

Impact of Assimilating Preconvective Upsonde Observations on Short-Term Forecasts of Convection Observed during MPEX

MICHAEL C. CONIGLIO

NOAA/National Severe Storms Laboratory, Norman, Oklahoma

STACEY M. HITCHCOCK

Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

KENT H. KNOPFMEIER

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

(Manuscript received 7 March 2016, in final form 12 August 2016)

ABSTRACT

This study examines the impact of assimilating preconvective radiosonde observations obtained by mobile sounding systems on short-term forecasts of convection. Ensemble data assimilation is performed on a mesoscale (15 km) grid and the resulting analyses are downscaled to produce forecasts on a convection-permitting grid (3 km). The ensembles of forecasts are evaluated through their depiction of radar reflectivity compared to observed radar reflectivity. Examination of fractions skill scores over eight cases shows that, for four of the cases, assimilation of radiosonde observations nearby to subsequent convection has a positive impact on the initiation and early evolution during the first 3–4 h of the forecasts, even for the smallest resolvable scales of the 3-km grid. For the four cases in which positive impacts near the smallest resolvable scales of the grid are not seen, analysis of the changes to the preconvective environment suggests that sub-optimal locations of the soundings compared to the location of convective initiation are to blame. The aggregate positive impacts on forecasts of convection is more clearly seen when spatial scales larger than individual thunderstorms are examined.

1. Introduction

Convection-permitting numerical weather prediction (NWP) models are useful to forecasters tasked with alerting the public of the threat for severe weather (Kain et al. 2006; Clark et al. 2012). Characteristics of convective storms are strongly tied to the meso- and synoptic-scale environment in which they develop, so it is important to continue to explore ways to improve the depiction of the mesoscale environment in model initial conditions, even for short-term forecasts (Stensrud et al. 2009; Benjamin et al. 2010; Wandishin et al. 2010; Romine et al. 2013). Over multiple days in 2013 the

mesoscale environment preceding severe convective events was sampled by balloon-borne radiosonde (upsonde) observations released from multiple ground-based mobile facilities as part of the Mesoscale Predictability Experiment (MPEX) [see Trapp et al. (2016) and Weisman et al. (2015) for details]. This study addresses a goal of MPEX to explore the impacts of assimilating afternoon preconvective upsonde observations on the analysis of the mesoscale environment, as well as their impacts on subsequent short-term (0–9 h) convection-permitting forecasts.

The ensemble Kalman filter (EnKF; Evensen 1994) is a popular choice for the assimilation of observations on convection-permitting (1–4 km) model grids compared to traditional variational methods because it is much easier to implement effectively at these scales and because it provides flow-dependent relationships among model variables. Numerous studies have shown that

Corresponding author address: Michael C. Coniglio, National Severe Storms Laboratory, 120 David L. Boren Blvd., Norman, OK 73072.
E-mail: michael.coniglio@noaa.gov

EnKF methods can improve forecasts of convection by assimilating radar data (Snyder and Zhang 2003; Dowell et al. 2004; Aksoy et al. 2009; Dawson et al. 2012; Marquis et al. 2014), surface data (Fujita et al. 2007; Wheatley et al. 2012; Knopfmeier and Stensrud 2013), satellite data (Jones et al. 2013, 2015), and various combinations of these data (Zhang et al. 2006; Snook et al. 2011; Romine et al. 2013; Yussouf et al. 2015; Wheatley et al. 2015; Jones et al. 2016).

Few studies have focused on the specific impacts of radiosonde observations on NWP model forecasts of convection. For the purpose of reducing initial condition (IC) errors and subsequent short-term forecast errors of convective precipitation, Fabry and Sun (2010) and Fabry (2010) suggest that midlevel¹ humidity is a particularly important variable, but uncertainties in temperature, humidity, and winds in low- and midlevels can greatly impact forecasts of convection. Upsondes, of course, provide an accurate source of simultaneous in situ temperature, humidity, and wind measurements over the depth of the troposphere.

The NWS radiosonde network has a mean spacing of about 350 km and the radiosondes are released only twice per day shortly after 2300 and 1100 UTC. As a result, the conditions a few hours before convective initiation (CI) over the central United States [typically peaking from 1800 to 2100 UTC; Carbone et al. (2002); Carbone and Tuttle (2008); Surcel et al. (2010)] are usually not sampled by upsondes. The NWS sometimes releases upsondes during the afternoon, but this only occurs on a few convective days per year and the locations are limited to the standard sites. The use of multiple mobile upsonde systems like those used in MPEX offer much more flexibility in sampling times and locations, effectively creating an observing network capable of sampling subsynoptic-scale conditions.

This study focuses on the impacts of assimilating the MPEX upsonde observations into an ensemble of full-physics NWP model forecasts. The assimilation method used herein is described in detail in a companion paper (Hitchcock et al. 2016, hereafter HCK16) but key points are summarized here. HCK16 also provide a detailed look at the impacts of MPEX upsonde observations on forecasts of the 31 May 2013 convective event. They show a reduction in spurious convection resulting from lower midlevel humidity and improvements in CI resulting from enhanced convergence along a front. Presented here is an analysis of additional MPEX cases for

which the preconvective environment was sampled by the mobile upsonde systems. The goal is to determine if the positive impacts seen on convective forecasts for the 31 May event are seen for other events and to examine if the physical reasons for the differences in forecasts are similar to those found in HCK16.

Like in HCK16, the focus of this study is on the mesoscale (20–200 km) evolution of convective forecasts [i.e., on the accuracy of small to large groups of storms in the forecasts and not on individual thunderstorm cells (although large supercell thunderstorms can approach the lower end of the mesoscale spectrum)]. This is partly because the convection-permitting model grid length used here (3 km) cannot fully resolve individual thunderstorms (Bryan et al. 2003) and partly because no radar or satellite data are assimilated, so convection develops entirely from the mesoscale forcing in the analysis. Forecasted storms are not expected to have a one-to-one correspondence with observed storms in the former case (Stensrud et al. 2009). Furthermore, this study focuses on the impacts of the MPEX upsondes over regional areas no more than $\sim 200 \text{ km} \times 200 \text{ km}$, and on CI and early convective evolution. Tying the changes to the initial environment to the differences in skill between forecasts becomes much more difficult after 4–5 h in these cases because of complex convective development and evolution, as is typical for model errors in the vicinity of deep moist convection (Zhang et al. 2007). Therefore, although comparisons in skill are made for the full 9-h forecasts, the focus is placed on the first 4–5 h of the forecasts when attempting to link the physical reasons to the differences in skill.

A description of the MPEX upsonde data, ensemble data assimilation and modeling system, experimental design, and methods of evaluating the forecasts is provided in section 2 and is meant to be supplemented with the more detailed description provided in HCK16. Section 3 presents a comparison of the ensemble analyses and forecasts between the control ensemble and an ensemble that assimilates the MPEX upsonde data for individual cases and in an aggregate sense. Section 4 presents a summary and discusses some implications of the results and future research directions.

2. Data and methods

a. MPEX upsonde data

The mobile upsonde systems in MPEX operated in the central United States between 15 May and 15 June 2013. On convectively active days, up to four mobile systems spaced ~ 20 –120 km apart released

¹ Fabry (2010) defined midlevel as the lower half of the free troposphere, which starts at 1.5 km AGL in their framework.

TABLE 1. Description of the eight MPEX cases examined in this study. High precipitation (HP; Moller et al. 1994) and mesoscale convective system (MCS; Parker and Johnson 2000). The forecasts are initialized approximately 1 h prior to observed CI in the area of interest.

Date	Region	Pre-CI sonde release times (UTC)	Forecast initial time (UTC)	Primary storm modes
18 May	Western KS	1704, 1711, 1802, 1815, 1856	1900	Tornadic HP supercells
19 May	Central OK	1858, 1900, 1857	1900	Tornadic supercells
20 May	Central OK	1714, 1714, 1718	1800	Tornadic and nontornadic supercells
23 May	Western TX	1755, 1759, 1801, 1805	1800	Weakly tornadic HP supercells
27 May	Central KS	2000, 2003, 2013, 2059, 2102	2100	Tornadic and nontornadic supercells
28 May	Central KS	2000, 2000, 2000, 2004	2000	Nontornadic supercells and disorganized MCS
31 May	Central OK	1610, 1801, 1920	2000	Tornadic supercells
8 Jun	Southwestern KS	1755, 1800, 1805	1800	Nontornadic supercells and small, bowing MCS

radiosondes a few hours prior to CI.² The locations for afternoon preconvective upsondes were chosen partly to provide validation observations for experimental NWP model forecasts in a separate study (see Weisman et al. 2015). However, to support a goal of MPEX to measure the upscale feedbacks of convection to the environment (see Trapp et al. 2016), the moist unstable air mass that supported the afternoon and evening storms, as well as either side of initiating boundaries in the environment, were sampled whenever possible. Although the locations of these afternoon upsondes were not guided by formal targeting methods to optimize impacts from data assimilation (e.g., Torn and Hakim 2008; Bednarczyk and Ancell 2015), they were obtained close enough to subsequent convection in space and time to expect some impacts from their assimilation on analyses and short-term forecasts of the convection.

The mobile upsonde systems operated on 17 days during MPEX [see Table 2 of Trapp et al. (2016)]. The cases analyzed in this study are those in which at least three upsondes were released from three separate vehicles at locations separated by at least 20 km, and were released at least 30 min prior to CI in the region of interest (roughly within 200 km of the soundings). There are 11 such cases that were the subject of the data assimilation experiments described in the next section. Three of these 11 cases (4, 11, and 12 June) were “correct nulls” meaning that all experiments correctly forecast the absence of strong convection in the region of interest and the soundings had minimal impact on forecasted convection more than ~200 km from the sounding locations. The remaining eight cases in which the mesoscale environment was sampled by at least three MPEX vehicles prior to nearby CI are the subject of this study (Table 1 and Fig. 1).

b. Data assimilation and NWP modeling system

WRF-ARW, version 3.4.1, configured with the Data Assimilation Research Testbed (DART) Lanai version (Anderson and Collins 2007; Anderson et al. 2009) is used for data assimilation and ensemble forecasting. The domain used for data assimilation covers the contiguous United States (Fig. 2) with 15-km horizontal grid spacing and 51 vertical levels topped at 50 hPa. The ICs for a 36-member ensemble are created by downscaling the 0000 UTC Global Ensemble Forecast System

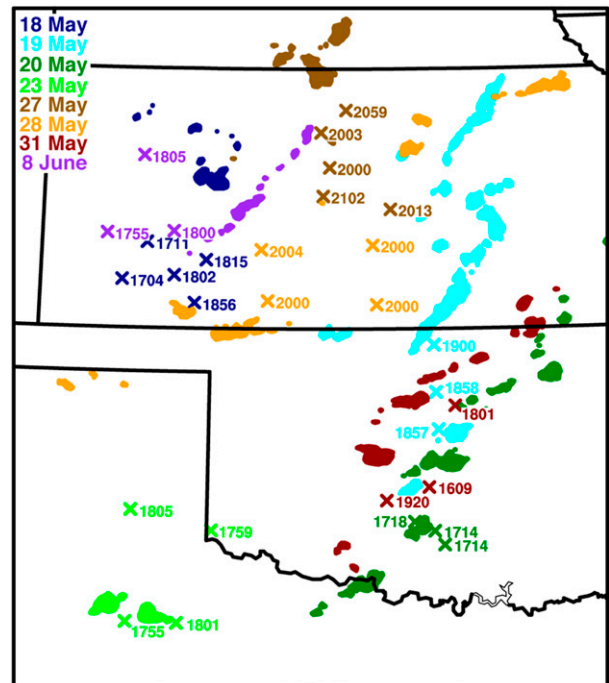


FIG. 1. Locations (×) and times (UTC) of pre-CI radiosonde releases that are assimilated for the eight cases examined in this study (color coded by day). Filled contours show gridded composite reflectivity ≥ 50 dBZ near the radiosonde locations approximately 3 h after the final radiosonde release for each day.

² Herein, composite reflectivity of at least 40 dBZ that persists for at least 30 min is a CI event.

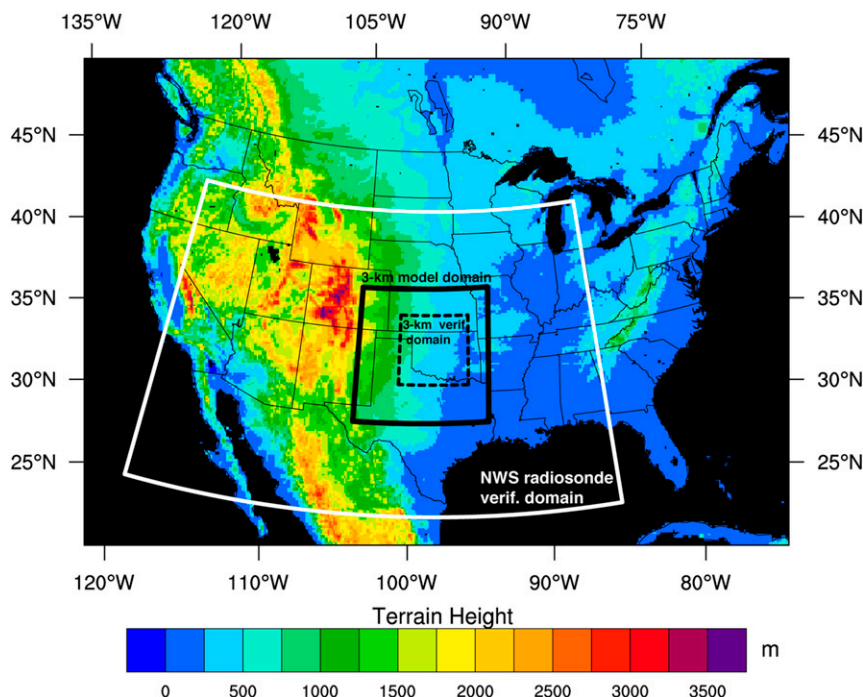


FIG. 2. The 15-km WRF-DART domain with terrain height (m) contoured used for all experiments and an example of the 3-km domain used for the 31 May 2013 case (enclosed by the solid black lines). The domain used for verification of 3-km model fields is enclosed by the black dashed lines. The area over which statistics for the ensemble analyses are compared to 1200 UTC NWS radiosondes is enclosed by the white lines. The sizes of the 3-km domain (331×331 grid points) and the verification domain (170×170 grid points) are the same in every experiment but are moved to cover the event in question.

(GEFS) 50-km analyses on the day of interest, and forecasts from this GEFS cycle are used as lateral boundary conditions (LBCs) for the 15-km grid. The first 18 members of the GEFS are used to create ICs for two sets of 18 ensemble members. Diversity is created among the full set of 36 members by altering the turbulence, radiation, and cumulus parameterization schemes for the second set of 18 members. This configuration is the same as that used in HCK16.

As in HCK16, observations are obtained from the Meteorological Assimilation Data Ingest System (MADIS; Miller et al. 2007) that include 1) mandatory and significant levels from the NWS radiosondes; 2) surface data from aviation routine weather reports (METARs), marine (ship and buoy) reports, the Oklahoma Mesonet, and the mesonet observations from a variety of networks available in the MADIS data; 3) Aircraft Meteorological Data Relay (AMDAR) reports for wind and temperature; and 4) atmospheric motion vectors (AMVs) derived from satellite observations. These observations are assimilated every hour from 0100 UTC up to ~ 5 h prior to observed CI using the ensemble adjustment Kalman filter (EAKF; Anderson

2001) encoded within the DART software. Between ~ 5 and 1 h prior to CI, observations are assimilated every half-hour. Adaptive inflation (Anderson 2007) is applied to the ensemble of forecasts prior to the assimilation step to help maintain spread. All data are assimilated onto the full 15-km grid except for the Oklahoma and MADIS mesonet observations, which are assimilated onto the 15-km grid only over the area covered by the 3-km grid (Fig. 2).

The 36-member ensemble of analyses valid at the final analysis time are downsampled to create ICs for an ensemble of forecasts on a convection-permitting grid ($\Delta = 3$ km; Fig. 2). Forecasts run on the 15-km grid serve as the LBCs for the 3-km grid. The initialization time of the ensemble of forecasts on the 3-km grid is allowed to vary among the cases (Table 1) and is chosen to allow for at least 30 min of integration time before CI occurs anywhere in the ~ 200 km by 200 km region of interest, and to allow for at least 1 h of integration before the CI that occurs closer to the location of the soundings. A set of 36 forecasts are then run for 9 h on the 3-km grid that use the same physics options as those for the parent grid, except no cumulus parameterization is used. The

resulting analyses and forecasts compose the control ensemble for each event. This approach makes the control ensemble representative of the best possible forecast that could have been made given all of the observations available in the operational data stream at the time.

The set of analyses and forecasts that assimilate the MPEX upsonde data using the procedures outlined next are referred to as the MPEX ensemble. Because the MPEX upsonde data contain observations every 1–2 s, the data are thinned by defining “significant” levels at which a substantive change in temperature, dewpoint, or wind occurs, similar to how the NWS radiosonde data are routinely thinned (e.g., see Fig. 6 of HCK16). This helps retain potentially meaningful meteorological features in the assimilation while mitigating the detrimental effects of correlated observation errors (Liu and Rabier 2002).

The effects of the assimilation on model variables are localized in the standard way by multiplying the error covariance estimate in the EAKF with a Gaussian-like weighted correlation function (Gaspari and Cohn 1999) that decreases to zero at a specified radius and is 20% of its original weight at half this radius (Hamill 2006). The horizontal and vertical radii used for the NWS soundings and for the MPEX upsondes are approximately 460 km and 8 km, respectively, the same as those used for a similar application in Wheatley et al. (2014) and in the MPEX ensembles detailed in HCK16.

Consideration of observation errors is required in the use of the EAKF. The specified observation errors for all observation types follow those used in Romine et al. (2013). For radiosondes, the specified standard deviation in temperature errors is 1.25 K at the surface and decreases to 0.75 K at 875 hPa. The temperature errors remain at 0.75 K up to 350 hPa before increasing again to 1.25 K at 250 hPa. The specified standard deviation in errors for winds is likewise a function of height, increasing from about 1.5 m s^{-1} at the surface to about 3 m s^{-1} near the tropopause. The humidity errors are specified through the dewpoint and are a function of RH (larger dewpoint errors are assigned to measurements with smaller RH) following the suggestions of Lin and Hubbard (2004).

Finally, the MPEX upsonde observations are binned into half-hour windows to accommodate the half-hour assimilation of data in the 4-h period prior to CI, and the sonde data are assimilated at its true position, which effectively takes balloon drift into account. Data from a single sonde profile may be split into two, occasionally three, assimilation time windows since more than an hour can elapse between the sonde release time and the time the sonde reaches the tropopause.

3. Results

a. Mesoscale ensemble evaluation

To first provide an overview of the assimilation of standard observations on the mesoscale grid prior to the assimilation of the MPEX upsonde data, time series of the ensemble mean bias (model – observation), root-mean-square error (RMSE) of the ensemble mean, ensemble spread, and consistency ratio (CR) are computed using assimilated 1200 UTC NWS radiosonde observations (Fig. 3) and are given by

$$\text{bias} = \frac{1}{N_o} \sum_{n=1}^{N_o} \overline{H(\mathbf{x}_n^f)} - y_n, \quad (1)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{n=1}^{N_e} [H(\mathbf{x}_n^f) - y_n]^2}{N_o}}, \quad \text{and} \quad (2)$$

$$\text{total spread} = \sqrt{\sigma_o^2 + \left\{ \frac{1}{(N_e - 1)} \sum_{n=1}^{N_e} [H(\mathbf{x}_n^f) - \overline{H(\mathbf{x}^f)}]^2 \right\}}, \quad (3)$$

where H is the forward operator that maps the model prior analysis (1-h forecast in this example) \mathbf{x}_n^f to the observation location and type y_n ; N_o is the number of observations; N_e is the number of ensemble members; and the overbar denotes the ensemble mean. The first term in Eq. (3) represents the specified observation error variance. The CR is the ratio of the ensemble prior variance [square of Eq. (3)] to the ensemble mean squared error [square of Eq. (2)]. A CR value close to 1 is indicative of optimal ensemble spread (Dowell et al. 2004), with values lower (higher) than 1 indicating an under- (over-) dispersive ensemble.

Except for near the ground, the ensemble mean is slightly too warm over most of the troposphere (Fig. 3a), similar to that seen in Romine et al. (2013). The moist bias in the prior mean (Fig. 3b) is typical for NWP model forecasts over the United States (Weisman et al. 2008; Fabry 2010; Coniglio 2012), but is smaller than the mid-to upper-level humidity bias seen in a similar application in Romine et al. (2013). The bias in winds is very small at all levels (Figs. 3c,d).

The CR for all variables at 1200 UTC ranges generally from 0.5 to 1 and is smallest for temperature from just above the ground to the lower troposphere. This indicates an ensemble that is somewhat underdispersive overall for variables above the surface at 1200 UTC. In these types of applications in which observations are assimilated over multiple cycles, the CR typically starts small and grows toward 1 as spread develops in the

12z NWS radiosondes

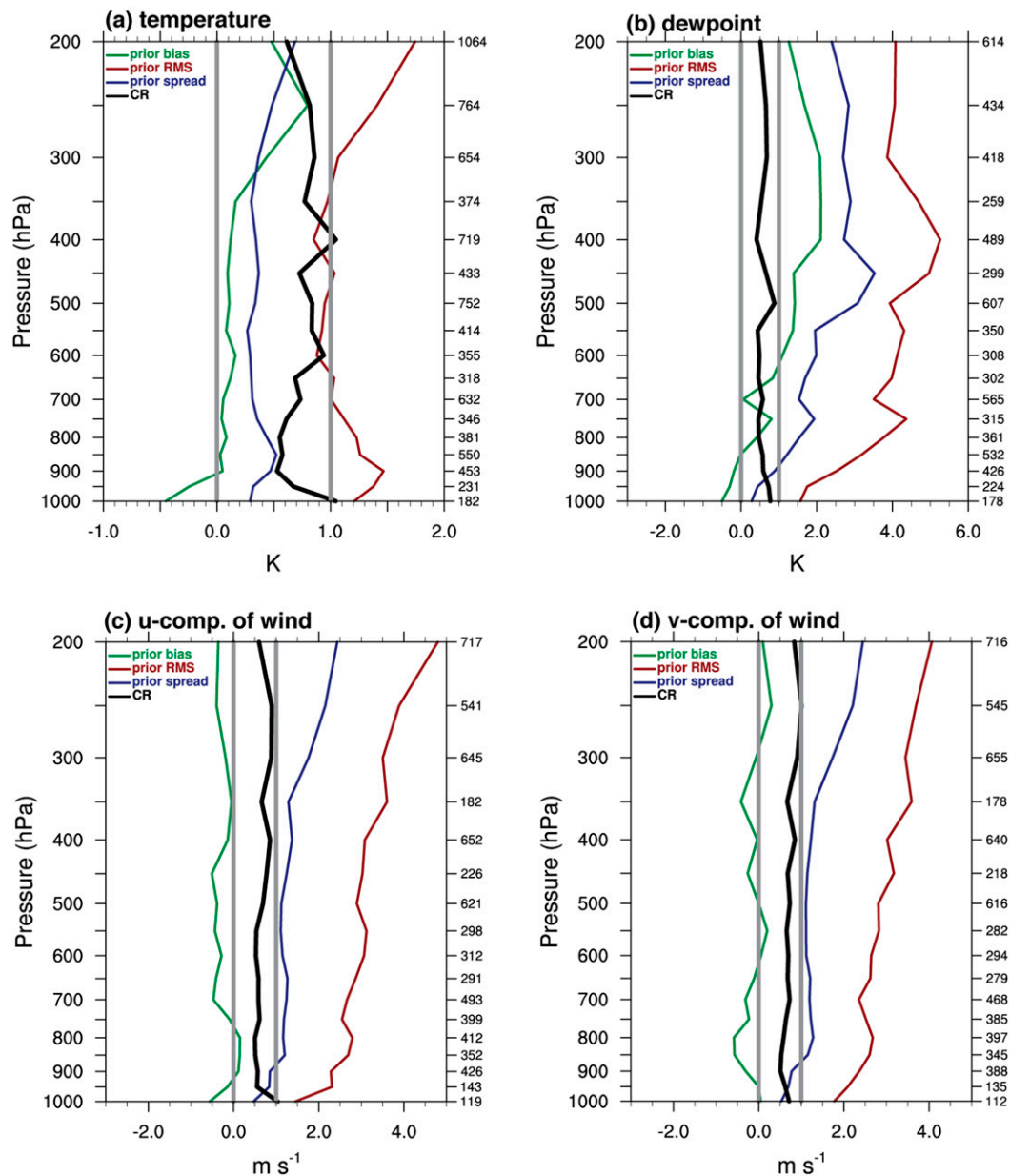


FIG. 3. Profiles of statistics for prior analyses (1-h forecasts) of (a) temperature (K), (b) dewpoint (K), (c) u component of wind (m s^{-1}), and (d) v component of wind (m s^{-1}) compared to assimilated 1200 UTC NWS radiosonde observations averaged over the area shown in Fig. 2 and binned to 50-hPa intervals. The sample size in each bin is shown on the right side of each panel. The model bias (mean error) is in green, the RMSE is in red, ensemble spread (not including the contribution of total spread from the assumed observation error standard deviation) is in blue, and the consistency ratio (CR) is in black. Reference lines for model bias and the CR are given by the gray lines at $x = 0$ and $x = 1$, respectively.

ensemble with each assimilation/forecast cycle. Profiles of the CR evaluated against the MPEX upsonde data, which were obtained in the 1800–2100 UTC period (1300–1600 local time) (Table 1), indeed show CR values larger than those found for the NWS soundings

that were taken in the morning (Fig. 4). The CR for afternoon lower-tropospheric temperature is generally between 0.7 and 1 from the surface to 750 hPa. Likewise, the underdispersion in the mid- to upper-tropospheric temperatures seen at 1200 UTC (Fig. 3a) becomes an

MPEX radiosondes

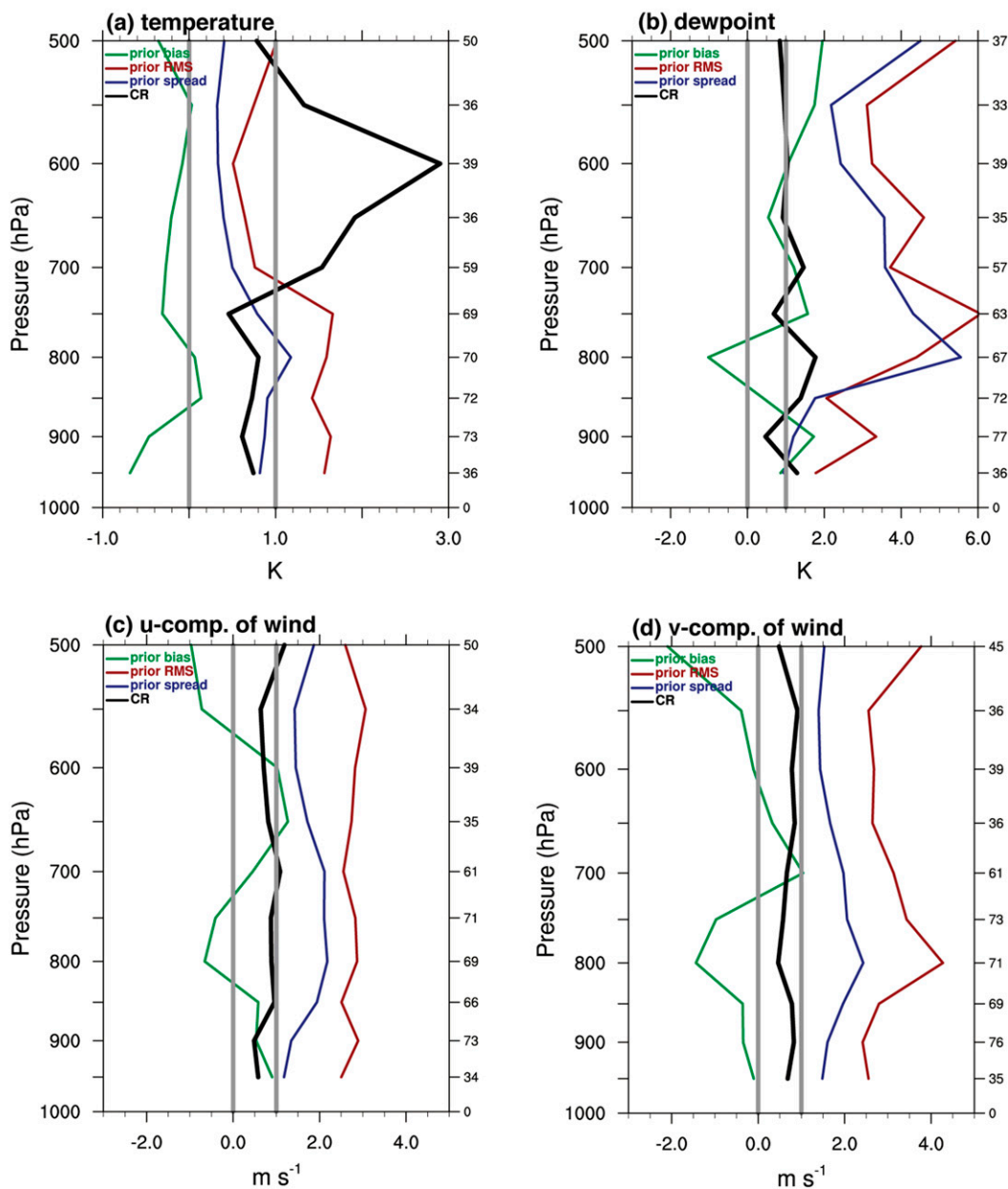


FIG. 4. As in Fig. 3, but compared to the 30 assimilated MPEX upsonde profiles. Note that profiles are only shown up to 500 hPa as bin sample sizes are less than 30 for all variables above this level.

overdispersion in the 700–550-hPa layer related to very small RMSE (Fig. 4a). Overall, most of the CR values range from 0.7 to 1.3, which is characteristic of ensemble systems for similar applications and various observation types (Dowell et al. 2004; Wheatley et al. 2012, 2014, 2015). This result gives confidence that the ensemble and assimilation design provide a sufficiently accurate background to later drive the convection-permitting forecasts.

b. Case-by-case skill of convection-permitting forecasts

Differences between the MPEX and control ensembles on the 3-km grid are assessed through their forecasts of simulated composite (or column maximum) reflectivity (referred to as reflectivity hereafter) compared to gridded reflectivity observations. The gridded reflectivity observations are created by interpolating the

0.01° latitude by 0.01° longitude analyses of reflectivity produced by the NSSL National Mosaic and Multi-Sensor Quantitative (NMQ) Precipitation Estimation system (Zhang et al. 2011) to the 3-km model grid. Following Schwartz et al. (2010), neighborhood probabilities are produced for each member, and are then averaged among all members of the ensemble to compute the neighborhood ensemble probabilities [NEPs; see Eqs. (2)–(5) in Schwartz et al. (2010)]. Neighborhoods are defined around the grid cells to give credit to forecasts of storms that may not overlap with but are close to observed storms, unlike traditional verification metrics [e.g., critical success index or equitable threat score; Schaefer (1990)]. In this study a square neighborhood³ is constructed around each 3-km grid cell and numerous neighborhood sizes and reflectivity thresholds are tested.

The NEPs are computed in order to calculate the fractions skill score (FSS; Roberts and Lean 2008) for each ensemble. FSS values range from 0 (no skill) to 1 (perfect), and values of ~ 0.5 and greater have been considered to represent forecasts with “useful” skill (Roberts and Lean 2008), but the absolute values of FSS are less important here than how the FSS compares between the ensembles since the goal is to determine how skill from the MPEX ensemble differs from the control ensemble. The domain over which FSS is computed is an inner subset of the 3-km grid (Fig. 2). Confidence intervals (95%) for the differences in FSS between the MPEX and control ensemble forecasts are computed following the bootstrap resampling technique described in Hamill (1999). In this application, the ensemble members serve as the independent samples.

FSS differences between the MPEX and control ensemble forecasts are first illustrated for a reflectivity threshold of 40 dBZ for a box width of 8Δ (~ 25 km \times 25 km). This box size reduces the influence from errors with spatial scales close to and smaller than the smallest resolvable scales of the grid ($\leq \sim 4 - 6\Delta$; Skamarock 2004), which allows forecast details on scales immediately larger than individual thunderstorms to be retained—a difficult test for convection-permitting models. The impacts from the MPEX upsonde data, and their effects on the FSS, are best understood by examining each case prior to examining the aggregate FSS over all the cases.

It should be noted that biased forecasts can sometimes inflate FSS for smaller spatial scales (Mittermaier and

Roberts 2010). Overall there is a high bias in simulated reflectivity at the 40-dBZ threshold compared to the NMQ gridded composite reflectivity. This bias arises from a tendency for too many forecasted storms (as in Burghardt et al. 2014) that tend to be slightly too large (Bryan and Morrison 2012), but the bias may also be related to the implementation of the Thompson microphysics scheme in WRF-ARW (Stratman et al. 2013). However, as can be seen next in the figures displaying differences in FSS, the differences in bias between the MPEX and control ensemble forecasts are small and generally affect each forecast set equitably, and, therefore, do not significantly impact the differences in FSS.

1) POSITIVE IMPACT CASES

HCK16 concentrates on FSS differences for the 31 May convective event. For that case, forecast improvements in the first 90 min (Fig. 5c) relate to reduced midlevel humidity that leads to far fewer spurious storms than in the control ensemble. In the following 90 min, the forecast improvements in the MPEX ensemble relate to more convection closer to where they were observed along a front that initiates abundant convection. The assimilation of the MPEX upsonde data increases convergence along the front leading to better placement of the storms. Positive improvements in FSS in at least the first 3 h of the forecasts are also seen for the 18 May (Fig. 5a), 23 May (Fig. 5b), and 8 June (Fig. 5d) experiments as shown next.

On 18 May, CI occurred at ~ 2030 UTC along the portion of a dryline that was sampled by the MPEX soundings a few hours earlier (Fig. 6a). The assimilation of the MPEX upsonde data adjusted the dryline northeastward in a bulge over southwestern Kansas and increased the water vapor mixing ratio near the top of the boundary layer to the northwest of the dryline bulge (Fig. 6b). The increased low-level convergence (not shown, but implied by the difference vectors in Fig. 6b) and increased low-level humidity there translated northward with time. As a result, more ensemble members produce CI closer to the time and location of observed CI (Fig. 6b), eventually leading to more storms forecasted to be closer to the small cluster of developing supercells observed at 2230 UTC (210-min forecasts; Fig. 6a). Later in the forecasts, the FSS for the MPEX ensemble stays significantly larger than that for the control ensemble (Fig. 5a) because many more members in the MPEX ensemble capture the upscale growth of the initial cluster of supercells into a small linear MCS (although both ensembles do not capture the development of a separate MCS farther south over southern Kansas) (Fig. 7).

Improvements in timing and location of CI are also seen for the 23 May experiment (Fig. 8). On 23 May,

³ Tests comparing the use of a square neighborhood versus a more strict distance criteria like that used in Hitchcock et al. (2016) showed negligible differences in the FSS values, as in Ebert (2009).

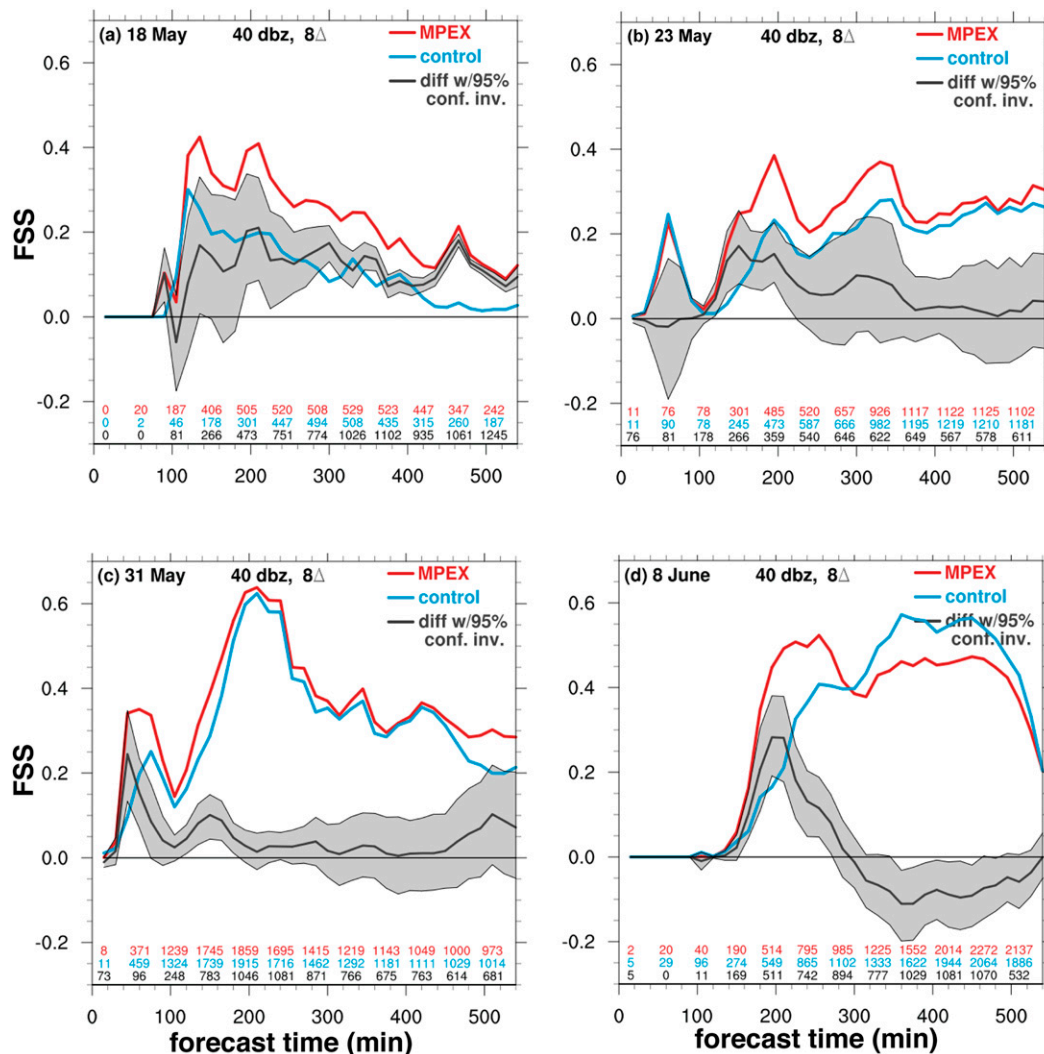


FIG. 5. FSS for the MPEX (red) and control (blue) forecasts, and their differences (MPEX – control; black line), for a reflectivity threshold of 40 dBZ and a 8 by 8 gridcell neighborhood (~25 km × 25 km) for the positive impact cases on (a) 18 May, (b) 23 May, (c) 31 May, and (d) 8 Jun experiments. The gray shading spans the 95% confidence interval on the FSS differences. The numbers along the bottom of each figure correspond to the number of grid cells with reflectivity ≥40 dBZ for the observed reflectivity (black text), the mean number of grid cells with reflectivity ≥40 dBZ for the MPEX ensemble forecasts (red text), and the mean number of grid cells with reflectivity ≥40 dBZ for the control ensemble forecasts (blue text).

about an hour prior to nearby CI, the MPEX upsondes sampled both sides of an outflow boundary (OFB) that was moving to the southwest across northwestern Texas (Fig. 8a). The OFB (and the associated maximum in low-level convergence) shifts to the southwest close to where it was observed as a result of assimilating the MPEX upsonde data (Fig. 8b). The MPEX ensemble mean more accurately represents the cool outflow air temperature in a comparison to 2-m temperature observations (Fig. 8b). Furthermore, the MPEX ensemble mean is warmer in the boundary layer (and more accurate at 2 m AGL) in the region between the dryline and

OFB (Fig. 8b). In this same region the MPEX ensemble mean water vapor mixing ratio is 1–2 g kg⁻¹ larger near the top of the boundary layer (Fig. 8b). These factors lead to more MPEX ensemble members producing CI closer to the time and location of observed CI (the location of observed CI is denoted on Fig. 8b), leading to more storms in the MPEX forecasts closer to the observed storms at 2100 UTC than in the control ensemble (Fig. 8a).

For the 8 June experiment, despite CI occurring over 150 min after the environment was sampled by the MPEX upsondes, the assimilation of the MPEX upsonde data produces many more storms closer to

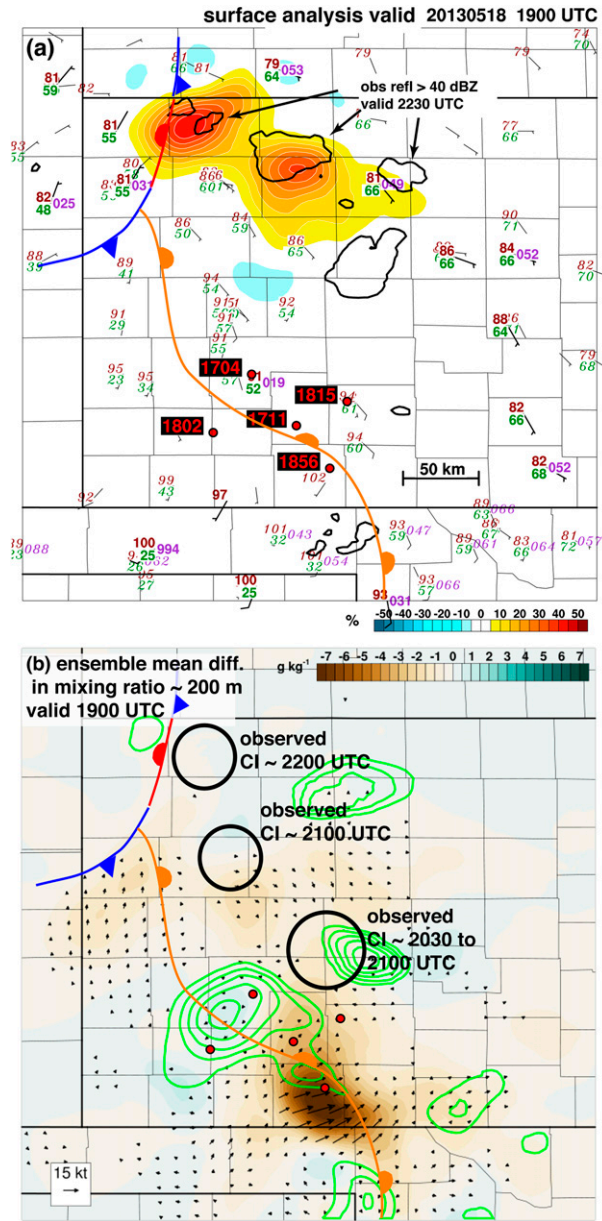


FIG. 6. (a) Surface observations and manually drawn front and dry-line valid 1900 UTC 18 May 2013 with temperature (red; °F), dewpoint (green; °F), and sea level pressure (purple; hPa and leading two digits removed) from NWS observations (bold text) and MADIS and Oklahoma Mesonet observations (italics). Black contours show NMQ gridded (observed) reflectivity ≥ 40 dBZ valid at 2230 UTC and color shading depicts the difference (MPEX - control) in neighborhood probabilities (%) of simulated reflectivity ≥ 40 dBZ using a $\sim 25 \text{ km} \times 25 \text{ km}$ neighborhood valid at 2230 UTC (210-min forecasts). Locations and times (UTC) of the MPEX radiosonde release times are also shown (red dots with accompanying red text). (b) Ensemble mean difference (MPEX - control) in water vapor mixing ratio (g kg^{-1}) and winds (kt; $1 \text{ kt} = 0.5144 \text{ m s}^{-1}$) at ~ 200 m AGL valid 1900 UTC (color shading) and ensemble mean difference in water vapor mixing ratio at ~ 800 m AGL (green contours every 0.5 g kg^{-1} starting at 1 g kg^{-1}). The approximate locations of observed CI in Kansas prior to 2200 UTC are circled. Red dots in both panels show the locations of the MPEX radiosonde releases.

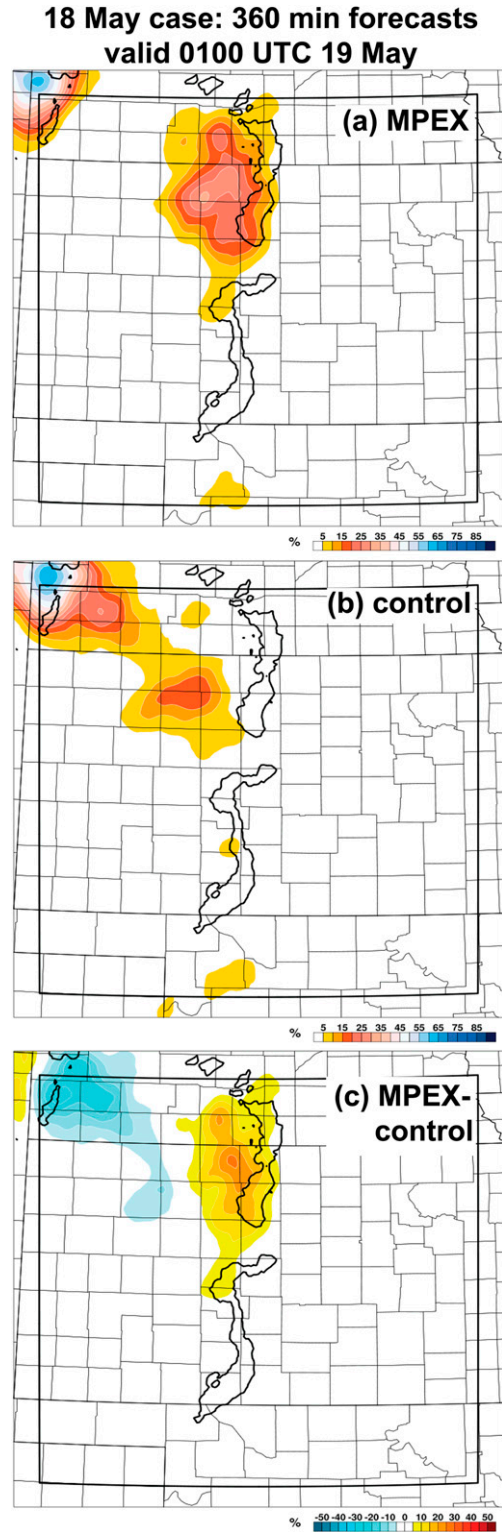


FIG. 7. Neighborhood probabilities of simulated reflectivity ≥ 40 dBZ using a $\sim 25 \text{ km} \times 25 \text{ km}$ neighborhood valid at 0100 UTC 19 May (360-min forecasts) for the (a) MPEX and (b) control ensemble, and (c) their difference (MPEX - control). Black contours show NMQ gridded (observed) reflectivity ≥ 40 dBZ. The black box encloses the area over which FSS is computed for the 18 May case.

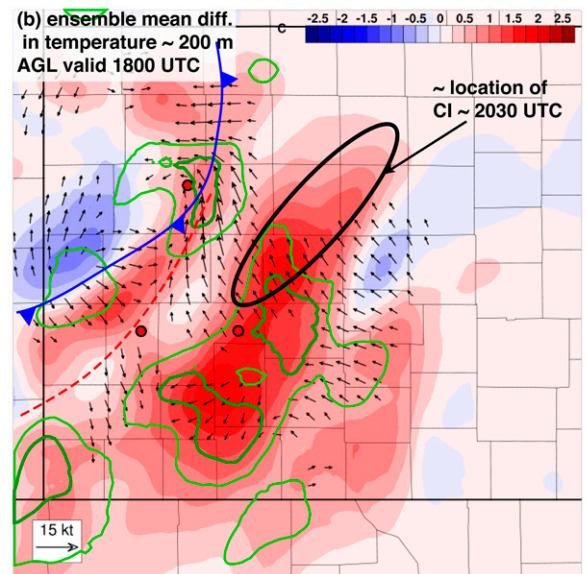
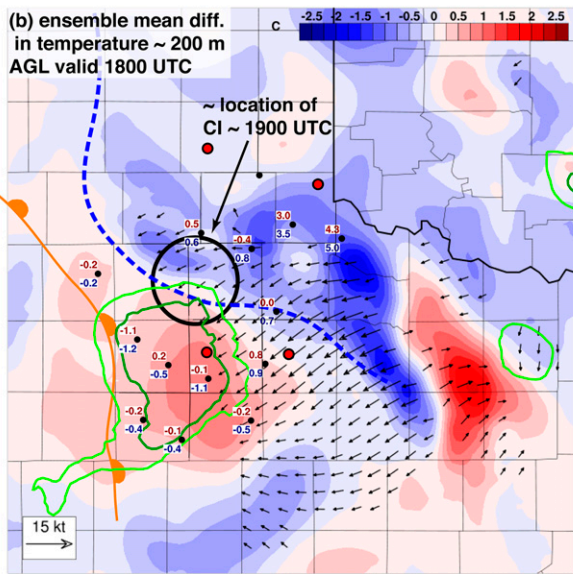
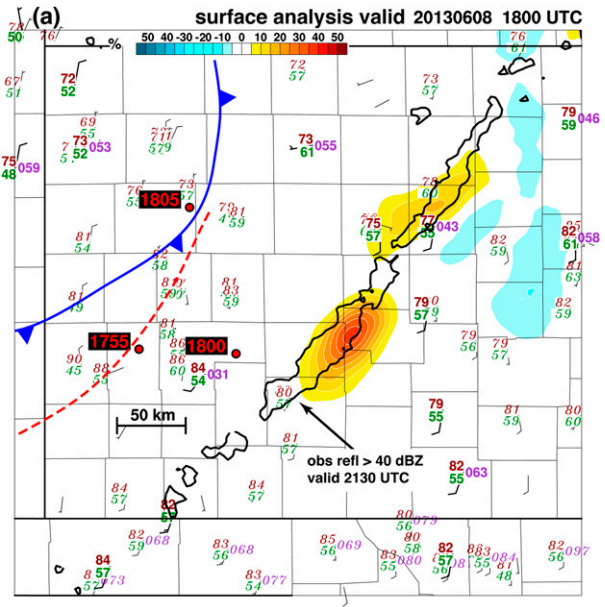
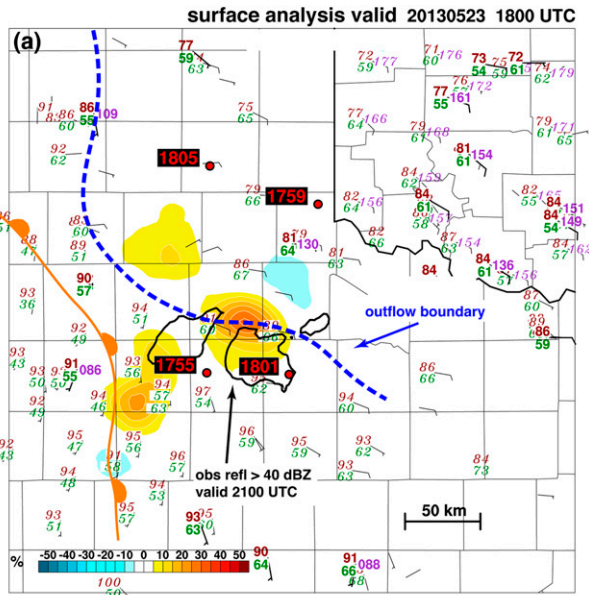


FIG. 8. As in Fig. 6, but for surface observations and analysis valid 1800 UTC 23 May 2013 and observed reflectivity and forecasts valid 2100 UTC (180-min forecasts) and ensemble mean difference (MPEX – control) in temperature ($^{\circ}\text{C}$) and winds (kt) at ~ 200 m AGL valid 1800 UTC (initialization time of the 3-km forecasts). The blue dashed line depicts the approximate position of an outflow boundary that emanated from a morning MCS over Oklahoma. Light (dark) green lines in (b) depict $+1$ (1.5) g kg^{-1} differences in ensemble mean water vapor mixing ratio near the top of the boundary layer (model level 10). Numbers in red (blue) adjacent to black dots depict the difference in MPEX (control) ensemble posterior analysis 2-m temperature ($^{\circ}\text{C}$) from the observation at that location.

FIG. 9. (a) As in Fig. 6a, but for surface observations and analysis valid at 1800 UTC 8 Jun 2013, observed reflectivity valid at 2130 UTC, and difference in neighborhood probabilities valid at 2130 UTC (210-min forecasts). (b) As in Fig. 6b, but for ensemble mean difference (MPEX – control) in temperature ($^{\circ}\text{C}$) and winds (kt) at ~ 200 m AGL valid 1800 UTC (initialization time of the 3-km forecasts). The red dashed line depicts a prefrontal wind shift line. Light (dark) green lines in (b) depict $+1$ (1.5) g kg^{-1} differences in ensemble mean water vapor mixing ratio near the top of the boundary layer (model level 8).

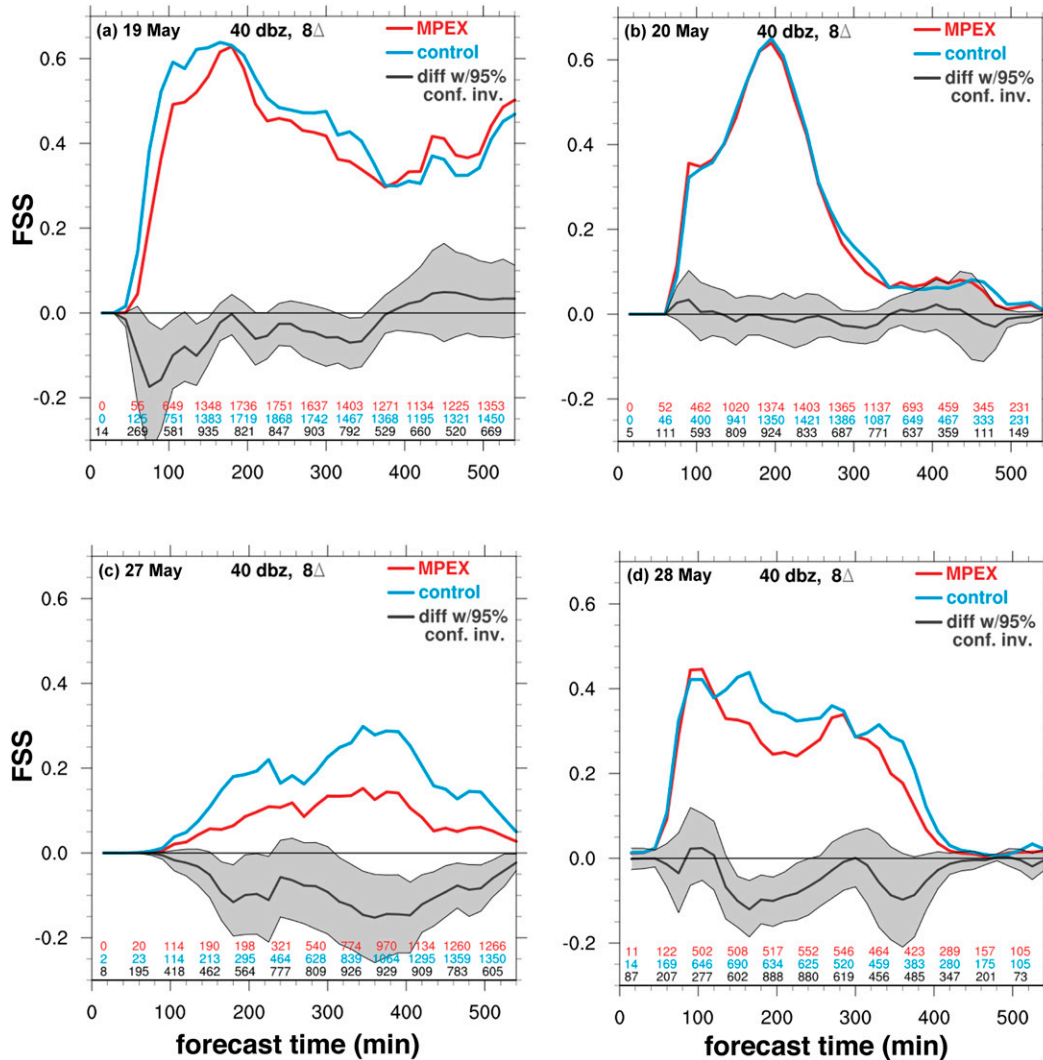


FIG. 10. As in Fig. 5, but for the neutral or negative impact cases on (a) 19, (b) 20, (c) 27, and (d) 28 May.

where they were observed (Fig. 9a) leading to significantly higher FSSs for the 150–240-min forecasts (Fig. 5d). The convective mode was linear in response to the strong frontal forcing, with two main convective lines developing by 2130 UTC that joined shortly thereafter. The MPEX ensemble depicts this development better than the control ensemble as seen with two areas of positive differences in neighborhood probability in the vicinity of the two main convective lines (Fig. 9a). In this case the improvements result from an improved analysis of low-level temperature and winds in the vicinity of a front–prefrontal trough (Fig. 9b). The humidity in the upper part of the boundary layer is 1–2 g kg⁻¹ higher in the MPEX ensemble mean near the intersection of the prefrontal trough and cold front and in the preconvective inflow environment over southwestern Kansas (Fig. 9b), and this increased humidity persists in the subsequent forecasts up

to the time of CI. Overall, the improved forecasts of CI result from both enhanced convergence in the area of the prefrontal trough–front intersection, and the warming and moistening of the boundary layer that increases the instability in the inflow air to the south and east of the front.

2) EVALUATION OF NEUTRAL OR NEGATIVE IMPACT CASES

While the four experiments discussed above (18 May, 23 May, 31 May, and 8 June) show some positive impacts from assimilation of the MPEX upsonde data, four of the experiments (19, 20, 27, and 28 May) show either little impacts or a degradation in FSS in the first 3–4 h of the forecasts (Fig. 10).

For the 19 May case three soundings spaced 50–60 km apart were taken in a north–south line ahead of a dryline and approaching cold front to the west (Fig. 11). Overall

differences in ensemble mean analysis temperature, water vapor mixing ratio, and winds between the MPEX and control ensembles are smaller than those seen for the positive impact cases. The ensemble mean analysis differences at 1900 UTC near 800 hPa shown in Figs. 12a and 12b represent the largest ensemble mean differences seen for this case at any level. The MPEX ensemble mean is cooler and more humid near the northernmost sounding (Figs. 12a and 12b), but the more substantial cool/moist adjustments are limited to mostly within ~50 km north of the sounding location. The MPEX ensemble mean analyses are warmer and drier than the control ensemble mean analyses near the location of the southern two soundings (Figs. 12a and 12b).

More importantly for the impacts on the forecasts for the 19 May experiments, the magnitude of the temperature and mixing ratio adjustments decrease substantially with time and most of the perturbations move out of the region where CI occurs over the subsequent few hours (Figs. 12c and 12d). The lack of improvements in the forecasts for this case could, therefore, be related to a suboptimal sampling strategy relative to where CI occurred. The MPEX soundings were obtained at a longitude about 75–150 km east of where CI occurred about an hour later (Figs. 12c and 12d). The 19 May MPEX ensemble forecasts approach a statistically significant degradation in forecast skill compared to the control ensemble (Fig. 10a) because CI is incorrectly delayed about 30 min in the MPEX ensemble forecasts in the two northern and westernmost areas of CI (Figs. 12c and 12d), which leads to a phase error in the cluster of storms a few hours later. Negative differences in neighborhood probability up to 30% are seen in the vicinity of storms that were undergoing upscale growth into a small convective system at the time (Fig. 11b). It is not clear why the small differences in temperature and mixing ratio in this region of CI lead to delayed initiation in the MPEX ensemble, but the delayed CI may be related to the slight warming and drying just above the boundary layer in the region of observed CI at 1945 UTC noted in Figs. 12c and 12d.

For the 20 May case there were three MPEX soundings taken relatively close to each other (about 20–25 km apart; Fig. 1) and all of them sampled the very unstable air mass to the east of a dryline and south of a stationary front, similar to the 19 May case [see Zhang et al. (2015) for details of the 20 May event]. Both ensemble mean temperature and dewpoint in about the lowest 2 km AGL are very similar between the MPEX and control ensembles because the errors in the control ensemble mean are already small (Fig. 13). While the assimilation of MPEX upsonde data produces 30-min forecasts that more accurately depict the dry midtropospheric

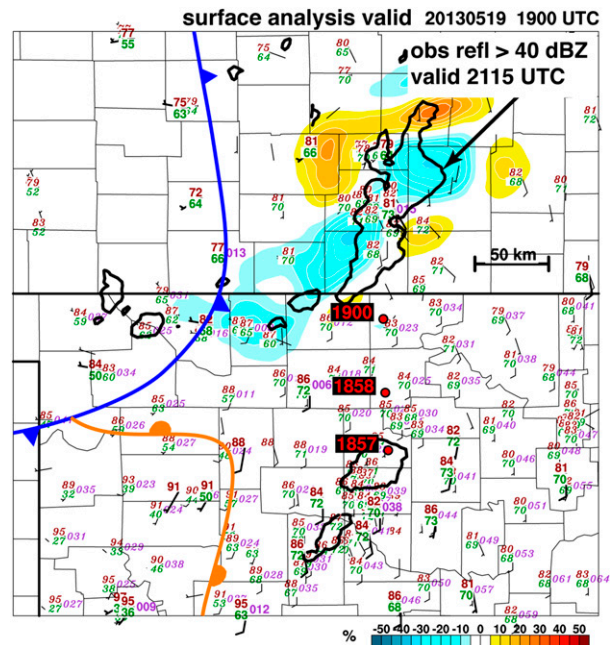


FIG. 11. As in Fig. 6a, but for surface observations and analysis valid at 1900 UTC 19 May 2013, observed reflectivity valid at 2115 UTC, and difference in neighborhood probabilities valid at 2115 UTC (135-min forecasts).

(~525–400 hPa) air, the ~800–700-hPa layer is too dry (Fig. 13). The assimilation of the humidity data below 700 hPa occurs in the 30-min window centered at 1730 UTC, which is 1 h before the comparison shown in Fig. 13. While the MPEX ensemble posterior analyses fit the entire humidity profile better at 1730 UTC as expected (not shown), the adjustments are quickly swept out of the area an hour later (by 1830 UTC; Fig. 13). Likewise, the MPEX ensemble analyses fit the wind profile better at 1730 UTC, but the differences in ensemble mean winds an hour later are small over most of the troposphere because the localized adjustments made at 1730 UTC move out of the area (hodograph inset to Fig. 13). The resulting small differences have little impact on the subsequent forecasts of convection (Fig. 10b). Again, the lack of FSS improvements for the MPEX ensemble in the 19 May and 20 May cases could, therefore, be related to a suboptimal sampling strategy of the soundings, but a forecast that already has small errors in temperature and humidity in the lowest few kilometers likely contributes to the lack of improvements in the 20 May case.

For the 27 May case five MPEX soundings sampled a moist, unstable air mass to the east of a diffuse dryline and to the south of a weak stationary front (Fig. 14). The degraded FSS in the first few hours of the MPEX ensemble forecasts (Fig. 10c) results from MPEX

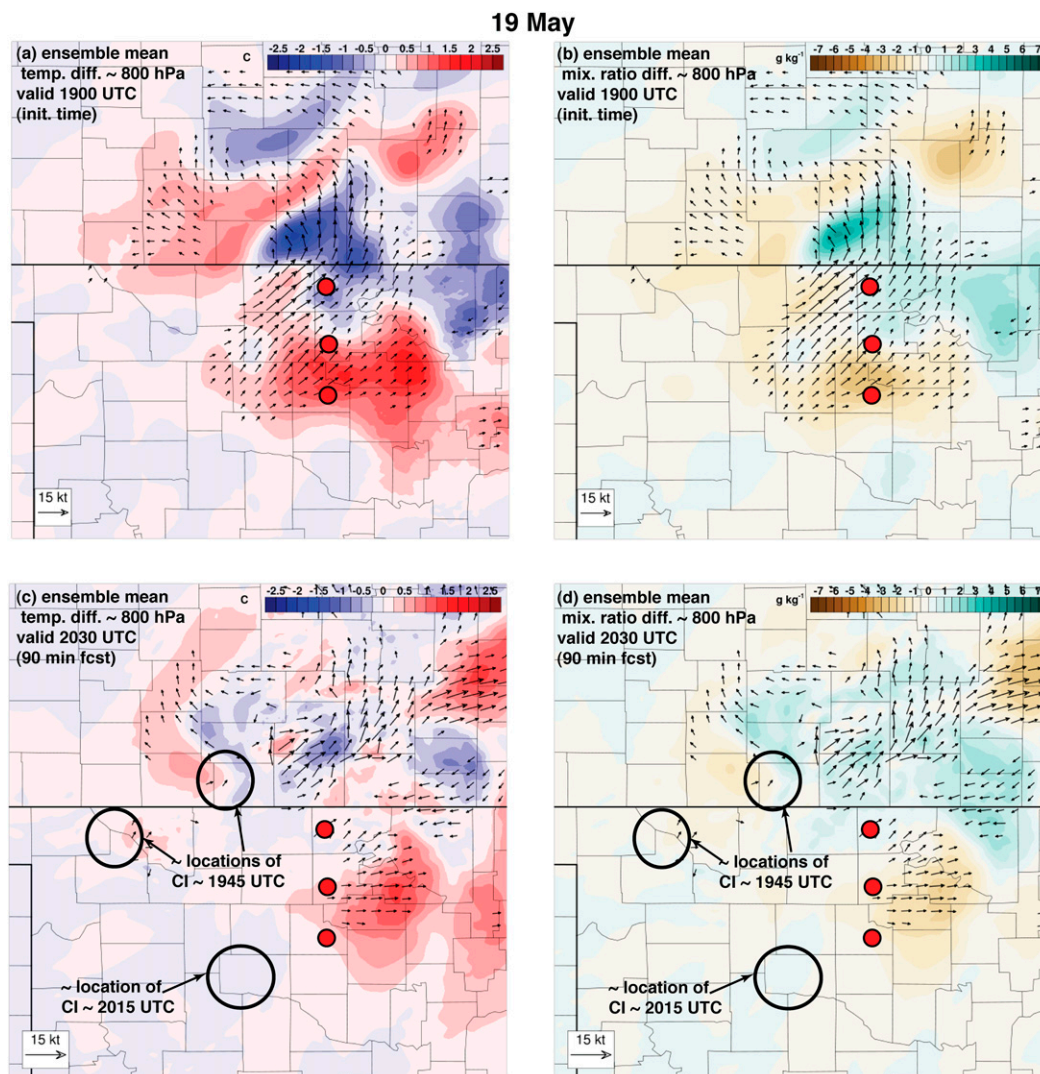


FIG. 12. (a) Ensemble mean difference (MPEX – control) in temperature ($^{\circ}\text{C}$) and winds (kt) and (b) water vapor mixing ratio (g kg^{-1}) and winds (kt) at ~ 800 hPa valid at 1900 UTC 19 May 2013 (initialization time of the 3-km forecasts). (c),(d) As in (a),(b), but for 90-min forecasts (valid at 2030 UTC) with approximate locations of observed CI circled in black (CI in the model forecasts occurred about 30–60 min later). Locations of the three MPEX soundings taken around 1900 UTC are shown with the red dots.

ensemble members producing storms about a county too far north from where the main cluster of storms was observed (Fig. 14). The MPEX ensemble also has slightly fewer storms (differences in neighborhood probabilities of 5%) in the vicinity where storms were observed (Fig. 14). This 3-h forecast error relates to errors in CI location that result from an area of moistening and cooling near the top of the boundary layer (i.e., low-level destabilization) in the final MPEX ensemble analysis at 2100 UTC (along the Kansas–Nebraska border in Figs. 15a and 15b). An hour into the forecasts, the area of moistening and cooling moves to the north-northeast into Nebraska

(Fig. 15c). The MPEX ensemble produces many more storms than the control ensemble along the northern edge of this moistening around 2215 UTC (area circled in Fig. 15c), whereas CI was observed about two counties to the southwest about a half-hour earlier (Fig. 15c).

The above analysis for the 27 May case highlights that verifying forecasts of convection that do not overlap with observed storms can be complex. An argument could be made that the MPEX ensemble has the better forecasts at this time because it produces abundant CI close to the time of observed CI, whereas only a few members of the control ensemble produce storms by this

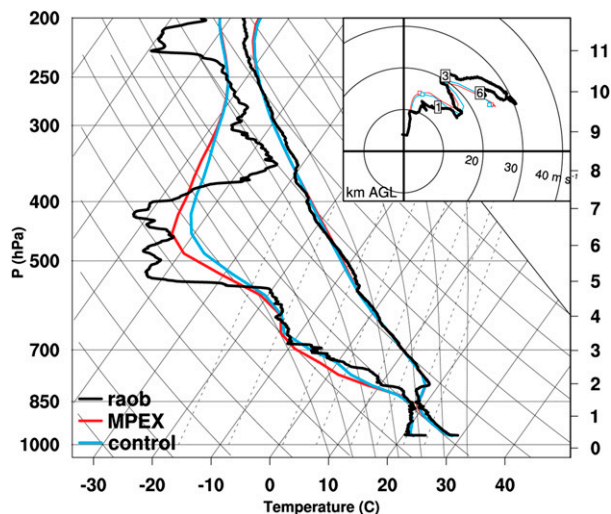


FIG. 13. Skew T -log p diagram of a nonassimilated MPEX sounding released at 1830 UTC 20 May 2013 (black lines) compared to ensemble mean 30-min forecasts from the MPEX (red) and control (blue) ensembles. The inset in the top right shows a comparison of the observed and ensemble mean winds in a hodograph over the lowest 6 km AGL.

time. But because the abundant CI in the MPEX ensemble is displaced about two counties away from observed CI, and is outside the 25 km \times 25 km box used to compute the FSS, the MPEX ensemble forecasts are penalized twice: once for producing storms where none are observed and again for not having storms where they are observed; the so-called double penalty (Gilleland et al. 2009) (as seen later, the use of a larger neighborhood alleviates some of this problem). In the first few hours, the control ensemble forecasts have larger FSS because they are only penalized for not having storms close to where they were observed (Fig. 10c).

After a few hours in the 27 May case, the MPEX ensemble has too few storms close to where the storms were observed compared to the control ensemble (Fig. 14). Both ensembles are too slow with developing the number and intensity of storms in the first few forecast hours, but the control ensemble initiates enough storms around 2315 UTC near the location of the evolving observed storms (Fig. 15d) to provide some overlap with the location of the observed storms at 0000 UTC (Fig. 14). These storms in the control ensemble initiate along a convergence zone associated with the stationary front (not shown). Through inspection of cloud and precipitation mixing ratio fields among the individual member forecasts, storms try to develop similarly in the MPEX ensemble forecasts but are delayed, likely because of warming (and some drying) in this region in the MPEX ensemble (Figs. 15c and 15d). As the storms mature and move east, the

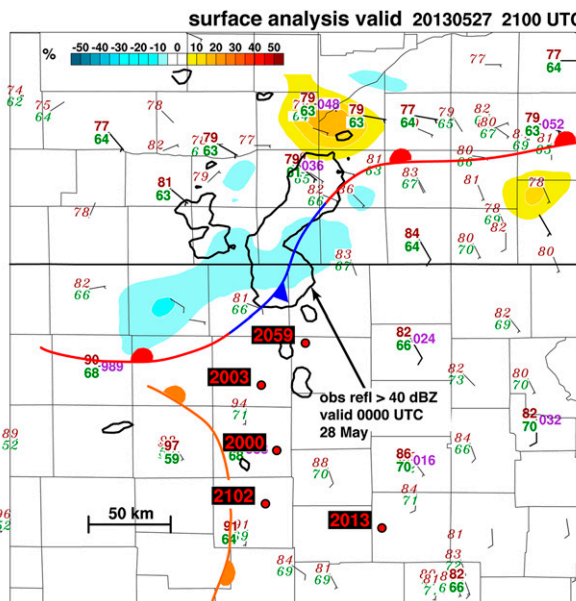


FIG. 14. As in Fig. 6a, but for surface observations and analysis valid at 2100 UTC 27 May 2013, observed reflectivity valid at 0000 UTC 28 May, and difference in neighborhood probabilities valid at 0000 UTC 28 May (180-min forecasts).

degradation in FSS for the MPEX ensemble (Fig. 10c) continues because of a phase error resulting from these errors in CI in the first few hours.

The convective scenario for the 28 May case was similar to that on 27 May with a weak southwest–northeast-oriented stationary front across Kansas with four MPEX soundings (Fig. 1) sampling the moist, unstable air mass to its south and east. The behavior of the differences in the preconvective environment is similar to that for the 19 May and 20 May experiments (not shown for brevity)—the adjustments decrease in magnitude with time and do not cover the areas where most of the storms developed and matured a few hours later. This is partly because few storms developed near the location of the MPEX soundings in the 3 h after the soundings were released (Fig. 1). Therefore, the lack of FSS improvements seen in the 28 May MPEX ensemble could again be related to a suboptimal sampling location.

c. Aggregation of fractions skill scores over the eight cases

1) 25-KM NEIGHBORHOODS

The results presented earlier suggest that when convection develops in the general region encompassed by the MPEX soundings, convection-permitting forecasts that are driven by a mesoscale ensemble of analyses that assimilate those soundings are improved (when considering scales close to the smallest resolvable scales of the

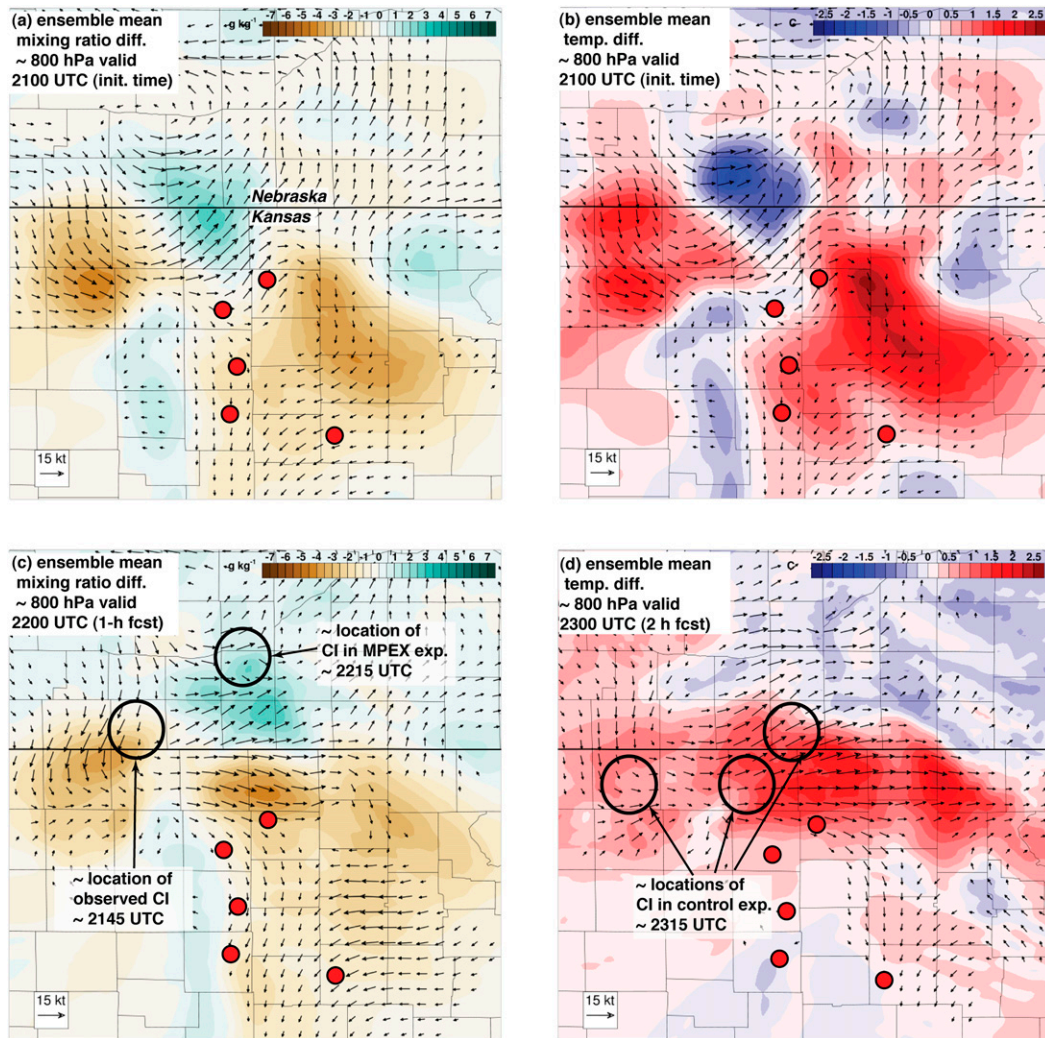


FIG. 15. (a) Ensemble mean difference (MPEX – control) in water vapor mixing ratio (g kg^{-1}) and winds (kt) at ~ 800 hPa valid at 2100 UTC 27 May 2013 (initialization time of the 3-km forecasts). (b) Ensemble mean difference (MPEX – control) in temperature ($^{\circ}\text{C}$) and winds (kt) at ~ 800 hPa valid at 2100 UTC (initialization time of the 3-km forecasts). (c) As in (a), but for 60-min forecasts (valid at 2200 UTC) with approximate locations of subsequent CI that occurs in several of the MPEX experiment members (but not in the control members) and approximate location of observed CI indicated. (d) As in (b), but for 120-min forecasts (valid at 2300 UTC) with approximate locations of subsequent CI that occurs in several of the control experiment members (but not in the MPEX members) indicated. Locations of the five MPEX soundings taken between 2000 and 2100 UTC are shown with the red dots.

grid). Adjustments to the wind and thermodynamic fields are either smaller or advected out of the region of CI in the four cases that lack improvements in FSS.⁴ Because of this case-dependent variability in relative

skill, aggregation of FSS⁵ for a reflectivity threshold of 40 dBZ and a $25 \text{ km} \times 25 \text{ km}$ neighborhood for all eight cases reveals no aggregate improvement in skill for the

⁴ At times, it was necessary to keep all the MPEX sounding vehicles relatively close to one another to address goals of the MPEX project other than data assimilation (see Weisman et al. 2015; Trapp et al. 2016), and this sometimes prevented the vehicles from sampling closer to the expected location of CI.

⁵ The aggregated FSS is computed by first aggregating the neighborhood probabilities rather than averaging the FSSs over the eight cases. Prior to aggregation, the forecasts are normalized by the time CI first occurs in either the MPEX or control experiments to account for different durations between the start time of the forecasts and the time of CI (the range of these durations is about 30–100 min).

MPEX experiments, except for a ~ 45 -min window about 2 h after storms develop in the model (Fig. 16).

Physical reasons for the FSS differences, like those discussed earlier, are much more difficult to determine 3–4 h after CI because of complex storm evolution and growth in convective coverage, along with an increase in spread among the convective forecasts. However, it is worth noting that the aggregated FSS shows that the MPEX upsonde data had an overall negative impact on the smallest resolvable scales of the model about 4–5 h after model CI (although statistical significance is not large; Fig. 16). This is partly because the positive impacts from the MPEX soundings seen for the 31 May and 8 June cases are largely lost about 3 h after CI; but the large reversal in relative skill between the 8 June MPEX and control ensemble forecasts about 3 h after CI (Fig. 5d) also plays a role.

2) 100-KM NEIGHBORHOODS

Recall that the FSSs presented up to this point are for a single reflectivity threshold (40 dBZ) and neighborhood size (~ 25 km \times 25 km). The use of larger neighborhoods results in larger FSS scores as the model is given more credit for correct forecasts for storms that are farther away from the observed storm (Schwartz et al. 2010). A consequence of using larger neighborhoods (and to a lesser extent, smaller reflectivity thresholds) is that it allows for a determination of the forecast skill of convection on different spatial scales (Ebert 2009; Mittermaier and Roberts 2010). The following examines the relative difference in skill when relaxing the neighborhood size to ~ 100 km \times 100 km and lowering the reflectivity threshold to 30 dBZ (Figs. 17 and 18), thus the focus shifts from the forecast skill of meso- γ -scale clusters of thunderstorms to meso- β -scale groups of thunderstorms or MCSs (e.g., Stratman et al. 2013; Pinto et al. 2015). Hereafter, the evaluations for a reflectivity threshold of 40 dBZ and a neighborhood size of ~ 25 km \times 25 km are referred to as the 40/25 evaluations and the evaluations for a reflectivity threshold of 30 dBZ and a neighborhood size of ~ 100 km \times 100 km are referred to as the 30/100 evaluations.

For the 18 May and 23 May cases, the MPEX FSS is larger than the control FSS for both the 40/25 and 30/100 evaluations (cf. Figs. 5a,b and 17a,b), particularly for the 18 May case. For the 27 May case, except for the 60-min forecasts, the MPEX FSS is smaller than the control FSS for both the 40/25 and 30/100 evaluations (cf. Figs. 10c and 18c).

For some cases, the 30/100 evaluation widens the time range of the improvement in FSS for the MPEX ensemble over the control ensemble over that seen for the

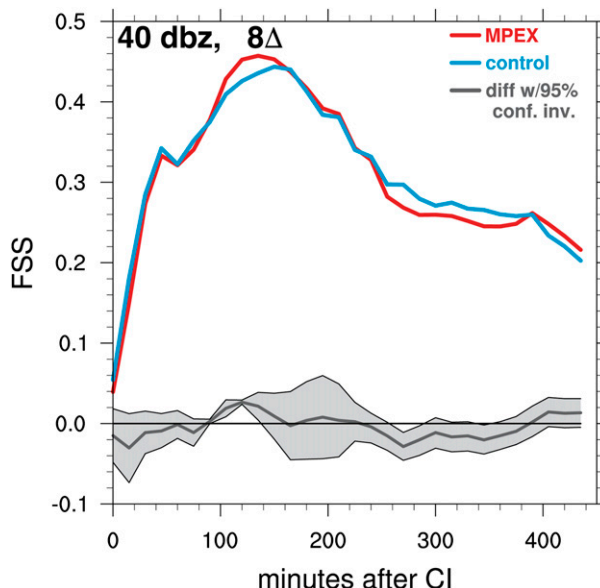


FIG. 16. Aggregation of FSS over the eight cases for a reflectivity threshold of 40 dBZ and a neighborhood size of ~ 25 km \times 25 km. Prior to aggregation, the forecast times among the cases are normalized by the time that CI first occurs in either the MPEX or control ensembles to account for different durations between the start time of the forecasts and the time of CI.

40/25 evaluations. For example, for the 28 May case, the MPEX FSS is larger than the control FSS at ~ 100 – 120 min for the 40/25 evaluation, but the time range of improved FSS increases to ~ 80 – 160 min and the peak FSS differences in these periods increase from 0.03 in the 40/25 evaluation to 0.10 in the 30/100 evaluation (cf. Figs. 10d and 18d). For the 8 June case, the relatively long time window with larger MPEX FSS in the 40/25 evaluation (~ 160 – 280 min) is extended even longer to ~ 100 – 360 min (cf. Figs. 5d and 17d).

Furthermore, for some forecast times when the MPEX FSS is smaller than the control FSS for the 40/25 evaluation, the opposite is true for the 30/100 evaluation. This reversal in relative FSS differences occurs at ~ 120 – 280 min for the 19 May case (cf. Figs. 10a and 18a), at ~ 120 – 160 min for the 28 May case (cf. Figs. 10d and 18d), and at ~ 300 – 360 min for the 8 June case (cf. Figs. 5d and 17d). There are no instances when the MPEX FSS is larger than the control FSS for the 40/25 evaluation but smaller for the 30/100 evaluation.

As expected from the above analysis of the relative differences in FSS for the two different threshold/neighborhood combinations, the aggregation of FSS among the eight cases for the 30/100 evaluation leads to a longer period when the aggregate MPEX FSS is larger than the control FSS (Fig. 19). For the

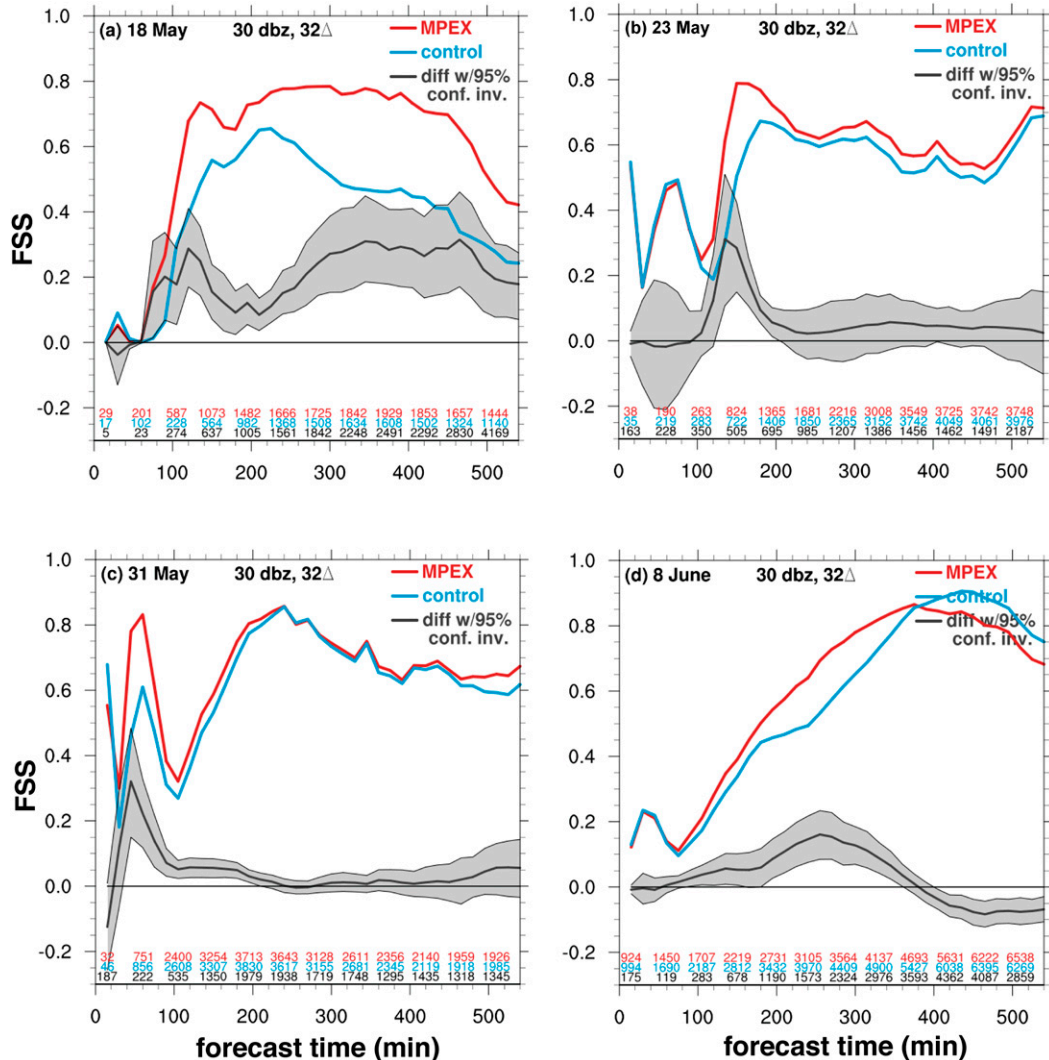


FIG. 17. As in Fig. 5, but for a reflectivity threshold of 30 dBZ and a neighborhood size of $\sim 100 \text{ km} \times 100 \text{ km}$. The numbers along the bottom of each figure correspond to the number of grid cells with reflectivity $\geq 30 \text{ dBZ}$ for the observed reflectivity (black text), the mean number of grid cells with reflectivity $\geq 30 \text{ dBZ}$ for the MPEX ensemble forecasts (red text), and the mean number of grid cells with reflectivity $\geq 30 \text{ dBZ}$ for the control ensemble forecasts (blue text).

aggregated FSS for the 40/25 evaluation, the MPEX FSS is larger than the control FSS for a relatively small period ($\sim 100\text{--}145 \text{ min}$ after CI; Fig. 16). However, the aggregated MPEX FSS for the 30/100 evaluation is larger than that for the control experiment from the time of CI to $\sim 260 \text{ min}$ after CI (Fig. 19).

4. Summary and conclusions

This study examines the impact of assimilating radiosonde observations (“upsondes”) taken by ground-based mobile sounding systems $\sim 1\text{--}4 \text{ h}$ prior to convective storm development on forecasts of convection

within a WRF-ARW ensemble using the DART data assimilation software. The upsondes were obtained as a part of MPEX (Trapp et al. 2016; Weisman et al. 2015). Focus is placed on the impacts of assimilating the MPEX upsonde data on the initial development and early evolution of convection. Including the 31 May 2013 convective event over Oklahoma examined in detail by Hitchcock et al. (2016), the impacts of the MPEX upsondes on ensemble forecasts of convection are examined for eight convective events observed during MPEX. At least three MPEX radiosondes sampled the environment in all eight cases.

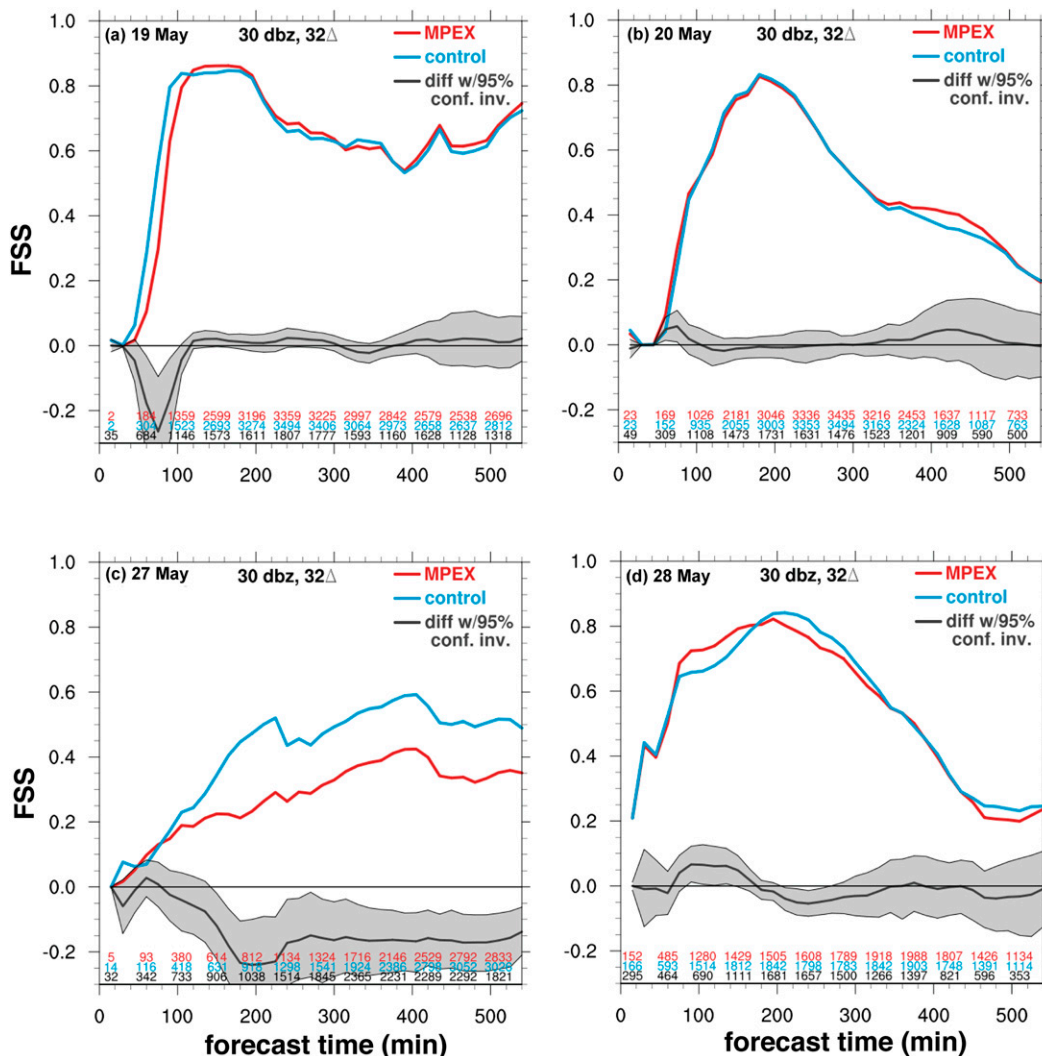


FIG. 18. As in Fig. 10, but for a reflectivity threshold of 30 dBZ and a neighborhood size of $\sim 100 \text{ km} \times 100 \text{ km}$. The numbers along the bottom of each figure correspond to the number of grid cells with reflectivity ≥ 30 dBZ for the observed reflectivity (black text), the mean number of grid cells with reflectivity ≥ 30 dBZ for the MPEX ensemble forecasts (red text), and the mean number of grid cells with reflectivity ≥ 30 dBZ for the control forecasts (blue text).

The results suggest that the assimilation of radiosonde observation profiles into a mesoscale ensemble generate coherent mesoscale perturbations in temperature, humidity, and winds in the posterior analyses that persist in subsequent forecasts. In four cases, these perturbations are found to have positive impacts on the skill of forecasts of convection even on scales approaching the limits of what the model can resolve (large supercells or small groups of cells). This is somewhat surprising given the rapid growth of errors at small scales induced by deep convection and the short time scales of intrinsic predictability (Ehrendorfer et al. 1999; Zhang et al. 2007; Melhauser and Zhang 2012; Cintineo and Stensrud 2013), but is a testament to the ability of

the mesoscale environment to raise the practical predictability of convection (Lorenz 1982; Zhang et al. 2007; Zhang et al. 2015). However, positive impacts are not seen for every case. Four of the cases (19, 20, 27, and 28 May) show no improvement overall and at times show a degradation in forecast skill of convection for the smallest resolvable scales of the grid.

Reasons for the differences in forecast impacts among the eight cases are explored, and a common characteristic of the four cases with positive impacts near the smallest resolvable scales of the grid is that the air mass on both sides of an initiating boundary was sampled by at least one of the upsondes. A variety of

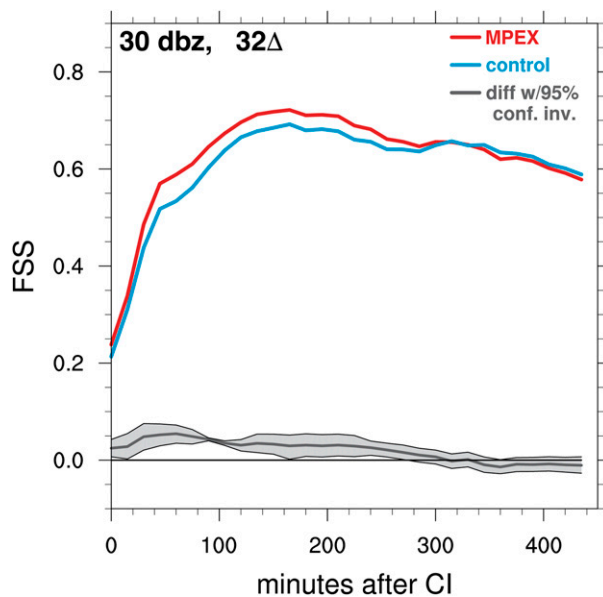


FIG. 19. As in Fig. 16, but for a reflectivity threshold of 30 dBZ and a neighborhood size of $\sim 100 \text{ km} \times 100 \text{ km}$.

boundary types were sampled among these four positive-impacts cases, including synoptic-scale fronts (31 May and 8 June), a dryline (18 May), and an outflow boundary emanating from a morning MCS (23 May). The cool or dry side of the boundaries was not sampled in the four cases that show no forecast improvements. Exploration of the correlation structures in the lower troposphere indeed show that higher correlations between variables tend not to cross significant air mass boundaries (e.g., Fig. 20). A hypothesis arising from these results is that it is important to sample the air mass on both sides of the initiating boundary if forecasts are to be improved. To address this hypothesis, experiments are performed on the four positive impact cases that are identical to the MPEX experiments except the sounding(s) that were taken on the cooler/drier side of the boundaries are not assimilated (the “MPEX-nopostbdy” ensemble). Results suggest that assimilating these postboundary data is not necessarily important as there are no significant differences in the MPEX-nopostbdy and MPEX ensembles, except for a short period in the 18 May case in which the smaller FSSs for the MPEX-nopostbdy ensemble approach statistical significance around 240 min (Fig. 21).

The main reason for the differences in relative skill among the cases appears to relate simply to differences in the magnitudes of the wind and thermodynamic adjustments among the cases and their location relative to CI, as shown in sections 3b(1) and 3b(2).

The convection in the cases with no forecast improvements developed farther away from the MPEX soundings than in the cases with positive impacts. Examination of the ensemble mean adjustments to variables over the low and midtroposphere show that 1) the ensemble mean fields tend to be adjusted little because the control ensemble mean state is already accurate (20 May in particular; e.g., Fig. 20b), 2) tend to advect away from the CI area (19 May and 20 May in particular), or 3) are removed from the main region of CI to begin with (28 May) (the lack of improvements seen in the 27 May case are complex and do not fall neatly into one of the three reasons above). Overall, the lack of positive impacts, therefore, could be related to suboptimal sampling strategies on these days (indeed, the preconvective sampling strategies on these days were chosen to address MPEX goals other than testing impacts from data assimilation).

It is also worth noting that the adjustments to the midlevel temperature and humidity do not seem to play as large a role in the improved forecasts for the 18 May, 23 May, and 8 June cases as they did for the 31 May 2013 case shown in Hitchcock et al. (2016). In the cases examined herein, the adjustments to temperature and humidity above 700 hPa are typically small or are advected out of the area in which CI occurs. In the 31 May case, the reduced humidity in the 700–500-hPa layer in the MPEX ensemble suppressed spurious storms that plagued the control experiment, but in the three positive impact cases shown in the current study, the improvements in the forecasts are mostly associated with improved depictions of the flow and thermodynamics in the lower troposphere and subsequent improvements in the timing and location of CI [as is part of the 31 May event as well as shown in Hitchcock et al. (2016)].

Despite the case dependence to the forecast skill, aggregate measures of skill (fractions skill score) still show systematic positive impacts in the first few hours after convection develops when $\sim 25\text{-km}$ neighborhoods are considered (representing the skill on the smallest resolvable scales of the grid) and for a longer period when $\sim 100\text{-km}$ neighborhoods are considered. In other words, given that CI occurs 1–2 h into the forecasts, these results (and verification using other reflectivity thresholds and neighborhood sizes not shown) show that positive impacts for the higher reflectivity thresholds and smallest resolvable scales tend to be lost after 3–4 h into the forecasts, but the positive impacts can continue much further into the forecasts if smaller reflectivity thresholds and/or larger spatial scales are considered in the verification.

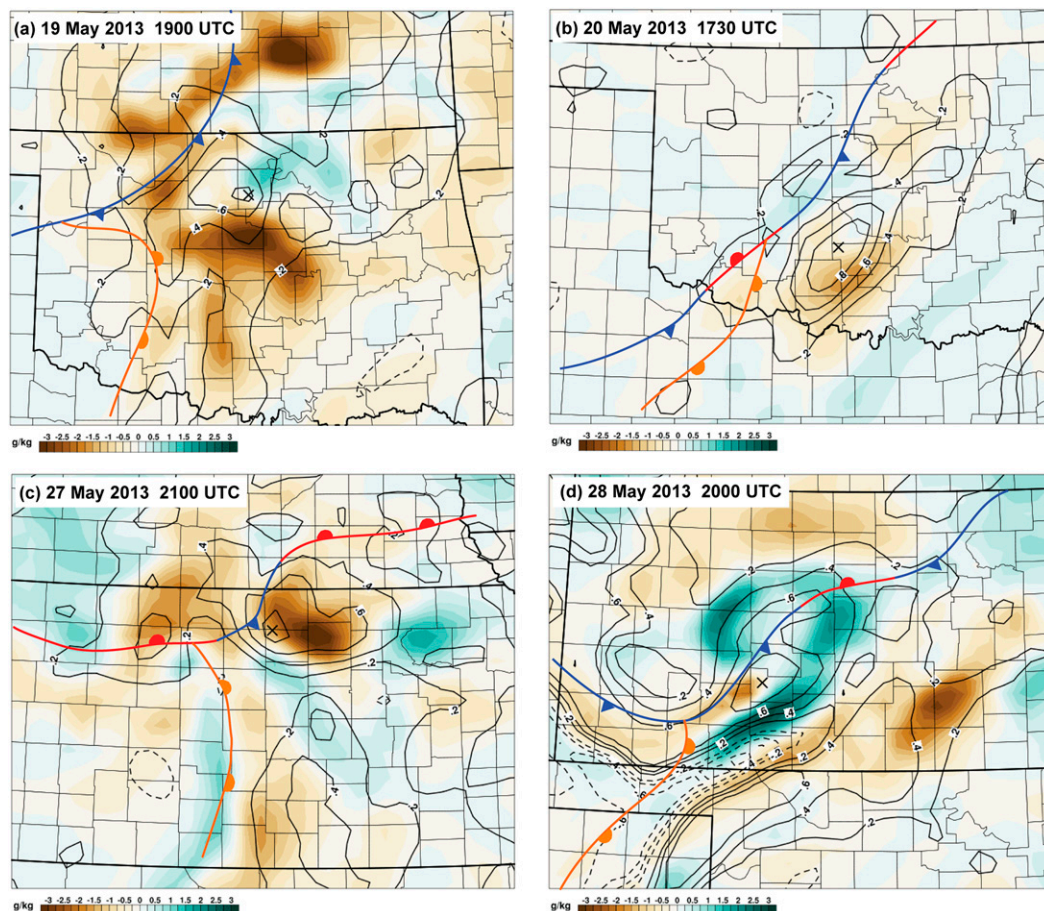


FIG. 20. Spatial correlation (black contours) between the 800-hPa mixing ratio at the points denoted by the \times (locations sampled by radiosondes) and the prior ensemble analyses (1-h forecasts) of 800-hPa mixing ratio valid at (a) 1900 UTC 19 May, (b) 1730 UTC 20 May, (c) 2100 UTC 27 May, and (d) 2000 UTC 28 May. Increments (posterior – prior) in 800-hPa mixing ratio at the valid time are shown by the color shading.

As a final note, we deliberately did not assimilate radar or other remotely sensed data into the initial model condition so that the impacts of the upsonde data assimilation on the mesoscale environment and convective forecasts could be isolated effectively. It is recognized that assimilation of clear-air radial velocity observations from radar data and retrievals of temperature and water vapor from ground-based profiling systems (Wulfmeyer et al. 2015) or satellites (e.g., Jones et al. 2015) may improve the mesoscale background environment further, and the relative impacts of upsonde data versus remotely sensed profiles of the preconvective atmosphere should be examined in future studies. Furthermore, if the goal is to closely emulate observed storms in model initial conditions, then radar data should be assimilated to introduce hydrometeors and wind perturbations that represent existing storms into the model initial conditions (e.g., Dowell et al. 2011;

Snook et al. 2012; Wheatley et al. 2014), as well as suppress spurious storm development in model forecasts. It is clear that assimilation of radar data is vital to increase the practical predictability of convection on scales of individual thunderstorms (Dowell et al. 2004; Stensrud et al. 2009; Snook et al. 2012; Jones et al. 2015; Yussouf et al. 2015). However, in order for radar data assimilation to be effective, key meteorological features in the mesoscale background that support the convection (e.g., fronts, drylines) need to be represented accurately (Aksoy et al. 2009; Stensrud and Gao 2010; Dong et al. 2011; Yussouf et al. 2015). Future studies should examine if the potential added benefits of nearby sounding assimilation identified in this study is seen in concurrent mesoscale/convective-scale assimilation systems (e.g., Johnson et al. 2015; Yussouf et al. 2015) that include radar data in the cycled assimilation procedure.

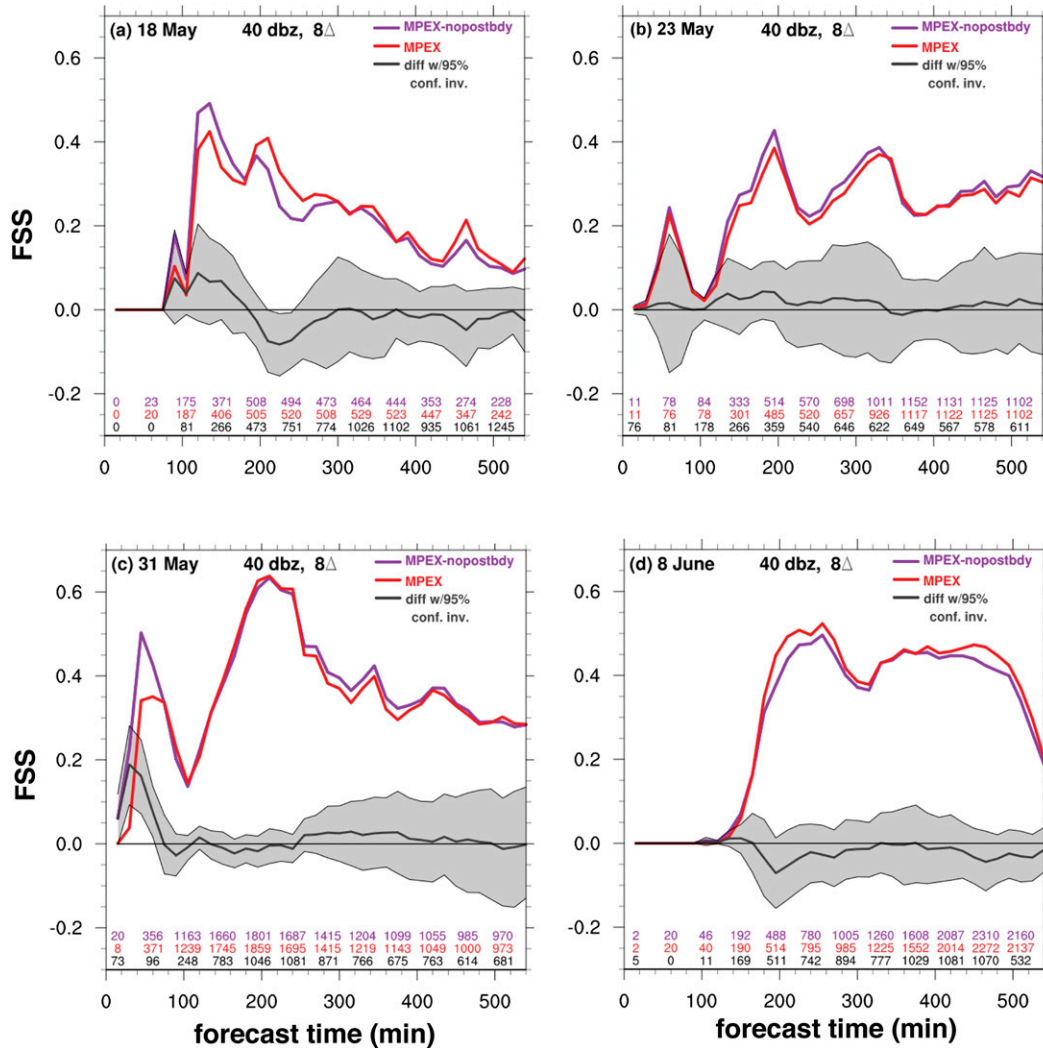


FIG. 21. As in Fig. 5, but for the MPEX (red) and MPEX-nopostbdy (purple) forecasts and their differences (MPEX – nopostbdy minus control; black line).

Acknowledgments. We thank the Field Observing Facilities Support team at NSSL, particularly Sean Waugh, for developing and maintaining the NSSL mobile sounding systems for MPEX. The NSSL sounding system was led in the field by the first author and Dr. David Stensrud. The other sounding teams were led by Dr. Russ Schumacher (CSU), Drs. Jeff Trapp and Mike Baldwin (Purdue), and Dr. Don Conlee (TAMU). Their leadership and collaboration are much appreciated, as well as the efforts of numerous students that helped collect the radiosonde data from all four systems. Dusty Wheatley provided helpful comments and suggestions on the experiment design and early versions of the manuscript. We thank two anonymous reviewers for their helpful comments and suggestions. We also thank the support from the Warn-on-Forecast group at NSSL,

especially Gerry Greager, for technical support. The efforts of EOL in quality controlling and formatting the data are much appreciated. A portion of the computing for this project was performed at the OU Supercomputing Center for Education and Research (OSCAR) at the University of Oklahoma (OU). This project was supported by funding from the NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA11OAR4320072 U.S. Department of Commerce, and by National Science Foundation Award 1230114.

REFERENCES

Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A multicaser comparative assessment of the ensemble Kalman filter for

- assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805–1824, doi:10.1175/2008MWR2691.1.
- 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.
- , 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus*, **59**, 210–224, doi:10.1111/j.1600-0870.2006.00216.x.
- , and N. Collins, 2007: Scalable implementations of ensemble filter algorithms for data assimilation. *J. Atmos. Oceanic Technol.*, **24**, 1452–1463, doi:10.1175/JTECH2049.1.
- , T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community facility. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296, doi:10.1175/2009BAMS2618.1.
- Bednarczyk, C. N., and B. C. Ancell, 2015: Ensemble sensitivity analysis applied to a southern plains convective event. *Mon. Wea. Rev.*, **143**, 230–249, doi:10.1175/MWR-D-13-00321.1.
- Benjamin, S. G., B. D. Jamison, W. R. Moninger, S. R. Sahn, B. E. Schwartz, and T. W. Schlatter, 2010: Relative short-range forecast impact from aircraft, profiler, radiosonde, VAD, GPS-PW, METAR, and Mesonet observations via the RUC hourly assimilation cycle. *Mon. Wea. Rev.*, **138**, 1319–1343, doi:10.1175/2009MWR3097.1.
- Bryan, G. H., and H. Morrison, 2012: Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Mon. Wea. Rev.*, **140**, 202–225, doi:10.1175/MWR-D-11-00046.1.
- , J. C. Wyngaard, and J. M. Fritsch, 2003: Resolution requirements for the simulation of deep moist convection. *Mon. Wea. Rev.*, **131**, 2394–2416, doi:10.1175/1520-0493(2003)131<2394:RRFTSO>2.0.CO;2.
- Burghardt, B. J., C. Evans, and P. J. Roebber, 2014: Assessing the predictability of convection initiation in the high plains using an object-based approach. *Wea. Forecasting*, **29**, 403–418, doi:10.1175/WAF-D-13-00089.1.
- Carbone, R., and J. Tuttle, 2008: Rainfall occurrence in the U.S. warm season: The diurnal cycle. *J. Climate*, **21**, 4132–4146, doi:10.1175/2008JCL2275.1.
- , —, D. Ahijevych, and S. Trier, 2002: Inferences of predictability associated with warm season precipitation episodes. *J. Atmos. Sci.*, **59**, 2033–2056, doi:10.1175/1520-0469(2002)059<2033:IOPAWW>2.0.CO;2.
- Cintineo, R. M., and D. J. Stensrud, 2013: On the predictability of supercell thunderstorm evolution. *J. Atmos. Sci.*, **70**, 1993–2011, doi:10.1175/JAS-D-12-0166.1.
- Clark, A. J., and Coauthors, 2012: An overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. *Bull. Amer. Meteor. Soc.*, **93**, 55–74, doi:10.1175/BAMS-D-11-00040.1.
- Coniglio, M. C., 2012: Verification of RUC 0–1-h forecasts and SPC mesoscale analyses using VORTEX2 soundings. *Wea. Forecasting*, **27**, 667–683, doi:10.1175/WAF-D-11-00096.1.
- Dawson, D. T., II, L. J. Wicker, E. R. Mansell, and R. L. Tanamachi, 2012: Impact of the environmental low-level wind profile on ensemble forecasts of the 4 May 2007 Greensburg, Kansas, tornadic storm and associated mesocyclones. *Mon. Wea. Rev.*, **140**, 696–716, doi:10.1175/MWR-D-11-00008.1.
- Dong, J., M. Xue, and K. Droegemeier, 2011: The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteor. Atmos. Phys.*, **112**, 41–61, doi:10.1007/s00703-011-0130-3.
- Dowell, D. C., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132**, 1982–2005, doi:10.1175/1520-0493(2004)132<1982:WATRIT>2.0.CO;2.
- , L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139**, 272–294, doi:10.1175/2010MWR3438.1.
- Ebert, E. E., 2009: Neighborhood verification: A strategy for re-warding close forecasts. *Wea. Forecasting*, **24**, 1498–1510, doi:10.1175/2009WAF2222251.1.
- Ehrendorfer, M., R. M. Errico, and K. D. Raeder, 1999: Singular-vector perturbation growth in a primitive equation model with moist physics. *J. Atmos. Sci.*, **56**, 1627–1648, doi:10.1175/1520-0469(1999)056<1627:SVPGIA>2.0.CO;2.
- 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.
- Fabry, F., 2010: For how long should what data be assimilated for the mesoscale forecasting of convection and why? Part II: On the observation signal from different sensors. *Mon. Wea. Rev.*, **138**, 256–264, doi:10.1175/2009MWR2884.1.
- , and J. Sun, 2010: For how long should what data be assimilated for the mesoscale forecasting of convection and why? Part I: On the propagation of initial condition errors and their implications for data assimilation. *Mon. Wea. Rev.*, **138**, 242–255, doi:10.1175/2009MWR2883.1.
- Fujita, T., D. J. Stensrud, and D. C. Dowell, 2007: Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainties. *Mon. Wea. Rev.*, **135**, 1846–1868, doi:10.1175/MWR3391.1.
- 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.
- Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of spatial forecast verification methods. *Wea. Forecasting*, **24**, 1416–1430, doi:10.1175/2009WAF2222269.1.
- Hamill, T. M., 1999: Hypothesis tests for evaluating numerical precipitation forecasts. *Wea. Forecasting*, **14**, 155–167, doi:10.1175/1520-0434(1999)014<0155:HTFENP>2.0.CO;2.
- , 2006: Ensemble-based atmospheric data assimilation. *Predictability of Weather and Climate*, T. Palmer and R. Hagedorn, Eds., Cambridge University Press, 124–156, doi:10.1017/CBO9780511617652.007.
- Hitchcock, S. M., M. C. Coniglio, and K. H. Knopfmeier, 2016: Impact of MPEX upsonde observations on ensemble analyses and forecasts of the 31 May 2013 convective event over Oklahoma. *Mon. Wea. Rev.*, **144**, 2889–2913, doi:10.1175/MWR-D-15-0344.1.
- Johnson, A., X. Wang, J. R. Carley, L. J. Wicker, and C. Karstens, 2015: A comparison of multiscale GSI-based EnKF and 3DVAR data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. *Mon. Wea. Rev.*, **143**, 3087–3108, doi:10.1175/MWR-D-14-00345.1.
- Jones, T. A., J. A. Otkin, D. J. Stensrud, and K. Knopfmeier, 2013: Assimilation of satellite infrared radiances and Doppler radar observations during a cool season observing system simulation experiment. *Mon. Wea. Rev.*, **141**, 3273–3299, doi:10.1175/MWR-D-12-00267.1.

- , D. Stensrud, L. Wicker, P. Minnis, and R. Palikonda, 2015: Simultaneous radar and satellite data storm-scale assimilation using an ensemble Kalman filter approach for 24 May 2011. *Mon. Wea. Rev.*, **143**, 165–194, doi:10.1175/MWR-D-14-00180.1.
- , K. Knopfmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikonda, 2016: Storm-scale data assimilation and ensemble forecasting with the NSSL experimental warn-on-forecast system. Part II: Combined radar and satellite data experiments. *Wea. Forecasting*, **31**, 297–327, doi:10.1175/WAF-D-15-0107.1.
- Kain, J. S., S. J. Weiss, J. J. Levit, M. E. Baldwin, and D. R. Bright, 2006: Examination of convection-allowing configurations of the WRF Model for the prediction of severe convective weather: The SPC/NSSL spring program 2004. *Wea. Forecasting*, **21**, 167–181, doi:10.1175/WAF906.1.
- Knopfmeier, K. H., and D. J. Stensrud, 2013: Influence of Mesonet observations on the accuracy of surface analyses generated by an ensemble Kalman filter. *Wea. Forecasting*, **28**, 815–841, doi:10.1175/WAF-D-12-00078.1.
- Lin, X., and K. Hubbard, 2004: Uncertainties of derived dewpoint temperature and relative humidity. *J. Appl. Meteor.*, **43**, 821–825, doi:10.1175/2100.1.
- Liu, Z.-Q., and F. Rabier, 2002: The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study. *Quart. J. Roy. Meteor. Soc.*, **128**, 1367–1386, doi:10.1256/003590002320373337.
- Lorenz, E., 1982: Atmospheric predictability experiments with a large numerical model. *Tellus*, **34**, 505–513, doi:10.1111/j.2153-3490.1982.tb01839.x.
- Marquis, J., Y. Richardson, P. Markowski, D. Dowell, J. Wurman, K. Kosiba, P. Robinson, and G. Romine, 2014: An investigation of the Goshen County, Wyoming, tornadic supercell of 5 June 2009 using EnKF assimilation of mobile mesonet and radar observations collected during VORTEX2. Part I: Experiment design and verification of the EnKF analyses. *Mon. Wea. Rev.*, **142**, 530–554, doi:10.1175/MWR-D-13-00007.1.
- Melhauser, C., and F. Zhang, 2012: Practical and intrinsic predictability of severe and convective weather at the mesoscales. *J. Atmos. Sci.*, **69**, 3350–3371, doi:10.1175/JAS-D-11-0315.1.
- Miller, P. A., M. Barth, L. Benjamin, R. Artz, and W. Pendergrass, 2007: MADIS support for UrbaNet. *14th Symp. on Meteorological Observation and Instrumentation/16th Conf. on Applied Climatology*, San Antonio, TX, Amer. Meteor. Soc., JP2.5. [Available online at https://ams.confex.com/ams/87ANNUAL/techprogram/paper_119116.htm.]
- Mittermaier, M., and N. Roberts, 2010: Intercomparison of spatial forecast verification methods: Identifying skillful spatial scales using the fractions skill score. *Wea. Forecasting*, **25**, 343–354, doi:10.1175/2009WAF2222260.1.
- Moller, A. R., C. A. Doswell III, M. P. Foster, and G. R. Woodall, 1994: The operational recognition of supercell thunderstorm environments and storm structures. *Wea. Forecasting*, **9**, 327–347, doi:10.1175/1520-0434(1994)009<0327:TOROST>2.0.CO;2.
- Parker, M. D., and R. H. Johnson, 2000: Organizational modes of midlatitude mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 3413–3436, doi:10.1175/1520-0493(2001)129<3413:OMOMMC>2.0.CO;2.
- Pinto, J. O., J. A. Grim, and M. Steiner, 2015: Assessment of the high-resolution rapid refresh model's ability to predict mesoscale convective systems using object-based evaluation. *Wea. Forecasting*, **30**, 892–913, doi:10.1175/WAF-D-14-00118.1.
- Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, doi:10.1175/2007MWR2123.1.
- Romine, G. S., C. S. Schwartz, C. Snyder, J. L. Anderson, and M. L. Weisman, 2013: Model bias in a continuously cycled assimilation system and its influence on convection-permitting forecasts. *Mon. Wea. Rev.*, **141**, 1263–1284, doi:10.1175/MWR-D-12-00112.1.
- Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570–575, doi:10.1175/1520-0434(1990)005<0570:TCSIAA>2.0.CO;2.
- Schwartz, C. S., and Coauthors, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting*, **25**, 263–280, doi:10.1175/2009WAF2222267.1.
- Skamarock, W. C., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Wea. Rev.*, **132**, 3019–3032, doi:10.1175/MWR2830.1.
- Snook, N., M. Xue, and Y. Jung, 2011: Analysis of a tornadic mesoscale convective vortex based on ensemble Kalman filter assimilation of CASA X-band and WSR-88D radar data. *Mon. Wea. Rev.*, **139**, 3446–3468, doi:10.1175/MWR-D-10-05053.1.
- , —, and —, 2012: Ensemble probabilistic forecasts of a tornadic mesoscale convective system from ensemble Kalman filter analyses using WSR-88D and CASA radar data. *Mon. Wea. Rev.*, **140**, 2126–2146, doi:10.1175/MWR-D-11-00117.1.
- Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677, doi:10.1175//2555.1.
- Stensrud, D. J., and J. Gao, 2010: Importance of horizontally inhomogeneous environmental initial conditions to ensemble storm-scale radar data assimilation and very short-range forecasts. *Mon. Wea. Rev.*, **138**, 1250–1272, doi:10.1175/2009MWR3027.1.
- , and Coauthors, 2009: Convective-scale warn-on-forecast system: A vision for 2020. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1499, doi:10.1175/2009BAMS2795.1.
- Stratman, D. R., M. C. Coniglio, S. E. Koch, and M. Xue, 2013: Use of multiple verification methods to evaluate forecasts of convection from hot-and cold-start convection-allowing models. *Wea. Forecasting*, **28**, 119–138, doi:10.1175/WAF-D-12-00022.1.
- Surcel, M., M. Berenguer, and I. Zawadzki, 2010: The diurnal cycle of precipitation from continental radar mosaics and numerical weather prediction models. Part I: Methodology and seasonal comparison. *Mon. Wea. Rev.*, **138**, 3084–3106, doi:10.1175/2010MWR3125.1.
- Torn, R. D., and G. J. Hakim, 2008: Ensemble-based sensitivity analysis. *Mon. Wea. Rev.*, **136**, 663–677, doi:10.1175/2007MWR2132.1.
- Trapp, R. J., D. J. Stensrud, M. C. Coniglio, R. S. Schumacher, M. E. Baldwin, S. Waugh, and D. T. Conlee, 2016: Mobile radiosonde deployments during the Mesoscale Predictability Experiment (MPEX): Rapid and adaptive sampling of upscale convective feedbacks. *Bull. Amer. Meteor. Soc.*, **97**, 329–336, doi:10.1175/BAMS-D-14-00258.1.
- Wandishin, M. S., D. J. Stensrud, S. L. Mullen, and L. J. Wicker, 2010: On the predictability of mesoscale convective systems: Three-dimensional simulations. *Mon. Wea. Rev.*, **138**, 863–885, doi:10.1175/2009MWR2961.1.

- Weisman, M. L., C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0–36-h explicit convective forecasts with the WRF-ARW Model. *Wea. Forecasting*, **23**, 407–437, doi:10.1175/2007WAF2007005.1.
- , and Coauthors, 2015: The Mesoscale Predictability Experiment (MPEX). *Bull. Amer. Meteor. Soc.*, **96**, 2127–2149, doi:10.1175/BAMS-D-13-00281.1.
- Wheatley, D. M., D. J. Stensrud, D. C. Dowell, and N. Yussouf, 2012: Application of a WRF mesoscale data assimilation system to springtime severe weather events 2007–09. *Mon. Wea. Rev.*, **140**, 1539–1557, doi:10.1175/MWR-D-11-00106.1.
- , N. Yussouf, and D. J. Stensrud, 2014: Ensemble Kalman filter analyses and forecasts of a severe mesoscale convective system using different choices of microphysics schemes. *Mon. Wea. Rev.*, **142**, 3243–3263, doi:10.1175/MWR-D-13-00260.1.
- , K. H. Knopfmeier, T. A. Jones, and G. J. Creager, 2015: Storm-scale data assimilation and ensemble forecasting with the NSSL experimental warn-on-forecast system. Part I: Radar data experiments. *Wea. Forecasting*, **30**, 1795–1817, doi:10.1175/WAF-D-15-0043.1.
- Wulfmeyer, V., and Coauthors, 2015: A review of the remote sensing of lower tropospheric thermodynamic profiles and its indispensable role for the understanding and the simulation of water and energy cycles. *Rev. Geophys.*, **53**, 819–895, doi:10.1002/2014RG000476.
- Yussouf, N., D. C. Dowell, L. J. Wicker, K. H. Knopfmeier, and D. M. Wheatley, 2015: Storm-scale data assimilation and ensemble forecasts for the 27 April 2011 severe weather outbreak in Alabama. *Mon. Wea. Rev.*, **143**, 3044–3066, doi:10.1175/MWR-D-14-00268.1.
- Zhang, F., Z. Meng, and A. Aksoy, 2006: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part I: Perfect model experiments. *Mon. Wea. Rev.*, **134**, 722–736, doi:10.1175/MWR3101.1.
- , N. Bei, R. Rotunno, C. Snyder, and C. C. Epifanio, 2007: Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics. *J. Atmos. Sci.*, **64**, 3579–3594, doi:10.1175/JAS4028.1.
- Zhang, J., and Coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) system: Description, results, and future plans. *Bull. Amer. Meteor. Soc.*, **92**, 1321–1338, doi:10.1175/2011BAMS-D-11-00047.1.
- Zhang, Y., F. Zhang, D. J. Stensrud, and Z. Meng, 2015: Practical predictability of the 20 May 2013 tornadic thunderstorm event in Oklahoma: Sensitivity to synoptic timing and topographical influence. *Mon. Wea. Rev.*, **143**, 2973–2997, doi:10.1175/MWR-D-14-00394.1.