Biometrics 63, 1207–1217
December 2007
DOI: 10.1111/j.1541-0420.2007.00809.x

Analysis of Pulsatile Hormone Concentration Profiles with Nonconstant Basal Concentration: A Bayesian Approach

Timothy D. Johnson

University of Michigan, Department of Biostatistics, 1420 Washington Heights, SPH II, Ann Arbor, Michigan 48109-2029, U.S.A.
email: tdjtdj@umich.edu

Summary. Many challenges arise in the analysis of pulsatile, or episodic, hormone concentration time series data. Among these challenges is the determination of the number and location of pulsatile events and the discrimination of events from noise. Analyses of these data are typically performed in two stages. In the first stage, the number and approximate location of the pulses are determined. In the second stage, a model (typically a deconvolution model) is fit to the data conditional on the number of pulses. Any error made in the first stage is carried over to the second stage. Furthermore, current methods, except two, assume that the underlying basal concentration is constant. We present a fully Bayesian deconvolution model that simultaneously estimates the number of secretion episodes, as well as their locations, and a nonconstant basal concentration. This model obviates the need to determine the number of events a priori. Furthermore, we estimate probabilities for all "candidate" event locations. We demonstrate our method on a real data set.

Key words: Bayesian analysis; BDMCMC; B-spline; Deconvolution; Dirichlet process prior; Ill-posed inverse problem; Latent point process; Marked point process; Non-stationary time series.

1. Introduction

Hormones are the main vehicle used in distant cell-to-cell communication throughout the human body. For example, gonadotropic hormones, produced and secreted into the circulatory system by the pituitary gland, signal distant target cells in the gonads. These, in turn, produce sex hormones that are responsible for the changes seen during puberty and for the monthly menstrual cycles in healthy, young, mature women. The timing and amount of hormone secreted into the circulatory system is thus vital to the normal functioning of many physiological processes.

For example, 26 clinically depressed women and 26 matched controls were enrolled in a study at the University of Michigan. Women were admitted to the General Clinical Research Center for a 24-hour period of observation, commencing at 9 a.m., during which time blood was drawn at 10-minute intervals. The serum was subsequently assayed for cortisol, a stress-related hormone. One question of interest was whether the number of pulse events is different in depressed women and controls (Young, Carlson, and Brown, 2001). Therefore, it is necessary to identify both the location and number of pulse events in study subjects. A second stage analysis, such as a two-sample t-test, is subsequently performed to test for significant differences between the groups. This article only addresses the first half of this analysis.

Secretion of hormones into the circulatory system can be broadly classified into two categories: oscillatory and pulsatile. An example of oscillatory hormone secretion is the diurnal pattern of dopamine. Other hormones, including cortisol, a stress-related hormone produced by the adrenal glands, are released in a pulsatile fashion, as can be seen from the concentration profile in the upper left panel of Figure 1. Pulsatile secretion is characterized by the release of a large amount of hormone in a short period of time (Berne and Levy, 1993). Furthermore, cortisol may have an underlying oscillatory component, on top of which, the pulsatile release of hormone is superimposed (Veldhuis et al., 1989).

Many methods for pulse identification have been proposed and subsequently classified into criterion- and model-based methods (Mauger and Brown, 1995). Criterion-based methods (see, e.g., Van Cauter et al., 1981; Merriam and Wachter, 1982; Oerter, Guardabasso, and Rodbard, 1986; Veldhuis and Johnson, 1986) are typically used to determine the location of pulses. Model-based methods assume some model for the data (see, e.g., O'Sullivan and O'Sullivan 1988; Diggle and Zeger, 1989; Kushler and Brown, 1991; Veldhuis and Johnson, 1992; Komaki, 1993; Guo, Wang, and Brown, 1999; Johnson, 2003; Yang, Liu, and Wang, 2006). In an extensive simulation study Mauger and Brown (1995) showed that model-based methods are preferred over criterion-based methods on the grounds of superior false positive and false negative error rates.

Many model-based methods use criterion-based methods to identify initial pulses or to select several competing models (three exceptions are Diggle and Zeger, 1989; Guo et al., 1999; Johnson, 2003). Thus, an error in the initial identification of

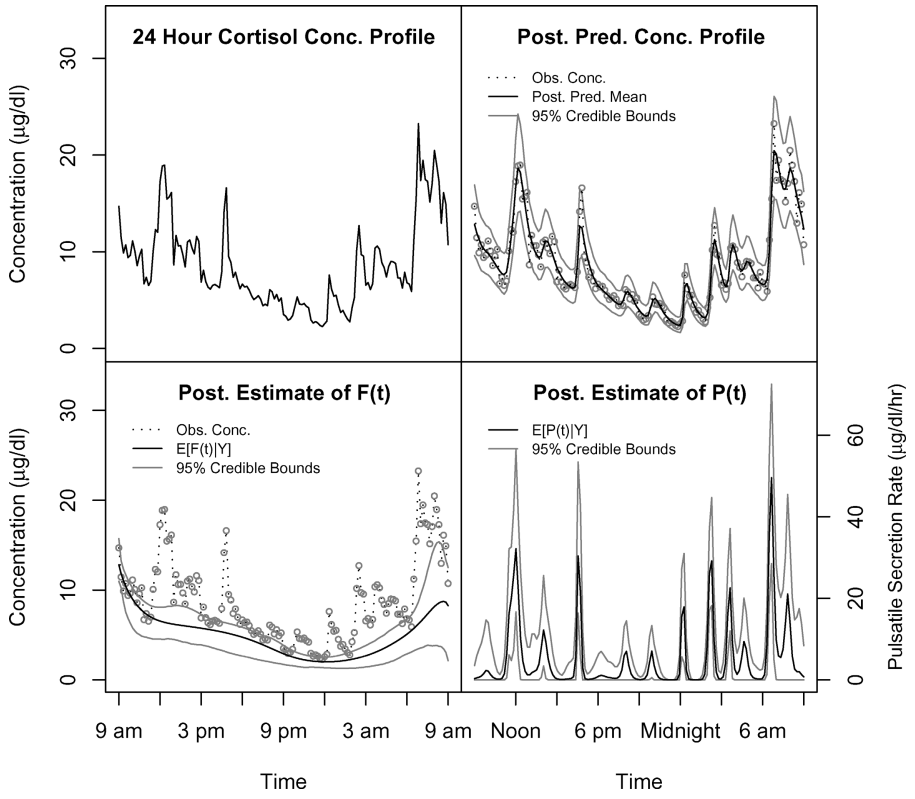

Figure 1. (UL) An example of cortisol concentrations from a female subject suffering from depression over a twenty-four-hour time period. Note the pulsatile nature of the concentration. Plasma concentration levels of cortisol were obtained at 10-minute intervals. (UR) Model fit to the data. (LL) Posterior mean of the nuisance function: $E(F(t)\,|\,Y)$. (LR) The posterior estimate of $P(t)\,|\,Y$.

a pulse is carried over to the model-based methods that rely on them. Further, only two previous methods have incorporated a changing basal concentration along with pulse identification/parameter estimation (Guo et al., 1999; Yang et al., 2006). All other methods assume a constant basal concentration. Thus, information such as the amount of hormone secreted per pulse will be overestimated. Lastly, a recent study (Keenan and Veldhuis, 2003) suggests that pulse size and shape vary throughout the day. Only two models have allowed for this biologic variation in pulse shape and size (Johnson, 2003; Yang et al., 2006).

To date, no model has been proposed that (i) does not rely on the initial identification of pulses; (ii) accounts for both pulsatile and oscillatory components of hormone release; and (iii) models variation in pulse shape and size. In this article, we present a Bayesian approach that addresses these three issues. Our method extends earlier work by Johnson (2003) and O'Sullivan and O'Sullivan (1988).

We present our model in Section 2. Then we briefly discuss some issues related to posterior simulation in Section 3. We analyze an example cortisol data set from the aforementioned study in Section 4. In Section 5, we discuss posterior estimation of event probabilities and compare results from our model with the nonlinear mixed effects partial spline model (NMPSM) developed by Yang et al. (2006). In Section 6, we present results from a small simulation study and discuss issues regarding sensitivity to priors. We conclude with a discussion in Section 7.

2. The Bayesian Deconvolution Model

2.1 *The Likelihood*

Let Y_j denote the observed concentration at time t_j, not necessarily equally spaced. Let ε_t denote the error and $C(t)$ denote the true concentration at time t. Then Y_j and $C(t_j)$ are related by

$$\ln(Y_j + 1) = \ln[C(t_j)] + \varepsilon_j$$
$$\text{with} \quad \varepsilon_j \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2), \quad j = 1, \ldots, n \tag{1}$$

where n is the number of observations in the time series. We modeled the log of the concentration for several reasons. First, both the true and observed concentration are nonnegative. Second, Rodbard, Rayford, and Ross (1970) argue that because the error in concentration is a combination of several different sources of error, including assay, biological, and dilution errors, a symmetric error structure is inappropriate. Equation (1) implies that the observed concentration has a lognormal distribution. Third, the assay error is proportional to the concentration level. An offset of 1 is added to the observed concentration to aid model fitting and to ensure that the log transformation is well defined. It is common to assume that $\varepsilon_t \sim N(0, \sigma^2)$, independently of one another.

Let Θ denote the set of all model parameters except σ^2. Concentration is modeled as a parametric function of Θ. The likelihood portion of our model is $\ln\,(Y_j + 1)\,|\,\Theta, \sigma^2 \sim N[\ln[C(t_j)], \sigma^2]$. (Note that we refer to the true concentration

as $C(t)$ and our model, presented below, of the true concentration also as $C(t)$. Although the model depends on Θ, we do not explicitly show this dependence.)

2.2 *The Concentration, $C(\cdot)$*

We assume a convolution model for hormone concentration. A goal is to deconvolve the concentration into a secretion function and a clearance function. Consider the first order, linear differential equation (with initial condition)

$$\frac{dC(t)}{dt} = kC(t) + S(t); \quad C(0) = C_0,\ k > 0. \tag{2}$$

This equation states that the change in hormone serum concentration at time t is proportional to both the concentration and the secretion rate, $S(t)$. It is solved by using the integrating factor $\exp(\int_0^t k\,dx) = \exp(kt)$. The solution is

$$C(t) = C_0 \exp(-kt) + \int_0^t S(x) \exp[-k(t-x)]\,dx. \tag{3}$$

The integral on the right-hand side (RHS) is the convolution integral of the hormone secretion and exponential decay (the clearance function).

We decompose $S(t)$ into the sum of basal secretion (the underlying oscillatory component), $B(t)$; pulsatile secretion, $P(t)$ and microscopic biologic fluctuation, $G(t)$. Thus (3) becomes (Keenan, Veldhuis, and Yang, 1998):

$$\begin{aligned} C(t) &= C_0 \exp(-kt) + \int_0^t [G(x) + B(x) \\ &\quad + P(x)] \exp[-k(t-x)]\,dx \\ &= \underbrace{C_0 \exp(-kt) + \int_0^t [G(x) + B(x)] \exp[-k(t-x)]\,dx}_{F(t)} \\ &\quad + \int_0^t P(x) \exp[-k(t-x)]\,dx \\ &= F(t) + \int_0^t P(x) \exp[-k(t-x)]\,dx. \end{aligned}$$

Often, the main interest lies in the number of pulsatile secretion events and the amount of hormone secreted during each event. Thus we consider $F(\cdot)$ a nuisance parameter. Johnson (2003) considers a similar model for the concentration; however, he only considers a constant function $F(t) \equiv c$ that was appropriate for that particular application.

2.2.1 The nuisance function, $F(\cdot)$. We believe that $F(\cdot)$ should vary smoothly over time and model it with a b-spline (de Boor, 1978) function with an a priori unspecified number of knots. Let N_K denote the number of interior knots at locations $\{\xi_i\}_{i=1}^{N_K}$ (throughout this article, we drop the indices on sets after their initial definition). Both N_K and $\{\xi_i\}$ are unknown quantities that are to be estimated. Conditional on N_K, let $\{\beta_i\}_{i=1}^{(N_K+4)}$ denote the set of b-spine coefficients and let $X \equiv X(\{\xi_i\})$ denote the design matrix whose rows are the basis functions evaluated at $\{\xi_i\}$ for the b-spline representation of $F(\cdot)$. Note that X depends on both the number and location of the knots. The dimension of X is $n \times (N_K + 4)$. Note that eight additional knots must be specified for a cubic b-spline with free end points. The first four knots are placed at the beginning of the data collection period, taken to be $t = 0$, and the last four knots at T, the end of the data collection period. Then $F(\cdot)$ can be approximated at the observation times t_j by $\hat{F}(t_j) = \langle X_{j,\cdot}^{\mathrm{T}},\ \beta_{N_K} \rangle$ where $X_{j,\cdot}^{\mathrm{T}}$ is the jth row of $X, \beta_{N_K} = (\beta_1, \ldots, \beta_{N_K+4})^{\mathrm{T}}$ and $\langle \cdot, \cdot \rangle$ denotes the inner product. To account for the offset of 1 added to the observed concentration, we must account for it in the estimation of the nuisance function. Otherwise, its estimate can be negative. We do so by adding 1 to its estimate.

2.2.2 The pulsatile secretion function, $P(\cdot)$. Parallel to the development in O'Sullivan and O'Sullivan (1988), we assume that the signaling of the release of hormone can be modeled by a nonhomogeneous point process $\{\tau_i\}_{i \geq 1}$ and its associated counting function $N((a,b]) = \sum_{i \geq 1} I_{(a,b]}(\tau_i)$, where $I_A(x)$ is the indicator function. We generalize their development by attaching marks, $\{\boldsymbol{\gamma}_i\}_{i \geq 1}$, to the process $\{\tau_i\}$ (independently of the event times) in a measurable marking space $(\mathcal{M}, \mathcal{F})$ with probability measure μ. We assume that μ is absolutely continuous with associated density $f(\boldsymbol{\gamma} \mid t)$. Then, $\{\tau_i,\ \boldsymbol{\gamma}_i\}_{i \geq 1}$ is also a point process with counting function: $N((a,b] \times A) = \sum_{i \geq 1} I_{(a,b] \times A}(\tau_i, \gamma_i)$ where $A \in \mathcal{F}$. Let $\lambda(t)$ denote the (marginal) rate function of $\{\tau_i\}$. Then the rate function of $\{\tau_i,\ \boldsymbol{\gamma}_i\}$ is $\lambda(t) f(\boldsymbol{\gamma} \mid t)$. Further, let $p(t; \boldsymbol{\gamma})$ represent a family of nonnegative functions parameterized by $\boldsymbol{\gamma}$. The pulsatile component, $P(t)$, of the secretion function is given by the convolution of $p(t; \boldsymbol{\gamma})$ and $N((0,t] \times \mathcal{M})$:

$$\begin{aligned} P(t) &= \int_{(\nu, \mathbf{y}) \in (0,t] \times \mathcal{M}} p(t - \nu;\ \gamma - \mathbf{y})\, dN((\nu, \mathbf{y})) \\ &\overset{\text{def}}{=} \sum_{i \geq 1} p(t - \tau_i;\ \gamma_i) I_{(0,t] \times \mathcal{M}}(\tau_i,\ \gamma_i) = \sum_{i=1}^{N(t)} p(t - \tau_i;\ \gamma_i), \end{aligned} \tag{4}$$

with $t \in (0,\ T]$. The pulse functions, $p(t - \tau;\ \boldsymbol{\gamma})$, are taken to be Gaussian shaped in this article: $p(t - \tau; \gamma) = \alpha \exp(-0.5(t-\tau)^2/\nu^2)/\sqrt{2\pi\nu^2}$ where $\boldsymbol{\gamma} = (\alpha,\ \nu^2)$, α represents the amount of hormone secreted by a gland due to the signal at time τ and ν^2 controls the width of the function $p(t; \boldsymbol{\gamma})$. Thus $\mathcal{M} = \mathbb{R}^+ \times \mathbb{R}^+$ and $\mathcal{F}$ are the Lebesque measurable rectangles in $\mathcal{M}$.

Remark 1: We take some notational liberty here. The process $\{\tau_i\}$ represents both the signaling mechanism arrival times and the centers of the Gaussian-shaped pulse functions $p(t - \tau_i;, \boldsymbol{\gamma}_i)$; implying that the gland releases $\alpha_i/2$ hormone prior to the signal. Technically, the signal should precede any hormone secretion. This would be the case, for instance, if $p(t; \boldsymbol{\gamma}) \propto \alpha\, t^{\gamma - 1} \exp(-\beta\, t)$ with $\boldsymbol{\gamma} = (\alpha, \beta, \gamma)$ for $t \geq 0$. Within the Bayesian framework, the pulse function family can be virtually any nonnegative family of functions. Henceforth, the process $\{\tau_i\}$ will represent the arrival times of (the centers of) the pulse functions $p(t - \tau_i; \boldsymbol{\gamma}_i)$ with the understanding that the signaling mechanism precedes any (significant amount of) hormone release: $\int_{-\infty}^{s_i} p(t - \tau_i;\ \gamma) \approx 0$, where $s_i (\leq \tau_i)$ is the ith signaling event.

Remark 2: Marking the process allows for biological variation in the amount of hormone released during each pulse

event. This biological variation may be due to several factors including the strength of the signaling mechanism and the amount of hormone in the gland available for secretion among other factors such as known or unknown positive and negative feedback mechanisms that are unobservable.

REMARK 3: Allowing the process to have a nonhomogeneous rate function gives flexibility in modeling the shape and size of a *pulsatile secretion event* (loosely defined as a large quantity of hormone released in a relatively short period of time). It also allows for alternating active and quiescent periods of hormone release throughout the day. A secretion event may consist of a single pulse function pulse $p(t - \tau_i; \boldsymbol{\gamma}_i)$ or of the superposition of several pulse functions from a rapid succession of signals.

2.2.3 The elimination function. There are several mechanisms for the elimination of hormone from the circulatory system including specific target cite binding, enzymatic cleavage, and glomerular filtration. Typically these mechanisms cannot be modeled individually and a single elimination function is used to model the overall clearance rate. The elimination function is typically chosen to be exponential or bi-exponential decay unless experimentation suggests some other form. In this article we will assume that clearance is exponential: $\exp(-kt)$. We model the half-life of the clearance function: $t_{1/2} = \ln(2)/k$, instead of the decay rate, k, which helps with convergence properties during Markov chain Monte Carlo (MCMC) simulation (see Section 2.3.4).

2.3 *Parameter Priors*

2.3.1 Prior factorization. We note here that the pulsatile function and the basal nuisance function are nearly confounded. It may be virtually impossible to separate out a small, broad pulse from a local maximum in the nuisance function. From the Bayesian perspective, this confounding can be controlled, in large part, by an appropriate prior specification of model parameters. Thus, an objective Bayesian approach is rather useless in a problem of this nature and informative priors are a necessity. Physiologically, a pulse is characterized by a large amount of hormone secreted into the circulatory system in a very short (almost as a bolus) period of time. Furthermore, the oscillations of the basal concentration, at least for cortisol, should have approximately a 24-hour period. In this section, we define the prior distributions used in our analyses. Cortisol hormone specific hyperprior values are defined in Section 4. A sensitivity analysis is presented in Section 6.

We assume the following factorization of the joint prior:

$$\begin{aligned}\pi\big[&\{\beta_i\}, \beta, \psi^2, \{\xi_i\}, N_K, \{\tau_j\}, \{\nu_j^2\}, \nu, \zeta^2, \{\alpha_j\}, \alpha, \phi^2, N(T), \\ &\times t_{1/2}, \sigma^2, \lambda\big] \\ &= \prod_{i=1}^{N_K} \big[\pi(\{\beta_i\} \mid N_K, \beta, \psi^2)\, \pi(\{\xi_i\} \mid N_K)\big]\ \pi(N_k)\, \pi(\beta)\, \pi(\psi^2) \qquad (5)\end{aligned}$$

$$\begin{aligned}&\times \Big\{ \prod_{j=1}^{N(T)} \big[\pi(\alpha_j \mid N(T), \alpha, \phi^2)\, \pi(\nu_j^2 \mid N(T), \nu, \zeta^2)\big] \\ &\times \pi(\tau_1, \ldots, \tau_{N(T)} \mid N(T)) \pi(N(T) \mid \lambda)\, \pi(\lambda)\, \pi(\nu)\, \pi(\zeta^2) \\ &\times \pi(\alpha)\, \pi(\phi^2)\Big\} \times \pi(t_{1/2})\, \pi(\sigma^2). \qquad (6)\end{aligned}$$

Equation (5) includes priors for parameters of the nuisance function, $F(\cdot)$, as well as hyperpriors on these parameters; the factors enclosed in parentheses in (6) are the priors (and hyperpriors) for the parameters of the pulsatile secretion function, $P(t)$. The last two factors are half-life and model variance priors, respectively.

2.3.2 Priors for the parameterized function $F(\cdot)$. The number of knots, N_K, is assigned a negative binomial prior with mean 3 and variance 6: $N_k \sim \text{Negbin}(3, 1)$. The number of knots can be thought of as a smoothing parameter in a b-spline representation (de Boor, 1978) with fewer knots resulting in a smoother estimate $\hat{F}(\cdot)$. Because we believe that $F(\cdot)$ should be a slowly varying function (smooth oscillations), we chose a small mean for N_K. The knots locations, $\{\xi_i\}$, are a priori distributed as N_K independent, uniform random variables over the data collection period. Their conditional joint prior density is thus $\pi(\xi_1, \ldots, \xi_{N_K} \mid N_K) = T^{-N_K}$.

The last set of parameters that require a prior, for the specification of $F(\cdot)$, is $\{\beta_i\}$. $F(\cdot)$ is necessarily a nonnegative function. Because the basis functions of the b-spline representation are nonnegative, a sufficient condition for a b-spline function to be nonnegative is that each b-spline coefficient β_i, $i = 1, \ldots, N_K$, be nonnegative (de Boor, 1978). Therefore, we model the natural log of the b-spline coefficients hierarchically: $\ln(\beta_i) \mid N_K, \beta, \psi^2 \sim N(\beta, \psi^2)$ with $\ln(\beta) \sim N(3, 1)$ and $\psi^2 \sim IG(2.1, 2)$, the inverse gamma distribution. The parameter values in the normal and inverse gamma distributions were selected, in large part, to give reasonable mixing properties in the birth-death MCMC (BDMCMC) algorithm (see Section 3).

2.3.3 Priors for the parameterized function $P(\cdot)$. Next, we specify the priors for parameters that define $P(\cdot)$ in (4). The priors are specified hierarchically. We begin with the number of pulse functions, $N(T)$. A priori we assume that the stochastic process, $\{\tau_i\}$, governing the arrival times of the pulse functions is a homogeneous conditional Poisson process, given λ, with intensity $\lambda^* = \lambda/T$ (any indications of a nonhomogeneous process that are present in the data will be reflected in the posterior). We reflect uncertainty in λ by placing a prior distribution on it as well: $\lambda \mid a_\lambda, b_\lambda \sim G(a_\lambda, b_\lambda)$. Hence $N(T) \mid \lambda \sim Pois(\lambda)$ (marginally $N(T)$ has a negative binomial distribution). For a homogeneous Poisson process, $\{\tau_i\}$, it is well known that the joint distribution of $\tau_1, \ldots, \tau_{N(T)}$ given $N(T) = n$ is distributed as the order statistics from n independent and identically distributed uniform random variables scaled to $[0, T]$: $\pi(\tau_1, \ldots, \tau_{N(T)} \mid N(T) = n) = n!/T^n$.

Conditional on $N(T)$, the remaining parameters are $\{\boldsymbol{\gamma}_i\}$. Recall that $\boldsymbol{\gamma}_i = (\alpha_i, \nu_i^2)$. We assume a priori that these parameters are independent of the event arrival times and that all of the elements of the sets $\{\alpha_i\}$ and $\{\nu_i^2\}$ are a priori conditionally independent and are all strictly positive. Thus, we specify the priors for α_i and ν_i^2 hierarchically on the log scale:

$$\ln(\alpha_i) \mid \alpha, \phi \overset{\text{i.i.d}}{\sim} N(\ln(\alpha), \phi^2), \quad i = 1, \ldots, N(T), \qquad (7)$$

$$\ln(\alpha) \sim N(m_\alpha, v_\alpha^2) \quad \text{and} \quad \phi^2 \sim IG(a_\phi, b_\phi), \qquad (8)$$

and

$$\ln(\nu_i^2) \mid \nu, \zeta \overset{\text{i.i.d}}{\sim} N(\ln(\nu), \zeta^2), \quad i = 1, \ldots, N(T), \qquad (9)$$

$$\ln(\nu) \sim N(m_\nu, v_\nu^2) \quad \text{and} \quad \zeta^2 \sim IG(a_\zeta, b_\zeta)., \tag{10}$$

We believe that there may be biological variation from pulse to pulse; however, we also believe that the pulse functions should be more or less similar. Modeling these pulse function shape parameters hierarchically reflects this belief. Note that this hierarchical modeling also reduces the influence of the priors on the posterior estimates. The hyperpriors a_λ, b_λ, a_ζ, b_ζ, a_ϕ, b_ϕ, m_ν, v_ν^2, and m_α, v_α^2 are all constant and will depend on the particular hormone under study. They are defined in Section 4 for the cortisol study presented in the Introduction.

2.3.4 Priors for $t_{1/2}$ and σ^2. The last two parameters that require prior specifications are the half-life and the model variance. We choose to model the decay function in terms of the hormone half-life instead of the decay rate, k. During MCMC simulation of the posterior, we observed that modeling hormone clearance from the system in terms of the half-life results in better simulation performance. When the decay rate was used, it often happened that both it and the $\{\alpha_i\}$ simultaneously escape to very large numbers, numbers that are biologically impossible, never to return. This is easy to control with a suitable prior under the half-life parameterization because as $k \to \infty, t_{1/2} \to 0$. Because the half-life is a strictly positive number we assign it a log-normal prior: $\ln(t_{1/2}) \sim N(m_t, v_t)$. The prior mean and variance will depend on the particular hormone under investigation. The model variance is given a vague, proper prior: $\sigma^2 \sim IG(0.001, 0.001)$.

3. Simulating from the Posterior

The full joint posterior is not analytic in form. Thus we resort to MCMC simulation of the posterior distribution. Furthermore, the number and location of both the b-spline knots as well as the number and location of the pulse function arrival times (and associated marks) are unknown. Standard MCMC techniques are not equipped to handle variable-dimension parameter space problems. Recent advances in Bayesian computation, however, have made variable-dimension problems feasible. We use the BDMCMC algorithm developed by Stephens (2000) to simulate the number of spline knots, their locations and associated coefficients. A separate BDMCMC algorithm is used to draw the number of pulse functions, their locations and associated marks. Both instances of the BDMCMC algorithm are imbedded within the MCMC simulation of the remaining parameters. An alternative algorithm to the BDMCMC algorithm that could also be used is the reversible-jump MCMC algorithm due to Green (1995).

The BDMCMC requires exchangeable priors and likelihoods (Stephens, 2000). The conditional prior $\pi(\tau_1, \ldots, \tau_{N(T)} \,|\, N(T))$ violates this requirement. Thus to implement the BDMCMC algorithm on the pulse function parameters, we further condition on a random permutation of the arrival time indices $(1, \ldots, N(T))$. Let p denote this random permutation with conditional prior $\pi(p \,|\, N(T)) = 1/N(T)!$. Then

$$\begin{aligned}&\pi(\tau_{p(1)}, \ldots, \tau_{p(N(T))} \,|\, N(T)) \\ &\quad = \pi(\tau_1, \ldots, \tau_{N(T)} \,|\, N(T))\, \pi(p \,|\, N(T)) = T^{-N(T)}.\end{aligned}$$

It is this joint distribution that is used in the BDMCMC algorithm. As for the $\{\alpha_i\}$ and $\{\nu_i^2\}$, they are all a priori independent. At the end of the simulation, we can easily convert back to the ordered arrival times. As an aid to chain mixing, after the BDMCMC algorithms return to the main MCMC chain, the parameters drawn from the BDMCMC algorithm are further updated with standard Gibbs or Metropolis–Hastings steps as suggested by Stephens (2000).

Remark 4: This permutation of the arrival times is justified by the fact that the simplest way to simulate the arrival times of a Poisson process with intensity λ in an interval $(a, b]$ conditional on $N((a, b])$ is to (i) draw $N((a, b])$ from a Poisson distribution with mean $\lambda(b - a)$; (ii) draw $N((a, b])$ independent uniform random variables on $(a, b]$; and (iii) order them. Conceptually, the BDMCMC algorithm executes (i) and (ii).

4. Example

All 52 cortisol data sets from the study described in the Introduction were analyzed. The same (hyper)prior distributions were used for all analyses. It is impractical to display results on all 52 data sets and so we report results on a randomly selected data set. Results are based on averaging over models indexed by $N(T)$. The data are shown in the upper left panel of Figure 1. The data collection took place during a 24-hour period ($T = 24$) and blood was drawn at 10-minute intervals. First we need to specify the remaining (hyper)prior parameters that are hormone specific. We chose values that result in prior means/modes that are reasonable from a biological point of view and variances that are moderately large. (An alternative approach would be to elicit the prior distribution for the half-life, then use the data and the convolution model to find reasonable, so that the concentrations profiles look similar to the data, priors for the pulse function parameters in a "quasi-empirical Bayes approach." The mean of the Poisson process could be obtained by asking an expert how many pulses they detect by eye.) To begin we specify our priors and hyperpriors for the pulse function parameters: $\alpha_\lambda = 10$ and $\beta_\lambda = 1$. Thus, $\lambda \,|\, \alpha_\lambda, \beta_\lambda \sim G(10, 1)$ and reflects our belief that the number of secretion events should be about 10 in a 24-hour period and thus, a priori we believe that each secretion event should be made up of one pulse function. However, it is also variable enough to allow fewer events and allow for the possibility that secretion events may be made up of several component functions. For the amount of hormone released, we set $m_\alpha = 3, v_\alpha^2 = 1, a_\phi = 5$, and $b_\phi = 2$. The prior mode of α, the fixed effect of the amount of hormone released, is 7.39 with a standard deviation of 43.4. Plugging in the prior means of α and ϕ^2 into (7) results in a prior mode of 12.2 for the random effects, α_i, with a standard deviation of 20.8. For the "width" of the pulse functions we set $m_\nu = -1, v_\nu^2 = 1, a_\zeta = 5$, and $b_\zeta = 2$. The fixed effect, ν, has mode 0.13 (about 8 minutes) and standard deviation 0.8 (about 47.7 minutes). Plugging in the prior means of ν and ζ^2 into (9) results in a prior mode of 0.22 (about 13.4 minutes) for the random effects ν_i with a standard deviation of 0.38 (about 22.8 minutes). For the half-life we set $\ln(t_{1/2}) \sim N(-1, 1)$. Upon exponentiating, this prior has 90% of its mass between 0.07 (4.2 minutes) and 1.9 (114 minutes)—a reasonable range for the elimination of cortisol. Only 1.6×10^{-4} of the mass is less than 0.01, thus controlling the half-life from becoming too small. We reiterate here that the priors in this model must be moderately

informative. Very diffuse, or uninformative, priors are inappropriate due to the potential confounding of the secretion function and the basal nuisance function and due to the inherent instability of deconvolution. (Deconvolution is a well-known ill-posed problem, see Wahba, 1990; Kaipio and Somersalo, 2005). In large part, the confounding of the secretion function and the basal nuisance function is controlled by the priors on $N(T)$ and N_K. The instability of deconvolution is controlled by the priors on $t_{1/2}$, α, $\{\alpha_i\}$, ν and $\{\nu_i\}$.

We ran the MCMC simulation for 225,000 iterations with a burn-in of 25,000. The chain was thinned by saving every 100th sample. The simulation took 19 minutes on a dual 2.7 GHz Power Mac G5. Convergence was assessed graphically on the parameters that are of fixed-dimension. Good mixing was realized for both BDMCMC algorithms with N_K taking on a different value approximately 4% of the time and $N(T)$ taking on a different value roughly 10% of the time.

Graphical results are given in Figure 1. The upper right panel of Figure 1 shows the expected value of the posterior predictive density (PPD) of the concentration (solid black line) as well as the 95% pointwise prediction bounds (solid gray lines). The lower left panel shows the posterior mean of the nuisance function $F(t) \,|\, Y$ (solid black line) as well as the 95% pointwise credible bounds (solid gray lines). The posterior estimated mean of the pulsatile secretion function $P(t)$ and 95% pointwise credible bounds are displayed in the lower right panel (black and gray lines, respectively). Fits were obtained by marginalizing the posterior over N_K and $N(T)$.

Overall, the model fits the data well. Bayesian deleted residuals (Gelfand, Dey, and Chang, 1992) show no structural pattern, indicating that the mean structure of the data has been captured by the model (this was true for all 52 data sets). There are only one or two outliers (absolute residual value greater than 3). Further evidence of model fit is ascertained by computing a Bayesian goodness-of-fit statistic (Johnson, 2004). The value of the statistic is 0.502. Under the null hypothesis that the model fits well, a value of 0.5 is expected. Values much smaller than 0.5 indicated overfitting of the data, while values much larger than 0.5 indicate a lack of fit.

5. Estimating the Posterior Rate Function

The number and location of secretion events are often of interest to the investigator. The posterior mean of the pulsatile secretion function, shown in the lower-right panel of Figure 1, and the 95% pointwise credible bounds give a good indication of the locations of these events and an indication of how likely it is that a secretion event occurs in a given interval of time. However, two questions are often asked that the figure does not address:

(i) What is the probability that a secretion event occurs at time t?
(ii) How many secretion events are there?

Johnson (2003) allocates pulse functions to secretion events by considering any group of pulse functions whose superposition results in a function with a single global maximum as a single secretion event. The resulting posterior distributions of the number of secretion events for the data sets he considers all have a dominant mode. That mode is taken as the number of secretion events and all further analyses are conditional on that number of events. No estimate of the probability that something is an event was attempted, nor was it necessary. For the data considered here, his method is unsatisfactory. A dominant secretion event mode does not exist; the mode occurs approximately 21% of the time (not shown) and thus the uncertainty in model choice is large. Furthermore, we want to estimate the posterior probability that a secretion event occurs in an arbitrary interval.

We propose a different solution that answers both questions: The procedure consists of the following three steps:

(i) Estimate the marginal posterior rate function $\lambda^*(t) \,|\, Y$ of the posterior process $\{\tau_i\} \,|\, Y$.
(ii) Estimate the location of secretion event j, SE_j, as the jth ordered local maximum, M_j, of $\lambda^*(t) \,|\, Y$.
(iii) For a neighborhood, $\mathcal{M}_j$, containing M_j, estimate the probability that at least one pulse arrival time is in $\mathcal{M}_j$. Take this estimate as the $\Pr(SE_j \in \mathcal{M}_j)$.

Step 1: Define the posterior rate function, $\lambda^*(t) \,|\, Y$, implicitly by $E(N(t) \,|\, Y) = \int_0^t \lambda^*(s) \,|\, Y \, ds$ for $t \in [0, T]$. Suppose $\lambda^*(t) \,|\, Y$ is bounded and continuous on $[0, T]$. Then

$$\lim_{h \to 0} \frac{1}{h} E[N(t+h) - N(t) \,|\, Y] = \lim_{h \to 0} \frac{1}{h} \int_t^{t+h} \lambda^*(s) \,|\, Y \, ds = \lambda^*(t) \,|\, Y.$$

This suggests a way to estimate $\lambda^*(t) \,|\, Y$. Partition $[0, T]$ into n subintervals $A_i = (t_{i-1}, t_i]$, $i = 1, \ldots, n$, $t_0 = 0$, $t_n = T$ each of length T/n. Then

$$\lambda^*(t_i) \,|\, Y \approx \frac{n}{T} E[N(t_i) - N(t_{i-1}) \,|\, Y], \quad i = 1, \ldots, n.$$

The expectation on the RHS can be approximated from the MCMC posterior samples. Suppose we save $\mathcal{S}$ samples. Then

$$E[N(t_i) - N(t_{i-1}) \,|\, Y] \approx \frac{1}{\mathcal{S}} \sum_{j=1}^{\mathcal{S}} \sum_{k=1}^{N(T)^{(j)}} I_{A_i}(\tau_k^{(j)}), \quad i = 1, \ldots, n, \tag{11}$$

where the superscript (j) on a parameter indicates its values from the jth MCMC sample. The RHS of (11) is a histogram estimate of the counts of $\{\tau_{p(i)}\} \,|\, Y$. We wish to have a smooth estimate of $\lambda^*(t) \,|\, Y$ and thus seek a smooth estimate of the histogram.

Consider the marginal posterior density

$$\pi(\{\tau_{p(i)}\} \,|\, Y) = \sum_{k=0}^{\infty} \pi(\{\tau_{p(i)}\} \,|\, N(T) = k, Y) \pi(N(T) = k \,|\, Y). \tag{12}$$

This can be estimated by dividing the RHS of (11) by $\sum_{j=1}^{\mathcal{S}} N(T)^{(j)}$. Therefore, $\lambda^*(t) \,|\, Y$ is proportional to this marginal posterior density. Thus, by estimating the marginal density, we estimate $\lambda^*(t) \,|\, Y$. We estimate this density with a mixture of normal density components whose number is unknown. We choose a Bayesian model for density estimation using mixtures of Dirichlet processes, MDP, (Ferguson, 1973; Antoniak, 1974; MacEachern and Müller, 1998) that was first proposed by Escobar and West (1995).

The MDP model is as follows. The data are the marginal posterior draws of the τ_j. The dependence on Y and on the

permutation p is dropped only for notational convenience. Each observation is allowed its own set of parameters. The parameters are assumed to be distributed, i.i.d., according to a distribution G, with additional uncertainty about the prior distribution G. In particular, G is a random distribution generated by a Dirichlet process with base measure $\alpha_0\ N(\mu,\ \sigma^2)$ and total mass parameter α_0:

$$\tau_j \sim N\big(\mu_j,\, \sigma_j^2\big), \quad \big(\mu_j,\, \sigma_j^2\big) \sim G, \quad G \sim DP[\alpha_0,\, N\big(\mu,\, \sigma^2\big)].$$

We also place priors on the parameters of the base measure in the Dirichlet process:

$$\mu \sim U(0, T), \quad \sigma^2 \sim IG(2.1, \beta), \quad \alpha_0 \sim G(1, 1)$$

with hyperprior $\beta \sim G(1,\, 1)$. The estimation of this model follows that given by Neal (2000, Algorithm 8). We estimate the posterior of α_0 by the method proposed in Escobar and West (1995). Let $\hat{\lambda}^*(t)\,|\,Y$ denote the estimated marginal posterior rate function. It is shown by the black curve in Figure 2. The gray region under the curve is the histogram based estimate of $\lambda^*(t)\,|\,Y$.

Step 2: The vertical lines at the bottom of Figure 2 show where all local maxima, $M_j,\ j = 1, \ldots, 34$, occur for the data set analyzed above (some of which are so small that they are not visible in $\hat{\lambda}^*(t)\,|\,Y$). The jth local maximum is taken to be a point estimate of the location of the jth secretion event, SE_j. These maxima are easily obtained because the mixture density is a mixture of normal density components. We can easily estimate the expected first and second derivatives of $\hat{\lambda}^*(t)\,|\,Y$ on a fine grid of [0, T] during the MCMC simulation of the MDP model and numerically estimate where the first derivative equals zero and the second derivative is negative.

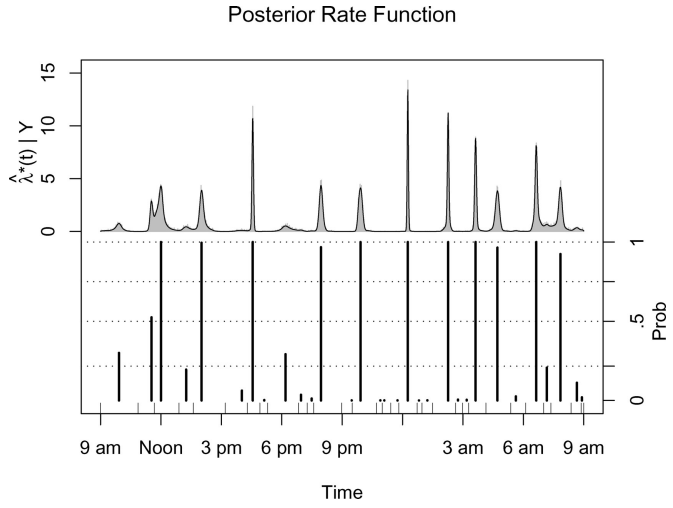

Figure 2. The top portion of this figure shows the estimated marginal posterior rate function, $\hat{\lambda}^*(t)\,|\,Y$, estimated from the MDP model. The vertical lines in the bottom portion show locations of all local maxima and the probabilities (14). Some maxima are so small, they cannot be seen. The rug at the bottom shows the intervals endpoints (the local minima) on which the probabilities were estimated.

Step 3: For any given interval, $\mathcal{I} \subset [0, T]$, it is easy to estimate the posterior expected number of pulse arrivals in $\mathcal{I}$:

$$E[N(\mathcal{I})\,|\,Y] \approx \int_{\mathcal{I}} (\hat{\lambda}^*(t)\,|\,Y)dt \approx \frac{1}{\mathcal{S}} \sum_{j=1}^{\mathcal{S}} \sum_{k=1}^{N(T)^{(j)}} I_{\mathcal{I}}\big(\tau_k^{(j)}\big). \quad (13)$$

Also the posterior probability that at least one pulse arrival occurs in $\mathcal{I}$ is easily estimated. First, let δ_i denote the Kronecker delta function ($\delta_i = 1$ if $i = 0$ and $\delta_i = 0$ if $i \neq 0$). Then

$$\Pr[N(\mathcal{I}) \geq 1\,|\,Y] \approx \frac{1}{\mathcal{S}} \sum_{j=1}^{\mathcal{S}} \left(1 - \delta_{N(\mathcal{I})^{(j)}}\right). \quad (14)$$

For certain intervals $\mathcal{I}$ that contain $M_j,\ \ j = 1, \ldots, J$, this probability can be taken to be the probability that SE_j takes place in $\mathcal{I}$.

To automate this procedure, define m_j as the jth local minima of $\lambda^*(t)\,|\,Y$ (also easily approximated on $\hat{\lambda}^*(t)\,|\,Y$). Define $m_0 = 0$ and $m_{J+1} = T$. Further, let $\mathcal{I}_j = (m_{j-1}, m_j]$, $j = 1, \ldots, J$. Now (13) and (14) can be estimated for each of the intervals $\mathcal{I}_j$. Furthermore, each $\mathcal{I}_j$ contains one and only one m_j. Therefore we tacitly assume that one and only one secretion event (SE_j) occurs in $\mathcal{I}_j$. Hence $\Pr[SE_j \in \mathcal{I}_j\,|\,Y] = \Pr[N(\mathcal{I}_j) \geq 1\,|\,Y]$ which, in turn, implies $E[N_{\mathrm{SE}}(24)\,|\,Y] = \sum_{j=1}^{J} \Pr[SE_j \in \mathcal{I}_j\,|\,Y]$ where $N_{\mathrm{SE}}(t)$ is the number of secretion events in (0, t].

The probabilities (14) for intervals $\mathcal{I}_j$, $j = 1, \ldots, 34$, are shown in the bottom half of Figure 2 for the data set examined in Section 4. The vertical lines denote the locations M_j, $j = 1, \ldots, J$ of the secretion events. The heights of the lines indicate the probabilities. The rug at the bottom of Figure 2 denotes the local minima. Applying this to the data set analyzed above we find $E[N(24)\,|\,Y] \approx 14.70$ while the expected number of secretion events is $E[N_{\mathrm{SE}}(24)\,|\,Y] \approx 12.67$.

Remark 5: We estimate the posterior rate function using the MDP model mainly to find the local maxima (which we take as the point estimates of the locations of secretion events) and local minima, and hence the intervals $\mathcal{I}_j$. Estimates of $E[N(\mathcal{I})\,|\,Y]$ can be obtained from the posterior sample without resorting to a smooth estimate of $\lambda^*(t)\,|\,Y$ [see (13), above]. The intervals $\mathcal{I}_j$ derived above may not be the optimal intervals in the following sense. Consider the fifth interval in Figure 2: $\mathcal{I}_5$ with end points 1:38 p.m. and 3:13 p.m. $\Pr(N(\mathcal{I}_5) \geq 1\,|\,Y) \approx 0.9955$. However, if $\mathcal{I}^*$ is the interval with end points 1:38 p.m. and 2:45 p.m., then $\Pr(N(\mathcal{I}^*) \geq 1\,|\,Y) \approx 0.9955$. Nevertheless, the intervals $\mathcal{I}_j$ are easily obtained in an automated way.

Remark 6: Parameter estimates from all other models to date are conditional parameter estimates. They are conditional on the number of secretion events. We have chosen to estimate parameters by model averaging [over models indexed by $N(T)$]. By doing so, we can estimate the posterior rate function as well as the expected number of pulse function arrivals and the expected number of secretion events. Furthermore, model averaging avoids underestimation of parameter uncertainty that results from choosing a single model.

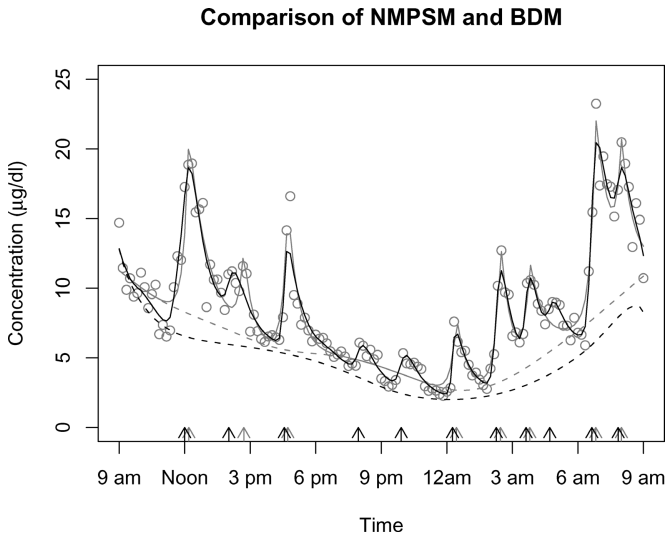

Figure 3. Comparison of the Bayesian deconvolution model and the NMPSM. Black denotes BDM, gray denotes NMPSM. The solid lines are the estimated concentration. The dashed lines are the estimated nuisance function from the two estimation procedures. The arrows at the bottom show point estimates of event locations.

We also compare the fit of our model/method with the recent model/method proposed by Yang et al. (2006). They model pulsatile hormone data with a NMPSM. To compare methods we need to threshold the probabilities of secretion events with an a priori threshold value. We choose 0.8. Thus if $\Pr[SE_j \in \mathcal{I}_j \,|\, Y] > 0.8$ we say that a secretion event has occurred at M_j in interval $\mathcal{I}_j$. Figure 3 displays fits from both models (black—BDM, gray—NMPSM). There are several notable differences. First, around 9:00 p.m., BDM finds two "secretion events that are not fit with NMPSM." There is one other "event," around 4:30 a.m., that is found by BDM and not by NMPSM. NMPSM does not detect the two around 9:00 p.m. because the heuristic procedure used to initialize event locations did not identify these two as events. The "event" at 4:30 a.m. was identified by the heuristic procedure as a potential event. After NMPSM is fit to the data, the final number of events is selected by either generalized cross validation, risk information criterion, Akaike information criterion, or Bayesian information criterion (BIC). Yang et al. (2006) recommends BIC. We chose BIC here, which did not select this event. The other difference with respect to events is the location of the "event" just prior to 3:00 p.m. BDM and NMPSM finds different locations for it. Note that the arrows at the bottom of this figure show point estimates of event locations. Point estimates from BDM always precede those of NMPSM. This is due to the fact that a point estimate from BDM is for the secretion event location and a point estimate from NMPSM is the MLE of the center of the pulse wave–form for the concentration (the convolution of secretion and elimination). As to what is the truth in this data set, we do not know. This comparison demonstrates what we said in the introduction. Models that rely on initial estimates of the number and location of events cannot recover from missed events.

6. Prior Sensitivity and a Simulation Study

To assess the sensitivity of the posterior distribution on the prior specification we doubled the standard deviations of the prior distributions while keeping the means fixed. The posterior mean (S.D.) of the half-life changed from 0.52 (0.13) to 0.63 (0.14). To compensate for this increase, the secretion rate function also changed. Numerical integration of the expected posterior secretion function, $\mathrm{E}\,(P(t)\,|\,Y)$, results in a change of the total amount of hormone secreted from 121.98 (23.17) to 104.36 (17.90). The expected posterior number of events changed from 12.67 to 12.95. These changes are not drastic but are expected because the deconvolution problem is ill-posed. We further changed the prior on the Poisson process mean from a $G(10,\ 1)$ to a $G(20,\ 1)$ and the prior on the number of knots, N_K, from a Negbin (3, 1) to a Negbin (6, 1). As expected, the posterior means of $N(T)$ and N_K increased. They changed from 14.70 (1.99) to 19.33 (2.96) and from 3.75 (1.25) to 5.94 (2.00), respectively. The posterior expected number of secretion events increased from 12.67 to 15.15. Given the a priori secretion event threshold of 0.8, one extra event was discovered. That event occurs just prior to noon (see Figure 2). The probability of that event increased from 0.52 to 0.92. These results are to be expected as the BDMCMC algorithm is sensitive to the prior specification of the number of entities (Stephens, 2000)—$N(T)$ and N_K in the present case.

We also conducted a small simulation study. We set the nuisance function $F(t) = 4\,\cos\,(2\pi\ t/24) + 6$. Ten event arrival times were drawn from the $3j$th, $j = 1, \ldots, 10$, order statistics from 32 uniform random variables on [0, 24]. Each of the ten pulses were then assigned an α_i and ν_i^2 by drawing $\ln\,(\alpha_i) \sim N(2,\ .25)$ and $\ln\,(\nu_i^2) \sim N(-3.5,\ .5)$. The half-life was set at 0.5. The resulting pulsatile function $P(t)$ was then convolved with exponential decay and added to $F(t)$. One hundred data sets were then generated by adding mean zero gaussian noise (S.D. = 0.1) to the log of the concentration at 10-minute intervals. All of these values were chosen to give realistic looking profiles.

For this set of simulated data sets, the sensitivity was 98.2% while the specificity was 99.7%, given the predetermined probability threshold of 0.8. We expect these numbers would go down as the model variance increases and/or the α_i decrease. An example of one simulated data set along with the true concentration profile is displayed in the upper left panel of Figure 4. The upper right panel shows the true concentration (black) and the posterior predictive means (gray) from 20 randomly selected simulations. The lower left panel shows $F(t)$ (black) and 20 randomly selected estimates (gray). The true secretion rate function (black) and 20 randomly selected estimates (gray) are shown in the lower right panel. Numerical results from the 100 simulations are displayed in Table 1 for various parameters of interest. We found that the marginal modes, or maximum a posteriori (MAP) estimates, resulted in less bias than the marginal means or medians due to skewness in the estimated posterior densities. In the interest of space, results for functions are given in terms of the integrated functions. Results from this simulation study suggest that the proposed model and estimation method perform well. Biases and root mean squared errors (RMSE) are not excessive. The

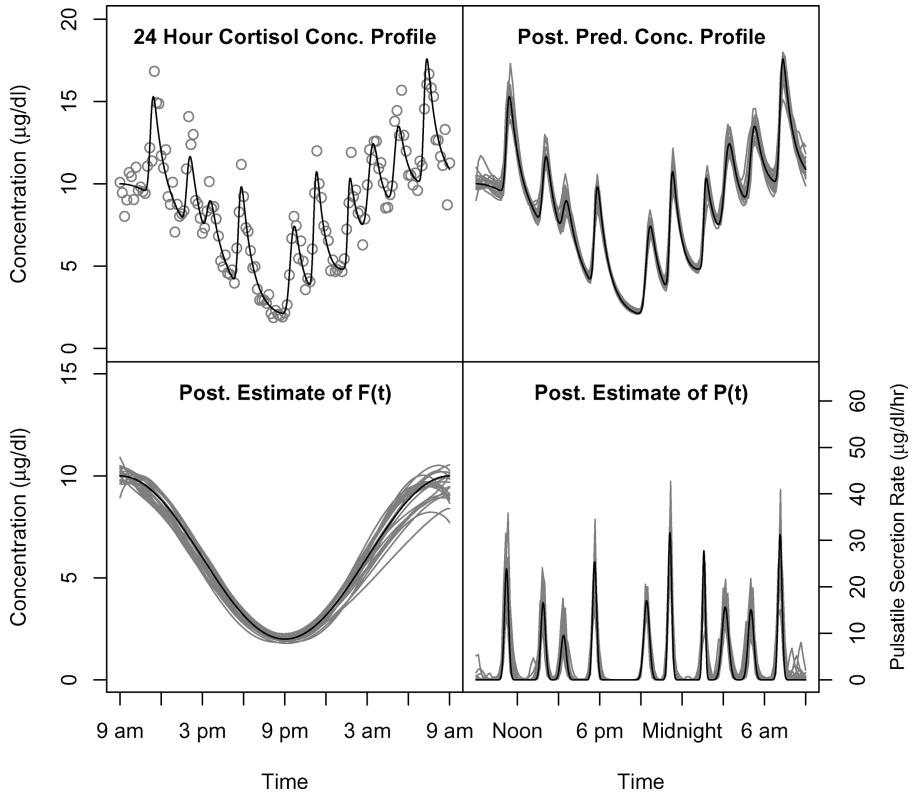

Figure 4. Simulation results. UL: The true concentration profile (black) and a randomly selected simulated concentration profile (gray). UR: True concentration and 20 randomly selected predicted concentration profiles (gray). LL: The true nuisance function (black) and 20 randomly selected nuisance functions (gray). LR: True secretion rate function (black) and all 20 randomly selected estimates (gray).

bias in the half-life translates to a 54-second bias in half-life. The bias in the number of secretion events $N_{\mathrm{SE}}(24)$ and in the integrated secretion rate function, $P(t)$, are to be expected because our estimation is based on model averaging and includes "false positives" in the estimate. However, the bias in the total amount of hormone secreted due to events is largely removed if we consider the average amount of hormone secreted per event—the last row of Table 1. We also note that all coverage rates of the 95% (equal tail area) credible intervals are within Monte Carlo simulation error (Monte Carlo 95% CI = (90.7,99.3)). Location and $N_{\mathrm{SE}}(24)$ are not amenable to the calculation of coverage rates and, thus, are not reported.

Table 1

Results from the simulation study for various parameters of interest. Results from event locations are averaged over the 10 *locations. The last row represents the average amount of hormone secreted per event. The last column represents the coverage rate of the* 95% *credible intervals. Location, $N_{SE}(24)$ and the amount of hormone secreted per event are not amenable to the calculation of coverage rates. The average MAP, RMSE, and bias are based on the* 100 *simulated data sets*

Parameter	Truth	Ave. MAP (S.D.)	Ave. RMSE	Ave. bias	% bias	% cov.
location	Varies	Varies	0.052	0.004	0.052	–
$t_{1/2}$	0.500	0.485 (0.060)	0.062	−0.015	−2.977	97.0
$N_{SE}(24)$	10.000	10.605 (0.379)	0.507	0.605	6.049	–
$\int_0^{24} C(t)\,dt$	204.376	205.382 (1.843)	2.091	1.006	0.492	96.0
$\int_0^{24} F(t)\,dt$	144.000	142.685 (6.011)	6.124	−1.315	−0.913	97.0
$\int_0^{24} P(t)\,dt$	84.586	89.115 (6.734)	8.088	4.529	5.354	91.0
$\int_0^{24} P(t)\,dt/N_{SE}(24)$	8.459	8.409 (0.652)	0.651	−0.049	−0.579	97.0

7. Discussion

Our method extends earlier work by Johnson (2003) in two significant aspects. First, we allow for a non-constant basal concentration. Second, we estimate probabilities of secretion events. We also extend the work by O'Sullivan and O'Sullivan (1988) by convolving a marked point process, we allow for biological variation in the size and shape of each pulse function. Third, O'Sullivan and O'Sullivan (1988) assume a zero basal concentration; we allow the basal concentration to be an arbitrary function that is approximated by a spline function. Lastly, our method contrasts with that of Yang et al. (2006): the BDM does not rely on an ad hoc method to initialize the number and location of events.

The BDM can easily incorporate other pulse function forms, such as a Laplacian or gamma. The only technical difficulty may be in the convolution of these two functions as it may not have an analytic solution. The convolution of a Gaussian and exponential decay does not have an analytic solution, but is proportional to the error function, erf, which can be efficiently computed (see Johnson, 2003).

A disadvantage of the BDM approach is that is it rather difficult to elicit prior information and to implement the advanced MCMC simulation (code is available from the author upon request). Further, because of the ill-posed nature of the deconvolution problem, uninformative prior specification and an objective Bayesian approach is not be feasible. Lastly, the Bayesian approach is computationally intensive. The MCMC simulation of the cortisol data set takes about 19 minutes on a dual 2.7 GHz PowerPC G5. Further the MDP postprocessing takes an additional 15 minutes.

We see the need to incorporate models that identify pulses and the subsequent second stage analyses (e.g., group comparison) into a single, coherent, joint modeling approach. A joint approach of this nature could go a long way in addressing some of the confounding issues that naturally arise in this problem. The idea of aligning the time-series by "time-warping" could be borrowed from functional data analysis. The warping would be used to align the oscillations in the basal concentration function as well as pulse locations between different data sets. Random effects could then be defined around fixed population effects. These effects would include the oscillatory nuisance function, the decay rate, number of pulses, and the amount of hormone release during each pulse. The down side, from a Bayesian perspective, would be the large amount of computing time necessary to simulate from the posterior. However, as large computing clusters become widely available, the solution of this problem will be feasible: the computation of subject-specific effects can be distributed across the nodes of the cluster. Population based effects can subsequently be updated.

Acknowledgements

I thank Dr Elizabeth A. Young of the Department of Psychiatry and Mental Health Research Institute, University of Michigan, for the use of the cortisol hormone data. I also thank Professor Yuedong Wang, Department of Applied Probability and Statistics, University of California, Santa Barbara, for analyzing the data set with NMPSM. Finally, I would like to thank the referee and associate editor for their helpful comments/suggestions. This work was funded by the NIH, grant P60 DK20572.

References

Antoniak, C. E. (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *The Annals of Statistics* **2,** 1152–1174.

Berne, R. M. and Levy, M. N., ed. (1993). *Physiology.* 3rd edition. St. Louis, Missouri: Mosby Year Book.

de Boor, C. (1978). *A Practical Guide to Splines.* New York: Springer-Verlag.

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.

Escobar, M. D. and West, M. (1995). Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association* **90,** 577–588.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *The Annals of Statistics* **1,** 209–230.

Gelfand, A. E., Dey, D. K., and Chang, H. (1992). Model determination using predictive distributions and implementation via sampling-base methods (with discussion). In *Bayesian Statistics 4*, J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (eds), 147–167. Oxford: Oxford University Press.

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.

Johnson, T. D. (2003). Bayesian deconvolution analysis of hormone concentration profiles. *Biometrics* **59,** 650–660.

Johnson, V. E. (2004). A Bayesian chi-squared test for goodness of fit. *Annals of Statistics* **32,** 2361–2384.

Kaipio, J. and Somersalo, E. (2005). *Statistical and Computational Inverse Problems.* New York: Springer.

Keenan, D. M. and Veldhuis, J. D. (2003). Cortisol feedback state governs adrenocorticotropin secretory-burst shapes, frequency, and mass in a dual-waveform construct: Time of day-dependent regulation. *American Journal of Physiology* **285,** R950–R961.

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.

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.

MacEachern, S. N. and Müller, P. (1998). Estimating mixtures of Dirichlet process models. *Journal of Computational and Graphical Statistics* **7,** 223–238.

Mauger, D. T. and Brown, M. B. (1995). A comparison of methods that characterize pulses in a time series. *Statistics in Medicine* **14,** 311–325.

Merriam, G. R. and Wachter, K. W. (1982). Algorithms for the study of episodic hormone secretion. *American Journal of Physiology* **243,** 310–318.

Neal, R. M. (2000). Markov chain sampling methods for Dirichlet process mixture models. *Journal of Computational and Graphical Statistics* **9,** 249–265.

Oerter, K. E., Guardabasso, V., and Rodbard, D. (1986). Detection and characterization of peaks and estimation of instantaneous secretory rate for episodic pulsatile hormone secretion. *Computers and Biomedical Research* **19,** 170–191.

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.

Rodbard, D., Rayford, P. L., and Ross, G. T. (1970). Statistical quality control. In *Statistics in Endocrinology*, J. W. McArthur, and T., Colton, (eds), 411–429. Cambridge, Massachusetts: The MIT Press.

Stephens, M. (2000). Bayesian analysis of mixture models with an unknown number of components—an alternative to reversible jump methods. *Annals of Statistics* **28,** 40–74.

Van Cauter, E. L. et al. (1981). Quantitative analysis of spontaneous variation in plasma prolactin in normal man. *American Journal of Physiology* **241,** 355–363.

Veldhuis, J. D. and Johnson, M. L. (1986). Cluster analysis: A simple, versatile and robust algorithm for endocrine pulse detection. *American Journal of Physiology* **250,** 486–493.

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., Iranamesh, A., Lizarralde, G., and Johnson, M. L. (1989). Amplitude modulation of a burstlike mode of cortisol secretion subserves the circadian glucocortocid rhythm in man. *American Journal of Physiology* **257,** 6–14.

Wahba, G. (1990). *Spline Models for Observational Data*. Philadelphia: Society for Industrial and Applied Mathematics.

Yang, Y., Liu, A., and Wang, Y. (2006). Detecting pulsatile hormone secretions using nonlinear mixed effects partial spline models. *Biometrics* **62**, 230–238.

Young, E. A., Carlson, N. E., and Brown, M. B. (2001). Twenty-four-hour ACTH and Cortisol pulsatility in depressed women. *Neuropsychopharmacology* **25,** 267–276.

Received April 2006. *Revised January* 2007.
Accepted January 2007.