STOCHASTICA: WG3: Bayesian Inference

Advancing multiscale and non-Gaussian data assimilation methodology for geophysical models

Yue Ying

(Nansen Environmental and Remote Sensing Center, Norway)

NERSC

About me - “Tropic to Arctic”

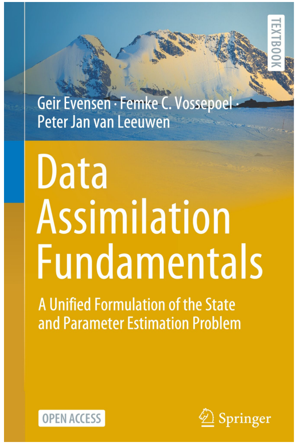
myying.github.io

- PhD in Meteorology, Penn State University, 2018
  - Tropical weather systems
  - Ensemble weather prediction and data assimilation (DA)
- Postdoc at NCAR “Advanced Study Program”
  - Multiscale DA methods (*Ying 2019, 2020*)
- Researcher at NERSC (2020-present)
  - Coupled ocean and sea ice models: addressing the challenges in DA
  - New software: NEDAS
  - New project: COMEDI

Background

Data assimilation seeks to optimally combine:

- Model state variables (the prior) $x^b \sim p(x)$
- Observations (likelihood) $y^o \sim p(y|x)$



Evensen et al. 2022 textbook

Bayes' rule gives the updated (analysis, posterior) distribution

$$x^a \sim p(x|y) \propto p(x)p(y|x)$$

Filtering (cycling DA): model forecasts are paused to be updated by available observations, then analysis is used as new condition for next forecasts.

The ensemble Kalman filter (EnKF)

Assume that error distribution are Gaussian

$$x^b \sim \mathcal{N}(\bar{x}^b, \mathbf{B}) \qquad y^o \sim \mathcal{N}(\bar{y}^o, \mathbf{R}) \qquad y = \mathbf{H}x$$

minimizing a cost function

$$J(x) = ||y^o - \mathbf{H}x||^2_{\mathbf{R}} + ||x - x^b||^2_{\mathbf{B}}$$

The ensemble Kalman filter (EnKF)

Assume that error distribution are Gaussian

$$x^b \sim \mathcal{N}(\bar{x}^b, \mathbf{B}) \qquad y^o \sim \mathcal{N}(\bar{y}^o, \mathbf{R}) \qquad y = \mathbf{H}x$$

a simple update equation is available (scalable to high dimensional problems)

$$\begin{aligned} x^a &= x^b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(y^o - \mathbf{H}x^b) \\ &= x^b + \frac{\mathrm{cov}(x^b, y^b)}{\mathrm{cov}(y^b) + \mathrm{cov}(y^o)}(y^o - y^b) \end{aligned}$$

… but the solution becomes suboptimal as nonlinearity increases

The particle filters

Fully nonlinear approaches:

represents error distribution with particles

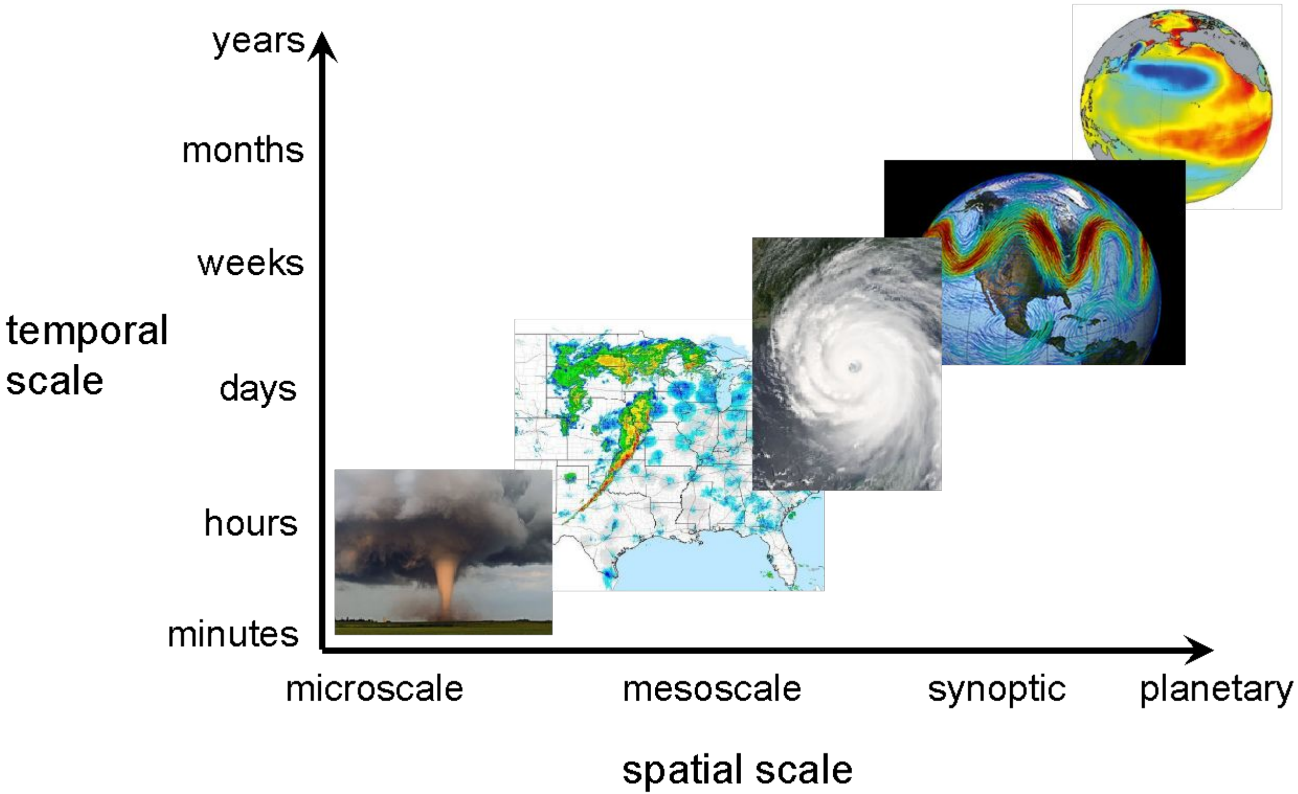
$$p(x) \approx \sum_{j=1}^{N} \frac{1}{N} \delta(x - x_j)$$

Importance sampling:

$$p(x|y) = \sum_{j=1}^{N} w_j \delta(x - x_j) \qquad w_j = \frac{p(y|x)}{p(y)}$$

… but suffer more from “curse of dimensionality”: weights collapse quickly in high dimensional problems.

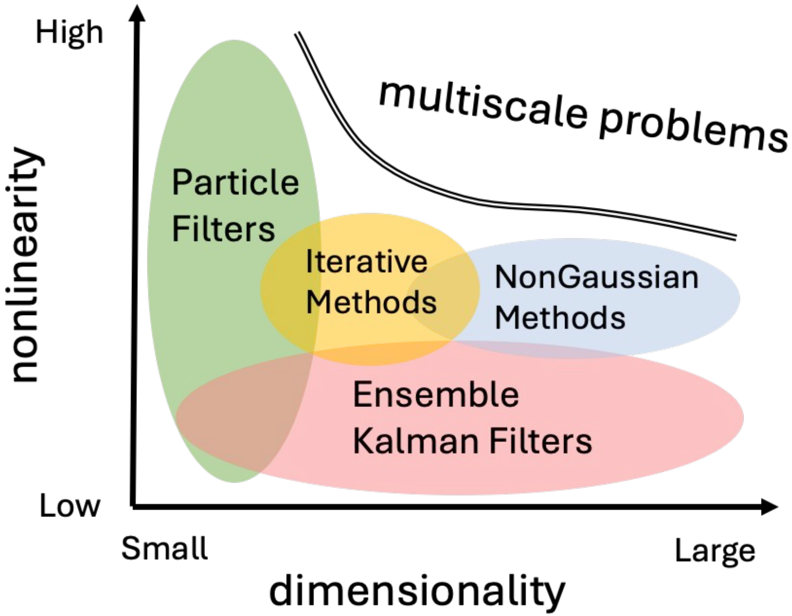
Geophysical systems



The DA research frontier: multiscale problems

Challenges in the analysis and prediction of geophysical systems (atmosphere, ocean, sea ice …)

- High dimensionality
  - $\mathcal{O}(10^7)$ model variables and $\mathcal{O}(10^5)$ observations
- High nonlinearity at small scales
  - Chaotic nature leads to rapid error growth
  - Non-Gaussian error distributions



A multiscale approach to DA

Decompose the state and (optionally) observations into scale components:

$$x_s^b = \mathbf{F}_s x^b \qquad\qquad x^b = \sum_{s=1}^{N_s} x_s^b$$

$$y_s^o = \mathbf{F}_s^o y^o \qquad y_s^b = \mathbf{F}_s^o \mathbf{H} x^b \qquad\qquad y^b = \sum_{s=1}^{N_s} y_s^b$$

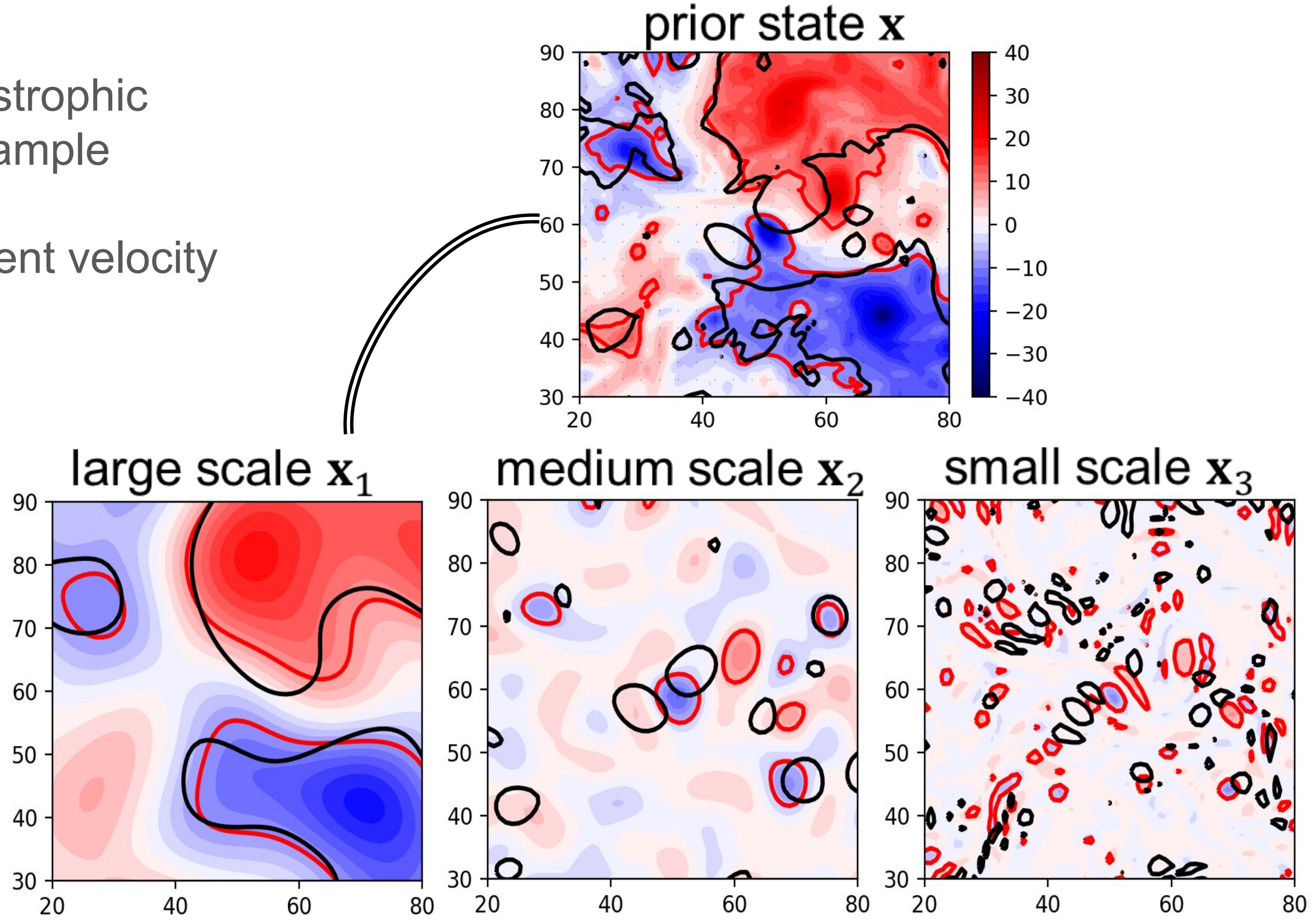
Iteratively update the scale components

$$x_s^a = x_s^b + \mathbf{L}_s \circ \frac{\mathrm{cov}(x_s^b, y_s^b)}{\mathrm{cov}(y_s^b) + \mathrm{cov}(y_s^o)} (y_s^o - y_s^b) \qquad s = 1, \ldots, N_s$$

… to get the final analysis

Quasi geostrophic system example

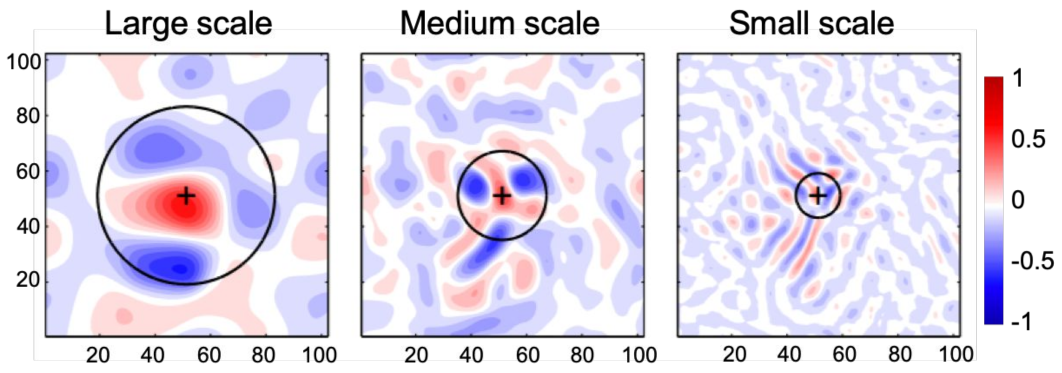
U-component velocity



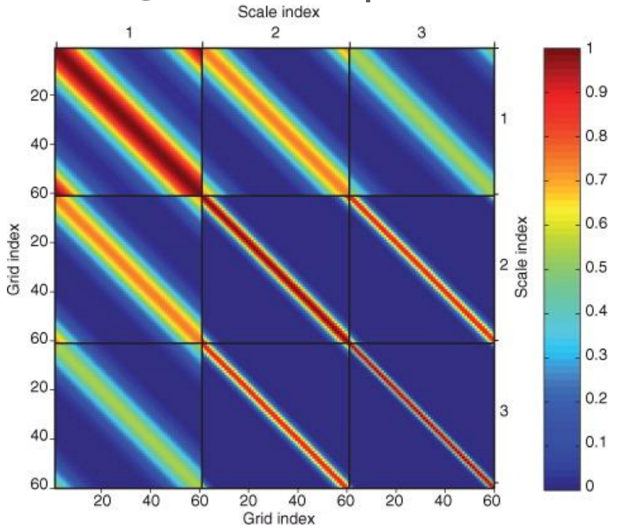
Multiscale localization

The localization function $\mathbf{L}$ tapers the covariance to remove spurious correlations

In a multiscale approach, different localization functions $\mathbf{L}_s$ can be specified



Optimal localization radius *(Ying et al. 2018)*



Buehner & Shlyaeva 2015

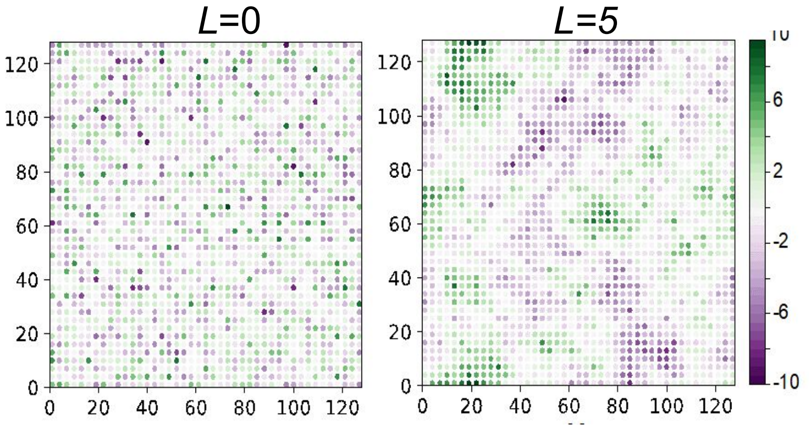
Multiscale observation

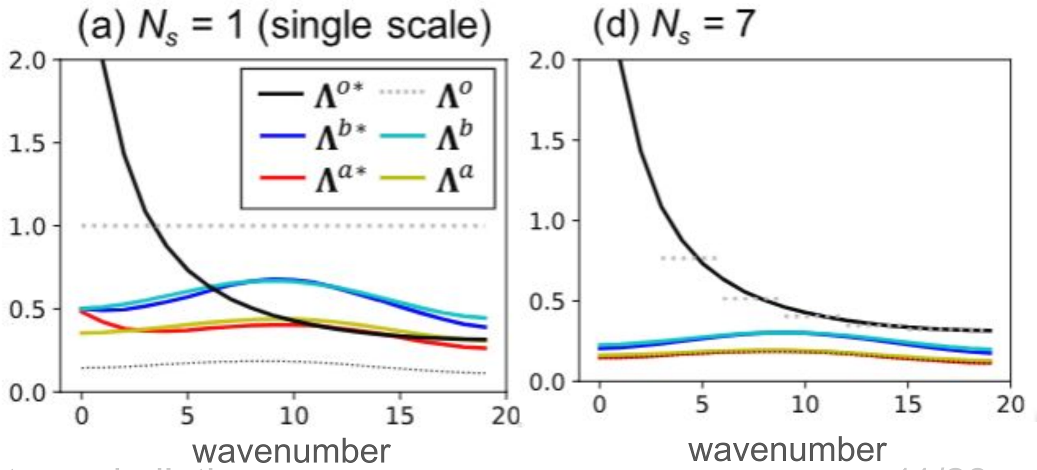
Ying 2020, Mon Wea Rev

For dense observation network with correlated errors $\mathbf{R}_{ij} = \sigma^2 \exp(-D_{ij}/L)$

EnKF update cannot utilize all the information (rely on thinning)

A multiscale approach can adapt to the varying spectral info content



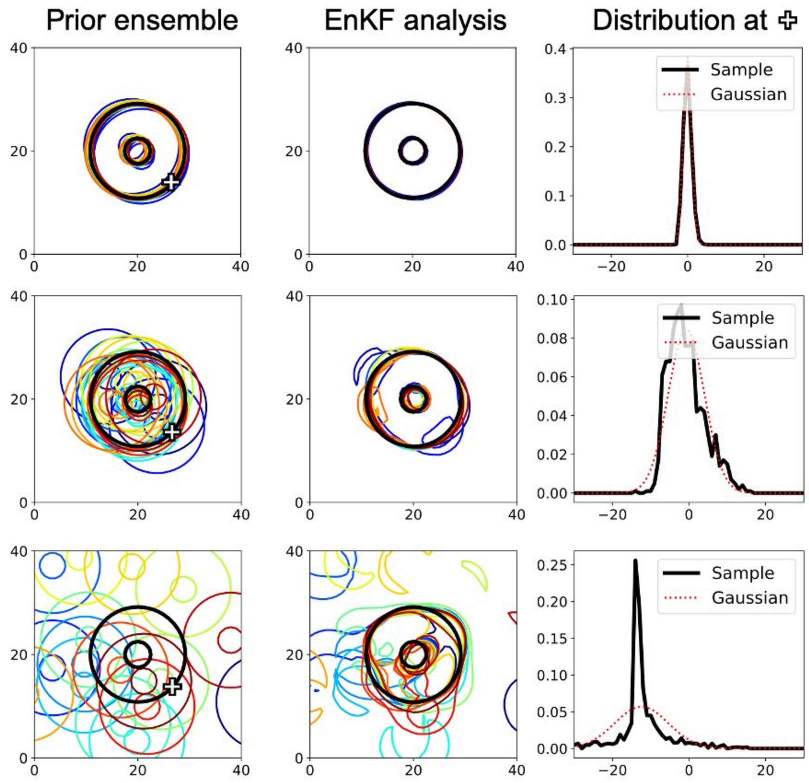


Nonlinearity due to position errors

$L_{\mathrm{sprd}}/R_{\mathrm{mw}}$ = 0.1

0.5

3



Rankine vortex (size R_{mw})

As position error L_{sprd} increases

- Error distribution more non-Gaussian
- EnKF analysis becomes more suboptimal (colored contours indicate analysis ensemble)

DA with nonlinear position error

Error model: $x^b = x^{true} + \varepsilon^d + \varepsilon^r$

$$\varepsilon^d = x^b - x^b(q)$$
$$\varepsilon^r = x^b(q) - x^{true}$$

Bayesian formulation with explicit position error

$$p(x, q|y) \propto p(x)p(x|q)p(y|x, q)$$

Cost function:

$$J(x, q) = ||y - H[x(q)]||^2_{\mathbf{R}} + ||x(q) - x^b(q)||^2_{\mathbf{B}(q)} + \ln(|B(q)|) + L(q)$$

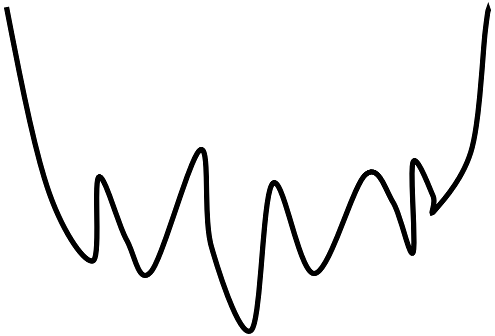
Two-step solver: 1. Derive displacement q, 2. EnKF update x

(Ravela et al. 2007; Nehrkorn et al. 2015)

Multiscale alignment

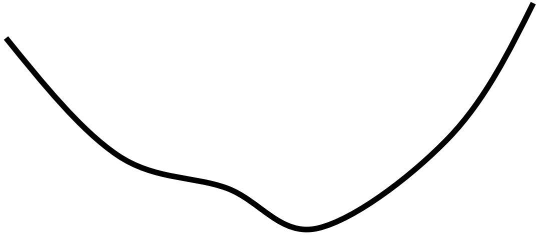
Topography of J(x,q):

Nonlinearity causes a lot of local minima



Idea:

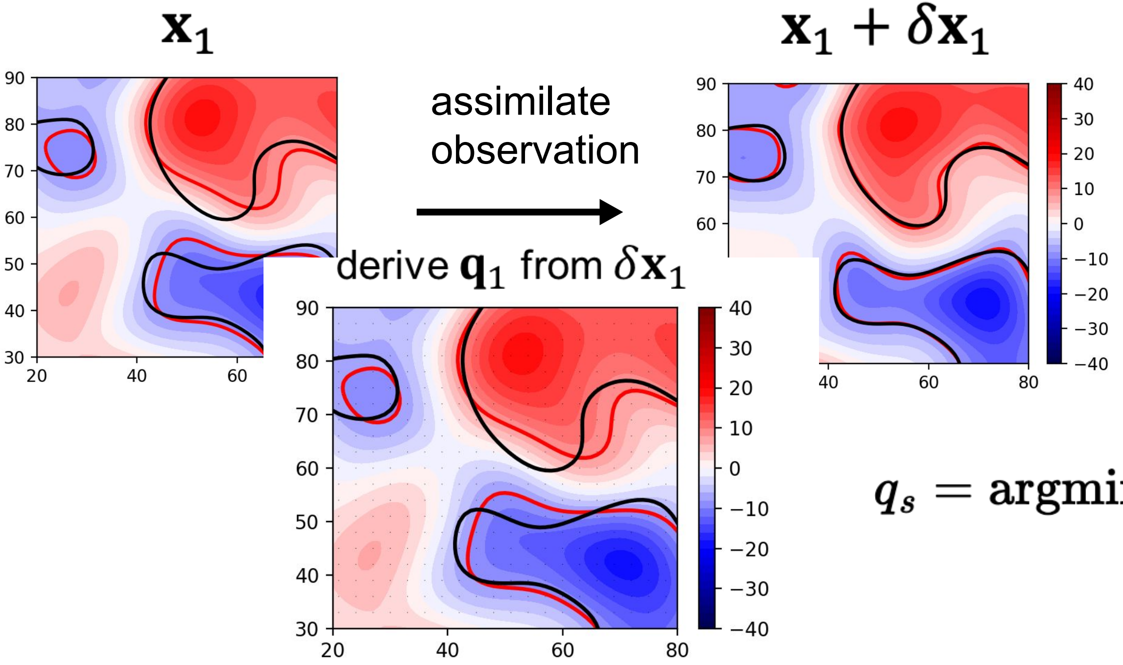
Use iterations over scale components (outer loops in 4DVar) to skip towards global minimum



Multiscale alignment approach

Ying 2019 Mon Wea Rev.

$$\mathbf{x} + \delta\mathbf{x} = \mathbf{x}_1 + \delta\mathbf{x}_1 + \mathbf{x}_2(\mathbf{q}_1) + \delta\mathbf{x}_2 + \mathbf{x}_3(\mathbf{q}_1 + \mathbf{q}_2) + \delta\mathbf{x}_3$$



$$\delta x_s = \mathbf{L}_s \circ \frac{\mathrm{cov}(x_s, y_s)}{\mathrm{cov}(y_s) + \mathrm{cov}(y_s^o)}(y_s^o - y_s)$$

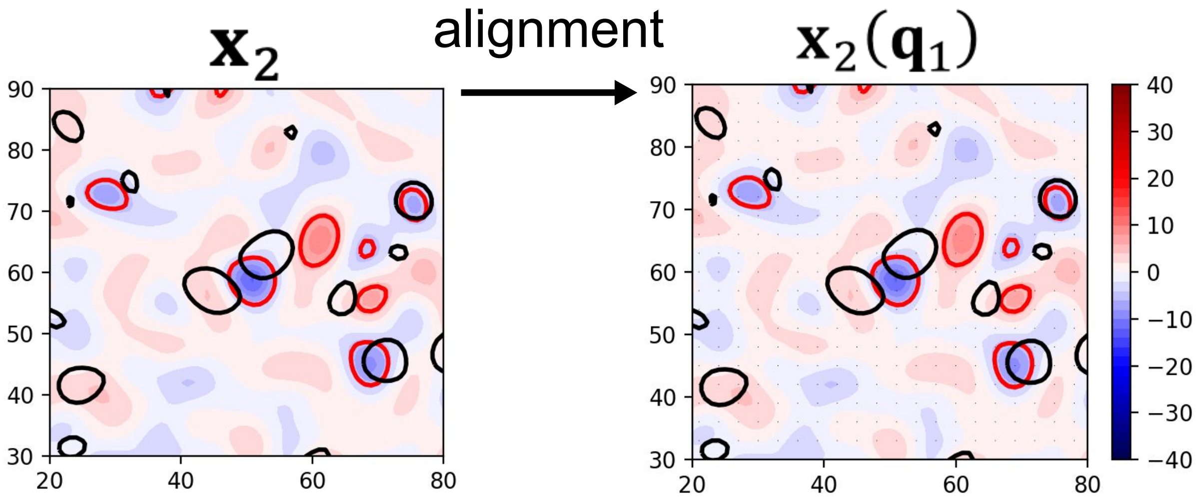
$$q_s = \mathrm{argmin} \quad ||\delta x_s||^2 + w||\nabla q||^2$$

Optical flows

Multiscale alignment approach

Ying 2019 Mon Wea Rev.

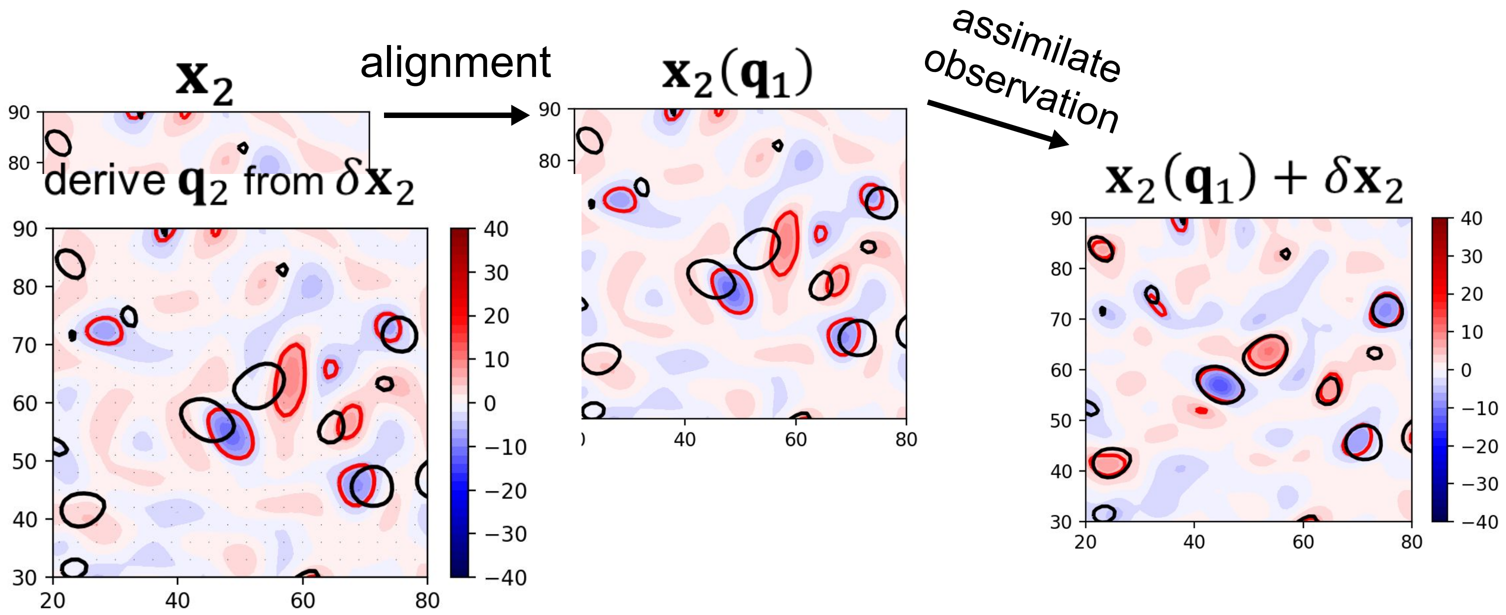
$$\mathbf{x} + \delta\mathbf{x} = \mathbf{x}_1 + \delta\mathbf{x}_1 + \mathbf{x}_2(\mathbf{q}_1) + \delta\mathbf{x}_2 + \mathbf{x}_3(\mathbf{q}_1 + \mathbf{q}_2) + \delta\mathbf{x}_3$$



Multiscale alignment approach

Ying 2019 Mon Wea Rev.

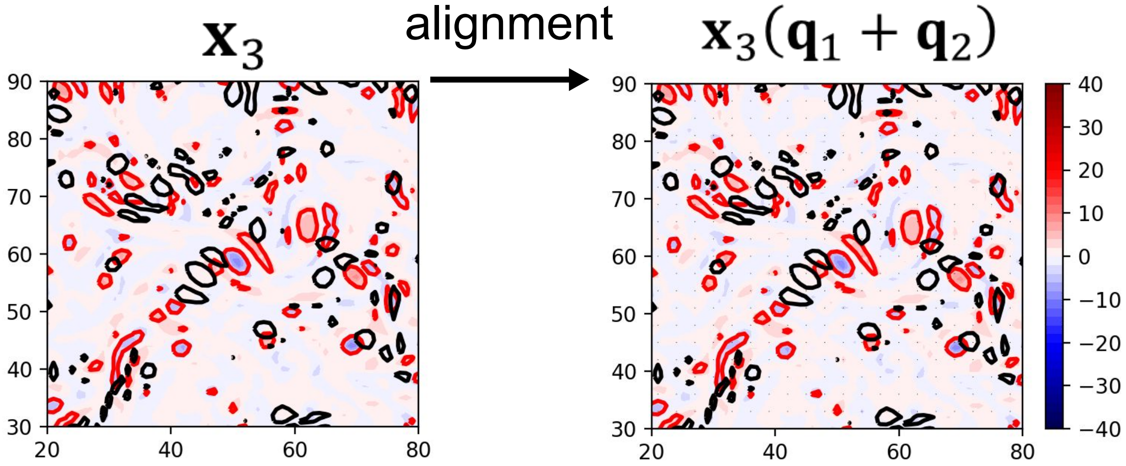
$$\mathbf{x} + \delta\mathbf{x} = \mathbf{x}_1 + \delta\mathbf{x}_1 + \mathbf{x}_2(\mathbf{q}_1) + \delta\mathbf{x}_2 + \mathbf{x}_3(\mathbf{q}_1 + \mathbf{q}_2) + \delta\mathbf{x}_3$$



Multiscale alignment approach

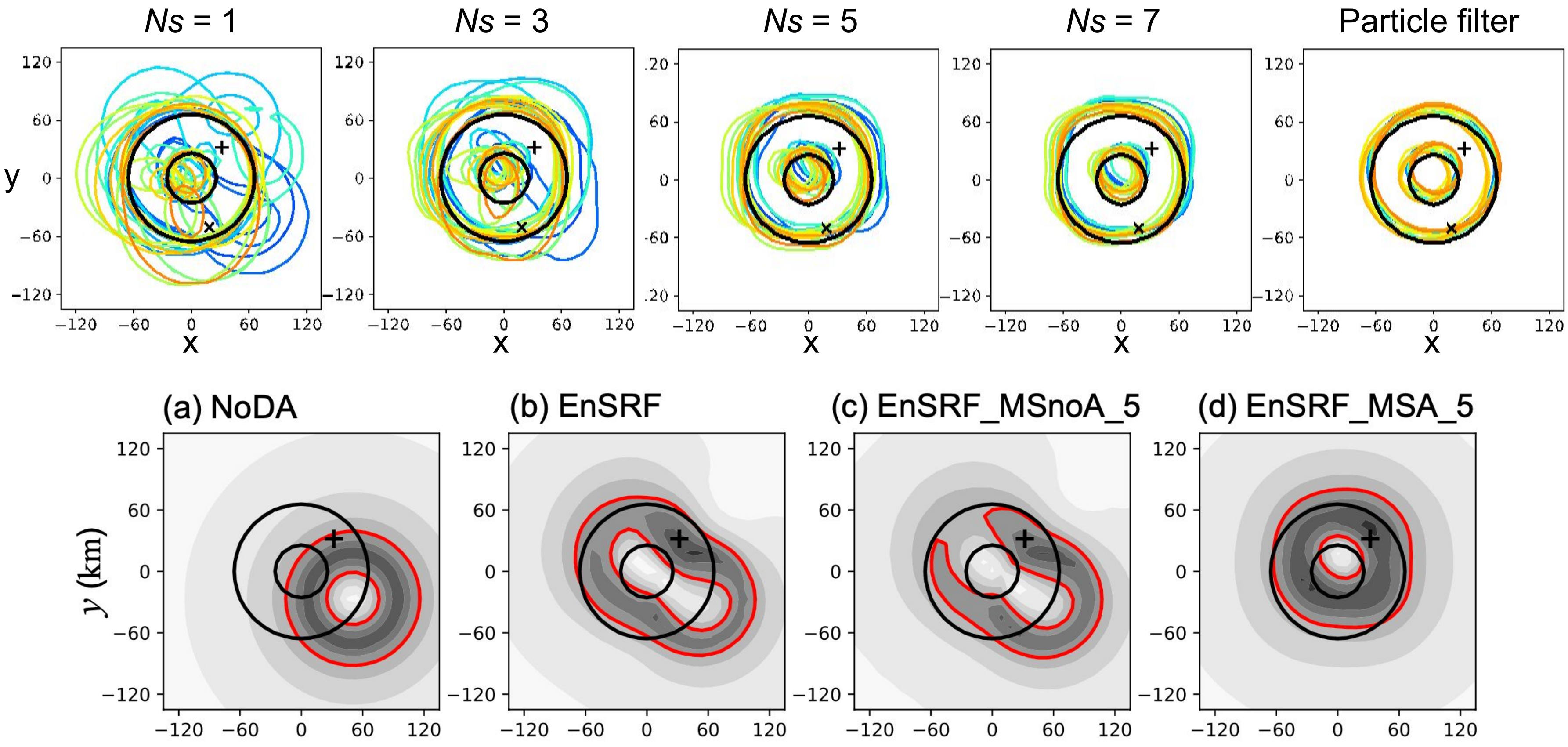
Ying 2019 Mon Wea Rev.

$$\mathbf{x} + \delta\mathbf{x} = \mathbf{x}_1 + \delta\mathbf{x}_1 + \mathbf{x}_2(\mathbf{q}_1) + \delta\mathbf{x}_2 + \mathbf{x}_3(\mathbf{q}_1 + \mathbf{q}_2) + \delta\mathbf{x}_3$$



Asymptotic behavior as *Ns* increases

Ying et al. 2023



Putting everything together…

for s in $1, \ldots, N_s$ **do**
  $\mathbf{x}^b_{n,s} = \mathsf{F}_s \mathbf{x}_n$
  $\mathbf{y}^b_n = h(\mathbf{x}_n)$
  if decompose_obs **then**
    $\mathbf{y}^o_s = \mathsf{F}^o_s \mathbf{y}^o$
    $\mathbf{y}^b_{n,s} = \mathsf{F}^o_s h(\mathbf{x}_n)$
    $\mathbf{x}^a_{n,s} = \mathbf{x}^b_{n,s} + \mathsf{L}_s \circ \frac{\mathrm{cov}(\mathbf{x}^b_s, \mathbf{y}^b_s)}{\mathrm{cov}(\mathbf{y}^b_s, \mathbf{y}^b_s) + \sigma^2_{o,s}\mathsf{I}} \left(\mathbf{y}^o_s - \mathbf{y}^b_{n,s}\right)$
  else
    $\mathbf{x}^a_{n,s} = \mathbf{x}^b_{n,s} + \mathsf{L}_s \circ \frac{\mathrm{cov}(\mathbf{x}^b_s, \mathbf{y}^b)}{\mathrm{cov}(\mathbf{y}^b, \mathbf{y}^b) + \sigma^2_{o}\mathsf{I}} \left(\mathbf{y}^o - \mathbf{y}^b_{n}\right)$
  end if
  if $s < N_s$ **then**
    $\mathbf{q}_{n,s} = \underset{\mathbf{q}}{\mathrm{argmin}} \left\|\mathbf{x}^b_{n,s}(\mathbf{q}) - \mathbf{x}^a_{n,s}\right\|^2 + w\left\|\nabla \mathbf{q}\right\|^2$
    $\mathbf{x}_n \leftarrow \mathbf{x}_n\left(\mathbf{q}_{n,s}\right) + \mathbf{x}^a_{n,s} - \mathbf{x}^b_{n,s}\left(\mathbf{q}_{n,s}\right)$
  else
    $\mathbf{x}_n \leftarrow \mathbf{x}_n + \mathbf{x}^a_{n,s} - \mathbf{x}^b_{n,s}$
  end if
end for

"Misc. transforms": scale components

"Assimilator":
EnKF algorithm

"Updator":
Alignment technique

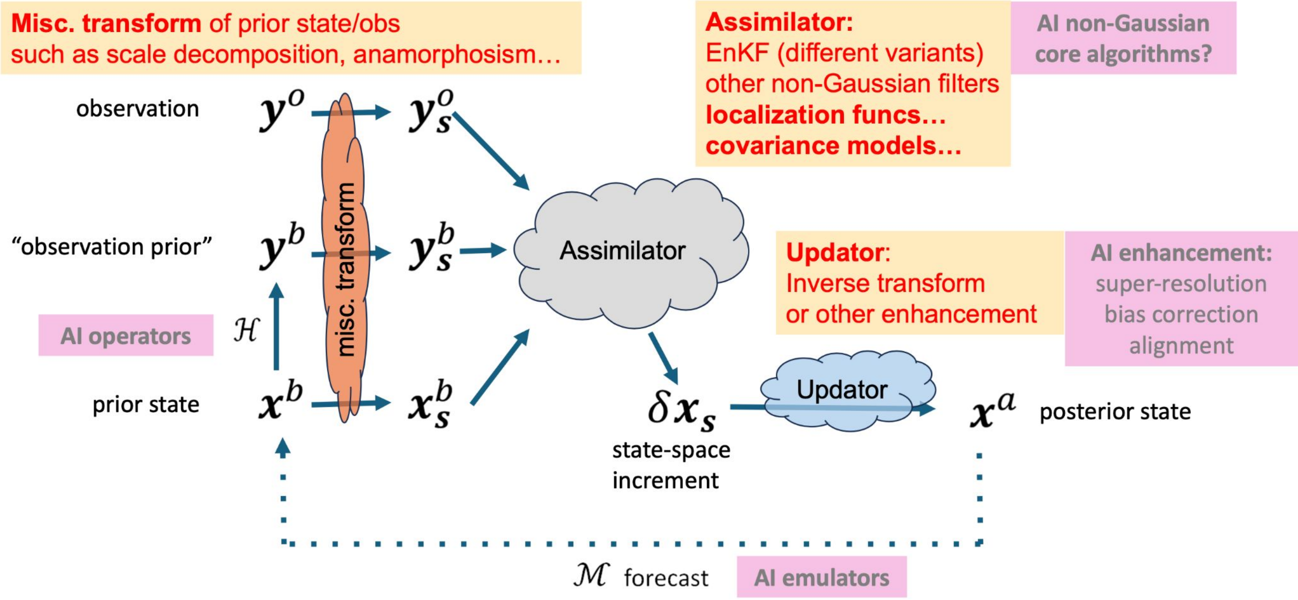
…or just additive

$n = 1, \ldots, N$ indexes ensemble members
$s = 1, \ldots, Ns$ indexes scale components (SC)

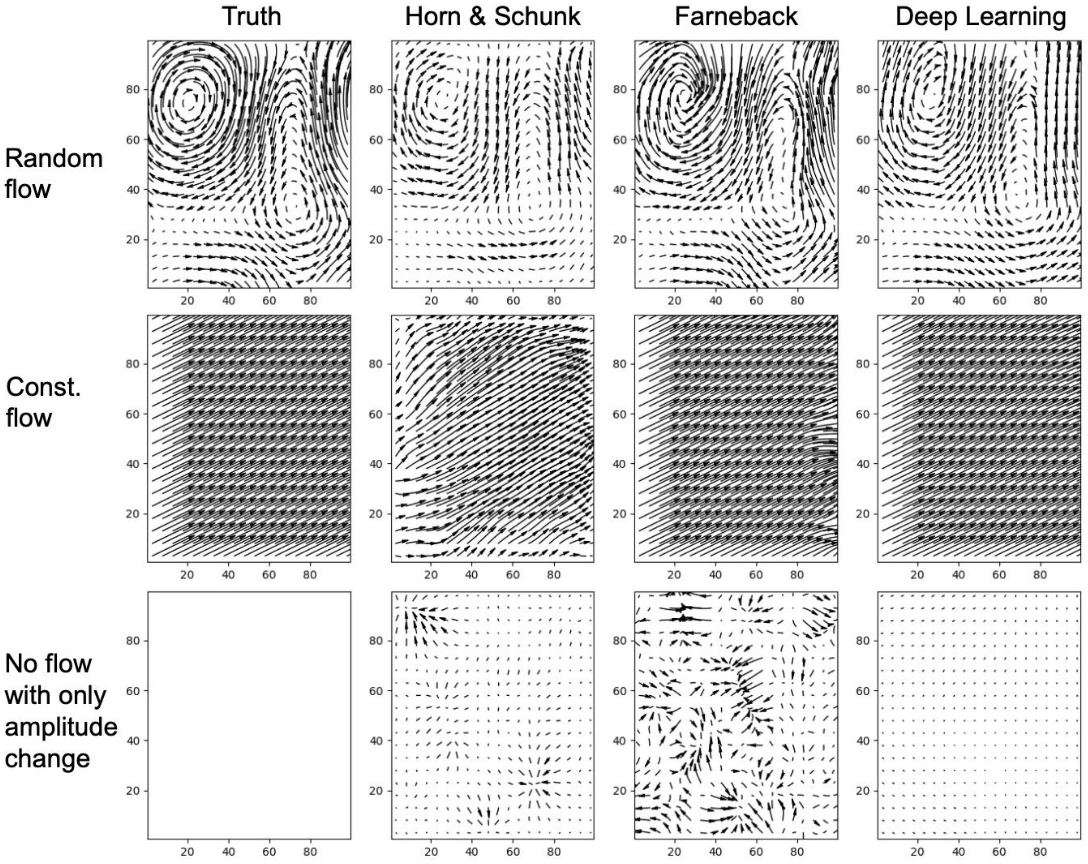
New software (written in Python)

Next-generation Ensemble Data Assimilation System

- Support many models at NERSC and beyond
- Test new ideas
- Enable AI integration (with python ecosystem)



AI enhancement in optical flow



Used in previous results:

- *Horn & Schunk 1980*

In OpenCV:

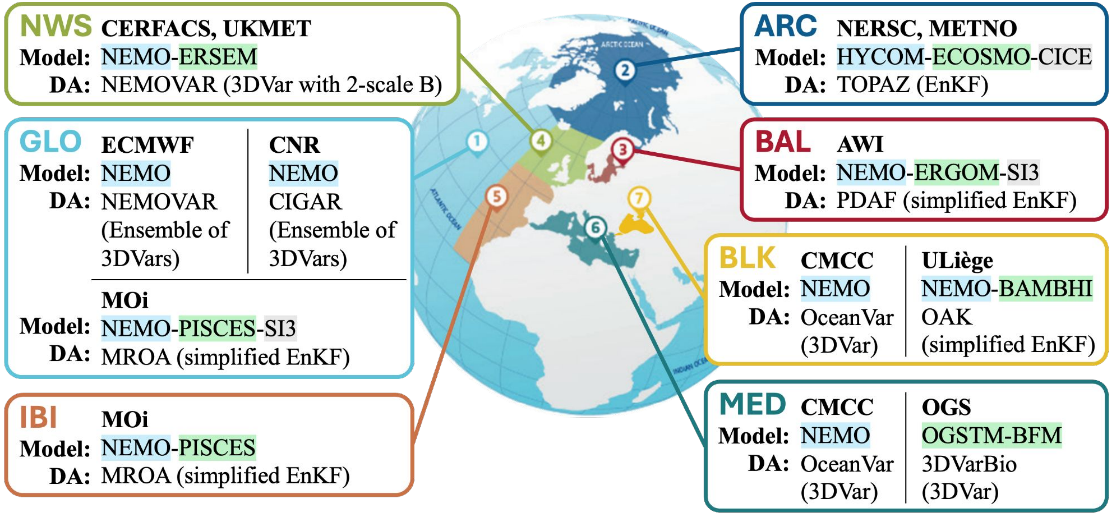
- *Farneback - dense parametric flow fields*
- *Deep Learning (DIS)*

Future work (new project)

COMEDI

COpernicus Marine service with Enhanced Data assimilation and improved Interoperability – funced by EU HORIZON

To improve the current DA practice in Monitoring and Forecasting Centres (MFCs)



COMEDI’s innovations

Implement correlated R for high-resolution SWOT data

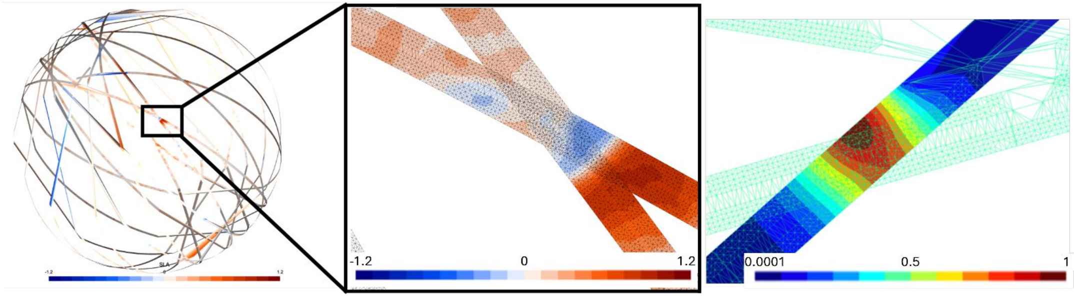


Fig. 1.10 *A 2D mesh constructed from sample SWOT data (subsampled to 20 km): (a) global SLA data for 1 day; (b) a zoom of the highlighted region; (c) an example of a diffusion-modelled horizontal error correlation function for simulated SWOT data on a similar mesh.*

Test non-Gaussian methods for ocean and sea ice DA

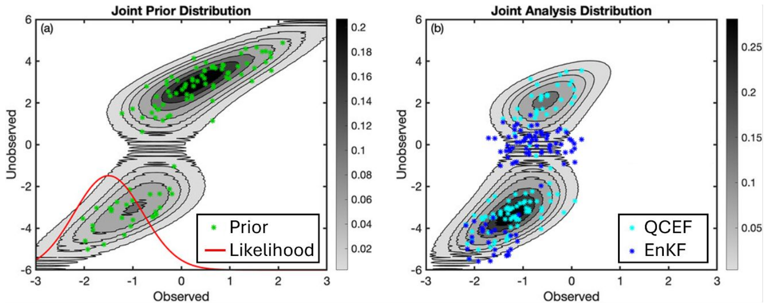


Fig. 1.7 *(a) Shading shows the prior non-Gaussian distribution, red curves shows the observation likelihood, green asterisks show a 100-member prior ensemble; (b) shading shows the true analytic posterior distribution, dark blue asterisks are updated members from EnKF, and light blue asterisks are those from QCEF. [from Anderson 2023]*

COMEDI's innovations

Use AI as nonlinear core DA algorithm – AI Joint Development Team

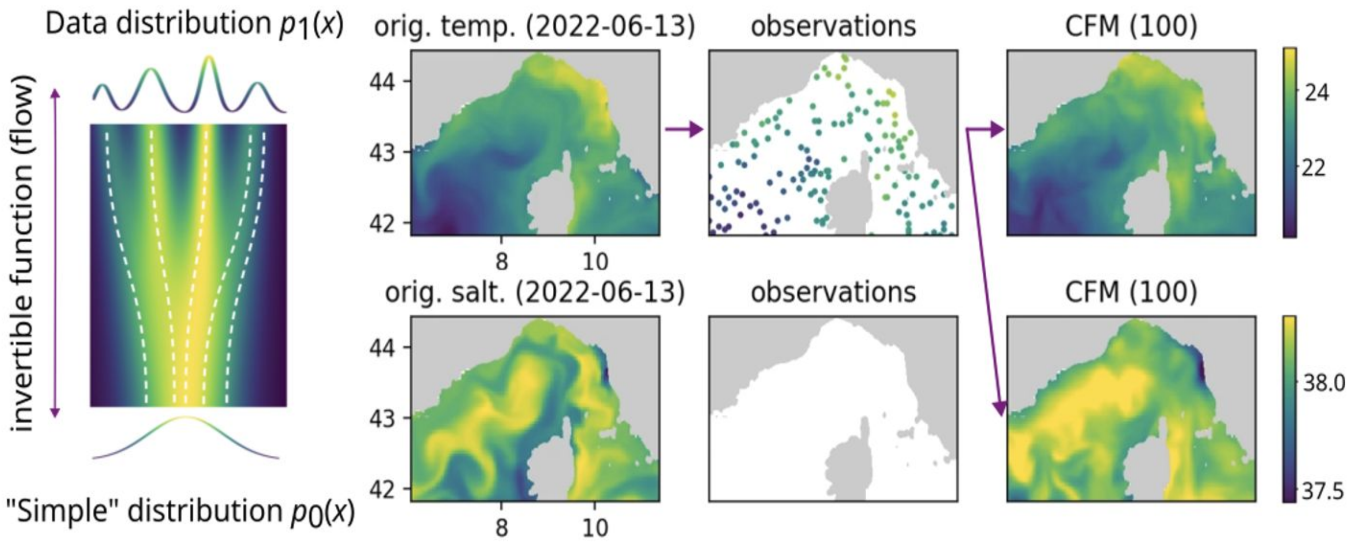


***Fig. 1.8** Conditional flow matching (CFM), and its application to inference of ocean temperature and salinity from temperature observations. The original temperature and salinity were not used during training of the CFM model. [unpublished results]*

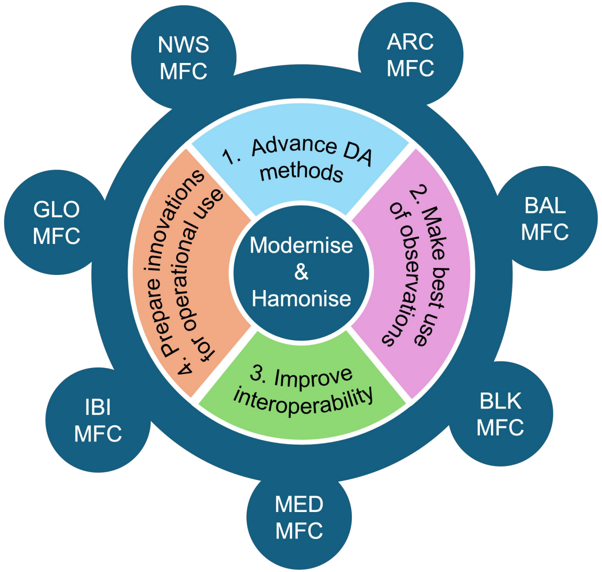
COMEDI's innovations

Interoperability of tools among MFCs

COMEDI innovations (categorised by their sharing paths)			ARC	BAL	BLK	GLO	IBI	MED	NWS	CIGAR	NEMOVAR	OceanVar	E-Med-BGC	MROA	NEDAS	OAK	PDAF
			MFCs							DA software: Variational			Ensemble-based				
Path 1	Core assim. method	Non-Gaussian filtering	X	X											X		X
Path 1	Core assim. method	AI assimilation core	X		X			X				X			X	X	
Path 1	Core assim. method	Automated tuning						X					X				X
Path 1	Core assim. method	Higher-order sampling						X					X				X
Path 2		AI super resolution						X					X				
Path 2		AI ensemble generation	X		X	X	X	X				X		X	X		
Path 2		Upgraded stochastic physics				X			X	X	X						
Path 2		AI bias correction				X				X							
Path 2		Mechanistic perturbation			X	X		X					X	X		X	
Path 2		Correlated **R**	X	X	X				X		X	X			X		X
Path 2	New H	Sentinel 3-NG WiSA							X		X						
Path 2	New H	CIMR Tb	X												X		
Path 2	New H	Sentinel 3 Reflectance			X	X		X					X	X		X	
Path 3		Dynamic ensemble	X	X	X	X	X	X	X	X	X		X	X	X	X	X
Path 3		Multiscale **B**		X		X	X		X		X			X			X
Path 3		Hybrid **B**			X	X		X	X	X	X	X					
Path 3		FSOI and its emulation	X			X					X				X		

Color code: (blue) an older version or different approach already exists; (dark green) developed and implemented in WPs 2-6; (light green) shared via a module / existing interface; (yellow) available via an interface developed in WP6; X: demonstrated in WP5 or WP7

COMEDI will start in Sep 2026…



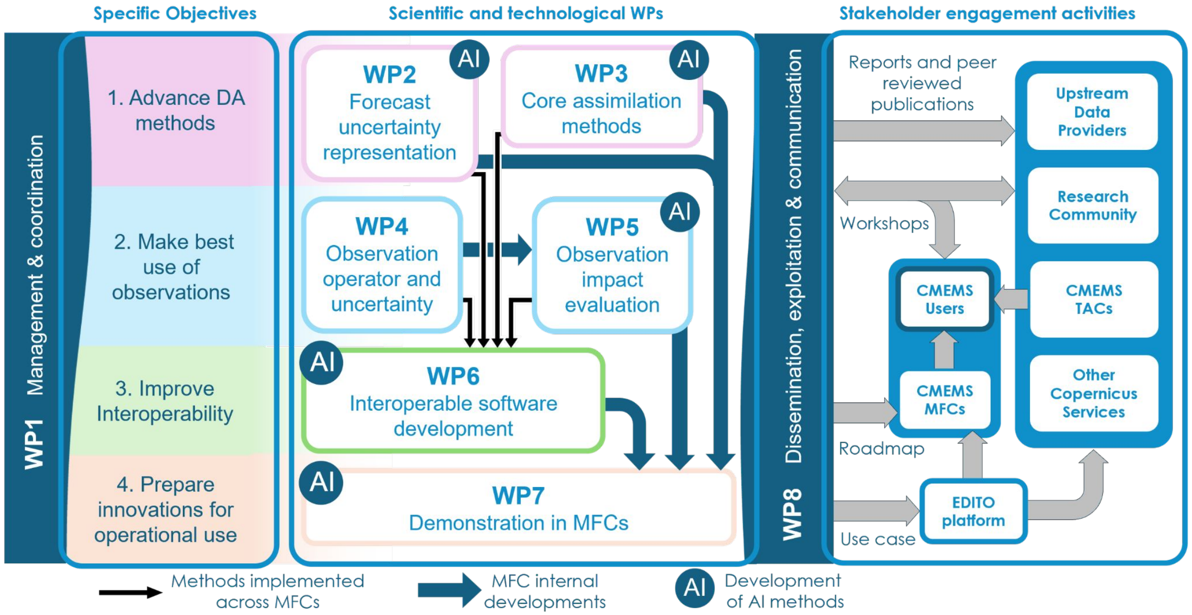


Fig. 1.4 COMEDI specific objectives, proposed work packages, and stakeholder engagement activities

References

Evensen et al. 2022 Data assimilation fundamentals.
https://link.springer.com/book/10.1007/978-3-030-96709-3

Ying et al. 2018 On the selection of localization radius in ensemble filtering for multiscale quasi-geostrophic dynamics. https://doi.org/10.1175/MWR-D-17-0336.1

Ying 2019 A multiscale alignment method for ensemble filtering with displacement errors.
https://doi.org/10.1175/MWR-D-19-0170.1

Ying 2020 Assimilating observations with spatially correlated errors using a serial ensemble filter with a multiscale approach. https://doi.org/10.1175/MWR-D-19-0387.1

Ying et al. 2023 Improving vortex position accuracy with a new multiscale alignment ensemble filter.
https://doi.org/10.1175/MWR-D-22-0140.1