

An iterative ensemble Kalman smoother

Marc Bocquet ¹,
Pavel Sakov ², Jean-Matthieu Haussaire ¹

¹CEREA, joint lab École des Ponts ParisTech and EdF R&D, Université Paris-Est, France

²Bureau of Meteorology, Australia

(bocquet@cerea.enpc.fr)



The IEnKS: at the crossroad between the EnKF and 4D-Var

- ▶ The IEnKS follows the scheme of the EnKF:

Deterministic analysis → Posterior ensemble generation → Ensemble forecast

- ▶ Except that:

- The analysis in ensemble space is variational [e.g. Zupanski, 2005] over a finite time windows. It may require several iterations in strongly nonlinear conditions [Gu & Oliver, 2007; Sakov et al., 2012; Bocquet and Sakov, 2012-2014].
- The gradient of the 4D cost function is computed with the ensemble [Gu & Oliver, 2007; Liu et al., 2008; Buehner et al. 2010]: no need for the tangent linear/adjoint.

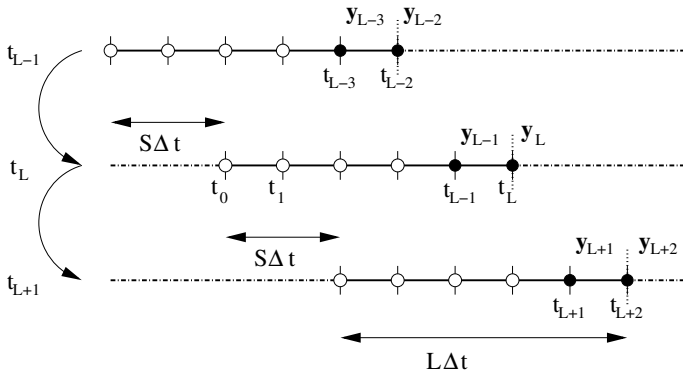
- ▶ It generalises the iterative extended Kalman filter/smoothen [Wishner et al., 1969; Jazwinski, 1970; Bell, 1994] to ensemble methods.

- ▶ It is a unified/straightforward scheme (no hybridisation so to speak).

- ▶ The IEnKS belongs to the class of **ensemble variational (EnVar)** methods. Other members of the family are: Hybrid methods, 4D-Var-Ben, 4DEnVar, Ensemble of data assimilation (EDA), etc.
- ▶ The IEnKF/IEnKS differ from the other ones in that they stem from a convincing (heuristic) **Bayesian derivation**.
This derivation says that, under the conditions that:
 - the prior is assumed Gaussian,
 - the statistics of the errors are modeled by an ensemble,and without worrying about practical implementation, the solution is unique.
- ▶ Besides, and very importantly, the complete cycle (i.e. including posterior ensemble generation) has been including from the very beginning (EnKF inheritance).
- ▶ The IEnKS has potential for parameter estimation, as it is variational but avoids the derivation of the adjoint.

The IEnKS: the cycling

- ▶ L : length of the data assimilation window,
- ▶ S : shift of the data assimilation window in between two updates.



The IEnKS: a variational standpoint

- Analysis IEnKS cost function in state space $p(\mathbf{x}_0|\mathbf{y}_L) \propto \exp(-\mathcal{J}(\mathbf{x}_0))$:

$$\mathcal{J}(\mathbf{x}_0) = \sum_{k=1}^L \frac{1}{2} (\mathbf{y}_k - H_k \circ \mathcal{M}_{k \leftarrow 0}(\mathbf{x}_0))^T \beta_k \mathbf{R}_k^{-1} (\mathbf{y}_k - H_k \circ \mathcal{M}_{k \leftarrow 0}(\mathbf{x}_0)) + \frac{1}{2} (\mathbf{x}_0 - \bar{\mathbf{x}}_0) \mathbf{P}_0^{-1} (\mathbf{x}_0 - \bar{\mathbf{x}}_0). \quad (1)$$

$\{\beta_0, \beta_1, \dots, \beta_L\}$ weight the observations impact within the window.

- Reduced scheme in ensemble space, $\mathbf{x}_0 = \bar{\mathbf{x}}_0 + \mathbf{A}_0 \mathbf{w}$, where \mathbf{A}_0 is the ensemble anomaly matrix:

$$\tilde{\mathcal{J}}(\mathbf{w}) = \mathcal{J}(\bar{\mathbf{x}}_0 + \mathbf{A}_0 \mathbf{w}). \quad (2)$$

- IEnKS cost function in ensemble space [Hunt et al., 2007; Bocquet and Sakov, 2012-2014]:

$$\tilde{\mathcal{J}}(\mathbf{w}) = \frac{1}{2} \sum_{k=1}^L (\mathbf{y}_k - H_k \circ \mathcal{M}_{k \leftarrow 0}(\bar{\mathbf{x}}_0 + \mathbf{A}_0 \mathbf{w}))^T \beta_k \mathbf{R}_k^{-1} (\mathbf{y}_k - H_k \circ \mathcal{M}_{k \leftarrow 0}(\bar{\mathbf{x}}_0 + \mathbf{A}_0 \mathbf{w})) + \frac{1}{2} (N-1) \mathbf{w}^T \mathbf{w}. \quad (3)$$

The IEnKS: minimisation scheme

► As a variational **reduced** method, one can use Gauss-Newton [Sakov et al., 2012], Levenberg-Marquardt [Bocquet and Sakov, 2012; Chen and Oliver, 2013], quasi-Newton, etc., minimisation schemes.

► Gauss-Newton scheme (the Hessian is approximate):

$$\begin{aligned}
 \mathbf{w}^{(p+1)} &= \mathbf{w}^{(p)} - \widetilde{\mathcal{H}}_{(p)}^{-1} \nabla \widetilde{\mathcal{J}}_{(p)}(\mathbf{w}^{(p)}), \\
 \mathbf{x}_0^{(p)} &= \mathbf{x}_0^{(0)} + \mathbf{A}_0 \mathbf{w}^{(p)}, \\
 \nabla \widetilde{\mathcal{J}}_{(p)} &= - \sum_{k=1}^L \mathbf{Y}_{k,(p)}^T \beta_k \mathbf{R}_k^{-1} \left(\mathbf{y}_k - H_k \circ \mathcal{M}_{k \leftarrow 0}(\mathbf{x}_0^{(p)}) \right) + (N-1) \mathbf{w}^{(p)}, \\
 \widetilde{\mathcal{H}}_{(p)} &= (N-1) \mathbf{I}_N + \sum_{k=1}^L \mathbf{Y}_{k,(p)}^T \beta_k \mathbf{R}_k^{-1} \mathbf{Y}_{(p)}, \\
 \mathbf{Y}_{k,(p)} &= [H_k \circ \mathcal{M}_{k \leftarrow 0}]'_{|\mathbf{x}_0^{(p)}} \mathbf{A}_0.
 \end{aligned} \tag{4}$$

► One solution to compute the 4D sensitivities: the **bundle** scheme. It simply mimics the action of the tangent linear by finite difference:

$$\mathbf{Y}_{k,(p)} \approx \frac{1}{\varepsilon} H_k \circ \mathcal{M}_{k \leftarrow 0} \left(\mathbf{x}^{(p)} \mathbf{1}^T + \varepsilon \mathbf{A}_0 \right) \left(\mathbf{I}_N - \frac{\mathbf{1} \mathbf{1}^T}{N} \right). \tag{5}$$

The IEnKS: ensemble update and the forecast step

- Generate an updated ensemble from the previous analysis:

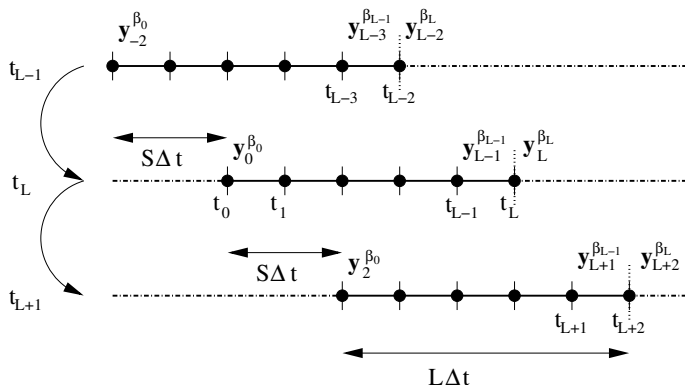
$$\mathbf{E}_0^* = \mathbf{x}_0^* \mathbf{1}^T + \sqrt{N-1} \mathbf{A}_0 \widetilde{\mathcal{H}}_*^{-1/2} \mathbf{U} \quad \text{where} \quad \mathbf{U} \mathbf{1} = \mathbf{1}, \quad \mathbf{u} = (1, 1, \dots, 1)^T. \quad (6)$$

- Just propagate the updated ensemble from t_0 to t_S :

$$\mathbf{E}_S = \mathcal{M}_{S \leftarrow 0}(\mathbf{E}_0). \quad (7)$$

- In the quasi-static, most efficient but costly case: $S = 1$.

IEnKS: single vs multiple data assimilation



- SDA IEnKS: The observation vector are assimilated once and for all. Exact scheme.
- MDA IEnKS: The observation vector are assimilated several times and weighted by β_k within the window. Approximate scheme in nonlinear conditions.

IEnKS: single vs multiple data assimilation

► Two flavours of Multiple Data Assimilation:

- The **splitting of observations**: Following the partition $1 = \sum_{k=1}^L \beta_k$, the observation vector \mathbf{y} with prior error covariance matrix is split into L partial observation \mathbf{y}^{β_k} , with prior error covariance matrix $\beta_k^{-1}\mathbf{R}$.
It is a consistent approach in the Gaussian/linear limit, and one hopes it is still so in nonlinear conditions [Emerick and Reynolds, 2012]. Maybe more stable over long windows.
- The **multiple assimilation of each observation** with its original weights. It is (heuristically) correct but the filtering/smoothing pdf (essentially) becomes a **power of the searched pdf!**

► An extra step in the analysis:

- MDA IEnKS does not approximate *per se* the filtering pdf, but a more complex pdf.
- To approach the correct filtering/smoothing pdf, one needs an extra step, that we called the **balancing step** which reweights the observations within the data assimilation window, and perform a final analysis.

Iterative ensemble Kalman filter: accounting for sampling errors

- ▶ Finite-size versions of the iterative filter are just defined by substituting the prior:

$$\frac{N-1}{2} \mathbf{w}^T \mathbf{w} \longrightarrow \frac{N}{2} \ln \left(\varepsilon_N + \mathbf{w}^T \mathbf{w} \right). \quad (8)$$

- ▶ It implicitly (primal) or explicitly (dual) defines an optimal multiplicative inflation meant to compensate for sampling errors, related to the argmin of a dual cost function

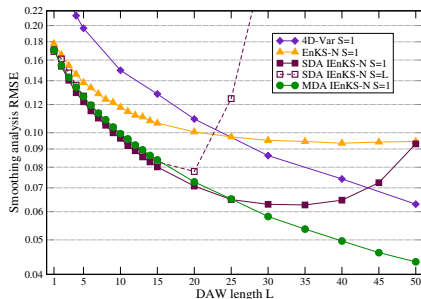
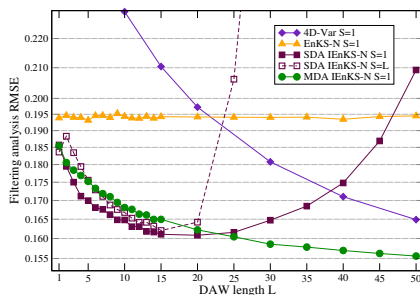
$$\lambda^* = \sqrt{\frac{N-1}{\zeta^*}}, \quad \zeta^* = \operatorname{argmin} \mathcal{D}(\zeta). \quad (9)$$

$$\mathcal{D}(\zeta) = \frac{1}{2} \delta^T \left(\mathbf{R} + \mathbf{Y} \zeta^{-1} \mathbf{Y}^T \right)^{-1} \delta + \frac{\varepsilon_N \zeta}{2} + \frac{N}{2} \ln \frac{N}{\zeta} - \frac{N}{2}. \quad (10)$$

- ▶ This adaptive scheme is quite useful since the IEnKS may require substantial inflation in strongly nonlinear regimes (because of subsequent sampling errors).
- ▶ Also (very) useful for extensive numerical experiments (no need to tune the inflation!).

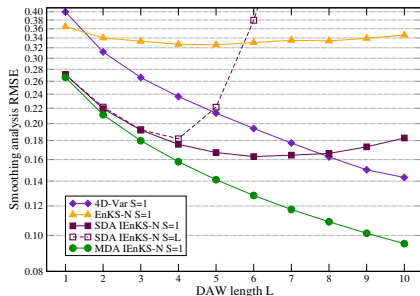
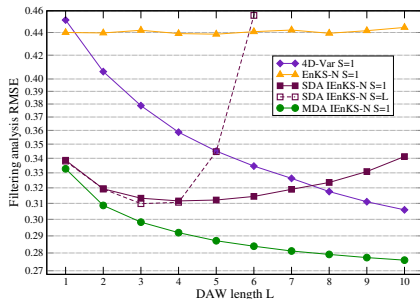
Application to the Lorenz-95 model

- ▶ Weakly nonlinear case: Lorenz-95, $M = 40$, $N = 20$, $\Delta t = 0.05$, $\mathbf{R} = \mathbf{I}$.
- ▶ Comparison of 4D-Var $S = 1$, EnKS $S = 1$, SDA IEnKS $S = 1$, SDA IEnKS $S = L$, and MDA IEnKS $S = 1$.



Application to the Lorenz-95 model

- Strongly nonlinear case: Lorenz-95, $M = 40$, $N = 20$, $\Delta t = 0.20$, $\mathbf{R} = \mathbf{I}$.
- Comparison of 4D-Var $S = 1$, EnKS $S = 1$, SDA IEnKS $S = 1$, SDA IEnKS $S = L$, and MDA IEnKS $S = 1$.

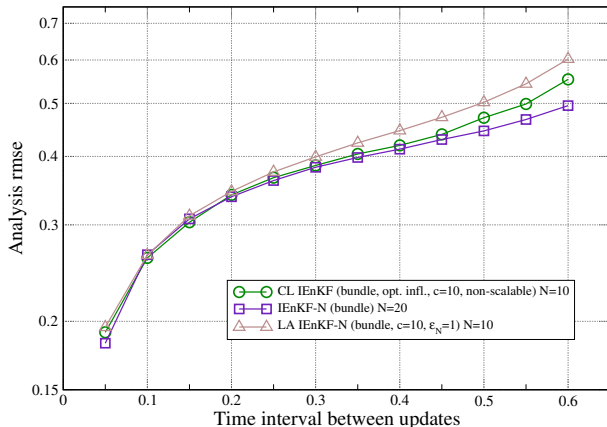


IEnKF/IEnKS: Localisation

- Localisation in an EnVar context is non-trivial because localisation and the evolution model do not commute:

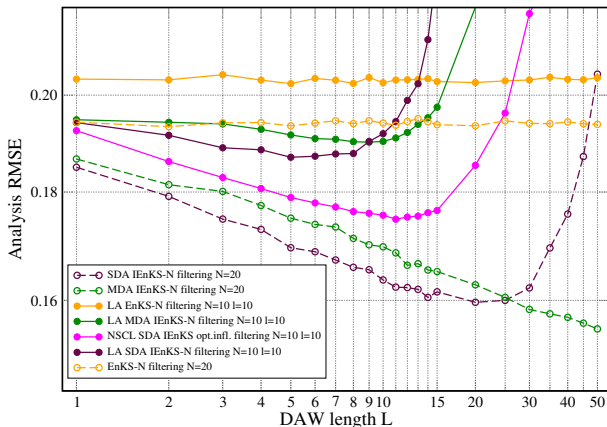
$$\mathbf{M}_{k \leftarrow 0} (\mathbf{C} \circ \mathbf{B}_0) \mathbf{M}_{k \leftarrow 0}^T \neq \mathbf{C} \circ (\mathbf{M}_{k \leftarrow 0} \mathbf{B}_0 \mathbf{M}_{k \leftarrow 0}^T). \quad (11)$$

- Local analysis of the IEnKF, and comparison with a non-scalable optimal approach.



IEnKF/IEnKS: Localisation

- Local analysis of the IEnKS, and comparison with a non-scalable optimal approach (filtering performance).



IEnKF/IEnKS: Augmented state formalism

- ▶ IEnKS treats parameters the way both 4D-Var and EnKF treat them.
- ▶ The state space is **augmented** from $\mathbf{x} \in \mathbb{R}^M$ to a vector

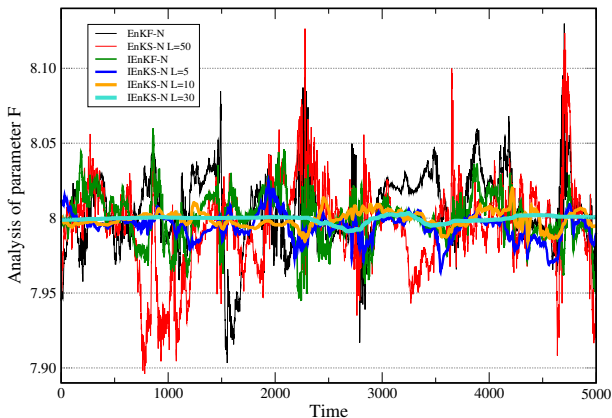
$$\mathbf{0} = \begin{pmatrix} \mathbf{x} \\ \boldsymbol{\theta} \end{pmatrix} \in \mathbb{R}^{M+P}, \quad (12)$$

Technically, there is nothing more to the joint state and parameter IEnKS than in the state IEnKS.

- ▶ A forward model needs to be introduced for the parameters:
 - For instance, the persistence model ($\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k$),
 - or some jittering such as a Brownian motion ($\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \boldsymbol{\varepsilon}_k$).

Estimation of the Lorenz-95 forcing parameter F

- F is **static** but **unknown**.

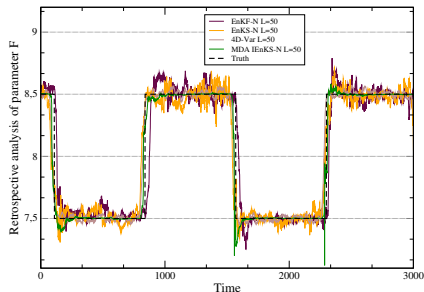
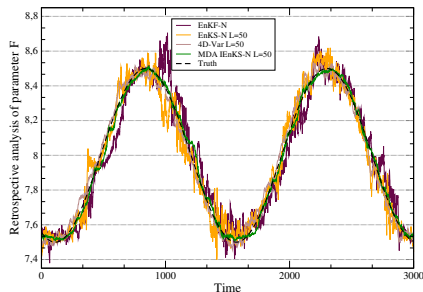


- Augmented state vector $\in \mathbb{R}^{41}$, $N = 20$. The forcing of the true model is $F = 8$.
- Evolution model for F : **persistence** assumption.

Estimation of the Lorenz-95 forcing parameter F

- ▶ The forcing parameter F is **time-varying** and **unknown**.
- ▶ **Internalised** model error (F is in the augmented state) + unaccounted **external** model error (the true F is time-varying \neq persistence assumption).

Method / F profile	Sinusoidal	Step-wise
EnKF-N	0.063	0.079
EnKS-N L=50	0.040	0.063
4D-Var L=50	0.030	0.045
MDA IEnKS-N L=50	0.020	0.031



Weak constraint IEnKS

► Weak constraint IEnKS analysis cost function: many options. But not many practical with potential. One elegant option is to perform the analysis in the following subspace:

$$k = 0, \dots, K : \quad \mathbf{x}_k = \mathbf{x}_k^j + \mathbf{A}_k^j \mathbf{w}_k \quad (16)$$

where the ensemble anomalies are defined as $\mathbf{A}_k^j = \mathbf{M}_{k \leftarrow k-1}^j \mathbf{A}_{k-1}^j$, $\mathbf{w} \in \mathbb{R}^{(K+1) \times N}$.

► Choice for the weak constraint IEnKS and comparison with the weak constraint 4D-Var and 4DEnVar:

- Weak constraint 4D-Var: $\mathbf{B} = \mathbf{L}\mathbf{D}\mathbf{L}^T$
- Weak constraint 4DEnVar: $\mathbf{B} = \mathbf{Z}\mathbf{Z}^T$, with $\mathbf{Z}^T = [\mathbf{Z}_0^T, \mathbf{Z}_1^T, \dots, \mathbf{Z}_K^T]$
- Weak constraint IEnKS: $\mathbf{B} = \mathbf{A}\mathbf{\Xi}\mathbf{A}^T$, with $\mathbf{\Xi} = \mathbf{G}\mathbf{\Gamma}\mathbf{G}^T$
 - $\mathbf{A} = \bigoplus_{k=0}^K \mathbf{A}_k$ (direct sum)
 - $\mathbf{G} = \mathbf{L}$ but with $\mathbf{M}_{k \leftarrow l} = \mathbf{I}$
 - $\mathbf{\Gamma} = (N-1)^{-1} \mathbf{I} \oplus \bigoplus_{k=1}^K q_k (\mathbf{A}_k \mathbf{A}_k^T + \mathbf{u}\mathbf{u}^T)$

► Weak constraint IEnKS somehow in-between **weak constraint 4D-Var** and **weak constraint 4DEnVar**.

Weak constraint IEnKS

- ▶ The weak constraint IEnKS has the proper limit when $\mathbf{Q}_k \rightarrow \mathbf{0}$, i.e., the IEnKS.
- ▶ Shares the same mathematical structure as 4D-EnVar but with the anomalies of the weak constraint 4D-Var $\mathbf{Z} = \mathbf{L}\mathbf{B}^{1/2}\mathbf{A}$ with \mathbf{A} a matrix of normal i.i.d., albeit in an ensemble space
- ▶ Convergence not proven yet.
- ▶ Trade-off between the singularity/smoothness of the $\mathbf{Q}_k \rightarrow \mathbf{0}$ limit (call it conditioning) and the use of $\mathbf{L}/\mathbf{L}^{-1}$ depending on whether the analysis is primal or dual. Mike Fisher's saddle point must be the right approach to follow. . . .
- ▶ Numerics so far relatively disappointing (but work in progress) on Lorenz-95. Not as performing as one would have expected. The adaptive inflation of IEnKS-N of its finite-size variant is correcting for most errors!
- ▶ If correct, then I might become sceptical about the interest of the weak-constraint formalism in general (temporal localisation better than weak constraint formalism ?)

Conclusions

- The **iterative ensemble Kalman smoother (IEnKS)** is a method to **seamlessly** combine the advantages of variational and ensemble Kalman filtering.
- The IEnKS is a generalisation of the iterative ensemble Kalman filter (IEnKF). It is an **EnVar** method. It is **flow-dependent, tangent linear and adjoint free**.
- The IEnKF/IEnKS have the potential to (significantly) outperform both the EnKF and the 4D-Var in all regimes. IEnKS already does so with toy-models.
- IEnKS is very well suited for parameter (or joint state/parameter) estimation, and does so in a very simple way via the augmented state formalism.
- **Weak constraint IEnKS** formalism in between **weak constraint 4D-Var** and **weak constraint 4DEnVar**.
- More complex reactive air quality toy-model under development in order to test the IEnKS on challenging atmospheric chemistry problems. → see Jean-Matthieu Haussaire's poster.

References

- ▶ Bocquet, M., Sakov, P., 2012. Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems. *Nonlin. Processes Geophys.* 19, 383–399.
- ▶ Bocquet, M., Sakov, P., 2014. **An iterative ensemble Kalman smoother.** *Q. J. R. Meteorol. Soc.*, 140, 1521–1535, doi: 10.1002/qj.2236.
- ▶ Bocquet, M., Sakov, P., 2013. **Joint state and parameter estimation with an iterative ensemble Kalman smoother.** *Nonlin. Processes Geophys.*, 20, 803–818.
- ▶ Gu, Y., Oliver, D. S., 2007. An iterative ensemble Kalman filter for multiphase fluid flow data assimilation. *SPE Journal* 12, 438–446.
- ▶ Desroziers, G., Camino, J.-T., Berre, L., 2014. **4DEnVar: link with 4D state formulation of variational assimilation and different possible implementations.** *Q. J. R. Meteorol. Soc.*, 140, 2097–2110, doi: 10.1002/qj.2325.
- ▶ Hunt, B. R., Kostelich, E. J., Szunyogh, I., 2007. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D* 230, 112–126.
- ▶ Sakov, P., Oliver, D., Bertino, L., 2012. An iterative EnKF for strongly nonlinear systems. *Mon. Wea. Rev.* 140, 1988–2004.
- ▶ Yang, S.-C., E. Kalnay, and B. Hunt, 2012. Handling nonlinearity and non-Gaussianity in the Ensemble Kalman Filter: Experiments with the three-variable Lorenz model. *Mon. Wea. Rev.*, 140, 2628–2646.