Enhanced Adaptive Inflation Algorithm for Ensemble Filters

Mohamad (Moha) E. Gharamti, NCAR, Boulder, CO

12TH International EnKF Workshop, Os, Norway

E-mail: gharamti@ucar.edu

DAReS Group: http://www.image.ucar.edu/DAReS/DART/

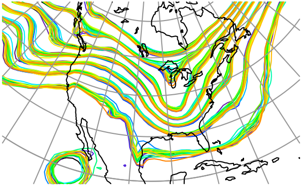
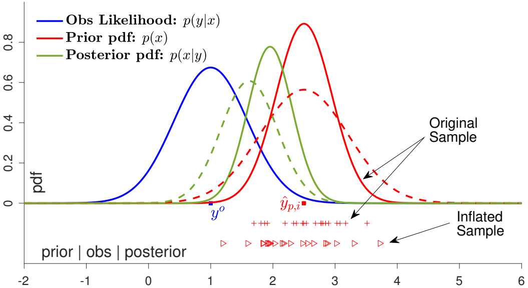


Table of Contents



1.1 Background

4 distinct inflation categories:

- Background covariance inflation
  1. Additive inflation, [Mitchell and Houtekamer 2000]
  2. Multiplicative inflation, [Anderson and Anderson 1999]

1.1 Background

4 distinct inflation categories:

- Background covariance inflation
  1. Additive inflation, [Mitchell and Houtekamer 2000]
  2. Multiplicative inflation, [Anderson and Anderson 1999]
- Observation error variance inflation
  - Moderation, [Sakov et al. 2012; Karspeck 2016]
  - Adaptive, [Minamide and Zhang 2017]

1.1 Background

4 distinct inflation categories:

- Background covariance inflation
  1. Additive inflation, [Mitchell and Houtekamer 2000]
  2. Multiplicative inflation, [Anderson and Anderson 1999]
- Observation error variance inflation
  - Moderation, [Sakov et al. 2012; Karspeck 2016]
  - Adaptive, [Minamide and Zhang 2017]
- Relaxation to prior
  - Perturbation (RTPP), [Zhang et al. 2004]
  - Spread (RTPS), [Whitaker and Hamill 2012; Ying and Zhang 2015]

1.1 Background

4 distinct inflation categories:

- Background covariance inflation
  1. Additive inflation, [Mitchell and Houtekamer 2000]
  2. Multiplicative inflation, [Anderson and Anderson 1999]
- Observation error variance inflation
  - Moderation, [Sakov et al. 2012; Karspeck 2016]
  - Adaptive, [Minamide and Zhang 2017]
- Relaxation to prior
  - Perturbation (RTPP), [Zhang et al. 2004]
  - Spread (RTPS), [Whitaker and Hamill 2012; Ying and Zhang 2015]
- Ensemble modification
  - its size, [Uzunoglu 2007]
  - its physical and model-based nature [Meng and Zhang 2007; Berner et al. 2009]

1.1 Background

4 distinct inflation categories:

- Background covariance inflation
  1. Additive inflation, [Mitchell and Houtekamer 2000]
  2. Multiplicative inflation, [Anderson and Anderson 1999]
- Observation error variance inflation
  - Moderation, [Sakov et al. 2012; Karspeck 2016]
  - Adaptive, [Minamide and Zhang 2017]
- Relaxation to prior
  - Perturbation (RTPP), [Zhang et al. 2004]
  - Spread (RTPS), [Whitaker and Hamill 2012; Ying and Zhang 2015]
- Ensemble modification
  - its size, [Uzunoglu 2007]
  - its physical and model-based nature [Meng and Zhang 2007; Berner et al. 2009]
- Others: EnTLHF [Luo and Hoteit 2011], EnKF-N [Bocquet et al. 2015]

More Fun Stuff

This is already getting boring! Forget about inflation, we have a new DART release ..

Native netCDF support, less filesystem I/O, better scaling, better computational performance, supports huge memory models. Simplified organizational layout. Supports ROMS, CICE, WRF-CHEM ...

http://www.image.ucar.edu/DAReS/DART/DART_download

1.2 Inflation and Innovation Statistics

Given a scalar variable with sample x_i and observation y

$$x_b = \frac{1}{N_e}\sum_{i=1}^{N_e} x_i, \qquad \widehat{\sigma_b}^2 = \frac{1}{N_e - 1}\sum_{i=1}^{N_e}(x_i - x_b)^2 \tag{1}$$

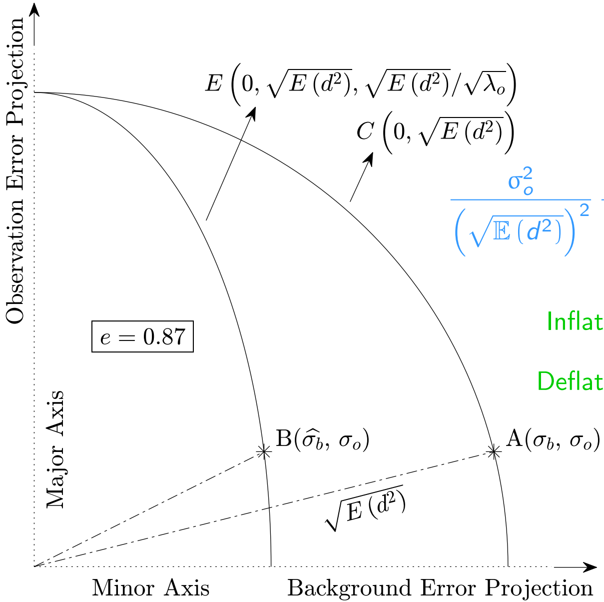
Following Desroziers et al. (2005)

$$\begin{aligned} d &= y - x_b = \varepsilon_o + (x_t - x_b) = \varepsilon_o + \varepsilon_b, & (2)\\ \mathbb{E}(d) &= \mathbb{E}(\varepsilon_o) + \mathbb{E}(\varepsilon_b) = 0, & (3)\\ \mathbb{E}\left(d^2\right) &= \mathbb{E}\left(\varepsilon_o^2\right) + \mathbb{E}\left(\varepsilon_b^2\right) + 2\mathbb{E}(\varepsilon_o\varepsilon_b) = \sigma_o^2 + \sigma_b^2. & (4) \end{aligned}$$

Impose $\sigma_b^2 = \lambda_o \widehat{\sigma_b}^2$. Assuming a correctly specified σ_o^2

$$\Rightarrow \boxed{\lambda_o = \frac{\mathbb{E}\left(d^2\right) - \sigma_o^2}{\widehat{\sigma_b}^2}} \tag{5}$$

1.3 Geometrical Interpretation



$$\frac{\sigma_o^2}{\left(\sqrt{\mathbb{E}\left(d^2\right)}\right)^2} + \frac{\widehat{\sigma_b}^2}{\left(\sqrt{\mathbb{E}\left(d^2\right)}/\sqrt{\lambda_o}\right)^2} = 1$$

Inflation $\rightarrow$ $e = \sqrt{1 - \frac{1}{\lambda_o}}$

Deflation $\rightarrow$ $e = \sqrt{1 - \lambda_o}$

1.4 Anderson (2009), A09 hereafter

$$p(\lambda|d) \propto p(d|\lambda) \cdot p(\lambda) \tag{6}$$

- Prior marginal distribution: $N\left(\lambda_b, \sigma^2_{\lambda_b}\right)$

1.4 Anderson (2009), A09 hereafter

$$p\left(\lambda|d\right) \propto p\left(d|\lambda\right) \cdot p\left(\lambda\right) \tag{6}$$

- Prior marginal distribution: $N\left(\lambda_b, \sigma^2_{\lambda_b}\right)$
- Likelihood: $d \sim N\left(0, \theta^2\right)$, with $\theta^2 = \lambda_o^k \widehat{\sigma_b}^2 + \sigma_o^2$
  - Spread the information across all variables

$$\begin{aligned} r &= corr(x^o, x^k) \quad k = 1, 2, \ldots, N_x && (7) \\ \lambda_o^k &= \left[\gamma\left(\lambda_b^k - 1\right) + 1\right]^2, \quad \gamma = \kappa|r| && (8) \end{aligned}$$

  - $p\left(d|\lambda\right)$ is not Gaussian in λ!

1.4 Anderson (2009), A09 hereafter

$$p(\lambda|d) \propto p(d|\lambda) \cdot p(\lambda) \tag{6}$$

- Prior marginal distribution: $N\left(\lambda_b, \sigma^2_{\lambda_b}\right)$
- Likelihood: $d \sim N\left(0, \theta^2\right)$, with $\theta^2 = \lambda_o^k \widehat{\sigma_b}^2 + \sigma_o^2$
  - Spread the information across all variables

$$\begin{aligned} r &= corr(x^o, x^k) \quad k = 1, 2, \ldots, N_x & (7) \\ \lambda_o^k &= \left[\gamma\left(\lambda_b^k - 1\right) + 1\right]^2, \quad \gamma = \kappa|r| & (8) \end{aligned}$$

  - $p(d|\lambda)$ is not Gaussian in λ!
- Posterior:

$$p(\lambda|d) \propto \frac{1}{2\pi\theta\sigma_{\lambda_b}} \exp\left[-\frac{(\lambda-\lambda_b)^2}{2\sigma^2_{\lambda_b}} - \frac{d^2}{2\theta^2}\right] \tag{9}$$

2.1 Enhanced Scheme: The Likelihood I

- What happens when σ_o^2 is incorrectly specified, or in general, when $\lambda_o < 0$?

2.1 Enhanced Scheme: The Likelihood I

- What happens when σ_o^2 is incorrectly specified, or in general, when $\lambda_o < 0$?
- The following relations should hold:

$$\begin{aligned} \sqrt{\mathbb{E}\left(d^2\right)} + \sqrt{\lambda_o \widehat{\sigma_b}} &> \sigma_o, \quad (10) \\ \sqrt{\mathbb{E}\left(d^2\right)} + \sigma_o &> \sqrt{\lambda_o \widehat{\sigma_b}}, \quad (11) \\ \sqrt{\lambda_o \widehat{\sigma_b}} + \sigma_o &> \sqrt{\mathbb{E}\left(d^2\right)} \quad (12) \end{aligned}$$

otherwise, these components *roughly* won't form a triangle!

2.1 Enhanced Scheme: The Likelihood I

- What happens when σ_o^2 is incorrectly specified, or in general, when $\lambda_o < 0$?
- The following relations should hold:

$$\begin{aligned} \sqrt{\mathbb{E}\left(d^2\right)} + \sqrt{\lambda_o \widehat{\sigma_b}} &> \sigma_o, & (10) \\ \sqrt{\mathbb{E}\left(d^2\right)} + \sigma_o &> \sqrt{\lambda_o \widehat{\sigma_b}}, & (11) \\ \sqrt{\lambda_o \widehat{\sigma_b}} + \sigma_o &> \sqrt{\mathbb{E}\left(d^2\right)} & (12) \end{aligned}$$

otherwise, these components *roughly* won't form a triangle!

- Li et al. (2009) estimated both σ_o^2 and σ_b^2
- Wang et al. (2007); $\mathbb{E}\left(d^2\right)$ not well represented with a small number of observations. The following assumption is imposed

$$\frac{1}{N_j} \sum_{j=1}^{N_j} d\left(t_j\right)^2 \approx \mathbb{E}\left(d^2\right) \quad (13)$$

2.1 Enhanced Scheme: The Likelihood II

- A sample distance (innovation) can be expanded as follows

$$
\begin{aligned}
d_i &= y - x_i^k = \varepsilon_o + \varepsilon_b - \widetilde{x}_i \\
d_i^2 &= (\varepsilon_o)^2 + (\varepsilon_b)^2 + \widetilde{x}_i^{\,2} + 2\varepsilon_o\varepsilon_b - 2\varepsilon_o\widetilde{x}_i - 2\varepsilon_b\widetilde{x}_i
\end{aligned}
$$

2.1 Enhanced Scheme: The Likelihood II

- A sample distance (innovation) can be expanded as follows

$$
\begin{aligned}
d_i &= y - x_i^k = \varepsilon_o + \varepsilon_b - \widetilde{x_i} \\
d_i^2 &= (\varepsilon_o)^2 + (\varepsilon_b)^2 + \widetilde{x_i}^2 + 2\varepsilon_o\varepsilon_b - 2\varepsilon_o\widetilde{x_i} - 2\varepsilon_b\widetilde{x_i}
\end{aligned}
$$

- Now, average over all ensemble members

$$
\frac{1}{N_e}\sum_{i=1}^{N_e} d_i^2 = (\varepsilon_o)^2 + (\varepsilon_b)^2 + \frac{N_e - 1}{N_e}\widehat{\sigma_b}^2 + 2\varepsilon_o\varepsilon_b
$$

2.1 Enhanced Scheme: The Likelihood II

- A sample distance (innovation) can be expanded as follows

$$
\begin{aligned}
d_i &= y - x_i^k = \varepsilon_o + \varepsilon_b - \widetilde{x}_i \\
d_i^2 &= (\varepsilon_o)^2 + (\varepsilon_b)^2 + \widetilde{x}_i^{\,2} + 2\varepsilon_o\varepsilon_b - 2\varepsilon_o\widetilde{x}_i - 2\varepsilon_b\widetilde{x}_i
\end{aligned}
$$

- Now, average over all ensemble members

$$
\frac{1}{N_e}\sum_{i=1}^{N_e} d_i^2 = (\varepsilon_o)^2 + (\varepsilon_b)^2 + \frac{N_e - 1}{N_e}\widehat{\sigma_b}^2 + 2\varepsilon_o\varepsilon_b
$$

- Take the expectation of both sides and workout the algebra

$$
\mathbb{E}\left(d^2\right) + \widehat{\sigma_b}^2 = \sigma_o^2 + \sigma_b^2 + \frac{N_e - 1}{N_e}\widehat{\sigma_b}^2
$$

2.1 Enhanced Scheme: The Likelihood II

- A sample distance (innovation) can be expanded as follows

$$\begin{aligned} d_i &= y - x_i^k = \varepsilon_o + \varepsilon_b - \widetilde{x}_i \\ d_i^2 &= (\varepsilon_o)^2 + (\varepsilon_b)^2 + \widetilde{x}_i^{\,2} + 2\varepsilon_o\varepsilon_b - 2\varepsilon_o\widetilde{x}_i - 2\varepsilon_b\widetilde{x}_i \end{aligned}$$

- Now, average over all ensemble members

$$\frac{1}{N_e}\sum_{i=1}^{N_e} d_i^2 = (\varepsilon_o)^2 + (\varepsilon_b)^2 + \frac{N_e - 1}{N_e}\widehat{\sigma_b}^2 + 2\varepsilon_o\varepsilon_b$$

- Take the expectation of both sides and workout the algebra

$$\mathbb{E}\left(d^2\right) + \widehat{\sigma_b}^2 = \sigma_o^2 + \sigma_b^2 + \frac{N_e - 1}{N_e}\widehat{\sigma_b}^2$$

- Introduce a new inflation factor, λ_o^*

$$\boxed{\lambda_o^* = \frac{\mathbb{E}\left(d^2\right) - \sigma_o^2}{\widehat{\sigma_b}^2} + \frac{1}{N_e} = \lambda_o + \frac{1}{N_e}} \qquad (14)$$

2.1 Enhanced Scheme: The Likelihood III

- Modified variance:

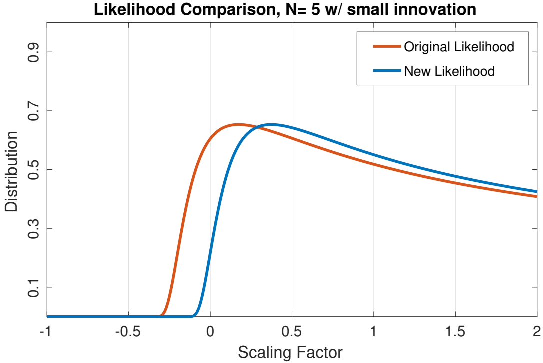
$$\theta^2 = \left(\left[\gamma \left(\lambda_b^k - 1 \right) + 1 \right]^2 - \frac{1}{N_e} \right) \widehat{\sigma_b}^2 + \sigma_o^2 \qquad (15)$$

2.1 Enhanced Scheme: The Likelihood III

- Modified variance:

$$\theta^2 = \left(\left[\gamma \left(\lambda_b^k - 1 \right) + 1 \right]^2 - \frac{1}{N_e} \right) \widehat{\sigma_b}^2 + \sigma_o^2 \qquad (15)$$

- Bocquet et al. (2015), predictive prior distribution
- Anderson et al. (2012), sampling error correction



2.2 Enhanced Scheme: The Prior I

- Instead of a Gaussian, describe the inflation prior by an inverse Gamma (IG) distribution. Why?
  - Restriction: to positive and not very close to zero values
  - More stable + cleaner code

2.2 Enhanced Scheme: The Prior I

- Instead of a Gaussian, describe the inflation prior by an inverse Gamma (IG) distribution. Why?
  - Restriction: to positive and not very close to zero values
  - More stable + cleaner code
- Best possible choices for α and β?

$$p(\lambda) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{-\alpha-1} \exp\left[-\frac{\beta}{\lambda}\right] \tag{16}$$

2.2 Enhanced Scheme: The Prior I

- Instead of a Gaussian, describe the inflation prior by an inverse Gamma (IG) distribution. Why?
  - Restriction: to positive and not very close to zero values
  - More stable + cleaner code
- Best possible choices for α and β?

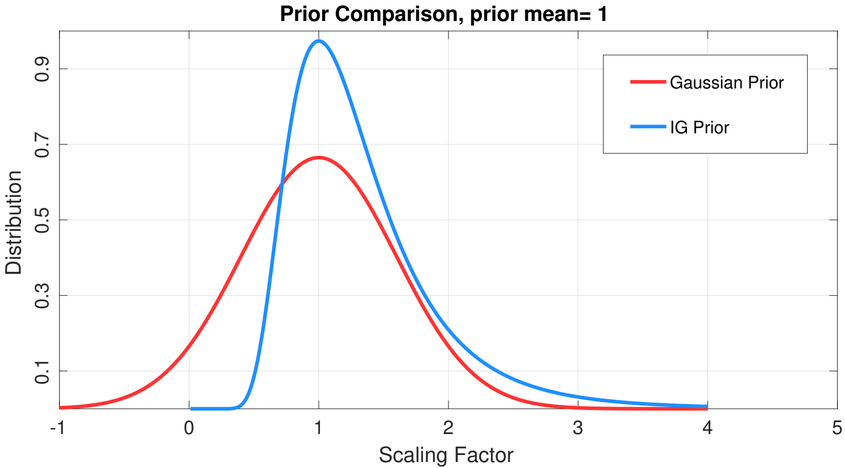
$$p(\lambda) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{-\alpha-1} \exp\left[-\frac{\beta}{\lambda}\right] \tag{16}$$

- In DART; impose minimal changes to current implementation and facilitate the life of the user
- Start with a Gaussian. Use λ_b and σ_λ to find α and β

$$\lambda_b = \frac{\beta}{\alpha+1} \equiv \mathrm{Mode}_{\mathrm{IG}} \tag{17}$$

$$\sigma^2_{\lambda_b} = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}, \qquad \alpha > 2 \tag{18}$$

2.2 Enhanced Scheme: The Prior II



$$\beta = \frac{\overline{b}}{3} + \frac{\lambda_b^3 + 7\sigma_{\lambda_b}^2 \lambda_b}{3\sigma_{\lambda_b}^2} + \frac{\lambda_b^2 \left(\sigma_{\lambda_b}^4 + 14\sigma_{\lambda_b}^2 \lambda_b^2 + \lambda_b^4\right)}{3\sigma_{\lambda_b}^4 \overline{b}} \tag{19}$$

2.3 Enhanced Scheme: The Posterior I

- The new posterior

$$\frac{\beta^{\alpha}\lambda^{-\alpha-1}}{\sqrt{2\pi}\theta\Gamma(\alpha)} \exp\left[-\frac{d^2}{2\theta^2} - \frac{\beta}{\lambda}\right] \tag{20}$$

2.3 Enhanced Scheme: The Posterior I

- The new posterior

$$\frac{\beta^{\alpha}\lambda^{-\alpha-1}}{\sqrt{2\pi}\theta\Gamma(\alpha)}\exp\left[-\frac{d^2}{2\theta^2}-\frac{\beta}{\lambda}\right] \tag{20}$$

- To find the updated inflation or the mode, i.e., λ_u

$$\left(1-\frac{\lambda_b}{\beta}\right)\lambda^2+\left(\frac{\bar{\ell}}{\ell'}-2\lambda_b\right)\lambda+\left(\lambda_b^2-\frac{\bar{\ell}}{\ell'}\lambda_b\right)=0 \tag{21}$$

2.3 Enhanced Scheme: The Posterior I

- The new posterior

$$\frac{\beta^{\alpha}\lambda^{-\alpha-1}}{\sqrt{2\pi}\theta\Gamma(\alpha)}\exp\left[-\frac{d^2}{2\theta^2}-\frac{\beta}{\lambda}\right] \tag{20}$$

- To find the updated inflation or the mode, i.e., λ_u

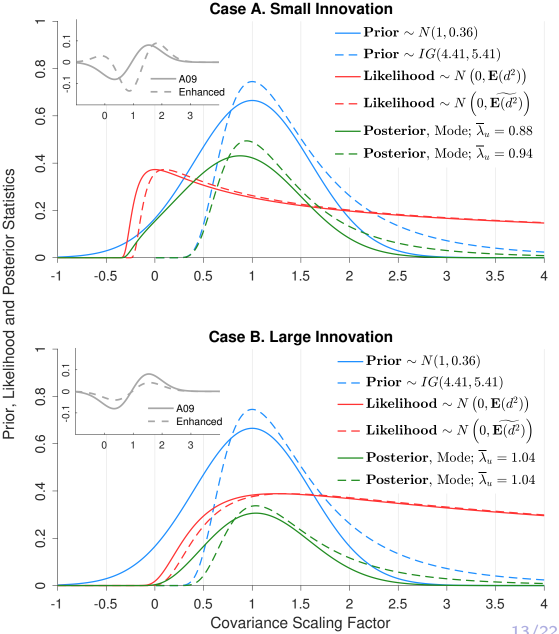
$$\left(1-\frac{\lambda_b}{\beta}\right)\lambda^2+\left(\frac{\bar{\ell}}{\ell'}-2\lambda_b\right)\lambda+\left(\lambda_b^2-\frac{\bar{\ell}}{\ell'}\lambda_b\right)=0 \tag{21}$$

- $p(\lambda|d)$ is assumed IG, numerically get the posterior variance

$$\begin{aligned} R &= \frac{p(\lambda|d)\big|_{\lambda=1}}{p(\lambda|d)\big|_{\lambda=2}}, & (22)\\ \omega &= \frac{1}{2}-\frac{\log(2)}{\lambda_u}, & (23)\\ \beta_u &= \omega^{-1}\log(R), & (24)\\ \alpha_u &= \frac{\beta_u}{\lambda_u}-1 & (25)\end{aligned}$$

2.3 Enhanced Scheme: The Posterior II

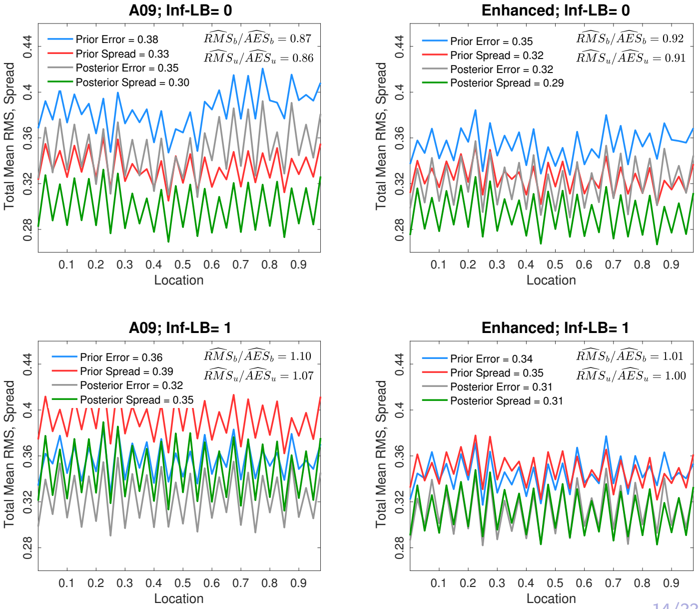
- A09 tends to deflate more when the ensemble mean and the obs. are close
- For very large distances, the enhanced scheme responds more aggresively



3.1 Experiments: L96

Inflation Cap: RMS & Spread

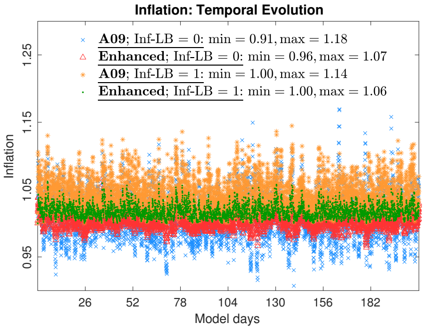
- $LB = 0 \rightarrow 8\%$
- $LB = 1 \rightarrow 4\%$
- Consistency

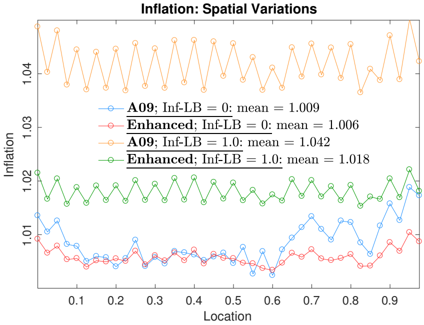


3.1 Experiments: L96

Inflation Cap: Inflation in space & time

- Range for λ
- λ for observed and unobserved variables
- Heterogeneity of λ





3.1 Experiments: L96

Robustness

I. Localization Sensitivity

	A09		**Enhanced**	
Half-width: c	$LB = 0$	$LB = 1$	$LB = 0$	$LB = 1$
0.10	0.4283	0.4032	0.4164	0.4015
0.20	0.3681	0.3436	0.3408	0.3315
0.30	0.3644	0.3267	0.3329	0.3191
0.40	0.3703	0.3260	0.3372	0.3200
0.50	0.3728	0.3211	0.3415	0.3210
0.60	0.3881	0.3302	0.3642	0.3288
0.70	0.3887	0.3336	0.3750	0.3315
0.80	0.3951	0.3411	0.3687	0.3361

3.1 Experiments: L96

Robustness

II. Forcing Sensitivity (Model Error)

	A09		**Enhanced**	
Forcing: F	$LB = 0$	$LB = 1$	$LB = 0$	$LB = 1$
1	2.2170	2.2080	2.1522	2.1318
3	1.7435	1.7463	1.6486	1.6503
5	1.3179	1.3129	1.2234	1.2168
7	0.8078	0.8030	0.7414	0.7111
9	0.8486	0.8319	0.7900	0.7437
11	1.5322	1.5244	1.4113	1.3978
13	2.1493	2.1666	1.9809	2.0519
15	2.8118	2.7636	2.6068	2.5739

3.2 Experiments: Idealized Atmospheric Model

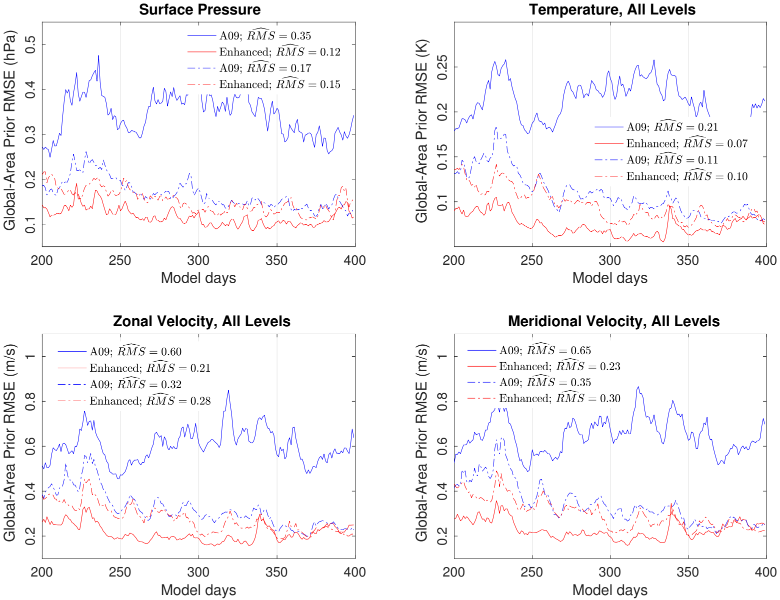
Model Setup

- Modified version of the dynamical core used in the GFDL global atmospheric model
- Prognostic variables: PS, T, U, V
- 30 latitudes & 60 longitudes
- 5 vertical layers
- $N_x \sim 3 \times 10^5$
- 1 hr time step

- Twin Experiments: 400 days
- Observe 300 PS locations every day
- Localization half-width = 0.2 rad
- 20 ensemble members, only the last 200 days are used for diagnostics

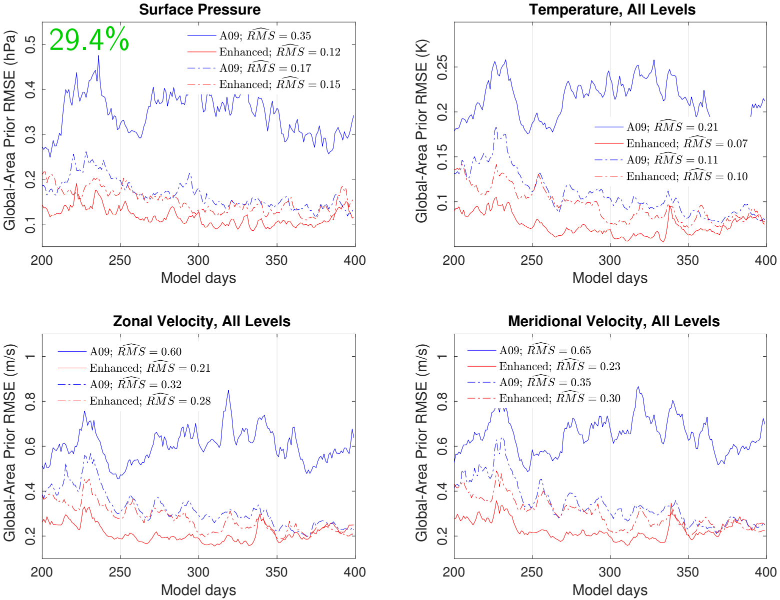
3.2 Experiments: Idealized Atmospheric Model

Fixed σ_λ: RMS



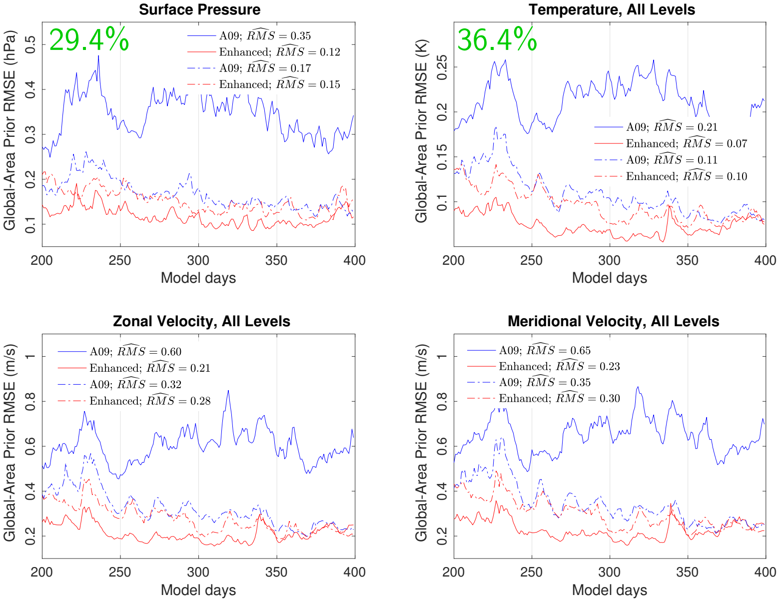
3.2 Experiments: Idealized Atmospheric Model

Fixed σ_λ: RMS



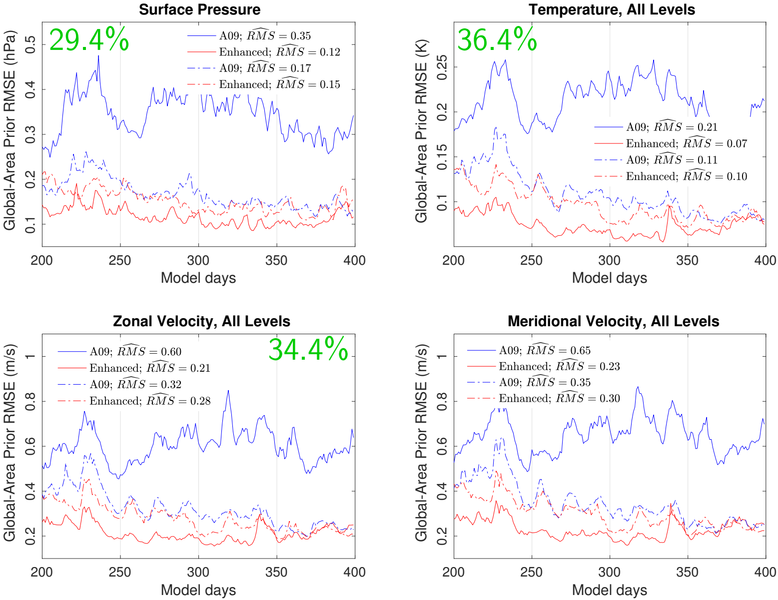
3.2 Experiments: Idealized Atmospheric Model

Fixed σ_λ: RMS



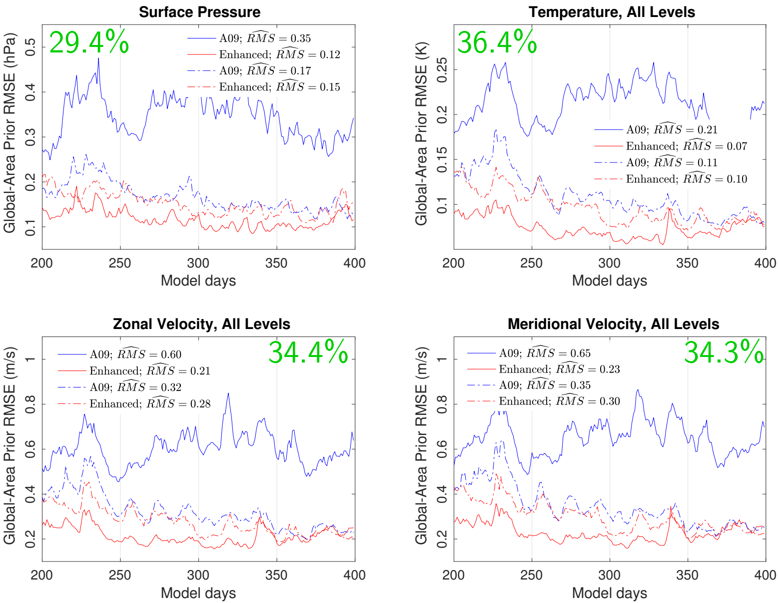
3.2 Experiments: Idealized Atmospheric Model

Fixed σ_λ: RMS



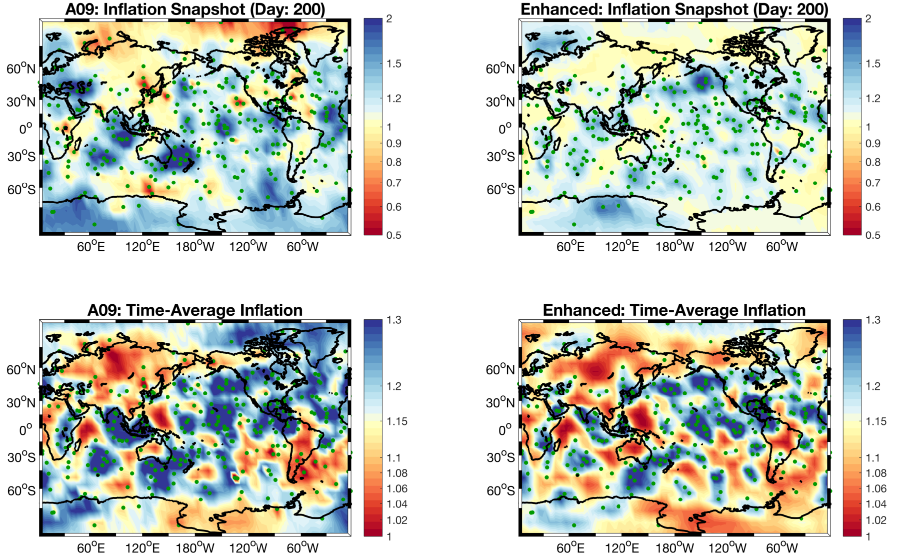
3.2 Experiments: Idealized Atmospheric Model

Fixed σ_λ: RMS



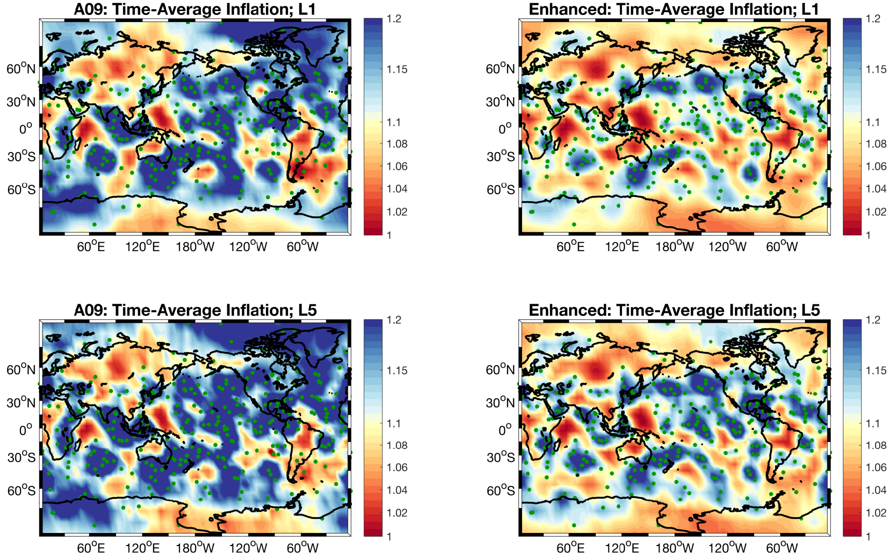
3.2 Experiments: Idealized Atmospheric Model

Fixed σ_λ: Inflation



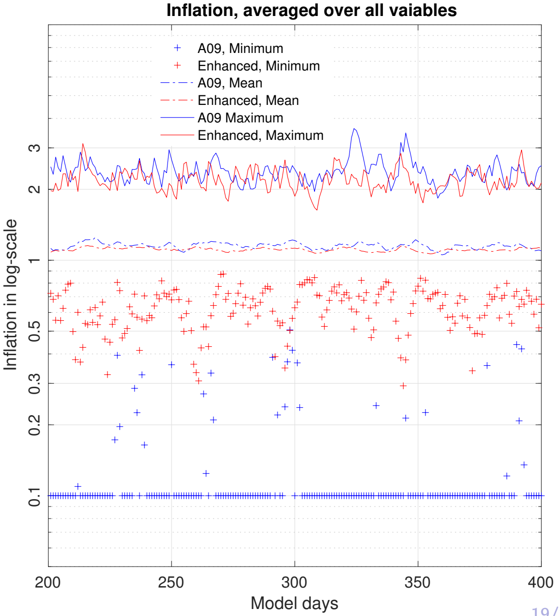
3.2 Experiments: Idealized Atmospheric Model

Fixed σ_λ: Inflation



3.2 Experiments: Idealized Atmospheric Model

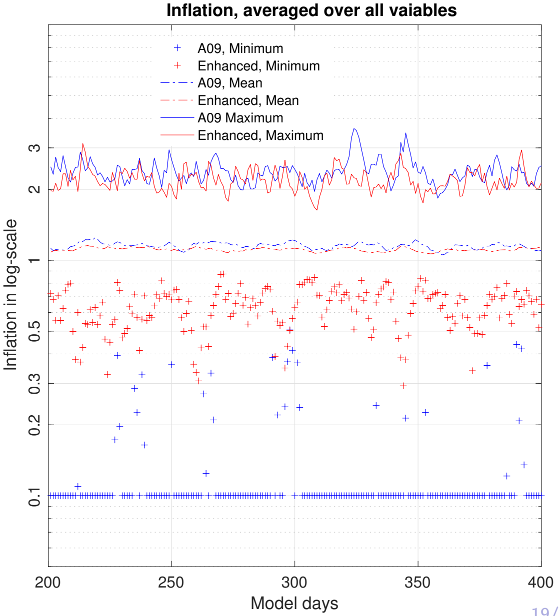
Fixed σ_λ: Inflation



3.2 Experiments: Idealized Atmospheric Model

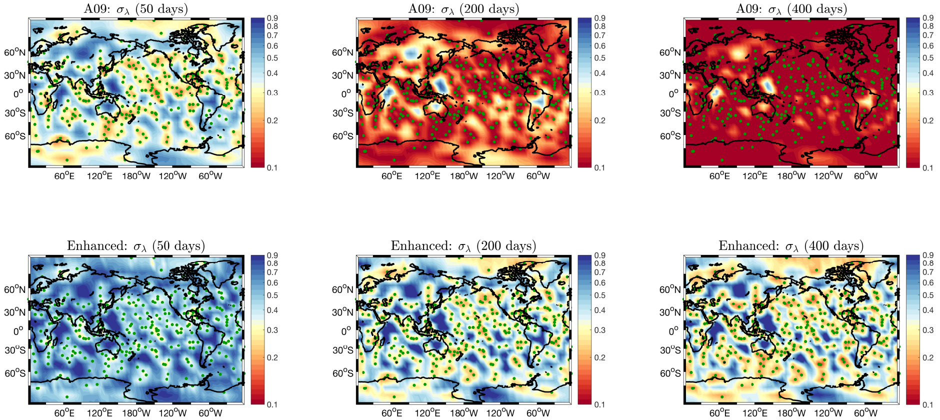
Fixed σ_λ: Inflation

- Ensemble statistics matching the innovations, is not longer satisfied
- Deflation helps get rid of the extra artificial spread when it's not needed anymore
- Enhanced scheme suggests a gentle and moderate use of deflation



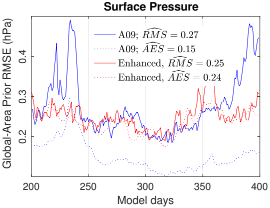
3.2 Experiments: Idealized Atmospheric Model

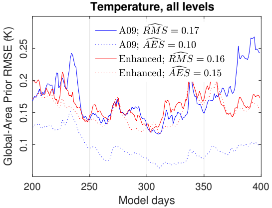
Varying inflation std: $0.1 < \sigma_\lambda < 0.9$

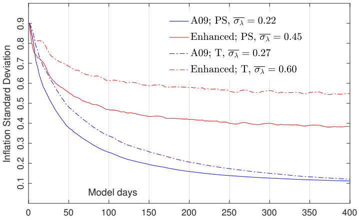


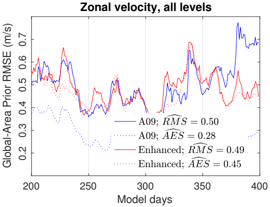
3.2 Experiments: Idealized Atmospheric Model

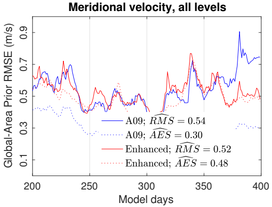
Varying inflation std: $0.1 < \sigma_\lambda < 0.9$





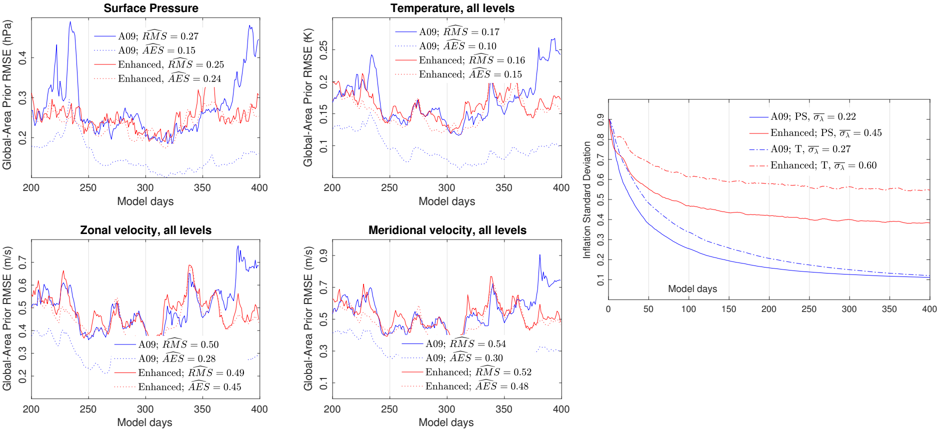






3.2 Experiments: Idealized Atmospheric Model

Varying inflation std: $0.1 < \sigma_\lambda < 0.9$



- A09's algorithm becomes ineffective after 200 days. Enhanced offers *more room* for correction
- Useful feature, especially if the observation network changes in time

4. Conclusion

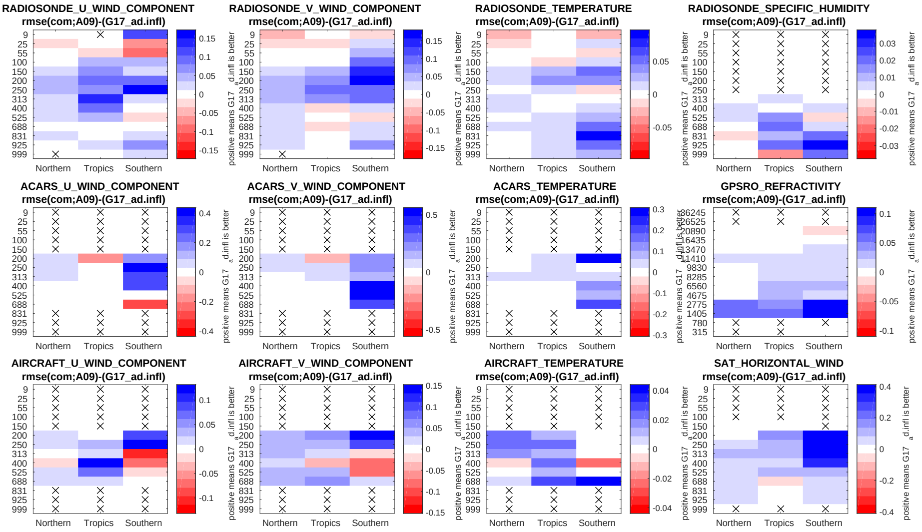
- An enhanced spatially and temporally varying adaptive prior covariance inflation
- With no cap on the inflation, the enhanced scheme produced the most accurate and best consistency of the estimates
- Original scheme's Bayes update lead to extreme tiny and large inflation values
- A moderate and not-so-frequent use of deflation can be useful
- Shortcomings: likelihood, forecast model, posterior inflation?

4. Conclusion

- An enhanced spatially and temporally varying adaptive prior covariance inflation
- With no cap on the inflation, the enhanced scheme produced the most accurate and best consistency of the estimates
- Original scheme's Bayes update lead to extreme tiny and large inflation values
- A moderate and not-so-frequent use of deflation can be useful
- Shortcomings: likelihood, forecast model, posterior inflation?

- 6-hour forecast experiments using the National Center for Atmospheric Research Community Atmospheric Model (CAM) are currently being conducted. Wind and temperature observations from radiosondes, ACARS, and aircraft along with GPS radio occultation observations are assimilated

Preliminaries, CESM-DART: CAM component



Preliminaries, CESM-DART: CAM component