Informing Policy via Dynamic Models: Cholera in Haiti

Jesse Wheeler AnnaElaine Rosengart Zhuoxun Jiang Kevin Hao
En Tan Noah Treutle Edward Ionides

2023-03-03

Section 1

Introduction

Partially Observed Markov Processes

- The models we focus on are partially observed Markov processes (POMP)

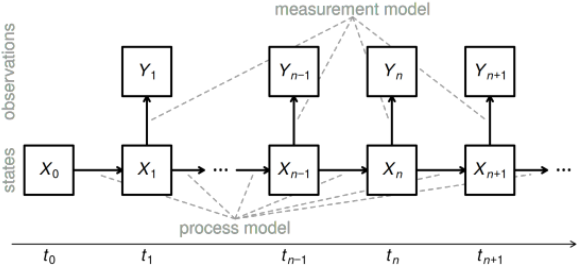


Figure 1: Credit: [1]

POMP Models

- Parameter vector θ that indexes the model
- One step transition density: $f_{X_n|X_{n-1}}(x_n|x_{n-1};\theta)$ that we can simulate from (may not be able to evaluate or express in closed form).
- Measurement Model: $f_{Y_n|X_n}(y_n|x_n;\theta)$.
- Initial density: $f_{X_0}(x_0;\theta)$

Advantages of Statistical Modeling

- Nonlinear-dynamic statistical models have proven to be a useful tool for modeling infectious disease outbreaks (TODO: CITE)
- Most common examples are SIR models and their various extensions.
- These models enable the modeling of scientifically meaningful states, prediction of the future of the outbreak, and modeling the potential effects of interventions (such as vaccinations) (TODO: CITE)

Concerns

- Despite their utility, there exist many cautionary warnings against the use of these types of models (TODO: CITE).
- Concerns include:
  - TODO
  - TODO
  - TODO
- Despite these warnings, there is very little practical advice on how to approach these issues.

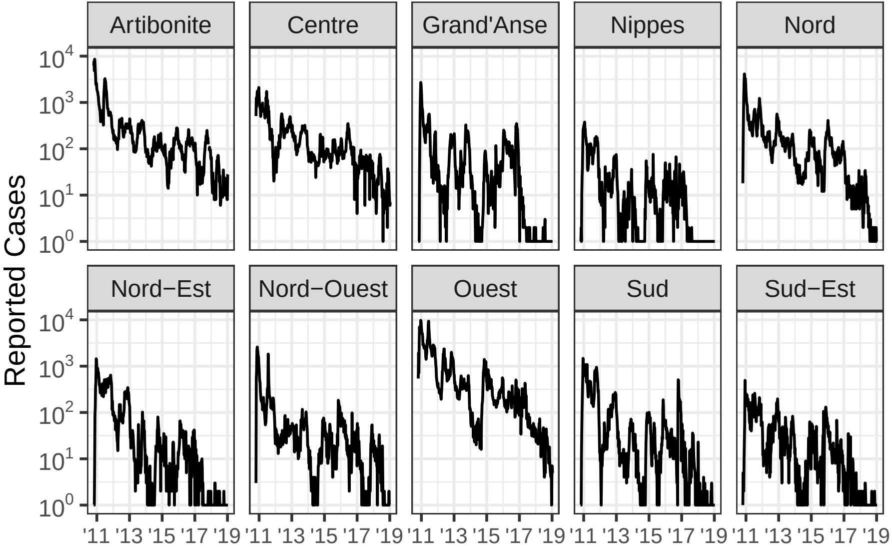
Cholera in Haiti

- We consider the 2010-2019 cholera outbreak in Haiti.
- Cholera was introduced to Haiti in 2010 following the devastating earthquake of the same year.
- Although some new cases have been detected, there were no recorded cholera cases in Haiti between February, 2019 and September 2022 (TODO: Cite).

Data

Reported Cases
10^4
10^3
10^2
10^1
2011
2012
2013
2014
2015
2016
2017
2018
2019

Data


Artibonite
Centre
Grand'Anse
Nippes
Nord
Nord−Est
Nord−Ouest
Ouest
Sud
Sud−Est
Reported Cases
10^4
10^3
10^2
10^1
10^0
'11 '13 '15 '17 '19

Models

- We build on the study by Lee et. al (2020), in which four independent teams built non-linear models to describe cholera dynamics.

	Model 1	Model 2	Model 3
Deterministic / Stochastic	Stochastic	Deterministic	Stochastic
Spatial Model	No	Yes	Yes
Fitting Method	IF2	Trajectory Matching	IBPF

Goals

Our goals in this study is to provide practical advice on how to use non-linear dynamic models to make inference on a dynamic system. We specifically focus on the following topics:

- Model Fitting
- Diagnostics
- Forecasting
- Interpreting Results
- Reproducibility

Section 2

Model Fitting

Computing

- Parameters of nonlinear dynamic models are often fit by finding their posterior distribution or by maximizing some objective measure.
- Because of the non-linear nature of the models, this can be computationally expensive.
- Great care should be taken to determine the necessary amount of computation needed to solve the problem at hand.

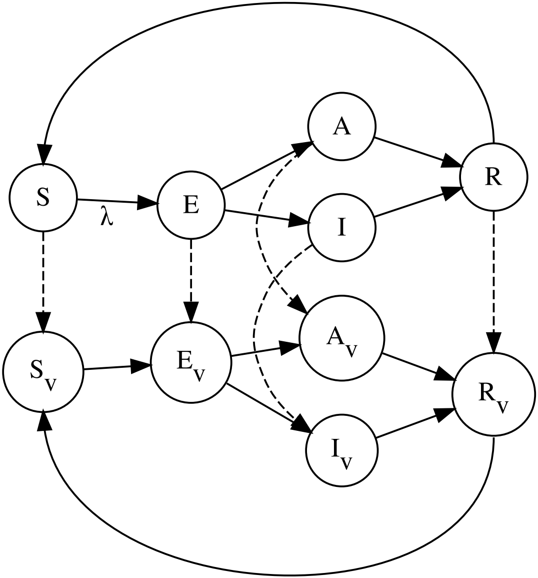
Example: Model 3

- We calibrate this model using the new IBPF algorithm (TODO: Cite), which enables parameter estimation via maximum likelihood for high-dimensional nonlinear POMP models.
- The original version of this model was fit using an alternative approach designed to maximize the model likelihood.
- Without modifications to the model, we obtained parameter estimates corresponding to hundreds of log-likelihood unit improvements to the original model fit.

Nested Models

- Non-linear dynamic models make assumptions about the dynamics of the system in question.
- Consider testing scientifically meaningful nested hypothesis.
- For example, we consider adding a linear trend in transmission to Model 1 in order to account for the apparent decrease in cholera cases.

Example: Model 1



Model 1 (Continued...)

- Individuals move $S \to E$ at time t with a rate of $\lambda(t)$, where:

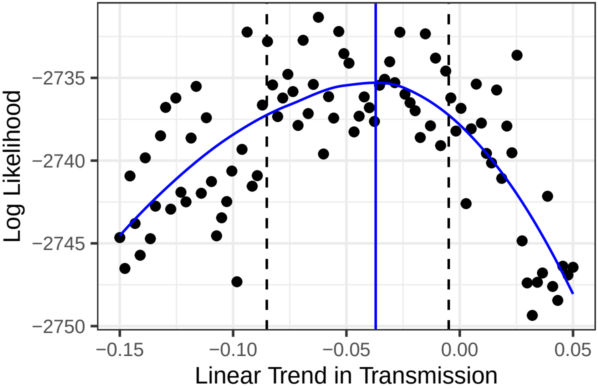
$$\lambda(t) = (I + \epsilon A)^{\nu} \frac{d\Gamma(t)}{dt} \beta(t)/N,$$

$$\log \beta(t) = \sum_{j=1}^{6} \beta_j s_j(t) + \xi \bar{t}$$

- $\frac{d\Gamma}{dt}$ is multiplicative Gamma white-process noise, $\epsilon, \nu, \xi, \beta_{1:6}$ are parameters to be estimated, $s_{1:6}(t)$ are a B-spline basis.

Testing a linear trend

- We test for a linear trend in transmission using a 95% Monte Carlo adjust profile confidence interval (MCAP-CI) (TODO: Cite).
- The test suggests a marginally significant negative linear trend in transmission, with confidence interval $\xi \in (-0.085, -0.005)$.



Section 3

Model Diagnostics

Don't just rely on simulations

text

Interpreting Results

text

Section 4

Model Forecasts

Model Forecasts

- One of the primary reasons dynamic models are fit to data is to obtain forecasts of the future state of the system.
- When done carefully, forecasts can be helpful in informing policy and may enable real-time testing of new scientific hypothesis [2].

Using Real-time data on the system

- Recent information about a dynamic system should be more relevant for a forecast than older information.
- While the previous statement seems self-evident, it is not the case for deterministic models which depend only on initial conditions, and is often not done in practice.
- Let $X_{0:N}$ and $Y_{1:N}$ be random vectors denoting the latent and observed states from times $t_0, t_1, \ldots, t_N$, and let $y^*_{1:N}$ be the observed data that we use to fit a model.
- Forecasts for future times $t_{N+1}, \ldots, t_{N+s}$ should be based on draws from $f_{Y_{N+1:N+S}|Y_{1:N}=y^*_{1:N}}(Y_{N+1:N+S}|Y_{1:N} = y^*_{1:N}; \hat{\theta})$, which can be done simulating starting from hidden states drawn from the filtering distribution at time t_N.

Accounting for All Uncertainty

- Deterministic models are useful for obtaining estimates of general trends of the system but can lead to over-confidence in model forecasts (TODO: Cite kings paper).
- Stochastic models account can include randomness in both the process and measurement models (TODO: Cite ionides paper on this)
- It is important to also account for uncertainty in parameter estimates, as uncertainty in just a single parameter can lead to drastically different forecasts [3].

Section 5

Reproducibility

Reproducibility

Text

Bibliography

[1] KING, I., A. (2022). Short course on simulation-based inference for epidemiological dynamics. *Summer Institute in Statistics and Modeling in Infectious Diseases (SISMID)*.

[2] LEWIS, A. S. L., ROLLINSON, C. R., ALLYN, A. J., ASHANDER, J., BRODIE, S., BROOKSON, C. B., COLLINS, E., DIETZE, M. C., GALLINAT, A. S., JUVIGNY-KHENAFOU, N., KOREN, G., MCGLINN, D. J., MOUSTAHFID, H., PETERS, J. A., RECORD, N. R., ROBBINS, C. J., TONKIN, J. and WARDLE, G. M. (2023). The power of forecasts to advance ecological theory. *Methods in Ecology and Evolution* **14** 746–56.

[3] SALTELLI, A., BAMMER, G., BRUNO, I., CHARTERS, E., DI FIORE, M., DIDIER, E., NELSON ESPELAND, W., KAY, J., LO PIANO, S., MAYO, D., et al. (2020). Five ways to ensure that models serve society: A manifesto.