Faculty and Researcher Publications   Faculty and Researcher Publications Collection

2015

The potential utility of high-resolution ensemble sensitivity analysis for observation placement during weak flow in complex terrain

Wile, Sean M.

American Meteorological Society

SPECIAL
MATERHORN
COLLECTION

The Potential Utility of High-Resolution Ensemble Sensitivity Analysis for Observation Placement during Weak Flow in Complex Terrain

SEAN M. WILE, JOSHUA P. HACKER,* AND KENNETH H. CHILCOAT

Department of Meteorology, Naval Postgraduate School, Monterey, California

(Manuscript received 6 June 2014, in final form 14 May 2015)

ABSTRACT

Expansion in the availability of relocatable near-surface atmospheric observing sensors introduces the question of where placement maximizes gain in forecast accuracy. As one possible method of addressing observation placement, the performance of ensemble sensitivity analysis (ESA) is examined for high-resolution ($\Delta x = 4$ km) predictions in complex terrain and during weak flow. ESA can be inaccurate when the underlying assumptions of linear dynamics (and Gaussian statistics) are violated, or when the sensitivity cannot be robustly sampled. A case study of a fog event at Salt Lake City International Airport (KSLC) in Utah provides a useful basis for examining these issues, with the additional influence of complex terrain. A realistic upper-air observing network is used in perfect-model ensemble data assimilation experiments, providing the statistics for ESA. Results show that water vapor mixing ratios over KSLC are sensitive to potential temperature on the first model layer tens of kilometers away, 6 h prior to verification and prior to the onset of fog. Potential temperatures indicate inversion strength in the Salt Lake basin; the ESA predicts southerly flow and strengthened inversions will increase water vapor over KSLC. Linearity tests show that the nonlinear response is about twice the expected response. Experiments with smaller ensembles show that qualitatively similar conclusions about the sensitivity pattern can be reached with ensembles as small as 48 members, but smaller ensembles do not produce accurate sensitivity estimates. Taken together, the results motivate a closer look at the fundamental characteristics of ESA when dynamics (and therefore correlations) are weak.

1. Introduction

Fog events in the Salt Lake basin in Utah, with impacts on aviation operations at the Salt Lake City International Airport (KSLC), arise in a range of flow scenarios. Typically, weak synoptic forcing and nonlinear water phase changes present challenges to numerical weather prediction (NWP) models when fog is possible. Because interactions between the land–water surface and the lower atmosphere can strongly modulate fog production and dissipation, near-surface shelter and anemometer-height observations contain potentially useful information for both forecasters and NWP model initialization. Surface observation networks could

conceivably be designed to improve fog forecasts in regions particularly susceptible. At the heart of the network design is an understanding of numerical forecast sensitivity to initial-condition analysis perturbations that result from assimilating proposed hypothetical observations.

One candidate method for quantifying forecast sensitivity to observations is ensemble sensitivity analysis (ESA). ESA was elucidated by Ancell and Hakim (2007), who showed a theoretical equivalence to adjoint sensitivity under linear dynamics, Gaussian statistics, and an infinite ensemble. Rather than linearizing about a deterministic model trajectory as in adjoint sensitivity, statistical linearization is performed about an ensemble mean trajectory by exploiting the distributions sampled by the ensemble. ESA has been used primarily for identifying dynamical relationships where strong dynamical signals can be expected, such as North Pacific synoptic storms (e.g., Ancell and Hakim 2007; Hakim and Torn 2008) and tropical cyclones (e.g., Torn and Hakim 2009; Torn 2010). ESA has also shown promise in the area of observation network design, where the

impact of hypothetical or proposed observations can be evaluated (Torn and Hakim 2008).

Uncertainty characterized by Lorenz (1963), which accompanies the growth of initial errors under nonlinear chaotic dynamics, is accounted for in ESA by the production of instantaneous and independent ensemble state estimates. Those are in turn provided in an ensemble data assimilation system (cf. Gombos and Hansen 2008), which produces the samples for ESA.

Before applying ESA to evaluate the hypothetical observation impact on a forecast, we must be confident that the sensitivity estimates are robust and meaningful. Past studies applied ESA primarily on the synoptic scale with relatively coarse model grids, and the potential for ESA in weak flows and complex terrain, and at high resolution, is relatively unexplored. Ancell and Mass (2006) reported that the accuracy of adjoint sensitivities suffers as resolution increases, for example, because nonlinearity is greater. Ensemble methods possibly have an advantage here because the ensemble mean retains linear error growth characteristics longer than a deterministic forecast. Several previous studies examined sensitivities for a wind-power forecast metric [e.g., Zach et al. (2011) and reports cited therein] at spatial scales similar to those here, finding some utility in sensitivities just upstream from a wind farm. Those studies did not evaluate the validity of the assumptions underlying the ESA nor did they quantify the effect of sampling error.

Fog in the Salt Lake basin is expected to challenge ESA in two ways. First, the forcing is weak. Variability is on relatively small scales, and spatial covariances are weak. We can expect the effects of sampling error resulting from a finite ensemble to be significant in these conditions. Using ensemble sensitivities to propose new observations relies on quantitatively accurate sensitivity estimates. Sampling error arises from using finite ensembles, and is inevitable. Weak correlations in the ensemble statistics can be difficult to estimate accurately. Yet many high-impact forecast problems occur under weak dynamics accompanied by weak correlations. Second, the condensation associated with fog formation is highly nonlinear, which can invalidate the linear perturbation assumption.

The goals of this work are twofold. The first goal is to determine whether ESA can provide useful information for a fog event characterized by weak synoptic forcing. The second is to expose weaknesses in ESA for weakly forced events and where nonlinearity may be substantial. To meet those goals, we evaluate the response to a range of perturbations to initial conditions, comparing the response given by the nonlinear forecast model to the linear ESA estimates derived from a 96-member ensemble. Later, experiments with smaller

ensembles provide context for how the sensitivity estimates can vary with ensemble size. Results reported herein are relevant to other flow scenarios defined by weak dynamics and nonlinearity, including mountain-valley flows and convection over terrain.

This paper reports upon experiments performed to validate ESA in complex terrain, under weak dynamics, and with fine model grid spacing. Section 2 reviews some previous work on fog in the Salt Lake valley, introduces the January 2009 dense fog case that is the focus of this work, and presents details about Weather Research and Forecasting (WRF) Model implementation and the data assimilation framework. Section 3 briefly reviews ESA and details the ESA experiments. Section 4 evaluates the ESA results. Section 5 presents results from three distinct methods for testing the validity of the response predicted by the sensitivity estimates. Section 6 presents an analysis of the ensemble size necessary to gain these results, and section 7 reviews conclusions.

2. Event and simulations

a. Fog in the Salt Lake basin

Though the basic processes governing fog formation in valley basins is reasonably well understood, in-depth analysis of fog formation in the Great Salt Lake (GSL) basin is sparse. Peer-reviewed papers on the subject are limited, but include a study on the effect of the size of the GSL on fog formation (Hill 1988) and a partial examination of fog formation mechanisms in deep stable layers (Wolyn and McKee 1989). Reports and data are more numerous, and include a study on dense fog initiation in the Salt Lake valley by Hogan (2013), the characteristics of fog there (Slemmer 2004), and a climatological database of fog events (Alder et al. 1998).

Hogan's (2013) examination of 30 years of Salt Lake valley dense fog events identified the role of low-level inversions. Hogan identified two characteristics of an inversion enabling dense fog formation: a shallow vertical extent, less than 3700 ft in height, and moderate strength, exhibiting greater than 1.5°C $(1000 \, \text{ft})^{-1}$ lapse rate. Without meeting these criteria, the formation of dense fog at KSLC is unlikely (Hogan 2013).

Slemmer (2004) studied dense fog at KSLC based on surface observations from July 1971 through June 2001, identifying and categorizing parameters typically observed during dense fog events. Results indicate the following: 1) dense fog forms primarily in early winter, with 77% of annual occurrences observed in December and January and 18% in February; 2) northwesterly and southeasterly winds are most likely during dense fog events when winds are >3 knots (kt; 1 kt = 0.51 m s^{-1});
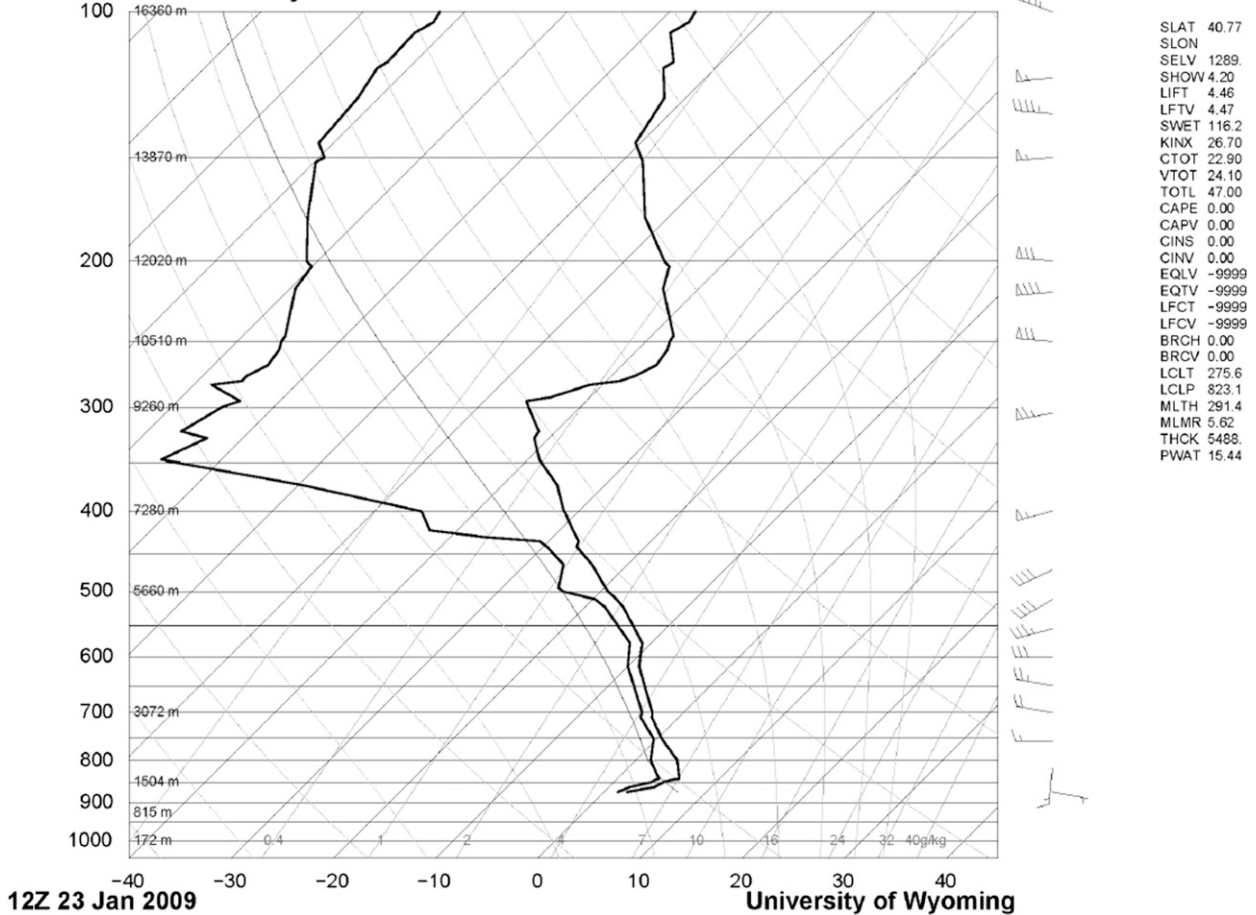
FIG. 1. The 1200 UTC 23 Jan 2009 sounding from KSLC indicating the presence of a strong low-level inversion and warm-air advection (http://weather.uwyo.edu/upperair/sounding.html).

and 3) less than 6% of all dense fog cases occurred with the temperature above freezing. Slemmer (2004) further categorized the dense fog events into three distinct event types: a persistent inversion case, or fog formation through radiative cooling at KSLC; precipitation falling into a weakened inversion and shallow cold pool case, where precipitation increased low-level moisture supporting fog formation; and shallow cold pool advecting from the GSL case, where radiative cooling over the lake enabled fog to form and later move over KSLC.

b. Dense fog event

At 2243 UTC 23 January 2009, dense fog developed over KSLC, forcing the closure of one runway and prompting the National Weather Service to issue a dense fog warning. Visibility remained ≤⅛ statutory miles for approximately 10 h during the event. A well-defined and slowly weakening omega block across the Rockies characterized the synoptic flow, and a weak

midlevel disturbance propagated through northern Utah on 23 January. This disturbance had two effects: it helped to weaken low-level inversions that had formed under subsidence across northern Utah, and it forced rain across the GSL that saturated the boundary layer. As this disturbance moved through northern Utah, synoptic gradients in the western United States induced southerly low-level flow and warm-air advection across much of the southern Intermountain West. Weak warm-air advection was evident at the surface as KSLC temperatures warmed from 1°C at 1500 UTC to 8°C at 2300 UTC during light rain. Stronger warm-air advection occurred at low levels above the surface (see Fig. 1), stabilizing the low-level inversion. The inversion prevented low-level moisture from mixing out, and the fog exhibited characteristics of the second fog scenario described by Slemmer (2004). Namely, precipitation fell into a weakening cold pool. The southerly flow contributed by persisting and strengthening the inversion.
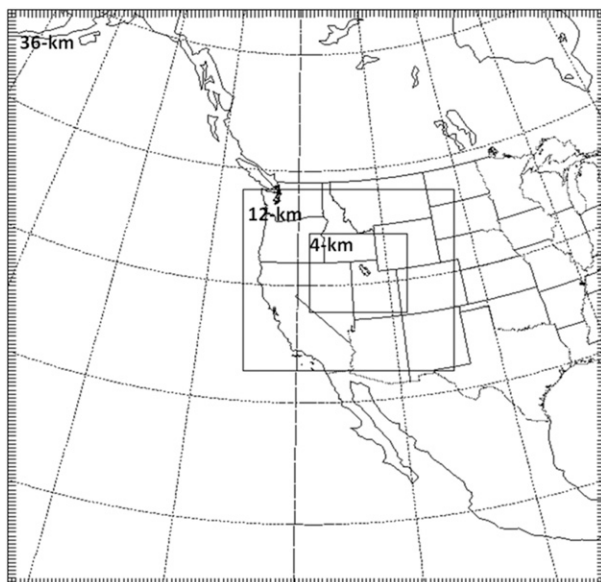
FIG. 2. Telescoping nested configuration used in this study. The outermost domain used 36-km horizontal grid spacing, the middle 12 km, and the innermost 4 km.

c. WRF Model setup

Version 3.2.1 of the Advanced Research version of WRF (Skamarock et al. 2008) was used to simulate the dense fog event. A 96-member ensemble was configured with a nested 36/12/4-km horizontal grid spacing configuration, centered over the Great Salt Lake (Fig. 2). Each grid contained 60 vertical η levels and an upper boundary at 100 hPa. All WRF simulations used the same physics suite as follows: the Noah land surface model (Chen and Dudhia 2001); the basic similarity theory surface layer scheme with the Beljaars (1995) convective velocity and stability functions from Paulson (1970), Dyer and Hicks (1970), and Webb (1970); the Yonsei University PBL scheme (Hong et al. 2006); the Rapid Radiative Transfer Model (Mlawer et al. 1997) for longwave radiation; the Dudhia (1989) scheme for shortwave radiation; and the WRF single-moment 5-class microphysics scheme (Hong et al. 2004). The Kain–Fritsch cumulus scheme (Kain 2004) was used on the 36- and 12-km outer and middle domains, and no cumulus scheme was in use on the 4-km domain.

The Data Assimilation Research Testbed (DART) facilitates ensemble data assimilation. DART is a community software environment created for ensemble data assimilation research, is principally developed by staff at the National Center for Atmospheric Research, and contains several varieties of ensemble filters (Anderson et al. 2009). The specific serial implementation of the ensemble adjustment Kalman filter (EAKF; Anderson 2001) is given by Anderson (2003). The EAKF is a deterministic approach to the Kalman filter, and differs from the original EnKF (Evensen 1994) primarily in that it does not rely on perturbed observations to inject noise into the system.

Perfect-model experiments, where a nature run with the WRF is sampled for synthetic observations (i.e., the "truth"), gives us results without the ambiguity introduced by unknown model errors. To produce the nature run, the WRF was initialized at 0000 UTC 19 January 2009 from the North American Regional Reanalysis (NARR), which also provided lateral boundary conditions. Soil temperature and moisture were reset every 3 h to the NARR, preventing land surface drift during the 10-day nature run. Synthetic observations were produced every 3 h; observation locations and physical quantities were identical to the actual radiosonde observations available in the National Centers for Environmental Prediction prepBUFR files at each assimilation time. The synthetic observations mirror the real observation frequency and spatial distribution of the radiosonde network, but the observation values differ because they are taken directly from the nature run.

Conditioning of the WRF ensemble began at 0000 UTC 20 January, and synthetic observations were assimilated every 3 h through 0000 UTC 24 January. As per the usual practice with the WRF in DART, observations can contribute to analysis increments on all domains within its influence (localization radius). At initialization time, 96 members of the ensemble system were formed by adding random, spatially consistent perturbations to the truth, drawn from the global static error background covariance field provided with the WRF variational data assimilation system (WRF-VAR; Barker et al. 2012). Lateral boundary condition perturbations to ensure sufficient spread at the boundaries were also drawn with WRF-VAR (e.g., Torn et al. 2006). Covariance localization with the fifth-order piecewise polynomial described by Gaspari and Cohn [(1999), their Eq. (4.10)] mitigated sampling errors from the finite ensemble. The localization length scale was approximately 1200 km. Adaptive covariance inflation following Anderson (2007) was also implemented to help mitigate the effect of bias, and insufficient ensemble spread, on the ensemble data assimilation. By not ingesting the NARR soil states directly, the ensemble can be biased in the lower boundary forcing (and consequently the lower atmosphere) compared to the nature run. Inflation helps to overcome the effects of bias on the ensemble filter by increasing the overlap between the
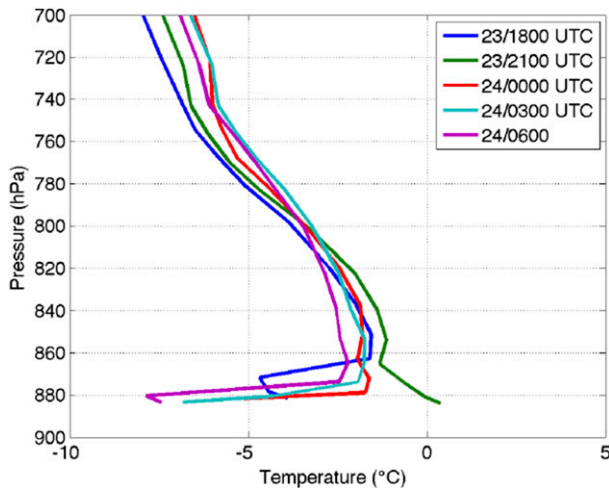
FIG. 3. Ensemble mean temperature from the surface to 700 hPa in 3-h intervals at KSLC from 1800 UTC 23 Jan to 0600 UTC 24 Jan.

distributions of background error and observation likelihood.

Ensemble initial conditions were then used to create 6-h hindcasts initialized at 1800 UTC 23 January and valid at 0000 UTC 24 January, approximately the onset time of dense fog. We interpret the hindcasts as forecasts here and refer to them as such.

d. Fog event simulation

Ensemble simulations show a shallow cool layer at the surface and southerly winds above. Southerly warm-air advection above the surface overran the cold pool and strengthened the inversion. Some light precipitation moistened the boundary layer and preceded observations of dense fog.

The WRF ensemble forecast simulated key aspects of the dense fog event. Figure 3 shows ensemble-mean 3-hourly profiles of temperature over KSLC, and Fig. 4 shows the model lowest-layer ensemble-mean water vapor and temperature through the period, overlaid with wind vectors on model layer 5 (approximately 270 m AGL). Although model surface temperature values are generally colder than observations, Fig. 3 shows a strong near-surface inversion that breaks up between 1800 and 2100 UTC 23 January, and reforms by 0000 UTC 24 January. The ensemble lacks substantive liquid water near the surface, consistent with systematic errors in the WRF liquid water simulations (e.g., Ryerson and Hacker 2014; Wilson and Fovell 2015). Instead, the water vapor mixing ratio q_v near the surface increases as moisture advects off the GSL with the colder air, and the inversion strengthens. Though the period of dense vapor concludes 3–6 h early compared

to the observations, the inversion and moisture advection appear to qualitatively agree with the observations, lending confidence to the interpretation of the ESA later.

3. Experiment

a. Ensemble sensitivity analysis (ESA)

The sensitivity of an arbitrary forecast metric J to an initial analysis state variable x_i was defined in Ancell and Hakim (2007) as

$$\frac{\partial J}{\partial x_i} \approx \frac{\text{cov}(J, x_i)}{\text{var}(x_i)}. \tag{1}$$

Covariance is denoted cov and quantifies the strength of the linear relationship between the two arguments. Variance is denoted var and quantifies analysis spread about the mean. Predictor x_i represents an analysis state variable at a single grid point and is an element of the analysis state vector. Samples of x_i and J are given by the ensemble analysis and forecast, respectively. The sensitivity in Eq. (1) then relates the expected change in a forecast metric J given a change to an analysis variable x_i. In this work, the forecast metric J quantifies water vapor in a small box over KSLC and x_i is an analysis gridpoint value on the first model layer (see section 3b).

Equation (1) is an approximation to the general form of the ensemble sensitivity. The full analysis covariance, as opposed to just the variance, is necessary for equivalence with adjoint sensitivity methods. To avoid inverting the large analysis ensemble covariance matrix, Ancell and Hakim (2007) proposed approximating the analysis ensemble covariance matrix with its diagonal, leading to the scalar approximation in Eq. (1). The full analysis covariance is necessary for equivalence with adjoint sensitivity methods for estimating gradients as in Eq. (1). The diagonal approximation is perfect in the limit that the analysis variables are all perfectly correlated predictors of the forecast metric, and is poor in the limit of analysis variables that are independent predictors of the forecast metric. In practice, the fidelity of the diagonal approximation to the complete sensitivity can vary with the scale of motion, which determines the spatial correlations. The approximate form has been used in the ensemble sensitivity literature to date. Equation (1) can easily be rewritten as a product of the linear (Pearson) correlation coefficient between J and x_i and the standard deviations of J and x_i.

The sensitivity is expected to be overestimated because of sampling error in the finite-sized ensemble. The analysis variance is expected to be systematically underestimated, and the covariance between the
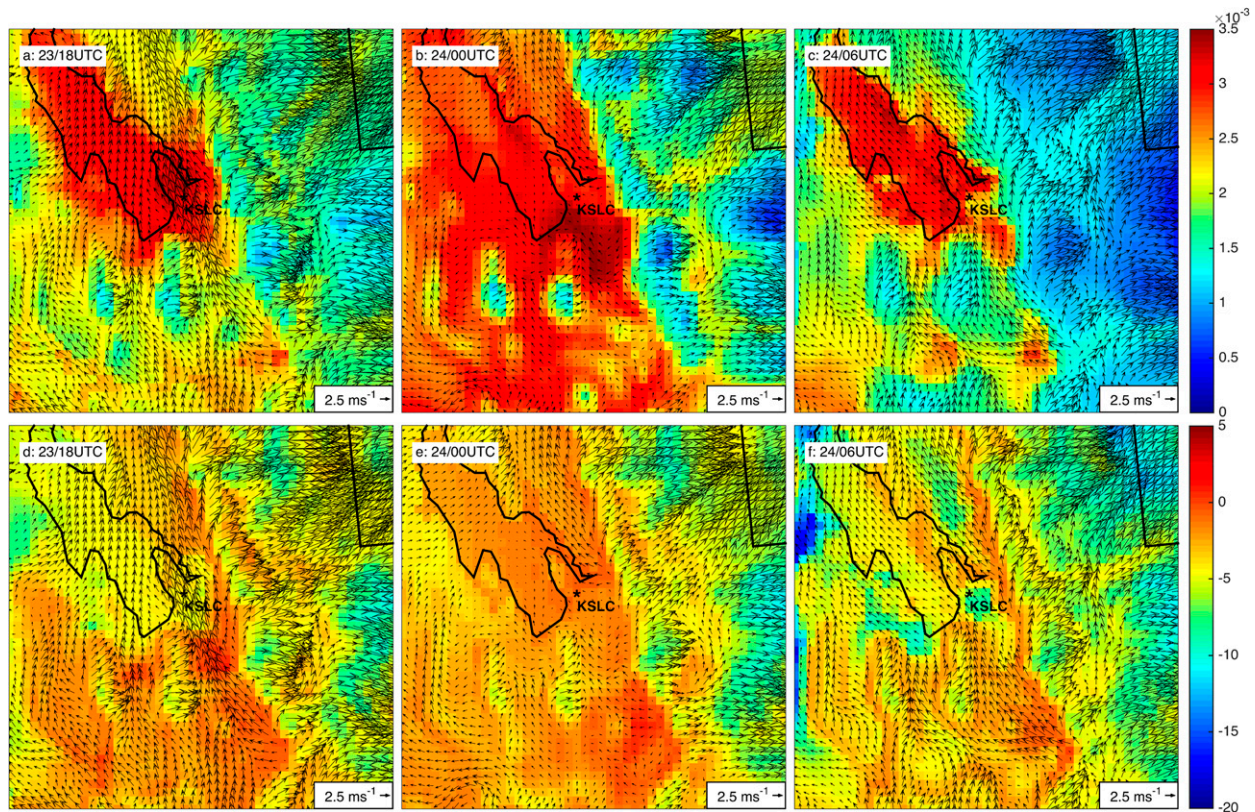
FIG. 4. (a)–(c) Lowest-layer ensemble-mean q_v (kg kg^{-1}) and (d)–(f) temperature (°C) in (from left to right) 6-h intervals from 1800 UTC 23 Jan to 0600 UTC 24 Jan. Vectors show winds on model level 5 (~270 m AGL).

analysis and the forecast is expected to be systematically overestimated. Sensitivity in Eq. (1) is then expected to be biased toward large magnitudes, and the predicted response to a known perturbation will also be most likely overestimated. To gain confidence in the ESA, we perform a Student's t test at the 95% significance level, for which the null hypothesis was that no linear relationship exists between J and x_i. We reject the null hypothesis when the absolute value of the correlation coefficient exceeds the 95% confidence bounds.

ESA is based on linear regression, which requires Gaussian distributions for proper interpretation. A test for Gaussian distribution can be used to eliminate sensitivities that should not be interpreted with confidence. Candidate sensitivity predictors (x) and predictands (J) are subject to a Lilliefors (1967) test. This test, a variation of the Kolmogorov–Smirnoff (K–S) test, is a two-sided goodness-of-fit check that measures the distribution of each variable set (empirical) to determine how well it compares against the theoretical normal distribution. We propose the null hypothesis that the empirical data come from a normally distributed population, with the alternative that the data do not come from a normal distribution. Candidate predictors and predictands that do not satisfy the K–S

goodness-of-fit test are rejected from further consideration. Predictors and predictands used in the remainder of this work cannot be rejected at the 0.95 confidence level from the K–S test.

b. Parameter selection

In addition to sampling error, characteristics of the analysis, its covariances, and the forecast metric distributions can affect sensitivity estimates. Careful selection of J ensures that sensitivities can be interpreted for the relevant physics of the flow. Because of some desirable characteristics described later, we chose J to be an average of the water vapor mixing ratio q_v within an $N = n_i \times n_j \times n_k$ box over KSLC.

After examining sensitivities to several analysis state variables, we focus on potential temperature θ at the first model layer (approximately 20 m AGL). The first model layer gives sensitivities relevant to surface observation stations that might be easily deployable, and θ is closely related to the temperature inversion often associated with dense fog formation in the GSL basin. Some additional state variables in our screening, including q_v and total dry air mass in the column, do not provide qualitatively coherent sensitivity relationships and were immediately rejected.
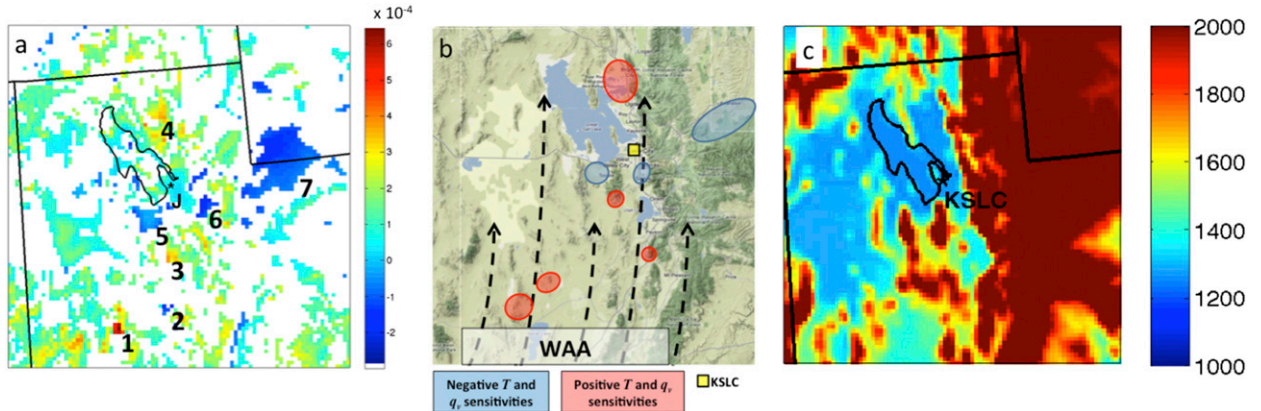
FIG. 5. (a) Sensitivities ($kg\,kg^{-1}\,K^{-1}$) using 1800 UTC 23 Jan analysis θ (K) as x_i, and forecasted box-mean q_v ($kg\,kg^{-1}$) as J, valid at 0000 UTC 24 Jan. White pixels are masked because they failed the hypothesis testing. Numbers indicate highest-ranking sensitivity magnitudes. (b) A conceptual diagram of the sensitivity field with pink (blue) areas indicating positive (negative) T and q_v sensitivity and arrows indicating the direction of the warm-air advection. (c) The WRF terrain with colored elevations from 1000 to 2000 m.

By choosing water vapor mixing ratio instead of liquid water to define the forecast metric J, we also avoid difficulties resulting from the highly nonlinear formation of fog. Compared with observations, the simulations adequately represent the ingredients of the fog, but do not condense water at the surface. The WRF Model's inability to reliably form fog is consistent with past studies (e.g., Ryerson and Hacker 2014); high values of q_v are, within this context, a better indicator that the model state should support fog formation. Conveniently, it also simplifies the sensitivity interpretation by avoiding the conversion to liquid. Because of inherent nonlinearity it is not straightforward to apply linear ensemble sensitivity with a forecast metric based on liquid water.

To ensure a Gaussian-distributed forecast parameter, and continuity between forecast lead times, N was allowed to vary depending on forecast variable and lead time. After trial and error, the smallest, yet still Gaussian J for 6- and 12-h forecast resulted from a $4 \times 4 \times 2$ box (approximately $16\,km \times 16\,km \times 70\,m$) centered on the grid point closest to KSLC. This box is used to define J in the remainder of the paper.

The sensitivity analysis and interpretation of perturbation experiments focus on the innermost domain in the WRF simulation. We next present the ESA and later test the WRF response to actual perturbations.

4. Sensitivity results

ESA is a statistical approach and cannot, by itself, prove that an analysis perturbation causes a specific forecast change. Even when subject to hypothesis testing, statistical relationships may be coincident and not causal. But in the absence of analytical solutions that

indicate cause and effect, spatial and temporal covariances in the sensitivity calculation can suggest a link between a forecast and an initial state. Intuition is helpful in proposing a plausible physical explanation and narrows down the meaningful sensitivities (e.g., Torn and Hakim 2008, 2009).

Results from the 6-h ESA with θ on the lowest model level as the independent variable are displayed in Fig. 5a. Broadly, positive sensitivities that cannot be rejected by the hypothesis testing are present at higher elevations where the first model layer is above the cold air and to the south of KSLC. Those regions indicate where a positive change in analysis potential temperature would result in an increase in q_v 6 h later at KSLC. Region 1, approximately 150 km south-southwest of the GSL, indicates the most positive sensitivity region within the innermost domain. Sensitivity in this area ranges from 4.43×10^{-4} to $6.62 \times 10^{-4}\,kg\,kg^{-1}\,K^{-1}$. At the most sensitive grid point (39.51°N, 112.92°W) the sensitivity predicts that a ± 0.052-K (i.e., the standard deviation of the analysis potential temperature; $\pm 1\sigma_\theta$) change in θ at the lowest model level will lead to an equivalent $\pm 3.41 \times 10^{-5}\,kg\,kg^{-1}$ change in the 6-h q_v at KSLC. While this value is rather small, an assumption allowing linear extrapolation predicts a $6.62 \times 10^{-4}\,kg\,kg^{-1}$ in 6-h q_v from a 1-K analysis change. A kelvin is a reasonable error in near-surface analysis θ, and $0.662\,g\,kg^{-1}$ is greater than 10% of the spatiotemporal variability evident in Fig. 4.

Other strong positive sensitivity responses appear in Fig. 5a across the central Utah desert ranges, in regions 2 and 3. Regions 1–3 are in the path of a plume of warm air pushing northward through the central Utah desert on 23 January (see temperature in Figs. 4d–f). The zonal
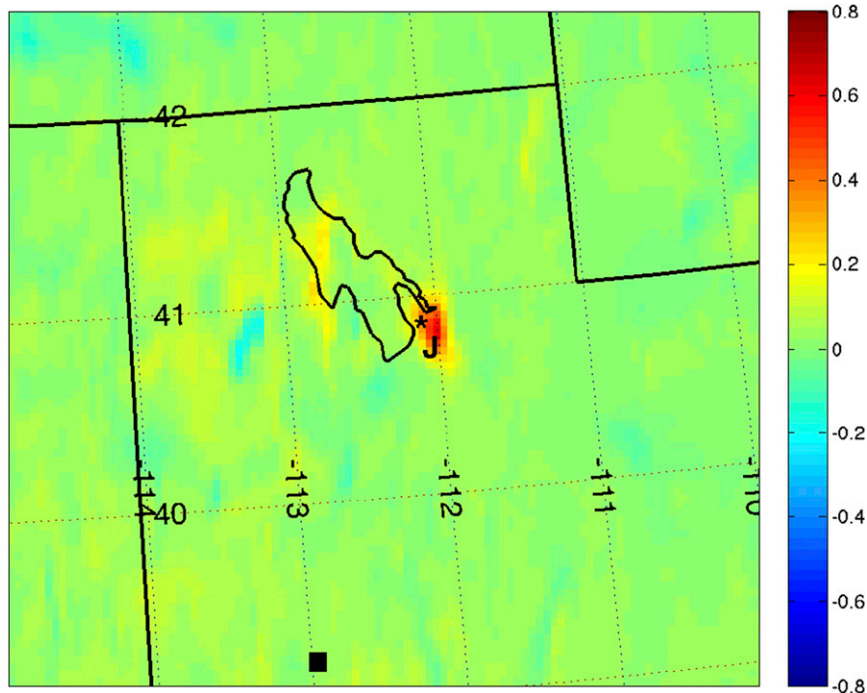
FIG. 6. Analysis of change in θ (K) after the addition of $+1\sigma_\theta$ perturbation applied at 39.51°N, 112.92°W (black square) on the first model level, regressed onto the entire analysis with the ensemble statistics. Positive (negative) values indicate an increase (decrease) in θ.

breadth of these features allows them to block or divert southerly flow. Because southerly flow in this region is likely associated with warm-air advection, spatial gradients in sensitivity to θ are apparent over the terrain that interacts most with the warm advection. Region 4, northeast of the GSL in a low-lying plain between the GSL and the Wasatch Range, indicates positive sensitivity as southerly warm-air advection is forced between the higher terrain of the Wasatch and the cold dome over the GSL.

The negative sensitivity regions most obvious in Fig. 5, region 5 (Tooele valley), region 6 (southeast Salt Lake valley), and region 7 (lee side of the Uinta Mountains), represent protected low-lying areas where cold dense air can pool during stable regimes, and the first model layer is in the cold air. They indicate where a negative perturbation to analysis potential temperature would result in a predicted increase in q_v 6 h later. Figure 5a shows several areas of negative sensitivity, with smaller magnitude relative to the positive regions; no values are less than -2.5×10^{-4} kg kg^{-1} K^{-1}.

The sensitivity regions offer a plausible path for fog formation (Fig. 5b). Warm-air advection prior to the event, as well as the importance of the low-level stability profile that supported fog, appear to be important according to the sensitivity estimates here and agree with

characteristics identified by Hogan (2013) and Slemmer (2004). Because KSLC is in a basin, negative sensitivities are most easy to link to fog at KSLC. Cooler and deeper layers in valleys, with less change to the temperature aloft, will strengthen the surface inversion. Assuming soil temperatures increase slowly compared to the first model layer, the first-layer temperature increases are indicative of a more stable temperature profile at the surface; positive sensitivities are consistent with a strengthening inversion over KSLC. The sensitivity field implies that advection of warmer air over the southern part of the domain, and more evident at higher elevations (see Fig. 5c for elevation), is correlated with cooler air in the low-lying valleys. A positive perturbation at the analysis time where the sensitivity is positive will correspond with cooling where the sensitivity is negative. Broadly, then, the sensitivity field is consistent with strengthening a near-surface inversion to support increased near-surface water vapor and the likelihood of fog over a large region.

Greater positive sensitivity to the south indicates the importance of the warm-air advection in dynamically strengthening the inversion. Warming to the south causes isentropes to slope upward toward the north. To the extent that the flow is isentropic, a positive sensitivity to the south will be correlated with a positive

sensitivity above the cold air over KSLC. The analysis statistics support these relationships. For brevity we do not show them here, but the vertical structure of the sensitivities also supports this conceptual model.

5. Tests of linear approximation

The sensitivities described in the last section, resulting from a linear statistical approximation, are meaningful only if the linear approximation is a good one. Three tests provide a basis to judge the effectiveness of ESA for identifying the sensitive points. Computational expense prevents an exhaustive exploration here, but the experiments suggest where further work is needed. A perturbation δx_i to state variable x_i, located at the point of greatest sensitivity on the first model layer (Fig. 5a), $X = 121$, $Y = 52$ (39.51°N, 112.92°W), is first applied. Each remaining state variable in the analysis is perturbed according to its linear relationship with the sensitivity point, estimated from the ensemble statistics.

The expected change in the forecast metric for a given scalar analysis perturbation δx_i is presented:

$$\delta J = \frac{\text{cov}(J, x_i)}{\text{var}(x_i)} \delta x_i, \qquad (2)$$

where δx_i is a perturbation to the ensemble mean (analysis state estimate) chosen differently for different experiments below. Comparing the predicted change in Eq. (2) to a forecast change realized by integrating the nonlinear model forward from a perturbed state is helpful in quantifying the accuracy of the expected change (i.e., the accuracy of the linear approximation). Differences between the predicted change and the actual forecast change may indicate nonlinearity, the effects of sampling error in the analysis statistics, or both.

The analysis state vector perturbation is formed by perturbing the remaining state elements according to each element's linear relationship with the selected x_i; here, x_i is chosen as the gridpoint potential temperature showing the greatest sensitivity. The magnitude of the perturbation assigned to that grid point also determines the vector magnitude of the analysis state perturbation.

Perturbing the remaining state according to each element's linear relationship with the sensitivity point accomplishes two things. First, it reduces the chance that resulting imbalances will cause the perturbation to immediately lose its energy and produce spurious gravity waves when the forecasts begin. Second, it builds in some part of the analysis ensemble covariances that are implicitly, but not explicitly, considered in the approximate sensitivity calculation [Eq. (1)]. The perturbations are imperfect because of sampling error in the analysis,

which was not accounted for in the diagonal sensitivity approximation; in general, the perturbation vector magnitude will be too large because analysis covariances are overestimated.

We could choose any analysis variable x_i to determine the magnitude of the analysis perturbation; by using the variable displaying the maximum sensitivity, we gain a relatively large perturbation and can more easily relate the nonlinear response to the predicted response. Sensitivity is determined by two things: 1) r_{xJ}, which explains the strength of the linear relationship between x_i and J, and 2) and the ratios of σ_J/σ_x, which quantifies the relative uncertainties (with units) of J and x_i. Maximum sensitivity occurs where for a fixed σ_J, a large correlation coefficient r_{xJ} is collocated with small spread (standard deviation) of an analysis state variable σ_x. Details of the analysis perturbation differ in different experiments explained below.

By perturbing the analysis at a point x_i, the slope of the regression line between x_i and J gives the expected forecast change. Further, assimilating an observation at a sensitivity point decreases σ_x, reflecting more certainty in the analysis. In this case the uncertainty in forecast J should also reduce as long as the linear approximation is valid, because to a linear approximation each member of the ensemble forecast is perturbed toward the ensemble mean of J.

a. Linearity tests with direct perturbation

A perturbation can be considered small if it is in the high-probability region of the analysis distribution at the sensitivity point, and we can expect the linearization to provide a good estimate of the actual response to small perturbations. First, a θ perturbation of $\delta x_i = \sigma_\theta = 0.0516$ K is applied to the first model layer at the greatest 6-h sensitivity point. The effect of this perturbation is spread to other model variables in 3D with univariate linear regression based on the analysis ensemble statistics. The regression step follows Torn and Hakim (2008) and others, and is a simple attempt at creating initial conditions that will retain the perturbation. Here, every ensemble member is perturbed identically so that the analysis spread is left unchanged. Using the analysis covariance to produce the initial vector perturbation introduces sampling error that is different from the expected error in the ESA, which does not explicitly use the analysis covariance in the diagonal approximation [Eq. (1)]. The covariances, and thus the regression coefficients, are likely to be erroneously large.

Figure 6 shows the 1800 UTC 23 January analysis perturbation θ (K) on the first model level resulting from the $+1\sigma_\theta$ perturbation applied to the most sensitive grid

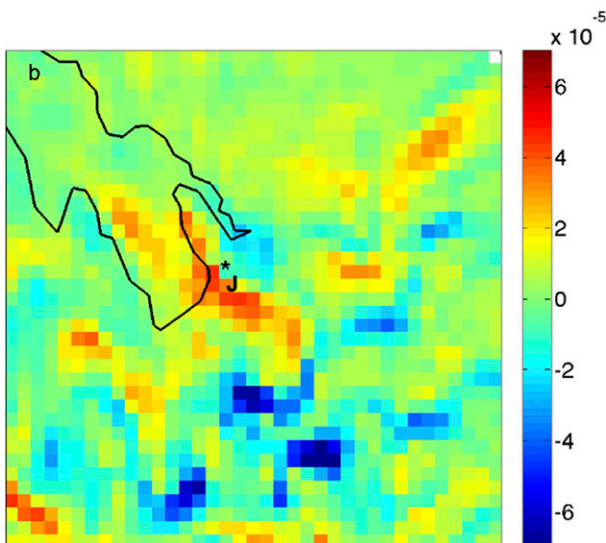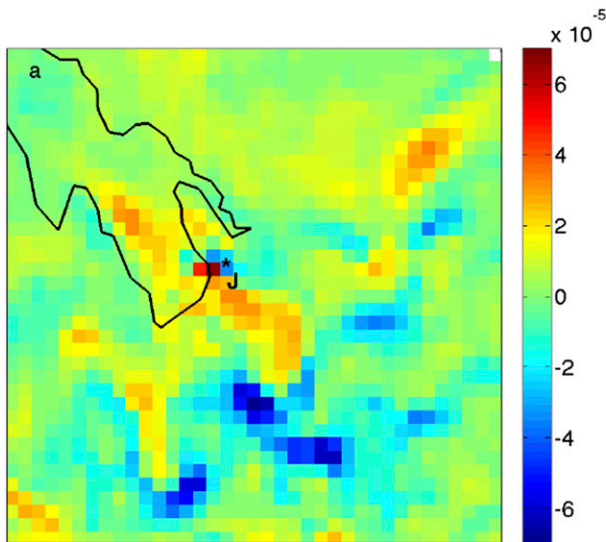FIG. 8. As in Fig. 7a, but for a $10\sigma_\theta$ perturbation.

FIG. 7. Forecast difference in q_v (kg kg^{-1}) at 0000 UTC 24 Jan at (a) the lowest model level and (b) the second model level following the introduction of a $+1\sigma_\theta$ perturbation to the 1800 UTC 23 Jan ensemble analysis at the most sensitive grid point.

point in the 6-h ESA (black square) and regressed onto the remaining state. Analysis θ increases slightly across the Salt Lake valley, with a stronger response at KSLC. Also evident in Fig. 6 is a region of decreased temperatures west of the GSL in the low-lying Great Salt Lake Desert. A warm perturbation introduced in the higher terrain south of the Great Salt Lake Desert could, with southerly flow, increase warm-air advection over the PBL in this region, creating or strengthening a low-level inversion. The cooler temperatures west of the GSL, introduced from the analysis statistics, are indicative
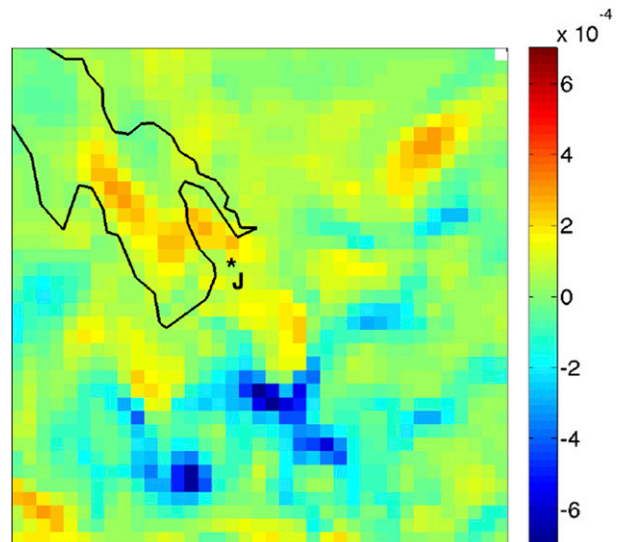
of a strengthening inversion and northward-sloping isentropes.

Integrating all members of the perturbed ensemble forward to the valid forecast time, 0000 UTC 24 January, gives the nonlinear forecast response. Figure 7 shows the difference in 6-h forecasted q_v (kg kg^{-1}) between the forecast from the 1800 UTC 23 January $+1\sigma_\theta$ perturbed analysis and the forecast from the 1800 UTC 23 January control analysis. Figures 7a and 7b show the first two model layers, respectively, illustrating the contributions to J, and both clearly show increased q_v above the KSLC area.

The slice across the region defining J shows increased q_v, compared to the control forecast, over the southern GSL and south of KSLC. The exact forecast difference in J is an increase in q_v so that the resulting $\delta J = 2.2 \times 10^{-5}$ kg kg^{-1}, which can be compared to the predicted change of 3.4×10^{-5} kg kg^{-1}. That is, the change from the nonlinear forecast is about 65% of the change predicted from the linear sensitivity.

A larger perturbation of $10\sigma_\theta = 0.516$ K is imposed to test whether a similar relationship between predicted and actual changes holds. The perturbation is an extreme value relative to the ensemble statistics and may be too large for the linear assumption to be valid. Because the analysis statistics for regressing that perturbation to the other state variables is unchanged, the analysis is perturbed as in Fig. 6 but scaled by a factor of 10. The resulting forecast (Fig. 8) shows q_v structures similar to those in Fig. 7a, but also greater magnitude. The details in Figs. 7 and 8 differ, including a broader swath of increased water vapor over the southern GSL and nearby. All grid points around J show increases,

contrasting [Fig. 7](#), which shows some localized decrease near J. Compared to a predicted $\delta J = 3.4 \times 10^{-4}\,\mathrm{kg\,kg^{-1}}$, the resulting forecast change is $1.3 \times 10^{-4}\,\mathrm{kg\,kg^{-1}}$ (about 40% of predicted).

The nonlinear forecast response from small and large perturbations, regressed onto the state with ensemble statistics, is half or less of the linear sensitivity predictions. This can result from either nonlinearity in the forecast or from sampling error in the analysis covariance. Both factors contribute here. Analysis regression statistics most likely lead to initial-condition perturbation magnitudes that are too large. We expect that the ESA gives sensitivity estimates that are too large, also because of sampling error. The effect of sampling error is multiplied by 10 when introducing the $10\sigma_\theta$ perturbation compared to the $1\sigma_\theta$ perturbation; the result is a larger error in analysis perturbation magnitude. If linear, the forecast response would reflect that. But the $10\sigma_\theta$ perturbation elicits a smaller forecast response compared to the predicted response, suggesting nonlinearity is also important. In the next section we test the trade-offs within a data assimilation context.

b. Linearity tests with assimilation

Covariance localization in ensemble data assimilation is meant to mitigate sampling error arising from a finite ensemble (e.g., [Houtekamer and Mitchell 1998](#)). Localization reduces the magnitude of the resulting analysis increments, from any single observation, as a function of distance from that observation. Compared to direct perturbation with linear regression, the smaller analysis perturbation would be expected to lead to a smaller δJ in the forecast. The smallest domain is the focus here, and a localization length scale of approximately 1200 km means the localization is generally weak at these scales, but it does have some effect. Inflation, to address underdispersion in the ensemble, is spun up to produce a filter that is closer to optimal as measured by the 3-h forecast agreement with observations. Assimilating an observation to impose a perturbation helps gauge the importance of using the assimilation system as the basis for sensitivity estimates.

The observation is perfect, with zero observation error, to isolate the effects of the localization. Normally temperature is assimilated, rather than potential temperature. The single synthetic observation that is equal to the analysis temperature plus one standard deviation (σ_T) at 1800 UTC 23 January at the 6-h sensitivity point is simply appended to the observation set for that time, and the assimilation is executed again.
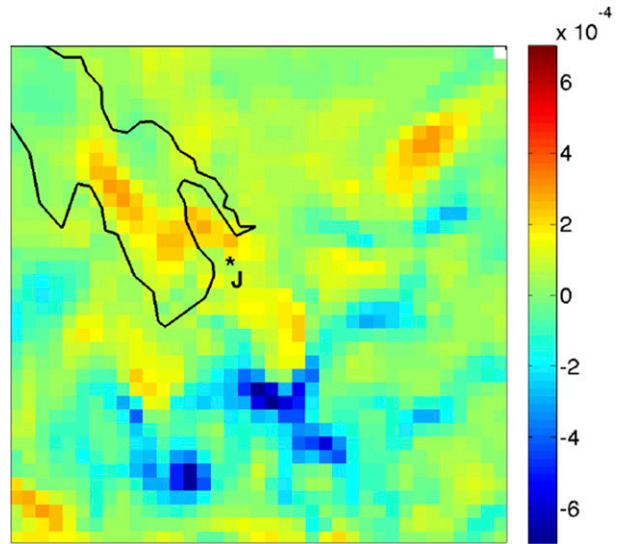

FIG. 9. Forecast difference in q_v ($\mathrm{kg\,kg^{-1}}$) at 0000 UTC 24 Jan at the lowest model level following the 1800 UTC 23 Jan assimilation of a synthetic observation that is σ_T greater than the analysis at the most sensitive grid point.

Assimilating the synthetic observation leads to an analysis perturbation that is qualitatively similar to that shown in [Fig. 6](#). The perturbation magnitude is smaller than for direct regression, and the smaller perturbation magnitude is easily quantified. The L2 norm of the first-layer θ perturbation measures the analysis perturbation magnitude. It is 3.7 K when assimilating, compared to 5.9 K when regressing. Differences in some details (not shown) arise because the ensemble spread is also reduced, where it is not changed when the direct regression is applied.

The resulting q_v forecast ([Fig. 9](#)) shows results qualitatively similar to the regression experiments ([Fig. 7a](#)), but with some details affecting the forecast metric J. Comparing [Figs. 8](#) and [9](#), forecasts from the regression shows less drying area and magnitude, and more areas of greater moistening around J. Here, $\delta J = 9.9 \times 10^{-6}\,\mathrm{kg\,kg^{-1}}$ from the nonlinear forecast compared to $\delta J = 2.7 \times 10^{-5}\,\mathrm{kg\,kg^{-1}}$ predicted from Eq. [(2)](#). The actual forecast change after assimilation is approximately 37%, compared to 65% when the perturbation is formed via direct regression, consistent with a smaller perturbation resulting from the data assimilation.

Similar to the regression experiments above, a large perturbation of approximately $10\sigma_T$ can be introduced via assimilation. The L2 norm of the θ perturbation on the first model layer is 30.9 K, compared to 59.2 K for the regression. The resulting forecast $\delta J = 8.8 \times 10^{-5}\,\mathrm{kg\,kg^{-1}}$, approximately 38% of $\delta J = 2.3 \times 10^{-4}\,\mathrm{kg\,kg^{-1}}$, as predicted from Eq. [(2)](#).

Results from the data assimilation experiments show the effect of mitigating sampling error in the analysis perturbation and retain the effects of nonlinearity. Because the localization reduces the overestimated covariances underlying the analysis perturbation, the initial perturbation and the resulting forecast response are smaller than when direct regression is used to form the perturbation. The ESA and its overprediction from sampling error are unchanged. A strategy for mitigating sampling error in the sensitivity calculation, which would reduce the sensitivity estimates, is complicated by dynamics evolving in space and time. That subject is addressed by Hacker and Lei (2015). As discussed next, these results are merely suggestive, and many more (computationally demanding) experiments are needed to confirm these results.

6. Effects of ensemble size

The number of ensemble members needed to perform a robust ESA is a primary focus of this study. Determining the necessary ensemble size helps establish computational requirements for a reliable ESA. Implementing ensemble data assimilation systems with smaller ensembles, alongside the full 96-member ensemble, forms the basis for comparison. At 0000 UTC 22 January, subsamples of 80, 64, and 48 ensemble members were randomly chosen from the 96-member ensemble valid. From that time forward, the smaller ensembles were maintained separately from the 96-member ensemble while assimilating the same observations. By running through several observation intervals, the smaller ensembles stabilize at a statistical state representative of that smaller ensemble. Assimilating the same observations keeps each of the ensembles close to each other in the mean, but the ensembles are not each sampling from the same distribution.

Sensitivities are computed as before and here again we focus on the forecast q_v sensitivity to θ. The three sensitivities at 6 h with greatest magnitude are identified, by grid point, for each ensemble. Sensitivities failing the 95% confidence test are masked from the results and are not considered in the ranking. We use the 96-member ensemble as the standard against which the smaller ensembles are judged.

One result is that the smaller ensembles show less spatial correlation in the sensitivity values (Fig. 10), diminishing with ensemble size. In the 96-member ensemble, the second and third most sensitive points are adjacent to the most sensitive point, used in the perturbation experiments above. Ensembles of 80 and 48 members produce the top three sensitivities in the same area as the greatest sensitivity in the 96-member

ensemble. Also in all of the smaller ensembles, the three greatest sensitivities are scattered rather than clumped (excepting the 64-member ensemble, where ranks one and three are adjacent).

The 80-member ensemble produces sensitivities that broadly agree with the positioning of the 96-member sensitivity features. The magnitude does not match as well, with only the strongest sensitivity grid point retaining its position. The sensitivity at the most sensitive grid point reduces from 6.62×10^{-4} to $3.02 \times 10^{-4} \, \mathrm{kg \, kg^{-1} \, K^{-1}}$. Ranked sensitivities 2 and 3 are not in the same locations as for the 96-member ensemble, suggesting that the additional sampling errors are large enough to affect the weaker sensitivities even if they pass the hypothesis test.

The 64-member ensemble also retains a similar geographic arrangement as the 96-member ensemble, with the most salient features qualitatively evident. But three of the most sensitive locations all differ from the 96-member sensitivities. The first-ranked sensitivity from the 96-member system is at the same grid point as the ninth-ranked sensitivity from the 64-member sensitivity. The deterioration in magnitude and different position sensitivity patterns suggest this is likely below the minimum size ensemble needed to accurately estimate a 6-h ESA in this case.

Finally, the 48-member ensemble retains the most salient sensitivity patterns from the 96-member ensemble, though most of the sensitivities fail the hypothesis test. The response is in some ways more like the response of the 96-member than 64-member ensemble; the second- and ninth-most sensitive grid points are both where the greatest sensitivity is in the 96-member ensemble. Although at first this seems counterintuitive, it is consistent with increased sampling error. An overestimation of the sensitivity is most likely, but some probability of an underestimation is possible. The estimates have lost little sensitivity magnitude, as the sensitivity at point 2 in Fig. 10d is $4.24 \times 10^{-4} \, \mathrm{kg \, kg^{-1} \, K^{-1}}$, which is greater than estimated from either the 80-member or 64-member ensemble at the same grid point.

Figure 11 shows both correlation and sensitivity from the 6-h ESA for each ensemble size. The green bars indicate the upper and lower boundaries of the 95% confidence interval for the correlation coefficient, and show the expected increase in sampling error associated with smaller ensembles. Figure 11 shows that ensembles smaller than 80 members exhibit reductions in sensitivity and linearity. The sampling error associated with the correlation coefficient (r value) of each ensemble member increases, which is also as expected. The green error bars for 80 members indicate variability
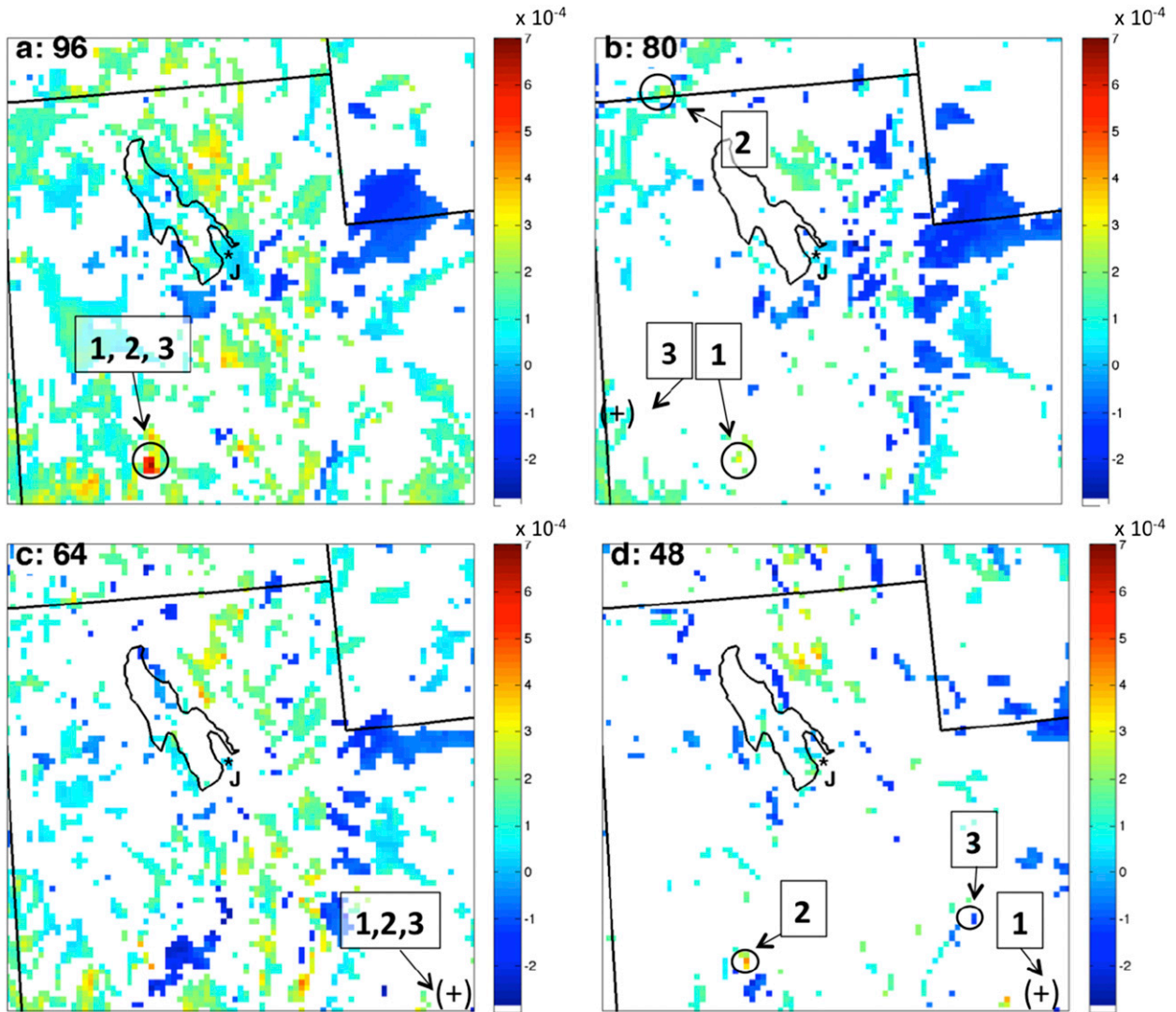
FIG. 10. The 6-h sensitivity $(\mathrm{kg\,kg^{-1}\,K^{-1}})$ for (a) 96-, (b) 80-, (c) 64-, and (d) 48-member ensemble systems using initial condition θ (K) from 1800 UTC 23 Jan 2009 and forecast q_v $(\mathrm{kg\,kg^{-1}})$ at KSLC at 0000 UTC 24 Jan 2009. Numbers identify grid points of the three sensitivities with the greatest magnitude for each ensemble. Markers indicate a positive (plus sign) or negative (minus sign) sensitivity that exists outside the cropped map. Note scales are not normalized.

$(r = \pm 0.20)$ in the correlation coefficients associated with this ensemble.

7. Summary and conclusions

This work examines the potential and behavior of ensemble sensitivity analysis during weak flow and in complex terrain. Most prior ensemble sensitivity studies have been completed on cases of stronger forcing, and at coarser scales, where we expect sensitivities to be stronger, smoother, and less susceptible to sampling error that arises from finite ensembles.

A case of dense fog that shut down the Salt Lake City International Airport (KSLC) is a useful experiment basis because the synoptic forcing is weak, though not completely absent, and the interactions between water vapor and thermal inversions present challenges to linear methods. A series of perfect-model data assimilation experiments with DART and the WRF Model enable the ensemble sensitivity analysis, where the forecast parameter (response function) is an average water vapor mixing ratio in a box over the airport. Forecasts are most sensitive to potential temperature, at analysis time, on the first model layer.

The overarching goal is to evaluate the ensemble sensitivity method for plausibility and robustness in these circumstances. Results address several questions.
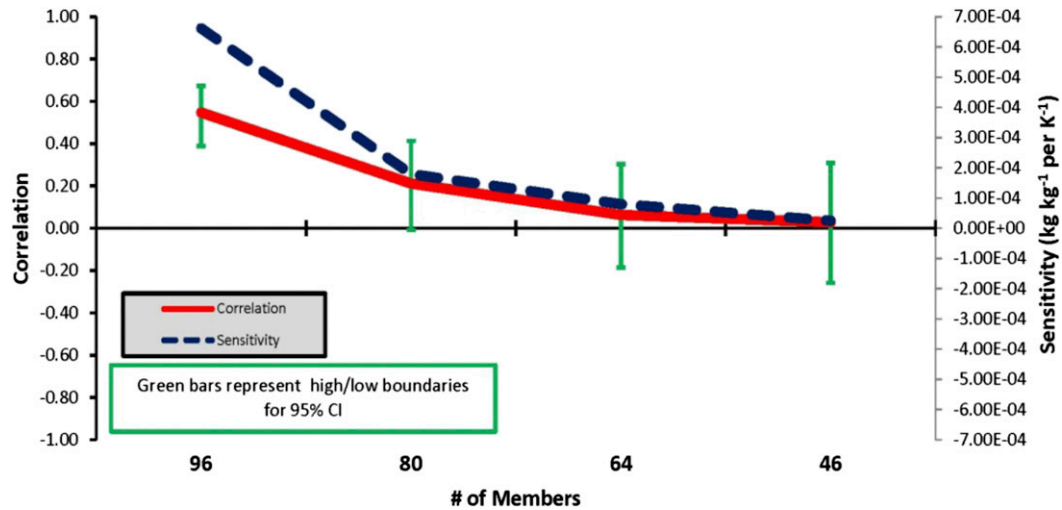
FIG. 11. Correlation (red) and sensitivity (blue) vs the number of ensemble members for a single point of positive sensitivities at 39.5°N, 112.9°W. Green bars indicate the high and low boundaries of the 95% confidence interval.

First, physically plausible sensitivity estimates are possible at lead times of 6 h. Patterns are consistent with southerly warm-air advection strengthening the inversion over KSLC, which has been observed as a mechanism for fog formation there.

Second, the forecast response magnitude is monotonic with initial perturbation magnitude, but the actual forecast response is consistently smaller than the response predicted by the ESA. Both nonlinearity and sampling error contribute to the disagreement. Sampling error leads to expected overprediction of both sensitivities and initial vector perturbations formed from regressing on the analysis covariances. Analysis perturbations formed by either assimilation of a synthetic T observation, or directly regressing a gridpoint θ perturbation to the remaining analysis, produce similar analysis perturbation structures. But covariance localization applied during assimilation reduces the perturbation magnitude.

Similar behavior for large perturbations shows that that the degree of nonlinearity does not change much over a large range of perturbation magnitude. Small and very large perturbations, of 1 and 10 standard deviations of the analysis distribution at the most sensitive grid point, are integrated with the nonlinear model in the ensemble. In all cases, forecast perturbations are half or less of that predicted by the sensitivity analysis.

Third, fairly large ensembles are needed to detect the weak correlations and provide reliable statistics during weak flow scenarios. Results here suggest that even 96 members may not be enough to clearly identify the most important sensitivities. In smaller ensembles, many of the patterns are intact, if weaker, but some of the details change.

Although the predicted responses are systematically too large, the results suggest that additional observations capturing the southerly warm-air advection may aid fog forecasting in the area around KSLC. This study has identified three regions. Region 1 in Fig. 12 represents the region of strongest sensitivity, and an observation there may provide advance warning of the impinging low-level warm air. Region 2 is where southerly warm-air advection meets the influence of the GSL cold dome. Additional observations placed in this region, the southeast Salt Lake valley, or along the Traverse Mountains to the south would help to define the warm-air advection as well as the inversion strength in the southeast GSL basin. Region 3 is a low-lying area northeast of the GSL, which exhibits strong sensitivity when warm-air advection is present across the GSL basin. Though a single observation exists on the northeast boundary of this region at Brigham City Airport (KBMC), additional observations across this plain would allow more thorough observation of warm air channeling between the Wasatch Range and a cold dome over the GSL. Combined, observations at these areas may help to characterize the near-surface conditions in the GSL known to lead to dense fog events at KSLC.

Finally, the results herein motivate a more focused study addressing the fundamental characteristics of ensemble sensitivities under sampling error. That is the primary subject of a forthcoming manuscript.
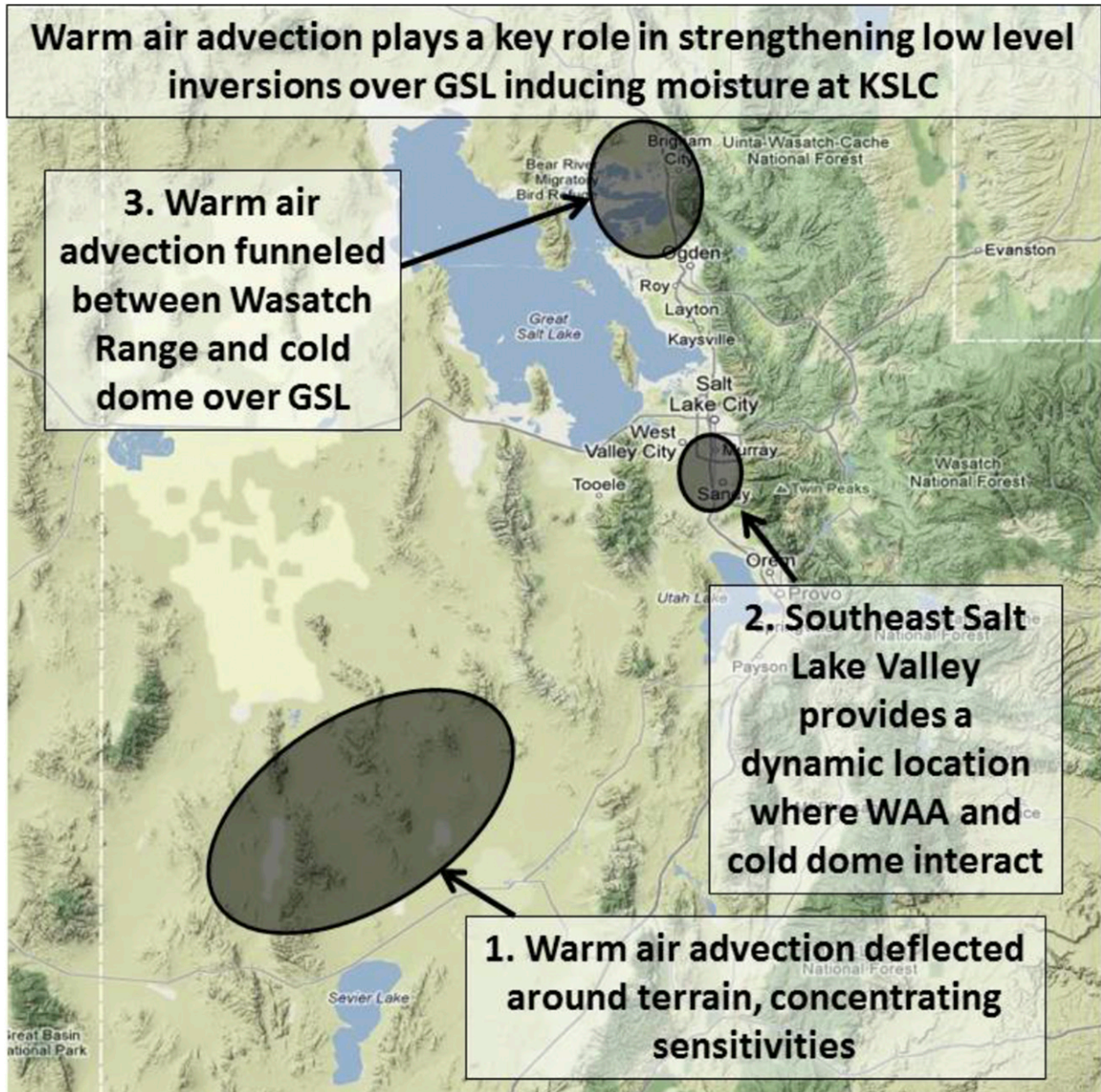
FIG. 12. Locations of strong and/or reliable θ sensitivity regions that are considered important in forecasting q_v at KSLC (map from Google Maps).

Terrain Atmospheric Modeling and Observations (MATERHORN) program. We are grateful for continued help from Jeff Anderson and the DART team at NCAR.

REFERENCES

Alder, W. J., L. S. Nierenberg, S. T. Buchanan, W. Cope, J. A. Cisco, C. C. Schmidt, A. R. Smith, and W. E. Figgins, 1998: Climate of Salt Lake City, Utah. NOAA Tech. Memo. NRS WR-152, 89 pp.

Ancell, B. C., and C. F. Mass, 2006: Structure, growth rates, and tangent linear accuracy of adjoint sensitivities to horizontal and vertical resolution. *Mon. Wea. Rev.*, **134**, 2971–2988, doi:10.1175/MWR3227.1.

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

——, 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: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus*, **59A**, 210–224, doi:10.1111/j.1600-0870.2006.00216.x.

——, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296, doi:10.1175/2009BAMS2618.1.

Barker, D., and Coauthors, 2012: The Weather Research and Forecasting Model's Community Variational/Ensemble Data Assimilation System: WRFDA. *Bull. Amer. Meteor. Soc.*, **93**, 831–843, doi:10.1175/BAMS-D-11-00167.1.

Beljaars, A. C. M., 1995: The parameterization of surface fluxes in large-scale models under free convection. *Quart. J. Roy. Meteor. Soc.*, **121**, 255–270, doi:10.1002/qj.49712152203.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model description and implementation. *Mon. Wea. Rev.*, **129**, 569–585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Dyer, A. J., and B. B. Hicks, 1970: Flux-gradient relationships in the constant flux layer. *Quart. J. Roy. Meteor. Soc.*, **96**, 715–721, doi:10.1002/qj.49709641012.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, doi:10.1029/94JC00572.

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.

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.

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

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., 147–162.

Hill, G. E., 1988: Fog effect of the Great Salt Lake. *J. Appl. Meteor.*, **27**, 778–783, doi:10.1175/1520-0450(1988)027<0778:FEOTGS>2.0.CO;2.

Hogan, D., 2013: Salt Lake Valley dense fog initiation study. National Weather Service/Western Region Headquaters. [Available online at http://www.wrh.noaa.gov/slc/projects/fogstud/FOGSTUDY.HTM.]

Hong, S.-Y., J. Dudhia, and S.-H. Chen, 2004: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Wea. Rev.*, **132**, 103–120, doi:10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2.

——, Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, doi:10.1175/MWR3199.1.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, doi:10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Lilliefors, H., 1967: On the Kolmogorov–Smirnov test for normality with mean and variance unknown. *J. Amer. Stat. Assoc.*, **62B**, 399–402, doi:10.1080/01621459.1967.10482916.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Mlawer, E. J., S. J. Taubman, P. D. Brown, and M. J. Iacono, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16 663–16 682, doi:10.1029/97JD00237.

Paulson, C. A., 1970: The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *J. Appl. Meteor.*, **9**, 857–861, doi:10.1175/1520-0450(1970)009<0857:TMROWS>2.0.CO;2.

Ryerson, W. R., and J. P. Hacker, 2014: The potential for mesoscale visibility predictions with a multimodel ensemble. *Wea. Forecasting*, **29**, 543–562, doi:10.1175/WAF-D-13-00067.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note, NCAR/TN-475+STR, 113 pp. [Available online at http://www2.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf.]

Slemmer, J., 2004: Study of dense fog at the Salt Lake City International Airport and its impacts to aviation. Western Region Tech. Attachment 04-01, 7 pp. [Available online at http://www.wrh.noaa.gov/media/wrh/online_publications/TAs/ta0401.pdf.]

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

——, and G. J. 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.

——, ——, and C. Snyder, 2006: Boundary conditions for limited-area ensemble Kalman filters. *Mon. Wea. Rev.*, **134**, 2490–2502, doi:10.1175/MWR3187.1.

Webb, E. K., 1970: Profile relationships: The log-linear range, and extension to strong stability. *Quart. J. Roy. Meteor. Soc.*, **96**, 67–90, doi:10.1002/qj.49709640708.

Wilson, T., and R. Fovell, 2015: Modeling the evolution and life cycle of cold pools. *Extended Abstracts, 15th Annual WRF Users' Workshop*, Boulder, CO, NCAR/MMM, 3.3. [Available online at http://www2.mmm.ucar.edu/wrf/users/workshops/WS2014/extended_abstracts/3.3.pdf.]

Wolyn, P. G., and T. B. McKee, 1989: Deep stable layers in the intermountain western United States. *Mon. Wea. Rev.*, **117**, 461–472, doi:10.1175/1520-0493(1989)117<0461:DSLITI>2.0.CO;2.

Zack, J., E. J. Natenberg, G. V. Knowe, J. Monobianco, K. Waight, D. Hanley, and C. Kamath, 2011: Use of data denial experiments to evaluate ESA forecast sensitivity patterns. Lawrence Livermore National Laboratory Rep. LLNL-TR-49916, 35 pp.