UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Advanced Data Assimilation
in strongly nonlinear systems

Youmin Tang and
Jaison Thomas Ambadan

Environmental Science & Engineering Program
Natural Resources & Environmental Science

Data assimilation workshop at Banff, Feb 3-8

8 February 2008 1

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Outline

- Introduction
- Sigma-point Kalman filters
- Simulations with Lorenz model
- A Reduced SPKF
- Summary & Conclusion.

8 February 2008 2

Introduction

Data assimilation – a dynamical state space estimate problem

State equation $x_k = f(x_{k-1}, u_k);$ (1)

Observation equation $y_k = h(x_k, v_k)$ (2)

where x_k is state vector at time *k*, *f* state transition function, and u_k process noise with known distribution; y_k is observations at time instant *k*, *h* observation function, and v_k observation noise with known distribution.

Statistical estimation

KF & EKF

Analysis step

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - E(h(\hat{x}_k^-, v_k))),$$

$$P_{x_k} = (I - K_k H_k) P_{x_k}^-$$

$$K_k = P_{x_k}^- H_k^T [R_k + H_k P_{x_k}^- H^T]^{-1}$$

The forecast step

$$\hat{x}_k^- = E[f(\hat{x}_{k-1}, u_{k-1})],$$

$$P_{x_k}^- = L_{k-1} P_{x_{k-1}} L_{k-1}^T + Q_{k-1}$$

$$Q_k = E[(u_k - \bar{u}_k)(u_k - \bar{u}_k)^T],$$

$$R_k = E[(v_k - \bar{v}_k)(v_k - \bar{v}_k)^T]$$

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Statistical estimation

EnKF--- Popular approach

Analysis step

$$\hat{\boldsymbol{x}}_{\boldsymbol{k}} = \hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-} + \boldsymbol{K}_{\boldsymbol{k}}(\boldsymbol{y}_{\boldsymbol{k}} - \boldsymbol{E}(\boldsymbol{h}(\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-}, \boldsymbol{v}_{\boldsymbol{k}}))),$$

$$K_k = P_{x_k}^{-} H_k^T [R_k + H_k P_{x_k}^{-} H^T]^{-1}$$

The forecast step

$$\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-} = \boldsymbol{E}[\boldsymbol{f}(\hat{\boldsymbol{x}}_{\boldsymbol{k}-1}, \boldsymbol{u}_{\boldsymbol{k}-1})],$$

$$\boldsymbol{P}_{\boldsymbol{x}_{\boldsymbol{k}}}^{-} = \boldsymbol{E}[(\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-} - \bar{\boldsymbol{x}})(\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-} - \bar{\boldsymbol{x}})^{\boldsymbol{T}}]$$

Linear assumption for measurement function.

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Statistical estimation

EnKF--- classic approach

Analysis step

$$\hat{\boldsymbol{x}}_{\boldsymbol{k}} = \hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-} + \boldsymbol{K}_{\boldsymbol{k}}(\boldsymbol{y}_{\boldsymbol{k}} - \boldsymbol{E}(\boldsymbol{h}(\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-}, \boldsymbol{v}_{\boldsymbol{k}}))),$$

$$P_{x_k} = P_{x_k}^{-} - K_k P_{y_k} K_k^T$$

$$K_k = P_{x_k y_k} P_{y_k}^{-1}$$

The forecast step

$$\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-} = \boldsymbol{E}[\boldsymbol{f}(\hat{\boldsymbol{x}}_{\boldsymbol{k}-1}, \boldsymbol{u}_{\boldsymbol{k}-1})],$$

$$\boldsymbol{P}_{\boldsymbol{x}_{\boldsymbol{k}}\boldsymbol{y}_{\boldsymbol{k}}} = \boldsymbol{E}[(\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-} - \bar{\boldsymbol{x}})(\boldsymbol{y} - \boldsymbol{E}(\boldsymbol{h}(\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-}, \boldsymbol{v}_{\boldsymbol{k}})))^{\boldsymbol{T}}]$$

$$\boldsymbol{P}_{\boldsymbol{y}_{\boldsymbol{k}}} = \boldsymbol{E}[(\boldsymbol{y} - \boldsymbol{E}(\boldsymbol{h}(\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-}, \boldsymbol{v}_{\boldsymbol{k}})))(\boldsymbol{y} - \boldsymbol{E}(\boldsymbol{h}(\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-}, \boldsymbol{v}_{\boldsymbol{k}})))^{\boldsymbol{T}}]$$

No linear assumption for measurement function.

Ref: (Gelb, 1974)

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Sigma-point Kalman filter

- **The SPKF makes use of a reformulated Kalman gain K and "chooses" the ensemble deterministically in such a way that it can capture the statistical moments of the nonlinear model accurately; in other words, the forecast error covariance equation is computed using deterministically chosen samples, called "sigma-points".**
- **The SPKF algorithm has been successfully implemented in many areas like robotics, artificial intelligence, natural language processing, and global positioning systems navigation.**

Ref: (Julier et al. 1995; Nørgad Magnus et al. 2000; Ito and Xiong 2000; Lefebvre et al. 2002;Wan and Van Der Merve 2000; Haykin 2001; Van der Merwe 2004, Van der Merwe and Wan4 April 2001,M;).

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

SP-Unscented Kalman filter

Unscented transformation (Julier et al. 1995; Julier 1998; Wan and Van Der Merve 2000; Julier 2002**).**

Consider the propagation pf a L-dimensional random variable x through an arbitrary nonlinear function:

$$\varphi = \boldsymbol{g}(\boldsymbol{x});$$

$$\boldsymbol{x} \text{ has mean } \bar{\boldsymbol{x}} \text{ and covariance } \boldsymbol{P_x}$$

$$S_i = \{w_i, \chi_i\}; \qquad w_0 = \frac{k}{L+k} \quad i = 0$$

$$\chi_0 = \bar{x};$$

$$\chi_i = \bar{x} + (\sqrt{(L+k)P_x})_i, \qquad w_i = \frac{1}{2(L+k)} \quad i = 1,\ldots,2L$$

$$\chi_i = \bar{x} - (\sqrt{(L+k)P_x})_i \qquad k \text{ is a scaling parameter}$$

$$\psi_i = \boldsymbol{g}(\chi_i);$$

the dimension of the state-space L increases, the radius of the sphere that bounds all the sigma-points increases as well.

UNIVERSITY OF NORTHERN BRITISH COLUMBIA

SP-UKF Algorithm

Analysis step

$$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_k(\boldsymbol{y}_k - \boldsymbol{h}(\hat{\boldsymbol{x}}_k^-)),$$

$$P_{x_k} = P_{x_k}^- - K_k P_{y_k} K_k^T$$

$$K_k = P_{x_k y_k} P_{y_k}^{-1}$$

Forecast step

$$\boldsymbol{\psi}_k^i = \boldsymbol{f}(\boldsymbol{\chi}_k^i),$$

$$\hat{\boldsymbol{x}}_k^- = \sum_{i=0}^{2L} \boldsymbol{w}^i \boldsymbol{\psi}_k^i$$

$$\boldsymbol{P}_{x_k}^- = \sum_{i=0}^{2L} \boldsymbol{w}^i (\boldsymbol{\psi}_k^i - \hat{\boldsymbol{x}}_k^-)(\boldsymbol{\psi}_k^i - \hat{\boldsymbol{x}}_k^-)^T$$

Measurement step

$$\boldsymbol{Y}_k^i = \boldsymbol{h}(\boldsymbol{\chi}_k^i),$$

$$\hat{\boldsymbol{y}}_k^- = \sum_{i=0}^{2L} \boldsymbol{w}^i \boldsymbol{Y}_k^i$$

$$\boldsymbol{P}_{y_k} = \sum_{i=0}^{2L} \boldsymbol{w}^i (\boldsymbol{Y}_k^i - \hat{\boldsymbol{y}}_k^-)(\boldsymbol{Y}_k^i - \hat{\boldsymbol{y}}_k^-)^T$$

$$\boldsymbol{P}_{x_k y_k} = \sum_{i=0}^{2L} \boldsymbol{w}^i (\boldsymbol{\psi}_k^i - \hat{\boldsymbol{x}}_k^-)(\boldsymbol{Y}_k^i - \hat{\boldsymbol{y}}_k^-)^T$$

UNIVERSITY OF NORTHERN BRITISH COLUMBIA

SP-Central Difference KF

- **In SP-CDKF the analytical derivatives in EKF are replaced by numerically evaluated central divided differences, based on Sterling's polynomial interpolation.**

(Ito and Xiong 2000; Nørga°d Magnus et al. 2000; Lefebvre et al. 2002).

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Lorenz model

$$\frac{dx}{dt} = \sigma(y - x) + q^x$$
$$\frac{dy}{dt} = \rho x - y - xz + q^y$$
$$\frac{dz}{dt} = xy - \beta z + q^z$$

True value: integrate the model over 4000 time steps using prescribed parameters and initial conditions.

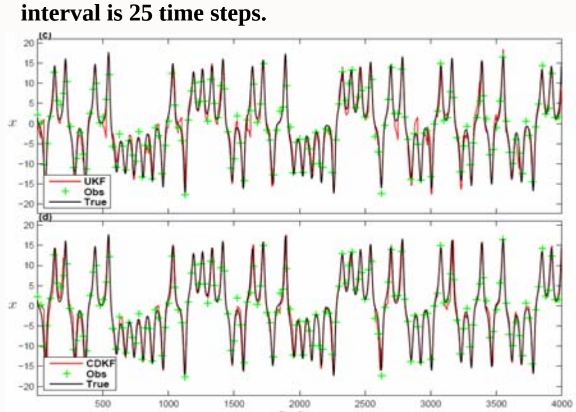
Observation: true value plus normal distribute noise;

Experimental conditions are the same as those used by Miller (Miller, 1994) and Evensen (Evensen 1997)

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

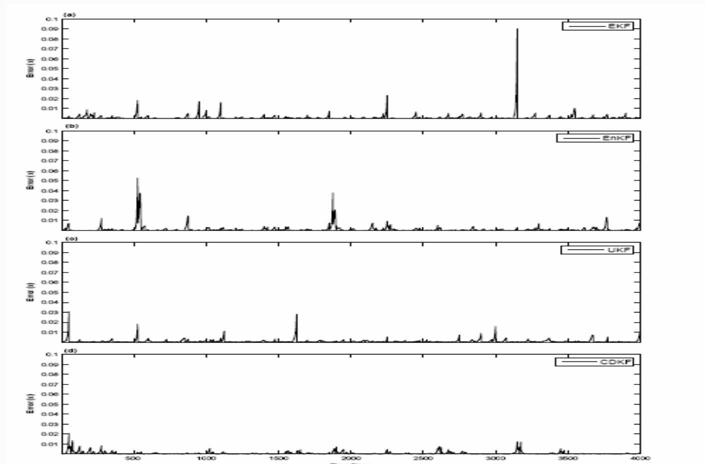
Lorenz model: state estimation

Observation and initial conditions: their true values plus normal distributed noise $N(0,\sqrt{2})$. The assimilation interval is 25 time steps.



UNIVERSITY OF NORTHERN BRITISH COLUMBIA

$$Error = \frac{1}{N}(x_k - x_k^{true})^2$$



UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Introduction

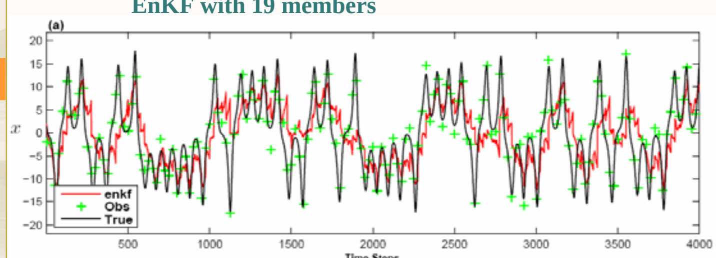
Sigma-point Kalman filters

Simulations with Lorenz model

A Reduced SPKF

Summary & Conclusion

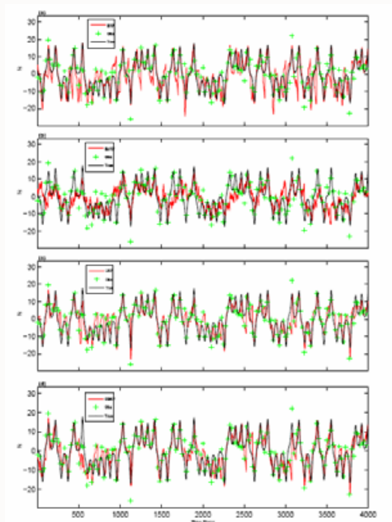
EnKF with 19 members



Assimilation Method	Computation Time (in Seconds)	RMSE
EKF	37.04	1.812
EnKF (with 1000 ensembles)	7143.57	1.987
EnKF (with 19 ensembles)	132.77	6.123
SP-UKF	133.91	1.640
SP-CDKF	90.42	1.592

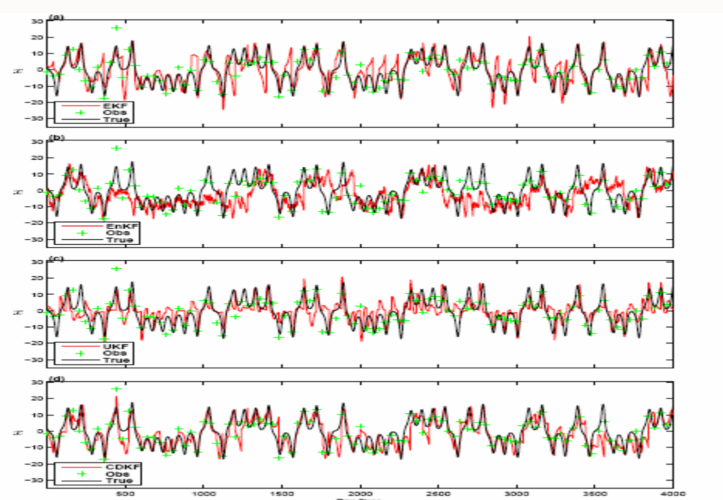
UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Stronger noise (10 times)



UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Strong noise + Less observations



UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Parameter estimation

- **If the measurement function is nonlinear, it has to be linearized in the EnKF.**

$$\Lambda_k = \Lambda_{k-1} + q_{k-1}$$
$$y_k = f(x_k, \Lambda_k) + r_k$$

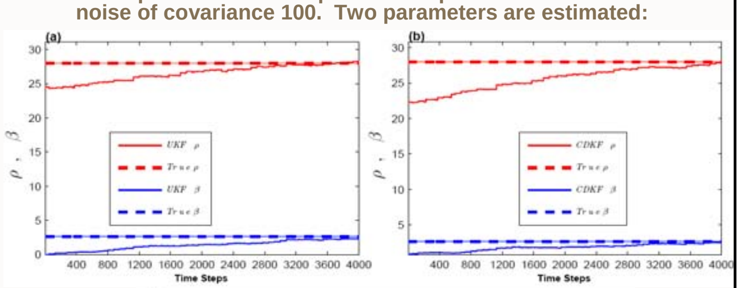
where $f(\cdot)$ is the nonlinear measurement model given by the Lorenz equations, and Λ is the parameter vector which constitutes the dynamical parameters σ, ρ and β

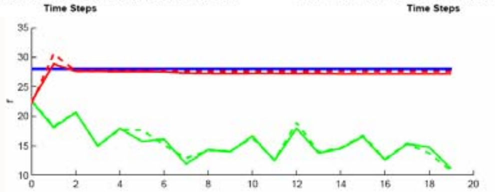
8 February 2008 17

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Parameter estimation

Initial parameters: true parameters plus normal distributed noise of covariance 100. Two parameters are estimated:



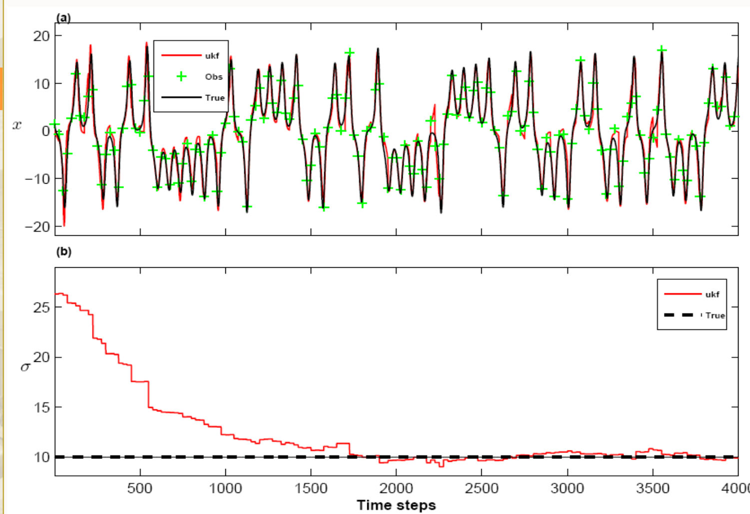


Kivman, G.A. 2003: Sequential parameter estimation for stochastic system, Nonlinear. Process. in Geophysics, 10, 253-259.

8 February 2008 18

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Joint estimation (state +parameters)



UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

A “big” drawback of SPKF

- For an L-dimensional system, the number of sigma-points required to estimate the true statistics is 2L+1

- 2L+1 sigma-point integration is impossible, when the dimension system is of the order of tens of millions as it happens often in GCMs

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

A Possible solution – reducing sigma-points

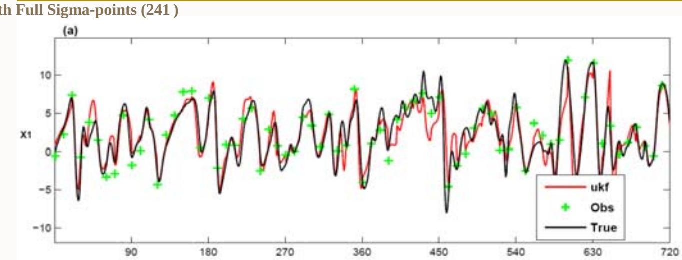
$$S_i = \{w_i, \chi_i\};$$
$$\chi_0 = \bar{x};$$
$$\chi_i = \bar{x} + (\sqrt{(L+k)P_x})_i,$$
$$\chi_i = \bar{x} - (\sqrt{(L+k)P_x})_i$$

8 February 2008 21

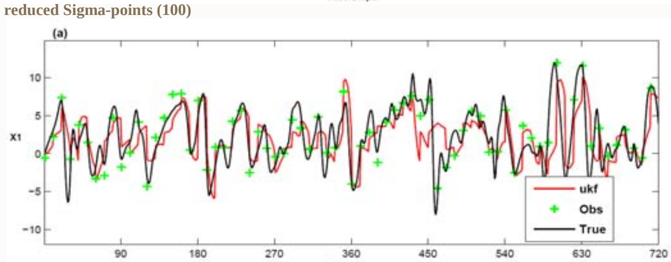
UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

A Reduced Sigma-point KF - (Lorenz '96 model)

with Full Sigma-points (241)



with reduced Sigma-points (100)

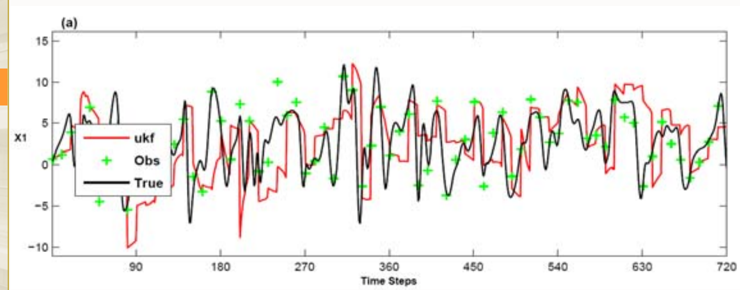


8 February 2008 22

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Reduced sigma-points

with reduced Sigma-points (20)



UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Conclusions

- The SPKF is a technique for a derivative-less optimal estimation using a deterministic sampling approach that ensures accurate estimation of error statistics.
- The SPKF is capable of estimating model state and parameters with better accuracy than EKF and EnKF for strong nonlinear systems.
- The SPKF is practically difficult for high dimensional systems. A possible solution is to reduce the number of sigma-points by "selecting a particular set of sigma-points".