School of Mathematics, Meteorology
and Physics

Department of Mathematics

Preprint MPS_2010_18

21 April 2010

State Estimation using the Particle Filter with Mode Tracking

by

J.A. Pocock, S.L. Dance, A.S. Lawless

A particle filter (PF) is a data assimilation scheme that employs a fully

Vaswani [16] suggests that the best results should be obtained from the particle filter with mode tracking when we mode track the maximum number of unimodal dimensions. In this paper we test this using the 3D stochastic Lorenz equations with a linear observation operator.

We begin, in section 2, by describing the algorithm for the particle filter with mode-tracking. We also consider how the state should be split to provide the best results from the particle filter with mode tracking. The simple nonlinear stochastic model used for our experiments is described in section 3. In section 4 we test one hypothesis for how to split the state to obtain the best results for the particle filter with mode tracking. We conclude in section 5 by summarising and discussing the main results.

2. The Particle Filter with Mode Tracking

In this section we consider the particle filter with mode tracking (PF-MT). We start by defining the notation used. Let $\psi_k \in \mathbb{R}^n$, $k = 0, 1, 2, 3, \ldots$, be a sequence of model states at discrete times k and assume the initial pdf of the state is given by $p(\psi_0)$. At subsequent times, ψ_k satisfies

$$\psi_k = M(\psi_{k-1}, w_k), \tag{1}$$

where $M : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is a possibly nonlinear function, and $w_k \in \mathbb{R}^n$ is a noise process sequence. This equation describes a Markov process with transition density $p(\psi_k|\psi_{k-1})$.

Let $d_k \in \mathbb{R}^p$ be the observation vector at time k which is related to the model state by the equation,

$$d_k = H(\psi_k, v_k), \tag{2}$$

Table 1: The PF-MT Algorithm

The PF-MT Algorithm [16]

Initialization:

- Set $k = 0$ and sample N times from the importance function $\pi(\psi_0)$ to give the initial ensemble $\{\psi_0^i\}_{i=1}^N$.
- Set the weights, $w_0^i = 1/N$.

For times $k = 1, 2, ...$

1. Importance Sample $\Psi_{k,s}$: For $i = 1, 2, \ldots, N$, sample

$$\Psi_{k,s}^i \sim p(\Psi_{k,s}^i | \psi_{k-1}^i)$$

.

2. Mode track $\Psi_{k,r}$: For $i = 1, 2, \ldots, N$, set $\Psi_{k,r} = m_k^i$ where

$$m_k^i(\psi_{k-1}^i, \Psi_{k,s}^i, d_k) = \arg\min_{\Psi_{k,r}}[-\log p(d_k|\Psi_{k,s}^i)p(\Psi_k$$ $k-$ i Ψ ks63]TJ-40.630.147034]TJ-20

part of the subspace. The conditional pdf for $\Psi_{k,r}$ may be written as

$$p(\Psi_{k,r}|\psi^i_{k-1}, \Psi^i_{k,s}, d_k) \propto p(d_k|\psi^i_{k-1}, \Psi^i_{k,s})p(\Psi_{k,r}|\psi^i_{k-1}, \Psi^i_{k,s}), \tag{8}$$

using Bayes' rule, since ψ_k is a Markov process and the observations are conditionally independent of the model state [16]. Following [16] we let

$$\begin{aligned} J^i(\Psi^i_{t,s}, \Psi_{t,r}) &\stackrel{def}{=} -\log p(\Psi_{k,r}|\psi^i_{k-1}, \Psi^i_{k,s}, d_k) \qquad (9) \\ &= -\log p(d_k|\psi^i_{k-1}, \Psi^i_{k,s}) \\ &\quad -\log p(\Psi_{k,r}|\psi^i_{k-1}, \Psi^i_{k,s}) + \text{const.} \qquad (10) \end{aligned}$$

If the pdf is unimodal we can set the constant term to zero and find the mode by minimizing the cost function J^i with respect to $\Psi_{t,r}$. The cost function for the example used in this paper follows on from (10) and can be written as

$$J^i(\Psi^i_{k,s}, \Psi_{k,r}) = \frac{1}{2}\left[J_o(\Psi^i_{k,s}, \Psi_{k,r}) + J_q(\Psi_{k,r})\right], \tag{11}$$

where, using (7),

$$\begin{aligned} J^i_o(\Psi^i_{k,s}, \Psi_{k,r}) &= -\log p(d_k|\psi^i_{k-1}, \Psi^i_{k,s}) \\ &= \Big(d_k - {}^i \quad] \quad \Big)\Big) \end{aligned}$$

weights. There are many resampling algorithms. Here we use a resampling scheme known as stratified resampling [8] as it is efficient and simple to implement [1]. We return to the forecast step to continue the assimilation cycle. We repeat the iteration until the final forecast time is reached.

We must now consider how to choose the state-space splitting. For the most efficient algorithm, $\Psi_{k,r}$ ostensibly should contain the maximum number of dimensions such that $p(\Psi_{k,r}|\psi^i_{k-1}, \Psi^i_{k,s}, d_k)$ is unimodal, so that the subspace that is modelled by the standard PF is as small as possible [16]. In principle, the unimodality of the mode-tracking subspace might change every time k and for each ensemble member. Vaswani [16] argues that the method should be successful if unimodality holds for most particles at most times. For our case the full pdf of the statim(3)(u)f8509(t)26.9419(y)-260.121(o)0.245057(f)-260.166(tf

when using the PF-MT. We tested the hypothesis that the best results should be obtained from the PF-MT when we mode track the maximum number of unimodal dimensions. When using a nonlinear model it was found that the best results from the PF-MT were not always obtained when the maximum number of unimodal dimensions was mode tracked. This was possibly due to the complicated nonlinear structure of the full problem. It was found that the best states to mode trackr5(t)-0.147034(o)0.245057]T(u)-0.447(l)0(7)0.245057(7(n653t530.

[8] G. Kitagawa. Monte carlo filter and smoother for non-gaussian nonlinear state space models. *o n o Co p on nd p c c*, 5:1–25, 1996.

[9] P. E. Kloden and E. Platen. *N e c o on o oc c D e en q on* . Springer, 1999.

[10] R. N. Miller, E. F. Carter, and S. T. Blue. Data assimilation into nonlinear stochastic models. *e* , 51A:167–194, 1999.

[11] J. A. Pocock. Ensemble data assimilation: How many members do we need? Master's thesis, Department of Mathematics, University of Reading, 2009.

[12] H. Salman. A hybrid grid/particle filter for Lagrangian data assimila-