Multivariate Ensemble Sensitivity with Localization

JOSHUA P. HACKER

National Center for Atmospheric Research, Boulder, Colorado*

LILI LEI

Cooperative Institute for Research in Environmental Sciences, Climate Diagnostics Center, University of Colorado, and Physical Sciences Division, NOAA/Earth System Research Laboratory, Boulder, Colorado

(Manuscript received 23 September 2014, in final form 13 January 2015)

ABSTRACT

Ensemble sensitivities have proven a useful alternative to adjoint sensitivities for large-scale dynamics, but their performance in multiscale flows has not been thoroughly examined. When computing sensitivities, the analysis covariance is usually approximated with the corresponding diagonal matrix, leading to a simple univariate regression problem rather than a more general multivariate regression problem. Sensitivity estimates are affected by sampling error arising from a finite ensemble and can lead to an overestimated response to an analysis perturbation. When forecasts depend on many details of an analysis, it is reasonable to expect that the diagonal approximation is too severe. Because spurious covariances are more likely when correlations are weak, computing the sensitivity with a multivariate regression that retains the full analysis covariance may increase the need for sampling error mitigation. The purpose of this work is to clarify the effects of the diagonal approximation, and investigate the need for mitigating spurious covariances arising from sampling error. A two-scale model with realistic spatial covariances is the basis for experimentation. For most problems, an efficient matrix inversion is possible by finding a minimum-norm solution, and employing appropriate matrix factorization. A published hierarchical approach for estimating regression factors for tapering (localizing) covariances is used to measure the effects of sampling error. Compared to univariate regressions in the diagonal approximation, skill in predicting a nonlinear response from the linear sensitivities is superior when localized multivariate sensitivities are used, particularly when fast scales are present, model error is present, and the observing network is sparse.

1. Introduction

Sensitivity analysis is a key methodology for understanding and tuning models. In numerical weather prediction (NWP), the goal is most often to understand the response to an arbitrary perturbation to initial conditions or a parameter in a model. Accurate a priori estimates of sensitivity to a perturbation or model change can save enormous computational expense by avoiding many integrations forward in time. So that sensitivity estimates can be interpreted correctly, it is critical to be sure that they give good approximations to the perturbation response realized in the full dynamical system. In the most common scenario, sensitivity estimated with a linearization about a nonlinear system evolution can be a good approximation as long as the perturbation remains small. Nonlinear growth of a perturbation limits the length of period that a sensitivity estimate is valid.

During the last two decades, sensitivity methods have been extended with the goal of quantifying how a new observation can affect a forecast by assimilating it at analysis time, or understanding the impacts of existing observation sets on forecast skill. Deploying a new observation based on the predicted response (e.g., error reduction) is usually called *targeting*. Singular vector

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Joshua P. Hacker, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307.
E-mail: hacker@ucar.edu

DOI: 10.1175/MWR-D-14-00309.1

targeting (Buizza and Montani 1999; Gelaro et al. 1999; Langland et al. 1999), which makes use of an adjoint, and ensemble-based targeting (Bishop and Toth 1999; Bishop et al. 2001; Hamill and Snyder 2002) were proposed for synoptic midlatitude weather forecasting. Langland and Baker (2004) also published the seminal paper showing how an adjoint model and a variational data assimilation system can together provide deterministic impact to observation sets. Later, Ancell and Hakim (2007) examined the relationship between adjoint sensitivities and ensemble sensitivities, pointing out that the two are equivalent in the limit of an infinite ensemble, Gaussian statistics, and linear perturbation growth. Around the same time, Hakim and Torn (2008) and Torn and Hakim (2008) used the regressions underlying ensemble sensitivities to identify dynamic links in evolving weather patterns.

Application of adjoint and ensemble sensitivity methods to high-resolution forecast problems (grid spacing less than 5 km) has to this point been sparse. One example is Wile et al. (2014, manuscript submitted to *Wea. Forecasting*, hereafter WHC) who applied ensemble sensitivities to a weakly forced fog case over the Great Salt Lake in Utah. Although the sensitivities between the initial conditions and the forecasts exposed physically plausible precursors for the fog, strong, coherent patterns of sensitivity were absent. The sensitivities led to systematic overprediction of response to perturbations, compared to the response measured from nonlinear model integrations. As suggested by Torn and Hakim (2008), an overpredicted response is one possible effect of sampling error. Namely, if an estimated covariance between the initial conditions and a forecast metric is spuriously large because of sampling error, and the analysis error statistics do not overestimate the covariance as severely, the sensitivity can be overestimated. Most results in the literature so far have used an approximation to the analysis error covariance, where it is assumed diagonal (e.g., Ancell and Hakim 2007; Torn and Hakim 2008). Across a broad range of problems, and in particular for mesoscale sensitivities lacking strong forcing or clear analytic balances such as geostrophy, the effects of a diagonal approximation are not clear. In related work using regressions with ensemble statistics to invert potential vorticity, Gombos and Hansen (2008) avoided the diagonal approximation and instead used singular value decomposition (SVD) to invert the covariance matrix, where the SVD was truncated to retain only the number of eigenvalues corresponding to the ensemble size. The truncated SVD stabilizes the pseudoinverse of the rank-deficient covariance. Later, Gombos et al. (2012) examined tropical cyclone track sensitivities, and projected the covariance onto the leading eigenvectors to improve the conditioning of the covariance matrix and stabilize the inversion.

This work examines both the potentially deleterious effects of the diagonal approximation and ignoring sampling error on ensemble sensitivity estimates. The Lorenz (2005, hereafter L05) two-scale model enables a large number of data assimilation cycles with an ensemble filter. Experiments with and without the fast scale in the model, and including model error, provide context for interpreting both synoptic and mesoscale sensitivities. An objectively estimated factor applied to reduce the regression coefficients is used to mitigate sampling error in ensemble sensitivity estimates.

Ordinary least squares provides a starting point for deriving ensemble sensitivities in section 2, which also clarifies where sampling error appears and how it can be mitigated. Section 3 provides experiment details, and section 4 results. Section 5 reviews the key results and suggests next steps.

2. Ensemble sensitivity

In this section a derivation of sensitivities from ordinary least squares provides the basis for estimating the impact of a hypothetical observation. It also elucidates the role of sampling error and where mitigation may be possible. The mathematical notion is given in Table 1.

a. Derivation from ordinary least squares

An expression for ensemble sensitivity is easily derived as an ordinary least squares (OLS) problem. Sensitivity of a forecast response function $J(\mathbf{x}^f)$ to perturbations in an analysis is often estimated from an *adjoint sensitivity*, expressed as

$$\frac{\partial J}{\partial \mathbf{x}^a}. \tag{1}$$

A probabilistic approach to estimating the sensitivity arises naturally within an ensemble context. An ensemble of forecasts from a sample of initial conditions provides a sample of response functions. The *ensemble sensitivity* can therefore be defined with respect to the ensemble means J_e and $\mathbf{x}^a$, and is simply a statistical linearization of Eq. (1).

Given a sample of J values and a sample of analyses $\mathbf{x}^a_k$, such as provided in a cycling ensemble data assimilation, the linear sensitivity can be estimated with the ensemble statistics. Collecting the ensemble of $J_{e,k}$ into a vector $\mathbf{J}_e$ and the ensemble of $\mathbf{x}^a_k - \mathbf{x}^a$ into a matrix $\mathbf{X}^a$, the sample statistics give the linear prediction equation:

$$\mathbf{J}_e = [\mathbf{X}^a]^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{2}$$

which is a system of K equations in N unknowns. The solution to Eq. (2) is an OLS problem that describes the

TABLE 1. Mathematical notation.

I	No. of observations available at a particular assimilation time
K	Ensemble size
N	Dimension of model state
J	Scalar deterministic forecast metric (i.e., from an adjoint)
J_e	Scalar forecast metric derived from an ensemble mean forecast
$\mathbf{J}_e$	$K \times 1$ response vector with elements $J_{e,k}$
$\mathbf{x}_i^a$	$N \times 1$ analysis state vector (ensemble mean) resulting from assimilating observations $1, \ldots, i$
$\mathbf{x}_{i+1}^a$	$N \times 1$ analysis state vector (ensemble mean) resulting from assimilating observations $1, \ldots, i+1$
$\mathbf{x}_k^a$	kth ensemble member analysis state vector
$\mathbf{X}^a$	$N \times K$ matrix containing the ensemble of state vector perturbations (columns)
$\delta\mathbf{x}^a$	$N \times 1$ analysis perturbation vector (change in ensemble mean)
$\mathbf{x}^f$	$N \times 1$ forecast (background) state vector
$\boldsymbol{\beta}$	$N \times 1$ vector of regression coefficients
$\boldsymbol{\varepsilon}$	$N \times 1$ vector of errors in the sensitivity prediction model
$\mathbf{H}$	$I \times N$ matrix, each row containing a linearized forward operator for observation $i \in [1, \ldots, I]$
$\mathbf{h}_i$	$1 \times N$ vector containing a linearized forward operator for observation i (i.e., rows of $\mathbf{H}$)
$\mathbf{P}^a$	$=\mathbf{X}^a[\mathbf{X}^a]^{\mathrm{T}}$, $N \times N$ ensemble-estimated analysis error covariance
$\mathbf{D}^a$	$=\mathrm{diag}(\mathbf{P}^a)$, $N \times N$ ensemble-estimated analysis error variance (diagonal)
$\mathbf{R}$	$I \times I$ observation error variance
$\boldsymbol{\rho}$	$N \times N$ covariance localization matrix
$\boldsymbol{\alpha}$	$N \times 1$ sensitivity localization vector
α	Scalar element of $\boldsymbol{\alpha}$

change in $\mathbf{x}^a$ needed to produce a given change in J. The solution $\hat{\boldsymbol{\beta}}$ gives a linear statistical estimate to the sensitivity:

$$\hat{\boldsymbol{\beta}} = \frac{\partial J_e}{\partial \mathbf{x}^a} \,. \tag{3}$$

Equation (2) is an underdetermined system because $K < N$. The matrix $[\mathbf{X}^a]^{\mathrm{T}}$ is sometimes called *fat*, and the associated system has an infinite number of solutions. Instead of the usual minimum-variance solution sought for overdetermined systems, we can seek a minimum-norm solution for the underdetermined system (Golub and Van Loan 1996). The solution arises from the optimization problem where $\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\beta}$ is minimized subject to the constraint $[\mathbf{X}^a]^{\mathrm{T}}\boldsymbol{\beta} = \mathbf{J}_e$. The method of Lagrange multipliers leads to the following form:

$$\hat{\boldsymbol{\beta}} = \mathbf{X}^a([\mathbf{X}^a]^{\mathrm{T}}\mathbf{X}^a)^{-1}\mathbf{J}_e = \mathbf{Q}\mathbf{R}^{-\mathrm{T}}\mathbf{J}_e, \tag{4}$$

where $\mathbf{X}^a = \mathbf{QR}$ is the $\mathbf{QR}$ factorization of $\mathbf{X}^a$. Note that $[\mathbf{X}^a]^{\mathrm{T}}\mathbf{X}^a$ is full rank and can thus be accurately inverted via the factorization. The quantity $\mathbf{X}^a([\mathbf{X}^a]^{\mathrm{T}}\mathbf{X}^a)^{-1}$ is the right pseudoinverse of $[\mathbf{X}^a]^{\mathrm{T}}$.

Solving the underdetermined system obviates the need for the principal approximation proposed in Ancell and Hakim (2007). They defined ensemble sensitivity by ignoring the off-diagonal components of $\mathbf{P}^a$. The approximation follows from the minimum-variance solution to Eq. (2):

$$\frac{\partial J}{\partial \mathbf{x}^a} = \mathbf{X}^a\mathbf{J}_e(\mathbf{X}^a[\mathbf{X}^a]^{\mathrm{T}})^{-1} = \mathbf{X}^a\mathbf{J}_e(\mathbf{P}^a)^{-1} \approx \mathbf{X}^a\mathbf{J}_e(\mathbf{D}^a)^{-1}. \tag{5}$$

The sensitivity is reduced to a univariate (scalar) regression for each diagonal element x^a, which avoids the matrix inversion.

The diagonal approximation can also be misleading or inadequate. By ignoring off-diagonal components, the sensitivity to individual state elements is overestimated. Sets of state elements that are individually and strongly correlated with the response function can still provide clues about dynamic links, but they cannot be quantitatively correct because they ignore contributions from all initial variables simultaneously. Further, when no coherent set of state elements are individually, and strongly, correlated with the response function, dynamic interpretation is much more difficult. This can easily be imagined under weakly forced or nonlinear scenarios (cf. WHC), or during periods of unbalanced dynamics common in mesoscale flows. Each analysis state component contributes to the change in J, with the change given by the linear combination of weighted predictors. For a given change in J, each sensitivity is necessarily smaller than when assuming no off-diagonal contributions.

Swapping the dependent and independent variables in Eq. (2) leads to an alternate strategy, and an alternate interpretation. That inverse sensitivity can answer the question of what is the initial time change needed for a given change in the response function. The linear model is for the change in J that results from a given change in $\mathbf{x}^a$, and the solution immediately reduces to a scalar problem. Torn and Hakim (2009) adopted this strategy, which also preserves in initial-condition covariances, for choosing initial-condition perturbations adhering to the sensitivities.

b. *Sensitivity to an observation*

Regression to solve Eq. (2) gives a linear prediction for how a forecast will respond to an initial-condition perturbation. Estimates are valid for the ensemble system used to derive the statistics, provided the distributions are Gaussian, the perturbation is small, and its evolution is linear. Besides using the sensitivities to infer dynamics, the sensitivity can be multiplied by an expected analysis increment from a new or hypothetical observation in a data assimilation system, providing an

estimate of the forecast response from including that observation.

Consider assimilating i observations, and proposing an additional observation $i+1$. After assimilating observations $1, \ldots, i$, the resulting analysis is $\mathbf{x}_i^a$ with error covariance estimated by $\mathbf{P}_i^a$. From the statistical analysis equation, introducing the new observation y_{i+1}^o will change the analysis as

$$\mathbf{x}_{i+1}^a - \mathbf{x}_i^a = \mathbf{K}_i(y_{i+1}^o - \mathbf{h}_{i+1}\mathbf{x}_i^a) \quad \text{and} \tag{6}$$

$$\mathbf{K}_i = \mathbf{P}_i^a\mathbf{h}_{i+1}^{\mathrm{T}}(\mathbf{h}_{i+1}\mathbf{P}_i^a\mathbf{h}_{i+1}^{\mathrm{T}} + \mathbf{R})^{-1}. \tag{7}$$

The inversion needed in Eq. (7) is trivial because it is a scalar. This is easily extended to consider multiple hypothetical observations simultaneously by using any of the myriad stable methods for solving Eq. (6).

The quantity of interest here is δJ_e, which is the response to an initial-condition perturbation. A linear prediction for the response arises naturally from the product of the sensitivity and $\delta\mathbf{x}_i^a = \mathbf{x}_{i+1}^a - \mathbf{x}_i^a$. From Eqs. (3), (4), and (6),

$$\delta J = \left(\frac{\partial J_e}{\partial \mathbf{x}^a}\right)^{\mathrm{T}} \delta\mathbf{x}_i^a, \tag{8}$$

$$= \left(\frac{\partial J_e}{\partial \mathbf{x}^a}\right)^{\mathrm{T}} \mathbf{K}_i(y_{i+1}^o - \mathbf{h}_{i+1}\mathbf{x}_i^a), \tag{9}$$

$$= \mathbf{J}_e^{\mathrm{T}}[\mathbf{X}_i^a([\mathbf{X}_i^a]^{\mathrm{T}}\mathbf{X}_i^a)^{-1}]^{\mathrm{T}}\mathbf{K}_i(y_{i+1}^o - \mathbf{h}_{i+1}\mathbf{x}_i^a). \tag{10}$$

Sampling error can be easily treated in this form, but it is also straightforward to simplify the expression above to show that the sampling error in estimates of the perturbation response arises from only the covariance between J and $\mathbf{h}_{i+1}\mathbf{x}^a$.

c. *Sampling error in ensemble sensitivities*

Finite ensembles lead to sampling error in both the sensitivity estimates and an analysis perturbation from assimilating an observation. Sampling error increases the probability that covariances are overestimated. The predicted response given by Eq. (8) is then overestimated. In the diagonal approximation Eq. (5) the overprediction is expected to be worse because the analysis variances are expected to be underpredicted when subject to sampling error.

Sampling error in the covariances underlying the sensitivity estimates has not been adequately addressed, but sampling error in ensemble data assimilation is typically addressed with a "localization" function. Most often the spatial correlation function given by the fifth-order piecewise polynomial documented by Gaspari and Cohn (1999) serves this purpose. That correlation function has proven useful at mitigating the sampling error, but has some drawbacks. In particular, covariances among different physical quantities differ and can upset physical balances such as geostrophy (Houtekamer and Mitchell 2005). Localization as a function of solely space does not account for nonzero sampling error between the collocated observations of different physical quantities. Other methods use the covariance statistics themselves to estimate sampling error, and derive an associated weight for any scalar covariance (e.g., Anderson 2007; Bishop and Hodyss 2009). Those methods require no assumptions about spatial relationships and are not explicitly distance dependent.

Localization in the assimilation is applied through a Hadamard, or Schür, product that reduces the elements of $\mathbf{P}^a$ in Eq. (7) and is denoted $\boldsymbol{\rho}\circ\mathbf{P}^a$. When solving the statistical analysis equation with scalar regressions as in Anderson (2003), the localization is applied as a factor to each regression coefficient. Because the ensemble statistics are derived from an ensemble filter including some form of localization, the analysis increment from a hypothetical observation would be subject to the same localization. Thus, in Eqs. (6) and (7) the same approach would be used, shortening $\delta\mathbf{x}^a$ and reducing the predicted response δJ in Eq. (8).

A spatial function is inappropriate to handle sampling error in the sensitivity estimates because the covariances are in both space and time. Predicting the forecast response from a hypothetical observation then requires a second factor for handling sampling error. Methods relying on the covariance statistics themselves are candidates. Regardless of the method chosen to estimate the weights on the sensitivity regressions, they can also be applied via a Hadamard product or an additional factor on the regressions. Although the factors applied to elements of the covariance between the analysis and response function are not necessarily a function of space, for lack of a better term we retain localization to describe them.

Instead of an $N\times N$ matrix of localization values in the data assimilation, either a scalar or an $N\times 1$ vector of localization values is appropriate. The resulting predicted, and localized, effect of a hypothetical observation on a forecast metric is

$$\delta J = \alpha\circ\{\mathbf{J}_e^{\mathrm{T}}[\mathbf{X}_i^a(\mathbf{X}_i^{a\mathrm{T}}\mathbf{X}_i^a)^{-1}]^{\mathrm{T}}\boldsymbol{\rho}\circ\mathbf{P}_i^a\mathbf{h}_{i+1}^{\mathrm{T}}(\mathbf{h}_{i+1}\boldsymbol{\rho}\circ\mathbf{P}_i^a\mathbf{h}_{i+1}^{\mathrm{T}} + \mathbf{R})^{-1}(y_{i+1}^o - \mathbf{h}_{i+1}\mathbf{x}_i^a)\}, \tag{11}$$

where α is a scalar localization weight (regression factor) applied to the ensemble sensitivity estimates.[1] When a multivariate regression is used, we can apply the factor α to the projection of the entire perturbation on the response. When the univariate form is adopted, the expression in Eq. (11) simplifies and a scalar regression factor appears anywhere in the product defining the predicted response. The method used here for choosing α is briefly described in section 3b.

Weights $\alpha \in \boldsymbol{\alpha}$; the localization reduces the predicted response. In the presence of sampling error, accuracy of the predicted response should improve from applying an appropriate $\boldsymbol{\alpha}$. Experiments described in the next section are designed to quantify its effectiveness, as well as the effects of considering multivariate sensitivity under different conditions.

3. Experiments

Experiments focus on validating the accuracy of the predicted response. A general strategy is to apply a perturbation $\delta\mathbf{x}^a$ to the initial state, then run the nonlinear model forward to the forecast time. The resulting change in response function can then be compared to the change predicted from the sensitivities and the known perturbation. A simple two-scale model provides a means for considering different types of dynamics, model error, and levels of noise in the system while generating adequate samples from which to draw robust conclusions.

a. Model

A relatively low-dimensional two-scale model described in L05 (model III) forms the basis for experiments. For an atmospheric analog, this model improves on the widely used models documented in Lorenz (1995) because the dominant waves in L05 result in strong spatial correlations between neighboring grid points. A summary of the most relevant parts of Lorenz (2005) follows, and we refer the interested reader to that paper for further details.

Using grid index n over N grid points, the model is written as

$$\frac{dZ_n}{dt} = [X,X]_{K,n} + b^2[Y,Y]_{1,n} + c[Y,X]_{1,n} - X_n - bY_n + F. \tag{12}$$

Here, Z is the prognostic variable, which has contributions from the slowly varying X and the quickly varying Y variables, defined below. The constant K is chosen to be much smaller than N, and an additional constant $J = K/2$ when K is even and $J = (K-1)/2$ when K is odd. The coefficient b determines the frequency and amplitude relationship between X and Y, and is chosen to be 10 so that Y varies 10 times faster than X and with one-tenth of the amplitude. The coupling coefficient c determines how strongly X and Y force each other. The forcing term F is chosen to be 15.

The advection terms denoted by the square brackets in Eq. (12) are formulated by introducing a special sum, Σ', which is like the usual sum except the first and last terms are divided by 2. Thus,

$$[X,X]_{K,n} = -W_{n-2K}W_{n-K} + \sum_{j=-J}^{J}{}' W_{n-K+j}X_{n+K+j}/K$$
$$W_n = \sum_{j=-J}^{J}{}' X_{n-i}/K, \tag{13}$$

and Σ' reverts to the usual sum when K is odd.

Separation of scales is achieved with a simple filter, constructed to ensure the conservation of quadratic quantities. Thus,

$$X_n = \sum_{i=-I}^{I}{}' (\gamma - \beta|i|)Z_{n+i}$$
$$Y_n = Z_n - X_n, \tag{14}$$

where γ, β, and I are chosen such that the X_n will equal Z_n whenever Z varies quadratically over the interval $n-I$ through $n+I$. The result follows, for free parameter I:

$$\gamma = (3I^2+3)/(2I^3+4I),$$
$$\beta = (2I^2+1)/(I^4+2I^2). \tag{15}$$

Variable Z_n can be thought of as a scalar quantity on N grid points around a latitude circle. We choose $N = 960$, as in L05, such that the grid spacing is 0.375°. Two scales, X and Y, are superimposed to produce Z. The choice of I determines the scale separation because it controls the length of the filter in Eq. (14). The choice of K, with implied J, determines the number of slow waves on the latitude circle. As pointed out by L05, this is a fundamental difference from earlier models, where the wavelength was fixed and adding grid points simply added more waves without changing the dynamics.

As for the Lorenz (1995) model, time normalization is such that each nondimensional time step of 0.001 is equivalent to 432 s. Equation (12) is integrated with the fourth-order Runge–Kutta scheme. Table 2 summarizes the parameter values for this model implementation.

An alternate model is easily created from model III by setting $Y = 0$ in Eq. (12), thereby eliminating the fast

[1] When more than one observation is included in Eq. (11), then the vector $\boldsymbol{\alpha}$ is needed.

TABLE 2. Summary of model parameters.

Parameter	Value
N	960
b	10
c	3
F	15
K	32
J	16
I	12
Δt	0.001

scales. Called model II in L05, it can be used in two experiment variants. First, model II can be considered to be perfect in experiments that consider only slow scales. Second, model III can be considered the truth, while model II can be a deficient model used for assimilation. In this case, the model lacks the energy exchanges between scales that are present in the truth, and observations of the truth will contain representativeness error relative to the model.

b. *Estimates for sensitivity localization*

Although we have freedom to choose a forecast response J, here we choose the root-mean-square error (RMSE) of the ensemble mean for relevance to real forecast problems. Minimizing forecast RMSE is a common goal when considering new observation siting, for example, and ensemble sensitivities can only be useful toward that goal if they are accurate.

Sampling error mitigation in the sensitivity estimates follows the Bayesian hierarchical filter proposed by Anderson (2007). It is straightforward to apply, and the goal here is simply to determine whether damping the sensitivity covariance is important. A factor $\alpha = [0, 1]$ is estimated by running multiple groups of ensemble filters over the experiment period. Comparing the regression coefficients among the groups to the mean of the coefficients across the groups is one measure of the relative sampling error based on signal-to-noise concepts. When the diversity of coefficients among the groups is the same order as the mean coefficient, it is noisy and leads to a small α. Conversely, when the diversity is small relative to the mean, we have confidence in the estimate of the coefficient and α approaches one. For each sensitivity estimate, α is obtained from running the hierarchical filter with ensemble distributions sampled at that particular location and time, and applied in Eq. (11). One could alternately choose to archive estimated regression factors for later averaging, producing an empirical set of factors to apply in all other cases, following Anderson (2007). For simplicity here we apply the regression factors estimated instantaneously and at each sensitivity point. Further details do not aid interpretation of the results herein, and we refer the reader to Anderson (2007).

c. *Design*

Initial conditions (ICs) for the nature run and N ensemble members are drawn from climatologies, which are separately generated for models II and III. Synthetic observations are generated every 6 h by applying the forward operator to the state-of-the-nature run, then adding random perturbations drawn from a normal distribution with mean zero and unit variance. Two observation networks are examined: one is the even grid points on the domain (480 observing locations), and the other is every grid point in half the domain (also 480 observing locations).

Three assimilation experiments provide context for each observation network. The first assimilation experiment uses the single-scale model II and observations generated from the nature run with model II; that is, it assumes a perfect model II. The second assimilation experiment is the same except it uses the two-scale model III. The third assimilation experiment uses the single-scale model II for assimilation, but observations contain two scales from model III; that is, the model is imperfect. Each assimilation experiment starts with ensemble initial conditions and assimilates the 6-hourly synthetic observations through the ensemble adjustment Kalman filter (EAKF; Anderson 2003).

Ensemble size is set to 40 for all assimilation experiments. Inflation and localization are used to mitigate model error and sampling error during the experiments. Following Hamill et al. (2001), a single value greater than 1.0 is used to inflate the prior ensemble spread, in order to maintain appropriate ensemble spread. The Gaspari and Cohn (1999) correlation function is used for localization. Inflation and localization parameters are manually tuned for each assimilation experiment to produce the lowest 6-h forecast RMSE.

The hierarchical filter, providing regression factors for the ensemble sensitivity, makes use of four concurrently cycling ensemble assimilation systems. Each identically implemented system has a unique set, or ''group,'' of ensemble members. The total number of members in the four groups together is $4 \times 40 = 160$. As described above, regression factors estimated from these groups are valid at individual locations and times.

All assimilation experiments are for 30 days (i.e., 120 data assimilation cycles). The first 40 cycles are discarded, and the last 80 cycles are used to explore the ensemble sensitivity. After all sensitivities are computed, perturbations to the initial conditions at each assimilation time are used to assess the accuracy of the sensitivities and the forecast response expected from

the perturbation. Experiments from forming perturbations based on each of 20 randomly selected grid points give a total of $80 \times 20 = 1600$ trials for each configuration.

Three perturbation approaches are tested, and comparisons aid in the interpretation of the results. First, a perturbation equal to one standard deviation is introduced to an individual grid point. The perturbation is regressed onto the remaining analysis state using the analysis statistics. The effect is a change in the ensemble mean, but not the spread, and is similar to the approach in Torn and Hakim (2008) to assess the linearity of the response. The second experiment is the same as the first except that the regression onto the analysis is spatially localized in the same way as it would be applied in the data assimilation. Again, the ensemble mean is perturbed, but not the spread. The purpose of this perturbation is to compare against the third perturbation method. Third, an observation of the truth is assimilated with the ensemble filter. The observation here is the truth value plus an error drawn from the specified observation error distribution. The first two perturbation approaches are characterized by a grid-point perturbation applied directly, and regressed to the analysis state without considering observation or forecast error statistics, as is done during the data assimilation. We refer to these first two as ''direct perturbation'' approaches.

4. Results and discussion

The accuracy of the predicted response is evaluated by comparison against the actual response measured from nonlinear simulations from the perturbed initial conditions. Plotting the predicted response versus the actual response from all 1600 trials on a scatter diagram gives a comprehensive summary. Greater distances from the 1:1 (diagonal) line indicate less accurate response predictions. Those are summarized by an RMSE value. The first tests assume model II in L05 is perfect; both the nature run and the assimilating model contain only synoptic-scale dynamics. Following that, we introduce faster scales then model error.

a. Slow-scale, perfect-model experiments

In all panels in Fig. 1 (and all scatterplots in the remainder of the paper), results from the univariate and multivariate regressions are shown in blue and red, respectively. The legend reports RMSE and linear correlation (r^2) values associated with the respective least squares best-fit lines of each. Because the actual response to the perturbation does not depend on the method for sensitivity calculation, the ordinate values for both univariate and multivariate regression results are the same. Further, in Fig. 1a the results from both regressions are identical because the product of the sensitivity and an arbitrary perturbation (here one standard deviation of the analysis) is the same. Localization affects the expected responses differently, but here no localization is used on either the sensitivity or the perturbation. Comparing results down a column of panels in Fig. 1 shows the dependence on perturbation method and amplitude ($\mathbf{1}\boldsymbol{\sigma}^a$ regressed onto analysis, $\mathbf{1}\boldsymbol{\sigma}^a$ perturbation regressed with localization $\boldsymbol{\rho}$, and assimilating an observation with localization). Comparing results across a row shows the effects of localization on the response predicted from sensitivities ($\boldsymbol{\alpha}$ applies in the right column).

Sensitivities when only slow dynamics are present demonstrate general robustness to regression approaches and localization (Fig. 1). RMSE values for the univariate and multivariate methods are similar in all cases. When a perturbation is constructed regressing a $\mathbf{1}\boldsymbol{\sigma}^a$ value applied at the observation location, localization has little effect (cf. Figs. 1a,c and Figs. 1b,d). Applying $\boldsymbol{\alpha}$ also appears to have little effect (left versus right panels). Using an observation error of 1.0 when assimilating a new observation leads to a small perturbation, which is clear from the axis ranges in Figs. 1e and 1f.

The positive shift of the cluster of points from the origin, when a perturbation is applied via assimilation (Figs. 1e,f), is a consequence of the experiment design. The observation error is specified too large (1.0) in these experiments, illustrating the importance of observation errors in predicting responses. This can be understood as follows. The predicted $\mathbf{J}$ (i.e., the forecast error) is positive definite, and will be predicted smaller in response to a perturbation only when both $\partial J/\partial x^a$ and δx^a are of opposite signs so that δJ is negative. Here, $\partial J/\partial x^a$ is given by the ensemble; δx^a results from observing a randomly selected discrete location in the true state, which can be either less than or greater than the ensemble mean analysis (it is a coin flip). Therefore, the marginal distributions of the predicted response are distributed randomly about zero, which can be seen along the abscissa in Figs. 1e and 1f. The response from integrating the nonlinear model is biased positive because a single observation with error variance greater than the analysis ensemble-mean error leads to that individual assimilated observation degrading the analysis. The marginal distribution of the actual response measured from nonlinear simulations shifts positively along the ordinate. Note that this will not occur when many randomly located observations are concurrently assimilated because the positive and negative observation increments will largely cancel. Direct perturbation at a single grid point (Figs. 1a–d) does not consider

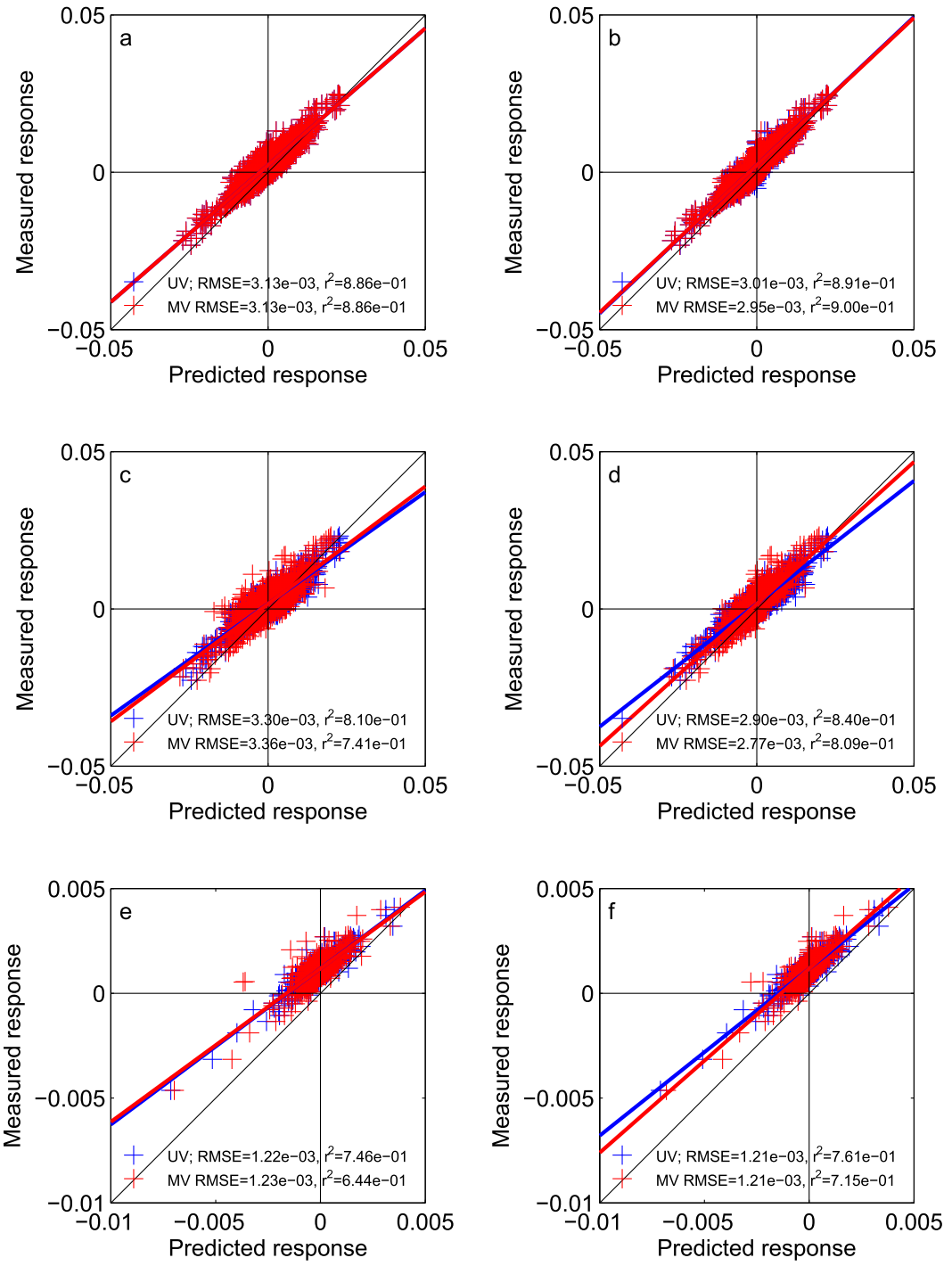

FIG. 1. Actual (nonlinear) perturbation responses against responses predicted from sensitivity estimates. Blue and red show univariate (scalar) and multivariate regressions in the sensitivity estimates, respectively. (a),(c),(e) Results without localization on the sensitivity estimates, and (b),(d),(f) results including localization. Shown are results taken from (a),(b) a perturbation equal to one standard deviation of the analysis; and (c),(d) a perturbation of one standard deviation of the analysis and the use of localization on the perturbation. (e),(f) Results from assimilating an observation of the truth, including localization in the assimilation step. Least squares best-fit lines, RMSE, and correlations (r^2) are also reported.

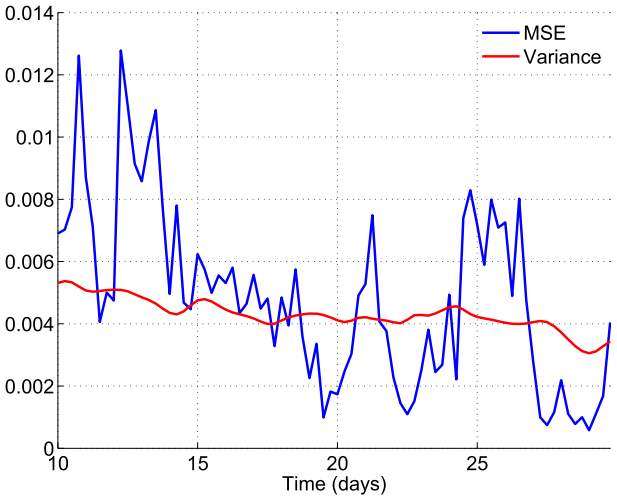

FIG. 2. Slow-scale perfect-model ensemble-mean analysis mean squared error (blue) and ensemble spread (red) at an arbitrary grid point.

observation error and, therefore, lacks the positive response bias.

In real weather prediction applications of ensemble data assimilations, forecast errors and observation errors are the same magnitude, but Fig. 2 shows that in this slow-scale, perfect-model experiment the forecast error is $\mathcal{O}(0.001)$, much less than the observation error variance of 1.0. Figure 3 shows that specifying an observation error variance of 0.01, still larger than the errors, but representing a much higher quality observation, effectively reduces (but does not eliminate) the positive bias in the actual response measured from nonlinear simulations. Perturbed observations of the true state are much less likely to make the forecast RMSE greater. The forecast RMSE response is more evenly distributed around zero. Predicted and actual responses also increase by an order of magnitude, both because the ensemble spread is smaller in the data assimilation and because the higher quality observations can result in a larger analysis increment. Because they are now the same order of magnitude, the behavior of the predicted response to perturbations from direct perturbation and assimilation is easier to compare.

Results from slow-scale, perfect-model experiments are consistent with published results that use the univariate sensitivity formulation. Ancell and Hakim (2007) and Torn and Hakim (2008) both showed that predicted responses verified reasonably well against nonlinear forecast responses. They did not address localization or directly test assimilation with the ensemble filter. Rather than choose random locations to form perturbations, they chose regions showing maximum sensitivity. In the assimilation experiments reported here, choosing only locations where the predicted response would favor reduced forecast error (negative δJ). Choosing those location would to some degree mask the effect of the overspecified observation errors. In the next section we test the response predictions under more demanding conditions, where faster scales are present.

b. *Two-scale, perfect-model experiments*

Model III from L05 is the dynamical system for these experiments; both slow and fast scales are included. Based on the results in the last section, a smaller observation error of 0.01 is retained. All other parameters and experiment design considerations are unchanged.

Results show that the faster scale elicits response predictions with a greater dependence on the sensitivity calculation method and the localization (Fig. 4). The multivariate sensitivity consistently provides a more accurate prediction of the response in the nonlinear model. Direct perturbation with localization, using the

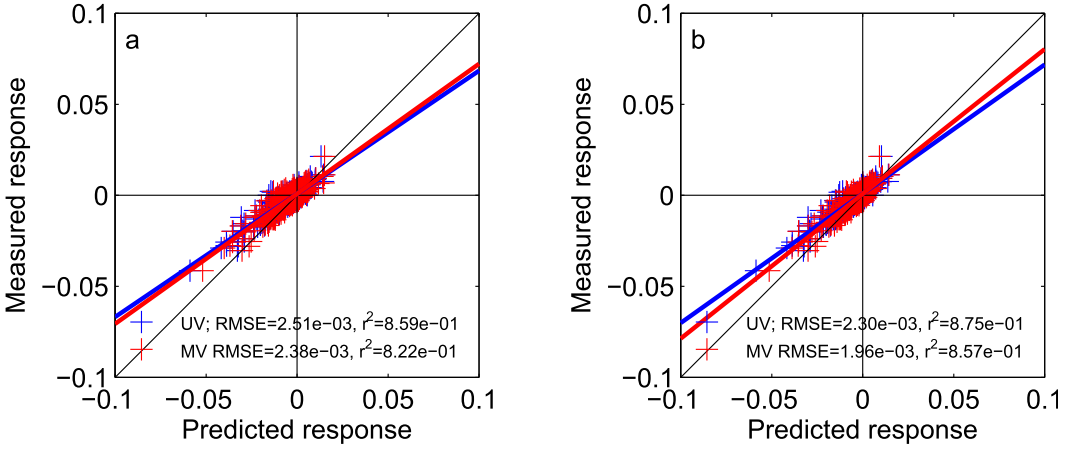

FIG. 3. As in Figs. 1e and 1f, but with observation error variance of 0.01.

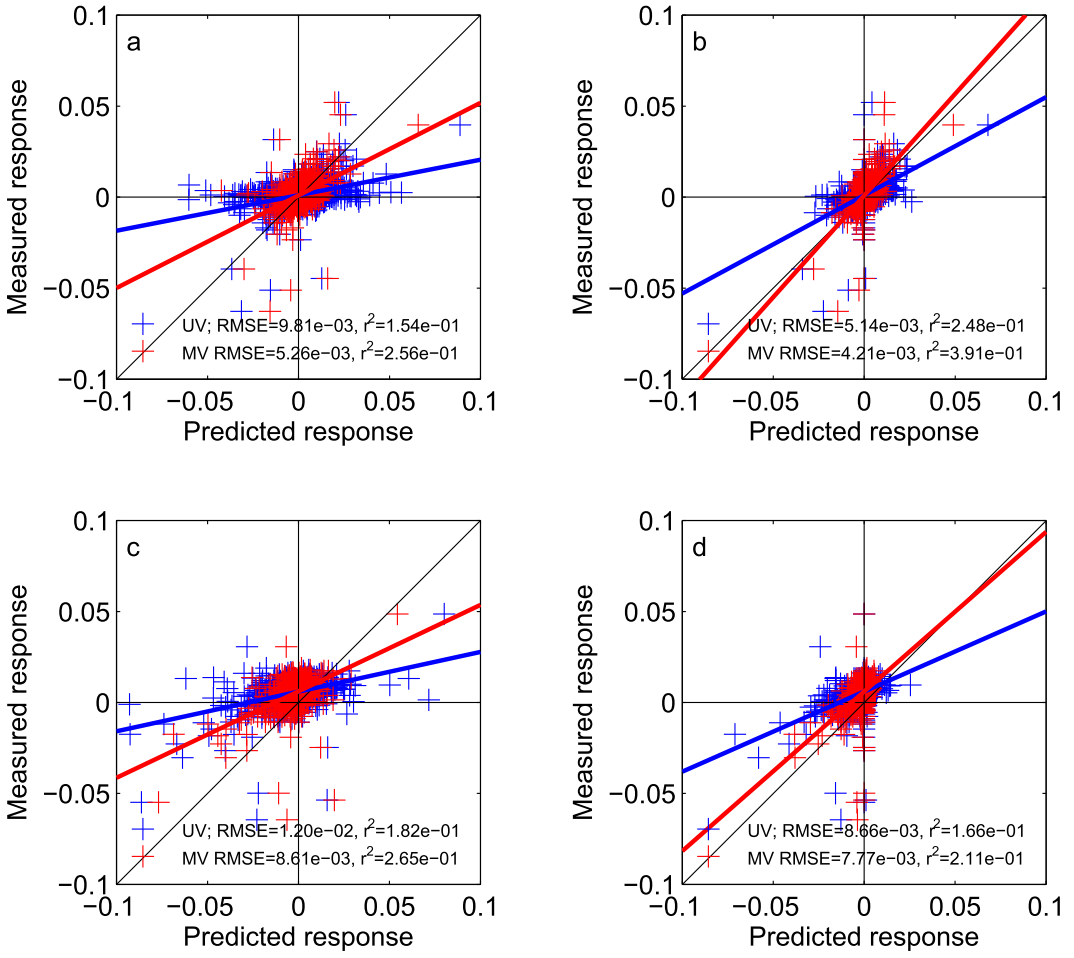

FIG. 4. As in Figs. 1c–f, but from a perfect model containing both slow and fast scales (model III), and using an observation error of 0.01.

nonlocalized sensitivity, results in a 46% reduction in the RMSE of the response predictions from the multivariate sensitivities compared to the univariate sensitivities (Fig. 4a). The corresponding assimilation experiment results show a 28% improvement (Fig. 4c). The best-fit lines to the univariate results in both Figs. 4a and 4c are clearly flatter than the diagonal one-to-one line that indicates a perfect prediction. This indicates that the univariate sensitivities overpredict the response magnitude compared to the multivariate sensitivities. Compared to the results from the single slow-scale model, greater scatter about the diagonal line here results from the faster scale. Nonlinear error growth, expected to be greater when the fast scale is included in the dynamics, can cause individual points to deviate from the diagonal. Because the ensemble sensitivity method examined here is linear, it is not clear that the sensitivity estimates can be improved for this particular choice of forecast metric. Spatial or temporal averaging in either the initial state or the forecast metric may produce a more linear response, and narrow the scatter in these results.

Localizing the predicted responses leads to more accurate response predictions. All of the RMSE values are

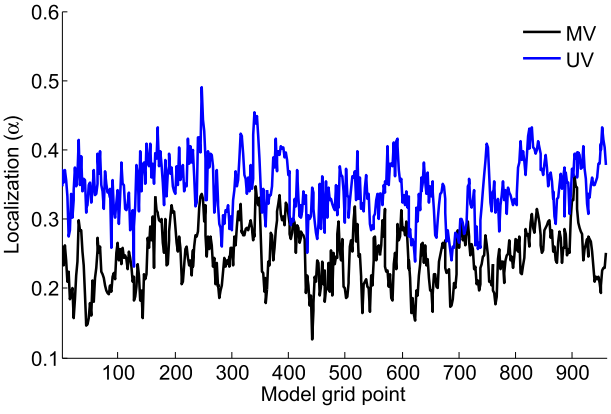

FIG. 5. Average localization factors for the scalar univariate regressions (blue) and the multivariate regressions (black).

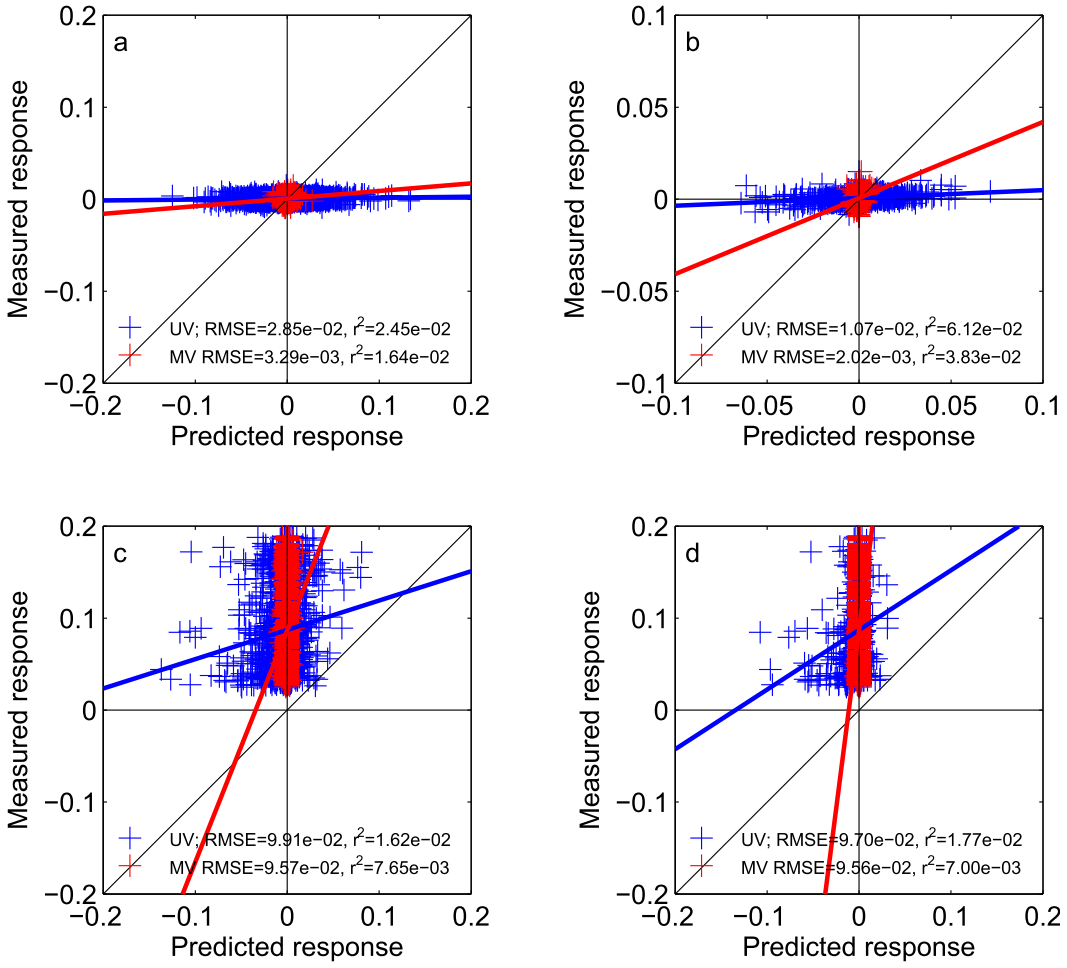

FIG. 6. As in Fig. 4, but for an imperfect model containing only slow scales, with a truth model containing slow and fast scales.

smaller in the right-hand panels, compared to the left in Fig. 4. Results in Figs. 4b and 4d also suggest further improvements in the sensitivity localization method may lead to improved response predictions. Figure 4d shows some points spread along the ordinate. These are points where the predicted response was nearly eliminated by the sensitivity localization, but the response in the nonlinear simulations remained. The RMSE values in Fig. 4d are sensitive to those points. Clearly, the localization eliminated too much in the predicted response for that handful of measurements.

The overpredicted responses from the univariate sensitivities have relatively more to gain from localization, and it is borne out in the results (e.g., Fig. 5), which shows the time-averaged localization factors valid at each grid point in the two-scale perfect-model experiments. Generally, small localization factors for both sensitivity calculation methods are clear, reducing the sensitivity to a perturbation at all grid points to less than half of what would be predicted from the directly sampled ensemble statistics. This results from weak sensitivities that are difficult to detect in the presence of sampling error. Because the relationship between $\mathbf{J}_e$ and any one predictor in the multivariate regression is small compared to the univariate case, the relative noise in the individual grid-point sensitivity estimates is greater. Thus, each requires more stringent localization than in the univariate case, where only the effects of a single grid point are considered. The single predictor in the univariate case requires less localization because the regression gives the stronger signal, which also manifests in the overpredicted sensitivities.

An experiment with a large ensemble confirms the expected effect of reducing sampling errors. Compared to 40 ensemble members used in most of these experiments, employing 1000 ensemble members reduces the RMSE of sensitivity estimates in Fig. 4 by an order of magnitude. Error in estimating the regression

factors is greater because the difference between groups of ensembles is much smaller, but in any case result in factors approximately twice the value of those in Fig. 5 to reflect greater confidence in the sensitivity regressions.

Torn and Hakim (2008) noted that failing to localize the perturbation (when regressing the perturbations as in Fig. 4a) can be expected to give overpredicted responses. Results here show that overprediction can persist even with localization on the perturbation, which reduces $\delta\mathbf{x}^a$. Applying $\boldsymbol{\alpha}$ also reduces the predicted response by lessening the sensitivities, but the details of the localization (regression factor) may need further attention.

c. *Imperfect model experiments*

Here, the assimilating model is fundamentally inadequate relative to the truth. Model III is truth; model II is the dynamical system used to assimilate data and provide samples for sensitivity estimates. The true dynamics contain scale interactions not present in the assimilating model. Representativeness error also exists when observing the true state. The experiment configuration is analogous to having a model of poorer quality than our current NWP models. The sensitivity itself is agnostic to the model from which samples are taken. Actual responses from the imperfect model, and over the perturbation magnitudes in section 4a (i.e., Fig. 3), should be qualitatively similar. Larger perturbations can result from assimilating high quality observations of a true system that differs from the model, and analysis ensemble spread can be large at unobserved locations. The analysis ensemble is also non-Gaussian (not shown) because model inadequacy leads to a suboptimal filter. The model is biased and lacks the small-scale variability, compared to the truth. The sensitivities, and also the imposed perturbations, cannot be expected to adhere to the linear theory.

Results depend strongly on both the method for calculating sensitivities and the method for introducing perturbations (Fig. 6). First, predicted responses from direct perturbations are much more accurately predicted from multivariate sensitivities; prediction errors are an order of magnitude smaller (Figs. 6a,b). The results indicate that sensitivities estimated with a poor model are better able to predict the response to a perturbation that contains very little noise. Here, the only noise introduced is from sampling error when regressing the perturbations onto the analysis; the localization cannot eliminate all the effects of sampling error. Univariate sensitivities strongly overpredict the response compared to multivariate sensitivities. Applying $\boldsymbol{\alpha}$ improves the results by a greater factor than for any of the perfect model experiments above.

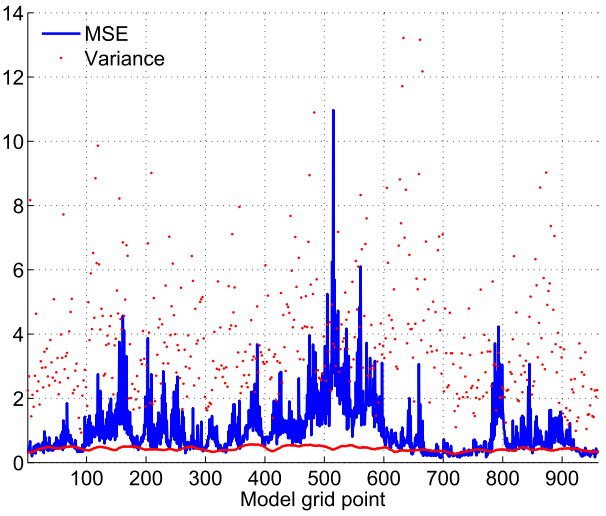

FIG. 7. Example forecast (background) spread (red dots) and ensemble-mean error (blue) for the imperfect model experiments.

Comparing the RMSE values reported in Fig. 6b to those in Fig. 6a, predictions from the univariate and multivariate sensitivities improve by 57% and 38%, respectively.

The positive bias in the response from the nonlinear model, pointed out in the last two sections, combines with the effects of the model error here (Figs. 6c,d). The analysis ensemble spread at unobserved grid points is large; the observed grid points show lower error and spread, while the unobserved grid points show much higher error and spread. Figure 7 provides an example. The large ensemble spread leads to a small sensitivity and thus a small predicted response. The larger error also leads to larger observation increments, and consequently large analysis increments, in the assimilation. The increments elicit a nonlinear (and positively biased) response where none was predicted.

In addition to demonstrating the advantages of multivariate sensitivities, results in this section highlight the need for a well-performing data assimilation system from which to estimate sensitivities. The extreme model errors in these experiments prevent an objective assimilation system that assumes Gaussian statistics, such as the ensemble filter, from approaching optimality. Fortunately, experience shows that our mesoscale models are probably not quite so poor as this example. We next present a complementary experiment that introduces a more realistic scenario of error growth in the ensemble, while examining both perfect and imperfect model scenarios.

d. *A data void*

Here, we consider a network such that half of the state (in space) is perfectly observed, while the other half is

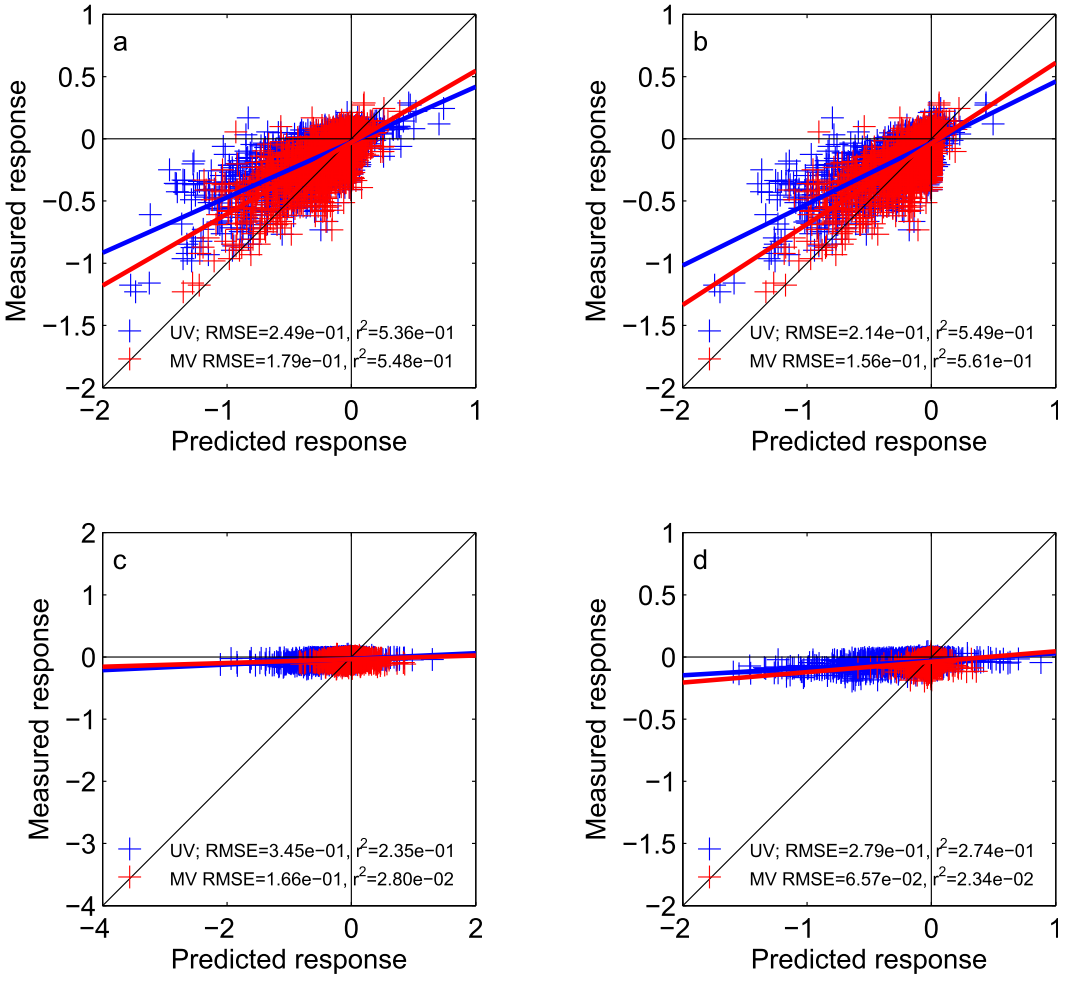

FIG. 8. As in Fig. 6, but for an observing network on half the domain and a data void on the other. Results for (a),(b) a perfect model with slow and fast scales and (c),(d) an imperfect model. As in earlier figures, the left side shows results without localizing the sensitivities, and the right side shows results with localization imposed.

completely void of observations. The idea is to allow the ensemble to produce more spread in half the domain, and identify how the sensitivities can predict the effect of any one observation. This observing problem is more analogous to the observing problems considered by Ancell and Hakim (2007), who examined the effect of surface pressure observations over the North Pacific data void. We present both perfect-model results with model III and imperfect model results.

Whether or not model error is present, the multivariate sensitivities clearly improve the response predictions (Fig. 8). The univariate sensitivities strongly overpredict the response; although much more accurate, the multivariate sensitivities also overpredict the perturbation response. Perfect-model results show some gain from localizing the predicted sensitivity response (cf. Figs. 8a,b), but the imperfect model results show a greater benefit (cf. Figs. 8c,d). Except for a few outliers, the multivariate sensitivities with localization lead to a much smaller overprediction problem. The localization reduces the impact from large-sensitivity elements in the well-observed part of the domain that are contributing to the sensitivity. The univariate sensitivity in the data void does not account for large-sensitivity contributions from the well-observed region and, instead, attributes all sensitivity to individual analysis locations. In both cases, localization certainly does not result in perfect response predictions.

Individual grid-point sensitivities are smaller over the data void because analysis uncertainty is larger in the absence of observations (Fig. 9a). The multivariate sensitivities account for it; although weighted, by combining information where analysis uncertainty is both small and large, the extremes of the univariate sensitivity distributions are pulled toward the origin in Fig. 8. The sensitivities are on average greater in the data void

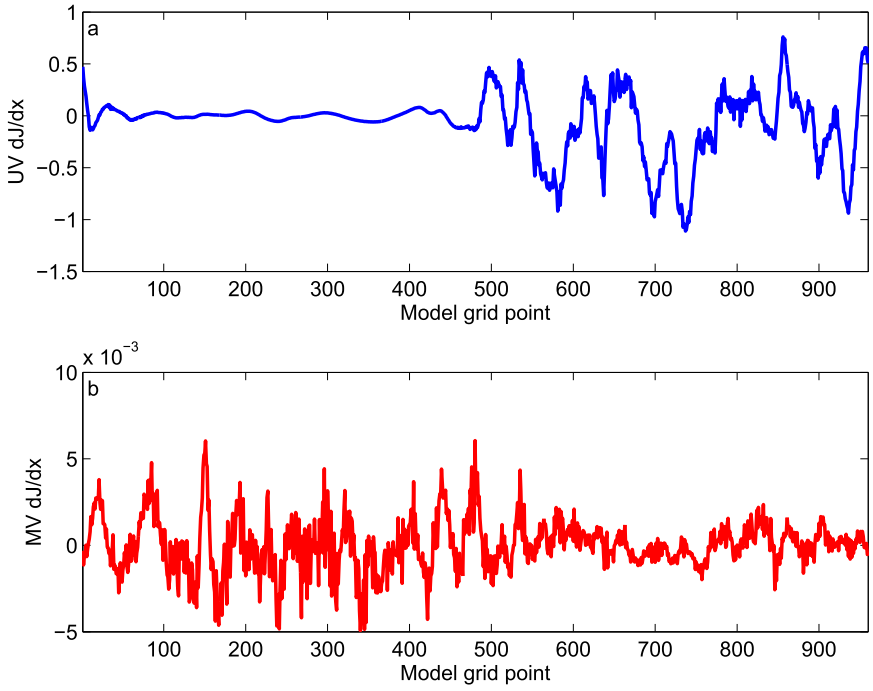

FIG. 9. (a) Scalar and (b) multivariate sensitivities averaged over 80 data assimilation cycles. Observations are on the right half of the domain, and a data void is on the left. Results are for the perfect two-scale model.

(Fig. 9b), which is intuitive because any single observation in the data void should reduce the forecast error more effectively.

5. Summary and conclusions

This work addresses two open issues through the use of ensemble sensitivities to estimate a perturbation response in mesoscale models: sampling error and the use of a diagonal approximation to the analysis covariance matrix in the regressions underlying the sensitivities. First, ensemble sensitivities are derived from ordinary least squares. It becomes apparent where sampling error appears in the regressions used to solve the least squares problem. It is also clear where the oft-used diagonal approximation in the predictor (covariance) appears and how it simplifies the problem. A regression factor, inspired by the hierarchical filter from Anderson (2007), is proposed to mitigate the sampling errors in the sensitivity estimates. Using the Lorenz (2005) two-scale model in a cycling ensemble data assimilation system, this study quantifies the effects of the regression factor (localization) and the univariate approximation to the multivariate regression on the effectiveness of the ensemble sensitivities on predicting a response to an analysis perturbation. Results demonstrate that damping poorly sampled covariances with a regression factor, and use of the complete multivariate regression, in the sensitivities can improve the perturbation response prediction under certain circumstances relevant to mesoscale problems.

Primary conclusions include the following:

- Under slow dynamics when covariances are strong and easily sampled, the diagonal approximation to the analysis covariances leads to skillful predictions of a forecast response. Localization of the sensitivities to mitigate the effects of sampling error has little effect.
- When fast scales are also present, sensitivities are not as easy to estimate. Individual correlations are weaker, and the multivariate sensitivity proves to be more effective at predicting responses. Sensitivity localization improves predicted responses from both univariate and multivariate sensitivities.
- Model error leads to a less optimal assimilation system from which to estimate sensitivities, and the univariate sensitivities are more prone to overpredicting the response.
- The effects of model error and fast scales are exacerbated by the large analysis ensemble spread in the data void. Multivariate sensitivities provide more accurate response predictions when proposing new observations in a data void, particularly when model error and fast scales are present.

The next step is to test these results in a real model. The weakly forced Great Salt Lake fog case of WHC presents one possibility. A strongly forced case such as one of the downslope winds analyzed by Reinecke and Durran (2009) would present a useful contrast. Cycling an ensemble data assimilation system long enough to produce samples for estimating sensitivity regression factors, which could be averaged and applied in a smaller ensemble, presents a computational challenge. But it is tractable on today's supercomputers.

Acknowledgments. This work was funded in part by the Office of Naval Research under the Mountain Terrain Atmospheric Modeling and Observation Program (MATERHORN) while the lead author was in residence at the Naval Postgraduate School and, in part, by the Army Test and Evaluation Command (ATEC) through an interagency agreement with the National Science Foundation. Discussions with Steven Thomas helped clarify the matrix response and implications for the minimum-norm matrix inversion.

REFERENCES

Ancell, B., and G. Hakim, 2007: Comparing adjoint- and ensemble-sensitivity analysis with applications to observation targeting. *Mon. Wea. Rev.,* **135,** 4117–4134, doi:10.1175/2007MWR1904.1.

Anderson, J. L., 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.,* **131,** 634–642, doi:10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2.

——, 2007: Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D,* **230,** 99–111, doi:10.1016/j.physd.2006.02.011.

Bishop, C. H., and Z. Toth, 1999: Ensemble transformation and adaptive observations. *J. Atmos. Sci.,* **56,** 1748–1765, doi:10.1175/1520-0469(1999)056<1748:ETAAO>2.0.CO;2.

——, and D. Hodyss, 2009: Ensemble covariances adaptively localized with ECO-RAP. Part 2: A strategy for the atmosphere. *Tellus,* **61A,** 97–111, doi:10.1111/j.1600-0870.2008.00372.x.

——, B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.,* **129,** 420–436, doi:10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.

Buizza, R., and A. Montani, 1999: Targeting observations using singular vectors. *J. Atmos. Sci.,* **56,** 2965–2985, doi:10.1175/1520-0469(1999)056<2965:TOUSV>2.0.CO;2.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.,* **125,** 723–757, doi:10.1002/qj.49712555417.

Gelaro, R., R. H. Langland, G. D. Rohaly, and T. E. Rosmond, 1999: As assessment of the singular vector approach to targeted observing using the FASTEX dataset. *Quart. J. Roy. Meteor. Soc.,* **125,** 3299–3327, doi:10.1002/qj.49712556109.

Golub, G. H., and C. F. Van Loan, 1996: *Matrix Computations.* The Johns Hopkins University Press, 694 pp.

Gombos, D., and J. A. Hansen, 2008: Potential vorticity regression and its relationship to dynamical piecewise inversion. *Mon. Wea. Rev.,* **136,** 2668–2682, doi:10.1175/2007MWR2165.1.

——, R. N. Hoffman, and J. A. Hansen, 2012: Ensemble statistics for diagnosing dynamics: Tropical cyclone track forecasts sensitivities revealed by ensemble regression. *Mon. Wea. Rev.,* **140,** 2647–2669, doi:10.1175/MWR-D-11-00002.1.

Hakim, G. J., and R. D. Torn, 2008: Ensemble synoptic analysis. *Synoptic–Dynamic Meteorology and Weather Analysis and Forecasting: A Tribute to Fred Sanders, Meteor. Monogr.,* No. 55, Amer. Meteor. Soc., 147–161.

Hamill, T. M., and C. Snyder, 2002: Using improved background error covariances from an ensemble Kalman filter for adaptive observations. *Mon. Wea. Rev.,* **130,** 1552–1572, doi:10.1175/1520-0493(2002)130<1552:UIBECF>2.0.CO;2.

——, J. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.,* **129,** 2776–2790, doi:10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.

Houtekamer, P. L., and H. L. Mitchell, 2005: Ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.,* **131,** 3269–3289, doi:10.1256/qj.05.135.

Langland, R. H., and N. L. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus,* **56A,** 189–201, doi:10.1111/j.1600-0870.2004.00056.x.

——, and Coauthors, 1999: The North Pacific Experiment (NORPEX-98): Targeted observations for improved North American weather forecasts. *Bull. Amer. Meteor. Soc.,* **80,** 1363–1384, doi:10.1175/1520-0477(1999)080<1363:TNPENT>2.0.CO;2.

Lorenz, E. N., 1995: Predictability—A problem partly solved. *Proc. Seminar on Predictability,* Vol. 1, Reading, United Kingdom, ECMWF, 1–18.

——, 2005: Designing chaotic models. *J. Atmos. Sci.,* **62,** 1574–1587, doi:10.1175/JAS3430.1.

Reinecke, P., and D. Durran, 2009: Initial-condition sensitivities and the predictability of downslope winds. *J. Atmos. Sci.,* **66,** 3401–3418, doi:10.1175/2009JAS3023.1.

Torn, R., and G. Hakim, 2008: Ensemble-based sensitivity analysis. *Mon. Wea. Rev.,* **136,** 663–677, doi:10.1175/2007MWR2132.1.

——, and ——, 2009: Initial-condition sensitivity of western Pacific extratropical transitions determined using ensemble-based sensitivity analysis. *Mon. Wea. Rev.,* **137,** 3388–3406, doi:10.1175/2009MWR2879.1.