2018

Ensemble Sensitivity Analysis for Targeted Observations of Supercell Thunderstorms

George L. Limpert

Adam L. Houston

Ensemble Sensitivity Analysis for Targeted Observations of Supercell Thunderstorms

GEORGE L. LIMPERT AND ADAM L. HOUSTON

Department of Earth and Atmospheric Sciences, University of Nebraska–Lincoln, Lincoln, Nebraska

(Manuscript received 10 February 2017, in final form 8 February 2018)

ABSTRACT

Ensemble sensitivity analysis (ESA) has been demonstrated for observation targeting of synoptic-scale and mesoscale phenomena, but could have similar applications for storm-scale observations with mobile platforms. This paper demonstrates storm-scale ESA using an idealized supercell simulated with a 101-member CM1 ensemble. Correlation coefficients are used as a measure of sensitivity and are derived from single-variable and multivariable linear regressions of pressure, temperature, humidity, and wind with forecast response variables intended as proxies for the strength of supercells. This approach is suitable for targeting observing platforms that simultaneously measure multiple base-state variables. Although the individual correlations are found to be noisy and difficult to interpret, averaging across small areas of the domain and over the duration of the simulation is found to simplify the analysis. However, it is difficult to identify physically meaningful results from the sensitivity calculations, and evaluation of the results suggests that the overall skill would be low in targeting observations at the storm scale solely based on these sensitivity calculations. The difficulty in applying ESA at the scale of an individual supercell is likely due to applying the linear model to an environment with highly nonlinear dynamics, rapidly changing forecast metrics, and autocorrelation.

1. Introduction

Observation targeting is the deployment of limited or mobile observing assets based on estimates of where they will provide the greatest benefit to forecasts (e.g., Lorenz and Emanuel 1998; Buizza and Montani 1999; Langland 1999; Morss et al. 2001; Szunyogh et al. 2002; Aberson 2003; Weissmann et al. 2005; Ancell and Hakim 2007; Wu et al. 2007; Torn and Hakim 2009; Garcies and Homar 2010; Torn 2010; Irvine et al. 2011; Chang et al. 2013; Xie et al. 2013; Zheng et al. 2013; Li et al. 2014; Wile et al. 2015; Romine et al. 2016). Ensemble sensitivity analysis (ESA; Ancell and Hakim 2007; Torn and Hakim 2008) provides a means to estimate the sensitivity of a dynamical model to perturbations in the initial state. This has applications for targeting observations where ESA predicts they have the potential for the greatest impact on the model solution (Ancell and Hakim 2007). Prior work with ESA has focused primarily on synoptic- through meso-β-scale phenomena in features such as midlatitude and tropical cyclones (e.g., Ancell and Hakim 2007; Hill et al. 2016). In the work presented here, ESA is applied to an idealized supercell to identify regions in which the forecast is most sensitive to perturbations of the initial state and to demonstrate statistical techniques that could improve upon prior applications of ESA.

Several techniques have been developed to assess the impact of adding additional observations; these include ESA (e.g., Ancell and Hakim 2007; Torn and Hakim 2008), adjoint sensitivity analysis or singular vector methods (e.g., Hall et al. 1982; Hall and Cacuci 1983; Hall 1986; Errico and Vukicevic 1992; Buizza and Montani 1999; Wu et al. 2007), observing system simulation experiments (OSSEs) that directly test the impact of assimilating a single additional observation (e.g., Lorenz and Emanuel 1998; Morss et al. 2001; Xie et al. 2013), and data denial experiments. Prior work has shown that ESA and adjoint sensitivity analysis are both useful for targeted observations, but ESA is preferred because of its relative simplicity (Ancell and Hakim 2007). The objective of this paper is to demonstrate storm-scale ESA around a supercell thunderstorm, with the later goal of using ESA to inform targeted observations with mobile platforms such as unmanned aircraft systems (UASs). Because multiple instruments can be attached to a single UAS, this study uses ESA to

Corresponding author: George Limpert, george.limpert@unl.edu

DOI: 10.1175/MWR-D-17-0029.1

estimate the collective impact of measuring temperature, humidity, pressure, and wind, on selected forecast response variables. The eventual goal is for additional storm-scale ESA simulations to be used to inform where to target observations to maximize the impact of UAS resources to improve nowcasting of supercell thunderstorms. Prior work (e.g., Bednarczyk and Ancell 2015; Hill et al. 2016; Yokota et al. 2016) applied ESA to thunderstorm forecasts generated by a convection-allowing ensemble. Despite Hill et al. (2016) noting that, although ensemble members frequently did not exhibit a Gaussian distribution, thus violating a fundamental assumption of ESA, the results were still useful. Thus, it is reasonable to assume that ESA could plausibly be extended to the scale of individual thunderstorms.

ESA assumes a linear relationship between perturbations at a location in the model analysis and the subsequent response, thus larger perturbations should be associated with greater forecast responses. ESA is performed through linear regression (e.g., Ancell and Hakim 2007; Torn and Hakim 2008), in which the slope of the regression line is a commonly used metric to estimate the impact of a given observation on the forecast response. In ESA, an ensemble, usually consisting of 50–100 members, is simulated to generate the ESA inputs. The forecast response is a scalar quantity such as the composite reflectivity associated with deep convection or the pressure at the center of a cyclone. The perturbations are relative to the mean state of a single variable at each point within the domain. The perturbations are the independent variable in a linear regression against the forecast response, and the resulting slope for each point in the domain is an estimate of the sensitivity to observations at that point. This paper extends ESA to estimate the combined sensitivity of a forecast response variable to multiple perturbation variables, which is useful for observing platforms that simultaneously measure several quantities (e.g., temperature, humidity, pressure, and wind).

The methods adopted herein are presented in section 2. Analysis of the results of single-variable and multivariable sensitivity analysis is presented in section 3. The veracity of the multivariable sensitivity calculations are evaluated in section 4. The implications of this work for storm-scale ESA are discussed in section 5.

2. Background and methods

The approach to storm-scale ESA using idealized simulations differs from prior applications of ESA to larger-scale phenomena. Prior work in observation targeting and ESA has used ensembles that were generated as a result of data assimilation performed with an ensemble Kalman filter (EnKF; Anderson 2001; Ancell and Hakim 2007; Torn and Hakim 2008; Hill et al. 2016). These approaches and associated assumptions are not necessarily suitable for idealized storm-scale simulations because no data assimilation is performed and the average model errors are unknown. Related to ensemble creation, ensuring ensemble spread requires a different tactic in an idealized model framework where horizontal homogeneity minimizes variance and model errors are unknown in the absence of observed storm evolution. Finally, unlike a large-scale phenomenon where member-to-member differences in the position of the phenomenon (e.g., position of deep convection initiation) relative to meso-β-scale/synoptic-scale far-field perturbations are small, storm-scale ESA focused on near-storm perturbations is very sensitive to member-to-member position differences. For example, grid point (i, j) may be located in the storm inflow in one ensemble member but in the precipitation core in another. This would dilute sensitivity. Because the intent of this work is to advance understanding toward a vision of targeted surveillance where targeting guidance is based on storm position and storm structure, sensitivity should be assessed in a storm-relative frame of reference.

a. Model configuration

Simulations were conducted using revision 18 of Cloud Model 1 (CM1; Bryan and Fritsch 2002) and executed on the high-performance computing resources of the Holland Computing Center at the University of Nebraska. The model domain had open lateral boundaries, a free-slip lower boundary, a 1-km horizontal grid spacing with a square 240-km domain, and 63 vertical levels stretched, with a 100-m vertical grid spacing at low levels. The base state is horizontally homogeneous and is perturbed with a warm bubble at the center of the domain to force the development of a storm. The Morrison microphysics parameterization was used, with hail as a hydrometeor class instead of graupel.

The ESA was performed on a 101-member ensemble. The control member was generated to have 2500 J kg^{-1} of CAPE and a 0–6-km bulk wind differential of 35 m s^{-1}. Spread was introduced by perturbing the base-state sounding of each member with a standard deviation of 250 J kg^{-1} of CAPE and 2.5 m s^{-1} in the 0–6-km bulk wind differential (Fig. 1). These perturbations are achieved by varying the state variables above 1 km while the ESA is performed using perturbations at 500 m, so the method of introducing perturbations should not directly bias the ESA results. Although the variance in the base-state sounding is significantly less than that of

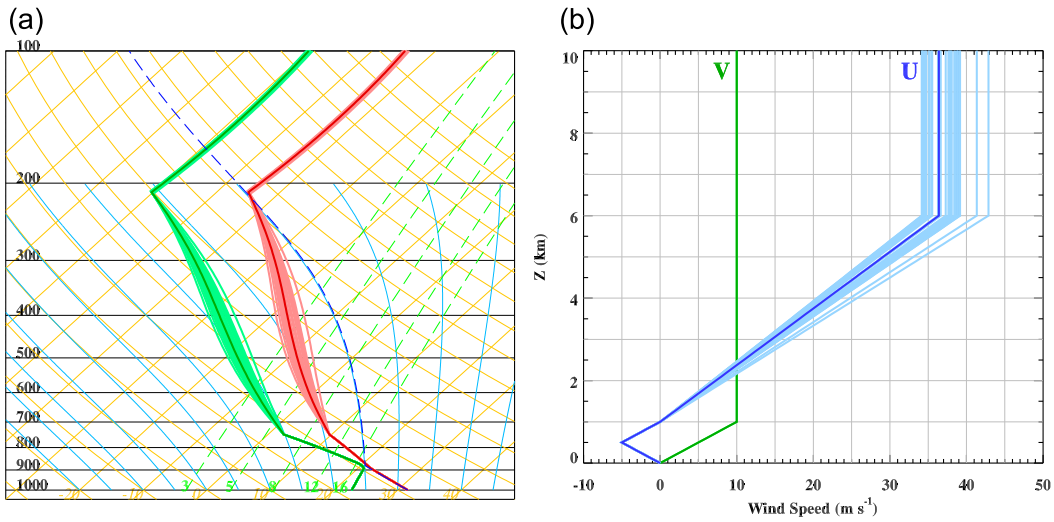

FIG. 1. Soundings used for the 101-member ensemble (control sounding illustrated with thick curves) (a) thermodynamic profile and (b) wind profile. The darker green, red, and blue colors are the profile for the control member, while the lighter shades are the individual profiles of the sounding for each ensemble member, in order to show the ensemble spread.

Cintineo and Stensrud (2013), the variance was intentionally small to ensure that all ensemble members produced supercells within reasonably similar environments for the purpose of demonstrating storm-scale ESA. Additional variance was introduced by adding random noise (±0.25 K) to the potential temperature fields throughout the domain, except near lateral boundaries, and varying the horizontal size and magnitude of the warm bubble. The warm bubbles were introduced at the center of the domain and 1.4 km above the surface, with a vertical semiaxis of 1.4 km, horizontal semiaxes of 10 km ± 25% with each horizontal semiaxis varying independently of the other, and a 4 K ± 25% potential temperature excess over the base state at the center of the bubble, modulated by a cosine function. The control member of the ensemble does not include random gridpoint perturbations or any perturbations to the size and magnitude of the warm bubble. Also, the gridpoint perturbations must be small enough to avoid initiating additional storms in the domain. However, absent these perturbations, the variance in the ensemble would be only a result of the perturbations to the initial sounding, which would be undesirable for obtaining meaningful ESA results and would likely provide less variance in the ensemble. Therefore, a combination of the aforementioned methods for introducing variance to the ensemble is likely to achieve the most useful results.

Perturbations and forecast responses were recorded starting at 45 min into the 2-h-long simulations to allow time for a storm to form initially. From that point, perturbation and forecast response variables were recorded every 5 min until the end of the simulations. Both the perturbations and forecast responses were recorded in a storm-relative coordinate system to allow for variations in storm motion and storm configuration among ensemble members. The storms were tracked with an automated system that followed the maximum value of a smoothed 2–5-km updraft helicity field. The perturbation variables were temperature, specific humidity, pressure, and the three wind components. The forecast response variables were the maximum 2–5-km updraft helicity, maximum lowest-model-level vertical vorticity, maximum composite reflectivity, and maximum lowest-model-level hail mixing ratio within the storm-relative forecast response area (Fig. 2).

The storm-relative frame of reference is also used for the perturbations. Because the focus here is on the sensitivity within easily identifiable storm-relative regions, the storm in each ensemble member is partitioned into four regions at each time: the forward-flank downdraft (FFD), forward-flank inflow (FFI), rear-flank downdraft (RFD), and rear-flank inflow (RFI; Fig. 3). These regions are defined according to mesocyclone location, storm orientation, and the position of thunderstorm outflow. Mesocyclone position is the location of the maximum 2–5-km updraft helicity, following the application of a Gaussian-smoothing operation implemented as the convolution of the 2–5-km updraft helicity on a 1-km grid with a 7 × 7 smoothing kernel where $\sigma = 1$ km. Storm orientation is determined with the 0–6-km shear vector. A line through the mesocyclone and perpendicular to the 0–6-km shear vector is used to partition the

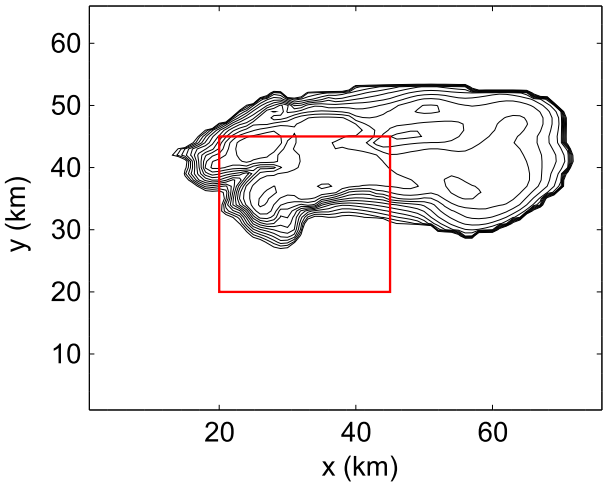

FIG. 2. Perturbation and forecast response domains plotted over reflectivity from the strong supercell control simulation. Simulated reflectivity is plotted in black, in units of 5 dBZ. The perturbation domain is the entire image and the forecast response domain is in red. Units are in km.

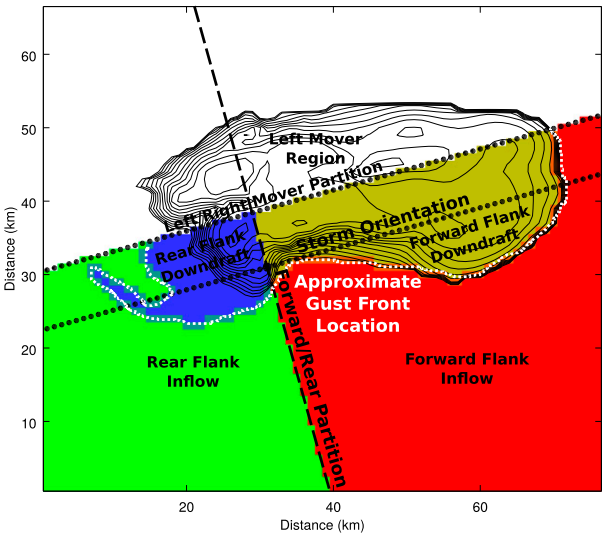

FIG. 3. The regions over which the correlations were averaged, shown for the storm 45 min into the simulation. Simulated reflectivity is plotted in black, in units of 5 dBZ. Units are in km. The southern dotted line is the orientation of the storm through the mesocyclone. The northern dotted line is the partition between the right- and left-moving storms. The dashed line is the partition between the forward and rear flanks of the storm. The green area is the rear-flank inflow and the red area is the forward-flank inflow. The blue area is the rear-flank downdraft and the yellow area is the forward-flank downdraft.

rear-flank and forward-flank regions. The split between the right- and left-moving storms is estimated as 8 km north of the mesocyclone and in the direction normal to the 0–6-km shear vector (Fig. 3); no points north of this line were included in the ESA calculations. Points are considered to be in thunderstorm outflow if they have $\theta' \leq -0.02$ K or if the reflectivity is at least 0 dBZ. Any "holes" in the outflow regions where the conditions are not met, but that are entirely surrounded by outflow points, are considered to be part of the outflow.

In addition to the four primary regions, analysis is also focused on perturbations within subregions within a transformed coordinate system defined by the gust fronts and the partition between the forward- and rear-flank regions (Fig. 3). Each subregion is a 2 km × 2 km box in the transformed system. The distance from the gust front and the partition line is calculated for each point in the domain, and the transformed coordinate system uses the distance from the partition line as the x coordinate and the distance from the gust front as the y coordinate. The θ, p, q_v, u, v, and w variables are then averaged within 2 km × 2 km bins in the transformed coordinate system. Analysis of these subregions enables a more detailed examination of areas of sensitivity that are washed out in the averaging within the primary regions (e.g., rear-flank downdraft, forward-flank inflow).

b. *Ensemble sensitivity analysis*

ESA is typically performed as a linear regression between a single perturbation variable and forecast response variable, with the sensitivity measured by the slope (e.g., Ancell and Hakim 2007; Torn and Hakim 2008; Hill et al. 2016). In the context of an observing platform such as a UAS that simultaneously measures temperature, humidity, pressure, and wind, the value of a targeted observation is the expected collective impact of all the quantities measured by the observing system. The approach to ESA used in prior work calculates the sensitivity of each measured variable independently, but does not account for multicollinearity between the perturbation variables. Hacker and Lei (2015) noted that this also does not account for the collective contributions to the variance from all of the perturbation variables simultaneously. These issues are addressed here by applying multiple linear regression in which the regression model is fitted through the ordinary least squares method. The sensitivity of the forecast response to the perturbation variables can be expressed through the coefficient of determination r^2, which is the proportion of variance in the forecast response (dependent variable) that is explained by the variance in the perturbations (independent variables). The coefficient of determination is closely related to the slope of the ordinary least squares regression line, especially in the case of a single independent and dependent variable. The correlation coefficient r is the slope of the ordinary least squares regression when the independent and

dependent variables are standardized such that $\sigma = 1$ for each variable. Even when performing ESA on a single perturbation and forecast response variable, using r as a metric of sensitivity may have desirable properties so that interpreting the results is easier.

At any given forecast time, the variance of the forecast response (dependent) variable is constant regardless of the location within the domain where the perturbation (independent) variable is obtained. If the forecast response is denoted as Y and the perturbation is denoted as x, the slope of the regression line is $\text{cov}(Y, x)/\text{var}(x)$. The slope of the regression line depends both on the amount of the variance in the forecast response that is explained by the perturbation variable and the actual variance of the perturbation variable. Because the variance of the perturbation variable can vary from one location to another, the slope will be biased to be higher in areas where the variance in the perturbations is small. Using the correlation coefficient in place of the slope does not alter the sensitivity metric except to remove this bias. In the case of multiple linear regression, the square root of r^2, which is equivalent to $|r|$, is a similar measure of sensitivity.

In this work, both of the aforementioned methods are used: 1) the sensitivity of the forecast response to a single perturbation variable using the aforementioned correlation coefficient, henceforth referred to as single-variable ensemble sensitivity; and 2) a multiple regression that calculates the sensitivity to all of the perturbation variables, henceforth referred to as multivariable ensemble sensitivity. The multivariable ensemble sensitivity extends ESA to fit a linear model between multiple perturbation variables and a single forecast response variable, calculating the sensitivity of the forecast response to the combination of all perturbation variables. This linear model takes the form of

$$\mathbf{y} = \boldsymbol{\beta} \times \mathbf{X} + \boldsymbol{\varepsilon}, \tag{1}$$

such that $\mathbf{X}$ is a matrix where each row contains the perturbations for the jth ensemble member and is a row vector $\mathbf{X}_j$ of the form $[1\ \theta'\ p'\ qv'\ u'\ v'\ w']$. In this model, $\mathbf{y}$ is a column vector containing the forecast responses for each ensemble member, $\boldsymbol{\beta}$ is a column vector fitting the perturbations to the forecast responses, and $\boldsymbol{\varepsilon}$ is a column vector of residuals. The model is fitted by minimizing the residual sum of squares $\text{RSS} = \sum_{j=1}^{n} \varepsilon_j$ for n ensemble members. The coefficient of determination, $r^2 = (\text{ESS/TSS}) = 1 - (\text{RSS/TSS})$ uses the explained sum of squares $\text{ESS} = \sum_{j=1}^{n} (\boldsymbol{\beta} \times \mathbf{X}_j - \bar{\mathbf{y}})^2$ and the total sum of squares $\text{TSS} = \sum_{j=1}^{n} (\mathbf{y}_j - \bar{\mathbf{y}})^2$ to compute the amount of the variance in the data explained by the linear model.

Even with random data where there is no sensitivity of the forecast response to the perturbations, r^2 and $|r|$ will almost certainly be greater than zero. The expected value of $|r|$ for random data, henceforth denoted as $|r|_{\text{rand}}$, is a baseline for how large $|r|$ needs to be to indicate actual sensitivity. Therefore, $|r| - |r|_{\text{rand}}$ is used as the preferred measure of sensitivity in this paper, such that a value of zero indicates no sensitivity. The expected value of $|r|_{\text{rand}}$ was estimated empirically, using the mean $|r|$ from 100 000 regressions with randomly ordered independent and dependent variables containing integers from 1 to 101. For one dependent variable and six independent variables (representing perturbations in temperature, humidity, pressure, and each component of the three-dimensional wind), $|r|_{\text{rand}}$ is approximately equal to 0.235. The maximum value of sensitivity is approximately 0.765, which occurs when $r^2 = 1$.

Lead time is defined as the time between when the perturbations and the forecast responses are recorded. For example, perturbations taken 45 min into the simulation for forecast responses an hour into a simulation would correspond to a lead time of 15 min. For a given lead time, sensitivity can have a large variability depending on when throughout the simulation the perturbations and forecast responses were recorded. To identify areas of consistently large sensitivity, a time-averaged forecast sensitivity is used, such that for a given location (or area) and lead time, the sensitivity throughout the simulation is averaged.

3. Ensemble sensitivity

This section presents results of ensemble sensitivity at 500 m above ground level, where an unmanned aircraft would sample the planetary boundary layer. Single-variable and multivariable sensitivity are presented together to assess whether there is a clearer signal of multivariable sensitivity arising from the combined sensitivity of the forecast response to multiple perturbation variables.

Sensitivity of composite reflectivity at a 20-min lead time to potential temperature is generally positive, with the exception of a corridor of negative sensitivity extending southeast of the mesocyclone (Fig. 4a). The pattern is generally inverted for water mixing ratio (Fig. 4c), with the exception of the rear-flank downdraft region, which also shows positive sensitivity. Additionally, there is generally negative sensitivity to pressure throughout the rear-flank region (Fig. 4e), that generally decreases with longer lead times (Fig. 4f). There is a positive sensitivity to the v-wind component across much of the inflow region, with the exception of the far

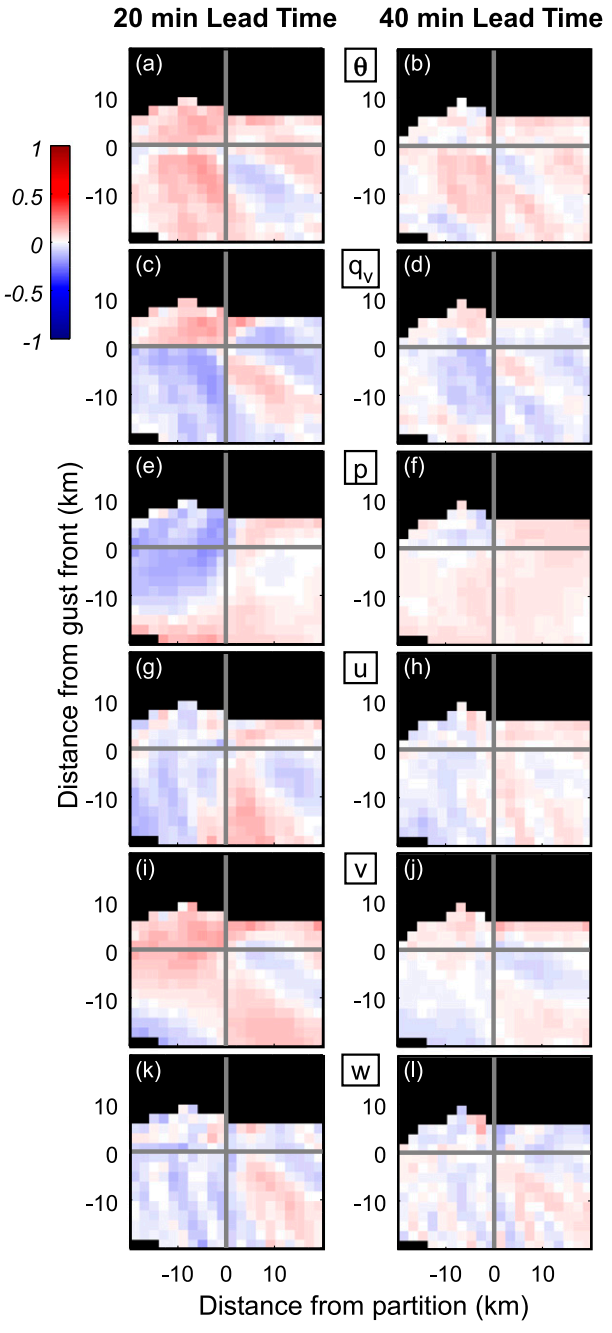

FIG. 4. Single-variable sensitivity of composite reflectivity at lead times of 20 and 40 min.

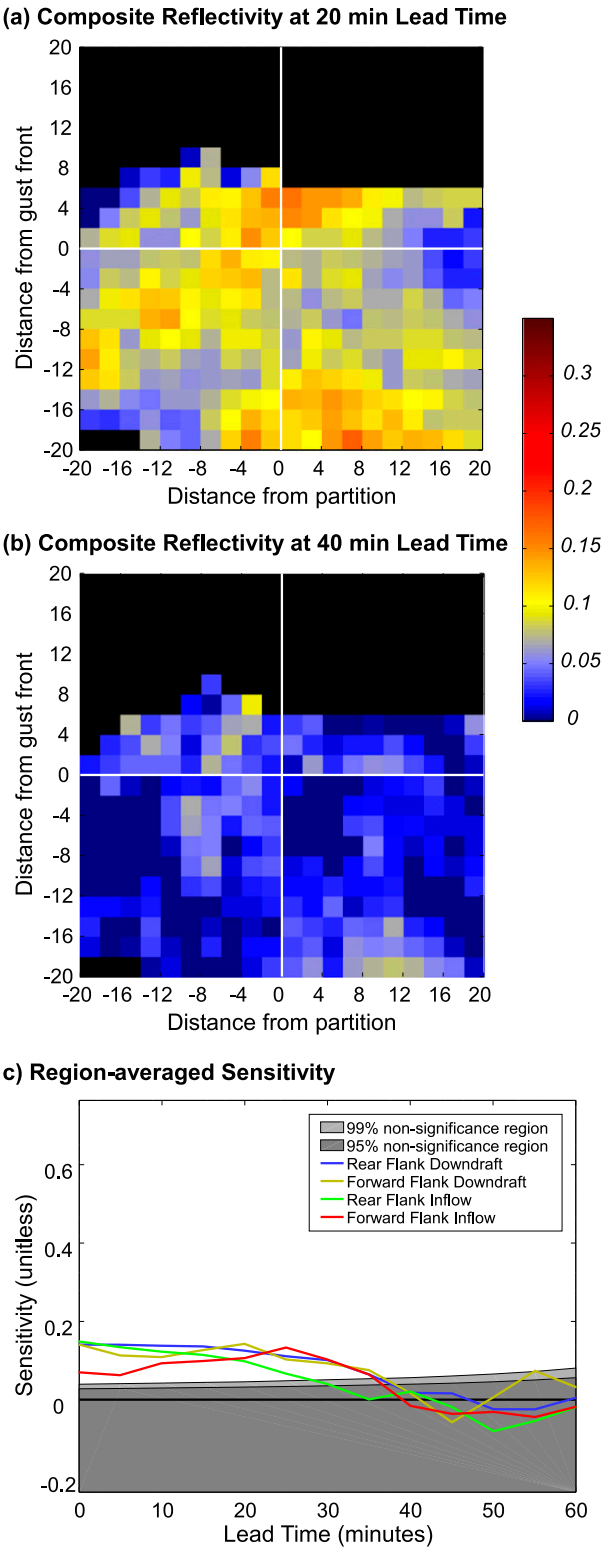

FIG. 5. Multivariable sensitivity of composite reflectivity shown at lead times of (a) 20 min, (b) 40 min, and (c) in a time series.

southwest area and along the forward-flank gust front (Fig. 4i). Across all variables considered, the sensitivity generally decreases with time and the correlations are low. However, the spatial patterns in sensitivity remain similar through lead times of exceeding 40 min. The multivariable sensitivity at a 20-min lead time (Fig. 5a) shows several maxima including near the mesocyclone and in the far south region of the forward-flank inflow.

However, the sensitivity is small relative to the maximum possible sensitivity (0.765) and decreases significantly in time (Figs. 5b,c) becoming statistically insignificant in all regions by a ~40-min lead time (Fig. 5c).

The overall spatial pattern for hail mixing ratio sensitivity at the lowest model level at a lead time of 20 min is similar to that of composite reflectivity in the potential temperature and water vapor fields (Figs. 6a,c), though with the signs of the correlation reversed (cf. Figs. 4a,c). In many respects, the spatial patterns of the u- and v-wind fields are similar to each other (Figs. 6g–j), with an area of positive sensitivity in the southwest half of the rear-flank inflow region and a band of negative or lower sensitivity to the northeast. There is no discernable pattern to sensitivity in the w-wind field at either a 20-min lead time (Fig. 6k) or 40-min lead time (Fig. 6l). Unlike composite reflectivity, the sensitivity for potential temperature and water vapor mixing ratio reverses sign in forward-flank inflow between lead times of 20 and 40 min (Figs. 6a–d). For the multivariable sensitivity (Fig. 7), the strongest sensitivity is in outflow air near the base of the mesocyclone and within the rear-flank inflow. Similar to composite reflectivity, both the single and multivariable sensitivities decrease substantially over time (Figs. 6 and 7c). The reasons for the patterns of sensitivity in composite reflectivity (Figs. 4–5) and hail mixing ratio (Figs. 6–7) are not clear, though the results suggest that there is weak sensitivity to the structure and characteristics of the inflow region of the storm.

Figure 8 shows the single-variable sensitivity of 2–5-km updraft helicity, with the most notable features at a lead time of 20 min being positive sensitivity to the v wind throughout the inflow region (Figs. 8i,j), positive sensitivity to u wind in the rear-flank regions (Figs. 8g,h), and negative sensitivity to u wind in the forward-flank regions (Figs. 8g,h). Additionally, there is generally positive sensitivity to potential temperature in the inflow region and negative sensitivity in the downdraft regions (Figs. 8a,b). Physically, these signals may indicate the sensitivity of the mesocyclone to convergence. For the multivariable sensitivity, there is a large area of strong sensitivity in the southwest corner of the rear-flank inflow region and a second area of sensitivity a few kilometers north of the forward-flank gust front at a 20-min lead time (Fig. 9a). These spatial patterns remain at a lead time of 40 min (Fig. 9b). The sensitivity of the mesocyclone to perturbations in the forward flank near the gust front is physically reasonable considering that the mesocyclone is likely impacted by vorticity generation in this region. Overall, sensitivity decreases with increasing lead time (Fig. 9c), but, unlike composite reflectivity (Fig. 5c) and lowest-model-level hail (Fig. 7c), overall sensitivity remains significant through lead times approaching 60 min, though values remain well below the maximum possible sensitivity.

The lowest-model-level vertical vorticity exhibits the strongest sensitivity of all forecasts response variables. Vertical vorticity exhibits particular sensitivity to water vapor near the mesocyclone at a lead time of 20 min (Fig. 10c) and also the near-storm rear-flank inflow at a lead time of 40 min (Fig. 10d). In general, at a lead time of 40 min, the single-variable sensitivity shows very similar structure to the composite reflectivity and hail mixing ratio, though the overall sensitivity is stronger. For multivariable sensitivity (Fig. 11), sensitivity is relatively large (cf. other forecast response variables) north of the gust front and near the mesocyclone at both 20- and 40-min lead times (Figs. 11a,b), corresponding with the strong sensitivity from water vapor in that region. This signal in sensitivity remains through a lead time of 60 min (not shown). There is also significant sensitivity near in the rear-flank inflow region, particularly at a lead time of 20 min (Fig. 11a). Overall, the sensitivity for the lowest-model-level vertical vorticity remains significant through all lead times of 60 min (Fig. 11c).

Although the sensitivity calculations appear to show some coherent patterns, particularly in the inflow region and near the base of the mesocyclone, the overall strength of the sensitivity is generally low relative to the maximum possible sensitivity (-1 and 1 for single-variable sensitivity, 0.765 for multivariable sensitivity). The generally weak sensitivity suggests that targeted observations based on the methods presented in this paper should produce small forecast improvements. The strongest and most coherent areas of sensitivity were in the forward-flank inflow, far-field rear-flank inflow, near the mesocyclone, and along the forward-flank gust front. However, areas of strong sensitivity are not necessarily due to physical processes that actually affect the strength of the storm at a later time and even large coherent reasons of sensitivity may be due to autocorrelation. Therefore, validation of the ensemble sensitivity is required and will be examined in section 4.

Autocorrelation may have a significant impact on the results even in the far field. Recall that analysis focuses on the sensitivity of forecast response variables to perturbations at 500 m above ground level where initial conditions across all ensemble members are identical, beyond random gridpoint noise. Thus, any perturbations in the inflow other than the random gridpoint noise must be induced by the storm.

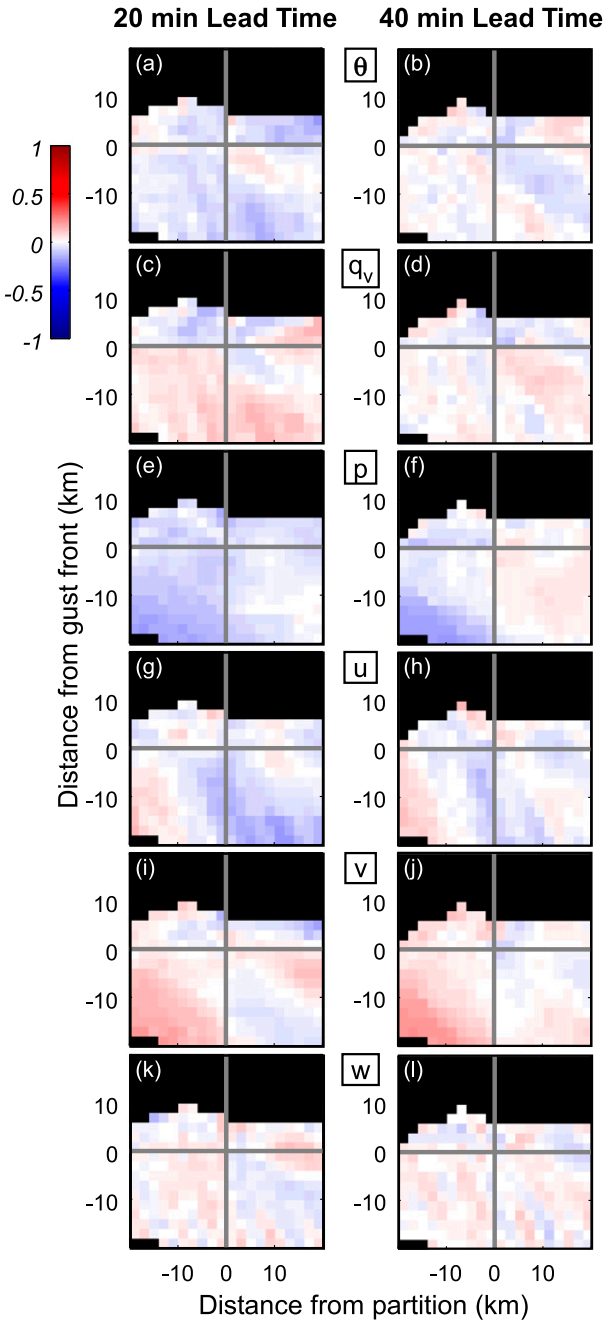

FIG. 6. Single-variable sensitivity of lowest-model-level hail mixing ratio at lead times of 20 and 40 min.

However, just because the storm induces perturbations in the inflow does not guarantee that these perturbations will be correlated to the future state of the storm. That these high correlations are found in this analysis can be attributed to one or both of the following explanations: 1) storm-induced perturbations cause changes in storm strength and/or 2) autocorrelation; that is, the storm perturbs the inflow with a

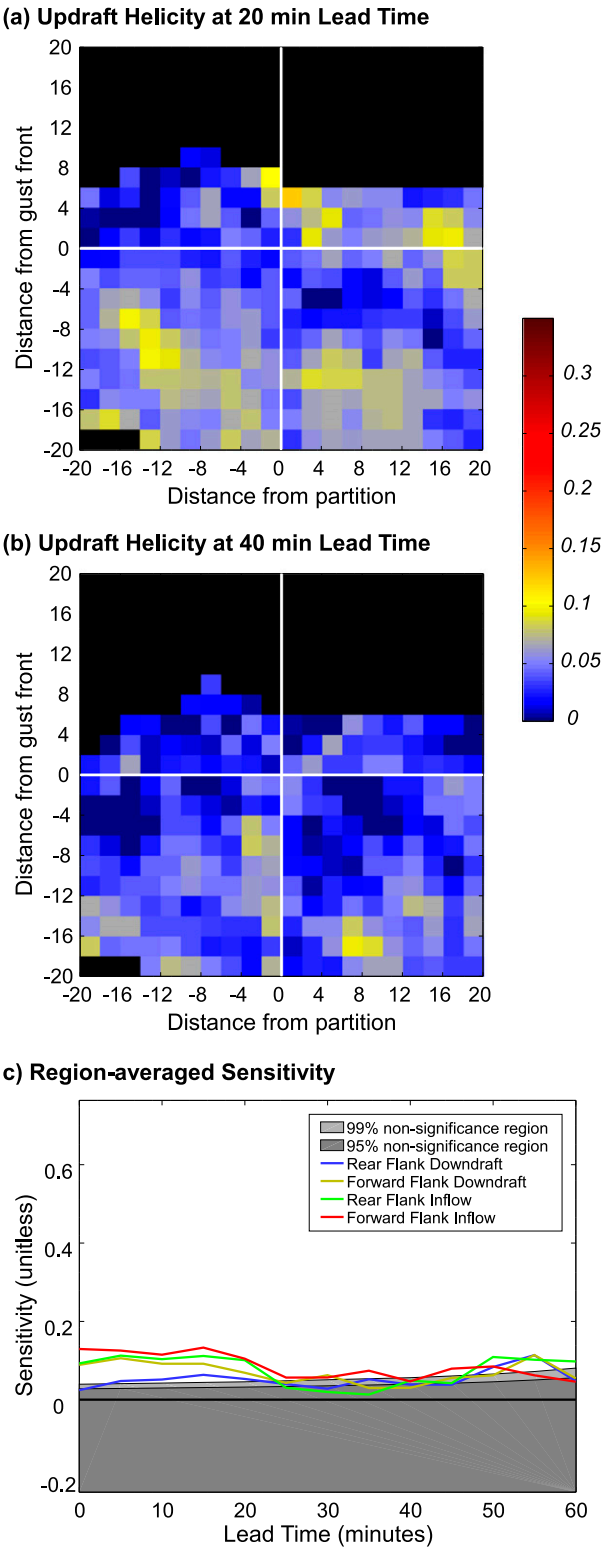

FIG. 7. Multivariable sensitivity of lowest-model-level hail mixing ratio shown at lead times of (a) 20 min, (b) 40 min, and (c) in a time series.

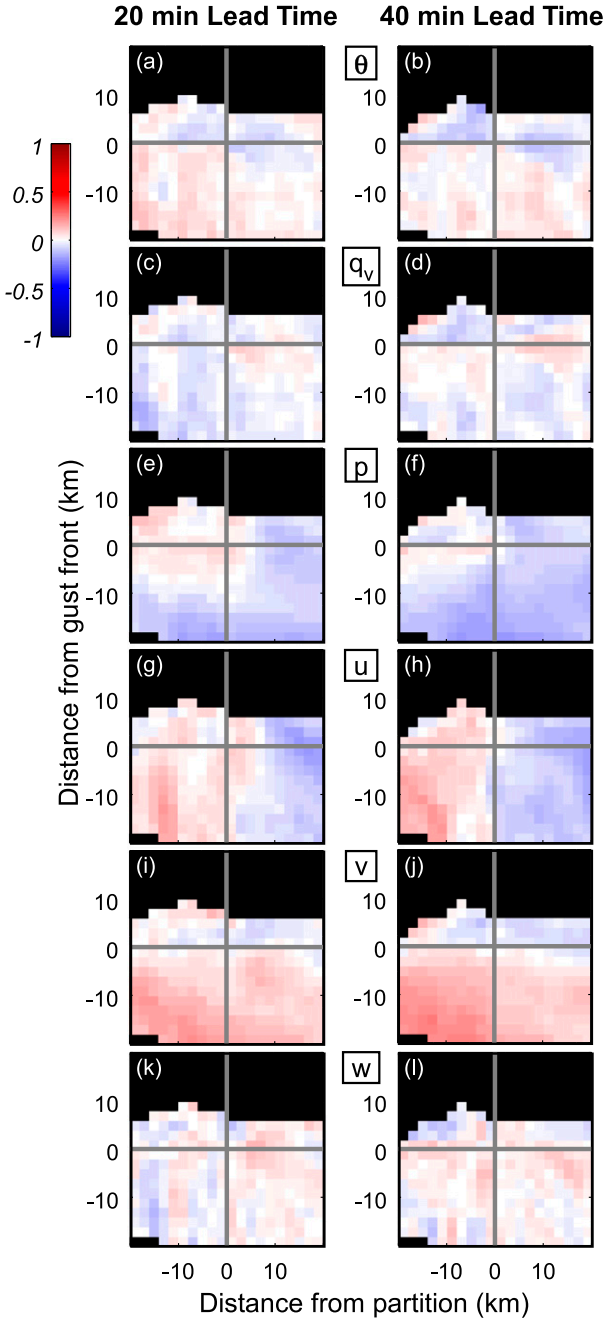

FIG. 8. Single-variable sensitivity of 2–5-km updraft helicity at lead times of 20 and 40 min.

magnitude that scales directly with storm strength, storm strength in the future is autocorrelated with current storm strength, thus, the inflow and future storm strength are correlated.

In an effort to evaluate the potential role of autocorrelation on the results, ensemble sensitivity was also calculated at negative lead times; that is, the same formulas presented earlier are used but with the

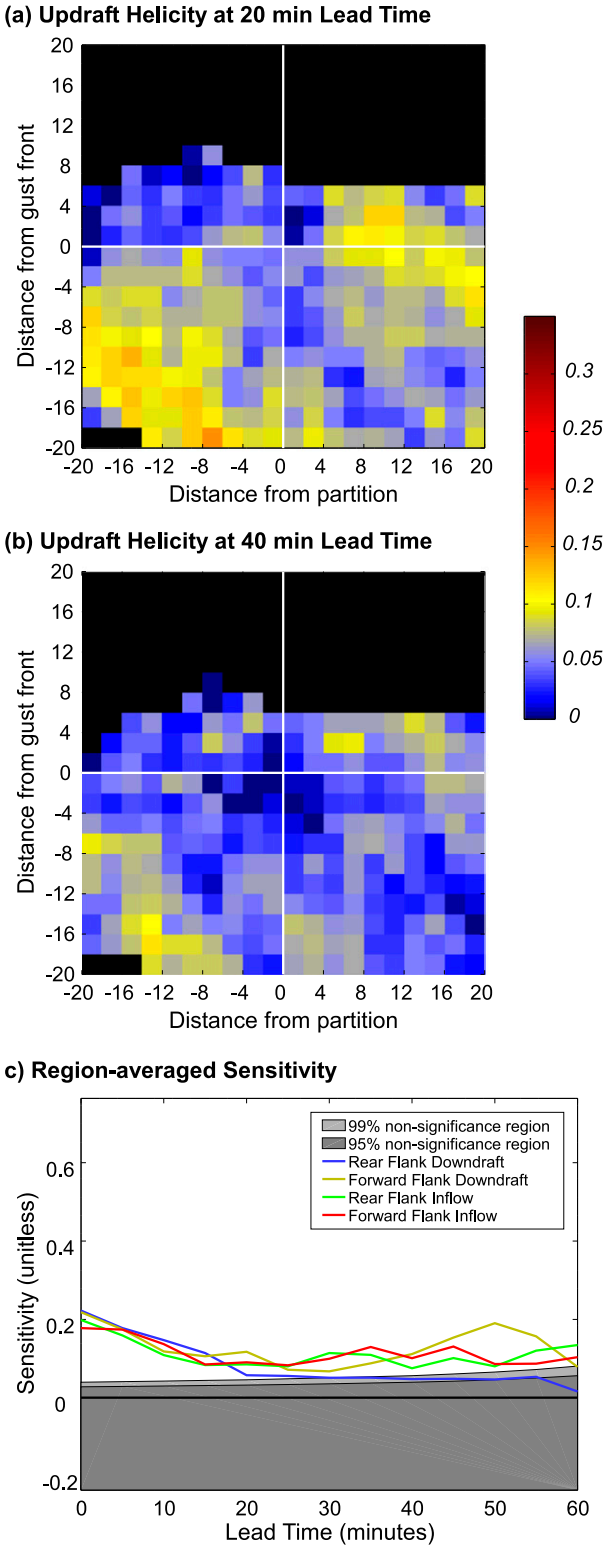

FIG. 9. Multivariable sensitivity of 2–5-km updraft helicity shown at lead times of (a) 20 min, (b) 40 min, and (c) in a time series.

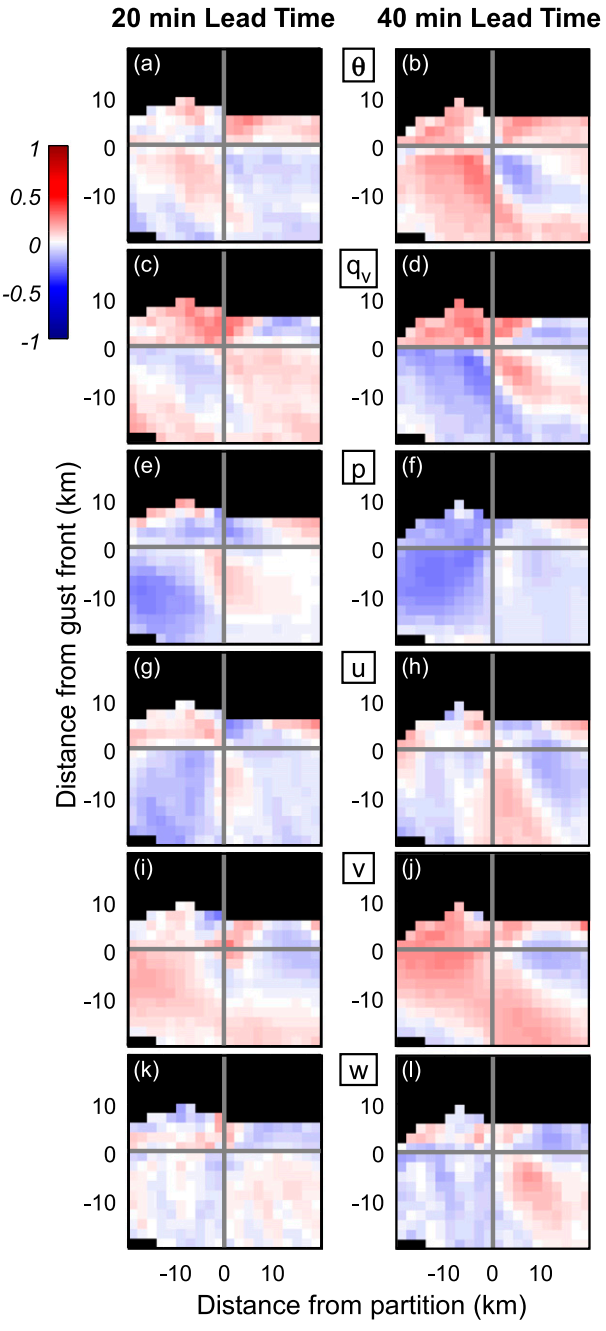

FIG. 10. Single-variable sensitivity of lowest-model-level vertical vorticity at lead times of 20 and 40 min.

forecast response at a time *prior* to the perturbations. In many instances, the sensitivity is at least as strong at negative lead times as at positive lead times. Figure 12 shows the single-variable sensitivity for composite reflectivity and lowest-level vertical vorticity at a lead time of −20 min. The spatial pattern of sensitivity of lowest-level vertical vorticity to all six perturbation variables are very similar, although the magnitudes of

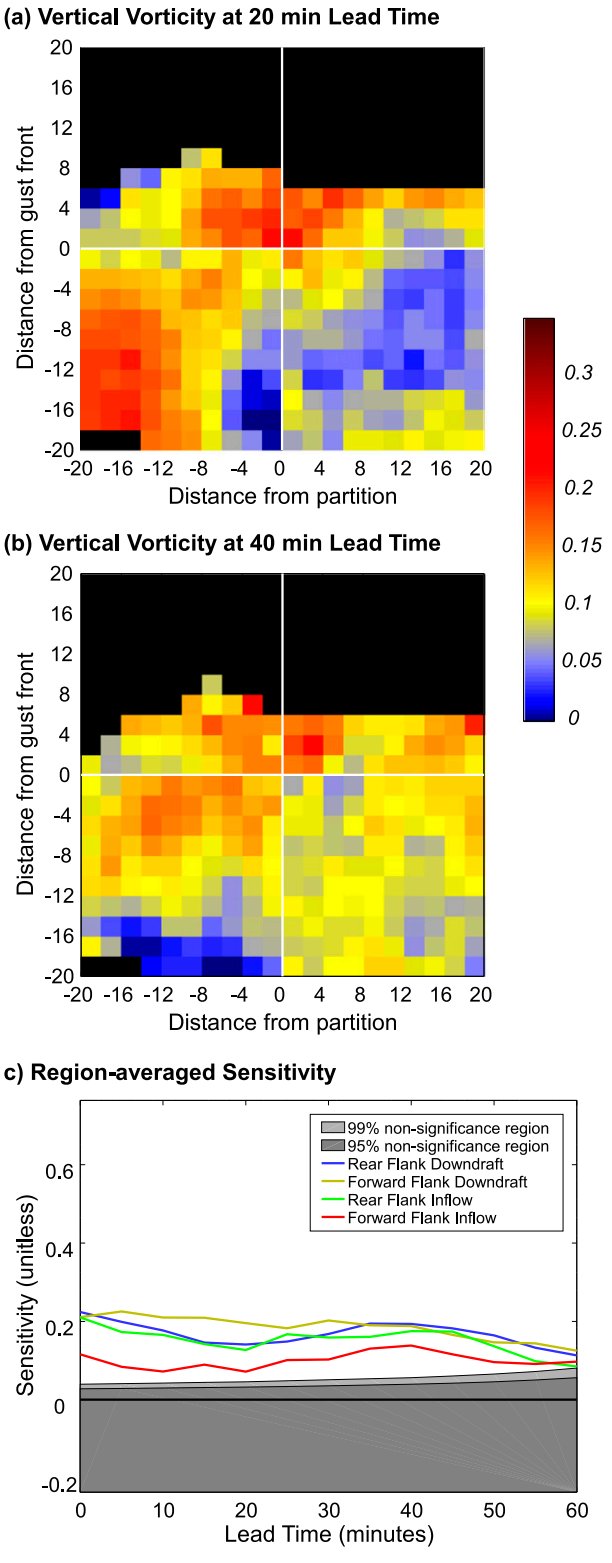

FIG. 11. Multivariable sensitivity of lowest-model-level vertical vorticity shown at lead times of (a) 20 min, (b) 40 min, and (c) in a time series.

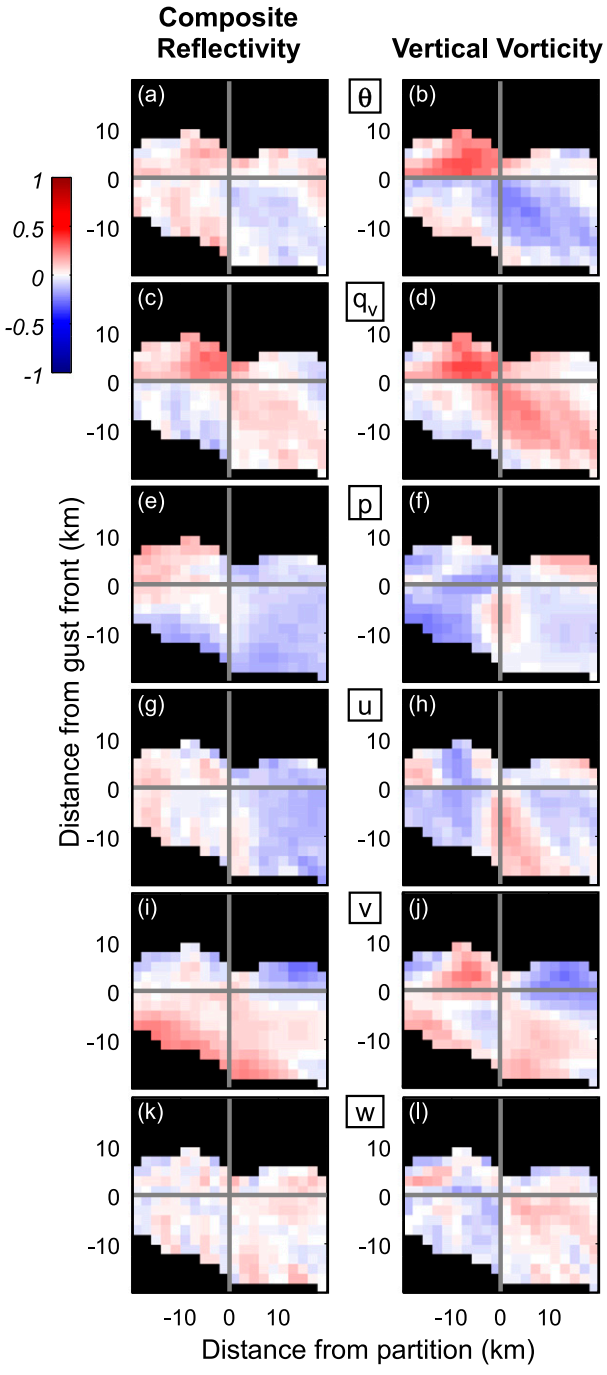

FIG. 12. Single-variable sensitivity of composite reflectivity and lowest-model-vertical vorticity at a lead time of −20 min.

the sensitivity are larger at −20 min than at 20 min (cf. Figs. 10 and 12). This differs for composite reflectivity, in which only the spatial patterns for potential temperature and humidity are similar at −20 min and 20 min (cf. Figs. 4 and 12). The similarity in spatial patterns of sensitivity between positive and negative lead times, particularly for lowest-model-level vertical vorticity, suggests that autocorrelation may be contributing to the sensitivity seen at positive lead times. As noted above, an autocorrelation signal alone does not necessarily devalue the possible impact of targeted observation of these regions on forecast response, but it raises questions about the degree to which ESA is a reliable measure of the magnitude this possible impact.

Further complicating interpretation of the sensitivity is the correlation between the perturbation variables (Fig. 13), which has implications on targeting observations. Although there are many overlapping areas of stronger sensitivity to single perturbation variables, these structures are generally not present in the multivariable calculations, which account for the covariances between the perturbation variables. For example, in the inflow region, warmer temperatures are correlated with less moisture (Fig. 13a) and, through much of the downdraft regions, warmer temperatures are associated with increasing humidity, likely related to evaporation in the downdraft (Fig. 13a). When assimilating these observations, updating potential temperature should also update moisture, which is strongly correlated, and therefore there is less value to the water vapor observation in improving the storm forecast. Likewise, warmer conditions are associated with increasing values of v wind in the downdraft region (Fig. 13j), probably indicating that warmer conditions mean a weaker northerly wind in the downdraft. There are other strong relationships such as a negative relationship between pressure and v wind through the entirety of the inflow region (Fig. 13f), a strongly positive relationship between pressure and u wind throughout the forward-flank inflow region, and correlations between w wind and other perturbation variables along the gust front (Figs. 13i,k–n). Also of note is that the pattern of v-wind sensitivity in the inflow region that appears in many of the single-variable ESA calculations is very similar to the covariance between v wind and both potential temperature (Fig. 13j) and water vapor mixing ratio (Fig. 13h). The presence of strong covariances between perturbation variables complicates the interpretation of the sensitivity calculations, and because the multivariable sensitivity accounts for the covariances, there may be value in using the multivariable sensitivity despite the noisy characteristics of the sensitivity.

4. ESA validation

In an effort to assess whether the subtle sensitivity signals are deterministic or merely statistical artifacts (i.e., whether or not sensitivity is associated with physical processes that affect storm strength), two validation methods are implemented. The first is statistics based

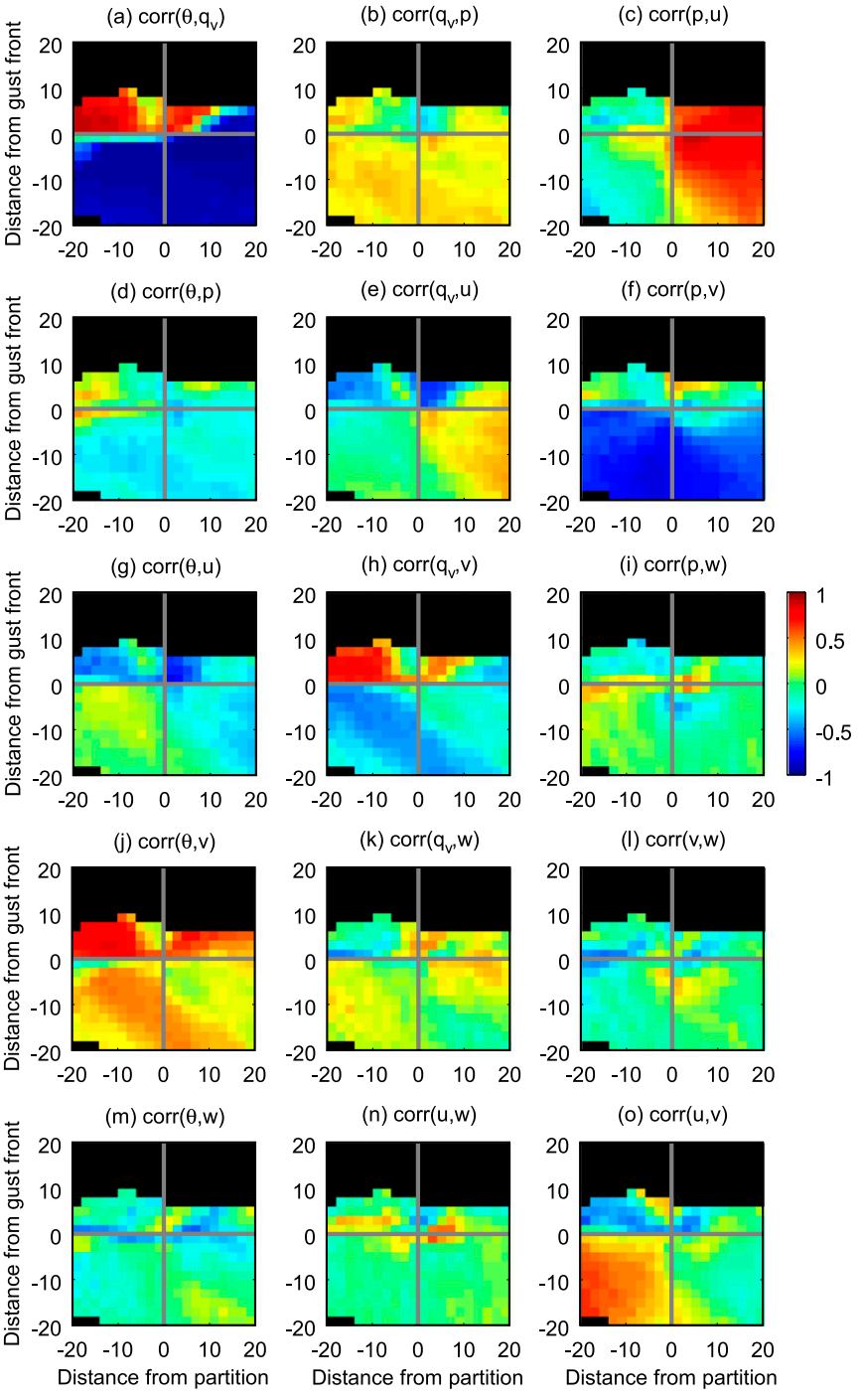

FIG. 13. Correlations between all of the perturbation variables.

while the second involves perturbed initial condition experiments that are somewhat similar to those performed by Wile et al. (2015).

The first approach involves withholding a single ensemble member, chosen at random, from the multivariable sensitivity calculations. The result is a linear model based on the 100 remaining ensemble members that relates the perturbations at a particular grid point to the forecast response, using the same methods discussed in section 2. The linear model is then used to predict the forecast response based on the actual perturbations of the withheld ensemble member. The predicted forecast

response is compared against the actual response, and it is hypothesized that if the ESA is valid, larger sensitivity in $|r|$ and $|r| - |r|_{\text{rand}}$ should be associated with smaller differences between the predicted and actual forecast response. For each 2 km × 2 km bin with at least four times contributing to the time-averaged sensitivity, the aforementioned approach was employed three times with different ensemble members chosen at random to compare predicted and actual sensitivity.

For Fig. 14, $|r|$ is binned into increments of 0.01, and within each bin, the distribution of the difference in predicted and actual response is shown. For composite reflectivity, there is no clear decrease in the difference between the predicted and actual forecast response with stronger correlations. This suggests that the areas of sensitivity for composite reflectivity are statistical noise rather than signal. Likewise, there is no clear improvement for lowest-model-level hail mixing ratio. For the 2–5-km updraft helicity, the errors do decrease with stronger sensitivity. There is a similar but weaker trend for the lowest-model-level vertical vorticity. This suggests that the sensitivity for composite reflectivity and lowest-model-level mixing ratio does not reflect physical processes affecting storm strength while the sensitivity for lowest-model-level vertical vorticity and 2–5-km updraft helicity is more likely to be associated with perturbations that do impact the storm at future times.

These results offer some support for the validity of ESA for updraft helicity and low-level vertical vorticity. However, the validity of individual regions can be tested by perturbing the initial conditions, or in this case, introducing perturbations during the simulation, to demonstrate that the perturbations do result in a change in forecast response. Although the approach of Wile et al. (2015) guided the formulation of this approach, there is not an obvious statistical technique to provide a single metric relating the correlation between six perturbation variables at one grid point and six perturbation variables at another grid point. However, the general approach of perturbing regions rather than grid points is employed by identifying a location of relatively strong sensitivity with respect to the surrounding area and delineating the boundaries of the region based on a minimum threshold. For experiments on vertical vorticity, a minimum sensitivity threshold of 0.18 is used; for updraft helicity, the threshold is 0.09. Perturbations are introduced for all six of the perturbation variables, with the relative magnitude for each variable determined by β [(1)] at the location of maximum sensitivity within the region that is delineated. The standard deviation σ of each perturbation variable is also calculated at that location. For each perturbation variable i, $k_i = \sigma_i/\beta_i$ is calculated as a scaling factor for the perturbations, and the largest $|k_i|$, denoted as k, is chosen as the scaling factor. Thus, the perturbation vector at the point of maximum sensitivity is defined by β/k. At each other point within the region, the perturbation vector is scaled by the ratio of the sensitivity at that point to the maximum sensitivity in the region. Because the level of perturbations is at 500 m, which occurs directly between two vertical levels, the perturbations are applied at the two surrounding vertical levels. The sensitivity is interpolated for the model points with u and v winds. Because the model levels for w winds include 500 m, to ensure that the same number of points is perturbed, the perturbations are also introduced at 400 m. The resulting set of perturbations is a three-dimensional matrix, denoted as $\mathbf{P}_{\text{pert}}$. A Gaussian smoothing kernel is applied to $\mathbf{P}_{\text{pert}}$, the result of which is henceforth referred to as $\mathbf{P}_{\text{smooth}} = \text{gauss}(\mathbf{P}_{\text{pert}})$, in order to avoid introducing discontinuities. The kernel has a horizontal width of 11 points and a vertical width of 21 points, with a σ of 1 in all dimensions. The final set of perturbations is the maximum value of the two matrices, $\mathbf{P} = \max(\mathbf{P}_{\text{pert}}, \mathbf{P}_{\text{smooth}})$. The ensemble control member, which should be close to the ensemble mean, is also used as a control simulation for these experiments. For each set of initial conditions, a pair of perturbed initial conditions are created, one by adding $\mathbf{P}$ to the initial conditions (which should yield an increase in the forecast response variable), and one by subtracting $\mathbf{P}$ from the initial conditions (which should yield a decrease in the forecast response variable).

Three sets of experiments are performed: two using vertical vorticity as the forecast response variable, and one with updraft helicity as the forecast response, using the sensitivity at lead times of 20 min. The two regions for vertical vorticity (Fig. 15a) are located near the mesocyclone (denoted "region 1") and on the southwest edge of the rear-flank inflow region (denoted "region 2"). For updraft helicity, an area of sensitivity east of the mesocyclone is used (Fig. 15b). Initial conditions are selected at 5-min intervals from 45 min into the simulation until 100 min, resulting in 12 pairs of initial conditions for region selected to be perturbed. All of the simulations are integrated forward 20 min to determine the impact of the perturbations on the forecast response.

These results are analyzed by subtracting the perturbed simulation forecast response from the control simulation forecast response and will be referred to as response difference. For each set of perturbations (negative or positive), the value of the response difference in many of the simulations is clustered around zero with significant variability (Table 1). This is unsurprising because of the significant variability from one time to the next in the ensemble sensitivity prior to applying the domain averaging. However, when the region east of the

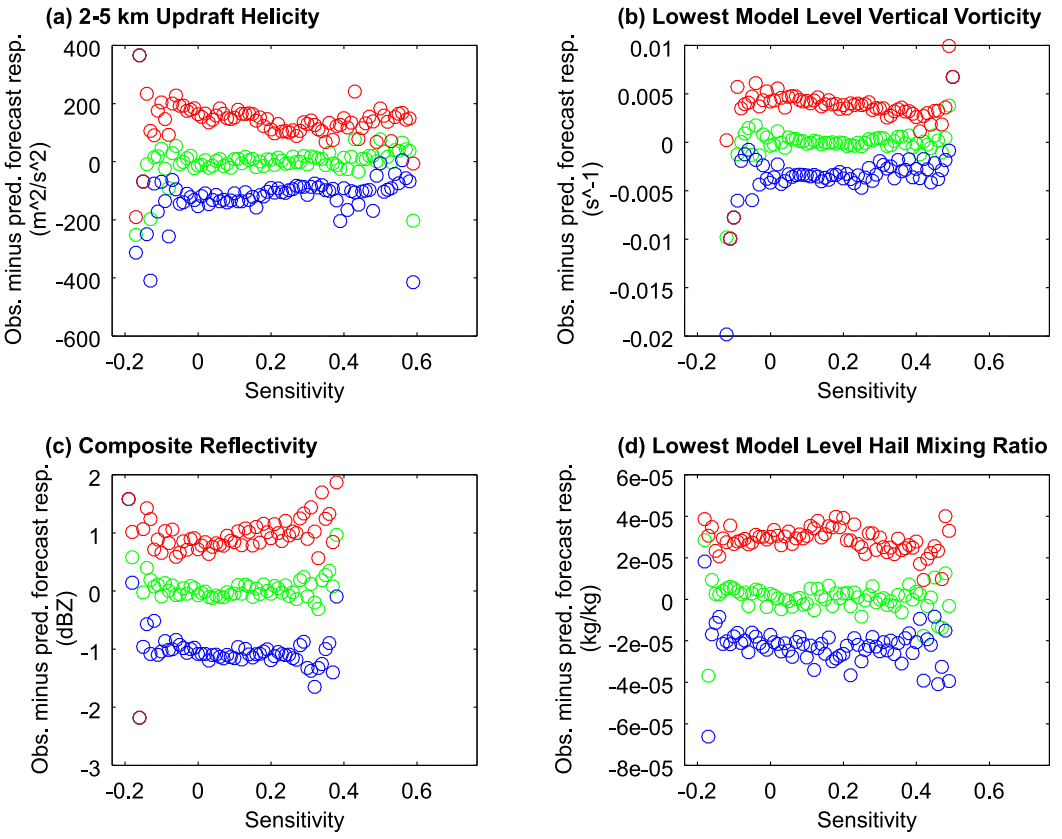

FIG. 14. The difference between predicted and actual forecast response depending on sensitivity across all (5–60 min) lead times for the forecast response variables. The blue scatterplot is the 25th percentile of the difference at that correlation level, the green scatterplot is the mean difference at that correlation level, and the red scatterplot is the 75th percentile at that correlation level.

mesocyclone is perturbed (region 1), the sign of the perturbation generally matches the sign of the mean vertical vorticity response difference, beyond a lead time of 5 min (Table 1). Moreover, the separation between the mean response difference for positive perturbations and the mean response difference for negative perturbations is largest at the longest lead time considered. The consistency between the sign of the perturbations and the sign of the response difference is generally not seen for vertical vorticity and perturbations imposed in region 2. Furthermore, the magnitudes of the mean response difference are much smaller than for region 1. The response difference for updraft helicity shows a consistent relationship between the sign of the perturbation and the sign of the mean response difference for lead times less than 15 (Table 1). Moreover, the largest mean response difference for positive perturbations is seen at the longest lead time considered.

The opposite signs of the response difference and the perturbations for region 1 and a lead time of 5 min (Table 1) would seem to be inconsistent with the finding that the highest average sensitivity in the rear-flank downdraft is found at shorter lead times not longer lead times (Fig. 11c). However, the pattern reflected in the perturbed initial condition experiments suggests that the sensitivity at shorter lead times may actually be due to autocorrelation.

Relating the actual and predicted forecast response to the correlation suggests that when considering all points in the domain, much of the sensitivity for the composite reflectivity forecast response variable is actually a statistical artifact. These results do indicate, however, that for the other forecast response variables, there may actually be some real signal in the sensitivity for those variables. The results from the perturbed initial condition experiments suggest that the sensitivity around the mesocyclone for low-level vertical vorticity and along the forward-flank gust front for updraft helicity may be associated with physical processes that affect storm strength. It is less clear whether there is an actual signal in the area of high sensitivity in the rear-flank inflow region for low-level vertical vorticity.

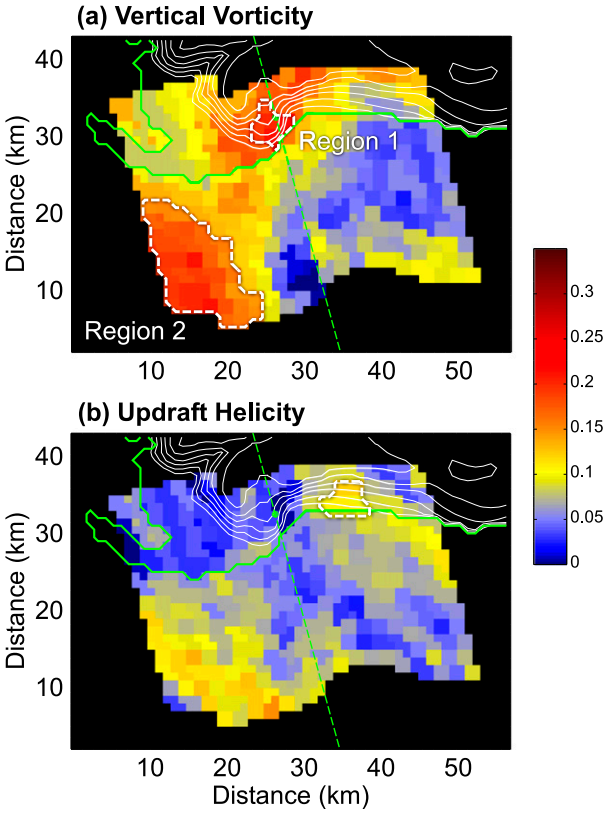

FIG. 15. Perturbations introduced to test the impact on lowest-model-level vertical vorticity and updraft helicity. Sensitivity is shaded, simulated reflectivity is plotted in solid white lines in increments of 10 dBZ, the mesocyclone is the green circle, the partition between the rear- and forward-flank regions is the dashed green line, and the solid green line is the gust front. The region perturbed is outlined by a dashed white line, though small perturbations may extend beyond this region because of the Gaussian smoothing.

5. Conclusions

This paper presented the results of ensemble-based sensitivity analysis applied at the scale of an individual supercell, as opposed to previous work at the mesoscale and synoptic scale. The impetus for conducting this research was to inform the use of unmanned aircraft observations for targeted observations of the planetary boundary layer near supercell thunderstorms, to improve prediction at lead times out to one hour. Because unmanned aircraft are capable of observing multiple variables simultaneously, this study investigated using the multiple regression to compute a single metric of sensitivity encompassing all of the observed variables.

Prior work applying ESA has used the slope of the linear regression line between a single perturbation variable and a single forecast response variable as a metric of sensitivity. For targeted observations where multiple variables are observed simultaneously, there are two advantages from using the multiple regression. First, the multiple regression provides a single metric that quantifies the sensitivity of the forecast response to all of the observed variables, which is easier to interpret than performing a separate linear regression for each observed variable. Second, the multiple regression accounts for covariances between perturbation variables, which is useful in estimating the true impact of assimilating those variables. Results presented from this work show that covariances between perturbation variables might explain some of the simulated single-variable sensitivities.

One of the primary challenges in storm-scale ESA is identifying which areas of sensitivity in the storm environment actually correspond to physical processes that affect the strength of the storm at a later time. Supercell thunderstorms are capable of inducing substantial modifications to their environment (e.g., Brooks et al. 1994; Potvin et al. 2010; Nowotarski and Markowski 2016) and these modifications can induce feedbacks that affect the strength of the storm. The results presented here suggest that the storm-induced environmental perturbations are sensitive to the strength of the storm at an earlier time, and some of those perturbations may impact the strength of the storm at a later time. The

TABLE 1. The mean and standard deviation of the difference between the perturbed and unperturbed simulations. Italicized text denotes where the sign of the mean is opposite to that of the perturbation.

	5 min	10 min	15 min	20 min
Lowest-model-level vertical vorticity (region 1) ($\times 10^{-5}$ s^{-1})				
Positive	*−28.14 ± 108.84*	53.87 ± 132.59	11.92 ± 390.21	139.97 ± 582.25
Negative	*26.72 ± 104.08*	−78.13 ± 131.35	−73.22 ± 365.86	−219.58 ± 514.78
Lowest-model-level vertical vorticity (region 2) ($\times 10^{-5}$ s^{-1})				
Positive	0.009 ± 0.058	0.045 ± 0.266	*−1.494 ± 3.637*	*−5.101 ± 11.825*
Negative	−0.020 ± 0.039	*0.098 ± 0.263*	−2.289 ± 5.067	−3.026 ± 8.753
2–5-km updraft helicity (m^2 s^{-2})				
Positive	0.075 ± 0.913	0.746 ± 2.836	*−0.627 ± 3.781*	3.588 ± 6.505
Negative	−0.132 ± 1.093	−0.807 ± 2.608	*0.383 ± 4.048*	*0.574 ± 3.142*

difficulty in interpreting these results is determining which areas of apparent sensitivity in the environment contribute to the strength of the storm at a later time and which do not.

To assess whether the sensitivity corresponds to a physically meaningful relationship, three regions of sensitivity were selected for perturbed initial condition experiments. In these experiments, pairs of perturbations were introduced based on the coefficients of the linear regression model, with one member of the pair predicted to increase the forecast response variable and the other member predicted to result in a decrease. Additionally, an experiment was performed in which one ensemble member was withheld from sensitivity calculations, a linear model was computed for the remaining ensemble members, and the linear model was tested using the perturbations and forecast response from the withheld ensemble member. Results suggest that targeting observations based on these sensitivity calculations could impact forecast accuracy. However, because of the high levels of spread in the results, it is likely that instances in which targeted observations yield forecast improvement will coexist with numerous instances in which such observations result in little to no improvement.

There is a bevy of prior work either identifying physical processes associated with regulating supercell strength and demonstrating sensitivity to low-level perturbations (e.g., Dahl et al. 2014; Markowski and Richardson 2014; Naylor and Gilmore 2014; Weiss et al. 2015; Orf et al. 2017), some of which even used the CM1 model also used in this study. This suggests that supercell strength should be sensitive to low-level perturbations and that the model is certainly capable of reproducing many of the processes that would modulate supercell strength. The results of the sensitivity calculations in this paper cannot be explained as an actual lack of sensitivity of supercell strength to low-level perturbations, nor a failure of the CM1 model to simulate these processes. Therefore, it is logical that the low values of sensitivity and highly uncertain inferred impact of targeted storm-scale observations are due to the limitations of the methods for calculating ensemble sensitivity. The results from experiments presented in this article, along with numerous others conducted by the authors but excluded for brevity, lead to the conclusion that making incremental improvements to the methodology employed in this paper will not produce significantly improved signals in the sensitivity. Instead, it is posited that applying a linear model similar to ensemble sensitivity analysis on the scale of a supercell thunderstorm is inherently limited because of the importance of highly nonlinear dynamics and error growth coupled with rapid changes in forecast metrics. These results should not be generalized to a conclusion that the problem of targeting observations on the scale of a supercell thunderstorm is an intractable one, but rather that linear models such as ensemble sensitivity analysis are suboptimal when used in isolation and that future work should explore superior approaches to informing targeted observations of supercells.

Acknowledgments. We would like to acknowledge high-performance computing support from the Holland Computing Center at the University of Nebraska. Funding support was supplied by the Nebraska Research Council's Maude Hammond Fling Faculty Research Fellowship, Air Force Office of Scientific Research Grant FA9550-12-1-0412, and NSF Grant IIS-1527113. We also thank Corey Potvin, Patrick Skinner, Christopher Weiss, Brian Ancell, and Aaron Hill for helpful discussions regarding this research. Additionally, we thank the editor, Ryan Torn, and two anonymous reviewers for their helpful feedback that helped to improve this manuscript.

REFERENCES

Aberson, S. D., 2003: Targeted observations to improve operational tropical cyclone track forecast guidance. *Mon. Wea. Rev.*, **131**, 1613–1628, https://doi.org/10.1175//2550.1.

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

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

Bednarczyk, C. N., and B. C. Ancell, 2015: Ensemble sensitivity analysis applied to a southern plains convective event. *Mon. Wea. Rev.*, **143**, 230–249, https://doi.org/10.1175/MWR-D-13-00321.1.

Brooks, H. E., C. A. Doswell, and J. Cooper, 1994: On the environments of tornadic and nontornadic mesocyclones. *Wea. Forecasting*, **9**, 606–618, https://doi.org/10.1175/1520-0434(1994)009<0606:OTEOTA>2.0.CO;2.

Bryan, G. H., and J. M. Fritsch, 2002: A benchmark simulation for moist nonhydrostatic numerical models. *Mon. Wea. Rev.*, **130**, 2917–2928, https://doi.org/10.1175/1520-0493(2002)130<2917:ABSFMN>2.0.CO;2.

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

Chang, E. K. M., M. Zheng, and K. Raeder, 2013: Medium-range ensemble sensitivity analysis of two extreme Pacific extratropical cyclones. *Mon. Wea. Rev.*, **141**, 211–231, https://doi.org/10.1175/MWR-D-11-00304.1.

Cintineo, R., and D. Stensrud, 2013: On the predictability of supercell thunderstorm evolution. *J. Atmos. Sci.*, **70**, 1993–2011, https://doi.org/10.1175/JAS-D-12-0166.1.

Dahl, J. M., M. D. Parker, and L. J. Wicker, 2014: Imported and storm-generated near-ground vertical vorticity in a simulated supercell. *J. Atmos. Sci.*, **71**, 3027–3051, https://doi.org/10.1175/JAS-D-13-0123.1.

Errico, R. M., and T. Vukicevic, 1992: Sensitivity analysis using an adjoint of the PSU–NCAR mesoscale model. *Mon. Wea. Rev.*, **120**, 1644–1660, https://doi.org/10.1175/1520-0493(1992)120<1644:SAUAAO>2.0.CO;2.

Garcies, L., and V. Homar, 2010: An optimized ensemble sensitivity climatology of Mediterranean intense cyclones. *Nat. Hazards Earth Syst. Sci.*, **10**, 2441–2450, https://doi.org/10.5194/nhess-10-2441-2010.

Hacker, J., and L. Lei, 2015: Multivariate ensemble sensitivity with localization. *Mon. Wea. Rev.*, **143**, 2013–2027, https://doi.org/10.1175/MWR-D-14-00309.1.

Hall, M. C. G., 1986: Application of adjoint sensitivity theory to an atmospheric general circulation model. *J. Atmos. Sci.*, **43**, 2644–2652, https://doi.org/10.1175/1520-0469(1986)043<2644:AOASTT>2.0.CO;2.

——, and D. G. Cacuci, 1983: Physical interpretations of adjoint functions for sensitivity analysis of atmospheric models. *J. Atmos. Sci.*, **40**, 2537–2546, https://doi.org/10.1175/1520-0469(1983)040<2537:PIOTAF>2.0.CO;2.

——, ——, and M. E. Schlesinger, 1982: Sensitivity analysis of a radiative-convective model by the adjoint method. *J. Atmos. Sci.*, **39**, 2038–2050, https://doi.org/10.1175/1520-0469(1982)039<2038:SAOARC>2.0.CO;2.

Hill, A. J., C. C. Weiss, and B. C. Ancell, 2016: Ensemble sensitivity analysis for mesoscale forecasts of dryline convection initiation. *Mon. Wea. Rev.*, **144**, 4161–4182, https://doi.org/10.1175/MWR-D-15-0338.1.

Irvine, E. A., S. L. Gray, J. Methven, and I. A. Renfrew, 2011: Forecast impact of targeted observations: Sensitivity to observation error and proximity to steep orography. *Mon. Wea. Rev.*, **139**, 69–78, https://doi.org/10.1175/2010MWR3459.1.

Langland, R., 1999: Workshop on targeted observations for extratropical and tropical forecasting. *Bull. Amer. Meteor. Soc.*, **80**, 2331–2338, https://doi.org/10.1175/1520-0477(1999)080<2331:WOTOFE>2.0.CO;2.

Li, J., J. Du, D. Zhang, C. Cui, and Y. Liao, 2014: Ensemble-based analysis and sensitivity of mesoscale forecasts of a vortex over southwest China. *Quart. J. Roy. Meteor. Soc.*, **140**, 766–782, https://doi.org/10.1002/qj.2200.

Lorenz, E. N., and K. A. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.*, **55**, 399–414, https://doi.org/10.1175/1520-0469(1998)055<0399:OSFSWO>2.0.CO;2.

Markowski, P. M., and Y. P. Richardson, 2014: The influence of environmental low-level shear and cold pools on tornadogenesis: Insights from idealized simulations. *J. Atmos. Sci.*, **71**, 243–275, https://doi.org/10.1175/JAS-D-13-0159.1.

Morss, R. E., K. A. Emanuel, and C. Snyder, 2001: Idealized adaptive observation strategies for improving numerical weather prediction. *J. Atmos. Sci.*, **58**, 210–232, https://doi.org/10.1175/1520-0469(2001)058<0210:IAOSFI>2.0.CO;2.

Naylor, J., and M. S. Gilmore, 2014: Vorticity evolution leading to tornadogenesis and tornadogenesis failure in simulated supercells. *J. Atmos. Sci.*, **71**, 1201–1217, https://doi.org/10.1175/JAS-D-13-0219.1.

Nowotarski, C. J., and P. M. Markowski, 2016: Modifications to the near-storm environment induced by simulated supercell thunderstorms. *Mon. Wea. Rev.*, **144**, 273–293, https://doi.org/10.1175/MWR-D-15-0247.1.

Orf, L., R. Wilhelmson, B. Lee, C. Finley, and A. Houston, 2017: Evolution of a long-track violent tornado within a simulated supercell. *Bull. Amer. Meteor. Soc.*, **98**, 45–68, https://doi.org/10.1175/BAMS-D-15-00073.1.

Potvin, C. K., K. L. Elmore, and S. J. Weiss, 2010: Assessing the impacts of proximity sounding criteria on the climatology of significant tornado environments. *Wea. Forecasting*, **25**, 921–930, https://doi.org/10.1175/2010WAF2222368.1.

Romine, G. S., C. S. Schwartz, R. D. Torn, and M. L. Weisman, 2016: Impact of assimilating dropsonde observations from MPEX on ensemble forecasts of severe weather events. *Mon. Wea. Rev.*, **144**, 3799–3823, https://doi.org/10.1175/MWR-D-15-0407.1.

Szunyogh, I., Z. Toth, A. V. Zimin, S. J. Majumdar, and A. Persson, 2002: Propagation of the effect of targeted observations: The 2000 winter storm reconnaissance program. *Mon. Wea. Rev.*, **130**, 1144–1165, https://doi.org/10.1175/1520-0493(2002)130<1144:POTEOT>2.0.CO;2.

Torn, R. D., 2010: Ensemble-based sensitivity analysis applied to African easterly waves. *Wea. Forecasting*, **25**, 61–78, https://doi.org/10.1175/2009WAF2222255.1.

——, and G. J. Hakim, 2008: Ensemble-based sensitivity analysis. *Mon. Wea. Rev.*, **136**, 663–677, https://doi.org/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, https://doi.org/10.1175/2009MWR2879.1.

Weiss, C. C., D. C. Dowell, J. L. Schroeder, P. S. Skinner, A. E. Reinhart, P. M. Markowski, and Y. P. Richardson, 2015: A comparison of near-surface buoyancy and baroclinity across three VORTEX2 supercell intercepts. *Mon. Wea. Rev.*, **143**, 2736–2753, https://doi.org/10.1175/MWR-D-14-00307.1.

Weissmann, M., R. Busen, A. Dörnbrack, S. Rahm, and O. Reitebuch, 2005: Targeted observations with an airborne wind lidar. *J. Atmos. Oceanic Technol.*, **22**, 1706–1719, https://doi.org/10.1175/JTECH1801.1.

Wile, S. M., J. P. Hacker, and K. H. Chilcoat, 2015: The potential utility of high-resolution ensemble sensitivity analysis for observation placement during weak flow in complex terrain. *Wea. Forecasting*, **30**, 1521–1536, https://doi.org/10.1175/WAF-D-14-00066.1.

Wu, C.-C., J.-H. Chen, P.-H. Lin, and K.-H. Chou, 2007: Targeted observations of tropical cyclone movement based on the adjoint-derived sensitivity steering vector. *J. Atmos. Sci.*, **64**, 2611–2626, https://doi.org/10.1175/JAS3974.1.

Xie, B., F. Zhang, Q. Zhang, J. Poterjoy, and Y. Weng, 2013: Observing strategy and observation targeting for tropical cyclones using ensemble-based sensitivity analysis and data assimilation. *Mon. Wea. Rev.*, **141**, 1437–1453, https://doi.org/10.1175/MWR-D-12-00188.1.

Yokota, S., H. Seko, M. Kunii, H. Yamauchi, and H. Niino, 2016: The tornadic supercell on the Kanto Plain on 6 May 2012: Polarimetric radar and surface data assimilation with EnKF and ensemble-based sensitivity analysis. *Mon. Wea. Rev.*, **144**, 3133–3157, https://doi.org/10.1175/MWR-D-15-0365.1.

Zheng, M., E. K. M. Chang, and B. A. Colle, 2013: Ensemble sensitivity tools for assessing extratropical cyclone intensity and track predictability. *Wea. Forecasting*, **28**, 1133–1156, https://doi.org/10.1175/WAF-D-12-00132.1.