# ⓐAn OSSE Study of the Impact of Micropulse Differential Absorption Lidar (MPD) Water Vapor Profiles on Convective Weather Forecasting

JUNKYUNG KAY,[a] TAMMY M. WECKWERTH,[a] WEN-CHAU LEE,[a] JENNY SUN,[a] AND GLEN ROMINE[a]

[a] National Center for Atmospheric Research, Boulder, Colorado

ABSTRACT: The National Center for Atmospheric Research (NCAR) and Montana State University jointly developed water vapor micropulse differential absorption lidars (MPDs) that are a significant advance in eye-safe, unattended, lidar-based water vapor remote sensing. MPD is designed to provide continuous vertical water vapor profiles with high vertical (150 m) and temporal resolution (5 min) in the lower troposphere. This study aims to investigate MPD observation impacts and the scientific significance of MPDs for convective weather analyses and predictions using observation system simulation experiments (OSSEs). In this study, the Data Assimilation Research Testbed (DART) and the Advanced Research version of the Weather Research and Forecasting (WRF-ARW) Model are used to conduct OSSEs for a case study of a mesoscale convective system (MCS) observed during the Plains Elevated Convection At Night (PECAN) experiment. A poor-performing control simulation that was drawn from a 40-member ensemble at 3-km resolution is markedly improved by assimilation of simulated observations drawn from a more skillful simulation that served as the nature run at 1-km resolution. In particular, assimilating surface observations corrected surface warm front structure errors, while MPD observations remedied errors in low- to midlevel moisture ahead of the MCS. Collectively, these analyses changes led to markedly improved short-term predictions of convection initiation, evolution, and precipitation of the MCS in the simulations on 15 July 2015. For this case study, the OSSE results indicate that a more dense MPD network results in better prediction performance for convective precipitation while degrading light precipitation prediction performance due to an imbalance of the analysis at large scales.

KEYWORDS: Deep convection; Lidars/Lidar observations; Profilers, atmospheric; Remote sensing; Numerical weather prediction/forecasting; Data assimilation

## 1. Introduction

Numerical weather prediction (NWP) models have advanced dramatically over the last two decades, due in part to continuous development of non-hydrostatic dynamics, more accurate representation of physical processes and data assimilation (DA) at convection-permitting resolutions (e.g., Bauer et al. 2015; Clark et al. 2016; Zhang et al. 2019). On the other hand, high-impact weather forecasting skill, including quantitative precipitation forecasting (QPF), has shown relatively slow improvement over the same period (Cherubini et al. 2002; Charba et al. 2003; Ralph et al. 2003; Morss and Ralph 2007; Brennan et al. 2008; Ralph et al. 2010; Schumacher and Davis 2010; Novak et al. 2011; Buehner and Jacques 2020). One of the reasons for relatively poor QPF skill is the lack of suitable thermodynamic profiling observations in the lower troposphere where nearly all high-impact weather occurs. This leads to difficulty capturing details of the rapidly changing water vapor and temperature, and their gradient in the lower atmosphere, which hinders the analysis of convective-scale weather systems and their environments (Dee et al. 2011; Wulfmeyer et al. 2015).

The sensitivity of atmospheric instability to water vapor is a critical factor for accurate prediction of convection initiation and evolution (Crook 1996; Ducrocq et al. 2002; Richard et al. 2007; Keil et al. 2008; Dierer et al. 2009; Wulfmeyer et al. 2011). It has been shown that small changes in the amount of water vapor within or just above the atmospheric boundary layer can determine whether or not convection initiates (Weckwerth 2000). In this context, several National Research Council reports noted that one of the significant gaps in observational capabilities for convective weather forecasting is the lack of sufficient water vapor measurements in the boundary layer (National Research Council 2009, 2010, 2012, 2018). To build observing networks that can fill the gap for both operational NWP and research purposes, the capabilities that water vapor observing instruments should possess include the following: (i) unattended continuous operations, (ii) accurate measurements with high vertical and temporal resolution, and (iii) low cost of purchasing and maintaining instruments (Crook 1996; Weckwerth et al. 1999; Weckwerth 2000; Lin et al. 2011; Turner and Löhnert 2014; Wulfmeyer et al. 2015; Weckwerth et al. 2016).

Weckwerth et al. (2016) summarizes detailed characteristics of water vapor profiling instruments such as radiosonde, microwave radiometer profiler (MWRP), atmospheric emitted radiance interferometer (AERI), and Raman lidar that that have been used for operational forecasting or atmospheric research. Unfortunately, the existing network of observing instruments have difficulty in achieving all of the measurement requirements simultaneously. For example, radiosondes typically require manual launching and are relatively costly. Furthermore, due to the low spatiotemporal resolution, radiosonde observations do not typically represent the mesoscale environment needed to improve forecasts of convective events. Some of the instruments also require periodic calibration which

---

ⓐ Denotes content that is immediately available upon publication as open access.

---

*Corresponding author*: Junkyung Kay, junkyung@ucar.edu

presents a challenge to achieving unattended operation (MWRP, Raman lidars) or they are limited by a coarse vertical resolution [MWRP, AERI above ~1 km above ground level (AGL)].

As a complement to the existing water vapor profiling instruments, the National Center for Atmospheric Research (NCAR) and Montana State University (MSU) have developed an advanced lidar-based remote sensing system for profiling water vapor in the lower atmosphere, called micropulse differential absorption lidar (MPD). The MPD satisfies the aforementioned requirements in a strategic way by utilizing low-cost, high-reliability lasers in a micropulse system with the differential absorption technique (Spuler et al. 2015). The MPD shows significant consistency with radiosondes and better resolves the vertical structure of moisture compared to AERI and MWRP (Weckwerth et al. 2016) although AERI and MWRP have the advantage of measuring temperature profiles in addition to water vapor profiles (Knuteson et al. 2004a,b; Turner and Löhnert 2014). In particular, the MPD can measure low-level moisture profiles with a uniform resolution up to 3.5 and 4.5 km AGL at day and night, respectively, while the vertical resolution of AERI decreases rapidly with altitude above 1-km height (Weckwerth et al. 2016). This advantage of MPD enhances the detection capabilities of elevated moist layers which play an important role for nocturnal convection initiation and evolution. The MPD does not require attended operation for continuous calibration and that is a significant benefit for operational use compared to Raman lidar.

Several previous studies have demonstrated that water vapor profiling data can improve short-term convective-scale prediction of various features (Wulfmeyer et al. 2006; Kawabata et al. 2007; Hitchcock et al. 2016; Degelia et al. 2019; Thundathil et al. 2021). For example, Hu et al. (2019) investigated the impact of assimilating thermodynamic profiles of AERI and horizontal wind retrievals of Doppler wind lidar (DWL), and found that AERI thermodynamic profiles had more positive impact on short-term probabilistic forecasts of convection initiation and the early evolution of supercell thunderstorms than the DWL for that case. Coniglio et al. (2019) showed that assimilation of DWL wind retrievals and AERI temperature and moisture profiles from a single profiling system often led to improvement in 1–6-h precipitation forecasts in terms of a neighborhood ensemble probability over the verification domain. Degelia et al. (2019) evaluated the impact of four AERI water vapor profiles deployed approximately 100 km apart during the Plains Elevated Convection At Night (PECAN; Geerts et al. 2017) field campaign and showed a positive impact on the analysis and forecast of deep moist convection for a nocturnal convection initiation case. Following the findings of Degelia et al. (2019), Degelia et al. (2020) expanded AERI DA experiments using 13 nocturnal convection initiation cases during the PECAN field campaign and confirmed systematic improvements in terms of precipitation and convection initiation forecasts and related various contingency metrics by assimilating AERIs. Chipilski et al. (2020, 2022) found convective-scale prediction benefits from assimilating AERI thermodynamic profiles at a high temporal frequency. Thundathil et al. (2021) assimilated water vapor profiles from

differential absorption lidar and temperature profiles from Raman lidar from a single site using the hybrid ensemble DA system and showed improvement in planetary boundary layer height, as well as thermodynamic fields, for a clear sky case compared to the result from the three-dimensional variational DA system (3DVAR). Despite these successes, considerable uncertainty remains on the optimal use of water vapor profiling systems when deployed in a nationwide network, similar to the operational radiosonde network. In addition, the fixed small number of observing sites in previous studies limits finding an optimal network design to maximize the value of thermodynamic profiling observations.

In this study, based on the encouraging capability of the MPD to continuously measure the time-height characteristics of water vapor profiles in the lower atmosphere, we first explore the potential impact of a network of MPD profilers on analyses and predictions of a convective weather event via observing system simulation experiments (OSSEs). The objective of OSSEs is to estimate the unknown potential impact of new observations on weather and climate in addition to the existing observing systems (Masutani et al. 2007, 2010). OSSEs have been applied to numerous convective scale prediction and assimilation studies (Hartung et al. 2011; Zhang et al. 2016; Hu et al. 2017; Liu et al. 2019). Because radiosondes provide profiling data that are routinely assimilated by operational weather prediction centers, we will evaluate the impact of MPDs being deployed at the current operational radiosonde sites and verify that MPD can complement the radiosondes. Furthermore, we will investigate spacing considerations for an MPD network configuration to demonstrate the potential of high-quality MPD measurements as a part of a nationwide network of thermodynamic profiling instruments.

The paper is organized as follows. Section 2 describes the MPD observation system in detail. In section 3, we describe the OSSE framework including the model configuration, the DA method, and the experimental configuration. Section 4 describes the OSSE results for a PECAN nocturnal convection system, and section 5 shows the OSSE results for various MPD network designs. Section 6 compares OSSE DA and real MPD DA results from a single observing site. Section 7 summarizes the results and discusses future work.

## 2. Description of the MPD

The MPD is a compact, lidar-based, eye-safe (class 1M), water vapor active remote sensing system that utilizes a commercially available micropulse laser (Spuler et al. 2015, 2021). The diode-laser based transmitter of the MPD instrument allows for a low-cost and continuous unattended operation, which differentiates MPD from existing water vapor profiling instruments. The transmitter of MPD generates pulses of two tunable wavelengths around 828 nm (Spuler et al. 2015). The online wavelength is located near the center of the water vapor absorption line, and the offline wavelength is chosen to be close in wavelength to the online (within 0.1 nm) but where it is minimally affected by water vapor. The ratio of returned online and offline signals and the timing of the return pulses give range-resolved water vapor transmission which can be
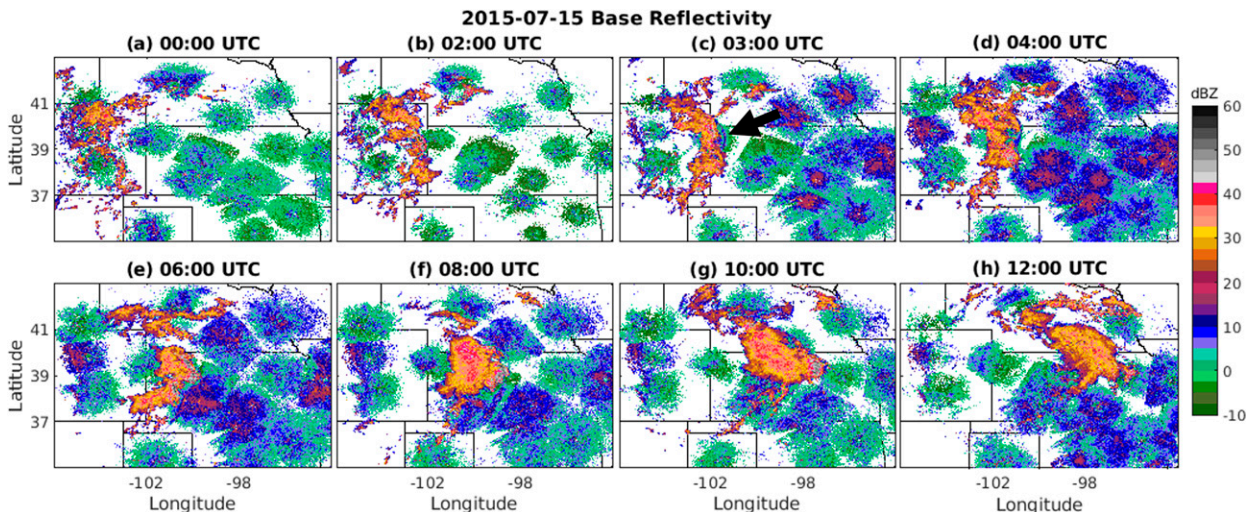
FIG. 1. (a)–(h) Evolution of the 15 Jul 2015 nocturnal convection event from composite reflectivity (dB$Z$). The black arrow in (c) indicates an outflow boundary (Available online at https://doi.org/10.5065/D6QR4VHM.)

solved for water vapor number density or absolute humidity. The elapsed time between transmitting and receiving back-scattered laser pulses determines the vertical range resolution of the MPD. The switching frequency between online and offline wavelengths determines the capability of the MPD to resolve the variability of the atmospheric water vapor in the vertical. The vertical range of MPD is primarily affected by the collection efficiency of the receiver that is a function of field-of-view and collection area (Spuler et al. 2015). In this OSSE study, the water vapor profiling of the MPD ranges from 300 m to 4 km AGL. This range is based on an intercomparison test between radiosondes and MPDs (Spuler et al. 2015; Weckwerth et al. 2016; Stillwell et al. 2020). Typical processed MPD data has a 150-m vertical resolution and 5-min temporal resolution. During daytime and cloudy conditions, the vertical range and resolution can be degraded because the high background signal can saturate the photon-counting receiver leading to large errors in retrievals of water vapor profiles (Spuler et al. 2015). However, the current OSSE uses the same MPD configuration for daytime, cloudy, nighttime, and clear sky conditions as a first trial to evaluate the MPD observation impact. The detailed design and specifications of the NCAR water vapor MPD instrument are described in Spuler et al. (2021).

One MPD instrument was autonomously operated at the PECAN field campaign from 1 June to 15 July 2015. The MPD provided continuous water vapor profiling data for 95% of the entire experimental period without performance loss. Recent efforts have been under way to develop additional MPD sensors, to now have a network of five instruments while extending the MPD architecture to incorporate aerosol properties and temperature profiling (Hayman and Spuler 2017; Stillwell et al. 2020). In this study, we consider a simulated network of MPDs that are hypothetically collocated with rawinsonde sites. In addition, OSSE experiments are conducted to determine the optimal density and spacing of

MPD sites to support a nationwide network of MPD sensors to improve QPF skill.

## 3. Description of the OSSE framework

### a. Case description: 15 July 2015

This study focuses on a nocturnal convection event that occurred on 15 July 2015 during the PECAN field campaign. Nocturnal elevated convection is less skillfully predicted than surface-based daytime convection; this is partially attributed to challenges with predicting elevated nocturnal convection initiation (Stelten and Gallus 2017; Johnson et al. 2018). Part of the issue is the lack of vertically resolved observations of moisture in the lower atmosphere and difficulties detecting sources of convective updrafts above the nocturnal stable surface layer at night (Weckwerth et al. 2019). In this study, we assess the impact of assimilating MPD observations on the prediction of a nocturnal mesoscale convective system (MCS). Here we focus on determining how an improved representation of the lower- to midlevel water vapor improves the prediction of elevated nocturnal convection initiation and its subsequent upscale growth into a high-impact MCS event.

Figure 1 shows the evolution of the nocturnal convection event using a composite radar reflectivity mosaic developed at NCAR (Dixon 2016). Convective cells developed after sunset in eastern Colorado on 14 July 2015 (not shown in figure). The cells organized into an MCS in eastern Colorado around 0200 UTC 15 July 2015 and propagated eastward (Fig. 1b). An outflow boundary was observed ahead of the MCS in northwestern Kansas at 0300 UTC (Fig. 1c). As the evening advanced, nocturnal convection developed along the outflow boundary of the MCS and merged with the MCS after 0300 UTC 15 July 2015 (Fig. 1). New convective cells developed along the outflow boundary after 0300 UTC and along the southern convergence zones, possibly caused by a bore, after 0600 UTC. The MCS propagated northeastward during its
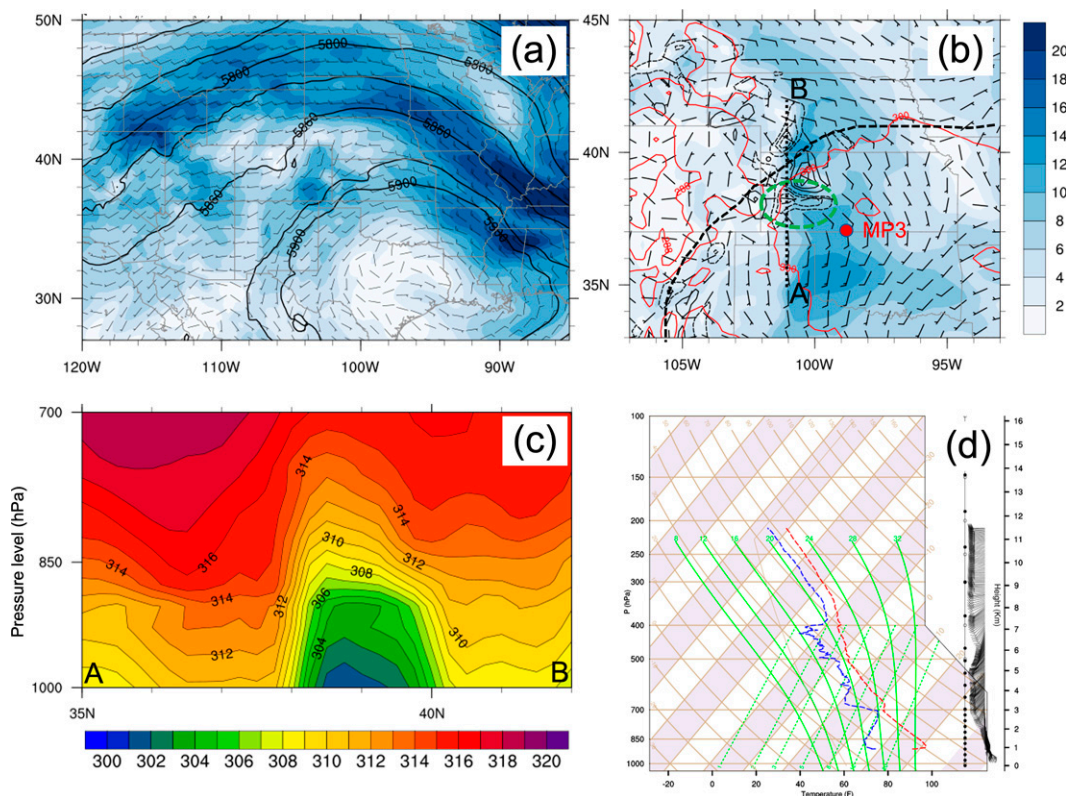
FIG. 2. ERA5 reanalysis of (a) 500-hPa geopotential height (m; black solid contours) and winds (m s$^{-1}$; color fill and barbs); (b) 875-hPa winds (m s$^{-1}$; color fill and barbs), 2-m temperature (K; red contours), and divergence (s$^{-1}$; positive: black contours; negative: black dashed contours); (c) height–latitude cross section of potential temperature (K; color fill) along the black dotted line A–B in (b); and (d) radiosonde sounding at MP3 valid at 0300 UTC 15 Jul 2015. The thick dashed black contour in (b) represents the location of a synoptic boundary, the dashed green circle in (b) indicates the approximate LLJ terminus, and the red dot in (b) indicates the MP3 site.

upscale growth and reached maximum intensity and size in western Kansas and southwestern Nebraska at 0800–1000 UTC (Figs. 1f,g). Large-scale features and weather conditions were favorable for convective development after sunset on 14 July 2015. As a ridge amplified over the Great Plains at 500 hPa at 0300 UTC 15 July (Fig. 2a), a southerly low-level jet (LLJ) developed across western Oklahoma and Kansas (Fig. 2b). The LLJ interacted with the synoptic boundary over southern Nebraska (shear line) to develop large-scale isentropic ascent over southern Kansas (Figs. 2a–c). The northern terminus of the LLJ enhanced mesoscale convergence in southwestern Kansas, and a sharp drop of temperature was found west of the convergence areas (Fig. 2c). Radiosonde sounding at a mobile PECAN integrated sounding array (PISA) (MP) site, ahead of the leading edge of the MCS, showed significant instability and a most unstable convective available potential energy (MUCAPE) of 1800 J Kg$^{-1}$at 700 hPa (Fig. 2d), indicating the possibility of elevated convection initiation. In addition, detailed analysis of this case (Grasmick et al. 2018) suggested a potential key role of the convergence zones in the southern portion of the MCS to maintain the convective system (Fig. 1f).

### b. Description of the model and data assimilation system (WRF-DART)

In this study, we run the Advanced Research version of the Weather Research and Forecasting (WRF-ARW) Model version 4.2.1 (Skamarock et al. 2008; Powers et al. 2017) to simulate nocturnal convection for the OSSEs. Three domains of dimensions 212 × 160 (D1), 411 × 321 (D2), and 744 × 655 (D3) are used for the simulation with horizontal grid spacings of 15, 3, and 1 km, respectively (Fig. 3). All domains have 71 vertical levels with the model top at 50 hPa with finer vertical grid spacing within the lower troposphere (Fig. 3b). The summary of model configurations for the nature run and the DA runs are presented in Table 1.

For DA, the ensemble adjustment Kalman filter (EAKF; Anderson 2001) system based on the DA Research Testbed (DART; Anderson et al. 2009) was used to perform the OSSEs on the 15-km (D1) and 3-km domain (D2). We did not assimilate observations on the 1-km domain (D3). This system has been applied to numerous convective-scale DA studies (Kerr et al. 2019; Keclik et al. 2017; Pan et al. 2018; Schwartz et al. 2015a, 2019). Ensemble Kalman filter–based DA methods generally employ both covariance localization

## (a) Observation distribution



## (b) Geopotential height
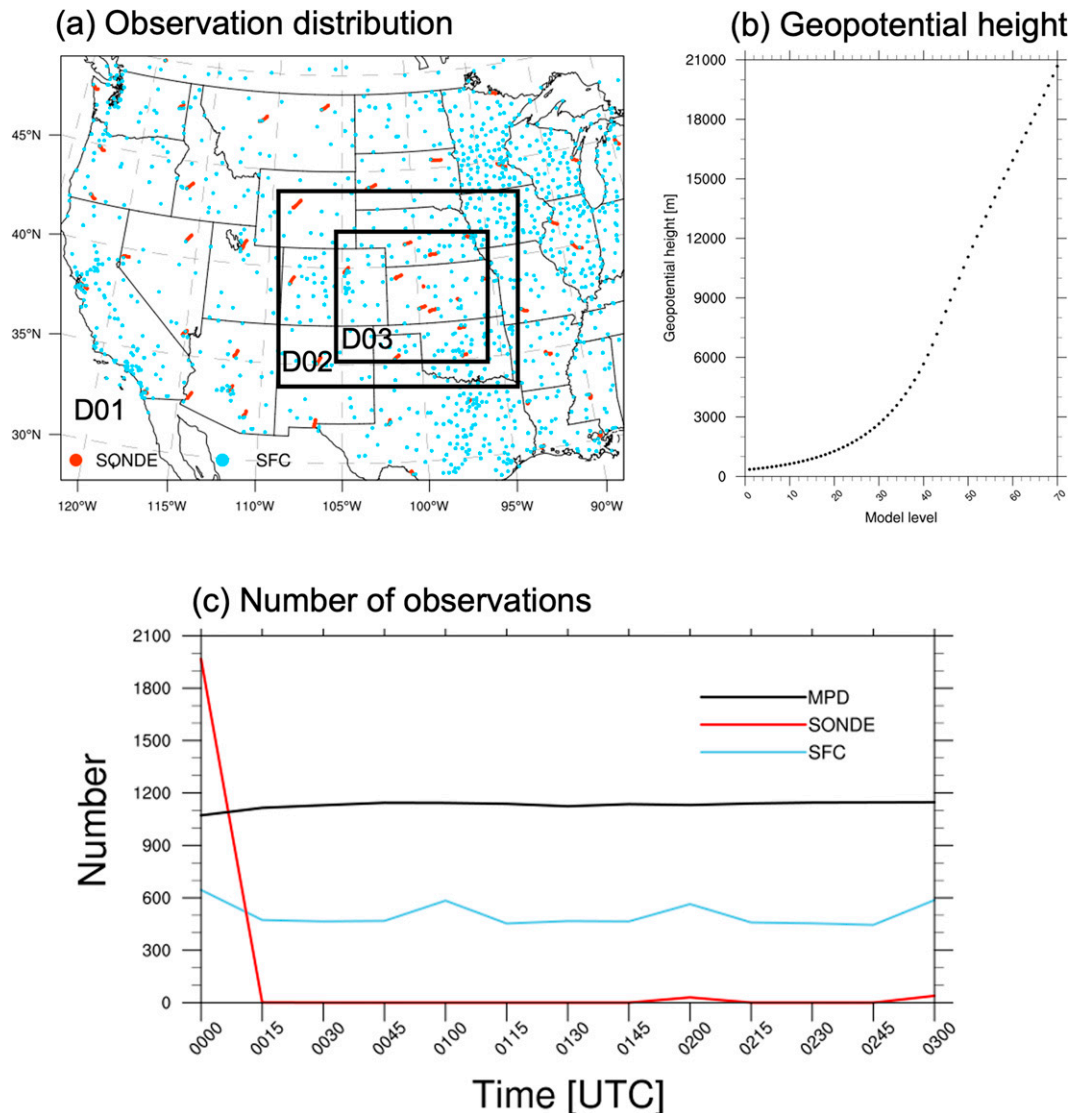


## (c) Number of observations



FIG. 3. (a) Nested model domains and horizontal distribution of conventional observations, (b) geopotential height (m) against vertical model levels, and (c) time series of total number of observations summed up over each DA cycle.

and inflation techniques owing to sampling error due to a limited ensemble size and to account for misrepresentation of model error, as these aspects can otherwise degrade the quality of the analysis and subsequent forecast. In this study we applied a covariance localization technique to all assimilated observations using the Gaspari and Cohn (1999) function to reduce the impact from spurious correlations. The localization technique tapers the analysis increment to zero if the distance between a grid point and an observation is larger than a specified threshold; here we used 300 km in the horizontal distance and ~2 km in the vertical distance for both the outer (D1) and inner domain (D2). The horizontal and vertical cutoff radii of localization were determined by experimentation to minimize the analysis error averaged over the first 8 cycles for the outer domain (D1; results not shown). These localization radii are comparable with previous convective scale DA studies (e.g.,

Degelia et al. 2019; Coniglio et al. 2019). To maintain comparable values of ensemble spread and the error of the ensemble mean, a three-dimensional multiplicative inflation factor is applied to the DART EAKF system. The inflation relaxes the posterior ensemble spread to be 85% of the priori ensemble spread (Anderson and Anderson 1999; Whitaker and Hamill 2012). Surface and MPD observations are assumed to be available every 15 min, rawinsonde observations are assumed to only be available at 0000 UTC (Fig. 3c).

### c. Generation of the nature run

In an OSSE study, a nature run is a numerical simulation that is assumed to be the "true" state of the atmosphere (Zeng et al. 2020). The nature run in this study was selected as it best resembled the targeted phenomena in terms of the initiation and evolution of the convective system described in

TABLE 1. Experiment configuration for the nature run, free run, and DA runs.

| | Nature run | Free run | DA runs |
|---|---|---|---|
| Domain | 15/3/1 km | 15/3 km | 15/3 km |
| ICs/BCs | ECMWF ERA5 | ECMWF ERA5 | ECMWF ERA5/perturbed using statistic forecast error covariance |
| Physics parameterization | | | |
|   Microphysical parameterization | Thompson microphysics (Thompson et al. 2008) | Thompson | Thompson |
|   Cumulus parameterization | Kain–Fritsch (Kain and Fritsch 1990, 1993; Kain 2004; only for outer domain) | Kain–Fritsch (D1 only) | Kain–Fritsch (D1 only) |
|   PBL parameterization | Mellor–Yamada–Janjić (MYJ) (Janjić 2002) | MYJ | MYJ |
|   Land surface model parameterization | Noah land surface model (Chen and Dudhia 2001) | Noah | Noah |
|   Longwave radiation parameterization | Rapid Radiative Transfer Model for Global Climate Models (RRTMG; Mlawer et al. 1997; Iacono et al. 2008) | RRTMG | RRTMG |
|   Shortwave radiation parameterization | RRTMG | RRTMG | RRTMG |
| Data assimilation | No | No | EAKF/conventional observations (surface, rawinsonde), MPD |

section 3a. Synthetic observations are then drawn from the nature run and are provided to the DA system. To generate the nature run, we conducted ensemble predictions using randomly perturbed initial conditions and boundary conditions to generate a 40-member ensemble at 1-km horizontal resolution. The configuration of the numerical simulation is described in section 3b and summarized in Table 1. The initial and lateral boundary perturbations were obtained from a multivariate normal distribution whose covariance is the pregenerated static forecast error covariance of the WRF 3DVAR (Barker et al. 2004; Torn et al. 2006; Wang et al. 2008). These random perturbations are added to the European Centre for Medium-Range Weather Forecasts (ECMWF) fifth-generation climate reanalysis (ERA5) to generate initial conditions (ICs) and lateral boundary conditions (LBCs) for the 15-km ensemble predictions. These ICs and LBCs are downscaled to the inner domains (D2 and D3) for the 3- and 1-km ensemble predictions, similar to Schwartz et al. (2015b). The ensemble prediction is initialized at 1200 UTC 14 July 2015 and integrated for 24 h valid at 1200 UTC 15 July 2015. We first evaluated the 40 members of the 1-km ensemble prediction according to objective and subjective criteria, then selected the best-performing member as the nature run. We also selected the worst-performing 3-km member to improve by assimilating the simulated observations (hereafter referred to as free run).

To meet an objective criterion for the selection, we applied the method for object-based diagnostic evaluation (MODE; Davis et al. 2006a,b, 2009) to the radar reflectivity of ensemble member predictions. MODE is a spatial verification method that can be applied to identify and match forecast and observed objects. When object pairs from the prediction field

and the observed field are identified, a fuzzy logic algorithm computes statistics of the similarities of the objects in the two datasets (Davis et al. 2006a,b, 2009). The MODE method can also be applied to select meteorological fields that have coherent spatial structure such as precipitation, clouds, and composite reflectivity (Davis et al. 2006a,b, 2009; Ahijevych et al. 2009; Clark et al. 2016). In this study, the process of selecting the nature run and free run by applying MODE is as follows: 1) the degree of similarity between the observed composite radar reflectivity field with each ensemble member prediction reflectivity field is assessed in terms of the intensity, shape, location, and timing of the convective system from 0000 UTC to 1200 UTC July 2015 every 1 h; and 2) based on a time-averaged similarity score from the MODE results, the five members with the highest scores (hereafter referred to as "good members") and five members with the lowest scores (hereafter referred to as "bad members") were selected. Figure 4 shows the time-averaged MODE similarity score of ensemble members. 3) As a first attempt to explore the impact of MPD on convective weather predictions, the convection initiation of the target event is the major concern for this study. Therefore, among the "good members," the one member that simulates the best timing of the convection initiation was subjectively selected as the nature run. Likewise, the member that was least able to represent the observed timing of the convection initiation and forecast evolution of the convective system was selected as the free run from among the "bad members." A comparison of the real observations, nature run, and free run illustrating the evolution of the MCS is shown in Fig. 5.

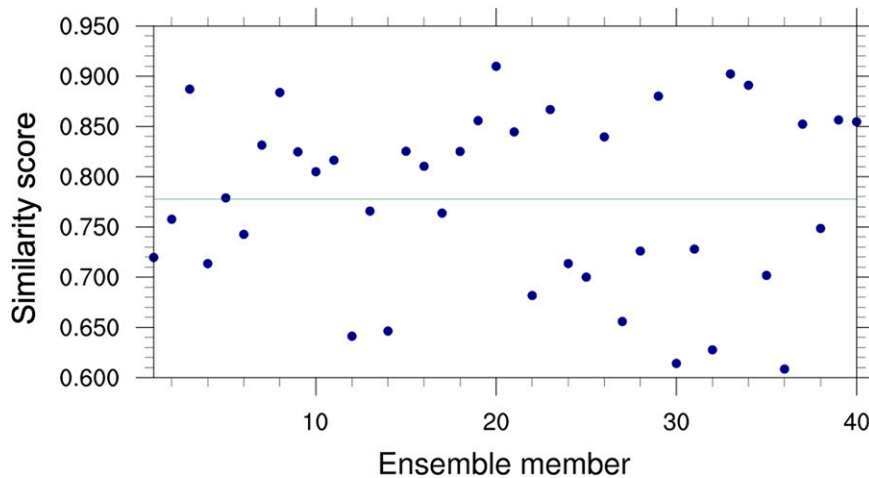The nature run should have similarities with actual atmospheric states regarding the evolution of the MCS and the

FIG. 4. The temporal-averaged values of the similarity score from MODE for each ensemble member between 0000 UTC 15 Jul and 1200 UTC 15 Jul 2015. The green line represents the averaged MODE for all the 40 ensemble members.

surrounding environment (Hoffman and Atlas 2016). Therefore, we have conducted additional verifications for the nature run. Figure 6a is the 875-hPa winds and convergence fields from the nature run similar to Fig. 2b. The nature run shows consistent environmental features with the ERA5 reanalysis, representing a well-developed synoptic boundary and the LLJ terminus in northern Oklahoma which is located south of the ERA5 reanalysis LLJ terminus location (Fig. 2b).

To evaluate the temporal evolution of the features related to the MCS from the nature run, we compare the time series of temperature and water vapor mixing ratio from the nature run and AERI profiling data at the PECAN fixed PISA station 3 (PS3) at Ellis, Kansas (black dot in Fig. 6a). Although there are detailed differences between the nature run and AERI, the nature run accurately represents the diurnal evolution of the thermodynamic states compared to the AERI profiles. In particular, the timing, magnitude and vertical extent
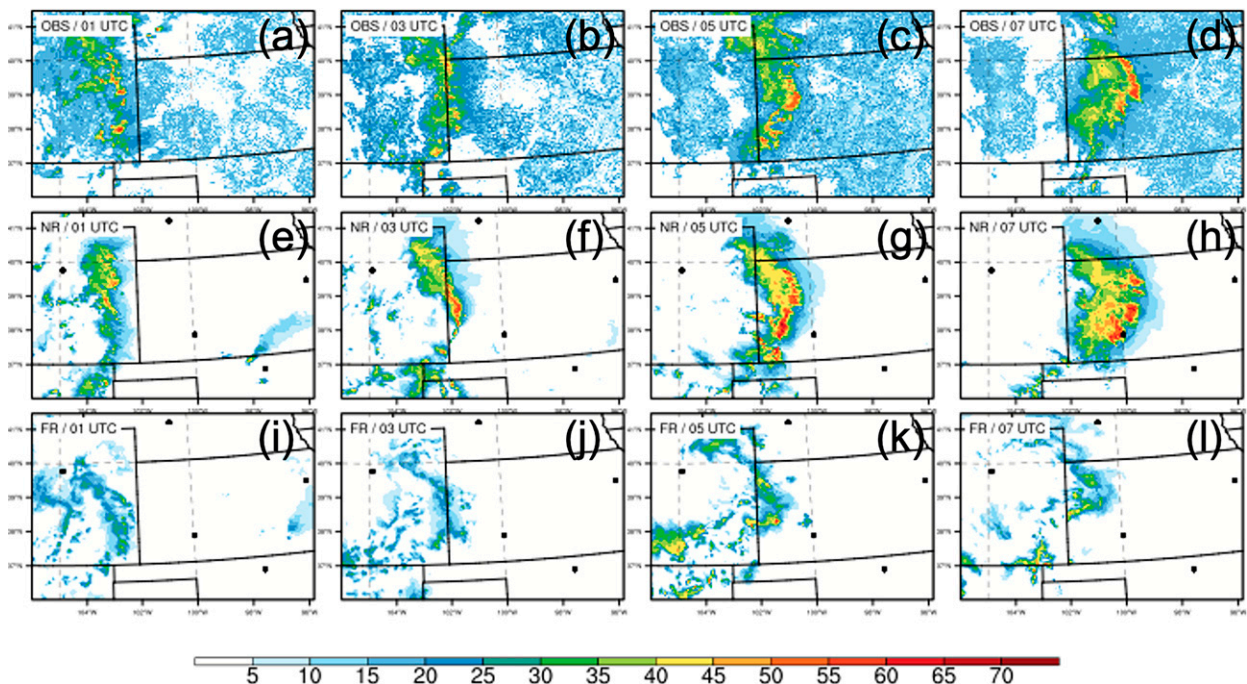


FIG. 5. Composite reflectivity (dBZ; color fill) of the nocturnal convection event from 0100 to 0700 UTC 15 Jul 2015 with 2-h intervals of (a)–(d) radar observations, (e)–(h) nature run, and (i)–(l) free run. The black dots represent the location of radiosonde sites.
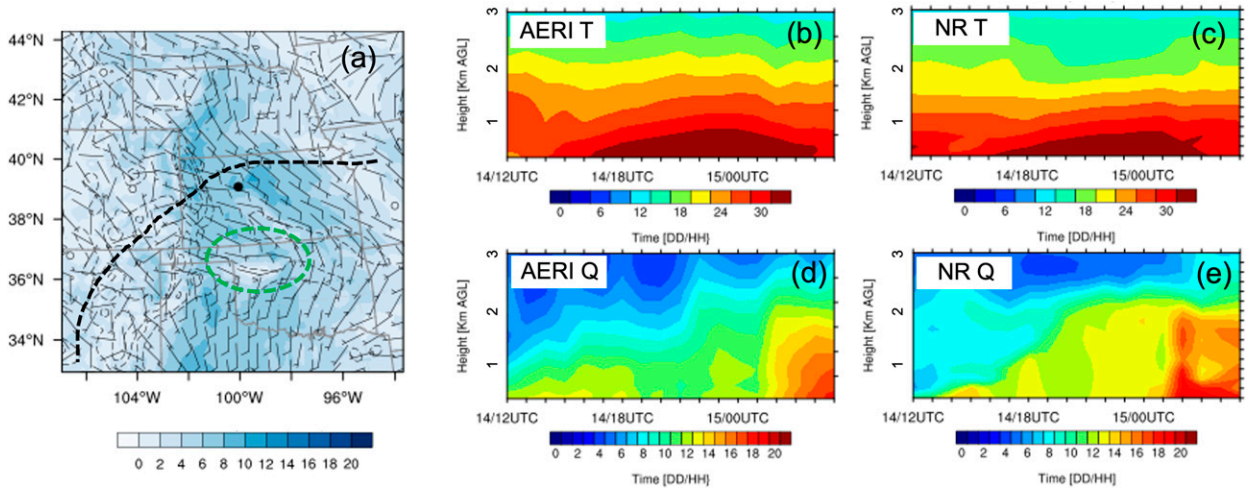
FIG. 6. (a) 875-hPa winds (m s$^{-1}$; color fill and barbs) and convergence (s$^{-1}$; black dashed contour) of nature run valid at 0300 UTC 15 Jul 2015, and time series of temperature from (b) AERI and (c) nature run, and water vapor mixing ratio from (d) AERI and (e) nature run from 1200 UTC 14 Jul to 0600 UTC 15 Jul 2015. The thick dashed black contour in (a) represents the location of a synoptic boundary and the dashed green circle indicates the approximate LLJ terminus. The black dot represents the location of PECAN fixed PISA station 3 at Ellis, KS, where the time series of the observations and nature run are calculated in (b)–(e).

of the heating associated with the growing daytime convective boundary layer are consistent between AERI and the nature run. Furthermore, the depth of the water vapor reaching ∼2 km after 0300 UTC is consistent, demonstrating the accuracy of the nature run to simulate the preconditioning of the MCS.

Previous studies have shown that increasing the resolution of numerical simulations is a critical factor to better represent convection initiation (Schumacher 2015; Tang and Kirshbaum 2020). To consider the model error associated with model resolution, we use higher resolution for the nature run (1 km) than the free run and DA runs (3 km). Despite this change in resolution, the nature run and DA runs have the same configuration of physical parameterizations (Table 1), resulting in near identical twin problems of OSSE (Privé and Errico 2013). Therefore, it is noted that the positive impact of assimilating the simulated MPD profiles in this OSSE study can be exaggerated compared to the impact of assimilating real MPD observations.

### d. Simulated observations and their errors

Simulated MPD water vapor profiling observations and conventional observations including rawinsonde data and surface station data are assimilated in this study. To obtain the synthetic observations for the OSSE, fields of observable variables in accordance with these instruments are extracted from the nature run, and random noise is then added to each simulated observation based on the magnitude of typical real-data observation errors. The random noise errors are drawn from a Gaussian distribution with zero mean and standard deviation equal to the observational errors that are further described below.

To synthesize absolute humidity profiling data of MPD, the set of WRF state variables of the nature run are used following Eqs. (1) and (2):

$$\rho_d = \frac{P}{R_d T \left( 1 + \frac{r}{R_d/R_v} \right)}, \qquad (1)$$

$$\rho_v = r\rho_d, \qquad (2)$$

where $\rho_v$ is the absolute humidity, $\rho_d$ is the density of dry air, $r$ is the water vapor mixing ratio, $T$ is temperature (K), $P$ is pressure (Pa), and $R_d/R_v$ is a gas constant for dry/moisture air. The MPD water vapor profiling data are assumed to be collocated with operational radiosonde sites during the experimental period (Fig. 3a). Because the radiosonde provides high vertical resolution water vapor profiling data that is routinely assimilated by operational NWP in this region of the United States, a comparison of radiosonde and MPD deployed at the same sites will reveal the benefits of the MPD as a potential candidate for continuously sampling water vapor profiling observation networks. As described in section 2, the MPD water vapor profiles assimilated in the study extend from 300 m to 4 km AGL with 150-m vertical resolution and a 5-min temporal resolution. It is noted that new hardware improvements have been made for the MPD to extend the lowest range down to 225 m AGL (Spuler et al. 2021).

In DA, the specified observation error determines the relative weight between an observation and background state, so estimation of an appropriate observation error and its covariance matrix is a critical factor in the successful assimilation of new observations. The observation error used in a DA system consists of an instrument error and representativeness error, where instrument error is typically much smaller than the representativeness error magnitude (Geer and Bauer 2011; Janjić et al. 2017). The algorithm of the MPD photon-counting system provides an instrument error profile for each observing
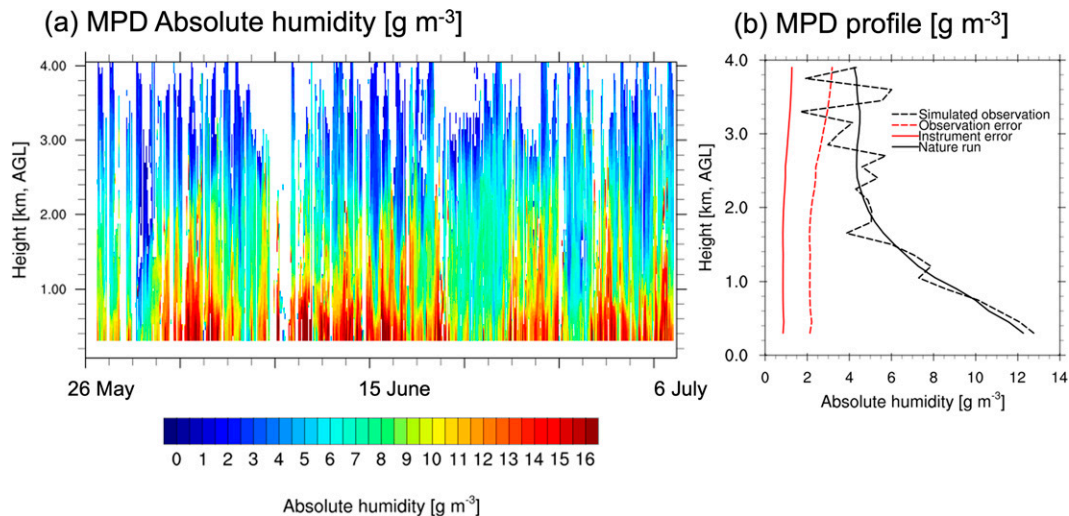
FIG. 7. (a) Time–height plot for actual MPD data (g m$^{-3}$) during the PECAN field campaign and (b) nature run of absolute humidity (g m$^{-3}$; black solid contour), simulated MPD absolute humidity observation (g m$^{-3}$; black dotted contour), standard deviation of the MPD instrument error (g m$^{-3}$; red solid contour), and standard deviation of the MPD observation error (g m$^{-3}$; red dotted contour) profile averaged for the duration of the PECAN field campaign.

time (Spuler et al. 2015). The error is a function of instrument parameters, atmospheric parameters, overlap function for receiver channels of the instrument, backscatter coefficients and water vapor number density. The temporal average of the instrument error profiles during the PECAN experiment is shown in Fig. 7b. The increasing error with height is related to strong backscatter gradients and low signal-to-noise ratio, while MPD observation confidence is also reduced near optically thick clouds in the analysis. The representativeness error component of the observation error is accounted for by multiplying an inflation factor to the instrument error, and the net observation error is calculated by adding the instrument error and representativeness error, similar to the estimation of the AERI observation error (Degelia et al. 2019). The only difference relative to Degelia et al. (2019) is that here we use a constant inflation factor with height rather than a linearly increasing inflation factor with height. Through a trial-and-error process varying the inflation factor from 1.0 to 3.0, we choose an inflation factor of 1.5 which was found to result in the best performance with this set of DA experiments and associated predictions. The resulting total observation error is used to generate the simulated observations and assimilate the simulated observations. Future studies will explore a more sophisticated method of estimating the MPD water vapor profiling observation error using time-averaged comparisons of observations and corresponding model backgrounds collected over a significantly longer period following several previous studies (Hollingsworth and Lönnberg 1986; Desroziers et al. 2005; Bormann and Bauer 2010).

For the conventional observations, we choose the same location, spatiotemporal resolution, and variables of the conventional observations from the real global telecommunication system (GTS) data. The distribution and number of conventional observations assimilated in this study are shown in Fig. 3c. Note

that the DA system also takes into account radiosonde drift during ascent. Because the conventional observations are routinely assimilated in operational DA systems, such as with Gridpoint Statistical Interpolation (GSI), we similarly assimilate simulated conventional observations using predefined observation error statistics provided by the GSI software (Shao et al. 2016; Hu et al. 2018).

*e. Data assimilation experimental design*

For the DA experiments, ICs and LBCs for ensemble predictions are generated by adding 50 initial ensemble perturbations to the free run at the beginning of the DA cycles valid at 1200 UTC 14 July 2015. The initial and lateral boundary ensemble perturbations for the outer domain (D1) are sampled from the static forecast error covariance available as part of the WRF 3DVAR system in the same manner as described in section 3b. These 15-km perturbed state fields were downscaled to the one-way nested 3-km inner domain to create ensemble perturbations of ICs for the inner domain (D2) as well. The initial ensemble conditions are then integrated for 3 h, and the 3-h ensemble prediction provides the background ensemble for DA of the conventional observations including surface station data and rawinsonde data starting from 1500 UTC 14 July 2015 to 0000 UTC 15 July 2015 with 3-hourly intervals on both D1 and D2 (Fig. 8). Rawinsonde and surface observations are assimilated within 1.5 h of the analysis time for 3-hourly DA cycles. After 0000 UTC 15 July, four sets of DA experiments with 15-min intervals, described in Table 2, are conducted to evaluate the effect of different combinations of observations on the nocturnal convection prediction. Conventional and MPD observations within 7.5 min from the analysis time are assimilated for 15-min DA cycles. DA_sfc experiment assimilates conventional surface station data including surface wind, temperature, specific humidity, and
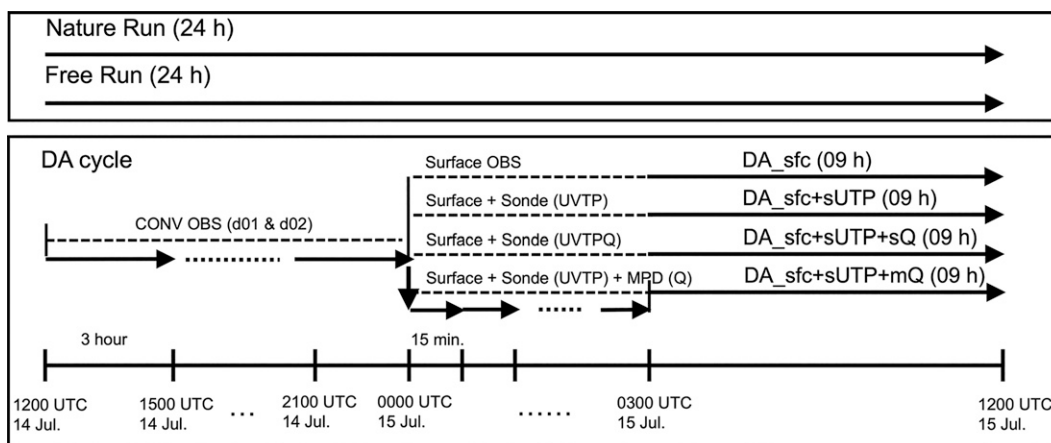
FIG. 8. Schematic illustration of the DA procedures.

pressure observations. DA_sfc+sUTP experiment, where the second lowercase "s" represents rawinsonde sounding profiles, assimilates conventional surface observations and rawinsonde wind, temperature, and pressure profile data. To discern the impact of water vapor profile data on the nocturnal convection prediction, two additional experiments, DA_sfc+sUTP+sQ and DA_sfc+sUTP+mQ, assimilate the rawinsonde specific humidity profile data or the MPD absolute humidity profile data in addition to DA_sfc+sUTP experiment, respectively. Lowercase "m" in DA_sfc+sUTP+mQ represents the MPD instruments. The MPDs are collocated with the operational rawinsondes so the comparison of DA_sfc+sUTP+sQ and DA_sfc+sUTP+mQ serves to facilitate the understanding of the impact of higher spatiotemporal resolution of MPD water vapor profiles compared to rawinsonde data.

## 4. Comparison of the impact of conventional water vapor rawinsonde data versus continuous MPD profiling data

### a. Overview of the analysis and prediction

In this section, we evaluate the overall impact of the MPD water vapor profiles on the analysis and subsequent prediction in terms of precipitation prediction performance relative to the nature run. Figure 9 shows the prediction evolution of reflectivity fields from the nature run and free run and the reflectivity prediction obtained when using the ensemble mean analyses for all DA experiments to initialize all DA

TABLE 2. Experimental configuration for DA runs. Abbreviations: "U" is horizontal wind, "T" is temperature, "P" is pressure, and "Q" is water vapor of either "s" (sounding specific humidity) or "m" (MPD absolute humidity).

| EXP | Surface | Upper air UVTP | Upper air Q |
|---|---|---|---|
| DA_sfc | Yes | No | No |
| DA_sfc+sUTP | Yes | Rawinsonde | No |
| DA_sfc+sUTP+sQ | Yes | Rawinsonde | Rawinsonde |
| DA_sfc+sUTP+mQ | Yes | Rawinsonde | MPD |

experiments at 0300 UTC 15 July. DA experiments assimilating conventional observations show similar features to each other in terms of convection initiation and evolution. All predictions struggle to accurately simulate the convection and subsequent evolution near the Colorado–Kansas border (labeled "extended MCS" in Fig. 9u) during the early prediction lead times (Figs. 9i,m,q). Consequently, the predictions of DA_sfc, DA_sfc+sUTP, and DA_sfc+sUTP+sQ show that the location of the MCS is too far north compared to the nature run after 0600 UTC 15 July (Figs. 9j–l,n–p,r–t). These results imply that assimilation of conventional surface station observations and rawinsonde observations with or without moisture profiles has little impact on the convection initiation prediction across the Colorado–Kansas border region in this case study. On the other hand, assimilating MPD improves the prediction of convection initiation and the development of a more organized convective structure within the southern part of the system compared to the other three DA experiments. Additional convection initiation is predicted in DA_sfc+sUTP+mQ over southwest Kansas between 0400 and 0900 UTC (Figs. 9u–w) that is not evident in the other DA experiments. Thereafter, predictions from DA_sfc+sUTP+mQ qualitatively show improved depiction of the location and structure of the predicted MCS as compared with that obtained in the other three DA experiments (Figs. 9u–x). We will primarily analyze these two convection events improved by the MPD DA; 1) the MCS extension into southwestern Kansas (Fig. 9u) and 2) newly initiated convection cells in southeastern Kansas (Fig. 9v).

Next, the 1-h accumulated precipitation predictions are compared through a time series of fractions skill score (FSS) over the D2 domain (Fig. 10). FSS is a neighborhood-based method for assessing spatial patterns in precipitation predictions (Roberts and Lean 2008). FSS has been widely used to quantify QPF of convection-permitting NWP models (Romine et al. 2013; Wilkinson 2017) because FSS can mitigate for the double-penalty problem in verifying the performance of high-resolution models (Nurmi 2003; Roberts and Lean 2008). Since all experiments show consistent patterns of FSS
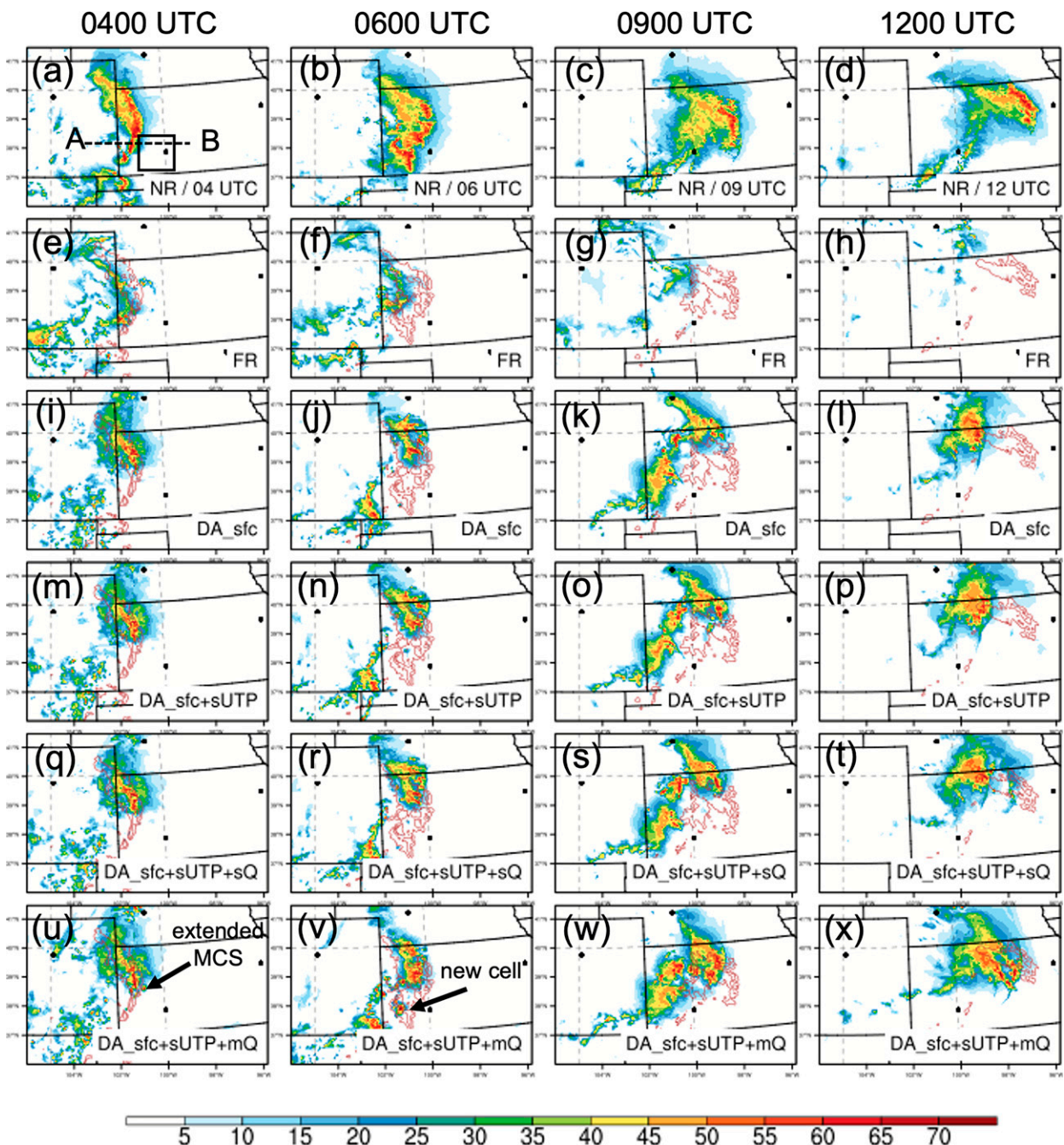
FIG. 9. Predictions of composite reflectivity for (a)–(d) nature run, (e)–(h) free run, (i)–(l) DA_sfc, (m)–(p) DA_sfc+sUTP, (q)–(t) DA_sfc+sUTP+sQ, and (u)–(x) DA_sfc+sUTP+mQ initialized at 0300 UTC 15 Jul 2015. The overlaid solid red lines represent reflectivity of nature run greater than 30 dBZ. The solid black box in (a) indicates the averaging regions for the profiles in Fig. 16. The dashed black line A–B indicates the location for the height–longitude cross section in Fig. 15. The black arrows in (u) and (v) indicate the extended MCS and new convection initiation of interest, respectively.

with neighborhood radii varied across a range from 9 to 48 km (not shown), we hereafter discuss the FSS with a 24-km neighborhood for light (0.5 mm h$^{-1}$) and heavy (5 mm h$^{-1}$) precipitation thresholds. Assimilating MPD water vapor profiles results in the highest skill scores for both light and heavy

precipitation predictions after 2-h prediction lead times. This is subjectively attributed to a more accurate prediction of the southward extension of the MCS and the new convection initiation in southwestern Kansas compared to the other DA experiments and free run (Fig. 9). Figure 10 shows that
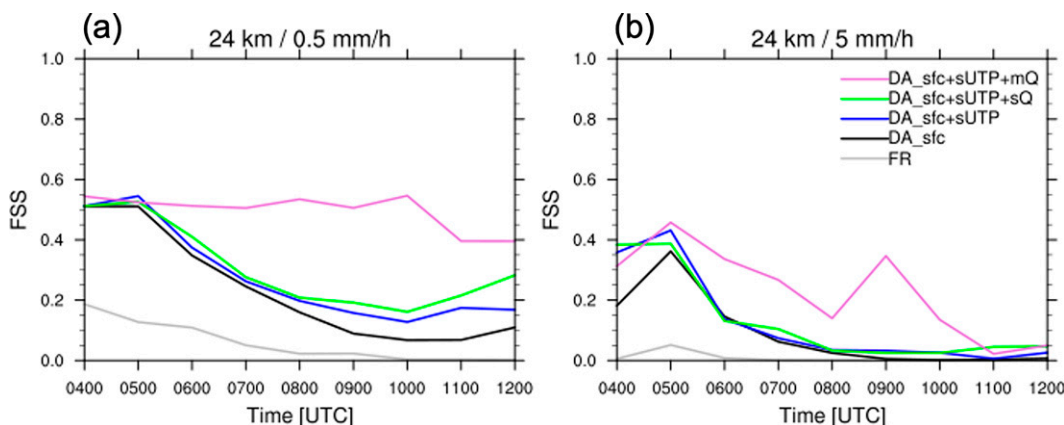
FIG. 10. Fractions skill score calculated with a 24-km neighborhood for rain rates of (a) 0.5 and (b) 5 mm h$^{-1}$.

DA_sfc+sUPT and DA_sfc+sUTP+sQ have similar FSS over the entire forecast lead times, with DA_sfc showing slightly lower scores after 5 h (i.e., 0800 UTC). This assimilation of the rawinsonde profiles improves the longer range precipitation prediction compared to assimilating only surface data (Fig. 10).

### b. Observation impacts on the dynamic and thermodynamic environment

As noted before, none of the DA experiments except for DA_sfc+sUTP+mQ are able to represent the evolution of precipitation associated with the southern part of the MCS over southwestern Kansas (Fig. 9u) and the convection initiation (Fig. 9v). In this section, we investigate how and why the assimilated MPD water vapor profiles led to the improvement of precipitation predictions compared to predictions from the other DA experiments. Previous studies have shown that the key ingredients for the development of deep convection are a sufficient moist layer in the low-to-mid troposphere, conditional instability, and a lifting mechanism to allow parcels to attain free convection (Johns and Doswell 1992; Holley et al. 2014); therefore, we investigate how MPD DA impacts these three contributions to the extended MCS (Fig. 9u) and convection initiation (Fig. 9v) for this case.

Figure 11 shows 850-hPa wind and its divergence analysis fields at 0300 UTC 15 July 2015. As the convective system moves to the east, the nature run illustrates a well-developed boundary layer convergence zone across the leading edge of the MCS (Fig. 11a) which is consistent with Figs. 1c and 2b. Elevated nocturnal convection occurs within a broader region of mesoscale lifting associated with the convergence zone within the nature run (Fig. 9a). The free run displaces a much weaker convergence zone to the east compared to the nature run (Fig. 11b). All DA experiments show similar location and structure of divergence fields with the nature run yet they all exhibit weaker divergence and convergence (Figs. 11c–f). Assimilation of surface station data plays an important role in improving the winds across west Kansas to be more easterly (Figs. 11b,c), but it does not improve the strength of the convergence. Assimilating MPD water vapor profiles modestly

enhances the convergence across west-central Kansas relative to the other DA experiments. The feedback between water vapor profiles by MPD and surrounding dynamic fields may contribute to improving the longer-range convection prediction over southwestern Kansas in the MPD DA experiment compared to other DA experiments (Figs. 9u,v).

Deep convection can be triggered by lifting a volume of air above its level of free convection. Because the lift from horizontal convergence is determined by the vertically integrated convergence, the vertical profiles of the convergence ahead of the cold pool are calculated over the box shown in Fig. 11a (Fig. 12). The nature run has stronger convergence than the other simulations below 2 km mean sea level (MSL) with a maximum around 1 km MSL. It is noted that the strong elevated horizontal convergence is favorable for elevated nocturnal convection. Assimilating a combination of surface station observations and rawinsonde profiles without moisture (DA_sfc+sUTP) leads to stronger convergence compared to assimilating only surface station data (DA_sfc) below 2 km MSL. While the moisture profiles of the rawinsonde (DA_sfc+sUTP+sQ) have little positive observation impact on the convergence profiles except for below 1.2 km MSL compared to DA_sfc+sUTP, MPD water vapor profiles (DA_sfc+sUTP+mQ) improve the convergence field below 1.6 km MSL. Combined with the horizontal fields of the convergence (Fig. 11), the vertical profiles of convergence imply that assimilation of MPD profiles has a positive impact on the lifting ingredient associated with convergence in the lower atmosphere ahead of the MCS but still produces much less convergence than that shown for the nature run.

Figure 13 shows the integrated precipitable water (PW) and 850-hPa wind analysis at 0300 UTC 15 July. The nature run shows a frontal boundary along the border between Oklahoma and Kansas, where the 850-hPa wind direction changes sharply and where there is a maximum in PW. North of the frontal boundary, the nature run exhibits southeasterly winds. These winds play an important role in transporting water vapor from northeastern Oklahoma to western Kansas and supplying water vapor to the convective system (Fig. 13a). On the other hand, the frontal boundary and associated
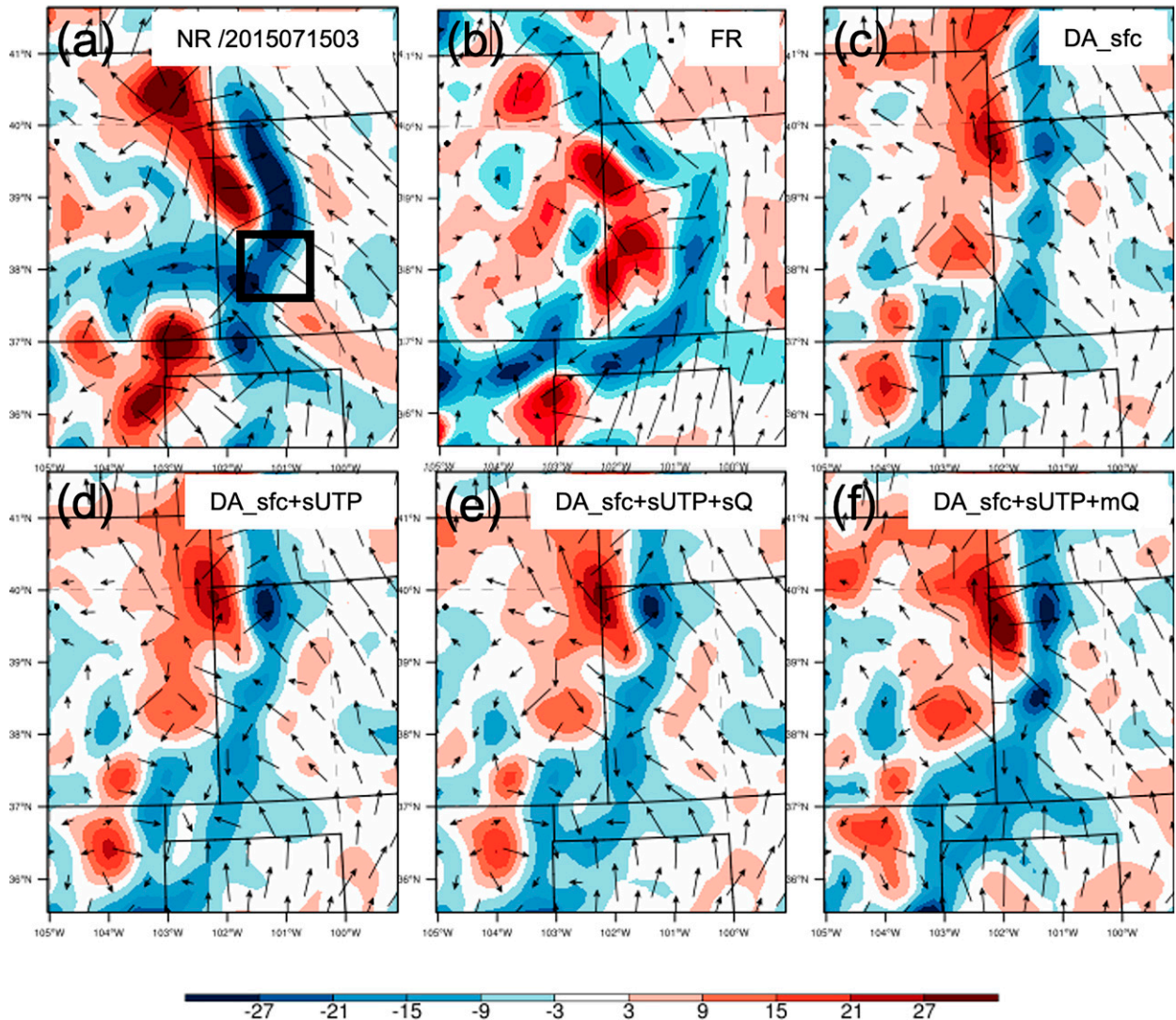
FIG. 11. 850-hPa winds (vector; m s$^{-1}$) and divergence (color fill; $10^{-5}$ s$^{-1}$) for (a) nature run, (b) free run, and ensemble mean analyses of (c) DA_sfc, (d) DA_sfc_sUTP, (e) DA_sfc+sUTP+sQ, and (f) DA_sfc+sUTP+mQ valid at 0300 UTC 15 Jul 2015. The positive and negative values indicate divergence (red) and convergence (blue), respectively. The solid black box indicates the location for the profiles of divergence in Fig. 12.

easterly wind field are not represented by the free run (Fig. 13b). Assimilation of the surface station data improves the easterly component of winds compared to the free run near the frontal boundary (Fig. 13c). The other DA experiments also produce similar wind structures associated with the frontal boundary primarily owing to corrections by assimilating the surface station observations (Figs. 13d–f).

Consistent with the positive contribution from the surface station data on the low-level wind analysis, assimilating surface station data improves the location of the maximum PW region compared to the free run (Figs. 13b,c). However, there remains an underestimation of the amount of vertically integrated water vapor compared to the nature run (Fig. 13c). It is notable that assimilation of rawinsonde profiles added to the surface station data have a minimal effect on the PW

analysis (Figs. 13d,e). Contrary to the rawinsonde results, the additional information provided by assimilating continuous MPD profiles best represents the horizontal structure and the placement and magnitude of the maximum PW (Fig. 13f). Rawinsonde water vapor profiles, which have a higher vertical depth and resolution than MPD, are available just once and at the beginning of the DA cycles (Fig. 3b). Compared to rawinsonde profiles, the improvements by assimilating MPD profiles imply that frequent and continuous updating of water vapor profiles is likely more effective in correcting rapidly changing water vapor fields than updating higher vertical resolution water vapor profiles from the rawinsonde when that data are only available at a single time. Figure 13g shows a time series of water vapor from nature run, free run, and mean analysis of all DA experiments
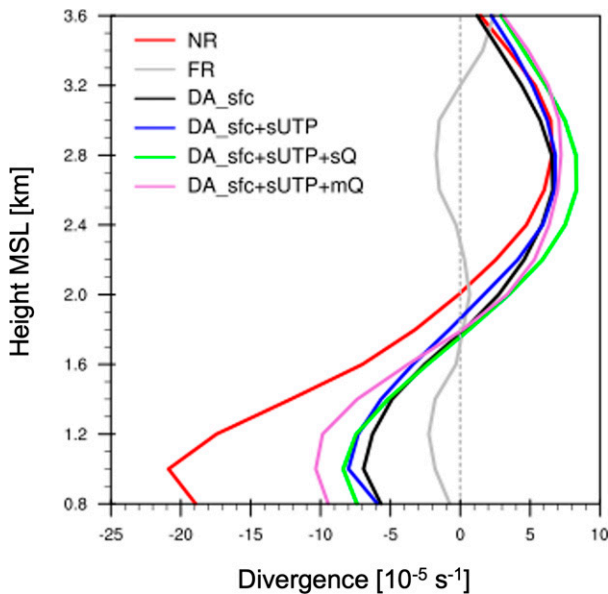
FIG. 12. Profiles of divergence ($10^{-5}$ s$^{-1}$) for nature run, free run, and ensemble mean analysis of each DA experiment valid at 0300 UTC 15 Jul 2015. The dotted gray line indicates zero divergence. The profiles are averaged over the black box in Fig. 11a.

averaged over the D2 domain during DA cycles. The nature run exhibits a continuous increase in vertically integrated moisture (Fig. 13g), possibly due to a strengthening southeasterly LLJ (not shown), over all of the DA cycles across western Kansas and western Oklahoma. This may have contributed to increasing storm-relative inflow and upscale growth of the convective system (e.g., Carroll et al. 2021). The gap between the nature run and DA_sfc+sUTP+mQ analysis decreases owing to frequent cycles with repeated moisture profile observations, while the gaps between the nature run and the other DA experiments remain constant without these frequent moisture profile observations (Fig. 13g). This result supports the potential benefits of frequent assimilation of continuous MPD water vapor profiles on the water vapor analyses and subsequent predictions.

Analyses suggest that there is a synoptic front to the north of the LLJ terminus (Fig. 6a). Convection forms ahead of the leading edge of MCS along the synoptic front to the west of the LLJ terminus (Fig. 5f). Although all DA experiments represent similar surface wind analyses, MPD profiles enhance the vertically integrated water vapor analysis to better match the nature run. To further investigate the impact of MPD water vapor profiles on the thermodynamic structure along the front, zonally averaged height-latitude cross-sections of potential temperature, water vapor mixing ratio and horizontal wind analysis are examined (Fig. 14). The potential temperature of the nature run sharply increases with height north of 36°N (Fig. 14a). The increase in the potential temperature which is associated with warm frontal overrunning develops to the north of the observed LLJ terminus (Fig. 6a) in the nature run (Fig. 14a). The cross-section of the nature run implies that the LLJ transports moist air upward reaching ~1.8 km

AGL and extending further northward along the sloping isentropic surface within the latitude range of 36.3° and 37.3°N, resulting in a favorable environment for the initiation and subsequent evolution of deep, moist convection in the region (Trier and Parsons 1993; Weckwerth et al. 2019). Along the ascent path, water vapor mixing ratio increases by 20 g kg$^{-1}$ from the surface to 1.8 km AGL where easterly flow is prevailing (Fig. 14a). In contrast, the isentropic ascent and related water vapor enhancement are not represented by the free run in the same region (Fig. 14b). Compared to the free run, assimilating surface station data corrects the wind structure and related vertical thermodynamic structure below 1.5 km AGL (Fig. 14c); however, moisture is constrained below 1.8 km AGL which is lower than the 2.2 km AGL in the nature run (Fig. 14c). Assimilating rawinsonde data in addition to the surface station data shows similar thermodynamic features to assimilating surface station data alone (Figs. 14c–e). This result therefore implies that the effect of rawinsonde data, which is assimilated only at the beginning of the DA cycle in this study (Fig. 3b), does not persist as DA cycles continue. Therefore, the rawinsonde profiles are shown to have little impact on the vertical structure of temperature and moisture analysis (Figs. 14d,e). This result is consistent with the time series of PW analysis in Fig. 13g. Conversely, the assimilation of the MPD water vapor profiles increases the vertical depth of the isentropic surface to 2.1 km MSL and results in the transport of moist air further upward compared to the other three DA experiments (Fig. 14f). This DA_sfc+sUTP+mQ result is more consistent with the nature run compared to the other DA experimental results. Furthermore, results with MPD DA also show a positive impact on representing the easterly component of the horizontal wind analysis with increasing depth of the moisture below 2.0 km AGL (Fig. 14f). The easterly winds, which extend to 2.0 km MSL in the nature run, also play an important role in transporting the elevated moist air to the convectively active area ahead of the outflow boundary.

Figure 15 shows vertical cross sections of water vapor mixing ratio from the nature run and difference fields between the three DA experiments and the nature run at 0300 UTC 15 July 2015. The nature run shows elevated moisture at around 2 km MSL in advance of the nocturnal convection (Fig. 15a). Figures 15b–f show the difference fields compared to the nature run. The free run has a prominent dry bias in advance of the region of the elevated convection (Fig. 15b). All DA experiments reduce the dry bias below 2 km compared to the free run (Figs. 15c–f). However, assimilating MPD water vapor profiles minimizes the dry bias more substantially than the DA experiments with other types of observations (Fig. 15f). In contrast to the rawinsonde water vapor profiles (Fig. 15e), assimilating MPD water vapor profiles improves the dry bias in the water vapor analysis through a deeper layer (Fig. 15f). At levels above 3 km, where the MPD lidar signal is relatively weak and MPD observation error starts to increase sharply, assimilation of MPD degrades the analysis in terms of bias compared to assimilation of rawinsonde profiles (Fig. 15f). The improved water vapor mixing ratio analysis in advance of the convective region is attributed to the continuously available MPD moisture profiles during the frequent 15-min
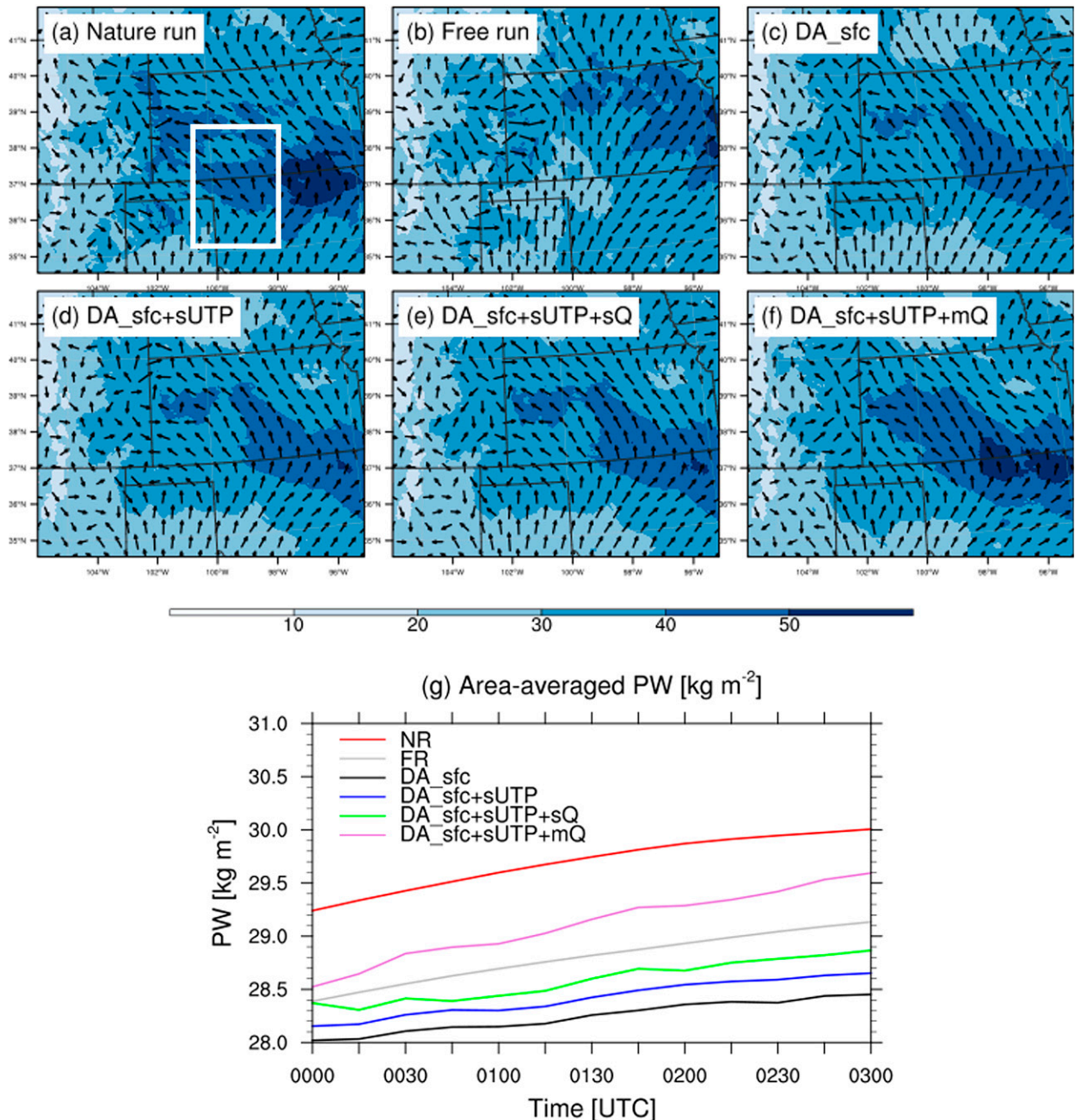
FIG. 13. Precipitable water (color fill; kg m$^{-2}$) and 850-hPa winds (vector; m s$^{-1}$) for (a) nature run, (b) free run, and ensemble mean analysis of (c) DA_sfc, (d), DA_sfc+sUTP, (e) DA_sfc+sUTP+sQ, and (f) DA_sfc+sUTP+mQ valid at 0300 UTC 15 Jul 2015. (g) Time series of precipitable water vapor (kg m$^{-2}$) averaged over the inner domain (D02) calculated from the nature run, free run, and ensemble mean analysis for each experiment. The solid white box in (a) indicates the averaging region for the height–latitude cross sections in Fig. 14.

DA cycling. The improvement in the analysis extends to predictions of the vertical moisture structure (not shown) and precipitation predictions (Figs. 9 and 10).

The MPD DA experiment leads to deep moist convection that is more consistent with the nature run. Here we examine how MPD DA leads to a more accurate evolution of the MCS. Figure 16 shows vertical profiles of thermodynamic fields in the region in advance of the MCS calculated over the box shown in Fig. 9a. The temperature and convective inhibition (CIN) profiles (Figs. 16b,d), together with convergence profiles (Fig. 12), are similar for all DA experiments showing improvement relative to the free run. The water vapor mixing
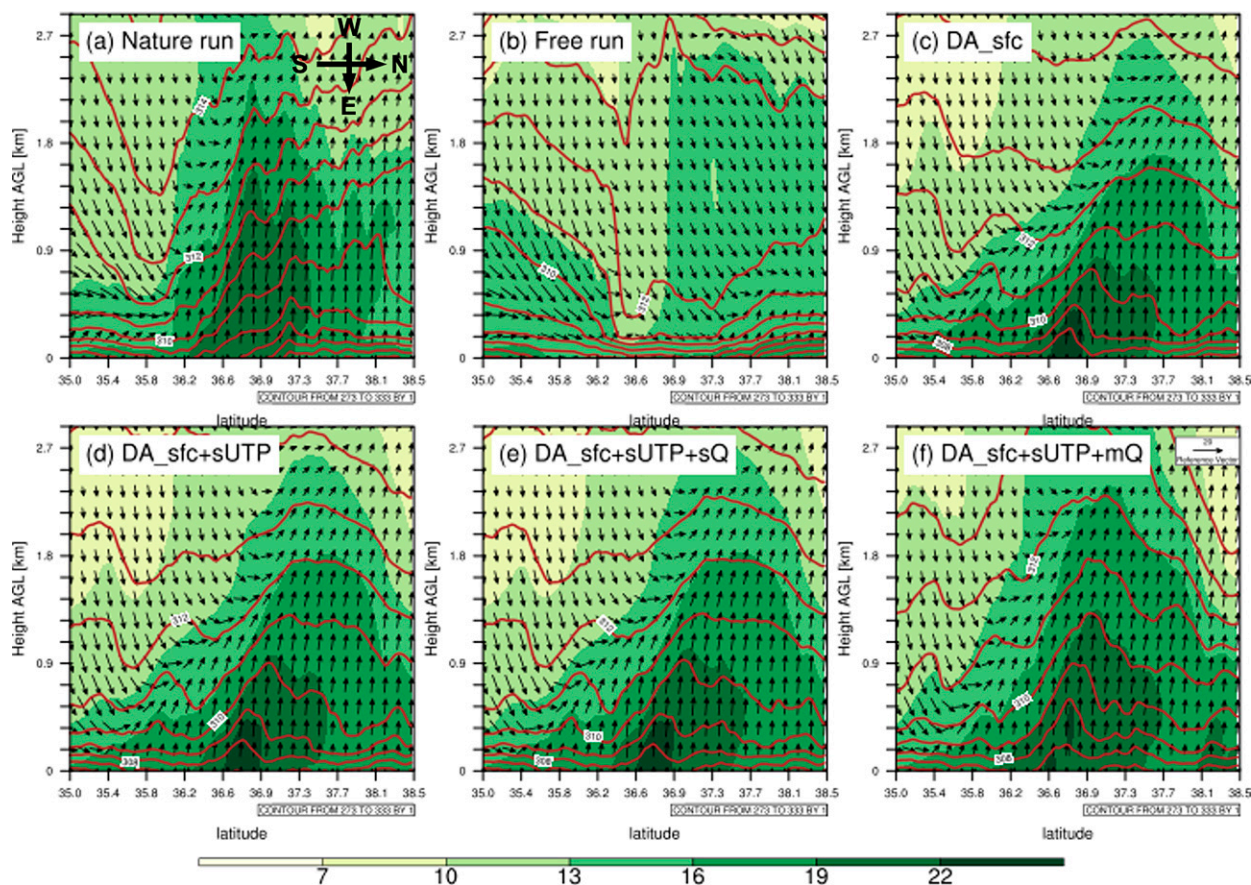
FIG. 14. Zonally averaged height–latitude cross section of water vapor mixing ratio (color fill; g kg$^{-1}$), potential temperature (red contour; K), and horizontal wind (vector; m s$^{-1}$) for (a) nature run, (b) free run, and ensemble mean analysis of (c) DA_sfc, (d), DA_sfc+sUTP, (e) DA_sfc+sUTP+sQ, and (f) DA_sfc+sUTP+mQ valid at 0300 UTC 15 Jul 2015. The cross sections are zonally averaged over the white box in Fig. 13a. The black arrow in the upper right of (a) represents the direction of the horizontal wind vector.

ratio (Fig. 16a), dewpoint temperature (Fig. 16b) and CAPE (Fig. 16c) are all impacted by assimilation of synthetic observations and are most improved by MPD DA. Because CAPE is a measure that incorporates moisture and instability, these results suggest that DA of MPD continuous water vapor profiles has the potential to markedly improve the moisture and instability in the convective storm environment, especially in midlevels in advance of the convection.

## 5. Additional effect of assimilating MPD data with randomly fixed MPD sites with varying number and distance

In the previous sections, we have collocated the MPD instruments with the operational rawinsonde sites and have verified the positive MPD impact of synthetic observations in comparison to conventional surface station and rawinsonde data for a case study. There are 45 rawinsonde sites within the outer domain and 15 sites within the inner domain with a separation of ~300 km (Fig. 3a). The MPD network examined in this study, collocated at the rawinsonde sites, was shown to improve the convective weather prediction skill. However,

prediction skill may be further improved by either more or fewer MPD observing sites. To further investigate the observational capability of a future operational nationwide network of instruments, we explore other potential configurations of an MPD network to complement the operational observing network.

To investigate the impact of changing the horizontal spacing of MPD sites, simulated MPDs are deployed at random locations with a variety of distances and total site counts from ~600 MPDs with ~70-km spacing to ~30 MPDs with ~380-km spacing within the outer domain (Fig. 17). The strategy to build the randomly fixed MPD network is as follows. An initial site is randomly chosen in the outer domain and a next site is randomly determined keeping a distance (e.g., for the 600-site experiment of ~70 km) from the previously chosen site. This process is repeated until the average distance between all sites is less than the designated spacing. This methodology is performed for each experiment with spacing of 70, 130, 200, 270, 320, and 380 km. To ensure a generalization in this random selection of the sites, the selection process is performed 10 times with different sets of randomly fixed MPD networks and the ten simulations are averaged for each MPD spacing.
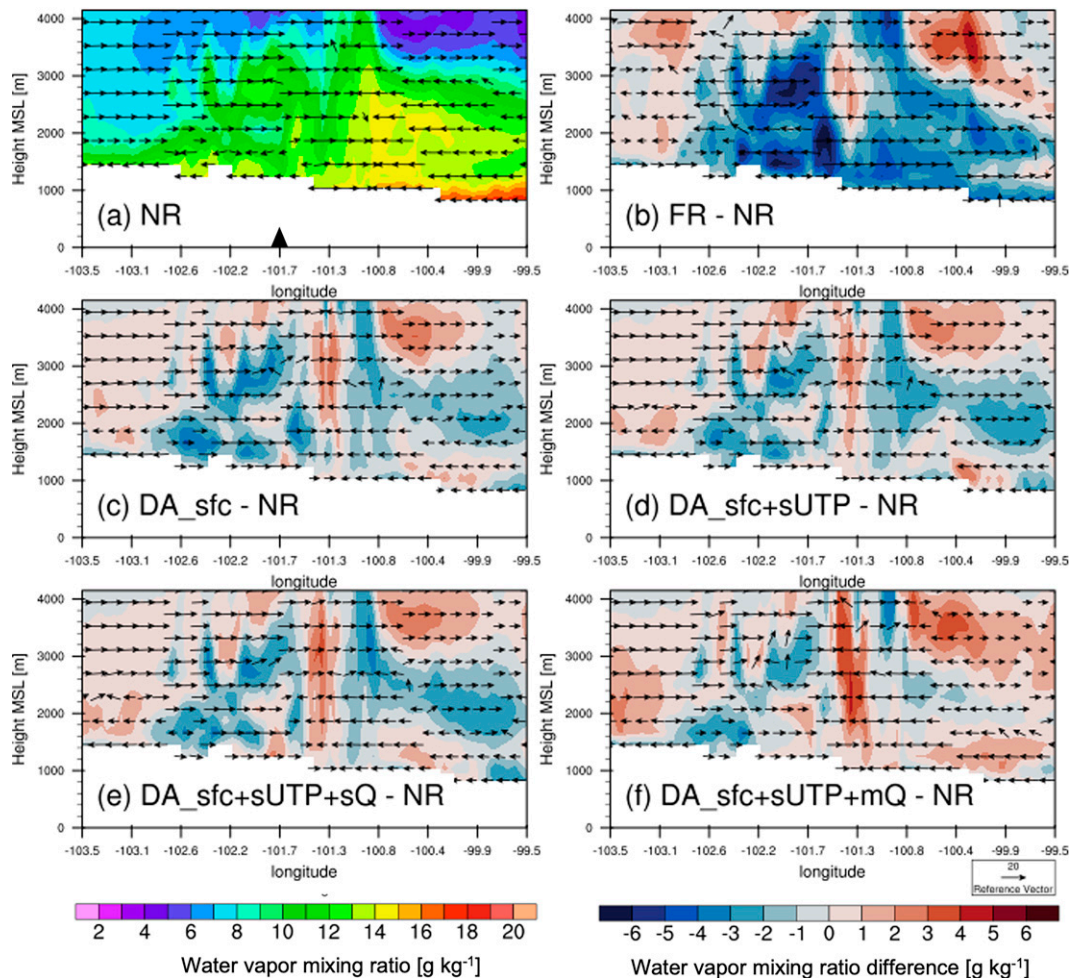
FIG. 15. (a) Height–longitude cross section of water vapor mixing ratio (color fill; g kg$^{-1}$) and wind (vector; m s$^{-1}$) for the nature run, and differences between nature run and (b) free run, (c) DA_sfc, (d), DA_sfc+sUTP, (e) DA_sfc+sUTP+sQ, and (f) DA_sfc+sUTP+mQ analysis along the line A–B in Fig. 9a valid at 0300 UTC 15 Jul 2015. The black triangle in (a) indicates the location of convection.

We compare the result of the 15 July case study with FSS (Fig. 18). To evaluate a potential nationwide network of MPDs, we broaden the verification region to the outer domain and increase sampling number for the verification over the United States (Fig. 3a). As the MPD spacing increases and the number of MPDs decrease, the experiments tend to predict light precipitation better (Fig. 18a). However, the difference of skills between the MPD spacing of 200 km and less are not statistically significant from the paired $t$ test with a significance level of 95% (not shown in figure). In contrast, as the MPD spacing decreases, the MPD has more impact on intense precipitation predictions except for the MPD spacing of 70 km (Fig. 18b). These contradictory results for different precipitation intensities imply that optimal spacing of MPDs is dependent on the scale of precipitation events. Because the intense convective precipitation results from complex multiscale interactions between the ambient flow and atmospheric convection (Majumdar et al. 2021), denser network of MPDs

can resolve fine scale variability of intense convective precipitation systems better than sparse MPD networks. The MPD spacings of 130 km is the optimal spacing in terms of FSS of convective precipitation for this case study and the difference between the other experiments are statistically significant. It is noted that MPD spacing of 70 km does not outperform the MPD spacing of 130 km in precipitation predictions for all forecast lead times and precipitation thresholds (Fig. 18). This is likely because frequent assimilation of noisy MPD water vapor profiles of increased density leads to an imbalance within the analyses and causes subsequent degraded predictions. On the other hand, light precipitation prediction is generally affected by large-scale forcing such as cyclonic and frontal systems. Therefore, improving the light precipitation prediction is mainly attributed to the more accurate analysis of the environment and is achieved with MPD spacing of 200 km. Figure 19 shows the temporal evolution of surface pressure tendency which is regarded as a measure of dynamic imbalance at
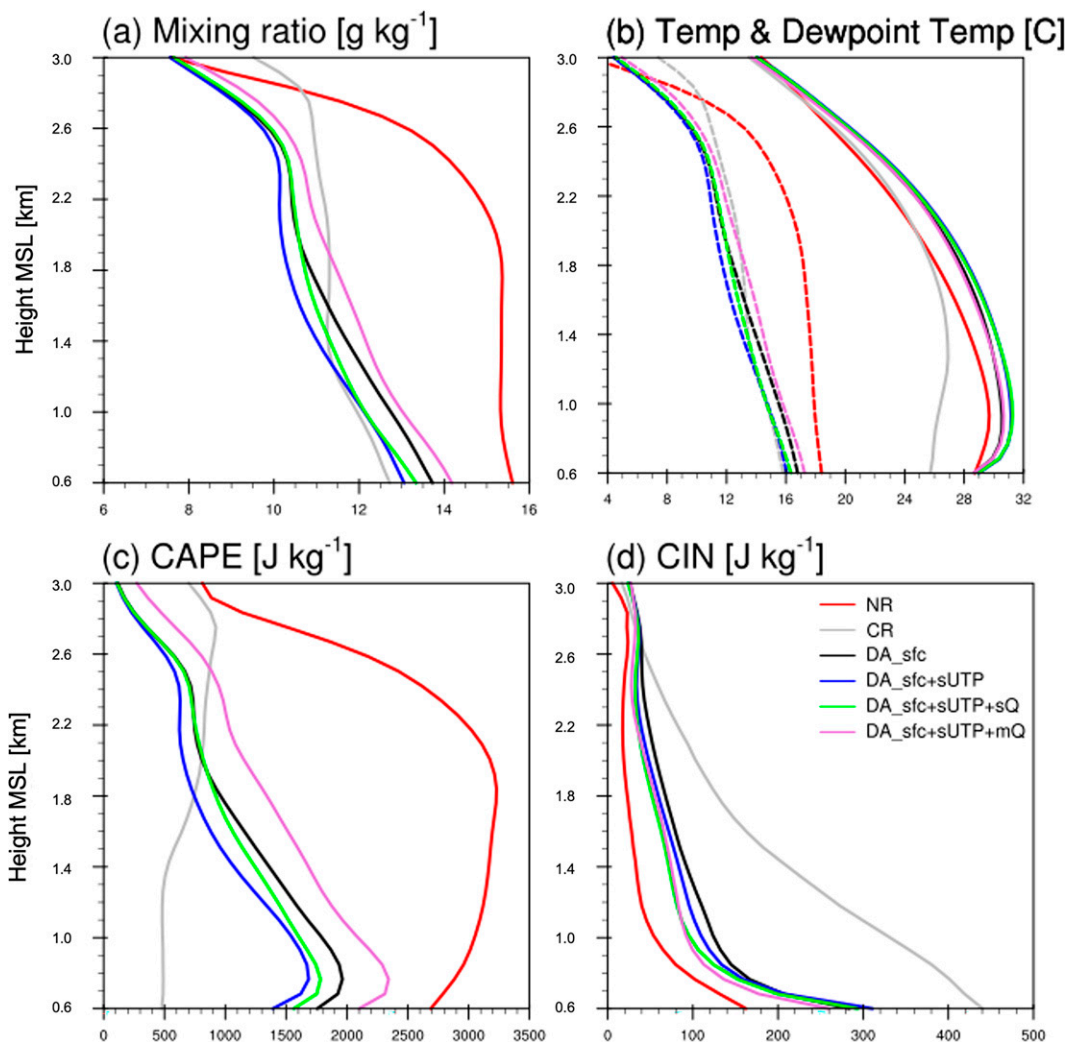
FIG. 16. (a) Profiles of water vapor mixing ratio (g kg$^{-1}$), (b) temperature (C), (c) CAPE (J kg$^{-1}$), and (d) CIN (J kg$^{-1}$) valid at 0300 UTC 15 Jul 2015. The profiles are averaged over the black box in Fig. 9a.

synoptic scales (Lynch and Huang 1992; Lei et al. 2012; Zeng et al. 2021). The surface pressure tendency calculated between 0- and 4-h predictions initialized at 0300 UTC 15 July is averaged over the outer domain. Consistent with FSS for light precipitation results (Fig. 18a), the decreasing MPD spacing results in larger values of surface pressure tendencies, indicating that assimilation of the high density water vapor profiles leads to less dynamically balanced predictions at large scales.

In this study, only MPD water vapor profiles are assimilated. Potentially, a network that additionally includes wind and temperature profiling data in the lower atmosphere could lead to more accurate analyses with less reliance on the ensemble covariance structure to update other state variables (Bannister et al. 2020). Inconsistent updates between the water vapor and temperature analyses can cause spurious precipitation forecasts and associated "spindown" (Sun et al. 2014; Clark et al. 2016). Therefore, a network consisting of dense MPD water vapor profiles and complementary temperature

and wind profiles is desired to improve precipitation prediction of various scales (Hartung et al. 2011). In addition, the ensemble DA system used in this study does not apply techniques to mitigate imbalance after DA, thus including an incremental analysis update technique (IAU; Bloom et al. 1996; Lei and Whitaker 2016) or digital filtering initialization technique (DFI; Lynch and Huang 1992) might lead to improved results.

Compared to the 3-km DA runs in section 4b, these results in this section could be influenced by a mixture of the MPD observation impact and the cumulus parameterization. Therefore, it should be noted that these results could be dependent upon the convection parameterization used in this study.

## 6. Comparison with the real MPD assimilation

During the PECAN field campaign, one MPD instrument was operated at Ellis, Kansas (fixed point 3; FP3), shown in Fig. 6a (Geerts et al. 2017). The insufficient spatial coverage
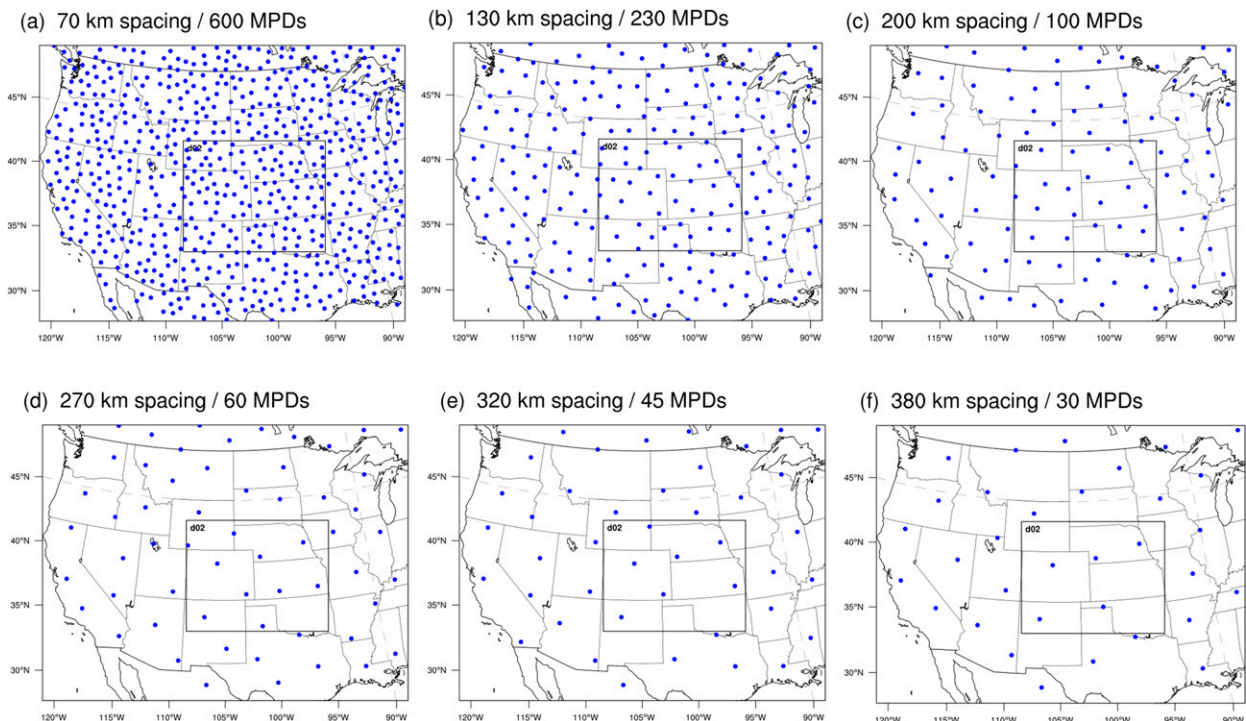
FIG. 17. Randomly fixed observation network (blue dots). The distance between observation sites ranges from ~70 to 380 km.

of the one MPD instrument limits the ability to assess meaningful impacts of the MPD data for this case study. But validity of the OSSE results in this study can be further assessed by comparing results of assimilating the real MPD and simulated MPD profiles at FP3. Therefore, we assimilate the real and simulated MPD data from the FP3 site in addition to the observations of the DA_sfc_sUTP experiment (hereafter referred to as "real_MPD" and "simulated_MPD" experiments).

Figure 20 shows the time series of the absolute humidity from the real_MPD and simulted_MPD, and their root-mean-square error (rmse) of analysis (O-A) and background (O-B) verified against real and simulated MPD observations, respectively.

Overall, real MPD has higher absolute humidity than the simulated MPD for all DA cycles, but the difference between them decreases with the cycles. The sawtooth patterns of O-A and O-B demonstrate the positive impact of the assimilation of both real and simulated MPD observations. These results imply that we can expect the positive MPD impact from assimilating real MPD datasets with similar configurations with the OSSE. However, it is noted that the background error and its correction by the real MPD observations are larger than the simulated_MPD. This is because the simulated_MPD experiment has smaller discrepancy between the true state (or nature run) and assimilated observations compared to the
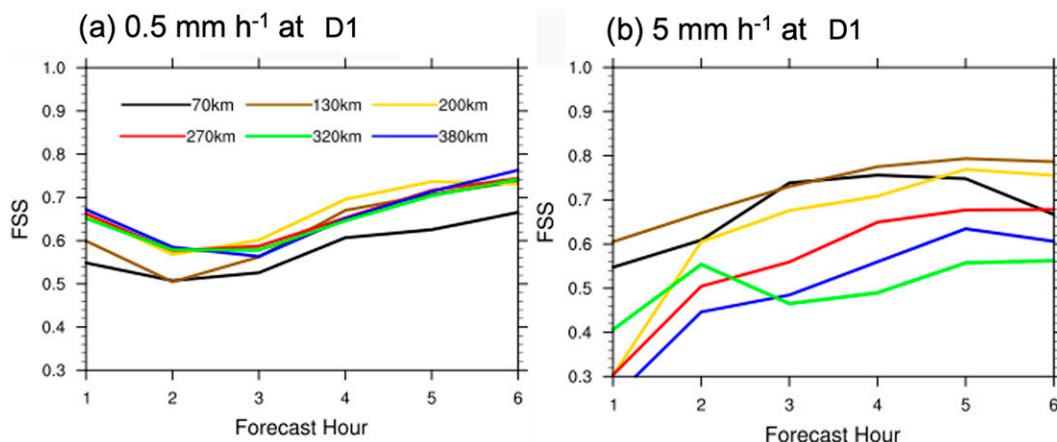


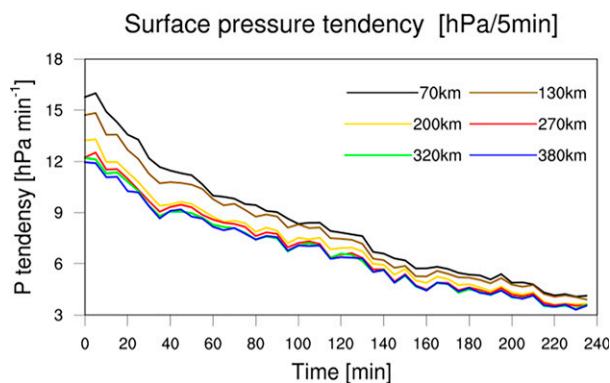FIG. 18. As in Fig. 10, but FSS of randomly located MPD with fixed station distances.

FIG. 19. Domain-averaged surface pressure tendency every 5 min during the 4-h deterministic prediction starting at 0300 UTC 15 Jul 2015 over the D1 domain calculated from MPD spacings of 70, 130, 200, 270, 320, and 380 km.
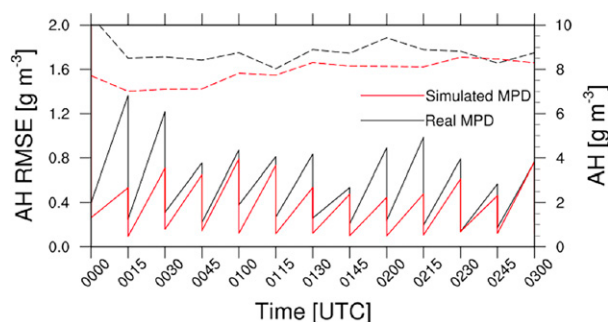


FIG. 20. A time series of absolute humidity (dotted line) and rmse of analysis (O–A) and background (O–B) for absolute humidity (solid line) for real MPD DA run (black line) and simulated MPD DA run (red line) from 0000 to 0300 UTC 15 Jun 2015.

real_MPD analysis. While continuous data assimilation in the simulated_MPD experiment efficiently corrects the background toward the true state (or nature run), the data assimilation of the real_MPD experiment generally includes larger errors associated with observations and numerical models than in the simulated_MPD. In particular, the identical twin problem of the OSSE discussed in section 3 can exaggerate the gap between simulated_MPD and real_MPD results. These results imply that we can expect a positive MPD impact from assimilating real MPD datasets with similar configurations with the OSSE, but their observation impact may decrease compared to the OSSE result.

## 7. Summary and discussion

NCAR and MSU jointly developed the MPD, which measures continuous, unattended, and accurate water vapor profiles. In addition, MPD has economic advantages in operation and manufacturing over several other water vapor profiling instruments, and thus may be suitable as a future network of instruments. To investigate the potential impact of a network of MPD water vapor profiles on convective weather forecasting, OSSEs are used to assimilate synthesized observational networks that are composed of various combinations of MPD profiles and conventional observations for a 15 July 2015 nocturnal convection case study. The NCAR MODE method is first used to provide objective criteria defining the nature run and free run from ensemble simulations for the OSSE.

It was found that the assimilation of frequent observations of high-resolution water vapor profiles in the lower troposphere provided by MPD can improve analyses of the preconvective environment for improved prediction of nocturnal convection. Surface station data are shown to be effective in correcting the frontal structure and winds near the surface. However, with DA of surface station data only, analyses show a limitation in improving the inflow of water vapor in the lower troposphere at ~2–3 km AGL where elevated convection can be rooted at night. Although rawinsondes provide water vapor profiles extending to high altitudes, the

low temporal resolution of rawinsondes is shown to have limited benefit to improve atmospheric water vapor fields in the vicinity of the nocturnal convective system. On the other hand, MPDs, here collocated with the rawinsondes, provide continuously updating water vapor profiles in the lower troposphere during each 15-min DA cycle. The MPD DA led to improved water vapor analyses ahead of the convective area, improving the representation of low-level convergence and instability and resulting in a more accurate precipitation prediction. Overall, DA of the MPD water vapor profiles led to improved predictions of convection initiation, evolution and precipitation of the MCS compared to the conventional observations for this particular case study.

To evaluate the impacts from the integrated DA efforts including MPD and rawinsonde data, we also conducted a DA run assimilating water vapor profiles from both the MPD and sounding (DA_sfc+sUTP+sQ+mQ). DA_sfc+sUTP+sQ+mQ and DA_sfc+sUTP+mQ show marginal differences in almost all metrics used in this study (not shown), which is likely due to the limited number of sounding data available at the beginning of the DA cycles (Fig. 3b). This result is consistent with the results showing the small difference between DA_sfc+sUTP and DA_sfc+sUTP+sQ (Figs. 9–16). All these results demonstrate the benefit of the high temporal resolution of MPD water vapor profiles in the lower atmosphere that complement the infrequent rawinsonde data.

Given the potential value of a nationwide network of water vapor profiles to improve weather forecast skill, comparisons were also performed to explore the impact of varied station spacing for a network of MPD sites using an OSSE framework. For this nocturnal case study, an MPD site spacing of 130 km (200 km), which corresponds to 300 (100) MPD sites across the outer domain, is found to be optimal in terms of intense (light) precipitation prediction skill. In this study, a denser MPD network than the current operational rawinsonde network resulted in better skillful prediction of intense precipitation, while a much sparser MPD network with greater than 200-km spacing had significant impact on light precipitation predictions for this event due to more balanced initial states.

This study is the first to explore the potential impact of an MPD water vapor profiling network on the skill of convective

weather predictions. As an initial investigation, the experiment is limited by addressing the MPD observation impact for a single nocturnal convection case. In this regard, further studies are planned to expand OSSEs to more convective cases to address the generality of these results. To determine the optimal combination of thermodynamic and wind profiles for improved convective weather forecasting, observation impact studies of wind and temperature profiles from other remote sensing instruments that can complement the MPD water vapor profiles will be explored through OSSEs. It is noted that the MPD effect in OSSEs could be exaggerated compared to the actual MPD observation effect because we underestimate model errors and overestimate observation errors in OSSEs. In the future, the observation impact of assimilating actual MPD water vapor profiles will also be investigated. Since the error in the MPD has been substantially reduced with a recent upgrade (Spuler et al. 2021), sophisticated methods to estimate more realistic observation errors will be explored to investigate DA of actual MPD profiles (e.g., Bormann et al. 2011; Minamide and Zhang 2017).

*Data availability statement.* Computing and data storage resources, including the Cheyenne supercomputer (https://doi.org/10.5065/D6RX99HX), were provided by the Computational and Information Systems Laboratory (CISL) at NCAR. The raw model/reanalysis data including NCEP GFS, FNL, and conventional observed data can be downloaded online at UCAR Research Data Archive (RDA) datasets (https://rda.ucar.edu/). The ERA5 data are retrieved from https://cds.climate.copernicus.eu. The MPD data can be downloaded online from https://data.eol.ucar.edu/dataset/597.001. The authors provide all the simulated datasets used in the manuscript through the following link: https://doi.org/10.5065/231h-jw06.

## REFERENCES

Ahijevych, D., E. Gilleland, B. G. Brown, and E. E. Ebert, 2009: Application of spatial verification methods to idealized and NWP-gridded model precipitation forecasts. *Wea. Forecasting*, **24**, 1485–1497, https://doi.org/10.1175/2009WAF2222298.1.

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

——, and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758, https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2.

——, 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, https://doi.org/10.1175/2009BAMS2618.1.

Bannister, R. N., H. G. Chipilski, and O. Martinez-Alvarado, 2020: Techniques and challenges in the assimilation of atmospheric water observations for numerical weather prediction towards convective scales. *Quart. J. Roy. Meteor. Soc.*, **146**, 1–48, https://doi.org/10.1002/qj.3652.

Barker, D. M., W. Huang, Y.-R. Guo, A. J. Bourgeois, and Q. N. Xio, 2004: A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Mon. Wea. Rev.*, **132**, 897–914, https://doi.org/10.1175/1520-0493(2004)132<0897:ATVDAS>2.0.CO;2.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, https://doi.org/10.1038/nature14956.

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271, https://doi.org/10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2.

Bormann, N., and P. Bauer, 2010: Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data. *Quart. J. Roy. Meteor. Soc.*, **136**, 1036–1050, https://doi.org/10.1002/qj.616.

——, A. J. Geer, and P. Bauer, 2011: Estimates of observation error characteristics in clear and cloudy regions for microwave imager radiances from numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **137**, 2014–2023, https://doi.org/10.1002/qj.833.

Brennan, M. J., J. Clark, and M. Klein, 2008: Verification of quantitative precipitation forecast guidance from NWP models and the Hydrometeorological Prediction Center for 2005–2007 tropical cyclones with continental U.S. rainfall impacts. *28th Conf. on Hurricanes and Tropical Meteorology*, Orlando, FL, Amer. Meteor. Soc., P2H.9, https://ams.confex.com/ams/pdfpapers/138022.pdf.

Buehner, M., and D. Jacques, 2020: Non-Gaussian deterministic assimilation of radar-derived precipitation accumulations. *Mon. Wea. Rev.*, **148**, 783–808, https://doi.org/10.1175/MWR-D-19-0199.1.

Carroll, B. J., B. B. Demoz, D. D. Turner, and R. Delgado, 2021: Lidar observations of a mesoscale moisture transport event impacting convection and comparison to Rapid Refresh model analysis. *Mon. Wea. Rev.*, **149**, 463–477, https://doi.org/10.1175/MWR-D-20-0151.1.

Charba, J. P., D. W. Reynolds, B. E. McDonald, and G. M. Carter, 2003: Comparative verification of recent quantitative precipitation forecasts in the National Weather Service: A simple approach for scoring forecast accuracy. *Wea. Forecasting*, **18**, 161–183, https://doi.org/10.1175/1520-0434(2003)018<0161:CVORQP>2.0.CO;2.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5

modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129**, 569–585, https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Cherubini, T., A. Ghelli, and F. Lalaurette, 2002: Verification of precipitation forecasts over the alpine region using a high-density observing network. *Wea. Forecasting*, **17**, 238–249, https://doi.org/10.1175/1520-0434(2002)017<0238:VOPFOT>2.0.CO;2.

Chipilski, H. G., X. Wang, and D. B. Parsons, 2020: Impact of assimilating PECAN profilers on the prediction of bore-driven nocturnal convection: A multiscale forecast evaluation for the 6 July 2015 case study. *Mon. Wea. Rev.*, **148**, 1147–1175, https://doi.org/10.1175/MWR-D-19-0171.1.

——, ——, ——, A. Johnson, and S. K. Degelia, 2022: The value of assimilating different ground-based profiling networks on the forecasts of bore-generating nocturnal convection. *Mon. Wea. Rev.*, **150**, 1273–1292, https://doi.org/10.1175/MWR-D-21-0193.1.

Clark, P., N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez, 2016: Convection-permitting models: A step-change in rainfall forecasting. *Meteor. Appl.*, **23**, 165–181, https://doi.org/10.1002/met.1538.

Coniglio, M. C., G. S. Romine, D. D. Turner, and R. D. Torn, 2019: Impacts of targeted AERI and Doppler Lidar wind retrievals on short-term forecasts of the initiation and early evolution of thunderstorms. *Mon. Wea. Rev.*, **147**, 1149–1170, https://doi.org/10.1175/MWR-D-18-0351.1.

Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.*, **124**, 1767–1785, https://doi.org/10.1175/1520-0493(1996)124<1767:SOMCFB>2.0.CO;2.

Davis, C., B. Brown, and R. Bullock, 2006a: Object-based verification of precipitation forecasts. Part I: Methods and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784, https://doi.org/10.1175/MWR3145.1.

——, ——, and ——, 2006b: Object-based verification of precipitation forecasts. Part II: Application to convective rain systems. *Mon. Wea. Rev.*, **134**, 1785–1795, https://doi.org/10.1175/MWR3146.1.

——, ——, ——, and J. Halley-Gotway, 2009: The method for object-based diagnostic evaluation (MODE) applied to numerical forecasts from the 2005 NSSL/SPC Spring Program. *Wea. Forecasting*, **24**, 1252–1267, https://doi.org/10.1175/2009WAF2222241.1.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

Degelia, S. K., X. Wang, and D. J. Stensrud, 2019: An evaluation of the impact of assimilating AERI retrievals, kinematic profilers, rawinsondes, and surface observations on a forecast of a nocturnal convection initiation event during the PECAN field campaign. *Mon. Wea. Rev.*, **147**, 2739–2764, https://doi.org/10.1175/MWR-D-18-0423.1.

——, ——, ——, and D. D. Turner, 2020: Systematic evaluation of the impact of assimilating a network of ground-based remote sensing profilers for forecasts of nocturnal convection initiation during PECAN. *Mon. Wea. Rev.*, **148**, 4703–4728, https://doi.org/10.1175/MWR-D-20-0118.1.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Quart. J. Roy. Meteor. Soc.*, **131**, 3385–3396, https://doi.org/10.1256/qj.05.108.

Dierer, S., and Coauthors, 2009: Deficiencies in quantitative precipitation forecasts: Sensitivity studies using the COSMO model. *Meteor. Z.*, **18**, 631–645, https://doi.org/10.1127/0941-2948/2009/0420.

Dixon, M., 2016: PECAN: Radar 3D mosaic DBZ and ZDR, version 1.0. UCAR/NCAR–Earth Observing Laboratory, accessed 11 May 2021, https://doi.org/10.5065/D6QR4VHM.

Ducrocq, V., D. Ricard, J. P. Lafore, and F. Orain, 2002: Storm-scale numerical rainfall prediction for five precipitating events over France: On the importance of the initial humidity field. *Wea. Forecasting*, **17**, 1236–1256, https://doi.org/10.1175/1520-0434(2002)017<1236:SSNRPF>2.0.CO;2.

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

Geer, A. J., and P. Bauer, 2011: Observation errors in all-sky data assimilation. *Quart. J. Roy. Meteor. Soc.*, **137**, 2024–2037, https://doi.org/10.1002/qj.830.

Geerts, B., and Coauthors, 2017: The 2015 Plains Elevated Convection At Night Field Project. *Bull. Amer. Meteor. Soc.*, **98**, 767–786, https://doi.org/10.1175/BAMS-D-15-00257.1.

Grasmick, C., B. Geerts, D. D. Turner, Z. Wang, and T. M. Weckwerth, 2018: The relation between nocturnal MCS evolution and its outflow boundaries in the stable boundary layer: An observational study of the 15 July 2015 MCS in PECAN. *Mon. Wea. Rev.*, **146**, 3203–3226, https://doi.org/10.1175/MWR-D-18-0169.1.

Hartung, D. C., J. A. Otkin, R. A. Petersen, D. D. Turner, and W. F. Feltz, 2011: Assimilation of surface-based boundary layer profiler observations during a cool-season weather event using an observing system simulation experiment. Part II: Forecast assessment. *Mon. Wea. Rev.*, **139**, 2327–2346, https://doi.org/10.1175/2011MWR3623.1.

Hayman, M., and S. Spuler, 2017: Demonstration of a diode-laser-based High Spectral Resolution Lidar (HSRL) for quantitative profiling of clouds and aerosols. *Opt. Express*, **25**, A1096–A1110, https://doi.org/10.1364/OE.25.0A1096.

Hitchcock, S. M., M. C. Coniglio, and K. H. Knopfmeier, 2016: Impact of MPEX observations on ensemble analyses and forecasts of the 31 May 2013 convective event over Oklahoma. *Mon. Wea. Rev.*, **144**, 2889–2913, https://doi.org/10.1175/MWR-D-15-0344.1.

Hoffman, R. N., and R. Atlas, 2016: Future observing system simulation experiments. *Bull. Amer. Meteor. Soc.*, **97**, 1601–1616, https://doi.org/10.1175/BAMS-D-15-00200.1.

Holley, D. M., S. R. Dorling, C. J. Steele, and N. Earl, 2014: A climatology of convective available potential energy in Great Britain. *Int. J. Climatol.*, **34**, 3811–3824, https://doi.org/10.1002/joc.3976.

Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A**, 111–136, https://doi.org/10.1111/j.1600-0870.1986.tb00460.x.

Hu, H., J. Sun, and Q. Zhang, 2017: Assessing the impact of surface and wind profiler data on fog forecasting using WRF 3DVAR: An OSSE study on a dense fog event over North China. *J. Appl. Meteor. Climatol.*, **56**, 1059–1081, https://doi.org/10.1175/JAMC-D-16-0246.1.

Hu, J., N. Yussouf, D. D. Turner, T. A. Jones, and X. Wang, 2019: Impact of ground-based remote sensing boundary layer observations on short-term probabilistic forecasts of a tornadic supercell event. *Wea. Forecasting*, **34**, 1453–1476, https://doi.org/10.1175/WAF-D-18-0200.1.

Hu, M., G. Ge, C. Zhou, D. Stark, H. Shao, K. Newman, J. Beck, and X. Zhang, 2018: Grid-point Statistical Interpolation (GSI) user's guide version 3.7. Developmental Testbed Center, 149 pp., http://www.dtcenter.org/com-GSI/users/docs/index.php.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113**, D13103, https://doi.org/10.1029/2008JD009944.

Janjić, T., and Coauthors, 2017: On the representation error in data assimilation. *Quart. J. Roy. Meteor. Soc.*, **144**, 1257–1278, https://doi.org/10.1002/qj.3130.

Janjić, Z. I., 2002: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model. NCEP Office Note 437, 61 pp., http://www.emc.ncep.noaa.gov/officenotes/newernotes/on437.pdf.

Johns, R. H., and C. A. Doswell, 1992: Severe local storms forecasting. *Wea. Forecasting*, **7**, 588–612, https://doi.org/10.1175/1520-0434(1992)007<0588:SLSF>2.0.CO;2.

Johnson, A., X. Wang K. R. Haghi, and D. B. Parsons, 2018: Evaluation of forecasts of a convectively generated bore using an intensively observed case study from PECAN. *Mon. Wea. Rev.*, **146**, 3097–3122, https://doi.org/10.1175/MWR-D-18-0059.1.

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

——, and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2784–2802, https://doi.org/10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2.

——, and ——, 1993: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models*, Meteor. Monogr., No. 24, Amer. Meteor. Soc., 165–170.

Kawabata, T., H. Seko, K. Saito, T. Kuroda, K. Tamiya, T. Tsuyuki, Y. Honda, and Y. Wakazuki, 2007: An assimilation and forecasting experiment of the Nerima heavy rainfall with a cloud-resolving nonhydrostatic 4-dimensional variational data assimilation system. *J. Meteor. Soc. Japan*, **85**, 255–276, https://doi.org/10.2151/jmsj.85.255.

Keclik, A. M., C. Evans, P. J. Roebber, and G. S. Romine, 2017: The influence of assimilated upstream, preconvective dropsonde observations on ensemble forecasts of convection initiation during the Mesoscale Predictability Experiment. *Mon. Wea. Rev.*, **145**, 4747–4770, https://doi.org/10.1175/MWR-D-17-0159.1.

Keil, C., A. Röpnack, G. C. Craig, and U. Schumann, 2008: Sensitivity of quantitative precipitation forecast to height dependent changes in humidity. *Geophys. Res. Lett.*, **35**, L09812, https://doi.org/10.1029/2008GL033657.

Kerr, C. A., D. J. Stensrud, and X. Wang, 2019: Diagnosing convective dependencies on near-storm environments using ensemble sensitivity analyses. *Mon. Wea. Rev.*, **147**, 495–517, https://doi.org/10.1175/MWR-D-18-0140.1.

Knuteson, R. O., and Coauthors, 2004a: The Atmospheric Emitted Radiance Interferometer. Part I: Instrument design. *J. Atmos. Oceanic Technol.*, **21**, 1763–1776, https://doi.org/10.1175/JTECH-1662.1.

——, and Coauthors, 2004b: The Atmospheric Emitted Radiance Interferometer. Part II: Instrument performance. *J. Atmos. Oceanic Technol.*, **21**, 1777–1789, https://doi.org/10.1175/JTECH-1663.1.

Lei, L., and J. S. Whitaker, 2016: A four-dimensional incremental analysis update for the ensemble Kalman filter. *Mon. Wea. Rev.*, **144**, 2605–2621, https://doi.org/10.1175/MWR-D-15-0246.1.

——, D. R. Stauffer, and A. Deng, 2012: A hybrid nudging-ensemble Kalman filter approach to data assimilation in WRF/DART. *Quart. J. Roy. Meteor. Soc.*, **138**, 2066–2078, https://doi.org/10.1002/qj.1939.

Lin, P.-F., P.-L. Chang, B. J.-D. Jou, J. W. Wilson, and R. D. Roberts, 2011: Warm season thunderstorm characteristics under weak synoptic-scale forcing over Taiwan Island. *Wea. Forecasting*, **26**, 44–60, https://doi.org/10.1175/2010WAF2222386.1.

Liu, C., M. Xue, and R. Kong, 2019: Direct assimilation of radar reflectivity data using 3DVAR: Treatment of hydrometeor background errors and OSSE tests. *Mon. Wea. Rev.*, **147**, 17–29, https://doi.org/10.1175/MWR-D-18-0033.1.

Lynch, P., and X.-Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019–1034, https://doi.org/10.1175/1520-0493(1992)120<1019:IOTHMU>2.0.CO;2.

Majumdar, S. J., and Coauthors, 2021: Multiscale forecasting of high-impact weather: Current status and future challenges. *Bull. Amer. Meteor. Soc.*, **102**, E635–E659, https://doi.org/10.1175/BAMS-D-20-0111.1.

Masutani, M., and Coauthors, 2007: Progress in Joint OSSEs: A new nature run and international collaboration. *12th Conf. on Integrated Observing and Assimilation Systems for Atmospheres, Oceans, and Land Surface (IOAS-AOLS)*, New Orleans, LA, Amer. Meteor. Soc., 12B.5.

——, and Coauthors, 2010: Observing system simulation experiments at the National Centers for Environmental Prediction. *J. Geophys. Res.*, **115**, D07101, https://doi.org/10.1029/2009JD012528.

Minamide, M., and F. Zhang, 2017: Adaptive observation error inflation for assimilating all-sky satellite radiance. *Mon. Wea. Rev.*, **145**, 1063–1081, https://doi.org/10.1175/MWR-D-16-0257.1.

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

Morss, R. E., and F. M. Ralph, 2007: Use of information by National Weather Service forecasters and emergency managers during the CALJET and PACJET-2001. *Wea. Forecasting*, **22**, 539–555, https://doi.org/10.1175/WAF1001.1.

National Research Council, 2009: *Observing Weather and Climate from the Ground Up: A Nationwide Network of Networks*. National Academies Press, 250 pp., https://doi.org/10.17226/12540.

——, 2010: *When Weather Matters: Science and Service to Meet Critical Societal Needs*. National Academies Press, 198 pp., https://doi.org/10.17226/12888.

——, 2012: *Weather Services for the Nation: Becoming Second to None*. National Academies Press, 86 pp., https://doi.org/10.17226/13429.

——, 2018: *The Future of Atmospheric Boundary Layer Observing, Understanding and Modeling*. National Academies Press, 58 pp., https://doi.org/10.17226/25138.

Novak, D. R., F. E. Barthold, M. J. Bodner, K. F. Brill, and M. Eckert, 2011: Quantifying extreme rainfall threats at the Hydrometeorological Prediction Center. *24th Conf. on Weather and Forecasting/20th Conf. on Numerical Weather Prediction*,

Seattle, WA, Amer. Meteor. Soc., 14A.4, https://ams.confex.com/ams/91Annual/webprogram/Paper179714.html.

Nurmi, P., 2003: Recommendations on the verification of local weather forecasts. ECMWF Tech. Memo. **430**, 19 pp., https://doi.org/10.21957/y1z1thg5l.

Pan, S., J. Gao, D. J. Stensrud, X. Wang, and T. A. Jones, 2018: Assimilation of radar radial velocity and reflectivity, satellite cloud water path, and total precipitable water for convective-scale NWP in OSSEs. *J. Atmos. Oceanic Technol.*, **35**, 67–89, https://doi.org/10.1175/JTECH-D-17-0081.1.

Powers, J. G., and Coauthors, 2017: The Weather Research and Forecasting (WRF) Model: Overview, system efforts, and future directions. *Bull. Amer. Meteor. Soc.*, **98**, 1717–1737, https://doi.org/10.1175/BAMS-D-15-00308.1.

Privé, N. C., and R. M. Errico, 2013: The role of model and initial condition error in numerical weather forecasting investigated with an observing system simulation experiment. *Tellus*, **65A**, 21740, https://doi.org/10.3402/tellusa.v65i0.21740.

Ralph, F. M., P. J. Neiman, D. E. Kingsmill, P. O. G. Persson, A. B. White, E. T. Strem, E. D. Andrews, and R. C. Antweiler, 2003: The impact of a prominent rain shadow on flooding in California's Santa Cruz Mountains: A CALJET case study and sensitivity to the ENSO cycle. *J. Hydrometeor.*, **4**, 1243–1264, https://doi.org/10.1175/1525-7541(2003)004<1243:TIOAPR>2.0.CO;2.

——, E. Sukovich, D. Reynolds, M. Dettinger, S. Weagle, W. Clark, and P. J. Neiman, 2010: Assessment of extreme quantitative precipitation forecasts and development of regional extreme event thresholds using data from HMT-2006 and COOP observers. *J. Hydrometeor.*, **11**, 1286–1304, https://doi.org/10.1175/2010JHM1232.1.

Richard, E., A. Buzzi, and G. Zängl, 2007: Quantitative precipitation forecasting in the Alps: The advances achieved by the mesoscale alpine programme. *Quart. J. Roy. Meteor. Soc.*, **133**, 831–846, https://doi.org/10.1002/qj.65.

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, https://doi.org/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, https://doi.org/10.1175/MWR-D-12-00112.1.

Schumacher, R. S., 2015: Resolution dependence of initiation and upscale growth of deep convection in convection-allowing forecasts of the 31 May–1 June 2013 supercell and MCS. *Mon. Wea. Rev.*, **143**, 4331–4354, https://doi.org/10.1175/MWR-D-15-0179.1.

——, and C. A. Davis, 2010: Ensemble-based forecast uncertainty analysis of diverse heavy rainfall events. *Wea. Forecasting*, **25**, 1103–1122, https://doi.org/10.1175/2010WAF2222378.1.

Schwartz, C. S., G. S. Romine, M. L. Weisman, R. A. Sobash, K. R. Fossell, K. W. Manning, and S. B. Trier, 2015a: A real-time convection-allowing ensemble prediction system initialized by mesoscale ensemble Kalman filter analyses. *Wea. Forecasting*, **30**, 1158–1181, https://doi.org/10.1175/WAF-D-15-0013.1.

——, Z. Liu, and X.-Y. Huang, 2015b: Sensitivity of limited-area hybrid variational-ensemble analyses and forecasts to ensemble perturbation resolution. *Mon. Wea. Rev.*, **143**, 3454–3477, https://doi.org/10.1175/MWR-D-14-00259.1.

——, G. S. Romine, R. A. Sobash, K. R. Fossell, and M. L. Weisman, 2019: NCAR's real-time convection-allowing ensemble project. *Bull. Amer. Meteor. Soc.*, **100**, 321–343, https://doi.org/10.1175/BAMS-D-17-0297.1.

Shao, H., and Coauthors, 2016: Bridging research to operations transitions: Status and plans of community GSI. *Bull. Amer. Meteor. Soc.*, **97**, 1427–1440, https://doi.org/10.1175/BAMS-D-13-00245.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

Spuler, S. M., K. S. Repasky, B. Morley, D. Moen, M. Hayman, and A. R. Nehrir, 2015: Field-deployable diode-laser-based differential absorption Lidar (DIAL) for profiling water vapor. *Atmos. Meas. Tech.*, **8**, 1073–1087, https://doi.org/10.5194/amt-8-1073-2015.

——, M. Hayman, R. A. Stillwell, J. Carnes, T. Bernatsky, and K. S. Repasky, 2021: MicroPulse DIAL (MPD)—A diode-laser-based lidar architecture for quantitative atmospheric profiling. *Atmos. Meas. Tech.*, **14**, 4593–4616, https://doi.org/10.5194/amt-14-4593-2021.

Stelten, S., and W. A. Gallus Jr., 2017: Pristine nocturnal convective initiation: A climatology and preliminary examination of predictability. *Wea. Forecasting*, **32**, 1613–1635, https://doi.org/10.1175/WAF-D-16-0222.1.

Stillwell, R. A., S. M. Spuler, M. Hayman, K. S. Repasky, and C. E. Bunn, 2020: Demonstration of a combined differential absorption and high spectral resolution lidar for profiling atmospheric temperature. *Opt. Express*, **28**, 71–93, https://doi.org/10.1364/OE.379804.

Sun, J., and Coauthors, 2014: Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Amer. Meteor. Soc.*, **95**, 409–426, https://doi.org/10.1175/BAMS-D-11-00263.1.

Tang, S. L., and D. J. Kirshbaum, 2020: On the sensitivity of deep-convection initiation to horizontal grid resolution. *Quart. J. Roy. Meteor. Soc.*, **146**, 1085–1105, https://doi.org/10.1002/qj.3726.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, https://doi.org/10.1175/2008MWR2387.1.

Thundathil, R., T. Schwitalla, A. Behrendt, and V. Wulfmeyer, 2021: Impact of assimilating Lidar water vapour and temperature profiles with a hybrid ensemble transform Kalman filter: Three-dimensional variational analysis on the convection-permitting scale. *Quart. J. Roy. Meteor. Soc.*, **147**, 4163–4185, https://doi.org/10.1002/qj.4173.

Torn, R. D., G. J. Hakim, and C. Snyder, 2006: Boundary conditions for limited area ensemble Kalman filters. *Mon. Wea. Rev.*, **134**, 2490–2502, https://doi.org/10.1175/MWR3187.1.

Trier, S. B., and D. B. Parsons, 1993: Evolution of environmental conditions preceding the development of a nocturnal mesoscale convective complex. *Mon. Wea. Rev.*, **121**, 1078–1098, https://doi.org/10.1175/1520-0493(1993)121<1078:EOECPT>2.0.CO;2.

Turner, D. D., and U. Löhnert, 2014: Information content and uncertainties in thermodynamic profiles and liquid cloud properties retrieved from the ground-based Atmospheric Emitted Radiance Interferometer (AERI). *J. Appl. Meteor. Climatol.*, **53**, 752–771, https://doi.org/10.1175/JAMC-D-13-0126.1.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008: A hybrid ETKF–3DVAR data assimilation scheme for the WRF Model. Part I: Observing system simulation experiment. *Mon. Wea. Rev.*, **136**, 5116–5131, https://doi.org/10.1175/2008MWR2444.1.

Weckwerth, T. M., 2000: The effect of small-scale moisture variability on thunderstorm initiation. *Mon. Wea. Rev.*, **128**, 4017–4030, https://doi.org/10.1175/1520-0493(2000)129<4017:TEOSSM>2.0.CO;2.

——, V. Wulfmeyer, R. M. Wakimoto, R. M. Hardesty, J. W. Wilson, and R. M. Banta, 1999: NCAR-NOAA lower tropospheric water vapor workshop. *Bull. Amer. Meteor. Soc.*, **80**, 2339–2357, https://doi.org/10.1175/1520-0477(1999)080<2339:NNLTWV>2.0.CO;2.

——, K. J. Weber, D. D. Turner, and S. M. Spuler, 2016: Validation of a water vapor micropulse differential absorption lidar (DIAL). *J. Atmos. Oceanic Technol.*, **33**, 2353–2372, https://doi.org/10.1175/JTECH-D-16-0119.1.

——, J. Hanesiak, J. W. Wilson, S. B. Trier, R. D. Roberts, W. A. Gallus Jr., S. K. Degelia, and X. Wang, 2019: Nocturnal convection initiation during PECAN 2015. *Bull. Amer. Meteor. Soc.*, **100**, 2223–2239, https://doi.org/10.1175/BAMS-D-18-0299.1.

Whitaker, J. S., and T. M. Hamill, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 3078–3089, https://doi.org/10.1175/MWR-D-11-00276.1.

Wilkinson, J. M., 2017: A technique for verification of convection-permitting NWP model deterministic forecasts of lightning activity. *Wea. Forecasting*, **32**, 97–115, https://doi.org/10.1175/WAF-D-16-0106.1.

Wulfmeyer, V., and Coauthors, 2011: The Convective and Orographically-induced Precipitation Study (COPS): The scientific strategy, the field phase, and research highlights. *Quart. J. Roy. Meteor. Soc.*, **137** (Suppl. 1), 3–30, https://doi.org/10.1002/qj.752.

——, and Coauthors, 2015: A review of the remote sensing of lower tropospheric thermodynamic profiles and its indispensible role for understanding and simulation of water and energy cycles. *Rev. Geophys.*, **53**, 819–895, https://doi.org/10.1002/2014RG000476.

——, H. S. Bauer, M. Grzeschik, A. Behrendt, F. Vandenberghe, E. V. Browell, S. Ismail, and R. A. Rerrare, 2006: Four-dimensional variational assimilation of water vapor differential absorption lidar data: The first case study within IHOP_2002. *Mon. Wea. Rev.*, **134**, 209–230, https://doi.org/10.1175/MWR3070.1.

Zeng, X., and Coauthors, 2020: Use of observing system simulation experiments in the United States. *Bull. Amer. Meteor. Soc.*, **101**, E1427–E1438, https://doi.org/10.1175/BAMS-D-19-0155.1.

Zeng, Y., A. de Lozar, T. Janjić, and A. Seifert, 2021: Applying a new integrated mass-flux adjustment filter in rapid update cycling of convective-scale data assimilation for the COSMO model (v5.07). *Geosci. Model Dev.*, **14**, 1295–1307, https://doi.org/10.5194/gmd-14-1295-2021.

Zhang, F., M. Minamide, and E. E. Clothiaux, 2016: Potential impacts of assimilating all-sky infrared satellite radiances from GOES-R on convection-permitting analysis and prediction of tropical cyclones. *Geophys. Res. Lett.*, **43**, 2954–2963, https://doi.org/10.1002/2016GL068468.

——, Y. Q. Sun, L. Magnusson, R. Buizza, S.-J. Lin, J.-H. Chen, and K. Emanuel, 2019: What is the predictability limit of midlatitude weather? *J. Atmos. Sci.*, **76**, 1077–1091, https://doi.org/10.1175/JAS-D-18-0269.1.