UNDERSTANDING & SEPARATING THE ROLES OF DYNAMICS & STATISTICS IN DATA ASSIMILATION

Malaquias Pena[1] and Zoltan Toth

Environmental Modeling Center
NCEP/NWS/NOAA

[1] SAIC at EMC/NCEP/NOAA

Acknowledgements:
Mozheng Wei, Takemasa Miyosi, & Roman Krzysztofowicz

DA Workshop
4-8 February 2008, Banff, Canada

OUTLINE / SUMMARY

- **STATE ESTIMATION**
  - Bayesian fusion of
    - New observations
    - Prior
- **PRIOR**
  - Dynamical forecast
    - Effect of all prior observations included
    - Dynamical constraints
- **FUSION**
  - Propagate information from observations to all state variables
    - Error covariance crucial
- **COVARIANCE ESTIMATION**
  - Climatological sample
    - Large sample BUT
    - Not representative of particular cases
  - Case dependent sample
    - Ensembles
      - How to reduce effect of sampling errors?
- **ENSEMBLE DA**
  - "Fully ensemble-based DA"
    - Analysis & forecast steps share full error covariance
      - Inflation/localization noise cycled => negative impact?
  - *ET + 3DVAR*
    - *Analysis step feeds error variance into forecast step*
    - *Forecast step feeds error correlation into analysis step*
      - *Noise from regularization in analysis step not cycled => better covariance => better state estimates?*

BACKGROUND ON DA

- **Goal**
  – Assess state of under-observed dynamical systems
- **Needed when**
  – Observations are
    - Erroneous
      – How to reduce errors in observational data?
    - Scarce
      – How to fill in gaps in observational data?
- **Types of constraints** used in DA
  – Past observational data
    - Climatology
    - Conditional climatology
    - Persistence
  – ***Dynamics***
    - Laws of nature

USE OF DYNAMICS IN DA

- **Concept**
  - Introduce dynamical constraint
    - Based on laws of nature
      - A priori info
- **Methodologies**
  - Use balance etc constraints
  - Use short range Numerical Prediction (NP)
- **“Vicious” cycle**
  - Prepare analysis
    - Needed as initial state for forecast
  - Run forecast
    - Needed to prepare analysis
- **Consequence**
  - Worry about convergence of DA cycles (not one step)
    - If & how fast convergence is?
    - How stable DA cycles are?
- **Practical solution**
  - ***DA cycled with Numerical Prediction*** (NP)
    - Forecast carries information from past
      - Ideally, all past info folded in?

DA CYCLED WITH NP

- **Components**
  - "First Guess" (FG) or Background
    - Short range numerical prediction
      - Prior information
  - Observations
    - To update prior information
- **Methodology**
  - Statistical combination of components
    - Bayesian principles
- **Algorithm**
  - Based on point-wise comparison of FG & observations
    - How to spread information in space?
      - ***Background error covariance*** (B)

HOW TO DEFINE BACKGROUND ERROR COVARIANCE?

- **Climatologically** – statistics of
  - Perceived error in FG (truth not known)
    - Laden with noise (due to noise in analysis)
  - Difference between lagged forecasts verifying at same time
    - “NMC method”
- **Dynamically** – statistics of
  - ***Ensemble forecasts***
    - Case dependent estimate
      - May help even in cases of linear error growth

PROCESS OF ENSEMBLE-BASED DA

- Project state into future
  - Numerical prediction
    - Cycles state estimate
      - Any noise in initial condition hurts state estimate
- Project initial info on covariance into future
  - Run ensemble
    - Cycles error covariance estimate
      - ***Any noise in initial info hurts covariance estimate***
- Estimate forecast error covariance
  - Based on finite sample of ensemble forecasts
    - Typically small sample due to high cost =>
      - ***How to limit noise in covariance estimate?***
- Collect new observations
  - Estimate observational error
- Combine FG & observations using error estimates
  - Noise in either projected state or covariance info hurts analysis

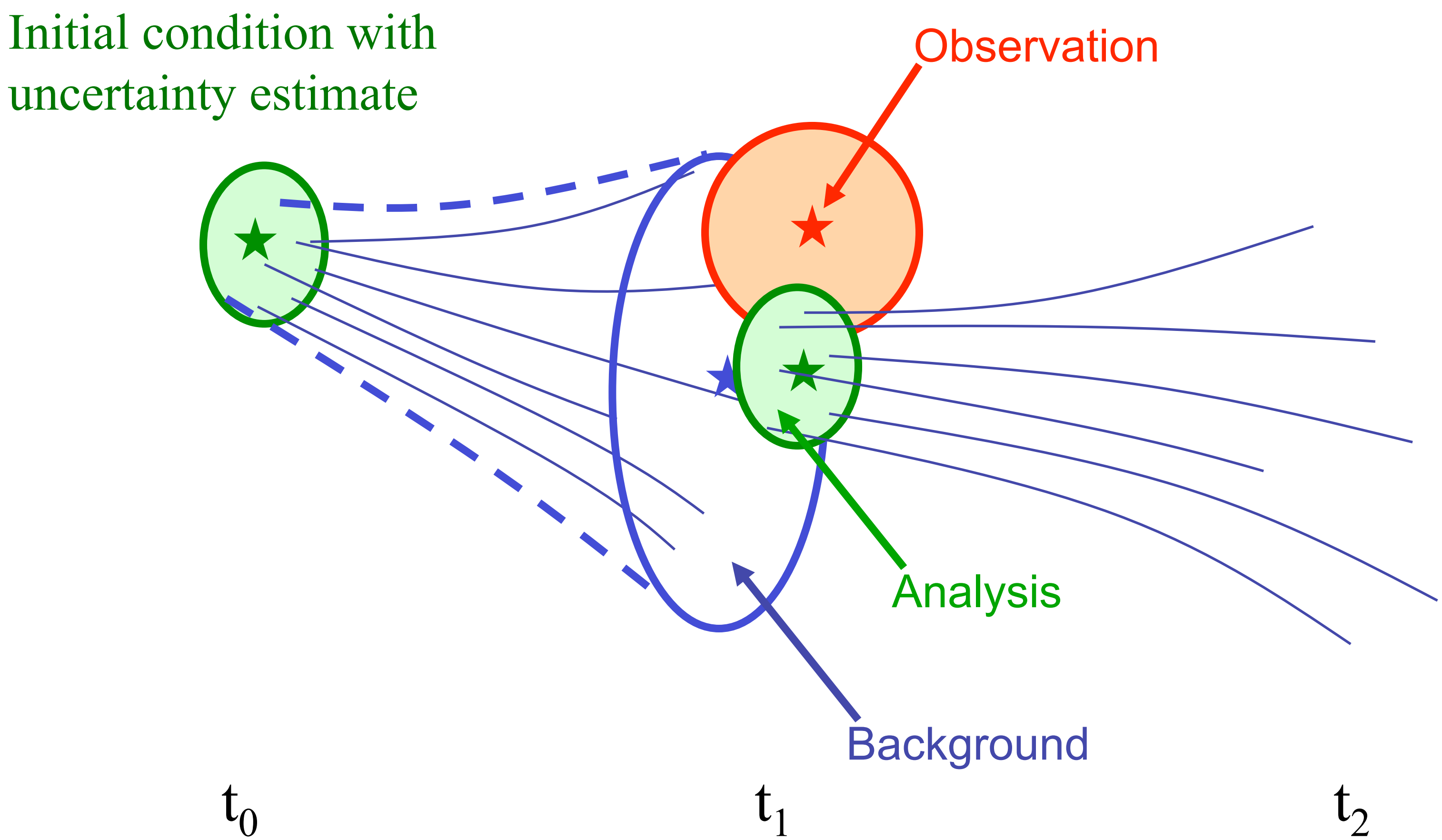
Dynamics

Statistics

ERROR COVARIANCE ESTIMATION

- Ingredients
  - Ensemble forecasts
    - Dynamical projection of prior covariance info
  - Statistical estimation
    - Small sample leads to filter divergence
- Methodology
  - Ensemble-based DA – ETKF-type methods
    - Modulate ens perts to avoid filter divergence
      - Covariance inflation – introduce noise
      - Localization
    - Cycle noisy covariance estimate
- Result
  - Noisy state and covariance estimates?
- Solution
  - ***Divorce dynamics from statistics***
    - Et + 3DVAR

Ensemble-based DA Experiments



How well the ensemble forecasts sample the background uncertainty? How much information the observations add?

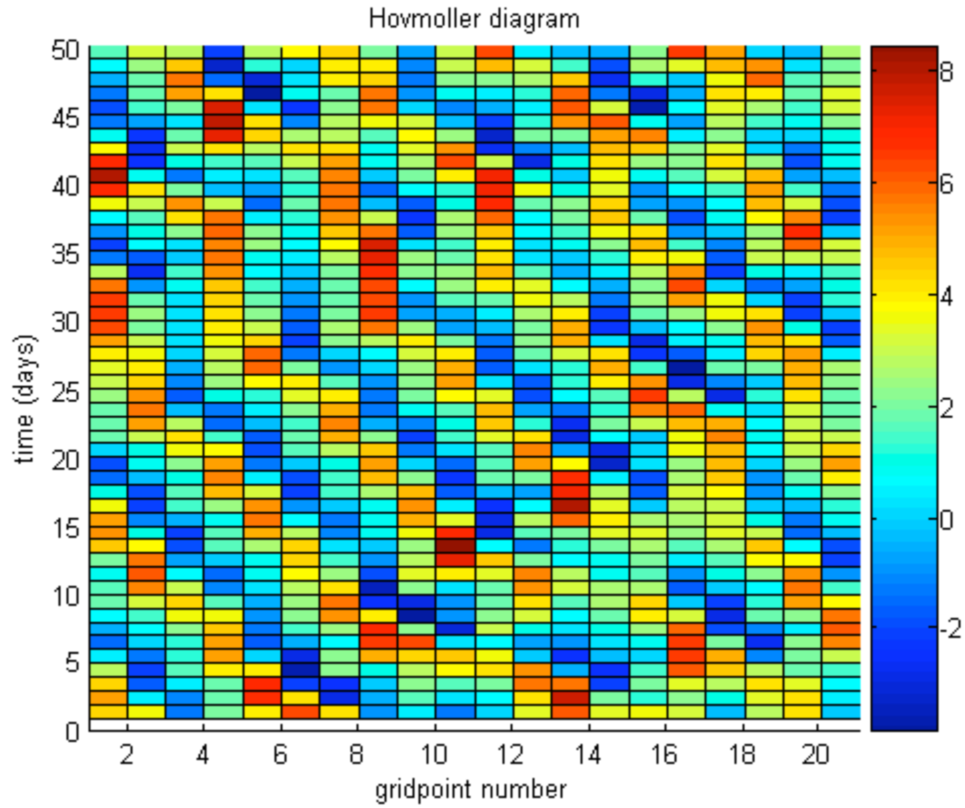
Lorenz 96 Model (with '07 pars)

$$\frac{\mathbf{dx_m}}{\mathbf{dt}} = (\mathbf{x}_{m+1} - \mathbf{x}_{m-2})\mathbf{x}_{m-1} - \mathbf{x}_m + \mathbf{F}$$

Where F= 5.1 and m=1,..,21

Experimental setting:

- Perfect model scenario
- One observation per grid-point
- Observational error: uncorrelated normally distributed random noise with unit variance (R=I)
- 6-hr assimilation cycle



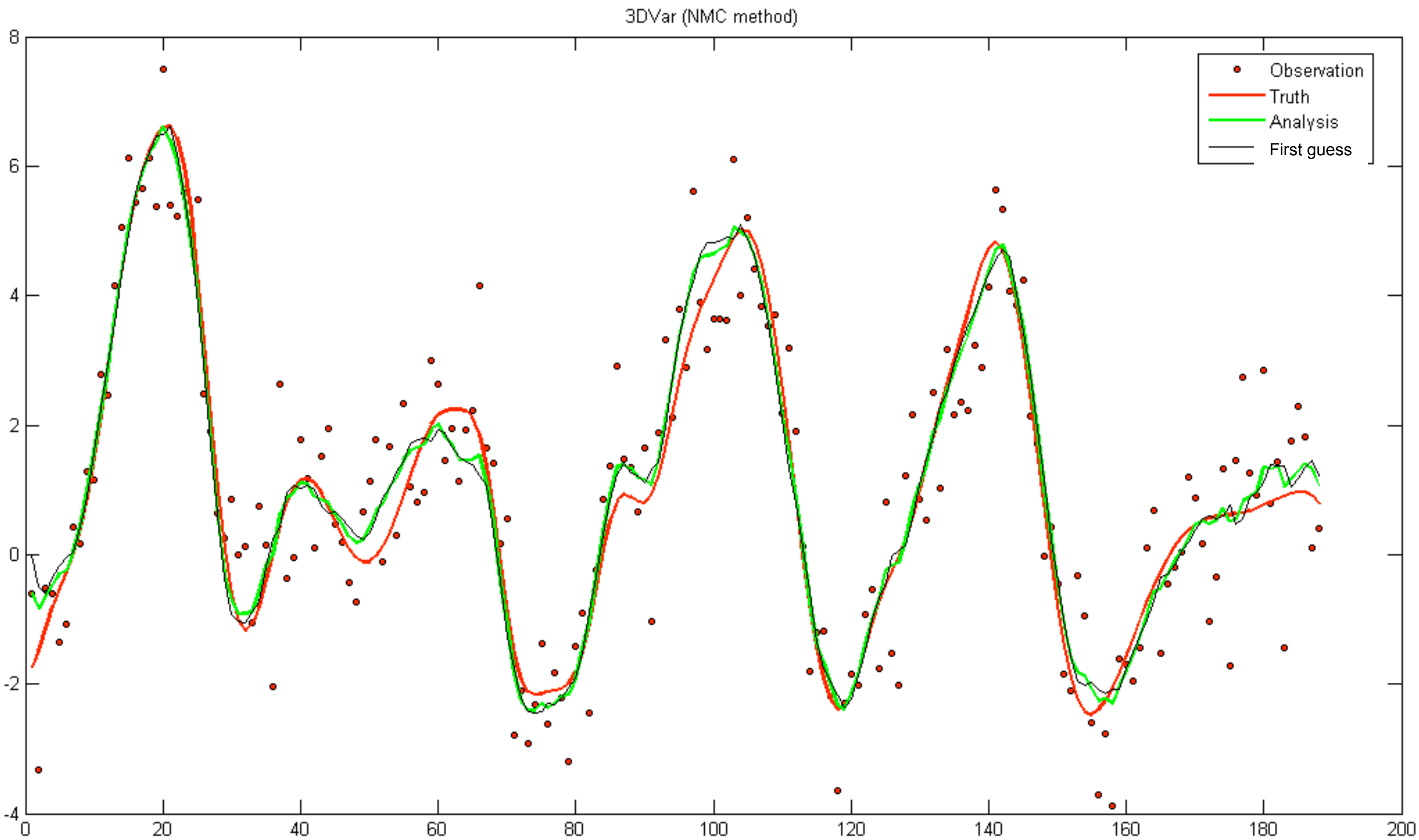
3DVar DA - Benchmark

- Minimizing the following cost function

$$\mathbf{J}(\mathbf{x}) = \frac{1}{2}[(\mathbf{x}-\mathbf{x}^{\mathbf{b}})\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^{\mathbf{b}}) + (\mathbf{y}^{\mathbf{o}}-\mathbf{H}\mathbf{x})^{\mathbf{T}}\mathbf{R}^{-1}(\mathbf{y}^{\mathbf{o}}-\mathbf{H}\mathbf{x})]$$

Inverse of background error covariance

Time series at one grid-point



Background error covariance: B= αB_o, where Bo is obtained from climatology and α a tuning parameter. α=0.05

ETKF

- $Z_a = Z_f T C^T$
- $T = C(G+I)^{-1/2}$

C, G eigenvector and eigenvalues of $Z_f^T H^T R^{-1} H Z_f$
(H=I used)

- $B = Z_f Z_f^T$

- $A = Z_a Z_a^T = Z_f T (Z_f T)^T$
- Full covariance shared between state & covariance update steps
- Covariance inflation and/or localization of B ***cycled***

Z is an MxK matrix whose columns are the K ensemble perturbations (departure from ensemble average) and M is the dimension of the state vector. Subindex a refers to analysis and f to forecast

ETKF

- $Z_a = Z_f T C^T$
- $T = C(G+I)^{-1/2}$

C, G eigenvector and eigenvalues of $Z_f^{\,T} H^T R^{-1} H Z_f$ (H=I used)

- $B = Z_f Z_f^T$

- $A = Z_a Z_a^T = Z_f T\,(Z_f T)^T$
- Full covariance shared between state & covariance update steps
- Covariance inflation and/or localization of B ***cycled***

ET + 3DVAR

- $Z_a = Z_f T C^T$
- $T = C G^{-1/2}$

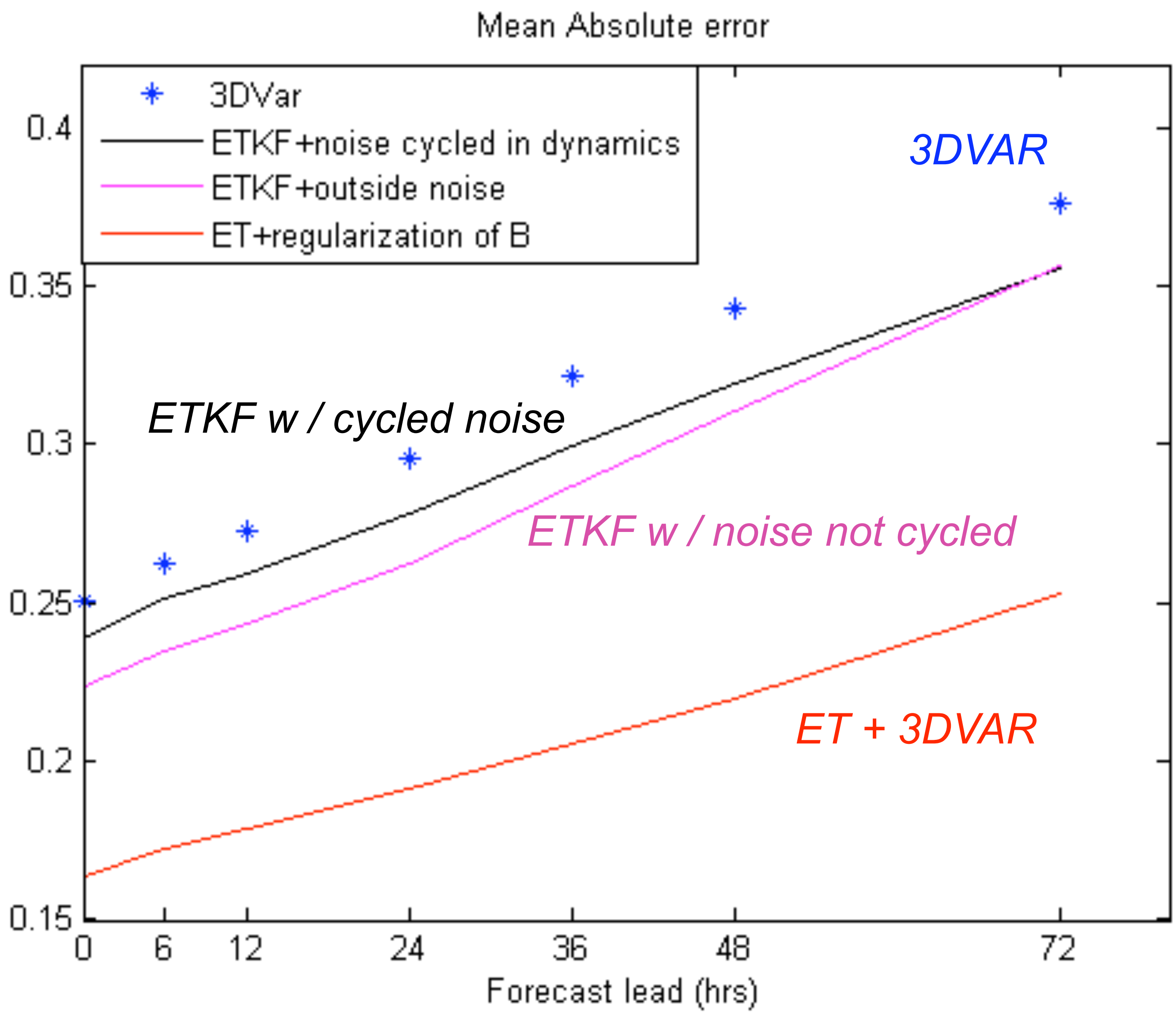
C and G are eigenvector and eigenvalues of $Z_f^{\,T} A^{-1} Z_f$

- $B = Z_f Z_f^T$ fed into analysis step

- $A^{-1} = B^{-1} + R^{-1}$ fed into ens pert generation step
- No noise is added into ens. perts. - “pure” dynamics
- Statistical manipulation of B not fed back into covariance – only variance affected

Z is an MxK matrix whose columns are the K ensemble perturbations (departure from ensemble average) and M is the dimension of the state vector. Subindex a refers to analysis and f to forecast

EFFECT OF SEPARATING ROLES OF DYNAMICS & STATISTICS



ETKF & other Ensemble-based DA methods

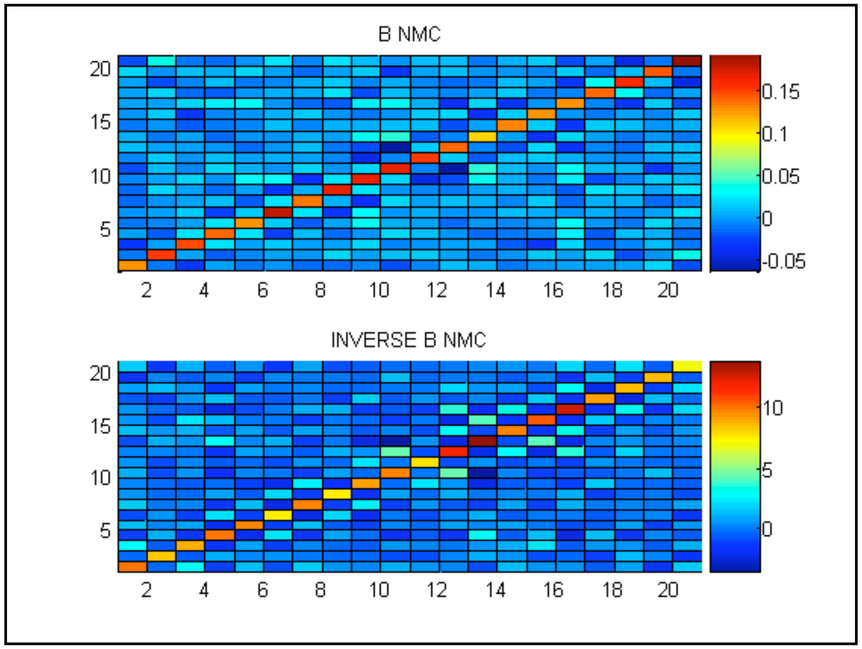
- Ensemble-based statistics of B are rank deficient and subject to sampling error
- Statistical regularization techniques to remedy prob lem
  - Add noise to analysis perturbations to avoid underestimation of B - Miller et al (1994) and Corazza et al (2002)
  - Blend ensemble B and 3DVar B – “hybrid” method - Hamill and Snyder (2000)
  - Localize effect of covariance
    - Shur product – Houtekamer et al
    - LETKF – Ott, Szunyogh et al., 2003

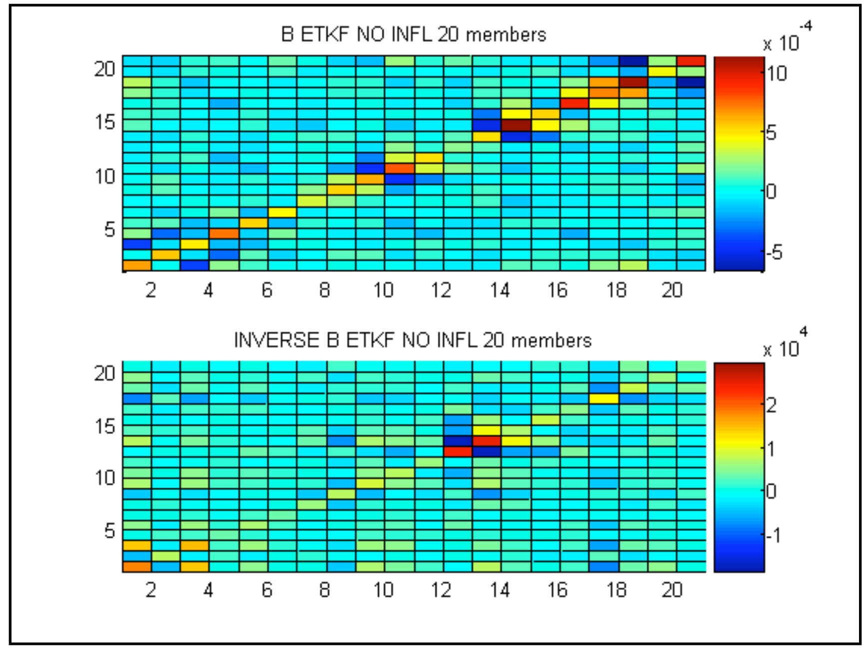
– Addition of noise used here with ETKF

Snapshot of B and B^{-1}

3DVAR – NMC method

ETKF without inflation:

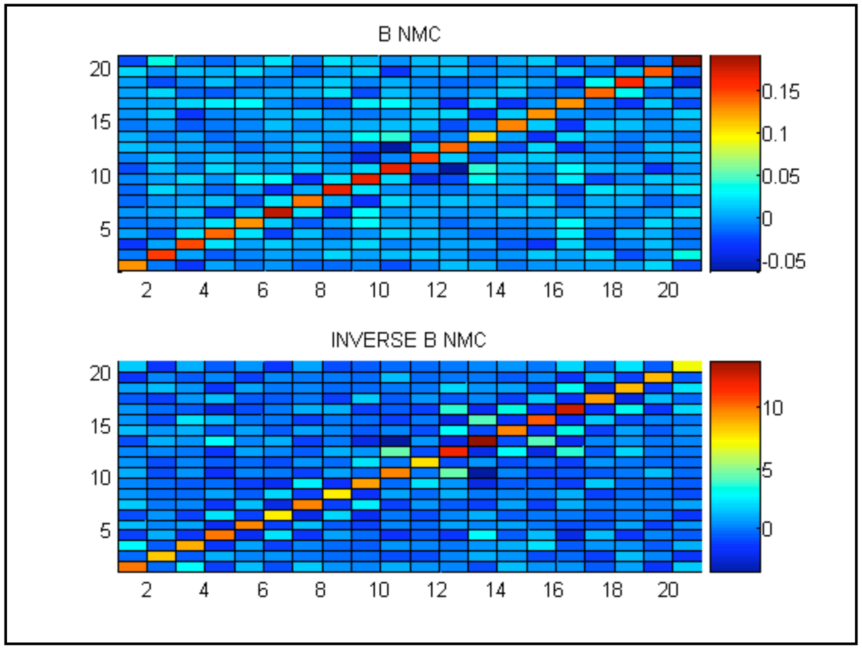




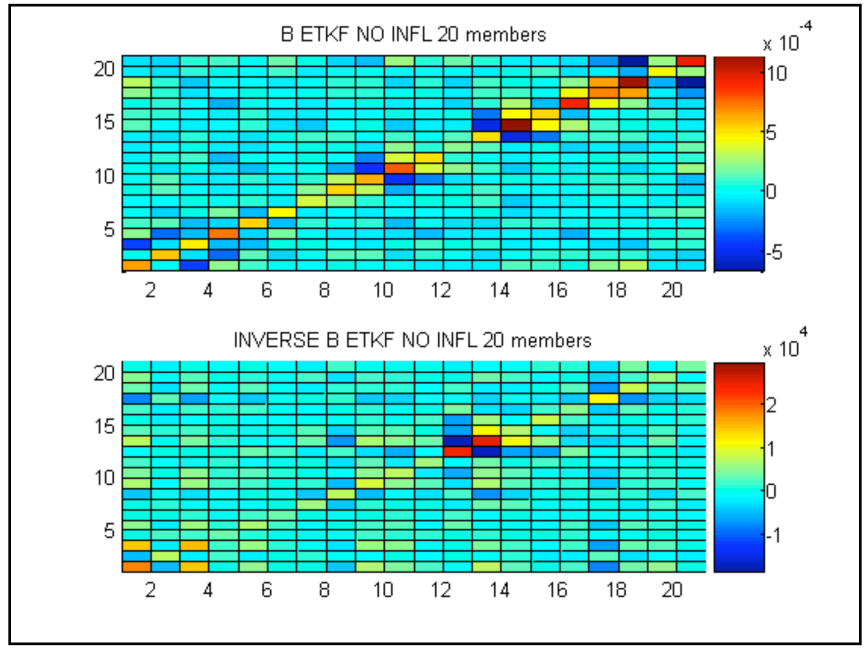
Well-conditioned, stable

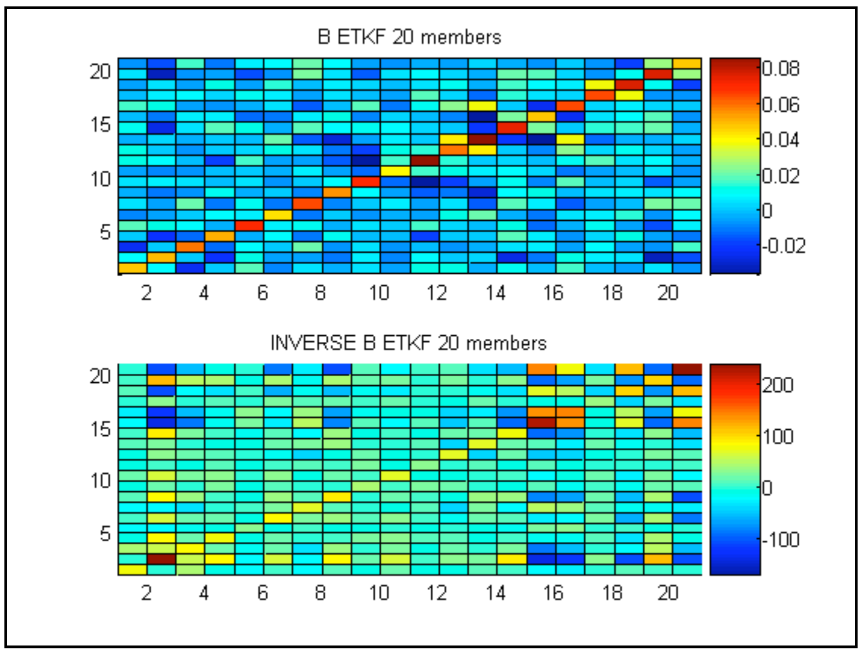
Very unstable

NMC method



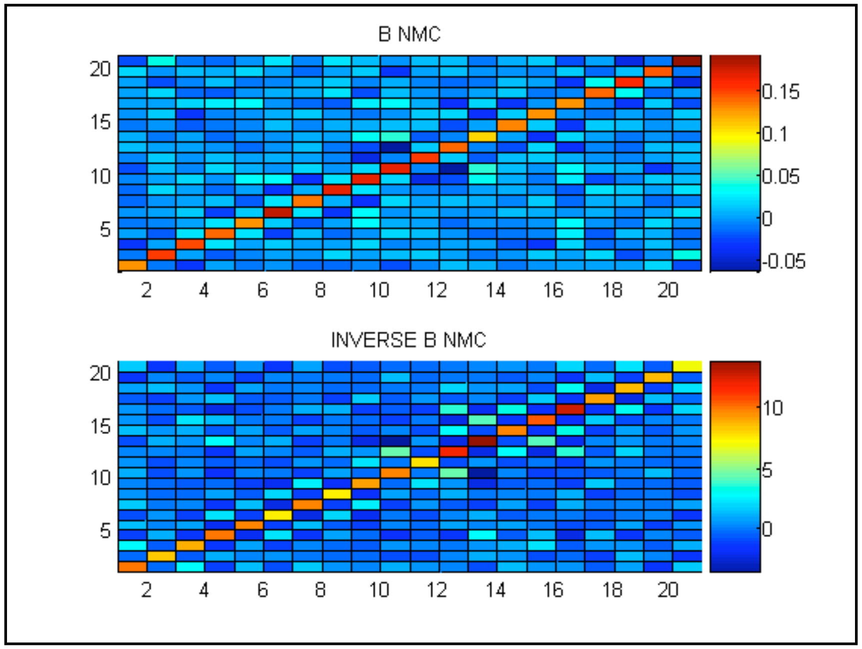
ETKF, no inflation



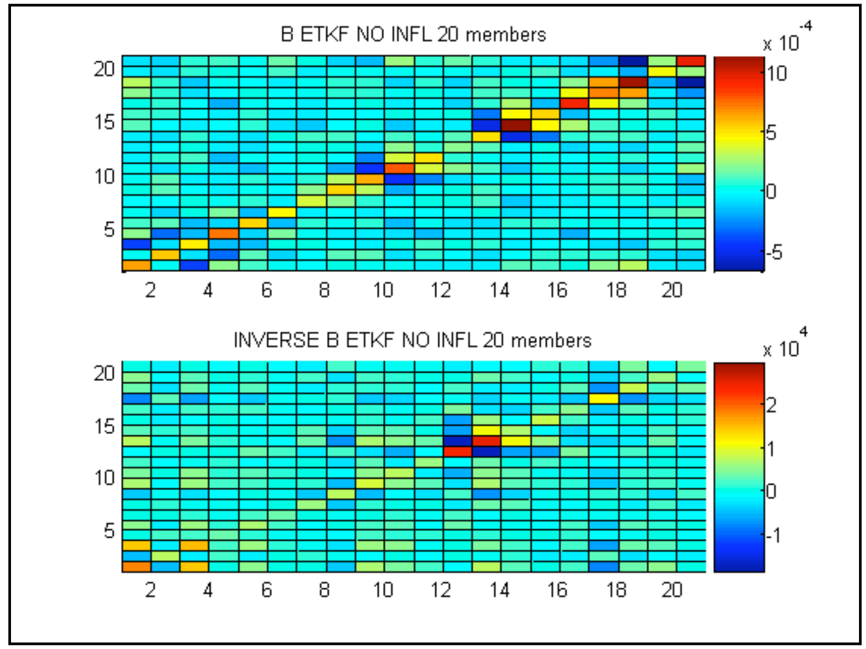


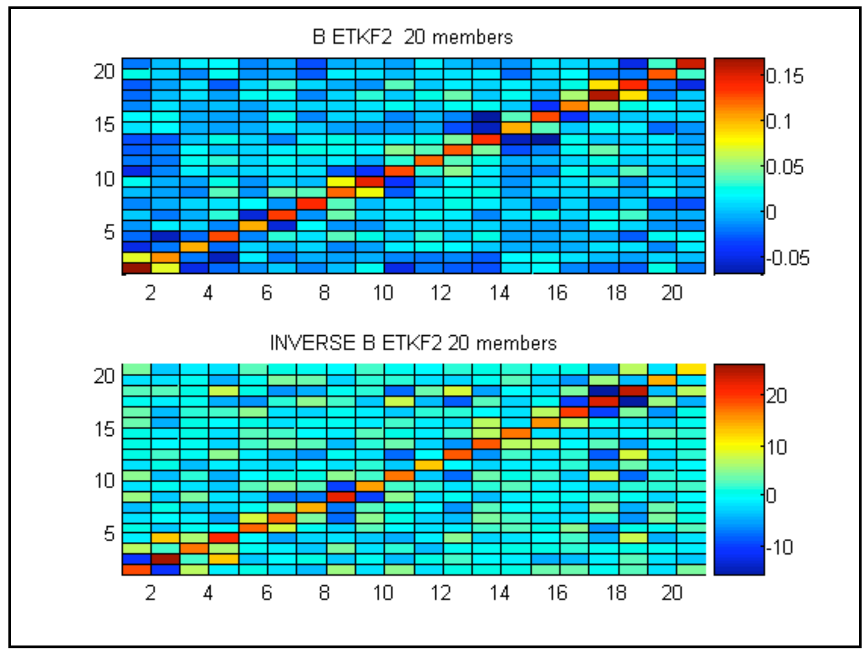
ETKF with random perturbations added. Inverse becomes stable; However, noise cycled

NMC method



ETKF, no inflation

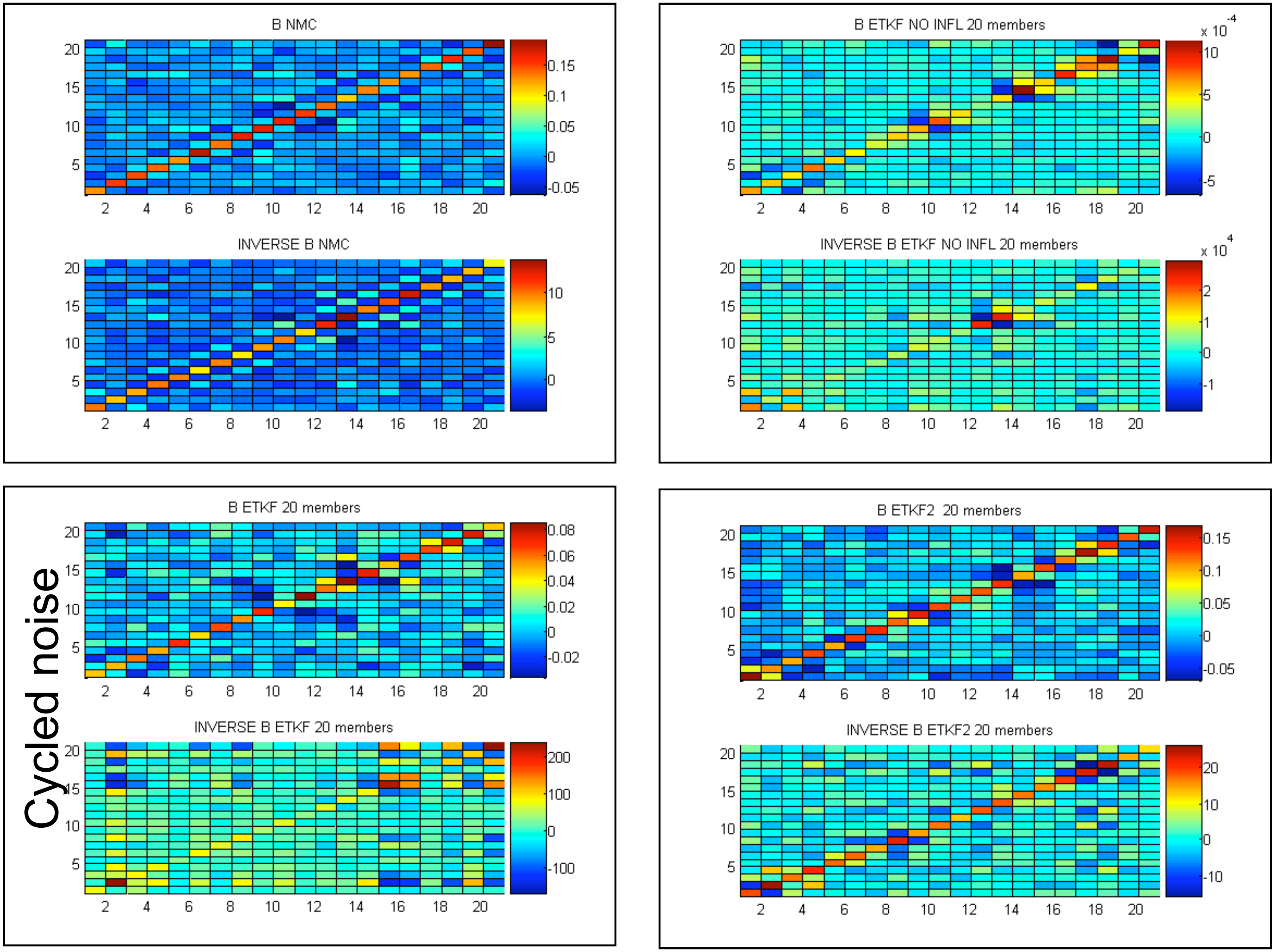




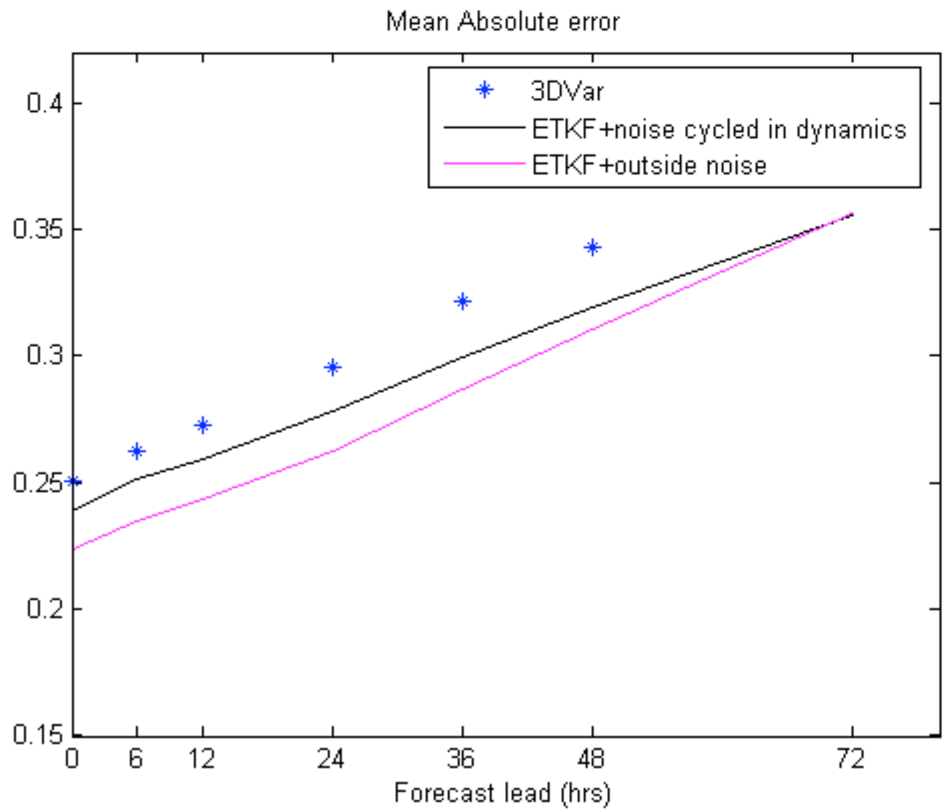
Two ensembles run with ETKF, one with (used for estimating B – regularization, then discarded), another without addition of noise (used for cycling covariance); Noise still impacts initial perturbs

NMC method

ETKF, no inflation



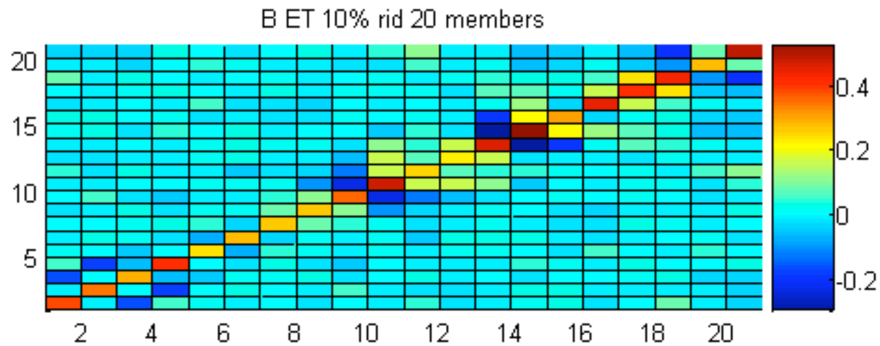
Noise for B only

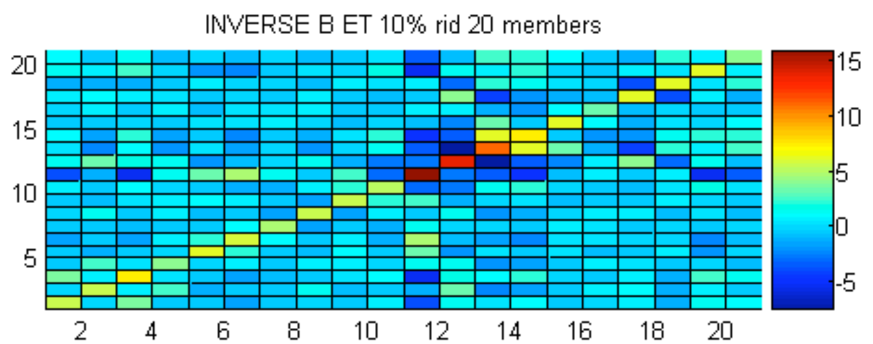

Mean Absolute error
3DVar
ETKF+noise cycled in dynamics
ETKF+outside noise
0.4
0.35
0.3
0.25
0.2
0.15
0
6
12
24
36
48
72
Forecast lead (hrs)

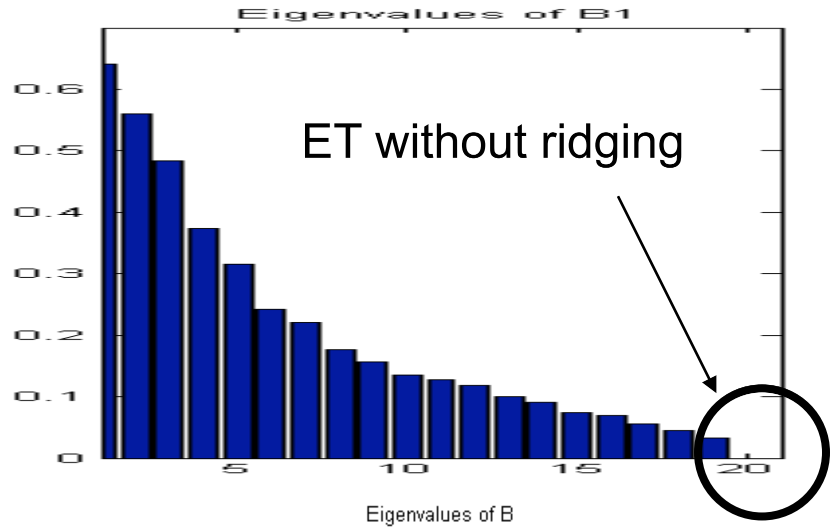
Ridging procedure in ET + 3DVAR

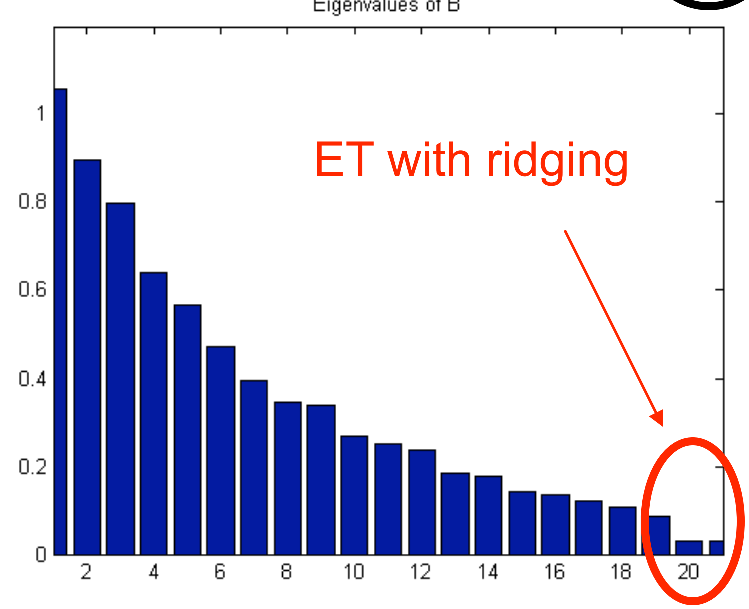
Statistics

Add small (10%) value to diagonal of B









OUTLINE / SUMMARY

- **STATE ESTIMATION**
  - Bayesian fusion of
    - Prior
    - New observations
- **PRIOR**
  - Dynamical forecast
    - Effect of all prior observations included
    - Dynamical constraints
- **FUSION**
  - Propagate information from observations to all state variables
    - Error covariance crucial
- **COVARIANCE ESTIMATION**
  - Climatological sample
    - Large sample BUT
    - Not representative of particular cases
  - Case dependent sample
    - Ensembles
      - How to reduce effect of sampling errors?
- **ENSEMBLE DA**
  - "Fully ensemble-based DA"
    - Analysis & forecast steps share full error covariance
      - Inflation/localization noise cycled => negative impact?
  - *ET + 3DVAR*
    - *Analysis step feeds error variance into forecast step*
    - *Forecast step feeds error correlation into analysis step*
      - *Noise from regularization in analysis step not cycled => better covariance => better state estimates?*

BACKGROUND

USE OF ENSEMBLES IN DA

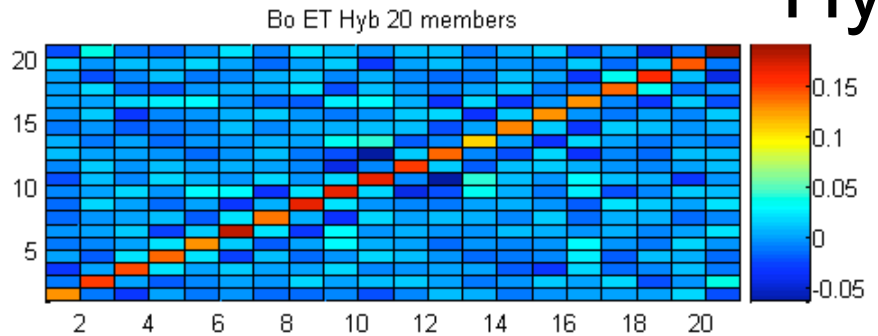
- Error covariance estimation
  - Needed even in quasi-linear regime?

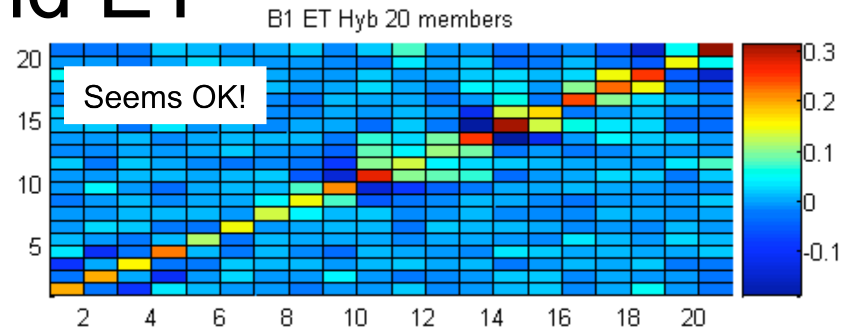
- State projection
  - Moderately non-linear regime
    - Use ensemble mean for estimating future state
  - Highly non-linear regime
    - Particle filtering needed?
      - Future study

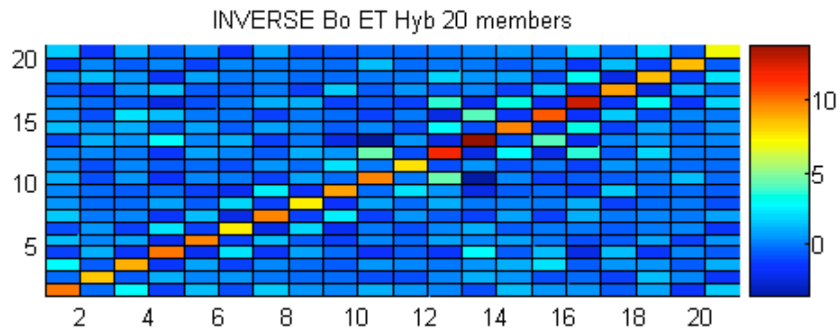
IMPERFECT NUMERICAL MODELS

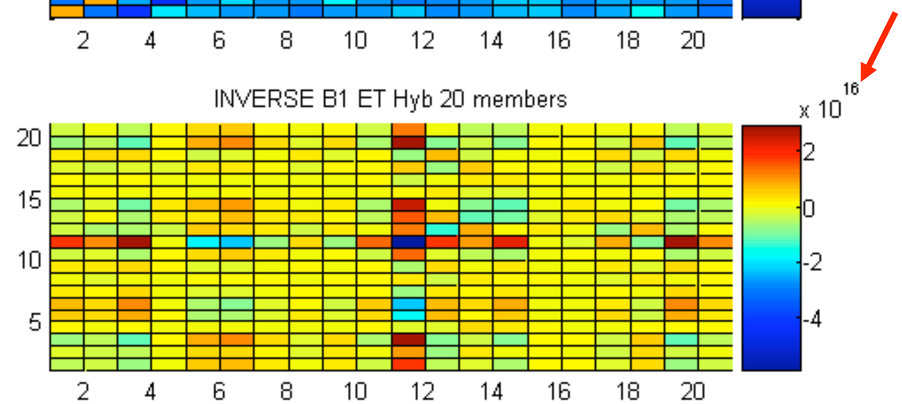
- Inconsistency between real & model systems
  - Transitional behavior if model started with real initial state
    - "Mapping paradigm" for reducing noise related to model drift
      - Physica D paper – Toth & Pena 2007

Hybrid ET



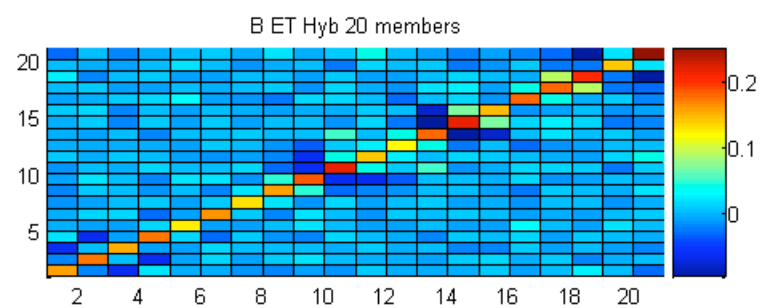




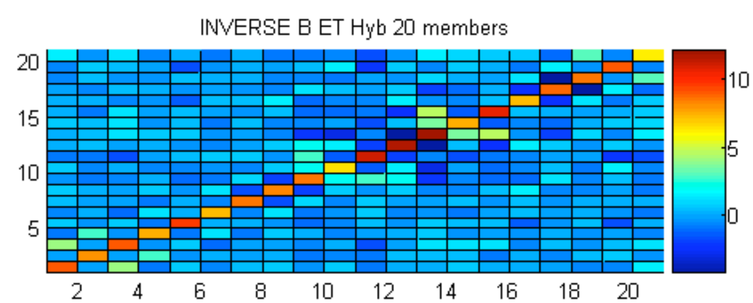


Bo From 3DVar

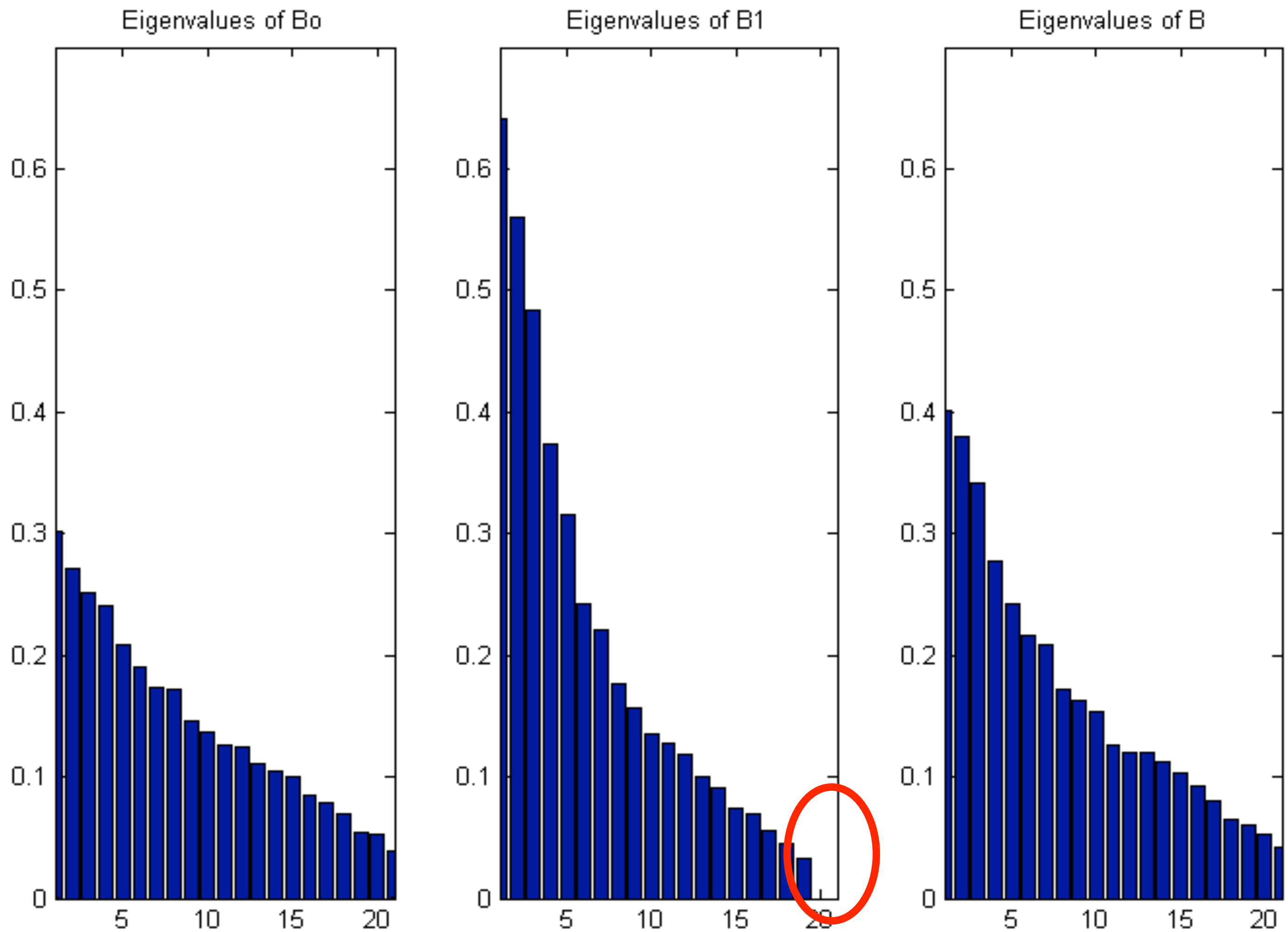
B1 from ET



B = (Bo+ B1)/2



Singular Value Decomposition of Bo, B1 and B:



Last 2 eigenvalues of B1 are zero! B1 is ill-conditioned

Is the hybrid approach a regularization strategy?

Remarks

- Ensemble DA may be able to enhance the covariance B by including "errors" of the day
- Ensembles may not be able to sample completely the uncertainty distribution of the background field. This is seen in the elements of B that are far away from the diagonal, which in principle should be small.
- Hybrid and sprinkling noise are two approaches that enhance B. In the latter the noise may get in cycled in the dynamics during the DA procedure, therefore may not be optimal. In the former, adding Bo appears to be a regularization approach.
- We introduced two approaches. One consists in adding noise to an ensemble only to enhance B. The second is a regularization procedure to make B well-posed.
- Local ETKF is also a promising approach since it populates the diagonal band of B, see diagram below for the case that the local vector includes 3 gridpoints:

$$\begin{array}{c} L \\ \mathbf{B} = \begin{pmatrix} \mathbf{b}_{11} & & & & \\ & \mathbf{b}_{22} & \mathbf{b}_{23} & \mathbf{b}_{24} & \\ & \mathbf{b}_{32} & \mathbf{b}_{33} & \mathbf{b}_{34} & \\ & \mathbf{b}_{42} & \mathbf{b}_{43} & \mathbf{b}_{44} & \\ & & & & \mathbf{b}_{55} \end{pmatrix} \end{array}$$

VARIANTS OF ENSEMBLE-BASED DA

- Perturbed observations
  - Represent all sources of forecast error at its source
    - Add noise to data representing observational uncertainty
      - Large amount of noise needed to avoid filter divergence
      - Cycling of noise makes state & error covariance estimates noisy
        - » Houtekamer, Anderson, etc, late 1990s
- ETKF & related methods
  - Reduce noise by eliminating perturbed observations
  - Covariance inflation needed to avoid filter divergence
    - Add noise to initial ensemble perturbations =>
      - Noisy covariance estimate
  - Cycle noisy covariance estimate
    - Negative effect on analysis state?
      - Anderson, Bishop, Szunyogh, Whitaker, etc, 2000+