Multi-move sampler for estimating non-Gaussian time series models: Comments on Shephard and Pitt (1997)

by Toshiaki Watanabe and Yasuhiro Omori*

Department of Economics, Tokyo Metropolitan University, Tokyo, Japan

twatanab@bcomp.metro-u.ac.jp omori@bcomp.metro-u.ac.jp

Abstract

This note points out a problem in the multi-move sampler proposed by Shephard and Pitt (1997) and corrects their method. The performance of the original Shephard and Pitt (1997) method and the correct method is examined by estimating stochastic volatility models using simulated data. It is found that the original method yields an estimation bias which increases with the number of blocks while the correct one performs well irrespective of the number of blocks.

Keywords: Markov chain Monte Carlo, Multi-move sampler, Simulation smoother, State variable, Stochastic volatility model.

*The authors thank Roman Liesenfeld and Jean-Francois Richard for sending comments on Watanabe (2000). Their comments stimulated us to write this note. Thanks are also due to Manabu Asai for his useful comments. Remaining errors are ours alone.

1 Introduction

Shephard and Pitt (1997) have proposed a method so called "multi-move sampler" for sampling latent state variables in non-linear and non-Gaussian state space models from their posterior density given the parameters. As shown in Shephard and Pitt (1997), a single move sampler which generates a single state variable at a time would produce a highly correlated sample sequence when state variables are highly autocorrelated. To obtain independent samples, we need to repeat the sampling a huge number of times. To reduce this inefficiency, Shephard and Pitt (1997) divide the state variables into several blocks and sample each block at a time. Another feature of their method is to sample state errors instead of sampling state variables directly. For this sampling, they use the Metropolis-Hastings acceptance-rejection algorithm proposed by Tierney (1994). To generate candidates, they first approximate a true posterior density for a block of state errors given the parameters and the other blocks by a linear Gaussian system. Then, using a simulation smoother (de Jong and Shephard, 1995), candidates for state errors are sampled.

Since their method is highly efficient and the speed of convergence is remarkably fast, it has been applied by some researchers such as Pitt and Shephard (1999) and Watanabe (2000). There is, however, a problem in their method. Specifically, they omit a term in the posterior density of a block of state errors given the parameters and the other blocks. This note corrects their method by taking account of this omitted term. To examine how this correction is important, we estimate stochastic volatility models using simulated data by the both original Shephard and Pitt (1997) and correct methods. We find that the original method yields an estimation bias which increases with the number of blocks while the correct one performs well irrespective of the number of blocks. We further find that this bias is severe when the latent variables are highly autocorrelated. In such a case, the original method yields a significant bias even if the number of blocks is small.

This note is organized as follows. Section 2 explains a problem in the Shephard and Pitt (1997) multi-move sampler and corrects their method. Section 3 examines the performance of the original Shephard and Pitt (1997) method and the correct method by estimating stochastic volatility models using simulated data. Section 4 gives conclusion.

2 A Correct Multi-Move Sampler

Suppose that the univariate time series y_t are distributed, conditionally on a univariate θ_t according to $f(y_t|\theta_t)$ $(t = 1, \ldots, n)$. Following Shephard and Pitt (1997), we consider the exponential family written as

$$\begin{aligned} \log f(y_t|\theta_t) &= y_t\theta_t - b(\theta_t) + c(y_t), \qquad (1) \\ h(\theta_t) &= d_t = z_t\alpha_t + x_t\beta \\ \alpha_{t+1} &= T_t\alpha_t + H_t u_t, \quad u_t \sim \mathrm{NID}(0, I), \qquad (2) \\ & \quad \alpha_1 \sim N(a_{1|0}, P_{1|0}), \end{aligned}$$

where α_t are state variables, x_t and W_t are regressors, and $b(\theta_t)$, $c(y_t)$, and $h(\theta_t)$ are known functions, The $\log f(y_t|\theta_t)$ and $h(\theta_t)$ are assumed to be continuously twice differentiable with respect to θ throughout.

Consider the density for $(u_{t-1}, \ldots, u_{t+k-1})$ conditional on $(\alpha_{t-1}, \alpha_{t+k+1}, y_t, \ldots, y_{t+k})$. Following Shephard and Pitt (1997), we write $l(\theta_t)$ to denote $\log f(y_t|\theta_t)$. Then, the logarithm of its density is given by

$$\begin{aligned} &\log f(u_{t-1}, \ldots, u_{t+k-1}|\alpha_{t-1}, \alpha_{t+k+1}, y_t, \ldots, y_{t+k}) \\ &= \text{constant} - \frac{1}{2}\sum_{s=t-1}^{t+k-1} u_s^2 + \sum_{s=t}^{t+k} l(\theta_s) \\ &\quad -\frac{1}{2}(\alpha_{t+k+1} - T_{t+k}\alpha_{t+k})'\Omega_{t+k}^{-1}(\alpha_{t+k+1} - T_{t+k}\alpha_{t+k}), \qquad (3) \end{aligned}$$

where $\Omega_{t+k} = H_{t+k}H'_{t+k}$. The problem with Shephard and Pitt (1997) is that the last term of the right-hand-side is omitted.

Applying a Taylor expansion around $\theta_s = \hat{\theta}_s$ to equation (3) without omitting the last term yields:

$$\begin{aligned} &\log f(u_{t-1}, \ldots, u_{t+k-1}|\alpha_{t-1}, \alpha_{t+k+1}, y_t, \ldots, y_{t+k}) \\ &\approx \text{constant} - \frac{1}{2}\sum_{s=t-1}^{t+k-1} u_s^2 + \sum_{s=t}^{t+k} l(\hat{\theta}_s) + \sum_{s=t-1}^{t+k-1} z_s(\alpha_s - \hat{\alpha}_s)l'(\hat{\theta}_s) + \frac{1}{2}\{z_s(\alpha_s - \hat{\alpha}_s)\}^2 l''(\hat{\theta}_s) \\ &\quad -\frac{1}{2}(\alpha_{t+k+1} - T_{t+k}\hat{\alpha}_{t+k})'\Omega_{t+k}^{-1}(\alpha_{t+k+1} - T_{t+k}\hat{\alpha}_{t+k}) \end{aligned}$$

$$+(\alpha_{t+k+1} - T_{t+k}\hat{\alpha}_{t+k})'\Omega_{t+k}^{-1}T_{t+k}(\alpha_{t+k} - \hat{\alpha}_{t+k})$$
$$-\frac{1}{2}(\alpha_{t+k} - \hat{\alpha}_{t+k})'T_{t+k}'\Omega_{t+k}^{-1}T_{t+k}(\alpha_{t+k} - \hat{\alpha}_{t+k}), \tag{4}$$

where $l'(\theta_s) = \partial l(\theta_s)/\partial d_s$, and $l''(\theta_s) = \partial^2 l(\theta_s)/\partial d_s^2$.

We write $\log g$ to denote the right hand side of equation (4). Then, the normalized version of g is a $(k+1)$ -dimensional normal density. To sample from this density, we must correct the definition of artificial variables $\hat{y}_s$ in Shephard and Pitt (1997) as follows. For $s = t, \ldots, t+k-1$ and $s = n$,

$$\hat{y}_s = z_s\hat{\alpha}_s + v_s l'(\hat{\theta}_s), \tag{5}$$

where

$$v_s = -1/l''(\hat{\theta}_s), \tag{6}$$

and for $s = t+k < n$,

$$\hat{y}_s = v_s\left[z_s'\left\{l'(\hat{\theta}_s) - l''(\hat{\theta}_s)z_s\hat{\alpha}_s\right\} + T_s'\Omega_s^{-1}\alpha_{s+1}\right], \tag{7}$$

where

$$v_s = \left[T_s'\Omega_s^{-1}T_s - l''(\hat{\theta}_s)z_s'z_s\right]^{-1}. \tag{8}$$

For the last block, i.e. t+k=n, α_{t+k+1} is removed from the condition of the left-hand-side of equations (3) and (4). Hence, $\hat{y}_n$ is defined by equation (5) instead of equation (7). Shephard and Pitt (1997) define $\hat{y}_s$ for all s by equation (5) because they omit the last term in equation (3).

Then, consider the following linear Gaussian model .

$$\hat{y}_s = \begin{cases} z_s\alpha_s + \epsilon_s, & \text{for } s = t, \ldots, t+k-1 \text{ and } s = n, \\ \alpha_s + \epsilon_s, & \text{for } s = t+k < n, \end{cases}$$

$$\alpha_{s+1} = T_s\alpha_s + H_s u_s, \text{ for all } s,$$

where $\epsilon_s \sim NID(0, v_s)$ and $u_s \sim NID(0, I)$. Notice that Sampling from g is the same as sampling $(u_{t-1}, \ldots, u_{t+k-1})$ given α_{t-1} and $(\hat{y}_t, \ldots, \hat{y}_{t+k})$ in the above model, which is possible by using the de Jong and Shephard (1995) simulation smoother.

For the other procedures, we may simply follow Shephard and Pitt (1997). The Metropolis-Hastings acceptance-rejection algorithm with g as a proposal density can be used to sample from the true density f given by equation (3). The values for $\hat{\theta}_s$'s around which the Taylor expansion is conducted are selected as the mode of the conditional density for θ_s 's, which can be found by using the moment smoother. We divide $\{\alpha_1, \ldots, \alpha_n\}$ into $K+1$ blocks, $(\alpha_{k_{i-1}+1}, \ldots, \alpha_{k_i})'$, $i = 1, \ldots, K+1$ with $k_0 = 0$, $k_{K+1} = n$. K knots, $(k_1, \ldots, k_K)$, are selected randomly such that

$$k_i = \text{int}[n \times (i + U_i)/(K+2)], \quad i = 1, \ldots, K,$$

where U_i 's are independent uniform random variables on $(0, 1)$.

3 Illustration using Stochastic Volatility Model

To examine the performance of the original Shephard and Pitt (1997) method and the correct one explained in the previous section, we analyze the following stochastic volatility model.

$$y_t = \epsilon_t \beta \exp(\alpha_t/2), \quad \epsilon_t \sim \text{NID}(0, 1), \tag{9}$$

$$\alpha_{t+1} = \phi\alpha_t + \eta_t, \quad \eta_t \sim \text{NID}(0, \sigma_\eta^2), \tag{10}$$

$$\alpha_1 \sim N(0, \sigma_\eta^2/(1-\phi^2)).$$

for $t = 1, \ldots, 1000$.

We simulate $\{y_1, \ldots, y_{1000}\}$ from the above model and, using this simulated data, estimate parameters $(\beta, \sigma_\eta, \phi)$ as well as latent variables $\{\alpha_1, \ldots, \alpha_{1000}\}$. Following Shephard and Pitt (1997), we use two parameter settings. One is $(\beta, \phi, \sigma_\eta^2) = (1, 0.9, 0.1)$, and the other is $(\beta, \phi, \sigma_\eta^2) = (1, 0.99, 0.01)$. We call the data simulated under the former parameter setting as "weekly data" and the data simulated under the latter parameter setting as "daily data" because the former is typical of weekly financial returns data and the latter is typical of daily financial returns data. For the parameters, we adopt the same prior as used in Shephard and Pitt (1997). For the number of knots, we use $K = 10, 20, 50, 100$. We can expect that the original Shephard and Pitt (1997) multi-move sampler performs worse with the number of knots because it omits the last term in equation (3)

every block except the last block in which the last term in equation (3) is not present. Notice that the original Shephard and Pitt (1997) multi-move sampler becomes correct when $K = 0$.

For the above stochastic volatility model, the correct definition of artificial variables is as follows. For $s = t, \ldots, t+k-1$ and $s = n$,

$$\hat{y}_s = \hat{\alpha}_s + v_s l'(\hat{\theta}_s),$$

where

$$v_s = \frac{2\beta^2}{y_s^2}\exp(\hat{\alpha}_s) \quad \text{and} \quad l'(\hat{\theta}_s) = \frac{1}{2}\left\{\frac{y_s^2}{\beta^2}\exp(-\hat{\alpha}_s) - 1\right\},$$

and for $s = t+k < n$,

$$\hat{y}_s = v_s\left[\left\{l'(\hat{\theta}_s) - l''(\hat{\theta}_s)\hat{\alpha}_s\right\} + \phi\sigma_\eta^{-2}\alpha_{s+1}\right],$$

where

$$v_s = \left(\phi^2\sigma_\eta^{-2} - l''(\hat{\theta}_s)\right)^{-1}.$$

The initial 1000 iterations are discarded to avoid the so-called "burn-in" period and the following $n = 10000$ iterations are recorded. Table 1 summarizes estimation results for the parameters and the middle state α_{500}. This table reports the posterior means, the 95% intervals, the standard errors of the posterior means, and the p-values for convergence diagnostic (CD) statistics. The posterior means are computed by averaging the simulated draws. The 95% intervals are calculated using the 2.5th and 97.5th percentiles of the simulated draws. The standard errors of the posterior means are computed using a Parzen window with a bandwidth of 1000. Geweke (1992) suggests assessing the convergence of the MCMC by comparing values early in the sequence with those late in the sequence. Let $X^{(i)}$ be the i th draw of a parameter in the recorded 10000 draws, and let $\bar{X}_A = \frac{1}{n_A}\sum_{i=1}^{n_A} X^{(i)}$ and $\bar{X}_B = \frac{1}{n_B}\sum_{i=10001-n_B}^{10000} X^{(i)}$. Using these values, Geweke (1992) proposes the following statistic called convergence diagnostics (CD).

$$\text{CD} = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\hat{\sigma}_A^2/n_A + \hat{\sigma}_B^2/n_B}}, \tag{11}$$

where $\sqrt{\hat{\sigma}_A^2/n_A}$ and $\sqrt{\hat{\sigma}_B^2/n_B}$ are standard errors of $\bar{X}_A$ and $\bar{X}_B$. If the sequence of $X^{(i)}$ is stationary, it converges in distribution to the standard normal. We set $n_A = 1000$ and $n_B = 5000$ and compute $\hat{\sigma}_A^2$ and $\hat{\sigma}_B^2$ using Parzen windows with bandwidth of 500.

According to the p-values for CD statistics, the null hypothesis that the sequence of 10000 draws is stationary is accepted at any standard level for all parameters. On one hand, the results based on the correct multi-move sampler are stable irrespective of the number of knots. The 95% intervals for all parameters and α_{500} include the true values for all K. On the other hand, as expected, the original multi-move sampler performs worse with the number of knots. Specifically, the estimates for ϕ decrease and the estimates for σ_η increase with the number of knots. For the weekly data, the 95% intervals for ϕ are below the true value and those for σ_η are above the true value when $K = 50$ and $K = 100$. For the daily data in which shocks to state variables α_t are highly persistent ($\phi = 0.99$), the performance of the original multi-move sampler is much worse. For all K, the 95% intervals for ϕ and σ_η do not include the true values.

Figures 1 and 2 plot the estimates for α_t along with the true values. The estimates based on the original multi-move sampler are more volatile than those on the correct one. The estimates based on the original one become more volatile with the number of knots while those on the correct one are stable. Table 2 shows the root mean square error (RMSE) of the estimates for α_t, i.e.

$$RMSE = \sqrt{\frac{1}{1000}\sum_{t=1}^{1000}(a_t - \alpha_t)^2}$$

where a_t is the posterior mean of α_t. The RMSEs based on the original multi-move sampler increase with the number of knots while those on the correct one are stable.

4 Conclusion

This note pointed out a problem in the multi-move sampler proposed by Shephard and Pitt (1997) and corrected their method. The performance of the original Shephard and Pitt (1997) sampler and our correct sampler was examined by estimating stochastic volatility models using simulated data. The results demonstrated that our correct sampler performs well irrespective of the number of knots while the original sampler performs worse with the number of knots. The correction explained in this note is important especially when shocks to state variables are highly persistent because, in such a case, the original sampler may yield a significant bias even though the number of knots is small.

References

de Jong, P. and Shephard, N. (1995) "The simulation smoother for time series models," *Biometrika*, **82**, 339-350.

Geweke, J. (1992), "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," in J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (eds.), *Bayesian Statistics*, 4, New York: Oxford University Press, pp169–188.

Pitt, M. K. and Shephard, N. (1999), "Time-Varying Covariances: A Factor Stochastic Volatility Approach," in J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (eds.), *Bayesian Statistics*, 6, New York: Oxford University Press, pp547–570.

Shephard, N. and Pitt, M. K. (1997) "Likelihood analysis of non-Gaussian measurement time series," *Biometrika*, **84**, 653–668.

Tierney, L. (1994), "Markov Chains for Exploring Posterior Distributions," *Annals of Statistics*, **22**, 1701–1728.

Watanabe, T. (2000), "Bayesian Analysis of Dynamic Bivariate Mixture Models: Can They Explain the Behavior of Returns and Trading Volume?" *Journal of Business & Economic Statistics,* **18,** 199–210.

Table 1. Estimation Results for Parameters and α_{500}

(a) Weekly Data ($\beta = 1$, $\phi = 0.9$, $\sigma_\eta^2 = 0.1$, $\alpha_{500} = -0.6819$)

	Correct				Original			
	Mean	95% Interval	S.E.	p-value	Mean	95% Interval	S.E.	p-value
K=10								
β	0.9705	[0.8533, 1.1059]	0.0032	0.86	0.9527	[0.8619, 1.0469]	0.0015	0.76
ϕ	0.9220	[0.8690, 0.9628]	0.0016	0.58	0.8658	[0.7831, 0.9249]	0.0028	0.85
σ_η	0.2489	[0.1816, 0.3284]	0.0034	0.62	0.3549	[0.2683, 0.4676]	0.0049	0.62
α_{500}	-0.3676	[-1.1838, 0.5066]	0.0082	0.01	-0.2570	[-1.2059, 0.7816]	0.0087	0.61
K=20								
β	0.9656	[0.8466, 1.0802]	0.0019	0.26	0.9431	[0.8563, 1.0287]	0.0013	0.87
ϕ	0.9225	[0.8693, 0.9641]	0.0018	0.35	0.8304	[0.7265, 0.9038]	0.0030	0.82
σ_η	0.2480	[0.1769, 0.3325]	0.0041	0.23	0.4116	[0.3073, 0.5346]	0.0051	0.84
α_{500}	-0.3569	[-1.1663, 0.5121]	0.0056	0.45	-0.2220	[-1.1965, 0.8456]	0.0074	0.57
K=50								
β	0.9685	[0.8480, 1.0910]	0.0040	0.86	0.9408	[0.8689, 1.0149]	0.0009	0.78
ϕ	0.9268	[0.8748, 0.9645]	0.0013	0.82	0.7386	[0.5965, 0.8419]	0.0039	0.85
σ_η	0.2391	[0.1793, 0.3252]	0.0031	0.97	0.5147	[0.4054, 0.6462]	0.0048	0.69
α_{500}	-0.3779	[-1.1665, 0.4605]	0.0111	0.39	-0.1428	[-1.2212, 1.0463]	0.0076	0.95
K=100								
β	0.9656	[0.8420, 1.0854]	0.0044	0.23	0.9404	[0.8713, 1.0098]	0.0009	0.77
ϕ	0.9241	[0.8713, 0.9631]	0.0016	0.68	0.6382	[0.4709, 0.7670]	0.0039	0.84
σ_η	0.2429	[0.1793, 0.3209]	0.0033	0.60	0.5980	[0.4792, 0.7360]	0.0038	0.58
α_{500}	-0.3600	[-1.1576, 0.4954]	0.0129	0.17	-0.0575	[-1.1970, 1.2111]	0.0070	0.69

NOTE: The first 1000 draws are discarded and then the next 10000 are used for calculating the posterior means, the 95% intervals, the standard errors of the posterior means, and the p-values for convergence diagnostic (CD) statistics proposed by Geweke (1992). The posterior means are computed by averaging the simulated draws. The 95% intervals are calculated using the 2.5th and 97.5th percentiles of the simulated draws. The standard errors of the posterior means are computed using a Parzen window with a bandwidth of 1000. The CD is computed using equation (11), where we set $n_A = 1000$ and $n_B = 5000$ and compute $\hat{\sigma}_A^2$ and $\hat{\sigma}_B^2$ using a Parzen window with bandwidth of 500.

(b) Daily Data ($\beta = 1$, $\phi = 0.99$, $\sigma_\eta^2 = 0.01$, $\alpha_{500} = -0.5463$)

	Correct				Original			
	Mean	95% Interval	S.E.	p-value	Mean	95% Interval	S.E.	p-value
K=10								
β	1.0294	[0.8260, 1.3085]	0.0101	0.04	1.0158	[0.8641, 1.1524]	0.0037	0.05
ϕ	0.9839	[0.9665, 0.9962]	0.0005	0.31	0.9452	[0.8981, 0.9768]	0.0015	0.17
σ_η	0.1050	[0.0760, 0.1420]	0.0019	0.84	0.2119	[0.1552, 0.2799]	0.0031	0.11
α_{500}	-0.2275	[-0.9160, 0.4489]	0.0216	0.02	-0.0169	[-0.7786, 0.7986]	0.0082	0.18
K=20								
β	1.0094	[0.7642, 1.2193]	0.0076	0.55	1.0160	[0.9030, 1.1269]	0.0015	0.81
ϕ	0.9833	[0.9655, 0.9958]	0.0004	0.92	0.9081	[0.8405, 0.9563]	0.0019	0.70
σ_η	0.1052	[0.0774, 0.1405]	0.0015	0.59	0.2795	[0.2055, 0.3762]	0.0036	0.69
α_{500}	-0.1872	[-0.8158, 0.4990]	0.0168	0.40	0.0292	[-0.8368, 0.9506]	0.0058	0.07
K=50								
β	0.9944	[0.7372, 1.2396]	0.0180	0.98	1.0198	[0.9365, 1.1044]	0.0011	0.52
ϕ	0.9841	[0.9660, 0.9963]	0.0006	0.77	0.8185	[0.7107, 0.8969]	0.0024	0.39
σ_η	0.1053	[0.0796, 0.1452]	0.0016	0.49	0.3912	[0.3106, 0.4870]	0.0032	0.45
α_{500}	-0.1548	[-0.8208, 0.5633]	0.0385	0.97	0.0121	[-1.0141, 1.1156]	0.0079	0.33
K=100								
β	1.0019	[0.7941, 1.2176]	0.0217	0.01	1.0276	[0.9549, 1.0982]	0.0008	0.62
ϕ	0.9846	[0.9677, 0.9960]	0.0005	0.65	0.7082	[0.5450, 0.8272]	.0034	0.39
σ_η	0.1036	[0.0783, 0.1346]	0.0016	0.73	0.4656	[0.3681, 0.5933]	0.0037	0.27
α_{500}	-0.1826	[-0.8219, 0.5085]	0.0439	0.03	-0.0171	[-1.0927, 1.1470]	0.0045	0.37

NOTE: The first 1000 draws are discarded and then the next 10000 are used for calculating the posterior means, the 95% intervals, the standard errors of the posterior means, and the p-values for convergence diagnostic (CD) statistics proposed by Geweke (1992). The posterior means are computed by averaging the simulated draws. The 95% intervals are calculated using the 2.5th and 97.5th percentiles of the simulated draws. The standard errors of the posterior means are computed using a Parzen window with a bandwidth of 1000. The CD is computed using equation (11), where we set $n_A = 1000$ and $n_B = 5000$ and compute $\hat{\sigma}_A^2$ and $\hat{\sigma}_B^2$ using a Parzen window with bandwidth of 500.

Table 2. Root Mean Square Error of State Variables Estimates

	Weekly		Daily	
K	Correct	Original	Correct	Original
10	0.4727	0.4724	0.2781	0.2871
20	0.4735	0.4786	0.2745	0.3049
50	0.4750	0.4967	0.2758	0.3404
100	0.4748	0.5207	0.2739	0.3812

NOTE: The table shows the root mean square error

$$RMSE = \sqrt{\frac{1}{1000}\sum_{t=1}^{1000}(a_t - \alpha_t)^2}$$

where a_t is the posterior mean of α_t.

Figure 1: Estimates of State Variable (Weekly)
(a) K = 10

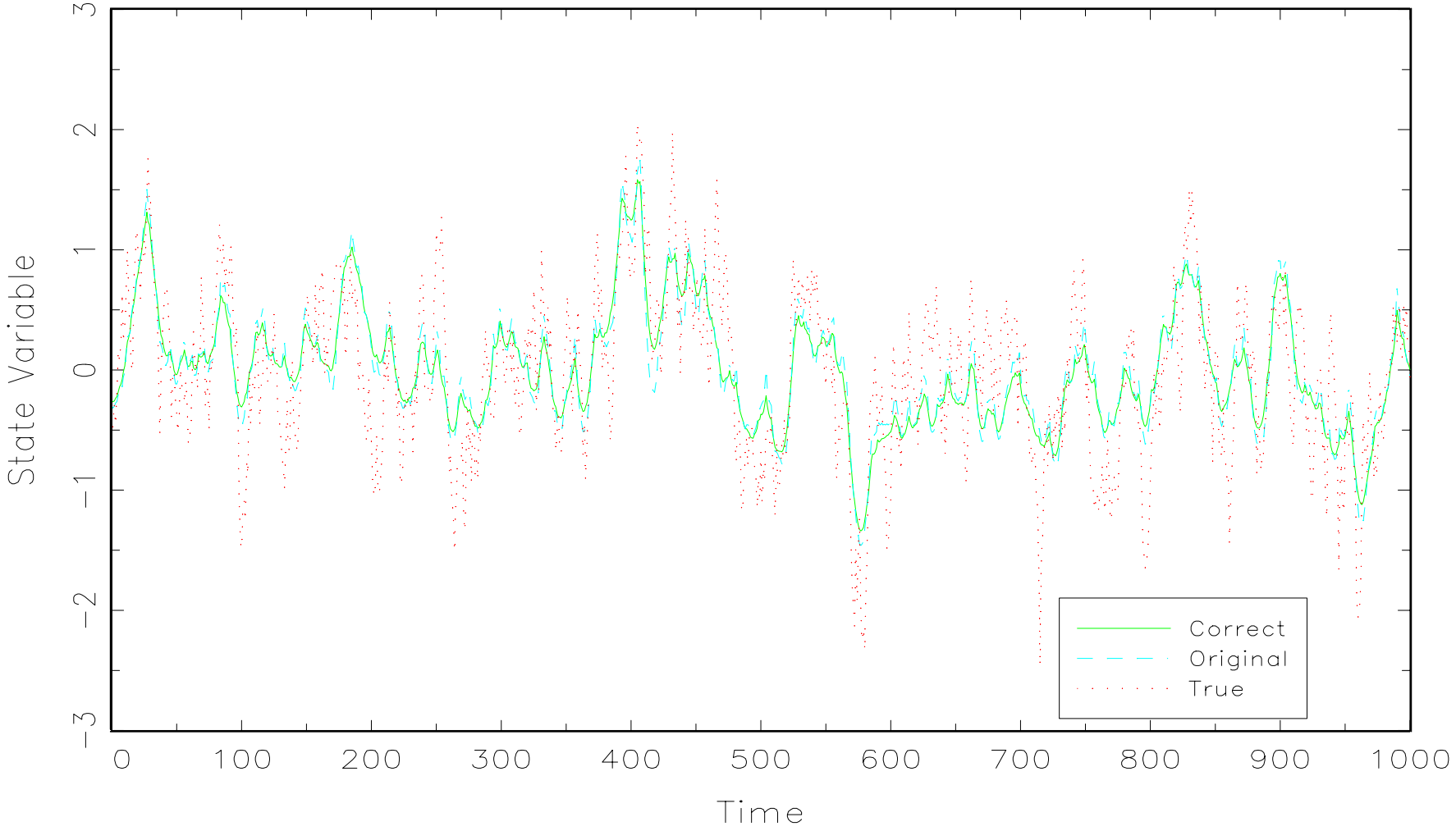

Figure 1: Estimates of State Variable (Weekly)
(b) K = 20

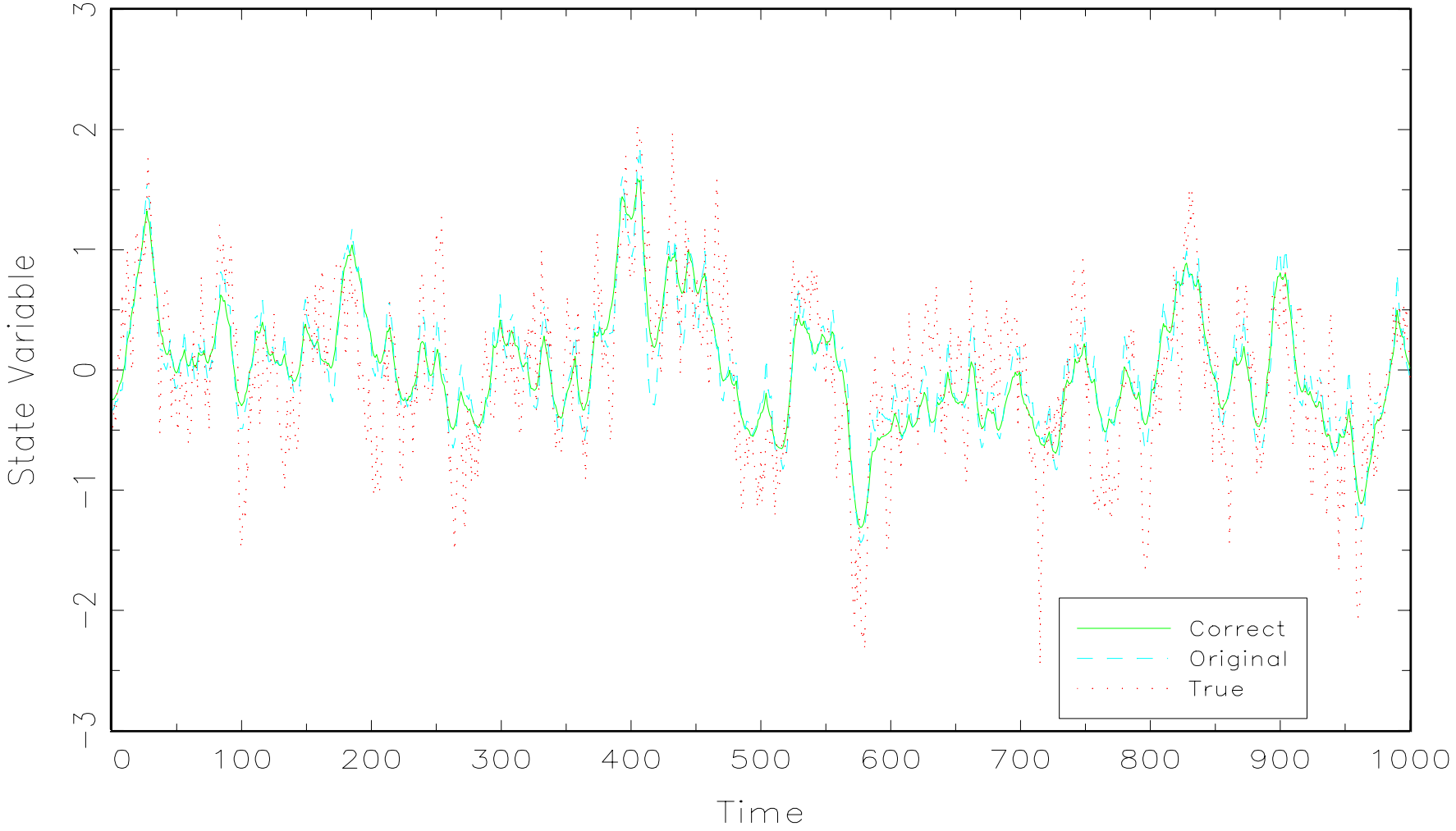

Figure 1: Estimates of State Variable (Weekly)
(c) K = 50

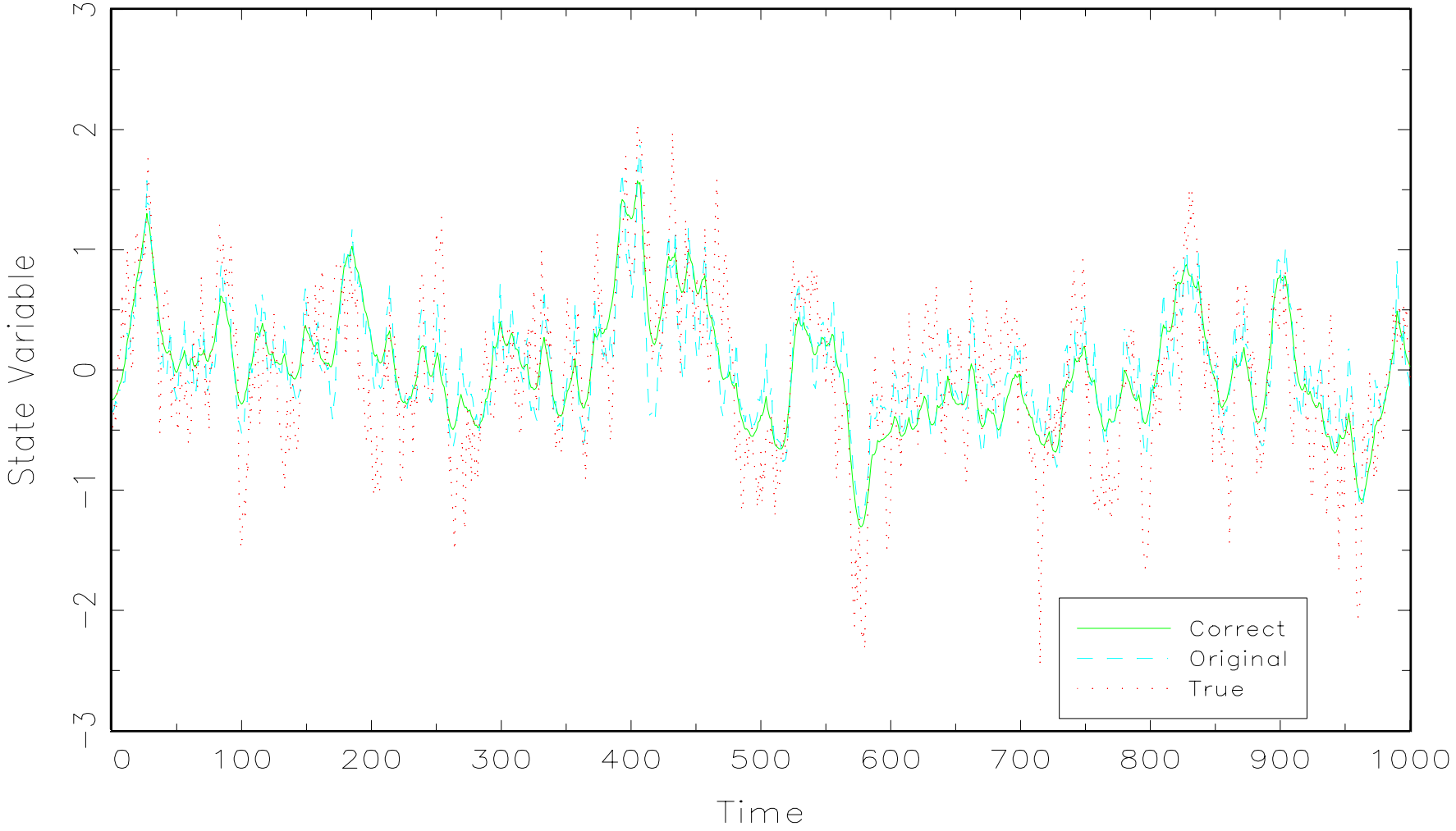

Figure 1: Estimates of State Variable (Weekly)
(d) K = 100

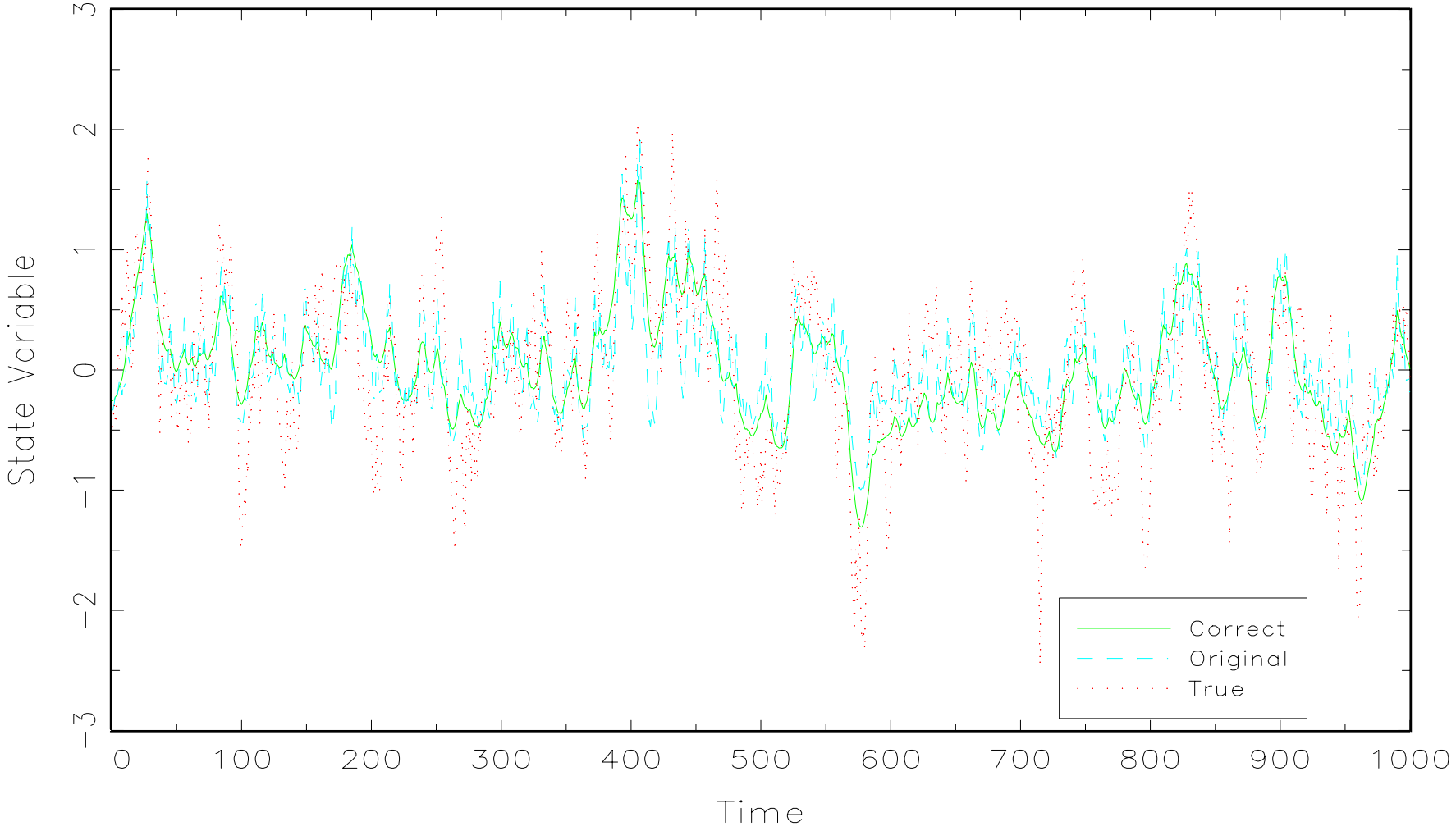

Figure 2: Estimates of State Variable (Daily)
(a) K = 10

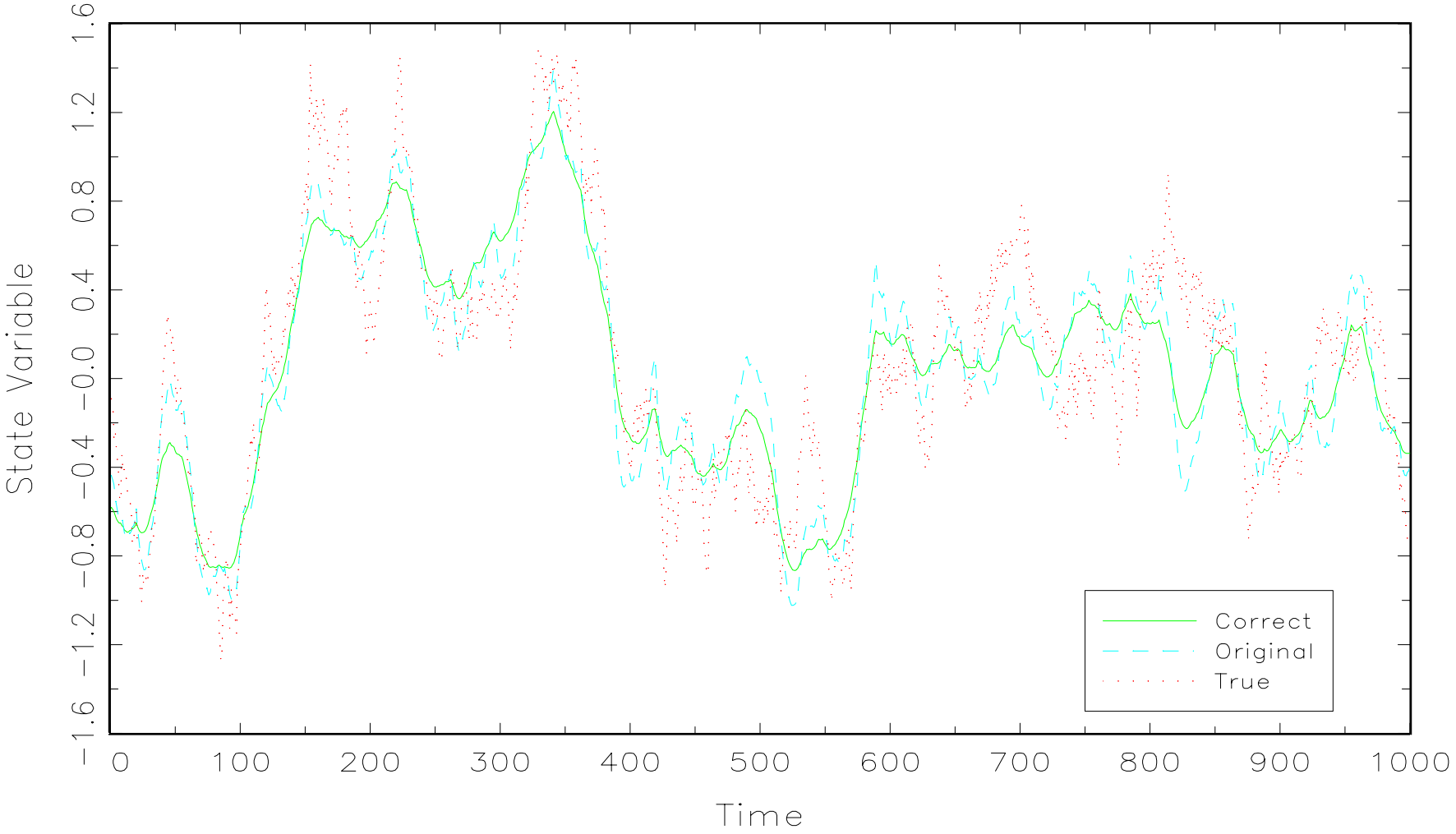

Figure 2: Estimates of State Variable (Daily)
(b) K = 20

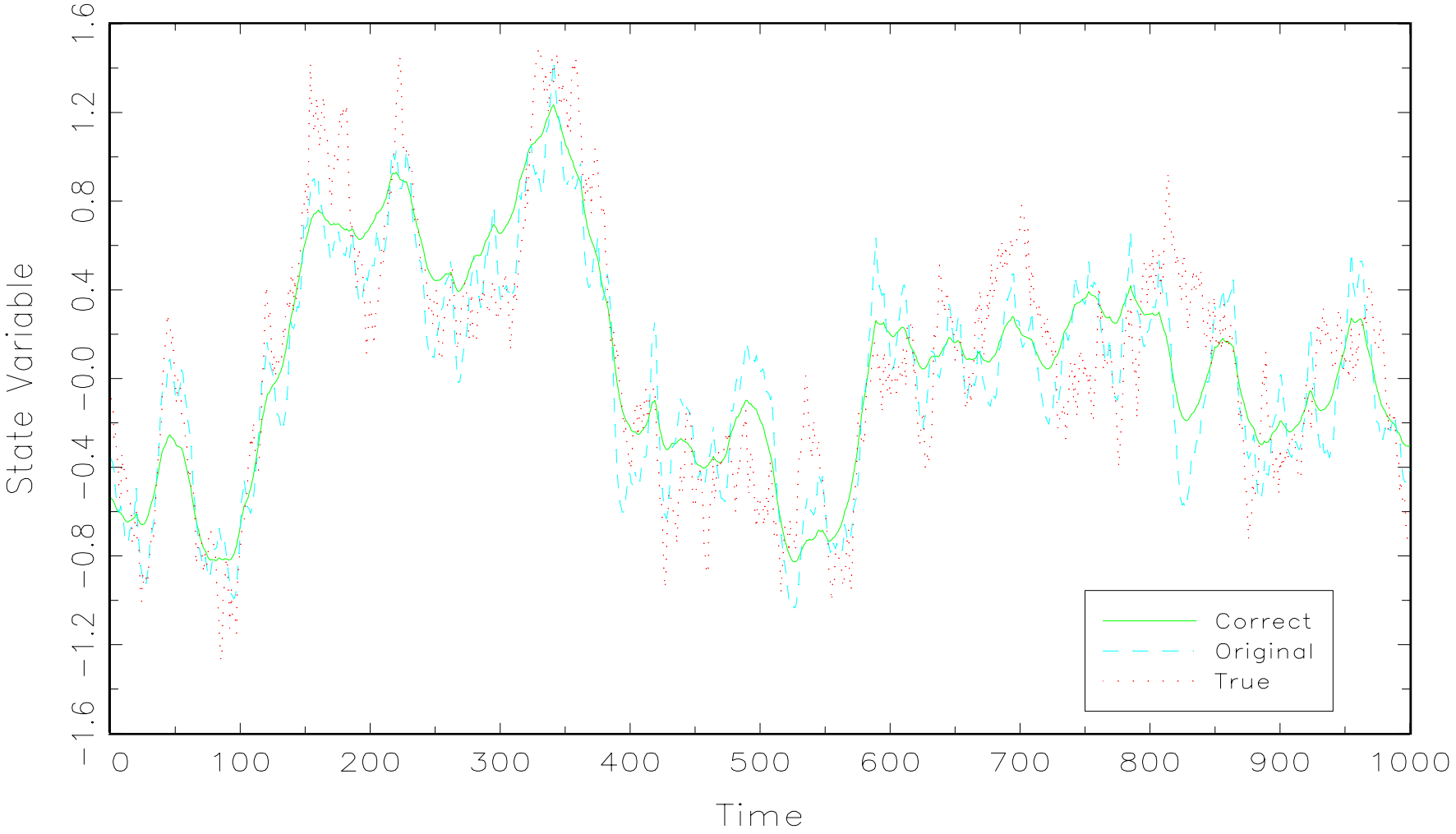

Figure 2: Estimates of State Variable (Daily)
(c) K = 50

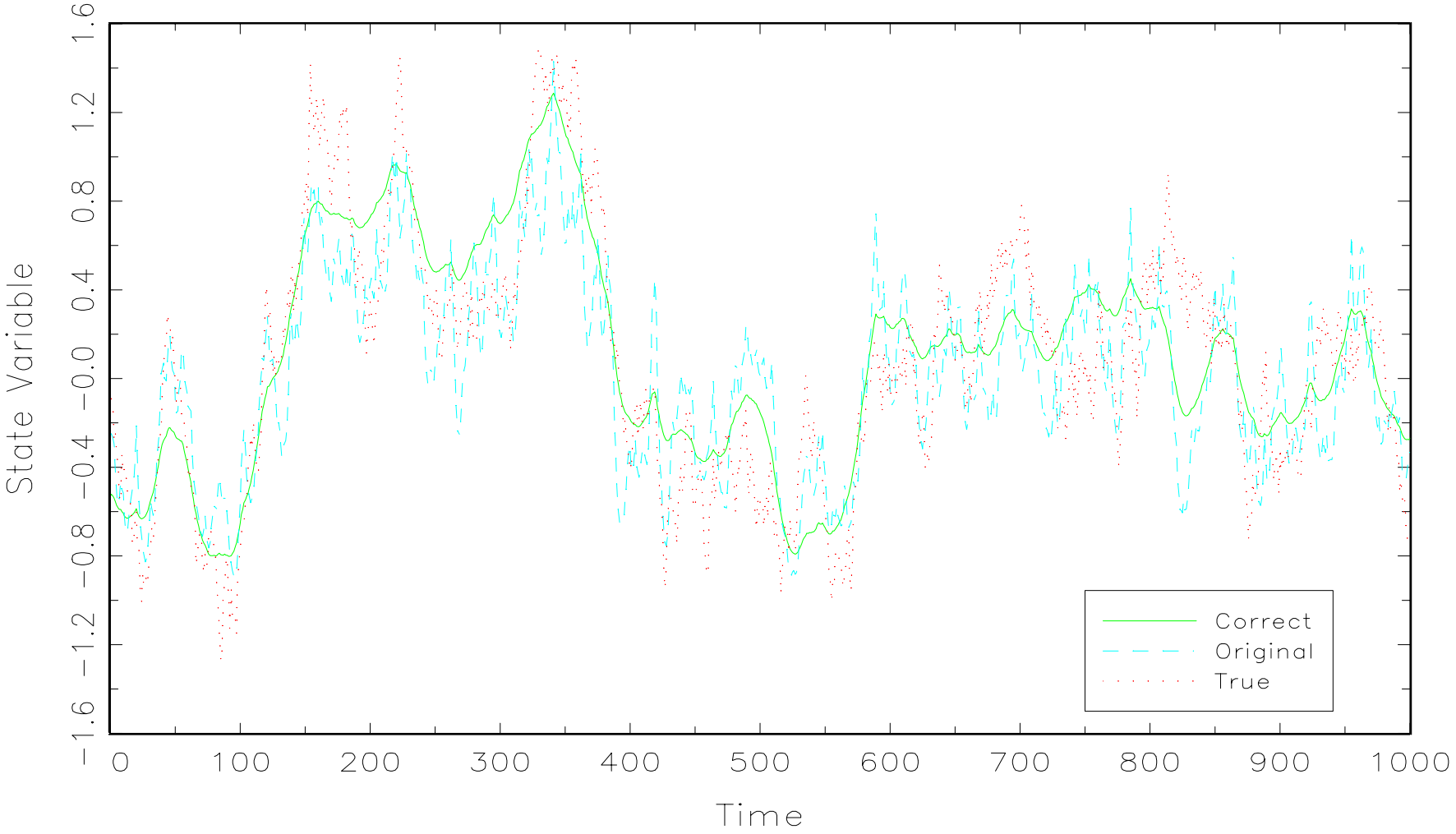

Figure 2: Estimates of State Variable (Daily)
(d) K = 100