Model Uncertainty Quantification for Data Assimilation in partially observed multi-scale systems

Sahani Pathiraja, with Peter Jan Van Leeuwen

Institut für Mathematik

Universität Potsdam

Workshop "Scales and scaling cascades in geophysical systems" – April 4-6, 2018, Haus des Sports, Hamburg

Motivation

- Main focus: Ensemble Data Assimilation
- Model Uncertainty due to unresolved sub-grid scale processes
- Uncertainty Quantification important for successful DA

What is Data Assimilation?

Problem Setting:

System states $\boldsymbol{x}_j$ evolve according to the following stochastic difference equation:

$$\boldsymbol{x}_j = M(\boldsymbol{x}_{j-1}) + \boldsymbol{\eta}_j$$

$\boldsymbol{\eta}_j$ is an additive stochastic model error.

What is Data Assimilation?

Problem Setting:

System states $\boldsymbol{x}_j$ evolve according to the following stochastic difference equation:

$$\boldsymbol{x}_j = M(\boldsymbol{x}_{j-1}) + \boldsymbol{\eta}_j$$

$\boldsymbol{\eta}_j$ is an additive stochastic model error.

Observations of the system are available in the form of:

$$\boldsymbol{y}_j = \mathbf{H}\boldsymbol{x}_j + \varepsilon_j$$
$$\varepsilon_j \sim N(0, \mathbf{R})$$

What is Data Assimilation?

Problem Setting:

System states $\boldsymbol{x}_j$ evolve according to the following stochastic difference equation:

$$\boldsymbol{x}_j = M(\boldsymbol{x}_{j-1}) + \boldsymbol{\eta}_j$$

$\boldsymbol{\eta}_j$ is an additive stochastic model error.

Observations of the system are available in the form of:

$$\boldsymbol{y}_j = \mathbf{H}\boldsymbol{x}_j + \varepsilon_j$$
$$\varepsilon_j \sim N(0, \mathbf{R})$$

Aim: estimate $p(\boldsymbol{x}_j|\boldsymbol{y}_j)$ (i.e. filtering).

Kalman Filter gives optimal solution for linear $\mathbf{M}, \mathbf{H}$ and zero mean time-uncorrelated Gaussian $\boldsymbol{\eta}$ and ε.

Aim

Develop method to estimate statistics of η for the following conditions:

- No knowledge of dynamics of fine scale process
- Only partial observations of coarse scale process available

Aim

Develop method to estimate statistics of η for the following conditions:

- No knowledge of dynamics of fine scale process
- Only partial observations of coarse scale process available

Many existing stochastic and deterministic parameterization techniques not amenable to the above:

e.g. Wilks (2005), Crommelin & Vanden-Eijnden (2008), Kwasniok (2012), Arnold et al. (2013), Lu et al. (2017)

Assumptions

1. States are directly but partially observed (i.e. $\mathbf{H}$ is a non-square $(0,1)$ matrix)

Assumptions

1. States are directly but partially observed (i.e. **H** is a non-square $(0, 1)$ matrix)
2. Model error $\boldsymbol{\eta}_j$ depends on some informative variable (e.g. $\boldsymbol{x}_{j-1}$ or some reduced form of it)

Assumptions

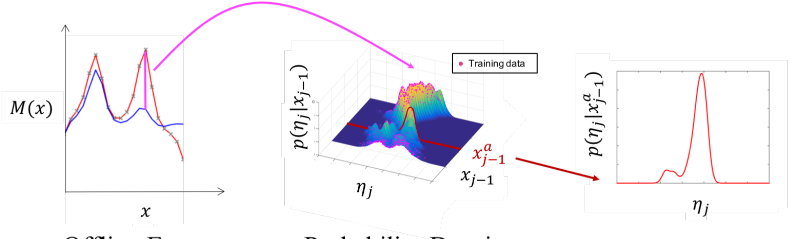
1. States are directly but partially observed (i.e. $\mathbf{H}$ is a non-square $(0,1)$ matrix)
2. Model error $\boldsymbol{\eta}_j$ depends on some informative variable (e.g. $\boldsymbol{x}_{j-1}$ or some reduced form of it)
3. $||\boldsymbol{\varepsilon}_j|| << ||\boldsymbol{\eta}_j||$

Assumptions

1. States are directly but partially observed (i.e. $\mathbf{H}$ is a non-square $(0,1)$ matrix)
2. Model error $\boldsymbol{\eta}_j$ depends on some informative variable (e.g. $\boldsymbol{x}_{j-1}$ or some reduced form of it)
3. $||\varepsilon_j|| << ||\boldsymbol{\eta}_j||$
4. Error statistics are the same at each point in time and space:

$$p(\boldsymbol{\eta}_j[k]|\boldsymbol{x}_{j-1}[k]) = p(\boldsymbol{\eta}_b[l]|\boldsymbol{x}_{b-1}[l]) \quad \forall k, j, b, l$$

Proposed Method



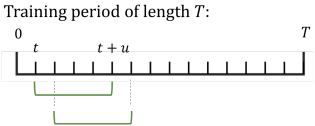
Offline Error Estimation

Partial Observations!

Probability Density Estimation

Assimilation

Proposed Method - Error Estimation



At any given time t, aim is to minimise:

$$J\left(\boldsymbol{\eta}^{\dagger}_{t:t+u}\right) = \sum_{i=1}^{m} \frac{1}{n_i - 1} \boldsymbol{v}_i^T \mathbf{C}_i^T \mathbf{C}_i \boldsymbol{v}_i$$

subject to:

$$\boldsymbol{y}_j = \mathbf{H}\boldsymbol{x}_j \quad \forall \quad j \in \{t, t+1, ..., t+u\}$$

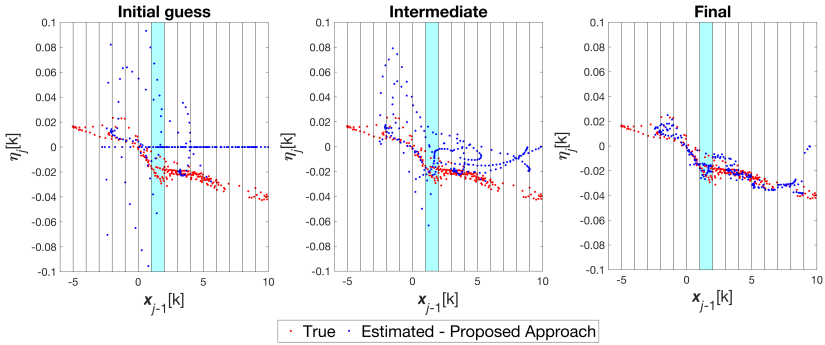
(negligible observation error assumption)

where:

$$\boldsymbol{x}_j = M(\boldsymbol{x}_{j-1}) + \boldsymbol{\eta}_j$$

$\boldsymbol{\eta}^{\dagger}_{t:t+u}$ = model errors on unobserved components of $\boldsymbol{x}_t, ..., \boldsymbol{x}_{t+u}$

Proposed Method - Error Estimation



Numerical Experiments

Two Layer Lorenz 96:

$$\frac{dX_k}{dt} = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F + \frac{h_x}{L}\sum_{l=1}^{L} Y_{l,k};\ k \in \{1, ..., K\}$$

$$\frac{dY_{l,k}}{dt} = \frac{1}{\xi}(-Y_{l+1,k}(Y_{l+2,k} - Y_{l-1,k}) - Y_{l,k} + h_y X_k; \quad l \in \{1, ..., L\}$$

Numerical Experiments

Two Layer Lorenz 96:

$$\frac{dX_k}{dt} = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F + \frac{h_x}{L}\sum_{l=1}^{L} Y_{l,k};\ k \in \{1, ..., K\}$$

$$\frac{dY_{l,k}}{dt} = \frac{1}{\xi}(-Y_{l+1,k}(Y_{l+2,k} - Y_{l-1,k}) - Y_{l,k} + h_y X_k; \quad l \in \{1, ..., L\}$$

Parameter	**Case Study 1 - large time scale sep.**	**Case Study 2 - small time scale sep.**
ξ	$\frac{1}{128} \approx 0.008$	0.7
h_x	-0.8	-2
h_y	1	1
J	128	20
K	9	9
F	10	14

Numerical Experiments

Observation Details:

- every 2nd X_k is measured
- 0.02 & 0.04 MTU for Case Study 1 and 2 respectively
- $R = 10^{-6} I$ (i.e. negligible)

Benchmark Method (B1)

Analysis Increment Based Method

- ETKF-TV (Mitchell & Carrassi, 2015)

$$\boldsymbol{x}_j^{fi} = M(\boldsymbol{x}_{j-1}^{ai}) - \alpha \boldsymbol{\eta}_j^i$$
$$\boldsymbol{\eta}_j^i \sim N(\overline{\boldsymbol{b}}_m, \overline{\mathbf{P}}_m)$$

Benchmark Method (B1)

Analysis Increment Based Method

- ETKF-TV (Mitchell & Carrassi, 2015)

$$\boldsymbol{x}_j^{fi} = M(\boldsymbol{x}_{j-1}^{ai}) - \alpha \boldsymbol{\eta}_j^i$$
$$\boldsymbol{\eta}_j^i \sim N(\overline{\boldsymbol{b}}_m, \overline{\mathbf{P}}_m)$$

where:

$$\alpha = \text{tuning parameter}$$

$$\overline{\boldsymbol{b}}_m = \frac{1}{N} \sum_{j=1}^{N} \delta \boldsymbol{x}_j^a$$

$$\overline{\mathbf{P}} = \frac{1}{N-1} \sum_{j=1}^{N} \left[\delta \boldsymbol{x}_j^a - \overline{\boldsymbol{b}}_m \right] \left[\delta \boldsymbol{x}_j^a - \overline{\boldsymbol{b}}_m \right]^T$$

$$\delta \boldsymbol{x}_j^a = \frac{1}{n} \sum_{i=1}^{n} \left(\boldsymbol{x}_j^{ai} - \boldsymbol{x}_j^{fi} \right)$$

Benchmark Method (B2)

Long Window Weak Constraint 4d-Var based

- Error estimation using Long Window Weak Constraint 4d-Var (Tremolet, 2006):

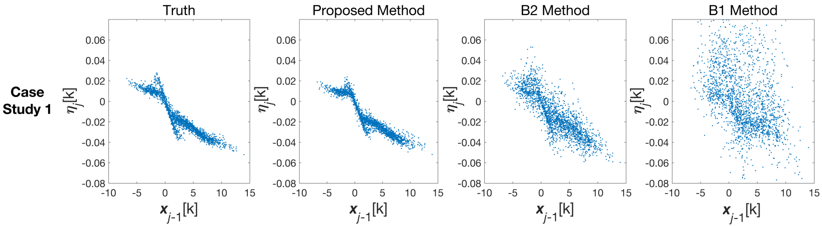
$$J(\boldsymbol{\eta}_{t:t+u}) = \frac{1}{2}\sum_{j=t+1}^{t+u} \boldsymbol{\eta}_j^T \mathbf{Q}_j^{-1} \boldsymbol{\eta}_j + \frac{1}{2}\sum_{j=t+1}^{t+u} (\mathbf{H}\boldsymbol{x}_j - \boldsymbol{y}_j)^T \mathbf{R}^{-1} (\mathbf{H}\boldsymbol{x}_j - \boldsymbol{y}_j)$$

  where:

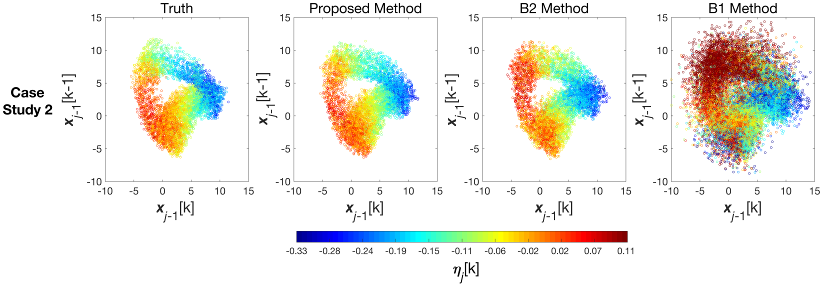
$$\boldsymbol{x}_j = M(\boldsymbol{x}_{j-1}) + \boldsymbol{\eta}_j \text{ for } j \in \{t+1, \ldots, t+u\}$$

- All other aspects same as proposed approach

Error Estimation Results - Case Study 1



Error Estimation Results - Case Study 2



Why is the conditional minimization approach better?

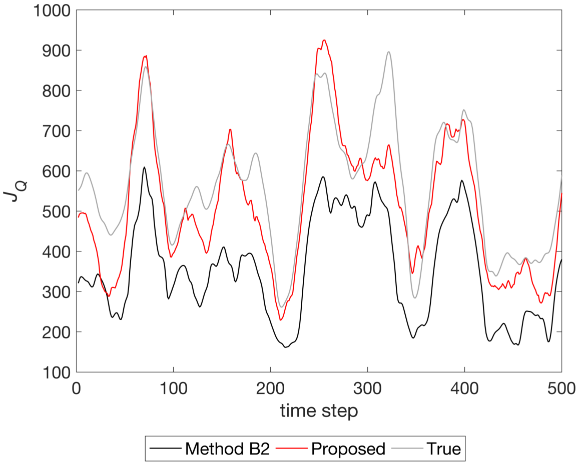


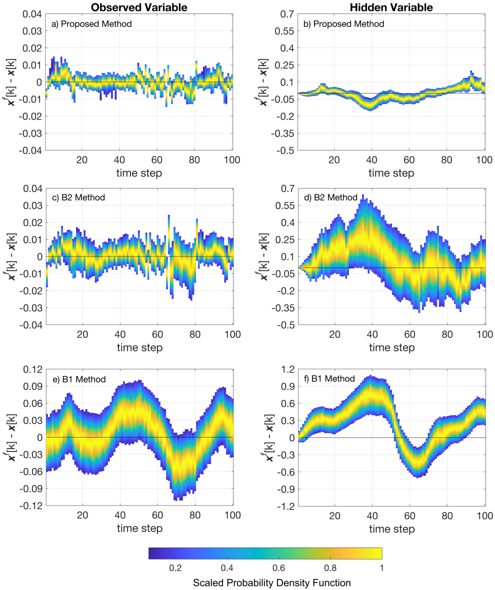
Figure: Snapshot of J_Q values for method B2, proposed and true data for Case Study 2

Impact on Assimilation

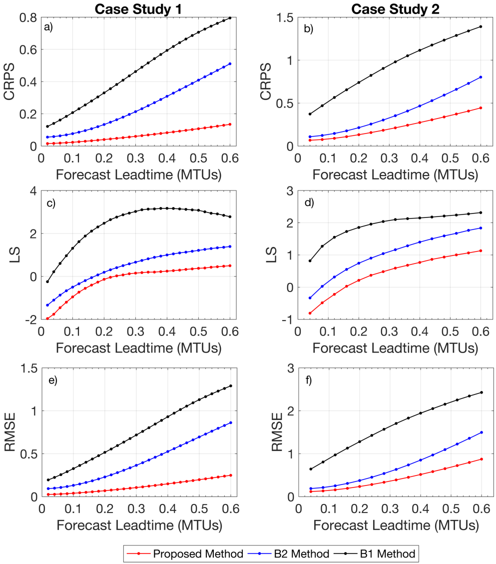
Data Assimilation setup:

- Ensemble Transform Kalman Filter (ETKF) (Wang & Bishop, 2004)
- ensemble size $(n) = 1000$
- observation frequency - as per estimation period
- assimilation length - 3000 observation intervals

One Step ahead forecast densities - Case Study 1



Forecast Skill



Summary

- Model UQ important for successful DA

Summary

- Model UQ important for successful DA
- Proposed method for quantifying model uncertainty due to unresolved sub-grid scale processes

Summary

- Model UQ important for successful DA
- Proposed method for quantifying model uncertainty due to unresolved sub-grid scale processes
- Difficult conditions: no knowlege of sub-grid scale dynamics and partial observations of coarse scale process

Summary

- Model UQ important for successful DA
- Proposed method for quantifying model uncertainty due to unresolved sub-grid scale processes
- Difficult conditions: no knowlege of sub-grid scale dynamics and partial observations of coarse scale process
- Numerical experiments with Lorenz 96 show improved error estimates and assimilation quality compared to benchmarks

Further Work

- Non-negligble observation error?
- Scalability?

Acknowledgements

- This research has been partially funded by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1294 ''Data Assimilation"
- We gratefully acknowledge Professor Georg Gottwald, Professor Sebastian Reich, and Dr. Jana De Wiljes for thoughtful discussions on this work.

References

- Arnold, H. M., Moroz, I. M. and Palmer, T. N. (2013) 'Stochastic parametrizations and model uncertainty in the Lorenz ' 96 system', Philosophical Transactions of the Royal Society A, 371. doi: dx.doi.org/10.1098/rsta.2011.0479.
- Crommelin, D. and Vanden-Eijnden, E. (2008) 'Subgrid-Scale Parameterization with Conditional Markov Chains', Journal of the Atmospheric Sciences, 65(8), pp. 2661–2675. doi: 10.1175/2008JAS2566.1.
- Kwasniok, F. (2012) 'Data-based stochastic subgrid-scale parametrization: an approach using cluster-weighted modelling', Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 370(1962), pp. 1061–1086. doi: 10.1098/rsta.2011.0384

References

- Lu, F., Tu, X., Chorin, A. J., Lu, F., Tu, X. and Chorin, A. J. (2017) 'Accounting for model error from unresolved scales in ensemble Kalman filters by stochastic parametrization', Monthly Weather Review, p. MWR-D-16-0478.1. doi: 10.1175/MWR-D-16-0478.1.
- Mitchell, L. and Carrassi, A. (2015) 'Accounting for model error due to unresolved scales within ensemble Kalman filtering', Quarterly Journal of the Royal Meteorological Society, 141(689), pp. 1417–1428. doi: 10.1002/qj.2451.
- Tremolet, Y. (2006) 'Accounting for an imperfect model in 4D-Var', Quarterly Journal of the Royal Meteorological Society, 132(621), pp. 2483–2504. doi: 10.1256/qj.05.224.

References

- Wang, X., Bishop, C. H. and Julier, S. J. (2004) 'Which Is Better, an Ensemble of Positive-Negative Pairs or a Centered Spherical Simplex Ensemble', Bulletin of the American Meteorological Society, pp. 2823–2829. doi: 10.1175/1520-0493(2004)132¡1590:WIBAEO¿2.0.CO;2.
- Wilks, D. S. (2005) 'Effects of stochastic parameterizations in the Lorenz 96 system', Quarterly Journal of the Royal Meteorological Society, 131. doi: 10.1256/qj.04.03.