A signal extraction approach to modeling hormone time series with pulses and a changing baseline

Wensheng Guo, Yuedong Wang and Morton B. Brown*

July 20, 1998

Abstract

Hormones serve as regulating signals for many biological processes. In recent years, it was determined that many hormones are secreted in a pulsatile manner and that the pulsatile secretion pattern, rather than the absolute concentration level, is important in regulating biological processes. Consequently, it is necessary to characterize the latent secretion patterns from measurements of concentration levels. The characterization is complicated by the presence of a biological circadian rhythm. When hormone concentrations are plotted over time, the resultant time series usually exhibits occasional short rises superimposed on a slowly changing baseline. This is a result of a mixture of pulsatile secretions and a circadian rhythm. In this paper, we present a signal extraction approach to simultaneously model a slowly changing component and a pulsatile component of a time series. A smoothing spline is used to model the baseline and a multiprocess dynamic linear model is used to model the pulsatile component. An additive structure is assumed and both components are estimated simultaneously using a multiprocess Kalman filter. The unknown parameters are estimated by approximate maximum likelihood. The locations and amplitudes of the pulses are also estimated as posterior means via the multiprocess Kalman filter. Bayesian confidence intervals can be constructed for the baseline. This approach is found to be robust in simulated data and effective in modeling hormone time series.

Keywords: Bayesian inference, Hormone model, Multiprocess dynamic linear model, Pulsatile time series, Smoothing spline, State-space model.

1 Introduction

This article considers the problem of modeling a time series consisting of pulses superimposed on a slowly changing component (baseline), which occurs in applications such as time

*Wensheng Guo is Assistant Professor, Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, PA 19104. Yuedong Wang is Assistant Professor, Department of Statistics and Applied Probability, University of California, Santa Barbara, CA 93106-3110. Morton B. Brown is Professor, Department of Biostatistics, University of Michigan, Ann Arbor MI 48109-2029. This research was supported by NIH grants P60 DK20572, P30 HD18258 and R01 EY09946. The authors would like to thank Dr. Leslie Crofford for allowing us to use the cortisol data. We also thank the Editor, the Associate Editor and the referees for their comments that substantially improved this article.

series of hormone concentration levels in blood, voltages in neurons and underground water levels. In hormone studies, the pulses may result from bursts of hormone secretion from a gland into the circulation system, while the baseline may represent a circadian rhythm in hormone concentration levels. For underground water levels, the pulses may result from precipitation and the baseline may represent a smooth seasonal trend. In these applications, the baseline may not be well-described by a parametric function. When the pulse locations are known, the time series can be modeled by a partial spline model (Wahba, 1990, Chapter 6). When both components are smooth and their variances are stationary, Kitagawa and Gersch's (1984) smoothness-prior approach can be used. However, there are no statistical methods available when the pulsatile component is irregular, changes abruptly, and has an unknown number of pulses and unspecified pulse locations. The difficulty is that the slowly changing component and the abruptly changing component are closely related and require simultaneous estimation.

The motivation for this paper comes from our interest in modeling hormone time series. Hormones play important roles in regulating biological processes, such as reproduction, metabolism and growth. Many hormones, such as gonadotropin releasing hormone (GnRH), luteinizing hormone (LH), follicle stimulating hormone (FSH), growth hormone (GH), adrenocortocotropic hormone (ACTH) and cortisol, are secreted in a pulsatile manner. In recent years, there is increasing evidence (Clarke et al., 1984; Wilson et al., 1984; Shupnik et al. 1990; Weiss et al., 1990) that the secretion patterns, rather than the absolute concentration level, are important in triggering the response of a target tissue. Therefore it is necessary to characterize the latent secretion patterns instead of the mean level which is the main focus in most current statistical models. While our eventual goal is to study effects on pulsatile patterns that various factors may have by comparisons across different groups, we use a two-stage approach because of the complexity of the problem. Our initial task is to characterize the underlying secretion by modeling the observed time series of hormone levels of a single subject. A second stage analysis can then be performed using the estimated parameters from the first stage as the new endpoints. Theoretically, it is possible to combine the two stages using a mixed effects model. However, the resultant model is usually too complicated to produce stable estimates. Since the second stage analysis uses standard statistical methods such as regression or mixed effects models after appropriate transformations of the new endpoints, in the current paper we focus on the first stage analysis that characterizes the secretion patterns for each time series.

In a typical endocrinological experiment, blood samples are collected sequentially from each subject and assayed for hormone concentrations. The resultant time series usually exhibits pulses superimposed on a slowly changing baseline, possibly due to a circadian rhythm. The top panel of Figure 1 (solid line with dots) shows a time series of cortisol concentration in plasma from a control subject in an ongoing stress study conducted in the medical center of the University of Michigan. Blood samples were taken every 10 minutes and assayed for cortisol concentrations for a 24 hour interval, which consists of 144 time points. Cortisol is a steroid hormone secreted by the adrenal cortex and plays an important role in regulating metabolism. Stress is believed to be an important activator of cortisol secretion. A circadian oscillator such as the wake/sleep cycle also has a major impact on its secretion. The pattern of secretion may be abnormal in stressed patients, and the goal of the current study is to identify whether there is an excessive number of pulses, smaller pulse amplitudes or dysruption of the circadian rhythm in stressed patients compared to normal controls. The first step of the analysis is to characterize the latent secretion patterns in terms of summary statistics, which serve as new endpoints in a second stage analysis to compare groups.

In recent years there have been several statistical models proposed for pulsatile time series without a baseline or with a constant baseline. A review can be found in Mauger, Brown and Kushler (1995). Among the model-based methods, O'Sullivan and O'Sullivan (1988) and Kushler and Brown (1991) chose pulse locations in a stepwise fashion. Diggle and Zeger (1989) and Komaki (1993), on the other hand, defined the probability of a pulse as a function of the concentration of the previous point and used maximum likelihood to estimate the model parameters. Because of their schemes to determine the pulse locations, these models break down when there is a changing baseline, which often occurs in hormone studies.

In this paper, we use structural modeling to decompose the observed time series into an abruptly changing component, a smoothly changing component and measurement errors. The abruptly changing component can be further decomposed into discrete pulsatile inputs convoluted with an exponential decay. The decomposition of the 24 hour cortisol time series is shown in Figure 1, which will be further explained in section 4. The abruptly changing component is modeled by a multiprocess dynamic linear model and the smoothly changing component is modeled by a smoothing spline. The structural parameters all have specific interpretations and can be further modeled as functions of covariates. The two components are jointly modeled using a state-space representation and simultaneously estimated using a multiprocess Kalman filter.

The multiprocess dynamic linear model was first introduced by Harrison and Stevens (1971, 1976) for robust Bayesian forecasting. This model along with their approximate estimation procedure (the multiprocess Kalman filter) has been widely used in various applications (e.g. Smith and West (1983), Bolstad (1988), West and Harrison (1997, Ch. 12), Gordon and Smith (1990) and Shumway and Stoffer (1991)). The basic idea of the multiprocess dynamic linear model is that there are a known number of underlying processes; at each time we only observe one of these processes through a certain observational mechanism. In the state space framework, the components of the latent processes are represented by the state vector and their transitions are characterized by a set of Markov transition equations; the observational equations define the observed signal as a linear combination of the state vector subject to measurement errors. In the applications to hormone time series with pulses, the two processes represent having and not having a pulse at time t.

In this paper, we adopt the multiprocess dynamic linear model framework and use it as a random effects model for the pulsatile component, where the pulse locations and magnitudes are modeled as random effects and estimated by their posterior means. Our main interest is to characterize the latent signals through the observed time series. The overall estimates of the structural parameters: mean pulse amplitude, pulse frequency and decay rate are usually more interesting to researchers than the individual pulses. These unknown parameters are estimated by approximate maximum likelihood. We also modify the multiprocess Kalman filter algorithm by Bolstad (1988) and the classic fixed interval smoothing algorithm by Rauch, Tung and Striebel (1965) to obtain posterior estimates of both the pulsatile input and the state vector conditioned on the entire series.

A smoothing spline is a natural choice for the changing baseline because of its flexibility. Traditionally, it has been derived from penalized least squares where the smoothing parameter controls the tradeoff between the variance and bias. Wahba (1978) showed that a smoothing spline can also be obtained through a signal extraction approach by introducing a stochastic smoothing spline model. Weinert, Byrd and Sidu (1980) and Wecker and Ansley (1983) expressed this stochastic smoothing spline model in a state-space form. We combine the stochastic smoothing spline model with the multiprocess dynamic linear model into a unified state-space model. This leads to an efficient $O(n)$ algorithm to jointly estimate the abruptly changing component and the slowly changing component. The smoothing parameter along with other unknown parameters can be estimated by maximizing the approximate joint likelihood, which can also be calculated efficiently using the multiprocess Kalman filter.

In recent years, simulation techniques (e.g. Carlin, Polson and Stoffer (1992), Carter and Kohn (1994, 1996), Shephard (1994) and De Jong and Shephard (1995)) have been proposed to obtain exact Bayesian inference for this class of models. However, these techniques are extremely computationally intensive, especially when there are many unknown parameters to be estimated. The choice of prior needs to be considered on a case-by-case basis. It is also very difficult to check the convergence of these algorithms. Therefore, the proposed approximate maximum likelihood estimation using the Harrison-Stevens approximation is an efficient method for practical applications, which is supported by our simulations.

In section 2 a general model is presented that decomposes a time series into a smooth trend and an abruptly changing signal. Its estimation is introduced in section 3. In section 4 a specific model for hormone time series is discussed and applied to real data. A simulation is used to assess the performance of the method in section 5. In section 6 we give a brief summary and discuss some generalizations.

2 The model

Without loss of generality, we assume $t \in [0, 1]$. Let $y(t)$ be a time series sampled at $t = t_1, t_2, ..., t_n$, where the times are not necessarily equally spaced. The observed series $y(t)$ is partitioned into a slowly changing baseline $f(t)$, a pulsatile series $x(t)$ and a measurement error $\epsilon(t)$:

$$y(t) = x(t) + f(t) + \epsilon(t), \tag{1}$$

where $x(t)$ is modeled as a multiprocess dynamic linear model, $f(t)$ is modeled as a polynomial smoothing spline, and $\epsilon(t)$ is modeled as Gaussian white noise, $\epsilon(t) \sim N(0, \sigma_\epsilon^2)$. The models for $x(t)$ and $f(t)$ are discussed in the following subsections.

2.1 The multiprocess model for $x(t)$

The multiprocess linear model for an unequally spaced setting, which includes data that are missing at random in an equally spaced setting as a special case, can be written as:

$$x(t_j) = A(t_j)x(t_{j-1}) + v(t_j), \quad j = 1, ..., n, \tag{2}$$

where $A(t_j)$ is determined by the system transfer function and $v(t_j)$ is the input. The system transfer function is usually determined by setting up an approximate physical model and then solving the derived differential equations. In the applications to hormone data, the circulation system is usually modeled by compartmental models. It is known that the ability to estimate the parameters of these models is limited when there are many compartments in the model, and a much faster sampling rate is required to estimate the extra parameters caused by the additional compartment. The one-compartmental model or two-compartmental model is usually used to derive the transfer function. The one compartmental model results in exponential decay (Kushler and Brown, 1991) while a two compartmental model leads to a mixture of two exponential decays (Hoffmann, 1989). In equally spaced cases, exponential decay is reduced to a first order autoregressive process, while the mixture of two exponential decays leads to a second order autoregressive process (see Mauger (1995) for details).

The input $v(t_j)$ is modeled by a multiprocess with both mean and variance indexed by the random variable $I(t_j)$:

$$v(t_j) \sim N(\mu_i(t_j), \sigma_i^2(t_j)) \quad \text{if } I(t_j) = i, \quad i = 0, ..., k-1,$$

where k is the number of potential states. The $I(t_j)$'s are assumed to be independent and therefore can be modeled as outcomes of a multinomial trial with prior probability $\pi_i(t_j) = Pr(I(t_j) = i)$ for $i = 0, ...k-1$. The parameters $\{\pi_i(t_j), \mu_i(t_j), \sigma_i^2(t_j)\}$ can be further modeled according to prior knowledge, such as functions of covariates including time.

2.2 The smoothing spline model for $f(t)$

Wahba (1978) showed that a smoothing spline of degree $2m-1$ can be obtained through a signal extraction approach by introducing the following stochastic model for $f(t)$:

$$f(t) = \sum_{k=0}^{m-1} \alpha_k \frac{t^k}{k!} + \lambda^{1/2}\sigma_e \int_0^t \frac{(t-h)^{m-1}}{(m-1)!} dW(h), \quad t \in [0,1],$$

where $W(h)$ is a Wiener process with unit dispersion parameter and λ is the smoothing parameter. The random vector $(\alpha_0, ..., \alpha_{m-1})'$ is assigned a diffuse prior distribution $N(\mathbf{0}, \xi I)$ with $\xi \to \infty$.

Wecker and Ansley (1983) modified this stochastic model and showed that the polynomial smoothing spline can be written in a state-space form. As the result of the state space representation, a smoothing spline can be calculated by efficient $O(n)$ procedures: Kalman filter and fixed interval smoother. Following their formulation, the system equation of $f(t)$ can be written in terms of $f(t_j)$ and its first $m-1$ derivatives, $\mathbf{F}(t_j) = (f(t_j), f^{(1)}(t_j), ..., f^{(m-1)}(t_j))'$:

$$\mathbf{F}(t_j) = T(t_j, t_{j-1})\mathbf{F}(t_{j-1}) + \mathbf{U}(t_j, t_{j-1}), \quad j = 1, ..., n, \tag{3}$$

where $T(t_j, t_{j-1})$ is a $m \times m$ matrix with element (i,k) being $(t_j - t_{j-1})^{k-i}/(k-i)!$ when $i \geq k$ and 0 otherwise. The perturbation $\mathbf{U}_{t_j,t_{j-1}}$ is a m dimensional vector with distribution $N(\mathbf{0}, \lambda\sigma_e^2 R(t_j, t_{j-1}))$, where $R(t_j, t_{j-1})$ is a $m \times m$ matrix with element (i,k) being $(t_j - t_{j-1})^{2m-i-k+1}/[(m-i)!(m-k)!(2m-i-k-1)!]$.

The most commonly used form of the equation (3) is the cubic spline ($m = 2$). If we set $U(t_j, t_{j-1})$ to zero, equation (3) represents a straight line. The perturbation component $U(t_j, t_{j-1})$ is modeled as a random effect and represents the smoothed deviation of the overall fitted curve from the straight line. The smoothness is controlled by the smoothing parameter λ.

2.3 The combined state-space model

Equations (1) – (3) can be combined into one unified state-space model:

$$y(t_j) = H\mathbf{Z}(t_j) + \epsilon(t_j) \tag{4}$$

$$\mathbf{Z}(t_j) = \Phi(t_j)\mathbf{Z}(t_{j-1}) + \mathbf{W}(t_j), \tag{5}$$

where $H = (1, 1, 0, ..., 0)$, $\mathbf{Z}(t_j) = (x(t_j), \mathbf{F}'(t_j))'$, $\Phi(t_j) = \text{diag}(A(t_j), T(t_j, t_{j-1}))$ and $\mathbf{W}(t_j) = (v(t_j), \mathbf{U}'(t_j, t_{j-1}))'$.

Equation (4) is called the observational equation which is a matrix representation of equation (1). It defines the way that we observe the latent components through an additive structure. Since the transition matrix $\Phi(t_j)$ is block diagonal, equation (5), the system level equation, is a matrix representation of equations (2) and (3) simultaneously. The system level equation is a Markov transition equation which defines the structures of the latent components $x(t)$ and $f(t)$. The advantage of this unified state-space representation is that it allows us to adopt existing efficient estimation procedures, such as the multiprocess Kalman filter and fixed interval smoother, to simultaneously estimate the latent components.

3 Estimation

Similar to other random effects models, estimation includes two steps: first we obtain maximum likelihood estimates for the unknown parameters; then based on these estimates, we calculate the posterior estimates of the latent signals $x(t_j)$, $f(t_j)$ and $v(t_j)$, which are the decomposition of the observed time series. The hidden Markov structure of state space models enables the likelihood and the posterior estimates of the latent signals conditioned on the history to be efficiently calculated in $O(n)$ operation in a recursive fashion by the Kalman filter (KF). A backward $O(n)$ recursive procedure, fixed interval smoothing algorithm, can then be applied to the results of the Kalman filter to produce posterior estimates of the state vector conditional on the whole series.

In Gaussian state space models, each step of the recursive estimation can be carried out in a closed form by the Kalman filter. The expository article by Meinhold and Singpurwalla (1983) provides an intuitive introduction to the Kalman filter. The KF is a sequential forward estimation procedure. At each step, the KF first uses the system level equation to calculate $P(\mathbf{Z}(t_j)|\mathbf{Y}(t_{j-1}))$ from the previous result, where $\mathbf{Y}(t_j) = \{y(t_1), ..., y(t_j)\}$ denotes the entire history up to t_j. It then uses the observational equation to calculate $P(y(t_j), \mathbf{Z}(t_j)|\mathbf{Y}(t_{j-1}))$. The $\mathbf{Z}(t_j)$ can then be integrated out in a closed form to obtain $P(y(t_j)|\mathbf{Y}(t_{j-1}))$, which is the contribution to the likelihood. By using some simple results from multivariate statistics, we then calculate $P(\mathbf{Z}(t_j)|\mathbf{Y}(t_j))$ from $P(y(t_j), \mathbf{Z}(t_j)|\mathbf{Y}(t_{j-1}))$, which prepares the KF for the next step. All these distributions are characterized by their first two moments.

In multiprocess dynamic linear models, estimation follows a similar procedure, but approximations or numerical methods are needed because the marginal distribution of $\mathbf{Z}(t_j)$ at each step is a mixture of normals. Harrison and Stevens (1971, 1976) proposed to use a normal approximation to the mixture of normals, which results in a multiprocess Kalman filter (MKF). The MKF can be used to calculate an approximate likelihood and posterior estimates of the state. We extend the MKF by Harrison and Stevens (1976) to obtain the posterior estimate of the input which is of particular interest in our applications.

Let $\boldsymbol{\theta}$ be the vector that includes all the unknown parameters. The likelihood can be obtained sequentially using the conditional densities $p(y(t_j)|\mathbf{Y}(t_{j-1}))$:

$$L(\boldsymbol{\theta}|\mathbf{Y}(t_n)) = p(y(t_n)|\mathbf{Y}(t_{n-1}), \boldsymbol{\theta})p(y(t_{n-1})|\mathbf{Y}(t_{n-2}), \boldsymbol{\theta})...p(y(t_1)|\mathbf{Y}(0), \boldsymbol{\theta}), \tag{6}$$

where $p(y(t_j)|\mathbf{Y}(t_{j-1}), \boldsymbol{\theta})$ is available at each step of the multiprocess Kalman filter. An optimization program can be used to obtain the maximum likelihood estimates for the parameters. Some transformations of the parameters to remove their constraints may also be needed, such as logarithm transformations for positive constraints and logit transformations for prior probabilities.

In order to fully specify the likelihood, the initial values of the state vector at time 0 need to be specified. Two approaches are available: the first is to include the $\mathbf{Z}(0)$ as parameters and estimate them by maximum likelihood; the other is to put a diffuse prior on $\mathbf{Z}(0)$ which reflects our lack of knowledge of the initial conditions. We use the second approach because it reduces the dimension of the optimization problem. In the Gaussian state space model, two approaches were introduced by Ansley and Kohn (1985) and De Jong (1988, 1991) to incorporate diffuse priors on initial states. These approaches can not be directly used in our model since they require knowledge about the pulse locations. We adopt the numerical realization approach by Harvey and Phillips (1979) which assumes that the elements of $\mathbf{Z}(0)$ have zero means and large variances. The criticism of this approach is that very large initial variances may cause numerical problems. From our experience, once the initial variances are greater than one hundred times the scale of the $y(t)$'s, different choices of the initial variances have little impact on the estimates. Since it is always possible to rescale the observations when the scale of $y(t)$ is truly large, the numerical values used for the initial variances are usually far from the boundary that causes numerical problems. We find that this numerical approach is an easy way to approximate the diffuse priors.

Our estimation procedure consists of a recursive forward algorithm, the multiprocess Kalman filter (MKF) and a recursive backward algorithm, the fixed interval smoother (FIS). The MKF starts from the initial values $\mathbf{Z}(0)$ and sequentially calculates $P(\mathbf{Z}(t_j)|\mathbf{Y}(t_j), \boldsymbol{\theta})$ for $j = 1, ..., n$. The FIS runs backward based on the results of the MKF to sequentially obtain the smoothed estimates $P(\mathbf{Z}(t_j)|\mathbf{Y}(t_n), \boldsymbol{\theta})$. Since each observation's contribution to the likelihood is available through the MKF, the approximate maximum likelihood estimation procedure only involves the MKF. The FIS needs to be run only once based on the final set of estimates.

In hormone time series, the input represents the hormone secretion from the gland and is the latent signal that we want to characterize. We propose to augment the input $v(t)$ into the state vector $\mathbf{Z}(t)$ and estimate it together with the state vector using the MKF and FIS. Define $\tilde{\mathbf{Z}}(t_j) = (v(t_j), \mathbf{Z}'(t_j))'$, $\tilde{H} = (0, H')'$, $\tilde{\Phi}(t_j) = \mathrm{diag}(0, \Phi(t_j))$. Let $\tilde{\mathbf{Z}}(t_j|t_k) = \mathrm{E}[\tilde{\mathbf{Z}}(t_j)|\mathbf{Y}(t_k)]$ and $\tilde{\Sigma}(t_j|t_k) = \mathrm{cov}[\tilde{\mathbf{Z}}(t_j)|\mathbf{Y}(t_k)]$. The MKF and FIS are then applied on $\tilde{\mathbf{Z}}(t_j)$ to obtain filtered estimates and smoothed estimates of both $v(t_j)$ and $\mathbf{Z}(t_j)$. Technically, all the conditional distributions should depend on $\boldsymbol{\theta}$, so do the means and variances. To simplify the notation, we omit $\boldsymbol{\theta}$ when it is obvious.

The MKF resembles that of Harrison and Stevens (1971, 1976) and Bolstad (1988), except that we augment the input $v(t)$ into the state vector $\mathbf{Z}(t)$. This enables us to obtain the posterior estimates of the input $v(t)$ along with those of the state vector $\mathbf{Z}(t)$. The algorithm is presented in Appendix A. Because of the approximation in step (5) of the algorithm, the resultant likelihood is only an approximation.

The MKF produces filtered estimates of the state variables $\tilde{\mathbf{Z}}(t_j)$ conditioned on the history $\mathbf{Y}(t_j)$. For the purpose of estimation, we are more interested in the smoothed estimates conditioned on the entire series $\mathbf{Y}(t_n)$. The classic fixed interval smoothing algorithm for the state vector was first introduced by Rauch, Tung and Striebel (1965). It can be directly extended to obtain the smoothed estimates for $\tilde{\mathbf{Z}}(t)$ (both the state vector and the input) based on the filtered estimates calculated previously. The algorithm is given in Appendix B. In the regular Kalman filter the smoothed estimates are the best linear unbiased predictions

(BLUPs) (Robinson 1991). In multiprocess models these are only approximate because of the normal approximation to the mixture of normal distributions at each step.

4 Application to hormone time series data

We apply our model to hormone time series data collected from a stress study conducted in the medical center of the University of Michigan. The study was approved by the Medical Center's Institutional Review Board for Human Subject Research and each subject provided written informed consent. Blood samples were taken every 10 minutes from each individual and assayed for cortisol concentrations for a 24 hour interval started at 9am; therefore, each series consists of 144 times points. The data contained less than 1% missing values due to bad assays. These missing values are treated as unobserved and the resultant time series is modeled as an unequally spaced time series in our analysis. The goal of the study is to identify abnormalities in the secretion patterns in stressed women compared to those in normal women. The first step of the analysis is to characterize the latent secretion patterns in terms of summary statistics. A second stage analysis can be performed using these summary statistics as the new endpoints. Since the methods used in the second stage analysis are usually standard, in this section we focus on the first stage analysis that is required for each time series.

The top panel of Figure 1 shows the data from a control subject, which consists of a pulsatile series superimposed on a slowly changing baseline. The baseline is due to a biological circadian rhythm and the pulses are results of bursts of cortisol from the adrenal cortex. The circulation system is modeled by a compartmental system which approximates the whole system by a few well-mixed compartments. In two compartmental models, it is known that the second compartment is usually much smaller than the main one. This leads to a mixture of two exponential decays with two very different half-lifes, which usually differ by ten to twenty times (e.g. Hoffmann (1989)). If we are not sampling fast enough, we do not have sufficient information to estimate the second decay component. In our particular application to cortisol data, the longer half-life is around 30 minutes, while the shorter half-life may be around $2-3$ minutes. In order to estimate the shorter half life, a sampling rate of one sample per minute may be needed, which is not feasible due to the amount of blood that would be needed from each subject. We therefore model the circulation system as a one compartment model. This leads to an exponential-decay-toward-baseline model for the pulsatile series (see Kushler and Brown (1991) for details). Equation (2) is then written as:

$$x(t_j) = e^{-\gamma(t_j - t_{j-1})} x(t_{j-1}) + v(t_j), \tag{7}$$

where γ is the decay rate and $v(t_j)$ is the net input of the hormone into the blood system from t_{j-1} to t_j. That is, $A(t_j) = e^{-\gamma(t_j - t_{j-1})}$. By modeling the circulation system as one compartment, and due to the removal effect from the body, the expected concentration at t_j is $e^{-\gamma(t_j - t_{j-1})} x(t_{j-1})$ if there is no secretion during the interval (t_j, t_{j-1}). The difference between $x(t_j)$ and $e^{-\gamma(t_j - t_{j-1})} x(t_{j-1})$ is the net input due to the secretion within $t_j - t_{j-1}$. We model this net input $v(t_j)$ by a multiprocess with two classes ($k = 2$) indexed by $I(t_j)$, where $I(t_j) = 1$ indicates pulsatile secretion between time t_{j-1} and t_j, and $I(t_j) = 0$ indicates no secretion. With $\mu_0(t_j)$ and $\sigma_0^2(t_j)$ set to zeros, $v(t_j)$ is zero with probability one when $I(t_j) = 0$. The structural parameters $\{\pi_1(t_j), \mu_1(t_j), \sigma_1^2(t_j)\}$ can be further modeled according to prior knowledge, such as functions of the history to allow a feedback mechanism. In the absence of prior knowledge, the net input $v(t_j)$ is modeled as:

$$v(t_j) \sim \begin{cases} N(\mu \times (t_j - t_{j-1}), \sigma_p^2 \times (t_j - t_{j-1})) & \text{if } I(t_j) = 1 \\ N(0, 0) & \text{if } I(t_j) = 0 \end{cases} .$$

with the prior probability $P(I(t_j) = 1) = \pi$. Theoretically, there can be multiple secretory events in the interval (t_j, t_{j-1}) and the prior probability $P(I(t_j) = 1)$ should also be a function of the length of this interval. However, if our sampling rate is much slower that the secretion rate, we do not have information to characterize an individual secretory event. The only information that we have is the net amount of secretion (total secretion minus decay) at the end of that interval, which is modeled by $v(t_j)$. Instead of defining a pulse as an underlying secretory event, we define a "pulse" as the net amount of secretion within the sampling interval. That is, each interval can only have one "pulse", and a longer interval may lead to a bigger "pulse" which may be the sum of multiple secretory events. We model the mean pulsatile input per unit time and its standard deviation as μ and σ_p. Assuming independence between two secretory events, the mean and variance of the net input in an interval is proportional to the length of the interval. The estimate of π has an interpretation as the average probability of a pulse occurring in each sampling interval. This is a measure of pulse frequency in equally spaced settings.

A cubic spline ($m = 2$) is usually adequate to model $f(t)$. When prior information on $f(t)$ is available, such as a circadian rhythm, the approach by Kohn, Ansley and Wong (1993) can be used. Equation (3) is simplified as follows for a cubic spline:

$$\begin{bmatrix} f(t_j) \\ f'(t_j) \end{bmatrix} = \begin{bmatrix} 1 & t_j - t_{j-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} f(t_{j-1}) \\ f'(t_{j-1}) \end{bmatrix} + \mathbf{U}(t_j, t_{j-1}) \tag{8}$$

where

$$\mathbf{U}(t_j, t_{j-1}) \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \lambda\sigma_e^2 \begin{bmatrix} (t_j - t_{j-1})^3/3 & (t_j - t_{j-1})^2/2 \\ (t_j - t_{j-1})^2/2 & (t_j - t_{j-1}) \end{bmatrix} \right).$$

The model specified by equations (8) and (9) was used to analyze the cortisol hormone data. For the purpose of illustration, we present the results of the analysis of the cortisol data from one of the control subjects. Figure 1 presents the original data along with the estimates of baseline, pulse locations and magnitudes. The estimate of the baseline matches our intuition and the variation in the baseline is significant because the 95% confidence interval does not include a horizontal line. The posterior probability of a pulse and the input are estimated at each time point. If a definitive classification is desired, an investigator can apply a threshold to the posterior probabilities according to the desired sensitivity and specificity. Another method of classification is to construct confidence intervals for the inputs and check whether they include zero.

Because we use approximate maximum likelihood estimates, the standard errors of these estimates can not be obtained directly from the naive Fisher information matrix which treats the approximate likelihood as the true likelihood. We use a parametric bootstrap approach which regenerates pulses and measurement errors based on the parameter estimates. We generated 1000 bootstrap samples and estimated the standard errors of the parameter estimates by bootstrap standard deviations. Since the MKF is a very efficient algorithm, this bootstrap procedure takes only about one hour on a Sun ULTRA-SPARC1. The parameter estimates for the time series shown in Figure 1 are $\{\hat{\mu} = 5.12(1.22)[0.62], \hat{\alpha} = 0.80(0.037)[0.030], \hat{\pi} = 0.15(0.056)[0.045], \hat{\sigma}_p = 1.67(1.33)[0.67], \hat{\sigma}_e = 0.85(0.18)[0.17]\}$, where the numbers in the parentheses are the estimates of standard errors by the parametric bootstrap and those in the brackets are estimates of standard errors obtained through the naive Fisher information matrix by treating the approximate likelihood as the true likelihood. Except for $\hat{\mu}$ and $\hat{\sigma_p}$, the estimates of the standard errors are similar for the two procedures. For $\hat{\mu}$ and $\hat{\sigma_p}$, the estimates of the standard errors by the naive procedure are about 1/2 those by the bootstrap; this may be due to the normal approximation to the mixture of normals.

5 Simulation

In order to assess the performance of our method, a set of simulations was carried out for the hormone model specified by equations (8) and (9). Since the exponential decay toward baseline model was used in many pulse detection models and has been proved to be a reasonable approximation of the circulation system, we generated our pulse sequence $x(t)$ using equation (8). The baseline was specified by a sine curve with a constant level shift to ensure that it was positive: $f(t) = 20 + Asin[2 \times 3.1415926 \times (t-1)/144]$, for $t = 1, ..., 144$, where A is the amplitude of the baseline. The settings are specified by parameters $\{\mu, \pi, \alpha = e^{-\gamma}, \sigma_e, \sigma_p, A\}$. Because estimation is invariant to rescaling, we fixed $\mu = 10$ and varied the other parameters. A factorial design with the following settings was used: π was 10% or 20%; α was 0.5, 0.7 or 0.84, which correspond to half-lifes equal to 1, 2 and 4 time points respectively; A was 2.5 or 7.5 which makes the fluctuation of the baseline $50\% \times \mu$ or $150\% \times \mu$; σ_p was chosen as 1, 2 or 3; and σ_e was 1 or 2. From our experience, these settings cover many real cases that arise from endocrinological studies.

For each setting, we generated 100 equally spaced time series, each with length of 144 time points which is the same as a time series sampled every 10 minutes for a 24 hour interval. Figure 2 presents a typical simulated series along with its fitted results. The settings of the parameters are $\{\mu = 10, \alpha = 0.7, \pi = 0.1, \sigma_p = 2, \sigma_e = 1, A = 15\}$. The estimates of the baseline, pulse locations and magnitudes are very close to their true values.

Boxplots of the maximum likelihood estimates are used to examine the performance of the estimation procedure at each setting. False-positive and false-negative rates for pulse identifications were calculated by applying a 0.5 threshold to the posterior probabilities. In the range of our simulations, different settings of the half-lifes and the fluctuations of the baseline did not have much influence on the parameter estimation. The choice of π has an effect of the estimation of μ and σ_p, but the patterns are very similar. We therefore only present results when $\alpha = 0.7$ and $\pi = 0.1$; we pooled the results across the two settings of A. The variance of the measurement error σ_e^2 was estimated well in all settings; however, it is usually of little interest. Figure 3 shows the estimates of π, μ, α, σ_p, and false positive and false negative rates, where each boxplot represents the results from 2×100 simulations. The variability of these estimates comes from both variation in generating the data and potential bias in estimation. The inter-quartile intervals for σ_p and μ are relatively wide because their effective sample sizes are the number of pulses, which is small. As π increases (not shown in Figure 3), the expected number of pulses increases and the inter-quartile intervals for μ and σ_p become narrower. The effect of π on the estimation of α is small, since it uses all the points on the decay sides of the pulses. In the simulation, the most important factors are σ_e and σ_p, which reflect the measurement error and the homogeneity of pulses. When $\sigma_e = 1$ (i.e. σ_e/μ is small), the inter-quartile intervals are very narrow and the false positive and false negative rates are essentially zero. As σ_e and σ_p increase, the inter-quartile intervals become wider, and the false positive rate and false negative rate increase. These results match our intuition that large measurement error and heterogeneity of pulses can cause misidentification of pulses.

To investigate whether the one compartmental model is applicable, we conducted another simulation which generated the data using a two compartmental model and re-estimated it with a one compartmental model. We set the longer half-life as 20 minute with a 10 minute sampling rate. We varied the ratios of the shorter half-life to the longer one by 1/5 and 1/10. In the first setting, the simulation showed very small bias in α, π, μ and σ_e. The bias vanished in the second setting. This result is intuitive, because the impact of faster decay after 5 or 10 half-lifes is negligible. The exponential-decay-toward-baseline model is therefore adequate for practical application.

6 Discussion

We have presented a signal extraction approach to model a time series that consists of a smoothly changing baseline and pulses, which can not be modeled by existing methods. Existing methods require that the baseline is constant or that the locations of the pulses are known. In our model, a smoothing spline is used to model the low frequency component and a multiprocess dynamic linear model is used to detect the abruptly changing signals. By using a state-space representation, the baseline and pulses are jointly modeled as stochastic components and simultaneously estimated by their posterior predictions conditioned on the whole series. An approximate likelihood method is used to estimate the parameters, such as pulse rate and mean pulse amplitude, which are often the primary interests of researchers. This approximate method produces very good estimates in the range of our simulations. The state-space representation also results in a very efficient $O(n)$ algorithm for the likelihood calculation.

We used an example of hormone time series with pulses to demonstrate this model, where the pulses are modeled as jumps followed by exponential decays. The pulses can be viewed broadly as any abrupt structural changes, such as jumps, spikes and even change-points. Equation (2) can have a vector form to allow a more complicated structure. Prior information, such as a feedback mechanism and covariates, can also be incorporated by modeling the structural parameters $\{\pi_i(t_j), \mu_i(t_j), \sigma_i^2(t_j)\}$.

The inputs $v(t_j)$'s in equation (2) are assumed to be independent of each other in the current model. Our method assumes that a surge with two points in the rising slope is a result of two consecutive independent inputs. This assumption is valid for time series with a slow sampling rate relative to the duration of the secretory event. However, in time series that are sampled rapidly, an input may continue over several time points. The dependency between the $v(t_j)$'s needs to be modeled explicitly and will be explored in future research.

We have focused on a times series model in the current paper. In many endocrinological studies, the eventual goal is to study the effects on secretion that various factors may have by comparison across series. Because of the complexity of the problem, we currently approach the problem by a two-stage approach. It may be desirable to combine the two stages using a mixed effects type of model. In the absence of pulses, several authors (e.g. Wang (1998); Brumback and Rice (1998)) have generalized smoothing splines to a mixed effects type of model which can model functional data from different subjects. When there are pulses superimposed on a smooth curve, the extension to multiple subjects can be complicated due to the complex model structure. We will explore this possibility in our future research.

Appendix A: The multiprocess Kalman filter (MKF)

For any given values of $\boldsymbol{\theta}$, a step of the algorithm MKF starts at t_{j-1} with $\tilde{\mathbf{Z}}(t_{j-1}|t_{j-1})$ and $\tilde{\Sigma}(t_{j-1}|t_{j-1})$.

(1.) Conditioned on $\{\mathbf{Y}(t_{j-1}), I(t_j) = i\}$, the distribution of $\tilde{\mathbf{Z}}(t_j)$ is characterized by its first two moments:

$$\tilde{\mathbf{Z}}(t_j|t_{j-1}, I(t_j) = i) = \tilde{\Phi}\tilde{\mathbf{Z}}(t_{j-1}|t_{j-1}) + \begin{bmatrix} \mu_i(t_j) \\ \mu_i(t_j) \\ 0 \\ 0 \end{bmatrix},$$

$$\tilde{\Sigma}(t_j|t_{j-1}, I(t_j) = i) = \tilde{\Phi}\tilde{\Sigma}(t_{j-1}|t_{j-1})\tilde{\Phi}' + \begin{bmatrix} \sigma_i^2(t_j) & \sigma_i^2(t_j) & \\ \sigma_i^2(t_j) & \sigma_i^2(t_j) & \\ & & \lambda\sigma_e^2 R(t_j, t_{j-1}) \end{bmatrix}.$$

(2.) The predicted distribution of $y(t_j)$ given $\mathbf{Y}(t_{j-1})$ conditioned on $I(t_j)$ is also characterized by its first two moments:

$$\begin{aligned} y(t_j|t_{j-1}, I(t_j) = i) &= \mathrm{E}(y(t_j)|\mathbf{Y}(t_{j-1}), I(t_j) = i) \\ &= \tilde{H}\tilde{\mathbf{Z}}(t_j|t_{j-1}, I(t_j) = i), \end{aligned}$$

$$\begin{aligned} V_y(t_j|t_{j-1}, I(t_j) = i) &= \mathrm{var}(y(t_j)|\mathbf{Y}(t_{j-1}), I(t_j) = i) \\ &= \tilde{H}\tilde{\Sigma}(t_j|t_{j-1}, I(t_j) = i)\tilde{H}' + \sigma_e^2. \end{aligned}$$

The marginal distribution of $(y(t_j)|\mathbf{Y}(t_{j-1}))$ is a mixture of k normal distributions:

$$(y(t_j)|\mathbf{Y}(t_{j-1})) \sim \sum_{i=0}^{k-1} \pi_i(t_j) N(y(t_j|t_{j-1}, I(t_j) = i), V_y(t_j|t_{j-1}, I(t_j) = i)).$$

The density $p(y(t_j)|\mathbf{Y}(t_{j-1}))$ calculated from the above distribution is the contribution to the likelihood in equation (6).

(3.) The posterior probability of the index variable $I(t_j) = i$ given $\mathbf{Y}(t_j)$ is then

$$q_i(t_j) = \pi_i(t_j) p(y(t_j)|\mathbf{Y}(t_{j-1}), I(t_j) = i)/p(y(t_j)|\mathbf{Y}(t_{j-1})), \tag{9}$$

which can be obtained from the conditional distributions given above.

(4.) Given $\mathbf{Y}(t_j)$, we can update the first two moments of $\tilde{\mathbf{Z}}(t_j)$ conditioned on $I(t_j)$ as:

$$\begin{aligned} \tilde{\mathbf{Z}}(t_j|t_j, I(t_j) = i) &= \tilde{\mathbf{Z}}(t_j|t_{j-1}, I(t_j) = i) \\ &\quad + \frac{\tilde{\Sigma}(t_j|t_{j-1}, I(t_j) = i)\tilde{H}'[y(t_j) - y(t_j|t_{j-1}, I(t_j) = i)]}{V_y(t_j|t_{j-1}, I(t_j) = i)}, \end{aligned}$$

$$\tilde{\Sigma}(t_j|t_j, I(t_j) = i) = \tilde{\Sigma}(t_j|t_{j-1}, I(t_j) = i) - \frac{\tilde{\Sigma}(t_j|t_{j-1}, I(t_j) = i)\tilde{H}'\tilde{H}\tilde{\Sigma}(t_j|t_{j-1}, I(t_j) = i)}{V_y(t_j|t_{j-1}, I(t_j) = i)}.$$

(5.) The marginal distribution of $\tilde{\mathbf{Z}}(t_j|t_j)$ is a mixture of k normal distributions. We adopt Harrison and Stevens' (1976) approach and approximate the marginal distribution by a normal distribution with the same first two moments, which minimizes the Kullback-Leibler distance (West and Harrison, 1997, chapter 12):

$$\tilde{\mathbf{Z}}(t_j|t_j) = \sum_{i=0}^{k-1} q_i(t_j)\tilde{\mathbf{Z}}(t_j|t_j, I(t_j) = i),$$

$$\begin{aligned}\tilde{\Sigma}(t_j|t_j) &= \sum_{i=0}^{k-1} q_i(t_j)[\tilde{\Sigma}(t_j|t_j, I(t_j) = i) \\ &+\{\tilde{\mathbf{Z}}(t_j|t_j, I(t_j) = i) - \tilde{\mathbf{Z}}(t_j|t_j)\}\{\tilde{\mathbf{Z}}(t_j|t_j, I(t_j) = i) - \tilde{\mathbf{Z}}(t_j|t_j)\}'].\end{aligned}$$

The algorithm is ready for the next step starting at t_j.

Appendix B: The fixed interval smoother (FIS)

The algorithm FIS starts with $\tilde{\mathbf{Z}}(t_n|t_n)$ and $\tilde{\Sigma}(t_n|t_n)$ from the last step of the MKF. For $j = n-1, ..., 1$, $\tilde{\mathbf{Z}}(t_j|t_n)$ and $\tilde{\Sigma}(t_j|t_n)$ are calculated recursively as:

$$\tilde{\mathbf{Z}}(t_j|t_n) = \tilde{\mathbf{Z}}(t_j|t_j) + C(t_j)[\tilde{\mathbf{Z}}(t_{j+1}|t_n) - \tilde{\mathbf{Z}}(t_{j+1}|t_j)],$$

$$\tilde{\Sigma}(t_j|t_n) = \tilde{\Sigma}(t_j|t_j) + C(t_j)[\tilde{\Sigma}(t_{j+1}|t_n) - \tilde{\Sigma}(t_{j+1}|t_j)]C'(t_j),$$

where

$$C(t_j) = \tilde{\Sigma}(t_j|t_j)\tilde{\Phi}'\tilde{\Sigma}^{-1}(t_{j+1}|t_j),$$

$$\tilde{\mathbf{Z}}(t_{j+1}|t_j) = \sum_{i=0}^{k-1} \pi_i(t_{j+1})\tilde{\mathbf{Z}}(t_{j+1}|t_j, I(t_{j+1}) = i),$$

$$\begin{aligned}\tilde{\Sigma}(t_{j+1}|t_j) &= \sum_{i=0}^{k-1} \pi_i(t_{j+1})[\tilde{\Sigma}(t_{j+1}|t_j, I(t_j) = i) \\ &+\{\tilde{\mathbf{Z}}(t_{j+1}|t_j, I(t_{j+1}) = i) - \tilde{\mathbf{Z}}(t_{j+1}|t_j)\}\{\tilde{\mathbf{Z}}(t_{j+1}|t_j, I(t_{j+1}) = i) - \tilde{\mathbf{Z}}(t_{j+1}|t_j)\}'].\end{aligned}$$

The "Bayesian confidence intervals" (Wahba, 1983) for $f(t)$ and its first $m-1$ derivatives can be constructed based on the covariance matrices $\Sigma(t_j|t_n)$.

Bibliography

Ansley, C. F. and Kohn, R. (1985). Estimation, filtering and smoothing in state space models with incompletely specified initial conditions. *Annal of Statistics* **13**, 1286-316

Bolstad, W. (1988) The multiprocess dynamic linear model with biased perturbations: A real time model for growth hormone level. *Biometrika* **75**, 685-92.

Brumback, B. A. and Rice, J. A. (1998) Smoothing spline models for the analysis of nested and crossed samples of curves. *Journal of American Statistical Association* (forthcoming)

Carlin, B. P., Polson, N. G. and Stoffer, D. (1992) A Monte Carlo approach to non-normal and nonlinear state space modeling. *Journal of American Statistical Association* **87**, 493-500.

Carter, C. K. and Kohn, R. (1994) On Gibbs sampling for state space models. *Biometrika* **81**, 541-53.

Carter, C. K. and Kohn, R. (1996) Markov chain Monte Carlo in conditionally Gaussian state space models. *Biometrika* **83**, 589-601.

Clarke, I. J., Cummins, J. T., Findlay, J. K., Burman, K. J. and Doughton, B. W. (1984) Effects on plasma luteinizing hormone and follicle-stimulating hormone of varying frequency and amplitude of gonadotropin-releasing hormone pulses in ovariectomized ewes with hypothalamo-pituitary disconnection. *Neuroendocrinology* **39(3)**, 214-221.

De Jong, P. (1988) The likelihood for state space model. *Biometrika* **75**, 165-9.

De Jong, P. (1991) The diffuse Kalman filter. *Annal of Statistics* **12**, 143-57.

De Jong, P. and Shephard, N. (1995) Simulation smoother for time series models. *Biometrika* **82**, 339-50.

Diggle, P. and Zeger, S. (1989) A non-Gaussian model for time series with pulses. *Journal of American Statistical Association* **84**, 354-59.

Gordon, A. E. and Smith, A. F. M. (1990) Monitoring and modeling biomedical time series. *Journal of American Statistical Association* **85**, 328-37.

Harrison, P. J. and Stevens, C. F. (1971) A Bayesian approach to short-term forecasting. *Operational Research Quarterly* **22**, 341-62.

Harrison, P. J. and Stevens, C. F. (1976) Bayesian forecasting (with discussion). *Journal of Royal Statistical Society, series B* **38**, 205-47.

Harvey, A. C. and Phillips, G.D.A (1979) Maximum likelihood estimation of regression models with autoregressive-moving average disturbances. *Biometrika* **66**, 49-58.

Hoffmann, R. G. (1989) Concordance and irregularity estimation for episodic hormone data, *Proceedings of the Statistical Computing Section, American Statistical Association*, 301-5.

Kitagawa, G. and Gersch, W. (1984) A smoothness priors state-space modeling of time series with trend and seasonality. *Journal of American Statistical Association* **79**, 378-89.

Kohn, R., Ansley, C.F. and Wong, D. (1993) Nonparametric spline regression with prior information. *Biometrika* **80**, 75-88.

Komaki, F. (1993) State-space modeling of time series sampled from continuous processes with pulses. *Biometrika* **80**, 417-29.

Kushler, R. H. and Brown, M. B. (1991) A model for the identification of hormone pulses. *Statistics in Medicine* **10**, 329-40.

Mauger, D. (1995) A new method for charactering hormone pulses in a time series. *unpublished dissertation, Department of Biostatistics, University of Michigan*

Mauger, D., Brown, M. B. and Kushler, R. H. (1995) A comparison of methods that characterize pulse in a time series. *Statistics in Medicine* **14**, 311-25.

Meihold, R. J. and Singurwalla, N. D. (1983) Understanding the Kalman Filter. *The American Statistician* **37**, 123-7.

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**, 339-53.

Rauch, H. E., Tung, F. and Striebel, C. T. (1965) Maximum likelihood estimates of linear dynamic systems. *AIAA Journal.* **3**, 1445-50.

Robinson, G.K. (1991) That BLUP is a good thing: the estimation of random effects (with discussion). *Statistical Science* **6**, 15-51.

Shephard, N. (1994) Partially non-Gaussian state space models. *Biometrika* **81**, 115-31.

Shumway, R. H. and Stoffer, D. S. (1991) Dynamic linear models with switching. *Journal of American Statistical Association* **86**, 763-9.

Shupnik, M. A. (1990) Effects of gonadotropin-releasing hormone on rat gonadotropin gene transcription in vitro: requirement for pulsatile administration for luteinizing hormone-beta gene stimulation. *Molecular Endocrinology* **4(10)**, 1444-1450.

Smith, A. F. M. and West, M. (1983) Monitoring renal transplants: an application of the multiprocess Kalman filter. *Biometrics*, **39**, 867-878.

Wahba, G. (1978) Improper priors, spline smoothing and the problem of guarding against model errors in regression. *Journal of Royal Statistical Society, series B* **40**, 364-72.

Wahba, G. (1983) Bayesian confidence intervals for the cross-validated smoothing spline. *Journal of Royal Statistical Society, series B* **45**, 133-150.

Wahba, G. (1990) *Spline models for observational data.* (CBMS-NSF Regional conference series in applied mathematics, Vol. 59) Philadelphia: Society of Industrial and Applied Mathematics.

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

Wecker, W. E. and Ansley, C. F. (1983) The signal extraction approach to nonlinear regression and spline smoothing. *Journal of American Statistical Association* **78**, 81-89.

Weinert, H. L., Byrd, R. H. and Sidhu, G. S. (1980) A stochastic framework for recursive computation of the spline functions: Part II. Smoothing spline. *Journal of Optimization Theory and Application* **30**, 255-268.

Weiss, J., Jameson, J. L., Burrin, J. M. and Crowley, W. F. (1990) Divergent responses of gonadotropin subunit messenger RNA's to continuous versus pulsatile gonadotropin-releasing hormone in vitro. *Molecular Endocrinology* **4(4)**, 557-564.

West, M. and Harrison, J. (1997) *Bayesian forecasting and dynamic models.* New York: Springer-Verlag.

Wilson, R. C., Kesner, J. S., Kaufman, J. M., Uemura, T., Akema, T. and Knobil, E. (1984) Central electro-physiologic correlates of pulsatile luteinizing hormone secretion in rhesus monkey. *Neuroendocrinology* **39(3)**, 256-260.

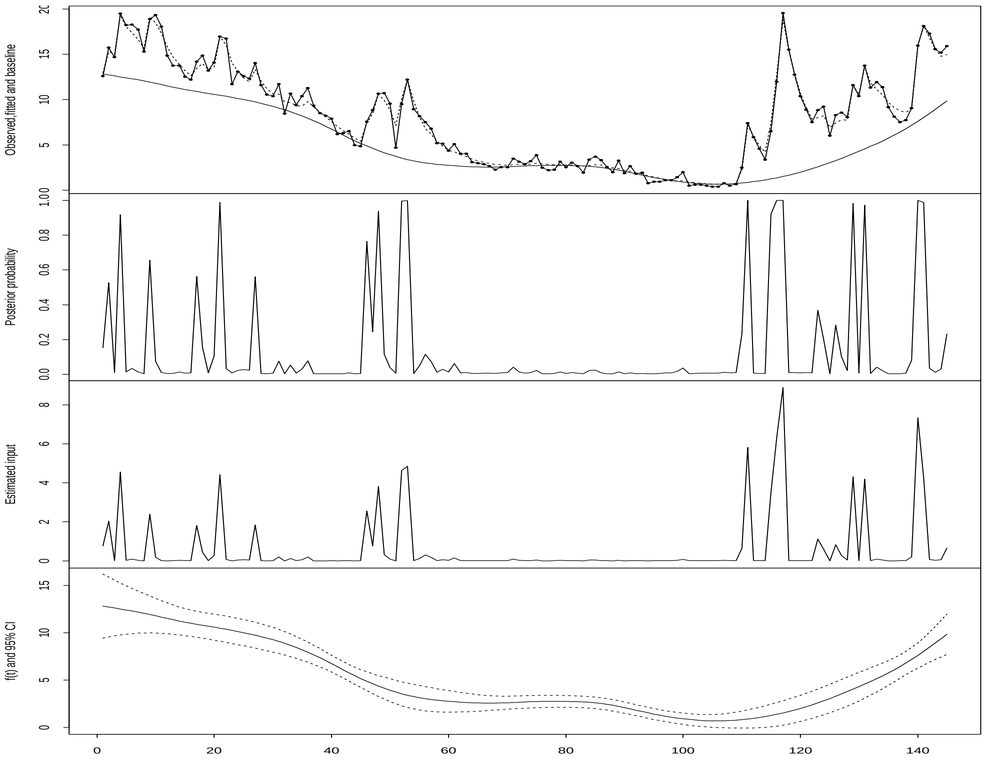

Figure 1: 24 hour cortisol data from a normal subject. The upper panel presents the original data (solid line with dots), its fitted values (dotted line) and the estimated baseline (solid line); the second panel shows the posterior probabilities of pulses; the third panel shows the estimated inputs; the bottom panel presents the estimated baseline (solid line) along with its 95% Bayesian confidence interval (dotted lines).

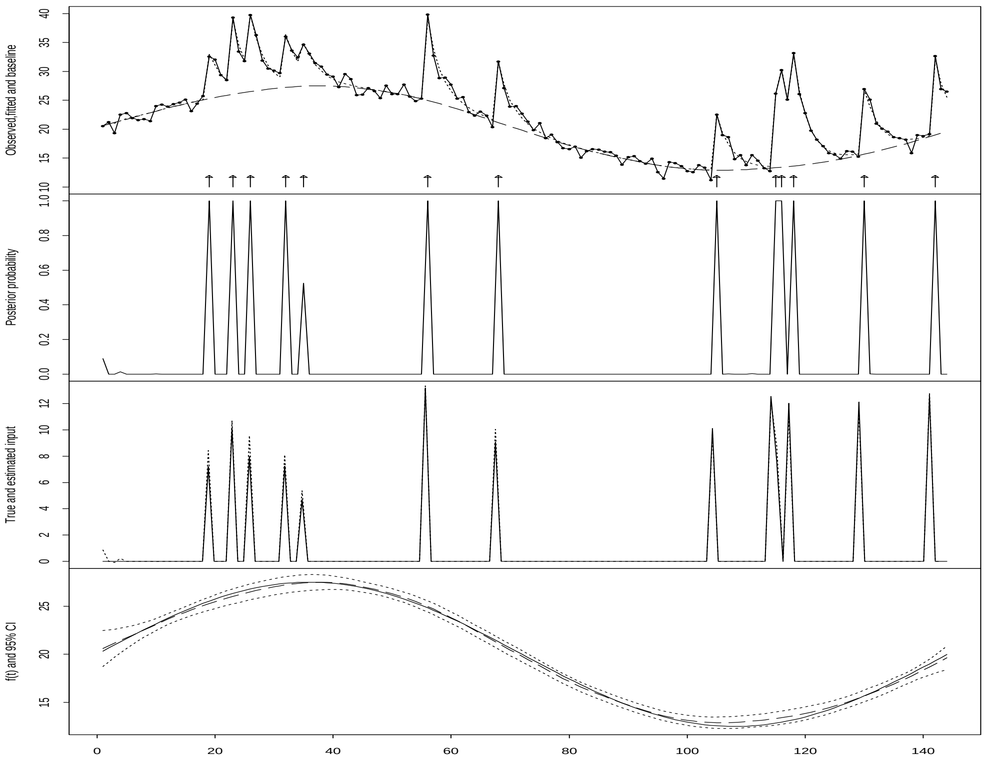

Figure 2: A simulated series. The upper panel presents the simulated data (solid line with dots), its fitted values (dotted line), the estimated baseline (broken line) and the true pulse locations (arrows); the second panel shows the posterior probabilities of pulses; the third panel shows the true inputs (solid line) and their estimates (dotted line); the bottom panel presents the true baseline (solid line), the estimated baseline (broken line) along with its 95% Bayesian confidence interval (dotted lines).

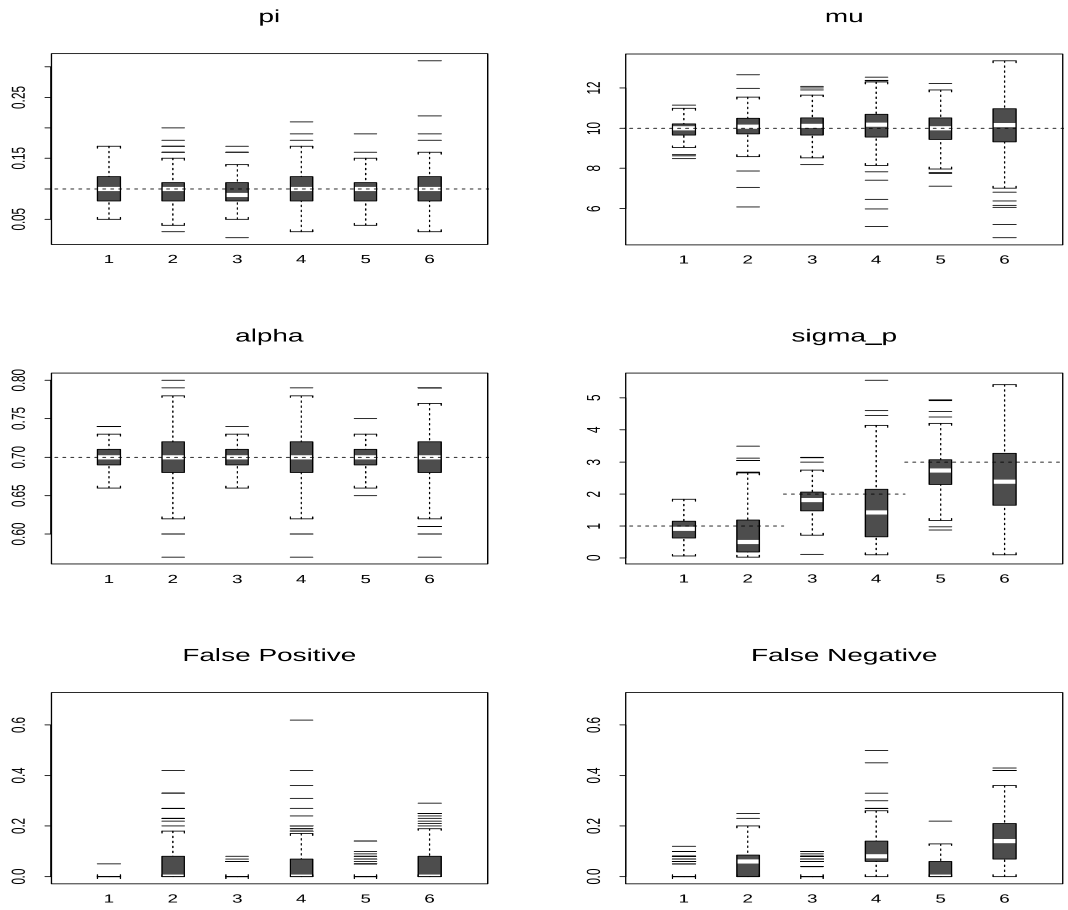

Figure 3: Boxplots of the parameter estimates. Plots from left to right, top to bottom are estimates of π, μ, α, σ_p, false positive rate and false negative rate. The dotted lines indicate the true values in each setting. $\pi = 0.1$, $\mu = 10$, $\alpha = 0.7$ and $A = 2.5$ & 7.5 across all 6 settings. $\{\sigma_e, \sigma_p\} = (1,1), (2,1), (1,2), (2,2), (1,3), (2,3)$ for settings 1 to 6 respectively.