Parameter estimation with EnKF in a QG model and subsequent application in TOPAZ

F. Counillon, L. Bertino, P. Sakov

Mohn-Sverdrup Center,

Nansen Environmental and Remote Sensing Center,

Bergen, Norway

NERSC

SANGOMA, Nov 26. 2012

Mohn-Sverdrup Center

Global Ocean Studies - Operational Oceanography

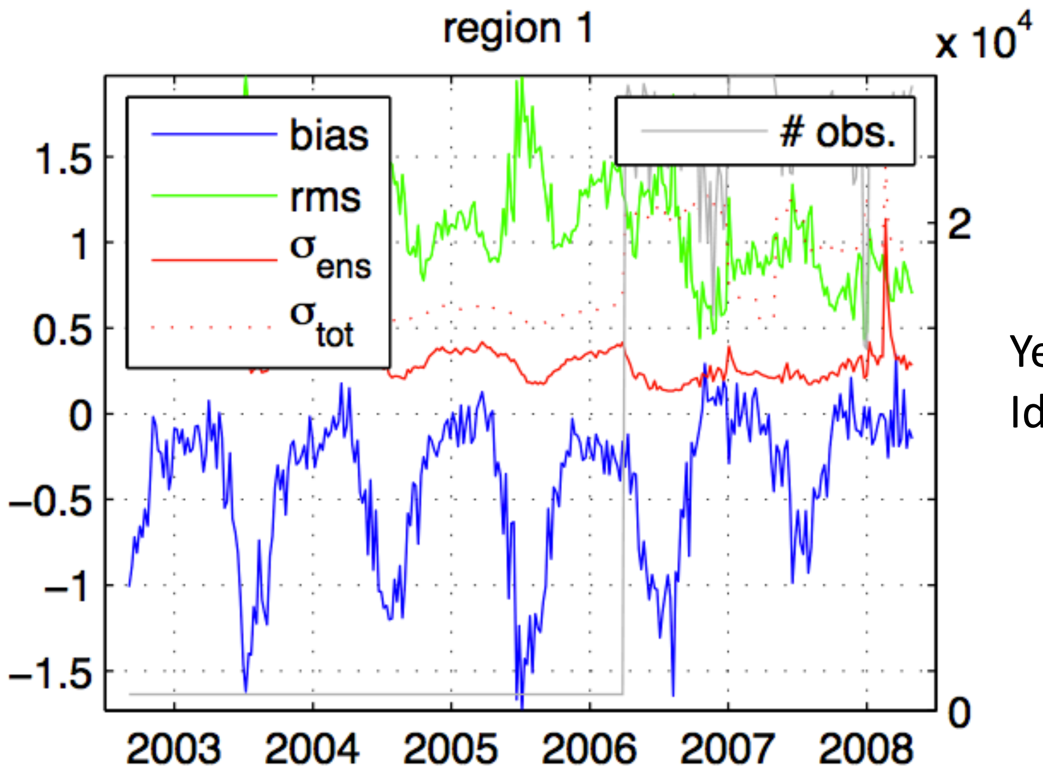
Motivation

- Bias in data assimilation are problematic :
  - Analysis state still contains part of the bias
  - Bias affects the error covariance matrix
  - Models are often attracted to their bias solution
    - Successive corrections may deteriorate model equilibrium
- Biases may be present both in observations and models:
  - **Observational bias** must be removed during assimilation
  - **Model bias** must be removed during assimilation and added to the prediction

- Biases can be estimated by extending the model state

What is the best approach and what happen if observational network covers partially the domain or vary with time?

Example for SST in the Gulf Stream area in TOPAZ pilot reanalysis

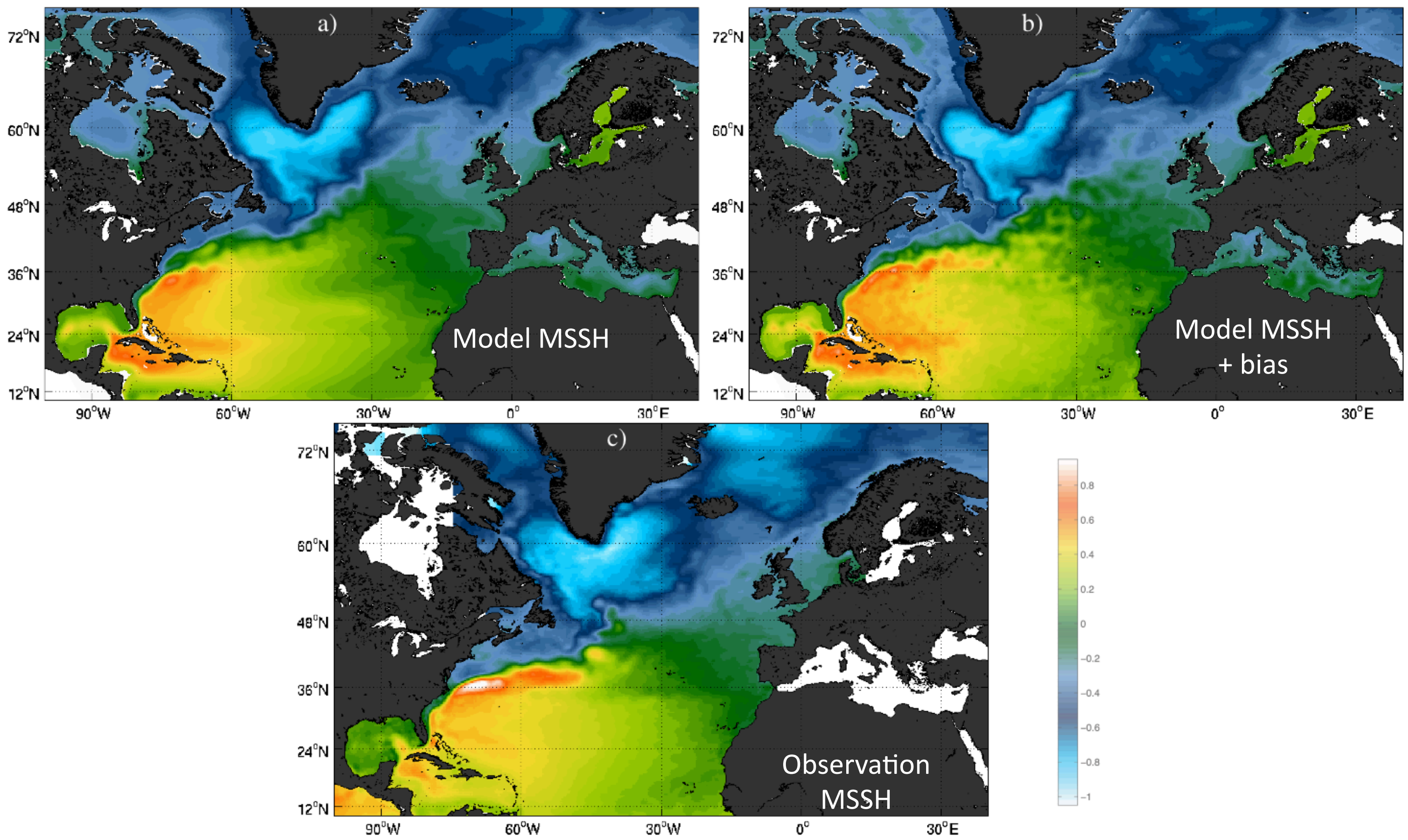


Yes we have bias and we should handle it
Ideally we should correct the model

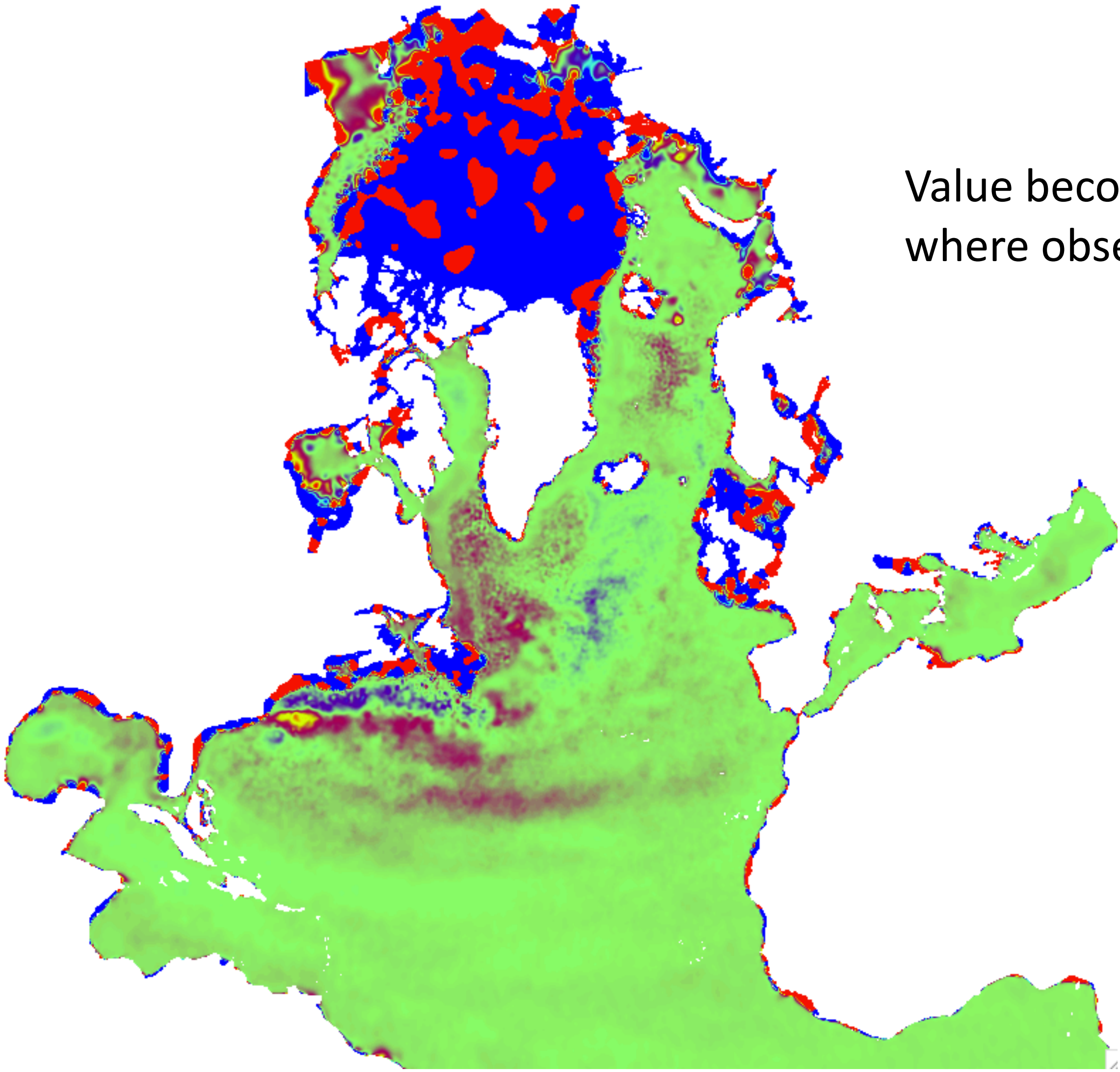
In TOPAZ system, bias of Sea Surface Temperature (SST) and Sea Surface Height (SSH) are estimated using **uniform ensemble inflation**

Example with SSH

- SSH is used to constrained the model dynamic
  - Observation based estimate contains large inaccuracy near the ice edge and near coastal area (land measurement and tidal signal pollution)
- During the TOPAZ pilot reanalysis, the bias of this quantity has been estimated (considered obs bias)



Lesson learned from reanalysis TOPAZ SSH

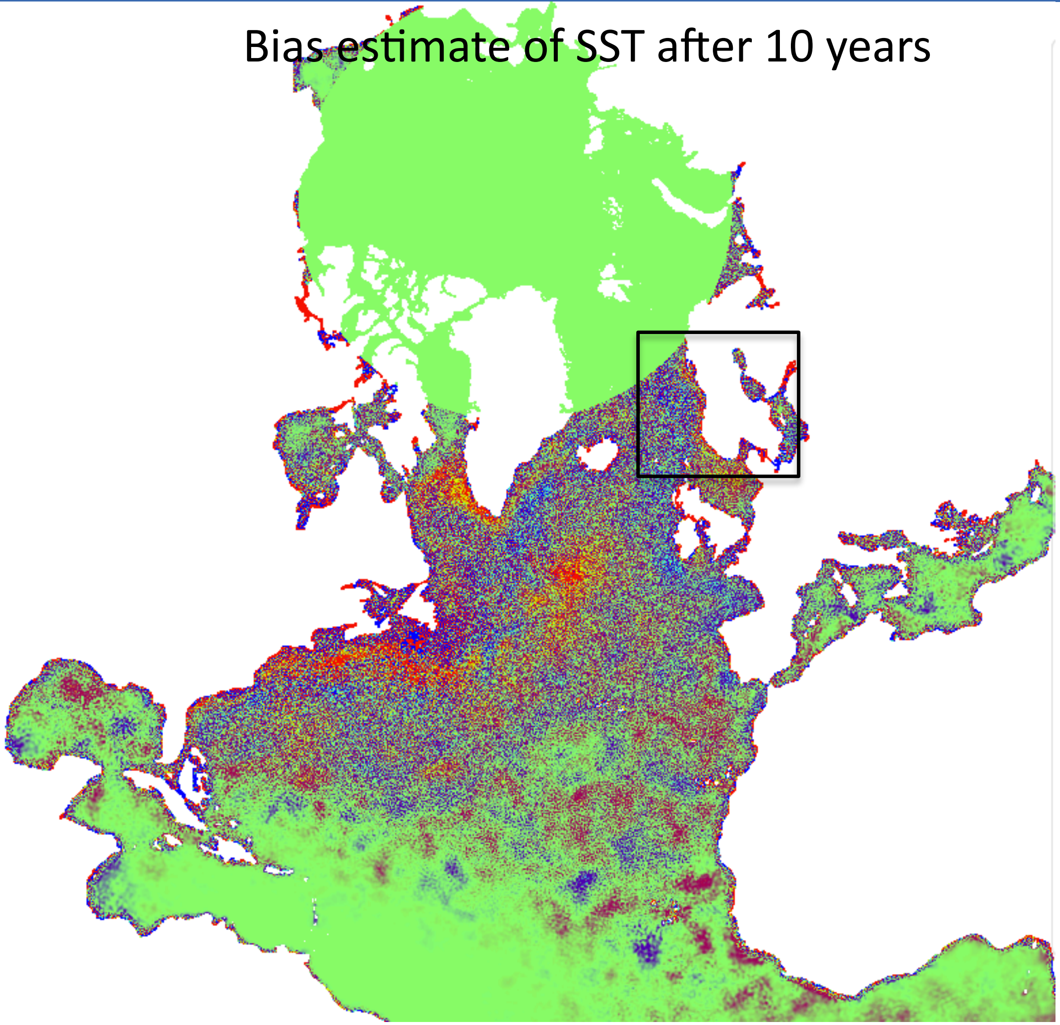


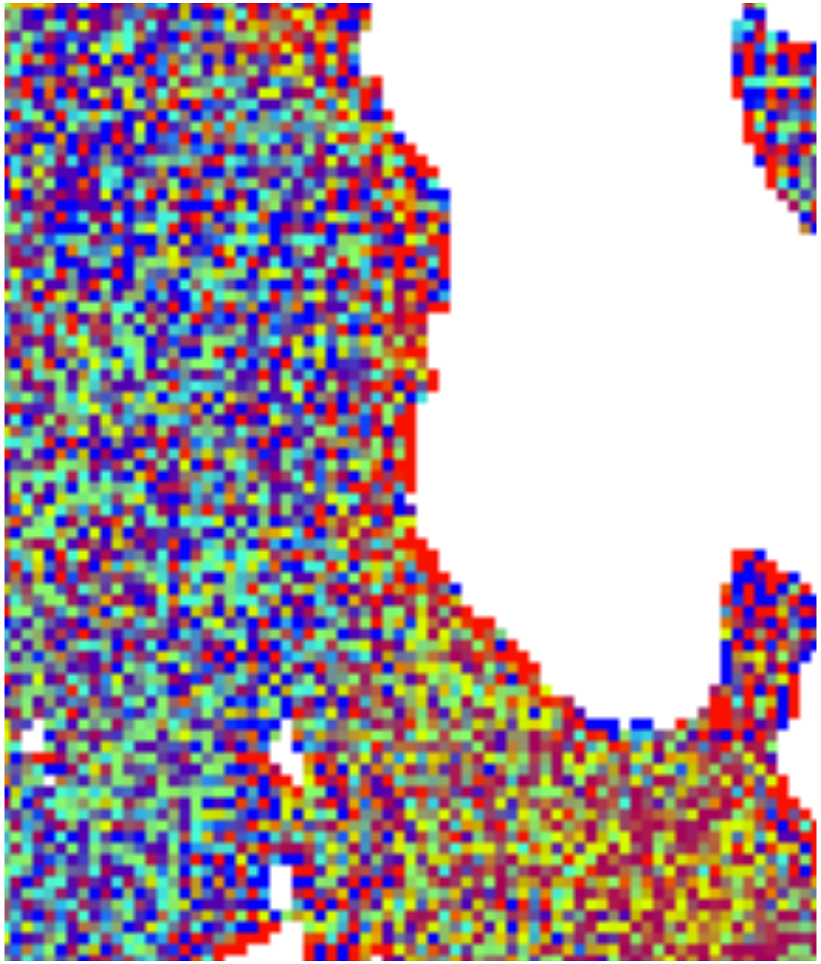
Value becomes completely unrealistic in places where observation are not used

Bias estimate of SSH after 10 years

Lesson learned from TOPAZ
SST

Bias estimate of SST after 10 years





- Unrealistic small scale structure
- Unrealistic feature near the coast
- Change of observation network leads to a model crash

Idealized study
QG model

How to estimate bias when observational network does not cover the domain uniformly. Which method, initialisation technique works best

Method:

- Inflation (*Andersen 2001*)
- Model error (*Evensen 2009 chap 12*), i.e. additive red noise

Observational network:

- Cover the whole domain
- Cover half of the domain

Initialisation:

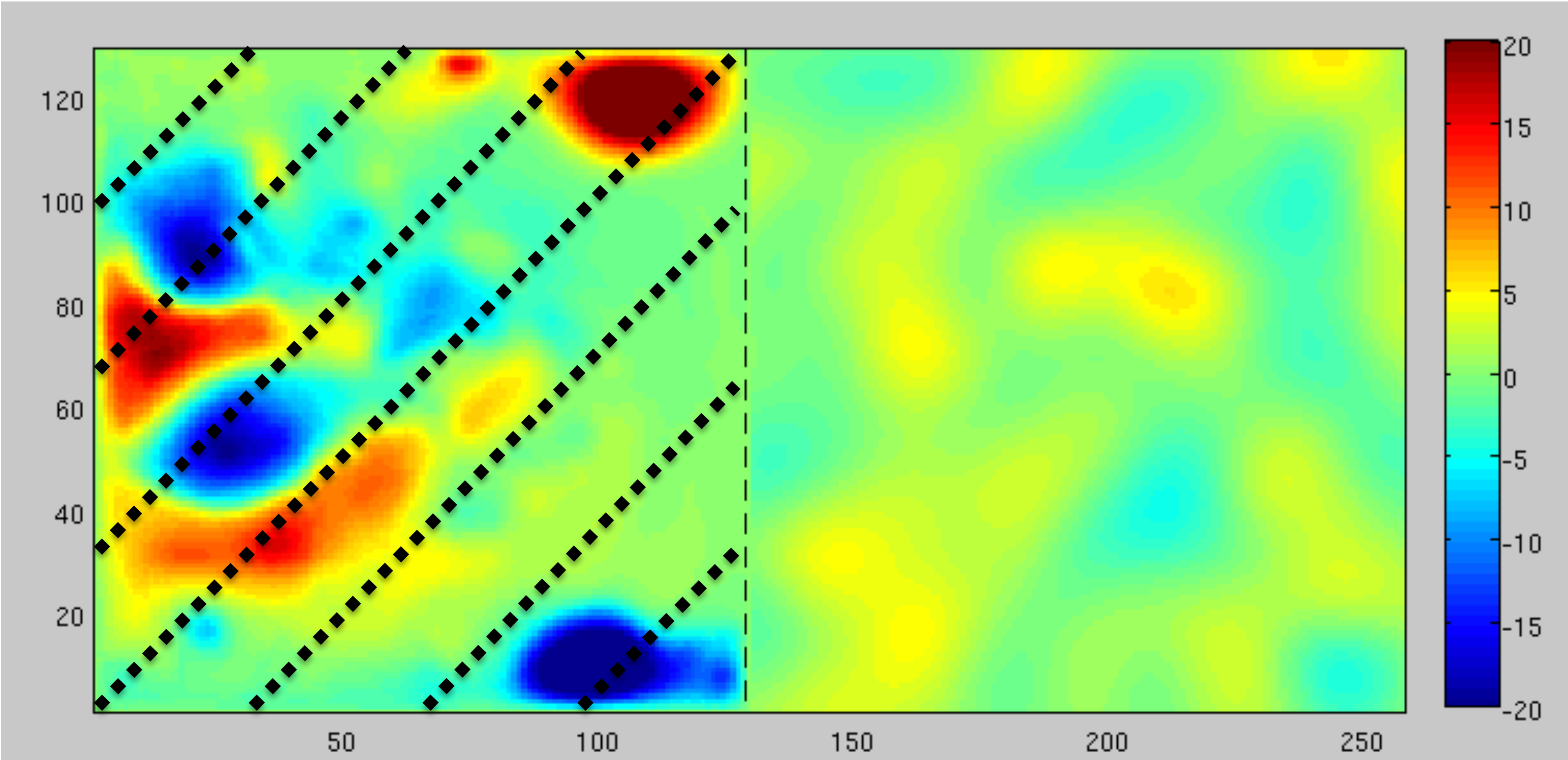
- Uniform for each member
- Red noise

Tools:

Quasi-geostrophic model available in the EnKF-matlab package

Part I:Uniform observational coverage
Bias constant in time

Model field **Bias=pseudo2D(3*randn,30) + 1**



Settings:

- •Ensemble size=25
- •Scheme DEnKF
- •Experiment run for 800 Δt
- •Inflation model field=1.1
- •300 observations (error var=4, 10 Δt)
- •Observation location is random
- •Loc radius 15 grid cell

Part I: Experiments

Want to analyze the performance of the method regarding the method and the initialisation

Exp 1:

•Uniform bias initialization

•Parameter inflation (from 1:0.025:1.2)

Exp 2:

•Red noise initialization [pseudo2D(3,30)]

•Parameter inflation (from 1:0.025:1.2)

Exp 3:

•Red noise initialization [pseudo2D(3,30)+rand(1)]

•Parameter inflation (from 1:0.025:1.2)

Exp4:

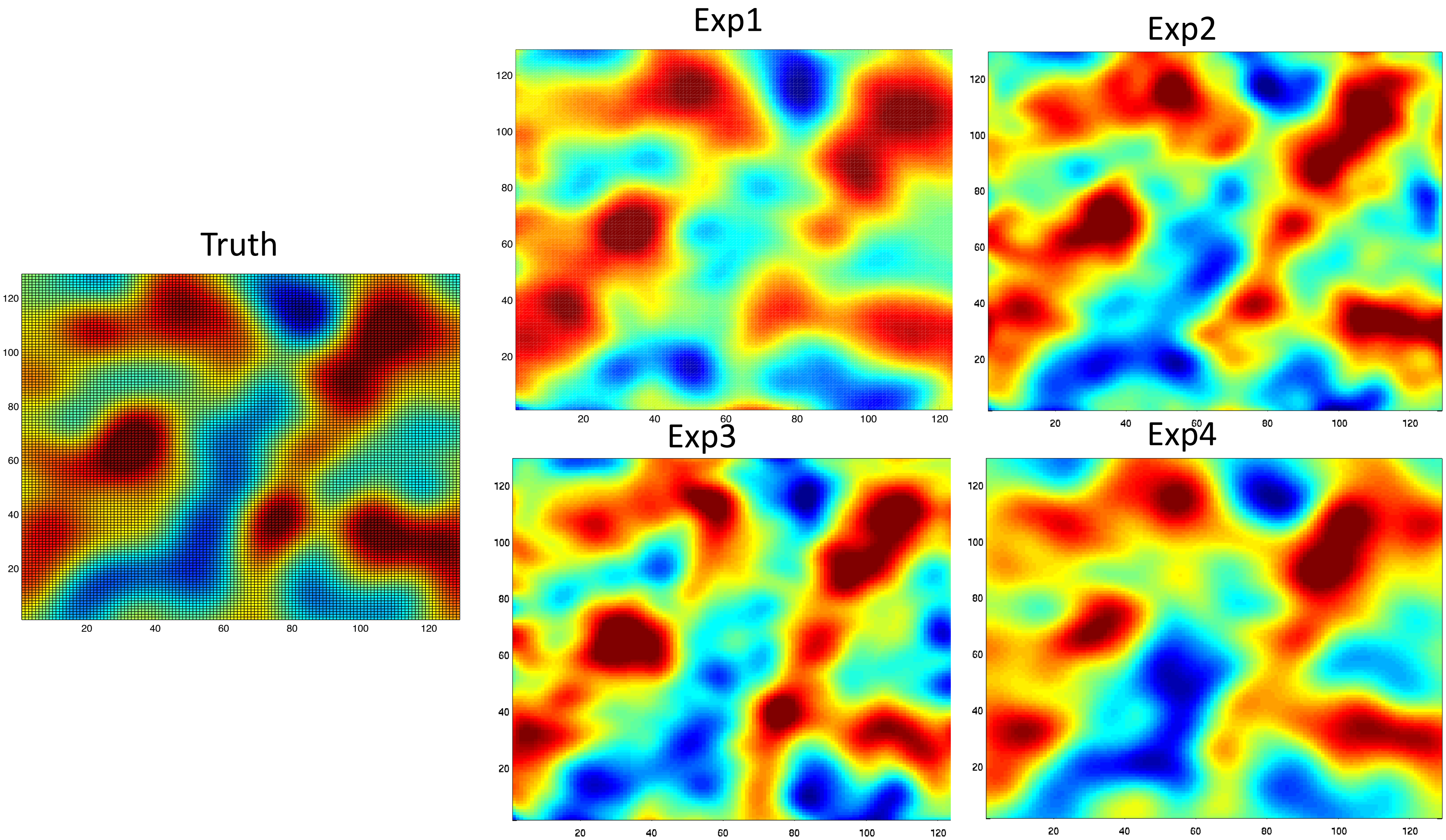
•Pseudo2D ensemble initialization [pseudo2D(3,30)]

•Additive red noise :

pseudo2D(α) with α=[0.1 0.15 0.2 0.3 0.4 0.5 0.6 0.8])

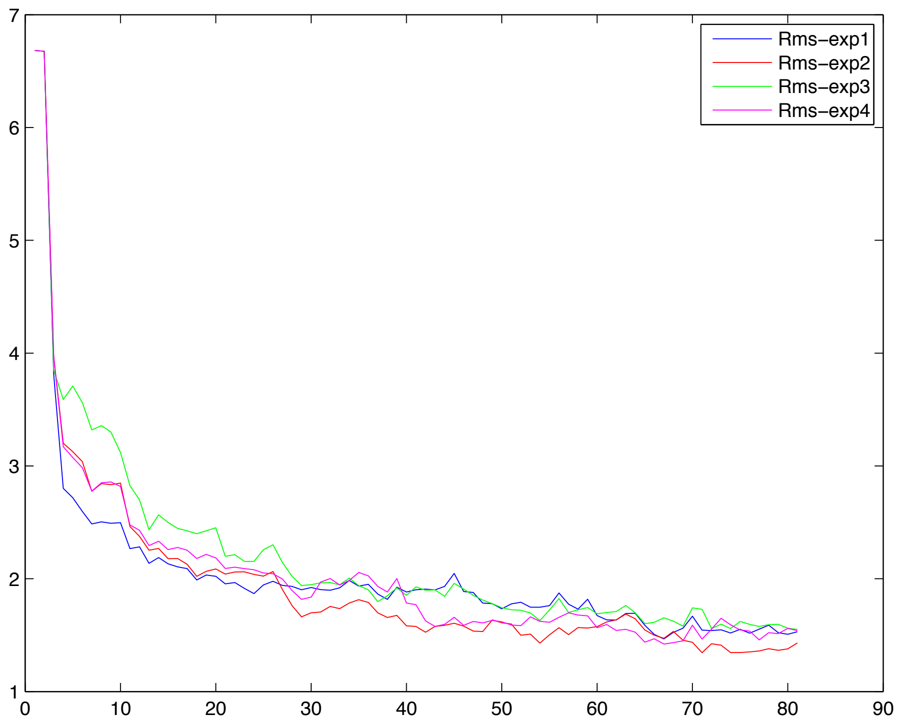
Optimal solution is the one that **minimize the RMSE of the parameter**

Part I: Qualitative comparison

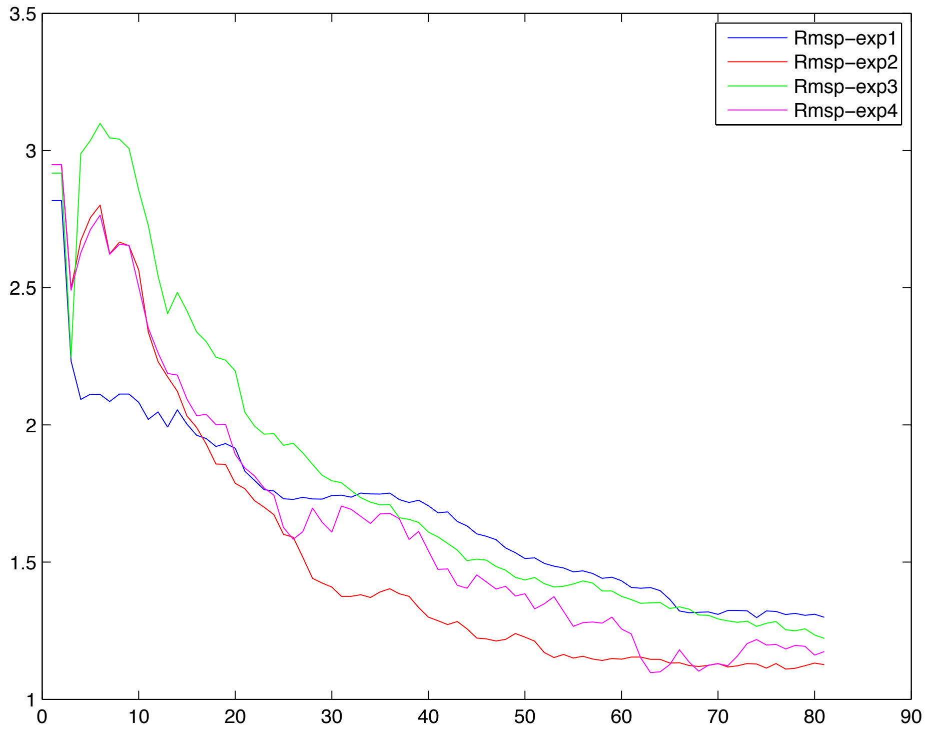


Part I: Quantitative comparison

Model state RMS



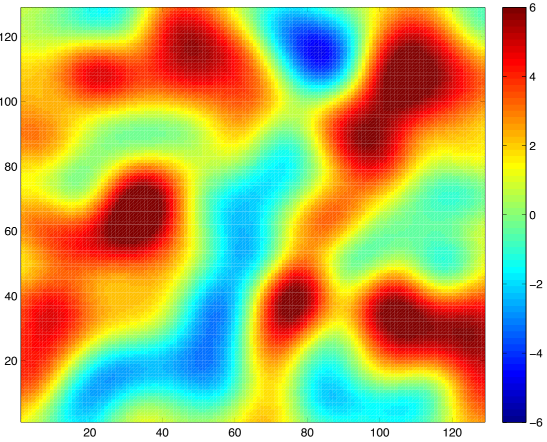
Param RMS



What happen if we continue beyond convergence ?

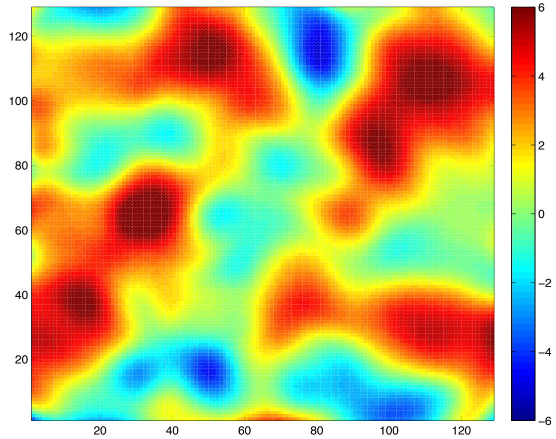
Long simulation (beyond convergence)

Truth



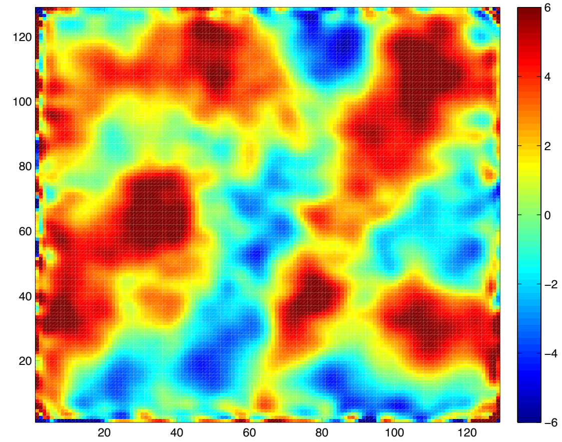
Exp1: Inflation

infl (800 tstep)



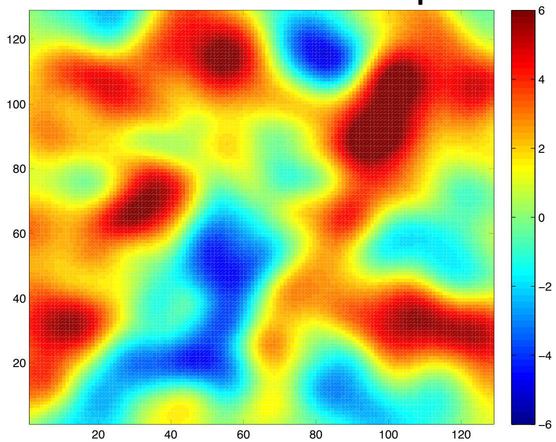
After 3875 time step

model crash

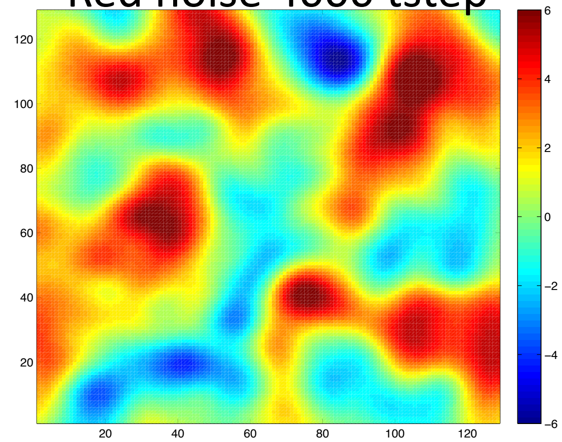


Exp2: Red noise

Red noise 800 tstep



Red noise 4000 tstep



Part I summary

Initialisation (uniform vs. red noise)

•Through localization, uniform initialization manages to catch small scale structures

•Red noise:

- initializing with smaller scale →divergence (Not shown)
- initializing with larger scale →convergence (Not shown)

Method (inflation vs. red noise)

•Both method successful initially until convergence

•When run beyond convergence inflation-like method seems unstable:

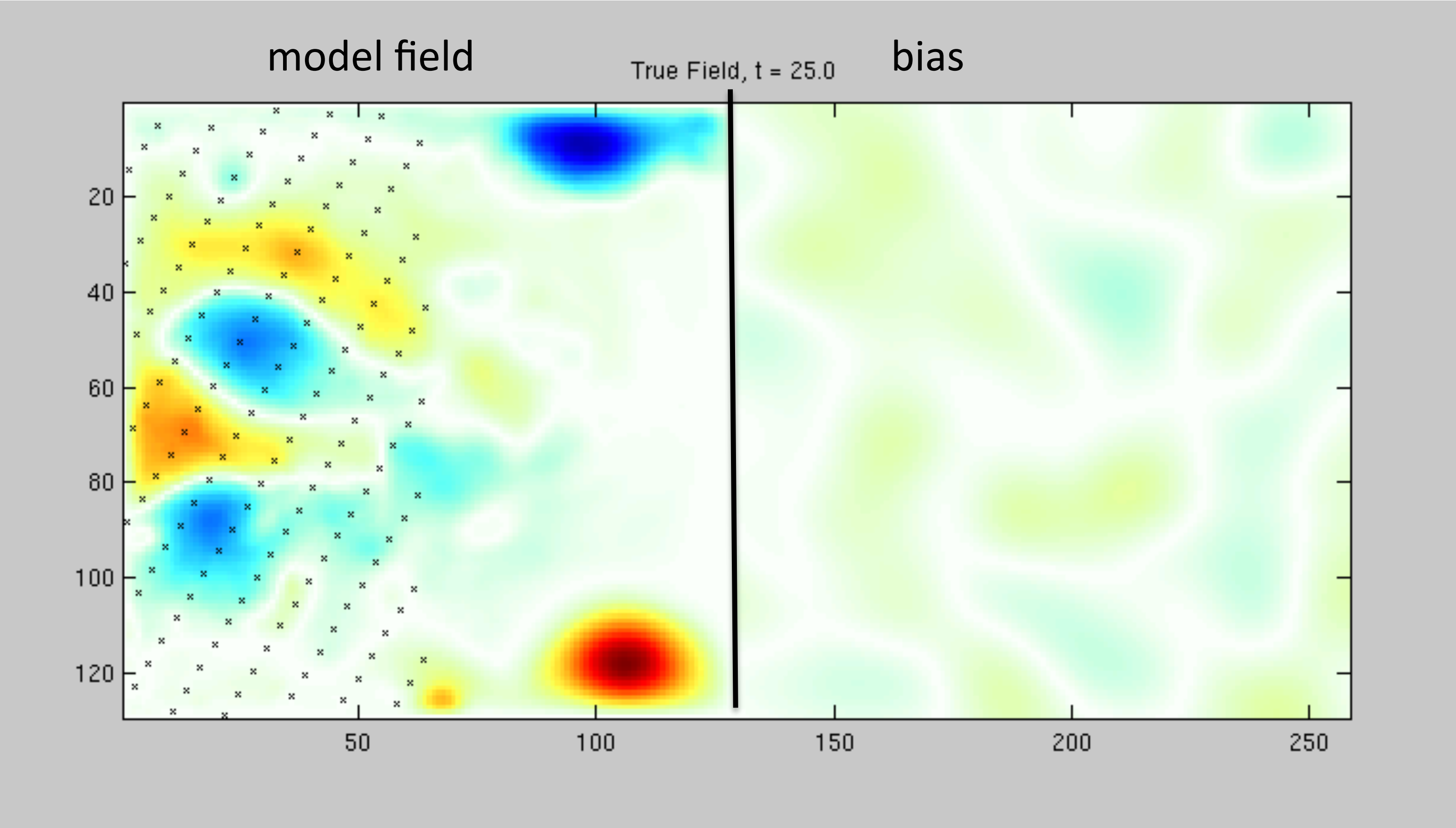
- seems to produce unrealistically small scale structure (~ effective loc radius ?)
- Some instabilities developed at observation mask boundary

•Red noise method seems more robust

Part II: Partial observation coverage

Bias is considered as red noise constant in time

Observation cover only half of the domain



method: inflation, additive red noise

initialization : uniform, red noise

Problems and description

•Model state variable inflation (1.1) makes the model unstable

→Variance in unobserved part increase infinitely until the model crashes

Without inflation ensemble collapse and the solution is inaccurate

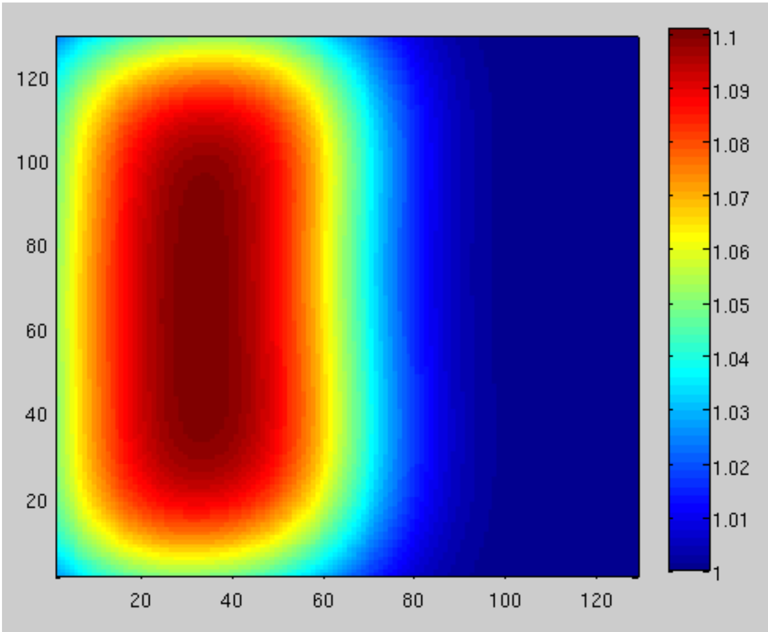
Use ad-hoc adaptive inflation

Inflation factor depends on #obs assimilated and their loc weight

*infl_mask(i,1) =1+(prm.inflation-1)*sum(coeffs)/max_obs;*

coeffs are weight return by G&C (within [0..1])

max_obs in the local window



Do we also need "adaptive" parameter inflation ?

Part II: Experiments

Exp1:

•Parameter “adaptive inflation” (1:0.025:1.2)

•Initialization uniform

Exp3 (used in TOPAZ):

•Parameter uniform inflation 1:0.025:1.2

•Initialization uniform

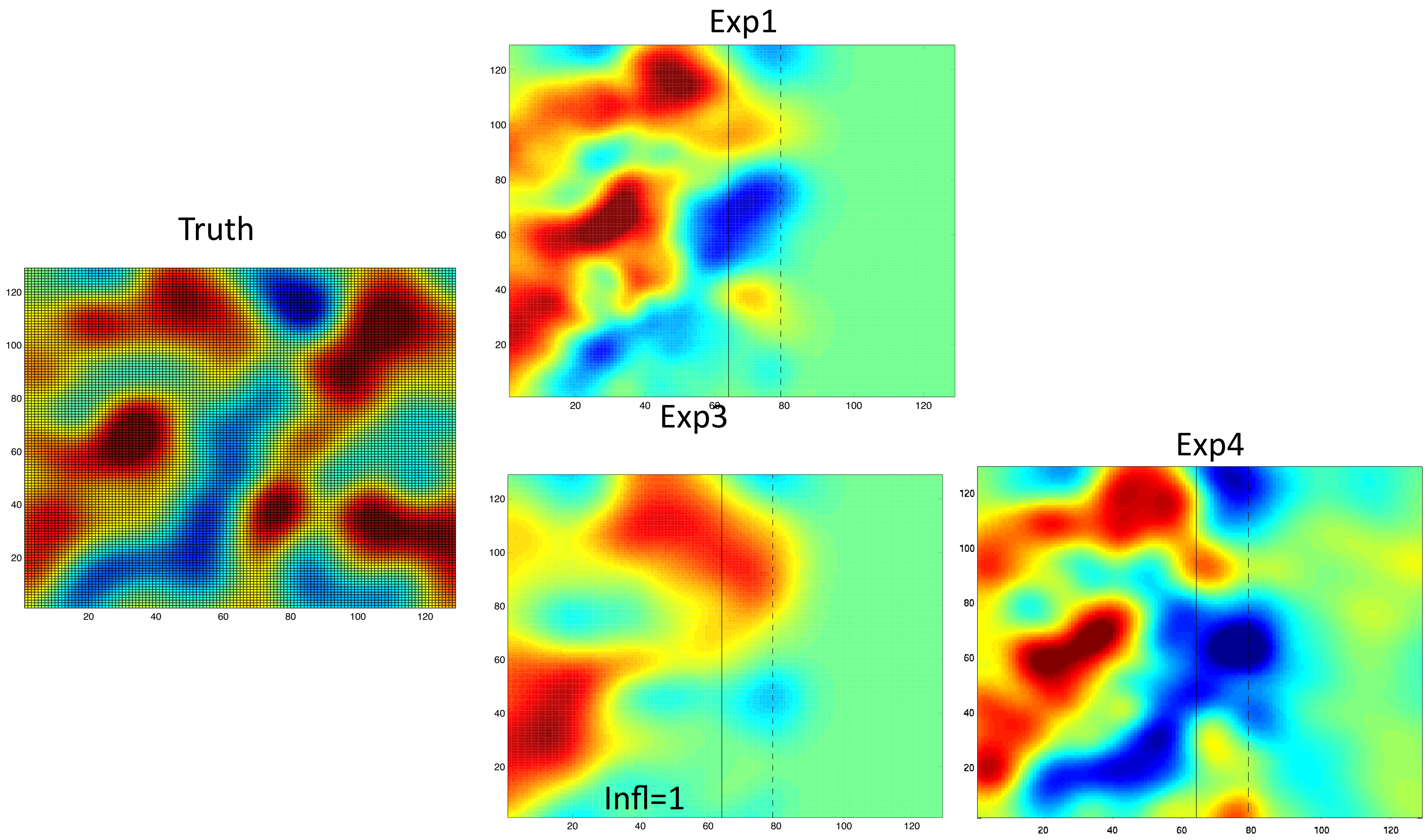
Exp4:

•Additive red noise:

pseudo2D(α) with α=[0.1 0.15 0.2 0.3 0.4 0.5 0.6 0.8])

Solution retained is the one that minimize the parameter RMSE on the full domain

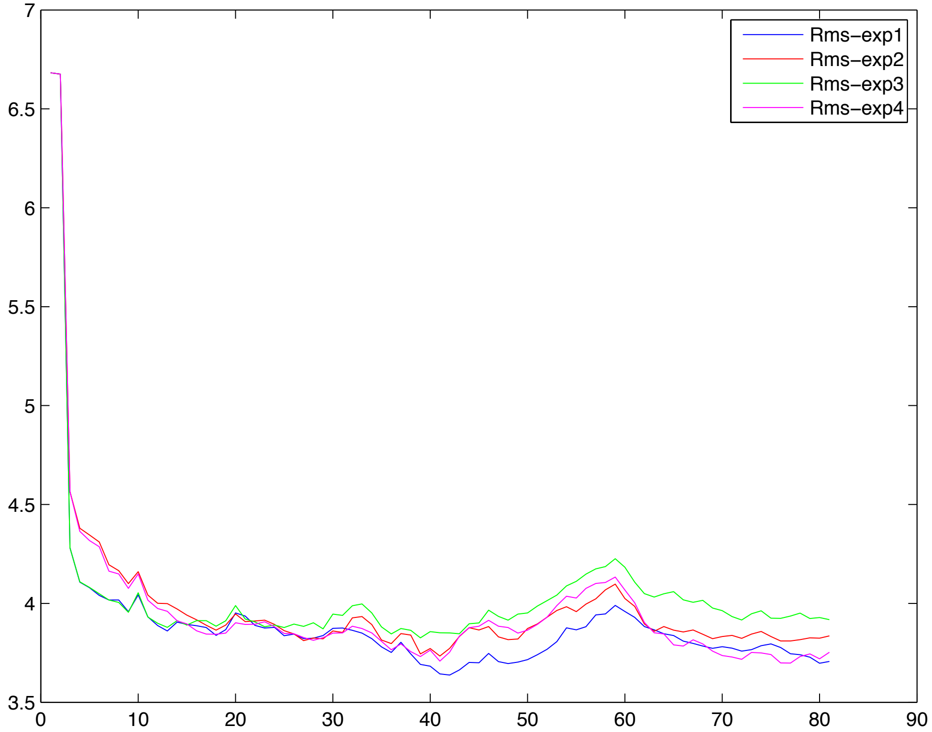
Part II: Partial coverage



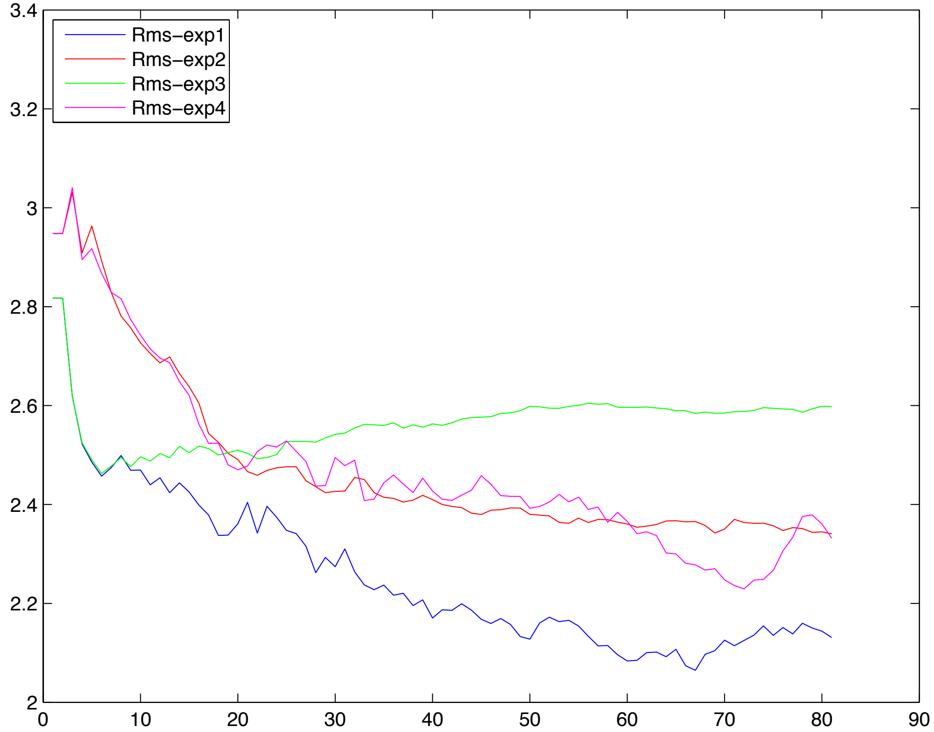
Part II: Partial coverage

Quantitative comparison *Random seed=2*

Model state RMS



Param RMS



Conclusion
Partial obs coverage

•Adaptive parameter inflation necessary when obs coverage is partial (Andersen 2001)

•Parameter estimation successful:

The match is good in area assimilated (+ localisation window)

Unrealistic values are not observed in areas not assimilated

•Red noise performs slightly poorer but:

- •Scale of the structure more realistic

•Would be interesting to test other approach that does not produce ensemble spread collapse (EnKF-N, Bocquet 2011)

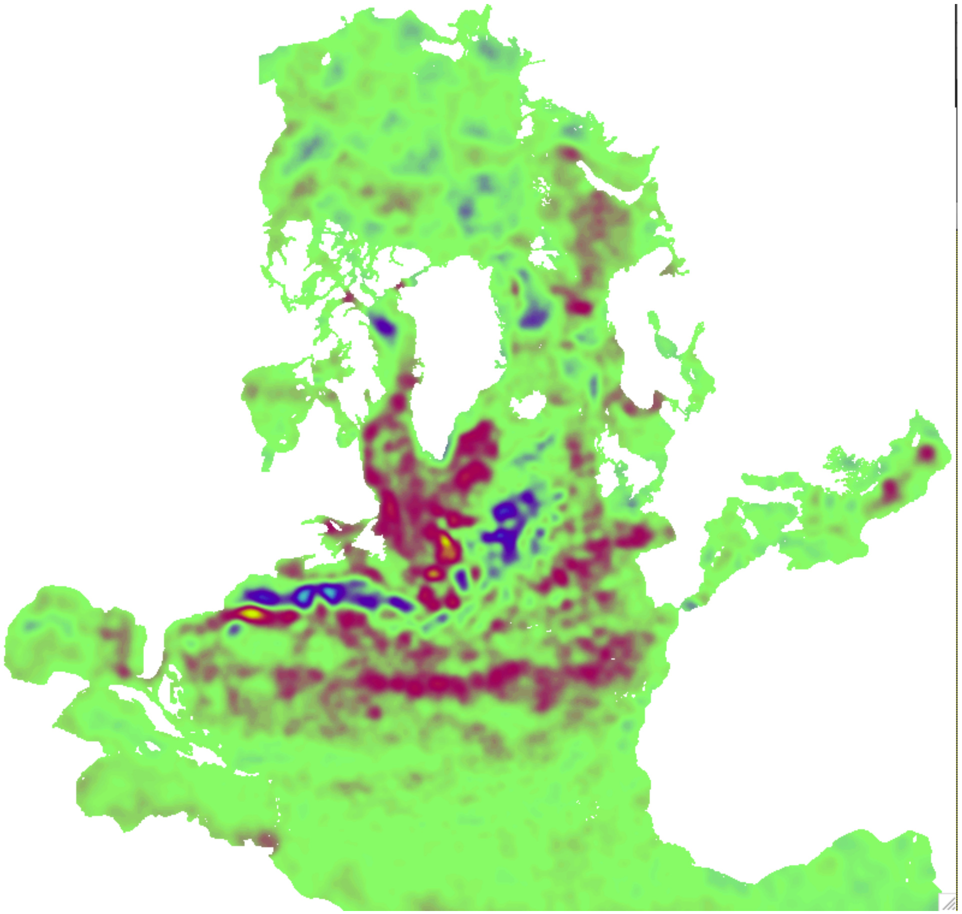
•What will happen with a fluctuating observation network ?

•Can we estimate adjective quantity (P*)

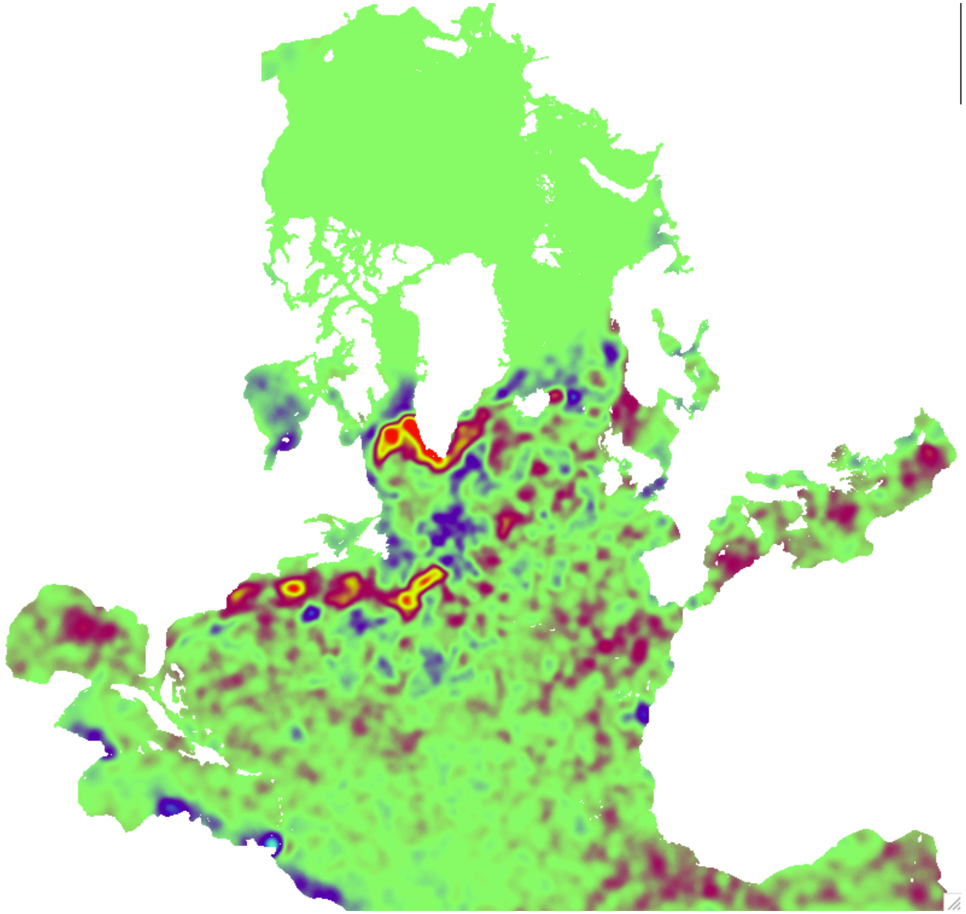
Application to TOPAZ Reanalysis

The Reanalysis was re-runned using additive spatially uniform noise and time correlation

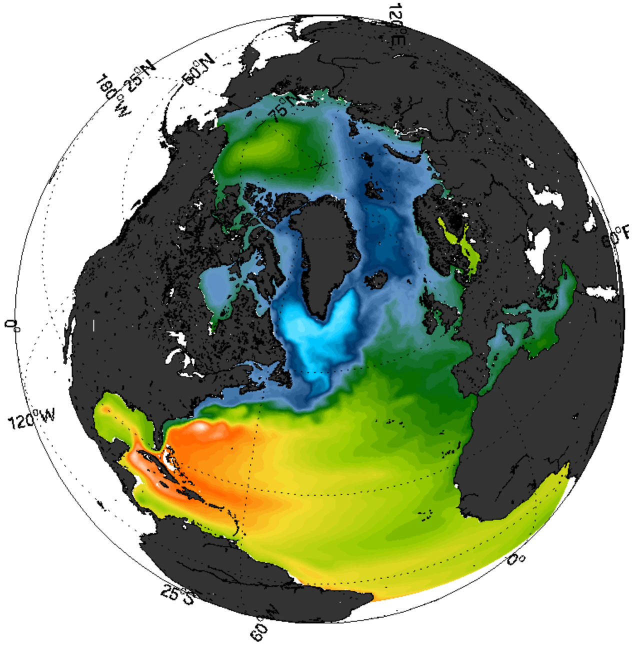
SSH bias



SST bias

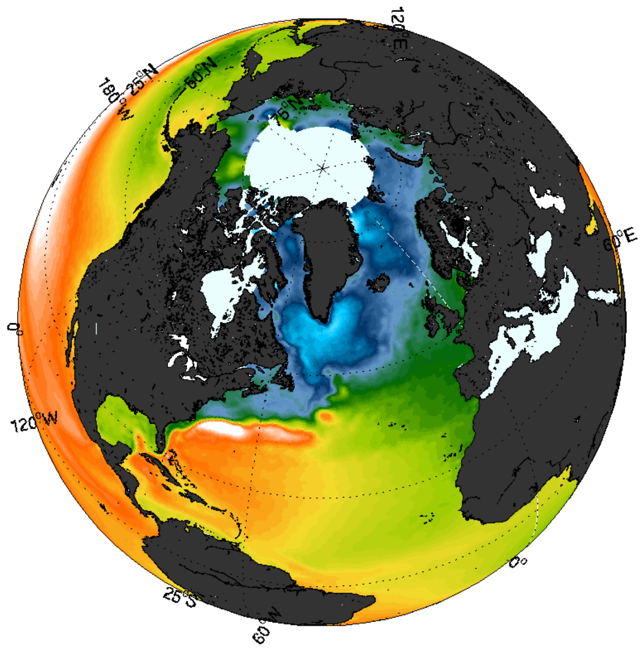


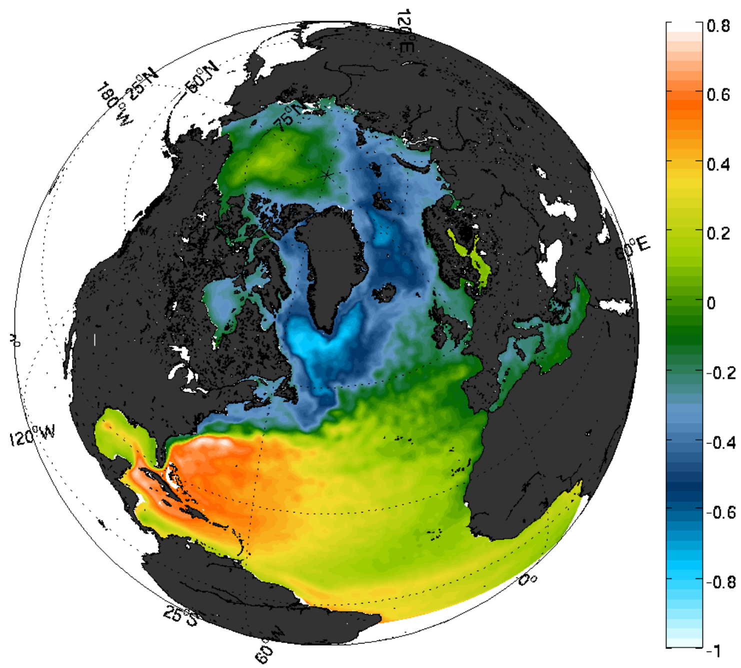
Application to TOPAZ Reanalysis



TOPAZ Free

CNES 2009





TOPAZ Free +
Bias estimated