University of Reading, July 2, 2013

Ensemble smoothing under the influence of nonlinearity

Lars Nerger

Alfred Wegener Institute for Polar and Marine Research
Bremerhaven, Germany

Svenja Schulte and Angelika Bunse-Gerstner

University of Bremen, Germany

Outline

- Ensemble smoothers
- Influence of nonlinearity
- Influence of localization
- Smoothing in a real model

Ensemble Smoothers

Smoothers

Filters (e.g. Ensemble Kalman filter)

- Estimate using observations until analysis time

Smoothers perform retrospective analysis

- Use future observations for estimation in the past

- Example applications:
  - Reanalysis
  - Parameter estimation

Ensemble smoothing

- Smoothing is very simple (ensemble matrix $\mathbf{X}^f_{k|k-1}$)
  (see e.g. Evensen, 2003)

  Filter: $\mathbf{X}^a_{k|k} = \mathbf{X}^f_{k|k-1}\mathbf{C}_k$

  Smoother: $\mathbf{X}^a_{k-1|k} = \mathbf{X}^a_{k-1|k-1}\mathbf{C}_k$

- Ensemble smoothing is cheap
  - e.g. E. Kalnay: “no-cost smoother”
  - weight matrix already computed in filter
  - just recombine previous ensemble states
    (actually the most costly part of the filter)
  - but: smoothing is recursive – application of each $\mathbf{C}_k$
    for all previous times within lag

Smoother with linear model

Smoother is optimal for linear systems:

➜ Forecast of smoothed state = filter analysis at later time

$$\mathbf{X}^a_{i|k} = \mathbf{X}^a_{i,i} \prod_{j=i+1}^{k} \mathbf{C}_j \qquad \mathbf{X}^a_{k|k} = \mathbf{M}_{k,i}\mathbf{X}^a_{i,i} \prod_{j=i+1}^{k} \mathbf{C}_j$$

➜ Based on ensemble cross-correlation between two time instances

$$\overline{\mathbf{x}^a_{i|k}} = \overline{\mathbf{x}^a_{i|k-1}} + \mathbf{X}'^a_{i|k-1} \left(\mathbf{X}'^a_{k|k}\right)^T \mathbf{E}$$

➜ Each additional lag reduces error
(if covariances are correctly estimated, Cohn et al. 1994)

(Ensemble perturbation matrix $\mathbf{X}' = \mathbf{X} - \bar{\mathbf{X}}$)

Smoother and Nonlinearity

Smoother and nonlinearity

- Optimality doesn't hold with nonlinear systems!

$$\overline{\mathbf{x}^a_{i|i+1}} = \overline{\mathbf{x}^a_{i|i}} + \mathbf{X}'^a_{i|i} \left(\mathbf{X}'^a_{i+1|i+1}\right)^T \tilde{\mathbf{E}}$$

influenced by nonlinear model

➜ What is the effect of the nonlinearity?

➜ Do ensembles just decorrelate?
(mentioned e.g. by Cosme et al. 2010)

➜ Consider smoother performance relative to filter
(Smoother reduces estimation error from the filter)

Numerical study with Lorenz-96

- Cheap and small model (state dimension 40)
- Local and global filters possible
- Nonlinearity controlled by forcing parameter *F*
  - Up to *F*=4: periodic waves; perturbations damped
  - *F*>4: non-periodic
- Nonlinearity of assimilation also influenced by forecast length

- Experiments over 20,000 time steps
- Use smoother with ESTKF (Nerger et al., 2012)
- Tune covariance inflation for minimal RMS errors
- Implemented in open source assimilation software PDAF (*http://pdaf.awi.de*)

The ESTKF: First compare ETKF and SEIK

Square root of covariance matrix (ensemble size *N*, state dim *n*)

$$\mathbf{Z} = \mathbf{X}^f \mathbf{T} \qquad \mathbf{P}^f = \mathbf{Z}\mathbf{Z}^T$$

Transformation matrix in ensemble space (small matrix)

$$\mathbf{A} = \left(\mathbf{G} + (\mathbf{HZ})^T \mathbf{R}^{-1} \mathbf{HZ}\right)^{-1}$$

Analysis state covariance matrix

$$\mathbf{P}^a = \mathbf{Z}\mathbf{A}\mathbf{Z}^T$$

Ensemble transformation based on square root of $\mathbf{A}$

$$\mathbf{X}^a \sim \mathbf{Z}\mathbf{L} \qquad \mathbf{L}\mathbf{L}^T = \mathbf{A}$$

Very efficient:
Transformation matrix computed in space of dim. *N or N-1*

L. Nerger et al., Monthly Weather Review 140 (2012) 2335-2345

The T matrix

Matrix **T** projects onto the error space spanned by ensemble

SEIK and ETKF use different projections **T**

$$\mathbf{Z} = \mathbf{X}^f \mathbf{T}$$

For identical forecast ensembles both filters

- yield identical analysis state
- perform slightly different ensemble transformations
- also: SEIK is slightly faster than ETKF

- ETKF provides minimum transformation
  - desirable for least disturbing ensemble states

- How to get minimum transformation into SEIK?

Error Subspace Transform Kalman Filter (ESTKF)

Combine advantages of SEIK and ETKF

Redefine **T:**

1. Remove ensemble mean from all columns
2. Subtract fraction of last column from all others
3. Drop last column

Features of the ESTKF:

- Same ensemble transformation as ETKF
- Slightly cheaper computations
- Direct access to ensemble-spanned error space

L. Nerger et al., Monthly Weather Review 140 (2012) 2335-2345

T-matrix in SEIK and ESTKF

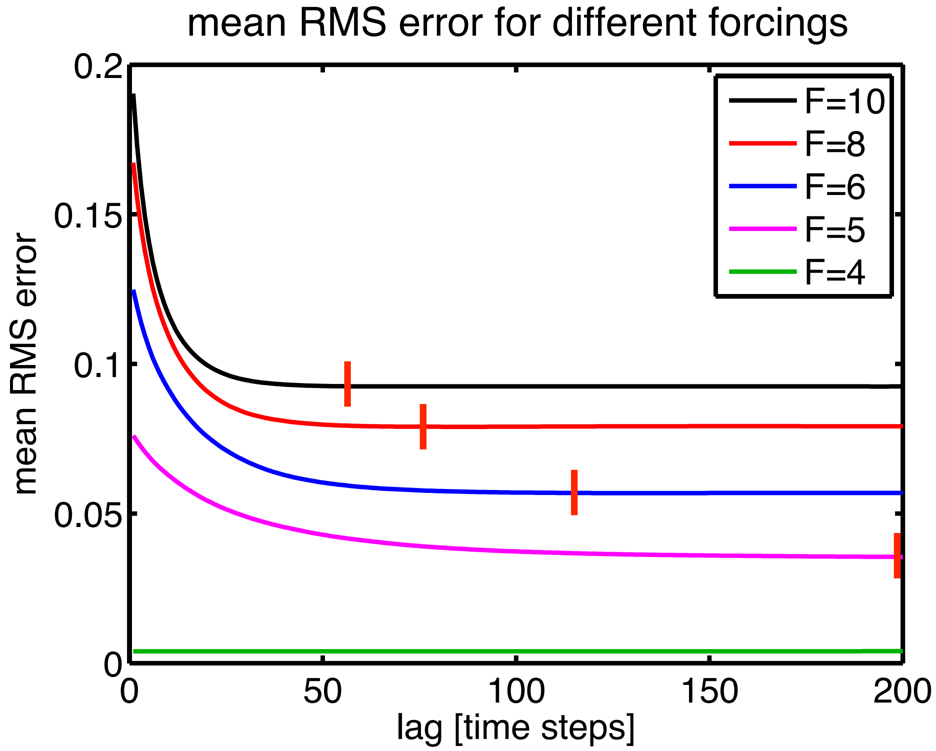
SEIK:
$$\mathbf{T}_{i,j} = \begin{cases} 1 - \frac{1}{N} & \text{for } i = j, i < N \\ -\frac{1}{N} & \text{for } i \neq j, i < N \\ -\frac{1}{N} & \text{for } i = N \end{cases}$$

ESTKF:
$$\hat{\mathbf{T}}_{i,j} = \begin{cases} 1 - \frac{1}{N}\frac{1}{\frac{1}{\sqrt{N}}+1} & \text{for } i = j, i < N \\ -\frac{1}{N}\frac{1}{\frac{1}{\sqrt{N}}+1} & \text{for } i \neq j, i < N \\ -\frac{1}{\sqrt{N}} & \text{for } i = N \end{cases}$$

- Efficient implementation as subtraction of means & last column
- ETKF: improve compute performance using a matrix **T**

Effect of forcing on the smoother – optimal lag

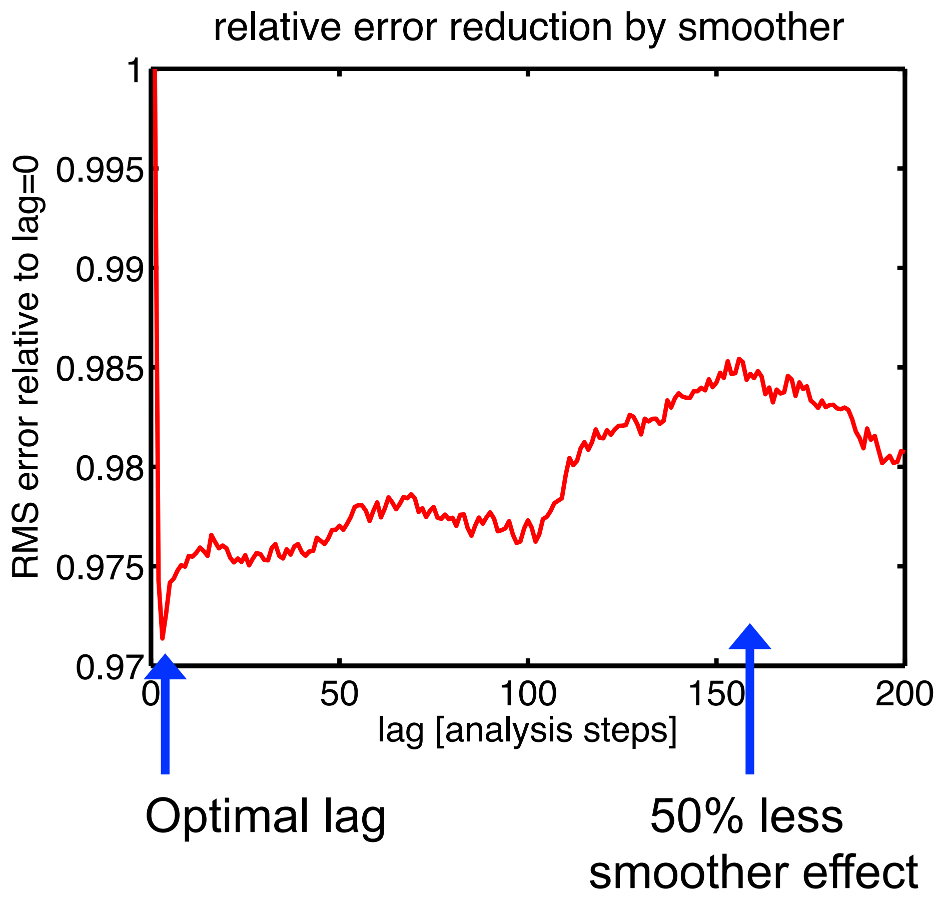
- Assimilate at each time step
- Ensemble size N=34
- Global ESTKF
- Inflation tuned for minimal RMS errors (account for inflation in smoother)



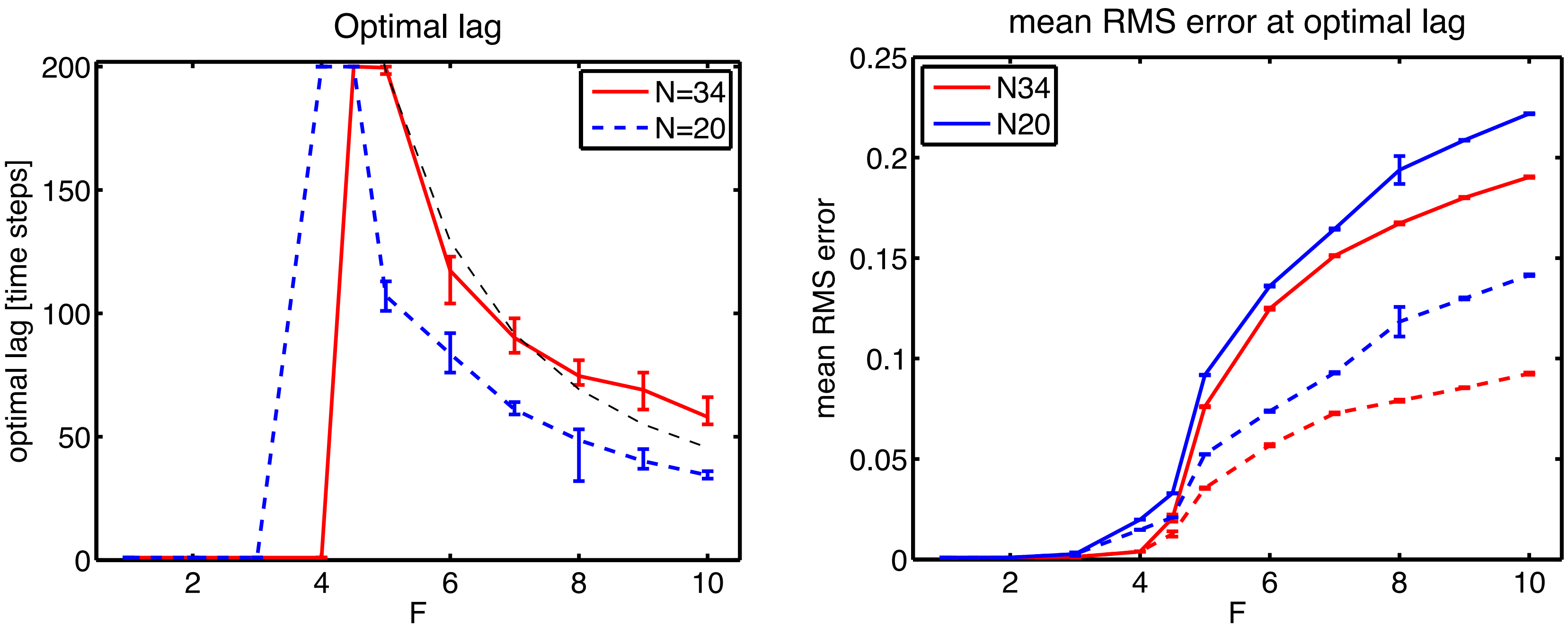
- Up to *F*=4
  - very small RMS errors
- *F*>4
  - Strong growth in RMS
  - Clear impact of smoother
  - Optimal lag: minimal RMS error (red lines)

Stronger nonlinearity

- *F*=7
- Forecast length: 9 steps
- Clear error minimum at lag=2 analysis steps
  ➜ the optimal lag
- Error increase beyond optimal lag (here 50%!)
  ➜ spurious correlations

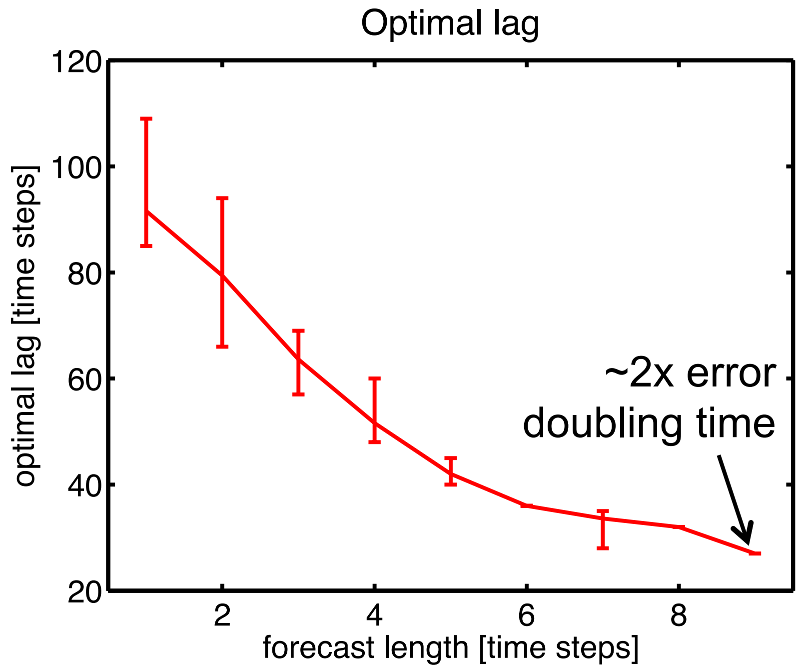


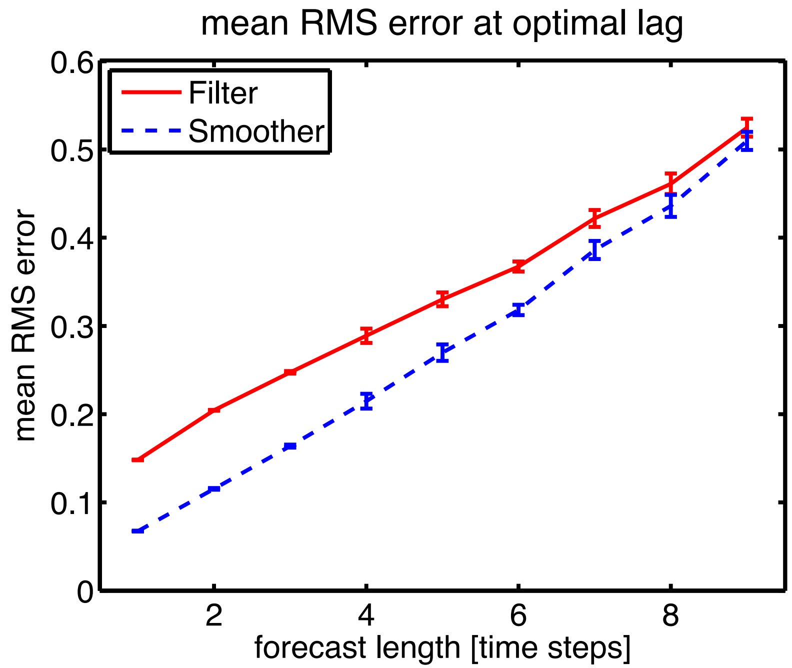
Impact of smoothing



- Optimal lag (minimal RMS error)
  - Behavior similar to error-doubling time
- RMS error at optimal lag
  - Smoother reduces error by 50% for all F>4
- Effect of sampling errors visible with smaller ensemble

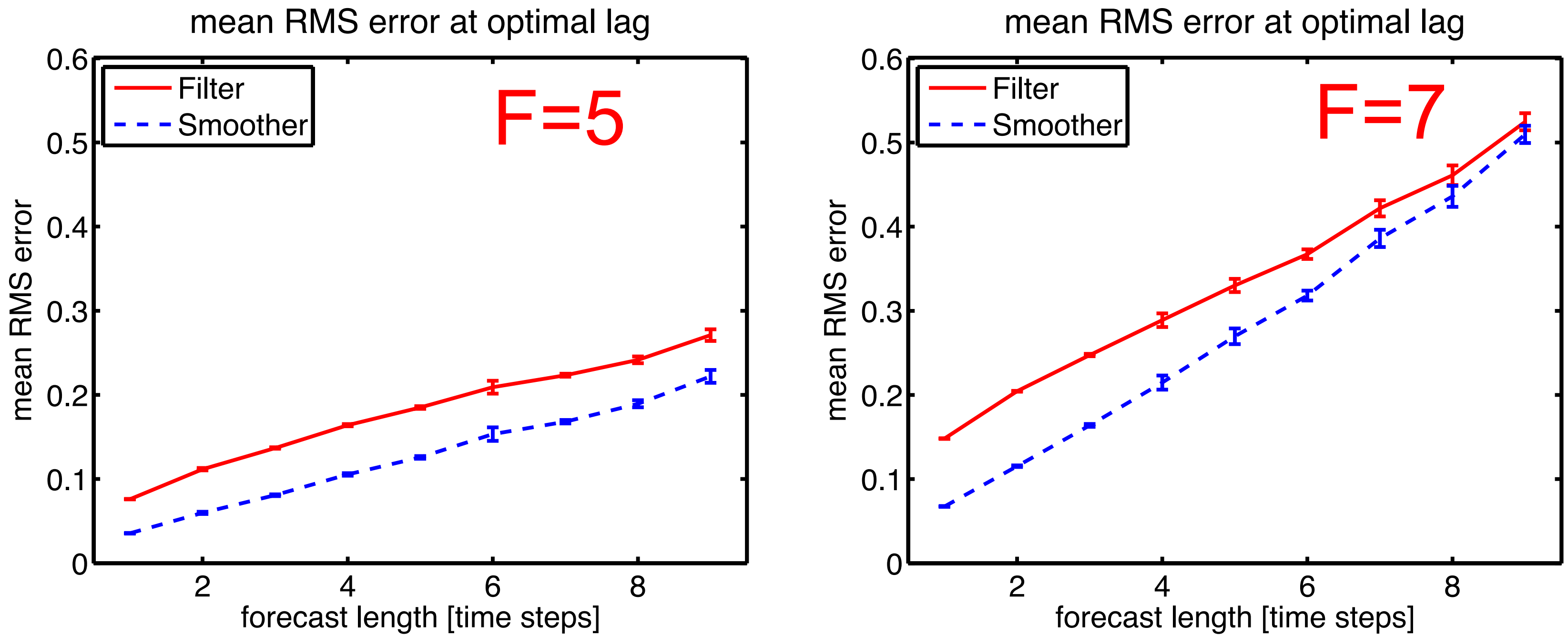
Vary forecast length (F=7)





- Forecast length = time steps over which nonlinearity acts on ensemble
- Longer forecasts:
  - ➜ Optimal lag shrinks
  - ➜ RMS errors grow for filter and smoother
  - ➜ Improvement by smoother shrinks (depends on forcing strength)

Vary forecast length – different forcing strength



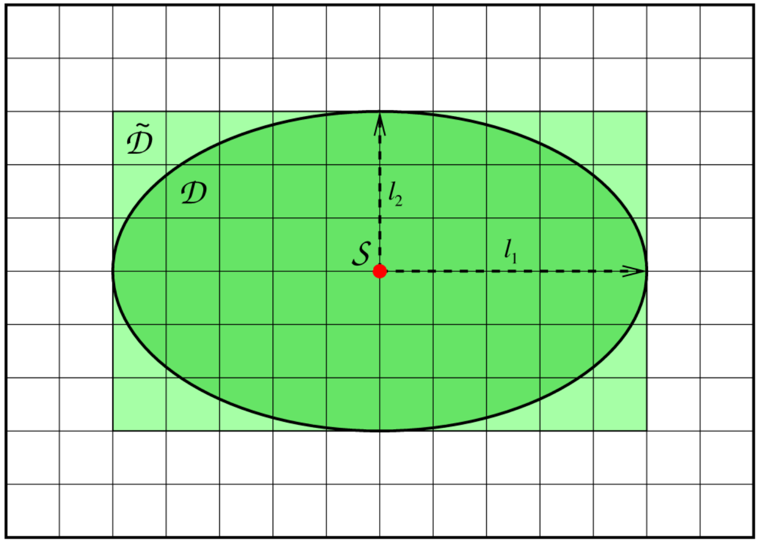
- Improvement by smoother depends on forcing strength
- Small forcing (F=5)
  - ➜ Approx. constant improvement by smoother
- Larger forcing (F=7)
  - ➜ Decreasing smoother effect

Impact of Localization

Domain & observation localization

Local Analysis:

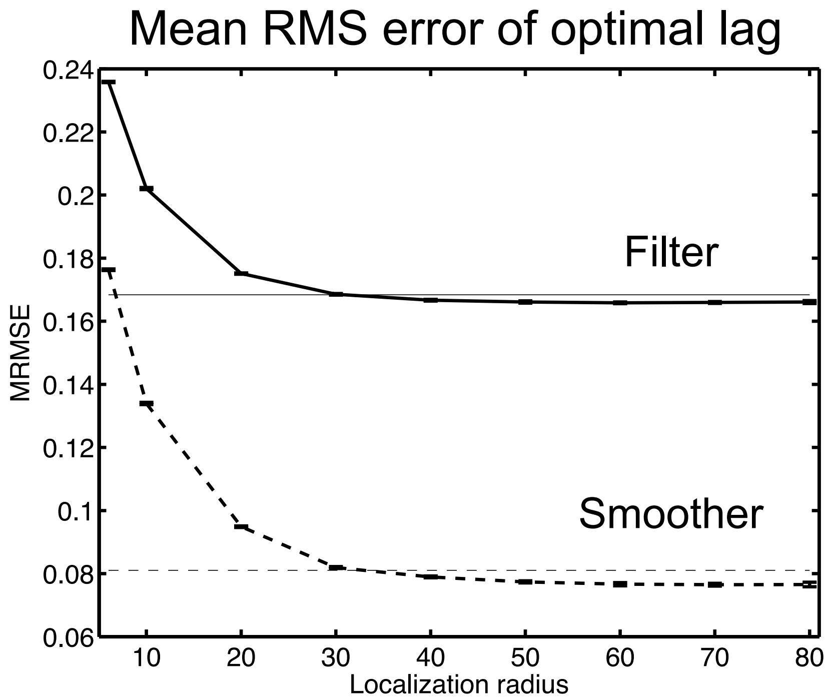
- Update small regions (like single vertical columns)
- Observation localizations: Observations weighted according to distance
- Consider only observations with weight >0
- State update and ensemble transformation fully local

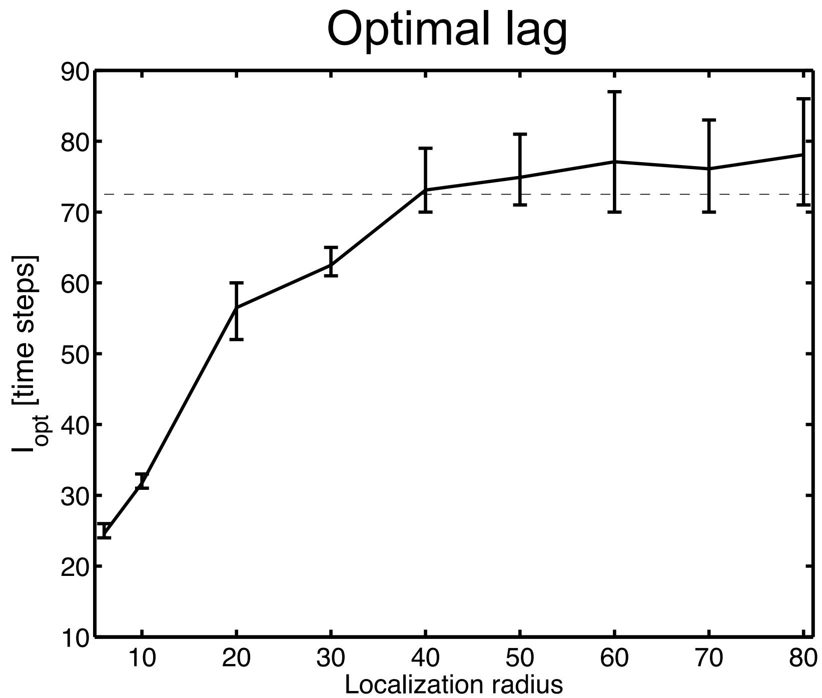


S: Analysis region
D: Corresponding data region

Similar to localization in LETKF (e.g. Hunt et al, 2007)

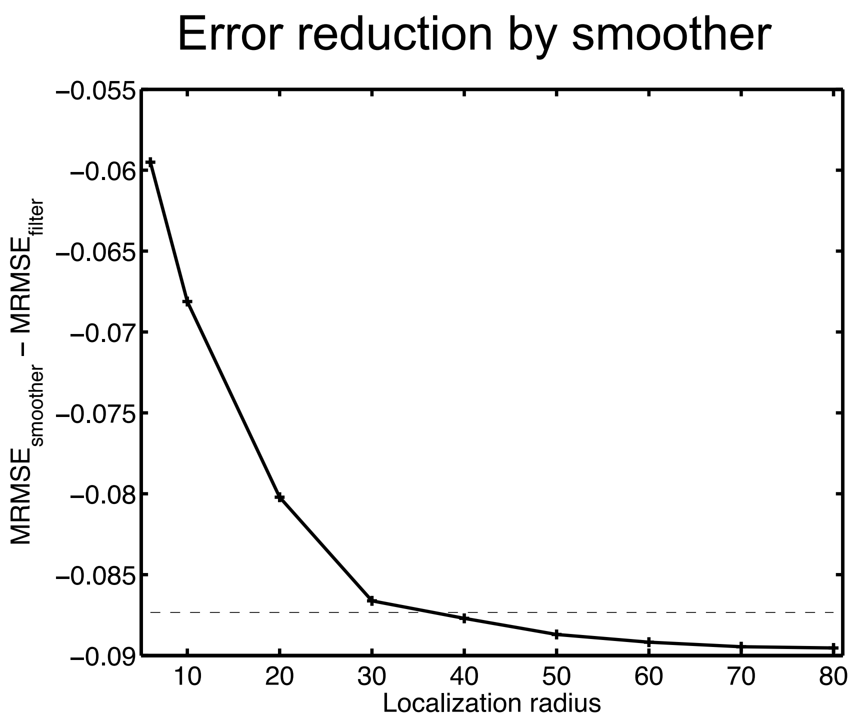
Influence of Localization on Smoothing





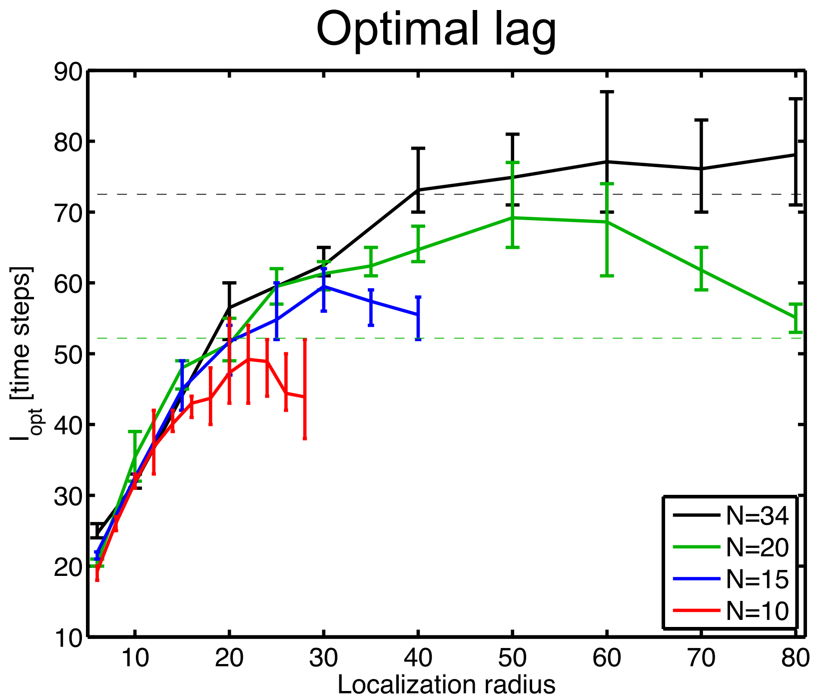
- Reduced RMS errors from filter and smoother by localization
  - localization is useful even for N=34
- Localization increases optimal lag
  - more observational information useable

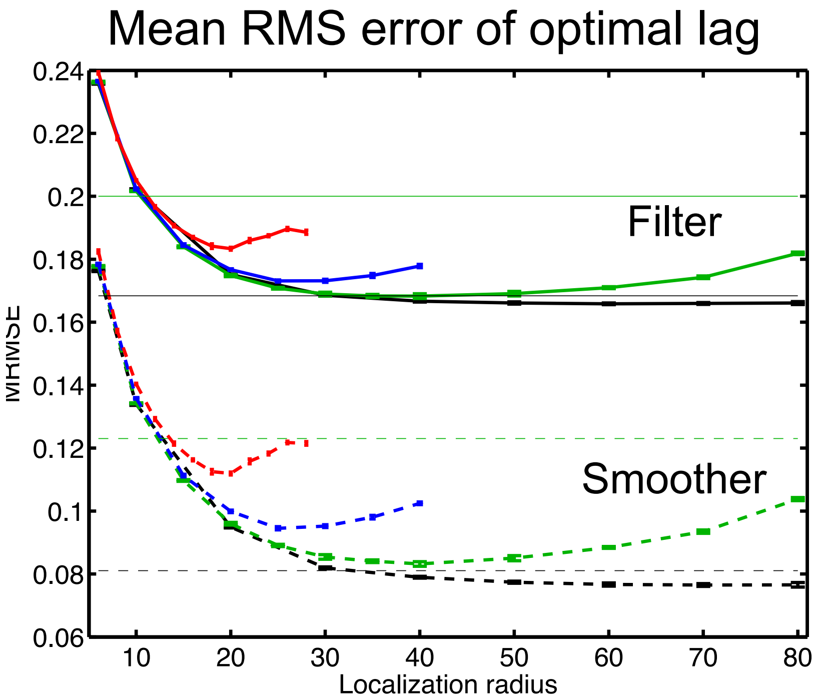
Influence of Localization on Smoothing (2)



- Use filter error as baseline
- Smoother results in additional reduction
- Smoother is more efficient with localization than for global filter

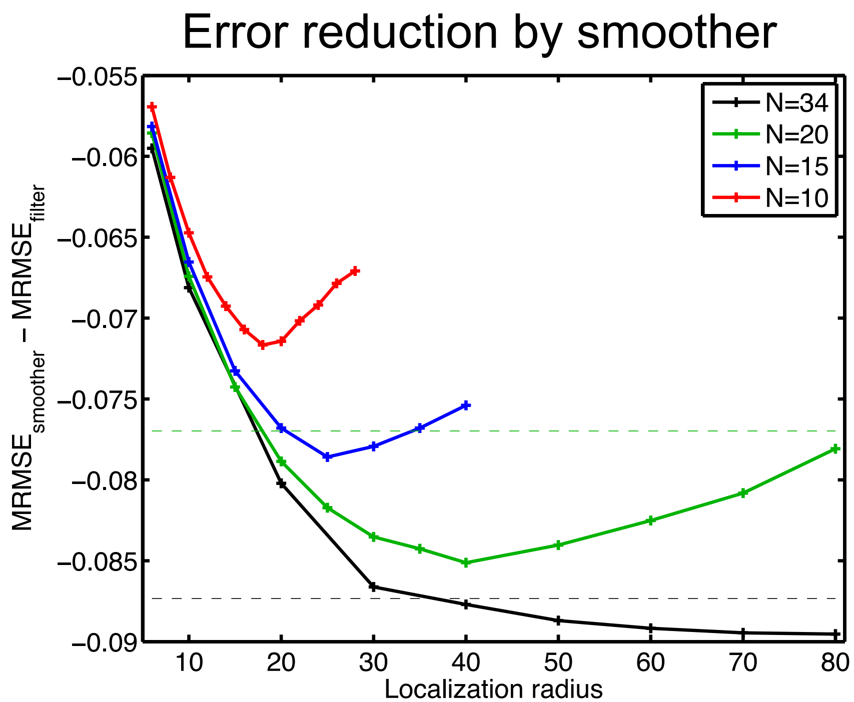
Smoothing with localization – smaller ensembles





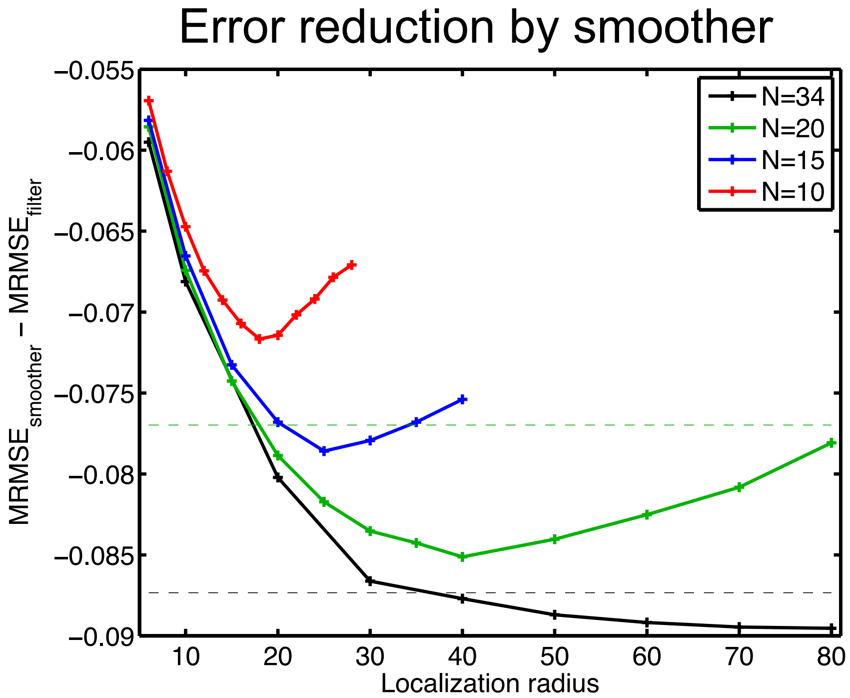
- Larger effect of localization with smaller ensembles
- Optimal lag shrinks (impact of sampling errors)
- Localization radius for maximum optimal lag slightly larger than for minimum RMS error

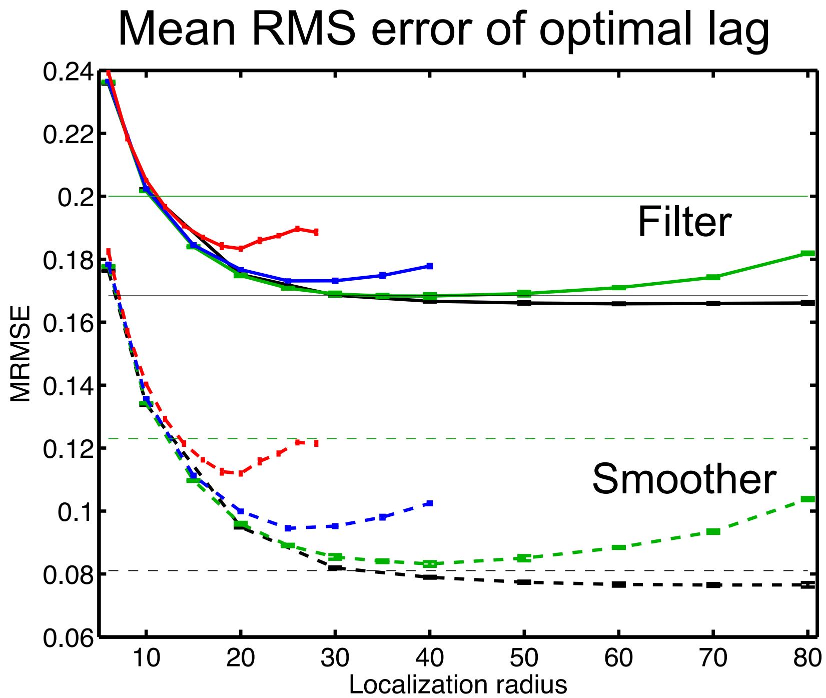
Smoother error reduction – smaller ensembles



- Smoother impact grows with ensemble size
  - Effect of sampling errors
- RMS error from smoother decreases faster than from filter
  - Amplification effect (multiple use of matrix **C**)

Optimal localization radius





Same localization radius for

- ➢ minimum filter RMS error
- ➢ largest smoother impact

➜ No re-tuning of localization radius for optimal smoothing!

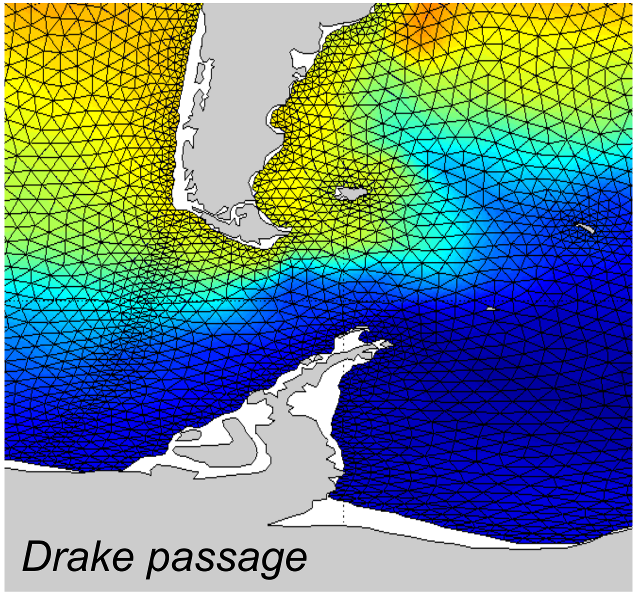
Smoothing in a Real Model

Global ocean model

FESOM (Finite Element Sea-ice Ocean model, Danilov et al. 2004)

Global configuration

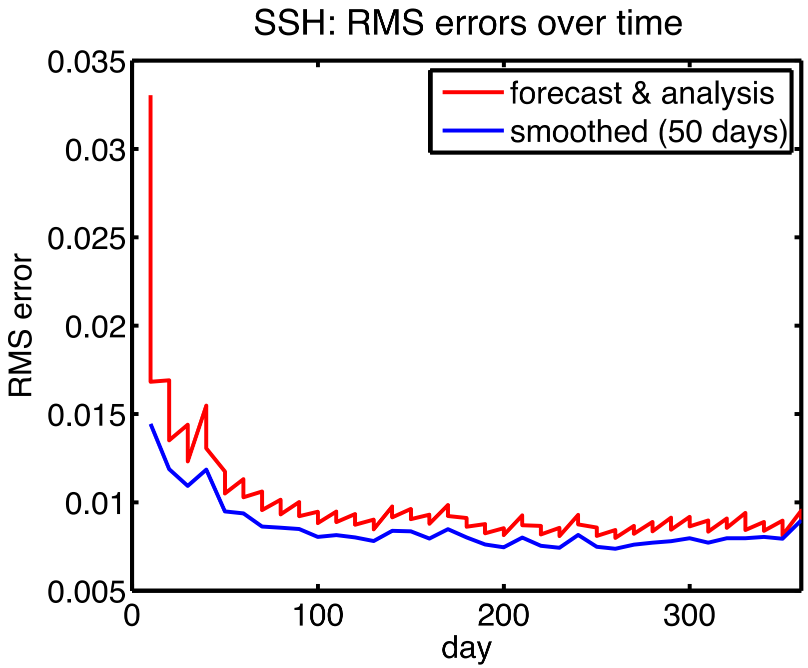
- 1.3° resolution, 40 levels
- horizontal refinement at equator
- state vector size 10^7
- weak nonlinearity (not easy to change)

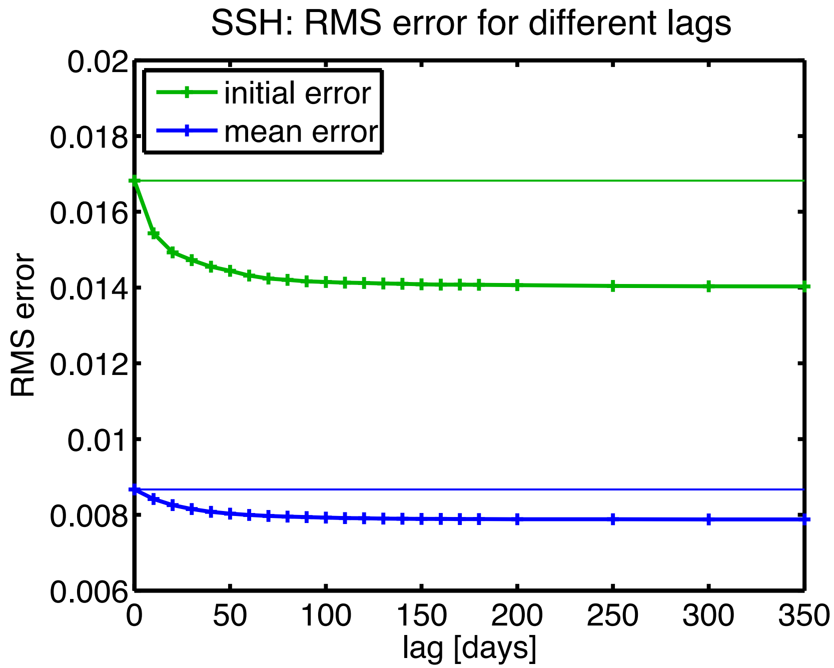


Twin experiments with sea surface height data

- ensemble size 32
- assimilate each 10th day over 1 year
- ESTKF with smoother extension and localization (using PDAF environment as single program)
- inflation tuned for optimal performance (ρ=0.9)
- run using 2048 processor cores (Timings: forecasts 8800s, filter+smoother 200s)

Effect of smoothing on global model





Typical behavior

- RMSe reduced by smoother

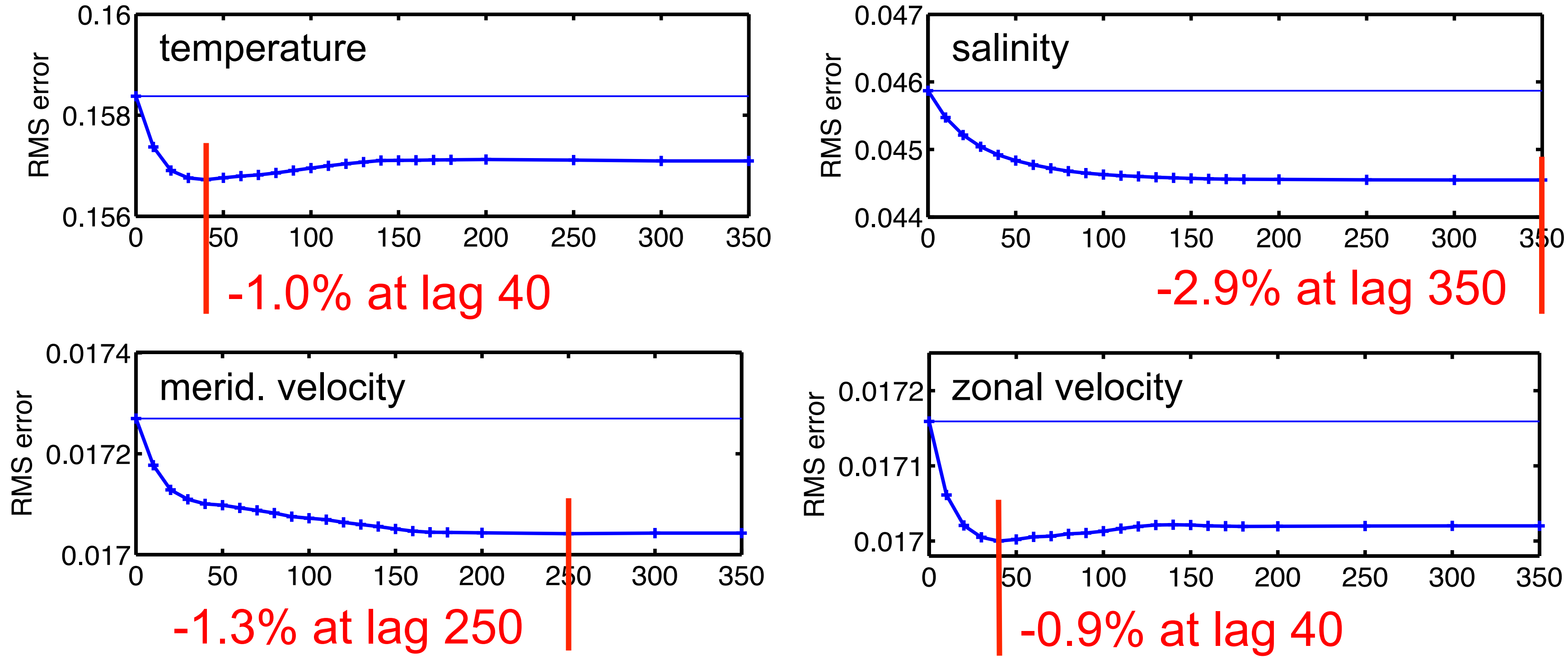
Error reductions:

~15% at initial time

~8% over the year

- Large impact of each lag up to 60 days
- Further reduction over full experiment (optimal lag = 350 days)

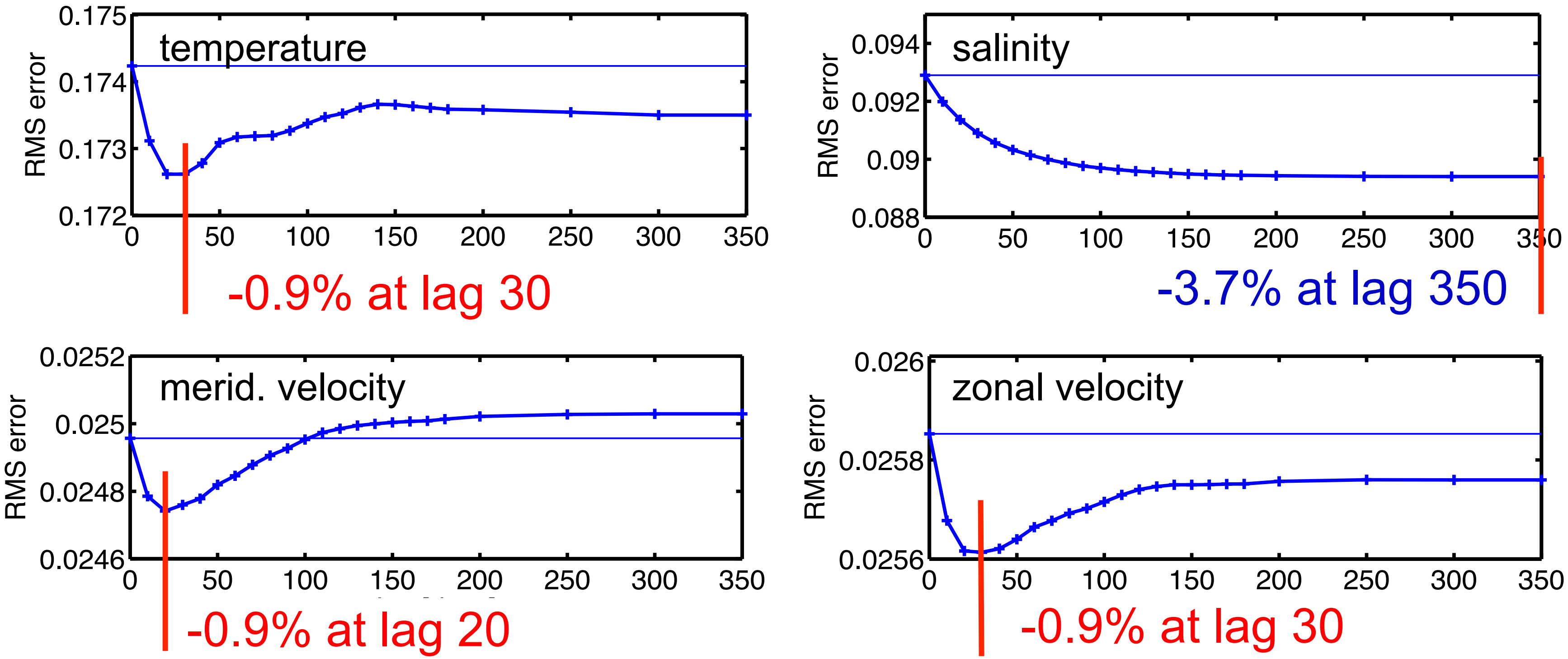
Multivariate effect of smoothing – 3D fields



3D fields:

- Multivariate impact smaller & specific for each field
- Optimal lag specific for field
- Optimal lag smaller than for SSH (e.g. temperature directly influenced by atmospheric forcing, Brusdal et al. 2003)

Multivariate effect of smoothing – surface fields



Ocean surface:

- Relative smoother impact not larger than for full 3D
- Deterioration for meridional velocity at long lags

➜ What is the optimal lag for multivariate assimilation?

Conclusion

- Multivariate assimilation:
  - ➜ Lag specific for field
  - ➜ Choose overall optimal lag or separate lags
  - ➜ Best filter configuration also good for smoother
- Nonlinearity:
  - ➜ Introduces spurious correlations in smoother
  - ➜ Error increase beyond optimal lag
  - ➜ Optimal lag: few times error doubling time
- Localization:
  - ➜ Increases smoother impact
  - ➜ Increases optimal lag

Thank you!

Web-Resources

SANGOMA www.data-assimilation.net

PDAF Parallel Data Assimilation Framework pdaf.awi.de