Storm-Scale Weather Analysis and Prediction at the NOAA National Severe Storms Laboratory Using a Localized Particle Filter

Jon Poterjoy and Louis Wicker

Cooperative Institute for Mesoscale Meteorological Studies
University of Oklahoma
NOAA/OAR/National Severe Storms Laboratory

Monday, 27^{th} February 2017

Introduction

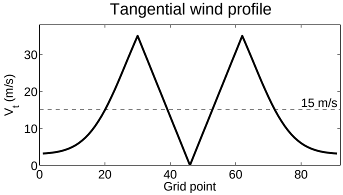
Data assimilation at the convective-scale is very difficult!

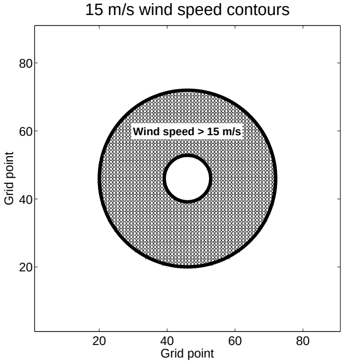
Challenges relevant to this talk:

1. Displacement errors in storms and clouds lead to non-Gaussian prior pdfs.
2. Cloud processes are very nonlinear.
3. Measurements often relate nonlinearly to model state.
4. Storm-environment interactions are likely nonlinear.

See van Lier-Walqui et al. (2012) and Posselt (2016) for examples.

Example of problem 1

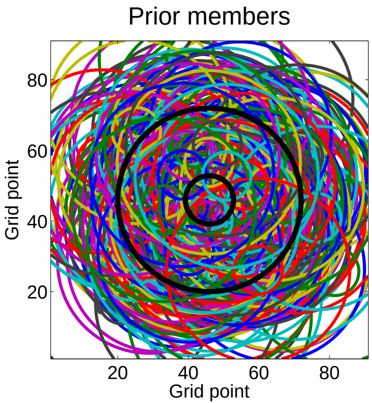


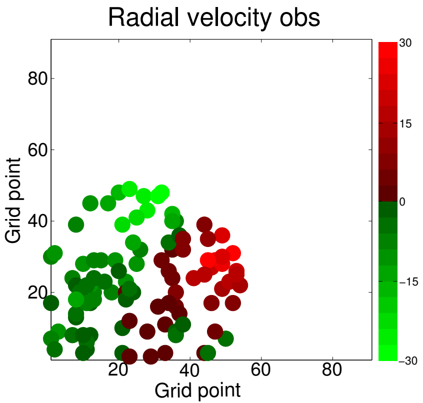


- A Rankine vortex reproduces the data assimilation problem posed by displacement errors.
- The 1-D wind profile (top panel) is interpolated spatially to produce a 2-D wind field (bottom panel).

$$\text{Model state: } \mathbf{x} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}$$

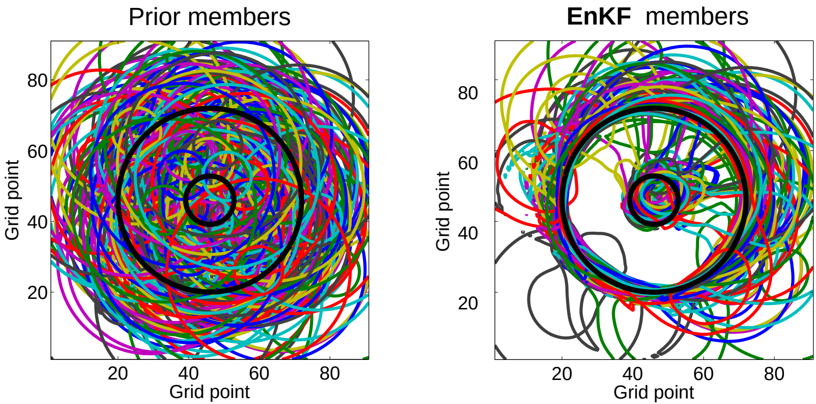
Experiments with Rankine vortex





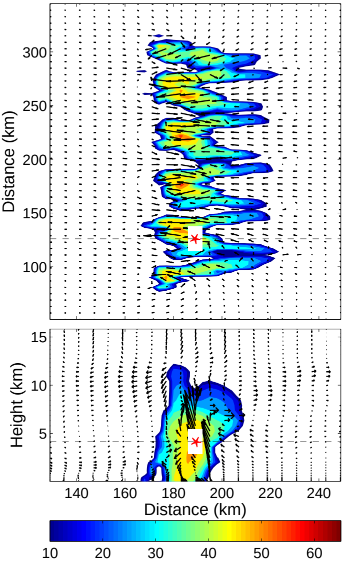
- **Prior**: 200 vortices with identical structure, but position error.
- **Observations** radar radial wind measurements, observing part of the vortex.

Experiments with Rankine vortex



- Linear/Gaussian methods are suboptimal when position error is comparable or larger than vortex size (Lawson and Hansen 2005; Chen and Snyder 2007).

Example of problems 2 and 3

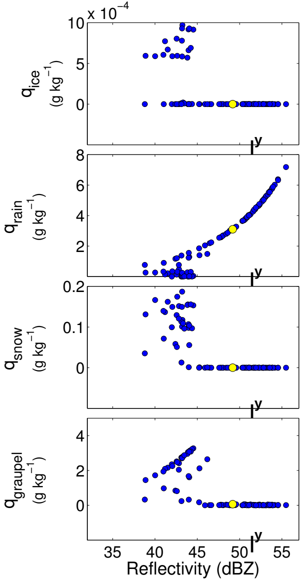


- **Prior**: 100-member ensemble forecasts for idealized squall line.
- **Observation:** Radar reflectivity measurement at $\star$.

Top and bottom panels show cross sections through true storm at observation location.

Reflectivity and storm-relative winds are plotted.

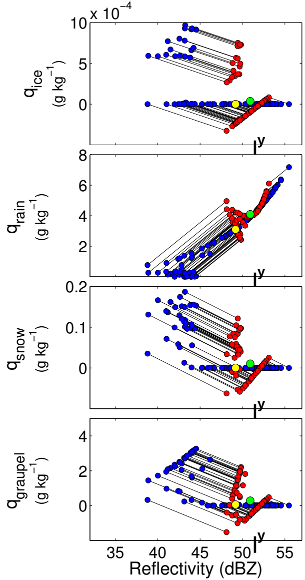
Prior ensemble at ⋆



$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$$

- **Blue markers**: prior samples from joint probability distribution of reflectivity and microphysics variables
- **Yellow markers**: true state
- **Black tickmarks**: observed reflectivity

EnKF update at ⋆



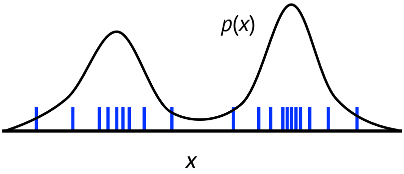
$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$$

- **Blue markers**: prior samples from joint probability distribution of reflectivity and microphysics variables
- **Yellow markers**: true state
- **Black tickmarks**: observed reflectivity
- **Red markers**: posterior samples
- **Green markers**: posterior mean

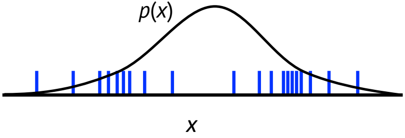
What about particle filters?

- **Particle filters (PFs)** use ensemble members, or “particles,” to form a general representation of probabilities.
- Unlike EnKFs, PFs do not approximate pdfs using Gaussians.

Samples drawn from a non-Gaussian pdf for x



Gaussian representation from samples

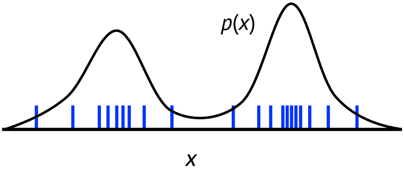


SIR PF

- Sequential importance resampling (Gordon et al. 1993) provides framework for technique used in this study.

SIR PF

STEP 1:



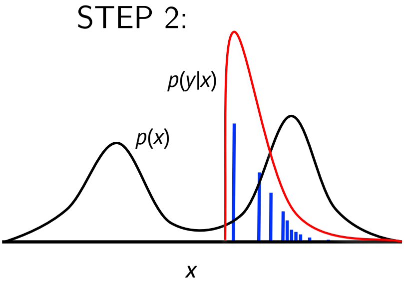
Prior pdf is estimated empirically by a sum of delta functions:

$$p(x) \approx \frac{1}{N_e}\sum_{n=1}^{N_e}\delta(x-x_n).$$

SIR PF

- Sequential importance resampling (Gordon et al. 1993) provides framework for technique used in this study.

SIR PF



Posterior pdf is estimated by sum of weighted delta functions:

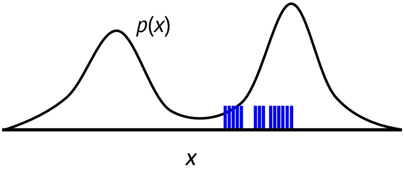
$$\begin{aligned} p(x|y) &\approx \sum_{n=1}^{N_e} w_n \delta(x - x_n), \\ w_n &\propto p(y|x_n). \end{aligned}$$

SIR PF

- Sequential importance resampling (Gordon et al. 1993) provides framework for technique used in this study.

SIR PF

STEP 3:



Particles are drawn from $p(x|y)$ by sampling with replacement based on $\{x_n, w_n\}$, n $= 1,2,...,N_e$

Localized SIR PF

The **Poterjoy (2016)** local PF:

- calculates weights for each state variable i in $\mathbf{x}$:
  $w_{n,i} \propto \prod_{j=1}^{N_y} \{[p(y_j|\mathbf{x}_n) - 1]c_{j,i}] + 1\}$, where $c_{j,i} = f(r, y_j, x_i)$.
- processes observations serially to merge sampled particles from SIR PF step with prior particles.

Advantages: computationally inexpensive and resembles SIR PF as $r \to \infty$.

Disadvantage: a sampling step is needed for each observation in sequence.

Localized SIR PF

Poterjoy and Anderson (2016) show that the local PF operates effectively for high-dimensional geophysical systems (tests using dry, coarse resolution GCM).

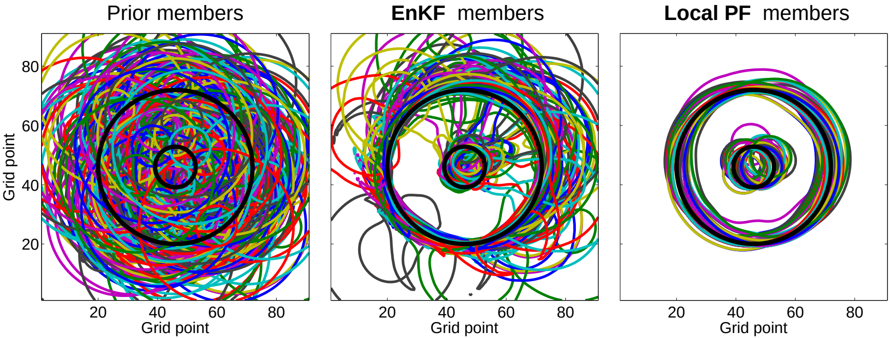
Poterjoy, Sobash, and Anderson (accepted) demonstrate advantages over EAKF for idealized convective-scale application in NCAR Weather Research and Forecasting (WRF) model.

Other localized PF-type methods

Not a new idea:

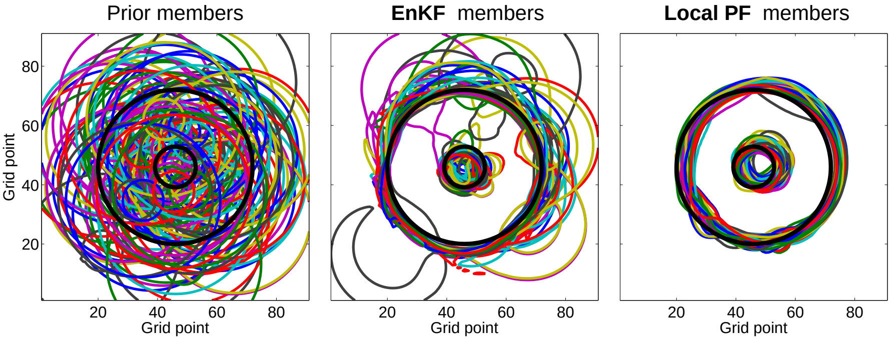
- **Bengtsson et al. (2003):** "Local local ensemble filter"
- **Lei and Bickel (2009):** "Moment matching ensemble filter"
- **And many others:** Chen and Reich (2015), Tödter et al. (2015), Chustagulprom et al. (2016), Penny and Miyoshi (2016), Lee and Majda (2016), Robert and Künsch (2017),...

The local PF: Problem 1



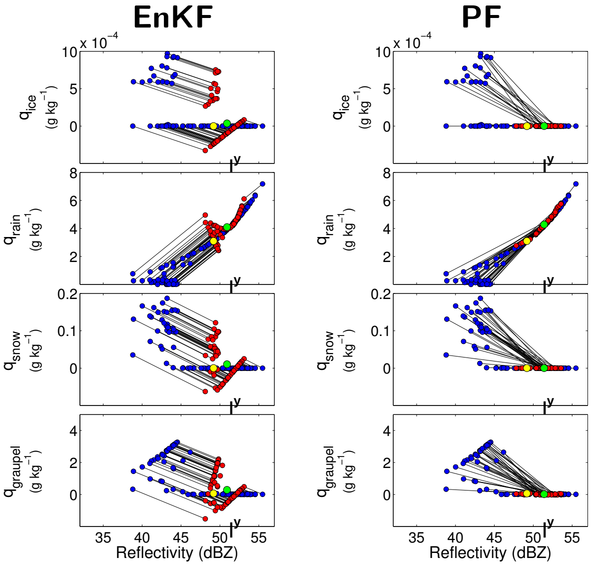
- 200-member local PF maintains vortex symmetry better than EnKF.
- Non-localized SIR PF requires over $1x10^4$ particles to prevent weight collapse.

The local PF: Problem 1



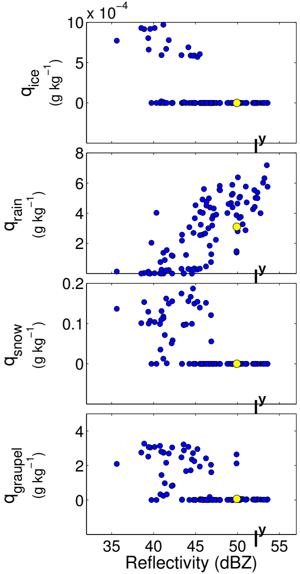
- 80-member local PF requires shorter correlation length scale for localization, which degrades vortex structure.
- Ensemble size comparable to operational numerical weather prediction centers.

The local PF: Problems 2 and 3



- PF provides more physically meaningful representation of posterior pdf for cloud variables.

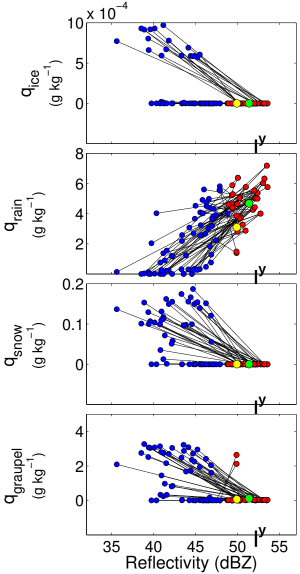
Prior with ob located 2 km lower



$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$$

- **Blue markers**: prior samples from joint probability distribution of reflectivity and microphysics variables
- **Yellow markers**: true state
- **Black tickmarks**: observed reflectivity

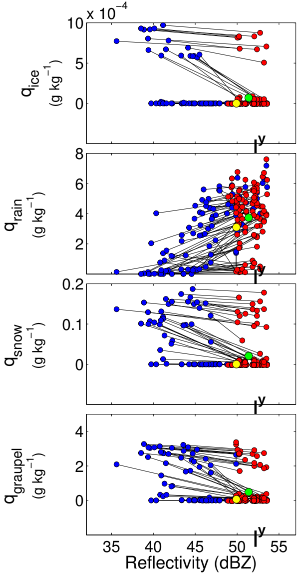
PF update



$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$$

- **Blue markers**: prior samples from joint probability distribution of reflectivity and microphysics variables
- **Yellow markers**: true state
- **Black tickmarks**: observed reflectivity
- **Red markers**: posterior samples
- **Green markers**: posterior mean

Local PF update



$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$$

- **Blue markers**: prior samples from joint probability distribution of reflectivity and microphysics variables
- **Yellow markers**: true state
- **Black tickmarks**: observed reflectivity
- **Red markers**: posterior samples
- **Green markers**: posterior mean

Research question

Can current convective-scale NWP efforts benefit from the local PF?

YES

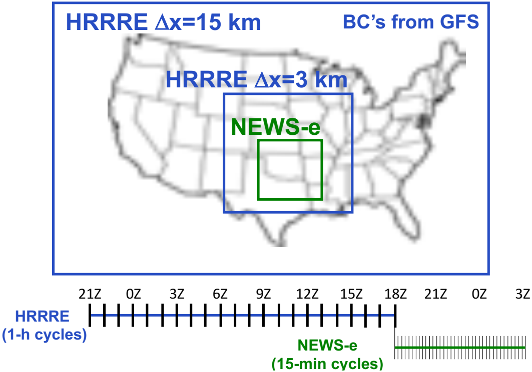
- Dynamical processes are very nonlinear.
- Measurements often relate nonlinearly to the model state or have non-Gaussian errors.

NO

- Models errors are still too much of a problem.
- Degrees of freedom are too large for current ensemble sizes.

Convective-scale forecasting at NOAA NSSL

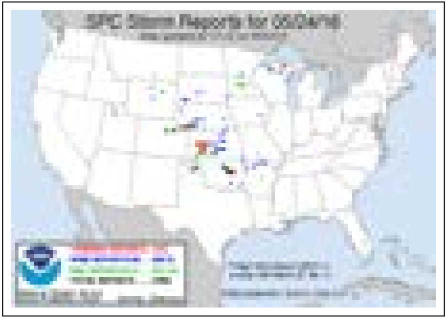
NSSL Experimental Warn-on-Forecast System for Ensembles (NEWS-e)



- **Model**: 3-km grid spacing NCAR WRF model.
- **Observations**: Radar, satellite, and surface data every 15-min
- **Data assimilation**: 36-member NCAR Data Assimilation Research Testbed (DART) EAKF

First test of local PF

24 May 2016 tornado outbreak



Loading video...

- Local PF is compared with EAKF for complex case where convective cells grow upscale into MCS.

90-min storm evolution from 00 UTC

EAKF members

(Loading video...)

- Composite reflectivity for first 10 members.
- Forecasts initialized at 00 UTC 25 May.
- 35 dBZ contour of observed composite reflectivity shown for reference.

Local PF members

(Loading video...)

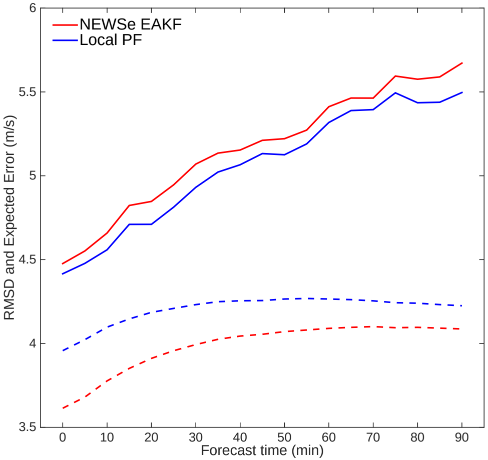
- Composite reflectivity for first 10 members.
- Forecasts initialized at 00 UTC 25 May.
- 35 dBZ contour of observed composite reflectivity shown for reference.

Forecast verification using radar winds

(Loading video...)

- Circles show absolute value of V_r observations - ensemble mean from forecasts initialized at 00 UTC 25 May.
- Contour of ensemble mean 20 dBZ composite reflectivity shown for reference.

Local PF vs. EAKF

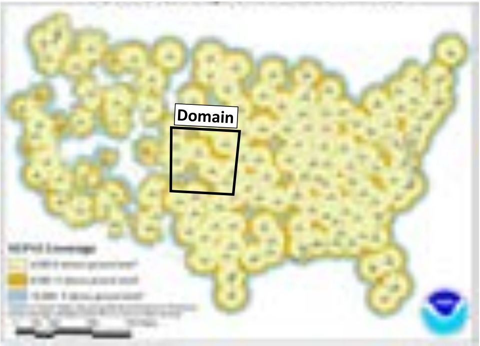


- Volume mean V_r RMSDs (solid lines) and expected error from ensemble and observation errors (dashed lines).
- Values averaged from forecasts initialized every 30 min from 2230 UTC to 0300 UTC.

More extensive obs coverage for verification

Radar coverage in Domain

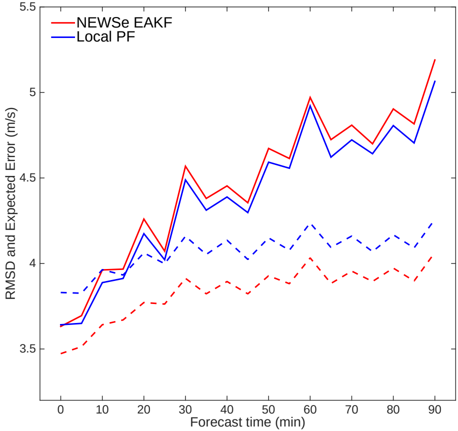
US NEXRAD Coverage



- Multiple radars in domain used to estimate forecast errors.
- Example shows radar obs within 5 min of 0000 UTC 24 May

More extensive obs coverage for verification

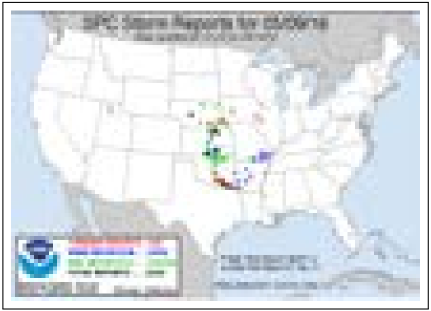
Multiple radar obs included



- Verification considers other storms in domain.
- Observation coverage changes with time because different scan volumes make it into each 5-min window.

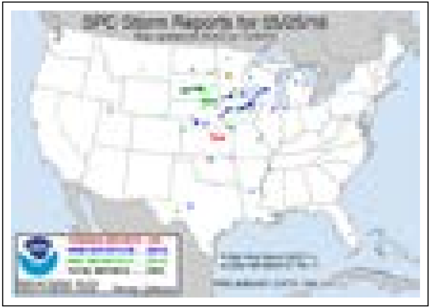
Additional cases

Case 2: 9 May 2016



- Storms grow upscale into organized MCSs.

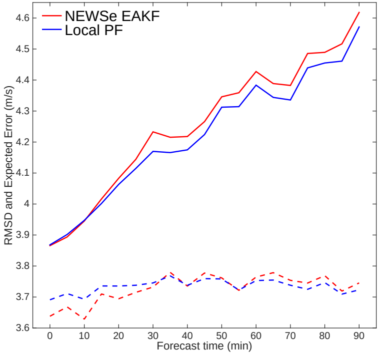
Case 3: 25 May 2016



- Long-lived tornadic supercell across domain.

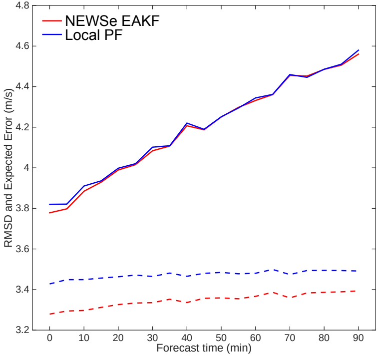
EAKF vs. Local PF

Case 2: 9 May 2016



- Storms grow upscale into organized MCSs.

Case 3: 25 May 2016



- Long-lived tornadic supercell across domain.

Discussion

A localized PF is tested in the NOAA National Severe Storms Laboratory realtime ensemble prediction system.

Despite large sampling error caused by small ensembles and an imperfect forecast model, the local PF provides comparable results to a thoroughly tested EAKF data assimilation system.

Other localized PF-based methods may show potential for data assimilation problems of this type; e.g., Walter Acevedo's talk tomorrow and current work by Robert and Künsch with hybrid EnKF/PF.

Looking forward

The flexibility provided by PFs allows for a less restricted treatment of observations than current framework.

- Observing systems are designed for linear/Gaussian data assimilation methods.
- Using non-Gaussian observation errors is trivial in PF framework.
- Raw observations can be used without rigorous post processing (extreme example: radar velocity unfolding).