Structural Time Series Models with Feedback Mechanisms

Wensheng Guo
Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, U.S.A.
email: wguo@ccgeb.upenn.edu

and

Morton B. Brown
Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48109-2029, U.S.A.

SUMMARY. Structural time series models have applications in many different fields such as biology, economics, and meteorology. A structural time series model can be represented as a state-space model where the states of the system represent the unobserved components and the structural parameters have clear interpretations. This paper introduces a class of structural time series models that incorporate feedback from the latent components of the history. An iterative procedure is proposed for estimation. These models allow flexible and robust feedback mechanisms, have clear interpretations, and have a computationally efficient estimation procedure. They are applied to hormone data to characterize hormone secretion and to explore a potential feedback mechanism.

KEY WORDS: Conditionally Gaussian model; Feedback; Multiprocess dynamic linear model; Pulsatile time series; State-space model.

1. Introduction

Hormone concentrations in the blood regulate many processes, including the secretion of other hormones. It is believed that there is a feedback mechanism such that, when the second hormone attains sufficient concentration in the blood, the secretion of the first hormone will be turned off. This pattern may cause a regularity in the pattern of secretion to be observed in either of the hormones in this feedback relationship. Often one of the two hormones may not be observable because its concentration in the blood is too low to be assayed. Therefore, the goal is to model the feedback relationship as observed in only one hormone. For example, Figure 1 presents the luteinizing hormone data studied by Midgley et al. (1997). The estimated results will be further explained in Section 5. Blood samples were drawn from the jugular vein of each of the six ewes every 5 minutes for 6 hours ($N = 72$) and assayed for the concentrations of luteinizing hormone (LH). LH is released from the pituitary in a pulsatile pattern and has a central role in regulating the reproductive cycle. It is known that the release of LH is regulated by gonadotropin releasing hormone (GnRH) produced by the hypothalamus. It is also believed that many hormones such as LH, cortisol, progesterone, and estradiol may provide feedback to GnRH. In spite of this complicated feedback system, the regularity in LH secretion is visible. The connection between regularity and negative feedback will be further explained in Section 6. Also, in the human, it is not possible to measure GnRH. When only LH is observed, we can still use a probabilistic feedback relationship to model the feedback mechanism of LH on itself.

Feedback mechanisms, especially negative feedback, exist in many time series applications across different fields, including biological signals. The feedback mechanism helps maintain a regularity that may still be visible even when the feedback system involves many intermediate factors. By modeling the feedback mechanism, we gain a better understanding of the underlying process and obtain more efficient estimates. In this article, we introduce a class of structural time series models that incorporate feedback from the history. The time series is modeled by a structural model and the structural parameters of interest are modeled as functions of the latent signals of the history. These models allow flexible and robust feedback mechanisms, have clear interpretations, and have a computationally efficient estimation procedure.

A structural model can be written as a state-space model with the state of the system representing the various unobserved components and the parameters (structural parameters) having clear interpretations. Discussion of the methodological and technical ideas underlying the structural time series models can be found in Harvey (1989) and Harvey and Shephard (1992). A Bayesian view of structural models can be found in West and Harrison (1997). Guo and Brown (2000, in press) introduced a bivariate structural time series model called cross-related time series models (CRSMs), where the connection across different time series is modeled through the

686
cross-regression level. Their model has the advantage of flexibility and robustness as a result of the use of a probabilistic relationship. In this article, we adopt a similar idea but in a univariate setting. We extend the feedback mechanism to allow the feedback signal to be a linear combination of the latent components (state vector).

Similar to the CRSMs, our models are conditionally Gaussian in general, and the multimover simulation technique (Shephard, 1994) or the approximate method by Harrison and Stevens (1976) can be extended to them. However, these estimation methods are usually computationally intensive and require special software. In this article, we propose a straightforward approximate estimation method that only uses the estimated mean of the latent history in the feedback equation. Conditional on the estimated latent signal, the conditional model is a regular state-space model with time-changing parameters and can be fitted by many existing programs. Our proposed method can therefore be implemented by iteratively calling these programs. This approach is computationally efficient and yields satisfactory results in our applications and simulations.

Diggle and Zeger (1989) introduced a first-order autoregressive model, AR(1), with a feedback mechanism from the previous observation for LH. Our model differs from theirs in that (1) the AR(1) is a special case of our state-space model, (2) our model allows feedback from any previous time point(s), and (3) our model allows feedback from the latent history, which can be the latent concentration or a function of the latent signal such as the rise from the baseline. The use of the latent history has two advantages, efficiency (not subject to measurement errors) and flexibility (any latent component).

In Section 2, we introduce an autofeedback structural time series model where we add a feedback equation to a Gaussian state-space model. In Section 3, we discuss some basic ideas in estimation but leave most of the technical details to the Appendix. In Section 4, we extend our model to the case in which the structural time series model itself is conditionally Gaussian. We then illustrate our models using the LH data in Section 5. Some simulation results are presented in Section 6. Discussion and remarks are in section 7.

2. The Model

The general model is specified as an observation-level equation, a system-level equation, and a feedback equation. The first two are essentially a state-space model with its structure specified so that the structural parameters all have clear interpretations. The state vector represents various unobserved components of the system, and the system-level equation defines the transition of these components. The feedback equation models the structural parameters of interest as functions of the latent history, which can be any linear combination of the state vector; therefore, the feedback signal may or may not be the same as the observation.

The model can be written as follows.

(A) The observation-level equation is

\[ y(t) = F(t)x(t) + v(t), \quad v(t) \sim N(0, \sigma_v^2(t)). \]  

(B) The system-level equation is

\[ x(t) = H(t)x(t-1) + w(t), \quad w(t) \sim N(z, W(t)), \]  

where \( y(t) (t = 1, 2, \ldots, N) \) is the observed time series and \( x(t) \) is the state vector, which can include dummy variables to allow more general model structures. The observational matrix \( F(t) \) and the transition matrix \( H(t) \) are of corresponding dimensions. All the matrices \( F(t), H(t), \) and \( W(t) \) can contain unknown parameters. These parameters along with \( \sigma_v^2(t) \) are called structural parameters, which are denoted as \( \theta(t) \).

The feedback equation can model either all or a subset of the structural parameters \( \theta(t) \) as functions of the latent history. Without loss of generality, we assume that \( \theta(t) = \{\gamma(t), \phi(t)\} \), where \( \gamma(t) \) is a scalar and the structural parameter of interest and \( \phi(t) \) denotes the rest of the structural parameters. We denote the latent signals in the feedback system as \( z(t) = G(t)x(t) \), where \( G(t) \) defines the feedback signal as a linear combination of the latent states and can be the same as \( F(t) \), the observation matrix.

(C) The feedback equation is

\[ \gamma(t) = g(\alpha + \beta z(t - \tau)), \]  

where \( \tau \) is the lag and can take values in \( \{1, \ldots, k\} \) with \( k < N \) and \( g(\cdot) \) is a link function and can be viewed as a transformation function to get rid of the domain restriction for the structural parameter \( \gamma(t) \). It should be chosen so that the parameters \{\( \alpha, \beta \)\} have nice interpretations. The new unknown
parameters are \( \tau \) and \( \Theta = \{ \phi(t), \alpha, \beta \} \). When \( \beta = 0 \), the model is reduced to a regular state-space model because \( \alpha \) is then a reparameterization of \( \gamma(t) \).

3. Estimation

In Gaussian state-space models, the likelihood can be efficiently calculated by the Kalman filter in \( O(n) \) operations. The Kalman filter is characterized by one-step-ahead prediction and filtering steps. The likelihood is numerically maximized to produce maximum likelihood estimates for the unknown parameters. Conditioned on the estimates of the unknown parameters, a smoothing (signal extraction) algorithm is applied based on the intermediate results of the Kalman filter to produce the posterior estimate of the latent signals (the state vector). The classical Kalman filter and smoothing algorithms can be found in Anderson and Moore (1979). Some modifications of the smoothing algorithm were given by Koopman (1993) and De Jong (1989).

First, to simplify the discussion, let us assume \( \tau \) is known. When the latent history is known in the feedback model, the model is reduced to a regular state-space model with time-changing parameters. The estimation of the parameters and the state vector can be efficiently calculated by the Kalman filter and the smoothing algorithm. The latent history can then be updated from the posterior estimate of the state vector. We therefore only need to iterate back and forth between the two steps.

First, we denote \( Y = \{y(1), \ldots, y(N)\} \) and \( Z = \{z(1), \ldots, z(N)\} \). The outline of the iterative algorithm is as follows:

1. Start with an initial estimate of \( \hat{Z} \).
2. Obtain \( \hat{\Theta} \) that maximizes the likelihood \( L(\Theta | Y, \hat{Z}) \), which only involves the Kalman filter.
3. Calculate \( \hat{Z} | Y, \hat{\Theta} \) using the smoothing algorithm.
4. Iterate between (2) and (3) until convergence.

This algorithm is a first-order approximate EM algorithm, where \( Z \) is treated as missing data. The \( i \)th step of the EM algorithm is as follows:

\[ E \text{ step.} \]

\[ Q(\Theta | \Theta^i, Y) = \int l(\Theta | Y, Z)f(Z | \Theta^i, Y)dZ \]

\[ \approx \int [l(\Theta | Y, \hat{Z}) + \frac{\partial l}{\partial Z} \hat{Z} (Z - \hat{Z})] \times f(Z | \Theta^i, Y)dZ \]

\[ = l(\Theta | Y, \hat{Z}). \]

\[ M \text{ step. Obtain } \Theta^{i+1} \text{ that maximizes } Q(\Theta | \Theta^i, Y). \]

Ideally, one can use simulation techniques (e.g., De Jong and Shephard, 1995) to implement equation (4). However, because of the large dimension of \( Z \), it is usually too computationally intensive to achieve the needed accuracy for the maximization step. In the non-Gaussian state-space model setting without feedback, Durbin and Koopman (1997) and Shephard and Pitt (1997) used importance sampling to reduce the simulation errors in calculating the simulated log likelihood. Their methods rely on the ability to derive an accurate approximate Gaussian model, which is difficult in our feedback model. Although the theoretical property of the first-order approximation in equation (5) requires further study, it produced satisfactory results in our simulations.

The initial estimate \( \hat{Z} \) in step 1 is important since a good initial value can lead to rapid convergence and avoid getting trapped in a local maximum. We propose estimating the initial value \( \hat{Z} \) using the model without feedback, which is a regular state-space model. This initial estimate worked well in our simulation.

So far, we have treated the lag as known. The lag can be estimated by maximizing the profile likelihood, \( l(\tau | Y) = \sup_{\Theta} l(\tau, \Theta | Y) \). In practice, researchers usually have a rough idea about the range of the lag. A grid search can be used to obtain the maximum likelihood estimate \( \hat{\tau} \).

4. Models with Mixture of Normals

We extend the model to allow the perturbations \( v(t) \) and \( w(t) \) in equations (1) and (2) to follow a mixture of normal distributions. Because of the similarity among these models, we will not formally introduce the general structure. An example is given in the next section in the context of hormone time series with pulses, where the system perturbation is modeled by a mixture of normals characterizing the two secretory stages (with and without a pulse in the previous interval).

The structural time series model itself is called a conditionally Gaussian model (Shephard, 1994) because, conditioned on the mixture parameters, the conditional model is a Gaussian state-space model. When the underlying structural time series model is conditionally Gaussian, the resultant feedback model is still conditionally Gaussian. If we condition on the latent history and the mixture parameters simultaneously, the conditional model is a Gaussian state-space model. Because of the conditionally Gaussian structure of the feedback model, the block Gibbs sampling and normal approximation methods described in Guo and Brown (2000, in press) can be extended here.

In the current paper, we focus on a straightforward iterative estimation procedure because it is easy to implement and yields satisfactory results. The iterative algorithm is essentially the same as the one outlined above except that we use a Harrison–Stevens approximation (Harrison and Stevens, 1976) at each filtering step, which approximates the mixture of normals by a single normal with the same first two moments. Details of the algorithm can be found in Harrison and Stevens (1976) and Guo, Wang, and Brown (1999).

5. Application to the LH Data

We apply our model to the LH data described in the introduction. LH is released as pulses into the blood system, which are important in regulating biological processes (cf., Weiss et al., 1990). A pulse can be modeled as a sudden jump followed by an exponential decay toward a baseline (Kushler and Brown, 1991). Bolstad (1988) and Guo et al. (1999) used multiprocess dynamic linear models (Harrison and Stevens, 1976) to model pulsatile hormone time series. This is equivalent to modeling the pulse locations and pulse amplitudes as random effects where the posterior probabilities of the process indicators provide estimates of pulse locations and the posterior estimates of the system-level input measure the amount of secretion. This approach has the advantage of having a clear interpretation and taking into account the uncertainty in classifying a pulse.

We extend this approach to allow feedback from the history in order to test whether the rise in LH concentration provides negative feedback to the release of LH. Because pulses may be more important than the absolute level in triggering the
response, we use the rises from the baseline instead of the estimated concentrations in the feedback equation. The model can be written as follows.

(A) The observation-level equation is

\[ y(t) = x(t) + b + v(t), \quad v(t) \sim N(0, \sigma_v^2), \]  

where \( b \) is the baseline, which can be changing over time (Guo et al., 1999), \( x(t) \) is the pulsatile component, and \( v(t) \) is the measurement error.

(B) The system-level equation is

\[ x(t) = ax(t - 1) + w(t), \]  

with the system-level input being modeled by a mixture of normal distributions,

\[ w(t) \sim N\left(\mu_j, \sigma_j^2\right) \quad \text{if} \quad i(t) = j, \quad j = 0, 1, \]

where \( a \) is the decay factor, which is a function of the half-life, and \( i(t) = 1 \) indicates a pulse at time \( t \) while \( i(t) = 0 \) indicates no pulse. With \( \mu_0 \) and \( \sigma_0^2 \) set to zero, \( w(t) \) is zero with probability one when \( i(t) = 0; \mu_1 \) and \( \sigma_1^2 \) are the mean pulse amplitude and its associated variance. The system-level perturbation \( w(t) \) has the interpretation as the net input (secretion minus removal) during the time interval \([t - 1, t]\).

(C) The feedback equation is

\[ \pi(t) = g\{\alpha + \beta z(t - \tau)\}, \]  

where \( \pi(t) = p\{i(t) = 1\} \) is the prior probability of being a pulse, the function \( g\{\cdot\} = \exp(\cdot)/(1 + \exp(\cdot)) \), and \( z(t) = x(t) \) in this case. The interpretation of the feedback model is that the rise of the concentration level may trigger \( (\beta > 0) \) or shutdown \( (\beta < 0) \) the secretion.

The unknown parameters besides \( \tau \) are \( \Theta = \{b, a, \sigma_v, \mu_1, \sigma_1, \alpha, \beta\} \). The initial state \( x(0) \) is numerically given a diffuse prior \( N(0, 10,000) \). When the \( z(t - \tau) \) is not yet available in the first few values, we assign \( \pi(t) \) to its average value, which

![Figure 2](image-url)
is obtained by running the univariate model without feedback ($\beta = 0$). The estimate of \( \hat{Z}(N) \) for the model without feedback also serves as the initial values for \( \hat{Z}(N) \) to start the iteration. A grid of one to eight (in terms of observations) is used to search for the optimal lag $\hat{\tau}$ using the approximate maximum profile likelihood.

The upper panels of Figure 1 show the observed (solid lines) versus fitted concentrations (dotted lines). The lower panels show the posterior probabilities $P\{i(t) = 1 \mid Y\}$. Because the signals are very clean, the fitted and the observed curves overlay each other very well and the posterior probabilities are almost all zeros and ones. Table 1 summarizes the estimates from the six ewes. We present half-life $t = 3$ are the same for all six ewes, which indicates that there is little variability in the lags, we choose to report the point estimates only.

The feedback parameter $\beta$ is negative in all cases, and their 95% confidence intervals do not include zero. This indicates a significant feedback relationship. The confidence intervals for $\alpha$ and $\beta$ are wide and not symmetric because we use a logit link that has a flat plateau at each end, which can also be seen in Figure 3 for the simulation.

In order to show that the estimates with feedback mechanisms are more efficient than the ones without, we fitted both models to each ewe and calculated the mean square errors (MSE) of the estimated signal $z(t)$. In Table 1, MSE1 is the mean square error without feedback and MSE2 is the mean square error with feedback. In all cases, MSE2 is smaller than MSE1 and the amount of reduction is proportional to the level of feedback indicated by the magnitude of $\beta$.

### 6. Simulations

In order to illustrate the connection between regularity and negative feedback, we plot in Figure 2 four simulated series with different levels of negative feedback ($\beta = 0, -0.5, -1, -2$). The rest of parameters are $\mu = 5, b = 5, \tau = 3, \sigma_e = 1, \sigma_w = 1, \alpha = 0, \alpha = 0.84$. It can be seen that the regularity increases with the level of negative feedback. The pulses are almost periodic in the last series, where the feedback is very strong. This indicates that biological regularity may be explained by a feedback mechanism, which needs to be confirmed by controlled experiments.

In order to assess the performance of our method, simulations were carried out using the hormone model (6)–(8). We set $\mu = 10, b = 10, \tau = 3, \sigma_w = 2, \alpha = 0, \alpha = 0.84$ (half-life = 4) and varied $\sigma_e = 1, 2$ and $\beta = -0.5, -1, -2$. For each setting, we generated 1000 series with length $N = 100$. Figure 3 shows the boxplots of the estimates. All the parameters except $\sigma_w$ are estimated reasonably well. The variance parameter $\sigma_w$ is slightly underestimated.

### 7. Discussion

We have presented a new class of structural time series models with flexible and robust feedback mechanisms. These models can provide an understanding of the underlying system if they are appropriately parameterized. In the current paper, we have focused on a straightforward iterative estimation procedure because it is easy to implement, is computationally efficient, and produces satisfactory results. More sophisticated methods can be adapted from Guo and Brown (2000, in press). The probabilistic feedback relationship in our model does not imply causality. Because of potential confounding with intermediate factors, the existence of the feedback relationship needs to be confirmed by controlled experiments.

For simplicity, we have presented the model with a simple feedback structure. Extensions to allow multiple feedbacks and feedbacks from multiple previous points are immediate. These models can be estimated by the same iterative estimation procedure. However, as in any model, the time to converge will increase with the expansion of unknown parameters.
ACKNOWLEDGEMENTS
This research was supported by NIH grants R01 CA84438-01, P60 DK20572, M01 RR00042, and P30 HD18258. The revision was partially supported by the McCabe Award from the University of Pennsylvania Medical School. The authors would like to thank Dr A. Rees Midgley for allowing us to use the hormone data. We also thank the editor and the referees for their comments, which substantially improved this article.

REFERENCES