Bayesian inference in a hidden stochastic two-compartment model for feline hematopoiesis

D. Golinelli*
RAND Statistics Group
Santa Monica, CA USA
and
P. Guttorp and J.A. Abkowitz
University of Washington
Seattle, WA USA

February 24, 2006

Abstract

In this paper we describe a hidden two-compartment stochastic process used to model the kinetics of feline hematopoietic stem cells in continuous time. Because of the experimental design and data collection scheme the inferential task presents numerous challenges. While the hematopoietic process evolves in continuous time, the observations are collected only at discrete irregular times and are a probabilistic function of the state of the process. In addition, the animals go through an experimental procedure such that their reserve of hematopoietic stem cells is severely depleted at the start of the observation period. This impedes any approximation of the hematopoietic process with a continuous state-space process (normal approximation of the transition probabilities would be inaccurate when the state of the process, i.e. the number of stem cells, is small). We implement an MCMC algorithm that allows us to estimate the posterior distribution of the parameters of the hematopoietic process while maintaining its state-space discrete (i.e. without using any approximation). We show the performance of the algorithm on simulated data. Finally, we apply the algorithm to data on multiple experimental cats and provide estimates of the rates of the fates of feline hematopoietic stem cells. The obtained estimates are in agreement with the estimates obtained with different methods published in the medical literature. However the proposed approach makes a more efficient use of the data and hence the parameter estimates are much more accurate than the one obtained with the methods previously proposed.

*daniela@rand.org

Keywords: Stochastic two-compartment model, Hidden Markov models, RJMCMC, Hematopoiesis.

1 Introduction

The main objective of this paper is to develop inferential tools for a class of hidden stochastic population processes. In particular, we will be focusing on a hidden stochastic two-compartment model proposed for studying hematopoiesis, the process of blood cell production (Abkowitz et al., 1996).

Stochastic two-compartment processes are used to model different phenomena, making the methods presented here applicable to a wider spectrum of problems. For example, the well-known SIR (Susceptibles-Infected-Removed) model used to describe the spread of infectious diseases is a stochastic two-compartment process. In addition, a stochastic two-compartment model very similar to the one described in this paper has been proposed for modelling the spread of malaria within a human host (Gravenor et al., 1998).

In most of the real applications these processes can only be partly observed. Partial observations may be due to questions of efficiency or to physical constraints. For example, in the case of the SIR model, usually it is only possible to observe the removal times but not the infection times (Gibson and Renshaw, 1998). In biology a scientist might be interested in the behavior of a population of cells that reside in a living body of which only a few subsets over time can be observed.

Partial observations make the computation of the likelihood function difficult. Such computation requires an integration step over the space of the hidden or missing data. This step presents the greatest difficulty. Advances in stochastic integration methods, like Markov chain Monte Carlo methods (MCMC) (Gilks et al., 1996), have led to the development of inferential methods that would not have been feasible a few years ago. In particular, when adopting a Bayesian approach MCMC methods have been extremely useful in determining the posterior distribution for the parameters of complicated stochastic systems.

This paper is organized as follows. In section 2 we describe the stochastic two-compartment model. In section 3 we describe the MCMC algorithm. In section 4 we illustrate the performance of the algorithm on a simulated dataset and in section 5 we apply the method to the available experimental data. Section 6 discusses the obtained results. The appendix contains computational details of the algorithm.

2 A stochastic two-compartment model for hematopoiesis

In this section we describe the stochastic two-compartment model used for hematopoiesis. This model represents an extension and an improvement with respect to the model proposed by Guttorp et al. (1990).

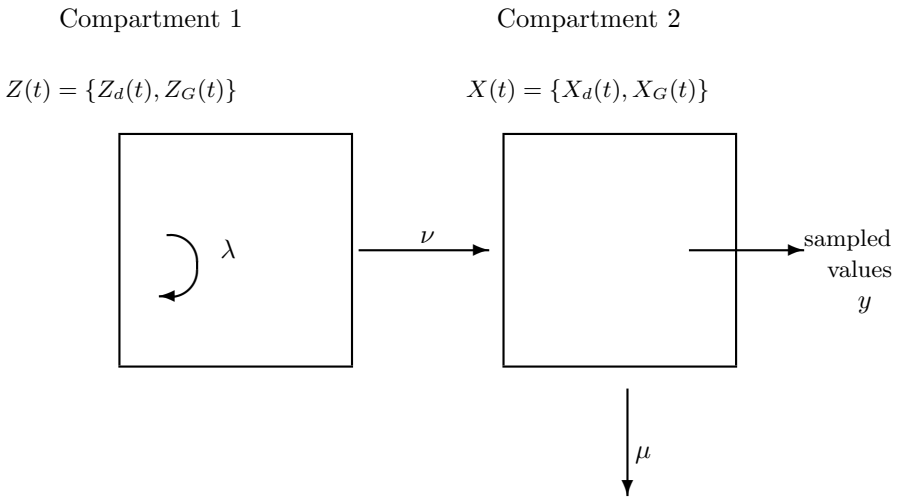

Figure 1: A stochastic two-compartment model for hematopoiesis.

Hematopoiesis is the process of blood cell production. More precisely, it is the process in which hematopoietic stem cells (a primitive blood cell), through sequential divisions, differentiate into progenitor cells. These last cells in turn can differentiate into white blood cells, red blood cells or platelets. While a lot is known about how progenitor cells differentiate, since their cell-cycle kinetics has been studied both in vivo and in vitro, very little is known about hematopoietic stem cells (HSC) behavior. This is due to the fact that HSCs are difficult to isolate, as they do not have a unique physical or antigenic phenotype. HSCs support the entire blood and immune system and reconstitute hematopoiesis after transplantation. Understanding their kinetics is of great importance. For example this could lead to new treatments for leukemia and more effective clinical HSC transplantation procedures.

Figure 1 shows a diagram of the model proposed by Abkowitz et al. (1996) to describe the hematopoiesis process. This model is an example of a stochastic two-compartment model. The first compartment is the reserve where the HSCs reside. The second compartment is the committed cells compartment and it is where the progenitor cells reside. The population of HSCs behaves as a simple linear birth-death process since HSCs can either give birth to other HSCs at a rate λ or differentiate into progenitor cells at a rate ν.

The population of progenitor cells, instead, behaves as a non-homogeneous immigration-death process. The population of progenitor cells increases when HSCs exit from the reserve compartment and immigrate to the second compartment. It decreases at a rate μ when the progenitor cells differentiate into more specialized blood cells. Since HSCs are not identifiable, observations are taken

from the second compartment. Because the population of interest resides in a living body, we can only observe subsets of the population of progenitor cells at discrete times.

In order to get an idea of the contributions of the feline stem cells to the progenitor cells compartment, researchers designed a specific experiment on female Safari cats. Safari cats are the offspring of matings between a domestic cat and the South American Geoffroy wild cat and for this reason they have an electrophoretically distinct phenotype of the X-chromosome-linked enzyme glucose-6-phosphate dehydrogenase (G6PD). During embryogenesis, since either the paternal or the maternal X-chromosome is inactivated, the female Safari cats have some somatic cells expressing the domestic-type G6PD (d G6PD) and other expressing the Geoffroy-type G6PD (G G6PD). The G6PD phenotype is retained after replication and differentiation and it is neutral. That is, the cells that express it do not have significant replication/differentiation advantages. Therefore, it provides a binary marker or label of each cell and its clones. It follows that the total number of cells both in compartment one and two in figure 1 can be seen as the sum of two independent and identically distributed population processes that differ only for a label: d G6PD or G G6PD. In short the population process in figure 1 is two-dimensional in both compartments. One dimension is the population of cells expressing the d phenotype, the other is the population of cells expressing the G phenotype. For more details on the experiment see Abkowitz et al. (1988, 1990, 1993).

However, observing the percentage of progenitor cells expressing the d G6PD phenotype over a period of almost 6 years in normal female Safari cats (with observations taken every 4 weeks approximately) did not seem to provide much information about the HSC behavior. In fact this percentage remained relatively constant during the six years of observation suggesting that hematopoiesis is a polyclonal and stable process.

Researchers believed that there should be more information in observing the hematopoiesis process when it is supported by a much smaller number of stem cells. For this reason, a number of female Safari cats were irradiated in order to kill their bone marrow (where HSCs reside) and a small number of bone marrow cells, collected prior to the radiation, was transplanted back. Since there are few HSCs in large animals, at the start of the experiment the transplanted cats are likely to contain a very small number of HSCs. For this reason the process modelling the HSCs behavior should be a discrete state-space process and not a continuous one. The behavior of the binary label (d G6PD versus G G6PD) within the progenitor cells was then monitored in samples taken every two to six weeks. Under this setting, the percentage of labelled cells is more variable over time. For example, some cats showed wide clonal fluctuations during the first year or so and then stabilized, suggesting that hematopoiesis was supported by only one or two clones.

Formally the Markov population process just described is denoted by the vector $W(t) = \{Z(t), X(t)\} = \{Z(t) = (Z_d(t), Z_G(t)), X(t) = (X_d(t), X_G(t))\}$, where $Z(t)$ represents the size of the reserve compartment at time t and $X(t)$ denotes the size of the committed cells compartment at time t. Note that both

$Z(t)$ and $X(t)$ are two-dimensional, where, for example, $Z_d(t)$ represents the number of stem cells expressing the phenotype d G6PD at time t.

The transition probabilities in a short time interval $(t, t+h)$ for the hidden component of the model introduced above are:

$$\begin{array}{lcl} P(Z(t+h) = z+1 | Z(t) = z) & = & \lambda z h + o(h) \\ P(Z(t+h) = z-1, X(t+h) = x+1 | Z(t) = z, X(t) = x) & = & \nu z h + o(h) \\ P(X(t+h) = x-1 | X(t) = x) & = & \mu x h + o(h). \end{array} \tag{1}$$

From these probabilities it follows that the distribution of the time to the next event is an exponential with rate $Z(t)(\lambda+\nu) + X(t)\mu$, and the probability that the event is a birth is $Z(t)\lambda / [Z(t)(\lambda+\nu) + X(t)\mu]$, or an emigration is $Z(t)\nu / [Z(t)(\lambda+\nu) + X(t)\mu]$ or a death is $X(t)\mu / [Z(t)(\lambda+\nu) + X(t)\mu]$.

The likelihood function for the process represented in figure 1 when it is observed continuously over a fixed interval of time $[0, T]$ and with $Z(0) = z_0$ and $X(0) = x_0$ can be written in the following way (the likelihood function is similar when considering the two dimensions d and G).

$$\begin{array}{rcl} L(\lambda, \nu, \mu) & = & p(w_{[0,T]} | \lambda, \nu, \mu) \\ & = & \left(\prod_{k=0}^{n-1} \frac{v_k \eta_k}{z_k(\lambda+\nu) + x_k \mu} \times [z_k(\lambda+\nu) + x_k\mu] \exp\{-[z_k(\lambda+\nu) + x_k\mu] t_k\} \right) \\ & & \times \exp\{-[z_n(\lambda+\nu) + x_n\mu] t_n\}, \end{array} \tag{2}$$

where $\eta_k = \lambda$ or ν or μ depending on whether the kth event is a birth, or an emigration or a death. $v_k = z_k$ or x_k depending on whether the kth event happened in the first or second compartment respectively. In essence, the likelihood function is given by the product of the probability of the n events that happened in the interval $[0, T]$, where the probability of a given event is equal to the product of the probability that the next event is either a birth, or an emigration or a death; times the distribution of the inter-arrival time between the previous event and the considered one (Basawa and Rao, 1980).

If we denote with B_T, E_T and D_T the number of births, the number of emigrations from the first to the second compartment and the number of deaths observed in the time interval $[0, T]$ and with $S_T^z = \sum_{k=0}^{n} z_k t_k$, $S_T^x = \sum_{k=0}^{n} x_k t_k$ the total time lived by the population in the first and second compartment respectively, then it can be seen that the likelihood function is proportional to

$$L(\lambda, \nu, \mu) \propto \lambda^{B_T} \nu^{E_T} \mu^{D_T} \exp\{-(\lambda+\nu) S_T^z - \mu S_T^x\}, \tag{3}$$

The likelihood is of exponential form and $(B_T, E_T, D_T, S_T^z, S_T^x)$ is the minimal sufficient statistic (Keiding, 1975). Therefore the maximum likelihood estimators for λ, ν and μ are $\hat{\lambda} = \frac{B_T}{S_T^z}$, $\hat{\nu} = \frac{E_T}{S_T^z}$ and $\hat{\mu} = \frac{D_T}{S_T^x}$.

It follows that estimation in a two-compartment model that has been observed continuously over a window of time $[0, T]$ is just a straightforward extension of the theory for continuously observed linear birth-death processes (Keiding, 1975). Similar to the case of the linear birth-death process the maximum

likelihood estimators for λ, ν and μ have some good asymptotic properties, conditional upon non-extinction, both for large initial population z_0 and for long periods of time (Catlin, 1997).

Inference in a two-compartment model becomes complicated when it is not observed continuously in time. In the hematopoiesis setting neither $Z(t)$ nor $X(t)$ can be observed. Instead, we observe $Y(t)$, a probabilistic function of $W(t)$. The observable process $Y(t)$ defines a hidden Markov model (HMM).

In the cat example, the observations $y_i = Y(t_i)$ are collected only at discrete points in time, usually every 2 to 6 weeks. Also, $Y(t)$ is only a function of the states/size of the second compartment $X(t)$. Specifically, if t_i for $i = 0, ..., M$ denotes the observation times, then the observations are assumed Binomial, so that, in a sample of N_i progenitor cells at time t_i, the distribution of the number of cells of d type is

$$[Y(t_i)|W(t_i) = (z_i, x_i)] \sim \text{Binomial}\left(N_i, \frac{x_{di}}{x_{di} + x_{Gi}}\right),$$

where the observations $Y(t_i)$ conditionally on $W(t_i)$ are assumed independent. Given these observations, the goal is to provide an estimate for the three parameters λ, ν and μ.

While the hidden two-compartment stochastic model is a HMM it presents some characteristics that are not shared by the HMMs typically seen in the literature. Usually these models assume that the hidden process is a stationary, irreducible Markov chain with a finite state space, all characteristics that are not satisfied by the two-compartment process.

The determination of the likelihood function or of the posterior distribution for $\theta = (\lambda, \nu, \mu)$ requires a huge integration step,

$$L(\theta) = p(\mathbf{y}|\theta) = \int_{w \in \mathcal{W}} p(\mathbf{y}|w, \theta) p(w|\theta) dw \tag{4}$$

where $\mathcal{W}$ is the space of all possible realizations of a stochastic two-compartment process observed over the interval $[0, T]$. That is we need to integrate over all the possible step functions starting at $w(0)$. Figure 2 shows how the population size of the second compartment might evolve over time. The two step functions (the solid and dashed ones) in figure 2 represent two possible realizations of the second compartment. The integration required by the equation above is actually more involved since we have two compartments and hence we need to integrate over two connected spaces of step functions.

In principle we should be able to determine the likelihood function for θ, $L(\theta)$, using similar algorithms to those used for HMMs (Baum, 1972), like the Forward-Backward algorithm. In practice that algorithm works well for HMMs with a hidden Markov chain that has a small number of states. The two-compartment stochastic process has a countably infinite number of states if there is no bound on the population size of the first compartment.

Another difficulty associated with the use of such an algorithm in this context is that it would require the discretization in time of the two-compartment

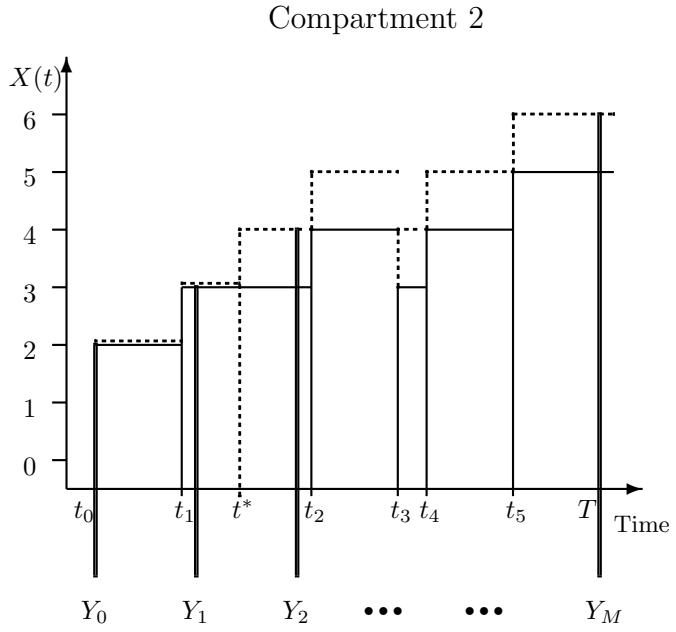

Figure 2: The solid and dashed step functions show how the population size $X(t)$ of the second compartment might evolve over time. The solid step function also represents the current state, while the dashed step function represents the proposed state/realization of the second compartment when an emigration is inserted at time t^*.

process, that is the computation of the transition probabilities from the observation time t_i to t_{i+1}. We have not been able to compute these transition probabilities explicitly, and even if we could, under the assumption that the population size of the first compartment is bounded, such computation would be very expensive and numerically unstable when the observation times t_i are far apart (Guttorp et al., 1990). In addition, since at the start of the experiment the population size of the first compartment is very small, approximations of the transition probabilities, like normal approximations, would be inaccurate (Catlin, 1997).

Two methods have been previously proposed to estimate the HSC rates: a simulation method and a method of moments. The simulation study in Abkowitz et al. (1996) searched the parameter space in a systematic way simulating 100 realizations of the hidden two-compartment process for each parameterization. The simulated data were then compared to the observed cats data with respect to five criteria/characteristics of the real data (such as the range of

the percentage of d cells over the entire observation interval and so on). With this method the data are used rather informally and the chosen criteria are somewhat subjective. The method of moments proposed by Catlin et al. (2001) consists in equating the analytical form of the variance of the observed proportion of d cells to the actual variance at three time points, providing one equation for each of the unknown parameters. This method requires many realizations to obtain good variance estimates and at the same time it discards information from other time points. Therefore the obtained parameter estimates tend to be inefficient.

Stochastic integration methods, such as Markov chain Monte Carlo and reversible jump Markov chain Monte Carlo provide a more complete solution to the inferential problem posed above. Because the two-compartment process is a continuous time Markov chain, any successful integration method for this problem should keep the time structure of the process continuous.

3 Bayesian inference in a hidden stochastic two-compartment model

In this section we introduce an MCMC algorithm for determining the posterior distribution of the unknown parameters λ, ν and μ sampling over $W_{[0,T]}$, the two-compartment process.

The goal is to develop an algorithm that simulates the posterior distribution $p(\theta|\mathbf{y})$ given a prior distribution $p(\theta)$, where $\theta = (\lambda, \nu, \mu)$. Determining $p(\theta|\mathbf{y})$ turns out to be a difficult task, since it requires integrating over all the possible realizations of $W_{[0,T]}$. However it is possible to devise an algorithm to sample from $p(\theta, w_{[0,T]}|\mathbf{y})$. In order to do that we build an irreducible Markov chain on the space $(\theta, w_{[0,T]})$ that has stationary distribution $p(\theta, w_{[0,T]}|\mathbf{y})$. The type of algorithm used is a two-stage algorithm with a Gibbsian outer structure. This means that at every iteration a new θ' is drawn from $p(\theta|w_{[0,T]}, \mathbf{y})$ and a new $w_{[0,T]}$ is sampled from $p(w_{[0,T]}|\theta', \mathbf{y})$ with a Metropolis-Hastings step. We call these two steps the parameter update and state update respectively.

3.1 Parameter update

Given the initial state $w(0)$, that is the population size of both compartments at time zero, the parameter update step turns out to be fully Gibbsian. We will make the assumption that $w(0)$ is known for the sake of describing the algorithm.

In order to perform the parameter update, we determine the full conditional distribution for the unknown parameters $p(\theta|w_{[0,T]}, \mathbf{y})$. If we assume that the parameters $\theta = (\lambda, \nu, \mu)$ a priori are independently Gamma-distributed (which is conjugate in this case) and we maintain the time structure of the hidden

process $W_{[0,T]}$ continuous, then the full conditionals are

$$\begin{array}{rcl} p(\lambda|w_{[0,T]},\mathbf{y}) & \sim & \text{Gamma}(\alpha_\lambda + B_T, \beta_\lambda + S_T^z), \\ p(\nu|w_{[0,T]},\mathbf{y}) & \sim & \text{Gamma}(\alpha_\nu + E_T, \beta_\nu + S_T^z), \\ p(\mu|w_{[0,T]},\mathbf{y}) & \sim & \text{Gamma}(\alpha_\mu + D_T, \beta_\mu + S_T^x). \end{array} \tag{5}$$

In order to get a realization for λ, ν and μ, it is enough to update the prior parameters of the Gamma distributions with the number of births B_T and the total time lived by the population in the first compartment S_T^z for the birth rate for example. We emphasize that the updating of the three rates is easy because we are maintaining the continuous time structure of the two-compartment process. Doing so allows us to compute the sufficient statistic $(B_T, E_T, D_T, S_T^z, S_T^x)$.

Note that we have two populations of cells (d and G) in both compartments, therefore we actually have two sufficient statistics, one for the d and one for the G population. It follows that in the updating for the birth rate, for example, B_T is actually given by the sum of the number of births for the d and for the G population. For simplicity of exposition we are assuming only one population of cells, but all the formulas are easily extended to the case of two populations.

3.2 State update

The state update is not as straightforward as the parameter update. In the observation period $[0, T]$ covered by the observation times, different realizations of $W_{[0,T]}$ can yield different numbers of events: B_T, E_T, and D_T. In order to simulate from $p(w_{[0,T]}|\theta', \mathbf{y})$ a regular Metropolis-Hastings algorithm is not sufficient since the dimension changes. Therefore reversible jump MCMC (RJMCMC) (Green, 1995) is used. The difficulty in implementing RJMCMC resides in the design of a set of moves that defines the proposal distribution.

We introduce an extension of the algorithm proposed by Gibson and Renshaw (1998) for making inference in a SIR model, when the infection times are not observed. We extend this algorithm to make inference in the more complicated model that we described in section 2, where the observations are only collected at discrete points in time. It is worth noticing that the observation/sampling scheme of our model produces observations that contain much less information. In the SIR model the removal times are observed exactly, in our example we do not observe the event times at all.

Before introducing in detail the proposed algorithm, it is helpful to describe how to parameterize $W_{[0,T]}$. Following the notation in Gibson and Renshaw (1998), the space $\mathcal{W}$ where $W_{[0,T]}$ lives can be decomposed in a countable union of subspaces $\mathcal{W}_{l,m,n}$, where $\mathcal{W}_{l,m,n}$ represents the set of all realizations of $W_{[0,T]}$ with l births, m emigrations and n deaths. Therefore, $\mathcal{W}_{l,m,n}$ can be expressed as

$$\begin{array}{rl} \mathcal{W}_{l,m,n} = & \{(t_1^b, ..., t_l^b) | t_j^b \in [0,T], t_j^b < t_{j+1}^b\} \times \\ & \{(t_1^e, ..., t_m^e) | t_j^e \in [0,T], t_j^e < t_{j+1}^e\} \times \\ & \{(t_1^d, ..., t_n^d) | t_j^d \in [0,T], t_j^d < t_{j+1}^d\}, \end{array} \tag{6}$$

where the three components specify the times of occurrence of events of type **B**, **E** and **D** respectively (where **B**, **E** and **D** stand for birth, emigration and death). Given such a representation of the process $W_{[0,T]}$, it becomes clear that we need to develop an MCMC algorithm that builds a Markov chain that moves within spaces of dimension $\mathcal{W}_{l,m,n}$ and between spaces of differing dimensions, e.g.: from $\mathcal{W}_{l,m,n}$ to $\mathcal{W}_{l+1,m,n}$ or to $\mathcal{W}_{l-1,m,n}$. From this observation, it follows that a proposal w' can be obtained by modifying w according to one of the following moves:

1. **Deletion move:** with probability $p_1 > 0$ delete a randomly chosen event (either **B** or **E** or **D**) from w;

2. **Insertion move:** with probability $p_2 > 0$ insert a new event at time t, where $t \sim \text{Uniform}(0, T)$ and the new event is of type **B**, **E** or **D** with equal probability;

3. **Shuffle move:** with probability $p_3 = 1 - p_1 - p_2$ move a randomly chosen event to a new time t, where $t \sim \text{Uniform}(0, T)$.

This set of moves defines a proposal distribution for the candidate state w' conditional on the current state w, $p(w'|w)$. Figure 2 illustrates how the algorithm works. For reasons of space we are showing only the second compartment in figure 2. The proposed state w' (dashed step function), given the current state w (solid step function), is obtained adding an emigration at the time $t^* \sim \text{Uniform}(t_0, T)$. w and w' differ from t^* on. The proposal w' is then accepted with probability $A = \min(1, R)$, where R (the acceptance probability ratio) assumes different forms depending on which move has been proposed (the mathematical details for computing R can be found in the Appendix).

4 Simulation results

In this section we discuss the implementation of the algorithm described in the previous section and analyze its performance on a simulated data set.

Figure 3 shows a simulated realization of the hidden stochastic two-compartment process. The process was simulated over a window of time of 100 weeks. We collected an observation every 2 weeks for a total of $M = 50$ observations. The top two plots show the evolution in continuous time of the d and G populations of cells in the first and second compartment. The left plot in the second row shows the d and G population sizes if we could observe the second compartment fully but only at discrete points in time. The right plot shows the true percentage of d cells in the second compartment at discrete points in time. Unfortunately, in reality we cannot observe any of the above. Instead our observations consist of Binomial samples of size $N = 70$ from the second compartment with probability of success given by the percentages given in the right plot of the second row. A set of these observations are shown in the bottom left plot of figure 3. The right bottom plot shows the observed percentage of d cells in the Binomial samples.

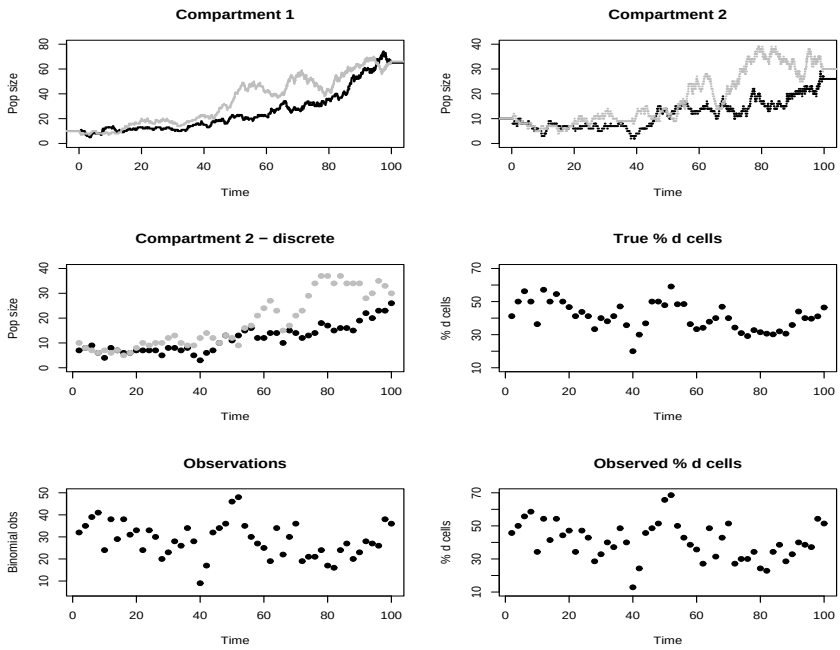

Figure 3: A simulated realization of the hidden stochastic two-compartment process. First row: the two plots show the evolution in continuous time of the d (black line) and G (grey line) populations of cells in the first and second compartments. Second row: the left plot shows the d and G population sizes if we could observe the second compartment fully but only at discrete points in time. The right plot shows the true percentage of d cells in the second compartment at discrete points in time. Third row: the left plot shows the Binomial observations, the right plot shows the observed percentage of the d cells in the Binomial samples.

The simulated process was obtained setting $\lambda = 0.1$, $\nu = 0.08$ and $\mu = 0.15$ and $w(0) = \{(z_d(0) = 10, z_G(0) = 10), (x_d(0) = 10, x_G(0) = 10)\}$. Both the observation scheme (i.e. the length of the observation interval and the number of observations in such interval) and the parameter values used in the simulation reflect the observation scheme adopted for the real cats and the likely values for the three rates. Given the observations in the bottom row of figure 3 and the value of $w(0)$ we would like to determine the posterior distribution of $\theta = (\lambda, \nu, \mu)$. In order to assess the sensitivity of the results to the prior distribution we adopted two prior distributions. In the first set of runs, we assumed that a priori all three parameters are independent and Gamma(5, 50) distributed. For the second set of runs we assumed that all three parameters a priori were uniformly distributed over the interval (0.0, 0.3).

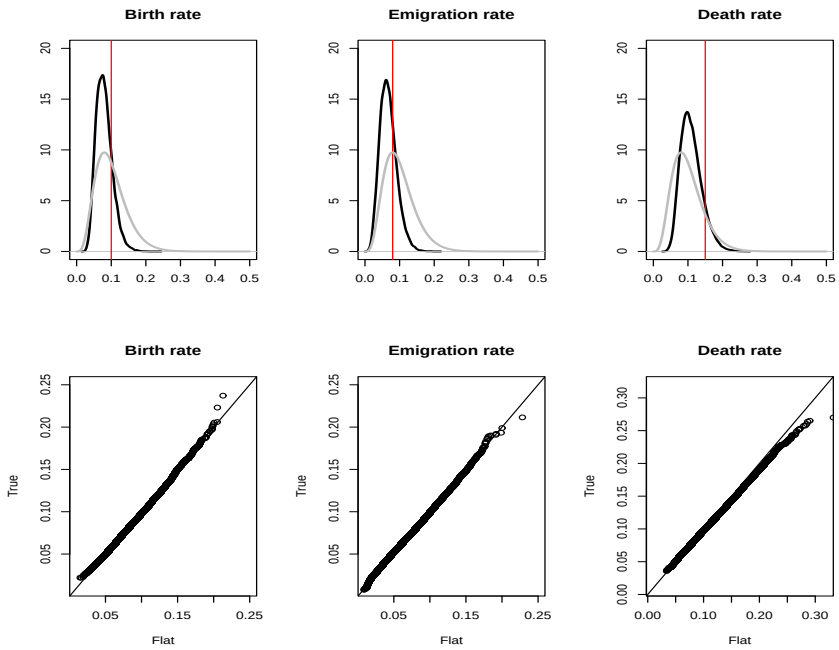

Figure 4: An estimate of the posterior density when using a Gamma(5, 50) prior. Top row: the prior distribution is superimposed. The vertical line represents the true parameter value. Bottom row: qqplot for two runs of the algorithm from two different starting values.

The top three plots of figure 4 show an estimate for the posterior distribution of λ, ν and μ, when the Gamma prior is adopted. The vertical line represents the true parameter value and the curve represents the prior distribution.

The three bottom plots of figure 4 show the qqplots for two runs of the algorithm. We ran the algorithm from two different starting values in order to check whether convergence occurred. It is important to note that what really matters is the starting value for the hidden process and not for the parameters. For these two plots the algorithm was run from the true hidden sequence (vertical axis) and from a degenerate two-compartment process (horizontal axis), i.e. a process that never jumps. The qqplots look extremely good showing that that there is a strong agreement between the two runs.

Figure 5 shows an estimate for the posterior distribution of λ, ν and μ, when the Uniform prior is adopted. The three bottom plots of Figure 5 show the qqplots for two runs of the algorithm. We used the same starting values used in figure 4. Again the qqplot look quite good, indicating that there is agreement between the two runs.

Looking at figure 4 and figure 5 we see that the shape of the posterior

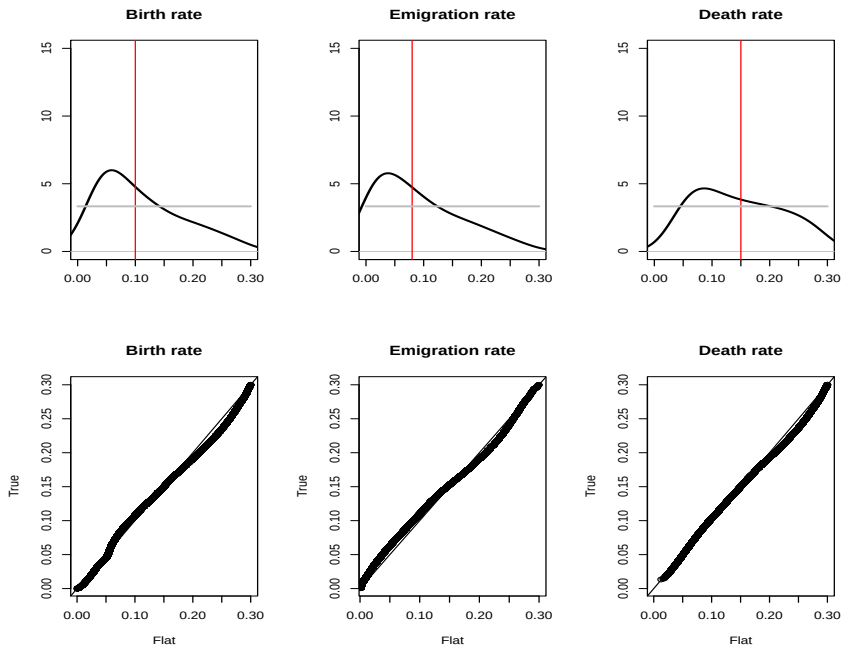

Figure 5: An estimate of the posterior density when using a Uniform(0.0, 0.3). Top row: the prior distribution is superimposed. The vertical line represents the true parameter value. Bottom row: qqplot for two runs of the algorithm from two different starting values.

distribution seem to depend on the adopted prior distribution. However, as we will discuss in the next section, this is mostly due to the small amount of information contained in the observations. The problem arises because a small number of observations are collected over a relatively long period of time and also because only one realization/cat is used to estimate the three rates. In the next section we will see that when we combine the data from multiple cats the sensitivity of the results to the prior distribution disappears.

Table 1 shows the sensitivity of the results to the two different adopted priors. It reports the posterior mean and standard deviation for the three parameters under the two priors and it also reports the 95% highest posterior density interval (HPD) and the posterior mode. The posterior means and standard deviations obtained with the uniform prior tend to be larger than the ones obtained with the gamma prior. Similarly the 95% HPD intervals are wider with the uniform prior. However, the difference between the posterior modes obtained using the two priors is not as large as the difference observed in the posterior means.

Table 1 shows the sensitivity of the results when the assumed initial value

Table 1: Posterior mean, standard deviation, 95% highest posterior density interval (HPD) and posterior mode for λ, ν and μ obtained with two different priors and different initial values $w(0)$. mle is the maximum likelihood estimate of the parameters if the process was not hidden.

	λ (sd)	ν (sd)	μ (sd)
true value	.100	.080	.150
mle	.097	.078	.143
$z_0 = 10$ $x_0 = 10$			
Uniform prior	.114 (.072)	.094 (.074)	.153 (.076)
95% HPD	(.011,.263)	(.001,.237)	(.037,.289)
Posterior Mode	.060	.038	.086
Gamma prior	.075 (.023)	.061 (.024)	.098 (.030)
95% HPD	(.037,.126)	(.022,.115)	(.057,.173)
Posterior Mode	.075	.060	.098
Gamma prior			
$z_0 = 5$ $x_0 = 10$	.087 (.025)	.065 (.024)	.095 (.030)
95% HPD	(.041,.139)	(.022,.113)	(.043,.155)
Posterior Mode	.079	.056	.082
$z_0 = 5$ $x_0 = 5$	.085 (.025)	.066 (.024)	.087 (.028)
95% HPD	(.039,.135)	(.022,.115)	(.038,.144)
Posterior Mode	.077	.057	.075
$z_0 = 15$ $x_0 = 10$	.076 (.024)	.069 (.025)	.122(.034)
95% HPD	(.034,.124)	(.023,.121)	(.062,.190)
Posterior Mode	.067	.062	.106

$w(0)$ differs from the "true" one (i.e. the one used for the simulation). We simulated the data with $w(0) = \{(z_d(0) = 10, z_G(0) = 10), (x_d(0) = 10, x_G(0) = 10)\}$ and using the gamma prior we performed three additional runs in which we assumed different values for $z(0)$ and/or for $x(0)$. In general different values for $x(0)$ do not seem to affect the parameter estimates. Different initial values for $z(0)$ have a slightly larger impact on the parameter estimates. However, the parameter that seems to be the most affected by the use of the "wrong initial" value is μ. In fact, when the initial value for $z(0) = 5$ and $x(0) = 5$, the 95% HPD interval for μ does not contain the true parameter value, the same is not true for the other two parameters. Catlin et al. (2001) show that the initial value $x(0)$ for the second compartment does not influence the parameter estimates. In addition, Catlin et al. (2001) find that the initial value for the first compartment $z(0)$ is non-identifiable. More precisely larger values of $z(0)$ are associated with larger parameter values. The results in table 1 support this finding. However, as we discuss in the next section, the non-identifiability is not really an issue for the cats study. Because of the experiment that the animals underwent to, we know that at time zero the cats must have a small number of

stem cells. In addition the results in table 1 suggest that small departures from the "true value" of $z(0)$ do not substantially affect the parameters estimates.

The algorithm was run for a total of 500,000,000 iterations. However the parameters were updated only every 200 state updates. The observations were further sub-sampled every 50 iterations. The algorithm was implemented in C++ and run on a Pentium 4 2.2GHz with 1 GB of RAM. Every run took approximately two days. The observations obtained using the algorithm introduced in the previous section tend to be correlated. This is not surprising given the nature of the process and the type of algorithm used to determine the posterior distribution of interest. However, the sampling scheme described above seems to reduce the correlation between the sampled values.

Roberts et al. (2004) proposed an MCMC algorithm for Bayesian inference for stochastic volatility processes, that are modelled by non-Gaussian Ornstein-Uhlenbeck processes. Their algorithm jointly parameterizes the hidden process and the model parameters. Being able to do so noticeably improves the performance of the MCMC algorithm. In particular, it reduces the correlation of successive iterations. Their algorithm is applicable whenever the driving Lévy process of the hidden process is a compound Poisson process. That is, the jump times of the hidden process form a Poisson process with finite rate and the corresponding jump sizes are IID from some distribution. Unfortunately, the jump times of the two-compartment process described in section 2 are not IID, as the probabilities of the three possible events indicate (see page 5). At the moment, the algorithm here proposed represents the best solution at hand to solve the inferential problem that we posed in this paper. In the future, we are hoping to be able to relax the assumption about the compound Poisson process and extend the methodology proposed by Roberts et al. (2004) also to our problem.

5 Analysis of the cat data

Figure 6 shows data for the six cats that underwent the experiment described in section 2. This figure shows the observed proportion of progenitor cells d in samples of bone marrow of varying size N_i, with an average of 60-80 progenitor cells per sample. In general the bone marrow samples were collected every two to six weeks and over intervals of time ranging from $[0, 77]$ to $[0, 100]$ weeks depending on the animal. The number and frequency of the observations varied also from animal to animal, ranging from a minimum of 23 to a maximum of 42 observations in the intervals of time reported above. For all the cats, we assumed that $w(0) = \{(z_d(0) = 10, z_G(0) = 10), (x_d(0) = 5, x_G(0) = 5)\}$. From the previous section, we saw that the parameter estimates are not sensitive to small departures from the true value of $z(0)$, while the initial size of the second compartment $x(0)$ has a negligible effect on the parameter estimates (Catlin et al., 2001). The choice of the value for $w(0)$ and more specifically for $z(0)$ was driven by two reasons. First, after radiating the bone marrow we know that the number of stem cells transplanted at time 0 must be extremely small. Second, research on other animal species shows that there are fewer stem cells

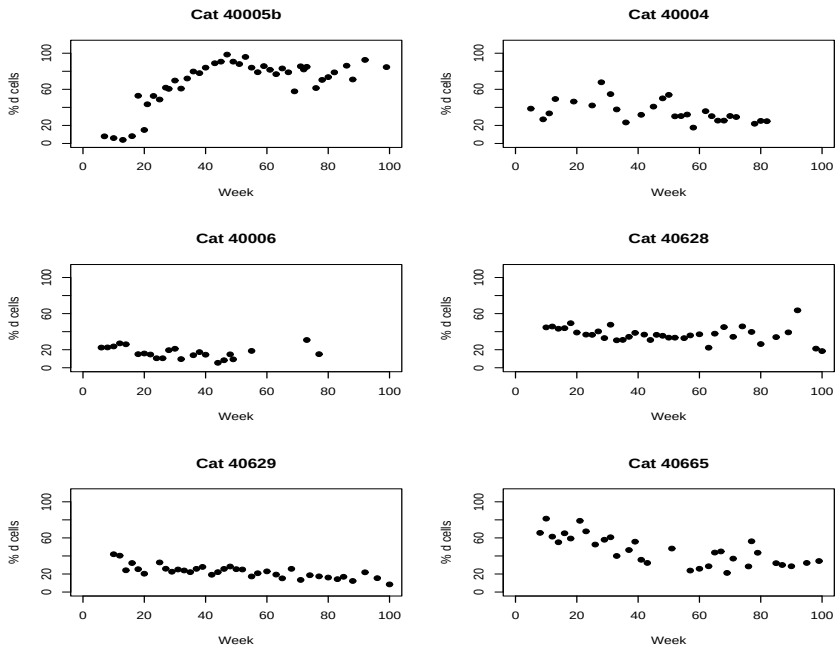

Figure 6: Observed percentage of progenitor cells d for six experimental cats.

in larger animals, such as cats or baboons, than in smaller animals such as mice (Abkowitz et al., 2000). In particular, research shows that both the frequency of the stem cells in the bone marrow and the stem cells replication rate vary inversely with respect to the average size and longevity of the animal species. Abkowitz et al. (2000) find that the average number of transplanted HSC in mice, that underwent a similar experiment as the cats did, is around 35 (i.e. $z(0) = z_d(0) + z_G(0) = 35$). This implies that the number of transplanted HSC in cats should be less than 35.

To check the sensitivity of the results to the prior distribution, we assumed that all three parameters were either Gamma(5,50) or Uniform(0.0,0.5) distributed a priori. The choice of the prior parameters and, hence the range of plausible values for the parameters, was again dictated by research on other animal species. As we indicated above, it is known that feline stem cells undergo replication at a slower rate than mice do. Therefore, the estimates for the murine rates represent upper bounds for the feline rates. In Abkowitz et al. (2000) the authors found that the best estimates for λ and ν are 0.4 and 0.3 respectively. Both adopted priors put a high probability mass on values smaller than 0.4, particularly so for the gamma prior. However, both priors also put some mass above 0.4.

We assumed that the data from the six cats are independent and that the

Table 2: Posterior mean, standard deviation and 95% highest posterior density interval (HPD) for λ, ν and μ obtained with two different priors.

	λ (sd)	ν (sd)	μ (sd)
Gamma prior	.125 (.020)	.104 (.022)	.147 (.026)
95% HPD	(.088,.169)	(.064,.149)	(.101,.201)
Uniform prior	.134 (.020)	.116 (.022)	.155 (.023)
95% HPD	(.098,.173)	(.077,.158)	(.113,.201)

three rates are the same for all the animals. While this last assumption might be in general quite strong, since it is likely that different subjects of the same species have different rates, we think that in this context it is quite appropriate. First, these cats are all female cats of the same species and coming from the mating of a domestic and a Geoffroy wildcat. Therefore, even if not genetically identical, they are likely to share a lot of genetic material. Second, since we have only six experimental cats assuming that the rates for these animals are exchangeable would not produce very different results from the ones obtained with the assumption that we made. Lastly, once we obtained the posterior estimates for the three rates we performed posterior predictive checks (results not shown) that did not contradict this assumption. Therefore, we combined the data from these six cats to obtain the posterior distribution for the three rates λ, ν and μ. From a methodological point of view the extension to the multiple realizations case is straightforward, from a computational point of view it presents several challenges. The extension of the algorithm described in section 3 requires that, at every iteration, the algorithm makes an RJMCMC step for every one of the six animals. These steps integrate over all six hidden two-compartment stochastic processes and compute for every one of the six cats the number of births, emigrations and deaths and total times lived by the population. These statistics update the parameters' full conditional for the parameter update step. This added complexity has consequences on the computing time. The running time for 400,000,000 iterations now takes about 10 days when using the same machine used in the previous section. On the other hand combining data from multiple animals has the advantage of producing better parameter estimates and in general reduces the amount of uncertainty around the parameters. In addition, as the results in table 2 and figure 7 show the sensitivity to the prior practically disappears.

Table 2 reports the posterior mean and the 95% HPD interval for two different runs of the algorithm: one using the Gamma(5,50) prior and one using the Uniform(0.0, 0.5) prior. The posterior means, the posterior standard deviations and the HPD intervals for the three rates are practically identical in the two runs.

Figure 7 shows the posterior distribution for the three rates obtained using the two different priors. The posterior distributions look virtually identical un-

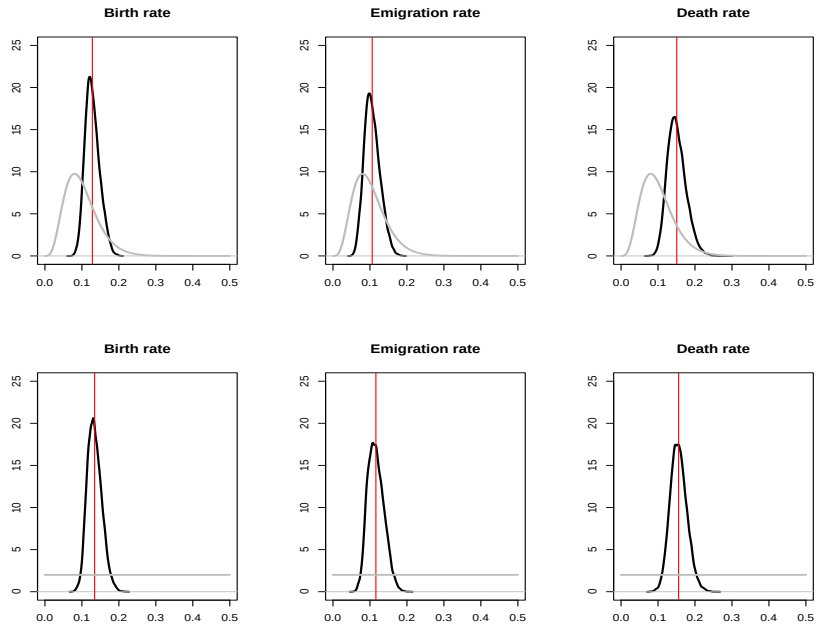

Figure 7: Posterior distribution for the three rates λ, ν and μ. Top row: posterior distribution obtained with a Gamma(5,50) prior. The lighter curve represents the prior density and the vertical line represents the posterior mean. Bottom row: posterior distribution obtained with a Uniform(0.0,0.5) prior.

der the two settings. In both settings, the posterior distribution differs sensibly from the prior, indicating that we have acquired a good amount of information about the 3 unknown rates.

6 Discussion

In this paper we described a hidden two-compartment stochastic process that has been adopted to study feline hematopoiesis. Because of the adopted experimental design that the cats underwent, the process describing the behavior of HSCs must be modeled as a discrete-state, continuous-time, stochastic process. Since the cats undergo radiation and only a small number of HSCs are transplanted back into the animals at the start of the observation period, approximating the hematopoietic process with a continuous state process would not be appropriate. Because of the adopted observation scheme, only a small sample of progenitor cells from the second compartment is collected at discrete points in time, making inference in the proposed model quite challenging.

We proposed a Bayesian approach to estimate the three parameters of interest: HSC replication rate, HSC differentiation rate, and progenitor cells differentiation rate. In particular we implemented a MCMC algorithm that allows us to maintain the continuous time structure of the hematopoietic process and does not require us to approximate the discrete state of the process with a continuous one.

We show with both simulated and real data that the proposed algorithm works quite well in estimating the three quantities of interest. We show that the adopted method is flexible, since it can be used both when data from only one animal are available and when data from multiple cats are available. When we combine the data from the six experimental cats we obtain estimates that are close to the ones obtained by Abkowitz et al. (1996) and Catlin et al. (2001). However, the estimates obtained with our approach are much more precise than the estimates obtained with the other two methods. In general, our approach makes a more efficient use of the data and therefore extracts more information about the parameters of interest from the observations.

Being able to extend the proposed algorithm to the case of multiple realizations has allowed us to fully analyze the available cat data and hence to obtain accurate estimates of the rates of the fates that hematopoietic stem cells undergo. These estimates, with the associated assessment of their variability, represent a huge advancement in the hematopoiesis research after much biological and statistical research.

Appendix

In this appendix we provide the mathematical details about the algorithm presented in section 3. In particular we give the form of and how to compute the acceptance probability ratio under the three different moves that define the proposal distribution.

In section 3.2 we said that the set of moves defining the proposal distribution was given by: a deletion, an insertion and a shuffle move.

If $w \in \mathcal{W}_{l,m,n}$ (i.e. the current realization of the hidden process is such that there are l births, n emigrations and m deaths and therefore a total number of events given by $l+m+n$), then the moves are proposed with probability:

1. **Deletion move:** $p_1/(l+m+n)$ (where p_1 is the probability of proposing a deletion move and $1/(l+m+n)$ is the probability of randomly selecting one of the $l+m+n$ existing events to be deleted);

2. **Insertion move:** $p_2/\{3T\}$ (where p_2 is the probability of proposing a deletion move, $1/3$ is the probably of selecting one of the three types of events to add and $1/T$ is the probability of selecting uniformly a time in the interval $[0,T]$ to insert the new event);

3. **Shuffle move:** $p_3/\{(l+m+n)T\}$ (where p_3 is the probability of proposing a shuffle move; $1/(l+m+n)$ is the probability of randomly selecting one

of the $l+m+n$ existing events to be moved to a different time and $1/T$ is the probability of selecting uniformly a time in the interval $[0,T]$ to move the selected event to).

The proposal w' is then accepted with probability $A = \min(1, R)$, where R assumes different forms depending on which move has been proposed. Move 3 does not alter the dimension of w, therefore R becomes

$$R = \frac{p(w'|\theta,\mathbf{y})p(w|w')}{p(w|\theta,\mathbf{y})p(w'|w)} = \frac{p(w',\theta,\mathbf{y})p(w|w')}{p(w,\theta,\mathbf{y})p(w'|w)} = \frac{p(w'|\theta)p(\mathbf{y}|w',\theta)p(w|w')}{p(w|\theta)p(\mathbf{y}|w,\theta)p(w'|w)}, \quad (7)$$

which can be further simplified since in this case the proposal distribution is symmetric, i.e. $p(w'|w) = p(w|w')$.

Move 1 and 2, instead, alter the dimension of w, so R now takes form

$$R = \frac{p(w'|\theta,\mathbf{y})p(w|w')}{p(w|\theta,\mathbf{y})p(w'|w)}|J|, \quad (8)$$

for an appropriate Jacobian J. It turns out that $|J| = 1$ for both moves 1 and 2, so the acceptance probabilities for all the move types can be calculated from (7).

R can be rewritten as $R = R_1R_2R_3$. Here, R_1 is the ratio of the likelihood for w' over w,

$$R_1 = \frac{p(w'_{[0,T]}|\theta)}{p(w_{[0,T]}|\theta)}, \quad (9)$$

that assumes different forms depending on the move that has been performed. R_2 is the ratio of the likelihood of the data conditional on w' and w respectively,

$$R_2 = \frac{p(\mathbf{y}|w'_{[0,T]},\theta)}{p(\mathbf{y}|w_{[0,T]},\theta)} = \frac{\prod_{i=1}^{M}\left(\frac{x'_{di}}{x'_{di}+x'_{Gi}}\right)^{y_i}\left(\frac{x'_{Gi}}{x'_{di}+x'_{Gi}}\right)^{N_i-y_i}}{\prod_{i=1}^{M}\left(\frac{x_{di}}{x_{di}+x_{Gi}}\right)^{y_i}\left(\frac{x_{Gi}}{x_{di}+x_{Gi}}\right)^{N_i-y_i}} \quad (10)$$

The ratio R_2 actually will have a simpler form since every proposal modifies either population d or G but not both. Also it is the same for all the moves except for those related to the insertion, deletion and shuffling of a birth. In this last case $R_2 = 1$, as the second compartment $x_{[0,T]}$ remains unchanged, this means that the observations are only indirectly linked to birth related moves.

Finally, R_3 is the ratio of the proposal distributions. Again its value depends on the type of move that has been proposed. We report the form of R_1 and R_3 for each of the move types below.

Deletion move

If $w \in R_{l,m,n}$, then the ratio of the proposal distributions is:

$$R_3 = \left(\frac{p_2}{3T}\right) / \left(\frac{p_1}{(l+m+n)}\right).$$

R_1, the ratio of the likelihood of the hidden process, instead, assumes different values whether a birth, an emigration or a death is deleted from the current state w.

- If a birth is deleted, then

$$R_1 = \frac{1}{\lambda}\frac{\prod z'_{j^*}}{\prod z_{j^*}} \exp\left\{-(\lambda+\nu)(S_T^{z'} - S_T^z)\right\} = \frac{1}{\lambda}\frac{\prod z'_{j^*}}{\prod z_{j^*}} \exp\left\{-(\lambda+\nu)(t_i - T)\right\},$$

  where the z''s and z's that are multiplied are those in correspondence to which a birth or an emigration was generated. t_i is the time of the birth to be deleted. Notice, also, that when a birth is added or deleted the second compartment is unchanged, so $R_2 = 1$. We were able to determine a general form for the difference between $S_T^{z'}$ and S_T^z and between $S_T^{x'}$ and S_T^x for the different moves. For example, when a birth is deleted, the difference between the total time lived by the population for the proposed state and the current one for the first compartment can be determined in the following way

$$S_T^{z'} - S_T^z = \left(\sum_{j=0}^{n'} z'_j(t'_{j+1} - t'_j)\right) - \left(\sum_{j=0}^{n} z_j(t_{j+1} - t_j)\right),$$

  observing that $\forall j < i-1$ $z'_j = z_j$ and $t'_{j+1} - t'_j = t_{j+1} - t_j$, while $\forall j > i$ $z'_j = z_j - 1$ and $t'_{j+1} - t'_j = t_{j+1} - t_j$. Also, noticing that $z'_{i-1} = z_{i-1}$ for the interval of time $t_{i+1} - t_{i-1}$, since the jump time t_i has been deleted, we have that

$$S_T^{z'} - S_T^z = (z'_{i-1} - z_i)(t_{i+1} - t_i) - (T - t_{i+1}) = t_i - T,$$

  where $z'_{i-1} - z_i = -1$.

  For the second compartment, instead, the difference $S_T^{x'} - S_T^x = 0$, since in this case the second compartment remains unchanged. These formulas make computations easier and faster, since at every iteration we do not have to recompute $S_T^{z'}$ or $S_T^{x'}$.

- If an emigration is deleted, then

$$\begin{aligned} R_1 &= \frac{1}{\nu}\frac{\prod z'_{j^*}}{\prod z_{j^*}}\frac{\prod x'_{\ell^*}}{\prod x_{\ell^*}} \exp\left\{-(\lambda+\nu)(S_T^{z'} - S_T^z) - \mu(S_T^{x'} - S_T^x)\right\} \\ &= \frac{1}{\nu}\frac{\prod z'_{j^*}}{\prod z_{j^*}}\frac{\prod x'_{\ell^*}}{\prod x_{\ell^*}} \exp\left\{-(\lambda+\nu)(T - t_i) - \mu(t_i - T)\right\}, \end{aligned}$$

  where the x_ℓ's that are multiplied are those that generate a death and where $S_T^{z'} - S_T^z = T - t_i$ and $S_T^{x'} - S_T^x = t_i - T$.

- If a death is deleted, then

$$R_1 = \frac{1}{\mu}\frac{\prod x'_{\ell^*}}{\prod x_{\ell^*}} \exp\left\{-\mu(S_T^{x'} - S_T^x)\right\} = \frac{1}{\mu}\frac{\prod x'_{\ell^*}}{\prod x_{\ell^*}} \exp\left\{-\mu(T - t_i)\right\},$$

where $S_T^{x'} - S_T^x = T - t_i$ and $S_T^{z'} - S_T^z = 0$, since when a death is deleted or added the first compartment remains unchanged.

Insertion move

If $w \in R_{l,m,n}$, then the ratio of the proposal distributions is

$$R_3 = \left(\frac{p_1}{(l+m+n)+1}\right) / \left(\frac{p_2}{3T}\right).$$

R_1 assumes different values whether a birth, an emigration or a death is added to the current state w.

- If a birth is added, then

$$R_1 = \lambda\frac{\prod z'_{j^*}}{\prod z_{j^*}} \exp\left\{-(\lambda+\nu)(S_T^{z'} - S_T^z)\right\} = \lambda\frac{\prod z'_{j^*}}{\prod z_{j^*}} \exp\left\{-(\lambda+\nu)(T - t^*)\right\},$$

where t^* is the time at which the new event, in this case a birth, is inserted. Also $S_T^{z'} - S_T^z = T - t^*$ and $S_T^{x'} - S_T^x = 0$.

- If an emigration is added, then

$$\begin{aligned} R_1 &= \nu\frac{\prod z'_{j^*}}{\prod z_{j^*}}\frac{\prod x'_{\ell^*}}{\prod x_{\ell^*}} \exp\left\{-(\lambda+\nu)(S_T^{z'} - S_T^z) - \mu(S_T^{x'} - S_T^x)\right\} \\ &= \nu\frac{\prod z'_{j^*}}{\prod z_{j^*}}\frac{\prod x'_{\ell^*}}{\prod x_{\ell^*}} \exp\left\{-(\lambda+\nu)(t^* - T) - \mu(T - t^*)\right\}. \end{aligned}$$

Here, $S_T^{z'} - S_T^z = T - t^*$ and $S_T^{x'} - S_T^x = T - t^*$.

- If a death is added, then

$$R_1 = \mu\frac{\prod x'_{\ell^*}}{\prod x_{\ell^*}} \exp\left\{-\mu(S_T^{x'} - S_T^x)\right\} = \mu\frac{\prod x'_{\ell^*}}{\prod x_{\ell^*}} \exp\left\{-\mu(t^* - T)\right\},$$

where $S_T^{x'} - S_T^x = t^* - T$ and $S_T^{z'} - S_T^z = 0$.

Shuffle move

If $w \in R_{l,m,n}$, then the ratio of the proposal distributions is:

$$R_3 = \left(\frac{p_3}{(l+m+n)}\right) / \left(\frac{p_3}{(l+m+n)}\right) = 1.$$

R_1 assumes different values whether a birth, an emigration or a death is moved from the time t_i at which it occurred in the current state w to a new time t^*.

- If a birth is moved to a new time t^*, then

$$R_1 = \frac{\prod z'_{j^*}}{\prod z_{j^*}} \exp\left\{-(\lambda+\nu)(S_T^{z'} - S_T^z)\right\} = \frac{\prod z'_{j^*}}{\prod z_{j^*}} \exp\left\{-(\lambda+\nu)(t_i - t^*)\right\},$$

where $S_T^{z'} - S_T^z = t_i - t^*$ and $S_T^{x'} - S_T^x = 0$. Notice that $S_T^{z'} - S_T^z$ is always equal to $t_i - t^*$ whether $t_i > t^*$ or $t_i < t^*$.

- If an emigration is moved to a new time t^*, then

$$\begin{aligned} R_1 &= \frac{\prod z'_{j^*}}{\prod z_{j^*}} \frac{\prod x'_{\ell^*}}{\prod x_{\ell^*}} \exp\left\{-(\lambda+\nu)(S_T^{z'} - S_T^z) - \mu(S_T^{x'} - S_T^x)\right\} \\ &= \frac{\prod z'_{j^*}}{\prod z_{j^*}} \frac{\prod x'_{\ell^*}}{\prod x_{\ell^*}} \exp\left\{-(\lambda+\nu)(t^* - t_i) - \mu(t_i - t^*)\right\}, \end{aligned}$$

where $S_T^{z'} - S_T^z = t^* - t_i$, while $S_T^{x'} - S_T^x = t_i - t^*$. The observation made above about the difference $S_T^{z'} - S_T^z$ holds also for the difference $S_T^{x'} - S_T^x$, which is always equal to $t_i - t^*$ no matter $t_i > t^*$ or $t_i < t^*$.

- If a death is moved to a new time t^*, then

$$R_1 = \frac{\prod x'_{\ell^*}}{\prod x_{\ell^*}} \exp\left\{-\mu(S_T^{x'} - S_T^x)\right\} = \frac{\prod x'_{\ell^*}}{\prod x_{\ell^*}} \exp\left\{-\mu(t^* - t_i)\right\},$$

where $S_T^{x'} - S_T^x = t^* - t_i$, while $S_T^{z'} - S_T^z = 0$.

References

Abkowitz, J., S. Catlin, and P. Guttorp (1996). Evidence that hematopoiesis may be a stochastic process in vivo. *Nature Medicine 2*, 190–197.

Abkowitz, J., D. Golinelli, D. Harrison, and P. Guttorp (2000). The in vivo kinetics of murine hematopoietic stem cells. *Blood 96*, 3399–3405.

Abkowitz, J., M. Linenberger, M. Newton, G. Shelton, R. Ott, and P. Guttorp (1990). Evidence for maintenance of hematopoiesis in a large animal by the sequential activation of stem cell clones. *Proceedings of the national academy of science 87*, 9062–9066.

Abkowitz, J., M. Linenberger, M. Persik, M. Newton, and P. Guttorp (1993). Behavior of feline hematopoietic stem cells years after busulfan exposure. *Blood 82*, 2096–2103.

Abkowitz, J., R. Ott, R. Holly, and J. Adamson (1988). Clonal evolution following chemotherapy-induced stem cell depletion in cats heterozygous for glucose-6-phosphate dehydrogenase. *Blood 71*, 1687–1692.

Basawa, I. and B. P. Rao (1980). *Statistical inference for stochastic processes*. Academic Press.

Baum, L. (1972). An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes. *Inequalities 3*, 1–8.

Catlin, S. (1997). *Statistical inference for partially observed Markov population processes*. PhD Thesis (unpublished), University of Washington.

Catlin, S., J. Abkowitz, and P. Guttorp (2001). Statistical inference in a two-compartment model for hematopoiesis. *Biometrics 57*, 546–553.

Gibson, G. and E. Renshaw (1998). Estimating parameters in stochastic compartmental models using Markov chain methods. *IMA Journal of Mathematics Applied in Medicine & Biology 15*, 19–40.

Gilks, W., S. Richardson, and D. Spiegelhalter (1996). *Markov chain Monte Carlo in practice*. Chapman & Hall.

Gravenor, M., M. van Hensboek, and D. Kwlatkoswski (1998). Estimating sequestered parasite population dynamics in cerebral malaria. *PNAS 95*, 7620–7624.

Green, P. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika 82*, 711–732.

Guttorp, P., M. Newton, and J. Abkowitz (1990). A stochastic model for haematopoiesis in cats. *IMA Journal of Mathematics Applied in Medicine & Biology 7*, 125–143.

Keiding, N. (1975). Maximum likelihood estimation in the birth-and-death process. *The Annals of Statistics 3*, 363–372.

Roberts, G., O. Papaspiliopoulos, and P. Dellaportas (2004). Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatility processes. *Journal of the Royal Statistical Society, Series B 66*, 369–393.