Ensemble Sensitivity Analysis: Applications and Challenges

Josh Hacker
National Center for Atmospheric Research, Research Applications Lab
Lili Lei
CIRES Climate Diagnostics Center, University of Colorado, and Physical Sciences Division, NOAA/Earth System Research Laboratory
Hank Chilcoat, Sean Wile, Paul Homan
USAF, various posts

NCAR

Ensemble sensitivity analysis (ESA)

NCAR

How does the change in a set of initial state variables x_s change a forecast metric J?

$$\frac{\partial J_e}{\partial x^a}$$

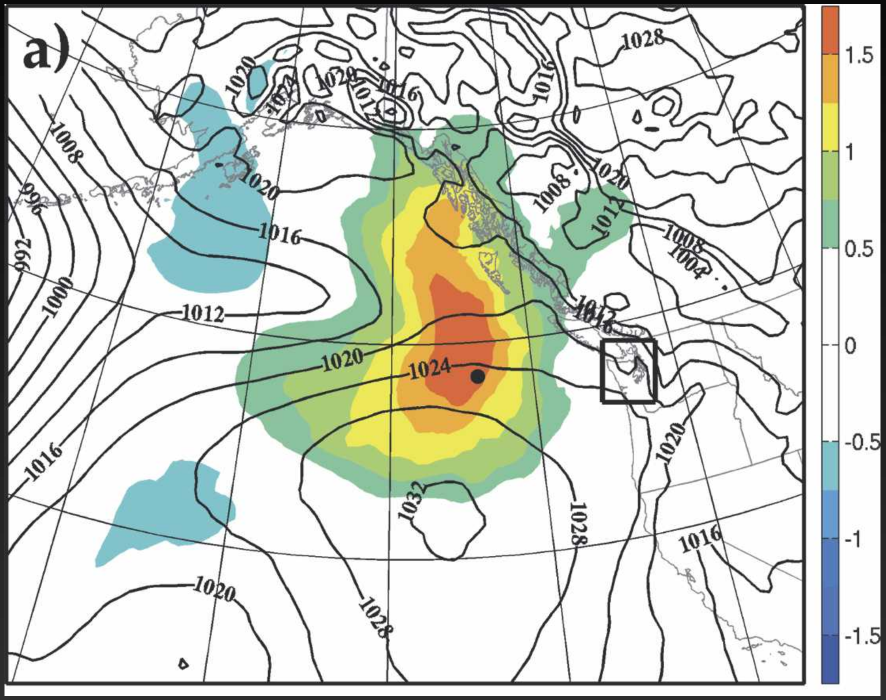
- Identify dynamically relevant covariance structures in space and time
- Propose observing strategies for mesoscale, short-range forecasts in complex terrain
- Sensitivity scales (time and space) to infer predictability scales
- Predictability of specific phenomena
- Open issues:
  - Sampling error
  - Linearity assumptions in complex terrain

Ensemble Sensitivity Background

NCAR

- Ancell and Hakim (2007) showed theoretical equivalence between adjoint and ensemble sensitivity for linear perturbations and Gaussian statistics
- Relies on linearization about an ensemble-mean trajectory
- Rigorous application has so far been limited to large-scale (smooth) and integrated processes where strong linear relationships are more likely

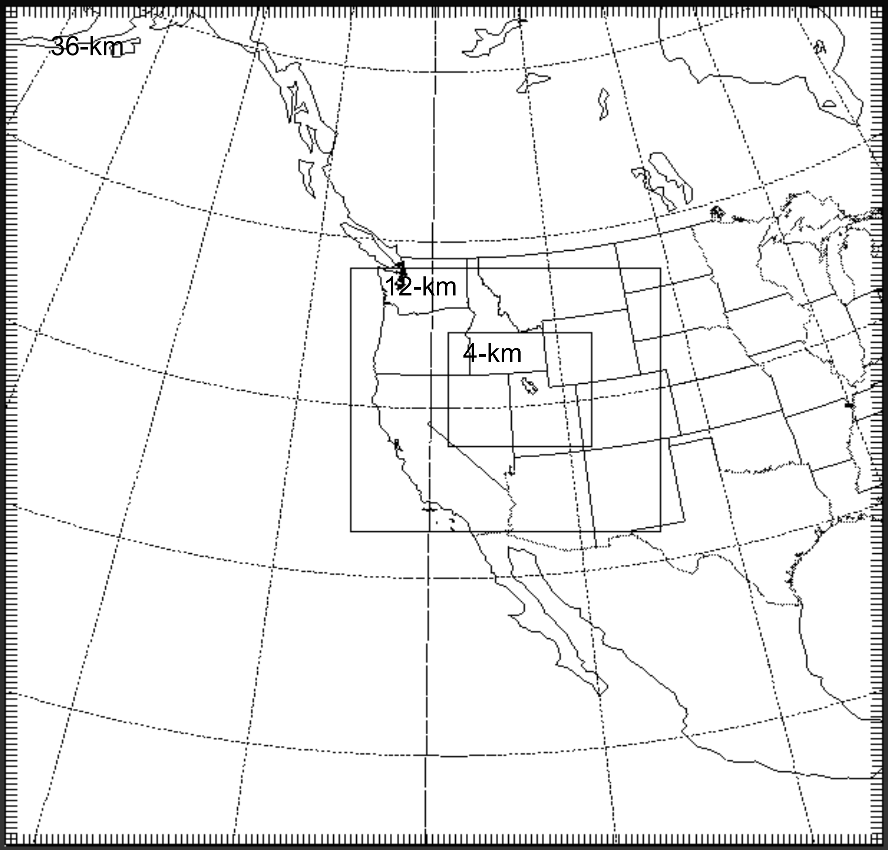
An optimal ensemble data assimilation system provides an appropriate sample



Sensitivity of 24-h sea-level pressure (SLP) over western Washington to SLP initial conditions, and ensemble-mean SLP (from Torn and Hakim 2008).

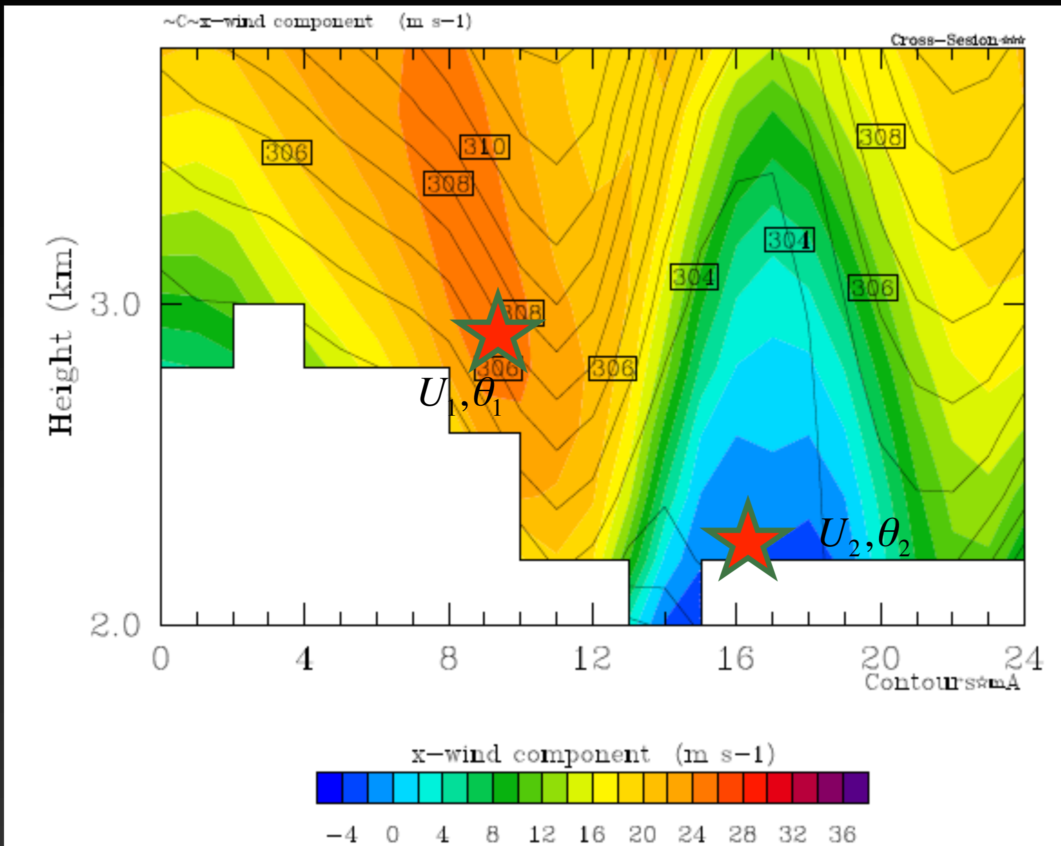
Experiment framework

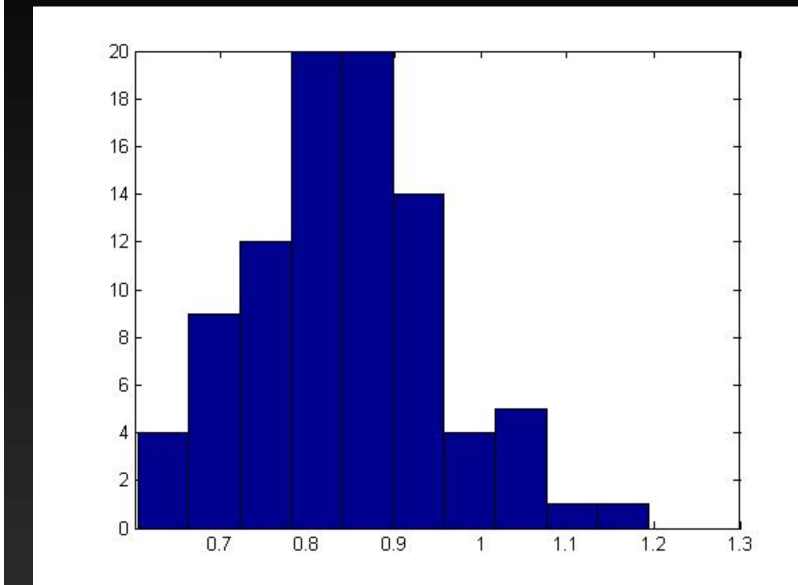
- 96-member ensemble data assimilation with the Data Assimilation Research Testbed (DART)
- Weather Research and Forecast (WRF) model
- Synthetic observations identical to rawinsonde network and surface altimeter
- 3-h cycling during Jan. 2009



Downslope winds at CO Springs

$$J_e = R_{ihB} = \frac{g}{\theta_2}\frac{(\theta_1-\theta_2)\Delta z}{(U_1-U_2)^2}$$

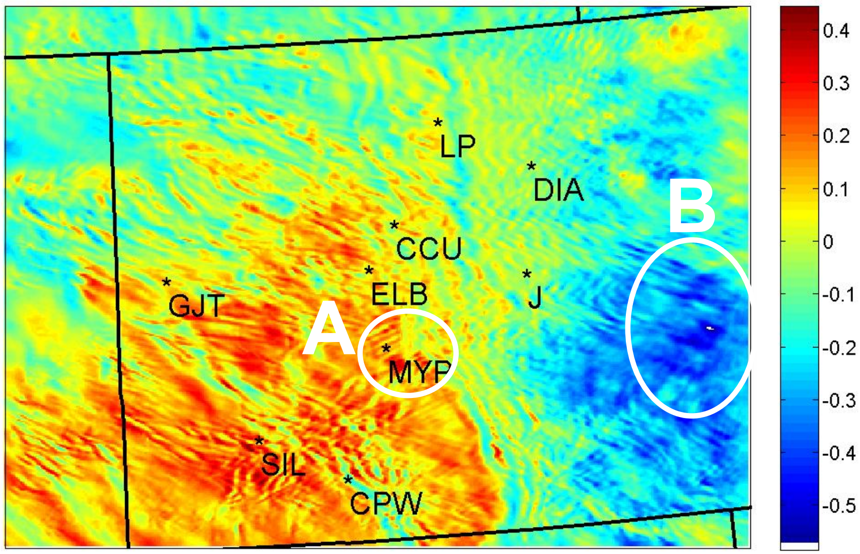




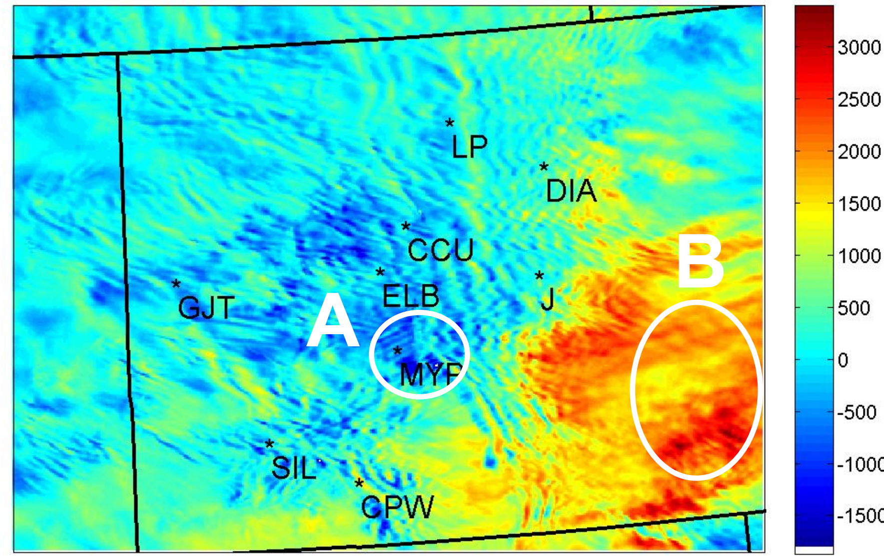
- Cross section looking north at U wind (shaded) and potential temperature; 3-h Ensemble mean forecast valid 30 Dec 2008 03Z
- J – analogous to the Bulk Richardson number to measures ratio of stability to shear across flow separation boundary
- Histogram of J (right) showing Gaussian distribution for metric

Predictors for wind storm

NCAR



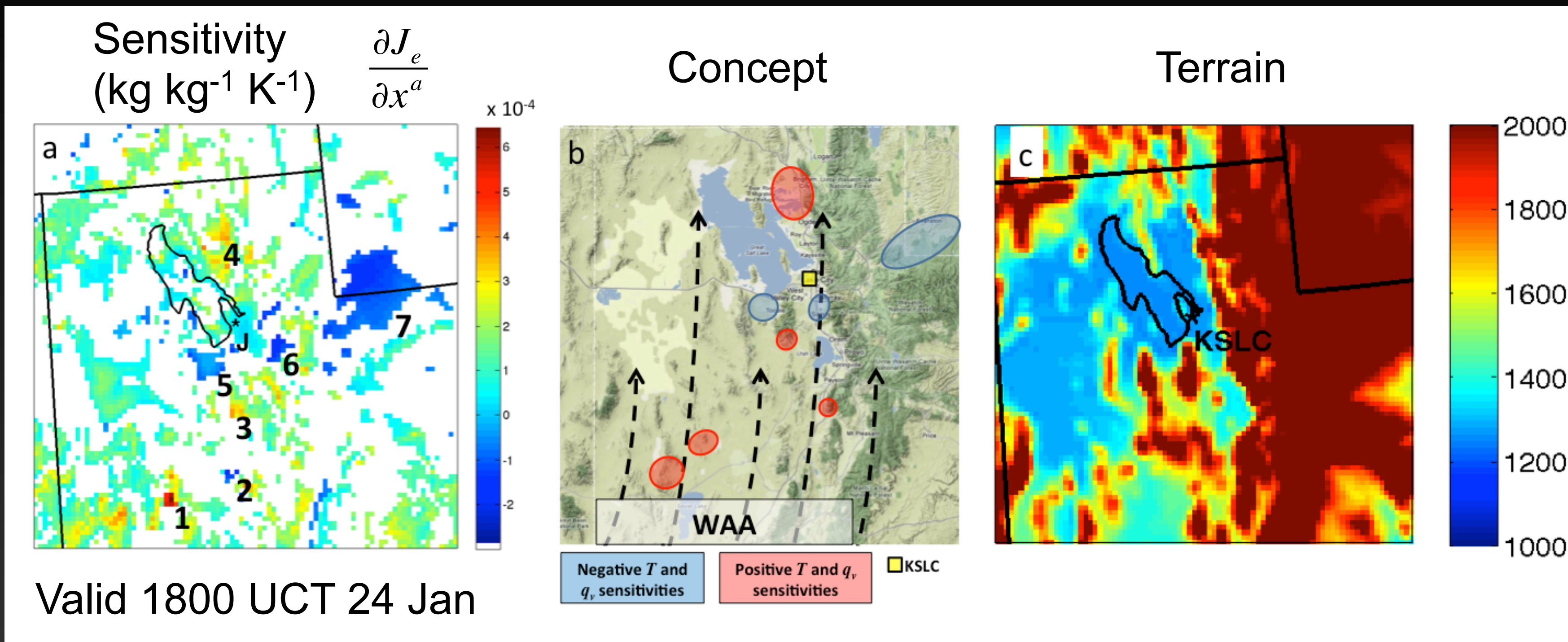
θ at model layer 14



Q_v at model layer 14

- Sensitivity of (dJ/dx) for 3-hr θ (left) and Q_v (right) at model level 14
- Strong dual sensitivities shown in both variables over plains and mountains
- Hypothesis – region A related to forcing and shear term in J, region B related to air mass characteristics over plains and stability term in J
- Good candidates for perturbations of IC for a new ensemble run

Moisture sensitivity to temperature

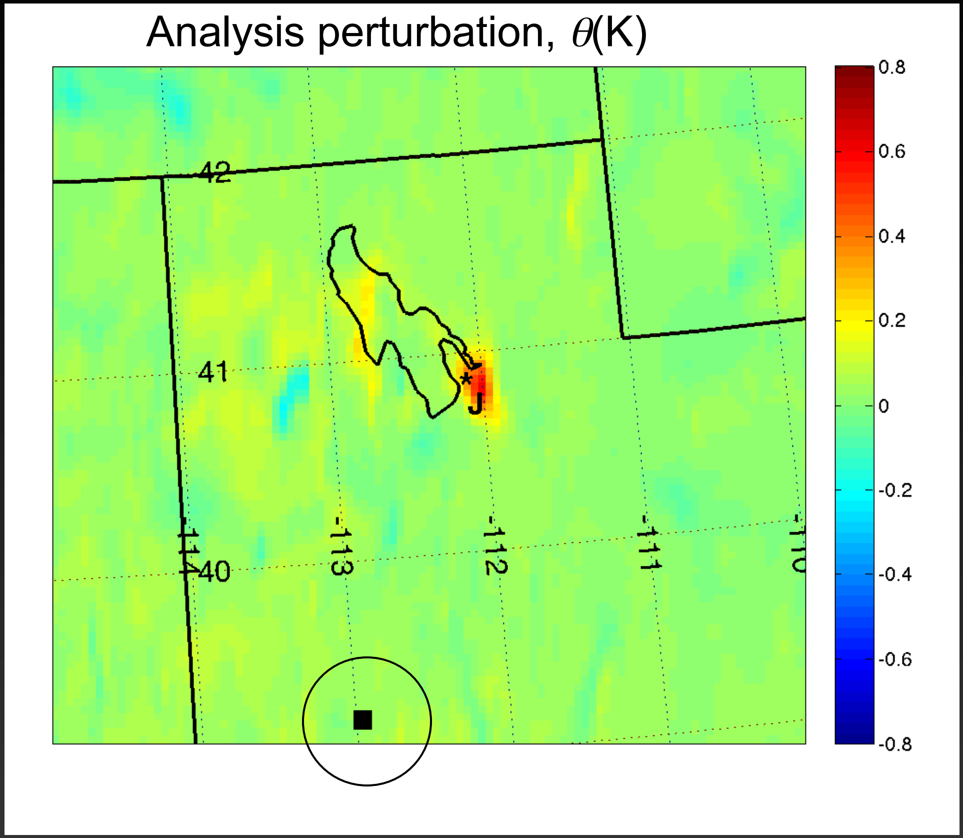


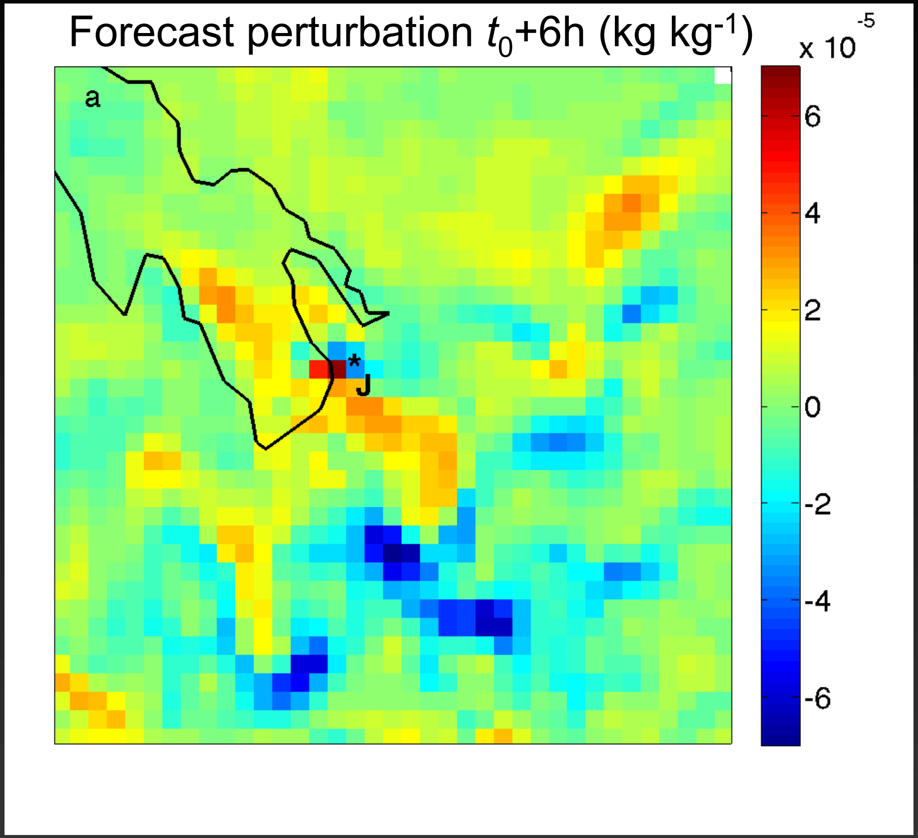
J = 2x2x2 box-mean water vapor mixing ratio over Salt Lake City airport

x=Potential temperature (here on model first layer)

Perturbation experiments

NCAR

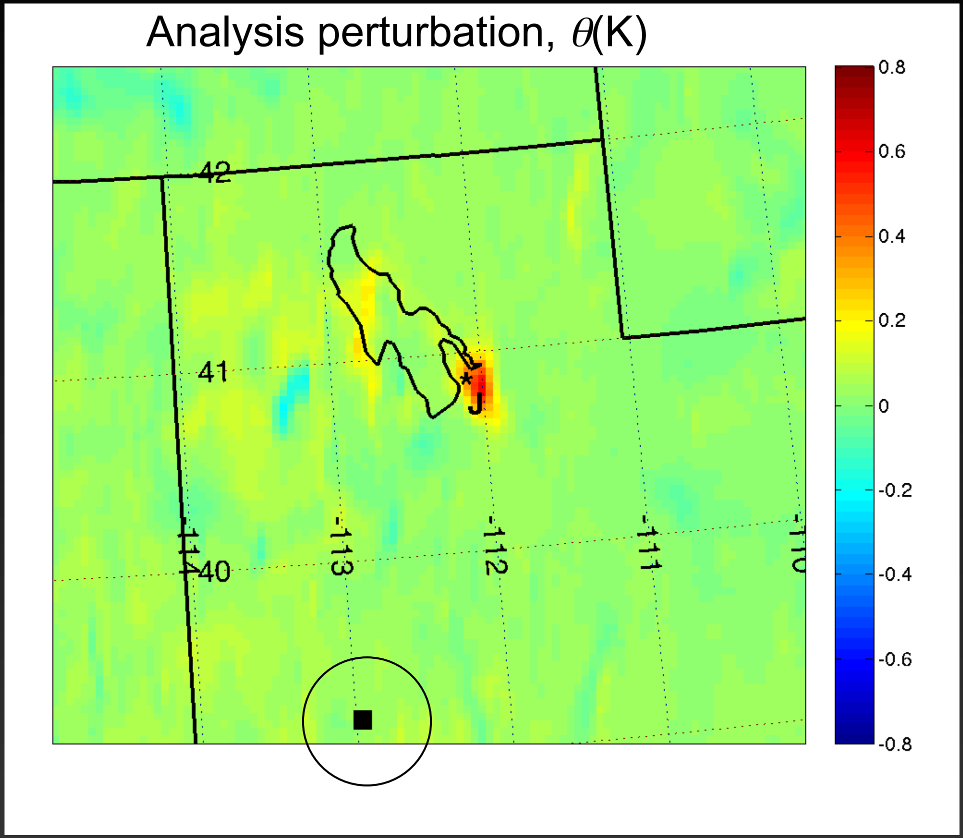


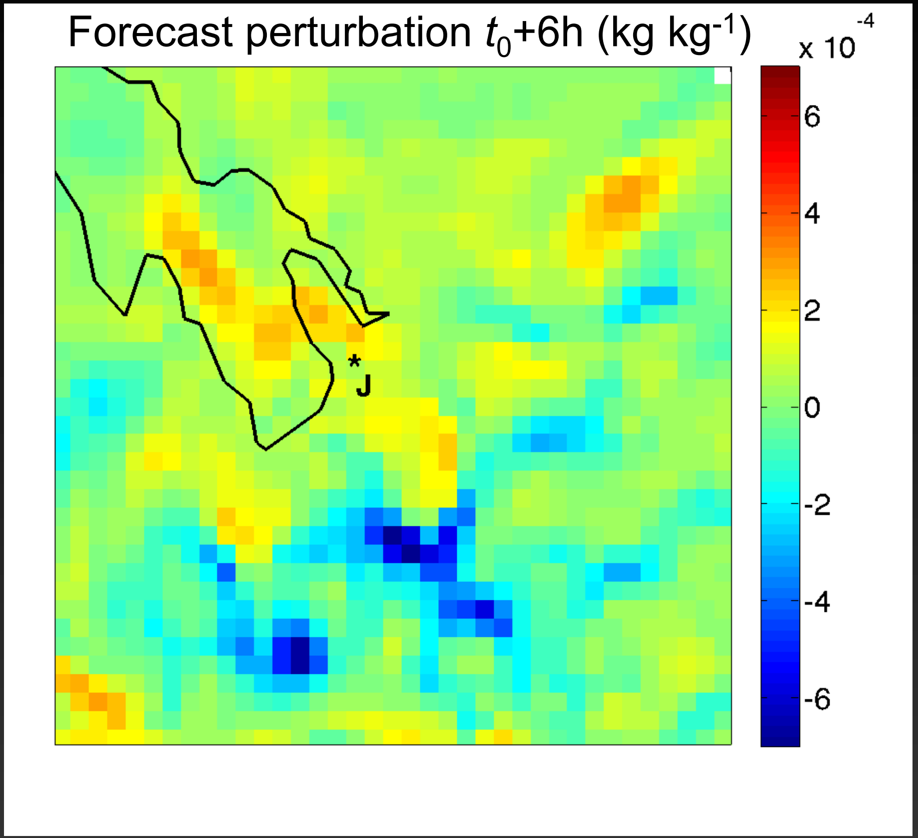


Perturbation of one analysis standard deviation in θ at the most sensitive location, *regressed* to remaining state elements.

Perturbation experiments

NCAR





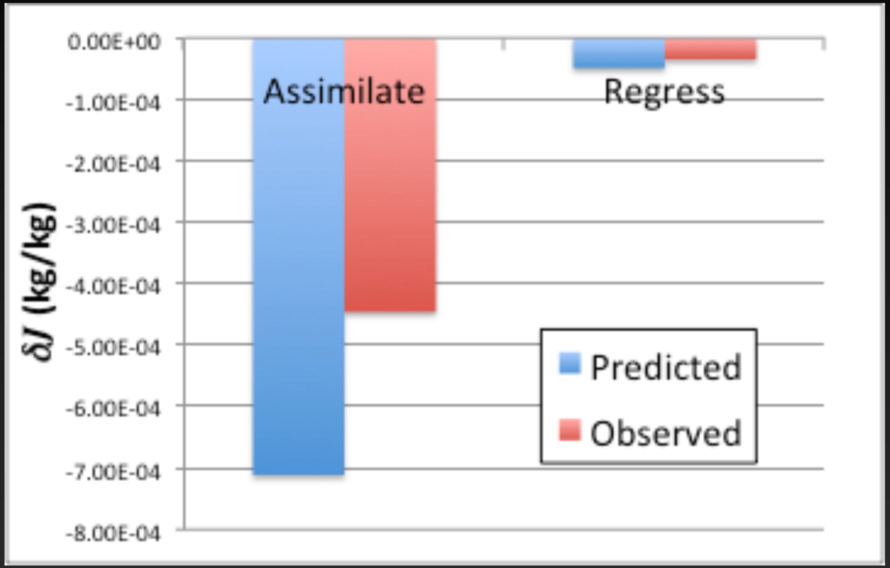
Perturbation of one analysis standard deviation in θ at the most sensitive location, *assimilated* with ensemble filter.

Effect of hypothetical θ observation

NCAR

$$\delta J_e = \frac{\partial J_e}{\partial x^a}\mathbf{K}(\mathbf{y}^o - \mathbf{h}\mathbf{x}^a)$$

$$\mathbf{K} = \mathbf{P}^a\mathbf{h}^{\mathrm{T}}\left(\mathbf{h}\mathbf{P}^a\mathbf{h}^{\mathrm{T}} + \mathbf{R}\right)^{-1}$$



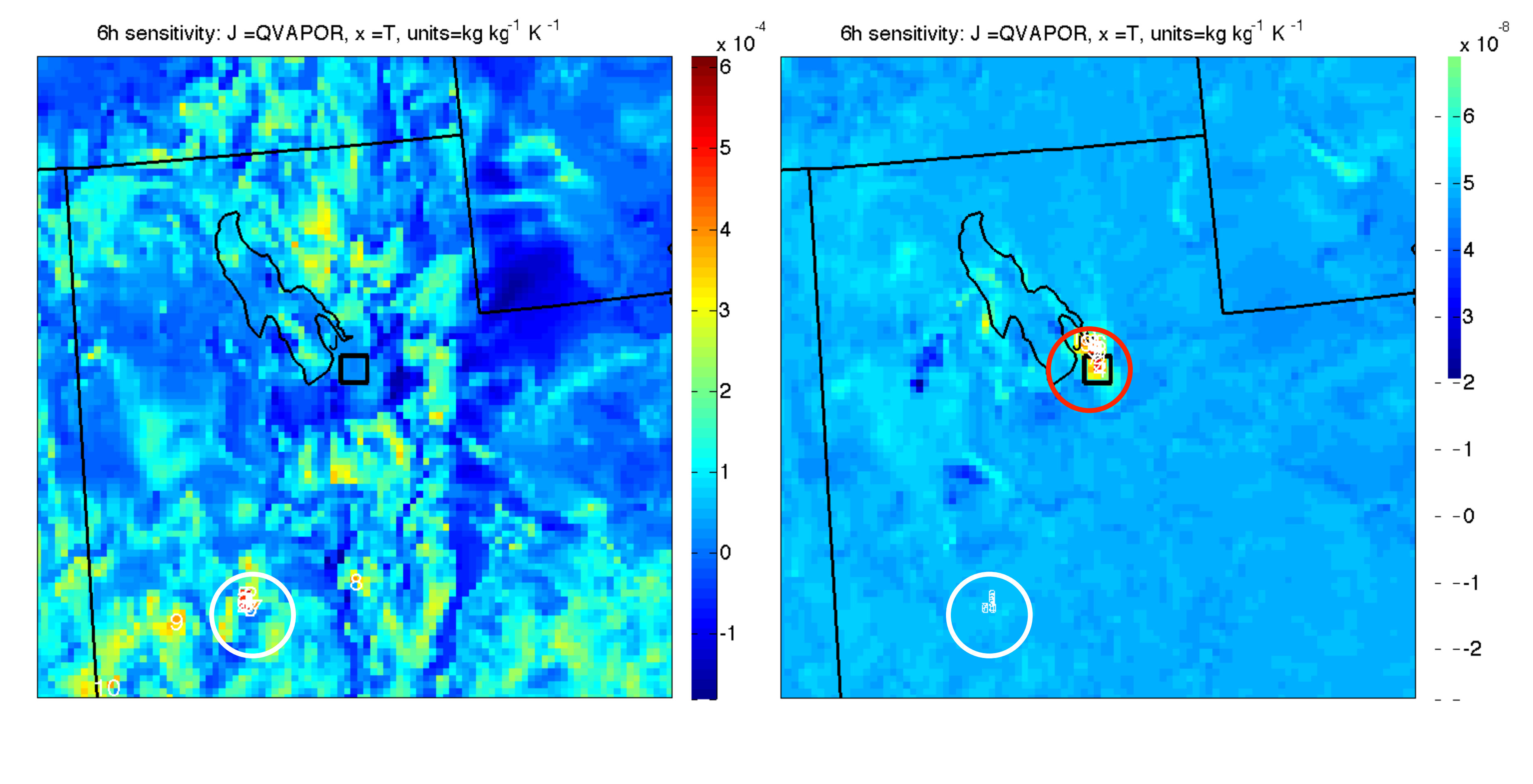
Can test use of sensitivities to predict the change in forecast metric resulting from a hypothetical observation. Analysis increment can come from:

- assimilating synthetic obs
- approximation with univariate linear regression

Effect of approximation

Diagonal approximation | Full covariance

NCAR



Approximation under-emphasizes sensitivities local to the response. Agreement on some sensitive points (numbered) to southwest of response.

Summary (1)

NCAR

- ESA appears promising for mesoscales and in complex terrain; hypothetical observations give qualitatively expected forecast change, but overestimated response.
- At mesoscales with weak sensitivity gradients, full covariance (and associated inversion) may be necessary.
- Linearity appears to hold for a variety of perturbations, possibly as large as 10 times the standard deviation of the analysis variable.

This research partially funded by Office of Naval Research Award # N00014-11-1-0709, Mountain Terrain Atmospheric Modeling and Observations (MATERHORN) Program.

Ensemble Sensitivity (1)

NCAR

$$\mathbf{J_e} = \left[\mathbf{X^a}\right]^{\mathrm{T}} \boldsymbol{\beta} + \varepsilon$$

$$\hat{\beta} = \frac{\partial J_e}{\partial \mathbf{x^a}}$$

$$\hat{\beta} = \mathbf{X^a}\left[\left(\mathbf{X^a}\right)^{\mathrm{T}} \mathbf{X^a}\right]^{-1} \mathbf{J_e} = \mathbf{QR^{-T}J_e}$$

- An ensemble sample (size K) of analysis perturbations and forecast metrics are assembled into matrix $\mathbf{X^a}$ and column vector $\mathbf{J_e}$, forming the regression equation.
- The solution, giving estimated regression coefficients, is the ensemble sensitivity defined as the gradient of the forecast metric relative to the analysis.
- Because K << N (state dimension), the system is extremely under-determined, but the minimum-norm solution is obtainable via a QR decomposition.

Ensemble Sensitivity (2)

NCAR

$$\mathbf{K}_i = \mathbf{P}_i^{\mathbf{a}} \mathbf{h}_{i+1}^{\mathbf{T}} \left(\mathbf{h}_{i+1} \mathbf{P}_i^{\mathbf{a}} \mathbf{h}_{i+1}^{\mathbf{T}} + \mathbf{R}_{i+1} \right)^{-1}$$

$$\delta J = \left(\frac{\partial J_e}{\partial \mathbf{x}^{\mathbf{a}}} \right)^{\mathbf{T}} \mathbf{K}_i \left(y_{i+1}^{o} - \mathbf{h}_{i+1} \mathbf{x}_i^{\mathbf{a}} \right)$$

$$= \left(\frac{\partial J_e}{\partial \mathbf{x}^{\mathbf{a}}} \right)^{\mathbf{T}} \delta \mathbf{x}^{\mathbf{a}}$$

$$\boxed{\mathbf{P}_i^{\mathbf{a}} = \mathbf{X}_i^{\mathbf{a}} \left(\mathbf{X}_i^{\mathbf{a}} \right)^{\mathbf{T}}}$$

$$= \mathbf{J}_{\mathbf{e}}^{\mathbf{T}} \left\{ \mathbf{X}_i^{\mathbf{a}} \left[\left(\mathbf{X}_i^{\mathbf{a}} \right)^{\mathbf{T}} \mathbf{X}_i^{\mathbf{a}} \right]^{-1} \right\}^{\mathbf{T}} \mathbf{K}_i \left(y_{i+1}^{o} - \mathbf{h}_{i+1} \mathbf{x}_i^{\mathbf{a}} \right)$$

Assimilation: a perturbation $\delta \mathbf{x}^{\mathbf{a}}$ resulting from assimilating an additional observation, multiplied by the sensitivities, gives the the expected forecast change resulting from assimilating that observation (i.e. the predicted response).

Ensemble Sensitivity (3)

$$\mathbf{K}_i = \mathbf{P}_i^{\mathbf{a}}\mathbf{h}_{i+1}^{\mathbf{T}}\left(\mathbf{h}_{i+1}\mathbf{P}_i^{\mathbf{a}}\mathbf{h}_{i+1}^{\mathbf{T}} + \mathbf{R}_{i+1}\right)^{-1}$$

$$\delta J = \left(\frac{\partial J_e}{\partial \mathbf{x}^{\mathbf{a}}}\right)^{\mathbf{T}} \mathbf{K}_i\left(y_{i+1}^o - \mathbf{h}_{i+1}\mathbf{x}_i^{\mathbf{a}}\right)$$

$$= \left(\frac{\partial J_e}{\partial \mathbf{x}^{\mathbf{a}}}\right)^{\mathbf{T}} \delta\mathbf{x}^{\mathbf{a}}$$

$$= \mathbf{J}_{\mathbf{e}}^{\mathbf{T}}\left\{\mathbf{X}_i^{\mathbf{a}}\left[\left(\mathbf{X}_i^{\mathbf{a}}\right)^{\mathbf{T}}\mathbf{X}_i^{\mathbf{a}}\right]^{-1}\right\}^{\mathbf{T}} \mathbf{K}_i\left(y_{i+1}^o - \mathbf{h}_{i+1}\mathbf{x}_i^{\mathbf{a}}\right)$$

Sampling error in ensemble data assimilation typically mitigated by reducing covariances with a function of distance; follows intuition that distant covariances must be small or zero.

Ensemble Sensitivity (4)

$$\mathbf{K}_i = \mathbf{P}_i^{\mathbf{a}} \mathbf{h}_{i+1}^{\mathbf{T}} \left(\mathbf{h}_{i+1} \mathbf{P}_i^{\mathbf{a}} \mathbf{h}_{i+1}^{\mathbf{T}} + \mathbf{R}_{i+1} \right)^{-1}$$

$$\delta J = \left(\frac{\partial J_e}{\partial \mathbf{x}^{\mathbf{a}}} \right)^{\mathbf{T}} \mathbf{K}_i \left(y_{i+1}^{o} - \mathbf{h}_{i+1} \mathbf{x}_i^{\mathbf{a}} \right)$$

$$= \left(\frac{\partial J_e}{\partial \mathbf{x}^{\mathbf{a}}} \right)^{\mathbf{T}} \delta \mathbf{x}^{\mathbf{a}}$$

$$= \mathbf{J}_{\mathbf{e}}^{\mathbf{T}} \left\{ \mathbf{X}_i^{\mathbf{a}} \left[\left(\mathbf{X}_i^{\mathbf{a}} \right)^{\mathbf{T}} \mathbf{X}_i^{\mathbf{a}} \right]^{-1} \right\}^{\mathbf{T}} \mathbf{K}_i \left(y_{i+1}^{o} - \mathbf{h}_{i+1} \mathbf{x}_i^{\mathbf{a}} \right)$$

Sampling error in sensitivities arise in spatio-temporal covariances. A few methods have been proposed in the ensemble assimilation literature. Here from a Bayesian hierarchical estimate (Anderson 2007).

Ensemble Sensitivity (5)

NCAR

$$\frac{\partial J}{\partial \mathbf{x}^{\mathbf{a}}} = \frac{\mathbf{X}^{\mathbf{a}}\mathbf{J}_{\mathbf{e}}}{\mathbf{P}^{\mathbf{f}}} \approx \frac{\mathbf{X}^{\mathbf{a}}\mathbf{J}_{\mathbf{e}}}{\mathbf{D}^{\mathbf{a}}}$$

$$\mathbf{D}^{\mathbf{a}} = \mathrm{diag}(\mathbf{P}^{\mathbf{f}})$$

Approximation: in the meteorology literature the inversion needed to solve the regression problem is always avoided by approximating the covariance with its diagonal. The result is a scalar (univariate) regression for each element in the state vector.

Ensemble sensitivity details

$$\mathbf{J}_e = \left[\mathbf{X}^a\right]^{\mathrm{T}} \boldsymbol{\beta} + \varepsilon$$

$$\hat{\boldsymbol{\beta}} = \frac{\partial J_e}{\partial \mathbf{x}^a} = \mathbf{X}^a \left(\left[\mathbf{X}^a\right]^{\mathrm{T}} \mathbf{X}^a \right)^{-1} \mathbf{J}_e = \mathbf{Q}\mathbf{R}^{-\mathrm{T}} \mathbf{J}_e$$

$\mathbf{J}_e$ are perturbations about J_e (scalars)

$\mathbf{X}^a$ are perturbations about x^a (vectors)

Sensitivity is multi-variate linear regression; coefficients can be estimated via a right pseudo-inverse.

$$\frac{\partial J_e}{\partial \mathbf{x}^a} = \left[\mathbf{P}^a\right]^{-1} \mathbf{X}^a \mathbf{J}_e \approx \left[\mathbf{D}^a\right]^{-1} \mathbf{X}^a \mathbf{J}_e$$

$$\mathbf{P}^a = \mathbf{X}^a \left[\mathbf{X}^a\right]^{\mathrm{T}}, \ \mathbf{D}^a = \mathrm{diag}\left(\mathbf{P}^a\right)$$

More common in the literature is to avoid an inversion by assuming covariances are zero, leading to a scalar problem for each state element.

Open questions

- Ensemble sensitivities in the presence of small, fast scales
  – May increase nonlinearity
  – Increases model error/inadequacy
  – Appear as noise in correlations/covariances
- Validity of diagonal approximation
- Need to account for sampling error arising from finite ensemble

Ensemble Sensitivity with Localization

$$\delta J = \alpha \circ \left\{ \mathbf{J}_{\mathbf{e}}^{\mathrm{T}} \left[\mathbf{X}_i^{\mathbf{a}} \left(\mathbf{X}_i^{\mathbf{a}\,\mathrm{T}} \mathbf{X}_i^{\mathbf{a}} \right)^{-1} \right]^{\mathrm{T}} \rho \circ \mathbf{P}_i^{\mathbf{a}} \mathbf{h}_{i+1}^{\mathrm{T}} \left(\mathbf{h}_{i+1} \rho \circ \mathbf{P}^{\mathbf{a}} \mathbf{h}_{i+1}^{\mathrm{T}} + \mathbf{R} \right)^{-1} \left(y_{i+1}^{o} - \mathbf{h}_{i+1} \mathbf{x}_i^{\mathbf{a}} \right) \right\}$$

$$= \alpha \circ \left\{ \mathbf{J}_{\mathbf{e}}^{\mathrm{T}} \left[\mathbf{X}_i^{\mathbf{a}} \left(\mathbf{X}_i^{\mathbf{a}\,\mathrm{T}} \mathbf{X}_i^{\mathbf{a}} \right)^{-1} \right]^{\mathrm{T}} \delta \mathbf{x}^{\mathbf{a}} \right\}$$

- Covariance localization, or tapering, can be applied
  - at the assimilation step with ρ
  - to the regressions with α
- ρ is typically a function of space alone
- α is function of space and time, here from a Bayesian hierarchical estimate (Anderson 2007)

Experiment Details (1)

NCAR

- Nature/truth from Lorenz (2005) one-scale Model II or two-scale Model III
  - Perfect-model experiments
  - Model error simulated by retaining fast scale in nature run/ truth and eliminating it in the assimilating model
- Ensemble-filter data assimilation every 6 h
  - 80 cycles
    - Network of every-other grid point; or
    - Network of one-half of domain totally observed
- Forecast metric (J) is root-mean square error (RMSE)

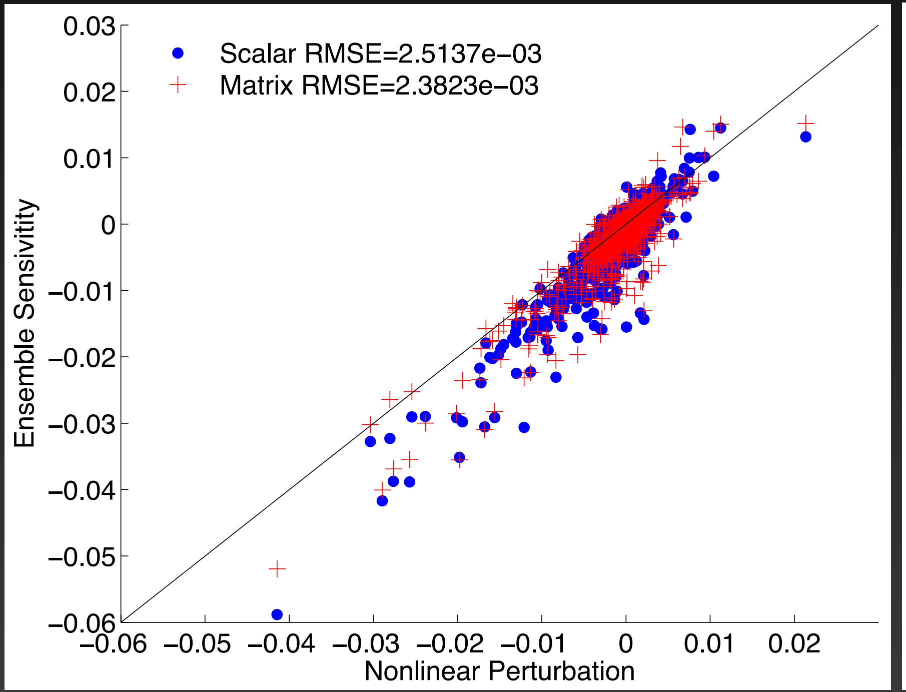
Experiment Details (2)

- Apply individual perturbations by assimilating individual observation at randomly-chosen unobserved gridpoints
- Evaluate 6-h forecast response with nonlinear model
- Compare to 6-h response as predicted by linear method: $$\delta J = \left(\frac{\partial J_e}{\partial \mathbf{x}^{\mathbf{a}}} \right)^{\mathbf{T}} \delta \mathbf{x}^{\mathbf{a}}$$

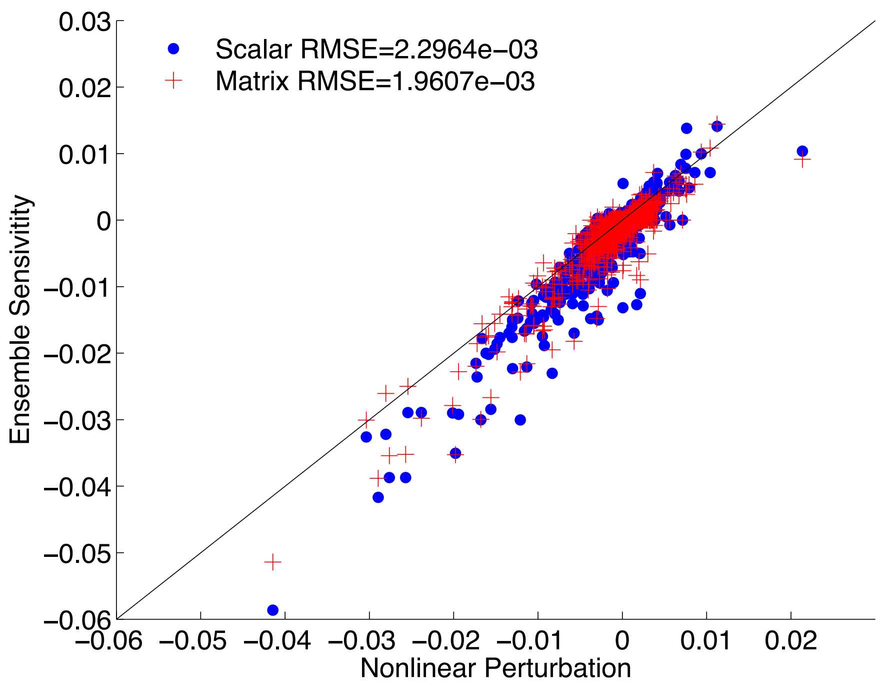
Perfect Model II

When only smooth/slow scales present, little difference between univariate (scalar) and multivariate (matrix) predictions of response to perturbation.

Sensitivity *without* localization



Sensitivity *with* localization

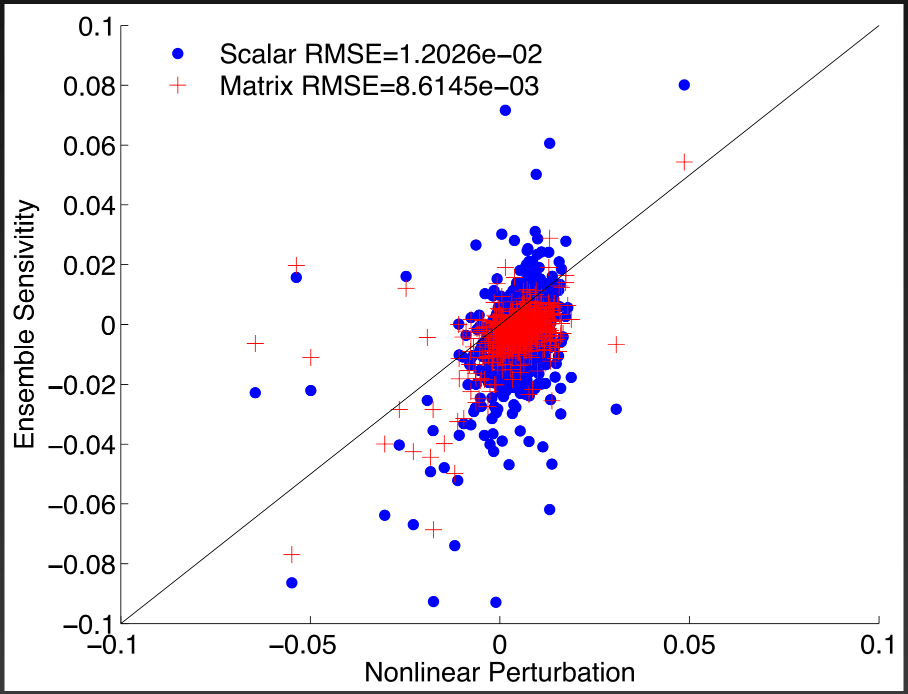


Here observations are randomly chosen from every other gridpoint (which are un-observed for sensitivity calculations).

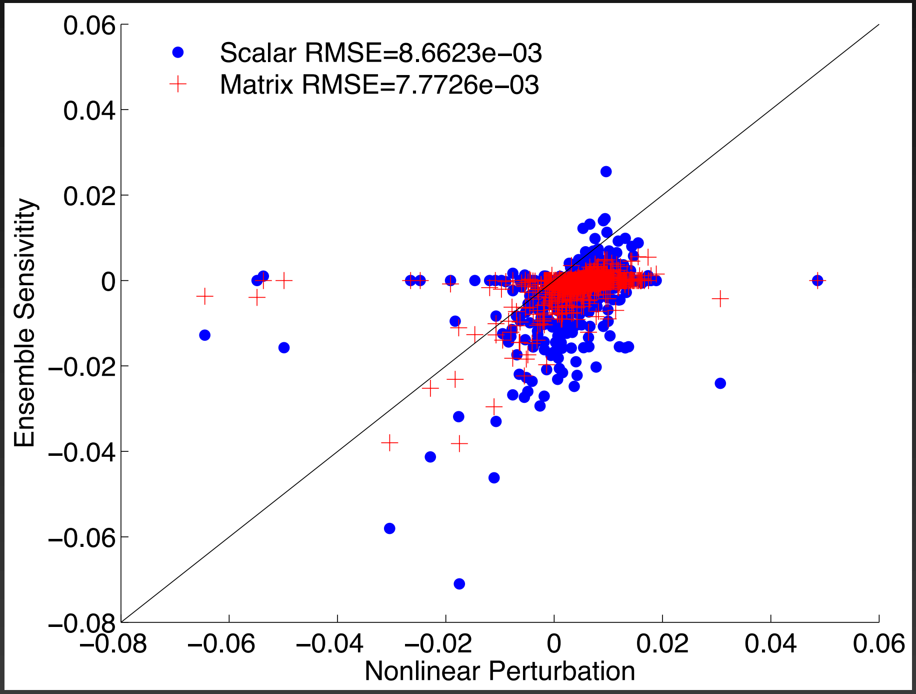
Perfect Model III

When both slow and fast scales are present, diagonal approximation is less accurate. Localization slightly improves predictions of response.

Sensitivity *without* localization



Sensitivity *with* localization



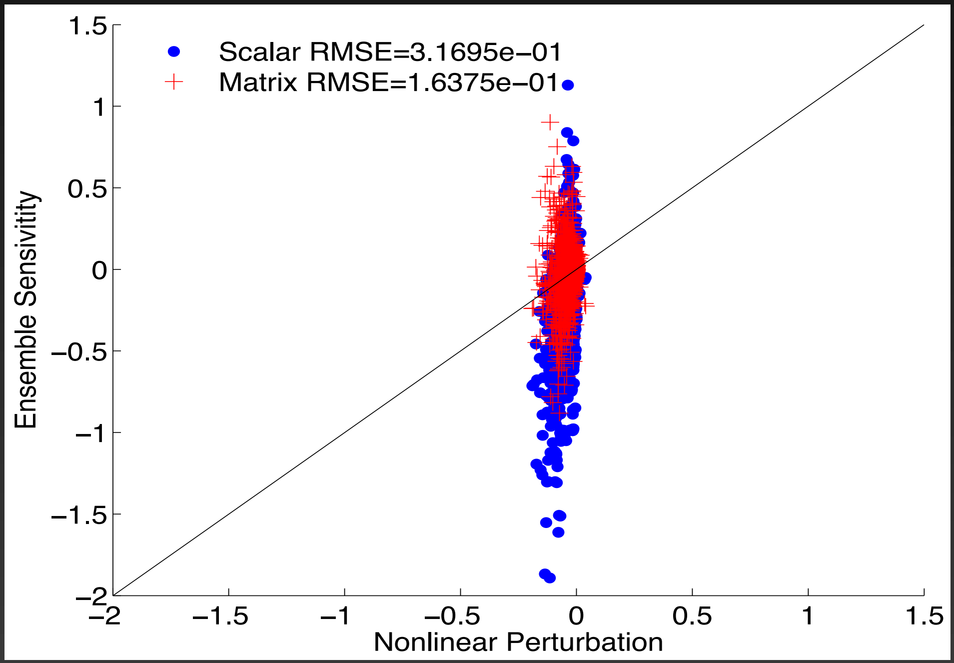
Here observations are randomly chosen from every other gridpoint (which are un-observed for sensitivity calculations).

Imperfect Model

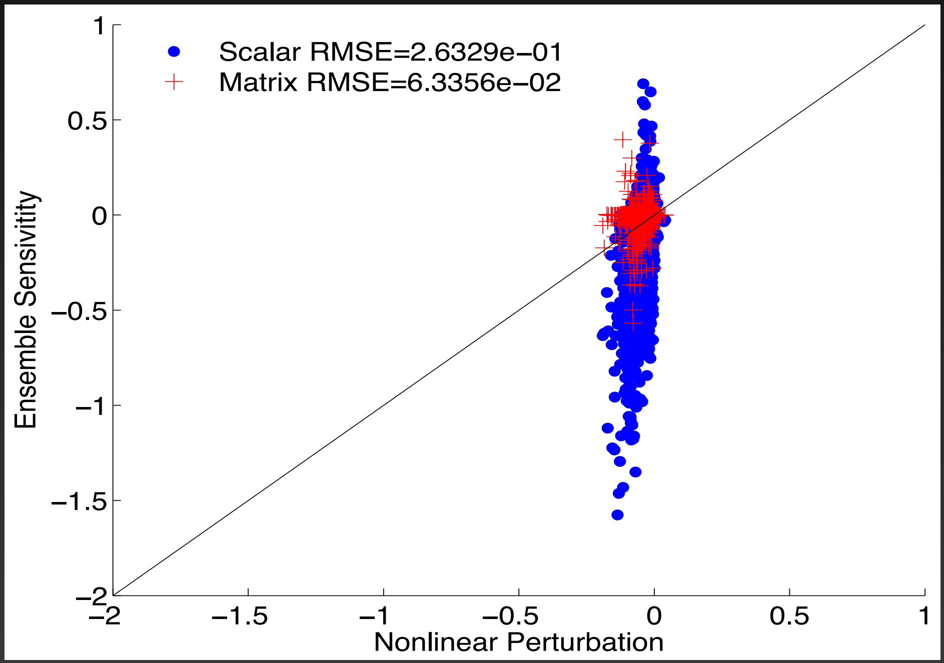
For imperfect model, diagonal approximation results in greater over-prediction of response; multivariate sensitivities account for presence of fast scales in real system, which appears as noise.

NCAR

Sensitivity *without* localization

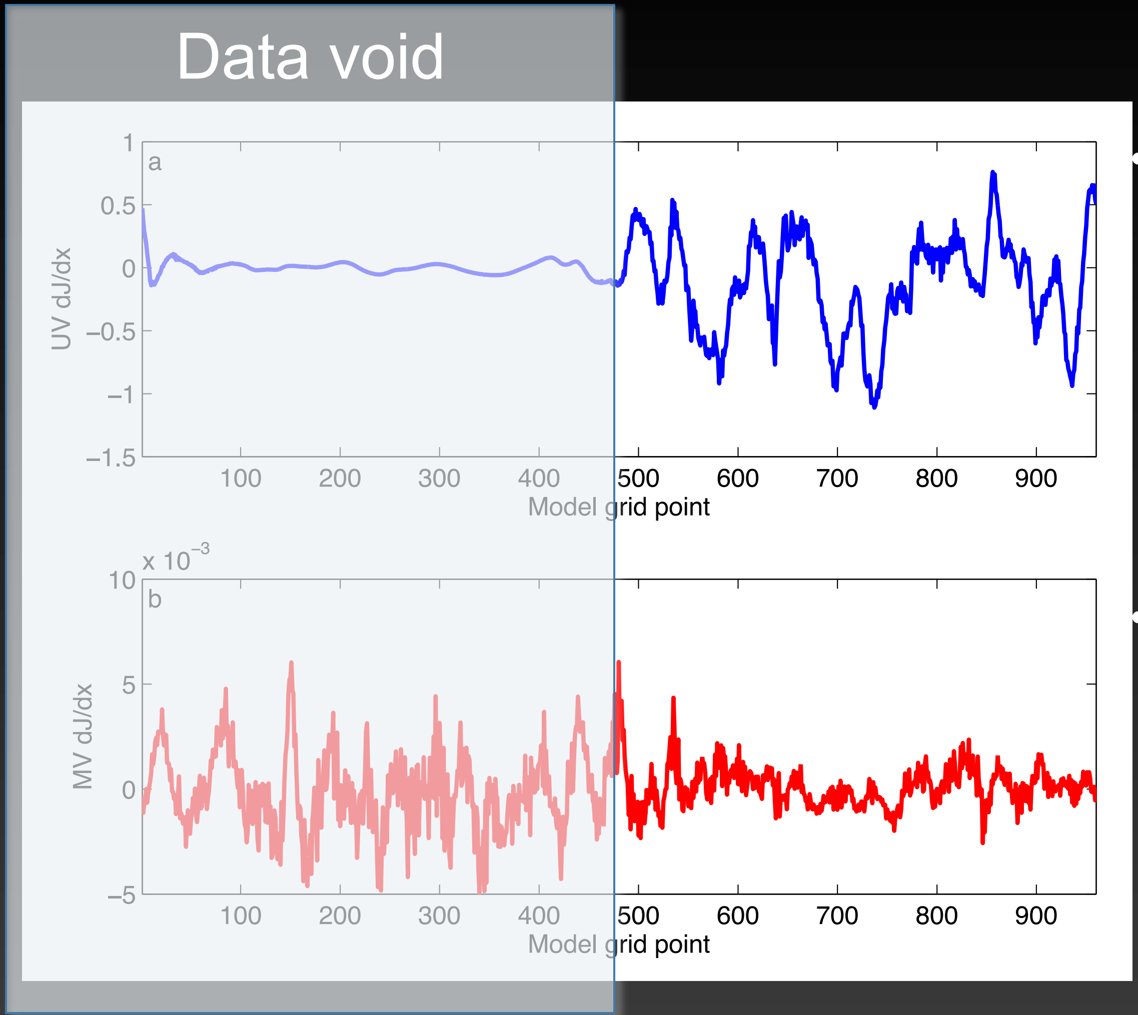


Sensitivity *with* localization



Here observations are assimilated on half of domain that is data void; more impact from observations because greater uncertainty in analysis.

Sensitivities in a data void



- Top: univariate sensitivities are small in the data void because analysis uncertainty is large
- Bottom: multivariate sensitivities larger over data void than over densely observed region, consistent with expectations

Summary (II)

NCAR

- Multivariate sensitivities are possible to estimate by finding a minimum-norm solution to the resulting underdetermined matrix problem.
- Whether univariate or multivariate methods are employed, sampling error is a problem.
- Sensitivities used to predict the perturbation response in the nonlinear system are more accurate when localized/tapered to account for sampling error.
- Multivariate sensitivities better predict the nonlinear response when:
  - Fast scales are present
  - Model error is present
  - Part of the state is poorly observed and can benefit from additional observations

Results suggest mesoscale sensitivities for real atmospheric problems will be more useful if using multivariate estimates.

This research partially funded by Office of Naval Research Award # N00014-11-1-0709, Mountain Terrain Atmospheric Modeling and Observations (MATERHORN) Program.

References

1. Anderson, J. L., 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634–642.
2. Anderson, J. L., 2007: Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D*, **230**, 99–111.
3. Ancell, B. and G. Hakim, 2007: Comparing adjoint- and ensemble-sensitivity analysis with applications to observation targeting. *Mon. Wea. Rev.*, **135**, 4117–4134.
4. Gaspari, G. and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.
5. Hacker, J. P. and L. Lei, 2015: Multivariate ensemble sensitivity with localization. Mon. Wea. Rev., **143**, 2013–2027.
6. Lorenz, E. N. 2005: Designing chaotic models. *J. Atmos. Sci.*, **62**, 1574–1587.
7. Torn, R. D., and G. J. Hakim, 2008: Ensemble-based sensitivity analysis. *Mon. Wea. Rev.*, **136**, 663–677.