Hybrid Ensemble Kalman Filtering and Optimal Interpolation

A new adaptive formulation

Moha Gharamti
https://dart.ucar.edu/
gharamti@ucar.edu

CU Boulder Nov. 19, 2021

National Center for Atmospheric Research
Data Assimilation Research Section (DAReS) - TDD - CISL

CISL Computational & Information Systems Lab

NCAR NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

1.1 Background

We want to find the state of a dynamical system using: [1] an *imperfect* **Model** and [2] a set of *sparse, noisy* **Observations**

1.1 Background

We want to find the state of a dynamical system using: [1] an *imperfect* **Model** and [2] a set of *sparse, noisy* **Observations**

Bayes Rule:

$$p\left(\mathbf{x}_k|\mathbf{y}_k,\mathbf{Y}_{k-1}\right) = \frac{p\left(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{Y}_{k-1}\right)\cdot p\left(\mathbf{y}_k|\mathbf{x}_k,\mathbf{Y}_{k-1}\right)}{p\left(\mathbf{y}_k|\mathbf{Y}_{k-1}\right)} \tag{1}$$

- $p\left(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{Y}_{k-1}\right) \rightsquigarrow$ Prior
- $p\left(\mathbf{y}_k|\mathbf{x}_k,\mathbf{Y}_{k-1}\right) \rightsquigarrow$ Likelihood
- $p\left(\mathbf{x}_k|\mathbf{y}_k,\mathbf{Y}_{k-1}\right) \rightsquigarrow$ Posterior

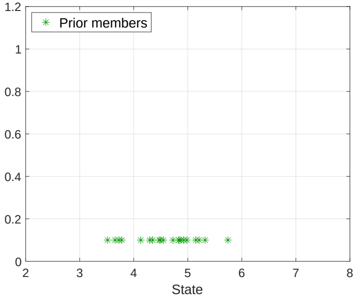
1.1 Background

We want to find the state of a dynamical system using: [1] an *imperfect* **Model** and [2] a set of *sparse, noisy* **Observations**

Bayes Rule:

$$p\left(\mathbf{x}_k|\mathbf{y}_k, \mathbf{Y}_{k-1}\right) = \frac{p\left(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{Y}_{k-1}\right) \cdot p\left(\mathbf{y}_k|\mathbf{x}_k, \mathbf{Y}_{k-1}\right)}{p\left(\mathbf{y}_k|\mathbf{Y}_{k-1}\right)} \tag{1}$$

- $p\left(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{Y}_{k-1}\right) \rightsquigarrow$ Prior
- $p\left(\mathbf{y}_k|\mathbf{x}_k, \mathbf{Y}_{k-1}\right) \rightsquigarrow$ Likelihood
- $p\left(\mathbf{x}_k|\mathbf{y}_k, \mathbf{Y}_{k-1}\right) \rightsquigarrow$ Posterior



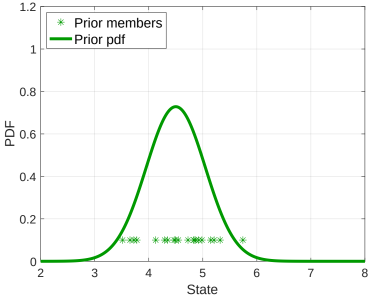
1.1 Background

We want to find the state of a dynamical system using: [1] an *imperfect* **Model** and [2] a set of *sparse, noisy* **Observations**

Bayes Rule:

$$p\left(\mathbf{x}_k|\mathbf{y}_k,\mathbf{Y}_{k-1}\right)=\frac{p\left(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{Y}_{k-1}\right)\cdot p\left(\mathbf{y}_k|\mathbf{x}_k,\mathbf{Y}_{k-1}\right)}{p\left(\mathbf{y}_k|\mathbf{Y}_{k-1}\right)} \quad (1)$$

- $p\left(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{Y}_{k-1}\right) \rightsquigarrow$ Prior
- $p\left(\mathbf{y}_k|\mathbf{x}_k,\mathbf{Y}_{k-1}\right) \rightsquigarrow$ Likelihood
- $p\left(\mathbf{x}_k|\mathbf{y}_k,\mathbf{Y}_{k-1}\right) \rightsquigarrow$ Posterior



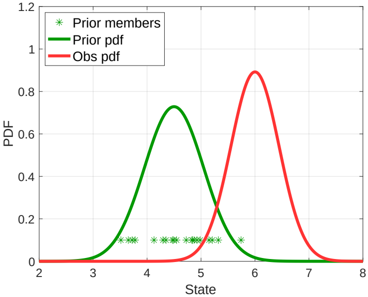
1.1 Background

We want to find the state of a dynamical system using: [1] an *imperfect* **Model** and [2] a set of *sparse*, *noisy* **Observations**

Bayes Rule:

$$p\left(\mathbf{x}_k|\mathbf{y}_k,\mathbf{Y}_{k-1}\right)=\frac{p\left(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{Y}_{k-1}\right)\cdot p\left(\mathbf{y}_k|\mathbf{x}_k,\mathbf{Y}_{k-1}\right)}{p\left(\mathbf{y}_k|\mathbf{Y}_{k-1}\right)} \tag{1}$$

- $p\left(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{Y}_{k-1}\right) \rightsquigarrow$ Prior
- $p\left(\mathbf{y}_k|\mathbf{x}_k,\mathbf{Y}_{k-1}\right) \rightsquigarrow$ Likelihood
- $p\left(\mathbf{x}_k|\mathbf{y}_k,\mathbf{Y}_{k-1}\right) \rightsquigarrow$ Posterior



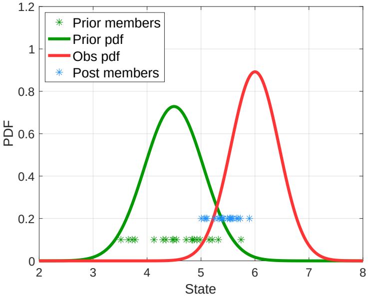
1.1 Background

We want to find the state of a dynamical system using: [1] an *imperfect* **Model** and [2] a set of *sparse, noisy* **Observations**

Bayes Rule:

$$p\left(\mathbf{x}_k|\mathbf{y}_k,\mathbf{Y}_{k-1}\right) = \frac{p\left(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{Y}_{k-1}\right)\cdot p\left(\mathbf{y}_k|\mathbf{x}_k,\mathbf{Y}_{k-1}\right)}{p\left(\mathbf{y}_k|\mathbf{Y}_{k-1}\right)} \tag{1}$$

- $p\left(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{Y}_{k-1}\right) \rightsquigarrow$ Prior
- $p\left(\mathbf{y}_k|\mathbf{x}_k,\mathbf{Y}_{k-1}\right) \rightsquigarrow$ Likelihood
- $p\left(\mathbf{x}_k|\mathbf{y}_k,\mathbf{Y}_{k-1}\right) \rightsquigarrow$ Posterior



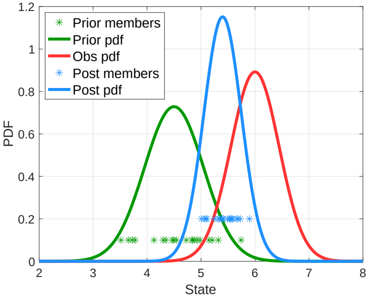
1.1 Background

We want to find the state of a dynamical system using: [1] an *imperfect* **Model** and [2] a set of *sparse, noisy* **Observations**

Bayes Rule:

$$p\left(\mathbf{x}_k \middle| \mathbf{y}_k, \mathbf{Y}_{k-1}\right) = \frac{p\left(\mathbf{x}_k \middle| \mathbf{x}_{k-1}, \mathbf{Y}_{k-1}\right) \cdot p\left(\mathbf{y}_k \middle| \mathbf{x}_k, \mathbf{Y}_{k-1}\right)}{p\left(\mathbf{y}_k \middle| \mathbf{Y}_{k-1}\right)} \tag{1}$$

- $p\left(\mathbf{x}_k \middle| \mathbf{x}_{k-1}, \mathbf{Y}_{k-1}\right) \rightsquigarrow$ Prior
- $p\left(\mathbf{y}_k \middle| \mathbf{x}_k, \mathbf{Y}_{k-1}\right) \rightsquigarrow$ Likelihood
- $p\left(\mathbf{x}_k \middle| \mathbf{y}_k, \mathbf{Y}_{k-1}\right) \rightsquigarrow$ Posterior



1.2 The Prior Distribution $p\left(\mathbf{x}_k | \mathbf{Y}_{k-1}\right)$

- Assumed normal $\sim \mathcal{N}\left(\mathbf{x}_k^f, \mathbf{P}_k^f\right)$; approximated using an EnKF

$$\mathbf{P}_k^f = \frac{1}{N-1} \sum_{i=1}^{N} \left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right) \left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right)^T \tag{2}$$

1.2 The Prior Distribution $p(\mathbf{x}_k | \mathrm{Y}_{k-1})$

- Assumed normal $\sim \mathcal{N}\left(\mathbf{x}_k^f, \mathbf{P}_k^f\right)$; approximated using an EnKF

$$\mathbf{P}_k^f = \frac{1}{N-1} \sum_{i=1}^{N} \left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right)\left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right)^T \tag{2}$$

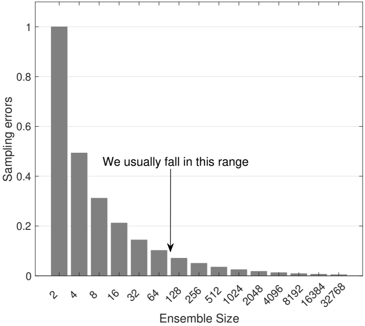
- Most EnKF errors are associated with the background covariance

1.2 The Prior Distribution $p(\mathbf{x}_k|\mathbf{Y}_{k-1})$

- Assumed normal $\sim \mathcal{N}\left(\mathbf{x}_k^f, \mathbf{P}_k^f\right)$; approximated using an EnKF

$$\mathbf{P}_k^f = \frac{1}{N-1}\sum_{i=1}^{N}\left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right)\left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right)^T \quad (2)$$

- Most EnKF errors are associated with the background covariance
  - Sampling errors $\rightsquigarrow$ variance underestimation



1.2 The Prior Distribution $p\left(\mathbf{x}_k | \mathbf{Y}_{k-1}\right)$

- Assumed normal $\sim \mathcal{N}\left(\mathbf{x}_k^f, \mathbf{P}_k^f\right)$; approximated using an EnKF

$$\mathbf{P}_k^f = \frac{1}{N-1} \sum_{i=1}^{N} \left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right) \left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right)^T \tag{2}$$

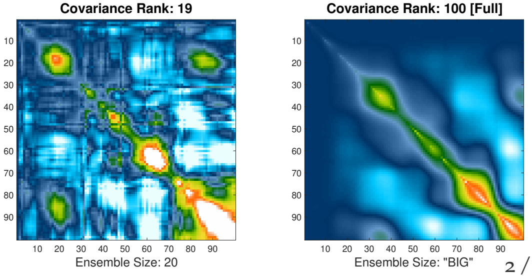
- Most EnKF errors are associated with the background covariance
  - Sampling errors $\rightsquigarrow$ variance underestimation
  - Model biases $\rightsquigarrow$ ensemble collapse

1.2 The Prior Distribution $p(\mathbf{x}_k | \mathbf{Y}_{k-1})$

- Assumed normal $\sim \mathcal{N}\left(\mathbf{x}_k^f, \mathbf{P}_k^f\right)$; approximated using an EnKF

$$\mathbf{P}_k^f = \frac{1}{N-1} \sum_{i=1}^{N} \left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right)\left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right)^T \tag{2}$$

- Most EnKF errors are associated with the background covariance
  - Sampling errors $\rightsquigarrow$ variance underestimation
  - Model biases $\rightsquigarrow$ ensemble collapse
  - Rank deficiency



1.2 The Prior Distribution $p\left(\mathbf{x}_k | \mathbf{Y}_{k-1}\right)$

- Assumed normal $\sim \mathcal{N}\left(\mathbf{x}_k^f, \mathbf{P}_k^f\right)$; approximated using an EnKF

$$\mathbf{P}_k^f = \frac{1}{N-1} \sum_{i=1}^{N} \left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right) \left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right)^T \qquad (2)$$

- Most EnKF errors are associated with the background covariance
  - Sampling errors $\rightsquigarrow$ variance underestimation
  - Model biases $\rightsquigarrow$ ensemble collapse
  - Rank deficiency
- (For now) we can't afford large ensembles and generally we can't do much about model biases

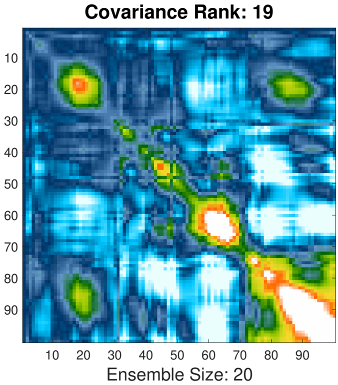
1.2 The Prior Distribution $p\left(\mathbf{x}_k | \mathbf{Y}_{k-1}\right)$

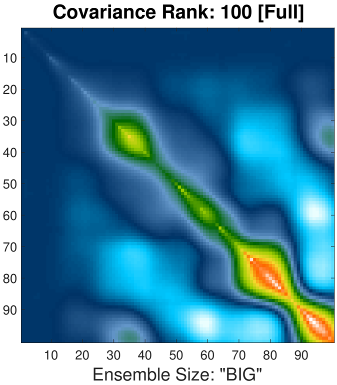
- Assumed normal $\sim \mathcal{N}\left(\mathbf{x}_k^f, \mathbf{P}_k^f\right)$; approximated using an EnKF

$$\mathbf{P}_k^f = \frac{1}{N-1} \sum_{i=1}^{N} \left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right) \left(\mathbf{x}_k^{f,i} - \mathbf{x}_k^f\right)^T \qquad (2)$$

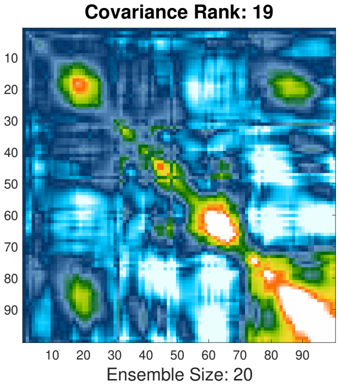
- Most EnKF errors are associated with the background covariance
  - Sampling errors $\rightsquigarrow$ variance underestimation
  - Model biases $\rightsquigarrow$ ensemble collapse
  - Rank deficiency
- (For now) we can't afford large ensembles and generally we can't do much about model biases
- Others known errors:
  nonGaussianity, nonlinearity, regression errors, ...

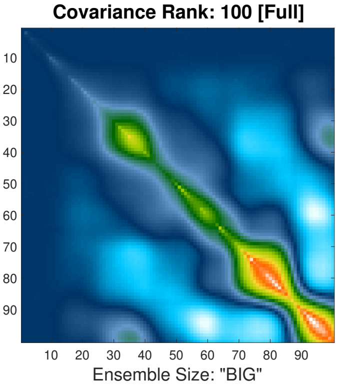
1.3 What can be done?





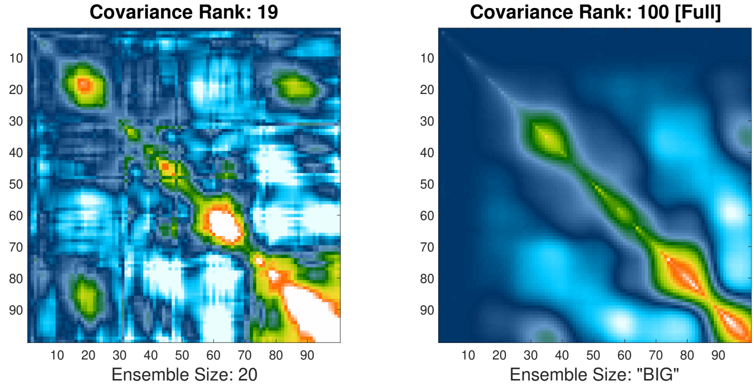
1.3 What can be done?





Holy Covariance: $\lim_{N\to\infty} \mathbf{P}^e \approx \mathbf{B}$

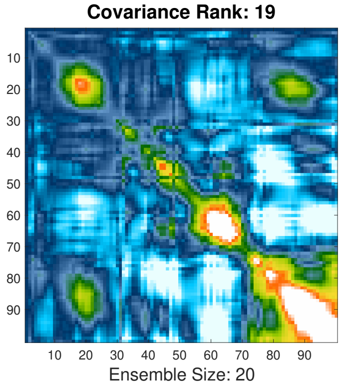
1.3 What can be done?

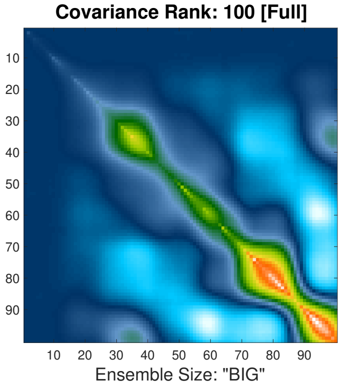


Holy Covariance: $\lim_{N \to \infty} \mathbf{P}^e \approx \mathbf{B}$

1. **Inflation:** increases the variance, rank is unchanged* $\mathbf{P}^f = \lambda \cdot \mathbf{P}^e$

1.3 What can be done?

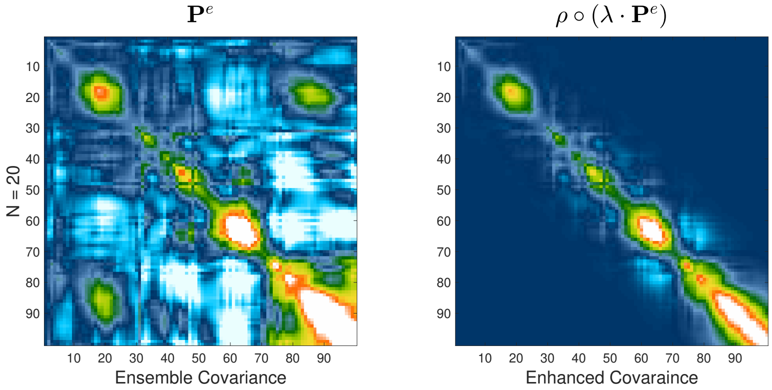




Holy Covariance: $\lim_{N\to\infty} \mathbf{P}^e \approx \mathbf{B}$

1. **Inflation:** increases the variance, rank is unchanged* $\mathbf{P}^f = \lambda \cdot \mathbf{P}^e$
2. **Localization:** removes spurious correlations, increases the rank $\mathbf{P}^f = \rho \circ \mathbf{P}^e$

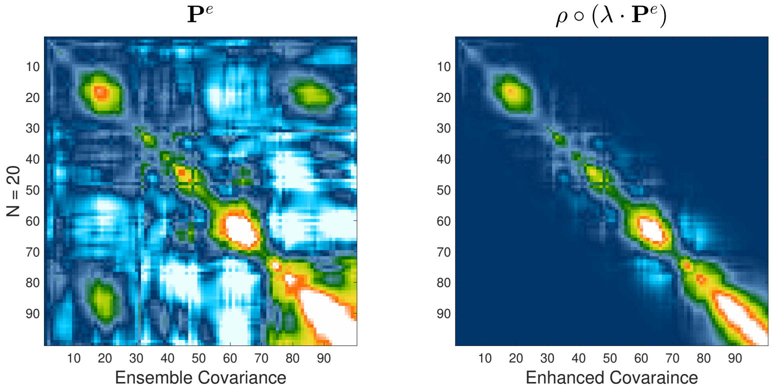
1.3 What can be done?



Holy Covariance: $\lim_{N \to \infty} \mathbf{P}^e \approx \mathbf{B}$

1. **Inflation:** increases the variance, rank is unchanged* $\mathbf{P}^f = \lambda \cdot \mathbf{P}^e$
2. **Localization:** removes spurious correlations, increases the rank $\mathbf{P}^f = \rho \circ \mathbf{P}^e$

1.3 What can be done?



Holy Covariance: $\lim_{N \to \infty} \mathbf{P}^e \approx \mathbf{B}$

1. **Inflation:** increases the variance, rank is unchanged* $\mathbf{P}^f = \lambda \cdot \mathbf{P}^e$
2. **Localization:** removes spurious correlations, increases the rank $\mathbf{P}^f = \rho \circ \mathbf{P}^e$
3. **Hybridization:** $\mathbf{P}^f = \alpha \mathbf{P}^e + (1 - \alpha)\mathbf{B}$

2.1 Hybrid EnKF-OI: Terminologies

- OI: Optimal Interpolation (essentially a KF with a prescribed invariant $\mathbf{P}^f$)
- Often referred to as EnKF-3DVar
- Initial effort by Hamill and Snyder (2000)

2.1 Hybrid EnKF-OI: Terminologies

- OI: Optimal Interpolation (essentially a KF with a prescribed invariant $\mathbf{P}^f$)
- Often referred to as EnKF-3DVar
- Initial effort by Hamill and Snyder (2000)

What's the idea?

Use a background covariance in the EnKF that is an "average" (weighted sum) of a flow-dependent background error covariance estimated from an ensemble and a predefined static covariance from a 3DVar or an OI system

2.1 Hybrid EnKF-OI: Terminologies

- OI: Optimal Interpolation (essentially a KF with a prescribed invariant $\mathbf{P}^f$)
- Often referred to as EnKF-3DVar
- Initial effort by Hamill and Snyder (2000)

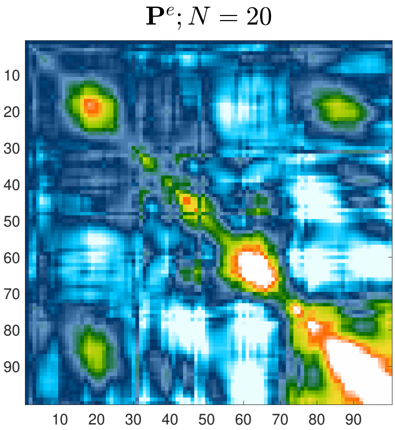
What's the idea?

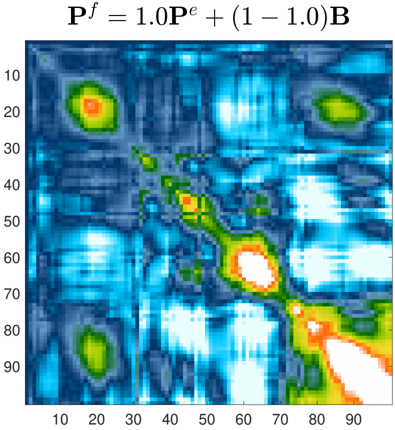
Use a background covariance in the EnKF that is an "average" (weighted sum) of a flow-dependent background error covariance estimated from an ensemble and a predefined static covariance from a 3DVar or an OI system

- Many different *hybrid* forms in the literature:
  - 4DEnVar: 4DVar with background covariance from an ensemble
  - En4DVar: Use an ensemble to approximate adjoint
  - hybrid 4(3)DVar: Var methods using a combination of climatological and ensemble covariances (e.g., α-control method in GSI)
  - EnVar: Term used for any of the previous hybrid forms

2.2 Hybrid EnKF-OI: in action

$$\mathbf{P}^f = \alpha \mathbf{P}^e + (1-\alpha)\mathbf{B}$$

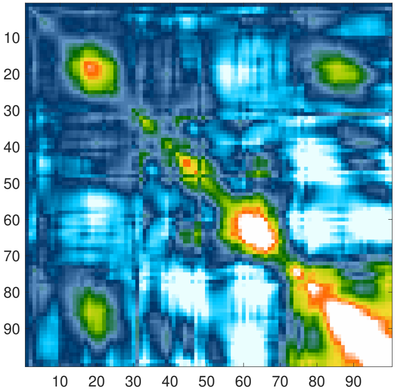




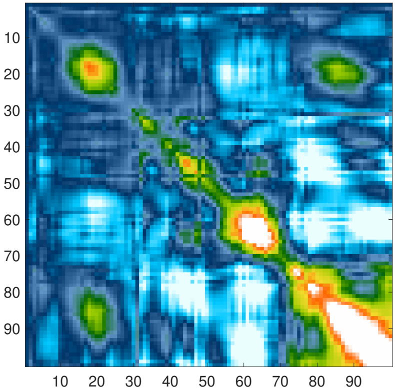
2.2 Hybrid EnKF-OI: in action

$$\mathbf{P}^f = \alpha \mathbf{P}^e + (1 - \alpha)\mathbf{B}$$

$\mathbf{P}^e; N = 20$



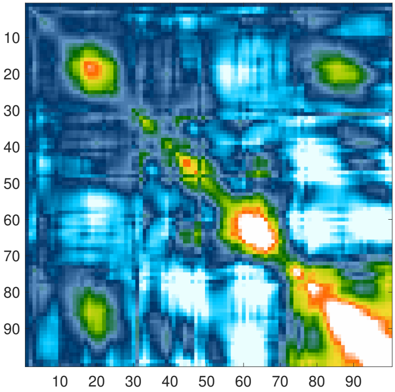
$\mathbf{P}^f = 0.9\mathbf{P}^e + (1 - 0.9)\mathbf{B}$



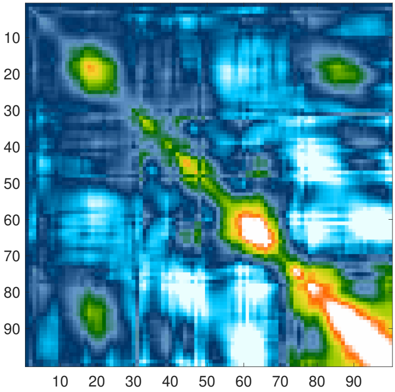
2.2 Hybrid EnKF-OI: in action

$$\mathbf{P}^f = \alpha \mathbf{P}^e + (1 - \alpha)\mathbf{B}$$

$\mathbf{P}^e; N = 20$



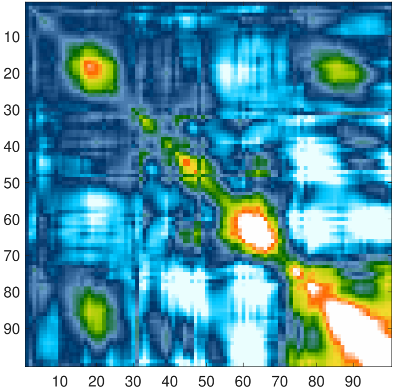
$\mathbf{P}^f = 0.8\mathbf{P}^e + (1 - 0.8)\mathbf{B}$



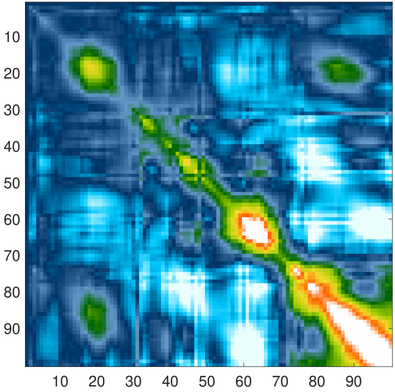
2.2 Hybrid EnKF-OI: in action

$$\mathbf{P}^f = \alpha \mathbf{P}^e + (1-\alpha)\mathbf{B}$$

$\mathbf{P}^e; N = 20$

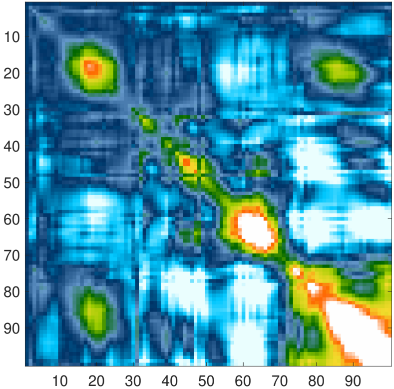


$\mathbf{P}^f = 0.7\mathbf{P}^e + (1-0.7)\mathbf{B}$

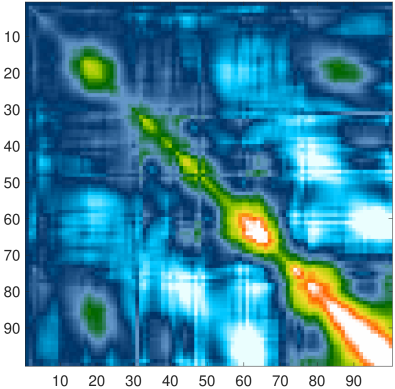


$$\mathbf{P}^f = \alpha \mathbf{P}^e + (1-\alpha)\mathbf{B}$$

$\mathbf{P}^e; N = 20$



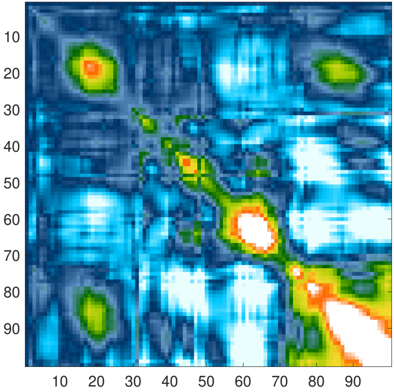
$\mathbf{P}^f = 0.6\mathbf{P}^e + (1-0.6)\mathbf{B}$



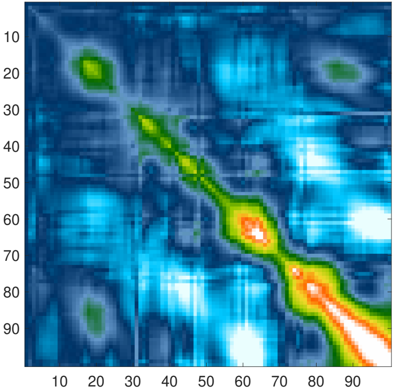
2.2 Hybrid EnKF-OI: in action

$$\mathbf{P}^f = \alpha \mathbf{P}^e + (1-\alpha)\mathbf{B}$$

$\mathbf{P}^e; N = 20$



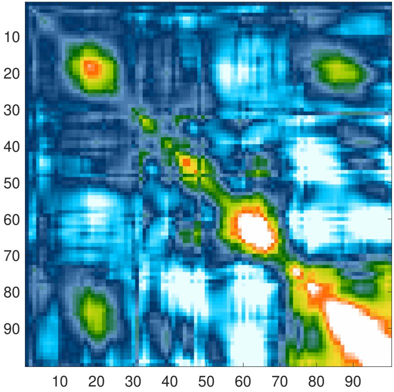
$\mathbf{P}^f = 0.5\mathbf{P}^e + (1-0.5)\mathbf{B}$



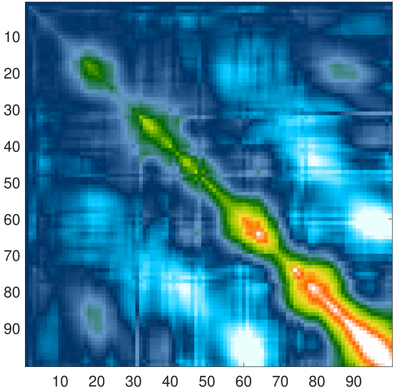
2.2 Hybrid EnKF-OI: in action

$$\mathbf{P}^f = \alpha \mathbf{P}^e + (1-\alpha)\mathbf{B}$$

$\mathbf{P}^e; N = 20$



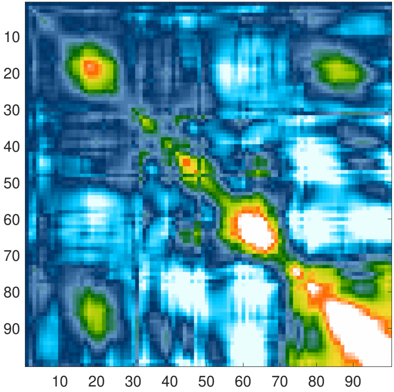
$\mathbf{P}^f = 0.4\mathbf{P}^e + (1-0.4)\mathbf{B}$



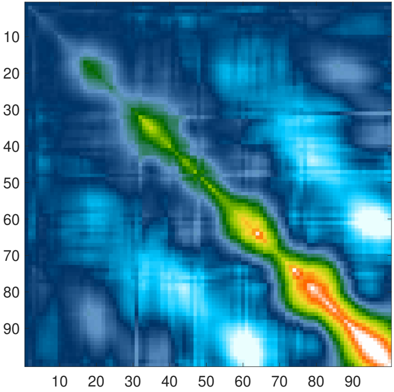
2.2 Hybrid EnKF-OI: in action

$$\mathbf{P}^f = \alpha \mathbf{P}^e + (1-\alpha)\mathbf{B}$$

$\mathbf{P}^e; N = 20$

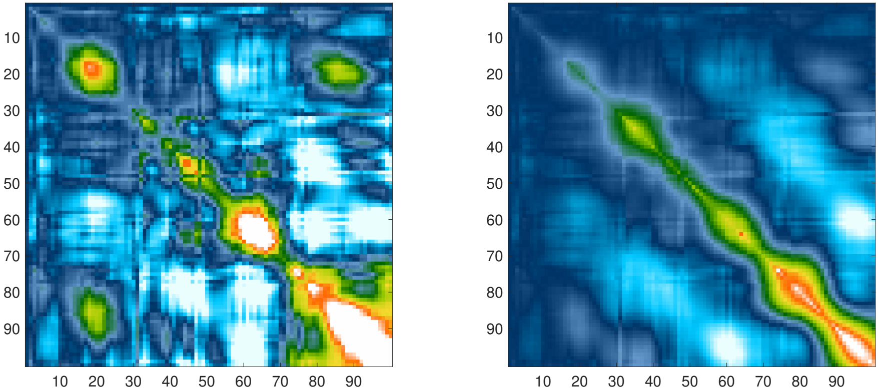


$\mathbf{P}^f = 0.3\mathbf{P}^e + (1-0.3)\mathbf{B}$



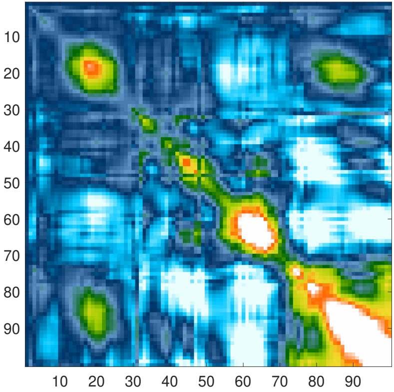
$$\mathbf{P}^f = \alpha \mathbf{P}^e + (1-\alpha)\mathbf{B}$$

$\mathbf{P}^e; N = 20$ $\quad\quad$ $\mathbf{P}^f = 0.2\mathbf{P}^e + (1-0.2)\mathbf{B}$

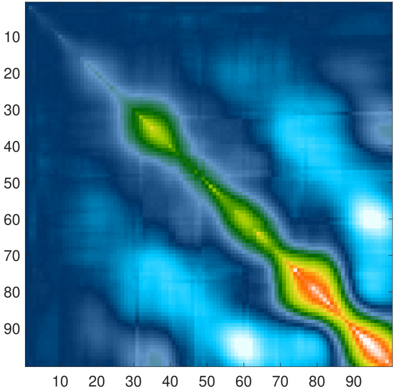


$$\mathbf{P}^f = \alpha \mathbf{P}^e + (1-\alpha)\mathbf{B}$$

$\mathbf{P}^e; N = 20$



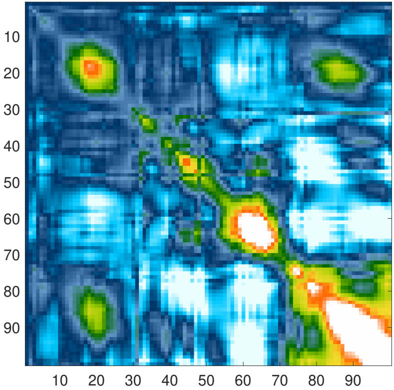
$\mathbf{P}^f = 0.1\mathbf{P}^e + (1-0.1)\mathbf{B}$



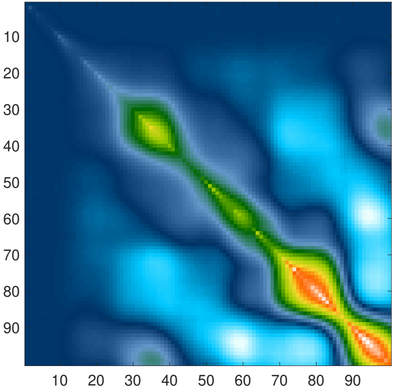
2.2 Hybrid EnKF-OI: in action

$$\mathbf{P}^f = \alpha \mathbf{P}^e + (1-\alpha)\mathbf{B}$$

$\mathbf{P}^e; N = 20$



$\mathbf{P}^f = 0.0\mathbf{P}^e + (1-0.0)\mathbf{B}$



Changes to the rank, variance, correlations, norm .. of the covariance

2.3 How to construct B

- Available from 3DVar systems (e.g., NMC method)
- Often formed from a large inventory of historical forecasts sampled over large windows (practical choice)

2.3 How to construct B

- Available from 3DVar systems (e.g., NMC method)
- Often formed from a large inventory of historical forecasts sampled over large windows (practical choice)
- Spectral decomposition is desirable

$$\mathbf{B} = \mathbf{S}\boldsymbol{\Omega}\mathbf{S}^T = \widehat{\mathbf{S}}\widehat{\mathbf{S}}^T, \tag{3}$$

where $\widehat{\mathbf{S}} = \mathbf{S}\boldsymbol{\Omega}^{\frac{1}{2}}$.

2.3 How to construct B

- Available from 3DVar systems (e.g., NMC method)
- Often formed from a large inventory of historical forecasts sampled over large windows (practical choice)
- Spectral decomposition is desirable

$$\mathbf{B} = \mathbf{S}\boldsymbol{\Omega}\mathbf{S}^T = \widehat{\mathbf{S}}\widehat{\mathbf{S}}^T, \tag{3}$$

  where $\widehat{\mathbf{S}} = \mathbf{S}\boldsymbol{\Omega}^{\frac{1}{2}}$.

- Succession of transform operators, $\mathbf{B} = \mathbf{B}^{1/2}\mathbf{B}^{T/2}$

$$\mathbf{B}^{1/2} = \mathbf{U}_p\mathbf{S}\mathbf{U}_v\mathbf{U}_h \tag{4}$$

2.3 How to construct B

- Available from 3DVar systems (e.g., NMC method)
- Often formed from a large inventory of historical forecasts sampled over large windows (practical choice)
- Spectral decomposition is desirable

$$\mathbf{B} = \mathbf{S}\boldsymbol{\Omega}\mathbf{S}^T = \widehat{\mathbf{S}}\widehat{\mathbf{S}}^T, \tag{3}$$

  where $\widehat{\mathbf{S}} = \mathbf{S}\boldsymbol{\Omega}^{\frac{1}{2}}$.

- Succession of transform operators, $\mathbf{B} = \mathbf{B}^{1/2}\mathbf{B}^{T/2}$

$$\mathbf{B}^{1/2} = \mathbf{U}_p\mathbf{S}\mathbf{U}_v\mathbf{U}_h \tag{4}$$

- Storage issue: $\mathbf{B}$ is of size $(N_x \times N_x)$; N_x is the state dimension

2.3 How to construct B

- Available from 3DVar systems (e.g., NMC method)
- Often formed from a large inventory of historical forecasts sampled over large windows (practical choice)
- Spectral decomposition is desirable

$$\mathbf{B} = \mathbf{S}\mathbf{\Omega}\mathbf{S}^T = \widehat{\mathbf{S}}\widehat{\mathbf{S}}^T, \tag{3}$$

  where $\widehat{\mathbf{S}} = \mathbf{S}\mathbf{\Omega}^{\frac{1}{2}}$.

- Succession of transform operators, $\mathbf{B} = \mathbf{B}^{1/2}\mathbf{B}^{T/2}$

$$\mathbf{B}^{1/2} = \mathbf{U}_p\mathbf{S}\mathbf{U}_v\mathbf{U}_h \tag{4}$$

- Storage issue: $\mathbf{B}$ is of size $(N_x \times N_x)$; N_x is the state dimension
  - Do we need to store **the entire B matrix?** May only need access to the historical (climatology) realizations

2.4 Hybrid EnKF-OI: Adaptive Form

How to choose α?

2.4 Hybrid EnKF-OI: Adaptive Form

How to choose α?

- The ensemble statistics should satisfy (Desroziers et al., 2005):

$$\mathbb{E}\left[\mathbf{d}\mathbf{d}^T\right] = \mathbf{R} + \mathbf{H}\mathbf{P}^f\mathbf{H}^T, \tag{5}$$

where $\mathbf{d} = \mathbf{y}^o - \mathbf{H}\mathbf{x}^f$. Substitute the hybrid covariance form in eq. (5):

$$\mathbb{E}\left[\mathbf{d}\mathbf{d}^T\right] = \mathbf{R} + \alpha\mathbf{H}\mathbf{P}^e\mathbf{H}^T + (1-\alpha)\mathbf{H}\mathbf{B}\mathbf{H}^T, \quad 0 \le \alpha \le 1 \tag{6}$$

How to choose α?

- The ensemble statistics should satisfy (Desroziers et al., 2005):

$$\mathbb{E}\left[\mathbf{d}\mathbf{d}^T\right] = \mathbf{R} + \mathbf{H}\mathbf{P}^f\mathbf{H}^T, \tag{5}$$

where $\mathbf{d} = \mathbf{y}^o - \mathbf{H}\mathbf{x}^f$. Substitute the hybrid covariance form in eq. (5):

$$\mathbb{E}\left[\mathbf{d}\mathbf{d}^T\right] = \mathbf{R} + \alpha\mathbf{H}\mathbf{P}^e\mathbf{H}^T + (1-\alpha)\mathbf{H}\mathbf{B}\mathbf{H}^T, \quad 0 \leq \alpha \leq 1 \tag{6}$$

  - Assume α to be a random variable
  - Start with a prior distribution for α: $p(\alpha) \sim \mathcal{N}, \mathcal{B}, ..$
  - Use the data to construct a likelihood function: $p(\mathbf{d}|\alpha)$
  - Use Bayes' rule to find an updated estimate of α:

$$p(\alpha|\mathbf{d}) \approx p(\alpha) \cdot p(\mathbf{d}|\alpha) \tag{7}$$

  - Posterior α can be used as the prior for the next DA cycle

2.4 Hybrid EnKF-OI: Adaptive Form cont.

```
switch Prior
    case 'Gaussian'
```

$$p(\alpha) = \mathcal{N}\left(\alpha_f, \sigma_{\alpha_f}\right) \equiv \frac{1}{\sqrt{2\pi\sigma^2_{\alpha_f}}} \exp\left[-\frac{\left(\alpha - \alpha_f\right)^2}{2\sigma^2_{\alpha_f}}\right]$$

```
    case 'Beta'
```

$$p(\alpha) = \mathcal{B}(\gamma, \beta) \equiv \alpha^{\gamma-1}(1-\alpha)^{\beta-1}\frac{\Gamma(\gamma+\beta)}{\Gamma(\gamma)\Gamma(\beta)}$$

```
end
```

2.4 Hybrid EnKF-OI: Adaptive Form cont.

```
switch Prior
    case 'Gaussian'
```

$$p(\alpha) = \mathcal{N}\left(\alpha_f, \sigma_{\alpha_f}\right) \equiv \frac{1}{\sqrt{2\pi\sigma_{\alpha_f}^2}} \exp\left[-\frac{\left(\alpha - \alpha_f\right)^2}{2\sigma_{\alpha_f}^2}\right]$$

```
    case 'Beta'
```

$$p(\alpha) = \mathcal{B}(\gamma, \beta) \equiv \alpha^{\gamma-1}(1-\alpha)^{\beta-1}\frac{\Gamma(\gamma+\beta)}{\Gamma(\gamma)\Gamma(\beta)}$$

```
end
```

Likelihood:

$$\begin{aligned} \theta(\alpha) &= \text{trace}(\mathbf{R}) + \alpha\,\text{trace}\left(\mathbf{H}\mathbf{P}^e\mathbf{H}^T\right) + (1-\alpha)\,\text{trace}\left(\mathbf{H}\mathbf{B}\mathbf{H}^T\right) \\ p(\mathbf{d}|\alpha) &= \frac{1}{\sqrt{2\pi\theta(\alpha)}} \exp\left[-\frac{\mathbf{d}^T\mathbf{d}}{2\theta(\alpha)}\right] \end{aligned}$$

Posterior: $p(\alpha|\mathbf{d})$ is either near Gaussian or near Beta

2.5 Hybrid EnKF-OI: Illustration

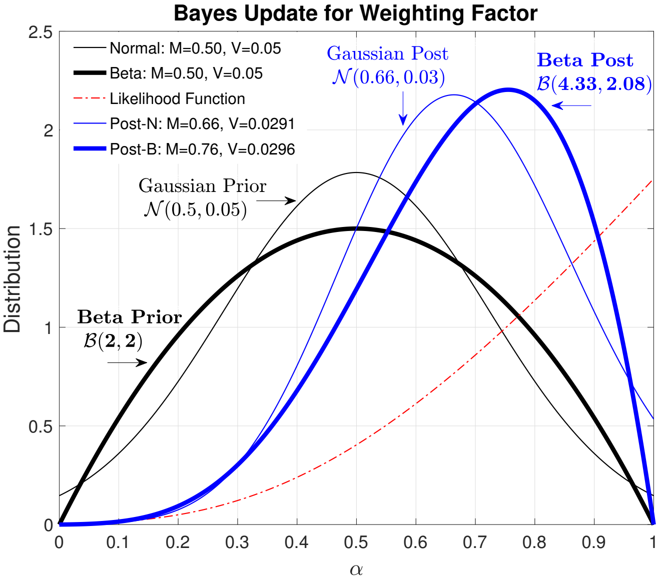
Scalar Example
6 parameters

$\mathbf{P}^e$	$\sigma_e^2 = 0.9$
$\mathbf{B}$	$\sigma_s^2 = 0.2$
$\mathbf{R}$	$\sigma_o^2 = 0.1$
$\mathbf{d}$	$d = 2.5$

2.5 Hybrid EnKF-OI: Illustration

Scalar Example
6 parameters

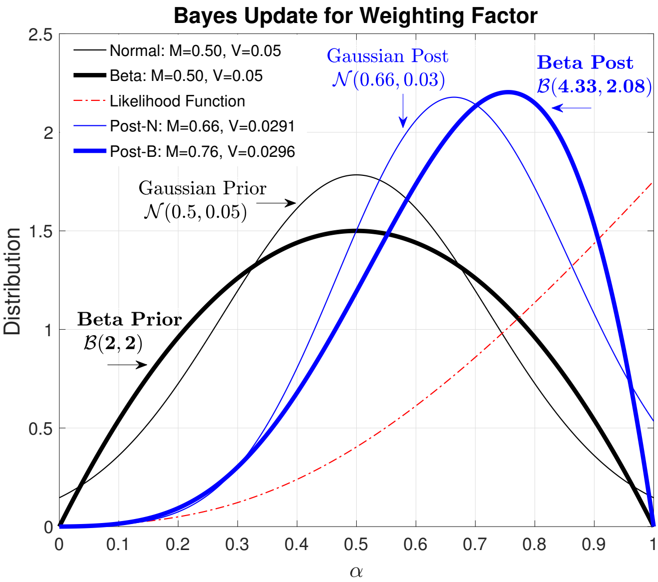
$\mathbf{P}^e$	$\sigma_e^2 = 0.9$
$\mathbf{B}$	$\sigma_s^2 = 0.2$
$\mathbf{R}$	$\sigma_o^2 = 0.1$
$\mathbf{d}$	$d = 2.5$



Scalar Example
6 parameters

$\mathbf{P}^e$	$\sigma_e^2 = 0.9$
$\mathbf{B}$	$\sigma_s^2 = 0.2$
$\mathbf{R}$	$\sigma_o^2 = 0.1$
$\mathbf{d}$	$d = 2.5$

Large bias causes α to increase (i.e., larger weight given to σ_e^2)



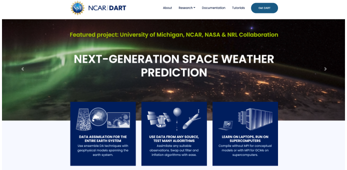
2.6 Hybrid EnKF-OI: Implementation

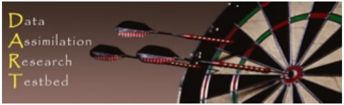
- Estimate moments of the hybrid weight pdf at each assimilation cycle using the data:
  - Maximizing the posterior requires finding cubic polynomial roots
  - Similar algorithm to existing adaptive inflation schemes

2.6 Hybrid EnKF-OI: Implementation

- Estimate moments of the hybrid weight pdf at each assimilation cycle using the data:
  - Maximizing the posterior requires finding cubic polynomial roots
  - Similar algorithm to existing adaptive inflation schemes
- DART implementation available

dart.ucar.edu
docs.dart.ucar.edu

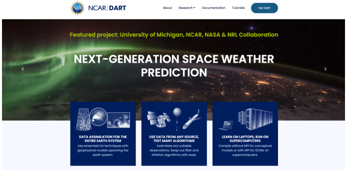


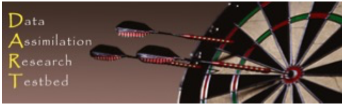


2.6 Hybrid EnKF-OI: Implementation

- Estimate moments of the hybrid weight pdf at each assimilation cycle using the data:
  - Maximizing the posterior requires finding cubic polynomial roots
  - Similar algorithm to existing adaptive inflation schemes
- DART implementation available

dart.ucar.edu
docs.dart.ucar.edu

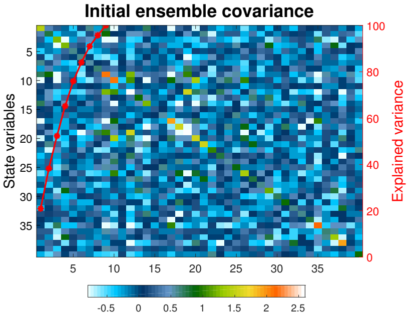


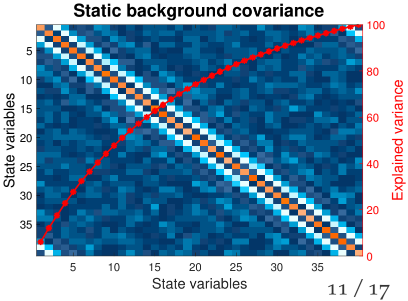


- Can assume the hybrid weight to be spatially-varying
  - Biases are not homogenous in space
  - Heterogenous observation networks (densely observed regions tend to have small ensemble spread)
  - Need to assimilate the observations serially

3.1 Experiments using L96

- L96: 40 variables
- Observe every other variable ($\mathbf{R} = 1$)
- Observe every 5 time steps
- **B** Climatological run (1000)
- $p(\alpha) \sim \mathcal{N}(0.5, 0.1)$





3.1 Experiments using L96

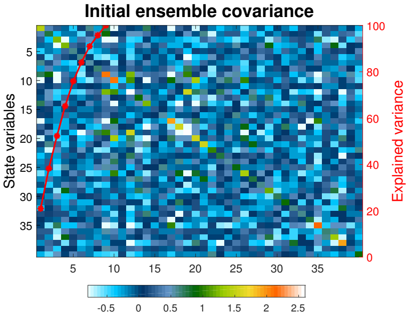
- L96: 40 variables
- Observe every other variable ($\mathbf{R} = 1$)
- Observe every 5 time steps
- **B** Climatological run (1000)
- $p(\alpha) \sim \mathcal{N}(0.5, 0.1)$

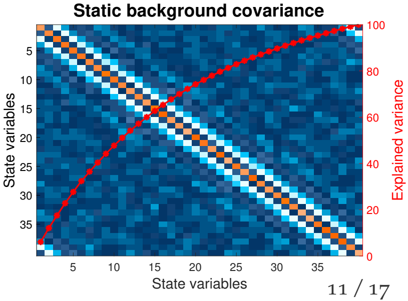
Sensitivity Tests [1]
Perfect OSSEs

- Ensemble size
- Obs. Network

Sensitivity Tests [2]
Model Errors

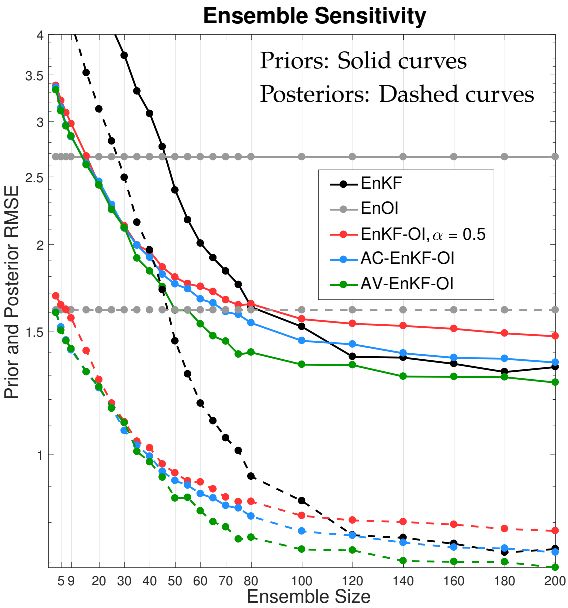
- Inflation
- Localization





3.2 Sensitivity Tests: Ensemble Size

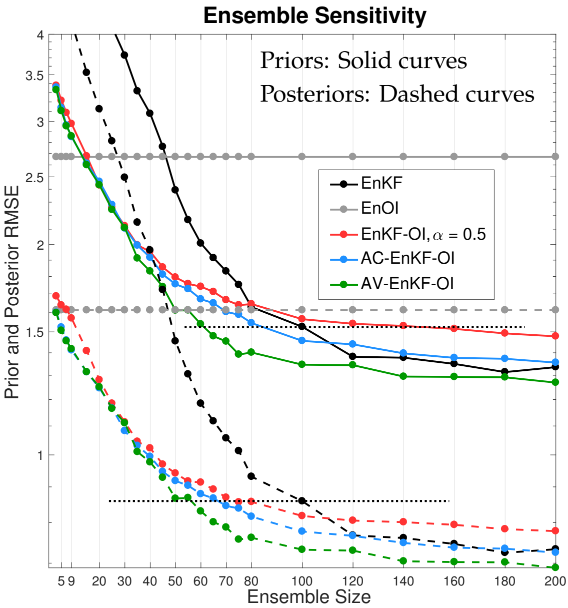
1. **EnKF**
2. **EnOI**: EnKF with fixed **B** (Hybrid; $\alpha = 0$)
3. **EnKF-OI**; $\alpha = 0.5$
4. **AC-EnKF-OI**: Adaptive, spatially-Constant EnKF-OI
5. **AV-EnKF-OI**: Adaptive, spatially-Varying EnKF-OI



3.2 Sensitivity Tests: Ensemble Size

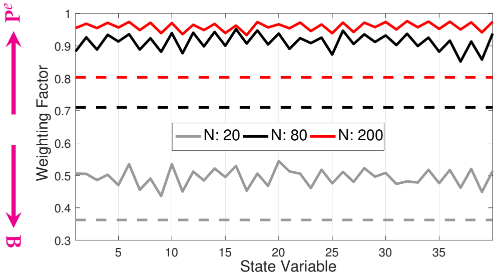
1. **EnKF**
2. **EnOI**: EnKF with fixed **B** (Hybrid; $\alpha = 0$)
3. **EnKF-OI**; $\alpha = 0.5$
4. **AC-EnKF-OI**: Adaptive, spatially-Constant EnKF-OI
5. **AV-EnKF-OI**: Adaptive, spatially-Varying EnKF-OI

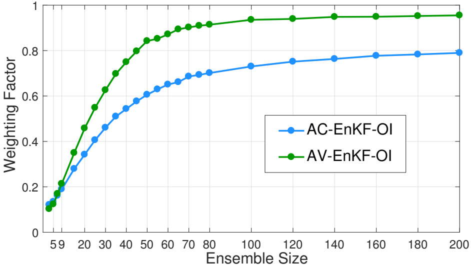
★ **EnKF**'s accuracy is reproduced by the hybrid schemes with 40 – 50% less ensemble members



3.2 Sensitivity Tests: Ensemble Size

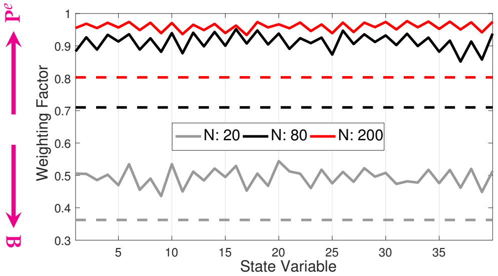
- **AC-EnKF-OI**: Dashed lines
- **AV-EnKF-OI**: Solid lines



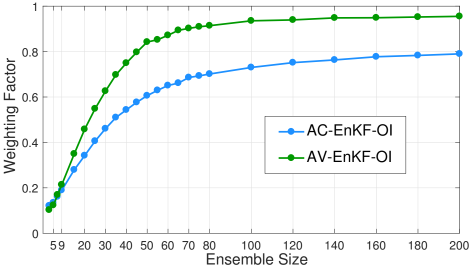


3.2 Sensitivity Tests: Ensemble Size

- **AC-EnKF-OI**: Dashed lines
- **AV-EnKF-OI**: Solid lines

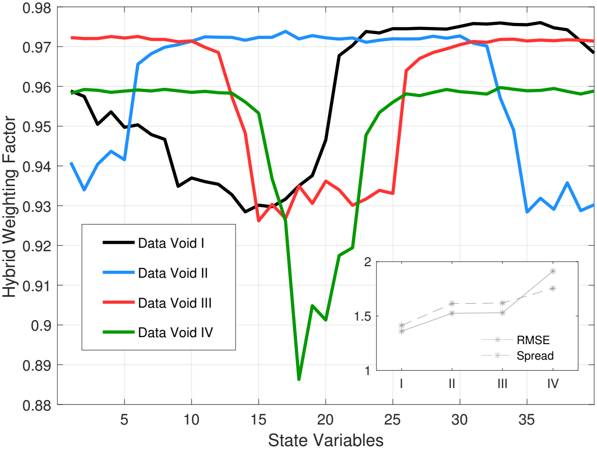


- For small ensembles, both adaptive spatially-constant and varying schemes behave the same
- AV-EnKF-OI responds more efficiently to changes in the ensemble



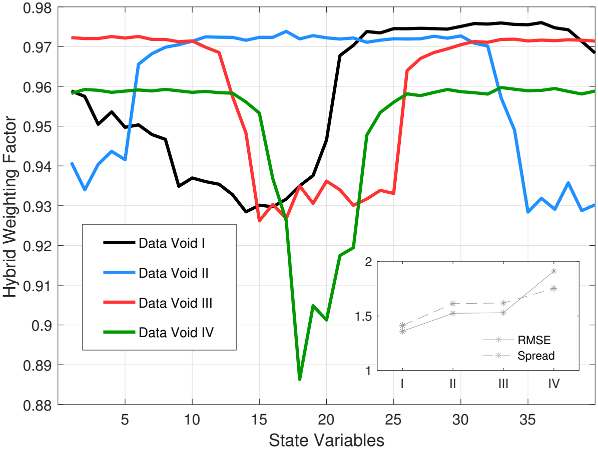
3.3 Sensitivity Tests: Observation Network

- **Data Void I:** Observe the first 20 variables
- **Data Void II:** Observe the first and last 5 variables
- **Data Void III:** Observe 10 center variables
- **Data Void IV:** Observe 5 center variables



3.3 Sensitivity Tests: Observation Network

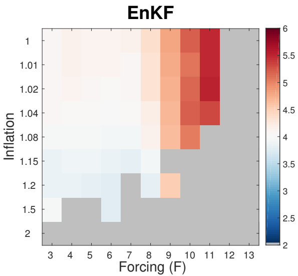
- **Data Void I:** Observe the first 20 variables
- **Data Void II:** Observe the first and last 5 variables
- **Data Void III:** Observe 10 center variables
- **Data Void IV:** Observe 5 center variables

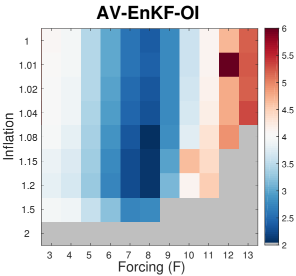


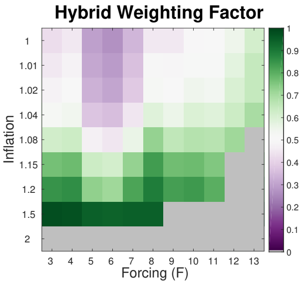
- In densely observed regions, the ensemble spread decreases
- Hybrid scheme places weight more on **B** to increase the variance, allowing better data fit

Sensitivity Tests: Model Errors + Inflation

- Ensemble size: 20
- Model error; vary $3 \leq F \leq 13$
- $\mathbf{B}$ is generated in each case using biased F
- No localization

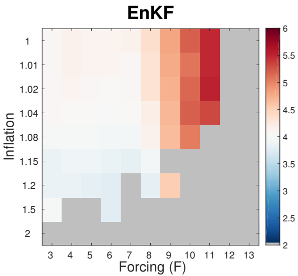


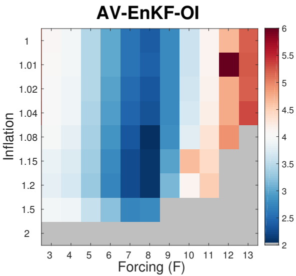




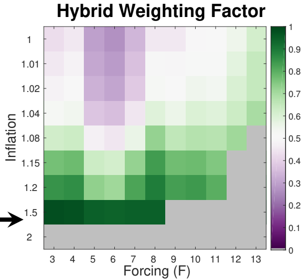
Sensitivity Tests: Model Errors + Inflation

- Ensemble size: 20
- Model error; vary $3 \leq F \leq 13$
- **B** is generated in each case using biased F
- No localization

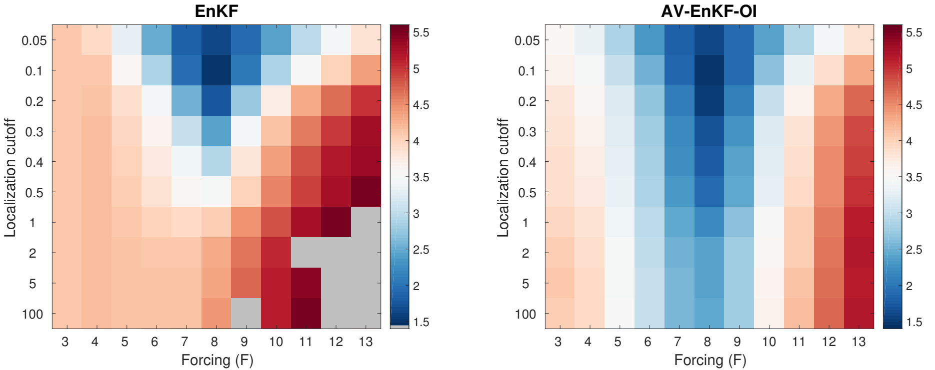




- Hybrid scheme: better stability and more accurate even in very biased conditions
- As inflation increases, adaptive α increases (more weight on the ensemble covariance)

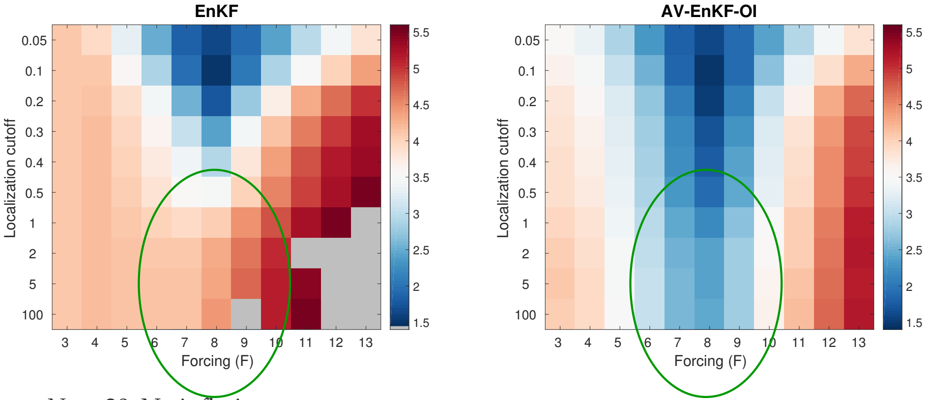


Sensitivity Tests: Model Errors + Localization



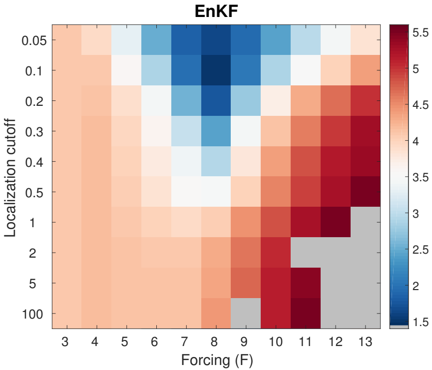
- $N_e = 20$, No inflation
- Vary both F and localization length scale
- Adaptive hybrid scheme is systematically better than the EnKF for all tested cases

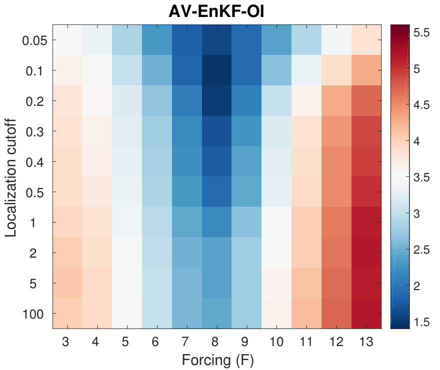
Sensitivity Tests: Model Errors + Localization

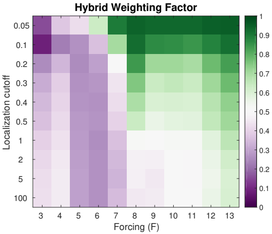


- $N_e = 20$, No inflation
- Vary both F and localization length scale
- Adaptive hybrid scheme is systematically better than the EnKF for all tested cases
- With very little to no localization, hybrid scheme still performs exceptionally well
- Does the climatological flavor from **B** mitigate spurious correlations?

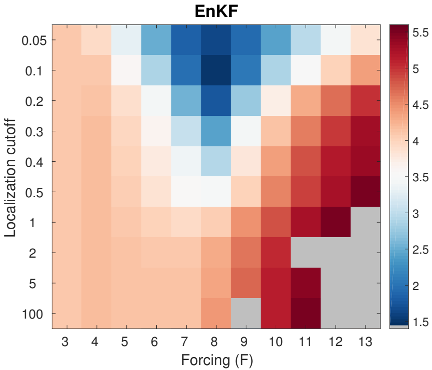
Sensitivity Tests: Model Errors + Localization

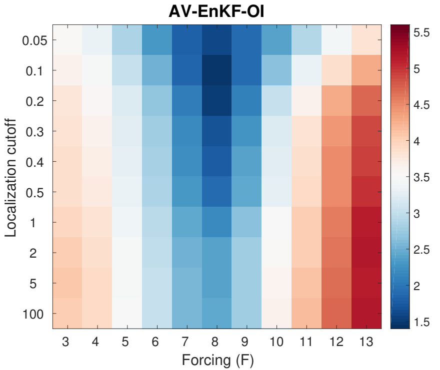




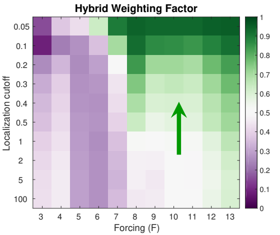


Sensitivity Tests: Model Errors + Localization

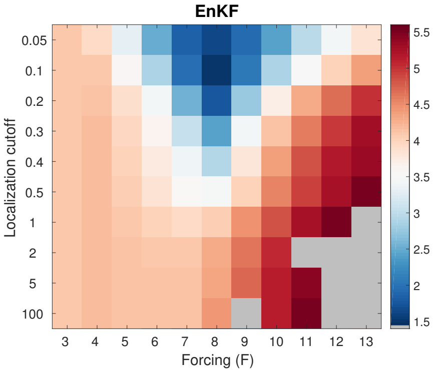


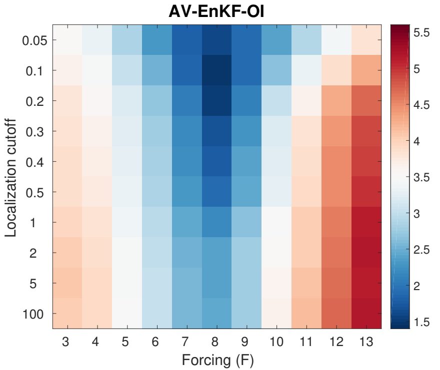


- For chaotic behaviour (i.e., $F \geq 8$): As localization increases, α increases

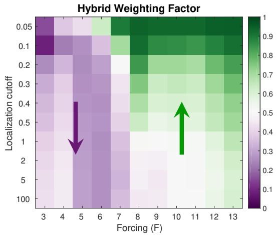


Sensitivity Tests: Model Errors + Localization





- For chaotic behaviour (i.e., $F \geq 8$): As localization increases, α increases
- Less chaotic (smaller ensemble variance): α decreases to *bring-in* variability from **B**



4.1 Concluding Remarks

- Prior (background) ensemble covariance **must** be enhanced
- On top of inflation and localization, hybridizing $\mathbf{P}^e$ with stationary OI-based background covariances can be helpful and perhaps crucial
- The adaptive scheme uses available data through Bayes rule to determine the relative weighting between the ensemble and the static covariance
- Lorenz-96 experiments show promising performance

4.1 Concluding Remarks

- Prior (background) ensemble covariance **must** be enhanced
- On top of inflation and localization, hybridizing $\mathbf{P}^e$ with stationary OI-based background covariances can be helpful and perhaps crucial
- The adaptive scheme uses available data through Bayes rule to determine the relative weighting between the ensemble and the static covariance
- Lorenz-96 experiments show promising performance

EnKF	**Adaptive Hybrid EnKF-OI**
– Only flow-dependent covariance	– OI flavor & flow-dependent information
– Requires a large ensemble size	– Works well with fairly small ensembles
– Fair computational cost	– Storage, additional IO cost
– Strong tuning (inf, loc, ..)	– Fully adaptive, requires less inf, loc, ..
– Strong biases cause divergence	– More stable; able to switch to EnOI

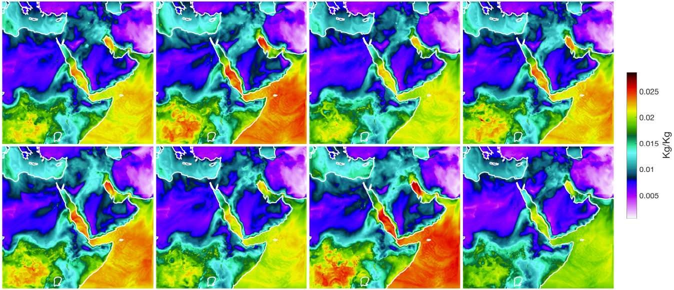
4.1 Concluding Remarks

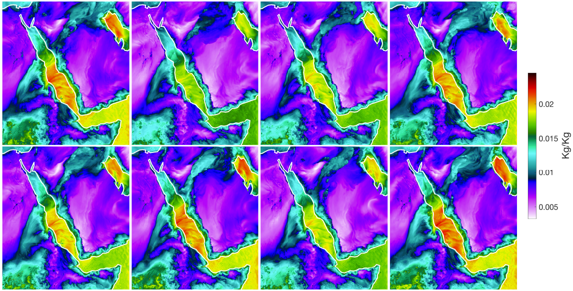
- Prior (background) ensemble covariance **must** be enhanced
- On top of inflation and localization, hybridizing $\mathbf{P}^e$ with stationary OI-based background covariances can be helpful and perhaps crucial
- The adaptive scheme uses available data through Bayes rule to determine the relative weighting between the ensemble and the static covariance
- Lorenz-96 experiments show promising performance

EnKF	**Adaptive Hybrid EnKF-OI**
– Only flow-dependent covariance	– OI flavor & flow-dependent information
– Requires a large ensemble size	– Works well with fairly small ensembles
– Fair computational cost	– Storage, additional IO cost
– Strong tuning (inf, loc, ..)	– Fully adaptive, requires less inf, loc, ..
– Strong biases cause divergence	– More stable; able to switch to EnOI

El Gharamti, M. (2021). Hybrid Ensemble-Variational Filter: A Spatially and Temporally Varying Adaptive Algorithm to Estimate Relative Weighting. Monthly Weather Review, 149(1), 65-76.

4.2 Future: Applications to Earth System Models





4.2 Future: Applications to Earth System Models