1. ptmc package performance

1.1 Background

Mathematically modelling the propagation of infectious diseases often involves solving complex non-linear SIR-type models. The evolution of these models are dictated by parameters values which are often difficult to accurately estimate due to biases in existing studies or surveillance data, leading to uncertainty in their estimates. The desire to formally incorporate such parameter uncertainty in epidemic models naturally leads to the use of Bayesian inference. However, combining Bayesian inference with complex non-linear systems forms intractable integrals which require Markov Chain Monte Carlo (MCMC) methods to solve. Given the rich, dense literature on MCMC samplers, it is often not clear which MCMC sampler is best to use (i.e. provides convincing convergence over a small time frame) for SIR-type infectious diseases models.

This analysis explores this issue, by comparing several popular MCMC samplers with a promising population-based MCMC sampler (parallel tempering), to sample from the posterior distribution for parameters in an age- and exposure-stratified epidemic model for Respiratory Syncytial Virus. The package finds that the population based parallel tempering MCMC sampler is significantly quicker and has improved mixing compared to popular MCMC samplers in existing packages. Therefore it is encouraged that mathematical modellers who require MCMC to calibrate epidemic models to use parallel tempering or explore similar population based samplers to ensure quick and convincing convergence.

1.2 Methods

1.2.1 Description of epidemic model

I modelled the number of individuals in six different epidemiological states, S_0, I_0, R_0, S_1, I_1, and R_0, where S, I, and R refer to susceptible, infectious and recovered individuals respectively. The subscript 0 relates to primary infection, and subscript 1 refers to secondary and subsequent infections. On average, the infectious period lasts for $1/\gamma$ days and the duration of immunity lasts $1/\omega$ days. The epidemiological states are stratified into two different age groups (0–2 years, and 2+ years), denoted by superscript a. To ensure the population of each age group remains constant, I assume that $\mu = 1861$ live births occur each day into the 0-2 years age group and that the same number of persons die each day from the 2 years and older group. The rate at which individuals acquire infection, λ_i^a depends on contacts between age groups and the number infectious persons in each age group. Thus, the RSV epidemic model is specified in a continuous time system for age group a is given by:

$$
\begin{array}{rcll}
 & & \overbrace{\qquad\qquad}^{\text{Transmission terms}} & \overbrace{\qquad\qquad}^{\text{Ageing terms}} \\
\frac{dS_0^a}{dt} & = & \overbrace{\mu \mathbb{1}_1(a) - \lambda^a S_0^a}^{} & \overbrace{-\eta^a S_0^a + \eta^{a-1} S_0^{a-1}}^{} \\
\frac{dI_0^a}{dt} & = & \lambda^a S_0^a - \gamma I_0^a & -\eta^a I_0^a + \eta^{a-1} I_0^{a-1} \\
\frac{dR_0^a}{dt} & = & \gamma I_0^a - \omega R_0^a & -\eta^a R_0^a + \eta^{a-1} R_0^{a-1} \\
\frac{dS_1^a}{dt} & = & \omega R_0^a + \omega R_1^a - \delta_1 \lambda^a S_1^a & -\eta^a S_1^a + \eta^{a-1} S_1^{a-1} \\
\frac{dI_1^a}{dt} & = & \delta_1 \lambda^a S_1^a - \gamma I_1^a & -\eta^a I_1^a + \eta^{a-1} I_1^{a-1} \\
\frac{dR_1^a}{dt} & = & \gamma I_1^a - \omega R_1^a & -\eta^a R_1^a + \eta^{a-1} R_1^{a-1} \\
\frac{dZ^a}{dt} & = & \lambda_0^a S_0^a + \lambda_1^a S_1^a &
\end{array}
\tag{1.1}
$$

where $\mathbb{1}_1(a)$ is the indicator function (non-zero at $a = 1$), η^a is the rate of ageing from age group a to age group $a+1$ (with $\eta^0 = 0$), and $\lambda^a(t)$ is the force of infection for age group a given by

$$
\lambda_i^a = \sum_{b=1}^{2} c^{a,b} \frac{I_0^b + I_1^b}{N^b} \tag{1.2}
$$

where N^b is the population of age group b, δ_1, is the relative susceptibility of the population to RSV infection after being infected, and $c^{a,b}$ is the number of daily contacts made between age group a and b, estimated from the POLYMOD study.

he cumulative number of infections in age group a is given by Z^a, so the number of new infections z_w^a per week w in age group a can be calculated:

$$
\left. \frac{dZ^a(t)}{dt} \right|_{7(w-1)}^{7w} = z_w^a \tag{1.3}
$$

1.2.2 Likelihood

The data, x, used the calibrate the model is the number of weekly (w) confirmed positive RSV samples per age group (a) from the Respiratory DataMart System (RDMS) at PHE between July 2010 and June 2017. This data is patient sensitive so cannot be shared publicly. In the likelihood I assume that the number of confirmed positive RSV samples per age group and week (x_w^a) is the realisation of the random variable X_w^a,

$$
X_w^a \sim \text{Bin}\left(z_w^a, \epsilon^a\right) \tag{1.4}
$$

where ϵ^a is the probability that someone with RSV is correctly identified in the RDMS dataset. By representing the set of all parameters to be fitted in the model as

$$
\Psi = \{\gamma, \omega, \delta_1, \alpha, a, b, \phi, \psi, I_1, I_2, I_3, \epsilon^1, \epsilon^2\}
$$

where I_1, I_2, I_3 are seeding parameters at $t = 0$ as defined in **Table 1.1** the likelihood calcualted through:

$$
\mathcal{L}(x, \Psi) = \prod_{t=1}^{52*7} \prod_{a=1}^{2} \mathcal{L}(x_w^a; \Psi) = \prod_{t=1}^{52*7} \prod_{a=1}^{2} \binom{z_w^a}{x_w^a} (\epsilon^a)^{x_w^a} (1 - \epsilon^a)^{z_w^a - x_w^a} \tag{1.5}
$$

1.2.3 Prior distributions

A summary of all the model parameters with their prior distributions are contained in **Table 1.1**. In implementation the parameters are logit-transformed.

	Parameter	Prior distribution	Source
$1/\gamma$	Duration of infection (days)	$\mathcal{U}(2, 20)$	
$1/\omega$	Duration of immunity (days)	$\mathcal{U}(60, 365)$	
δ_1	Relative proportion of population susceptibility RSV after primary infection	$\mathcal{U}(0, 1)$	—
α	Relative infectiousness of RSV after primary infection	$\mathcal{U}(0, 1)$	—
Parameters associated with transmission			
a	Rate of infection	$\mathcal{U}(0, 1)$	—
b	Amplitude	$\mathcal{U}(0, 1)$	—
ϕ	Peak transmission (offset) (days)	$\mathcal{U}(100, 250)$	RDMS
ψ	Width of seasonal peak	$\mathcal{U}(0, 36.5)$	—
Initial parameters (at $t = 0$)			
I_1	Initial proportion infected RSV	$\mathcal{U}(0, 1)$	—
I_2	Initial proportion recovered from RSV	$\mathcal{U}(0, 1)$	—
$1/I_3$	Average duration between infections (years)	$\mathcal{U}(1, 4)$	—
Detection rates			
ϵ^1	0–2 years	$\mathcal{U}(0.0, 0.01)$	—
ϵ^2	2+ years	$\mathcal{U}(0.0, 0.01)$	—

Table 1.1: Prior distributions of the parameters in the mathematical model.

1.2.4 MCMC samplers

Overview of sampling methods

To sample values from the posterior distribution of the fitted parameters, we use Bayes rule

$$\pi(\Psi) = p(D|\Psi) \propto \mathcal{L}(x, \Psi)p(\Psi) \tag{1.6}$$

which simplifies the problem to sampling from the previously-defined likelihood and priors. To generate the random samples we use the Metropolis-Hastings algorithm (a MCMC method). We consider four different types of Metropolis-Hastings algorithms; a normal Metropolis-Hastings (MH), an adaptive Metropolis-Hastings (AMH), a Metropolis-Hastings with differential evolution (DE) and a parallel tempering (PT). For the MH algorithm the proposal distribution assumes no covariance between parameters. The AMH algorithm in which the proposal distribution is a normal distribution with a covariance matrix estimated from previous values in the chain. The DE algorithm is one in which XXXX. The PT algorithm runs various chains which have an likelihood which is altered so that some (hotter) chains accept many proposals and other (colder) chains are similar to an unaltered likelihood.

Implementation

The MH, AMH, and DE samplers are implemented via the BayesianTools R package. Explicit details about their implementation is best explained within the documentation. These three samplers for 1,000,000 steps with thinning occurring every 10th step. The PT sampler is implemented via the ptmc R package () and is run for 100,000 steps with a temperature ladder of 10 chains per run with thinning every 10th step (only the coldest chain is used to calculate the posterior distribution). The burn in is half of the total run time and each of the four samplers and each are ran ten times so that convergence via Gelman-Rubin statistics and manual inspection of trace plots can be performed. An R Markdown file with all code used for this analysis can be found XX.

1.3 Results

The trace plots for the log(-LL) for 10 runs of each of the four MCMc samplers are given below.

1.3.1 MH

For the standard random-walk Metropolis-Hastings algorithm, only one of the ten chains appears to reach the convergent state after one millions samples.

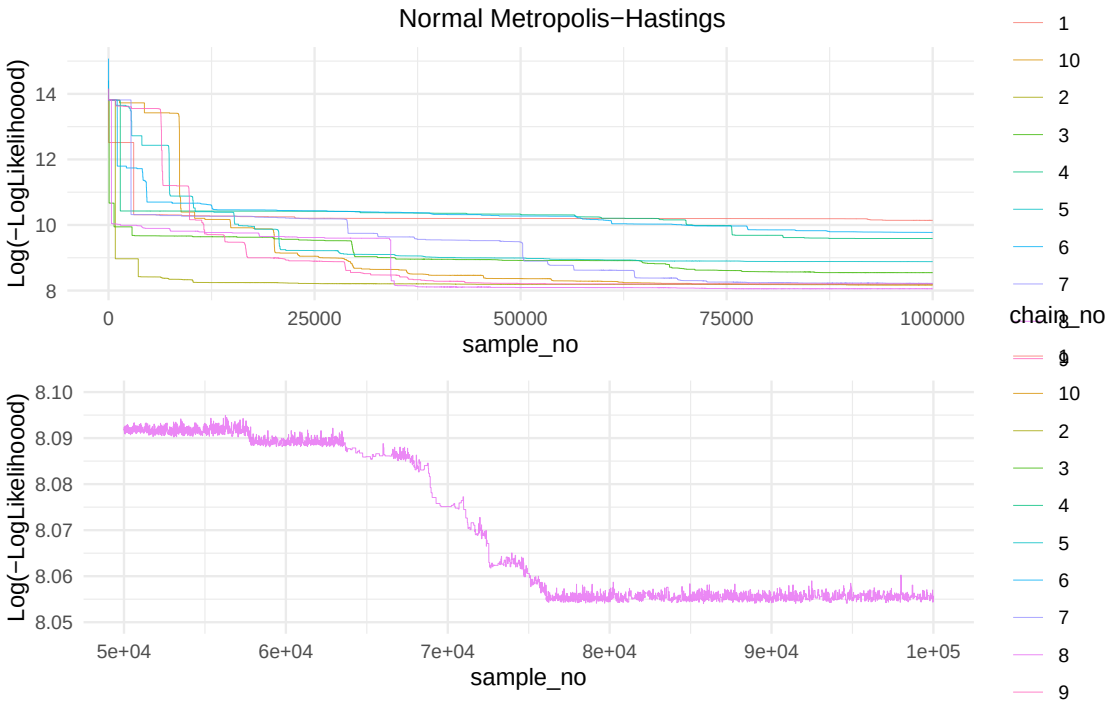

Figure 1.1

1.3.2 AM

For the adaptive Metropolis-Hastings algorithm, only two of the ten chains appears to reach the convergent state after one millions samples.

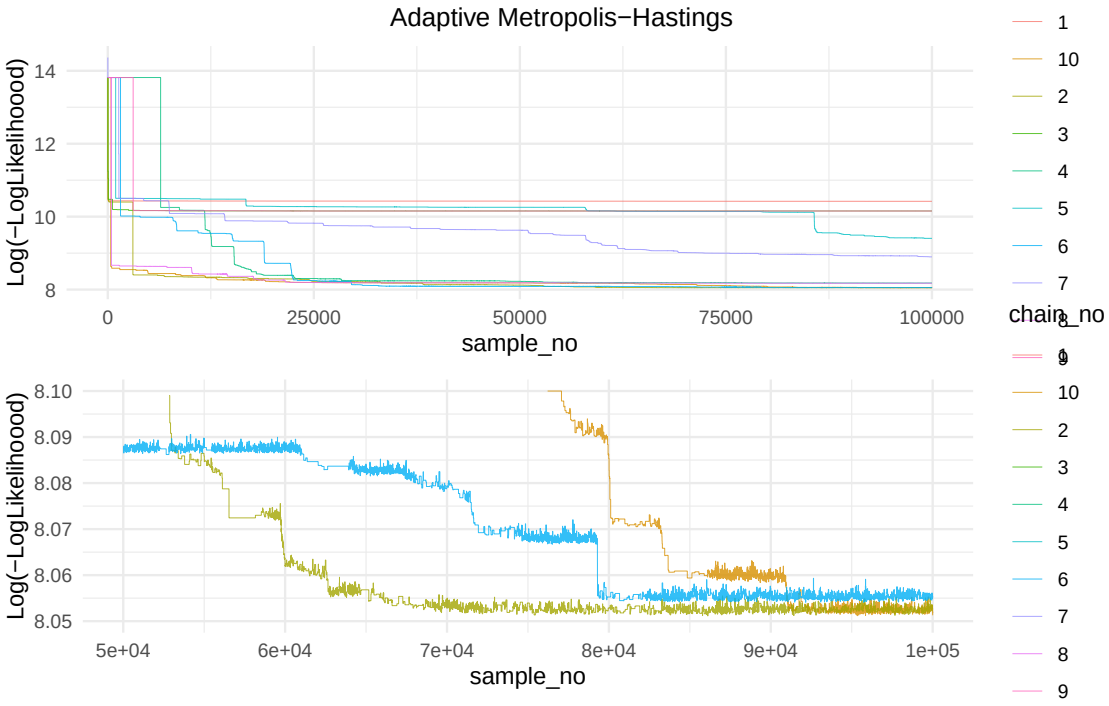

Figure 1.2

1.3.3 DE

For the adaptive Metropolis-Hastings algorithm, only one of the ten chains appears to reach the convergent state after one millions samples.

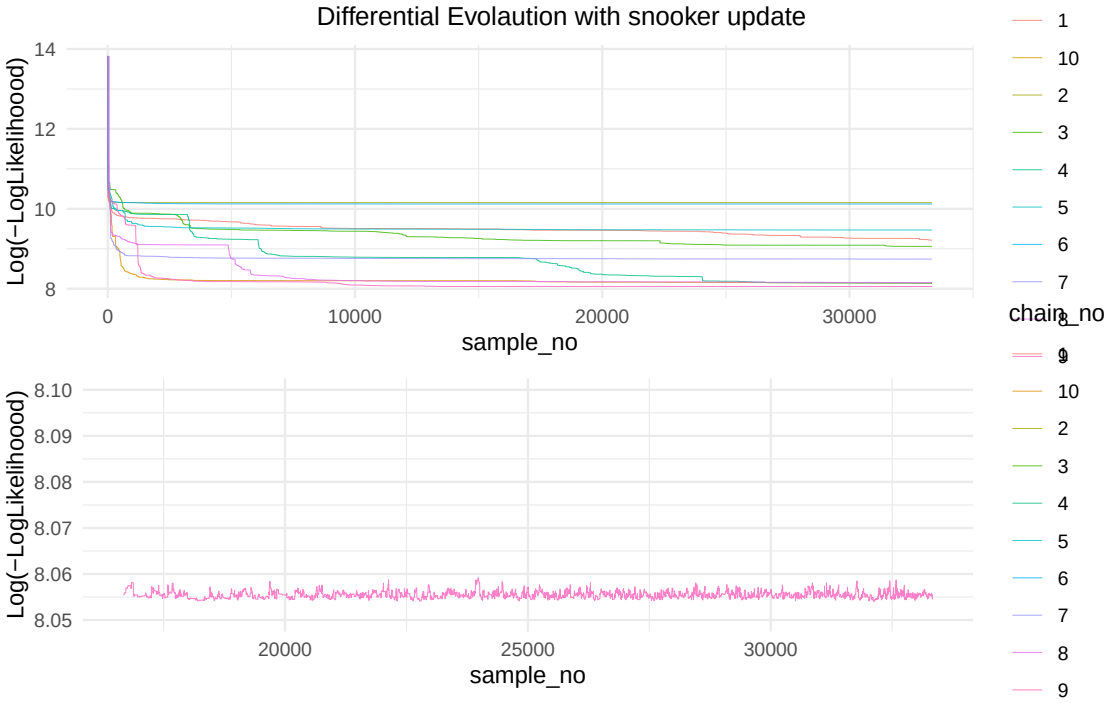

Figure 1.3

1.3.4 PT

or the adaptive Metropolis-Hastings algorithm, all ten chains appear to reach the convergent state after 100,000 samples.

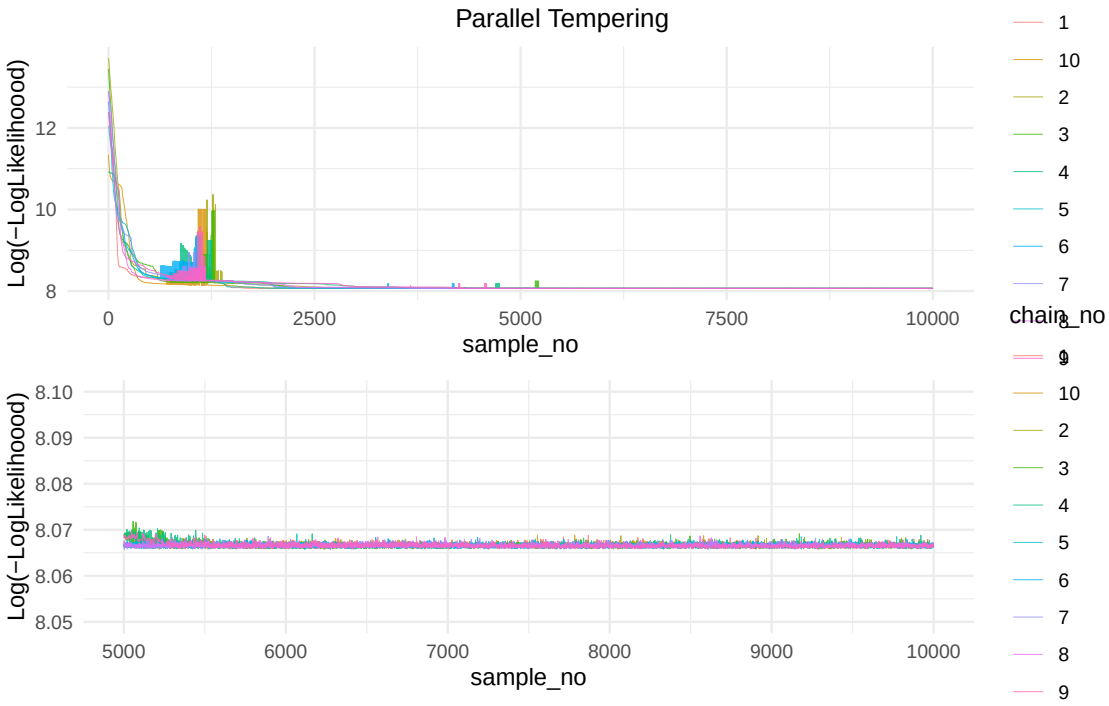

Figure 1.4

1.4 Discussion

The MCMC samplers are run for a fixed duration to see which provides the most reassuring convergence in a given space of time. The parallel tempering has the most convincing convergence compare to the other 3 samplers, with all 10 of the chains converging in contrast to the other MCMC samplers. The multivariate potential scale reduction factor is 1.03 and less than 1.08 for all the fitted parameters for the parallel tempering, suggesting convergence has occurs.

It is encouraged that future epidemic models will use the parallel tempering MCMC sampler (through the ptmc R package) to sample from posterior distributions. Using such pre-built algorithms will assuage mathematical modellers concerns surrounding convergences of epidemic models, allowing more time for analysis more relevant to policy makers (i.e. impact of intervention programmes or cost-effectiveness analysis) to be performed.

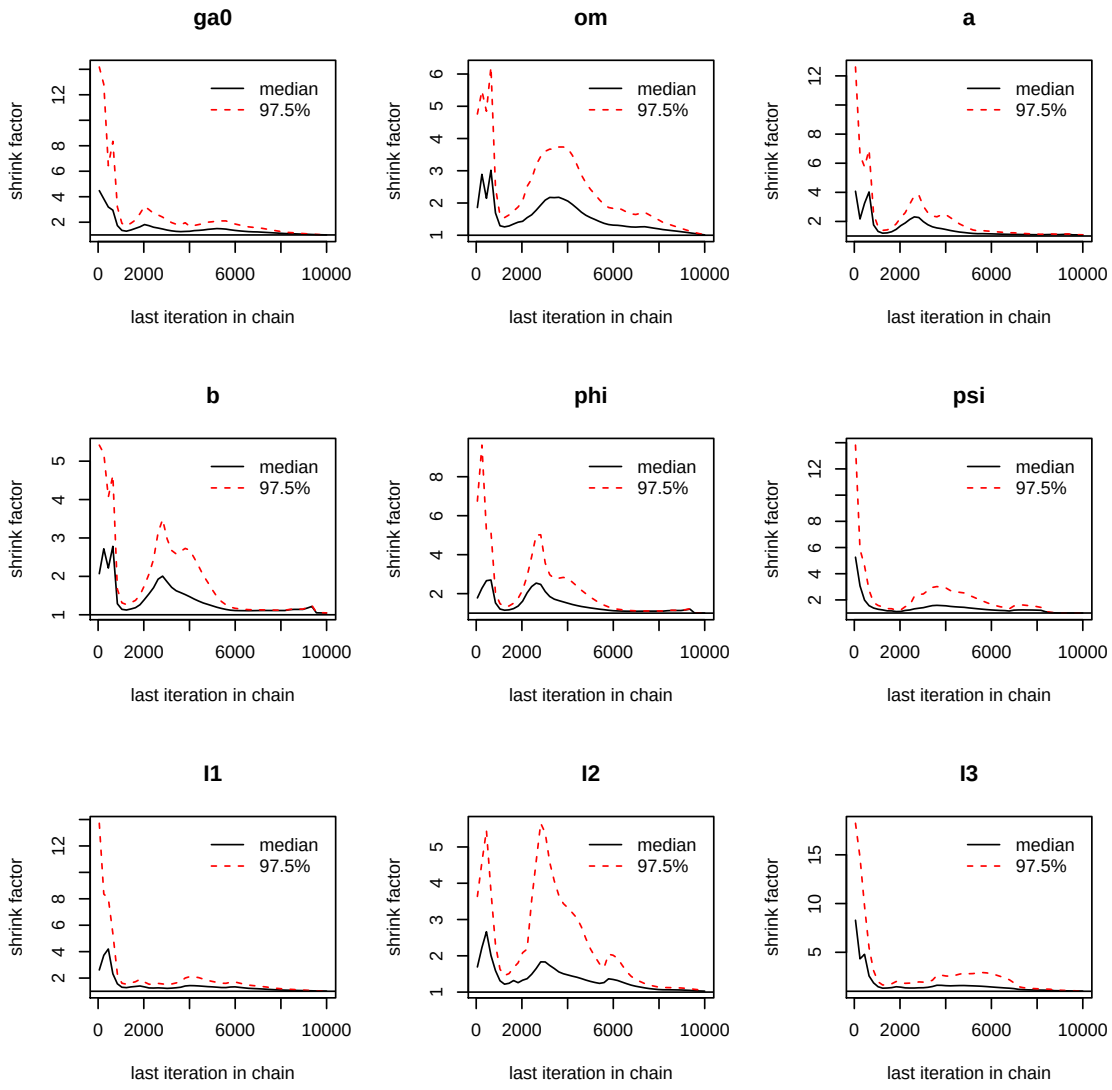
ga0
om
a
b
phi
psi
l1
l2
l3
shrink factor
median
97.5%
last iteration in chain
0
2000
6000
10000

Figure 1.5