

# Plug-and-play Bayesian inference for compartmental models in PLOM

Joseph Dureau, Sébastien Ballesteros

July 6, 2013

## 1 Introduction

## 2 Compartmental models

### 2.1 Definition

Compartmental models are a general framework used to represent the state of a countable population (of humans, animals, molecules, etc) and its evolution. At a given time  $t$ , the population is described by the number of individuals in each of  $n$  possible states: the ensemble of individuals in a same state defines what is termed as a compartment. Each individual belongs to one and only one compartment. Individuals within a same compartment are considered indistinguishable. We will consider in this document that  $n$  is known, fixed, and finite.

Each compartment can correspond to very diverse characterisations, depending on the context. They can be used to track the status of an individual with regards to a given disease in a human or animal population (susceptible or infected, for example), their age or their geographical location. Additionnally, compartments can be used to track the number of specimens of different animal species in an ecosystem, as in the Lotka-Volterra predator-prey model. They can also be used in physics and chemistry to determine molecule types, their electronic charge or radioactive state, for example. Less classical illustrations of the use of compartmental models include tracking the spread of rumors among a population, or the propagation of economic difficulties among countries following a financial crisis.

We note  $C_t^{(i)}$  the size of compartment  $i$  ( $1 \leq i \leq n$ ) at time  $t$ , and  $z_t = [C_t^{(1)}, \dots, C_t^{(n)}]$ . A model is defined by a (finite) number  $m$  of transformations of the system called reactions (the ensemble of all indexes is noted  $\mathcal{R}$ ). These reactions correspond to one or several individuals passing from one compartment to another, or arriving or leaving the total population. In any case, each reaction  $k$  is characterised by its effect on the structure of the population corresponding to a vector  $l^{(k)} \in \mathbb{Z}^n$ , and its intensity of occurrence  $r_t^{(k)}$ . We allow for these rates to depend on time, to reflect the influence of potential forcing of external factors on the system, and to depend on a finite set of constant quantities gathered in a parameter vector  $\theta$ . In addition, we consider transition rates to be density-dependent, i.e. that they can be written as  $r_t(z_t, \theta) = N\dot{r}(z_t/N, \theta) = N\dot{r}(\dot{z}_t/N, \theta)$  when  $N$  is the total size of the population. In the remaining of this document, we will define compartmental models using the following formalism:

Reaction	Effect	Rate
reaction 1	$z_t \rightarrow z_t + l^{(1)}$	$r^{(1)}(z_t, \theta)$
...	...	...
reaction $k$	$z_t \rightarrow z_t + l^{(k)}$	$r^{(k)}(z_t, \theta)$
...	...	...
reaction $m$	$z_t \rightarrow z_t + l^{(m)}$	$r^{(m)}(z_t, \theta)$

Formally, this framework leads to the definition of a Markovian jump process, which dynamic can be expressed in the following way:

Markovian jump process compartmental model

$$\begin{aligned} P(z_{t+\delta} = z_t + l^{(k)} | z_t) &= r_t^{(k)}(z_t, \theta) \delta + o(\delta) \quad \text{for any } k \in \mathcal{R} \\ P(z_{t+\delta} = z_t | z_t) &= \left(1 - \sum_{k \in \mathcal{R}} r_t^{(k)}(z_t, \theta) \delta\right) + o(\delta) \end{aligned} \quad (1)$$

Under some regularity conditions detailed in Ethier and Kurtz (1986), Fuchs (2013) or Guy et al. (2013), and due to the density-dependance of transition rates, the dynamic of the system converges to a deterministic behaviour as the population size tends to infinity. For finite populations, the additional stochastic behaviour is termed *demographic* stochasticity.

## 2.2 Environmental stochasticity

Extensions of the compartmental modeling framework introduced in the previous section have been proposed by the authors of Breto et al. (2009), to account for additional sources of uncertainty related to fluctuations in extrinsic determinants of the epidemic. This uncertainty is reflected through additional sources of stochasticity, termed *environmental* stochasticity. The approach suggested in Breto et al. (2009) is to consider stochastic transition rates  $\tilde{r}_t^{(k)}$  for a subset  $\mathcal{R}^e$ , under the following constraints for all  $t$ :

$$\begin{aligned} \mathbb{E}(\tilde{r}_t^{(k)}) &\sim r_t^{(k)} \\ \tilde{r}_t^{(k)} &\geq 0 \end{aligned}$$

However, uncertain variations of extrinsic factors cannot always be modeled through high-frequency independent fluctuations. The evolution of climate, for example, has been shown to exhibit complex seasonal and inter-annual variations that influence epidemic dynamics Viboud et al. (2004). Following the work of Cazelles and Chau (1997) and Cori et al. (2009), the authors of Dureau et al. (2013) have proposed a general inferential framework for time-varying parameters, that is extended in the present document. Under this approach, parameters are modeled through stochastic differential equations or extensions thereof. The state vector is extended with additional components  $x_t^{\theta_t}$  which dynamic is determined by the following equation:

$$dx_t^{\theta_t} = \mu^{\theta_t}(x_t^{\theta_t}, \theta) dt + L^{\theta_t} dB_t^{Q^{\theta_t}} \quad (2)$$

Note that some constraints as positivity or boundedness generally need to be preserved when allowing parameters to vary over time, which is achieved by defining  $x_t^{\theta_t}$  as the respectively the log or logit transformation of the quantity of interest.

## 2.3 Examples

## 2.4 Tractable approximations of compartmental models

### 2.4.1 Ordinary differential equations

The simplest and most stringent approximation of compartmental models are ordinary differential equations. Under this formalism, the number of individuals in each compartment takes continuous values, and varies continuously (and in a differentiable manner) over time. More specifically, all kind of demographic or environmental stochasticity are neglected, leading the state of the system to evolve deterministically. From a practical perspective, the use of ordinary differential equations drastically simplifies the process of Bayesian inference, mainly due to the deterministic one-to-one mapping between trajectories  $z_{0:T}$  and parameters  $\theta$ .

This formalism can be legitimately used for large populations and when all significant environmental factors have been explicitly incorporated in the deterministic skeleton of the model. However, in alternative cases results should be treated with caution, and the use of alternative formalisms accounting for sources of stochasticity may be required.

### 2.4.2 Stochastic differential equations

### 2.4.3 Poisson process with stochastic rates

## 3 Plug-and-play Bayesian inference

### 3.1 The Bayes rule, and quantities of interest

### 3.2 Conditional state exploration

### 3.3 Full inference of paths and parameters

#### 3.3.1 Particle Marginal Metropolis Hastings algorithm

#### 3.3.2 Efficiently starting the PMMH close to the global mode

#### 3.3.3 Efficient calibration of the PMMH sampling covariance matrix

## 4 Examples

## 5 Perspectives: call for contributions

## References

- Breto, C., He, D., Ionides, E., King, A., Ghosal, S., Lember, J., and van der Vaart, A. (2009). Time series analysis via mechanistic models. *Annals*, 3(1):319–348.
- Cazelles, B. and Chau, N. (1997). Using the kalman filter and dynamic models to assess the changing hiv/aids epidemic. *Mathematical biosciences*, 140(2):131–154.
- Cori, A., Boelle, P., Thomas, G., Leung, G., and Valleron, A. (2009). Temporal variability and social heterogeneity in disease transmission: The case of sars in hong kong. *PLoS computational biology*, 5(8):e1000471.

- Dureau, J., Kalogeropoulos, K., and Baguelin, M. (2013). Capturing the time-varying drivers of an epidemic using stochastic dynamical systems. *Biostatistics*.
- Ethier, S. and Kurtz, T. (1986). *Markov processes: characterization and convergence*, volume 6. Wiley.
- Fuchs, C. (2013). *Inference for Diffusion Processes: With Applications in Life Sciences*. Springer.
- Guy, R., Larédo, C., and Vergu, E. (2013). Approximation of epidemic models by diffusion processes and their statistical inference. *arXiv preprint arXiv:1305.3492*.
- Viboud, C., Pakdaman, K., Boelle, P., Wilson, M., Myers, M., Valleron, A., and Flahault, A. (2004). Association of influenza epidemics with global climate variability. *European Journal of Epidemiology*, 19(11):1055–1059.