Bayesian inference for diffusion driven mixed-effects models

G.A. Whitaker*, A. Golightly*, R.J. Boys* and C. Sherlock†

* School of Mathematics & Statistics, Newcastle University, UK
† Department of Mathematics and Statistics, Lancaster University, UK

g.a.whitaker@ncl.ac.uk (https://www.students.ncl.ac.uk/g.a.whitaker/)

The Problem

- When repeated measurements on a system are made, differences between individuals or experimental units can be incorporated through random effects. Quantification of both system (intrinsic) variation and variation between units leads to a stochastic differential mixed-effects model (SDMEM).
- Unfortunately, analytic intractability of SDEs governing most nonlinear multivariate diffusions can make likelihood-based inference methods problematic.
- We consider a data augmentation approach that adopts an Euler-Maruyama approximation of unavailable transition densities and augments low frequency data with additional time points over which the approximation is satisfactory. The approach is flexible, and is not restricted to reducible diffusions.
- Furthermore, we make use of a novel bridging scheme that allows for observations made sparsely in time, as the process of interest may exhibit nonlinear dynamics between measurement times.
- As is well documented, care must be taken in the design of the MCMC sampler due to dependence between the parameters entering the diffusion coefficient and the latent process. We therefore adapt the reparameterisation technique (known as the modified innovation scheme) to the SDMEM framework.

Stochastic Differential Mixed-effects models

- Consider N experimental units, and associated with each unit i is a continuous-time d-dimensional Itô process $\{X_t^i, t \geq 0\}$ governed by the SDE

$$dX_t^i = \alpha(X_t^i, \theta, b^i)\,dt + \sqrt{\beta(X_t^i, \theta, b^i)}\,dW_t^i, \quad X_0^i = x_0^i, \quad i = 1, \ldots, N. \tag{1}$$

- Here, α is a d-vector of drift functions, the diffusion coefficient β is a $d \times d$ positive definite matrix with a square root representation $\sqrt{\beta}$ such that $\sqrt{\beta}\sqrt{\beta}^T = \beta$ and W_t^i is a d-vector of (uncorrelated) standard Brownian motion processes.
- $\theta = (\theta_1, \ldots, \theta_p)^T$ is common to all units whereas $b^i = (b_1^i, \ldots, b_q^i)^T$, $i = 1, \ldots, N$, are unit-specific effects, which may be fixed or random.
- We assume that each experimental unit cannot be observed exactly, but observations $y^i = (y_{t_0}^i, y_{t_1}^i, \ldots, y_{t_n}^i)^T$ are available and these are conditionally independent (given the latent process). We link the observations to the latent process via

$$Y_t^i = F^T X_t^i + \epsilon_t, \qquad \epsilon_t | \Sigma \overset{indep}{\sim} N(0, \Sigma). \tag{2}$$

- Together (1) and (2) completely specify the stochastic differential mixed-effects model.
- Since, in general, the form of the SDE in (1) will not permit an analytic solution, we work with the Euler-Maruyama approximation over a small interval of length $\Delta\tau$, obtained by partitioning $[t_j, t_{j+1}]$ as $t_j = \tau_{j,0} < \tau_{j,1} < \tau_{j,2} < \ldots < \tau_{j,m-1} < \tau_{j,m} = t_{j+1}$. Such an approach introduces $m - 1$ intermediate time points.

Bayesian Inference

- The joint posterior is given by

$$\pi(\theta, \psi, \Sigma, b, x | y) \propto \pi(\theta)\pi(\psi)\pi(\Sigma)\pi(b|\psi)\pi(x|\theta, b)\pi(y|x, \Sigma), \tag{3}$$

where

$$\pi(x|\theta, b) = \prod_{i=1}^{N}\prod_{j=0}^{n-1}\prod_{k=1}^{m} \pi(x_{\tau_{j,k}}^i | x_{\tau_{j,k-1}}^i, \theta, b^i), \qquad \pi(y|x, \Sigma) = \prod_{i=1}^{N}\prod_{j=0}^{n} \pi(y_{t_j}^i | x_{t_j}^i, \Sigma),$$

$$\pi(x_{\tau_{j,k}}^i | x_{\tau_{j,k-1}}^i, \theta, b^i) = N\left(x_{\tau_{j,k}}^i;\, x_{\tau_{j,k-1}}^i + \alpha(x_{\tau_{j,k-1}}^i \theta, b^i)\Delta\tau,\ \beta(x_{\tau_{j,k-1}}^i, \theta, b^i)\Delta\tau\right)$$

and $\pi(y_{t_j}^i | x_{t_j}^i, \Sigma) = N(y_{t_j}^i;\, x_{t_j}^i, \Sigma)$.

- Given the intractability of (3) we construct a Gibbs sampler which generates realisations from this posterior.
- We use the modified innovation scheme to update θ and b to negate the dependence issue.
- Parameters are updated conditional on an innovation process $\{Z_t, t \geq 0\}$ with unit diffusion coefficient, thereby side-stepping the problem of dependence between (θ, b^i) and x. This sampler is known as the innovation scheme.
- The innovation process can be determined using the Brownian increments of the Euler-Maruyama approximation. This is likely to be inefficient and we therefore use the Brownian increments driving the modified diffusion bridge Durham & Gallant (2002) as the innovation process. The resulting sampler is known as the modified innovation scheme.

Improved Bridge Construct

- The form of (3) suggests a scheme where unit-specific paths are updated separately. We update x in overlapping blocks of size $2m + 1$.
- We require the ability to (approximately) generate a discrete-time realisation of a diffusion process between two time points at which the process is either observed exactly or subject to Gaussian noise. The resulting trajectory is typically referred to as a diffusion bridge.
- We propose a novel bridge construct that requires no tuning parameters, is simple to implement, computationally efficient and explicitly allows for the effect of the drift governing the target SDE.
- Partition X_t as $X_t = \eta_t + R_t$, where $\{\eta_t, t \in [t_j, t_{j+1}]\}$ is a deterministic process satisfying the ODE

$$\frac{d\eta_t}{dt} = \alpha(\eta_t), \quad \eta_{t_j} = x_{t_j},$$

and $\{R_t, t \in [t_j, t_{j+1}]\}$ is a residual stochastic process satisfying

$$dR_t \equiv dX_t - d\eta_t = \{\alpha(X_t) - \alpha(\eta_t)\}dt + \sqrt{\beta(X_t)}\,dW_t.$$

- Sample $\{R_t, t \in [t_j, t_{j+1}]\}$ using a local linearisation of the Brownian bridge SDE

$$dR_t = \frac{R_{t_{j+1}} - R_t}{t_{j+1} - t}\,dt + \sqrt{\beta(X_t)}\,dW_t.$$

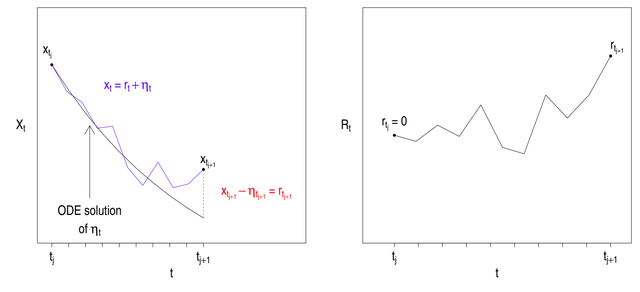


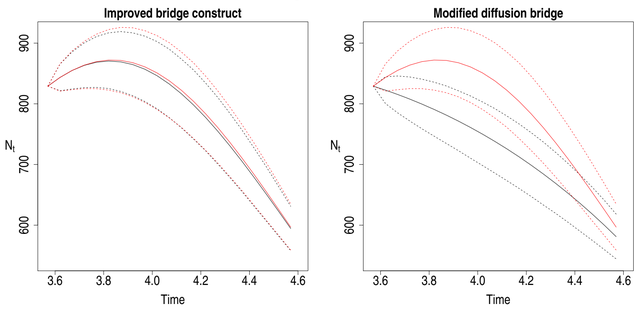
Illustration of the residual bridge.
Left: The full bridge.
Right: A sample path of R_t.

Example: Aphid Growth

- The model concerns both the aphid population size (N_t) and cumulative population size (C_t).
- The data were collected in July 2004 in Lamesa, Texas and consist of twenty-seven treatment-block combinations (Matis et al. 2008).
- We formulate an appropriate SDMEM model for $X_t^{ijk} = (N_t^{ijk}, C_t^{ijk})^T$ with $i, j, k \in \{1, 2, 3\}$ as

$$dX_t^{ijk} = \begin{pmatrix} \lambda^{ijk}N_t^{ijk} - \mu^{ijk}N_t^{ijk}C_t^{ijk} \\ \lambda^{ijk}N_t^{ijk} \end{pmatrix} dt + \begin{pmatrix} \lambda^{ijk}N_t^{ijk} + \mu^{ijk}N_t^{ijk}C_t^{ijk} & \lambda^{ijk}N_t^{ijk} \\ \lambda^{ijk}N_t^{ijk} & \lambda^{ijk}N_t^{ijk} \end{pmatrix}^{1/2} dW_t^{ijk}.$$

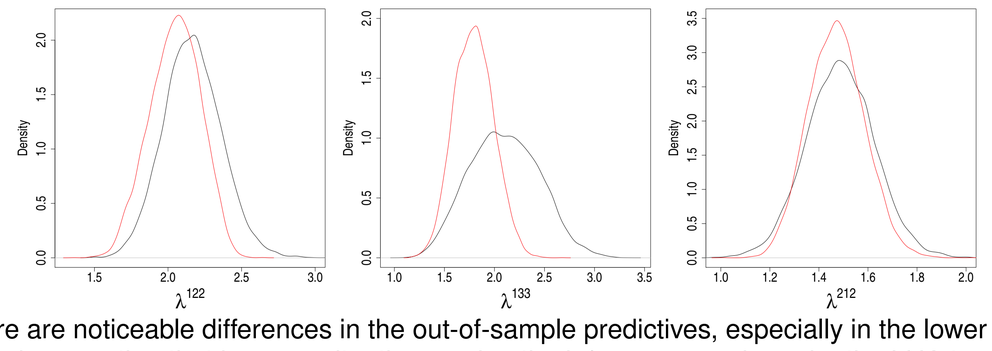
- Do we need the improved bridge construct? YES!



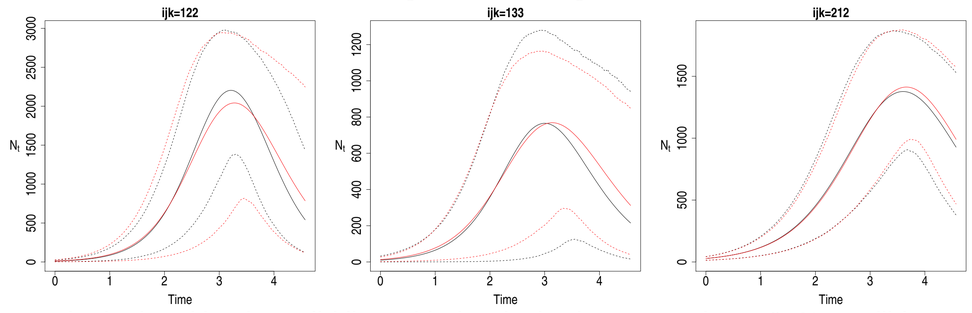
95% credible region (dashed line) and mean (solid line) of the true conditioned aphid population (red) and two competing bridge constructs (black).

For a full discussion of different bridging methods see Whitaker et al. (2016*b*).

- Marginal posterior densities of the overall birth rates (λ^{ijk}). We see distinct differences between posteriors obtained under the Bayesian imputation approach (black) and a linear noise approximation (LNA) based approach (red) (see Fearnhead et al. (2014)). The posteriors displayed are indicative of those obtained for all treatment combinations. Moreover, similar patterns are evident in the overall death rates (μ^{ijk}).



- There are noticeable differences in the out-of-sample predictives, especially in the lower credible bound suggesting that in some situations, using the inferences made under the LNA to predict the outcome of future experiments can give misleading results.



The mean is depicted by the solid line with the dashed representing a 95% credible region. Bayesian imputation (black). LNA (red).

- See Whitaker et al. (2016*a*) for a full analysis of this dataset.

Summary and References

- We have provided a framework that permits (simulation-based) Bayesian inference for a large class of multivariate SDMEMs.
- By adopting a Bayesian imputation approach, we have shown how the modified innovation scheme, which is necessary for overcoming the problematic dependence between the latent process and any parameters that enter the diffusion coefficient, can be applied to SDMEMs.
- Fundamental to our approach is the development of a novel bridge construct that can be used to sample a discretisation of a conditioned diffusion process, and does not break down when the process exhibits strong nonlinearity.
- We fitted the diffusion approximation of a Markov jump process description of aphid dynamics. Whilst both the imputation and LNA approaches provided a reasonable fit to the aphid data, differences were found between the parameter posteriors, leading to differences in the out-of-sample predictive distributions.

Durham, G. B. & Gallant, A. R. (2002), 'Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes', *Journal of Business and Economic Statistics* **20**, 297–338.

Fearnhead, P., Giagos, V. & Sherlock, C. (2014), 'Inference for reaction networks using the linear noise approximation', *Biometrics* **70**(2), 457–466.

Matis, T. I., Parajulee, M. N., Matis, J. H. & Shrestha, R. B. (2008), 'A mechanistic model based analysis of cotton aphid population dynamics data', *Agricultural and Forest Entomology* **10**(4), 355–362.

Whitaker, G. A., Golightly, A., Boys, R. J. & Sherlock, C. (2016*a*), 'Bayesian inference for diffusion driven mixed-effects models', *Bayesian Analysis*. To appear.

Whitaker, G. A., Golightly, A., Boys, R. J. & Sherlock, C. (2016*b*), 'Improved bridge constructs for stochastic differential equations', *Statistics and Computing*. To appear.