Detecting Pulsatile Hormone Secretions Using Nonlinear Mixed Effects Partial Spline Models

Yu-Chieh Yang
Department of Statistics, National Taichung Institute of Technology, Taiwan.
email: yuchieh@ntit.edu.tw
Anna Liu
Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003.
email: anna@math.umass.edu
Yuedong Wang
Department of Statistics and Applied Probability, University of California, Santa Barbara, CA 93106.
email: yuedong@pstat.ucsb.edu
September 27, 2004

Abstract

Neuroendocrine ensembles communicate with their remote and proximal target cells via an intermittent pattern of chemical signaling. The identification of episodic releases of hormonal pulse signals constitutes a major emphasis of endocrine investigation. Estimating the number, temporal locations, secretion rate and elimination rate from hormone concentration measurements is of critical importance in endocrinology. In this paper, we propose a new flexible statistical method for pulse detection based on nonlinear mixed effects partial spline models. We model pulsatile secretions using biophysical models and investigate biological variation between pulses using random effects. Pooling information from different pulses provides more efficient and stable estimation for parameters of interest. We combine all nuisance parameters including a non-constant basal secretion rate and biological variations into a baseline function which is modeled nonparametrically using smoothing splines. We develop model selection and parameter estimation methods for the general nonlinear mixed effects partial spline models and a R package for pulse detection and estimation. We evaluate performance and the benefit of shrinkage by simulations and apply our methods to data from a medical experiment.

Key words: endocrinology, hormone data, model selection, random effects, semi-parametric nonlinear mixed effects model, shrinkage, smoothing spline.

1 Introduction

Hormones play an important role in regulating biological processes. Through secretions of hormones, signals are sent to other organs enabling interaction within the human body (Keener and Sneyd 1998). There are two types of secretions: pulsatile secretions which are bursts of hormone from glands to bloodstream, and basal secretion which is a tonic pattern secretion (Merriam and Wachter 1982, Veldhuis, Iranmanesh, Lizarralde and Johnson 1989, Hartman, Vance, Johnson, Thorner and Veldhuis 1991, Keenan and Veldhuis 1997, Winer, Shaw and Baumann 1990, Guo, Wang and Brown 1999). Since pulses act as signals to target organs for physiological communication within the endocrine system, it is biologically and clinically important to investigate the occurrence and/or frequency of pulses. The identification of discrete hormonal pulse signals constitutes a major emphasis of endocrine investigation (Van Cauter, L'Hermite, Copinschi, Refetoff, Desir and Robyn 1981, Merriam and Wachter 1982, Clifton and Steiner 1983, Veldhuis and Johnson 1986, Oerter, Guardabasso and Rodbard 1986, Veldhuis, Carlson and Johnson 1987, Munson and Rodbard 1989, Diggle and Zeger 1989, Bolstad 1988, O'Sullivan and O'Sullivan 1988, Kushler and Brown 1991, Komaki 1993, Guo et al. 1999, Johnson 2003).

Experiments are typically conducted in such a way that some hormone concentrations are measured from blood samples withdrawn at regular time intervals, say every 10 minutes, for a period of time, say 24 hours, from a group of normal (or sick) human subjects (or animals). For example, in an experiment conducted in the University of Michigan, 10-minute sampling for hormones adrenocorticotropic (ACTH) and cortisol was performed for 24 hours in 36 patients with fibromyalgia and/or chronic fatigue syndrome and 36 age-matched controls (Crofford, Young, Engleberg, Korszun, Brucksch, McClure, Brown and Demitrack 2004). Figure 1 shows the profile of ACTH concentrations over time from a patient. Pulse locations and a baseline function are estimated by the methods proposed in this paper.

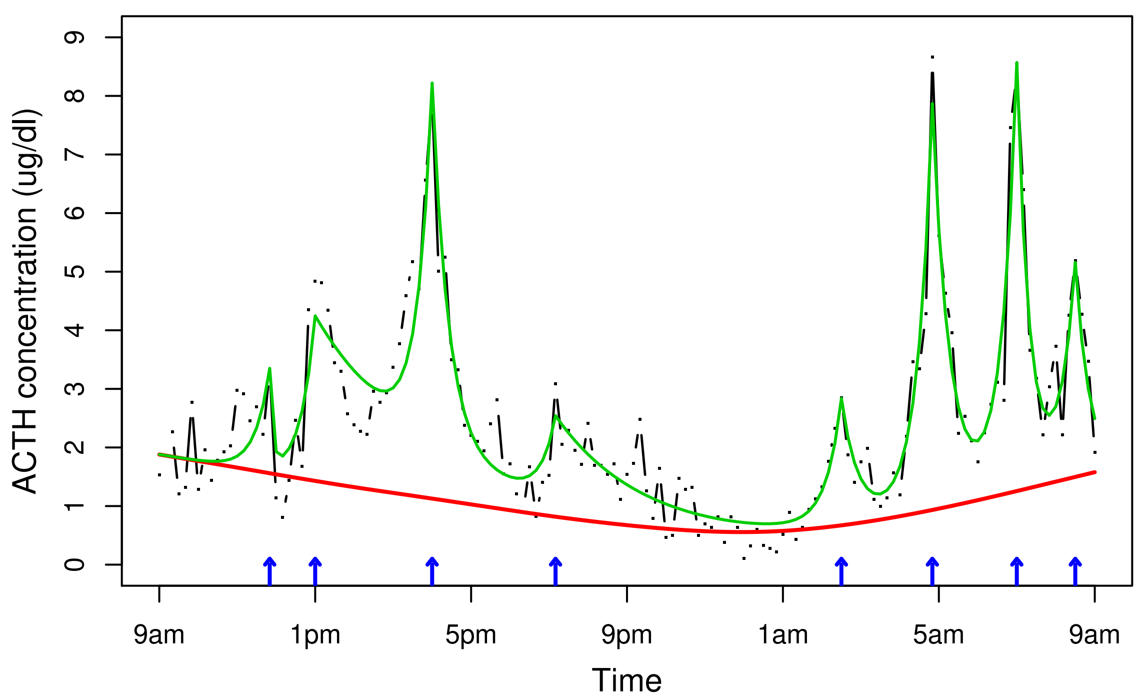


Figure 1: Profile of ACTH concentration of a patient with chronic fatigue syndrome (broken line). Pulse locations identified by our method with the BIC criterion are marked below as vertical arrows. The overall fit and estimated baseline function are plotted as solid lines.

The goal of the study was to investigate disease effects, if any, on the secretion pattern. Statistical problems at the first stage of the data analysis are to estimate the number and locations of pulses, parameters such as the mass (amplitude) and half-life associated with each pulse and the baseline (Crofford et al. 2004). These problems are technically challenging due to indirect observations, near confounding among several components and multiple sources of variation. "There are many proposed pulse-detection algorithms, all based upon trying to detect a point of rapid increase, but none has proven to be completely acceptable" (Keenan, Sun and Veldhuis 2000).

Existing methods for pulse identification and characterization fall into two categories: criterion-based methods which use test statistics to identify rises and/or falls in hormone concentration, and model-based methods which assume statistical models to approximate the secretion pattern. Among criterion-based methods, CLUSTER compares the concentrations at a peak location with concentrations at the nearest nadir using a two-sample t-test (Veldhuis and Johnson 1986). PULSAR removes the baseline by a robust linear regression and identifies pulses as outliers from the residuals (Merriam and Wachter 1982). Other criterion-based methods include DETECT (Oerter et al. 1986) and ULTRA (Van Cauter et al. 1981). In general, most criterion-based methods use the assay's coefficient of variance (CV) as the true CV. Other sources of variation such as biological noises are ignored. Therefore estimates of quantities related to variation such as the threshold are biased which leads to over-identifying the numbers of pulses. Among the model-based methods, Veldhuis et al. (1987) used a biophysical model that represents hormone concentration as the convolution of a secretion rate with an elimination function (see Section 2.1 for more details). Diggle and Zeger (1989) used a non-Gaussian autoregressive model that incorporates feedback. O'Sullivan and O'Sullivan (1988) represented hormone concentration as a convolution of individual pulses with their locations following a non-homogeneous Poisson process. Guo et al. (1999) proposed a state-space model that incorporates a non-constant baseline. Other model-based methods include Bolstad (1988), Kushler and Brown (1991), Komaki (1993) and Johnson (2003). In general, the model-based methods are preferred to criterion-based methods based on the false positive and false negative error rates (Mauger, Brown and Kushler 1995). Model-based methods also provide estimates for the parameters of interest. Criterion-based methods are often used to identify initial pulses for model-based methods.

In this paper we propose nonlinear mixed effects partial spline models to detect pulse locations and estimate parameters. All current model-based methods except Guo et al. (1999) assume a constant or zero baseline. These restrictive assumptions may lead to biases in estimates of the parameters. We combine all nuisance parameters into a baseline function and model it nonparametrically using smoothing splines. All current model-based methods assume that parameters such as the decay rate are fixed and common for all pulses. Thus all these methods ignore biological variations between pulses within a subject which may be of scientific interest (Keenan, Roelfsema, Biermasz and Veldhuis 2003, Keenan and Veldhuis 2003). We introduce a general second stage mixed effects model for parameters which allow us to model variation between pulses and incorporate covariate effects and/or feedback mechanisms. Pooling information from different pulses, our estimates have smaller mean-squared errors (MSE). We also allow random errors to be correlated. We develop an estimation procedure for a general form of pulse shape function. Therefore our methods and software can be applied to fit models with several different pulse shape functions in the literature.

The article is organized as follows. Section 2 introduces the nonlinear mixed effects partial spline model. Section 3 describes methods for pulse detection and parameter estimation. Section 4 presents simulation results. Section 5 presents the analysis of the data set introduced in Section 1. Section 6 concludes the article with a brief discussion.

2 Nonlinear Mixed Effects Partial Spline Model

2.1 Biophysical models for hormonal secretions and measurements

The basic construct consists of four components at different levels: *secretion*, *elimination*, *concentration* and *observation*.

(a) *Secretion.* In vivo and in vitro studies of endocrine glands suggest the existence of two physiologically distinguishable modes of secretion: slowly varying *basal* release and secretagogue-driven *pulsatile* release. *Basal* secretion represents a slow rate of granule leakage or non-regular release not directly triggered by a pulse stimulus. It is continuous and changes slowly over time. It is usually modeled by a constant (Keenan and Veldhuis 1997) or a smooth function (Guo et al. 1999). A *pulse* is defined by its *onset time* , *waveform* and *mass*:

1. *Onset times.* Conceptually, they represent the times that the neuroendocrine gland sends signals to activate the biological synthesis of hormones and to transport granules and individual molecules to the cell membrane for secretion. Usually observations are taken in a time period, typically 24 hours. Without loss of generality, we assume that the time period has been transformed into an interval $[0, 1]$. We denote the number of pulsatile secretions as K and successive onset times as $\tau_1 < \tau_2 < \cdots < \tau_K$. Note that they are unknown and usually are the main interest of estimation.
2. *Waveform.* Conceptually, a burst reflects abrupt exocytotic discharge of hormone-containing granules followed by less rapid secretion of newly synthesized molecules (Keenan and Veldhuis 1997). Different functions have been used to model the waveform: a Gaussian density when it is symmetric, a convolution of a Gaussian with an exponential, Weibull and Gamma families (Keenan and Veldhuis 1997). The most popular choice is the following generalized-Gamma family of densities (Keenan and Veldhuis 1997),

$$\psi(t) = \frac{\rho_3}{\Gamma(\rho_1)\rho_2^{\rho_1\rho_3}} t^{\rho_1\rho_3-1} \mathrm{e}^{-(t/\rho_2)^{\rho_3}}, \qquad t \geq 0,\ \rho_1 > 1,\ \rho_2 > 0,\ \rho_3 > 0, \tag{1}$$

   with $\psi(t) = 0$ for $t < 0$. This family of models is very flexible, and includes most of the above functions as special cases. The function (1) is shown in Figure 2(a) with a particular combination of parameters ρ_1, ρ_2 and ρ_3.
3. *Mass* represents the amount of mass accumulated from the last pulse. It equals the area under the waveform curve (Figure 2(a)). Denote α_k as the mass of pulse k, $k = 1, \cdots, K$.

The total secretion rate is then

$$S(t) = \rho(t) + \sum_{k=1}^{K} \alpha_k \psi(t - \tau_k), \tag{2}$$

where $\rho(t)$ is the rate of basal secretion. The main interest is usually in the pulsatile secretions. Nevertheless, "accurately partitioning basal versus pulsatile hormone secretion, albeit technically challenging, should aid in separating normal physiology (low basal release), pathophysiology (jointly increased basal and pulsatile secretion), and pathology (elevated basal hormone production)" (Keenan and Veldhuis 2000).

(b) *Elimination* represents hormone distribution within the vascular space and metabolic removal of the hormone from the blood. This component is usually modeled by a monoexponential model,

$$E(t) = \exp(-\xi t), \tag{3}$$

or a biexponential model,

$$E(t) = \delta \exp(-\xi_1 t) + (1 - \delta) \exp(-\xi_2 t), \tag{4}$$

where two modes of elimination due to diffusion (rapid) and metabolic (slow) are separated with δ as the proportion. The monoexpontial function is shown in Figure 2(b) with a particular choice of ξ.

(c) *Concentration* of a hormone at time t, $X(t)$, can be represented by (Keenan, Veldhuis and Yang 1998)

$$\begin{aligned} X(t) &= X(0)E(t) + \int_0^t S(u)E(t-u)du + g(t) \\ &= X(0)E(t) + \int_0^t \rho(u)E(t-u)du + g(t) + \sum_{k=1}^{K} \alpha_k \int_0^t \psi(u - \tau_k)E(t-u)du, \end{aligned} \tag{5}$$

where $X(0)$ is the concentration at time 0, and g represents microspic biological variation. The central part of the model is convolutions of pulse waveforms with elimination functions. The convolution of two curves in Figures 2(a) and 2(b) is shown in Figure 2(c).

(d) *Observations* are measurements from blood samples drawn at discrete time points, say every 10 minutes. To account for measurement errors and other source of variations,

$$y_j = X(t_j) + \epsilon_j, \quad j = 1, \cdots, n, \tag{6}$$

where ϵ_j are usually assumed to be independent normal with mean zero and a constant variance or a constant coefficient of variation. We allow random errors to be correlated in this paper.

Technical challenges include: (1) the number and locations of onset times are not observed, and (2) two modes of secretions and elimination are near confounded. Even for the special case

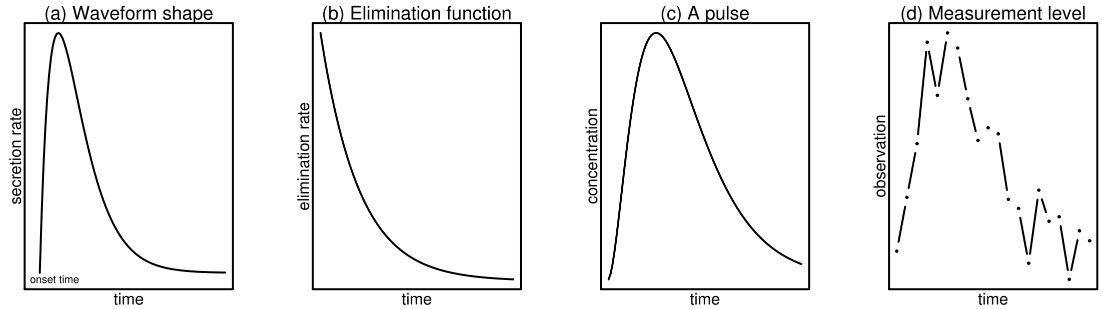


Figure 2: An illustration of the waveform, elimination function, convolution and measurements.

with $X(0) = 0$, $g(t) = 0$, $\rho(t) = 0$, $\alpha_k = 1$, and the assumption that onset times follow a non-homogeneous Poisson process with intensity function $h(t)$, the expected concentration, $\mathrm{E}(X(t)) = \int_0^t \left(\int_0^u \psi(u-v)h(v)dv\right) E(t-u)du$, involves two layers of convolutions. Thus the estimation of the intensity function h involves two layers of deconvolutions where the filter functions ψ and E depend on unknown parameters. The one-layer deconvolution with a single known filter function is a well-known ill-posed problem (Wahba 1990).

2.2 Nonlinear partial spline models

All current methods except Guo et al. (1999) assume that the basal secretion rate $\rho(t)$ is a constant function. Some methods even require that $\rho(t) = 0$. We assume that $\rho(t)$ is a smooth function and treat it as a nuisance parameter. Keenan et al. (1998) used an one fold integrated Wiener process to model microspic biological variation g in (5) which is equivalent to a linear spline (Wahba 1990). We note that both $X(0)$ and g are unknown. They are nuisance parameters and are ignored by most of the current methods. These restrictive assumptions and omissions may lead to large bias in estimates of the parameters. We combine all three nuisance parameters into a *baseline function*

$$f(t) = X(0)E(t) + \int_0^t \rho(u)E(t-u)du + g(t). \tag{7}$$

We then consider the following general class of nonlinear partial spline models

$$y_i = f(t_i) + \sum_{k=1}^{K} \alpha_k p(\boldsymbol{\gamma}_k; t_i - \tau_k) + \epsilon_i, \qquad i = 1, \cdots, n, \tag{8}$$

where y_i is the concentration measurement at time t_i, f is the baseline function, $p(\boldsymbol{\gamma}; \cdot)$ is the *pulse shape* function with parameters $\boldsymbol{\gamma}$, K is the number of pulses, α_k, $\boldsymbol{\gamma}_k$ and τ_k are the mass (or amplitude), pulse shape parameters and onset (or peak) times associate with the kth pulse, and ϵ_i's are random errors. We allow random errors to be correlated. Specifically, let $\boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_n)^T$. We assume that $\boldsymbol{\epsilon} \sim N(0, \sigma^2\boldsymbol{\Lambda})$. We now discuss how to model the pulse shape p and the baseline function f.

We will consider general pulse shape function p in our estimation procedure and software implementation. Several prototype pulse shape functions are used in the literature. For example, from the basic construction described in Section 2.1,

$$p(\boldsymbol{\gamma}; t-\tau) = \int_0^t \psi(u-\tau)E(t-u)du = \int_\tau^t \psi(u-\tau)E(t-u)du = \int_0^{t-\tau} \psi(u)E(t-\tau-u)du, \tag{9}$$

where τ is the most recent onset time before t, $\boldsymbol{\gamma}$ collects parameters ρ_1, ρ_2, ρ_3 in (1) and ξ in (3) or ξ_1, ξ_2 in (4). Another simple and useful pulse shape function is the following double exponential pulse function (O'Sullivan and O'Sullivan 1988)

$$p(\boldsymbol{\gamma}; t-\tau) = \begin{cases} \exp\{\gamma_1(t-\tau)\}, & t < \tau, \\ \exp\{-\gamma_2(t-\tau)\}, & t \geq \tau. \end{cases} \tag{10}$$

For the double exponential pulse function, τ_k and α_k represent peak time and amplitude of the kth pulse. In practice the ability to distinguish between different pulse shape functions is limited by the sampling rate. The double exponential pulse functions usually provide good approximations. Therefore, even though our methods apply to the general pulse functions, we use the double exponential pulse function in our simulations and data analysis.

As indicated in (7) that the baseline function f combines all nuisance parameters. It is reasonable to assume that f vary slowly over time. However, it is usually difficult, if not impossible, to specify a parametric model for f. Thus we model it nonparametrically using a polynomial spline with the model space (Wahba 1990, Green and Silverman 1994)

$$W_m = \left\{ f: \ f, f', \cdots, f^{(m-1)} \text{ absolutely continuous}, \int_0^1 (f^{(m)})^2 dt < \infty \right\}. \tag{11}$$

$m = 2$ corresponds to the well-known cubic spline which is used in our simulations and data analysis. We note that our methods apply to general spline models defined in Wahba (1990).

2.3 Mixed effects models for parameters

Keenan et al. (1998) provided biological justifications for modeling the mass parameter as random effects:

$$\alpha_k = \beta_1 + b_k, \quad b_k \overset{iid}{\sim} N(0, \sigma_b^2), \tag{12}$$

where random effects b_k model the biological variation. To allow mass to depend on the preceding inter-pulse interval, we may add $\beta_2 \times (\tau_k - \tau_{k-1})$ in model (12) which assumes a constant rate of mass accumulation (Keenan et al. 1998, Keenan et al. 2003). Pulses may also be modulated by circadian rhythms (Keenan and Veldhuis 1997). Specifically, masses vary in a systematic circadian pattern. To quantify the underlying pulsatile secretion-generating mechanisms, we may add a simple periodic function such as $\beta_3 \sin 2\pi\tau_k + \beta_4 \cos 2\pi\tau_k$ to model (12). An alternative approach is to regard b_k as observations of a stochastic process, $b(t)$, evaluated at time points τ_k. For example, we may assume $b(t)$ as a stochastic process corresponds to the periodic spline (Wahba 1990). More complicated mixed effects models require more observations on pulses to get accurate estimates. In this paper we consider the simple model (12) only. We will develop integrated models for all subjects with more complicated mixed effects models in the future.

All existing methods ignore variations in the shape parameters and assume that $\boldsymbol{\gamma}_k = \boldsymbol{\gamma}$ for all pulses within a subject. Recent studies indicate that $\boldsymbol{\gamma}_k$ may vary during the day (Keenan et al. 2003, Keenan and Veldhuis 2003). It is of scientific interest to model the variation between pulses. Random effects models discussed above for the mass parameter can also be constructed similarly for $\boldsymbol{\gamma}_k$.

In this paper, we consider a general second stage model for the mass α_k and shape parameters $\boldsymbol{\gamma}_k$ which include models discussed above as special cases. Let $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_K)^T$, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^T, \cdots, \boldsymbol{\gamma}_K^T)^T$, and $\boldsymbol{\phi} = (\boldsymbol{\alpha}^T, \boldsymbol{\gamma}^T)^T$. Then we assume the following linear mixed model for all parameters $\boldsymbol{\phi}$:

$$\boldsymbol{\phi} = \boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{B}\boldsymbol{b}, \qquad \boldsymbol{b} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{D}), \tag{13}$$

where $\boldsymbol{\beta}$ and $\boldsymbol{b}$ are fixed and random effects, and $\boldsymbol{A}$ and $\boldsymbol{B}$ are design matrices for the fixed and random effects respectively.

3 Pulse Detection and Estimation

The nonlinear mixed effects partial spline model (NMPSM) is the combination of the first stage model (8) and the second stage model (13). Let $\boldsymbol{y} = (y_1, \cdots, y_n)^T$, $\boldsymbol{f} = (f(t_1), \cdots, f(t_n))^T$, and $\boldsymbol{\eta} = (\sum_{k=1}^K \alpha_k p(\boldsymbol{\gamma}_k; t_1 - \tau_k), \cdots, \sum_{k=1}^K \alpha_k p(\boldsymbol{\gamma}_k; t_n - \tau_k))^T$. Then the NMPSM can be written in a matrix form

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{f} + \boldsymbol{\eta} + \boldsymbol{\epsilon}, \qquad \boldsymbol{\epsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{\Lambda}), \\ \boldsymbol{\phi} &= \boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{B}\boldsymbol{b}, \qquad \boldsymbol{b} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{D}). \end{aligned} \tag{14}$$

In the following we assume that $\boldsymbol{\Lambda}$ and $\boldsymbol{D}$ depend on an unknown parameter vector $\boldsymbol{\theta}$. We need to estimate the number of pulses K, pulses locations $\boldsymbol{\tau} = (\tau_1, \cdots, \tau_K)^T$, $\boldsymbol{\beta}$, f, $\boldsymbol{\theta}$, σ^2 and $\boldsymbol{b}$. Since the total number of parameters depends on the unknown parameter K, it is difficult to estimate all the parameters simultaneously. Our estimation procedure consists of two stages:

- *pulse detection*: estimate K and $\boldsymbol{\tau}$.
- *parameter estimation*: conditioned on the estimates of K and $\boldsymbol{\tau}$, estimate $\boldsymbol{\beta}$, f, $\boldsymbol{\theta}$, σ^2 and $\boldsymbol{b}$.

3.1 Parameter estimation and inference

We present the second stage of our estimation procedure first. At this stage, we assume that K and $\boldsymbol{\tau}$ are known and develop methods for estimating $\boldsymbol{\beta}$, f, $\boldsymbol{\theta}$, σ^2 and $\boldsymbol{b}$. The NMPSM (14) is a special case of the semi-parametric nonlinear mixed effects models (SNMM) proposed in Ke and Wang (2001). Thus the same estimation method can be used. Specifically, the estimation procedure iterates between two steps. At the first step, for fixed σ^2 and $\boldsymbol{\theta}$, we estimate $\boldsymbol{\beta}$, f and $\boldsymbol{b}$ by minimizing the following double-penalized log-likelihood

$$\min_{f \in W_m, \boldsymbol{\beta}, \boldsymbol{b}} \left\{ (\boldsymbol{y} - \boldsymbol{f} - \boldsymbol{\eta})^T \boldsymbol{\Lambda}^{-1} (\boldsymbol{y} - \boldsymbol{f} - \boldsymbol{\eta}) + \boldsymbol{b}^T \boldsymbol{D}^{-1} \boldsymbol{b} + n\lambda \int_0^1 (f^{(m)}(u))^2 du \right\}, \tag{15}$$

where the first two terms are the Laplace approximation to the log-likelihood of the NMPSM, the third term is a penalty to the roughness of the function f, and λ is a smoothing parameter which controls the trade-off between the goodness-of-fit and the smoothness of the function f. We choose λ using a data-adaptive criterion such as the generalized cross validation (GCV) and generalized maximum likelihood (GML) methods (Wahba 1990, Ke and Wang 2001, Wang and Ke 2002).

At the second step, fixing $\boldsymbol{\beta}$, f and $\boldsymbol{b}$ as their current estimates $\boldsymbol{\beta}_{\_}$, $f_{\_}$ and $\boldsymbol{b}_{\_}$, we estimate $\boldsymbol{\theta}$ and σ^2 by maximizing the approximate profile-likelihood

$$\log|\sigma^2 \boldsymbol{V}_{\_}| + \sigma^{-2}(\boldsymbol{y} - \boldsymbol{f}_{\_} - \boldsymbol{\eta}_{\_} + \boldsymbol{Z}_{\_}\boldsymbol{b}_{\_})^T \boldsymbol{V}_{\_}^{-1}(\boldsymbol{y} - \boldsymbol{f}_{\_} - \boldsymbol{\eta}_{\_} + \boldsymbol{Z}_{\_}\boldsymbol{b}_{\_}), \tag{16}$$

where $\boldsymbol{V}_{\_} = \boldsymbol{\Lambda} + \boldsymbol{Z}_{\_}\boldsymbol{D}\boldsymbol{Z}_{\_}^T$ and $\boldsymbol{Z}_{\_} = \partial\boldsymbol{\eta}/\partial\boldsymbol{b} \mid_{\boldsymbol{\beta}_{\_},\boldsymbol{b}_{\_}}$. Detailed implementation of this procedure can be found in Ke and Wang (2001).

Inferences on parameters, random effects and the non-parametric baseline function are based on a linear mixed effects partial spline approximation at convergence. Specifically, denote $\hat{\boldsymbol{\phi}}$ as the estimate of $\boldsymbol{\phi}$ and $\hat{\boldsymbol{X}} = \partial\eta/\partial\boldsymbol{\phi}|_{\hat{\boldsymbol{\phi}}}$. Then at convergence, we approximate model (14) by $\boldsymbol{y} \approx \boldsymbol{f} + \eta(\hat{\boldsymbol{\phi}}) + \hat{\boldsymbol{X}}(\boldsymbol{\phi} - \hat{\boldsymbol{\phi}}) + \boldsymbol{\epsilon}$. Let $\tilde{\boldsymbol{y}} = \boldsymbol{y} - \eta(\hat{\boldsymbol{\phi}}) + \hat{\boldsymbol{X}}\hat{\boldsymbol{\phi}}$. Then we approximate the original NMPSM (14) by the following linear mixed effects partial spline model

$$\tilde{\boldsymbol{y}} = \boldsymbol{f} + \hat{\boldsymbol{X}}\boldsymbol{A}\boldsymbol{\beta} + \hat{\boldsymbol{X}}\boldsymbol{B}\boldsymbol{b} + \boldsymbol{\epsilon}. \tag{17}$$

Combining the parametric fixed effects, $\hat{\boldsymbol{X}}\boldsymbol{A}\boldsymbol{\beta}$, with the bases of the null space of f, model (17) is a special case of the nonparametric mixed effects model in Wang (1998) and Wang and Ke (2002). Covariance matrices of the BLUP estimates given in Theorem 1 of Wang (1998) are used for inferences.

3.2 Pulse detection

We now present the first stage of our estimation procedure. The number of pulses is never known in practice. We propose methods for estimating K and $\boldsymbol{\tau}$ in this subsection. This stage of our estimation procedure consists of two phases:

phase 1: identify potential pulse locations.

phase 2: creat a nested sequence by eliminating pulses one-by-one and then decide the final model.

At the first phase, we want to find all possible pulses and not to be concerned with false identifications. Many detection methods are available in the endocrinology and statistical literature. One may use any existing pulse detection method such as the CLUSTER method (Veldhuis and Johnson 1986). When pulse locations are peaks of the double exponential function (10), the mean function has change points in the first derivative at these positions. Thus existing methods for detecting change points in the first derivative can also be used (Yang 2002). Simulations (not shown) indicate that both CLUSTER and change points methods perform well. The change point method tend to have smaller false negative rate. Therefore, it is used in our simulations and data analysis. Other methods such as wavelet and local polynomials may also be used (Yang 2002). We note that users can always add or eliminate pulse locations at this phase based on visual inspection. More details, R functions and examples can be found in Yang, Liu and Wang (2004).

Let the number of potential pulses identified in the first step as K_{max}. Denote the minimal number of pulses as K_{min}. A simple choice of K_{min} is zero. In phase 2, we create a nested sequence of pulse locations by fitting the NMPSM (14) and eliminating the least significant pulse location one-by-one from K_{max} to K_{min}. We then select the final model using a model selection criterion.

We now discuss model selection methods involved phase 2 in some detail. For a fixed K, $K_{min} \leq K \leq K_{max}$, we fit the NMPSM (14) using the method discussed in Section 3.1. We define t-statistics

$$t_k = \hat{\alpha}_k / \sqrt{\hat{\text{var}}(\hat{\alpha}_k)}, \quad k = 1, \cdots, K,$$

where $\hat{\text{var}}(\hat{\alpha}_k)$ is the approximate variance of $\hat{\alpha}_k$ after linearization. Specially, we compute $\hat{\alpha}_k$ using Theorem 1 in Wang (1998) based on the approximated linear mixed effects partial spline model (17). We then eliminate the pulse location with the smallest $|t_k|$. Simulations in Section 4 indicate that this simple procedure works very well: false pulse locations are correctly eliminated before true pulse locations in most simulations.

Denote models correspond to the resulting nested sequence of pulse locations as $\mathcal{M}_{K_{min}}, \cdots, \mathcal{M}_{K_{max}}$. We need to select the final model among this sequence of models. Model (8) contains two additive components: the nonparametric baseline function and the parametric pulses. Usually as K increases, the complexity of the parametric component increases while the complexity required for f decreases. Therefore both λ and K act as tuning parameters and they usually compensate for each other. Although pulse locations of the sequence created in the first step are nested, $\mathcal{M}_{K_{min}}, \cdots, \mathcal{M}_{K_{max}}$ are not necessary nested since λ are different for different K. Consequently, the residual sum of squares is not necessarily decreasing as K increases. We will select the final model using a model selection criterion such as the Akaike information criterion (AIC) (Akaike 1973), Bayesian information criterion (BIC) (Schwarz 1978), risk inflation criterion (RIC) (Foster and George 1994) and GCV (Craven and Wahba 1979). To be able to use these model selection procedures, we need to define a measures of complexity for the model $\mathcal{M}_K$, $K = K_{min}, \cdots, K_{max}$. For an additive model, it is reasonable to take the addition of degrees of freedom for each component as a measure of complexity. However, when a selection procedure is involved in the estimation, extra degrees of freedom are required (Hinkley 1971, Friedman and Silverman 1989, Friedman 1991, Luo and Wahba 1997). Let $\tilde{\boldsymbol{H}}(\hat{\lambda})$ be the smoother matrix for the nonparametric function f where $\hat{\lambda}$ is an estimate of λ by the GCV or the GML method (Wahba 1990, Wang and Ke 2002). A commonly used measure of complexity for f is $\text{tr}\tilde{\boldsymbol{H}}(\hat{\lambda})$. Let $df_P(K)$ be the number of parameters associated with the pulses. As in Luo and Wahba (1997), we define an inflated degrees of freedom, IDF, to account for the extra cost for selecting pulse locations. Specifically, we define the total degrees of freedom for $\mathcal{M}_K$ as

$$df_K \equiv \text{tr}\tilde{\boldsymbol{H}}(\hat{\lambda}) + \text{IDF} \times df_P(K). \tag{18}$$

Simulations show that a good choice of IDF is around 1.2, the same value suggested in Luo and Wahba (1997). IDF = 1, that is no inflation, leads to poor performance.

Let $RSS(K)$ be the residual sum of squares of model $\mathcal{M}_K$. Note that $RSS(K)$ depends on both K and λ. For a fixed K, as discussed in Section 3.1, we estimate λ by a data-driven method such as the GCV or GML method. Therefore, $\hat{\lambda}$ depends on K and λ is essentially profiled in $RSS(K)$. Now consider the following selection criteria

$$RSS(K) + a\sigma^2 df_K, \tag{19}$$

where $a = 2$, $a = \log n$, and $a = 2\log df_{K_{max}}$ correspond to the AIC, BIC and RIC criteria respectively. We estimate σ^2 based on the biggest model with $K = K_{max}$. The GCV criterion is defined as

$$RSS(K)/(1 - df_K/n)^2.$$

Estimate of K is the minimizer of one of those criteria which also decides $\boldsymbol{\tau}$ and the final model. Simulations show that all four model selections procedures work very well. BIC and RIC perform slightly better.

3.3 Algorithm

Combining all steps in two stages, we have the following algorithm.

1. *Initialize*: identify potential pulse locations and provide initial values. Denote the total number of potential pulses as K_{max}. Specify a low bound for the number of pulses K_{min}.

2. *Pulse detection*:

   (a) For $K = K_{max}, K_{max} - 1, \cdots, K_{min}$, repeat

      i. fit the model (14) and compute t-statistics t_k, $k = 1, \cdots, K$.

      ii. delete the location with the smallest $|t_k|$.

   (b) Select the final model using one the AIC, BIC, RIC and GCV criteria.

3. *Parameter estimation*. Fit the final model.

We use the estimation methods for the general SNMM to accomplish step 3. However, the R function developed for fitting SNMM (Wang and Ke 2002) can not be applied directly due to the complicated structure of the parametric part in (14). Therefore, we developed a new user-friendly R package, `PULSE`, for hormone pulse detection and estimation. `PULSE` consists of three main functions, `pulini`, `puldet` and `pulest`, for steps 1, 2 and 3 respectively. The manual of `PULSE` contains more details and examples. It can be downloaded from *http://www.pstat.ucsb.edu/faculty/yuedong/software.html*.

Due to the complexity of the NMPSM, there may be multiple local optimal solutions. Good initial values are critical to the performance of our algorithm. We have developed several methods and R functions for finding good initial values (Yang 2002, Yang et al. 2004).

When desirable, a fixed effect model can be assumed for all parameters $\boldsymbol{\phi}$. Then the second stage model (13) contains the fixed effects part only. Estimation and software can be developed similarly

(Yang 2002). R functions in the `PULSE` package allow parameters to be speficied as fixed, random or mixed. Even when $\boldsymbol{\phi}$ is considered as deterministic, it may be advantageous to estimate them using the penalized likelihood (15). For example, model (12) corresponds to shrinking α_k towards the common mean. Pooling data from different pulses, the resulting shrinkage estimates have smaller variances. This is especially important for the estimation of decay rates. Usually there are only a few observations on each pulse which make the maximum likelihood estimates of the decay rates unreliable. All existing methods are forced to assume a common decay rate for all pulses. Our simulations in Section 4 indicate that the shrinkage estimates are more efficient.

4 Simulation

4.1 Performance of pulse detection

In this subsection we conduct simulations to evaluate the performance of our methods for pulse detection. We generate data from model (8) with $n = 144$, $t_i = i/n$, and $f(t) = .5\cos(2\pi t) + 2$. We consider two choices for the number of pulses K: $K = 5$ or $K = 10$. For a fixed K, we generate pulse locations according to a nonhomogeneous Poisson process with intensity function $\lambda(t) = 35(.26 - (t - .5)^2)$. We use the double exponential function (10) as the pulse shape function with a fixed infusion rate, $\gamma_1 = 100$, and random amplitudes α_k and random decay rates γ_{2k}. Specifically, we generate pulse amplitudes such that $\log\alpha_k \overset{iid}{\sim} N(1, \sigma_1^2)$ and pulse decay rates such that $\log\gamma_{2k} \overset{iid}{\sim} N(3.66, \sigma_2^2)$. We generate random errors according to $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$. We consider four settings for variance parameters σ, σ_1 and σ_2: (I) $(\sigma, \sigma_1, \sigma_2) = (0.3, 0.3, 0.18)$; (II)$(\sigma, \sigma_1, \sigma_2) = (0.5, 0.3, 0.18)$; (III) $(\sigma, \sigma_1, \sigma_2) = (0.3, 0.5, 0.27)$; and (IV) $(\sigma, \sigma_1, \sigma_2) = (0.5, 0.5, 0.27)$. We repeat 100 times for each simulation setting.

We use the change point method to identify initial pulse locations and then apply our elimination procedure with $K_{min} = K/2$. We assume the second stage models $\log\alpha_k \overset{iid}{\sim} N(\beta_1, \sigma_\alpha^2)$ and $\log\gamma_{2k} \overset{iid}{\sim} N(\beta_2, \sigma_{\gamma_2}^2)$. Note that instead of using model (12) which were assumed in Keenan et al. (1998), we use log transformations to relax positive constraints on α_k and γ_{2k}. For $K = 10$ and Setting III, Figure 3 shows profiles of the AIC, GCV, BIC and RIC criteria with IDF=1.2 and histograms of the estimated K based on these four criteria. Plots for other simulation settings are similar. Figure 4 plots the false positive rates and false negative rates for each setting. All four criteria provide good estimates of pulse numbers and pulse locations while BIC and RIC perform slightly better except for the Setting IV where variances are large. The performance depend on the choice of the IDF: a larger IDF may improve the performance of the AIC and GCV. Overall, we recommand BIC and RIC with IDF=1.2. We note that the performances of our methods in terms of false positive rates and false negative rates are comparable to those in Mauger et al. (1995) even though our simulation settings are more difficult with a slower sampling rate, multiple sources of variations and a non-constant baseline.

We treat the baseline function f as a nuisance parameter. Nevertheless, MSEs of $\hat{f}$ listed in Table 1 indicate that our methods estimate the baseline function very well.

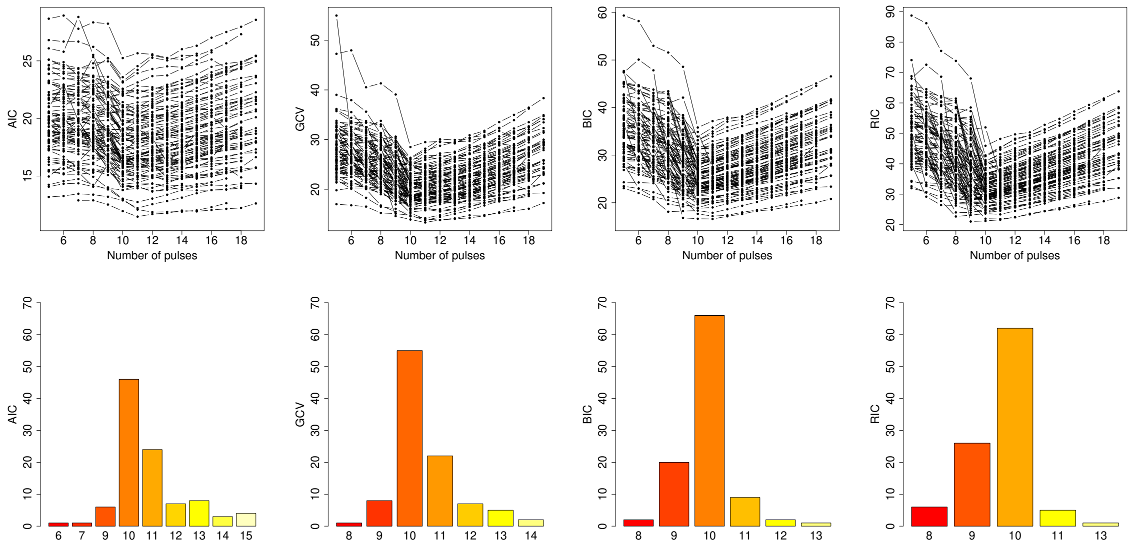
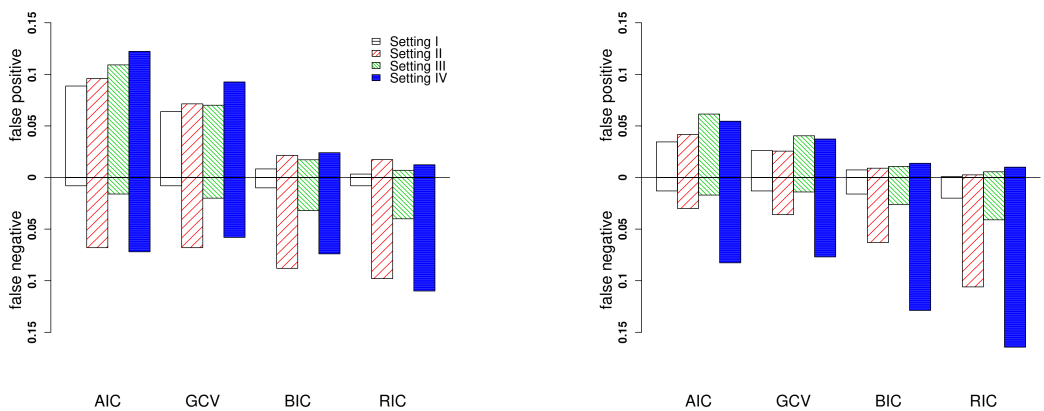


Figure 3: Four columns correspond to the AIC, GCV, BIC, and RIC criteria. The upper panel plots scores of these four criteria versus the number of pulses. The lower panel plots histograms of the estimated K.



(a) $K = 5$

(b) $K = 10$

Figure 4: False positive and false negative rates.

4.2 Efficiency of shrinkage estimates

For linear regression models, it is well-known that the shrinkage (also known as ridge) estimators reduces variance while increases bias. With the right amount of shrinkage, it is always possible to reduce the MSE (Efron and Morris 1975, Draper and Van Nostrand 1979, Gruber 1998). In this section we evaluate performance of the shrinkage methods for our nonlinear models. The simulation settings are the same as in Section 4.1. Instead of estimating the pulse locations, we now use true

	K=5				K=10			
	I	II	III	IV	I	II	III	IV
AIC	1.776	5.162	1.697	4.443	2.191	5.079	2.176	4.286
GCV	1.620	4.816	1.922	4.042	1.971	4.648	1.577	4.453
BIC	0.983	3.606	1.216	2.708	1.940	4.646	1.507	7.406
RIC	0.959	3.700	1.252	2.775	1.905	4.733	1.522	9.255

Table 1: MSEs of the baseline estimates.

pulse locations and evaluate our estimation methods.

In this subsection we consider all parameters including α_k and γ_{2k} as deterministic and our estimates based on the NMPSM as the shrinkage estimates. Specifically, the shrinkage estimates are minimizers of the following penalized least squares

$$||\boldsymbol{y}-\boldsymbol{f}-\boldsymbol{\eta}||^2+\lambda_1\sum_{k=1}^K(\log\alpha_k-\bar{\beta}_1)^2+\lambda_2\sum_{k=1}^K(\log\gamma_{2k}-\bar{\beta}_2)^2+n\lambda\int_0^1(f^{(m)}(u))^2du, \tag{20}$$

where $\bar{\beta}_1=\sum_{k=1}^K\log\alpha_k/K$, $\bar{\beta}_2=\sum_{k=1}^K log\gamma_{2k}/K$, and λ_1 and λ_2 are two shrinkage parameters. We shrink $\log\alpha_k$ and $\log\gamma_{2k}$ towards their grand means. Now consider a nonlinear mixed effect model with (8) as the first stage model and random effects $\log\alpha_k\overset{iid}{\sim}N(\beta_1,\sigma^2/\lambda_1)$ and $\log\gamma_{2k}\overset{iid}{\sim}N(\beta_2,\sigma^2/\lambda_2)$. Then it is not difficult to check that the penalized least squares (20) is equivalent to the penalized likelihood (15) with $\bar{\beta}_1$ and $\bar{\beta}_2$ replaced by β_1 and β_2. Therefore, we approximate the shrinkage estimates by the estimates of the corresponding nonlinear mixed effects model.

For comparison, we also calculate the least squares (LS) estimates of α_k and γ_{2k} using the `gnls` function in the `nlme` package. We repeat the simulation 100 times. Define MSE of $\hat{\alpha}_k$ as MSE$=\sum_{s=1}^S\sum_{k=1}^K(\hat{\alpha}_k^{(s)}-\alpha_k^{(s)})^2/(SK)$ where $\hat{\alpha}_k^{(s)}$ are the LS or shrinkage estimates of the true parameters $\alpha_k^{(s)}$ in the sth simulation and S is the number of simulations. The MSEs of $\hat{\gamma}_{2k}$ are defined similarly. Table 2 lists the MSEs of the LS and shrinkage estimates of α_k and γ_{2k}. The efficiencies, ratios between the MSE of the LS estimates and the MSE of the shrinkage estimates (Efron and Morris 1975), are also listed in Table 2. Shrinkage estimators are obviously more efficient, especially for the decay rates. From our experiments, the comparative superiority becomes less obvious when the variation of a parameter becomes larger or error variance becomes smaller.

5 Application

We now show the analysis of the data introduced in Section 1. We used the double exponential function (10) to model the pulse shape function and a cubic spline to model the baseline function. Amplitudes and pulse decay rates are modeled using random effects. Specifically, we assume that $\log\alpha_k\overset{iid}{\sim}N(\beta_1,\sigma_\alpha^2)$, $\log\gamma_{2k}\overset{iid}{\sim}N(\beta_2,\sigma_{\gamma_2}^2)$, and they are mutually independent. We used the simple random effects model since the number of pulses for each subject is relatively small which hinders more complex models. We will investigate more complicated integrated models for all subjects in the future.

	parameters	setting	shrinkage	standard	efficiency
K=5	amplitudes	I	0.06	0.09	1.55
		II	0.14	0.22	1.62
		III	0.10	0.14	1.38
		IV	0.18	0.25	1.37
	decay rates	I	0.12	0.42	3.52
		II	0.32	1.17	3.66
		III	0.20	0.67	3.41
		IV	0.30	1.28	4.22
K=10	amplitudes	I	0.15	0.45	3.00
		II	0.32	0.88	2.72
		III	0.23	0.59	2.55
		IV	2.73	4.24	1.55
	decay rates	I	0.26	2.23	8.61
		II	0.50	3.82	7.65
		III	0.41	2.23	5.45
		IV	1.75	10.51	6.02

Table 2: MSEs and efficiencies of the estimates for the amplitudes and decay rates.

The change point method identified $K_{max} = 13$ and potential pulse locations at 9:50am, 11:00am, 11:50am, 1:00pm, 4:00pm, 5:40pm, 7:10pm, 9:20pm, 10:30pm, 2:30am, 4:50am, 7:00am and 8:30am. We then applied our elimination procedure with $K_{min} = 5$. Table 3 shows the resulting sequence of pulse locations which are eliminated one-by-one. The estimates of the number of pulses based on the BIC, RIC, AIC and GCV criteria are 8, 8, 10 and 10 respectively. The identified pulse locations using BIC, overall fit and estimate of the baseline are shown in Figure 1. For comparison, we also show the identified pulse locations using AIC, overall fit and estimate of the baseline in Figure 5. It is interesting to note how pulses and the baseline function compensate for each other. Table 4 lists estimates of the amplitudes and decay rates based on the final model selected by the BIC criterion.

6 Discussion

The NMPSM provides more efficient and stable estimates of parameters by allowing different shape parameters for pulses within each subject and combining all nuisance parameters into a baseline function. The general form of the second stage mixed effects model will also allow researchers to investigate patterns of biological variation including circadian rhythms and feedback/feedforward control mechanisms (Keenan et al. 2003, Keenan and Veldhuis 2003), and further construct integrated models for all subjects. The challenge lies in detecting the number and locations of pulses masked by indirect observations and multiple sources of variations. It requires a sophisticated model selection procedure such as the one proposed in this article.

# pulses	BIC	RIC	AIC	GCV	DROP	DF
13	82.673	106.897	52.847	62.965	10:30pm	40.334
12	80.275	102.983	52.316	61.358	9:50am	37.809
11	78.252	99.593	51.975	60.219	9:20pm	35.534
10	76.133	96.032	51.631	59.087	5:40pm	33.134
9	73.760	90.818	52.757	59.767	11:00am	28.402
8	72.349	87.297	53.945	60.596	7:10pm	24.889
7	73.652	89.228	54.475	61.680	11:50am	25.934
6	75.068	89.202	57.666	65.513	2:30am	23.534
5	79.926	92.618	64.298	73.722	1:00pm	21.134

Table 3: Summary of the elimination procedure and criteria scores. The column DROP lists initial locations eliminated at each iteration. DF is defined in equation (18) with $IDF = 1.2$.

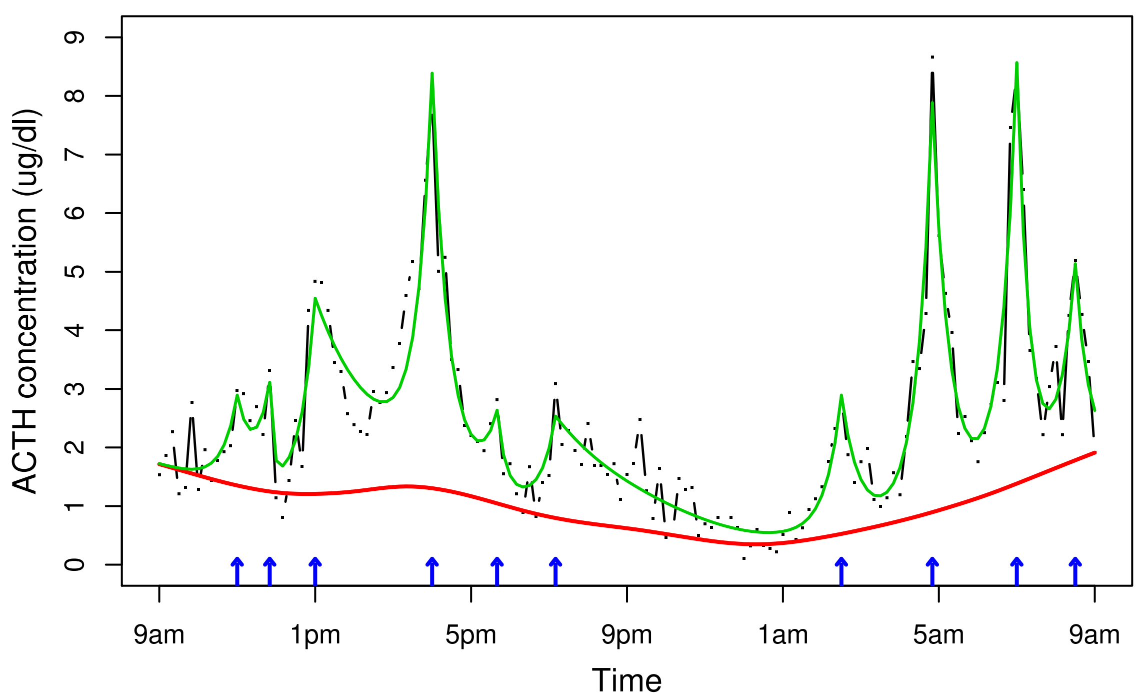


Figure 5: Pulse locations identified by the AIC criterion are marked below as vertical arrows. The overall fit and estimated baseline function are plotted as solid lines.

location	α_k	SE	γ_{2k}	SE
11:50am	0.641	0.056	5.770	0.496
1:00pm	1.074	0.012	2.272	0.076
4:00pm	1.827	0.005	4.003	0.023
7:10pm	0.458	0.041	2.632	0.172
2:30am	0.801	0.040	4.038	0.140
4:50am	1.934	0.004	3.961	0.021
7:00am	1.987	0.005	4.268	0.026
8:30am	1.351	0.028	4.117	0.113

Table 4: Estimates of α_k and γ_{2k} on log scale and their standard errors.

We limited our discussions to the problem of hormone pulse detection. However, as a general model, the NMPSM has other potential applications when observations are in a form of signals plus slow changing baseline (Hunt 1998, McBride 2002, Yang 2002). With slight modifications, our methods can be applied to these situations.

Acknowledgement

This research was supported by NIH Grant R01 GM58533. The authors would like to thank Leslie J. Crofford, MD, and Professor Morton B. Brown from the University of Michigan for allowing us to use the hormone data.

References

Akaike, H. (1973). Information theory and the maximum likelihood principle, *International Symposium on Information Theory*, eds. V. Petrov and F. Csáki, Budapest: Akademiai Kiádo, pp. 267–281.

Bolstad, W. (1988). The multiprocess dynamic linear model with biased perturbations: a real time model for growth hormone level, *Biometrka* **75**: 685–692.

Clifton, D. K. and Steiner, R. A. (1983). Cycle detection: a technique for estimating the frequency and amplitude of epsiodic fluctuations in blood hormone and substrate concentrations, *Endocrinology* **112**: 1057–1064.

Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions, *Numer. Math.* **31**: 377–403.

Crofford, L. J., Young, E. A., Engleberg, N. C., Korszun, A., Brucksch, C. B., McClure, L. A., Brown, M. B. and Demitrack, M. (2004). Basal circadian and pulsatile ACTH and cortisol secretion in patients with fibromyalgia and/or chronic fatigue syndrome, *Brian, Behavior and Immunity* **18**: 314–325.

Diggle, P. J. and Zeger, S. L. (1989). A non-gassian model for time series with pulses, *Journal of the American Statistical Association* **84**: 354–359.

Draper, N. R. and Van Nostrand, R. C. (1979). Ridge regression and James-Stein estimation: Review and comments, *Technometrics* **21**: 451–466.

Efron, B. and Morris, C. (1975). Data analysis using Stein's estimator and its generalizations, *Journal of the American Statistical Association* **70**: 311–319.

Foster, D. P. and George, E. I. (1994). The risk inflation criterion for multiple regression, *Annals of Statistics* **22**: 1947–1975.

Friedman, J. H. (1991). Multivariate adaptive regression splines (with discussion), *Annals of Statistics* **19**: 1–67.

Friedman, J. H. and Silverman, B. W. (1989). Flexible parsimonious smoothing and additive modeling (with discussion), *Technometrics* **31**: 3–39.

Green, P. J. and Silverman, B. W. (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*, London: Chapman and Hall.

Gruber, M. H. J. (1998). *Improving Efficiency by Shrinkage*, Marcel Dekker, New York.

Guo, W., Wang, Y. and Brown, M. B. (1999). A multiprocess state-space model for time series with pulse and changing baseline, *Journal of the American Statistical Association* **94**: 746–756.

Hartman, M. L. Faria, A. C., Vance, M. L., Johnson, M. L., Thorner, M. O. and Veldhuis, J. D. (1991). Temporal structure of in vivo growth hormone secretory events in man, *Am. J. Physiol. (Endocrinol. Metab. 23)* **260**: E101–E110.

Hinkley, D. (1971). Inference in two-phase regression, *Journal of the American Statistical Association* **66**: 736–743.

Hunt, M. A. (1998). *Quantitative pattern recognition and non-linear model based analysis*, PhD thesis, The University of Tennessee, Knoxville.

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

Ke, C. and Wang, Y. (2001). Semi-parametric nonlinear mixed effects models and their applications (with discussion), *Journal of the American Statistical Association* **96**: 1272–1298.

Keenan, D. M. and Veldhuis, J. D. (1997). Stochastic model of admixed basal and pulsatile hormone secretion as modulated by a deterministic oscillator, *Am. J. Physiol. (Regulatory Integrative Comp. Physiol. 42)* **273**: R1173–R1181.

Keenan, D. M. and Veldhuis, J. D. (2000). Explicating hypergonadotropism in postmenopausal women: A statistical model, *American Journal of Physiology* **278**: R1247–R1257.

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., Roelfsema, F., Biermasz, N. and Veldhuis, J. D. (2003). Physiological control of putuitary hormone secretory-burst mass, frequency, and waveform: A statistical formulation and analysis, *American Journal of Physiology* **285**: R664–R673.

Keenan, D. M., Sun, W. and Veldhuis, J. D. (2000). A stochastic biomathematical model of the male reproductive hormone system, *SIAM Journal on Applied Mathematics* **61**: 934–965.

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.

Keener, J. and Sneyd, J. (1998). *Mathematical Physiology*, Springer, New York.

Komaki, F. (1993). State-space modeling of time series sampled from continuous processes with pulses, *Biometrika* **80**: 417–429.

Kushler, R. H. and Brown, M. B. (1991). A model for identification of hormone pulses, *Statistics in Medicine* **10**: 329–340.

Luo, Z. and Wahba, G. (1997). Hybrid adaptive splines, *Journal of the American Statistical Association* **92**: 107–116.

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

McBride, S. J. (2002). A marked point process model for the source proximity effect in the indor environment, *Journal of the American Statistical Association* **97**: 683–691.

Merriam, G. R. and Wachter, K. W. (1982). Algorithms for the study of episodic hormone secretion, *American Journal of Physiology* **243**: E310–E318.

Munson, P. J. and Rodbard, D. (1989). Pulse detection in hormone data: simplified efficient algorithm, *Proceedings of the Statistical Computing Section, American Statistical Association* pp. 295–300.

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. 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.

Schwarz, G. (1978). Estimating the dimension of a model, *Annals of Statistics* **12**: 1215–1231.

Van Cauter, E. L., L'Hermite, M., Copinschi, G., Refetoff, S., Desir, D. and Robyn, C. (1981). Quantitative analysis of spontaneous variations in plasa prolactin in normal man, *American Journal of Physiology* **241**: E355–E363.

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**: E486–E493.

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 simutaneous multiple-parameter deconvolution of plasma hormone concentration, *Proceedings of the National Academy of Science* **84**: 7686–7690.

Veldhuis, J. D., Iranmanesh, A., Lizarralde, G. and Johnson, M. L. (1989). Amplitude modulation of a burstlike mode of cortisol secretion subserves the circadian glucocortocoid rhythm in man, *Am. J. Physiol. (Endocrinol. Metab. 20)* **257**: E6–E14.

Wahba, G. (1990). *Spline Models for Observational Data*, SIAM, Philadelphia. CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 59.

Wang, Y. (1998). Mixed-effects smoothing spline ANOVA, *Journal of the Royal Statistical Society B* **60**: 159–174.

Wang, Y. and Ke, C. (2002). ASSIST: A suite of s-plus functions implementing spline smoothing techniques, Manual for the ASSIST package. Available at `http://www.pstat.ucsb.edu/faculty/yuedong/software`.

Winer, L. M., Shaw, M. A. and Baumann, G. (1990). Basal plasma growth hormone levels in man: new evidence for rhythmicity of growth hormone secretion, *Journal of Clinical Endocrinology* **70**: 1678–1686.

Yang, Y. (2002). Detecting change points and hormone pulses using partial spline models, Ph.D. Thesis, University of California-Santa Barbara, Dept. of Statistics and Applied Probability.

Yang, Y., Liu, A. and Wang, Y. (2004). PULSE: A suite of R functions for detecting pulsatile hormone secretions, Manual for the PULSE package. Available at `http://www.pstat.ucsb.edu/faculty/yuedong/software`.