

A Particle Filter Approach to Multiprocess Dynamic Models with Application to Hormone Data

Ziyue Liu

Department of Biostatistics, Fairbanks School of Public Health and School of Medicine, Indiana University, 410 W. 10TH St., Suite 3000, Indianapolis, IN 46202, USA.

Tel.: +001-317-2786734

Fax: +001-317-2742678

E-mail: ziliu@iu.edu

Abstract We extend the multiprocess dynamic models to the general non-Gaussian and nonlinear setting. Under this framework, we propose specific models to simultaneously model hormone smooth basal trend and pulsatile activities. The pulse input is modeled by two processes: one as a point mass at zero and one as a gamma distributed random variable. This gamma-driven approach ensures the pulse estimates to be nonnegative, which is an intrinsic characteristic of hormone dynamics. The smooth trend is modeled by smoothing splines. Both additive and multiplicative observational errors are investigated. Parameters are estimated by maximizing the marginal likelihood. Baseline and pulses are estimated by posterior means. For implementation, particle filter is adopted. Unlike the traditional condensation method where a single distribution is used to approximate a mixture of distributions, this particle filter approach allows the model components to be accurately evaluated at the expense of computational resources. The specific models are applied to a cortisol series. The finite sample performance is evaluated by a simulation. The data application and the simulation show that the biological characteristics can be incorporated and be accurately estimated under the proposed framework.

Keywords Hormone pulses · Multiprocess dynamic models · Particle filter · Smooth baseline · State space

1 Introduction

Simultaneous modeling of hormone smooth basal trends and pulsatile activities has been a scientifically important but statistically challenging task. For many hormones, their secretions have both a basal oscillatory component that varies slowly and a pulsatile component that causes abrupt rises in circulating concentrations (e.g. Veldhuis et al 1989; Keenan and Veldhuis 1997). The secretion combined with distribution and clearance creates complex circulating hormone profiles, where both smooth trends and pulsatile activities are apparent in many cases even by simple visual examinations. Additionally, the effects of circulating hormones on their target tissues can have both a short term component and a long term component. For example, the short term effect of adrenocorticotrophic hormone (ACTH) on the adrenal glands is in minutes and depends on rapid ACTH concentration changes. The long term effect is in hours and depends on the average ACTH levels (e.g. Hanukoglu et al 1990; Gallo-Payet and Payet 2003). Naturally to many researchers, decomposing the hormone dynamics into smooth trends and pulsatile activities and modeling them simultaneously are of great interest.

In the statistical literature, several simultaneous modeling methods have been developed. Komaki (1993) modeled small deviations from a constant baseline by a Wiener process, which can lead to non-smooth baseline estimate. He modeled the pulses by dividing the observation intervals into a finer grid and assuming each small interval to be either secretion mode or elimination mode. However, this partition contradicts the fact that hormones are continuously eliminated. Guo et al (1999) developed a smooth baseline plus pulses model where the baseline was modeled by

smoothing splines. The pulses were modeled by a multiprocess dynamic linear model and the pulse inputs were modeled as Gaussian random variables. This Gaussian distribution may lead to negative pulse input estimates, which by definition should be nonnegative. Yang et al (2006) proposed a two-stage nonlinear mixed effects model for the pulse parameters. The nonlinear mixed effects component allows the pulse estimates to be different from pulse to pulse. However, the first-stage pulse identification methods were originally developed for constant baselines and may fail when the basal trends have big curvatures. Johnson (2007) proposed to combine smooth curves and deconvolution of hormone secretion and elimination for simultaneous modeling. The locations of pulses were modeled by point processes and the implementation was by Markov chain Monte Carlo (MCMC). However, deconvolution is known to be computationally ill-posed (e.g. Wahba 1990, Chapter 8). Adding point processes and smooth curves may further increase the numeric difficulties. A modeling framework that can incorporate the biological properties of hormone dynamics and at the same time can be implemented straightforwardly is still unavailable.

Multiprocess dynamic models (MDMs) present a natural framework for complex hormone dynamics. In the literature, two classes of multiprocess dynamic models (MDMs) have been proposed (Harrison and Stevens 1971, 1976), among which we will focus on Class II in this paper. Instead of assuming that a system evolves over time according to a single process, Class II MDMs allow the system to follow one of several processes at each time point. Consequently, MDMs are extremely flexible and have wide applications such as forecasting, capturing abrupt events and outliers, and modeling complex dynamics (e.g., West and Harrison 1997, Chapter12). Appli-

cations to hormone pulses have been explored under Gaussian and linear structures (Bolstad 1988; Guo et al 1999). Specifically, the pulse input was modeled by two processes: a point mass distribution at zero indicating no pulse and a Gaussian random variable indicating a pulse. The chance of being a pulse was captured by a prior probability to be estimated from the data. The exponential decay of the pulses was characterized by an autoregressive process of order 1 (AR(1)). The observations of the overall hormone levels were modeled using additive Gaussian errors.

One major challenge for MDMs is the computational complexity. For a single-process linear Gaussian dynamic model which is also known as a state space model, its components can be exactly evaluated in the order of $O(T)$ for T time points by the Kalman filtering and smoothing algorithms. For MDMs even with Gaussian and linear structures, such exact algorithms no longer exist. This is because that the multiple processes introduce a mixture of distributions and the number of possible combinations grows exponentially with respect to the number of time points. Harrison and Stevens (1971, 1976) proposed to approximate this mixture by a single distribution at each time point, hence $O(T)$ algorithms can be obtained. Such collapsing or condensation methods have also been extended to several specific non-Gaussian and nonlinear cases (Bolstad 1995; Sohn and Kang 1996; Oikonomou 1997). On the other hand, these approximation methods confine the obtainable accuracies of the estimations. The evaluated likelihood is not the true likelihood. Additionally, hormone dynamics like other biological processes are likely be non-Gaussian or nonlinear or both. How to construct MDMs according to the biological properties and how to calculate the true likelihood and accurately estimate the model components are still

unclear. Note that an alternative is to adopt a fully Bayesian approach and use MCMC for implementation (e.g., Frühwirth-Schnatter 2001), which path we will not pursue in this paper.

In this paper, we adopt a particle filter approach to MDMs to simultaneously model hormone smooth basal trends and pulsatile activities. Particle filter is also known as sequential Monte Carlo. It is a set of algorithms to numerically evaluate integrations where a workable analytic form does not exist (Gordon et al 1993). In this method, the model components including the marginal likelihood are approximated by sequential draws from the underlying distributions. These draws are known as particles. The accuracies of the approximations can be increased to any desired levels at the expense of more particles. As the number of the particles goes to infinity, these approximations converge to the true values (e.g. Del Moral 2004). Particle filter is well developed and has been widely applied to nonlinear and non-Gaussian models. However, the potential applications of particle filter to MDMs have not been explored. By adopting particle filter, MDMs can be accurately evaluated. Additionally, nonlinear and non-Gaussian structures can be incorporated with straightforward implementations. Thus, we first extend MDMs to a general nonlinear and non-Gaussian situation. We then construct specific models for hormone dynamics. The smooth basal trend is viewed as a shared component for all the processes and is modeled by smoothing splines. The pulse inputs are modeled by gamma random variables to ensure non-negativity. Multiplicative observational errors are explored in addition to additive Gaussian errors. In the particle filter implementations, we first draw the indicators of the underlying processes according to their one-step-ahead probabilities. We then

draw one random variable corresponding to each indicator. The numerically evaluated marginal likelihood is used for parameter estimation and inference. Other model components are numerically evaluated by posterior values, which can be obtained by several algorithms such as Godsill et al (2004).

The paper is organized as follows. The model is presented in Section 2. Estimation and inference are in Section 3. A small simulation is performed in Section 4. A cortisol data example is given in Section 5. Discussion and concluding remarks are given in Section 6.

2 The Model

2.1 The General Model

We extend MDMs to a general nonlinear and non-Gaussian situation as follows

$$y(t_j) | \alpha(t_j) \sim p_k \{y(t_j) | \alpha(t_j)\}, \quad (1)$$

$$\alpha(t_j) | \alpha(t_{j-1}) \sim p_k \{\alpha(t_j) | \alpha(t_{j-1})\}, \quad (2)$$

$$\alpha(0) \sim p\{\alpha(0)\}, \quad (3)$$

$$k = 0, \dots, K - 1,$$

for $j = 1, \dots, T$. Equation (1) is the observation equation. The state vector $\alpha(t_j)$ is related to the observation $y(t_j)$ through the probability density function $p_k(\cdot)$. It can be the traditional linear observational equation with Gaussian noise, or linear but with a non-Gaussian noise, or a nonlinear observation equation such as in generalized linear models. The subscript k indicates which process the system follows at time t_j among

K candidates. The probability of being the k th process is $\pi_k(t_j)$ with $\sum \pi_k(t_j) = 1$, which can be a function of the system history up to t_{j-1} . Given all the state vectors, the observations are independent. Equation (2) is the state equation. It defines how the system evolves over time. At each time point, the system can follow one of the K possible processes. The state process is Markovian. Equation (3) defines the initial condition of the system. It can be obtained as a mixture of K initial conditions corresponding to the K processes.

The proposed model can be viewed as an integration of MDMs and nonlinear non-Gaussian state space models. The development of MDMs has mainly focused on linear relationships and Gaussian perturbations. On the other hand, the nonlinear non-Gaussian state space models have focused on a single process. The combination of the two will inherit strengths from both approaches. Especially, the well-developed particle filter in nonlinear non-Gaussian state space models can be adopted for computation. Note that the K processes do not need to be totally different. Some of them may share a common component.

2.2 A Specific Model for Hormone Dynamics

The true circulating hormone concentration $\mu(t_j)$ is decomposed as follows

$$\mu(t_j) = x(t_j) + f(t_j), \quad (4)$$

where $x(t_j)$ is the pulsatile process and $f(t_j)$ is the smooth basal trend. Without loss of generality, time is scaled to $t \in [0, 1]$. The basal hormone level $f(t_j)$ is modeled by a smoothing spline as a member of Sobolev space \mathscr{W}_2 , which comprises functions

that are absolutely continuous up to the first derivative and have square integrable second derivative. It is incorporated into the model as the following dynamic linear model (Wecker and Ansley 1983)

$$f(t_j) = Fv(t_j), \quad (5)$$

$$v(t_j) = H(t_j)v(t_{j-1}) + \eta(t_j). \quad (6)$$

The observation matrix is $F = (1 \ 0)$. The state vector is $v(t_j) = \{f(t_j) \dot{f}(t_j)\}^\top$ and $\dot{f}(t_j)$ is the first derivative with respect to time. The state transition matrix is $H(t_j) = \{1 \ \Delta t_j; 0 \ 1\}$ with $\Delta t_j = t_j - t_{j-1}$. The innovation is $\eta(t_j) \sim N(0, \lambda \Sigma(t_j))$ with $\Sigma(t_j) = \{\frac{1}{3}\Delta t_j^3 \ \frac{1}{2}\Delta t_j^2; \frac{1}{2}\Delta t_j^2 \ \Delta t_j\}$ and λ is the smoothing parameter. Equations (5) and (6) are initialized at time zero as $v(0) \sim N(0, \kappa I_2)$ with $\kappa \rightarrow \infty$ and I_2 is the 2×2 identity matrix.

The pulsatile activities are modeled as

$$x(t_j) = \phi_j x(t_{j-1}) + \xi(t_j). \quad (7)$$

The coefficient ϕ_j captures the exponential elimination. In one-compartment models, elimination takes the form of $\exp(-a \times \Delta t)$ for some parameter $a > 0$. For equally spaced data, it simplifies to a constant coefficient ϕ . The pulse input $\xi(t_j)$ can take one of the two statuses as follows

$$\xi(t_j) = 0, \quad \text{if } I(t_j) = 0, \quad (8)$$

$$\xi(t_j) \sim \text{gamma}(\alpha, \beta), \quad \text{if } I(t_j) = 1. \quad (9)$$

When $I(t_j) = 0$, it indicates that there is no pulse during $(t_{j-1}, t_j]$. When $I(t_j) = 1$, it indicates that there is one or more pulses. The probability of $I(t_j) = 1$ is $\pi_1(t_j) =$

$\Pr\{I(t_j) = 1\}$. For simplicity considerations, we assume that $I(t_j)$'s are serially independent and $\pi_1(t_j) = \pi$ for all j 's. Innovations $\eta(t_j)$ and $\xi(t_j)$ are serially and mutually independent.

The initialization has two parts: $v(0)$ and $x(0)$. For $x(t)$ with zero innovations, it eventually decreases to zero as its stationary distribution. For $x(t)$ with $\text{gamma}(\alpha, \beta)$ innovations, its stationary mean and variance are $\alpha/\{\beta(1-\phi)\}$ and $\alpha/\{\beta^2(1-\phi^2)\}$, respectively (Grunwald et al 2000). Consequently, we initialize $x(t)$ by treating $x(0)$ as a mixture of the two where the later one has a weight of π . For $v(0)$, exact diffuse initialization is well developed in the linear Gaussian case (see, e.g. Durbin and Koopman 2012, Chapter 5). It is, however, much less studied in the particle filters. We thus adopt a numeric diffuse by setting κ to a large number such as 1000.

The true hormone concentration $\mu(t_j)$ can be observed with either additive errors or multiplicative errors as follows

$$y(t_j) = \mu(t_j) + \varepsilon(t_j), \quad \text{or} \quad y(t_j) = \mu(t_j) e(t_j), \quad (10)$$

where $y(t_j)$ is the observed value. The additive error is usually modeled as $\varepsilon(t_j) \sim N(0, \sigma_\varepsilon^2)$. The multiplicative error can be modeled by a nonnegative random variable with unit mean. In this paper, we will investigate lognormal multiplicative errors parameterized as $\log\{e(t_j)\} \sim N(-\sigma_\varepsilon^2/2, \sigma_\varepsilon^2)$. This parameterization ensures that $E\{e(t_j)\} = 1$. Both the additive errors and the multiplicative errors are assumed to be serially independent and identically distributed. Overall, this specific model has six parameters $\theta = (\lambda \ \phi \ \pi \ \alpha \ \beta \ \sigma_\varepsilon^2)^\top$ or $\theta = (\lambda \ \phi \ \pi \ \alpha \ \beta \ \sigma_\varepsilon^2)^\top$.

3 Estimation and Inference

The parameters θ are estimated by maximizing the marginal likelihood which can be decomposed as a product of sequential conditional densities

$$L(\theta | y) = p\{y(t_1) | \theta\} \prod_{j=2}^N p\{y(t_j) | y_{1:j-1}, \theta\}, \quad (11)$$

where y is the collection of all observed data and $y_{1:j} = \{y(t_1), \dots, y(t_j)\}$. The maximized marginal likelihood can be used for inference purpose such as model selection. The model components are estimated by the posteriors. In particle filter, these are approximated by discrete weighted random draws. A large body of literature on particle filter has been produced. Interested readers can see Creal (2012) for a review. For the filtering step, we adapt the classical bootstrap filter (Gordon et al 1993), which utilizes the transition density to generate random draws. For a given parameter vector θ , the algorithm is outlined in the follows for N particles.

Algorithm 1 *Bootstrap filter (Gordon et al 1993)*

1. At $j = 0$, for $i = 1, \dots, N$, draw $\alpha^{(i)}(0)$ according to its initial distributions.
2. For $j = 1, \dots, T$:
 - (1) For $i = 1, \dots, N$:
 - (i) Draw $I^{(i)}(t_j)$ from a multinomial distribution with parameters $n = 1$ and (π_1, \dots, π_K) .
 - (ii) Based on the drawn indicator $I^{(i)}(t_j) = k$, draw $\alpha^{(i)}(t_j) \sim p_k\{\alpha^{(i)}(t_j) | \alpha^{(i)}(t_{j-1})\}$.
 - (iii) Calculate the weight $w^{(i)}(t_j) = p_k\{y(t_j) | \alpha^{(i)}(t_j)\}$.

(2) For $i = 1, \dots, N$, normalize the weights as

$$\tilde{w}^{(i)}(t_j) = \frac{w^{(i)}(t_j)}{\sum_{b=1}^N w^{(b)}(t_j)}$$

(3) The filtered probability of being the k th process is

$$\frac{\sum_{I^{(a)}(t_j)=k} \tilde{w}^{(a)}(t_j)}{\sum_{b=1}^N \tilde{w}^{(b)}(t_j)}$$

(4) Resample N particles $\tilde{\alpha}^{(i)}(t_j)$'s with replacement from $\{\alpha^{(i)}(t_j)\}_{i=1}^N$ according to weights $\{\tilde{w}^{(i)}\}_{i=1}^N$.

(5) For $i = 1, \dots, N$, let $\alpha^{(i)}(t_j) = \tilde{\alpha}^{(i)}(t_j)$.

Resampling is an essential step of the particle filter. Without resampling, the sample path will soon be dominated by a few or even only one particle. This is the well-known degeneracy problem. Resampling can be performed at each time point, or only when the effective sample size (ESS), defined as $ESS = [\sum_{i=1}^N \{\tilde{w}^{(i)}(t_j)\}^2]^{-1}$, becomes below certain threshold (Liu and Chen 1998). The sample before and after resampling are both random samples from the filtered distribution. Before resampling, particles have different weights. After bootstrap resampling, particles are equally weighted. In calculating filtered items, it is usually the sample before resampling is used. The conditional density that can be used for likelihood evaluation is approximated as

$$\hat{p}\{y(t_j) \mid y_{1:j-1}, \theta\} = \frac{1}{N} \sum_{i=1}^N w^{(i)}(t_j). \quad (12)$$

To generate posterior samples, Godsill et al (2004)'s algorithm is adopted.

Algorithm 2 *Godsill et al (2004)*

For $l = 1, \dots, L$:

1. Sample $\hat{\alpha}^{(l)}(t_T)$ from $\{\alpha^{(i)}(t_T)\}_{i=1}^N$ with equal weight $1/N$.

2. For $j = T - 1, \dots, 1$:

(1) For $i = 1, \dots, N$, calculate weight

$$v^{(i)}(t_j | t_{j+1}) \propto p\left\{\hat{\alpha}^{(i)}(t_{j+1}) | \alpha^{(i)}(t_j)\right\}.$$

(2) Sample $\hat{\alpha}^{(i)}(t_j)$ from $\{\alpha^{(i)}(t_j)\}_{i=1}^N$ according to weights $\{v^{(i)}(t_j | t_{j+1})\}_{i=1}^N$.

3. $\hat{\alpha}_{1:T}^{(i)} = \{\hat{\alpha}^{(i)}(t_1), \dots, \hat{\alpha}^{(i)}(t_T)\}$ is a realization from $p(\alpha_{1:T} | y)$.

From the posterior sample, we can estimate the means, medians and variations for all the model components. The posterior probability of being a particular process at time t_j can be calculated as the proportion of $\{\hat{\alpha}^{(l)}(t_j)\}_{l=1}^L$ that come from the k th process.

4 Simulation

In this section we evaluate the finite sample performance of the proposed method and compare it with Guo et al (1999) and Yang et al (2006), which are hereafter referred as GWB and YLW for the simplicity of presentation. Gaussian additive errors are adopted for the reasons given in the next section. Three scenarios are used.

- (S1). The baseline function is $f(t) = 20 + 15 \sin(2\pi t)$. For each time point, the probability of being a pulse is 0.10. The magnitude, if a pulse, follows a normal distribution $\xi_j \sim N(10, 2^2)$, and $\phi = 0.7$. Serially independent errors follow $\varepsilon_j \sim N(0, 1^2)$.
- (S2). The baseline function is $f(t) = 0.5 \cos(2\pi t) + 2$, $t \in [0, 1]$. Pulse locations are generated from an inhomogeneous Poisson process with intensity function $\lambda(t) = 35(0.26 - (t - 0.5)^2)$. For each pulse, its magnitude follows a lognormal distribu-

tion $\log(\alpha_i) \sim N(1, 0.5^2)$. The pulse function is double exponential as

$$p(t - \tau_i) = \begin{cases} \exp\{100(t - \tau_i)\}, & t < \tau_i, \\ \exp\{-\gamma_i(t - \tau_i)\}, & t \geq \tau_i, \end{cases}$$

where γ_i follows a lognormal distribution as $\log(\gamma_i) \sim N(3.66, 0.27^2)$. Serially independent errors follow $\varepsilon_j \sim N(0, 0.5^2)$.

- (S3). The baseline function is $f(t) = 5 \sin(4\pi t + \pi) + 10$. For each time point, the probability of being a pulse is 0.15. The magnitude, if a pulse, follows a gamma distribution $\xi_j \sim \text{gamma}(15.70, 0.33)$, and $\phi = 0.8$. Serially independent errors follow $\varepsilon_j \sim N(0, 0.85^2)$.

Scenario 1 is adopted from GWB. The baseline function has big curvatures. The pulses have big magnitudes and small variations. This means we have a high signal-to-noise ratio in this scenario. Scenario 2 is adopted from YLW. With the given intensity function, the number of overall pulses follows a Poisson distribution $M \sim \text{Poisson}(28/3)$. We do not fix M at 5 or 10 as in YLW. The baseline function has small curvatures, which is important for the first stage pulse detection to work properly. Scenario 3 has a gamma pulse input as the true signals. The baseline function also has more curvature than Scenario 2. For each scenario, 100 series are generated. Each series has $T = 145$ time points with $t_j = (j - 1)/144$. The proposed method, GWB and YLW are applied to each series. For the proposed method, $N = 1000$ particles and $L = 100$ posterior draws are adopted.

The proposed method produces reliable estimates for ϕ . For Scenario 1 where $\phi = 0.7$, the median of the parameter estimates is 0.69. For Scenario 2 where the equivalent ϕ 's are around 0.76, the median is 0.75. For Scenario 3 where $\phi = 0.8$,

the median is 0.80. This suggests that the numerically evaluated marginal likelihood by particle filter works well in parameter estimate. We further compare the proposed method to GWB and YLW using four criteria: 1) mean squared errors (MSEs) of the estimated baseline, which are calculated as the average squared difference between the true baseline and the estimated baseline on the observational grid; 2) false positive pulse numbers, which are the numbers of observations falsely identified as pulses; 3) false negative pulse numbers, which are the numbers of true pulses that are not identified; 4) MSEs of the pulse inputs. For the proposed method and GWB, a pulse is classified if its posterior probability is equal to or bigger than 0.5.

The results are displayed in Table 1 as median and interquartile range (IQR). In estimating the baseline functions, GWB has the smallest medians and IQRs of MSE for Scenario 1 and 2, and the proposed method has the smallest median and IQR for Scenario 3. YLW, on the other hand, fails to reasonably estimate the baseline functions for Scenario 1 and 3. In identifying the pulses, GWB produces the fewest false negatives and false positives for Scenario 1 and 3, but it generates the most false positives for Scenario 2. The proposed method performs similarly to GWB for Scenario 1 and 3, but it leads to fewer false positives compared to GWB for Scenario 2. YLW generates the most false positives and false negatives for Scenario 1 and 3, while it does produce the fewest false positives for Scenario 2. For pulse input magnitude estimates, GWB performs best for Scenario 1 and 3, while the proposed method performs best for Scenario 2. One major disadvantage of GWB is that it may lead to negative pulse estimates. For example for Scenario 3 within the 100 series, GWB leads to 12 negative pulse estimates with posterior pulse probability ≥ 0.5 , and

Table 1 Simulation results displayed as (median, interquartile range)

	Data	Method		
		GWB	YLW	Proposed
Baseline MSE	S1	0.22, 0.22	12.09, 20.34	0.41, 0.29
	S2	0.08, 0.15	0.28, 0.66	0.33, 0.35
	S3	0.63, 0.59	10.78, 11.86	0.61, 0.56
False positive	S1	0, 0	2, 2	0, 0
	S2	6, 6	2, 2	2.5, 3
	S3	0, 1	2, 2	0, 1
False negative	S1	0, 1	9, 4	0, 1
	S2	2, 2.5	4, 2.5	4, 3
	S3	2, 2	17, 4.5	3, 3
Pulse MSE	S1	0.84, 0.53	19.66, 14.40	2.93, 1.99
	S2	1.86, 1.70	3.74, 4.18	1.82, 2.26
	S3	0.28, 0.20	3.49, 0.91	0.50, 0.26

23 negative pulse estimates with posterior probability within $[0.25, 0.5)$. For Scenario 1, these numbers are 2 and 3. For Scenario 2, these numbers are 19 and 95.

5 Application to a Cortisol Series

In this section, we apply the proposed method to a cortisol series. We then compare it to two other methods: GWB (Guo et al 1999) and YLW (Yang et al 2006). The data came from a study conducted at the University of Michigan Medical Center. The study aimed to evaluate and compare the hypothalamic-pituitary-adrenal (HPA) axis between patients with chronic fatigue syndrome or fibromyalgia or both and healthy controls. The cortisol series used in this paper came from a healthy control. Blood

samples were collected at 10-minute intervals over a 24-hour period beginning at 9am, hence 145 observations for each series. Details of the study can be found in Crofford et al (2004). As the end product of the HPA axis, cortisol is a well-known stress-related hormone and is of particular interest to many researchers. To fully understand the dynamics of cortisol, a model that can incorporate both its smooth baseline and pulsatile activities according to the biological characteristics is needed.

5.1 Additive Errors versus Multiplicative Errors

The model in Section 2.2 with additive Gaussian errors was first applied to the cortisol series. Parameter estimates are $\hat{\phi} = 0.80$, $\hat{\pi} = 0.34$, $\hat{\alpha} = 1.48$ for the gamma shape parameter, $\hat{\beta} = 1.05$ for the gamma scale parameter, and $\hat{\sigma}_\varepsilon^2 = 0.21$. Figure 1 displays the fitting results. The figure shows that the overall fitting is reasonably well. The baseline and pulse inputs are estimated as posterior means from $N = 10000$ particles and $L = 1000$ posterior draws. The estimated baseline tracks the overall trend. The estimated pulse inputs are all nonnegative and agree with the abrupt jumps in the observed values. The proposed method generates posterior pulse probabilities that have many nonzero values. One choice is to classify each point based on whether the posterior pulse probability is bigger than 0.5. In this data example, the identified pulses using this threshold all have nonzero magnitudes, while the identified non-pulses have at most negligible magnitudes.

We then compared additive Gaussian errors with multiplicative lognormal errors. In practice, hormone data have been analyzed both on the original scale and after logarithmic transform. The rationale for logarithmic transform is that the measure-

ment errors may be proportional to the true hormone levels. Indeed, the accuracies of hormone assays are commonly quantified by coefficient of variation. Consequently, a multiplicative error term on the original scale with unit mean seems worth investigating. The same particle filter approach can be applied, except that the observational density is lognormal. Figure 2 displays the fitting results. The estimated baseline from multiplicative errors is very similar to the one from additive errors, hence not displayed. At low overall hormone levels, the multiplicative error approach tends to attribute small variations to pulses, such as the two pulses identified 5pm–6pm and 8pm–9pm. On the other hand, the additive error approach characterizes these variations as noises. For high overall hormone levels, the multiplicative error approach tends to attribute relatively large variations to noises, such as 11am–1pm and 5am–9am. But the additive error approach characterizes these variations by more pulses.

For this particular cortisol example, we prefer the additive error approach because of two reasons. Firstly, the multiplicative error approach leads to suspicious overall estimates, e.g., a peak between 11am–12noon, another peak around 4pm and a quick dip between 3pm–4pm. On the other hand, the additive model approach fits these time periods reasonably well. Secondly, high cortisol levels following sharp rise are usually considered to be the results of multiple cortisol pulses. For example for the period around 6am, the additive error approach identifies multiple pulses that can explain these patterns well. But the multiplicative error approach identifies too few pulses to explain these patterns. However, the multiplicative error approach would be preferred when high resolution is needed at low hormone levels.

5.2 Comparisons with Two Other Methods

Figure 3–5 display the comparisons of the proposed additive-error model with GWB and YLW. Figure 3 displays the estimated baselines, where three methods agree with each other for 1pm–3am. The proposed method led to lower estimates for 9am–1pm and 3am–9am. It also generated baseline that roughly follows a 24-hour period. This would be preferred because cortisol basal trends are supposed to exhibit a 24-hour circadian rhythm. Figure 4 displays the estimated pulse inputs. YLW estimated one pulse during 9am–1pm and four pulses during 3am–9am, which are less likely to generate the observed patterns especially for the morning period. Both the proposed method and GWB estimated multiple pulses for these two time periods, which explain the observed patterns well. This suggests that YLW may underestimate the number of pulses. GWB led to multiple small negative pulse input estimates during 9am–11am and a big negative pulse input estimate around 8am, which are at odds with the biological properties of hormone pulses. Figure 5 displays the estimated pulse probability. The proposed method produces estimates closer to zero than GWB for locations where pulses are unlikely. One possible reason is that the estimated pulse probability in GWB is only conditional on up-to-current observations $y_{1:j}$, while in the proposed method it is conditional on all the observations $y_{1:T}$.

6 Discussion

We have extended MDMs to a general nonlinear and non-Gaussian setting. The proposed particle filter approach allows the model components to be accurately evalu-

ated. In the literature, several specific nonlinear or non-Gaussian MDMs have been developed, where condensation was used for implementation. For example, Bolstad (1995) extended the condensation method to an MDM with Poisson observations which allows shifts in the mean level and outliers. Sohn and Kang (1996) adopted a nonlinear state equation with zero-mean perturbations and used linearization and condensation for implementation. Oikonomou (1997) extended the condensation method to a system with gamma-distributed observations which has a variance inflation state equation. The proposed particle filter approach is suitable for the general nonlinear and non-Gaussian setting. This would greatly broaden the types of MDMs that can be constructed and applied. However, these advantages do require more computational resources. For a simulated series in Section 4, it takes less than a minute for Guo et al (1999)'s method on a personal computer with 3.00GHz Intel Core i7 3540M CPU and 8.00GB RAM. For Yang et al (2006)'s method, it takes 3–5 minutes. For the proposed method, it takes around 30 minutes.

For hormone dynamics, we have proposed a gamma-driven approach for hormone pulses and allow additive errors and multiplicative errors. The data application and simulation show that the parameters can be reliably estimated and the hormone pulse estimates agree with their biological properties. In the literature, there are mainly two approaches in modeling the hormone pulses. One approach is linear Gaussian MDMs. The other approach is the deconvolution method developed by Veldhuis and Johnson (e.g. Veldhuis and Johnson 1992). This method assumes that the circulating hormone concentrations are the results of a convolution of pulsatile secretions and exponential eliminations and aims to disentangle these two components. Both MDMs and the

deconvolution method come from compartment models. The deconvolution method emphasizes characterizing the hormone secretion patterns, while MDMs emphasize capturing the circulating hormone patterns. In this paper, we have focused on the MDM approach. MDMs are numerically easier to work with, and they also provide sufficient information for hormone circulating patterns.

Conceptually, all the methods for simultaneous modeling of hormone smooth trend and pulsatile activities can be recast as MDMs. The smooth baseline plus pulses model proposed by Guo et al (1999) is already in the form of MDM. The state space approach proposed by Komaki (1993) can be viewed as a special case of MDMs. The two-stage nonlinear mixed effects model proposed by Yang et al (2006) and the Bayesian approach by Johnson (2007) can also be reformulated as MDMs, if the secretion phase is assumed to be finite in time. Therefore, the proposed method provides a general and unified approach for modeling complex hormone dynamics.

Conflict of Interest Disclosure Statement

The author has no conflict of interest.

References

- Bolstad WM (1988) The multiprocess dynamic linear model with biased perturbations: a real time model for growth hormone level. *Biometrika* 75:685–692
- Bolstad WM (1995) The multiprocess Poisson model. *Journal of the American Statistical Association* 90:227–232
- Creal D (2012) A survey of sequential Monte Carlo methods for economics and finance. *Economics Reviews* 31:245–296

- Crofford LJ, Young EA, Engleberg NC, Korszun A, Brucksch CB, McClure LA, Brown MB, Demitrack MA (2004) Basal circadian and pulsatile ACTH and cortisol secretion in patients with fibromyalgia and/or chronic fatigue syndrome. *Brain, Behavior, and Immunity* 18:314–325
- Del Moral P (2004) Feynman-Kac formulae: genealogical and interacting particle systems with applications. Springer-Verlag, New York
- Durbin J, Koopman SJ (2012) Time series analysis by state space methods, 2nd edn. Oxford University Press
- Frühwirth-Schnatter S (2001) Markov chain Monte Carlo estimation of classical and dynamic switching and mixture models. *Journal of the American Statistical Association* 96:194–209
- Gallo-Payet N, Payet MD (2003) Mechanism of action of ACTH: beyond cAMP. *Microscopy Research and Technique* 61:275–287
- Godsill SJ, Doucet A, West M (2004) Monte Carlo smoothing for nonlinear time series. *Journal of the American Statistical Association* 99:156–168
- Gordon N, Salmond DJ, Smith AFM (1993) A novel approach to nonlinear and non-Gaussian Bayesian state estimation. *IEE Proceedings Part F: Radar and Sonar Navigation* 140:107–113
- Grunwald GK, Hyndman RJ, Tedesco L, Tweedie RL (2000) Non-Gaussian conditional linear AR(1) models. *Australian & New Zealand Journal of Statistics* 42:479–495
- Guo W, Wang Y, Brown MB (1999) A signal extraction approach to modeling hormone time series with pulses and a changing baseline. *Journal of the American Statistical Association* 94:746–756
- Hanukoglu I, Feuchtwanger R, Hanukoglu A (1990) Mechanism of corticotropin and cAMP induction of mitochondrial cytochrome P450 system enzymes in adrenal cortex cells. *The Journal of Biological Chemistry* 265:20,602–20,608
- Harrison PJ, Stevens CF (1971) A Bayesian approach to short-term forecasting. *Operational Research Quarterly* 22:341–362
- Harrison PJ, Stevens CF (1976) Bayesian forecasting. *Journal of the Royal Statistical Society, Series B (Methodological)* 38:205–247
- Johnson TD (2007) Analysis of pulsatile hormone concentration profiles with nonconstant basal concentrations: a Bayesian approach. *Biometrics* 63:1207–1217

- Keenan DM, Veldhuis JD (1997) Stochastic model of admixed basal and pulsatile hormone secretion as modulated by a deterministic oscillator. *American Journal of Physiology (Regulatory, Integrative and Comparative Physiology)* 273:R1173–R1181
- Komaki F (1993) State-space modelling of time series sampled from continuous processes with pulses. *Biometrika* 80:417–429
- Liu J, Chen R (1998) Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association* 93:1032–1044
- Oikonomou KN (1997) Prediction with the dynamic Bayesian gamma mixture model. *IEEE Transaction on Systems, Man, and Cybernetics, Part A* 27:529–542
- Sohn JK, Kang SG (1996) Bayesian estimation procedure in multiprocess non-linear dynamic generalized model. *Communications in Statistics: Theory and Methods* 25:2281–2296
- Veldhuis JD, Johnson ML (1992) Deconvolution analysis of hormone data. In: Brand L, Johnson ML (eds) *Methods in Enzymology*, Academic Press, San Diego, pp 539–575
- Veldhuis JD, Iranamesh A, G L, Johnson ML (1989) Amplitude modulation of a burstlike mode of cortisol secretion subserves the circadian glucocorticoid rhythm in man. *American Journal of Physiology* 257:6–14
- Wahba G (1990) Spline models for observational data. *CBMS-NSF Regional Conference Series in Applied Mathematics*, SIAM, Philadelphia
- Wecker WE, Ansley CF (1983) The signal extraction approach to nonlinear regression and spline smoothing. *Journal of the American Statistical Association* 78:81–89
- West M, Harrison J (1997) *Bayesian forecasting and dynamic models*, 2nd edn. Springer-Verlag, New York
- Yang YC, Liu A, Wang Y (2006) Detecting pulsatile hormone secretion using nonlinear mixed effects partial spline models. *Biometrics* 62:230–238

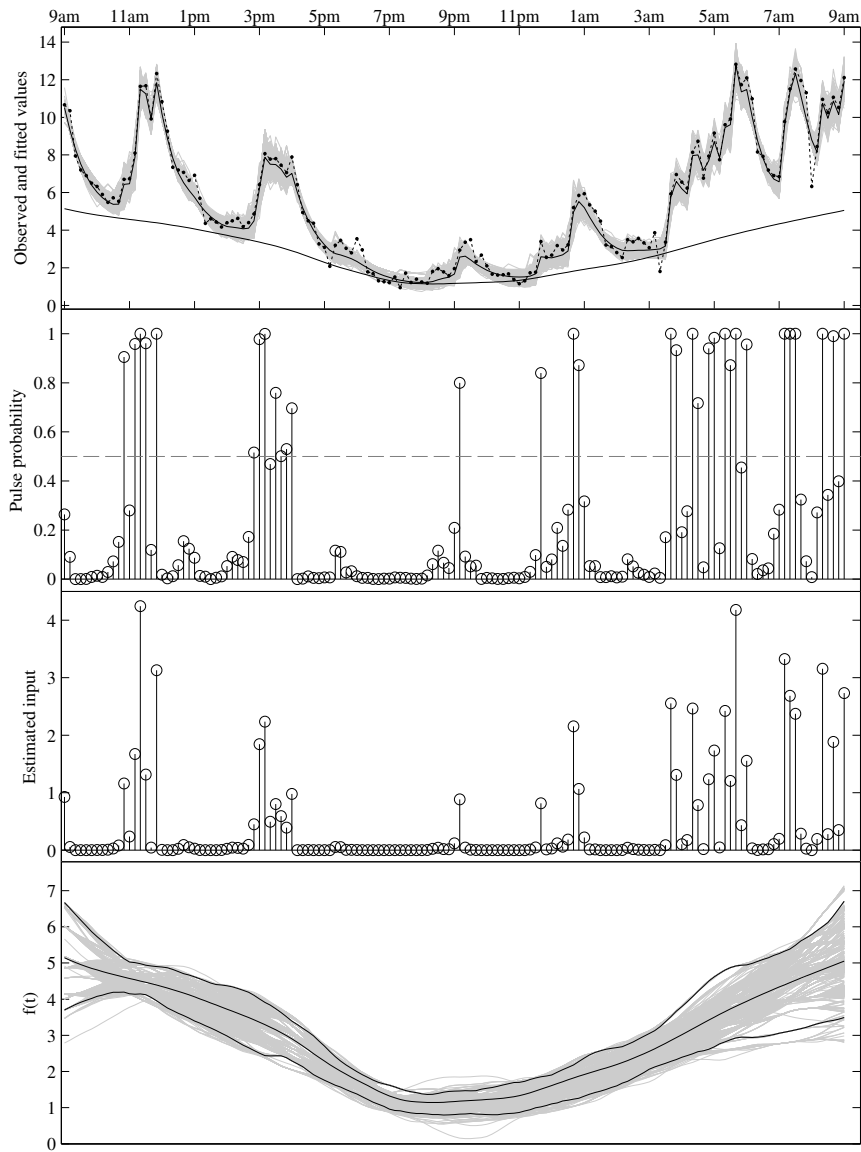


Fig. 1 Additive-error model fitting results. The top panel displays the observed data as black dots connected by a dotted line, the posterior draws of the overall fit as gray lines, the point-wise mean of the posterior draws as a solid black line, and the fitted baseline as a solid black line. The second panel displays the posterior pulse probabilities. The gray horizontal line is at value 0.5. The third panel displays the estimated pulse inputs. The bottom panel displays the draws from the baseline posterior as gray lines, the point-wise mean values and 95% confidence intervals as solid black lines.

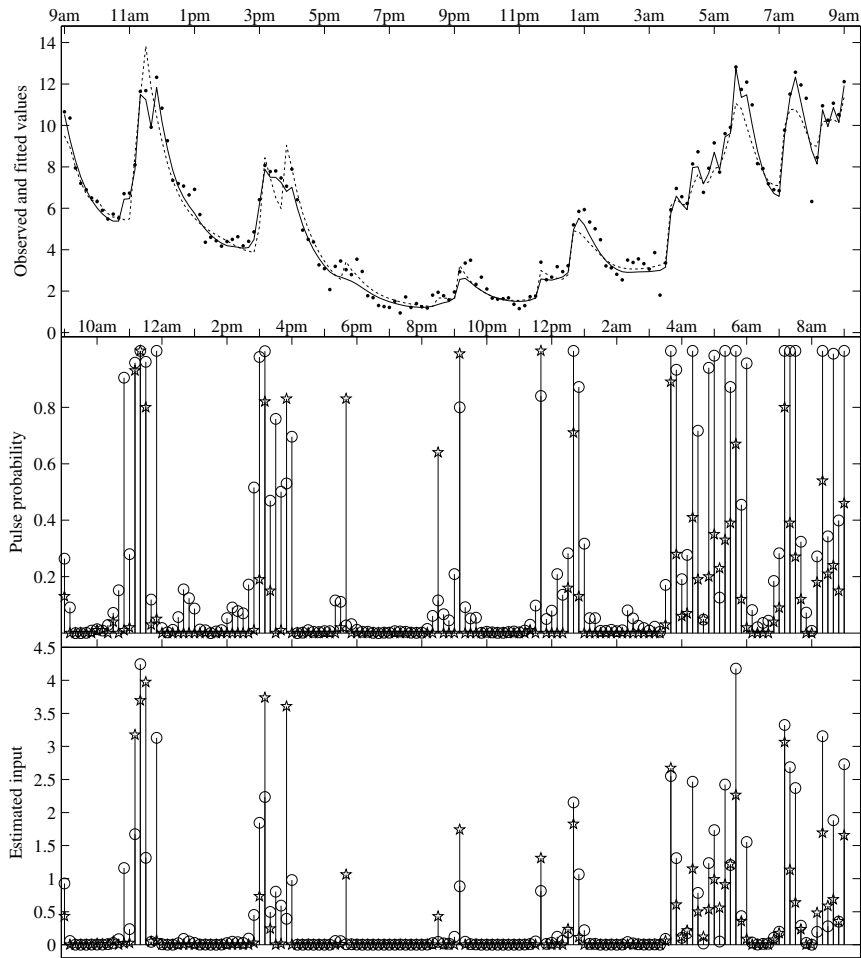


Fig. 2 Comparing additive errors with multiplicative errors. The top panel displays the observed data as black dots and the point-wise mean of the posterior draws from the proposed models, with the solid line for additive errors and the dotted line for multiplicative errors. The middle panel displays the posterior pulse probabilities, with circles for additive errors and pentagrams for multiplicative errors. The bottom panel displays the estimated pulse inputs, with circles for additive errors and pentagrams for multiplicative errors.

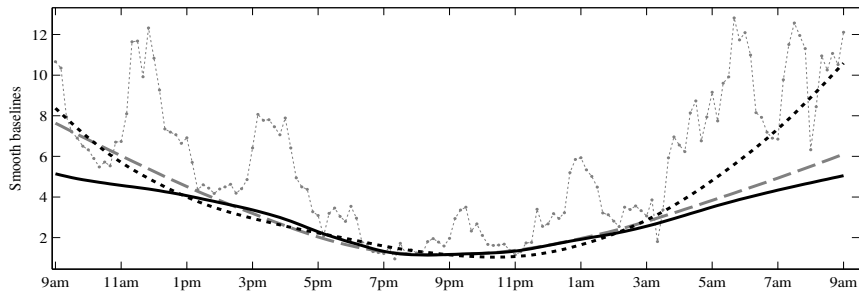


Fig. 3 Estimated baselines displayed as: Guo et al (1999) by the dashed curve, Yang et al (2006) by the dotted curve, and the proposed method by the solid curve. The observed data are displayed as dots connected by a dotted line.

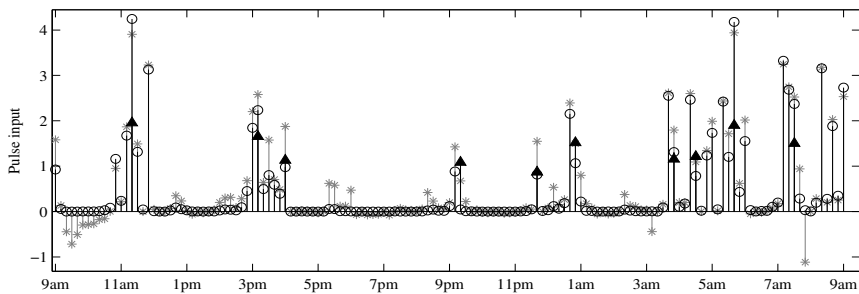


Fig. 4 Estimated pulse inputs displayed by stems with markers as: Guo et al (1999) by asterisks, Yang et al (2006) by solid triangles, and the proposed method by circles.

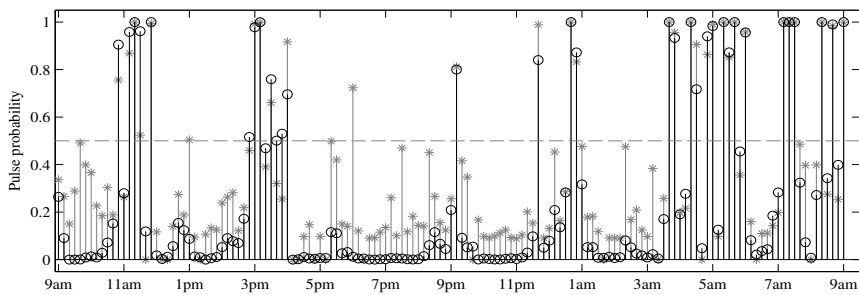


Fig. 5 Estimated pulse probabilities displayed by stems with markers as: Guo et al (1999) by asterisks, and the proposed method by circles. The gray horizontal line is at value 0.5.