Fitting simple epidemic models: an approximate approach

DES Postgraduate Presentation Competition 2020

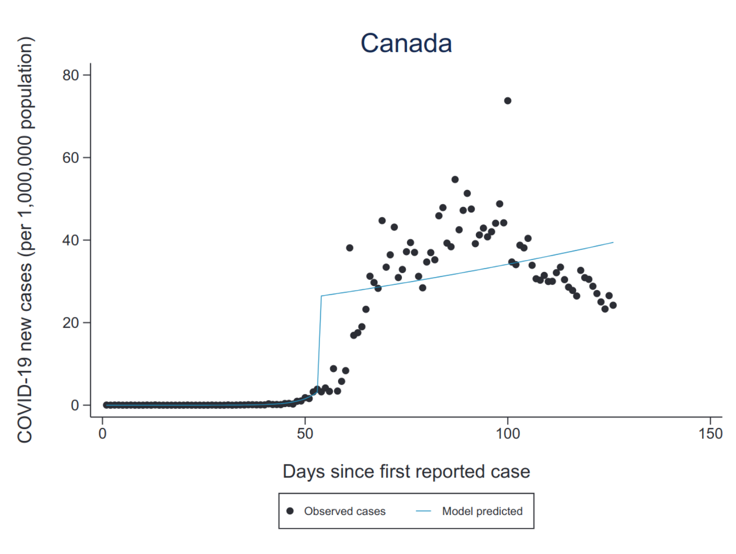
David Wu

December 2020

Department of Engineering Science
University of Auckland

Predicting epidemics is hard.

Prediction is hard



Islam et. al. (2020) *Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries.* BMJ 2020;371:m3939

Prediction is hard

Predicting an epidemic trajectory is difficult

Claus O. Wilke and Carl T. Bergstrom

PNAS November 17, 2020 117 (46) 28549-28551; first published November 3, 2020; https://doi.org/10.1073/pnas.2020200117

Prediction is hard

Predicting an epidemic trajectory is difficult

Claus O. Wilke and Carl T. Bergstrom

1073

The turning point and end of an expanding epidemic cannot be precisely forecast

Mario Castro, Saúl Ares, José A. Cuesta, and Susanna Manrubia

PNAS October 20, 2020 117 (42) 26190-26196; first published October 1, 2020; https://doi.org/10.1073/pnas.2007868117

Prediction is hard

Predicting an epidemic trajectory is difficult

Claus O. Wilke and Carl T. Bergstrom

1073

The turning point and end of an expanding epid

Mario

PNAS Octob
/pnas.2007

KeAi
CHINESE ROOTS
GLOBAL IMPACT

Infectious Disease Modelling
Volume 5, 2020, Pages 271-281

Infectious Disease Modelling

Why is it difficult to accurately predict the COVID-19 epidemic?

Weston C. Roda [a], Marie B. Varughese [b], Donglin Han [a], Michael Y. Li [a]

Basic Problem Statement

Fit the model

$$x = F(t; \theta)$$

to some data

$$y = g(t, x; \theta)$$

Basic Problem Statement

Model is typically
not explicit

Fit the model

$$\cancel{x = F(t;\theta)}$$

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

to some data

$$y = g(t, x; \theta)$$

Basic Problem Statement

Model is typically
not explicit

Fit the model

Highly nonlinear

$$\cancel{x = F(t; \theta)}$$

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

to some data

$$y = g(t, x; \theta)$$

Basic Problem Statement

Model is typically
not explicit

Fit the model

Highly nonlinear

$$\cancel{x = F(t;\theta)}$$

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

to some data

$$y = g(t, x; \theta)$$

Data is partially
observed

Basic Problem Statement

Fit the model

Model is typically *not* explicit

Highly nonlinear

$$\cancel{x = F(t; \theta)}$$

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

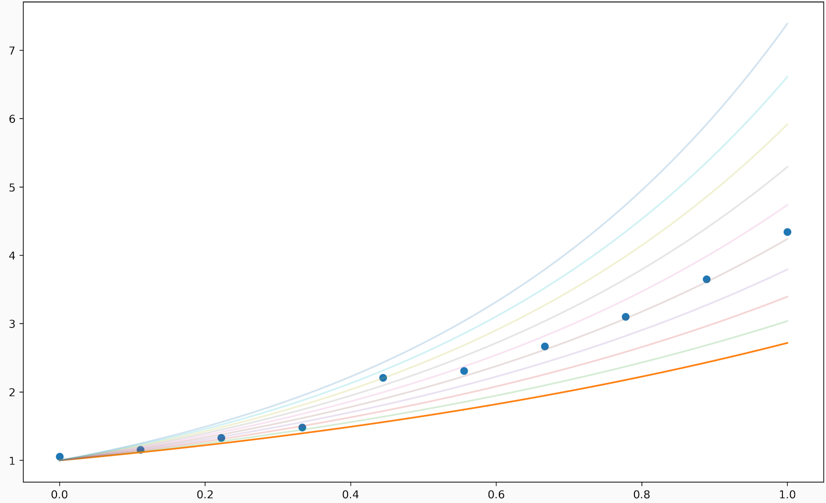
to some data

$$y = g(t, x; \theta)$$

Data is partially observed

Model is probably not (quite) correct

Standard Approach



Standard Approach

Initialise with some guess $x_0^{(0)}, \theta^{(0)}$.
At each iteration i until convergence,

1. Integrate the model f from $x_0^{(i)}$ with parameters $\theta^{(i)}$ to produce solution $x^{(i)}$
2. Compute residuals $r^{(i)} = y - g(x^{(i)}, \theta^{(i)})$
3. Step to reach a next iterate $x_0^{(i)}, \theta^{(i)}$
    - Typically a Newton step

A Useful Fitting Method

1. Solves relatively quickly
2. Allows for model misspecification
3. Infers (or otherwise deals with) non-observed states

We can achieve these goals by:

- using collocation methods to avoid integration, and
- relaxing the condition of matching the model.

Generalised Profiling

$$\min_{x,\theta} \quad \underbrace{\|L_d\,(y - g(x))\|^2}_{\substack{\text{state estimation}\\ \text{data misfit}}} + \underbrace{\left\|L_m \left(\frac{dx}{dt} - f(x;\theta)\right)\right\|^2}_{\substack{\text{parameter estimation}\\ \text{model misfit}}}$$

where

$$\begin{aligned} L_d^T L_d &= \Gamma_d^{-1} && \left.\right\} \text{Inverse covariance of observation error} \\ L_m^T L_m &= \frac{1}{\Delta t}\Gamma_m^{-1} && \left.\right\} \text{Inverse covariance of process error} \end{aligned}$$

and x is projected on a spline basis

$$x(t) = \Phi(t)c$$

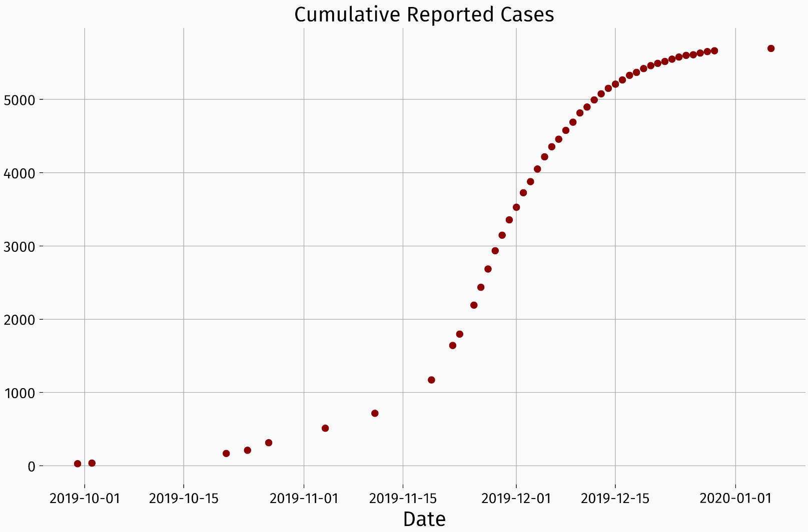
Ramsay et al. (2007) *Parameter estimation for differential equations: a generalized smoothing approach*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69: 741-796.

Case Study: Samoan Measles Outbreak

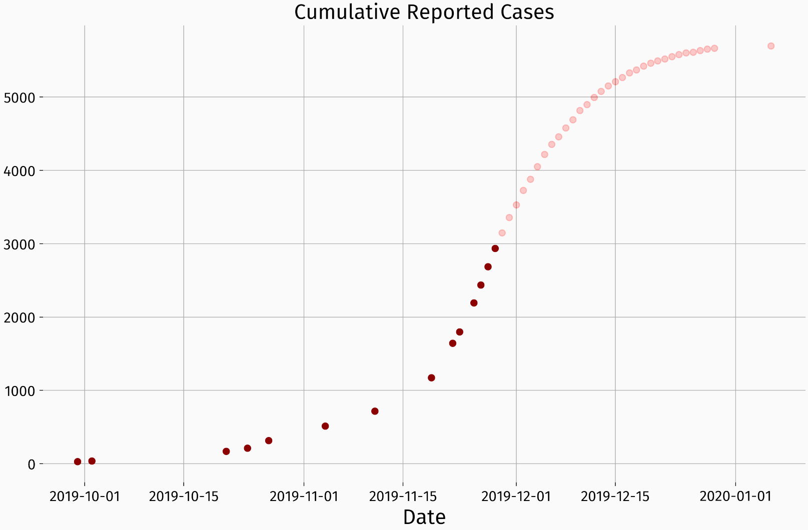
Applied these concepts to a compartmental model for predicting the Samoan measles outbreak of 2019.

- Lasted over 3 months (Sep 2019 – Jan 2020)
- Over 5700 cases (pop. 200,000)
- 83 deaths
- Anomalously low vaccination at time of outbreak
    - 40% first-dose MCV
    - 28% second-dose MCV

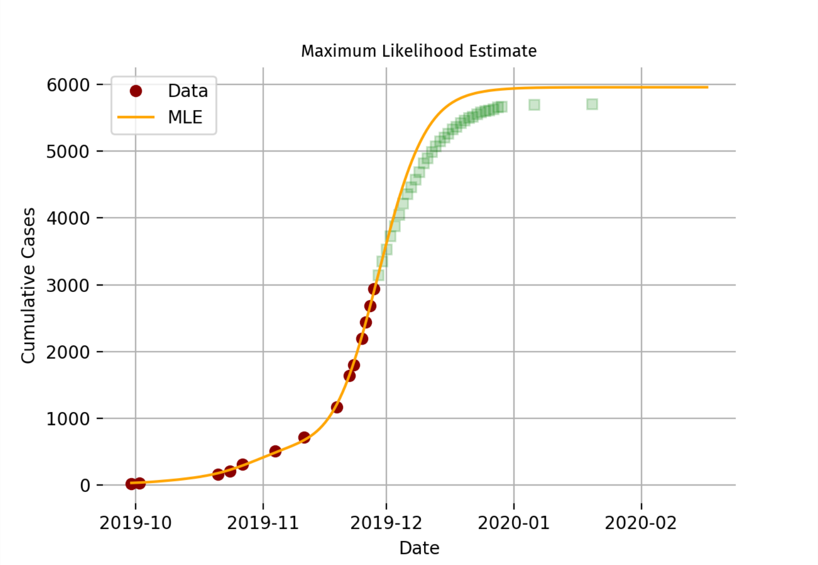
Case Study: Samoan Measles Outbreak



Case Study: Samoan Measles Outbreak



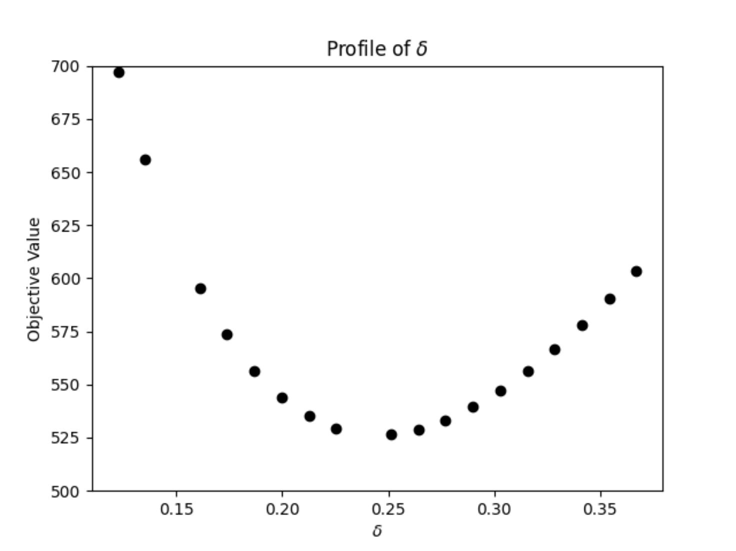
Case Study: Samoan Measles Outbreak



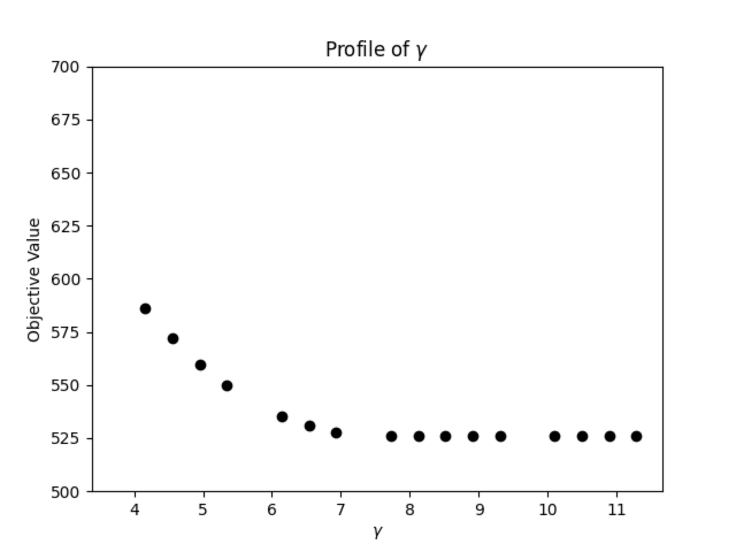
We can profile the **sensitivity** of the objective function **with respect to the parameters** of the model **by gridding** over them, **and re-solving**, letting all other parameters be free.

If other parameters can compensate for the change in the parameter of interest, we'll get a **flat profile**, signalling *nonidentifiability*.

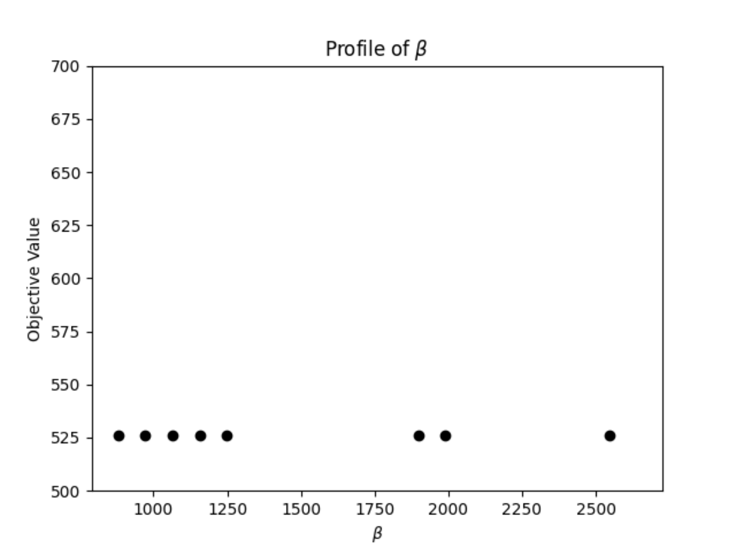
Parameter Uncertainty



Parameter Uncertainty

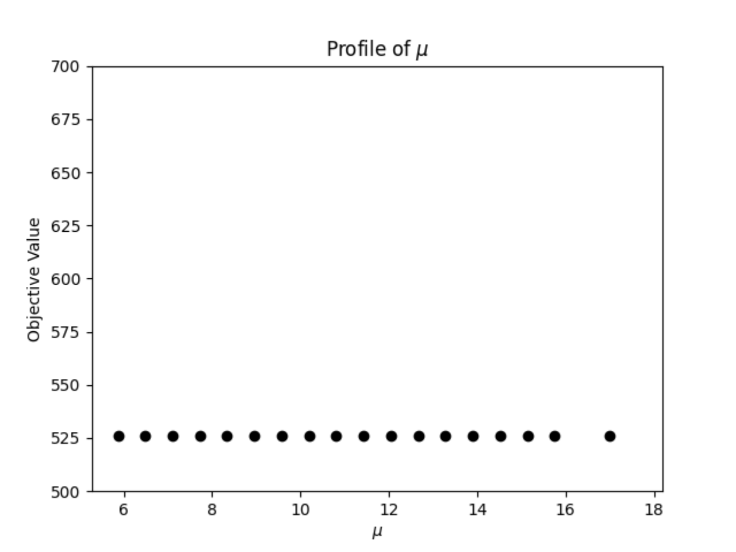


Parameter Uncertainty



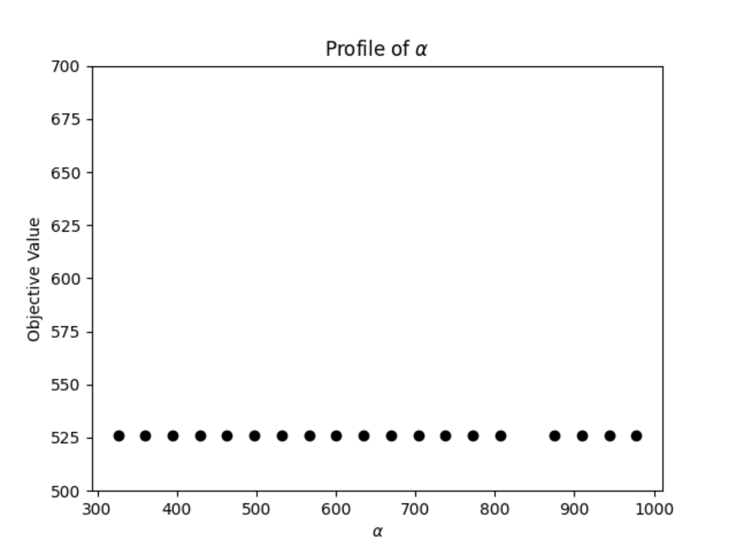
$$R_0 = \beta/(\mu + \alpha + \eta)$$

Parameter Uncertainty



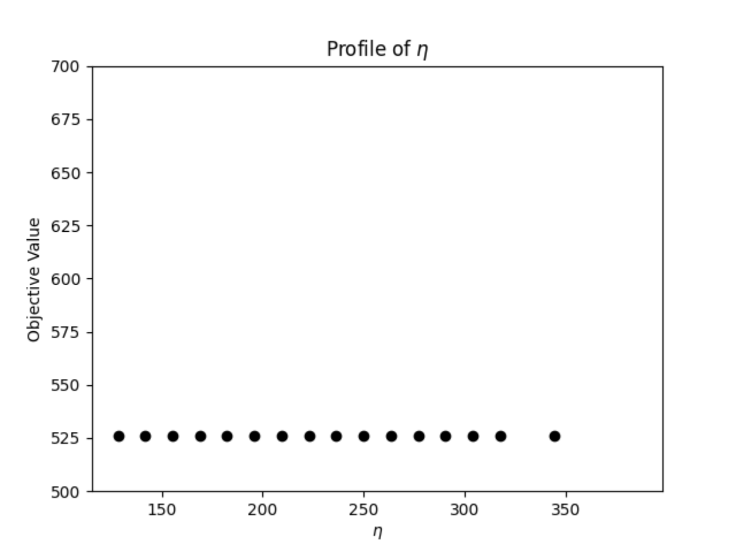
$R_0 = \beta/(\mu + \alpha + \eta)$

Parameter Uncertainty



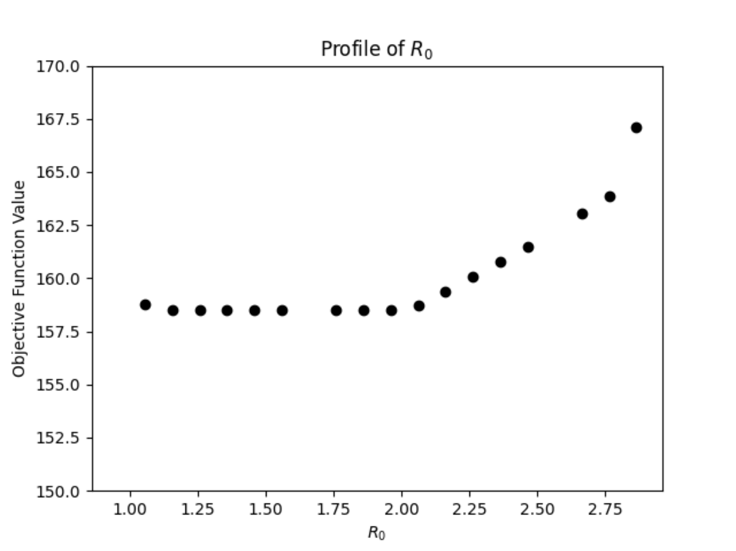
$$R_0 = \beta/(\mu + \alpha + \eta)$$

Parameter Uncertainty



$$R_0 = \beta/(\mu + \alpha + \eta)$$

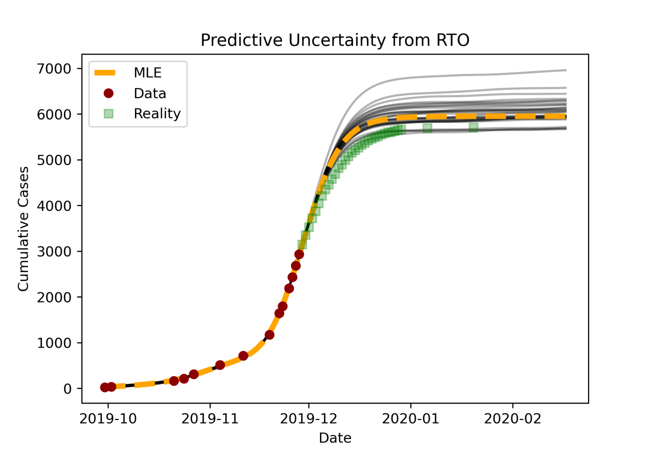
Parameter Uncertainty



$$R_0 = \beta/(\mu + \alpha + \eta)$$

Predictive Uncertainty

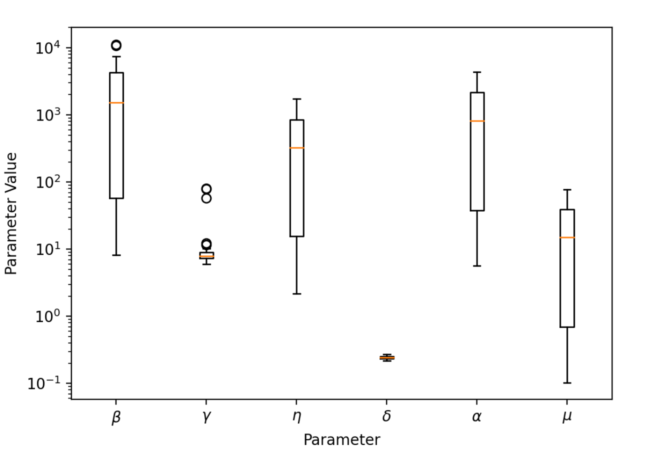
We derive predictive uncertainty under the Randomise-Then-Optimise framework (RTO). This involves perturbing the data and re-solving about the MLE.



Bardsley, J. M., Solonen, A., Haario, H., & Laine, M. (2014). *Randomize-Then-Optimize: A Method for Sampling from Posterior Distributions in Nonlinear Inverse Problems.* SIAM Journal on Scientific Computing, 36(4), A1895–A1910.

Parameter Uncertainty from RTO

We can also use the samples from RTO to infer parameter sensitivity.



Ionides, E. L., Breto, C., Park, J., Smith, R. A., & King, A. A. (2017). *Monte Carlo profile confidence intervals for dynamic systems.* Journal of the Royal Society, Interface, 14(132).

Concluding Remarks

- Framework for inference that does not require integration
- With parameter and predictive uncertainty
- Works for a case study with sparse data

Concluding Remarks

- Framework for inference that does not require integration
- With parameter and predictive uncertainty
- Works for a case study with sparse data
- Still a need for regularisation
- Still problems with convergence

Thanks

Oliver Maclaren

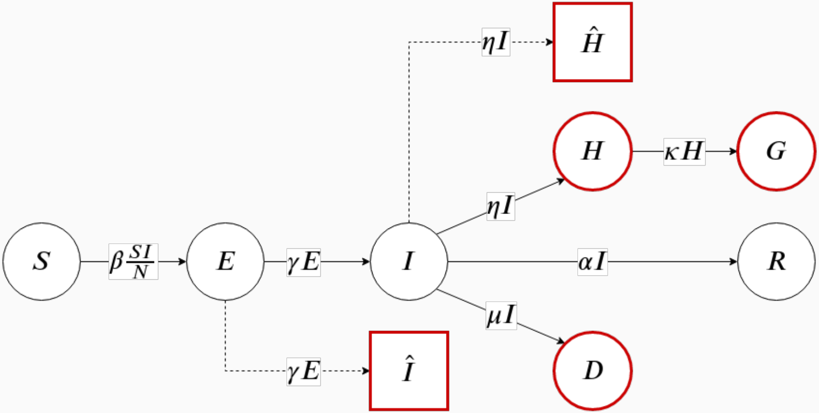
Vinod Suresh

Helen Petousis-Harris

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

samoaobserver

Samoan Measles Model



Model (Equation Form)

$$\begin{aligned}
\frac{dS}{dt} &= -\beta \frac{SI}{N} \\
\frac{dE}{dt} &= \beta \frac{SI}{N} - \gamma E \\
\frac{dI}{dt} &= \gamma E - \alpha I - \mu I \\
\frac{dR}{dt} &= (1-\eta)\alpha I
\end{aligned}$$

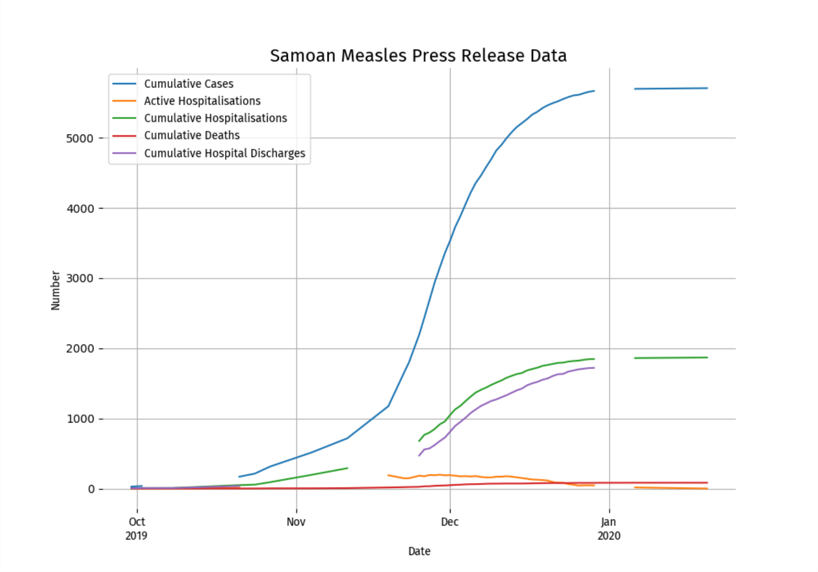
$$\begin{aligned}
\frac{dH}{dt} &= \eta \alpha I - \kappa H \\
\frac{dG}{dt} &= \kappa H \\
\frac{dD}{dt} &= \mu I \\
\frac{d\hat{I}}{dt} &= \gamma E \\
\frac{d\hat{H}}{dt} &= \eta \alpha I
\end{aligned}$$

Covariance Structure for Samoan Measles Model

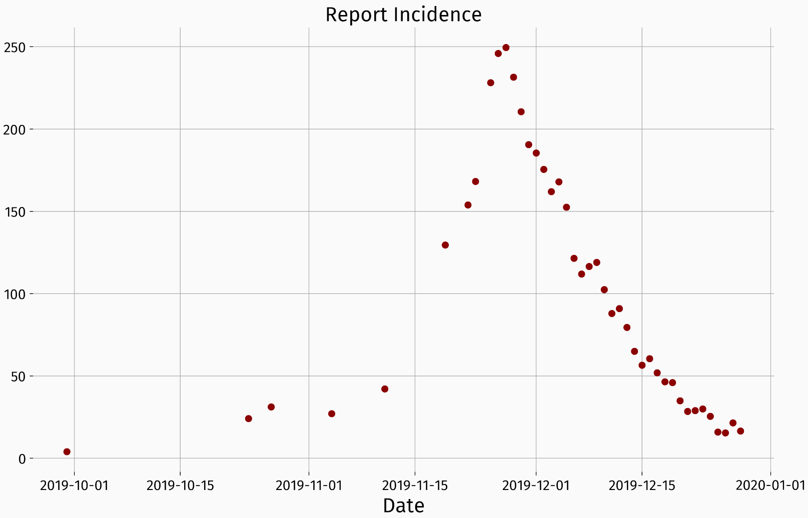
$$\Gamma_d = \begin{bmatrix} \sigma_1 I & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 I & 0 & 0 & 0 \\ 0 & 0 & \sigma_3 I & 0 & 0 \\ 0 & 0 & 0 & \sigma_4 I & 0 \\ 0 & 0 & 0 & 0 & \sigma_5 I \end{bmatrix} \Leftrightarrow \begin{pmatrix} H \\ G \\ D \\ \hat{I} \\ \hat{H} \end{pmatrix} \binom{\text{observed}}{\text{states}}$$

$$\Gamma_m = \begin{bmatrix} s_1 I & 0 & \cdots & 0 & 0 \\ 0 & s_2 I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & s_{N-1} I & 0 \\ 0 & 0 & \cdots & 0 & s_N I \end{bmatrix} \Leftrightarrow \begin{pmatrix} S \\ E \\ \vdots \\ \hat{I} \\ \hat{H} \end{pmatrix} \binom{\text{all}}{\text{states}}$$

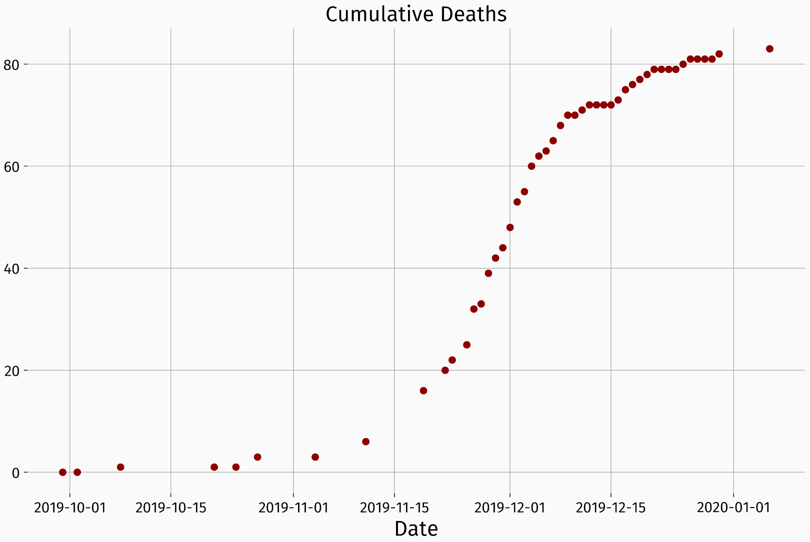
Available Data



Report Incidence

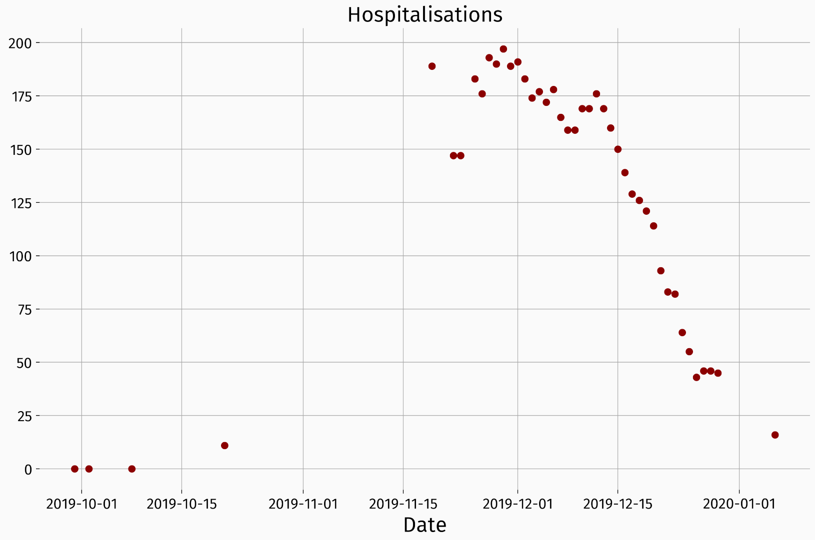


Deaths



Cumulative Deaths
80
60
40
20
0
2019-10-01
2019-10-15
2019-11-01
2019-11-15
2019-12-01
2019-12-15
2020-01-01
Date

Hospitalisations



Full Objective

Assuming Gaussian error structures

$$l(x,\theta) = \|y - g(x,\theta)\|_{L_d}^2 + \|\mathcal{D}x - f(x,\theta)\|_{L_m}^2$$
$$- \log|\Gamma_d| - \log|\Gamma_m|$$

where

$$L_d^T L_d = \Gamma_d^{-1},$$
$$L_m^T L_m = \Gamma_m^{-1}\Delta t$$

Full Objective

We conceptualise the objective function as of the form

$$\sum_i \left\{ \|u_i - v_i\|_{L_i}^2 + s(L_i) \right\}$$

where u_i are some "data", v_i some function of the unknowns (parameters, states), and s some function on the covariance. For our particular formulation, we take:

$$\begin{aligned} u_0 &= y & u_1 &= 0 \\ v_0 &= g(x, \theta) & v_1 &= \mathcal{D}x - f(x, \theta) \\ L_0 &= L_d & L_1 &= L_m \end{aligned}$$

Regularisation can be interpreted as "extra data" on parameters, a la the Bayesian framework.

Estimating Γ

Ramsay and Hooker suggest a simple grid search to determine appropriate Γ when they take the form $\sigma I, \ \sigma \in \mathbb{R}$, but this is both slow and not extensible to more general coviariance structures.

Instead of gridding, we can iterate through candidate Γ by utilising the iterated reweighted least squares method.

This involves sequentially computing optimal values of (x, θ) and (Γ_d, Γ_m) until convergence.

We found that this method typically need to be regualrised via early stopping — if we consider that $\Gamma_m \to \infty$ is the naïve nonlinear least squares formulation, this makes sense.

Interpreting Model/Process Error

$$\|\mathcal{D}x - f(x;\theta)\|_{L_m}^2$$

We can interpret this as an enforcement of an (Euler-Maruyama) discretisation of the SDE

$$\Delta x = f(x;\theta)\,\Delta t + \Delta W$$

with $\mathrm{Var}[\,\Delta W\,] = \Gamma_m \Delta t$

Traditional Confidence Intervals

Traditional confidence intervals use **quadratic approximations** of the log-likelihood (our objective function) about the MLE.

If the parameters are nonidentifiable, this can become problematic — **the Fisher information may be singular** (leading to the (typically) correct inference that finite intervals do not exist), or the approximation may **fail to detect one-sided nonidentifiability** (leading to the incorrect inference that there is a bounded interval).

Predictive Uncertainty (RTO)

Our implementation of RTO works by perturbing the data in the log-likelihood

$$l(x,\theta) = \Big\| \underbrace{y}_{\text{data}} - g(x,\theta) \Big\|_{L_d}^2 + \Big\| \underbrace{0}_{\text{data}} - (\mathcal{D}x - f(x,\theta)) \Big\|_{L_m}^2$$

to

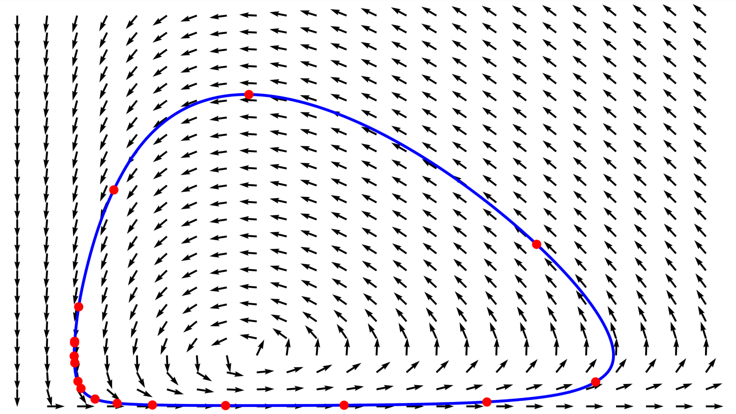
$$l^*(x,\theta) = \left\| y^* - g(x,\theta) \right\|_{L_d}^2 + \left\| r^* - (\mathcal{D}x - f(x,\theta)) \right\|_{L_m}^2,$$

where

$$\begin{aligned} y^* &= y + \epsilon_y, & \epsilon_y &\sim \mathcal{N}(0, \Gamma_d), \\ r^* &= 0 + \epsilon_r, & \epsilon_r &\sim \mathcal{N}(0, \Gamma_m \Delta t). \end{aligned}$$

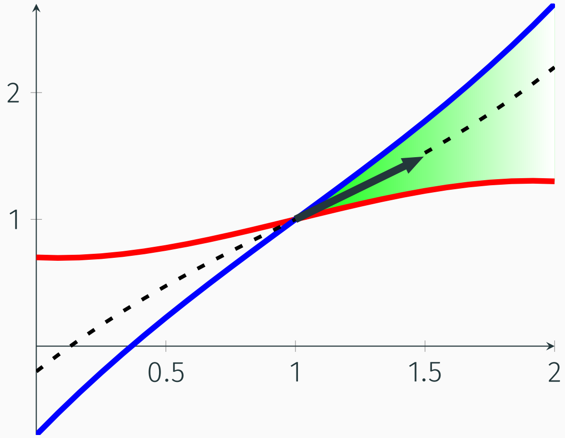
Collocation for model misfit

an example on the Lotka Volterra model in the phase plane



Generalised Profiling

Relaxing for model misspecification



(some of) It is on github.

`dwu402/pypei`

NEWS

National Science Foundation: Science Hard

6/05/02 3:00PM • SEE MORE: SCIENCE & TECHNOLOGY