ADAPTIVE KALMAN FILTERING WITH MULTIVARIATE GENERALIZED LAPLACE SYSTEM NOISE

Journal:	*Communications in Statistics - Simulation and Computation*
Manuscript ID:	LSSP-2010-0379.R1
Manuscript Type:	Original Paper
Date Submitted by the Author:	28-Feb-2011

Complete List of Authors:	Khawsithiwong, Pairoj; Faculty of Science, Silpakorn University, Department of Statistics Yatawara, Nihal; Curtin University of Technology, Mathematics and Statistics Pongsapukdee, Veeranun; Faculty of Science, Silpakorn University, Department of Statistics

Keywords:	adaptive filter, multivariate generalized Laplace distribution, system noise outlier

Abstract:	Adaptive Kalman filter is proposed to estimate the states of a system where the system noise is assumed to be a multivariate generalized Laplace random vector. In the presence of outliers in the system noise, it is shown that improved state estimates can be obtained by using an adaptive factor to estimate the dispersion matrix of the system noise term. A Monte-Carlo investigation is carried out to access the performance of the proposed filters and other robust filters. The results show that the proposed filter is superior to other filters when the magnitude of system change is moderate or large.

Adaptive Kalman Filtering with Multivariate Generalized Laplace System Noise

Khawsithiwong, P.

Department of Statistics, Faculty of Science, Silpakorn University, Nakorn Phatom, Thailand

Yatawara, N.

Department of Mathematics and Statistics, Curtin University of Technology, Perth, Western Australia

Pongsapukdee, V.

Department of Statistics, Faculty of Science, Silpakorn University, Nakorn Phatom, Thailand

Keywords: Adaptive filter, Multivariate generalized Laplace distribution, System noise outlier

Mathematics Subject Classification 62M; 62M20.

Abstract

An adaptive Kalman filter is proposed to estimate the states of a system where the system noise is assumed to be a multivariate generalized Laplace random vector. In the presence of outliers in the system noise, it is shown that improved state estimates can be obtained by using an adaptive factor to estimate the dispersion matrix of the system noise term. For the implementation of the filter, an algorithm which includes both single and multiple adaptive factors is proposed. A Monte-Carlo

investigation is also carried out to access the performance of the proposed filters in comparison with other robust filters. The results show that, in the sense of minimum mean squared state error, the proposed filter is superior to other filters when the magnitude of a system change is moderate or large.

1. Introduction

The well-known Kalman filter (Kalman 1960) was introduced to deal with problems of linear estimation and prediction for a linear Gaussian system defined by

$$\mathbf{X}_{t+1} = \mathbf{A}_t \mathbf{X}_t + \mathbf{W}_t \tag{1.1}$$

$$\mathbf{Y}_t = \mathbf{C}_t \mathbf{X}_t + \mathbf{V}_t \tag{1.2}$$

where $\mathbf{W}_t$ and $\mathbf{V}_t$ are respectively system noise and measurement noise sequences which are mutually uncorrelated. $\mathbf{W}_t$ and $\mathbf{V}_t$ are distributed as zero mean Gaussian random vectors with covariance matrices $\boldsymbol{\Sigma}_{\mathbf{W}}$ and $\boldsymbol{\Sigma}_{\mathbf{V}}$, respectively. By means of orthogonal projections, the original Kalman filter was derived in the sense of minimum mean squared state error. It consists of estimates of the state and its covariance matrix given by

$$\hat{\mathbf{X}}_{t|t} = \hat{\mathbf{X}}_{t|t-1} + \mathbf{K}_t \left(\mathbf{Y}_t - \mathbf{C}_t \hat{\mathbf{X}}_{t|t-1} \right) \tag{1.3}$$

and

$$\mathbf{P}_{t|t} = \left(\mathbf{I} - \mathbf{K}_t \mathbf{C}_t \right) \mathbf{P}_{t|t-1} \tag{1.4}$$

where

$$\hat{\mathbf{X}}_{t|t-1} = \mathbf{A}_{t-1} \hat{\mathbf{X}}_{t-1|t-1} \tag{1.5}$$

$$\mathbf{P}_{t|t-1} = \mathbf{A}_{t-1} \mathbf{P}_{t-1|t-1} \mathbf{A}'_{t-1} + \boldsymbol{\Sigma}_{\mathbf{W}} \tag{1.6}$$

and $$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{C}_t'\left(\mathbf{C}_t\mathbf{P}_{t|t-1}\mathbf{C}_t' + \frac{1}{2}\mathbf{\Sigma}_{\mathbf{V}}\right)^{-1} \quad (1.7)$$

However, the presence of outliers in the system noise term can cause a shift in the mean to a new level. In this situation, the traditional Kalman filter is no longer optimal due to the non-Gaussian noise term. To accommodate the outlier, the noise term is frequently assumed to have a symmetric heavy-tailed distribution such as a mixture of Gaussian components (Sorenson & Alspach 1971, Pena & Guttman 1988, Yatawara et al. 1991), a mixture of Student-t distributions (Meinhold & Singpurwalla 1989) or a univariate generalized Gaussian distribution (Niehsen 2002).

The construction of the adaptive filter proposed in this manuscript depends on a covariance matrix of the system noise which is estimated sequentially in time by exploiting system knowledge based on the discrepancy between the predicted state estimate and the measurement at a given time. There are several techniques available to estimate the noise covariance matrix such as Bayesian estimation, maximum likelihood estimation, correlation method, and covariance matching technique (Mehra 1972). In addition, a single adaptive factor (see, Yang et al. (2001)) which acts as a weighting factor between the predicted state estimates and the measurements can also be used in conjunction with the estimated noise covariance matrix to obtain an improved adaptive filter. A drawback of the latter technique is that all variables of the state vector are weighed by a single adaptive factor at the same time. Yang and Cui (2008) introduced an alternative adaptive filter with multi adaptive factors. This approach could reduce the state estimation errors more significantly than the filter with a single adaptive factor.

In this manuscript, an adaptive Kalman filter is developed by assuming that the system noise term is multivariate generalized Laplace distributed whose shape depends upon a shape parameter. In the presence of outliers, the shape parameter can be easily manipulated to estimate the covariance matrix of the system noise term in real time which enhances the performance of the filter. It is also shown that further gains in efficiency of the filters are achievable through the introduction of single and multi adaptive factors.

The paper is organized as follows. In Section 2, a brief description of the multivariate generalized Laplace distribution is given. The adaptive Kalman filtering as well as the algorithms with single and multi adaptive factors are developed in Section 3. In Section 4, the performances of the proposed adaptive filters are compared with the traditional Kalman filter and other robust filters. Finally, the summary and conclusions are provided in Section 5.

2. Multivariate Generalized Laplace Distribution

The multivariate generalized Laplace (MGL) distribution was introduced by Ernst (1998) as a class of multivariate models consisting of several distributions depending on the value of a shape parameter. This shape parameter λ distinguishes between members of the family such as the multivariate Laplace $(\lambda = 1)$, the multivariate normal $(\lambda = 2)$ and the multivariate uniform $(\lambda \to \infty)$ distributions as shown in Figure 1.

Let $\mathbf{Y}$ be a $k \times 1$ random vector, $\boldsymbol{\mu}$ be a $k \times 1$ vector of constants, and $\boldsymbol{\Sigma} = \left[\left(\sigma_{ij}\right)\right]$ be a $k \times k$ non-negative definite matrix. Suppose the random vector $\mathbf{Y}$ has an MGL distribution with the mean vector $\boldsymbol{\mu}$, the scale parameter matrix $\boldsymbol{\Sigma}$, and the shape parameter λ, denoted by $\mathbf{Y} \sim MGL_k\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda\right)$, with the joint density of $\mathbf{Y}$ defined as

$$f\left(\mathbf{y}\right) = \frac{\lambda \Gamma\left(\frac{k}{2}\right)}{2\pi^{\frac{k}{2}}\Gamma\left(\frac{k}{\lambda}\right)} \left|\boldsymbol{\Sigma}\right|^{-\frac{1}{2}} \exp\left\{-\left[\left(\mathbf{y}-\boldsymbol{\mu}\right)' \boldsymbol{\Sigma}^{-1}\left(\mathbf{y}-\boldsymbol{\mu}\right)\right]^{\frac{\lambda}{2}}\right\} \tag{2.1}$$

where $\Gamma(\cdot)$ denotes a gamma function. The mean vector and the covariance matrix of the MGL random vector $\mathbf{Y}$ are given respectively by

$$E\mathbf{Y} = \boldsymbol{\mu} \tag{2.2}$$

and

$$Cov(\mathbf{Y}) = \frac{\Gamma\left(\frac{k+2}{\lambda}\right)}{k\Gamma\left(\frac{k}{\lambda}\right)}\boldsymbol{\Sigma} \tag{2.3}$$

Also, let a partitioned random vector $\mathbf{Y}'_{k\times 1} = \left(\mathbf{Y}_{1(1\times k_1)}, \mathbf{Y}_{2(1\times(k-k_1))}\right)$ with a mean vector $\boldsymbol{\mu}'_{k\times 1} = \left(\boldsymbol{\mu}_{1(1\times k_1)}, \boldsymbol{\mu}_{2(1\times(k-k_1))}\right)$, and a scale parameter matrix $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11(k_1\times k_1)} & \boldsymbol{\Sigma}_{12(k_1\times(k-k_1))} \\ \boldsymbol{\Sigma}_{21((k-k_1)\times k_1)} & \boldsymbol{\Sigma}_{22((k-k_1)\times(k-k_1))} \end{bmatrix}$ be given.

Then, the conditional MGL density of $\mathbf{Y}_1 \middle| \mathbf{Y}_2 = \mathbf{y}_2^*$ is defined as

$$f\left(\mathbf{y}_1 \middle| \mathbf{y}_2 = \mathbf{y}_2^*\right) = \frac{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{k-k_1}{\lambda}\right)}{\pi^{\frac{k_1}{2}}\Gamma\left(\frac{k}{\lambda}\right)\Gamma\left(\frac{k-k_1}{2}\right)}\left|\boldsymbol{\Sigma}_{11.2}\right|^{-\frac{1}{2}}$$
$$\exp\left\{-\left[\left(\mathbf{y}_1-\boldsymbol{\mu}_{1.2}\right)'\boldsymbol{\Sigma}_{11.2}^{-1}\left(\mathbf{y}_1-\boldsymbol{\mu}_{1.2}\right)+\left(\mathbf{y}_2^*-\boldsymbol{\mu}_2\right)'\boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{y}_2^*-\boldsymbol{\mu}_2\right)\right]^{\frac{\lambda}{2}}+\left[\left(\mathbf{y}_2^*-\boldsymbol{\mu}_2\right)'\boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{y}_2^*-\boldsymbol{\mu}_2\right)\right]^{\frac{\lambda}{2}}\right\} \tag{2.4}$$

where $\boldsymbol{\mu}_{1.2} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{y}_2^*-\boldsymbol{\mu}_2\right)$ and $\boldsymbol{\Sigma}_{11.2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$. The mean vector and the covariance matrix of $\mathbf{Y}_1 \middle| \mathbf{Y}_2 = \mathbf{y}_2^*$ are given respectively by

$$E\left(\mathbf{Y}_1 \middle| \mathbf{Y}_2 = \mathbf{Y}_2^*\right) = \boldsymbol{\mu}_{1.2} \tag{2.5}$$

and

$$Cov\left(\mathbf{Y}_1 \middle| \mathbf{Y}_2 = \mathbf{Y}_2^*\right) = \frac{\Gamma\left(\frac{k_1+2}{\lambda}\right)}{k_1\Gamma\left(\frac{k_1}{\lambda}\right)}\boldsymbol{\Sigma}_{11.2} \tag{2.6}$$

See further details in Fang, Kotz & Ng (1990).

3. Adaptive Kalman Filtering

Consider a linear discrete-time stochastic system which possesses properties of observability and reachability given by equations (1.1) and (1.2) where $\mathbf{W}_t$ is assumed to be distributed as an zero mean MGL random vector with a scale parameter $\boldsymbol{\Sigma}_{\mathbf{W}}$ and a shape parameter $\lambda_{\mathbf{W}}$, denoted by $\mathbf{W}_t \sim MGL_r\left(0, \boldsymbol{\Sigma}_{\mathbf{W}}, \lambda_{\mathbf{W}}\right)$. $\mathbf{V}_t$ is a Gaussian random vector, denoted by $\mathbf{V}_t \sim MGL_k\left(0, \boldsymbol{\Sigma}_{\mathbf{V}}, 2\right)$ and the scale parameter matrices $\boldsymbol{\Sigma}_{\mathbf{W}}$ and $\boldsymbol{\Sigma}_{\mathbf{V}}$ are assumed to be known positive definite matrices. Then,by means of the least squares technique (Kalman 1960, Duncan & Horn 1972), an unbiased minimum variance state estimate can be derived leading to an adaptive Kalman filter with recursive estimates of the state and its covariance matrix given by

$$\hat{\mathbf{X}}_{t|t} = \hat{\mathbf{X}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{Y}_t - \mathbf{C}_t\hat{\mathbf{X}}_{t|t-1}\right) \tag{3.1}$$

and

$$\mathbf{P}_{t|t} = \left(\mathbf{I} - \mathbf{K}_t\mathbf{C}_t\right)\mathbf{P}_{t|t-1} \tag{3.2}$$

where

$$\hat{\mathbf{X}}_{t|t-1} = \mathbf{A}_{t-1}\hat{\mathbf{X}}_{t-1|t-1} \tag{3.3}$$

$$\mathbf{P}_{t|t-1} = \mathbf{A}_{t-1}\mathbf{P}_{t-1|t-1}\mathbf{A}'_{t-1} + \frac{\Gamma\left(\dfrac{r+2}{\lambda_{\mathbf{W}}}\right)}{r\Gamma\left(\dfrac{r}{\lambda_{\mathbf{W}}}\right)}\boldsymbol{\Sigma}_{\mathbf{W}} \tag{3.4}$$

and

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{C}'_t\left(\mathbf{C}_t\mathbf{P}_{t|t-1}\mathbf{C}'_t + \frac{1}{2}\boldsymbol{\Sigma}_{\mathbf{V}}\right)^{-1} \tag{3.5}$$

The use of MGL system noise term makes the adaptive Kalman filter, referred to as the MGLF here after, more versatile as it can accommodate outliers appearing in a system. However, to implement the MGLF, the shape parameter of the system noise term $\lambda_{\mathbf{W}}$ in (3.4) should be estimated at each point in time.

Define a pseudo state innovation, $\mathbf{Z}_t = \hat{\mathbf{X}}_{t|t} - \hat{\mathbf{X}}_{t|t-1} = \mathbf{K}_t\left(\mathbf{Y}_t - \mathbf{C}_t\hat{\mathbf{X}}_{t|t-1}\right)$, which is assumed to be an MGL distributed random vector with time-varying shape parameter, denoted by $\mathbf{Z}_t \sim MGL_r\left(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{Z}}, \lambda_{\mathbf{Z}(t)}\right)$, with the covariance matrix $\mathbf{\Sigma}_{\mathbf{Z}(t)} = \dfrac{\Gamma\left(\dfrac{r+2}{\lambda_{\mathbf{Z}(t)}}\right)}{r\Gamma\left(\dfrac{r}{\lambda_{\mathbf{Z}(t)}}\right)}\mathbf{\Sigma}_{\mathbf{Z}} = \mathbf{K}_t\left(\mathbf{C}_t\mathbf{P}_{t|t-1}\mathbf{C}'_t + \dfrac{1}{2}\mathbf{\Sigma}_{\mathbf{V}}\right)\mathbf{K}'_t$. Also, define a partitioned random vector $\mathbf{Z}_t = \left(Z_{jt}, \mathbf{Z}^*_t\right)'$ where $\mathbf{Z}^*_t = \left(Z_{1t}, Z_{2t}, \ldots, Z_{(j-1)t}, Z_{(j+1)t}, \ldots, Z_{rt}\right)$ and the corresponding covariance parameter matrix $\mathbf{\Sigma}_{\mathbf{Z}} = \begin{bmatrix} \sigma_{jj} & \mathbf{\Sigma}^{*\prime}_j \\ \mathbf{\Sigma}^*_j & \mathbf{\Sigma}^* \end{bmatrix}$ for $j = 1, 2, \ldots, r$. Then, the conditional density function of Z_{jt} given $\mathbf{Z}^*_t = \mathbf{z}^*_t$ is also distributed as a zero mean univariate generalized Laplace random variable with a scale parameter $\sigma^2_j = \sigma_{jj} - \mathbf{\Sigma}^{*\prime}_j\mathbf{\Sigma}^{*-1}\mathbf{\Sigma}^*_j$ and a shape parameter $\lambda_{Z_j(t)}$, denoted by $Z_{jt}\left|\mathbf{Z}^*_t = \mathbf{z}^*_t\right. \sim MGL_1\left(0, \sigma^2_j, \lambda_{Z_j(t)}\right)$, where the conditional variance of $Z_{jt}\left|\mathbf{Z}^*_t = \mathbf{z}^*_t\right.$ is given by $\sigma^2_{Z_{jt}|\mathbf{Z}^*_t=\mathbf{z}^*_t} = \dfrac{\Gamma\left(\dfrac{3}{\lambda_{Z_j(t)}}\right)}{\Gamma\left(\dfrac{1}{\lambda_{Z_j(t)}}\right)}\sigma^2_j$. The MGLF can be implemented using the following algorithm.

i) Enter the initial estimates $\hat{\mathbf{X}}_{t|t-1}$ and $\mathbf{P}_{t|t-1}$.

ii) Collect a new measurement $\mathbf{Y}_t$.

iii) Compute the filter gain $\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{C}_t'\left(\mathbf{C}_t\mathbf{P}_{t|t-1}\mathbf{C}_t' + \frac{1}{2}\mathbf{\Sigma}_\mathbf{V}\right)^{-1}$.

iv) **State vector update**

Update the state estimate $\hat{\mathbf{X}}_{t|t}$ and state error covariance matrix $\mathbf{P}_{t|t}$ by

$$\hat{\mathbf{X}}_{t|t} = \hat{\mathbf{X}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{Y}_t - \mathbf{C}_t\hat{\mathbf{X}}_{t|t-1}\right)$$
$$\mathbf{P}_{t|t} = \left(\mathbf{I} - \mathbf{K}_t\mathbf{C}_t\right)\mathbf{P}_{t|t-1}.$$

v) Approximating the shape parameter

Compute a pseudo state innovation, $\mathbf{Z}_t = \hat{\mathbf{X}}_{t|t} - \hat{\mathbf{X}}_{t|t-1}$, by assuming that $\mathbf{Z}_t$ is distributed as an MGL random vector, denoted by $\mathbf{Z}_t \sim MGL_r\left(\mathbf{0}, \mathbf{\Sigma}_\mathbf{Z}, \lambda_{\mathbf{Z}(t)}\right)$.

a. **MGLF with single adaptive factor (MGLF-S)**

Obtain a maximum likelihood estimate $\hat{\lambda}_{\mathbf{Z}(t)}$ with the scale parameter matrix $\mathbf{\Sigma}_\mathbf{Z}$ defined by $\mathbf{\Sigma}_\mathbf{Z} = 2\delta_t^2\mathbf{K}_t\left(\mathbf{C}_t\mathbf{P}_{t|t-1}\mathbf{C}_t' + \frac{1}{2}\mathbf{\Sigma}_\mathbf{V}\right)\mathbf{K}_t'$ where $\delta_t = \frac{\hat{\lambda}_{\mathbf{Z}(t-1)}}{2}$ is a time-varying adaptive factor.

b. **MGLF with multi adaptive factors (MGLF-M)**

Obtain maximum likelihood estimators $\hat{\lambda}_{Z_j(t)}$ for $j = 1, 2, \ldots, r$ from the conditional distribution of Z_{jt} given $\mathbf{Z}_t^* = \mathbf{z}_t^*$ with the scale parameter matrix σ_j^2 defined by $\sigma_j^2 = 2\delta_{jt}^2\left(\sigma_{jj} - \mathbf{\Sigma}_j^{*\prime}\mathbf{\Sigma}^{*-1}\mathbf{\Sigma}_j^*\right)$ where $\delta_{jt} = \frac{\hat{\lambda}_{Z_j(t-1)}}{2}$ is a time-varying adaptive factor of variable j at time t.

vi) Time update

Compute the one step ahead state estimate $\hat{\mathbf{X}}_{t+1|t}$ and state forecast error covariance matrix $\mathbf{P}_{t+1|t}$ given by

a. For MGLF-S,

$$\hat{\mathbf{X}}_{t+1|t} = \mathbf{A}_t \hat{\mathbf{X}}_{t|t}$$

and $$\mathbf{P}_{t+1|t} = \mathbf{A}_t \mathbf{P}_{t|t} \mathbf{A}_t' + \frac{\Gamma\left(\frac{r+2}{\hat{\lambda}_{\mathbf{Z}(t)}}\right)}{r\Gamma\left(\frac{r}{\hat{\lambda}_{\mathbf{Z}(t)}}\right)} \boldsymbol{\Sigma}_{\mathbf{W}}.$$

b. For MGLF-M,

$$\hat{\mathbf{X}}_{t+1|t} = \mathbf{A}_t \hat{\mathbf{X}}_{t|t}$$

and $$\mathbf{P}_{t+1|t} = \mathbf{A}_t \mathbf{P}_{t|t} \mathbf{A}_t' + \boldsymbol{\Lambda}^{\frac{1}{2}} \boldsymbol{\Sigma}_{\mathbf{W}} \boldsymbol{\Lambda}^{\frac{1}{2}},$$

where $$\boldsymbol{\Lambda} = diag\left(\frac{\Gamma\left(\frac{r+2}{\hat{\lambda}_{Z_1(t)}}\right)}{r\Gamma\left(\frac{r}{\hat{\lambda}_{Z_1(t)}}\right)}, \frac{\Gamma\left(\frac{r+2}{\hat{\lambda}_{Z_2(t)}}\right)}{r\Gamma\left(\frac{r}{\hat{\lambda}_{Z_2(t)}}\right)}, \ldots, \frac{\Gamma\left(\frac{r+2}{\hat{\lambda}_{Z_r(t)}}\right)}{r\Gamma\left(\frac{r}{\hat{\lambda}_{Z_r(t)}}\right)} \right).$$

vii) Let $t = t+1$ and go to step 2.

The algorithms of both MGLF-S and MGLF-M are similar. However, as shown in the simulation study their performances are significantly different. Suppose a bivariate linear discrete-time stochastic system with a shift in the first variable is considered. In Figure 2(a), both the MGLFs and the traditional Kalman filter are presented to illustrate their performances. It is clear that the state estimates of MGLFs can describe the system more accurately than those of the traditional Kalman filter. Furthermore, the state estimates from MGLF-M are much smoother than those of MGLF-S in the second variable as shown in Figure 2(b).

Consider a particular system shift at time 51 in the first variable. For this case, the evolution of the estimated shape parameter values of the first variable, calculated by MGLF-M, are given in Figure 2(d). These values are similar to those of MGLF-S as in Figure 2(c). However, unlike in this case the shape parameter values of MGLF-M in the second variable are rarely affected by a change in the first variable as shown in Figure 2(e). Hence, it can be argued that the use of multi adaptive factors tend to protect against over adjustment providing state estimates which are more reliable for all variables.

4. Effect of adaptive factor

Essentially the adaptive factor ranging between 0 and 1 is used as a weighting factor in the covariance matrix of the state innovations for calculating the shape parameter. Its role is pivotal to the efficiency of the MGLF algorithm and facilitates a spontaneous response to a large system change by assuming a value less than 1 in a short period of time. In the estimation of adaptive state estimates, the adaptive factor and the adaptively estimated covariance matrix of the system noise term function together to significantly improve the ability of the MGLFs.

Figure 3 shows the mean squared state error (MSSE) for each value of the adaptive factor δ_t in a range of 0.1 to 4 stepped up by 0.1 when the systems are subjected shifts of various magnitudes 0, 1, 3, and 5 times the standard deviation. In the absence of a shift, the value of δ_t becomes approximately equal to one when the filter attains a minimum MSSE value. However, when the magnitude of a system shift increases, δ_t tends to zero when MSSE is minimum. Therefore, the adaptive factor should be varied over time depending upon the magnitude of the pseudo state innovation. This further suggests that the use of a time-varying adaptive factor could considerably improve the performance of the MGLF. These are explored in the next section.

5. Performance Study

Consider the 5-variate linear discrete-time Gaussian system defined by equations (1.1) and (1.2) where the matrices $\mathbf{A}_t$ and $\mathbf{C}_t$ are set to be the identity matrices, $\mathbf{I}$. A Monte Carlo simulation consisting of 2,000 iterations with MSSE as a preferred criterion for comparison was conducted under following conditions.

i) To compare with the traditional Kalman filter (KALMAN), the robust filter with mixture Gaussian noises (MIXTURE) (see Yatawara (1986)), the robust filter with generalized Gaussian noise (GGAUSSIAN) (see Masreliez (1975) and Niehsen (2002)), the MGLF-S and the MGLF-M are selected.

ii) Measurement noise variances $\sigma_{\mathbf{V}}^2 = 1$ and correlation coefficients of the measurement noise terms $\rho_{\mathbf{V}} = 0,\ 0.4,$ and 0.8,

iii) System noise variances $\sigma_{\mathbf{W}}^2 = 0.01$ and correlation coefficients of system noise terms $\rho_{\mathbf{W}} = 0,\ 0.4,$ and 0.8,

iv) Magnitude of system shift $\delta_M = 0,\ 0.5,\ 1,\ 2,\ 3,\ 4,$ and 5,

v) A mean shift in the first variable at time 26 of the multivariate time series of length 50 is introduced, by means of $W_{1(26)}$ which is distributed as Gaussian random variable with mean $\mu_1 = \mu_0 + \delta_M \sigma_{\mathbf{V}}$ where $\mu_0 = 0$ and variance $\sigma^2_{\mathbf{W}}$,

vi) MSSE using 10 time points after a system shift is calculated by,

$$\text{MSSE} = \frac{1}{10} \sum_{t=26}^{35} \left(\mathbf{X}_t - \hat{\mathbf{X}}_{t|t} \right)' \left(\mathbf{X}_t - \hat{\mathbf{X}}_{t|t} \right).$$

The structure of both noise covariance matrices is based on $\mathbf{\Sigma} = \sigma^2 (1-\rho)\mathbf{I} + \sigma^2 \rho \mathbf{J}$ where $\mathbf{J}$ is a matrix of one and $\mathbf{I}$ is an identity matrix.

The results in Table 1 reveal that MSSE values of KALMAN tend to increase rapidly and are considerably higher than those of the robust filters when δ_M increases for all cases of correlation of the system and the measurement noise terms. Evidently, the KALMAN is not optimal following an occurrence of a system shift. In comparison, the MSSE values of the robust filters also increase but at a slower rate and with significantly less magnitudes than the KALMAN. This implies that robust filters provide more consistent state estimates than the KALMAN.

When a system is not subjected to a shift $(\delta_M = 0)$, KALMAN clearly provides the smallest MSSE values for all combinations of $\rho_{\mathbf{W}}$ and $\rho_{\mathbf{V}}$ showing its optimality in this situation. However, for $\delta_M = 0.5$, MSSE values of MIXTURE become the smallest in comparison to all other filters. When δ_M is moderate or large, MGLF-M shows superiority over other filters in the minimum MSSE sense. This implies that MGLF-M gives more precise state estimates than other filters for all significant δ_M and $(\rho_{\mathbf{W}}, \rho_{\mathbf{V}})$ combinations.

All filters excluding GGAUSSIAN are affected by $\rho_{\mathbf{W}}$ and $\rho_{\mathbf{V}}$ and the MSSE values of the filters tend to increase proportional to $\rho_{\mathbf{W}}$. Unfortunately, a large value of $\rho_{\mathbf{W}}$ depreciates the precision of the state estimates. In contrast, MSSE values decrease as $\rho_{\mathbf{V}}$ increases. But, the effects of $\rho_{\mathbf{W}}$ and $\rho_{\mathbf{V}}$ are not overwhelming as shown by the results of the comparisons. It's also noticeable that the MSSE values of GGAUSSIAN are consistently higher than for other robust filters when $\rho_{\mathbf{V}}$ =0.4 and 0.8 but are insensitive to the value of $\rho_{\mathbf{W}}$.

The rate of convergence in MSSE of all filters excluding GGAUSSIAN are illustrated in Figure 4. For $\delta_M = 0$ and 0.5, MGLF-S and MGLF-M provide larger MSSE values than KALMAN and MIXTURE. When the magnitude of a shift is significant $(\delta_M \geq 1)$, MSSE values of MGLF-M are the smallest values and decrease more rapidly than in other filters after a shift occurs. In contrast, KALMAN yields a gradual decrease in MSSE values and then becomes excessively large when δ_M is large. MSSE values of MIXTURE become closer to those of MGLF-M as δ_M increases. Clearly, the larger δ_M, the slower the convergence rate in MSSE of MGLF-S is, in comparison to MGLF-M.

6. Conclusion

The adaptive Kalman filtering approach proposed in this paper mainly deals with the presence of an outlier in the system noise term. The filters are developed by assuming an MGL system noise distribution to adaptively estimate the system noise covariance matrix requiring only information of the current measurement. In addition, a time-varying adaptive factor is also used to enhance the performance of the MGLFs. This leads to a more accurate recursive estimation procedure to represent the evolution of a system.

Moreover, algorithms to implement the MGLFs are given in both situations of the single and multi adaptive factors. The adaptive factor is mainly used as a weight in the scale parameter matrix of the state innovation to approximate the shape parameter. The single adaptive factor suggested treats all variables equally. This was shown to be a drawback of the method as it tends to adjust all variables whether they are changing or not. To eliminate this drawback, the single adaptive factor was replaced by multi adaptive factors which were determined by considering magnitudes of innovations corresponding to each state variable. This leads to an efficient recursive filter with reduced estimation error of the state and provides more reliable information of each variable.

The simulation results show that when a system shift is moderate or large, the MGLFs are more effective than the traditional Kalman filter and other investigated robust filters for all combinations of ρ_{W} and ρ_{V}. Also, MGLF-M provides the most rapid reduction in MSSE. Moreover, MGLFs are not more cumbersome to implement than the algorithm of the traditional Kalman filter. However, it is far superior to track the evolution of a system than the traditional one.

Acknowledgements

This work is supported by Office of the Higher Education Commission and Thailand Research Fund which are greatly acknowledged, through project number MRG5180370. The anonymous reviewers are acknowledged. With their critical comments and helpful revisions, the manuscript has been greatly improved.

References

Duncan, D. B., Horn, S. D. (1972). Linear dynamic recursive estimation from the viewpoint of regression analysis. *Journal of the American Statistical Association* 67:815-821.

Ernst, M. D. (1998). A multivariate generalized Laplace distribution. *Computational Statistics* 13:227-232.

Fang, K.-T., Kotz, S., Ng, K. W. (1990). *Symmetric Multivariate and Related Distributions.* New York: Chapman and Hall.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82(D):35-45.

Masreliez, C. J. (1975). Approximate non-Gaussian filtering with linear state and observation relations. *IEEE Transactions on Automatic Control* AC-20,107-110.

Meinhold, R. J., Singpurwalla, N. D. (1989). Robustification of Kalman filter models. *Journal of the American Statistical Association* 84:479-486.

Mehra, R. K. (1972). Approaches to adaptive filtering. *IEEE Transactions on Automatic Control* 17:693-698.

Niehsen, W. (2002). Robust Kalman filtering with generalized Gaussian measurement noise. *IEEE Transactions on Aerospace and Electronic System* 38:1409-1412.

Pena, D., Guttman, I. (1988). Bayesian approach to robustifying the Kalman filter. In: Spall, J. C., ed., *Bayesian Analysis of Time Series and Dynamic Models.* New York:Marcel Dekker, pp. 227-253.

Sorenson, H. W., Alspach, D. L. (1971). Recursive Bayesian estimation using Gaussian sums. *Automatica* 7:465-479.

Yatawara, N. (1986). Detection of outliers and random events in time series, PhD thesis, Department of Statistics and Actuarial Science, University of Waterloo, Canada.

Yatawara, N., Abraham, B., MacGregor, J. F. (1991). A Kalman filter in the presence of outliers. *Communications in Statistics-Theory and Methods* 20:1803-1820.

Yang, Y., He, H., Xu, G. (2001). A new adaptively robust filtering for kinematic geodetic positioning. *Journal of Geodesy* 75:109-116.

Yang, Y., Cui, X. (2008). Adaptively robust filter with multi adaptive factor. *Survey Review* 40:260-270.

Figures

Figure 1 Density plots of the bivariate MGL distribution with $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathbf{I}$ when values of the shape parameter are 1, 2, 5, and ∞

Figure 2 Performances and shape parameters of MGLF-S and MGLF-M: (a) the state estimates of the first variable $\left(\hat{X}_{1t|t}\right)$, (b) the state estimates of the second variable $\left(\hat{X}_{2t|t}\right)$, (c) the estimated shape parameter $\hat{\lambda}_{\mathbf{Z}(t)}$ of MGLF-S, (d) the estimated shape parameter $\hat{\lambda}_{\mathbf{Z}_1(t)}$ of MGLF-M for the first variable, (e) the estimated shape parameter $\hat{\lambda}_{\mathbf{Z}_2(t)}$ of MGLF-M for the second variable.

Figure 3 The magnitude of adaptive factor δ_t corresponding to a minimum mean of squared state error for various magnitudes of system shift.

Figure 4 Rate of convergence in MSSE for various magnitudes of system shift.

Tables

Table 1 MSSE of the filters for various magnitudes of system shifts and combinations of correlation coefficients $\rho_{\mathbf{W}}$ and $\rho_{\mathbf{V}}$.

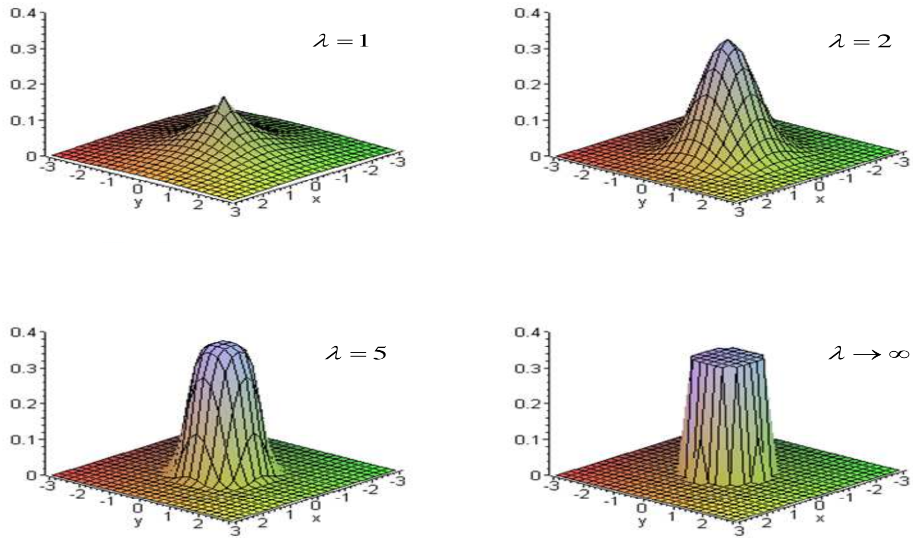


Figure 1 Density plots of the bivariate MGL distribution with $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathbf{I}$ when values of the shape parameter are 1, 2, 5, and ∞

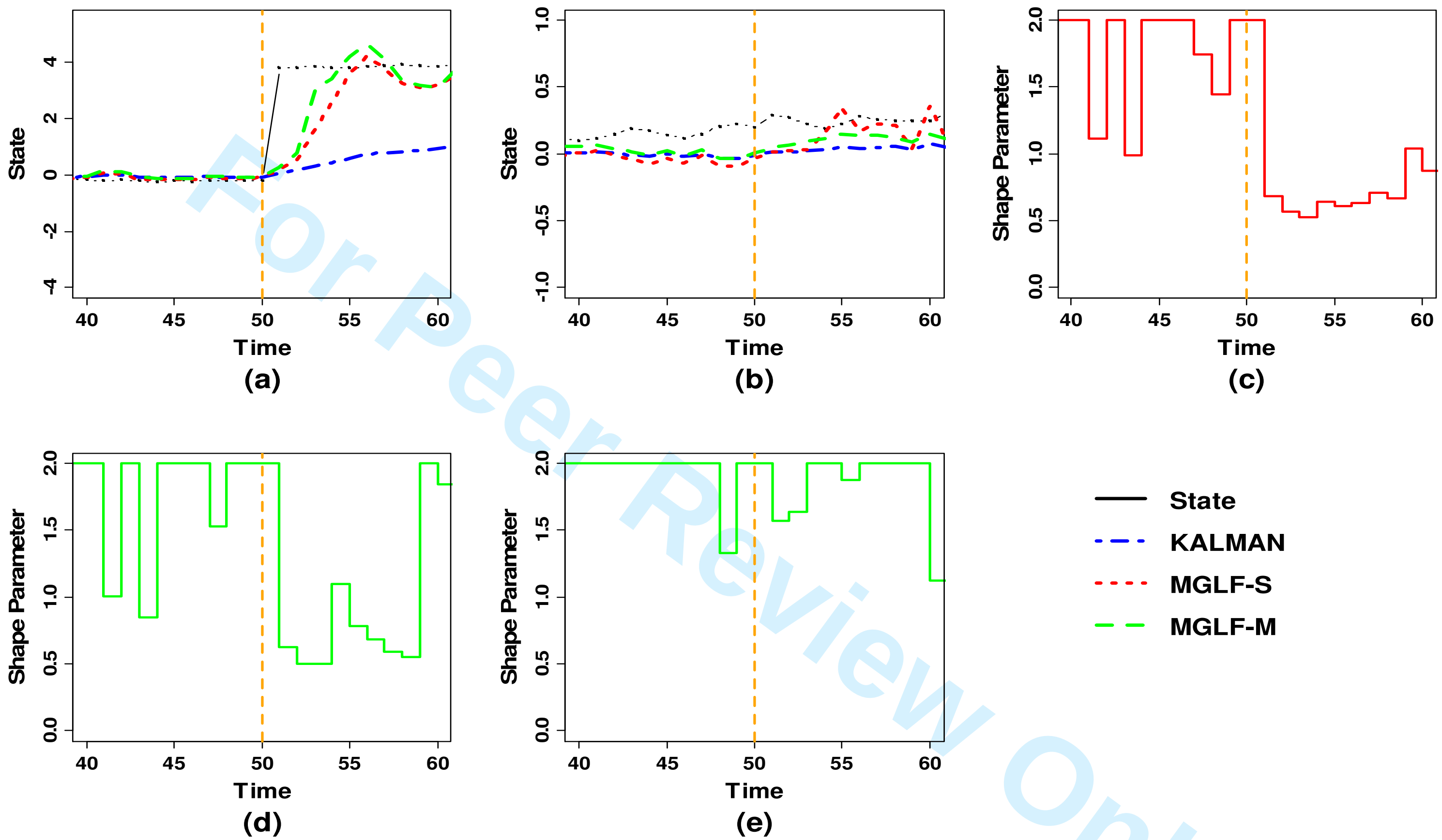


Figure 2 Performances and shape parameters of MGLF-S and MGLF-M: (a) the state estimates of the first variable $\left(\hat{X}_{1t|t}\right)$, (b) the state estimates of the second variable $\left(\hat{X}_{2t|t}\right)$, (c) the estimated shape parameter $\hat{\lambda}_{\mathbf{Z}(t)}$ of MGLF-S, (d) the estimated shape parameter $\hat{\lambda}_{\mathbf{Z}_1(t)}$ of MGLF-M for the first variable, (e) the estimated shape parameter $\hat{\lambda}_{\mathbf{Z}_2(t)}$ of MGLF-M for the second variable.

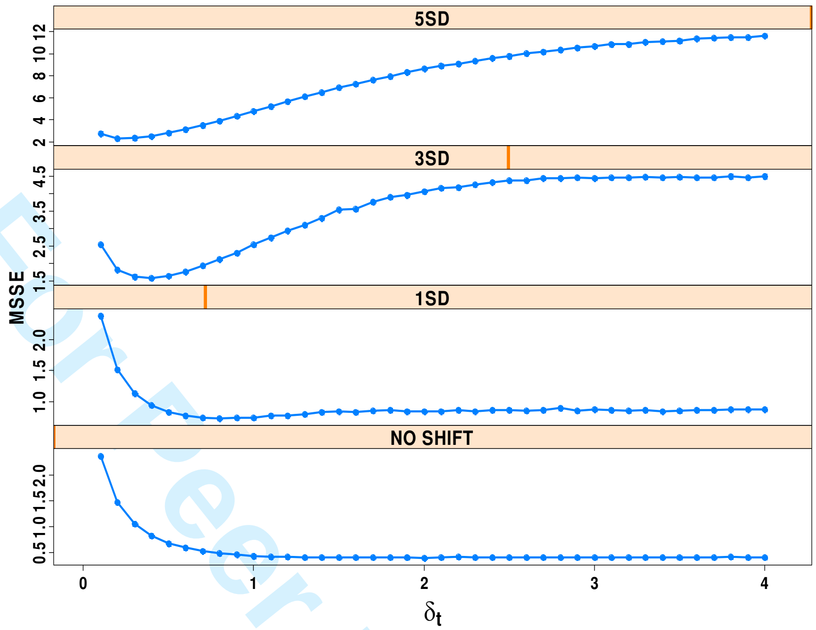


Figure 3 The magnitude of adaptive factor δ_t corresponding to a minimum mean of squared state error for various magnitudes of system shift.

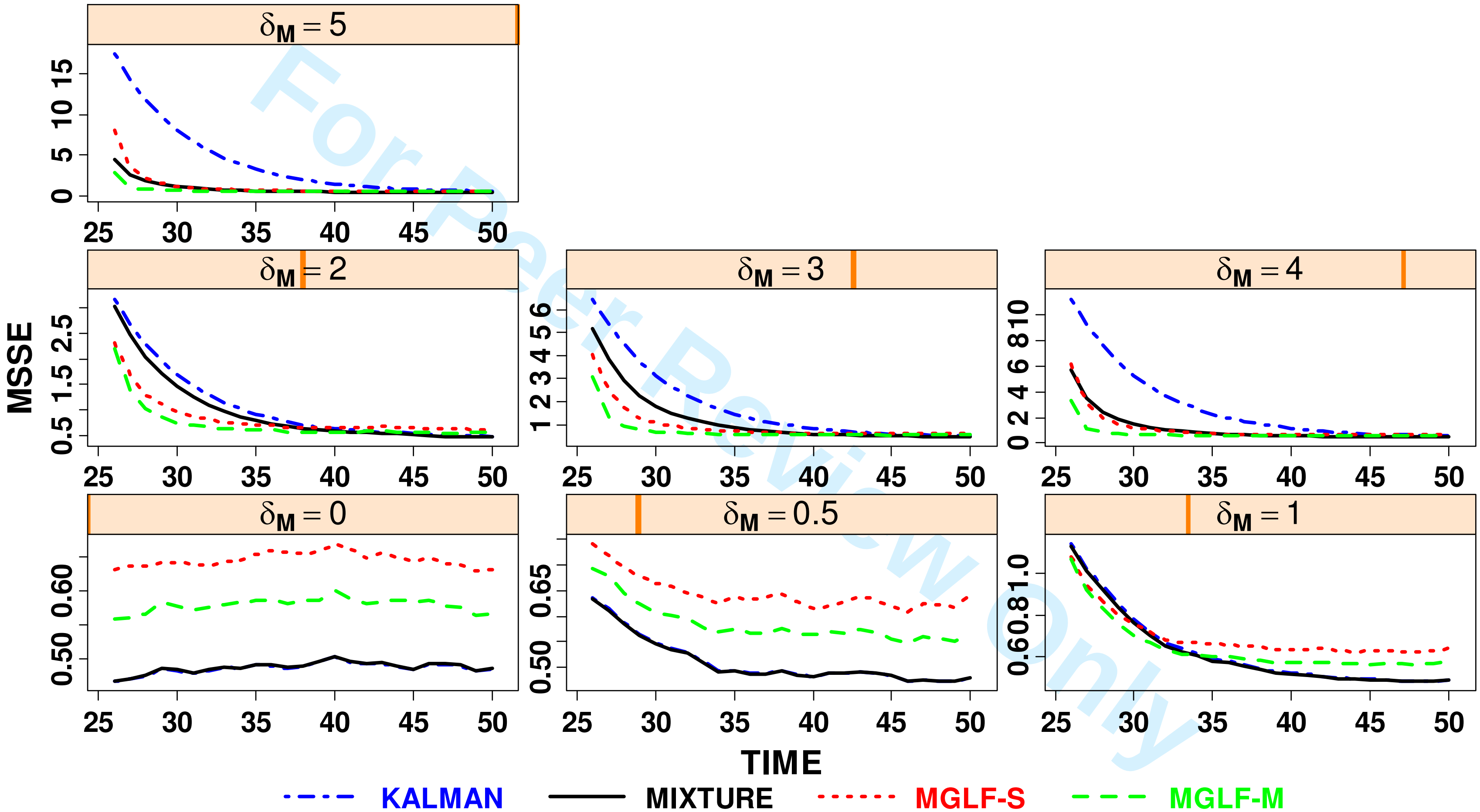


Figure 4 Rate of convergence in MSSE for various magnitudes of system shift.

Table 1

MSSE of the filters for various magnitudes of system shifts and combinations of correlation coefficients

ρ_{W} and ρ_{V} .

			FILTER				
ρ_{W}	ρ_{V}	δ_M	KALMAN	MIXTURE	GGAUSSIAN	MGLF-S	MGLF-M
0	0	0	**0.47306**	0.47351	1.56170	0.62759	0.56687
		0.5	0.54853	**0.54632**	1.57122	0.66713	0.60918
		1	0.80302	0.79079	1.63258	0.77635	**0.74448**
		2	1.76552	1.65742	1.65939	1.12863	**1.04947**
		3	3.35479	2.75679	1.67633	1.56973	**1.21407**
		4	5.66575	3.33769	1.61848	1.98901	**1.24630**
		5	8.52175	2.86450	1.64129	2.31587	**1.21557**
	0.4	0	**0.44519**	0.44571	16.67773	0.58836	0.52951
		0.5	0.50892	**0.50693**	16.35445	0.61494	0.56322
		1	0.74167	0.72835	16.38809	0.72911	**0.69935**
		2	1.59116	1.44404	16.28529	1.03751	**0.91943**
		3	3.11503	2.19364	15.90009	1.43165	**1.01234**
		4	5.11525	2.07543	15.74169	1.74691	**0.99905**
		5	7.72958	1.65953	15.60685	2.03978	**0.96043**
	0.8	0	**0.33066**	0.33162	45.71241	0.43776	0.39486
		0.5	0.38418	**0.38405**	45.66550	0.46858	0.43362
		1	0.55562	0.54411	45.73173	0.56040	**0.51760**
		2	1.15418	0.86298	45.65861	0.78913	**0.56478**
		3	2.21408	0.90197	45.99350	1.03920	**0.59230**
		4	3.65632	1.05598	45.67210	1.27708	**0.60305**
		5	5.55024	1.36908	45.60713	1.48399	**0.64258**

Table 1

continued

			FILTER				
ρ_W	ρ_V	δ_M	KALMAN	MIXTURE	GGAUSSIAN	MGLF-S	MGLF-M
0.4	0	0	**0.43686**	0.43730	1.56489	0.58704	0.53526
		0.5	0.53176	**0.52939**	1.58837	0.63899	0.59534
		1	0.80035	0.78652	1.60171	0.74960	**0.72858**
		2	1.89129	1.77652	1.64419	1.18731	**1.07480**
		3	3.74580	3.05963	1.65336	1.73108	**1.29312**
		4	6.25678	3.62481	1.63848	2.25208	**1.34365**
		5	9.57008	3.07329	1.65172	2.69843	**1.33901**
	0.4	0	**0.47704**	0.47748	16.59604	0.62938	0.56642
		0.5	0.55529	**0.55270**	16.51674	0.67367	0.61205
		1	0.79237	0.77821	16.52560	0.77053	**0.72434**
		2	1.74497	1.56634	16.31735	1.11680	**0.94753**
		3	3.38400	2.23078	15.94353	1.50707	**1.01349**
		4	5.60411	2.01532	15.77101	1.83702	**0.99266**
		5	8.57257	1.56396	15.74194	2.09234	**0.95245**
	0.8	0	**0.43708**	0.43813	45.94689	0.57344	0.50412
		0.5	0.49812	**0.49774**	46.00221	0.61934	0.55478
		1	0.65673	0.64157	45.75734	0.68350	**0.60001**
		2	1.28498	0.91278	45.64017	0.87591	**0.62314**
		3	2.36920	0.91058	45.94054	1.06529	**0.59882**
		4	3.87421	1.05157	45.78245	1.21801	**0.59907**
		5	5.76528	1.28089	45.67897	1.29987	**0.60281**

Table 1

Continued

			FILTER				
ρ_W	ρ_V	δ_M	KALMAN	MIXTURE	GGAUSSIAN	MGLF-S	MGLF-M
0.8	0	0	**0.33839**	0.33870	1.57574	0.47722	0.47394
		0.5	0.46595	**0.46404**	1.57797	0.54466	0.54371
		1	0.85827	0.84862	1.64532	0.76548	**0.75477**
		2	2.43277	2.32896	1.67108	1.52171	**1.24368**
		3	5.01866	4.26382	1.67592	2.50489	**1.53625**
		4	8.65884	5.18856	**1.64364**	3.50083	1.65543
		5	13.47593	4.27368	**1.63858**	4.49276	1.78279
	0.4	0	**0.42751**	0.42793	16.56262	0.58113	0.54110
		0.5	0.55203	**0.54940**	16.53059	0.65677	0.62284
		1	0.89735	0.88085	16.51947	0.83868	**0.77381**
		2	2.31026	2.05126	16.34804	1.43225	**1.07659**
		3	4.63855	2.81756	15.54790	2.10278	**1.17461**
		4	7.90570	2.33690	15.53645	2.70485	**1.20841**
		5	12.18559	1.65301	15.65867	3.22789	**1.19132**
	0.8	0	**0.47304**	0.47354	46.08451	0.64046	0.56837
		0.5	0.55684	**0.55543**	45.43531	0.68207	0.60777
		1	0.80269	0.77461	45.88378	0.80331	**0.68161**
		2	1.77115	1.05534	45.63807	1.14735	**0.76462**
		3	3.37481	0.91471	46.11275	1.38656	**0.74408**
		4	5.65228	1.05940	45.86809	1.60624	**0.75353**
		5	8.54040	1.25765	45.69582	1.68928	**0.74807**

Department of Statistics
Faculty of Science, Silpakorn University
Rajchamakchanai Road Aumphour Maung
Nakorn Phatom Thailand 73000

14 December 2009

Communications in Statistics
Prof. N. Balakrishnan, Editor-in-Chief
Department of Mathematics & Statistics
McMaster University
1280 Main Street West
Hamilton, Ontario L8S 4K1, CANADA

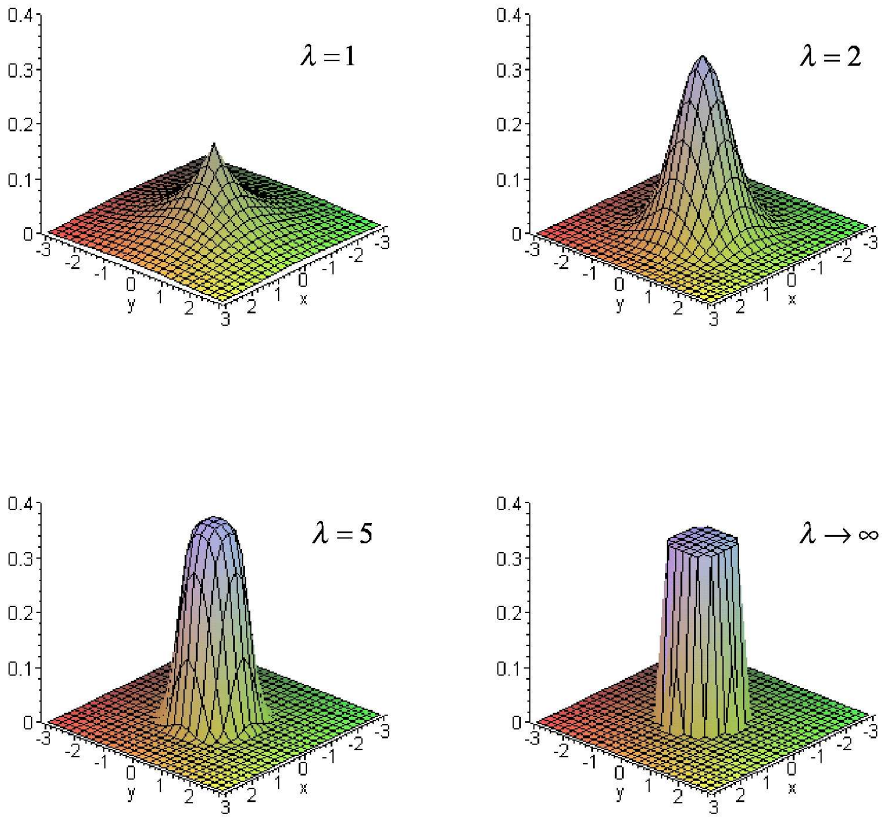
Dear Sir/Madam

Please find our revision manuscript (LSSP-2010-0379) titled "ADAPTIVE KALMAN FILTERING WITH MULTIVARIATE GENERALIZED LAPLACE SYSTEM NOISE" that has been corrected as editor's comments giving in a decision letter from the Communications in Statistics—Simulation and Computation.

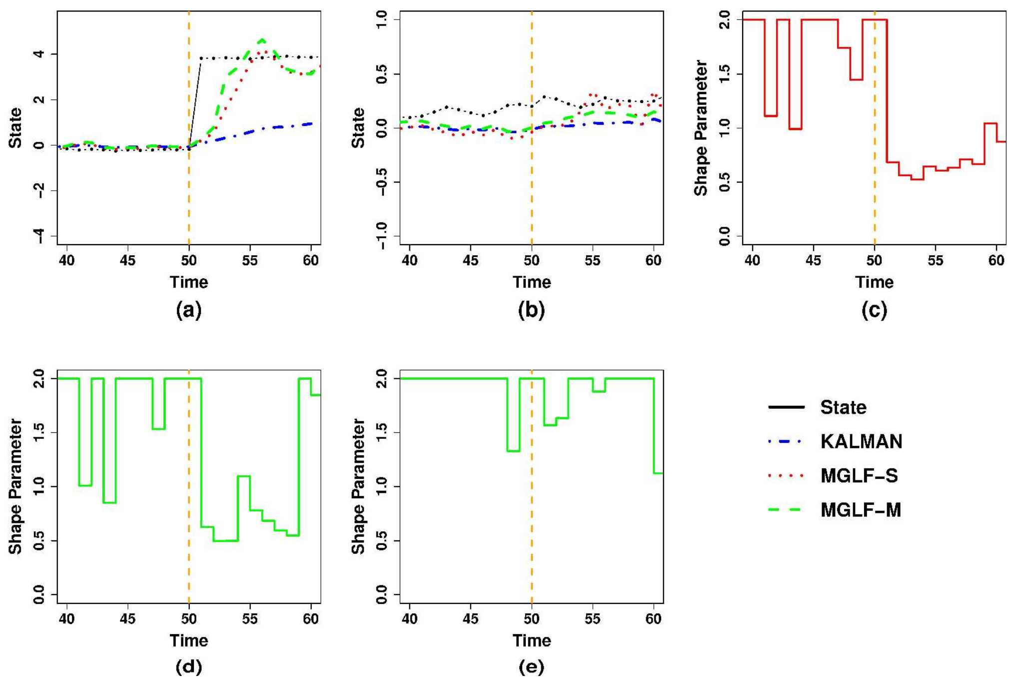
Thank you.

Yours sincerely

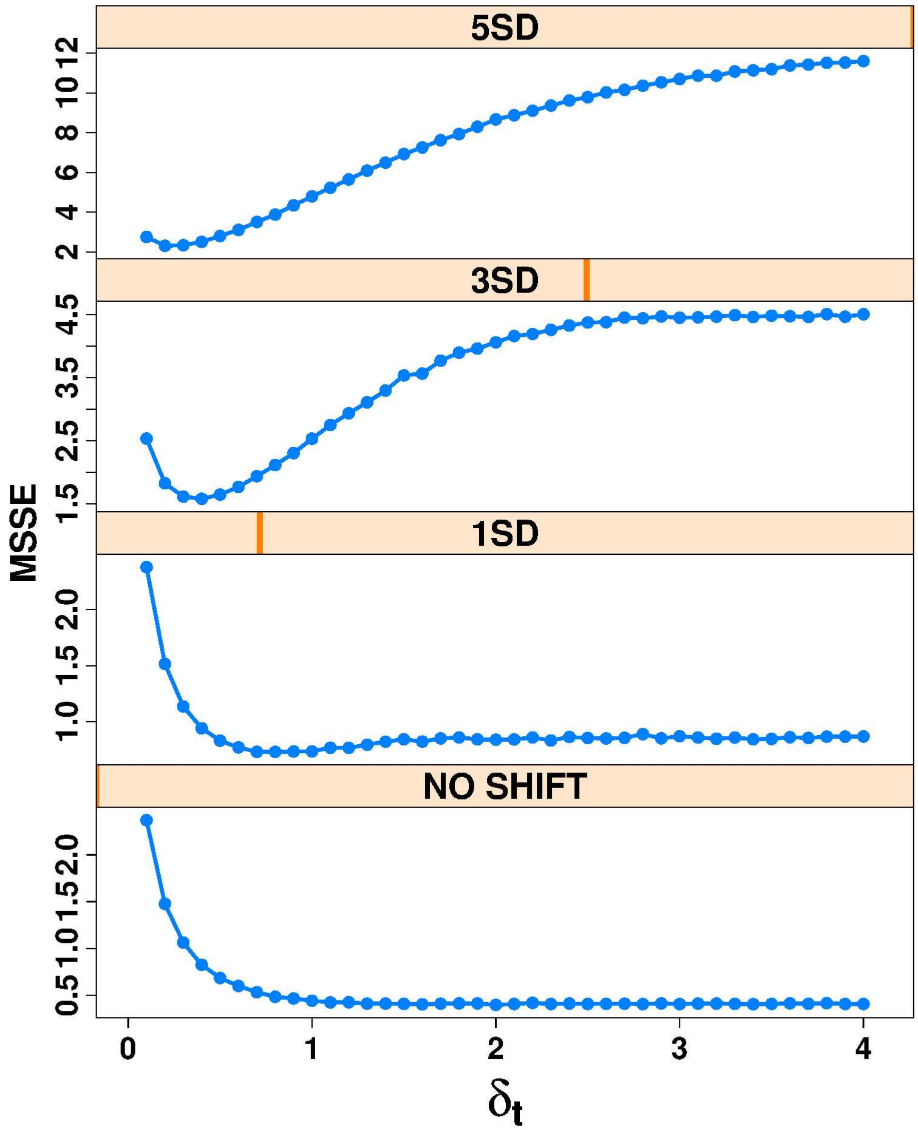
Dr. Pairoj Khawsithiwong



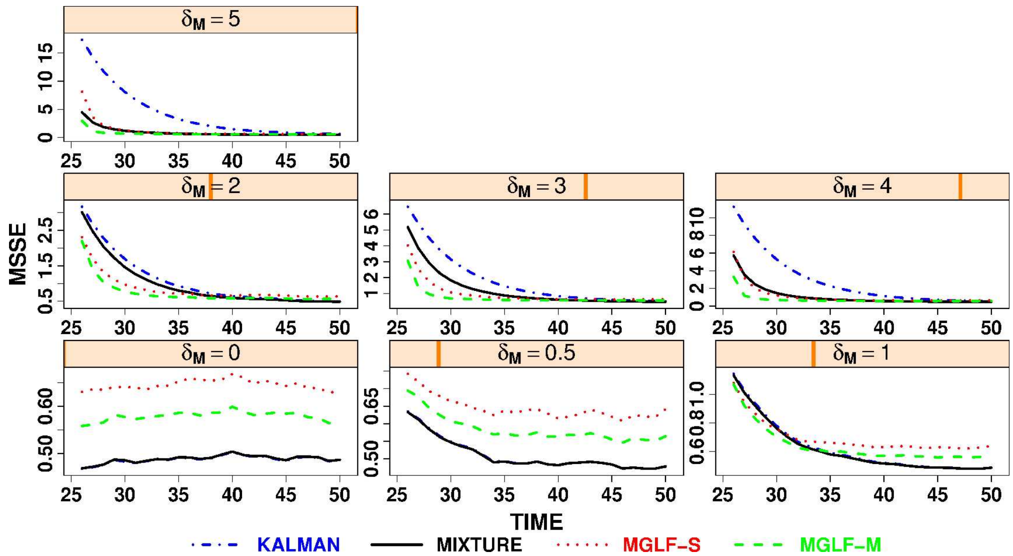
215x214mm (600 x 600 DPI)



288x191mm (600 x 600 DPI)



197x251mm (600 x 600 DPI)



290x158mm (600 x 600 DPI)