

## Bayesian Deconvolution Analysis of Pulsatile Hormone Concentration Profiles

Timothy D. Johnson

Department of Biostatistics, School of Public Health, University of Michigan,  
Ann Arbor, Michigan 48109, U.S.A.  
*email:* tdjtdj@umich.edu

**SUMMARY.** Many hormones are secreted into the circulatory system in a pulsatile manner and are cleared exponentially. The most common method of analyzing these systems is to deconvolve the hormone concentration into a secretion function and a clearance function. Accurate estimation of the model parameters depends on the number and location of the secretion pulses. To date, deconvolution analysis assumes the number and approximate location of these pulses are known *a priori*. In this article, we present a novel Bayesian approach to deconvolution that jointly models the number of pulses along with all other model parameters. Our method stochastically searches for the secretion pulses. This is accomplished by viewing the set of parameters that define the pulses as a point process. Pulses are determined by a birth-death process which is embedded in Markov chain Monte Carlo algorithm. This idea originated with Stephens (2000, *Annals of Statistics* **28**, 40–74) in the context of finite mixture model density estimation, where the number of mixture components is unknown. There are several advantages that our model enjoys over the traditional frequentist approaches. These advantages are highlighted with four datasets consisting of serum concentration levels of luteinizing hormone obtained from ovariectomized ewes.

**KEY WORDS:** Bayesian analysis; Birth-death process; Deconvolution analysis; Markov chain Monte Carlo; Mixture model; Model choice.

### 1. Introduction

Serum concentration levels of many hormones, such as luteinizing hormone, show distinct pulsatile patterns (see Figure 3). Clearance of the hormone, via mechanisms such as cellular binding, enzymatic cleavage, and glomerular filtration, has typically been modeled using the one- or two-compartment pharmacokinetic model for elimination (i.e., exponential or biexponential decay) and it has been shown that these elimination models fit experimental data (e.g., Veldhuis et al., 1986).

In principle, one can reconstruct the serum concentration levels of the hormone over time through the convolution of the (known) secretion and clearance functions. Deconvolution is the inverse problem: determine the parameters of the functional forms of the secretion events and the hormonal clearance given the concentration. Many aspects are of interest to the endocrinologist, including pulse frequency, the average hormone mass released per event, and the hormone clearance rate. Reasonable estimates of these parameters in healthy individuals can assist investigators in determining causes of disease processes that affect hormone levels and secretion patterns.

Deconvolution of serum concentration levels of hormone into a secretory input function and a clearance function has been used for over a decade since its introduction by Veldhuis, Carlson, and Johnson (1987) (see also Veldhuis and Johnson, 1992). Veldhuis et al. (1987) use nonlinear least squares

to perform the deconvolution conditional on the number and locations of the secretory inputs (a separate program, or prior knowledge, must be used to determine the number and approximate location of secretion events). Alternatively, one could fit a small subset of models with a different number of secretion events and use generalized cross-validation or Akaike's information criteria (AIC) to select the "best" model, as do O'Sullivan and O'Sullivan (1988).

Several other model-based approaches for the analysis of pulsatile hormone data have been proposed, including works by O'Sullivan and O'Sullivan (1988), Diggle and Zeger (1989), Kushler and Brown (1991), Komaki (1993), Keenan, Veldhuis, and Yang (1998), and Guo, Wang, and Brown (1999). However, these have not been as extensively used in the field of endocrinology as deconvolution—a five-minute literature search revealed over 100 published journal articles using deconvolution in the analysis of pulsatile hormone data. We do not claim that the deconvolution model is superior to these other models, but only that a method that jointly models the number (and location) of secretion pulses and the other parameters of the deconvolution model would make a significant contribution to the analysis of pulsatile hormone data and, in particular, to the deconvolution model. (Through personal communication, Dr Veldhuis, arguably the world's preeminent authority on neurohormonal secretion, agrees with this conclusion.)

We present a novel Bayesian method for jointly modeling parameters of the deconvolution model and the number and location of the pulsatile secretion events. The idea is to stochastically search for the secretion events, embedding the search in a Bayesian analysis of the deconvolution model, conditional on the number of events, via Markov chain Monte Carlo (MCMC). The stochastic search is performed using a birth-death process proposed by Stephens (2000a), who uses this methodology in mixture models with an unknown number of components. The method in Stephens (2000a), birth-death MCMC (BDMCMC), is an alternative to the reversible jump Markov chain Monte Carlo (RJCMCMC) algorithm developed by Green (1995) and extended to mixture models by Richardson and Green (1997). BDMCMC was chosen over RJCMCMC because of its ease of implementation relative to RJCMCMC.

Our approach to this problem offers several advantages over the traditional deconvolution approach. First, it provides a way to select the “best” model out of a much larger set of models than the typically few competing models using generalized cross-validation. Second, it provides a coherent way to combine results over different models. For example, several competing models may fit the data equally well and we may want to “model average” over a subset of models. Third, we can assign probabilities that secretion events are indeed events and not noise. Fourth, a secretion event does not have to assume a particular parametric form. Our approach allows for secretion events to be a mixture of component functions.

We present our Bayesian deconvolution model with an unknown number of secretion events in Section 2. The birth-death process is introduced in Section 3 and we outline how we simulate from the posterior of the joint model in Section 4. We then demonstrate our method on luteinizing hormone sampled from ewes, in Section 5. A small simulation study is presented in Section 6, and we end the article with a brief discussion in Section 7.

## 2. The Bayesian Model

Suppose the concentration  $C$  of a substrate is the convolution of an input (secretion) function  $S(t)$  and a clearance (elimination) function  $E(t)$ , where the convolution of  $S$  and  $E$  is defined by  $S * E = \int_{-\infty}^{\infty} S(v)E(t-v)dv$ .

We model the secretion profile  $S(t)$  as the superposition of  $k$  component functions,  $f_i(t)$ , where  $k$  is given. That is,  $S(t) = \sum_{i=1}^k f_i(t)$ . The component functions all come from a parameterized family and we assume that their parameters can be subdivided into a set of parameters,  $\eta$ , common to every component, and a set of component specific parameters  $\theta_i$ ,  $i=1, \dots, k$ . We set  $\theta = \{\theta_1, \dots, \theta_k\}$ . The elimination function will be parameterized by a set of parameters  $\phi$ . Let  $f(t; \theta_i, \eta)$  denote the  $i$ th component and let  $E(t; \phi)$  denote the clearance function. Then the concentration at time  $t$  is the sum of the convolutions of the  $f(t; \theta_i, \eta)$  and  $E(t; \phi)$  plus a baseline concentration level,  $\mu$ , and is given by

$$C(t; \mu, \theta, \eta, \phi, k) = \mu + \int_{-\infty}^t \left\{ \sum_{i=1}^k f(v; \theta_i, \eta) \right\} E(t-v; \phi) dv \quad (1)$$

$$= \mu + \sum_{i=1}^k \int_{-\infty}^t f(v; \theta_i, \eta) E(t-v; \phi) dv. \quad (2)$$

We digress here for a moment to explain that for the most commonly used deconvolution program (Veldhuis et al., 1987), secretion events take a parametric form, typically Gaussian, and  $k$  represents the number of secretion events. In our representation,  $k$  is the number of component functions that make up the entire secretion profile over time, and  $k$  may be larger than the number of secretion events. A secretion event may be the superposition of two or more component functions. We will see this below in Section 5.

Veldhuis et al. (1987) assumes that the observed values  $y(t)$  are equal to the concentration plus error,  $y(t) = C(t; \mu, \theta, \eta, \phi, k) + \varepsilon_t$ , with  $\varepsilon_t \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma^2)$ . The errors include biological and measurement errors. However, hormone levels are strictly nonnegative, as are the assays that measure the observed concentrations. Hence, an observed value less than zero is impossible. Further, the error in the concentration profile is a combination of several different sources of error, including assay, biological, and dilution errors. A normal, or any symmetric, error structure is inappropriate (Rodbard, Rayford, and Ross, 1970). We therefore assume a multiplicative error structure and model the log of the observed values,  $y(t_i)$ ,  $i=1, \dots, n$  at times  $t_i \in [0, T]$  (the range of the data):

$$\begin{aligned} \ln\{y(t_i)\} &= \ln \left[ \left\{ \mu + \sum_{i=1}^k \int_{-\infty}^{t_i} f(v; \theta_i, \eta) E(t_i-v; \phi) dv \right\} \varepsilon_{t_i} \right] \\ &= \ln\{C(t_i; \mu, \theta, \eta, \phi, k)\} + \ln(\varepsilon_{t_i}), \end{aligned} \quad (3)$$

where  $\ln(\varepsilon_{t_i}) \sim N(0, \sigma^2)$  independently of one another and independently of the log concentration  $\ln\{C(t_i; \mu, \theta, \eta, \phi, k)\}$ . Thus,

$$\ln\{y(t)\} \sim N[\ln\{C(t; \mu, \theta, \eta, \phi, k)\}, \sigma^2]. \quad (4)$$

Alternatively, one may use a constant coefficient of variation model.

We further assume that some of the model parameters depend on a set of hyperparameters  $\omega$ . We factor the conditional prior according to

$$p(\mu, \theta, \eta, \omega, \phi, \sigma^2 | k) = p(\theta | k, \eta, \omega) p(\mu, \phi, \sigma^2, \eta, \omega). \quad (5)$$

## 3. The Birth-Death Process

Thus far, we have specified the likelihood (4) and the conditional prior (5) of our joint model. Now we define the virtual birth-death process that is embedded in the MCMC algorithm. Stephens (2000a) proposed an alternative to the reversible jump MCMC algorithm developed by Green (1995). Stephens (2000a) views the components of a mixture model as point processes in an appropriate parameter space. Then, using the theory of point processes, he shows how to construct a virtual birth-death process, combining it with standard fixed-dimension MCMC updates to create a Markov chain whose stationary distribution is the desired posterior distribution. Stephens (2000a) also shows how the algorithm can be adapted to change-point models and how it can be used for variable selection in regression models. Here, we show how his method can be adapted to deconvolution analysis.

Before proceeding, we point out that our model is not a mixture of distributions. The secretion components are not distributions and the observations are not samples from a mixture of distributions. Rather, one should think of the secretion over time as a superposition of the secretion component functions.

We require that the posterior distribution of  $(k, \theta | \mathbf{y}, \eta, \omega)$  be exchangeable and thus ignore the labeling of the components in  $\theta$ . The posterior will then be a point process on the appropriate parameter space and, given the following conditions on the birth and death rates, we have a continuous time Markov birth-death process with stationary distribution  $p(k, \theta | \mathbf{y}, \eta, \omega)$ . It is essential for the theory developed in Stephens (2000a) that this conditional posterior be exchangeable. For our particular application (see Section 5),  $\theta$  is the collection of the  $k$  component-specific parameters. We believe the exchangeability assumption is reasonable for our application. Secretion of luteinizing hormone is not dependent on time over the 6-hour period in which data were collected, and there is no evidence that secretion events are dependent on past secretion events. On a time scale of a month or so, luteinizing hormone depends on time, due to the menstrual cycle in animals with intact ovaries. However, the data presented in Section 5 were collected on ovariectomized ewes, thus removing the negative feedback loop of the sex steroid hormones released from the ovaries, which change in accord with the menstrual cycle.

Let  $p(k | \eta, \omega)$  be the prior of  $k$  given  $\eta$  and  $\omega$ . We factor the full prior according to

$$p(k, \mu, \theta, \eta, \omega, \phi, \sigma^2) = p(k | \eta, \omega) p(\theta_1, \theta_2, \dots, \theta_k | \eta, \omega) \times p(\mu, \phi, \sigma^2, \eta, \omega). \tag{6}$$

In order for the posterior distribution,  $p(k, \theta | \mathbf{y}, \eta, \omega)$ , to be exchangeable, it is necessary that  $p(\theta_1, \theta_2, \dots, \theta_k | \eta, \omega)$  be invariant with respect to random permutations of the labels for any  $k$ . Henceforth, we assume that this assumption is satisfied.

Births of secretion components will occur at a constant rate  $\lambda_b$  according to density  $b(\theta | \eta, \omega)$ . Deaths of secretion components occur independently of one another. The death rate for component  $j$ , given  $k \geq j$  components currently in the model, is

$$\delta_j = \lambda_b \frac{L_{(-j)} p(k-1 | \eta, \omega)}{L kp(k | \eta, \omega)} \frac{p(\theta_1, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_k | \eta, \omega)}{p(\theta_1, \dots, \theta_{j-1}, \theta_j, \theta_{j+1}, \dots, \theta_k | \eta, \omega)} \times b(\theta_j | \eta, \omega), \tag{7}$$

where  $L$  is the full likelihood, including all  $k$  components, and  $L_{(-j)}$  is the likelihood with the  $j$ th secretion component removed.  $\Delta = \sum_j \delta_j$  is the total death rate. Next, we simulate the waiting time,  $\tau_w$ , to the next birth or death from an exponential distribution with mean  $(\lambda_b + \Delta)^{-1}$ . A birth occurs at  $\tau_w$  with probability  $\lambda_b / (\lambda_b + \Delta)$  and a death occurs with probability  $\Delta / (\lambda_b + \Delta)$ . If a birth is to occur, we draw a point  $\theta$  from  $b(\theta | \eta, \omega)$ . If a death is to occur, we remove a secretion component according to probabilities  $\delta_j / \Delta, j = 1, \dots, k$ . This process is performed iteratively, conditional on  $\eta$  and  $\omega$ . The resulting birth-death process has stationary distribution

$p(k, \theta | \mathbf{y}, \eta, \omega)$ , since the conditions of Theorem 3.1 in Stephens (2000a) are satisfied.

**4. Simulating from the Joint Model**

Given the current state,  $\{k, \theta, \eta, \omega, \mu, \phi, \sigma^2\}^j$ , of the model at iteration  $j$ , simulate a value for  $\{k, \theta, \eta, \omega, \mu, \phi, \sigma^2\}^{j+1}$  at iteration  $j+1$  according to the following steps:

- Step 1:** Sample  $\{k, \theta\}^{j+1}$  by running the virtual birth-death process for a fixed time  $t_0$  with  $\eta$  and  $\omega$  fixed.
- Step 2:** Sample  $\{\mu, \phi, \eta, \omega, \sigma^2\}^{j+1}$  by Gibbs, Metropolis-Hastings, or any other appropriate, fixed-parameter space MCMC algorithm.
- Step 3:** Sample  $\theta^{j+1}$  by the appropriate MCMC algorithm. (Note this step is not necessary for convergence, but improves mixing (Stephens, 2000a).)

The parameters sampled in steps 2 and 3 above may be jointly sampled or sampled one at a time. In either case, they are sampled conditional on all other parameters in the model not currently being sampled.

**5. Application: Luteinizing Hormone in Ewes**

Luteinizing hormone (LH) is a gonadotropic hormone that regulates growth, development, reproductive processes, pubertal maturation, and sex steroid hormone secretion of the gonads in both of the sexes (Berne and Levy, 1993). We analyze the data collected from four ovariectomized ewes. Blood samples were drawn from the jugular vein of the ewes every 5 minutes for approximately 6 hours. The samples were then assayed for LH concentration by radioimmunoassay. (See Midgley et al. (1997) for details.)

We assume that hormone elimination follows exponential decay:

$$E(t; \phi) = e^{-\phi t}, \tag{8}$$

where the parameter set  $\phi$  is now a single parameter. The secretion components can, in theory, be any nonnegative function. We assume a Gaussian form as do Veldhuis et al. (1987). Our algorithm is not dependent on this particular form of the secretion components and, under a Bayesian framework, it is not difficult to incorporate any nonnegative function that is deemed appropriate.

Let  $\beta_{j1}$  denote the hormonal mass secreted from component  $j$ ,  $\beta_{j2}$  denote its scale parameter, and  $\tau_j$  denote its center of mass. Continuing with our notation from Section 2, we set  $\theta_j = \{\beta_{j1}, \beta_{j2}, \tau_j\}$  and  $\eta = \{\emptyset\}$ . Further let  $\beta_j = (\beta_{j1}, \beta_{j2})^T$ . Then

$$f(t; \theta_j) = \beta_{j1} \exp\{-0.5\beta_{j2}^{-1}(t - \tau_j)^2\} / \sqrt{2\pi\beta_{j2}}. \tag{9}$$

Substituting (8) and (9) into (1), performing a little algebra and a change of variable, the concentration at time  $t$  can be written as

$$C(t; \mu, \phi, \theta, k) = \mu + 0.5 \sum_{j=1}^k \beta_{j1} \exp\{(\tau_j - t)\phi + 0.5\phi^2\beta_{j2}\} \{1 + \text{erf}(x_j)\} \tag{10}$$

where  $\text{erf}(z) = (2/\pi^{1/2}) \int_0^z \exp(-v^2) dv$  and  $x_j = (t - \tau_j - \phi\beta_{j2}) / (2\beta_{j2})^{1/2}$ . The error function,  $\text{erf}(z)$ , is a special

case of the well-known incomplete gamma function:  $\int_0^x \exp(-t)t^{a-1}dt/\Gamma(a)$  (set  $a=1/2$ , make the change of variable  $z=t^2$ , and let  $x^2=z$ ). This integral can rapidly be approximated (see Press et al. (1990) for details).

### 5.1. Priors and Initial Values

We believe that each component,  $f(t; \theta_j)$ , should be somewhat similar in size and shape, so we adopt a random effects model for the mass and scale of each component. *A priori*, the random effects  $\beta_j$  are given a truncated multivariate student- $t$  distribution with four degrees of freedom,  $t_4(\beta, \Sigma)$ . The distribution is restricted, or truncated, to the first quadrant of two-dimensional Euclidean space,  $\mathbb{R}^{2+} = \{(x, y)^T : x > 0, y > 0, x, y \in \mathbb{R}\}$ . We chose the  $t_4$  distribution as a robust alternative to the normal distribution. The multivariate  $t_s$  distribution can be obtained as a scale mixture of the multivariate normal distribution (Kleinman and Ibrahim, 1998). If  $x | \kappa \sim N(\mu, \kappa^{-1}\Sigma)$  and  $\kappa \sim \text{Gamma}(r/2, r/2)$ , then the marginal distribution of  $x$  is the multivariate  $t_r$ .

Let  $\omega = \{\beta, \Sigma, \kappa\}$  with  $\kappa = \{\kappa_1, \dots, \kappa_k\}$ . We further decompose our prior (6) into:

$$p(k)p(\mu)p(\phi)p(\sigma^2)p(\beta)p(\Sigma)p(\tau_1, \dots, \tau_k | k) \\ \times \prod_{j=1}^k p(\beta_j | \beta, \Sigma, \kappa)p(\kappa_j).$$

In this section, we set up two prior models on the number of components,  $k$ , their locations,  $\tau_j$ , and on the random effects prior precision matrix,  $\Sigma^{-1}$ . These priors are displayed side by side in Table 1. The first prior model reflects the belief that each secretion event is Gaussian-shaped and therefore there is a one-to-one correspondence between component functions and secretion events. The prior random effects covariance structure is diagonal (*a priori* independent random effects) with relatively small common variance, 5, which reflects our belief that each event should be similar in size. The prior on  $k$  is Poisson with a mean of 10, which reflects the belief that there are about 10 distinct secretion events. We also believe that the events are somewhat spread apart in time and more or less periodic. Thus, our prior for the event locations, given  $k$ , are the  $(3j)$ th,  $j=1, \dots, k$ , order statistics, randomly permuted, from  $3k+2$  i.i.d. uniform random variates over the interval of interest,  $(-40, T+10)$ . For both prior models we extended our search for secretion components outside the range of the data (by 40 minutes prior to data col-

lection and 10 minutes past data collection) so that we can detect a secretion event whose center occurs before collection of data or whose center occurs after collection of data. Let  $p(\tau_{(1)}, \dots, \tau_{(k)} | k)$  denote the conditional density of these order statistics; then the secretion event locations have joint density

$$p(\tau_1, \dots, \tau_k | k) = \frac{1}{k!} p(\tau_{(1)}, \dots, \tau_{(k)} | k) \\ = \frac{1}{k!} \frac{(3k+2)!}{2^{k+1}(T+50)^{2(k+1)}} \\ \times (\tau_{(1)}+40)^2 (\tau_{(2)}-\tau_{(1)})^2 \dots (T+10-\tau_{(k)})^2,$$

which is invariant to any relabeling of the components, as required.

The second prior we consider reflects a belief that secretion events may not be Gaussian-shaped and, hence, we allow each event to be a superposition of an unknown number of component functions. The random effects variance structure is again diagonal; however, we increased the common variance from 5 to 25. The prior on  $k$  is now uniform on the integers from 1 to 60, and the prior on the component function locations is now i.i.d.  $U(-40, T+10)$ . With this i.i.d. uniform prior on the component function locations, the conditional posterior distribution of  $(k, \theta)$  is exchangeable, as the theory requires.

All remaining parameters are given priors that are common to both prior models. Let  $b = (15, 15)^T$ ,  $F = \text{diag}(10,000, 10,000)$ . We adopt the following heirarchical structure for our priors on the random effects:

$$p(\beta_j | \beta, \Sigma, \kappa) \propto |\kappa_j^{-1}\Sigma|^{-1/2} \\ \times \exp\{-0.5\kappa_j(\beta_j - \beta)^T \Sigma^{-1}(\beta_j - \beta)\} \mathbb{I}_{\mathbb{R}^{2+}}(\beta_j), \\ p(\beta) \propto |F|^{-1/2} \exp\{-0.5(\beta - b)^T F^{-1}(\beta - b)\} \mathbb{I}_{\mathbb{R}^{2+}}(\beta), \\ \kappa_j \sim \text{Gamma}(2, 2),$$

where  $\text{Gamma}(\alpha, \beta)$  is the gamma distribution with mean  $\alpha/\beta$  and variance  $\alpha/\beta^2$ , and  $\mathbb{I}_A(x)$  is the indicator function equal to 1 if  $x \in A$  and 0 if  $x \notin A$ . Thus, the  $\beta_j$  are random effects heirarchically centered on  $\beta$  with dispersion matrix  $\kappa_j^{-1}\Sigma$ . The remaining priors are  $\mu \sim \text{Expon}(1000)$ ,  $t_{1/2} \equiv \ln(2)/\phi \sim \text{Expon}(1000)$  and  $\sigma^{-2} \sim \text{Gamma}(0.0001, 0.0001)$ . The baseline concentration  $\mu$  and the half-life,  $t_{1/2} = \ln(2)/\phi$ , have *a priori* exponential distributions with a mean of 1000. This gives quite a flat prior over the range of values expected. The model precision,  $\sigma^{-2}$ , is given a vague gamma prior with mean 1 and variance 10,000.

Death rates are given in (7). The virtual simulation time  $t_0$  and the birth rate  $\lambda_b$  are inversely proportional to one another. Without loss of generality, then, we set  $t_0 = 1$ . The birth rate was chosen to be  $\lambda_b = 1$ . We also investigated larger birth rates and found that the posterior distribution did not depend on the birth rate (as indeed it should not). However, more time was spent searching for components in the birth-death process.

We started the MCMC algorithm by setting  $k = 1$ . We also found that the initial values for  $\mu$ ,  $t_{1/2}$ , and  $\beta$ , which all have vague priors, must be near the mode of the posterior distribution. When we tried extreme values, the number of

**Table 1**

*Priors on  $k$ ,  $\tau_j$ , and  $\Sigma^{-1}$  for the two prior models considered in our analysis. All other priors are common to both prior models.*

Parameter	Prior 1	Prior 2
$k$	$\mathcal{P}(10)$ ,	Uniform on $\{1, 2, \dots, 60\}$
$\tau_j$	$(3j)$ th order statistics, $j=1, \dots, k$ from $3k+2$ i.i.d. $U[-40, T+10]$ , randomly permuted	$U[-40, T+10]$
$\Sigma^{-1}$	Wishart $_4[S]$	Wishart $_4[S]$
$S^{-1}$	diag(5, 5)	diag(25, 25)

component functions got stuck at 1 with a mass near 0 and an extremely large variance. The baseline secretion rate rose to the overall mean of the data and the decay rate was driven to zero. So, for each dataset, we set  $\mu$  equal to the minimum observed concentration,  $t_{1/2} = 15$  and  $\beta = (15, 15)^T$ . These numbers were obtained from subject-specific knowledge. Without this knowledge, we recommend an initial run conditional on the number of secretion events (derived empirically) and their approximate locations (i.e., skip the search for location and number of events/component functions) and use the posterior mean of the parameter values. All other parameter values were drawn from their prior distributions.

5.2. Posterior Simulation

We simulated the parameters, steps 2 and 3 of Section 4, in the following order:  $\mu$ ,  $t_{1/2}$ ,  $\beta$ ,  $\Sigma$ ,  $\kappa_j$ ,  $\beta_j$ ,  $\tau_j$ , and  $\sigma^2$ . We sampled

$\sigma^2$ ,  $\kappa_j$ , and  $\Sigma$  directly from their posterior distributions:

$$\sigma^{-2} \sim \text{Gamma} \left[ a_\sigma + 0.5k, b_\sigma + 0.5 \sum_{i=1}^n \{ \ln(y_i) - \ln(C_i) \}^2 \right],$$

$$\kappa_j \sim \text{Gamma} \left[ 0.5(r + 1), 0.5 \{ r + (\beta_j - \beta)^T \Sigma^{-1} (\beta_j - \beta) \} \right],$$

$$\Sigma^{-1} \sim \text{Wishart}_{n+k} \left[ \left\{ S^{-1} + \sum_{j=1}^k \kappa_j (\beta_j - \beta)(\beta_j - \beta)^T \right\}^{-1} \right],$$

All other parameters were sampled using a Metropolis-Hasting step. The chain was run for 275,000 iterations with the first 25,000 iterations discarded as a burn-in. Every 25th iteration was kept (for a total of 10,000) to reduce storage requirements. Convergence was assessed graphically. Each simulation took approximately 50 minutes to run.

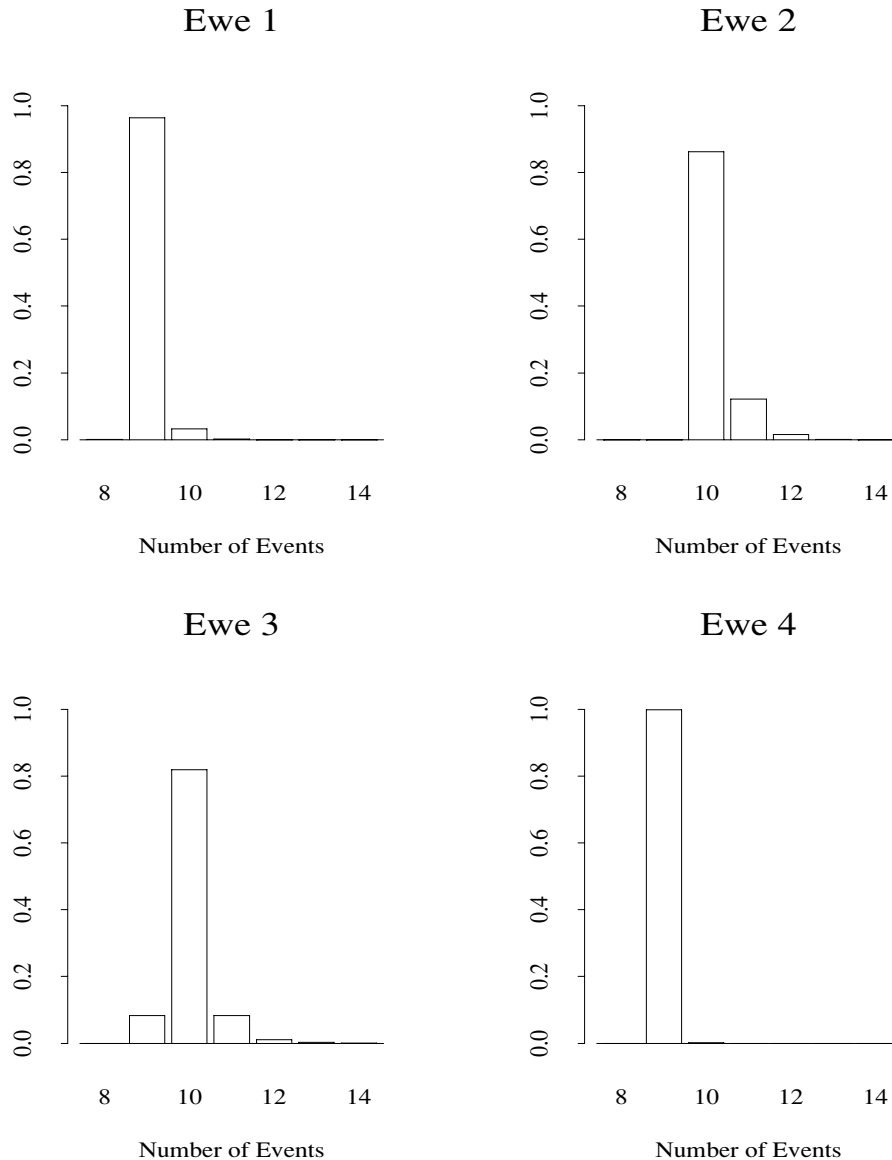


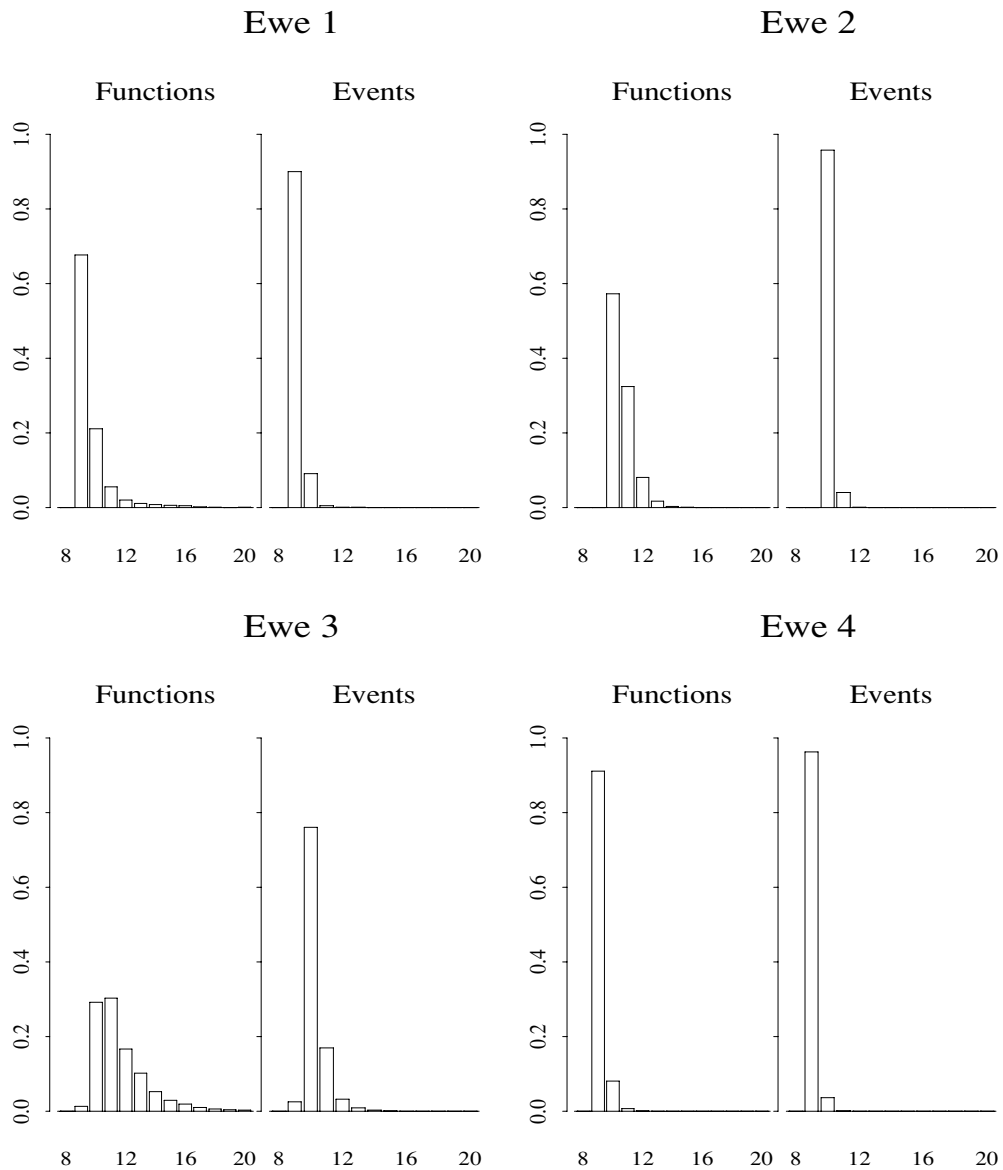
Figure 1. Marginal posterior distributions of the number of secretion events: prior 1.

5.3. Results

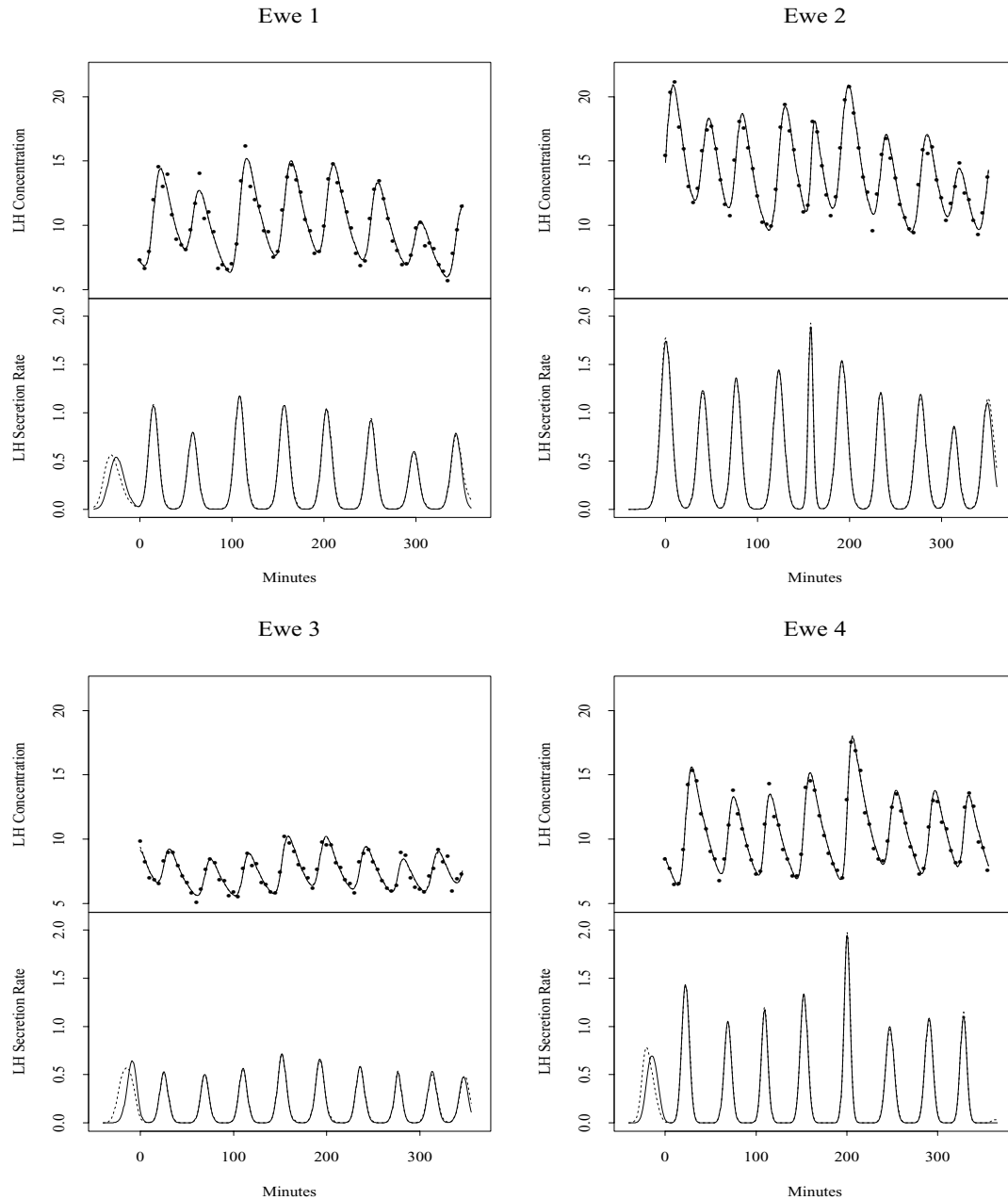
We now present the results and show that the two prior models considered can give quite different posterior distributions for  $k$ , the number of component functions. With regard to the first prior model, inference on  $k$  is necessary, since there is a one-to-one correspondence between component functions and secretion events (i.e.,  $k$  is the number of secretion events). Thus, our analysis using the first prior is conditional on the mode of  $k$ . With regard to the second prior model,  $k$  does not directly correspond to the number of secretion events and may not have any physical interpretation. Thus, inference on  $k$  is not of direct interest and so we perform a marginal analysis, integrating over  $k$ . Identification of the actual secretion events in the second prior model requires appropriate post-simulation processing. In particular, we allocated component

functions to secretion events, thus allowing estimates of the marginal posterior distribution of the number of secretion events, as well as their masses. Aside from  $k$ , the parameters of interest are insensitive to our two choices of prior information, including inference on the number of secretion events.

The posterior distributions of  $k$  for each animal are shown in Figure 1 (prior 1) and Figure 2 (prior 2). One can see that the posterior distribution of  $k$  is more variable for prior 2, but that the center of the distribution is not much affected. In fact, given our interpretation of the second prior in Section 5.1, we can allocate component functions to secretion events (see the Appendix for details). Figure 2 also shows the posterior distribution of the number of events under prior 2 after postprocessing. Comparing Figure 1 with



**Figure 2.** Marginal posterior distributions of the number of secretion component functions (left panels) and the number of secretion events after postprocessing (right panels): prior 2.



**Figure 3.** Expected concentration and secretion rate profiles. *Top panels:* concentration profiles. *Bottom panels:* secretion profiles. Solid profiles correspond to prior 1 and the dashed profiles to prior 2.

Figure 2, one sees that there are only minor differences in the distribution of the number of secretion events. Posterior means and standard deviations of the other parameters of interest are given in Table 2. The secretion events' masses were computed after conditioning on the mode of the number of secretion events. The events were sorted in ascending order according to their locations. We note here that the events are separated far enough apart in time that very little, if any, label switching occurs, as it does when one is using a mixture model to estimate a density. Other, more sophisticated, methods that deal with label switching have been developed (see, for example, Celeux, Hurn, and Robert (2000) and Stephens

(2000b)), but we feel that they are unnecessary here because of the distance between events.

Figure 3 shows the conditional (on the mode of the number of secretion events) expected values of the concentration

$$E\{C(t; \mu, \phi, \theta, k) | y, k\} = \int C(t; \mu, \phi, \theta, k) p(\mu, \phi, \theta | y, k) d\mu d\phi d\theta$$

and secretion profiles

$$E\{S(t; \theta, k) | y, k\} = \int S(t; \theta, k) p(\theta | y, k) d\theta$$

**Table 2**  
 Posterior means and standard deviations (in parenthesis) for model parameters under the two different prior models. Masses are per secretion event.

	Ewe 1		Ewe 2	
	Prior 1	Prior 2	Prior 1	Prior 2
$\mu$	2.40 (1.18)	2.80 (1.24)	5.82 (1.39)	5.33 (1.82)
$t_{1/2}$	19.82 (3.49)	18.56 (3.68)	13.36 (2.69)	14.1 (3.60)
$\sigma^2$	4.11 (0.86)	4.20 (0.91)	1.94 (0.48)	1.92 (0.46)
# events	9	9	10	10
Event 1, mass	11.05 (2.49)	12.33 (3.32)	25.84 (1.59)	26.32 (1.77)
Event 2, mass	14.16 (1.06)	14.52 (1.27)	16.90 (1.52)	16.99 (1.87)
Event 3, mass	10.62 (0.99)	10.77 (1.15)	17.19 (1.46)	17.32 (1.71)
Event 4, mass	15.83 (1.15)	16.16 (1.37)	20.18 (1.46)	20.36 (1.76)
Event 5, mass	14.59 (1.21)	14.96 (1.47)	12.60 (1.06)	12.67 (1.21)
Event 6, mass	13.87 (1.21)	14.23 (1.44)	22.42 (2.02)	22.50 (2.53)
Event 7, mass	12.11 (0.99)	12.31 (1.19)	14.43 (1.18)	14.58 (1.35)
Event 8, mass	8.08 (0.84)	8.06 (0.96)	16.55 (1.19)	16.79 (1.32)
Event 9, mass	10.86 (1.52)	11.51 (2.87)	10.28 (0.94)	10.52 (1.01)
Event 10, mass			15.57 (4.62)	18.15 (7.36)

	Ewe 3		Ewe 4	
	Prior 1	Prior 2	Prior 1	Prior 2
$\mu$	3.71 (0.98)	3.43 (1.26)	1.69 (0.95)	2.06 (0.95)
$t_{1/2}$	16.41 (4.78)	17.59 (6.41)	21.77 (2.55)	20.66 (2.53)
$\sigma^2$	3.61 (0.85)	3.5 (0.83)	1.64 (0.38)	1.57 (0.35)
# events	10	10	9	9
Event 1, mass	8.42 (1.35)	11.00 (2.75)	10.64 (1.60)	11.87 (1.90)
Event 2, mass	5.93 (0.84)	5.96 (1.07)	14.74 (0.57)	14.93 (0.57)
Event 3, mass	5.54 (0.65)	5.58 (0.70)	10.87 (0.50)	10.96 (0.49)
Event 4, mass	6.23 (0.69)	6.27 (0.76)	10.80 (0.48)	10.84 (0.47)
Event 5, mass	8.27 (0.92)	8.36 (1.21)	13.71 (0.57)	13.87 (0.57)
Event 6, mass	7.83 (1.02)	7.94 (1.40)	16.55 (0.70)	16.76 (0.73)
Event 7, mass	6.74 (0.83)	6.80 (1.02)	11.41 (0.55)	11.56 (0.56)
Event 8, mass	5.14 (0.64)	5.13 (0.70)	11.69 (0.51)	11.82 (0.51)
Event 9, mass	6.36 (0.74)	6.47 (0.89)	10.02 (0.47)	10.06 (0.47)
Event 10, mass	5.83 (2)	6.86 (3.71)		

corresponding to prior 1 and the marginal expected value of the concentration

$$E\{C(t; \mu, \phi, \theta, k) | y\} = \int_k \left\{ \int C(t; \mu, \phi, \theta, k) p(\mu, \phi, \theta | y, k) d\mu d\phi d\theta \right\} p(k | y) dk$$

and secretion profiles

$$E\{S(t; \theta, k) | y\} = \int_k \left\{ \int S(t; \theta, k) p(\theta | y, k) d\theta \right\} p(k | y) dk$$

corresponding to prior 2 for each of the four datasets. The concentration profiles are shown only within the range of the data, whereas the secretion profiles extend out to the left by 40 minutes and the right by 10 minutes. The expected concentration profiles corresponding to the two priors are virtually indistinguishable, as are the secretion profiles, except for perhaps the first and last events, where there are some minor differences.

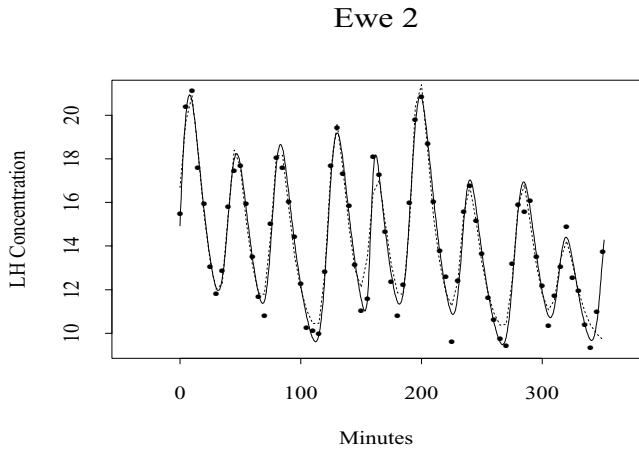
Figure 4 shows a comparison, for one of the datasets, of our method and that of Veldhuis et al. (1987). The concen-

tration profile is similar between the two methods, with one noteworthy exception. The last two data points appear to define the beginning of a secretion event. Our algorithm detected this as an event, whereas the nonlinear least-squares algorithm did not (to be precise, the algorithm used to locate secretion events prior to the nonlinear least-square fit did not detect this as a secretion event). This is a consequence of our search for component functions outside the range of the data. Thus, we are able to estimate secretion events and masses based on partial pulse information.

### 6. Simulation Study

The reader may note that the four datasets analyzed above are quite clean and the question arises about the performance of our method on noisier data. To answer this, we simulated 20 datasets. Four hundred minutes of data were simulated at 5-minute intervals. Nine secretion event positions per data set were simulated from the 3rd-, 6th-, up to the 27th-order statistics, from 29 uniform random variates on [0, 400]. The mass of each event was simulated from a  $t_4$  distribution with mean 12 and scale 5. The width, or scale, of each event was

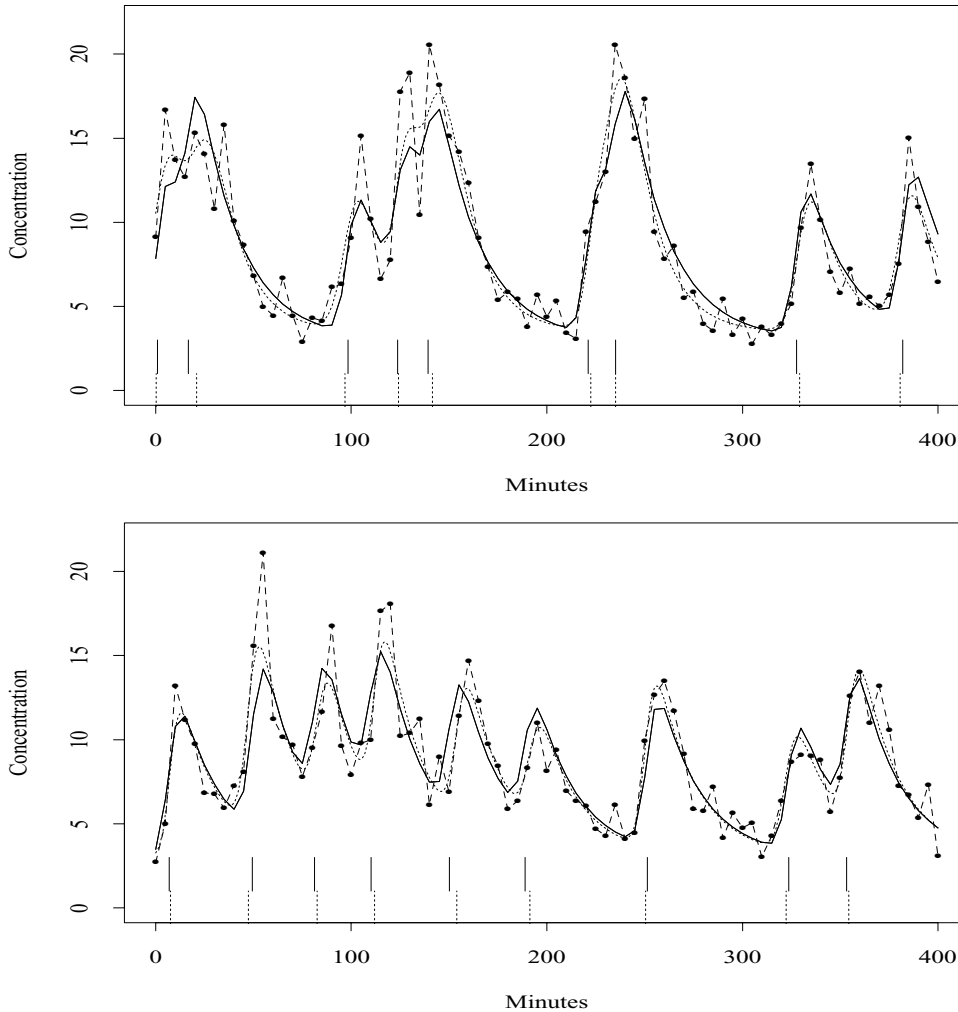




**Figure 4.** Comparison of our BDMCMC model and the deconvolution method of Veldhuis et al. (1987). Our method is the solid line.

simulated from a  $t_4$  distribution with mean 15 and scale 5. The baseline concentration was taken to be 3 and the removal half-half was 15 minutes. These parameters are comparable to those obtained from the real datasets above. Noise was added according to equation (4) above. The largest mean scale parameter,  $\sigma^2$ , from the four real datasets analyzed above was 0.004. In our simulated data, we increased this by a factor of 10 to 0.04. This gives us quite noisy data, but not so noisy that the signals are lost in the noise.

Two of the 20 simulated concentration profiles, along with estimates of the concentration profile, are displayed in Figure 5. We chose to show the simulated concentration profile in the top panel of Figure 5, because there appears to be only one large pulse between 200 and 300 minutes. However, there are two secretion events close in time at 221.1 and 235.1 minutes that, when convolved with exponential decay, result in a unimodal concentration profile. Our method detected both of these secretion events. The other 18 results were similar to the two shown. The MCMC output from these 20



**Figure 5.** Two simulated concentration profiles. The dashed line is the simulated data with noise. The solid black line is the simulated data and the dotted line is the expected concentration. The vertical lines at the bottom of each graph show the secretion event location (truth is solid) and the expected location from our model (dotted).

datasets were pooled together and 78% of these detected the nine secretion events. The mean baseline concentration was 2.82 (s.d. 0.63), the mean fixed effects event mass was 12.52 (s.d. 2.11), the mean fixed effects event scale was 16.52 (s.d. 8.89), the mean half-life was 15.55 (s.d. 3.33), and the model mean variance was 0.0398 (s.d. 0.01)—all quite close to the true parameter values.

## 7. Discussion

We saw above that the posterior distribution of  $k$  is dependent on the prior model chosen. For the particular datasets analyzed above, we would recommend using the first prior model, based on parsimony alone. We showed that the parameters of interest are not sensitive to the priors; however, the second prior requires postsimulation processing where the first does not. For other datasets, we would strongly recommend either using prior information elicited from an expert in endocrinology or trying several different prior models and investigating parameter sensitivity to prior information.

For the second prior, we used the posterior distribution to allocate component functions to events, where we defined an *event* as a mixture of the component functions with a unique maximum. An alternative, more principled, approach would be to search for event locations and masses in a birth-death process, then nest a second birth-death process that, conditional on a particular event, would determine the mixture of Gaussian component functions that make up that event. The search for the location of component functions would necessarily be on a smaller scale (centered on the event locations) than the scale on which to search for the event location. With this approach, we would expect to have better mixing over the number of component functions that make up each event. We would also expect event locations to be more evenly spread over the observed time interval, and thus the i.i.d. uniform prior would not be appropriate for event locations. Rather, some type of inhibition process would be more appropriate, for example, the  $m$ th,  $2m$ th,  $\dots$ ,  $km$ th order statistics from  $km + (m - 1)$  i.i.d. uniform random variates.

We have assumed in this article that the model errors are independent of one another. If this is not an appropriate assumption, one may wish to impose a correlation structure on the error terms. In this case, care must be taken in assuring that the exchangeability assumption is satisfied. If it is not, then RJMCMC should be used instead of BDMCMC.

We have demonstrated how one can jointly model pulsatile hormone data when the number of secretion events is unknown (which is the typical situation) using a Bayesian approach. The strengths of our model are: (1) It jointly models the number of events and parameters of interest. Other deconvolution models to date use statistical methods external to the model to determine the number of events and thus do not have the coherence of the Bayesian framework. (2) The subset of competing models is much larger than practically allowed using cross-validation or AIC. (3) It provides a coherent way to model average when one particular model is not substantially better than another. (4) We can estimate the probabilities of the number of events. (5) One can easily investigate any functional form (including generalized functions such as the Dirac delta function) for the secretion events. (6) Specific events may be more complex than Gaussian and, un-

der the model used here, single events can assume Gaussian shape or be a superposition (mixture, if you will) of two or more Gaussian-shaped functions.

## ACKNOWLEDGEMENTS

This work was funded by National Institute of Health grant P60 DK20572. The author would also like to thank Dr A. R. Midgely of the University of Michigan for supplying the data and Nichole Carlson, a graduate student in the Department of Biostatistics, University of Michigan, for analyzing the data with the nonlinear least-squares deconvolution program supplied by Dr Johannes Veldhuis and Dr Michael Johnson of the University of Virginia Health Sciences Center. I would also like to thank the referees for their insightful comments and suggestions.

## RÉSUMÉ

De nombreuses hormones sont sécrétées dans le système circulatoire de façon pulsatile, et sont éliminées exponentiellement. La méthode la plus habituelle pour analyser ces systèmes est de faire une déconvolution entre une fonction de sécrétion et une fonction d'élimination. L'estimation précise des paramètres du modèle dépend du nombre et de la position des pulses sécrétoires. A ce jour, l'analyse par déconvolution suppose connus a-priori et le nombre et la position approximative de ces pulses. Dans cet article nous présentons une approche bayésienne de déconvolution pour un modèle joint pour le nombre de pulses avec tous les autres paramètres du modèle. Notre méthode effectue une recherche stochastique des pulses sécrétoire. Pour cela nous envisageons l'ensemble des paramètres comme un processus ponctuel. Les pulses sont déterminés selon un processus de naissance-et-mort inséré dans un algorithme MCMC. Cette idée tire son origine de Stephens (2000a) dans le contexte de l'estimation d'un modèle de mélange fini de distributions où le nombre de composants du mélange est inconnu. Notre modèle présente plusieurs avantages par rapport aux approches fréquentistes traditionnelles. Ces avantages sont mis en évidence par quatre ensembles de données relatifs aux niveaux de concentration sérique de l'hormone LH, obtenus à partir de brebis ovariectomisées.

## REFERENCES

- Berne, R. M. and Levy, M. N. (eds). (1993). *Physiology*, 3rd edition. St. Louis: Mosby Year Book.
- Celeux, G., Hurn, M., and Robert, C. P. (2000). Computational and inferential difficulties with mixture posterior distributions. *Journal of the American Statistical Association* **95**, 957–970.
- Diggle, P. J. and Zeger, S. L. (1989). A non-Gaussian model for time series with pulses. *Journal of the American Statistical Association* **84**, 354–359.
- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**, 711–732.
- Guo, W., Wang, Y., and Brown, M. B. (1999). A signal extraction approach to modeling hormone time series with pulses and a changing baseline. *Journal of the American Statistical Association* **94**, 746–756.
- Keenan, D. M., Veldhuis, J. D., and Yang, R. (1998). Joint recovery of pulsatile and basal hormone secretion by

- stochastic nonlinear random-effects analysis. *American Journal of Physiology* **275**, R1939–R1949.
- Kleinman, K. P. and Ibrahim, J. G. (1998). A semiparametric Bayesian approach to the random effects model. *Biometrics* **54**, 921–938.
- Komaki, F. (1993). State-space modelling of time series sampled from continuous processes with pulses. *Biometrika* **80**, 417–429.
- Kushler, R. H. and Brown, M. B. (1991). A model for the identification of hormone pulses. *Statistics in Medicine* **10**, 329–340.
- Midgley, A. R., McFadden, K., Ghazzi, M., Karsch, F.J., Brown, M. B., Mauger, D. T., and Padmanabhan, V. (1997). Nonclassical secretory dynamics of LH revealed by hypothalamo-hypophyseal portal sampling of sheep. *Endocrine* **6**, 133–143.
- O'Sullivan, F. O. and O'Sullivan, J. (1988). Deconvolution of episodic hormone data: An analysis of the role of season on the onset of puberty in cows. *Biometrics* **44**, 339–353.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1990). *Numerical Recipes in C*. Cambridge: Cambridge University Press.
- Richardson, S. and Green, P. J. (1997). On Bayesian analysis of mixtures with an unknown number of components (with discussion). *Journal of the Royal Statistical Society, Series B* **59**, 731–792.
- Rodbard, D., Rayford, P. L., and Ross, G. T. (1970). Statistical quality control. In *Statistics in Endocrinology*, McArthur, J. W. and Colton, T. (eds), 411–429. Cambridge, Massachusetts: MIT Press.
- Stephens, M. (2000a). Bayesian analysis of mixture models with an unknown number of components: An alternative to reversible jump methods. *Annals of Statistics* **28**, 40–74.
- Stephens, M. (2000b). Dealing with label-switching in mixture models. *Journal of the Royal Statistical Society, Series B* **62**, 795–809.
- Veldhuis, J. D. and Johnson, M. L. (1992). Deconvolution analysis of hormone data. In *Methods in Enzymology*, L. Brand and M. L. Johnson (eds), 539–575. San Diego: Academic Press.
- Veldhuis, J. D., Fraioli, F., Rogol, A. D., and Dufau, M. L. (1986). Metabolic clearance of biologically active luteinizing hormone in man. *Journal of Clinical Investigation* **77**, 1122–1128.
- Veldhuis, J. D., Carlson, M. L., and Johnson, M. L. (1987). The pituitary gland secretes in bursts: Appraising the nature of glandular secretory impulses by simultaneous multiple-parameter deconvolution of plasma hormone concentrations. *Proceedings of the National Academy of Science* **84**, 7686–7690.

Received March 2002. Revised December 2002.  
Accepted January 2003.

## APPENDIX

### Postsimulation Processing

We can find the marginal posterior distribution of the number of secretion events if we are willing to accept that an event is the superposition of  $r \leq k$  component functions if this superposition has a unique maximum. Note that this precludes two events occurring too close in time. Since the entire secretion rate function  $S(t)$  is the superposition of all  $k$  component functions, we can search  $S(t)$  for local maxima. Each local maximum will be considered the location of a secretion event, or pulse. Let  $N_E$  denote the number of local maxima (number of secretion events) in  $S(t)$  and let  $t_i$  denote the position of the  $i$ th local maximum of  $S(t)$ . The posterior distribution of the number of secretion events can be estimated by

$$p(N_E = n | y) \approx \sum \mathbb{I}_{\{n\}}(N_E^{(i)}) / M$$

where the summation is over all saved iterations and  $M$  is the number of saved iterations. The superscript in parenthesis indicates the iteration number. The distributions of the number of secretion events for each ewe are presented in Figure 2. There are nine secretion events for the hormone from ewes 1 and 4 and 10 events for ewes 2 and 3.

Given the number of secretion events,  $N_E$ , we can estimate the mass of hormone released per event. Let  $M_i$  denote the hormone mass released during secretion event  $i$ . To compute  $M_i$ , we must allocate the component functions to secretion events.

For each  $j \leq k$ , draw  $t \sim p(\tau_j | y, N_E)$ . Then  $f_j$  is a component of event  $E_i$  if

$$\operatorname{argmax}_\ell \{g_\ell(t | y, N_E)\} = i, \quad (\text{A.1})$$

where  $g_i(t | y, N_E)$  is the density of the location of event  $E_i$ . Let  $\mathcal{A}_i = \{\ell : (\text{A.1}) \text{ holds}\}$ . Then  $M_i = \sum_{j \in \mathcal{A}_i} \beta_{j1}$  and

$$E(M_i | y, N_E) \approx \sum_{\ell \in \mathcal{I}_{N_E}} M_i^{(\ell)} / |\mathcal{I}_{N_E}|,$$

where  $|\mathcal{S}|$  denotes the cardinality of the set  $\mathcal{S}$  and  $\mathcal{I}_{N_E}$  is the set of iteration indices in which there were  $N_E$  secretion events.