King Abdullah University of Science and Technology
Earth Science & Engineering

International workshop on coupled data assimilation

Ensemble Kalman Filtering with One-Step-Ahead Smoothing for Efficient Data Assimilation into One-Way Coupled Models

Naila Raboudi, B. Ait-El-Fquih, A. Subramanian, I. Hoteit

October, 2016

Context

- Consider a discrete-time state-parameter dynamical system

$$\begin{cases} x_n = \mathcal{M}_{n-1}(x_{n-1}, \theta) + \eta_{n-1}; & \eta_{n-1} \sim \mathcal{N}(0, Q_{n-1}) \\ y_n = H_n x_n + \varepsilon_n; & \varepsilon_n \sim \mathcal{N}(0, R_n) \end{cases},$$

Context

- Consider a discrete-time state-parameter dynamical system

$$\begin{cases} x_n = \mathcal{M}_{n-1}(x_{n-1}, \theta) + \eta_{n-1}; & \eta_{n-1} \sim \mathcal{N}(0, Q_{n-1}) \\ y_n = H_n x_n + \varepsilon_n; & \varepsilon_n \sim \mathcal{N}(0, R_n) \end{cases},$$

- Standard solutions
  - Joint EnKF: Apply EnKF on the augmented state $z_n = \left[x_n^T \ \theta^T\right]^T$
    - Subject to instability and intractability (e.g. Moradkhani et al., 2005)
    - May lead to inconsistency between state and parameters estimates (Wen and Chen, 2006)

Context

- Consider a discrete-time state-parameter dynamical system

$$\begin{cases} x_n = \mathcal{M}_{n-1}(x_{n-1}, \theta) + \eta_{n-1}; & \eta_{n-1} \sim \mathcal{N}(0, Q_{n-1}) \\ y_n = H_n x_n + \varepsilon_n; & \varepsilon_n \sim \mathcal{N}(0, R_n) \end{cases},$$

- Standard solutions
  - Joint EnKF: Apply EnKF on the augmented state $z_n = \left[x_n^T \ \theta^T\right]^T$
    - Subject to instability and intractability (e.g. Moradkhani et al., 2005)
    - May lead to inconsistency between state and parameters estimates (Wen and Chen, 2006)
  - Dual-EnKF: Operates as a succession of two separate EnKFs updating first the parameters and then the state.
    - The separation of the update steps was shown to provide more consistent estimates
    - Heursitic algorithm, not fully Bayesian consistent

The classical Dual-EnKF for state-parameter estimation

The classical Dual-EnKF for state-parameter estimation

EnKF θ
EnKF x

$$\{x_{n-1}^{a,(i)}(\theta_{|n-1}^{(i)})\}_{i=1}^{N_e}$$

The classical Dual-EnKF for state-parameter estimation

EnKF θ

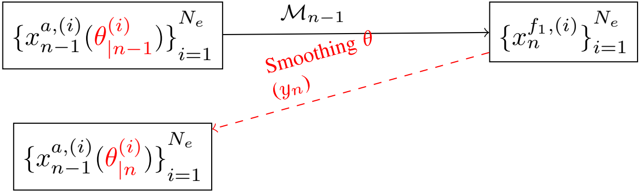
EnKF x

$$\{x_{n-1}^{a,(i)}(\theta_{|n-1}^{(i)})\}_{i=1}^{N_e} \xrightarrow{\mathcal{M}_{n-1}} \{x_n^{f_1,(i)}\}_{i=1}^{N_e}$$

The classical Dual-EnKF for state-parameter estimation

EnKF θ

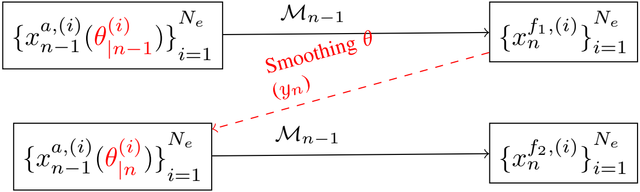
EnKF x



The classical Dual-EnKF for state-parameter estimation

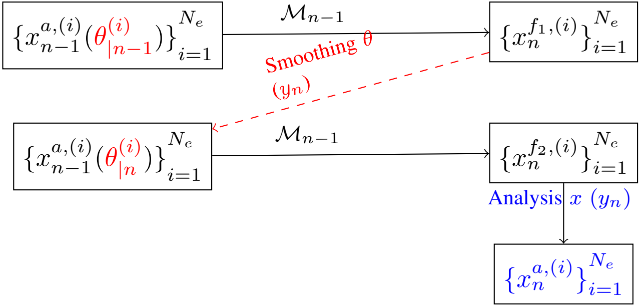
EnKF θ

EnKF x



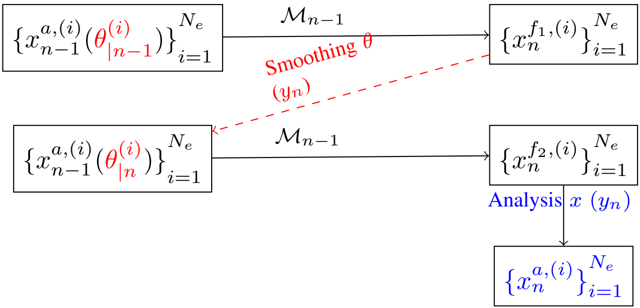
The classical Dual-EnKF for state-parameter estimation

EnKF θ
EnKF x



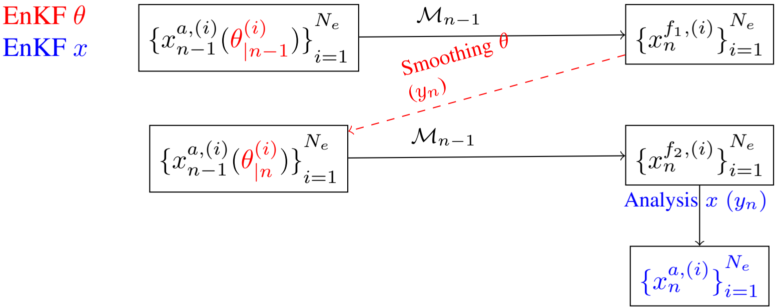
The classical Dual-EnKF for state-parameter estimation

EnKF θ
EnKF x



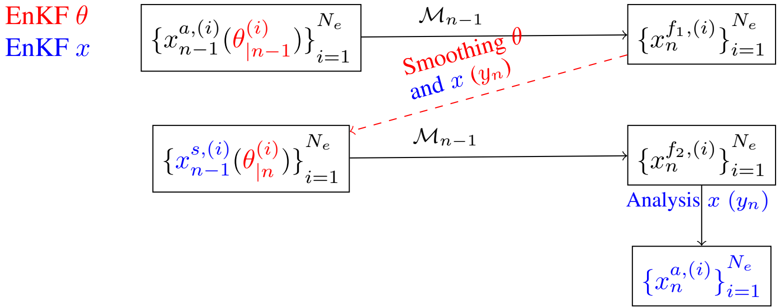
- Heuristic, not fully Bayesian consistent.

The classical Dual-EnKF for state-parameter estimation



- Heuristic, not fully Bayesian consistent.
- Recently, Ait-El-Fquih et al., (2016) proposed a new Bayesian consistent dual-like filtering scheme, which introduced a new smoothing step between two successive analysis steps.

A Dual-EnKF with OSA smoothing for state-parameter estimation



- Fully Bayesian consistent
- Recently, Ait-El-Fquih et al., (2016) proposed a new Bayesian consistent dual-like filtering scheme based on a one-step-ahead (OSA) formulation, which introduces a new smoothing step between two successive analysis steps.

- Successfully tested with a groundwater flow model to estimate the hydraulic head and the conductivity field at KAUST, and with a marine ecosystem model at NERSC (Gharamti et al, 2015; 2016)

context

- Successfully tested with a groundwater flow model to estimate the hydraulic head and the conductivity field at KAUST, and with a marine ecosystem model at NERSC (Gharamti et al, 2015; 2016)
- One-way coupled filtering problems could be considered as a generalization of the state-parameter estimation problem where the parameter θ evolves according to a dynamical model and is observed.

- Successfully tested with a groundwater flow model to estimate the hydraulic head and the conductivity field at KAUST, and with a marine ecosystem model at NERSC (Gharamti et al, 2015; 2016)
- One-way coupled filtering problems could be considered as a generalization of the state-parameter estimation problem where the parameter θ evolves according to a dynamical model and is observed.

Objective

Build an efficient Dual-like EnKF scheme for data assimilation into one-way coupled systems: Separate the updates to tackle inconsistencies, and inject back the ”lost” information through a ”re-forecast” step with the nonlinear model

Outline

Standard solutions for coupled systems

- One-way coupled systems

$$\begin{cases} x_n = \mathcal{F}_{n-1}(x_{n-1}) + \eta^x_{n-1} \\ y^x_n = H^x_n x_n + \varepsilon^x_n \end{cases} \rightarrow \begin{cases} w_n = \mathcal{G}_{n-1}(w_{n-1}, x_{n-1}) + \eta^w_{n-1} \\ y^w_n = H^w_n w_n + \varepsilon^w_n \end{cases}$$

x free variable — *w forced variable*

- Standard solutions

Standard joint (strong) EnKF

$$y^z_n = \begin{pmatrix} y^x_n \\ y^w_n \end{pmatrix}$$

$$z^f_n = \begin{pmatrix} x^f_n \\ w^f_n \end{pmatrix} \longrightarrow \boxed{\text{Standard EnKF}} \longrightarrow z^a_n = \begin{pmatrix} x^a_n \\ w^a_n \end{pmatrix}$$

- Cross-correlations between coupled variables are considered in the update
- Cross-correlations may not be well estimated with small ensembles and not very representative in presence of strong nonlinearities

Standard solutions for coupled systems

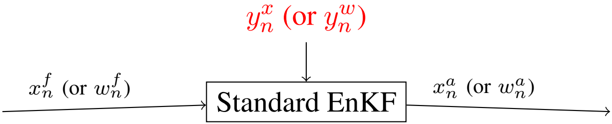
- One-way coupled systems

$$\left\{\begin{array}{lcl} x_n & = & \mathcal{F}_{n-1}(x_{n-1}) + \eta^x_{n-1} \\ y^x_n & = & H^x_n x_n + \varepsilon^x_n \end{array}\right. \rightarrow \left\{\begin{array}{lcl} w_n & = & \mathcal{G}_{n-1}(w_{n-1}, x_{n-1}) + \eta^w_{n-1} \\ y^w_n & = & H^w_n w_n + \varepsilon^w_n \end{array}\right.$$

x free variable *w forced variable*

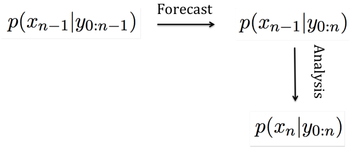
- Standard solutions

Standard weak EnKF

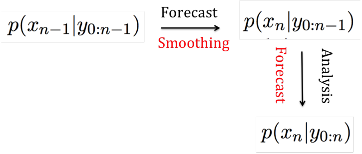


- Separate updates are more practical in real applications
- Lost of information from neglecting cross-correlations

- Standard path

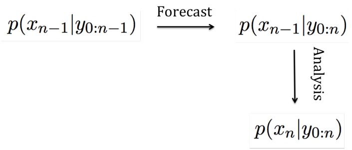


- OSA smoothing based path

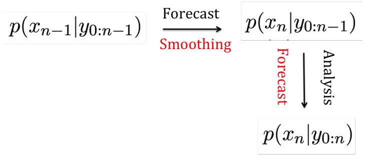


The concept of filtering with OSA smoothing

- Standard path



- OSA smoothing based path



- *Smoothing step:*
  $p(x_{n-1}|y_{0:n})$ is first computed using the likelihood $p(y_n|x_{n-1}, y_{0:n-1})$

$$p(x_{n-1}|y_{0:n}) \propto p(y_n|x_{n-1}, y_{0:n-1})p(x_{n-1}|y_{0:n-1})$$

- *Analysis step:*
  $p(x_n|y_{0:n})$ is computed using the *a posteriori* transition $p(x_n|x_{n-1}, y_{0:n})$

$$p(x_n|y_{0:n}) = \int p(x_n|x_{n-1}, y_{0:n})p(x_{n-1}|y_{0:n})dx_{n-1},$$

The concept of filtering with OSA smoothing

- Two classical stochastic sampling properties are used to derive the joint EnKF-OSA algorithm

Property 1 (Hierarchical sampling)
Assuming that one can sample from $p(\mathbf{x}_1)$ and $p(\mathbf{x}_2|\mathbf{x}_1)$, then a sample, $\mathbf{x}_2^*$, from $p(\mathbf{x}_2)$ can be drawn as follows:

1. $\mathbf{x}_1^* \rightsquigarrow p(\mathbf{x}_1)$;
2. $\mathbf{x}_2^* \rightsquigarrow p(\mathbf{x}_2|\mathbf{x}_1^*)$.

Property 2 (Conditional sampling)
Consider a Gaussian pdf, $p(\mathbf{x}, \mathbf{y})$, with $\mathbf{P}_{xy}$ and $\mathbf{P}_y$ denoting the cross-covariance of $\mathbf{x}$ and $\mathbf{y}$ and the covariance of $\mathbf{y}$, respectively. Then a sample, $\mathbf{x}^*$, from $p(\mathbf{x}|\mathbf{y})$, can be drawn as follows:

1. $(\widetilde{\mathbf{x}}, \widetilde{\mathbf{y}}) \rightsquigarrow p(\mathbf{x}, \mathbf{y})$;
2. $\mathbf{x}^* = \widetilde{\mathbf{x}} + \mathbf{P}_{xy}\mathbf{P}_y^{-1}[\mathbf{y} - \widetilde{\mathbf{y}}]$.

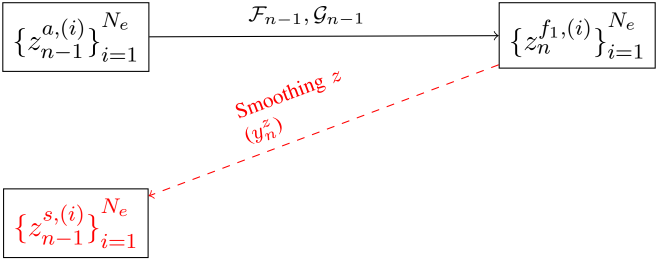
A joint EnKF-OSA for one-way coupled systems

$$\boxed{\{z_{n-1}^{a,(i)}\}_{i=1}^{N_e}}$$

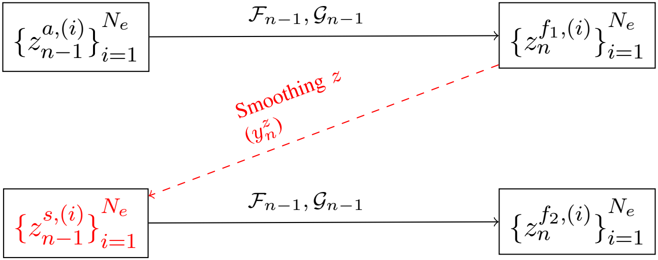
A joint EnKF-OSA for one-way coupled systems

$$\left\{z_{n-1}^{a,(i)}\right\}_{i=1}^{N_e} \xrightarrow{\mathcal{F}_{n-1},\,\mathcal{G}_{n-1}} \left\{z_{n}^{f_1,(i)}\right\}_{i=1}^{N_e}$$

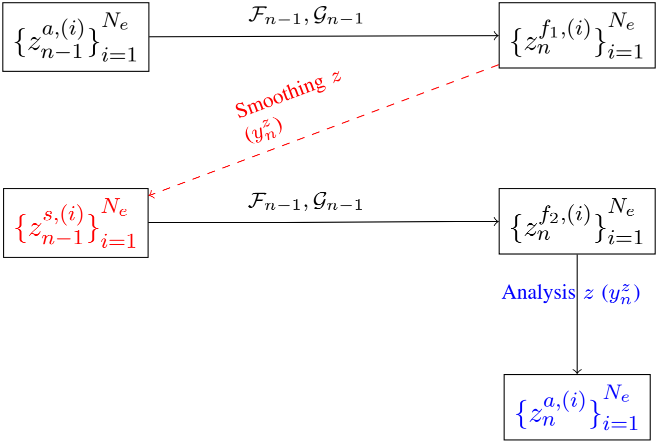
A joint EnKF-OSA for one-way coupled systems



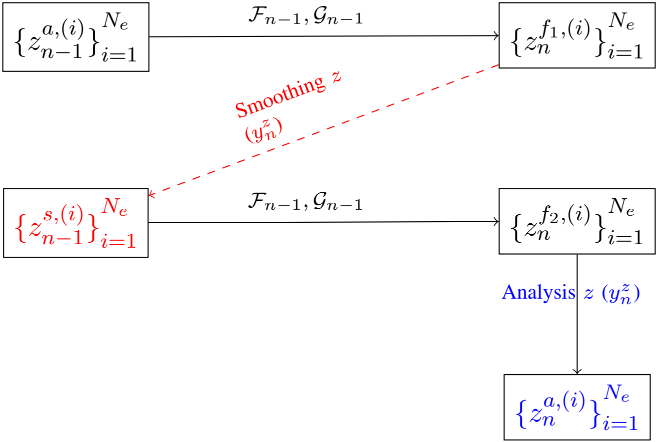
A joint EnKF-OSA for one-way coupled systems



A joint EnKF-OSA for one-way coupled systems



A joint EnKF-OSA for one-way coupled systems

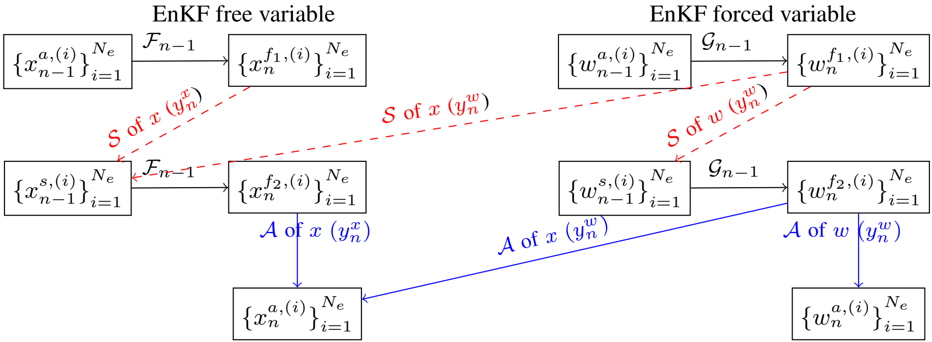


- OSA adds one smoothing step and another re-forecast step
- Data are exploited twice in a fully Bayesian framework

An EnKF-OSA with separate updates for one-way coupled systems

Assumption to separate the update steps

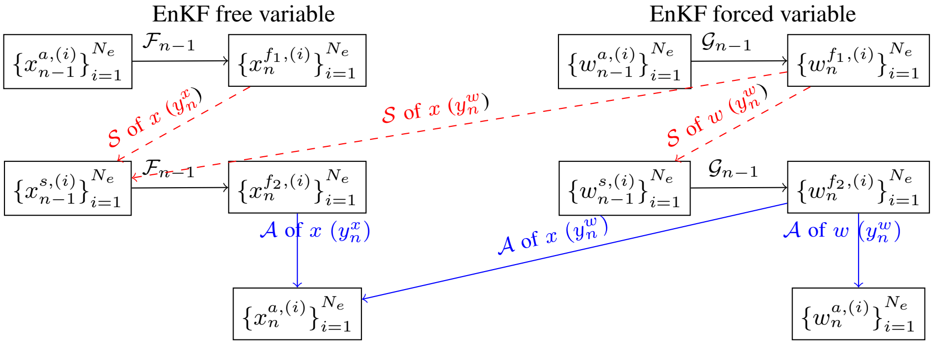
Given the past observations of both x and w, w_n needs to be independent of the current and future observations of the variable x.



An EnKF-OSA with separate updates for one-way coupled systems

Assumption to separate the update steps

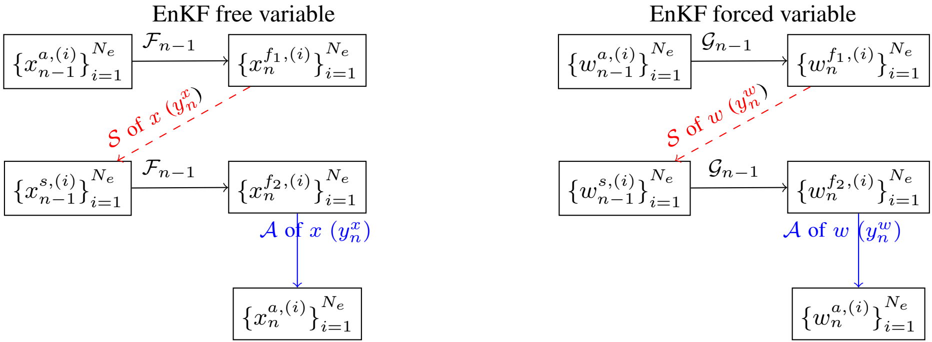
Given the past observations of both x and w, w_n needs to be independent of the current and future observations of the variable x.



- The free variable x is updated by both y_n^x and y_n^w while the forced variable w is updated by y_n^w only.

Assumption to separate the update steps

Given the past observations of both x and w, w_n needs to be independent of the current and future observations of the variable x.



- The weak version is derived by simply neglecting the cross-correlations and updating each variable by its own observations using an EnKF-OSA on each model

EnKF $\mathcal{V}s$ EnKF-OSA

- EnKF OSA filtering uses an alternative path to better exploit the data (useful in non-optimal ensemble implementation)
- Constraining the state with future observations in the smoothing step should provide improved background for the next analysis

  $\Longrightarrow$ Mitigate some of the background limitations of EnKF-like methods
- EnKF-OSA conditions the ensemble resampling with future data
- May help mitigating for the loss of information in a weak formulation, by better exploiting of the data and re-forecasting with the nonlinear model.
- OSA algorithm is computationally (almost twice) more expensive.

Numerical experiments with Lorenz-96

- Governing equations

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})\, x_{i-1} - x_i + F, \qquad i = 1, \cdots, N_x$$

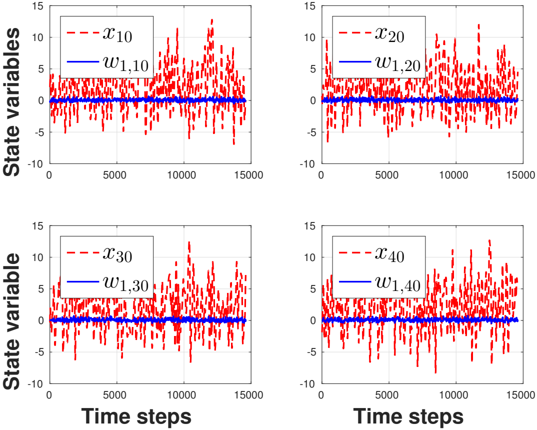
$$\frac{dw_{j,i}}{dt} = (w_{j-1,i} - w_{j+2,i})\, cbw_{j+1,i} - cw_{j,i} + \frac{hc}{b} x_i, \qquad j = 1, \cdots, N_w$$

$$N_x = 40; \quad N_w = 1;$$

- Experimental setup
  - Twin experiments
  - 5-years simulation period
  - Time step: 0.005
- Assimilation scenarios
  - 3 different observation scenarios: all (40), half (20), and quarter (10) model state variables are observed
  - 3 different ensemble sizes: 10, 20 and 40
  - Observations are assimilated every 1 day

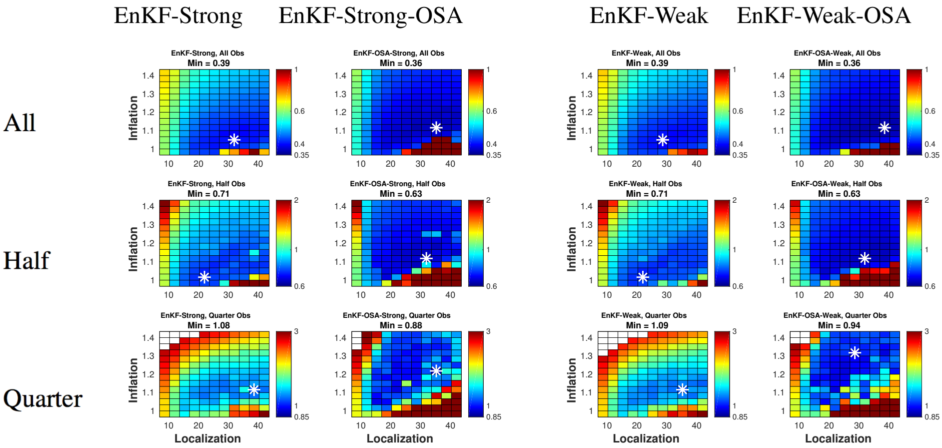
Numerical experiments with Lorenz-96

- Reference states for both models



- We only analyze the free variable x: assimilation results are consistent for both variables, but improvements are more pronounced with x

40 members and assimilation every $24h$



	EnKF-Strong (OSA/Reg)	EnKF-Weak (OSA/Reg)
	$N_e = 40$	$N_e = 40$
all	8.2 %	8.1 %
half	10.5 %	10.4 %
quarter	19.2 %	14.1 %

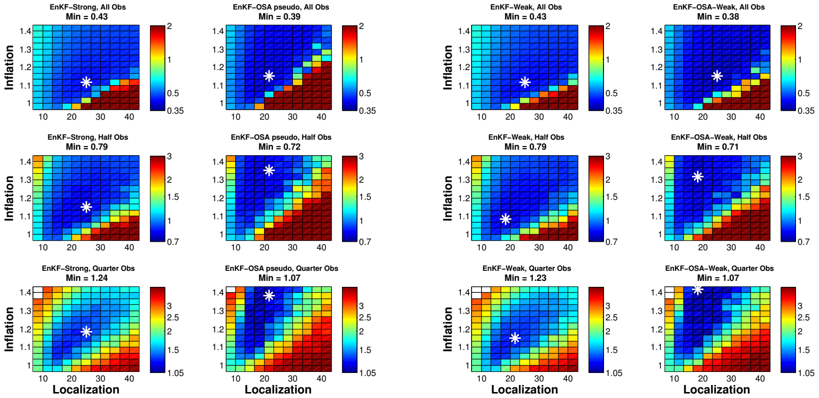
20 members and assimilation every $24h$

EnKF-Strong | EnKF-Strong-OSA | EnKF-Weak | EnKF-Weak-OSA

All

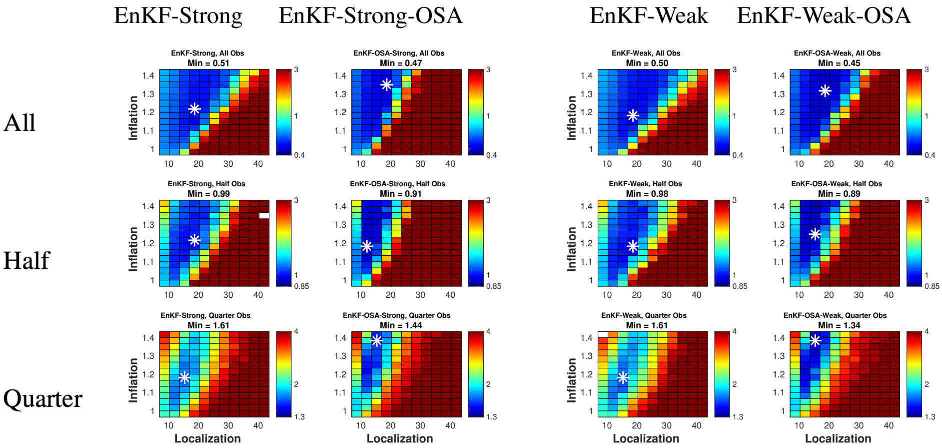
Half

Quarter



	EnKF-Strong (OSA/Reg)	EnKF-Weak (OSA/Reg)
	$N_e = 20$	$N_e = 20$
all	9 %	11.2 %
half	10.2 %	10.8 %
quarter	13.5 %	13.6 %

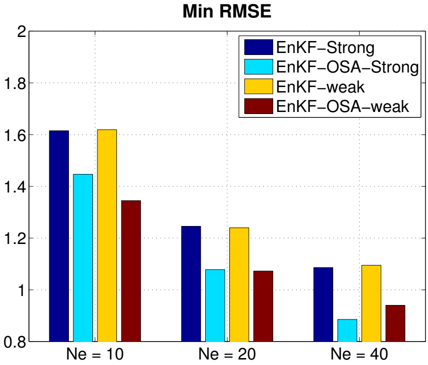
10 members and assimilation every $24h$



	EnKF-Strong (OSA/Reg)	EnKF-Weak (OSA/Reg)
	$N_e = 10$	$N_e = 10$
all	6.3 %	9.6 %
half	8 %	9.8 %
quarter	10.5 %	17 %

Sensitivity to ensemble size

- Quarter of the observations are assimilated



- Slight improvement with same computational cost; need to examine larger ensembles

Summary

- EnKF-OSA (weak and strong formulations) is found beneficial for one-way coupled DA compared to the standard EnKF
- EnKF-OSA produces better estimates when the number of observations is smaller
- The weak formulation is more beneficial when cross-correlations are not well estimated, otherwise, better use a strong formulation
- The strong OSA version was shown to be more beneficial than the weak OSA when the ensemble is enough representative
- At the same computational cost and large enough ensemble, EnKF-OSA seems to outperform the standard EnKF, but, more tests are needed

Thank you for your attention