Hacettepe Journal of Mathematics and Statistics
Volume 32 (2003), 109 – 120

GIBBS SAMPLING ON A STEADY MODEL

Gül Ergün*

Received 04. 04. 2003 : Accepted 30. 09. 2003

Abstract

In this study Gibbs sampling, a widely used simulation method, is applied to the steady model, a simple variation of the dynamic linear model, and the model parameters are estimated. The estimates obtained from Gibbs sampling and the results for the standard Kalman filter are compared and are found to be close. These similarities in the results indicate the success of the stochastic simulation. In this study, a variance modulation on the steady model is also applied and Gibbs sampling is proposed to overcome analytic problems. In the variance adaptation, defined as $a\mu_t^b$ $(a, b > 0)$, estimates for the model parameters are obtained for different values of a and b.

Keywords: Bayesian approach, BUGS, Gibbs sampling, Steady model, Variance modulation

1. Introduction

Statistical inference is concerned with drawing conclusions about quantities that are not observed. Once a model is built, there are many ways to proceed with inference. The Bayesian approach considers all unknown quantities as random variables. Obtaining the posterior distribution for the unknown parameter is an important step, but not the final one. One must be able to extract meaningful information from this distribution. This is usually achieved by evaluation of point estimates such as mean, mode or interval summaries given by probability intervals. This extraction or summary can be performed analytically, that is, an exact appraisal of the situation can be made. In most cases, however, the complexity of the model prevents the analytical solution. There are many examples that fall into the category of large dimensional models, such as dynamic models, hierarchical models and random effects models [6].

Dynamic linear models provide a flexible and fairly simple tool for modelling time series data. Estimators are computed using the Kalman filter, which gives the optimal solution under some assumptions. When the models are complicated, it is no longer possible to perform exact Bayesian inference. Therefore, it is necessary to use alternative approaches for analysing time series data withinthe Bayesian framework in real life applications. A stochastic simulation method, Gibbs Sampling, is considered in this study

*Hacettepe University, Faculty of Science, Department of Statistics, Ankara, Turkey

as an alternative to an analytical solution. The steady model is taken as an example in which to implement Gibbs sampling for dynamic models. The BUGS (Bayesian Inference Using Gibbs Sampling) program is used to implement the model. Spielgelhalter, Thomas and Best [21] argued that Gibbs sampling gives generally good results, but that it can produce wrong answers. Therefore, it is also aimed to compare the results from the stochastic simulation with results from the updating equations to see possible similarity of the estimates.

There are strong practical reasons why it is necessary to carry out a variance modulation for the dynamic linear models [18]. The study is mainly interested in the class of power variance laws. It has been suggested by West and Harrison [22] that the variance modulated model is intractable, so the prior mean, $\mathrm{E}(\mu_t/y^{t-1})$ should be used as a substitute for μ_t. However, in this study, variance modulation is simply accommodated within the steady model and Gibbs sampling is applied to the variance modulated process.

The steady model is defined in Section 2. Gibbs sampling is described in Section 3. The outputs of the model from updating equation and Gibbs sampling are presented in Section 4. A conclusion is given in Section 5.

2. The Steady Model

Dynamic linear models were first introduced by Harrison and Stevens [15]. The formal definition of the model for time t is as follows [22].

(2.1) Observation Equation : $y_t = \mu_t + v_t,\ v_t \sim \mathrm{N}(0, V_t)$

(2.2) State (System) Equation : $\mu_t = \mu_{t-1} + w_t,\ w_t \sim \mathrm{N}(0, W_t)$

Initial Information : $\mu_0 \sim \mathrm{N}(m_0, C_0)$,

where y_t is the process observation at time t, μ_t the process parameter at time t (also referred to as the level), v_t and w_t are error terms and the third component here is the probabilistic representation of the forecaster's beliefs and information about the level at time $t = 0$. Also, m_0 is an estimate of the level μ_0 and C_0 a measure of uncertainty about m_0. The values of m_0 and C_0 are known, and the error sequences V_t and V_t are independent of μ_0.

Many important underlying concepts and analytical features of dynamic linear models are apparent in the simplest and most widely used case of the *steady model*, which is also called the *first order polynomial model*. This model has been used effectively in numerous applications, particularly in short term forecasting for production planning and stock control.

2.1. Updating Equations. Let us assume that the information about μ_{t-1} at time $t-1$ conditional on y^{t-1}, can be described by

(2.3) $\mu_{t-1}/y^{t-1} \sim \mathrm{N}(m_{t-1}, C_{t-1})$,

where $y^{t-1} = y_1, y_2, \ldots, y_{t-1}$ and "/" is used to indicate conditioning here. The states can be updated using Bayes' Theorem. The posterior distribution of μ_t is obtained as follows:

(2.4) $f(\mu_t/y^t) \propto f(\mu_t/y^{t-1}) f(y_t/\mu_t)$

where, from equation (2.2), the prior distribution of μ at time t is

(2.5) $\mu_t/y^{t-1} \sim \mathrm{N}(m_{t-1},\ C_{t-1} + W_t)$.

From a standard result on the multivariate normal, the steady model satisfies

(2.6) $\mu_t/y^t \sim \mathrm{N}(m_t, C_t)$, where,

$$\begin{aligned} m_t &= A_t y_t + (1 - A_t)m_{t-1} \\ &= m_t - 1 + A_t e_t \end{aligned}$$

and

$$\begin{aligned} e_t &= y_t - m_{t-1}, \\ C_t^{-1} &= (C_{t-1} + W_t)^{-1} + V_t^{-1}, \\ A_t &= C_t/V_t \quad \text{(the adaptive coefficient)}. \end{aligned}$$

The observation y_t and the state parameter are normally distributed, from which it follows that

$$f(\mu_t/y^{t-1}) \propto (f(\mu_{t-1}/yt - 1)^{k_{t-1}},$$

where $k_{t-1} = \frac{C_{t-1}}{C_{t-1}+W_t}$.

The one-step forecast distribution can be calculated from the following equation:

$$f(y_t/y^{t-1}) = \int f(y_t/\mu_t) f(\mu_t/y^{t-1})\, d\mu_t.$$

Thus,

$$y_t/y^{t-1} \sim \mathrm{N}(f_t, Q_t), \tag{2.7}$$

where $f_t = m_{t-1}$ and $Q_t = k_{t-1}C_{t-1} + V_t$.

Equations (2.5), (2.6) and (2.7) are known as the Kalman filter results.

2.2. Smoothing. The joint distribution of $y^n = (y_1, y_2, \ldots, y_n)'$ and $\mu = (\mu_1, \mu_2, \ldots, \mu_n)'$ has the density

$$f(y^n, \mu) = \prod_{t=1}^{n} f(y_t/\mu_t) \prod_{t=2}^{n} f(\mu_t/\mu_{t-1}) f(\mu_1).$$

Therefore, the full conditional density of μ_t is

$$f(\mu_t/y^n) \propto f(y_t/\mu_t) f(\mu_{t+1}/\mu_t) f(\mu_t/\mu_{t-1}) \tag{2.8}$$

Using the equations (2.1) and (2.2), it is easy to obtain the distributions on the right hand side of equation (2.8):

$$\begin{aligned} (y_t/\mu_t) &\sim \mathrm{N}(\mu_t, V_t), \\ (\mu_{t+1}/\mu_t) &\sim \mathrm{N}(\mu_t, W_{t+1}), \\ (\mu_t/\mu_{t-1}) &\sim \mathrm{N}(\mu_{t-1}, W_t). \end{aligned}$$

Thus, the full conditional density of μ_t is summarized by

$$(\mu_t/y^n) \sim \mathrm{N}(b_t, B_t), \tag{2.9}$$

where $b_t = B_t(V_t y_t + W_{t+1}^{-1}\mu_{t+1} + W_t^{-1}\mu_{t-1})$ and $B_t = (V_t + W_{t+1}^{-1} + W_t^{-1})^{-1}$, for $t = 2, 3, \ldots, n-1$.

The endpoint parameters μ_1 and μ_n also have full conditional distributions $N(b_1, B_1)$ and $N(b_n, B_n)$, where $b_1 = B_1(V_1 y_1 + W_2^{-1}\mu_2 + W_1^{-1}\mu_0)$, $B_1 = (V_1 + W_2^{-1} + W_1^{-1})^{-1}$; $b_n = B_n(V_n y_n + W_n^{-1}\mu_{n-1})$ and $B_n = (V_n + W_n^{-1})^{-1}$ [9].

The distribution of the model parameters at time t can be revised after data at times subsequent to t becomes available. A set of distributions of $f(\mu_t/y^{t+k})$, for k an integer, can be considered. When $k > 0$, they are called *smoothed* or *filtered* distributions of the parameters. When $k = 0$, we have the updated distribution, and when $k < 0$, a prior distribution. In dynamic models, the smoothed distribution $f(\mu/y^n)$ is most commonly used. It has density

$$(2.10)\qquad \begin{aligned} f(\mu/y^n) &= f(\mu_n/y^n)\prod_{t=1}^{n-1} f(\mu_t/\mu_{t+1},\dots,\mu_n,y^n) \\ &= f(\mu_n/y^n)\prod_{t=1}^{n-1} f(\mu_t/\mu_{t+1},y^t), \end{aligned}$$

where the last equality follows from the fact that μ_{t+1}, μ_t are given as independent of all quantities indexed by times larger than t. Integrating (2.10) with respect to $\mu_1,\mu_2,\dots,\mu_{t-1}$ gives

$$f(\mu_t,\dots,\mu_n/y^n) = f(\mu_n/y^n)\prod_{k=t}^{n-1} f(\mu_k/\mu_k+1,y^t) \text{ for } t=1,2,\dots,n-1 \text{ and}$$

$$(2.11)\qquad f(\mu_t,\mu_{t+1}/y^n) = f(\mu_{t+1}/y^n)f(\mu_t/\mu_{t+1},y^t) \text{ for } t=1,2,\dots,n-1.$$

Equation (2.11) provides a simple and recursive form to obtain the marginal posterior distributions of μ_t/y^n. After sequentially obtaining the updated distributions of μ_t/y^t for $t=1,2,\dots,n$, time orientation is reversed from the distribution of μ_n/y^n so as to successively obtain the distributions of μ_t/y^n for $t=n-1,n-2,\dots,1$.

3. Gibbs Sampling

The 1990's have witnessed a burst of activity in applying Bayesian methods. Most of these application have used Markov chain Monte Carlo (MCMC) methods to simulate posterior distributions. There is a large literature on Bayesian analysis with MCMC methods (for example, Brooks [1], Chib and Greenberg [4], Clayton [5], Congdon [6], Cowles and Carlin [7], Fearnhead [8] and Gamerman [9] provide an excellent introduction to MCMC). It is not common to directly obtain samples from the joint posterior distribution with complicated models. Multivariate normal or t–approximations about the modes or importance resampling techniques can improve the approximation. Unfortunately, the simulation from these approximations may not be adequate for the inferential task [12]. The idea of Markov chain simulation is to simulate a random walk in the space of $\boldsymbol{\theta}$ which converges to a stationary distribution that is the joint posterior (target) distribution, $f(\boldsymbol{\theta}/y)$.

The practical virtue of simulation methods in general, including MCMC, is that given a set of random draws $\theta^{(1)},\theta^{(2)},\theta^{(3)},\dots,\theta^{(n)}$ from the posterior distribution, one can estimate virtually all summaries of interest from the posterior distribution directly from the simulation. MCMC methods have been successful because they allow one to draw simulations from a wide range of distributions [16]. There are two basic methods of MCMC, Gibbs sampling and The Metropolis-Hastings algorithm. Gibbs sampling is the one which is used in this study. The reason for preferring this method is that the simulation program used in the application of the steady model is based on Gibbs sampling. Gibbs sampling originated in the context of image processing. Gelfand and Smith [11] were the first authors to successfully introduce the sampling scheme devised by Geman and Geman [13] to the statistical community. It has been applied in a wide array of problems. It is a MCMC scheme where the transition kernel is formed by the full conditional distributions. Assume that the distribution of interest is $f(\boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\theta_1,\theta_2,\dots,\theta n)'$. Each one of the components θ_i can be a scalar or a vector. Suppose that the full conditional distributions, $f_i(\theta_i) = f(\theta_i/\theta_{i-1}), i=1,2,\dots,n$, are available, which means that they are known and can be sampled from.

The problem to be solved is to make a draw from $f(\,.\,)$, in the case where direct generation schemes are costly, complicated or unavailable, while generations from the $f_i(\,.\,)$ are possible. Gibbs sampling provides an alternative generation scheme based on successive generations from the full conditional distributions. Three steps are required for the basic algorithm. It can be described in the following way [9]:

(1) Initialize the iteration counter of the chain to $j = 1$, and set the initial values of $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \ldots, \theta_n^{0})'$;

(2) Obtain a new value $\boldsymbol{\theta}^{(j)} = (\theta_1^{(j)}, \theta_2^{(j)}, \ldots, \theta_n^{(j)})'$ from $\theta^{(j-1)}$ through successive generation of values

$$\theta_1^j \sim f(\theta_1/\theta_2^{(j-1)}, \ldots, \theta_n^{(j-1)}),$$
$$\theta_2^j \sim f(\theta_2/\theta_1^{(j)}, \theta_3^{(j-1)}, \ldots, \theta_n^{(j-1)}),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\theta_n^j \sim f(\theta_n/\theta_1^{(j)}, \ldots, \theta_{n-1}^{(j-1)});$$

(3) Change the counter j to $j+1$ and return to step 2 until convergence is reached.

When convergence is reached, the resulting value $\theta^{(j)}$ is a draw from $f(\,.\,)$. When the number of iteration increases the chain approaches its limiting condition.

4. Application of Gibbs Sampling to a Steady Model

The steady model considered here is given by (1) and (2), and observational and system variances are taken as constant and assumed to be unknown for the application of the model. The specification of the model can be taken as a basis for blocking parameters. So, the natural choice is to form blocks $\mu_1, \mu_2, , \ldots, \mu_n$, V and W. The full conditional distributions of the μ_t are given by equation (9) in subsection 2.2. Assume that $\phi = V^{-1}$ and $\psi = W^{-1}$. It is not possible to obtain analytic expressions for the posterior densities of μ, ϕ and ψ. But it is easy to obtain the full densities of ϕ and ψ. The prior distributions of $\phi = V^{-1} \sim Gamma(\alpha, \beta)$ and $\psi = W^{-1} \sim Gamma(\gamma, \vartheta)$ are independent and the full conditional distributions of ϕ and ψ are given as follows

$$\begin{aligned} f(\phi/\mu, \psi) &\propto \prod_{t=1}^{n} f(y_t/\mu_t, \phi) f(\phi/\mu, \psi) \\ &\Rightarrow \phi/\mu, \psi \sim \Gamma(\alpha + n, \beta + \sum_t (y_t - \mu_t)^2), \\ f(\psi, \mu, \phi) &\propto \prod_{t=2}^{n} f(\mu_t/\mu_{t-1}, \psi) f(\psi/\mu, \phi) \\ &\Rightarrow \psi/\mu, \phi \sim \Gamma(\gamma + n - 1, \vartheta + \sum_{t=2}^{n} (\mu_t - \mu_{t-1})^2). \end{aligned}$$

It is seen that the parameters are conditionally conjugate. These full conditional distributions complete a cycle of the Gibbs sampler. It is clear that the μ_t should be included in the Gibbs sampler, but this may be done either through the distributions

$$(\mu_t/y^t, \phi, \psi, \mu_k(k \neq t)), (\phi/y^t, \mu_t, \psi), (\psi/y^t, \mu_t, \phi) \tag{4.12}$$

or through the distributions

$$(\mu_1, \mu_2, ..., \mu_n/y^t, \phi, \psi), (\psi/y^t, \mu_t, \psi), (\psi/y^t, \mu_t, \phi) \tag{4.13}$$

The two samplers differ in the way they simulate the μ_t's. In (4.12) the levels are simulated from their individual full conditional distribution, while in (4.13) they are sampled from their joint full conditional distribution. Because the μ_t are correlated, the blocking in (4.13) will lead to faster convergence to target distribution as is indicated by Chib and Greenberg [3]. Blocking is possible by using (2.10). Incorporating explicitly the conditional on V and W, each term in (2.10) is given by Bayes's theorem as

$$f(\mu_t/\mu_{t+1}, y^t, V, W) \propto f(\mu_{t+1}/\mu_t, y^t, V, W) f(\mu_t/y^t, V, W),$$

where $(\mu_{t+1}/\mu_t, y^t, V, W) \sim \mathrm{N}(\mu_t, W)$ from (2.11) and $(\mu_t/y^t, V, W) \sim \mathrm{N}(m_t, C_t)$ from (2.4). Thus it is easy to obtain that

$$(4.14) \quad (\mu_t/\mu_{t+1}, y^t, V, W) \sim \mathrm{N}[W^{-1} + C_t^{-1})^{-1}(W^{-1}\mu_{t+1} + C_t^{-1}m_t), (W^{-1} + C_t^{-1})^{-1}]$$

for $t = 1, \ldots, n-1$.

The Gibbs sampling algorithm for drawing samples from the full conditional of the block μ for the steady model is given by the following steps:

(1) Sample μ_n from its posterior distribution using (2.6) and set $t = n-1$.
(2) Sample μ_t from the distribution using (4.14).
(3) Decrease t to $t-1$ and return to step 2 until $t = 1$

In most cases, the proposed Gibbs algorithms for the models are difficult to implement. The *WinBUGS1.4* program is suitable for carrying out Gibbs sampling for the steady model. BUGS has been written by Spiegelhalter, Thomas and Best [20, 21] in the C-Language. The user of the program has to code his model into BUGS. The code developed for the steady model is given in Appendix A.

It can be seen that the language provides a fairly direct translation of the original description of the steady model. A couple of data set was examined in this study. But the results are given for only one data set, which is stated in the BUGS code. The data sets were simulated from the equations (2.1) and (2.2) with V=0.1 and W=0.01 precisions to have series with steady-space structures. The BUGS program was first run for 10000 iterations after 1000 burn-in for the model. The posterior means (mu[.]) and standard errors of the posterior distribution of the μ_t's; one-step ahead forecasts (y.new[.]), the standard errors of the y_t's from Gibbs sampling and the simulation error (MC error) are shown in Table 1.

The MC errors for the estimations are seen to be very small, which indicate the success of these approximations. The values of the mean and median for each node in Table 1 are also very close to one another.

The Kalman filter equations were also applied to this sample data that was simulated from equations (2.1) and (2.2), in order to analyze the difference between the above mentioned stochastic simulation results and the Kalman filter estimations that depend on analytic methods. The Kalman filter estimations were obtained by the formulae in Section 2.1 for $V = 0.1$ and $W = 0.01$, and are presented in Table 2, where, y_t is the process observation, m_t is the posterior mean of μ_t, C_t is the variance of m_t, f_t is the mean of the one-step ahead forecast distribution of y_t, and Q_t is the variance of the one-step ahead distribution. It can be seen from Tables 1 and 2 that the results are very similar. The posterior means obtained from Gibbs sampling are quite close to the values of the m_t's, and the same is true for the values of the one-step ahead forecast's means. This shows that Gibbs sampling can be used as an alternative to time series modelling, and indicates that one can trust the estimates obtained from the stochastic simulation if the model is analytically intractable.

Table 1. Estimates from the Gibbs Sampling

NODE	y_t	MEAN	S. ERROR	MC ERROR	MEDIAN	97.5% INTERVAL
mu[1]	10.70	10.74	0.6438	0.007119	10.74	9.47 – 11.99
mu[2]	10.71	10.91	0.6021	0.007536	10.91	9.68 – 12.07
mu[2]	8.56	9.41	0.6021	0.007536	9.41	9.68 – 12.07
mu[4]	9.87	9.722	0.5707	0.005537	9.731	8.59 – 10.81
mu[5]	10.50	10.091	0.5929	0.008299	9.729	8.57 – 10.88
mu[6]	7.37	8.341	0.5946	0.007174	8.121	6.96 – 9.31
mu[7]	8.47	8.108	0.5772	0.005697	8.109	6.99 – 9.23
mu[8]	6.79	7.287	0.5857	0.007277	7.279	6.15 – 8.47
mu[9]	7.80	7.514	0.569	0.006726	7.516	6.40 – 8.64
mu[10]	6.76	7.098	0.5811	0.007108	7.089	5.96 – 8.28
mu[11]	7.42	7.392	0.5767	0.007525	7.393	6.26 – 8.51
mu[12]	8.33	7.653	0.6043	0.008135	7.651	6.45 – 8.82
mu[13]	6.68	7.092	0.5831	0.006832	6.556	5.41 – 7.70
mu[14]	5.33	5.831	0.5802	0.006905	5.276	4.14 – 6.4
mu[15]	3.05	3.873	0.6292	0.009699	3.864	2.65 – 5.13
mu[16]	4.29	4.164	0.6618	0.00699	4.169	2.85 – 5.48
NODE	y_t	MEAN	S. ERROR	MC ERROR	MEDIAN	97.5% INTERVAL
y.new[1]	10.70	10.00	2.503	0.01705	10.00	7.15 – 12.60
y.new[2]	10.71	10.73	1.503	0.01605	10.72	7.75 – 13.66
y.new[3]	8.56	10.9	1.462	0.01208	10.88	8.05 – 13.79
y.new[4]	9.87	9.412	1.484	0.01454	9.426	6.49 – 12.31
y.new[5]	10.50	9.723	1.477	0.01473	9.73	6.80 – 12.73
y.new[6]	7.37	10.031	1.458	0.01504	9.697	6.91 – 12.61
y.new[7]	8.47	8.132	1.472	0.01469	8.14	5.22 – 11.02
y.new[8]	6.79	8.117	1.487	0.01368	8.111	5.16 – 11.01
y.new[9]	7.80	7.286	1.454	0.0157	7.282	4.41 – 10.16
y.new[10]	6.76	7.521	1.468	0.01462	7.511	4.67 – 10.47
y.new[11]	7.42	7.087	1.467	0.01543	7.098	4.18 – 9.92
y.new[12]	8.33	7.396	1.483	0.01349	7.416	4.44 – 10.32
y.new[13]	6.68	7.639	1.487	0.01589	7.609	4.75 – 10.63
y.new[14]	5.33	6.576	1.474	0.01313	6.562	3.66 – 9.47
y.new[15]	3.05	6.021	1.461	0.01368	5.28	2.42 – 8.17
y.new[16]	4.29	3.87	1.508	0.0173	3.868	0.93 – 6.85

Table 2. Kalman Filter Estimates

t	y_t	m_t	C_t	F_t	Q_t
1	10.70	10.48	0.67	10.00	3.00
2	10.71	11.25	0.58	10.48	2.67
3	8.56	9.59	0.57	11.25	2.64
4	9.87	9.77	0.59	9.59	2.63
5	10.50	10.25	0.61	9.77	2.63
6	7.37	8.47	0.59	10.25	2.63
7	8.47	8.47	0.60	8.47	2.63
8	6.79	7.43	0.63	8.47	2.63
9	7.80	7.66	0.62	7.43	2.63
10	6.76	7.10	0.61	7.66	2.63
11	7.42	7.30	0.61	7.10	2.63
12	7.42	7.30	0.61	7.10	2.63
13	6.68	7.16	0.61	7.93	2.63
14	5.33	6.03	0.62	7.16	2.63
15	3.05	4.19	0.63	6.03	2.63
16	4.29	4.25	0.62	4.19	2.63

4.1. Variance Modulation. In the previous application, the variance of observation equation in the state-space model has been assumed to be constant, and not to be dependent on the level, μ_t. But it usually happens for positive data that the variation may increase with level. Thus the variation needs to be formulated in terms of the observational variance. Although a theory of the law of variance is not properly completed and gets messy, it is necessary to use variance modulation in practice. In this study, variance modulation is considered for the steady model. There are several forms of the variance law, which are used in practice. The power law is used in this paper. The power law is given below:

$$\mathrm{Var}(V_t/\mu_t) = a\mu_t^b.$$

The observation equation and the state equation in (1) and (2) are rewritten as follows,

$$\begin{aligned} y_t/\mu_t &\sim \mathrm{N}(\mu_t, a\mu_t^b), \\ \mu_t/\mu_{t-1} &\sim \mathrm{N}(\mu_{t-1},\ W_t). \end{aligned}$$

The BUGS code used for obtaining estimations by Gibbs sampling with a variance modulated steady model is presented in Appendix B. The data set for the previous model has also been used in the application involving a steady model with a power variance law. For a situation in which the quantity "a", used in variance modulation, is not known, the distribution of uniform and gamma nominees have been examined. The value of b is considered to be fixed and the model estimations for $b = 1, 2$, and 3 have been found. Using the gamma and uniform distributions for an unknown quantity "a" did not lead to a significant difference. The estimations for "a" that were obtained from two different distributions were found to be similar. Table 3 has been presented for $b = 2$ and $\boldsymbol{a} \sim$ Uniform [0–2].

Table 3. Gibbs Sampling on a Variance Modulated Steady Model

NODE	y_t	MEAN	S. ERROR	MC ERROR	MEDIAN	97.5% INTERVAL
mu[1]	10.70	10.39	0.6479	0.002349	10.54	8.90 – 11.99
mu[2]	10.71	10.83	0.9880	0.005504	11.05	8.70 – 12.18
mu[2]	8.56	9.205	0.8391	0.004020	9.068	7.804 – 10.94
mu[4]	9.87	9.67	0.7323	0.000233	9.776	7.988 – 11.12
mu[5]	10.50	9.788	0.9114	0.004399	10.01	7.709 – 11.13
mu[6]	7.37	8.149	0.9307	0.004922	8.010	6.605 – 10.01
mu[7]	8.47	8.384	0.8203	0.002441	8.432	6.513 – 9.902
mu[8]	6.79	7.601		0.005568	7.458	5.92 – 9.688
mu[9]	7.80	8.182	0.8376	0.03390	8.072	6.536 – 9.936
mu[10]	6.76	10.42	2.068	0.14360	10.26	6.619 – 12.99
mu[11]	7.42	8.01	0.5767	0.007525	7.393	6.26 – 8.51
mu[12]	8.33	8.166	0.8068	0.002506	8.26	6.359 – 9.738
mu[13]	6.68	7.359	0.9604	0.009456	7.214	5.718 – 9.333
mu[14]	5.33	7.674	0.8407	0.03269	7.552	6.089 – 9.423
mu[15]	3.05	4.873	0.8816	0.03109	5.245	3.234 – 9.941
NODE	y_t	MEAN	S. ERROR	MC ERROR	MEDIAN	97.5% INTERVAL
y.new[1]	10.70	10.00	2.503	0.01705	10.00	7.15 – 12.60
y.new[2]	10.71	10.38	2.234	0.0287	10.38	5.90 – 14.82
y.new[3]	8.56	10.82	2.41	0.0553	10.77	6.106 – 15.77
y.new[4]	9.87	9.211	2.367	0.0466	9.28	4.365 – 13.78
y.new[5]	10.50	9.697	2.331	0.02964	9.71	5.05 – 14.36
y.new[6]	7.37	9.781	2.483	0.050	9.76	4.90 – 14.80
y.new[7]	8.47	8.131	2.512	0.055	8.199	3.001 – 12.88
y.new[8]	6.79	8.340	2.477	0.0306	8.371	3.228 – 13.12
y.new[9]	7.80	7.567	2.543	0.00596	7.62	2.403 – 12.42
y.new[10]	6.76	8.178	2.431	0.00366	8.234	3.221 – 12.89
y.new[11]	7.42	10.42	2.543	0.01546	10.30	4.563 – 16.62
y.new[12]	8.33	8.017	2.413	0.04366	8.04	3.026 – 12.84
y.new[13]	6.68	8.176	3.096	0.00312	8.198	3.171 – 13.17
y.new[14]	5.33	7.354	2.472	0.00529	7.383	2.233 – 12.24
y.new[15]	3.05	7.654	2.542	0.0370	7.718	2.576 – 12.46
y.new[16]	4.29	7.117	2.518	0.00392	7.654	3.356 – 11.41

As a conclusion, for b$\leq$ 2 the results from Gibbs sampling and the standard Kalman filter have been found to be quite similar. These comparisons with the Kalman filter results are presented in Table 4. When b=3, bimodality of kernel densities and deviations in the estimations might occur. In Table 4, mean absolute values (MAD), posterior means of the quantity "**a**" and related standard deviations with simulation errors for different values of b are presented for 10000 iterations after a 1000 step burn-in.

Model results for different value of b ($a \sim$ Uniform [0–2])

	Mean value of a	S. Error	MC Error	MAD	S. Error	MC Error69
b=1	0.4059	0.1862	0.0084	1.514	0.4166	0.002961
b=2	0.1937	0.1277	0.01141	1.895	0.4442	0.0027
b=3	0.1702	0.364	0.01683	2.954	0.5938	0.01832

5. Conclusion

In this study, the Gibbs sampling algorithm was applied in a straightforward way to the steady model. A comparison of Gibbs sampling and the Kalman filter was performed, and the estimates from the stochastic simulation and Kalman filter found to be very similar. Therefore, the Gibbs sampling algorithm can be recommended as an alternative to the analytic solutions in dynamic models. Stochastic simulation techniques can be applied to any kind of high dimensional model with a short computational time due to the excellent development of computers and software. There are several studies considering different aspects and sampling schema on dynamic models. Carter and Kohn [2] use Gibbs sampling on a linear model with errors that are a mixture of normals, and Kunsch [17] uses hidden Markov models. They compare their proposed algorithm to Gibbs sampling. Other alternative sampling schema for dynamic models is given by Gamerman [10]. He mainly focuses on the Metropolis-Hasting algorithm. MCMC sampling may be too slow in problems involving a large number of posterior (target) distributions in dynamic modelling. A simulation technique for tracking moving target distribution, known as particle filters, which combine importance sampling, importance resampling and MCMC sampling, suggested for nonlinear state-space models, is studied by Radford [19]. He uses embedded hidden Markov models for parameter estimations. It is pointed out by Spielgelhalter, Thomas and Best [21] that *WinBUGS* simulates each node in turn: this may cause slow convergence for the models with strongly related parameters, such as hidden-Markov and other time series structures. But the simplicity of the steady model prevents such problems.

In this study, a variance law has also been considered for the steady model. As is stated by West and Harrion [22], letting the variance depend on an unknown parameter μ_t definitely makes the posterior distributions and forecast distribution of dynamic linear models analytically intractable. Therefore, to eliminate this problem, Gibbs sampling has been proposed in this study and the posterior means and one-step ahead forecasts have been obtained for this situation. The results show that the more robust models are those with small values of a and b.

Appendix A

```
model {# STEADY MODEL WITH FIXED OBSERVATIONAL AND
STATE VARIANCES
W ~ dgamma(0.1,1)
V ~ dgamma(1,1)
# observation model
for (t in 1:N){y[t] ~ dnorm(teta[t],V);
```

```
teta[t] <- mu[t];}
# state model
for (t in 2:N) mu[t] ~ dnorm(mu[t-1],W);}
# one-step ahead forecasts
for ( t in 1:N) mu.new[t] ~ dnorm(mu[t],W);}
for (t in 2:N)  teta.new[t] <- mu.new[t-1];
y.new[t] ~ dnorm(teta.new[t],V)}
# settings for year 1
mu[1] ~ dnorm(10,0.01) }
Data
y[]
10.7 11.7 8.56 9.87 10.5 7.37 8.47 6.79
7.8 6.76 7.42 8.33 6.68 5.33 3.05 4.29 ...
Inits
list(mu=c(10,10,10,10,10,...,10,10,10))
```

Appendix B

```
model {#STEADY MODEL WITH A POWER VARIANCE MODULATION
W ~ dgamma(1,2)
a ~ dunif(0,2)
b <- 1
# observation model
for (t in 1:N) y[t] ~ dnorm(μ[t],V[t]);
μ[t] <- β[t]
V[t] <- a*pow(β[t],b)}
# state model
for (t in 2:N) β[t] ~ dnorm(β[t-1],W)}
# one-step ahead forecasts
for ( t in 1:N) β.new[t] ~ dnorm(β[t],W)}
for (t in 2:N) μ.new[t] <- β.new[t-1];
y.new[t] ~ dnorm(μ.new[t],V[t]);
residual[t] <- y.new[t] - y[t];
absdev[t] <- abs(residual[t])}
MAD <- sum(absdev[2:N])/(N)
# settings for year 1
β[1] ~ dnorm(10,10) }
Data
list(y=c(10.7,11.7,8.56,9.87,10.5,
7.37,8.47,6.79,7.8,6.76,7.42,8.33,6.68,5.33,3.05,4.29,...))
Inits
list(beta=c(10,10,10,...,10,10))
```

References

[1] Brooks, S. P. *Markov Chain Monte Carlo Method and its Application*, The Statistician **47**, 69–100, 1998.

[2] Carter, C. K. and Kohn, R. *On Gibbs Sampling for State Space Models*, Biometrika **81** (3), 541–553, 1994.

[3] Chib, S. and Greenberg, E. *Markov Chain Monte Carlo Simulation Methods in Econometrics*, Econometric Theory **12**, 409–431, 1996.

[4] Chib, S. and Greenberg, E. *Understanding the Metropolis-Hasting Algorithm*, The American Statistician **49**, 327–335, 1995.
[5] Clayton, D. G. *Generalized Linear Mixed Models, in Markov Chain Monte Carlo in Practice* (eds. W. R. Gilks, S. Richardson and D. J. Spiegelhalter), (Chapman &Hall, 1996), 275–301.
[6] Congdon, P. *Bayesian Statistical Modelling* (Wiley, London, 2001).
[7] Cowles, M. K. and Carlin, B. P. 1996, *Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review*, J. Amer. Statist. Assoc. **91**, 883–904, 1996.
[8] Fearnhead, P. *MCMC, Sufficient Statistics*, Journal of Computational and Graphical Statistics **11** (4), 848–862, 2002.
[9] Gamerman, D. *Markov Chain Monte Carlo*, (Chapman & Hall, 1997).
[10] Gamerman, D. *Sampling from the Posterior Distribution in Generalized Linear Mixed Models*, [1] Statistics and Computing **7**, 57–68, 1998.
[11] Gelfand, A. E. and Smith, A. F. M. *Sampling-Based Approaches to Calculating Marjinal Densities*, J. Amer. Stats. Assoc. **85**, 398–409, 1990.
[12] Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. *Bayesian Data Analysis*, (Chapman & Hall, London, 1998).
[13] Geman, S. and Geman, D. *Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images*, IEEE Transaction on Pattern Analysis and Machine Intelligence **6**, 721–741, 1984.
[14] Gilks, W. R. and Berzuini, C. *Following a Moving Target – Monte Carlo Inference for Dynamic Bayesian Models*, Journal of Royal Statistical Society, Series B. **63**, 1–20, 2001.
[15] Harrison, P. J. and Stevens , C. F. *Bayesian Forecasting (with discussion)*, Journal of Royal Statistical Society, Series B **38**, 205–247, 1976.
[16] Kass, R. E. *Markov Chain Monte Carlo in Practice: A Roundtable Discussion.* Panel Discussion at Joint Statistical Meeting, 1997.
[17] Kunsch, H. R. 2001, *State-Space and Hidden Markov Models*, in Complex Stochastic Systems, O. E. Barndorff-Nielsen, D. R. Cox and C. Kluppelberg, Eds., Number **87** in Monographs on Statistics and Applied Probability, (Chapman and Hall, 2001), 1009–173.
[18] Pole, A., West, M. and Harrison, J. *Applied Bayesian Forecasting and Time Series Analysis*, (Chapman & Hall, New York, 1994).
[19] Radford, M. N. 2003, *Markov Chain Sampling for Nonlinear State-Space Models Using Embedded Hidden Markov Models*, Technical Report, No:0304, 9 Pages; Dept. of Statistics and Dept. of Computer Science, (Universtiy of Toronto, 2003)
[20] Spiegelhalter, Thomas and Best. *WinBUGS User Manual, Version 1.3*, `www.mrc-bsu.cam.ac.uk./bugs`, 1995.
[21] Spielgelhalter, Thomas and Best. *WinBUGS User Manual, Version 1.4*, `www.mrc-bsu.cam.ac.uk./bugs`, 2002.
[22] West, M. and Harrison, P. J., 1997, *Bayesian Forecasting and Dynamic Models*, (Springer Verlag, New York, 1997).