Prediction and inference for epidemic models using likelihood-based methods

2nd New Zealand Workshop on Uncertainty Quantification and Inverse Problems

David Wu[1]
with Helen Petousis-Harris[2], Janine Paynter[2], Vinod Suresh[1,3], and Oliver J. Maclaren[1]

November 2021

[1]Dept. Engineering Science, Engineering, University of Auckland
[2]Dept. of General Practice and Primary Health Care, FMHS, University of Auckland
[3]Auckland Bioengineering Institute, University of Auckland

Outline

1. Epidemic Models (brief)

Outline

1. Epidemic Models (brief)
2. Circumventing Forward Model Simulations

Outline

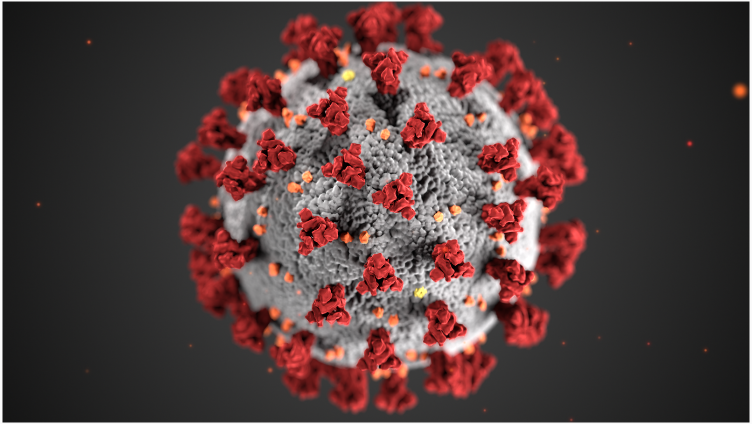
1. Epidemic Models (brief)
2. Circumventing Forward Model Simulations
3. Likelihood-based UQ techniques

Outline

1. Epidemic Models (brief)
2. Circumventing Forward Model Simulations
3. Likelihood-based UQ techniques
4. Case Study

Epidemic Modelling

Epidemic modelling is a *severely neglected* field of applied mathematics, due to its *low economic and societal impacts*.



https://www.fda.gov/files/Coronavirus_3D_illustration_by_CDC_1600x900.png

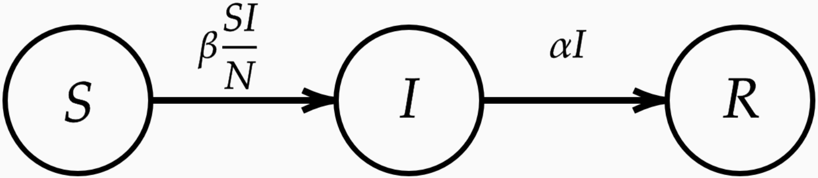
Various ways to model epidemic outbreaks:

- Phenomenological Models
- Deterministic ODEs
- Stochastic SDEs
- Branching Processes
- Agent-based Simulations

Various ways to model epidemic outbreaks:

- Phenomenological Models
- Deterministic ODEs
- Stochastic SDEs
- Branching Processes
- Agent-based Simulations

SIR Model



$$\frac{dS}{dt} = -\beta\frac{SI}{N}$$
$$\frac{dI}{dt} = \beta\frac{SI}{N} - \alpha I$$
$$\frac{dR}{dt} = \alpha I$$

Deterministic ODE Models

In general:

$$y = g(x) + \epsilon$$
$$\frac{dx}{dt} = f(x; \theta)$$

Two sources of information:

- the data, y
- the differential equation model, f

How do we use both?

Traditional Methods

Integrate the differential equation for proposal θ

$$\theta \xmapsto{\frac{dx}{dt} = f(x;\theta)} x$$

Traditional Methods

Integrate the differential equation for proposal θ

$$\theta \xmapsto{\frac{dx}{dt}=f(x;\theta)} x$$

Form a likelihood function based on the observation function

$$x(\theta) \xmapsto{y=g(x)+\epsilon} p(y|\theta)$$

Traditional Methods

Integrate the differential equation for proposal θ

$$\theta \xmapsto{\frac{dx}{dt}=f(x;\theta)} x$$

Form a likelihood function based on the observation function

$$x(\theta) \xmapsto{y=g(x)+\epsilon} p(y|\theta)$$

Then apply your chosen statistical philosophy to this likelihood

$$p(y|\theta) \xmapsto{\text{magic}} \text{results}$$

Gradient Matching/2-stage NLS

If we had some estimate of x, we could find an estimate of θ pretty easily, without inegration:

$$\frac{d}{dt}x_{\text{est}} = f(x_{\text{est}}; \theta)$$

Gradient Matching/2-stage NLS

If [...]thout

in[...]

We[...]

An example of G

We can do this, for example, using smoothing splines.

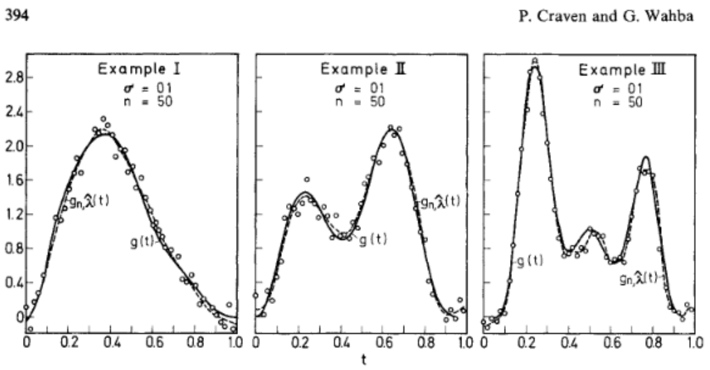


Fig. 2. Examples I, II, and III, $g, g_{n,\hat{\lambda}}$, and the data

from Craven and Wahba (1979), Smoothing Noisy Data with Spline Functions, Numer. Math, 31,377-403

Gradient Matching/2-stage NLS

If we had some estimate of x, we could find an estimate of θ pretty easily, without inegration:

$$\frac{d}{dt} x_{\text{est}} = f(x_{\text{est}}; \theta)$$

We might be able to approximate x from y

$$y = g(x) + \epsilon$$
$$\downarrow$$
$$x_{\text{est}} = G(y, \epsilon)$$

Gradient Matching/2-stage NLS

If we had some estimate of x, we could find an estimate of θ pretty easily, without inegration:

$$\frac{d}{dt}x_{\text{est}} = f(x_{\text{est}}; \theta)$$

We might be able to approximate x from y

$$y = g(x) + \epsilon$$
$$\downarrow$$
$$x_{\text{est}} = G(y, \epsilon)$$

If we can't find an appropriate G (e.g. when the x is only partially observed), then we cannot use this method.

Idea of Ramsay et al (2007):

We can get around that restriction, if we estimate both x and θ simultaneously.

The differential equation "makes up for" the data we are missing.

Generalised Profiling

Idea of Ramsay et al (2007):

We can get around that restriction, if we estimate both x and θ simultaneously.

The differential equation "makes up for" the data we are missing.

Can project x onto some smooth basis ($x = \Phi c$), to reduce the dimension of the problem.

Generalised Profiling

Two-level process:

Inner: for a given θ, compute a optimal smooth $\hat{c}$

$$\hat{c}(\theta) = \min_{c|\theta} \left\{ \|y - g(\Phi c)\|^2 + \lambda \left\| \frac{d}{dt}(\Phi c) - f(\Phi c, \theta) \right\|^2 \right\}$$

Outer: Optimise over θ, using the inner to eliminate nuisance parameters (c)

$$\min_{\theta} \left\{ \|y - g(\Phi \hat{c}(\theta))\|^2 \right\}$$

Aside: Profile Likelihood

Given a likelihood:

$$\mathcal{L}(x_1, x_2)$$

Marginal likelihood:

$$\mathcal{L}_{X_1}(x_1) = \int \mathcal{L}(x_1, x_2) dx_2$$

Profile likelihood:

$$\mathcal{L}_{P,X_1}(x_1) = \max_{x_2} \mathcal{L}(x_1, x_2)$$

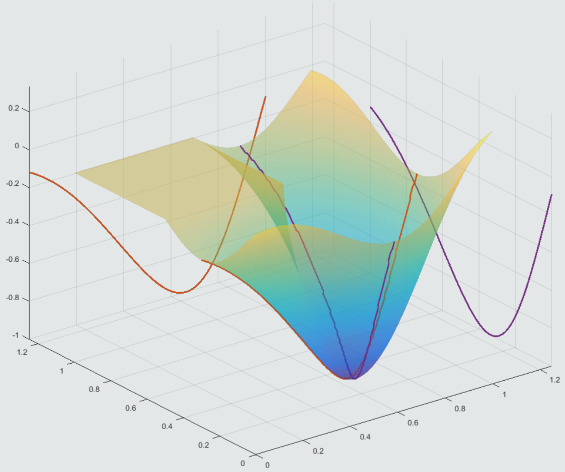
Aside: Profile Likelihood

Gi

Ma

Pr

Profiling



Aside: Profile Likelihood

Given a likelihood:

$$\mathcal{L}(x_1, x_2)$$

Marginal likelihood:

$$\mathcal{L}_{X_1}(x_1) = \int \mathcal{L}(x_1, x_2) dx_2$$

Profile likelihood:

$$\mathcal{L}_{P,X_1}(x_1) = \max_{x_2} \mathcal{L}(x_1, x_2)$$

The profile likelihood is more natural under a frequentist lens, since the maximum likelihood estimator is left unchanged:

$$\max_{x_1, x_2} \mathcal{L}(x_1, x_2) = \max_{x_1} \mathcal{L}_{P,X_1}(x_1)$$

GP Likelihood we need a better name

The generalised profiling method feels a lot like taking the profile likelihood of the inner objective.

Using ideas from mixed estimation (Theil & Golberger 1961), we can motivate the inner objective as a likelihood.

$$\mathcal{L}(\theta, c) = \left\| L(y - g(x)) \right\|^2 + \left\| W \left(\frac{d}{dt}(\Phi c) - f(\Phi c; \theta) \right) \right\|^2$$

*Can also add regularisation to this in a justified manner

GP Likelihood

$$\mathcal{L}(\theta, c) = \left\|L(y - g(\Phi c))\right\|^2 + \left\|W\left(\frac{d}{dt}(\Phi c) - f(\Phi c; \theta)\right)\right\|^2$$

Also estimating covariances $(LL^T)^{-1}, (WW^T)^{-1}$.

Can't do this all at once, so choose to use an iterative method (iteratively reweighted least squares).

1. Estimate (x, c) for a fixed (L, W).
2. Estimate (L, W) for a fixed (x, c).
3. Return to Step 1.

UQ (Likelihood-based)

We can optimise the (log-)likelihood to get a point estimate (the MLE).

We really also want confidence sets to go with this.

Profile Likelihood

Claim: The likelihood contains all the information about the uncertainty in the problem.

Profile Likelihood

Claim: The likelihood contains all the information about the uncertainty in the problem.

All we need to do is to understand the behaviour of the likelihood with respect to quantities of interest.

Profile likelihood is a good way of doing this.

Also gives us information about identifiability.

Bootstrapping

If we have model $y = X(\theta) + \epsilon$, and we estimate MLE as $\hat{\theta}$ and the error as e, we can get an idea of the uncertainty in the estimate by sampling from fits of θ_{BS} for realisations of e.

$$X(\hat{\theta}) + e = X(\theta_{BS})$$

Bootstrapping

If we have model $y = X(\theta) + \epsilon$, and we estimate MLE as $\hat{\theta}$ and the error as e, we can get an idea of the uncertainty in the estimate by sampling from fits of θ_{BS} for realisations of e.

$$X(\hat{\theta}) + e = X(\theta_{BS})$$

We can also use the randomised maximum likelihood interpretation of this, and sample from the perturbed data instead:

$$y + e = X(\theta_{RML})$$

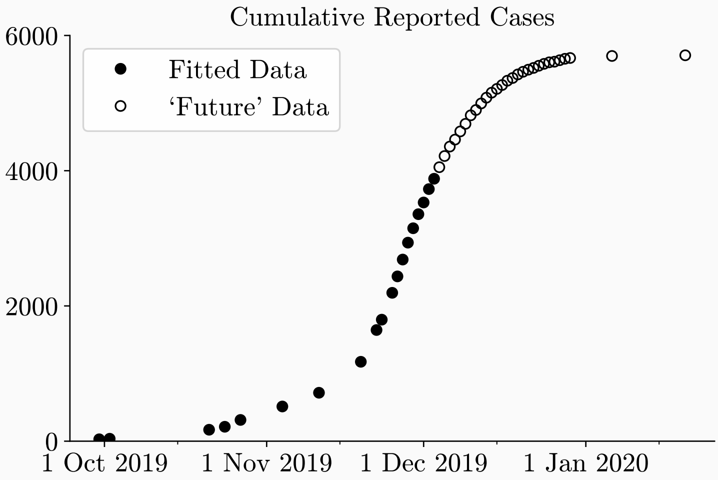
Case Study

Case Study: 2019 Samoan Measles Outbreak

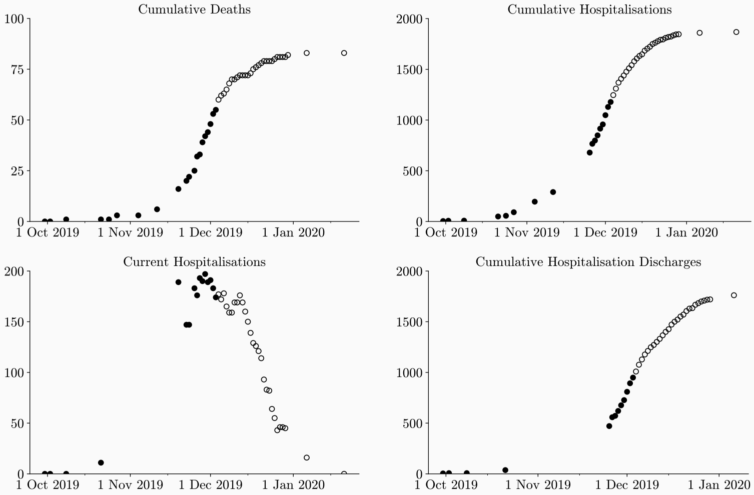
Measles outbreak in the small South Pacific island nation of Samoa

- September 2019 – January 2020
- Over 5700 cases (pop. 200 000)
- 83 deaths
- Anomalously low vaccination coverage in infants
    - 40% MCV1
    - 28% MCV2

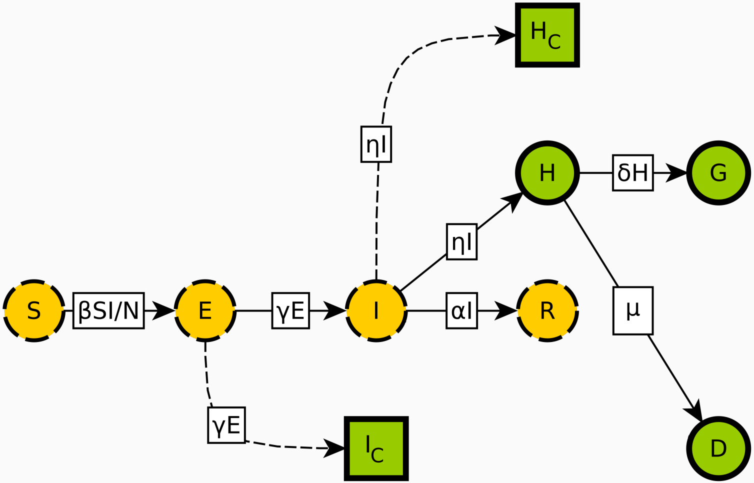
Case Study: 2019 Samoan Measles Outbreak



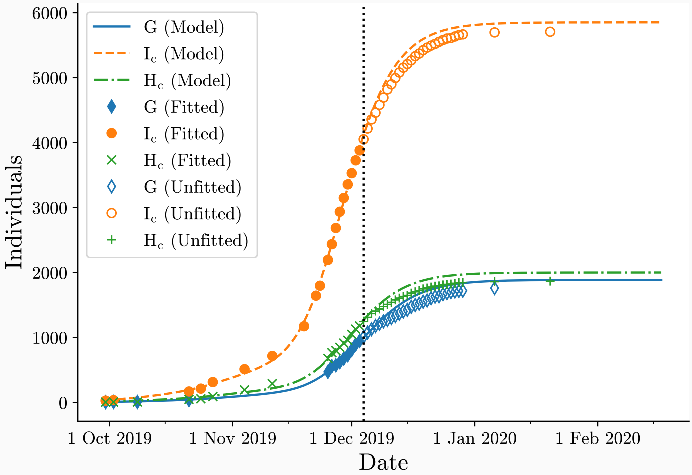
Case Study: 2019 Samoa Measles Outbreak



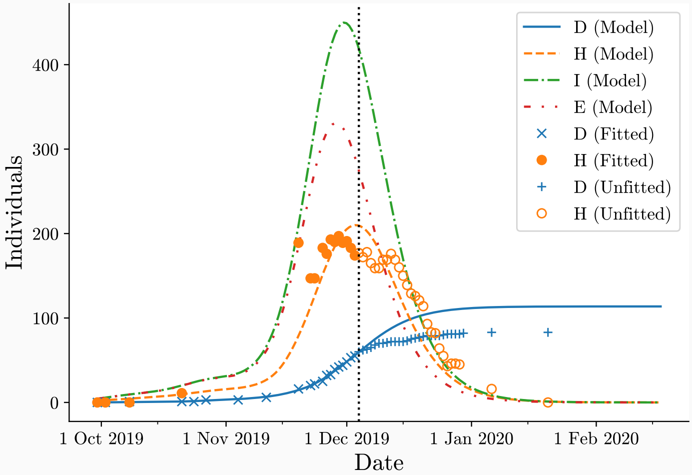
Case Study: Model



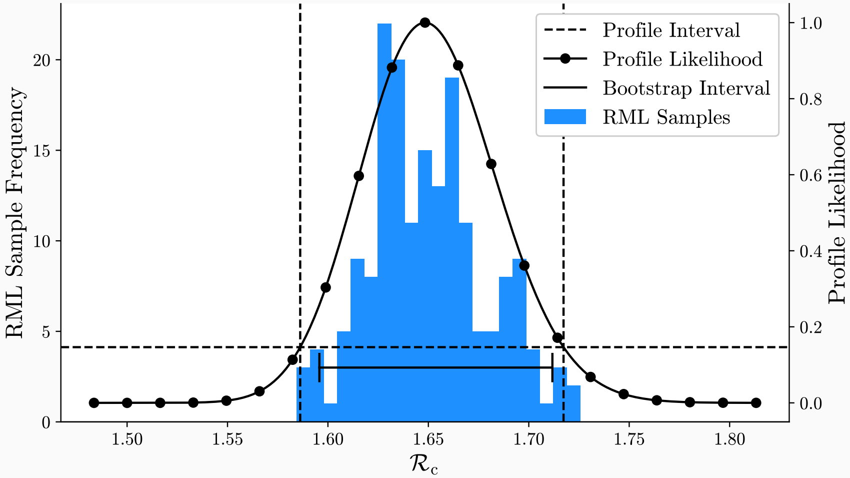
Case Study: Results (MLE)



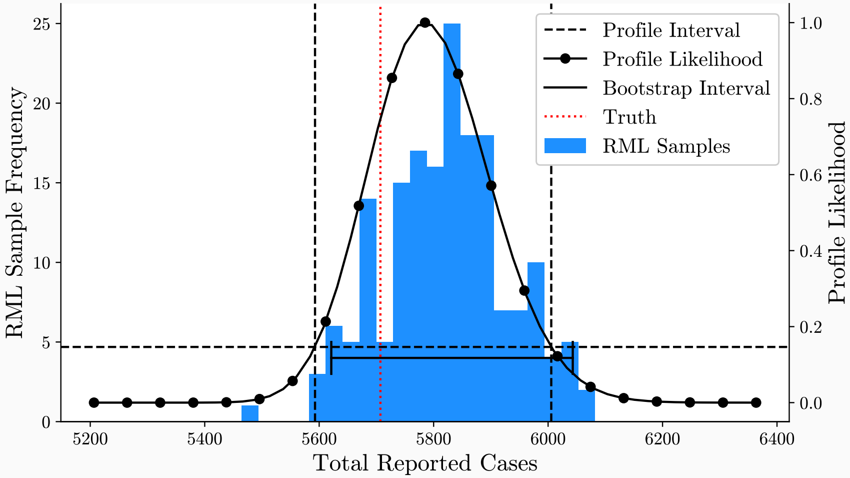
Case Study: Results (MLE)



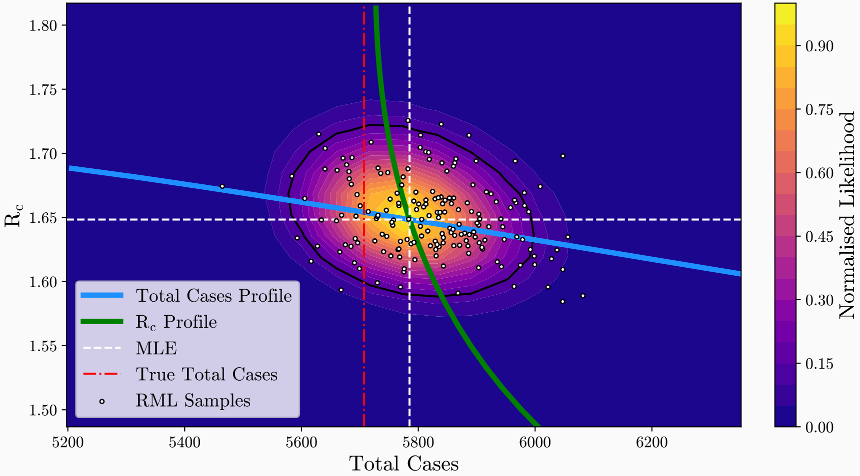
Case Study: Results (UQ)



Case Study: Results (UQ)



Case Study: Results (UQ)



Closing Remarks

- Can avoid the integration/forward problem solve step by estimating the state and the parameters simultaneously

Closing Remarks

- Can avoid the integration/forward problem solve step by estimating the state and the parameters simultaneously
- There exist frequentist methods (profile likelihood, bootstrapping) to perform UQ on likelihood formulations

Closing Remarks

- Can avoid the integration/forward problem solve step by estimating the state and the parameters simultaneously
- There exist frequentist methods (profile likelihood, bootstrapping) to perform UQ on likelihood formulations
- This all seems to work on a case study of real outbreak data

Closing Remarks

- Can avoid the integration/forward problem solve step by estimating the state and the parameters simultaneously
- There exist frequentist methods (profile likelihood, bootstrapping) to perform UQ on likelihood formulations
- This all seems to work on a case study of real outbreak data
- Though there is some tuning that needs to be done by hand

Thanks for coming!

Questions?

arXiv: 2103.16058

References

Bardsley, Johnathan M et al. (2014). “Randomize-Then-Optimize: A Method for Sampling from Posterior Distributions in Nonlinear Inverse Problems”. In: *SIAM Journal on Scientific Computing* 36.4, A1895–A1910. doi: 10.1137/140964023.

Efron, Bradley and R.J. Tibshirani (1994). *An Introduction to the Bootstrap*. en. Monographs on Statistics and Applied Probability 57. Chapman and Hall/CRC. isbn: 978-0-412-04231-7.

Hooker, Giles et al. (July 2011). “Parameterizing state-space models for infectious disease dynamics by generalized profiling: measles in Ontario.”. In: *Journal of the Royal Society, Interface* 8.60, pp. 961–74. issn: 1742-5662. doi: 10.1098/rsif.2010.0412.

Ramsay, J. O. et al. (Nov. 2007). “Parameter estimation for differential equations: a generalized smoothing approach”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 69.5, pp. 741–796. issn: 13697412. doi: 10.1111/j.1467-9868.2007.00610.x.

Ramsay, James and Giles Hooker (2017). *Dynamic Data Analysis*. Springer Series in Statistics. New York, NY: Springer New York. isbn: 978-1-4939-7188-6. doi: 10.1007/978-1-4939-7190-9.

Raue, A. et al. (Aug. 2009). “Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood”. In: *Bioinformatics* 25.15, pp. 1923–1929. issn: 1460-2059. doi: 10.1093/bioinformatics/btp358.