TAMPERE UNIVERSITY OF TECHNOLOGY

Author(s) Pesonen, Henri; Piché, Robert

Title Bayesian positioning using Gaussian mixture models with time-varying component weights

Citation Pesonen, Henri; Piché, Robert 2011. Bayesian Positioning Using Gaussian Mixture Models with Time-varying Component Weights. JSM 2011 Joint Statistical Meetings 2011, Miami Beach, Florida, USA, July 30-August 4, 2011. Joint Statistical Meetings JSM Miami Beach, FL, American Statistical Association. 4516-4524.

Year 2011

Version Post-print

URN http://URN.fi/URN:NBN:fi:tty-201311011410

All material supplied via TUT DPub is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorized user.

Bayesian Positioning Using Gaussian Mixture Models with Time-varying Component Weights

Henri Pesonen[*] Robert Piché[†]

Abstract

Gaussian mixture models are often used in target tracking applications to take into account maneuvers in state dynamics or changing levels of observation noise. In this study it is assumed that the measurement or the state transition model can have two plausible candidates, as for example in positioning with line-of-sight or non-line-sight-signals. The plausibility described by the mixture component weight is modeled as a time-dependent random variable and is formulated as a Markov process with a heuristic model based on the Beta distribution. The proposed system can be used to approximate some well-known multiple model systems by tuning the parameter of the state transition distribution for the component weight. The posterior distribution of the state can be solved approximately using a Rao-Blackwellized particle filter. Simulations of GPS pedestrian tracking are used to test the proposed method. The results indicate that the new system is able to find the true models and its root mean square error-performance is comparable to filters that know the true models.

Key Words: Bayesian filtering, multiple model filtering, model uncertainty, Rao-Blackwellization

1. Introduction

Positioning and tracking are often carried out by modeling the problem as discrete time stochastic systems [2]. By modeling the motion of the mobile station (MS) and the relationship of the state (position, velocity, acceleration, etc, ...) and some observable quantities stochastically, we can solve the posterior distribution of the state using Bayesian framework optimally. Furthermore, under certain assumptions the computations can be carried out recursively. However, in positioning systems the chosen models may not be valid at every time step. For example, in satellite positioning system such as global positioning system (GPS) signals obtained from visible or non-visible satellites have different statistical qualities [8]. As another example, MS moving with constant velocity and heading has very different model as opposed to maneuvering target.

Two common solutions to the problem is to use mixture distributions to describe the systems, or describe the system using switching models. In the case of Gaussian linear systems, the posterior distribution can be evaluated analytically with the celebrated Kalman filter [7]. When using Gaussian mixture distributions to describe the stochastic components, Gaussian mixture filter (GMF) evaluates the posterior distribution analytically, at least in theory [10]. GMF gives generally only a theoretical solution to the problem, as it requires an exponentially growing number of mixture components to describe the posterior distribution. In practice, approximations such as multiple model filters (MMFs) are used [2].

Problem when using mixture distributions to describe the models, is that the weights, or the probabilities of the models, are static. However, in general case the probabilities will change with time. For example, often a two-component mixture distribution is used

[*]Tampere University of Technology, Department of Mathematics, Korkeakoulunkatu 1, 33720, Tampere, Finland

[†]Tampere University of Technology, Department of Mathematics, Korkeakoulunkatu 1, 33720, Tampere, Finland

to describe the observation error, with one component describing a 'good' observation and the other describing a 'bad' observation [3, 9]. In the case of GPS positioning, obtaining a bad observation in an urban canyon is more likely than in a forest and much more likely than on a highway. We propose a method where the mixture weight of a two-component mixture distribution is modeled as a stochastic process, and is solved jointly with the state parameters with a Bayesian filter.

The paper is organized as follows. In Section 2 we describe the considered linear discrete time stochastic system, and in Section 3 we present a heuristic motion model for the time-varying mixture component weight based on the Beta-distribution. In Section 4 we formulate the Rao-Blackwellized particle filter (RBPF) for the described system. In Section 5 we test the performance of the filters based on the discussed models and conclude the study in Section 6.

2. Problem formulation

We consider a linear discrete time stochastic system

$$x_{k+1} = F_k x_k + w_k \tag{1}$$

$$y_k = H_k x_k + v_k. \tag{2}$$

$$x_0 \sim \mathsf{N}\left(\bar{x}_{0|0}, P_{0|0}\right), \tag{3}$$

where $\mathsf{N}(\mu, \Sigma)$ is a Gaussian distribution with mean μ and covariance Σ. It is well-known that in the case of white Gaussian noises w_k and v_k the Kalman filter (KF) can be used to compute recursively the Gaussian posterior filtering distribution parameters. However, as well-known is the sensitivity of the KF against unaccounted large observation errors .

A common approach for making the system less sensitive against modeling errors is to assume a set of models from which a member generated the state or the observations. These are often referred to multiple model (MM) systems. We consider the case where there are two possible models, although in general case any number of models could be considered as plausible. The model is

$$\begin{aligned} p(x_{k+1}|x_k) &= \mathsf{N}\left(x_{k+1} \mid F_k x_k, Q_k\right) \\ p(y_k|x_k, \lambda_k) &= \mathsf{N}\left(y_k \mid H_k x_k, (1-\lambda_k)R_k^{(1)} + \lambda_k R_k^{(2)}\right), \end{aligned} \tag{4}$$

where the state model is Gaussian but observation error is modeled a Gaussian mixture (GM) distributed random variable. This could model for example the presence of line-of-sight/non-line-of-sight (LOS/NLOS) signals between MS and a satellite.

The parameter $\lambda_k \in \{0, 1\}$ is the model parameter, and it remains to be defined. In simplest case it would be a Bernoulli distributed random variable independent of previous values

$$\begin{aligned} p(\lambda_k = 0 \mid \lambda_{k-1}) &= p(\lambda_k = 0) = 1 - \epsilon \\ p(\lambda_k = 1 \mid \lambda_{k-1}) &= p(\lambda_k = 1) = \epsilon \end{aligned} \tag{5}$$

The probability ϵ represents the uncertainty of the model. Often the probability of the model depends on the time. One approach to take this into account is consider the multiple model approach for switching models. Now $\lambda_k \in \{0, 1\}$ is a Markov chain with the transition probabilities

$$\begin{aligned} p(\lambda_k = 0 \mid \lambda_{k-1} = 0) &= 1 - p(\lambda_k = 1 \mid \lambda_{k-1} = 0) = 1 - \epsilon^{(1)} \\ p(\lambda_k = 0 \mid \lambda_{k-1} = 1) &= 1 - p(\lambda_k = 1 \mid \lambda_{k-1} = 1) = 1 - \epsilon^{(2)} \end{aligned} \tag{6}$$

This adaptive formulation of the multiple model state-space problem is sometimes called the jump markov linear system (JMLS) [5]. Notice that (5) is a special case of (6) where $\epsilon^{(1)} = \epsilon^{(2)} = \epsilon$. In theory, the posterior distribution can be found analytically for (4) with the Gaussian mixture filter (GMF) [10, 1]. GMF evaluates the posterior distribution using a bank of Kalman filters increasing the number of filters at each time step. Computationally the complexity of GMF increases exponentially with time, and complexity reduction methods are needed [9]. Two main methods are the pruning and the merging, in which some mixture components are removed or merged together, respectively. Monte Carlo simulation-based Rao-Blackwellized Particle Filter (RBPF) is an example of the pruning approach. RBPFs automatically cut of improbable branches of the mixture posterior [5]. Different multiple model filters such as the interacting multiple model filter and the generalized pseudo-Bayesian approaches are popular examples of merging algorithms [2].

In real life applications where the environment has a major impact on the models, the assumption that the switching probability is known is not always very realistic. For example in a target tracking application with GPS data, the environment has a major impact on the data quality because objects such as trees and buildings block visibility to satellites and degrade the data quality significantly. In applications of this nature, the uncertainty which model is generating the data is changing. In these kind of situations, a more realistic model could be one where the model uncertainty at certain time is similar to the model uncertainty within a small time interval. In the next section we describe a state model for the model uncertainty with a continuous probability density.

3. Time-varying model uncertainty variable

We consider a hierarchical model for the MM filter where the switching probability ϵ_k in (5) and (6) is a time-dependent variable. The probability mass function of λ_k is now defined by ϵ_k

$$p(\lambda_k = 0 \mid \epsilon_k) = 1 - p(\lambda_k = 1 \mid \epsilon_k) = 1 - \epsilon_k \tag{7}$$

Constructing a model for the evolution of the model uncertainty parameter we have to take into account that the uncertainty $\epsilon_k \in [0, 1]$. We model the parameters as a Markov process and take the density $p(\epsilon_{k+1}|\epsilon_k)$ to be unimodal, with the mode near to the value of ϵ_k. A probability density fulfilling these criteria would be a Beta density

$$\mathsf{Beta}(\xi \mid \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \xi^{\alpha-1}(1-\xi)^{\beta-1}. \tag{8}$$

The mode and variance of a Beta distributed random variable are

$$\mathsf{mode}\,(\xi) = \frac{\alpha - 1}{\alpha + \beta - 2}, \quad \mathsf{V}\,(\xi) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta-1)}. \tag{9}$$

The Beta density function is unimodal when $\alpha, \beta > 1$. The variance of a Beta distributed random variable depends on α and β, with variance $\to 0$ as $\alpha, \beta \to \infty$.

We use a state-transition density

$$p(\epsilon_{k+1} \mid \epsilon_k, S) = \mathsf{Beta}(\epsilon_{k+1} \mid \epsilon_k(S-2)+1, (1-\epsilon_k)S + 2\epsilon_k - 1), \tag{10}$$

where S is a tuning parameter. The mode and variance of (10) are

$$\mathsf{mode}\,(\epsilon_{k+1} \mid \epsilon_k, S) = \epsilon_k, \quad \mathsf{V}\,(\epsilon_{k+1} \mid \epsilon_k, S) = \frac{(1-\epsilon_k)\epsilon_k}{S-1}. \tag{11}$$

A larger tuning parameter S reduces the variance of $\epsilon_{k+1} \mid \epsilon_k$, and the expected model uncertainty is the previous model uncertainty. The hierarchical state-space model is illustrated by the directed acyclic graph (DAG) in the Figure 2.

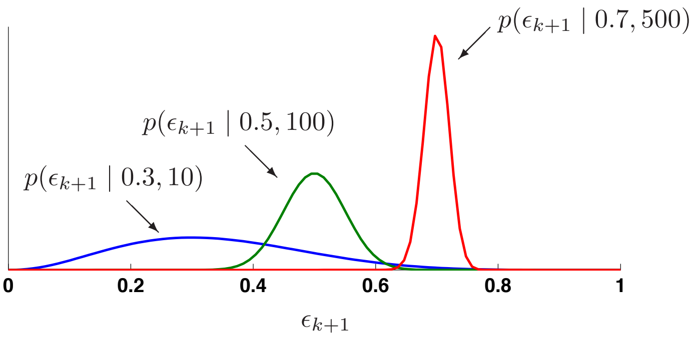

Figure 1: DAG of the hierarchical state-space model

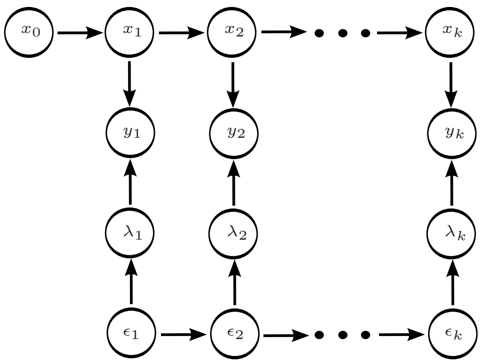

Figure 2: DAG of the hierarchical state-space model

4. Bayesian estimation

Bayesian filtering framework can be used to solve the state-space model with the time-varying model uncertainty variable. The posterior distribution $p(x_k|y_{1:k})$ can in theory be solved recursively [6]. In practice it can be approximated empirically using sequential Monte Carlo methods (SMC) [4]. The empirical representation of the posterior is a sum of N support points $\{(x_{0:k}^{(i)}, \lambda_{0:k}^{(i)}, \epsilon_{0:k}^{(i)}) : i = 1, \ldots, N\}$ with the corresponding weights $\{\omega_{1:k}^{(i)} : i = 1, \ldots, N\}$

$$\widehat{p_N}(x_{0:k}, \lambda_{0:k}, \epsilon_{0:k} \mid y_{1:k}) \approx \sum_{i=1}^{N} \omega_{0:k}^{(i)} \delta\left((x_{0:k}^{(i)}, \lambda_{0:k}^{(i)}, \epsilon_{0:k}^{(i)}) - (x_{0:k}, \lambda_{0:k}, \epsilon_{0:k})\right). \tag{12}$$

The posterior filtering distribution can be obtained as a marginal distribution of (12). The performance of the SMC method can be enhanced in our system by using the fact that conditioned on $\lambda_{0:k}$, problem (4) can be solved optimally with the Kalman filter algorithm. This enables us to approximate the posterior using RBPF, where we need only approximate empirically $p(\epsilon_{0:k}, \lambda_{0:k} \mid y_{1:k})$. If we write

$$\begin{aligned} p(x_k \mid y_{1:k}) &= \int_{[0,1]^{k+1}} \sum_{j=1}^{2^{k+1}} p(x_k \mid y_{1:k}, \epsilon_{0:k}, \lambda_{0:k}^{(j)}) p(\epsilon_{0:k}, \lambda_{0:k}^{(j)} \mid y_{1:k}) \mathrm{d}\epsilon_{0:k} \\ &= \int_{[0,1]^{k+1}} \sum_{j=1}^{2^{k}} p(x_k \mid y_{1:k}, \epsilon_{0:k}, \lambda_{0:k}^{(j)}) \frac{p(\epsilon_{0:k}, \lambda_{0:k}^{(j)} \mid y_{1:k})}{\pi(\epsilon_{0:k}, \lambda_{0:k}^{(j)} \mid y_{1:k})} \pi(\epsilon_{0:k}, \lambda_{0:k}^{(j)} \mid y_{1:k}) \mathrm{d}\epsilon_{0:k}, \end{aligned} \tag{13}$$

then the posterior can be approximated based on the strong law of large numbers as

$$\begin{aligned}
&p(x_k \mid y_{1:k}) \\
&\approx \sum_{i=1}^{N} p(x_k \mid y_{1:k}, \epsilon_{0:k}, \lambda_{0:k}) \frac{p(\epsilon_{0:k}, \lambda_{0:k} \mid y_{1:k})}{\pi(\epsilon_{0:k}, \lambda_{0:k} \mid y_{1:k})} \delta\left((\epsilon_{0:k}^{(i)}, \lambda_{0:k}^{(i)}) - (\epsilon_{0:k}, \lambda_{0:k})\right) \\
&= \sum_{i=1}^{N} p(x_k \mid y_{1:k}, \lambda_{0:k}^{(i)}) \frac{p(\epsilon_{0:k}^{(i)}, \lambda_{0:k}^{(i)} \mid y_{1:k})}{\pi(\epsilon_{0:k}^{(i)}, \lambda_{0:k}^{(i)} \mid y_{1:k})}
\end{aligned} \tag{14}$$

where $\{\epsilon_{0:k}^{(i)}, \lambda_{0:k}^{(i)} : i = 1, \ldots, N\}$ is a sample drawn from the importance sampling distribution $\pi(\epsilon_{0:k}, \lambda_{0:k} \mid y_{1:k})$. In sequential importance sampling (SIS), potential importance distributions are restricted to form

$$\pi(\lambda_{0:k}, \epsilon_{0:k} \mid y_{1:k}) = \pi(\lambda_0, \epsilon_0) \prod_{i=1}^{k} \pi(\lambda_i, \epsilon_i \mid y_{1:i}, \lambda_{0:i-1}, \epsilon_{0:1:i-1}). \tag{15}$$

For which $\pi(\lambda_{0:k-1}, \epsilon_{0:k-1} \mid y_{1:k-1})$ is a marginal distribution at time $k-1$. This enables a sample set from $\pi(\lambda_k, \epsilon_k \mid y_{1:k})$ to be estimated by replacing the samples from $\pi(\lambda_{k-1}, \epsilon_{k-1} \mid y_{1:k-1})$ with a sample from $\pi(\lambda_k, \epsilon_k \mid y_{1:k}, \lambda_{0:k-1}^{(i)}, \epsilon_{0:1:k-1}^{(i)})$. Because we can write

$$\begin{aligned}
&p(\epsilon_{0:k}, \lambda_{0:k} \mid y_{1:k}) = \frac{p(y_k \mid \epsilon_{0:k}, \lambda_{0:k}, y_{1:k-1}) p(\epsilon_{0:k}, \lambda_{0:k} \mid y_{1:k-1})}{p(y_k \mid y_{1:k-1})} \\
&\propto p(y_k \mid \lambda_{0:k}, y_{1:k-1}) p(\epsilon_k, \lambda_k \mid \epsilon_{0:k-1}, \lambda_{1:k-1}, y_{1:k-1}) p(\epsilon_{0:k-1}, \lambda_{0:k-1} \mid y_{1:k-1}) \\
&= p(y_k \mid \lambda_{0:k}, y_{1:k-1}) p(\lambda_k \mid \epsilon_k, \epsilon_{0:k-1}, \lambda_{0:k-1}, y_{1:k-1}) \times \\
&\times p(\epsilon_k \mid \epsilon_{0:k-1}, \lambda_{0:k-1}, y_{1:k-1}) p(\epsilon_{0:k-1}, \lambda_{0:k-1} \mid y_{1:k-1}) \\
&= p(y_k \mid \lambda_{0:k}, y_{1:k-1}) p(\lambda_k \mid \epsilon_k) p(\epsilon_k \mid \epsilon_{k-1}) p(\epsilon_{0:k-1}, \lambda_{0:k-1} \mid y_{1:k-1}),
\end{aligned} \tag{16}$$

we have a recursive update formula for the empirical distribution weights

$$\begin{aligned}
&\omega_{0:k} \triangleq \frac{p(\epsilon_{0:k}, \lambda_{0:k} \mid y_{1:k})}{\pi(\epsilon_{0:k}, \lambda_{0:k} \mid y_{1:k})} \\
&\propto \frac{p(y_k \mid \lambda_{0:k}, y_{1:k-1}) p(\lambda_k \mid \epsilon_k) p(\epsilon_k \mid \epsilon_{k-1}) p(\epsilon_{0:k-1}, \lambda_{0:k-1} \mid y_{1:k-1})}{\pi(\lambda_0, \epsilon_0) \prod_{i=1}^{k} \pi(\lambda_i, \epsilon_i \mid y_{1:i}, \lambda_{0:i-1}, \epsilon_{0:i-1})} \\
&= \frac{p(y_k \mid \lambda_{0:k}, y_{1:k-1}) p(\lambda_k \mid \epsilon_k) p(\epsilon_k \mid \epsilon_{k-1})}{\pi(\lambda_k, \epsilon_k \mid y_{1:k}, \lambda_{0:k-1}, \epsilon_{0:k-1})} \\
&\times \frac{p(\epsilon_{0:k-1}, \lambda_{0:k-1} \mid y_{1:k-1})}{\pi(\lambda_0, \epsilon_0) \prod_{i=1}^{k-1} \pi(\lambda_i, \epsilon_i \mid y_{1:i}, \lambda_{0:i-1}, \epsilon_{0:i-1})} \\
&= \frac{p(y_k \mid \lambda_{0:k}, y_{1:k-1}) p(\lambda_k \mid \epsilon_k) p(\epsilon_k \mid \epsilon_{k-1})}{\pi(\lambda_k, \epsilon_k \mid y_{1:k}, \lambda_{0:k-1}, \epsilon_{0:k-1})} \times \omega_{0:k-1}.
\end{aligned} \tag{17}$$

In practice, after few time steps, all but one particle will have nonzero weight. A procedure called resampling is required [5].

5. Tests

We test the introduced method in target tracking application, where we are able to observe horizontal coordinates MS at each time step, for example from a GPS receiver. The state consists of two-dimensional position x_k^{pos} and velocity vectors x_k^{vel}

$$x_k = \begin{bmatrix} x_k^{\text{pos}T} & x_k^{\text{vel}T} \end{bmatrix}.$$

The problem is modeled using a state-space model (1) – (3), with

$$F_k = \begin{bmatrix} I_2 & I_2 \\ \mathbf{0}_2 & I_2 \end{bmatrix}, \quad H_k = \begin{bmatrix} I_2 & \mathbf{0}_2 \end{bmatrix}. \tag{18}$$

I_n and $\mathbf{0}_n$ denote $n \times n$ identity and null matrices respectively. We use the constant velocity model

$$\mathsf{E}\left(w_k\right) = 0, \quad \mathsf{V}\left(w_k\right) = Q_k = 0.1^2 \cdot \begin{bmatrix} \frac{1}{3} I_2 & \frac{1}{2} I_2 \\ \frac{1}{2} I_2 & I_2 \end{bmatrix} \tag{19}$$

to describe the target motion [2]. As the observation model, we use the mixture distribution

$$p(y_k \mid x_k, \lambda_k) = \mathsf{N}\left(y_k \mid H_k x_k, (1-\lambda_k) 5^2 I_2 + \lambda_k 25^2 I_2\right) \tag{20}$$

We use 100 runs of three different tests to test the performance of the introduced model. The evolution of the uncertainty parameter with time is changed with each test. This is illustrated in the Figure 3. In each test the averaged uncertainty is $\frac{1}{k_{\max}} \sum_{k=1}^{k_{\max}} \epsilon_k = 0.5$.

The RBPF algorithms with 50 particles using different models are compared. GMFs with model (5) and $\epsilon = 0.5$ (GMF$_1$) and model (6) with $\epsilon^{(1)} = 1 - \epsilon^{(2)} = 0.02$ (GMF$_2$) are compared to GMF employing the introduced hierarchical model for time-evolution of the model uncertainty pararameter ϵ_k (GMF$_3$). Tuning parameter $S = 100$ is fixed in all the tests.

The performance of the algorithms is compared through the average ability to estimate the parameter ϵ_k and root mean square error (RMSE) of the position coordinate estimates.

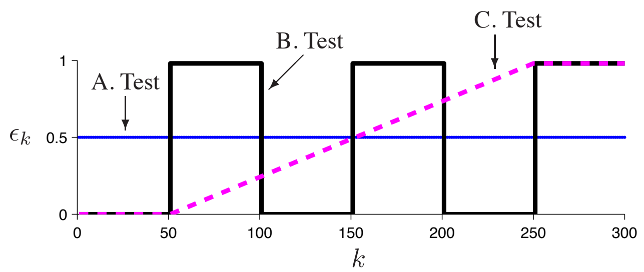

Figure 3: The evolution of the model uncertainty parameter with time in different tests.

A. Test

Observations are simulated from (20) with constant probability $p(\lambda_k = 0 \mid \epsilon_k) = 0.5$. GMF$_1$ is the filter based on the correct model. The results of the tests are reported in Figure 4. The RMSE performances of GMF$_1$ and GMF$_3$ are virtually identical, as is the estimations of ϵ_k. GMF$_2$ has the worst performance due to the small switching probabilities of λ_k.

B. Test

Observations are simulated from (20) with three changes of $\lambda_k = 0$ to $\lambda_{k+1} = 1$ and three changes of $\lambda_k = 1$ to $\lambda_{k+1} = 0$. GMF$_2$ is the filter based on model closest to the simulation model. The results of the tests are reported in Figure 5. The RMSE performance and the estimation accuracy of ϵ_k given by GMF$_2$ is the best. GMF$_3$ has an improved performance compared to GMF$_1$. The effect of the transition model for ϵ_k in GMF$_3$ algorithm can be seen in the estimation of ϵ_k as well as in the decrease of RMSE in time steps closer to the change points $k_c \in \{100, 200, 300\}$.

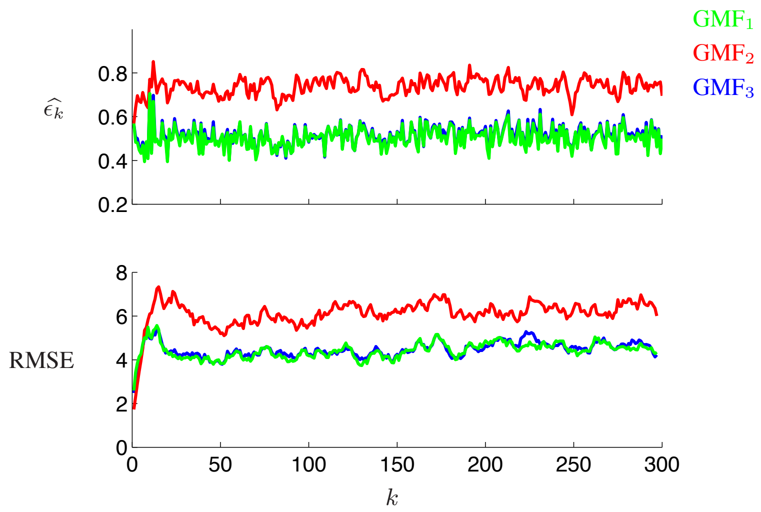

Figure 4: Results of A. Test.

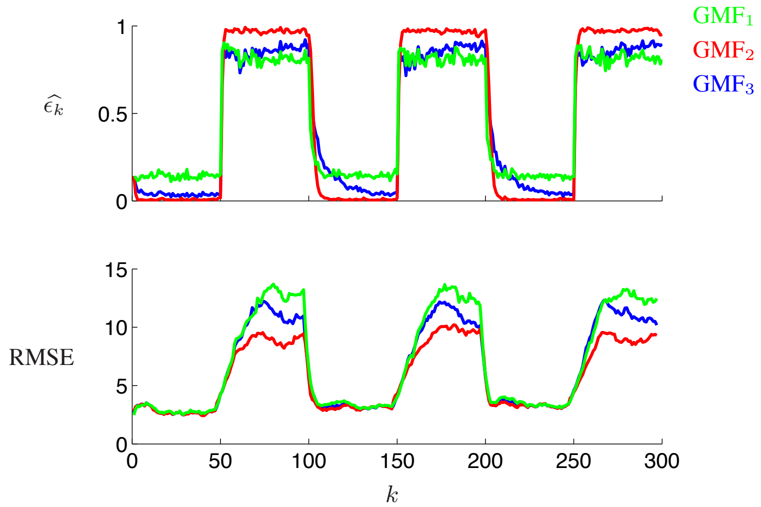

Figure 5: Results of B. Test.

C. Test

Observations are simulated from (20) such that there is a linear increase in the probability of $\lambda_k = 1$ in the interval $k \in [50, 250]$. The overall RMSE performance of GMF_3 is the best as is the estimation capability of ϵ_k. GMF_2 perform well when ϵ_k is close to 0 or 1 but has degraded performance otherwise. This is consistent with its performance in the previous tests.

6. Conclusions

A heuristic hierarchical model for the time-evolution of the model uncertainty parameter has been constructed and a RBPF-based algorithm for solving the resulting problem has been provided. Through simulations we showed that the new proposed method is able to approximate well the uncertainty parameter and it has RMSE performance comparable to

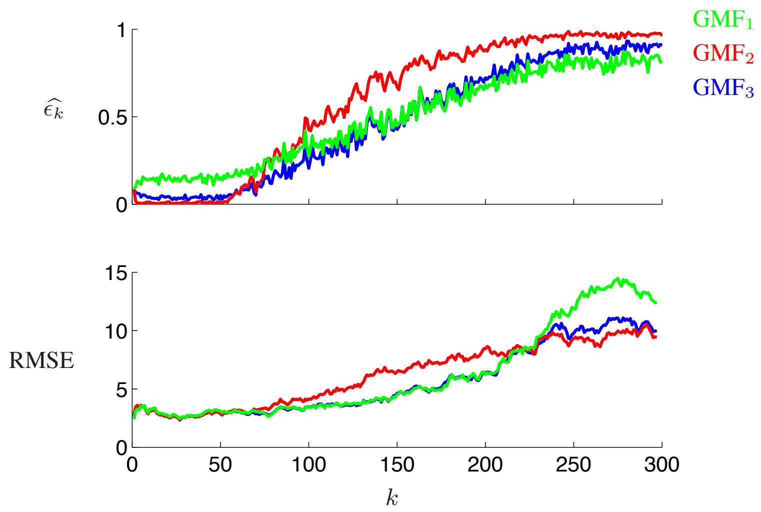

Figure 6: Results of C. Test.

the situations where an optimal model would be used to approximate the state.

It is expected that employing this additional level of hierarchy will be beneficial for applications where there are several models from which other are more plausible at certain times. Further study is required to expand the algorithm to handle more than two competing models.

References

[1] D. L. Alspach and H. W. Sorenson. Nonlinear Bayesian estimation using Gaussian sum approximations. *IEEE Transactions on Automatic Control*, AC-17(4):439–448, August 1972.

[2] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., 2001.

[3] G.E.P. Box and G.C. Tiao. A Bayesian approach to some outlier problems. *Biometrika*, 55(1):119–129, 1968.

[4] A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag New York, Inc., 2001.

[5] A. Doucet, N. J. Gordon, and V. Krishnamurthy. Particle filters for state estimation of jump Markov linear systems. *IEEE Transactions on Signal Processing*, 49(3):2001, March 2001.

[6] A. H. Jazwinski. *Stochastic Processes And Filtering Theory*. Academic Press, Inc., 1970.

[7] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering*, 82, 1960.

[8] J. Marais and B. Godefroy. Analysis and optimal use of GNSS pseudo-range delays in urban canyons. In *Proceedings of IMACS Multiconference on Computational Engineering in System Applications (CESA), October 4-6, 2006, Beijing, China*, 2006.

[9] D. Peña and I. Guttman. Optimal collapsing of mixture distributions in robust recursive estimation. *Communications in Statistics: Theory and Methods*, 18(3):817–833, 1989.

[10] H.W. Sorenson and D.L. Alspach. Recursive Bayesian estimation using Gaussian sums. *Automatica*, 7:465–479, 1971.