



A Novel Approximate Bayesian Inference Method for Compartmental Models in Epidemiology

CPS 07
Epidemiological Modelling

Ms. Xiahui Li, Dr Fergus Chadwick, Dr Ben Swallow
University of St Andrews
Monday 06 October, 4:00PM - 5:00PM

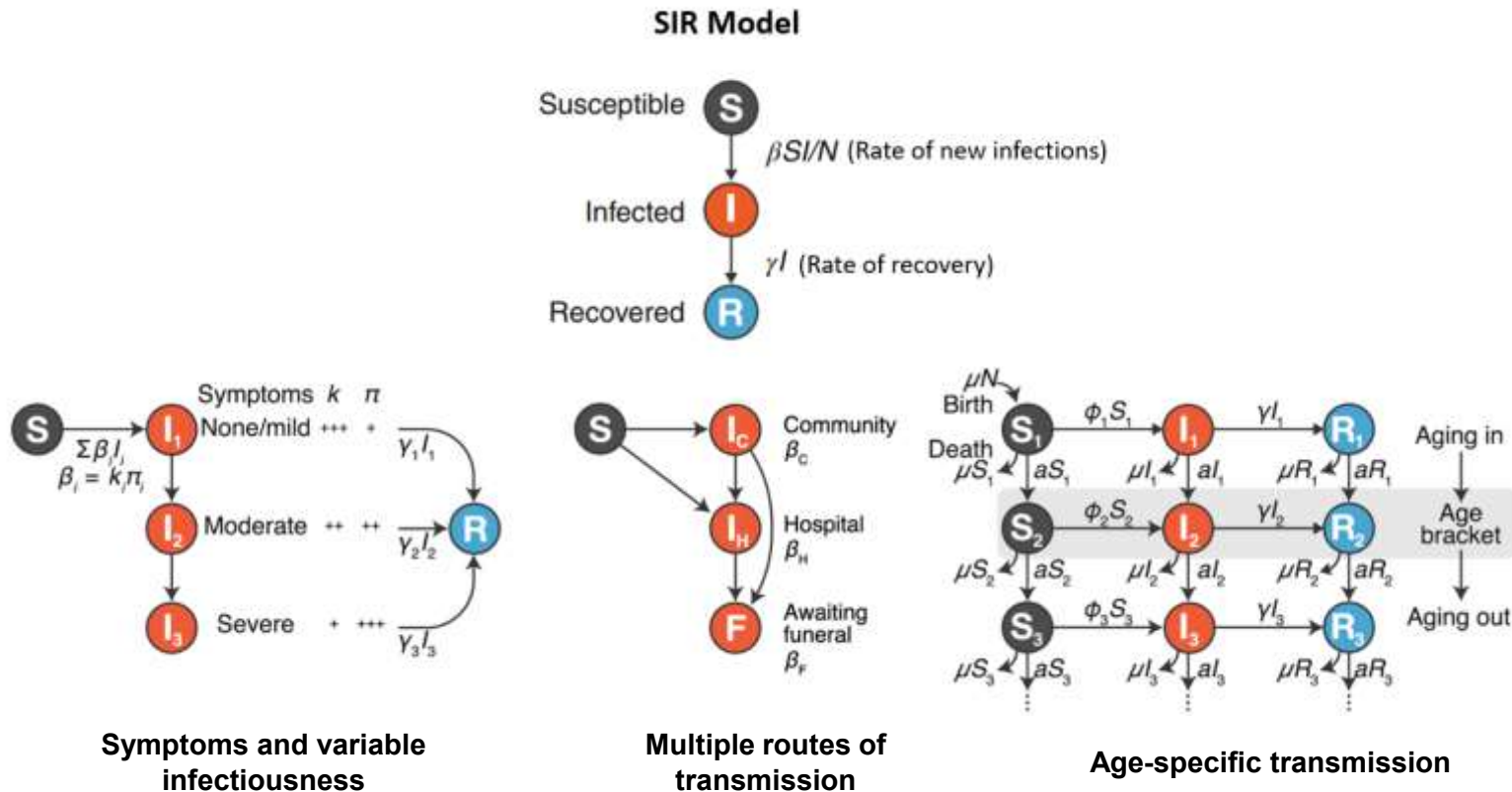


Figure 1: SIR model extensions. Adapted from "The SEIRS model for infectious disease dynamics," by Bjørnstad, O. N., Shea, K., Krzywinski, M., & Altman, N. , 2020.

High-dimensional parameter spaces

(e.g., transmission rates,
contact rates, recovery rates,
birth and death)

Latent variables

(e.g., infection time,
transmission events)

**Challenges in
Parameter
Inference**

Incomplete or noisy data

(e.g., Missing case reports
Underreporting
Imprecise observations)

Uncertainties in model structure

(e.g., model assumptions, mixing
patterns, inclusion of environmental
factors)

Bayes' rule

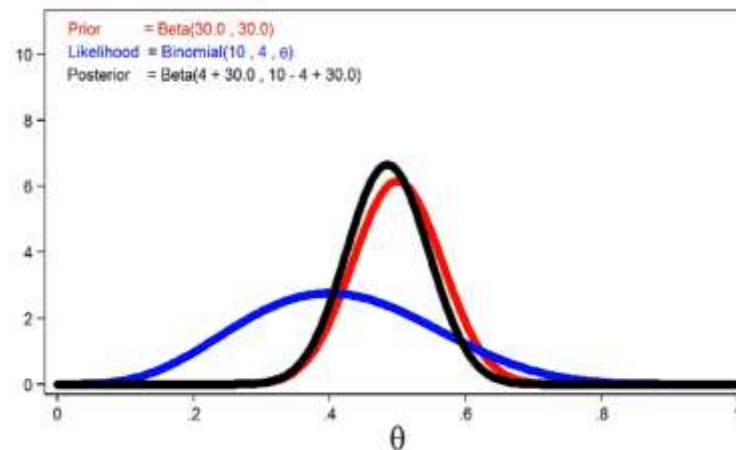
$$\pi(\theta|y_{obs}) \propto \pi(y_{obs}|\theta)\pi(\theta)$$

↓
↓

Posterior
Prior

↑
↑

Likelihood



GIF Animation 1: The effect of larger sample sizes on the posterior distribution. Adapted from “Introduction to Bayesian statistics,” by Chuck Huber, 2016.

<https://blog.stata.com/2016/11/01/introduction-to-bayesian-statistics-part-1-the-basic-concepts/>

1.4 Asymptotically Exact and Approximate Bayesian Method

$$\pi(\theta|y_{obs}) \propto \pi(y_{obs}|\theta)P(\theta) \text{ in Complex Model}$$

Computationally Prohibitive

Analytically Impossible

TRACTABLE Likelihood

INTRACTABLE Likelihood

Asymptotically Exact Bayesian Method

Directly sample from the true posterior distribution

Popular Examples

Metropolis-Hastings Algorithm
Hamiltonian Monte Carlo (HMC)

Approximate Bayesian Method

Bypass direct likelihood computation by matching simulated data to observations.

Popular Examples

Approximate Bayesian Computation
Synthetic Likelihood Method

2.1 Approximate Bayesian Computation

Input

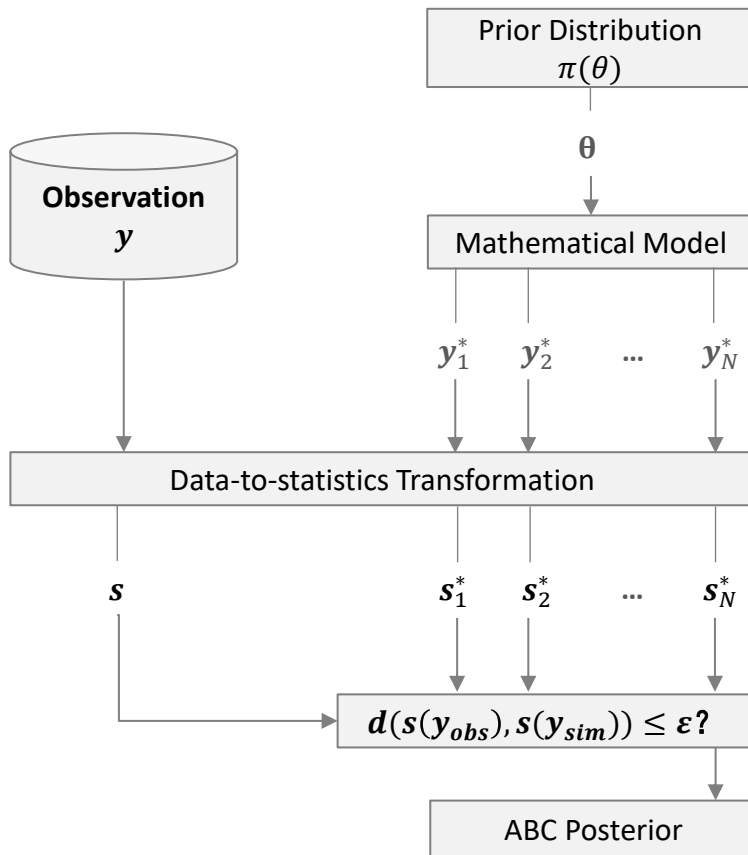
Observed data y_{obs}

Prior $\pi(\theta)$

Summary Statistics $s(\cdot)$

Distance $d(\cdot, \cdot)$

Tolerance ε



Step 1: Draw candidate parameter from prior

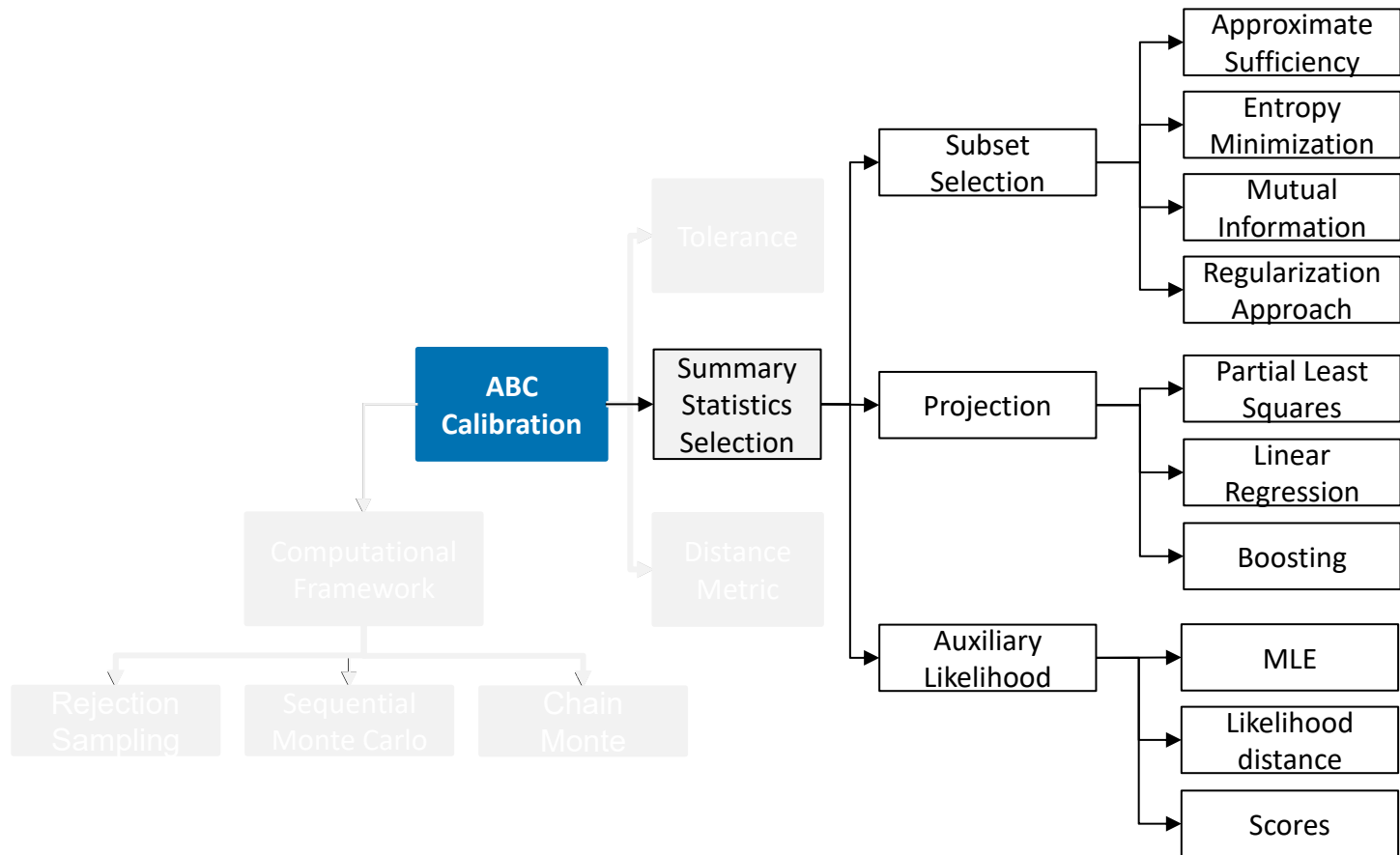
Step 2: Simulate data given parameters

Step 3: Summarize data

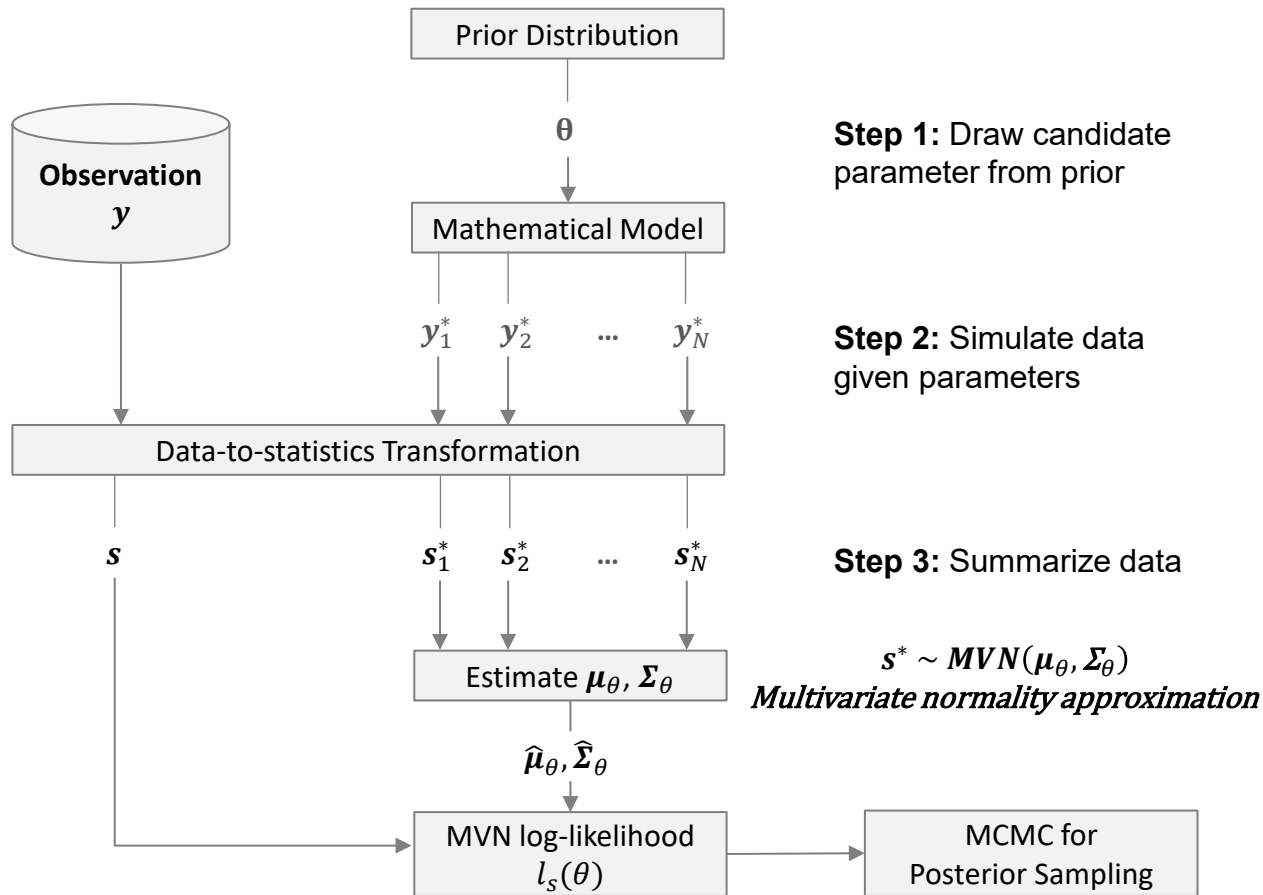
Step 4: Accept or reject candidate parameter

Step 5: Repeat until posterior samples is obtained.

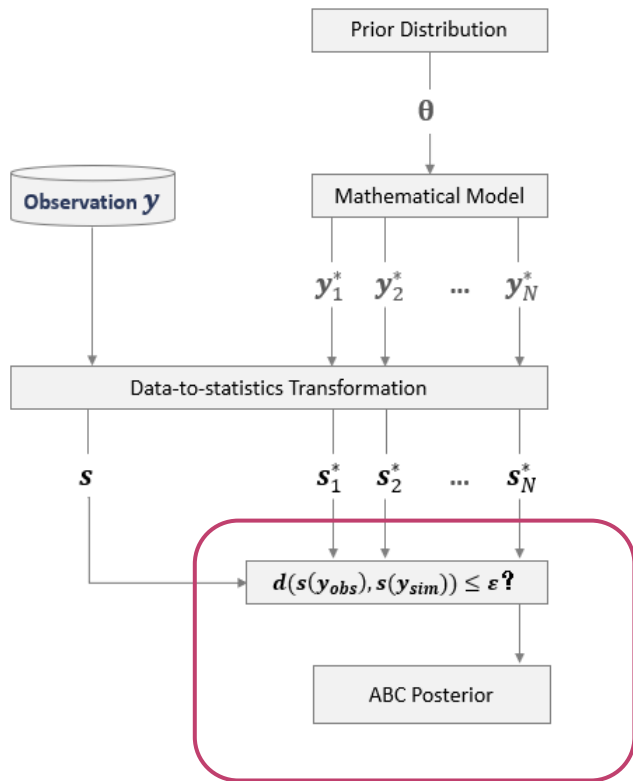
2.1 Approximate Bayesian Computation



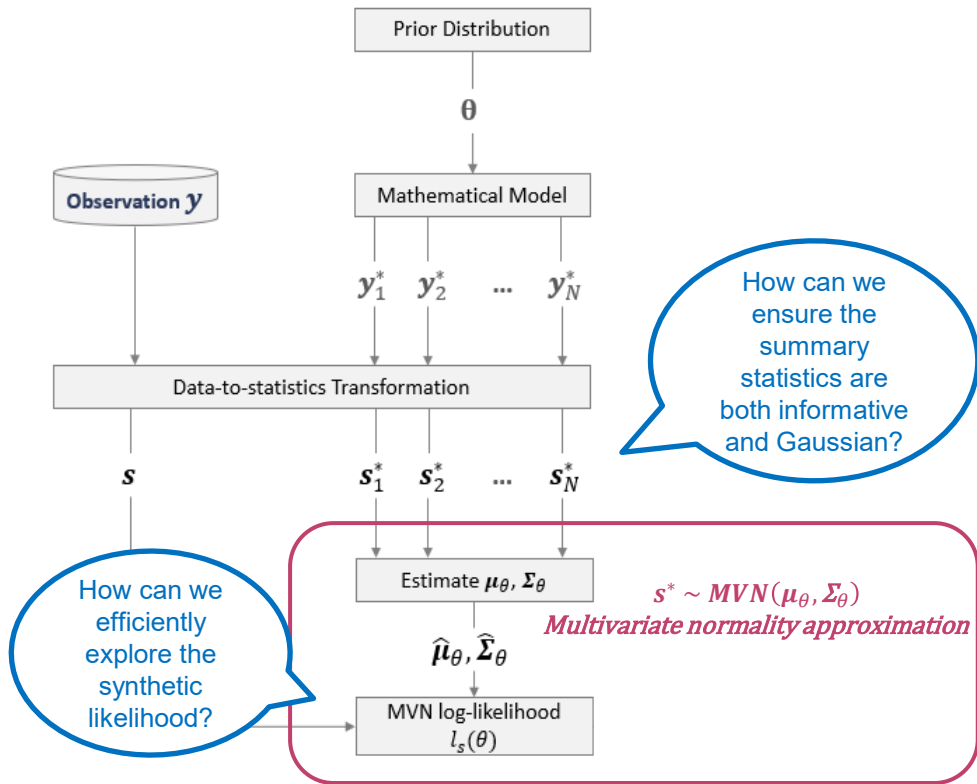
2.2 Synthetic Likelihood Method (Wood, 2010)

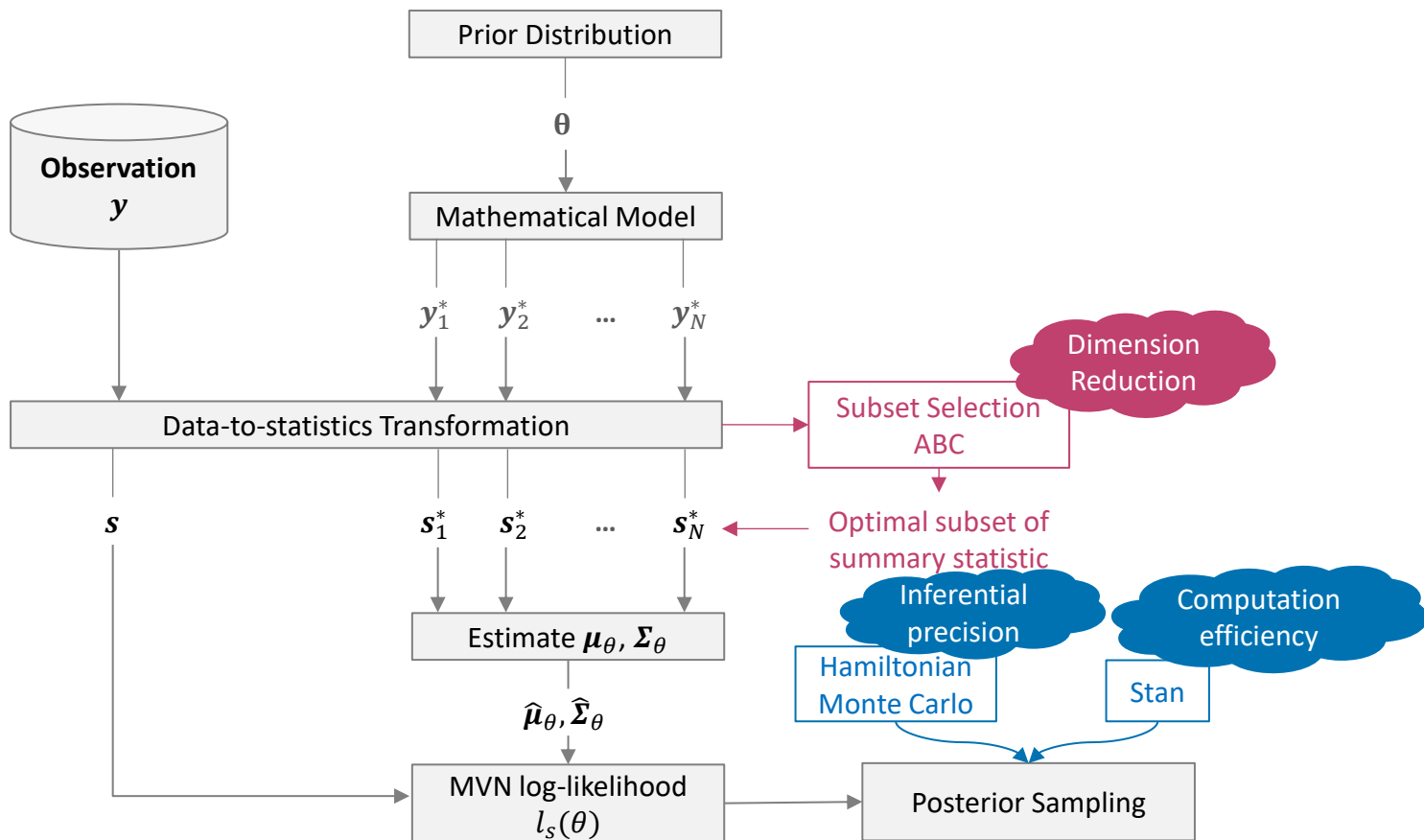


ABC Workflow

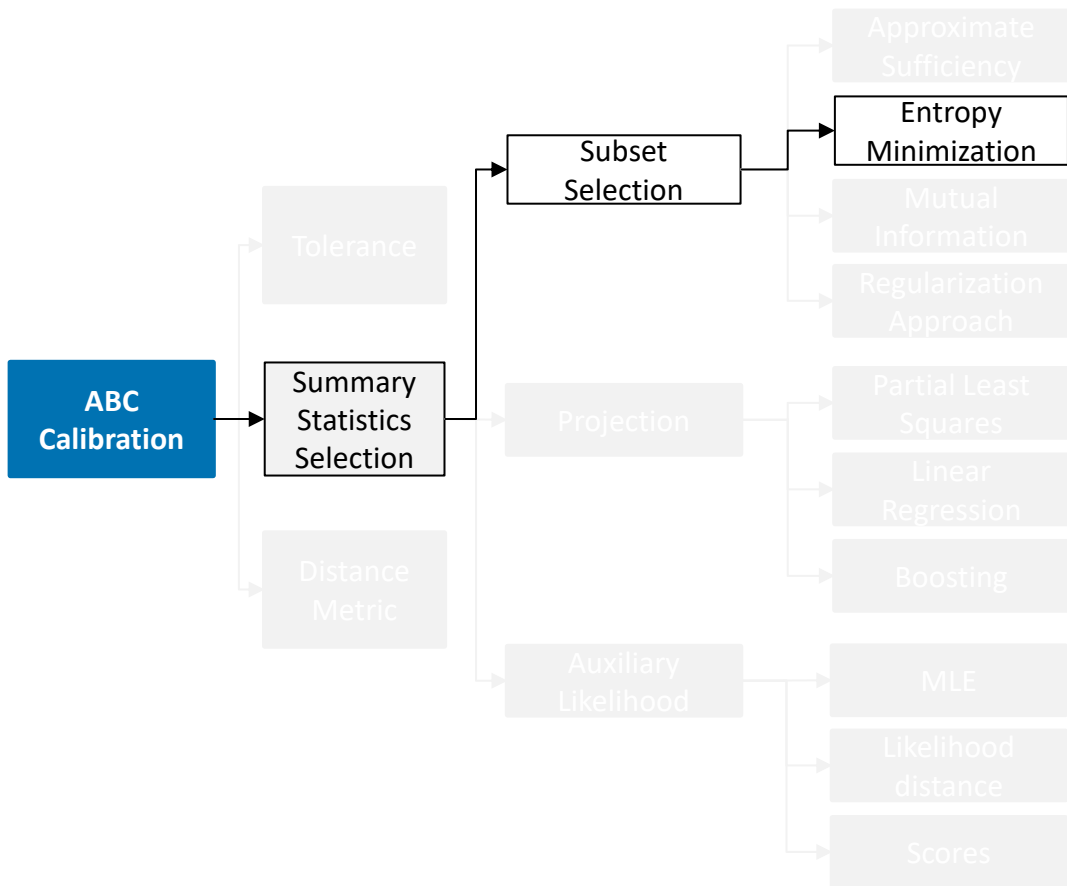


BSL Workflow





3.1 Key Innovation 1: ABC for Summary Statistics Selection



Entropy Minimisation

- Perform rejection-ABC and compute the k th nearest neighbour of entropy on the ABC posterior sample.

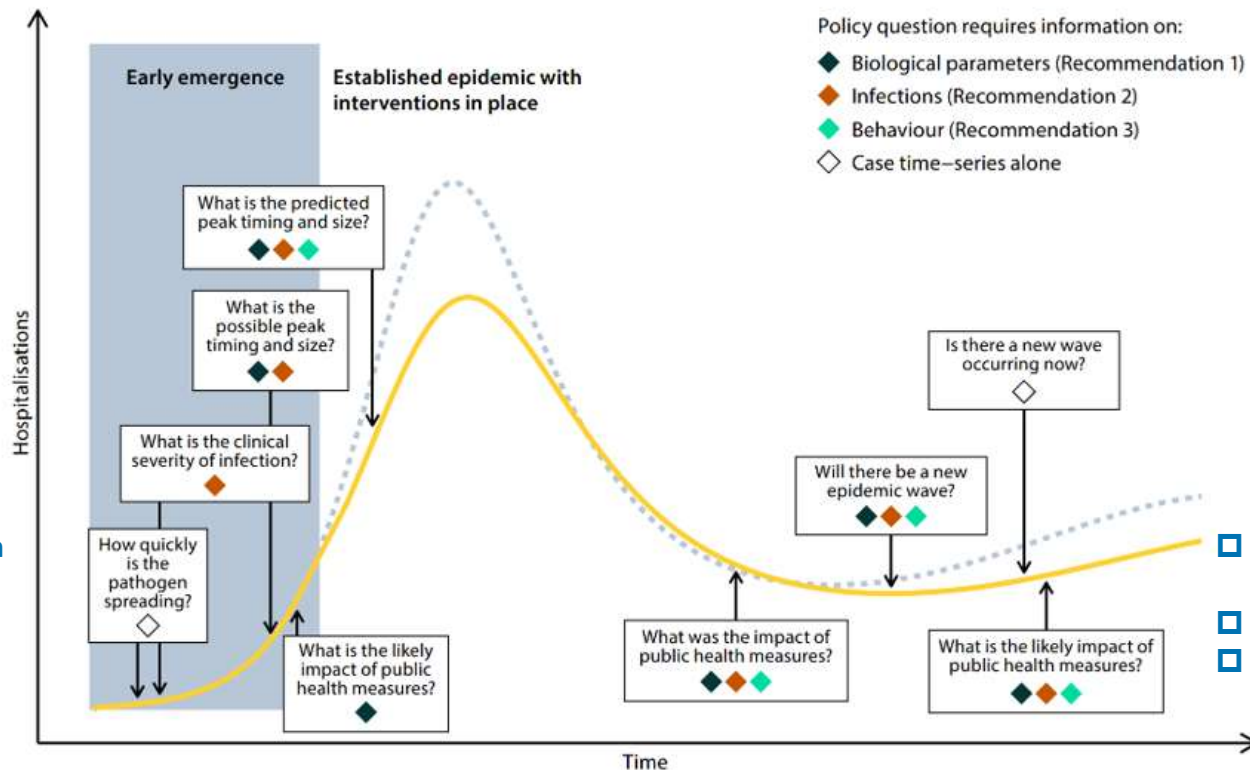
$$\hat{E} = \log \left[\frac{\pi^{p/2}}{\Gamma(\frac{p}{2} + 1)} \right] - \Psi(\kappa) + \log(n) + \frac{p}{n} \sum_{i=1}^n \log(R_{i,\kappa})$$

Why this method?

- balance of flexibility and interpretability—especially when data are limited during early outbreak

3.2 Key Innovation 2: Policy-informed Candidate Statistics

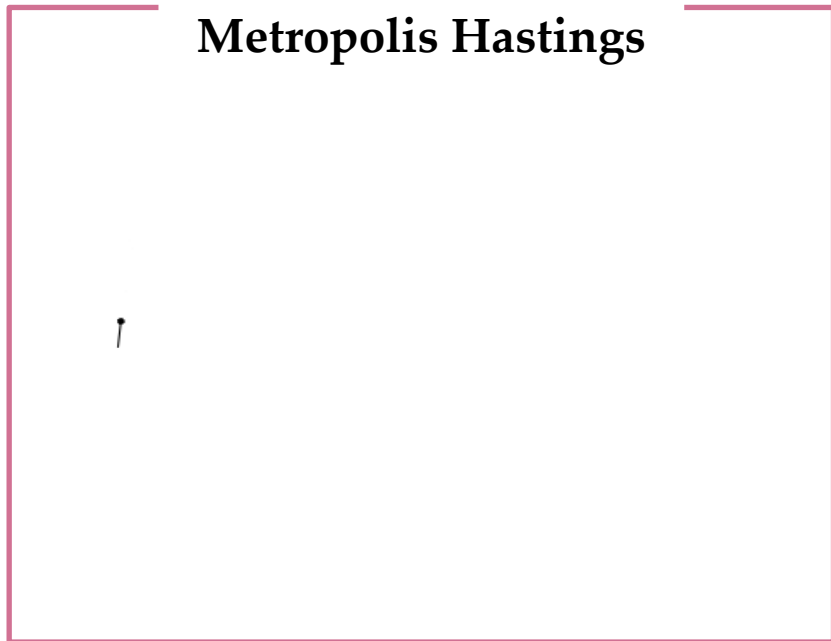
Figure 1: Exemplar policy questions and transmission-related surveillance needs



- Basic reproduction number (R_0)
- Growth rate
- Doubling time

- Effective reproduction number (R_e)
- Incidence rate ratios
- Hospitalization trends

Random Walk Metropolis Hastings

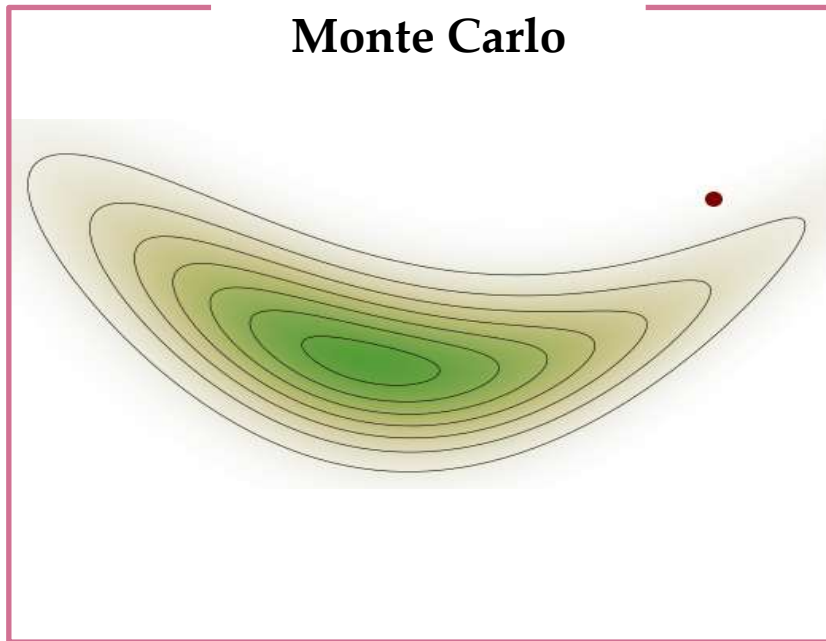


GIF Animation 1: Random walk Metropolis Hastings

Adapted from

https://bookdown.org/danbarch/psy_207_advanced_stats_I/MCMC-methods.html

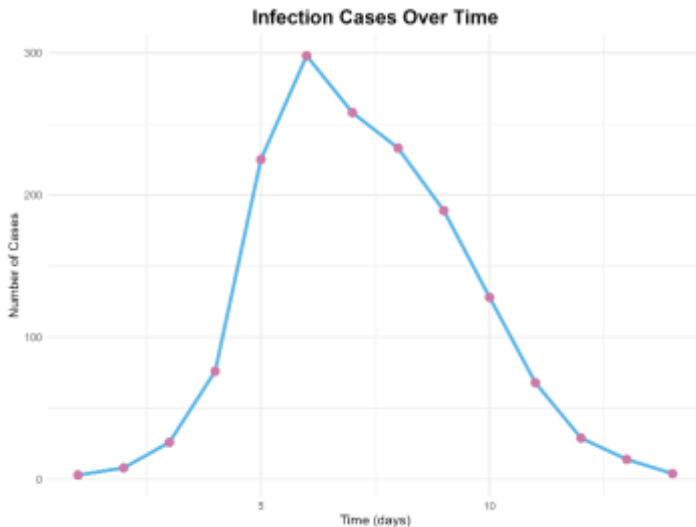
Hamiltonian Monte Carlo



GIF Animation 2: Hamiltonian Monte Carlo sampling

Adapted from Justinkunimune - Own work

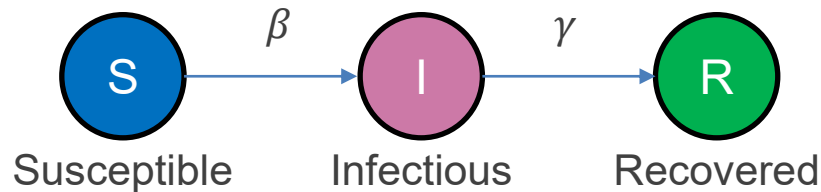
using: github.com/jkunimune/hamiltonian-mc, CC0



About data

- 1978 influenza A outbreak at a British boarding school.
- During the outbreak, 512 out of 763 students became ill.

Model



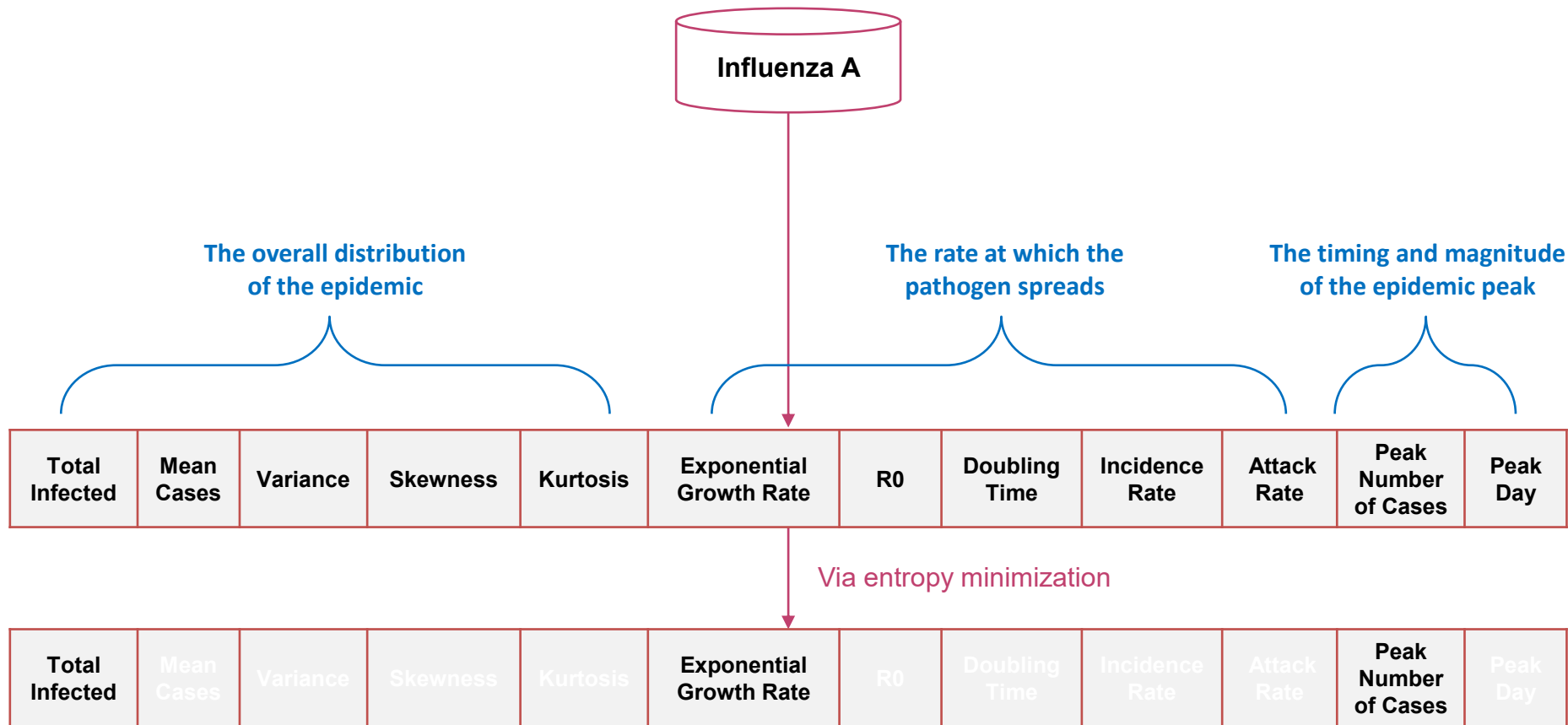
$$\frac{dS}{dt} = -\beta \frac{I(t)}{N} S(t)$$

$$\frac{dI}{dt} = \beta \frac{I(t)}{N} S(t) - \gamma I(t)$$

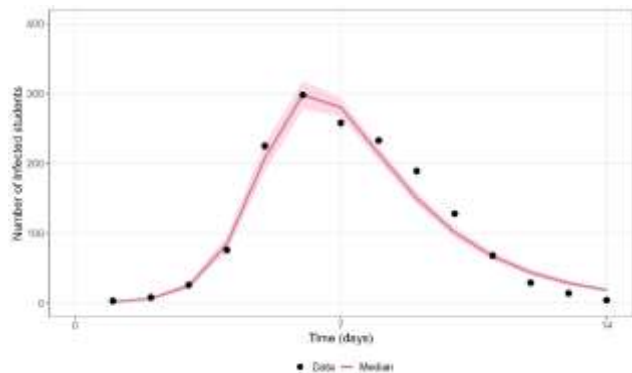
$$\frac{dR}{dt} = \gamma I(t)$$

$$Y_t \sim \text{Poisson}(\lambda_t)$$

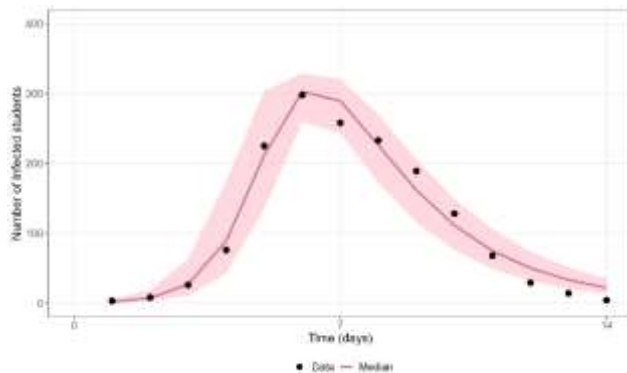
4.2 Application – Summary Statistics Selection using ABC



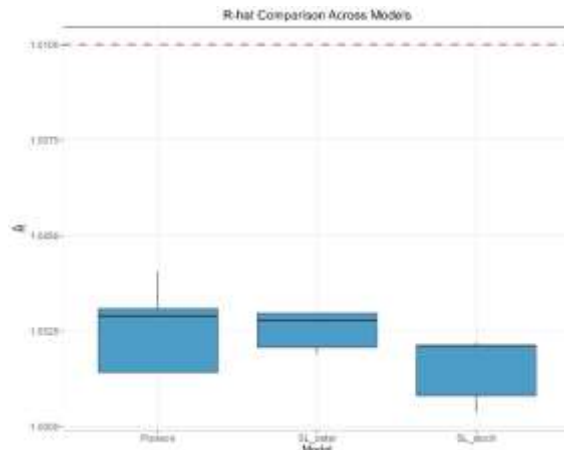
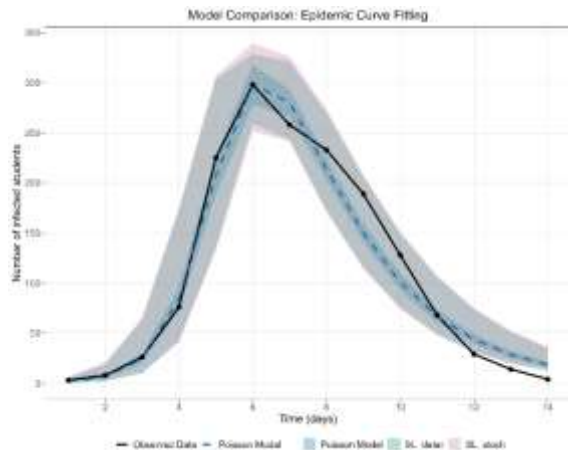
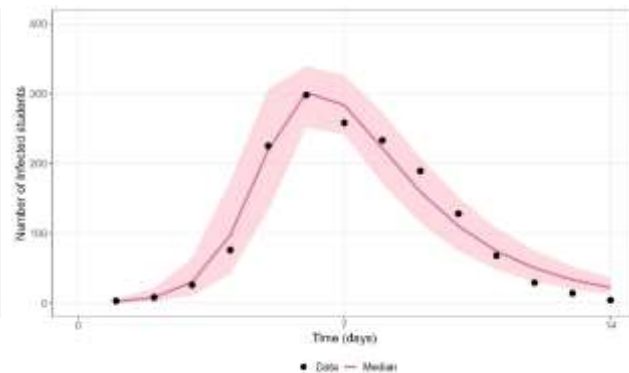
Poisson Model



SL Deterministic Model



SL Stochastic Model



Summary:

- Introduced a **hybrid inference framework** for parameter estimation in epidemiological models with intractable likelihoods.
- Addressed challenges in Bayesian inference for compartmental models, by integrating
 - **ABC-based entropy minimization** for summary statistics selection
 - **BSL** for flexible likelihood approximation
 - **HMC** for posterior sampling



Further Research:

- Current numerical studies only focused on **compartmental models**, can extend to **more complex models**.
- **Gaussian approximation** might be restrictive for extreme cases, can explore **alternative assumptions** to improve **robustness**.
- Apply to **large-scale, real-world datasets** to test scalability and performance.
- **Theoretical investigation** on method under varying degrees of model complexity and data quality.



THANK YOU.

- Beaumont, M. A., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, 162(4), 2025-2035.
- Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. arXiv preprint arXiv:1701.02434.
- Bjørnstad, O. N., Shea, K., Krzywinski, M., & Altman, N. (2020). The SEIRS model for infectious disease dynamics. *Nature methods*, 17(6), 557-559.
- Brauer, F. (2008). Compartmental models in epidemiology. *Mathematical epidemiology*, 19-79.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, 76.
- Fearnhead, P., & Prangle, D. (2012). Constructing summary statistics for approximate Bayesian computation: semi-automatic approximate Bayesian computation. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 74(3), 419-474.
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). Bayesian data analysis. Chapman and Hall/CRC.
- Li, X., Swallow, B., & Chadwick, F. J. (2024). A Novel Approximate Bayesian Inference Method for Compartmental Models in Epidemiology using Stan. *arXiv preprint arXiv:2408.03415*.
- Lopes, J. S., & Beaumont, M. A. (2010). ABC: a useful Bayesian tool for the analysis of population data. *Infection, Genetics and Evolution*, 10(6), 825-832.
- Minter, A., & Retkute, R. (2019). Approximate Bayesian Computation for infectious disease modelling. *Epidemics*, 29, 100368.
- Nunes, M. A., & Prangle, D. (2015). abctools: an R package for tuning approximate Bayesian computation analyses. *The R Journal*.
- Shearer, F. M., Edwards, L., Kirk, M., Eales, O., Golding, N., Hassall, J., ... & McCaw, J. M. (2024). Opportunities to strengthen respiratory virus surveillance systems in Australia: lessons learned from the COVID-19 response. *Communicable Diseases Intelligence*, 48.
- Toni, T., Welch, D., Strelkowa, N., Ipsen, A., & Stumpf, M. P. (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*, 6(31), 187-202.
- Wood, S. N. (2010). Statistical inference for noisy nonlinear ecological dynamic systems. *Nature*, 466(7310), 1102-1104.