On the selection of localization radius in ensemble filtering for multi-scale quasi-geostrophic dynamics

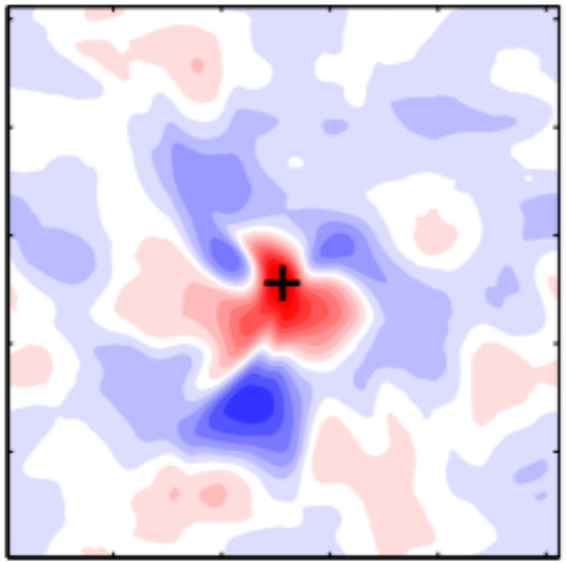
Yue (Michael) Ying, Fuqing Zhang, *PennState/ADAPT*

Jeffery Anderson, *NCAR/DAReS*

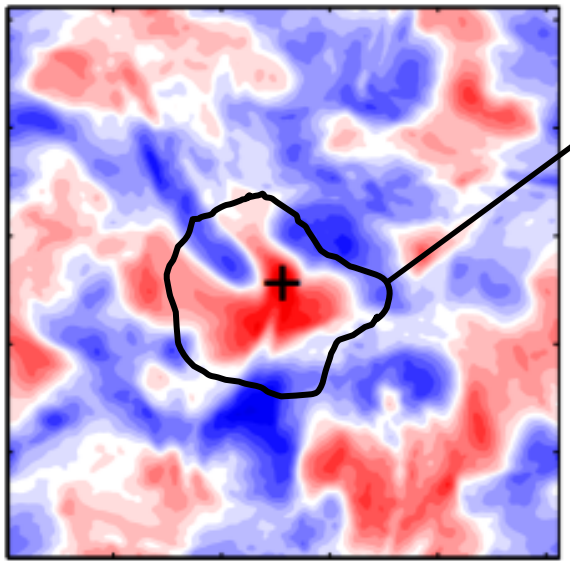
Group meeting, May 1 2017

In pursuit of an optimal localization:

"True" correlation
(N=2000)



sample-estimated
correlation (N=20)



localized
analysis increment

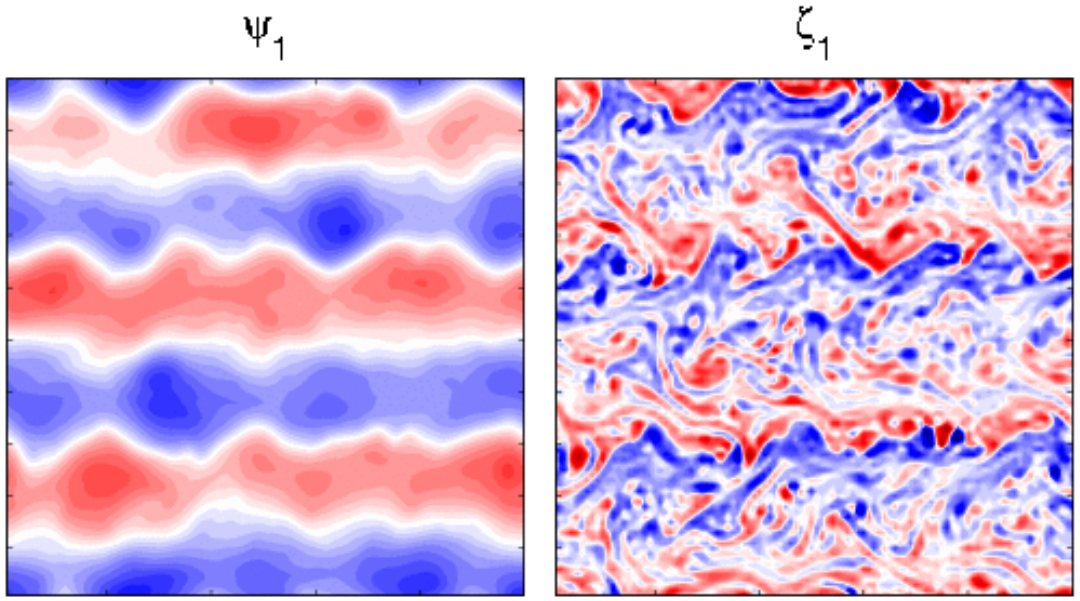
adaptive algorithms:
Zhen and Zhang 2014
Perianez et al. 2014
Kirchgessner et al. 2014
Flowerdew 2015

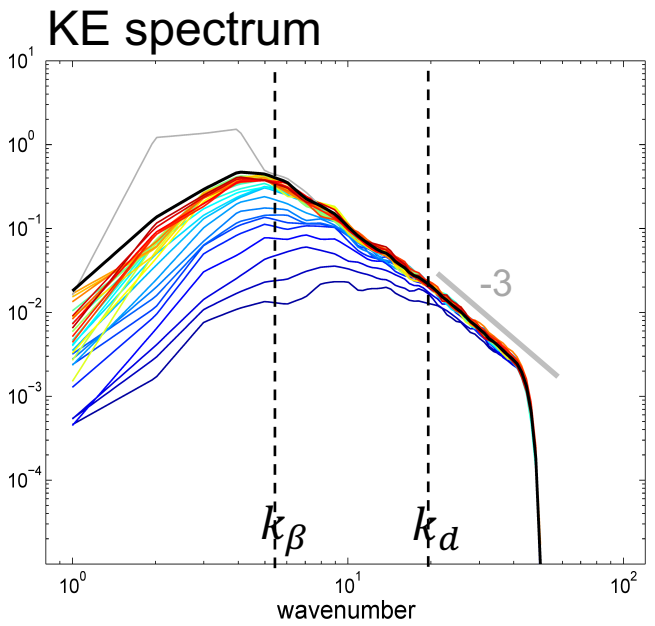
localization radius depends on:

- ensemble size
- observing network properties (density, frequency, accuracy)
- underlying covariance structure, flow-dependency (Anderson 2007)

What is their relative importance for multi-scale dynamics?

Two-layer Quasi-Geostrophic Model





colored lines: error evolution
gray: reference energy

governing equations:

$$\frac{\partial q_1}{\partial t} + J(\psi_1, q_1) + U\frac{\partial q_1}{\partial x} + (\beta + k_d^2 U)\frac{\partial \psi_1}{\partial x} = 0,$$

$$\frac{\partial q_2}{\partial t} + J(\psi_2, q_2) - U\frac{\partial q_2}{\partial x} + (\beta - k_d^2 U)\frac{\partial \psi_2}{\partial x} + r\nabla^2\psi_2 = 0,$$

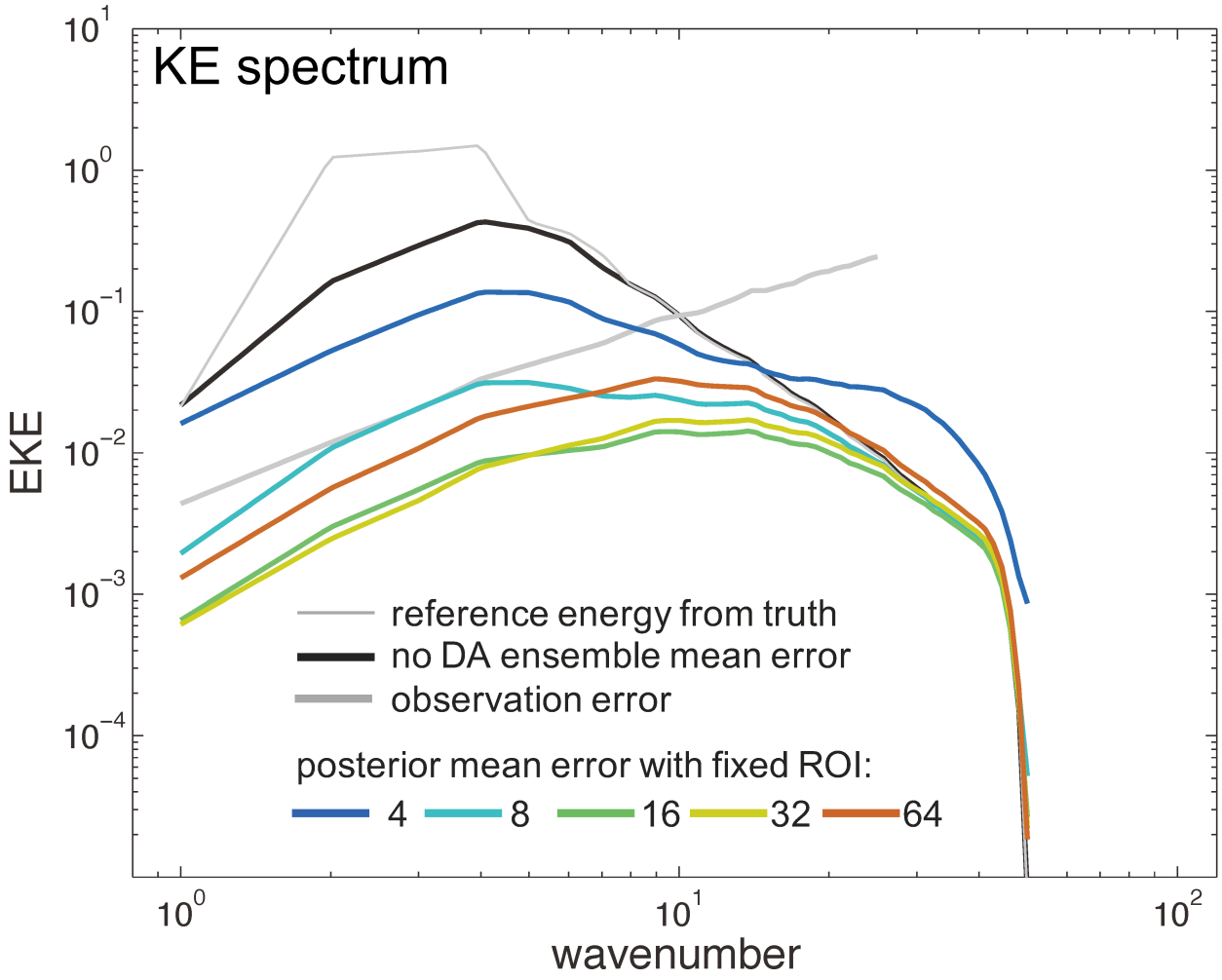
$$q_1 = \nabla^2\psi_1 + \frac{k_d^2}{2}(\psi_2 - \psi_1),$$

$$q_2 = \nabla^2\psi_2 - \frac{k_d^2}{2}(\psi_2 - \psi_1),$$

capture the essence of large-scale atmospheric dynamics

simple enough for sensitivity experiments

Finding best radius of influence (ROI) by trial-and-error



CNTL experiment

ensemble size N = 40

observing network
$\Delta x = 2\ dx$

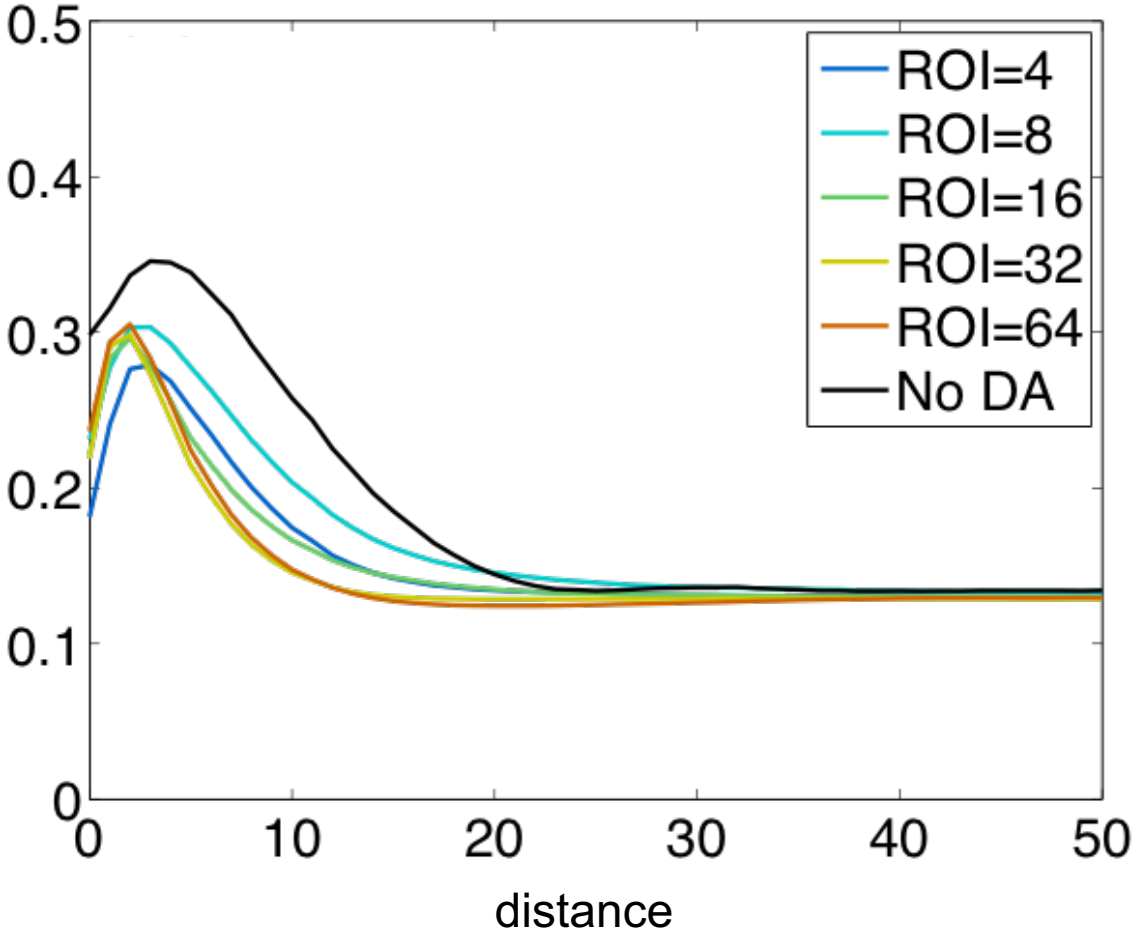
model grid resolution
$L = n\ dx,\ n = 100$

observation: u,v
state variable: ψ

- best ROI is 16
- larger ROI is suboptimal for all scales
- smallest ROI introduce noise in the unobserved scales

Mean absolute correlation (MAC) as a function of distance

$|\mathrm{corr}(\psi, u)|$ averaged over space and time

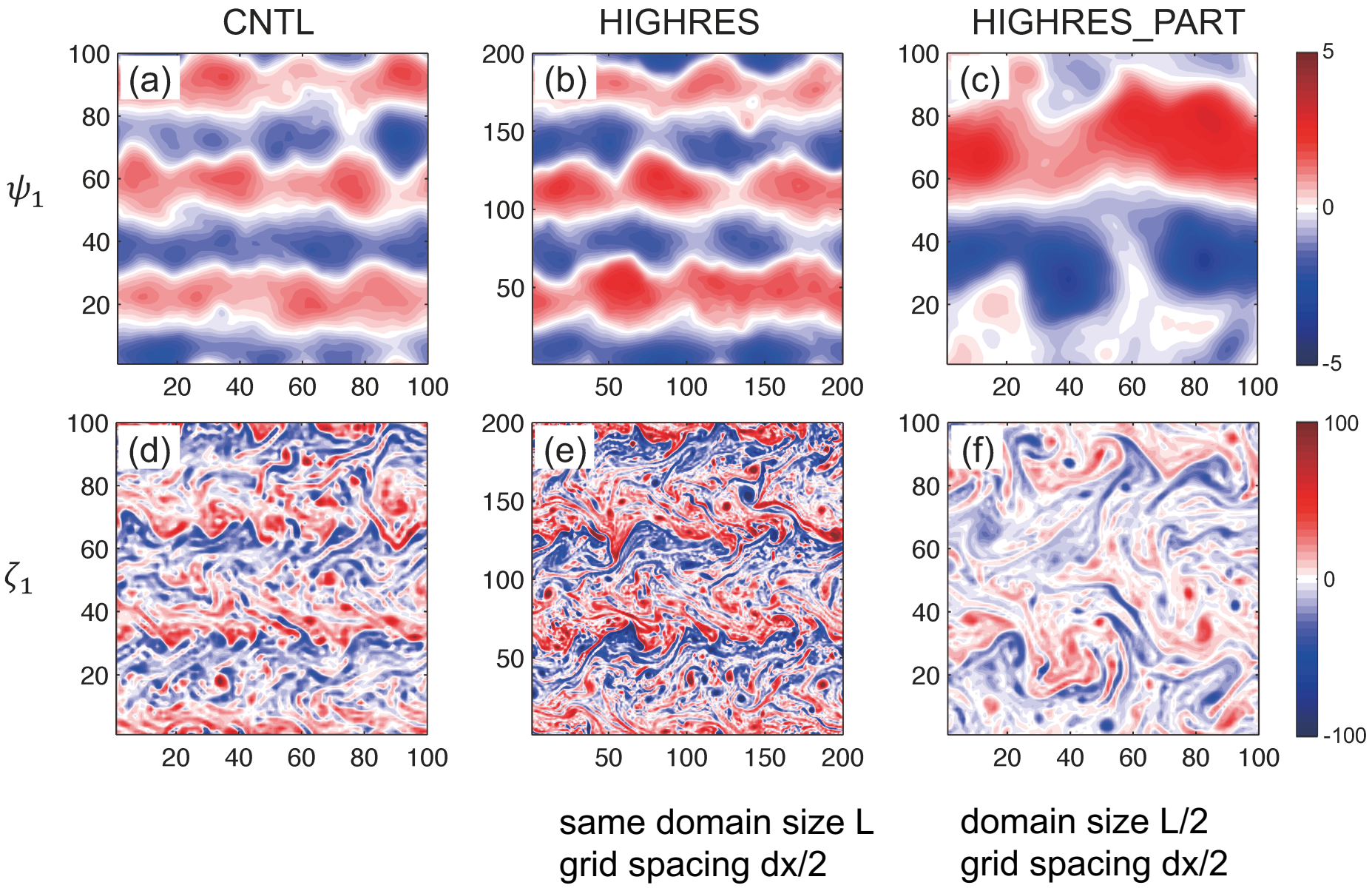


No DA correlation scale remains large

Asymptotic value of MAC is determined by ensemble size N.

For the best ROI = 16, at distance = 16 the MAC is close enough to the asymptotic value.

Sensitivity to model resolution

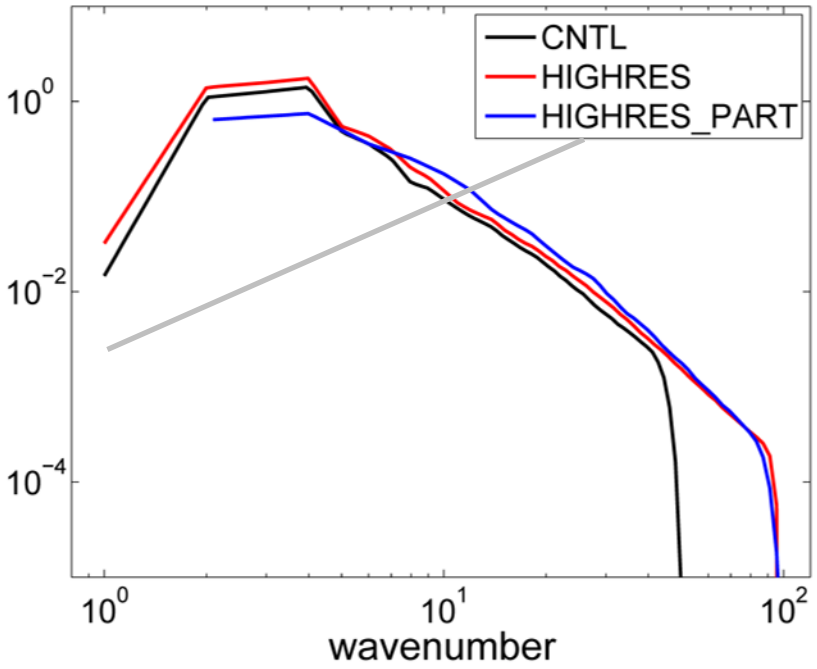


same domain size L
grid spacing dx/2

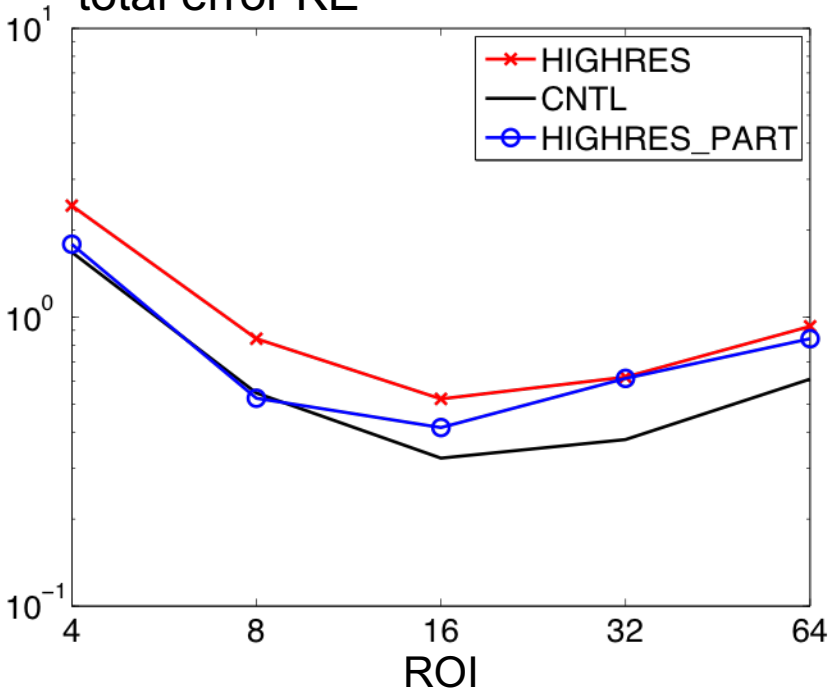
domain size L/2
grid spacing dx/2

Sensitivity to model resolution

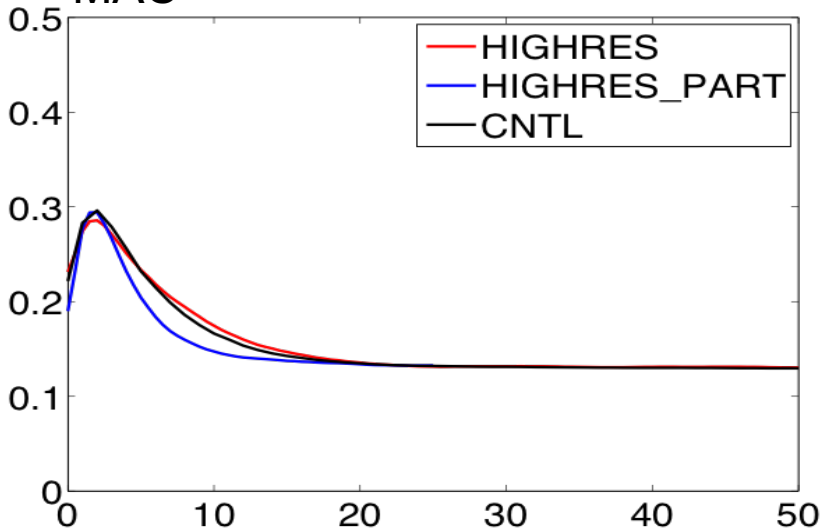
reference KE spectrum



total error KE

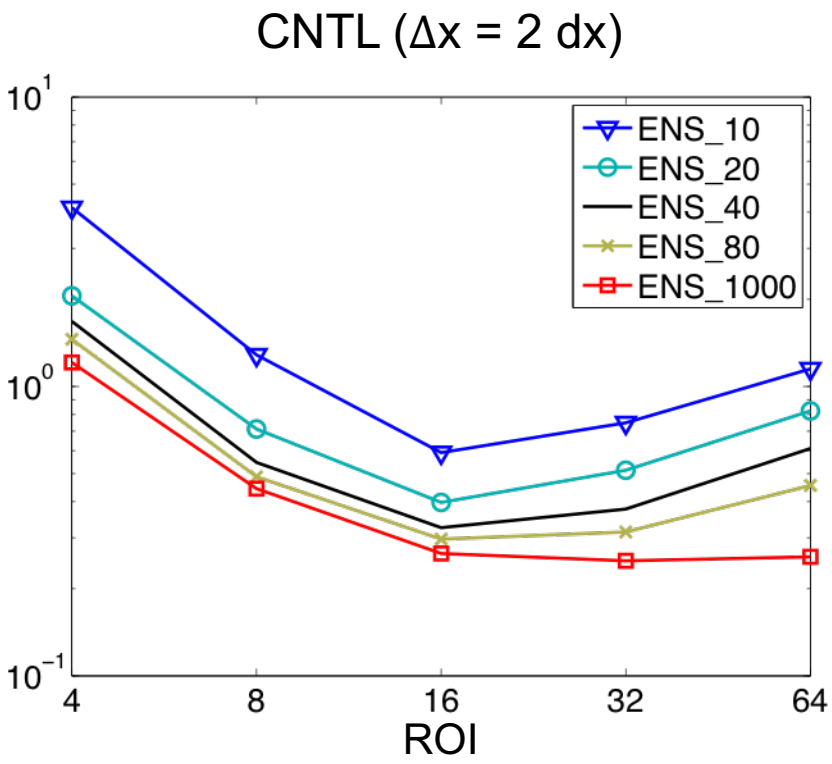


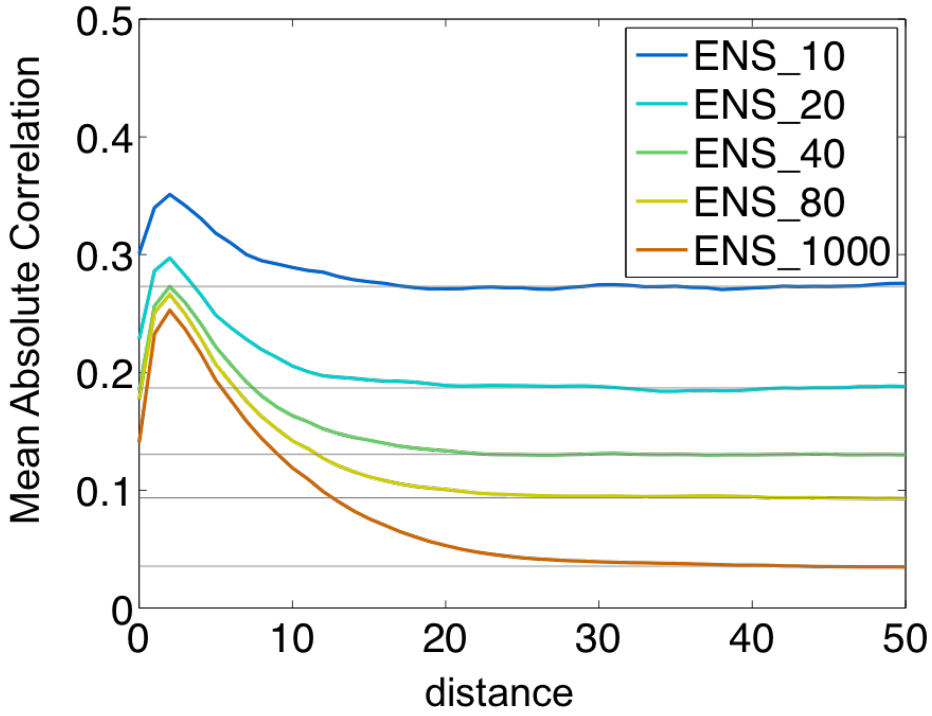
MAC



Increasing model resolution does not change the characteristic correlation scales, therefore the best ROI remains 16.

Sensitivity to ensemble size



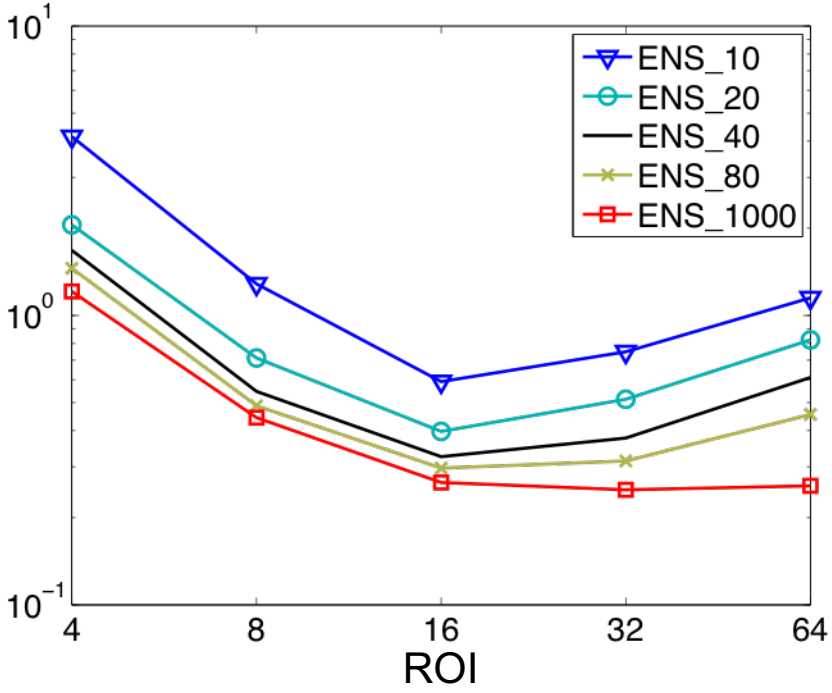


sampling error decreases as ensemble size increase

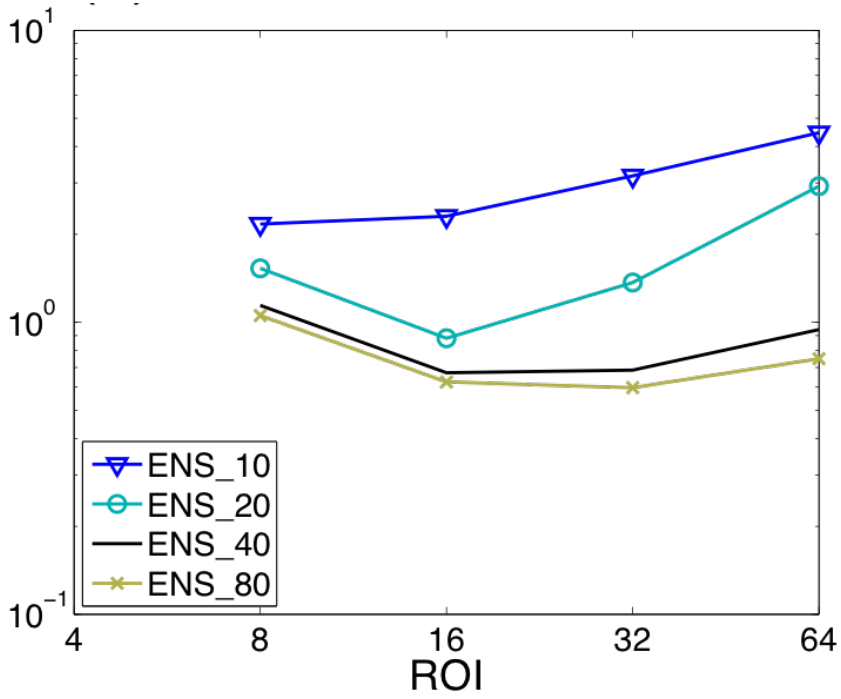
the favorable range of ROIs is wider for larger ensemble size

Sensitivity to ensemble size

CNTL (Δx = 2 dx)



sparse network (Δx = 4 dx)

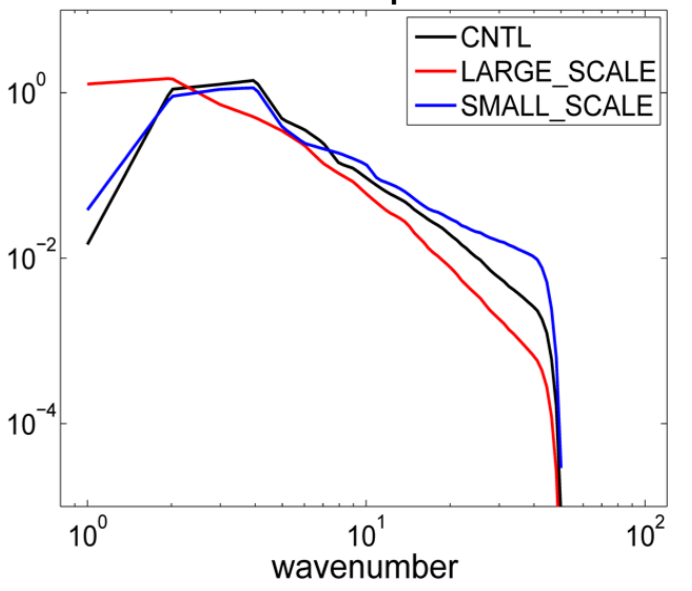


shorter ROI should be used for smaller ensemble size (Anderson 2007 and many others)

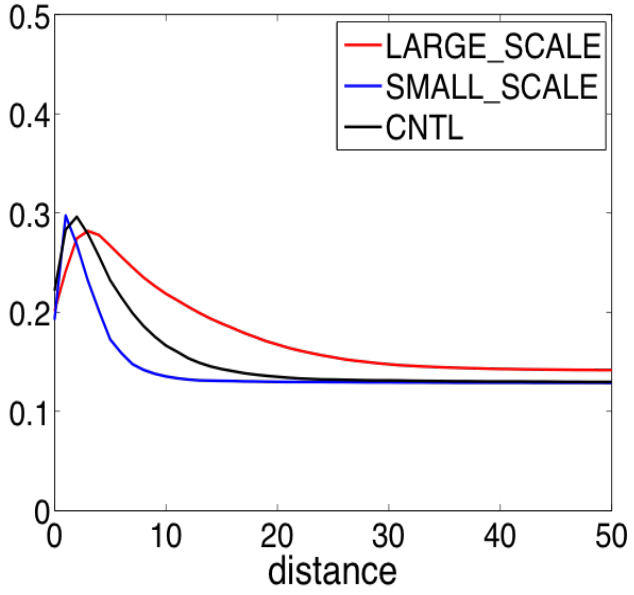
For denser observing network (left), the small ensemble size seems to have more tolerance to sampling error.

Sensitivity to characteristic correlation scales

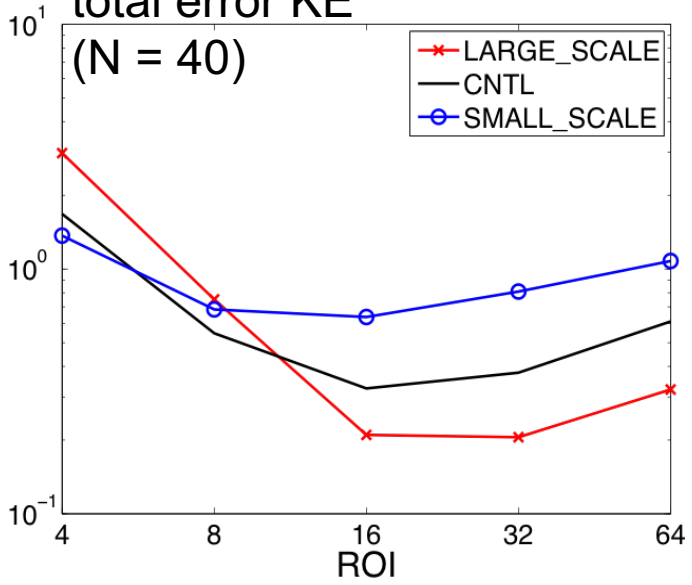
reference KE spectrum



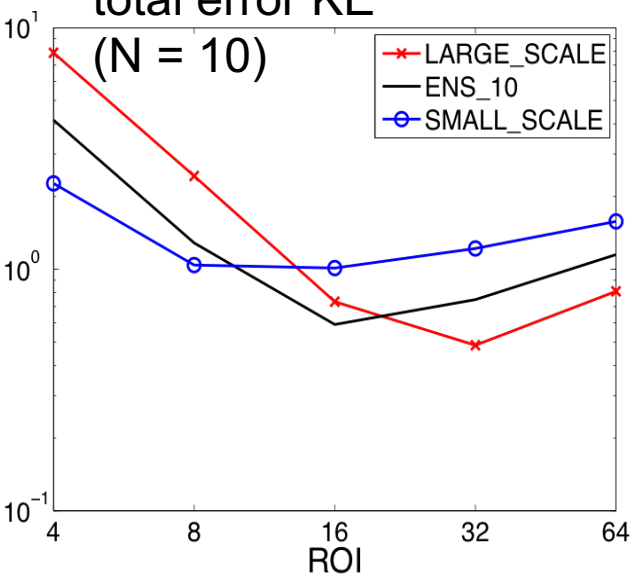
MAC



total error KE (N = 40)



total error KE (N = 10)



Changing characteristic correlation scale by shifting spectral peak,

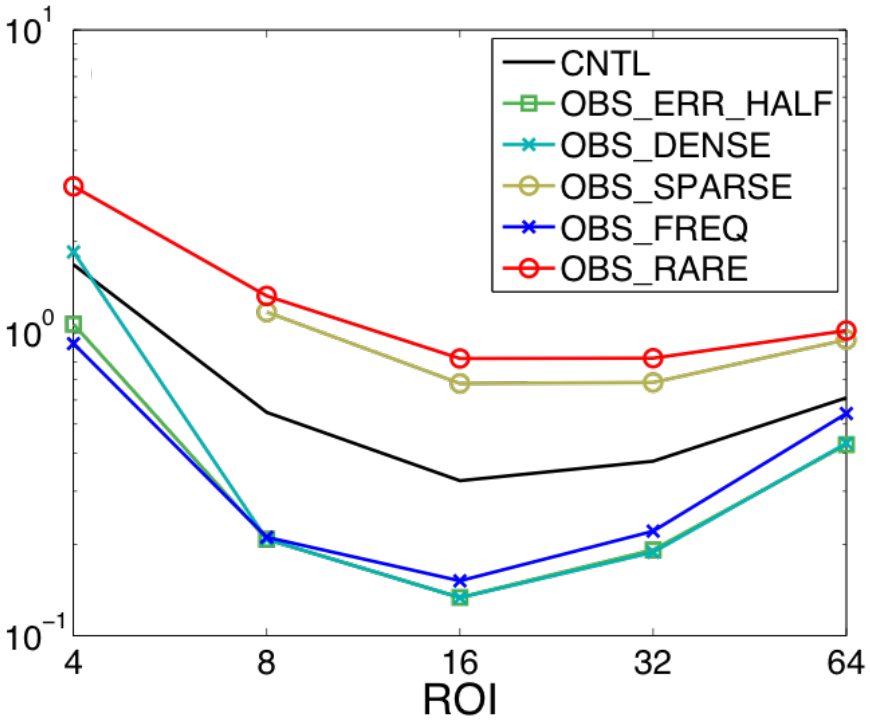
more energy in small scale => shorter correlation scale.

The best ROI shifts to larger value if correlation scale is larger.

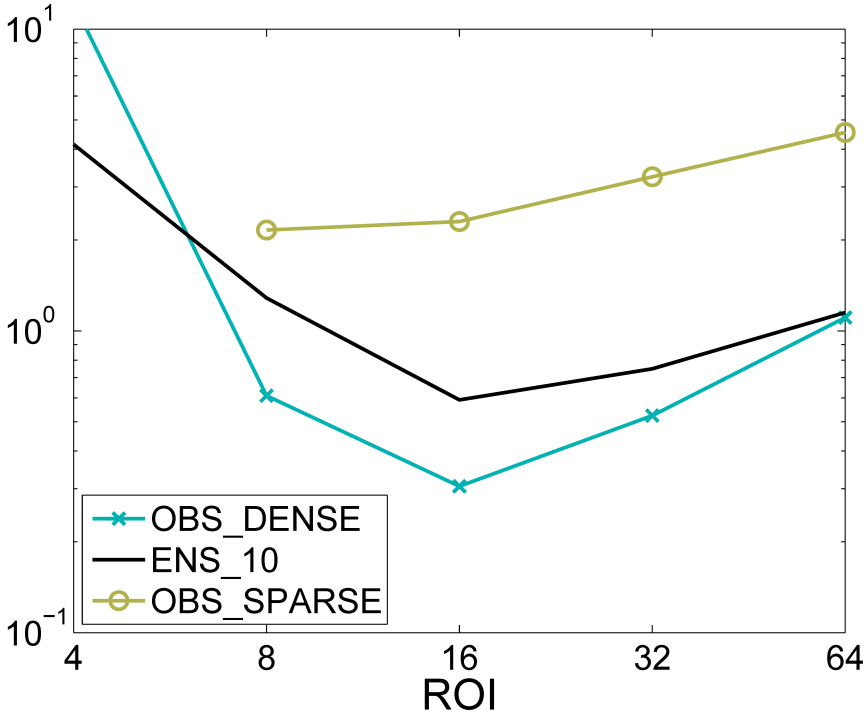
This effect is more evident for smaller ensemble

Sensitivity to observing network

CNTL (N = 40)



small ensemble (N = 10)



Contradicting results:

- Denser network requires shorter ROIs (Fuqing's experience, Dong et al. 2011, Kirchgessner et al. 2014, Perianez et al. 2014)
- Denser network has broader optimal localization function (Anderson 2007)

Similar effect of increasing observation density (OBS_DENSE), time frequency (OBS_FREQ), or reducing observation error (OBS_ERR_HALF).

Concluding remarks

- A comprehensive sensitivity experiment is conducted using QG model to investigate how best-performing localization radius changes with:

- model resolution
(not sensitive)

- ensemble size
(small N requires shorter ROI, especially for sparse network)

- observing network
(sparser network requires shorter ROI if N is small; but can use longer ROI if N is large enough?)

- characteristic correlation scale (determined by dynamics)
(an important factor, especially for small N)