



# Adaptive (Prior|Posterior?) Inflation for Ensemble Kalman Filters

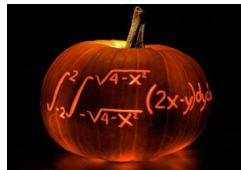
**Moha Gharamti**, NCAR, Boulder, CO

**University of Colorado Boulder**

Data Assimilation [APPM 5510]

[gharamti@ucar.edu](mailto:gharamti@ucar.edu)

Nov. 1st, 2019



# Table of Contents

- 1 **The Ensemble Kalman Filter**
- 2 **Covariance Inflation**
  - Prior Inflation
  - Posterior Inflation
- 3 **Applications**
  - Low-order Models: Lorenz-96|63
  - High-order Model: CAM
- 4 **Conclusions**

# The EnKF

- State estimation tool
- Given an observation  $y$  of state  $x$ , use Bayes:

$$p(x_k|y_k, Y_{k-1}) \approx p(x_k|Y_{k-1}) \cdot p(y_k|x_k, Y_{k-1}) \quad (1)$$

$Y_k: \{y_1, y_2, \dots, y_{k-1}, y_k\}$ ,       $k$  is time index

# The EnKF

- State estimation tool
- Given an observation  $y$  of state  $x$ , use Bayes:

$$p(x_k|y_k, Y_{k-1}) \approx p(x_k|Y_{k-1}) \cdot p(y_k|x_k, Y_{k-1}) \quad (1)$$

$Y_k: \{y_1, y_2, \dots, y_{k-1}, y_k\}$ ,  $k$  is time index

- Successive Forecast and Update (Analysis) stages:

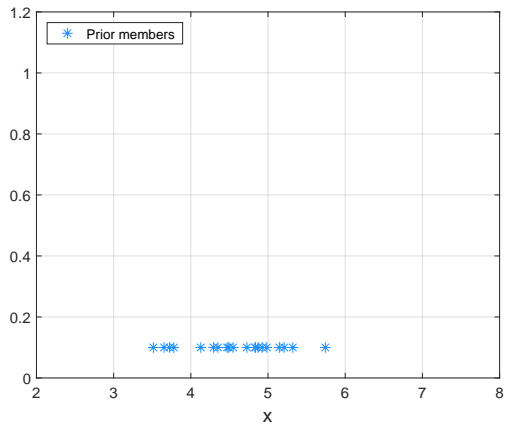
$$x_f^i = \mathcal{M}(x_a^i)$$

$$\bar{x}_f = \frac{1}{N} \sum_{i=1}^N x_f^i, \quad \hat{\sigma}_f^2 = \frac{1}{N-1} \sum_{i=1}^N (x_f^i - \bar{x}_f) (x_f^i - \bar{x}_f)^T$$

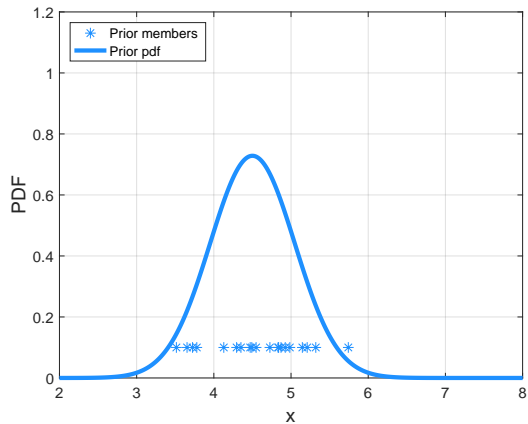
$$x_a^i = x_f^i + \frac{\hat{\sigma}_f^2}{\sigma_o^2 + \hat{\sigma}_f^2} (y^i - x_f^i)$$

$N$  is the ensemble size

# The EnKF cont.

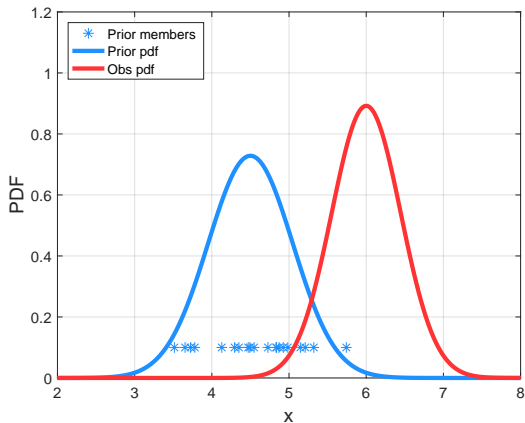


# The EnKF cont.



$$\bar{X}_f = 4.5, \hat{\sigma}_f^2 = 0.30$$

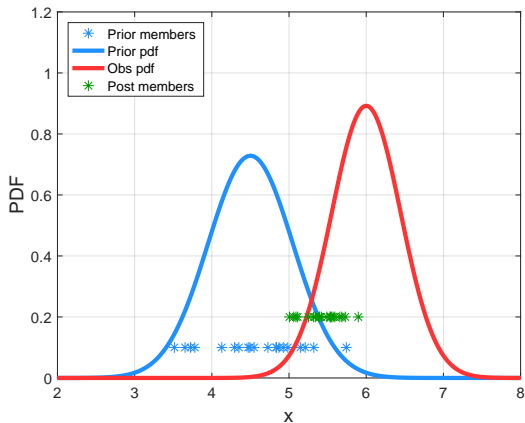
# The EnKF cont.



$$\bar{x}_f = 4.5, \hat{\sigma}_f^2 = 0.30$$

$$y = 6.0, \sigma_o^2 = 0.20$$

# The EnKF cont.

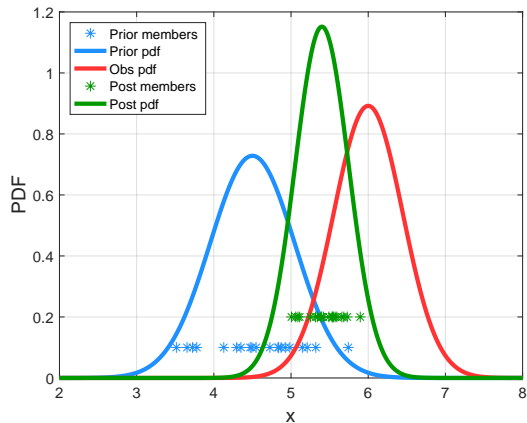


$$\bar{x}_f = 4.5, \hat{\sigma}_f^2 = 0.30$$

$$y = 6.0, \sigma_o^2 = 0.20$$



# The EnKF cont.



$$\bar{x}_f = 4.5, \hat{\sigma}_f^2 = 0.30$$

$$y = 6.0, \sigma_o^2 = 0.20$$

$$\bar{x}_a = 5.43, \hat{\sigma}_a^2 = 0.06$$

## Some Drawbacks

### 1. Sampling Errors:

- Ideal scenario:  $N = \infty$ ;  $\sigma_f^2$  is out-of-reach!
- $\hat{\sigma}_f^2$  depends on the ensemble size

$$\lim_{N \rightarrow \infty} \hat{\sigma}_f^2(N) = \sigma_f^2$$

- When  $N$  is small,  $\hat{\sigma}_f^2$  may **underestimate**  $\sigma_f^2$

## Some Drawbacks

### 1. Sampling Errors:

- Ideal scenario:  $N = \infty$ ;  $\sigma_f^2$  is out-of-reach!
- $\hat{\sigma}_f^2$  depends on the ensemble size

$$\lim_{N \rightarrow \infty} \hat{\sigma}_f^2(N) = \sigma_f^2$$

- When  $N$  is small,  $\hat{\sigma}_f^2$  may **underestimate**  $\sigma_f^2$

### 2. Model Errors:

- Biased model produces ensemble predictions that are far from the observations. May lead to **ensemble collapse**

## Some Drawbacks

### 1. Sampling Errors:

- Ideal scenario:  $N = \infty$ ;  $\sigma_f^2$  is out-of-reach!
- $\hat{\sigma}_f^2$  depends on the ensemble size

$$\lim_{N \rightarrow \infty} \hat{\sigma}_f^2(N) = \sigma_f^2$$

- When  $N$  is small,  $\hat{\sigma}_f^2$  may **underestimate**  $\sigma_f^2$

### 2. Model Errors:

- Biased model produces ensemble predictions that are far from the observations. May lead to **ensemble collapse**

The goal is to maintain *enough* spread in the ensemble

# Inflation

- One way to increase of the variance of the ensemble is to inflate:

$$x^i \leftarrow \sqrt{\lambda} (x^i - \bar{x}) + \bar{x} \quad (2)$$

while preserving the ensemble mean.

# Inflation

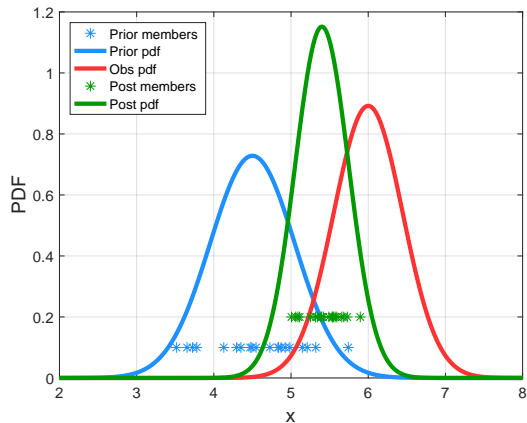
- One way to increase of the variance of the ensemble is to inflate:

$$x^i \leftarrow \sqrt{\lambda} (x^i - \bar{x}) + \bar{x} \quad (2)$$

while preserving the ensemble mean.

- Which variance to inflate: Prior (after the forecast) or posterior (after the update)?
- What to choose for  $\lambda$ ? 1.02 (2%), 1.04, 1.2, 10, ... ?
- Why this is useful?

# Inflation

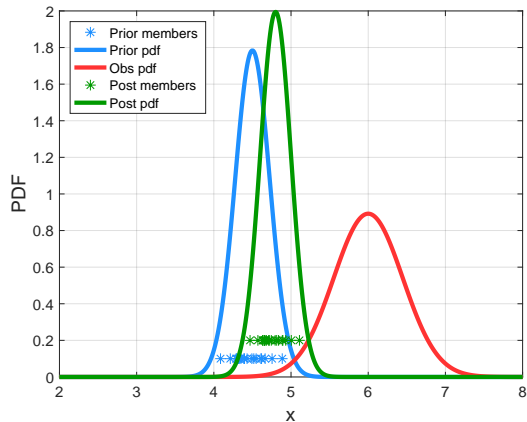


$$\bar{x}_f = 4.5, \sigma_f^2 = 0.30$$

$$y = 6.0, \sigma_o^2 = 0.20$$

$$\bar{x}_a = 5.43, \sigma_a^2 = 0.06$$

# Inflation



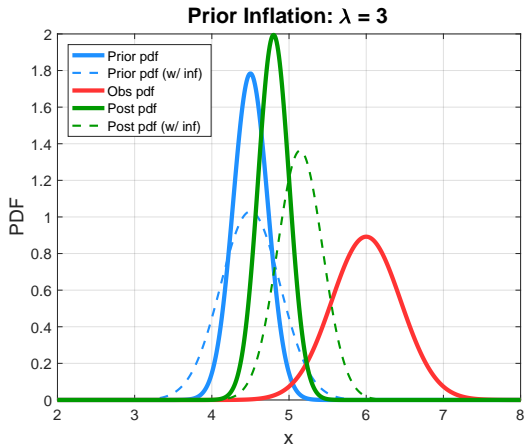
$$\bar{X}_f = 4.5, \sigma_f^2 = \mathbf{0.05}$$

$$y = 6.0, \sigma_o^2 = 0.20$$

$$\bar{X}_a = \mathbf{4.75}, \sigma_a^2 = \mathbf{0.02}$$



# Inflation

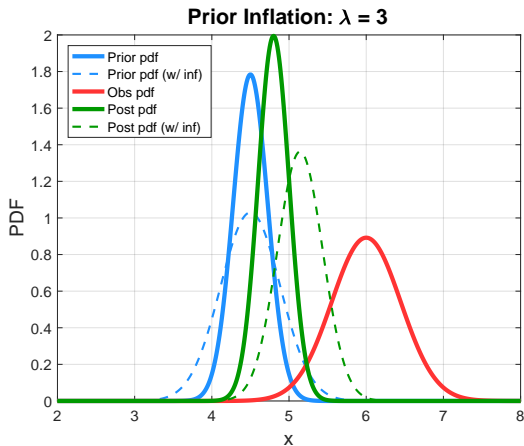


$$\sigma_f^2 : 0.05 \rightarrow 0.15$$

$$\bar{X}_a = 4.75 \rightarrow 5.18$$

$$\sigma_a^2 = 0.02 \rightarrow 0.07$$

# Inflation



$$\sigma_f^2 : 0.05 \rightarrow 0.15$$

$$\bar{X}_a = 4.75 \rightarrow 5.18$$

$$\sigma_a^2 = 0.02 \rightarrow 0.07$$

Higher spread: Larger uncertainty (less confidence in the estimates)!

# Adaptive Prior Inflation I

- Idea is to use the data to estimate the inflation parameter.

# Adaptive Prior Inflation I

- Idea is to use the data to estimate the inflation parameter.

For a single variable case, denote the following:

$$\text{Forecast (Background) error: } \varepsilon_f \sim \mathcal{N}\left(0, \sigma_f^2\right), \quad (3)$$

$$\text{Observation error: } \varepsilon_o \sim \mathcal{N}\left(0, \sigma_o^2\right), \quad (4)$$

$$\text{Analysis error: } \varepsilon_a \sim \mathcal{N}\left(0, \sigma_a^2\right); \sigma_a^2 \stackrel{\text{Kalman}}{=} \sigma_o^2 \sigma_f^2 / (\sigma_o^2 + \sigma_f^2). \quad (5)$$

# Adaptive Prior Inflation I

- Idea is to use the data to estimate the inflation parameter.

For a single variable case, denote the following:

$$\text{Forecast (Background) error: } \varepsilon_f \sim \mathcal{N}\left(0, \sigma_f^2\right), \quad (3)$$

$$\text{Observation error: } \varepsilon_o \sim \mathcal{N}\left(0, \sigma_o^2\right), \quad (4)$$

$$\text{Analysis error: } \varepsilon_a \sim \mathcal{N}\left(0, \sigma_a^2\right); \sigma_a^2 \stackrel{\text{Kalman}}{=} \sigma_o^2 \sigma_f^2 / (\sigma_o^2 + \sigma_f^2). \quad (5)$$

The forecast innovation (discrepancy):

$$d_f = y - \bar{x}_f = y - x_t + x_t - \bar{x}_f = \varepsilon_o - \varepsilon_f, \quad (6)$$

where  $x_t$  is the true value of the variable.

# Adaptive Prior Inflation I

- Idea is to use the data to estimate the inflation parameter.

For a single variable case, denote the following:

$$\text{Forecast (Background) error: } \varepsilon_f \sim \mathcal{N}(0, \sigma_f^2), \quad (3)$$

$$\text{Observation error: } \varepsilon_o \sim \mathcal{N}(0, \sigma_o^2), \quad (4)$$

$$\text{Analysis error: } \varepsilon_a \sim \mathcal{N}(0, \sigma_a^2); \sigma_a^2 \stackrel{\text{Kalman}}{=} \sigma_o^2 \sigma_f^2 / (\sigma_o^2 + \sigma_f^2). \quad (5)$$

The forecast innovation (discrepancy):

$$d_f = y - \bar{x}_f = y - x_t + x_t - \bar{x}_f = \varepsilon_o - \varepsilon_f, \quad (6)$$

where  $x_t$  is the true value of the variable.

**Bayesian Approach** [Anderson 2007, Anderson 2009, El Gharamti 2018]

$$p(\lambda|d_f) \propto p(\lambda) \cdot p(d_f|\lambda) \quad (7)$$

# Adaptive Prior Inflation II

- **Prior:**  $p(\lambda)$  prescribed .. a Gaussian, an inverse-Gamma

# Adaptive Prior Inflation II

- **Prior:**  $p(\lambda)$  prescribed .. a Gaussian, an inverse-Gamma
- **Likelihood:**  $p(d_f|\lambda) \sim \mathcal{N}(\mathbb{E}(d_f), \text{var}(d_f))$



# Adaptive Prior Inflation II

- **Prior:**  $p(\lambda)$  prescribed .. a Gaussian, an inverse-Gamma
- **Likelihood:**  $p(d_f|\lambda) \sim \mathcal{N}(\mathbb{E}(d_f), \text{var}(d_f))$

$$\mathbb{E}(d_f) = \mathbb{E}(\varepsilon_o - \varepsilon_f) = 0, \quad (8)$$

$$\text{var}(d_f) = \theta^2 = \mathbb{E}[(d_f - \mathbb{E}(d_f))^2] \quad (9)$$

$$= \mathbb{E}[\varepsilon_o^2 - \varepsilon_o\varepsilon_f - \varepsilon_f\varepsilon_o + \varepsilon_f^2] \quad (10)$$

$$= \sigma_o^2 + \sigma_f^2 \quad (11)$$

$$\Rightarrow \theta^2(\lambda) = \sigma_o^2 + \lambda \hat{\sigma}_f^2 \quad (12)$$

# Adaptive Prior Inflation II

- **Prior:**  $p(\lambda)$  prescribed .. a Gaussian, an inverse-Gamma
- **Likelihood:**  $p(d_f|\lambda) \sim \mathcal{N}(\mathbb{E}(d_f), \text{var}(d_f))$

$$\mathbb{E}(d_f) = \mathbb{E}(\varepsilon_o - \varepsilon_f) = 0, \quad (8)$$

$$\text{var}(d_f) = \theta^2 = \mathbb{E}[(d_f - \mathbb{E}(d_f))^2] \quad (9)$$

$$= \mathbb{E}[\varepsilon_o^2 - \varepsilon_o\varepsilon_f - \varepsilon_f\varepsilon_o + \varepsilon_f^2] \quad (10)$$

$$= \sigma_o^2 + \sigma_f^2 \quad (11)$$

$$\Rightarrow \theta^2(\lambda) = \sigma_o^2 + \lambda \hat{\sigma}_f^2 \quad (12)$$

## Assumptions:

1. Observation and forecast errors are uncorrelated;  $\mathbb{E}[\varepsilon_o\varepsilon_f] = 0$

# Adaptive Prior Inflation II

- **Prior:**  $p(\lambda)$  prescribed .. a Gaussian, an inverse-Gamma
- **Likelihood:**  $p(d_f|\lambda) \sim \mathcal{N}(\mathbb{E}(d_f), \text{var}(d_f))$

$$\mathbb{E}(d_f) = \mathbb{E}(\varepsilon_o - \varepsilon_f) = 0, \quad (8)$$

$$\text{var}(d_f) = \theta^2 = \mathbb{E}[(d_f - \mathbb{E}(d_f))^2] \quad (9)$$

$$= \mathbb{E}[\varepsilon_o^2 - \varepsilon_o\varepsilon_f - \varepsilon_f\varepsilon_o + \varepsilon_f^2] \quad (10)$$

$$= \sigma_o^2 + \sigma_f^2 \quad (11)$$

$$\Rightarrow \theta^2(\lambda) = \sigma_o^2 + \lambda \hat{\sigma}_f^2 \quad (12)$$

## Assumptions:

1. Observation and forecast errors are uncorrelated;  $\mathbb{E}[\varepsilon_o\varepsilon_f] = 0$
2. Errors are associated only with the forecast variance. The observation error variance  $\sigma_o^2$  is known and is right!

# Adaptive Prior Inflation II

- **Prior:**  $p(\lambda)$  prescribed .. a Gaussian, an inverse-Gamma
- **Likelihood:**  $p(d_f|\lambda) \sim \mathcal{N}(\mathbb{E}(d_f), \text{var}(d_f))$

$$\mathbb{E}(d_f) = \mathbb{E}(\varepsilon_o - \varepsilon_f) = 0, \quad (8)$$

$$\text{var}(d_f) = \theta^2 = \mathbb{E}[(d_f - \mathbb{E}(d_f))^2] \quad (9)$$

$$= \mathbb{E}[\varepsilon_o^2 - \varepsilon_o\varepsilon_f - \varepsilon_f\varepsilon_o + \varepsilon_f^2] \quad (10)$$

$$= \sigma_o^2 + \sigma_f^2 \quad (11)$$

$$\Rightarrow \theta^2(\lambda) = \sigma_o^2 + \lambda \hat{\sigma}_f^2 \quad (12)$$

## Assumptions:

1. Observation and forecast errors are uncorrelated;  $\mathbb{E}[\varepsilon_o\varepsilon_f] = 0$
2. Errors are associated only with the forecast variance. The observation error variance  $\sigma_o^2$  is known and is right!
3. Inflating the ensemble variance by  $\lambda$  will match the theoretical (hidden) forecast variance

# Adaptive Prior Inflation III

- **Posterior:**

$$p(\lambda|d_f) = \mathcal{JG}(\alpha, \beta) \cdot \mathcal{N}\left(0, \theta^2(\lambda)\right), \quad (13)$$

# Adaptive Prior Inflation III

- **Posterior:**

$$p(\lambda|d_f) = \mathcal{JG}(\alpha, \beta) \cdot \mathcal{N}\left(0, \theta^2(\lambda)\right), \quad (13)$$

$$\begin{aligned} &= \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{-\alpha-1} \exp\left(-\frac{\beta}{\lambda}\right) \cdot \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{d_f^2}{2\theta^2}\right), \\ &= \frac{\beta^\alpha \lambda^{-\alpha-1}}{\sqrt{2\pi\theta} \Gamma(\alpha)} \exp\left(-\frac{\beta d_f^2}{2\lambda\theta^2}\right), \end{aligned} \quad (14)$$

# Adaptive Prior Inflation III

- **Posterior:**

$$p(\lambda|d_f) = \mathcal{JG}(\alpha, \beta) \cdot \mathcal{N}\left(0, \theta^2(\lambda)\right), \quad (13)$$

$$\begin{aligned} &= \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{-\alpha-1} \exp\left(-\frac{\beta}{\lambda}\right) \cdot \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{d_f^2}{2\theta^2}\right), \\ &= \frac{\beta^\alpha \lambda^{-\alpha-1}}{\sqrt{2\pi\theta} \Gamma(\alpha)} \exp\left(-\frac{\beta d_f^2}{2\lambda\theta^2}\right), \end{aligned} \quad (14)$$

- Form very close to being inverse-gamma

# Adaptive Prior Inflation III

- **Posterior:**

$$p(\lambda|d_f) = \mathcal{JG}(\alpha, \beta) \cdot \mathcal{N}\left(0, \theta^2(\lambda)\right), \quad (13)$$

$$\begin{aligned} &= \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{-\alpha-1} \exp\left(-\frac{\beta}{\lambda}\right) \cdot \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{d_f^2}{2\theta^2}\right), \\ &= \frac{\beta^\alpha \lambda^{-\alpha-1}}{\sqrt{2\pi\theta} \Gamma(\alpha)} \exp\left(-\frac{\beta d_f^2}{2\lambda\theta^2}\right), \end{aligned} \quad (14)$$

- Form very close to being inverse-gamma
- To find the posterior mode, need to maximize  $p(\lambda|d_f)$ . Not easy!



# Adaptive Prior Inflation III

- **Posterior:**

$$p(\lambda|d_f) = \mathcal{JG}(\alpha, \beta) \cdot \mathcal{N}\left(0, \theta^2(\lambda)\right), \quad (13)$$

$$\begin{aligned} &= \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{-\alpha-1} \exp\left(-\frac{\beta}{\lambda}\right) \cdot \frac{1}{\sqrt{2\pi}\theta} \exp\left(-\frac{d_f^2}{2\theta^2}\right), \\ &= \frac{\beta^\alpha \lambda^{-\alpha-1}}{\sqrt{2\pi}\theta\Gamma(\alpha)} \exp\left(-\frac{\beta d_f^2}{2\lambda\theta^2}\right), \end{aligned} \quad (14)$$

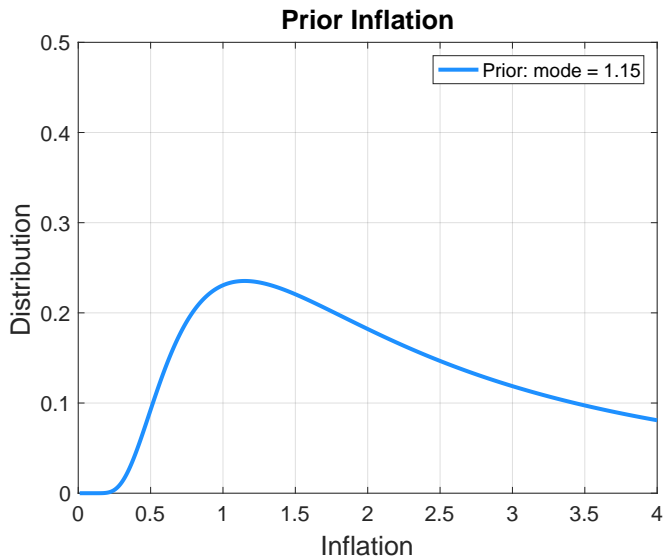
- Form very close to being inverse-gamma
- To find the posterior mode, need to maximize  $p(\lambda|d_f)$ . Not easy!
- Linearizing the Likelihood will simplify the problem:

$$p(d_f|\lambda) \approx \underbrace{p(d_f|\lambda_f)}_{\bar{\ell}} + \underbrace{\frac{\partial p(d_f|\lambda)}{\partial \lambda}\bigg|_{\lambda_f}}_{\ell'} (\lambda - \lambda_f), \quad (15)$$

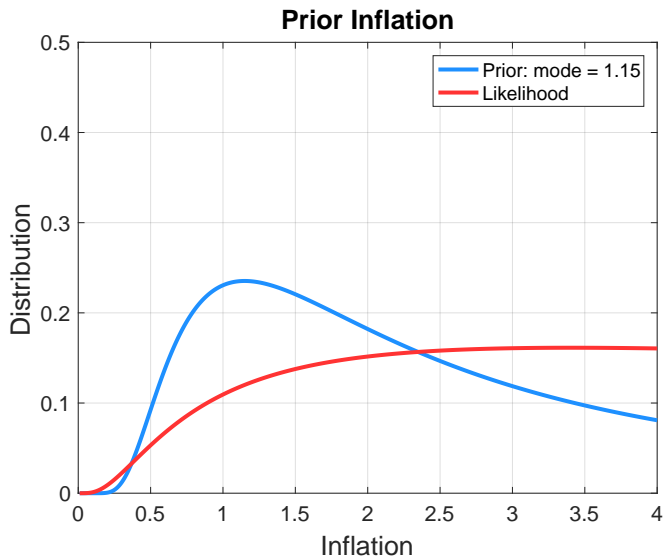
$\lambda_f$  is the mode of the prior distribution. Set  $\ell = \bar{\ell}/\ell'$ :

$$(1 - \lambda_f/\beta) \lambda^2 + (\ell - 2\lambda_f) \lambda + (\lambda_f^2 - \lambda_f \ell) = 0 \quad (16)$$

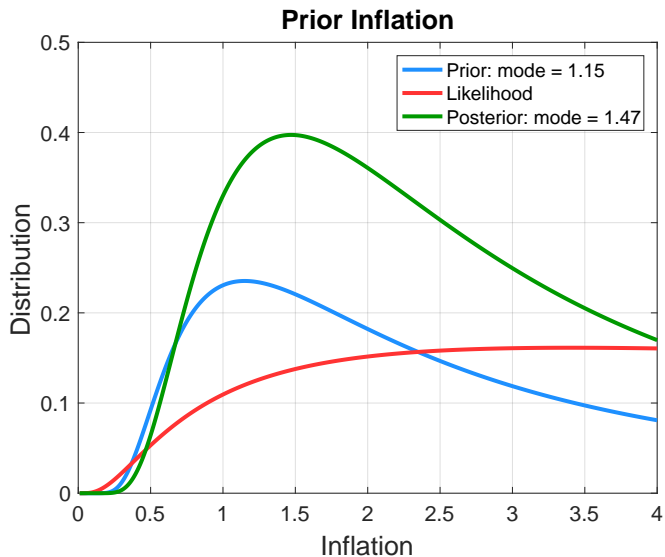
# Adaptive Prior Inflation IV



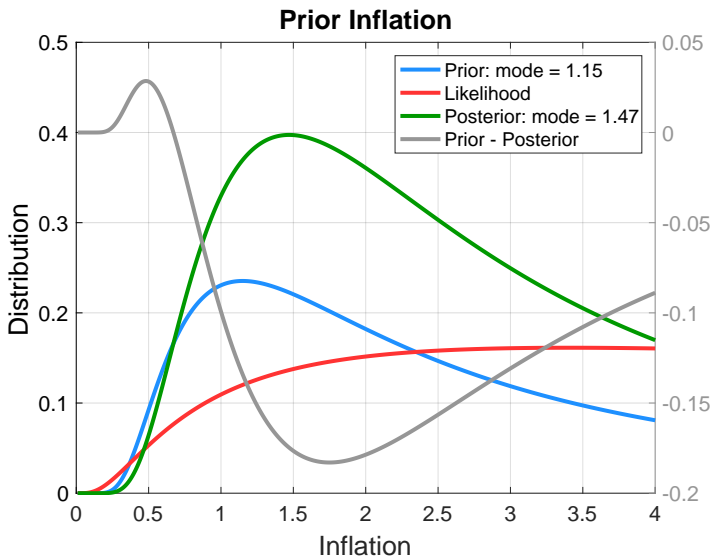
# Adaptive Prior Inflation IV



# Adaptive Prior Inflation IV



# Adaptive Prior Inflation IV



Can the inflation variance increase after the update?

# Adaptive Posterior Inflation

Follow similar Bayes' formulation:

$$p(\lambda|d_a) \propto p(\lambda) \cdot p(d_a|\lambda) \quad (17)$$

# Adaptive Posterior Inflation

Follow similar Bayes' formulation:

$$p(\lambda|d_a) \propto p(\lambda) \cdot p(d_a|\lambda) \quad (17)$$

Need to compute the likelihood,  $p(d_a|\lambda) \sim \mathcal{N}(\mathbb{E}(d_a), \text{var}(d_a))$

$$\mathbb{E}(d_a) = \mathbb{E}(\varepsilon_o) - \mathbb{E}(\varepsilon_a) = 0, \quad (18)$$

# Adaptive Posterior Inflation

Follow similar Bayes' formulation:

$$p(\lambda|d_a) \propto p(\lambda) \cdot p(d_a|\lambda) \quad (17)$$

Need to compute the likelihood,  $p(d_a|\lambda) \sim \mathcal{N}(\mathbb{E}(d_a), \text{var}(d_a))$

$$\begin{aligned} \mathbb{E}(d_a) &= \mathbb{E}(\varepsilon_o) - \mathbb{E}(\varepsilon_a) = 0, \\ \text{var}(d_a) &= \mathbb{E}(\varepsilon_o^2) + \mathbb{E}(\varepsilon_a^2) - 2 \underbrace{\mathbb{E}(\varepsilon_o \varepsilon_a)}_{\neq 0 \text{ correlated errors}}, \end{aligned} \quad (18)$$



# Adaptive Posterior Inflation

Follow similar Bayes' formulation:

$$p(\lambda|d_a) \propto p(\lambda) \cdot p(d_a|\lambda) \quad (17)$$

Need to compute the likelihood,  $p(d_a|\lambda) \sim \mathcal{N}(\mathbb{E}(d_a), \text{var}(d_a))$

$$\begin{aligned} \mathbb{E}(d_a) &= \mathbb{E}(\varepsilon_o) - \mathbb{E}(\varepsilon_a) = 0, \\ \text{var}(d_a) &= \mathbb{E}(\varepsilon_o^2) + \mathbb{E}(\varepsilon_a^2) - 2 \underbrace{\mathbb{E}(\varepsilon_o \varepsilon_a)}_{\neq 0 \text{ correlated errors}}, \\ &= \sigma_o^2 + \sigma_a^2 - 2\mathbb{E}[(1-K)\varepsilon_o\varepsilon_f + K\varepsilon_o^2], \\ &\dots \\ &= \sigma_o^2 - \sigma_a^2 \equiv \sigma_o^2 - \lambda\hat{\sigma}_a^2, \end{aligned} \quad (18)$$

where  $K = \sigma_f^2 (\sigma_o^2 + \sigma_f^2)^{-1}$ . Thus,

# Adaptive Posterior Inflation

Follow similar Bayes' formulation:

$$p(\lambda|d_a) \propto p(\lambda) \cdot p(d_a|\lambda) \quad (17)$$

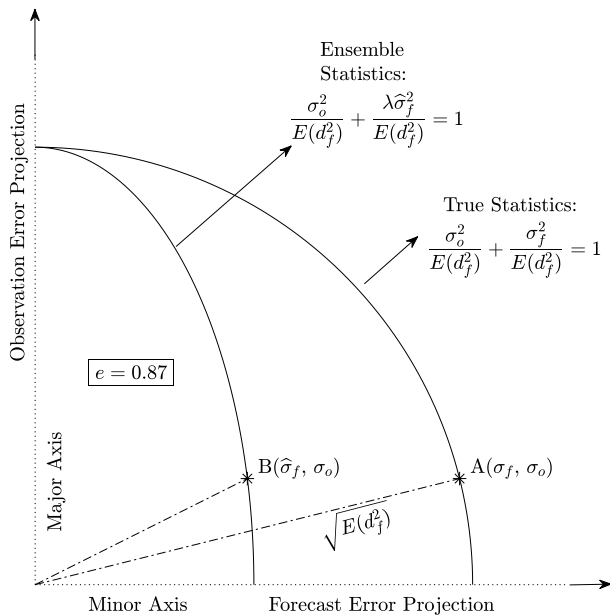
Need to compute the likelihood,  $p(d_a|\lambda) \sim \mathcal{N}(\mathbb{E}(d_a), \text{var}(d_a))$

$$\begin{aligned} \mathbb{E}(d_a) &= \mathbb{E}(\varepsilon_o) - \mathbb{E}(\varepsilon_a) = 0, \\ \text{var}(d_a) &= \mathbb{E}(\varepsilon_o^2) + \mathbb{E}(\varepsilon_a^2) - 2 \underbrace{\mathbb{E}(\varepsilon_o \varepsilon_a)}_{\neq 0 \text{ correlated errors}}, \\ &= \sigma_o^2 + \sigma_a^2 - 2\mathbb{E}[(1-K)\varepsilon_o\varepsilon_f + K\varepsilon_o^2], \\ &\dots \\ &= \sigma_o^2 - \sigma_a^2 \equiv \sigma_o^2 - \lambda\hat{\sigma}_a^2, \end{aligned} \quad (18)$$

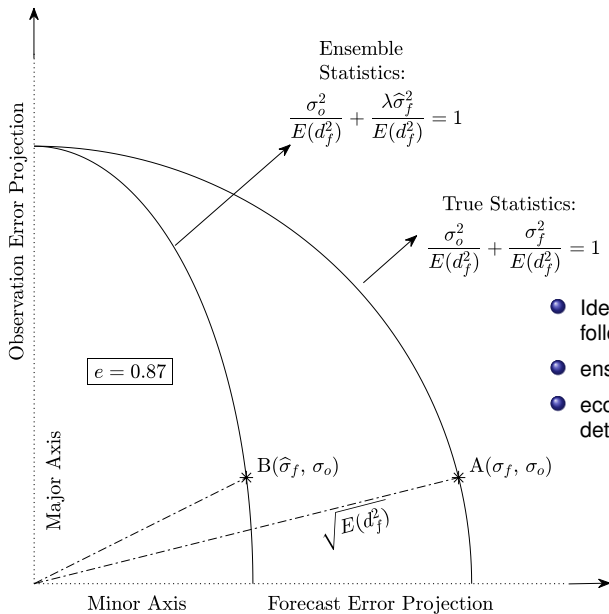
where  $K = \sigma_f^2 (\sigma_o^2 + \sigma_f^2)^{-1}$ . Thus,

$$p(d_a|\lambda) = (2\pi)^{-\frac{1}{2}} \exp\left[-\frac{d_a^2}{2}(\sigma_o^2 - \lambda\hat{\sigma}_a^2)^{-1}\right] (\sigma_o^2 + \lambda\hat{\sigma}_a^2)^{-\frac{1}{2}}. \quad (20)$$

# Geometrical Interpretation (Prior vs Posterior)

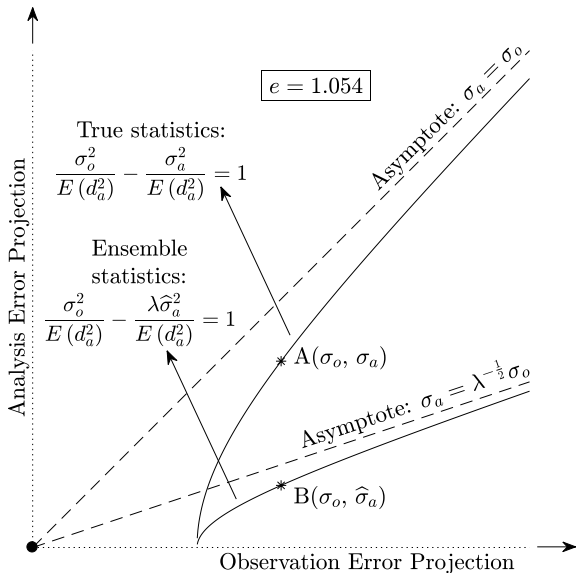


# Geometrical Interpretation (Prior vs Posterior)

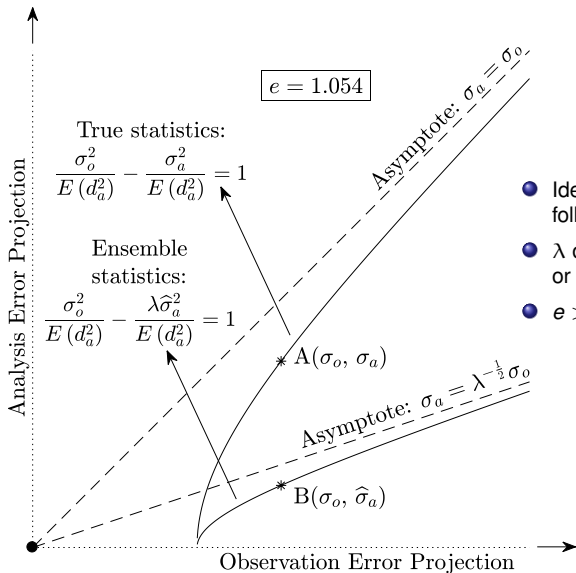


- Ideal state and observation statistics follow a circle
- ensemble ones satisfy an ellipse
- eccentricity,  $0 < e < 1$ , a measure to determine the deviation from circle

# Geometrical Interpretation (Prior vs Posterior)



# Geometrical Interpretation (Prior vs Posterior)

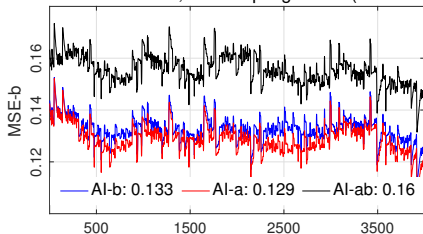


- Ideal and ensemble-based statistics follow hyperbolas
- $\lambda$  determines the degree of expansion or contraction of the hyperbola
- $e > 1$  also a measure of deviation

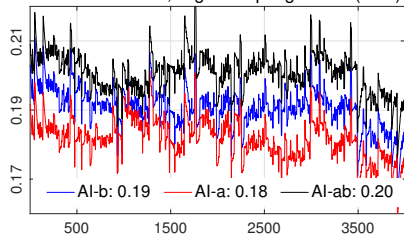


# Low-Order Models: Lorenz-63

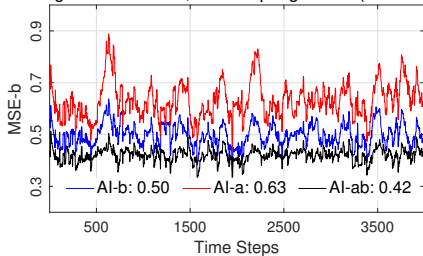
No model errors; No sampling errors (N=5000)



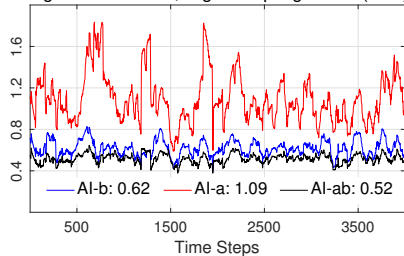
No model errors; High sampling errors (N=5)



High model errors; No sampling errors (N=5000)



High model errors; High sampling errors (N=5)



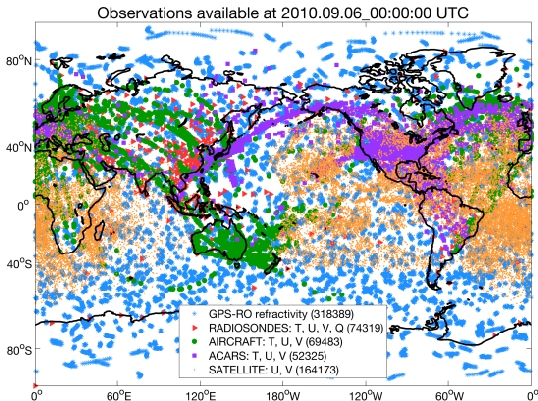


# CAM (The Community Atmosphere Model)

- version: CESM2\_0\_beta05
- resolution:  $1.9^\circ \times 1.9^\circ$  FV core;  
LAT: 96, LON: 144, LEV: 26
- State variables: surface pressure (PS), sensible temperature (T), wind components (U and V), specific humidity (Q), cloud liquid water (CLDLIQ) and cloud ice (CLDICE).

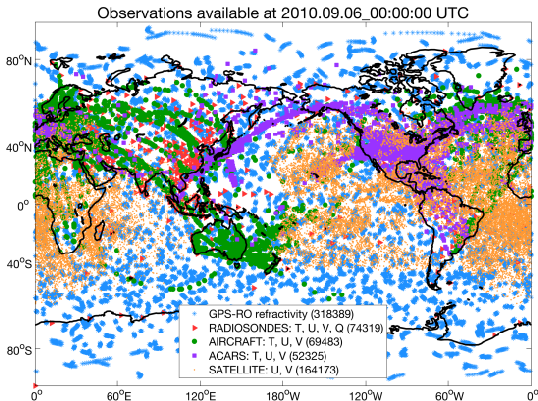
# CAM (The Community Atmosphere Model)

- version: CESM2\_0\_beta05
- resolution:  $1.9^\circ \times 1.9^\circ$  FV core;  
LAT: 96, LON: 144, LEV: 26
- State variables: surface pressure (PS), sensible temperature (T), wind components (U and V), specific humidity (Q), cloud liquid water (CLDLIQ) and cloud ice (CLDICE).
- DA: 08.16.2010 to 09.30.2010
- 80 members
- data available every 6 hours

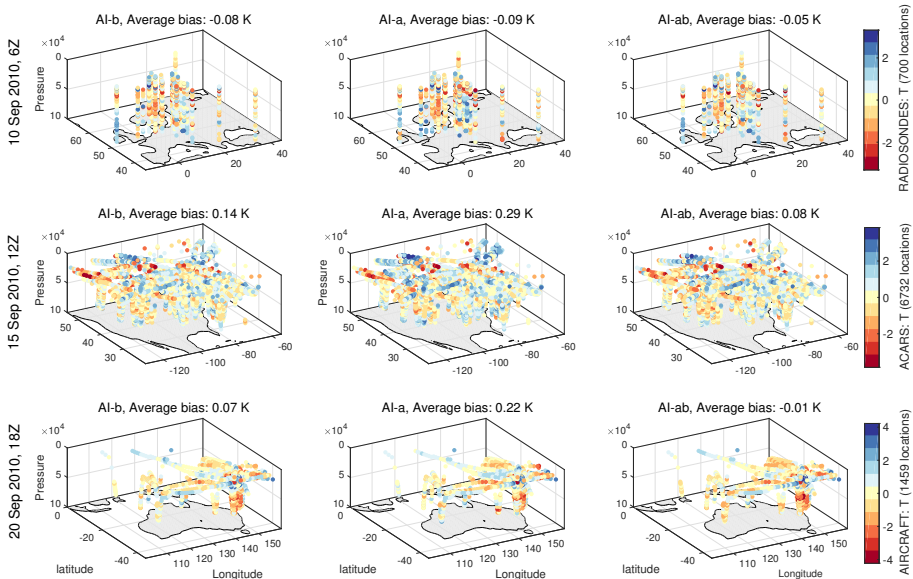


# CAM (The Community Atmosphere Model)

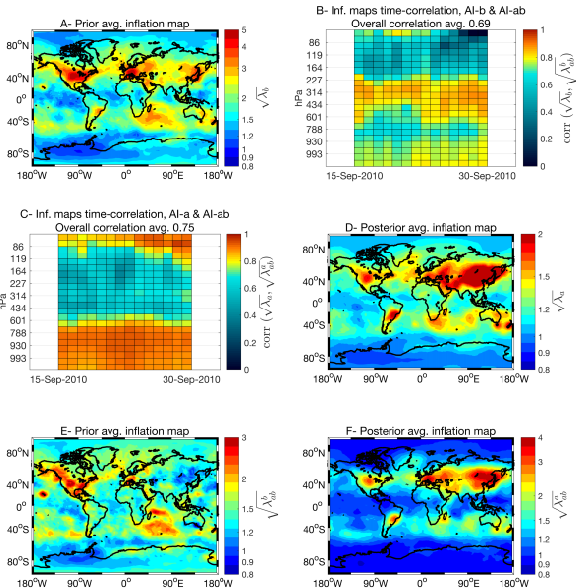
- version: CESM2\_0\_beta05
- resolution:  $1.9^\circ \times 1.9^\circ$  FV core;  
LAT: 96, LON: 144, LEV: 26
- State variables: surface pressure (PS), sensible temperature (T), wind components (U and V), specific humidity (Q), cloud liquid water (CLDLIQ) and cloud ice (CLDICE).
- DA: 08.16.2010 to 09.30.2010
- 80 members
- data available every 6 hours
- Horizontal localization:  $\approx 960$  km
- DART: latest 'Manhattan' release



# CAM Assimilation Results: Bias Treatment



# CAM Assimilation Results: Inflation Fields



# Conclusion

- Inflation is an important tool for ensemble Kalman filters
- The adaptive algorithm is based on Bayes' and uses forecast/analysis innovations to update the inflation
- With no model errors, posterior inflation produces higher quality estimates than prior inflation (better treatment of sampling errors)
- When model errors are dominant, as in CAM4, posterior inflation is found less useful
- Compelling results obtained by combining both prior and posterior inflation

# Conclusion

- Inflation is an important tool for ensemble Kalman filters
- The adaptive algorithm is based on Bayes' and uses forecast/analysis innovations to update the inflation
- With no model errors, posterior inflation produces higher quality estimates than prior inflation (better treatment of sampling errors)
- When model errors are dominant, as in CAM4, posterior inflation is found less useful
- Compelling results obtained by combining both prior and posterior inflation

## Papers

- **Gharamti, M. E.** (2018) "Enhanced Adaptive Inflation Algorithm for Ensemble Filters." *Monthly Weather Review*, 2, 623-640
- **Gharamti, M. E.**, Raeder, K., Anderson, J. and Wang, X. (2019) "Comparing Adaptive Prior and Posterior Inflation for Ensemble Filters Using an Atmospheric General Circulation Model." *Monthly Weather Review*, 147, 2535-2553

# THANK YOU

DART webpage: <https://dart.ucar.edu/>

NCAR | DART

National Center for  
Atmospheric Research

DOCUMENTATION RESEARCH ABOUT US SUPPORT RELEASES

## WELCOME TO DART

DART has been reformulated to better support the ensemble data assimilation needs of researchers who are interested in native netCDF support, less filesystem I/O, better computational performance, good scaling for large processor counts, and support for the memory requirements of very large models. Manhattan has support for many of our larger models (*WRF, POP, CAM, CICE, CLM, ROMS, MPAS\_ATM, ...*) with many more being added as time permits.

DOWNLOAD

## THE DATA ASSIMILATION RESEARCH TESTBED (DART)

DART is a community facility for ensemble DA developed and maintained by the Data Assimilation Research Section (DAReS) at the National Center for Atmospheric Research (NCAR). DART provides modelers, observational scientists, and geophysicists with powerful, flexible DA tools that are easy to implement and use and can be customized to support efficient operational DA applications. DART is a software environment that makes it easy to explore a variety of data assimilation methods and