness Models



- Adding importation can yield randomly located outbreaks.
- Importation keeps the infectious population from staying at zero.

Small-N Awareness Models



- The location of the outbreaks depend on the pattern of importation.
- The long period without outbreaks is caused by no importation.

Small-N Awareness Models

```
awareness_model = mp_tmb_model_spec(
  ...
  , mp_per_capita_flow("S", "E"
      , "beta * I / N / (1 + (convolution(death, kernel)/delta_c)^k))"
      , "infection"
    )
  ...
)
```

Small-N Awareness Models

```
importation_awareness_model = mp_tmb_insert(awareness_model
  , expressions = list(
        importation ~ rbinom(1, importation_prob)
      , I ~ I + importation
    )
  ...
)
```

Open-Source Tools

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

Comparing 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://canmod.github.io/macpan2/

Conclusions

- Translating pandemic modeling experience into open-source software for future preparedness.
- The `macpan2` project is part of this effort.
- Ongoing incremental improvements.
- Focus on model specification, simulation, and calibration.
- Training modellers on `macpan2` (Interested?)