

Design and Implementation of a GSI-Based Convection-Allowing Ensemble Data Assimilation and Forecast System for the PECAN Field Experiment. Part I: Optimal Configurations for Nocturnal Convection Prediction Using Retrospective Cases

AARON JOHNSON

Cooperative Institute for Mesoscale Meteorological Studies, and School of Meteorology, University of Oklahoma, Norman, Oklahoma

XUGUANG WANG

School of Meteorology, University of Oklahoma, Norman, Oklahoma

(Manuscript received 22 May 2016, in final form 4 November 2016)

ABSTRACT

A real-time GSI-based and ensemble-based data assimilation (DA) and forecast system was implemented at the University of Oklahoma during the 2015 Plains Elevated Convection at Night (PECAN) experiment. Extensive experiments on the configuration of the cycled DA and on both the DA and forecast physics ensembles were conducted using retrospective cases to optimize the system design for nocturnal convection. The impacts of radar DA between 1200 and 1300 UTC, as well as the frequency and number of DA cycles and the DA physics configuration, extend through the following night. Ten-minute cycling of radar DA leads to more skillful forecasts than both more and less frequent cycling. The Thompson microphysics scheme for DA better analyzes the effects of morning convection on environmental moisture than WSM6, which improves the convection forecast the following night. A multi-PBL configuration during DA leads to less skillful short-term forecasts than even a relatively poorly performing single-PBL scheme. Deterministic and ensemble forecast physics configurations are also evaluated. Thompson microphysics and the Mellor–Yamada–Nakanishi–Niino (MYNN) PBL provide the most skillful nocturnal precipitation forecasts. A well thought out multiphysics configuration is shown to provide advantages over evenly distributing three of the best-performing microphysics and PBL schemes or a fixed MYNN/Thompson ensemble. This is shown using objective and subjective verification of precipitation and nonprecipitation variables, including convective initiation. Predictions of the low-level jet are sensitive to the PBL scheme, with the best scheme being variable and time dependent. These results guided the implementation and verification of a real-time ensemble DA and forecast system for PECAN.

1. Introduction

It has long been known that the precipitation climatology over the Great Plains region of the United States has a nocturnal maximum (Kincer 1916; Wallace 1975; Easterling and Robinson 1985; Heideman and Fritsch 1988; Colman 1990). Nocturnal convection is typically elevated (i.e., rooted well above ground level) and is often initiated, organized, and maintained by different processes than diurnal convection such as bores and other disturbances on the stable boundary layer (Parker 2008; Trier et al. 2011; Marsham et al. 2011; Coleman and Knupp 2011; Erlingsson and Barros 2014) and low-level

jets (LLJs; Carbone et al. 1990; Trier and Parsons 1993; Trier et al. 2006; Tuttle and Davis 2006; Coniglio et al. 2010; Trier et al. 2010; French and Parker 2010). The notoriously low skill of warm-season quantitative precipitation forecasting is due in large part to the low predictability of nocturnal convective initiation (CI) and subsequent evolution of nocturnal convection (Davis et al. 2003; Fritsch and Carbone 2004; Weisman et al. 2008). Improvements to the predictability of nocturnal convection will therefore require better representation of the processes across multiple scales that influence nocturnal convection.

The Plains Elevated Convection at Night (PECAN) field experiment was designed to comprehensively observe summer nocturnal convection and related features over the Great Plains (Geerts et al. 2017). One of the

Corresponding author e-mail: Dr. Aaron Johnson, ajohns14@ou.edu

primary goals of PECAN was to improve the prediction of nocturnal CI, nocturnal mesoscale convective systems (MCSs), atmospheric bores, and nocturnal LLJs “with a particular focus on the next generation convective-permitting models and advanced assimilation techniques” (Geerts et al. 2017).

Much of the literature on model configurations for predicting convection in the Great Plains focuses primarily on daytime convection or discusses only the forecast lead time, with little or no mention of the differences between daytime and nighttime results (e.g., Stensrud et al. 2000; Hou et al. 2001; Mittermaier 2007; Clark et al. 2008; Schwartz et al. 2010; Johnson et al. 2011, 2014; Johnson and Wang 2012, 2013). However, nocturnal convection and related phenomena are unique in that they are typically dependent on processes that are distinct from daytime convection. For example, idealized (Parker 2008) and real-data (Trier et al. 2006, 2011; Erlingis and Barros 2014) case studies have demonstrated the sensitivity to features such as atmospheric bores, LLJs, elevated frontal zones, mesoscale regions of elevated ascent, and land-atmosphere interactions during the previous day. Therefore, the goal of this study is to determine the sensitivity of nocturnal convection forecasts to different aspects of the ensemble data assimilation (DA) and forecast system design in order to improve its configuration for the unique foci of PECAN. In particular, this paper will focus on the configuration of the cycled ensemble Kalman filter (EnKF) radar DA (i.e., number of cycles, interval between cycles, and total length of DA) and the physics configuration of the DA and forecast ensembles. This paper presents the first work specifically focused on a systematic evaluation of convection-permitting modeling system configurations specifically for nocturnal convection. This evaluation of retrospective forecasts from 2014 focuses on morning (i.e., 1300 UTC) initialized forecasts for consistency with the real-time PECAN forecasts from 2015, which are guided by these results (Johnson et al. 2016, manuscript submitted to *Wea. Forecasting*, hereafter Part II).

Early studies using cycled EnKF for radar DA have shown a 60-min period of DA to be sufficient to analyze convective systems (e.g., Aksoy et al. 2009; Dowell et al. 2011; Yussouf et al. 2013). Yussouf and Stensrud (2012) have also shown promising results with only 30 min of radar DA. Since these past studies have taken place in a research setting where the time constraints of a real-time setting are not present, there has not been much careful examination of the marginal benefit of 60 versus 30 min of radar DA in the published literature. This study will therefore conduct such an examination. It has

been shown that cycling the radar DA at too frequent of an interval does not give the increments enough time to adjust to the model while cycling at too infrequent of an interval can degrade analysis quality if four-dimensional assimilation techniques are not used (Wang et al. 2013). This study will therefore also examine the impact of the time interval between radar DA cycles.

Another important aspect of the DA configuration is the physics parameterizations used during DA, which may be different than the optimal configurations during the longer free forecasts (e.g., Yussouf and Stensrud 2012). The impact of adding similar parameter perturbations to the microphysics scheme during radar DA as in Yussouf and Stensrud (2012) will therefore also be examined within the context of larger-scale and longer-lead-time forecasts of nocturnal convection than the observation system simulation experiment (OSSE) forecasts of a discrete supercell in Yussouf and Stensrud (2012). The impacts of using an entirely different microphysics scheme (i.e., Thompson; Thompson et al. 2008) during DA will also be evaluated, as well as the impacts of different boundary layer parameterization scheme configurations during DA.

Convective precipitation forecasts are highly sensitive to uncertainties in the parameterization schemes representing microphysical and boundary layer processes (e.g., Johnson et al. 2011; Duda et al. 2014). However, there is still limited theoretical or experimental guidance on how to optimally sample these sources of uncertainty in the convection-permitting ensemble forecast design (e.g., Duda et al. 2014). Such guidance is even sparser within the context of nocturnal convective systems. This study will use the many physics options in the Weather Research and Forecasting (WRF) Model to examine not only the best physics options for predicting nocturnal convection but also the sampling of physics error in the ensemble configuration to generate forecast spread. Emphasis will be placed on features relevant for the unique PECAN foci of nocturnal MCSs, nocturnal CI, and nocturnal LLJs. Atmospheric bore predictions will be evaluated in Part II only, given the availability of observations on sufficient time and space scales only during the 2015 PECAN field experiment.

This paper is part one of a two-part study on the design and implementation of a multiscale Gridpoint Statistical Interpolation analysis system (GSI) based EnKF DA (Johnson et al. 2015) and ensemble forecast system during PECAN. In the remainder of this paper, section 2 describes the setup of the different experiments that were conducted to study the optimal system configuration for nocturnal convection prediction. Results are presented and discussed in section 3 while section 4 contains a summary and conclusions. Part II will describe

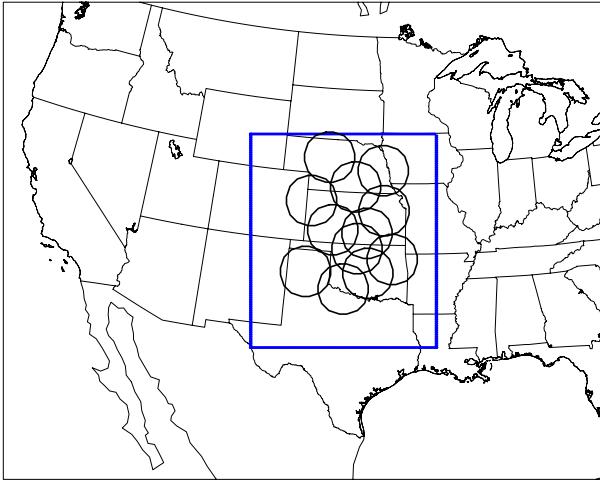


FIG. 1. Model domains used in this study. The outermost domain has a grid spacing of 12 km. The domain enclosed by the blue box has a grid spacing of 4 km.

the real-time implementation of the system for daily forecasts in support of PECAN field operations and some preliminary verification of the real-time forecasts.

2. Methods and experiments

a. GSI-based multiscale data assimilation system

The GSI-based DA system that was described and extended to a multiscale framework including direct assimilation of radar data in Johnson et al. (2015) is used in this study. The GSI-based DA system uses a 40-member ensemble based on the Advanced Research version of WRF (ARW) version 3.6.1 with the EnKF configuration described in Johnson et al. (2015). Conventional surface and upper-air observations from the data stream of the operational North American Mesoscale Forecast System at the National Centers for Environmental Prediction (NCEP), including surface and mesonet stations, Aircraft Communication, Addressing, and Reporting System (ACARS), NOAA wind profiler platforms, and radiosondes, are assimilated every 3 h on a domain approximately covering the continental United States with 12-km grid spacing (Fig. 1). For the outer domain, a 3-h observation window is used with 4D EnKF, which uses first-guess and ensemble covariance values valid at the same time as the observation, which can be different than the analysis time (Johnson et al. 2015). Additional quality control checks are also performed with the GSI part of the system to check for gross observation errors (e.g., excessive departure from first guess) and to adjust the observation error statistics of questionable observations. Increasing the observation error of a questionable observation reduces the impact that it has on the analysis

because it causes more weight to be given to the first guess. Before making such adjustments, the observation error statistics are based on observation type and altitude and are provided by a static table within the GSI package.

Nested within the 12-km domain is a convection-permitting domain with 4-km grid spacing covering the region of interest for PECAN operations on which radar observations of reflectivity and radial velocity are assimilated (Fig. 1). An observation error standard deviation of 5 dBZ and 2 m s^{-1} is assumed for reflectivity and velocity, respectively. The quality control of radar observations is described in detail in Johnson et al. (2015). In short, the Warning Decision Support System–Integrated Information (WDSS-II; Lakshmanan et al. 2007) software (available online at www.wdssii.org) is used to dealias the velocity observations and remove suspected nonmeteorological reflectivity observations. The checks for remaining gross observation errors are then also performed for the radar observations within GSI. Covariance inflation and localization are typically used with EnKF to account for sampling errors of the finite ensemble size and underrepresentation of the model error. The inflation and covariance localization parameters follow the settings described in Johnson et al. (2015).

b. Experiments on radar data assimilation configuration

Unlike the 0000 UTC initialization time in Johnson et al. (2015), an initialization time of 1300 UTC is used for the 2014 forecasts to follow the initialization time of the 2015 real-time PECAN forecasts (the subject of Part II of this study). The 1300 UTC initialization time is chosen as a compromise between assimilating as much up-to-date data as possible and producing forecast products in time to be useful to PECAN forecasters during the late morning and early afternoon. Downscaling from the outer domain is performed at 1200 UTC, followed by the inner domain radar DA for 1 h. While the motivation for this work is an evaluation of the nocturnal period of approximately 0300–0900 UTC, the forecast period between 1300 and 0300 UTC is also considered for completeness.

Given the 1300 UTC initialization time and nocturnal focus of these experiments, past studies using cycled EnKF radar DA with afternoon initialization for daytime storm prediction (e.g., Aksoy et al. 2009; Dowell et al. 2011; Yussouf and Stensrud 2012; Yussouf et al. 2013; Wang et al. 2013) may not provide directly applicable guidance on the optimal number of, and time interval between, radar DA cycles. In addition to the primarily diurnal focus of such earlier studies, the length of forecast time (~ 15 – 21 h) preceding the forecast

TABLE 1. Deterministic forecast experiments used to evaluate aspects of the radar data assimilation configuration. The asterisk after WSM6 indicates the parameter perturbations as described in the text. The “10min.6cycle” experiment is also referred to as the control (CTL) experiment.

Expt	Cycling interval (min)	Length of DA (min)	Start time of DA (UTC)	Downscale time from outer domain (UTC)	MP scheme during DA
5min.12cycle	5	60	1200	1200	WSM6*
10min.6cycle (i.e., CTL)	10	60	1200	1200	WSM6*
15min.4cycle	15	60	1200	1200	WSM6*
10min.3cycle.1200downscale	10	30	1230	1200	WSM6*
10min.3cycle.1230downscale	10	30	1230	1230	WSM6*
NOPERT	10	60	1200	1200	WSM6
THOM	10	60	1200	1200	Thompson

features of interest is also much greater than in the earlier studies (typically minutes to a few hours). Furthermore, there has not been extensive investigation of the many different combinations of these settings in the published literature, with the exception of Wang et al. (2013), who used a simplified observation system simulation experiment framework. Such an investigation is needed for implementation of the real-time cycled EnKF radar DA system because the added time and computational cost of additional cycles or more frequent cycles may not be justified within the context of other design trade-offs that may need to be made to meet operational deadlines. Therefore, the first set of experiments is designed to determine the sensitivity to, and optimal configuration of, the radar DA cycling.

The first three experiments use a fixed length of DA (60 min), with cycling intervals of 5, 10, and 15 min (5min.12cycle, 10min.6cycle, 15min.4cycle, respectively; Table 1). A 5-min observation window is used in all experiments, as also done in Wheatley et al. (2015). Two additional experiments are then conducted with only a 30-min length of DA, with and without a 30-min spinup period (10min.3cycle.1200downscale and 10min.3cycle.1230downscale, respectively; Table 1). The experiment 10min.3cycle.1200downscale is similar to 10min.6cycle in that the analyses are downscaled from the outer domain at 1200 UTC, with the difference that radar DA does not begin until after 1230 UTC in 10min.3cycle.1200downscale. In 10min.3cycle.1230downscale, the downscaling from the outer domain is done at 1230 UTC. Comparison of the two experiments with 30 min of radar DA controls for the greater spinup time of convective-scale features on the convection-allowing domain in 10min.6cycle than 10min.3cycle.1230downscale and would reveal any advantages of a 30-min spinup period before beginning radar DA.

The above experiments are all conducted with the WRF single-moment 6-class (WSM6; Hong and Lim 2006) microphysics scheme with parameter perturbations during DA. The parameter perturbations were applied to the variables listed in Table 2 of Yussouf and

Stensrud (2012) using values randomly selected from a uniform distribution. The parameter values were randomly selected once and remained fixed for all forecast days and all experiments. Yussouf and Stensrud (2012) showed the advantage of such perturbations within an OSSE context for short-term predictions of a splitting supercell. However, the suitability of such perturbations for more complex real-data analyses used to initialize forecasts over larger space and time scales remains unknown. Two additional experiments are conducted to evaluate the sensitivity to this design choice. In these experiments the WSM6 scheme with parameter perturbations (i.e., WSM6*) is compared with WSM6 without parameter perturbations (i.e., WSM6) and a different microphysics scheme, also without perturbations (Thompson microphysics). These last two experiments are labeled NOPERT and THOM in Table 1, respectively.

All of the above experiments result in analyses valid at 1300 UTC, which are used to initialize a deterministic forecast from the ensemble mean analysis. The quality of the analyses is evaluated by verification of the forecasts, which all use the same physics configuration and differ only in their initial conditions. The forecast physics configuration is Thompson’s microphysics scheme, the Mellor–Yamada–Nakanishi–Niino (MYNN; Nakanishi and Niino 2009) planetary boundary layer (PBL) scheme, the Goddard shortwave (Tao et al. 2003) and Rapid Radiative Transfer Model for GCMs (RRTMG; Mlawer et al. 1997) longwave radiation schemes, and the Noah (Ek et al. 2003) land surface model. The forecasts are on the convection-permitting 4-km domain (Fig. 1), so no cumulus parameterization is used.

In addition to the deterministic forecasts described above, three experiments are performed using different PBL configurations during DA and 20-member ensemble forecasts to evaluate the analysis quality (Table 2). The three experiments are configured with a fixed MYNN PBL scheme, a fixed quasi-normal scale elimination (QNSE; Sukoriansky et al. 2005) PBL scheme,

TABLE 2. Ensemble forecast experiments used to evaluate the PBL scheme configuration during the DA period.

Expt	Radar DA configuration	Forecast physics	PBL scheme for DA
QNSE	10min.6cycle	See Table 3	QNSE
MYNN	10min.6cycle	See Table 3	MYNN
MULTIPBL	10min.6cycle	See Table 3	See Table 3

and a multi-PBL configuration (Table 2). Ensemble forecasts are used for evaluation so that the full advantage of the multi-PBL configuration might be seen. The forecast physics configuration is the same for all three of these experiments, selected based on the forecast physics experiments below, and is shown in Table 3. The multi-PBL DA configuration also follows the PBL column of Table 3, extended to the 40-member DA ensemble by configuring the PBL schemes of members 21–40 the same as members 1–20 in Table 3.

c. Experiments on forecast physics configuration

Experiments are also conducted to understand the impacts of using different physics parameterizations during the forecast period. One goal of these experiments is to understand which physics configurations provide the best predictions of nocturnal convection and related features. A second goal of these experiments is to understand which combination of physics configurations best represents the forecast uncertainty through ensemble spread. Convective-scale forecasts can be sensitive to many aspects of the model physics, including the microphysics scheme, PBL scheme, land surface model, radiation scheme, and cumulus scheme of the model driving the lateral boundary conditions. To keep the number of experiments manageable, the focus of these experiments is on the microphysics (MP) and PBL physics parameterizations based on their demonstrated importance in convection-permitting forecasts (Johnson et al. 2011). For all of these experiments, the forecasts are initialized with the analyses from the 10min.6cycle experiment so the only differences among them are the forecast physics (Table 4).

The relative biases of different PBL schemes in WRF depend strongly on the time of day at which they are evaluated (Coniglio et al. 2013). Past studies of forecast sensitivities to the PBL scheme have focused on very specific features or applications such as dryline location (Coffer et al. 2013), tropical cyclones (Hill and Lackmann 2009), the southwest U.S. monsoon (Bright and Mullen 2002), or air quality forecasting (Hu et al. 2010). Erlingis and Barros (2014) showed differences in nocturnal MCS simulation attributable to whether Mellor–Yamada–Janjić (MYJ; Janjić 1994, 2001) or

TABLE 3. Forecast configuration for the experiments described in Table 2 and the experiment labeled MULTI in Table 4. The second through fourth columns show the MP, PBL, and cumulus parameterization (CP) schemes, respectively. The last column is also used during outer-domain DA for all experiments, and the third column is also used during DA for the experiment labeled MULTIPBL in Table 2.

Member	MP scheme	PBL scheme	CP scheme (outer domain only)
001	Thompson	MYNN	Grell-3
002	WDM6	QNSE	Kain–Fritsch
003	Lin	YSU	Kain–Fritsch
004	Thompson	ACM2	Grell–Freitas
005	WDM6	MYJ	Grell-3
006	Morrison	MYNN	Kain–Fritsch
007	Thompson	QNSE	Kain–Fritsch
008	WDM6	YSU	Grell–Freitas
009	Lin	ACM2	Grell-3
010	Thompson	MYJ	Kain–Fritsch
011	WDM6	MYNN	Kain–Fritsch
012	Morrison	QNSE	Grell–Freitas
013	Thompson	YSU	Grell-3
014	WDM6	ACM2	Kain–Fritsch
015	Thompson	MYJ	Kain–Fritsch
016	Morrison	MYNN	Grell–Freitas
017	Thompson	QNSE	Grell-3
018	Thompson	YSU	Kain–Fritsch
019	Thompson	ACM2	Kain–Fritsch
020	Thompson	MYJ	Grell–Freitas

Yonsei University (YSU; Noh et al. 2003) was the PBL scheme for a single case study. However, there has not been a systematic study over many cases and many PBL schemes of how the PBL scheme impacts nocturnal convection. One way that nocturnal convection differs from diurnal convection is that it typically occurs in the presence of a very stable PBL but may also be influenced by PBL processes from the previous day (Erlingis and Barros 2014). This study systematically evaluates the impact of PBL scheme configuration, within both deterministic and ensemble contexts, on forecasts of nocturnal convection and related features in the Great Plains. Toward this end, five deterministic forecasts are initialized from the ensemble mean analysis with the physics configuration described in section 2b (i.e., CTL in Table 4), except that each forecast uses either the MYJ, MYNN (i.e., CTL), QNSE, the Asymmetric Convective Model version 2 (ACM2; Pleim 2007), or YSU PBL scheme, all with the Thompson MP scheme.

Past studies have also shown a strong sensitivity of convection-permitting forecasts to the MP parameterization scheme (e.g., Gilmore et al. 2004; Clark et al. 2008; Johnson et al. 2011; Yussouf and Stensrud 2012; Cintineo et al. 2014; Duda et al. 2014). However, these studies have also primarily emphasized daytime, surface-based convection, which is strongly impacted by

TABLE 4. Deterministic and ensemble forecast experiments used to evaluate the physics configurations during the forecast period.

Expt	No. of members	MP scheme	PBL scheme	PBL scheme during DA
YSU	1	Thompson	YSU	MYNN
MYJ	1	Thompson	MYJ	MYNN
ACM2	1	Thompson	ACM2	MYNN
QNSE	1	Thompson	QNSE	MYNN
CTL	1	Thompson	MYNN	MYNN
MORR	1	Morrison	MYNN	MYNN
LIN	1	Lin	MYNN	MYNN
WSM6	1	WSM6	MYNN	MYNN
WDM6	1	WDM6	MYNN	MYNN
FIXED	20	Thompson	MYNN	QNSE
MULTI	20	See Table 3	See Table 3	QNSE
MULTI2	20	See Table 6	See Table 6	QNSE

the cold pool evolution. The cold pool evolution is in turn strongly influenced by the MP scheme (e.g., Dawson et al. 2015). For nocturnal convection which tends to be elevated and often owes its initiation and maintenance to features such as gravity waves and bores, rather than surface-based cold pools, the forecast sensitivity to the MP scheme has not been as well studied. Therefore, this study systematically evaluates the impact of the MP scheme configuration in a similar manner as the PBL configuration. A second set of deterministic forecasts are identical to the first set except that each forecast uses either the Thompson (i.e., CTL), WSM6, Morrison (Morrison et al. 2009), WRF double-moment 6-class; (WDM6; Lim and Hong 2010), or Lin (Lin et al. 1983) MP scheme, all with the same MYNN PBL scheme.

The experiments employing deterministic forecasts are used to choose a fixed-physics ensemble configuration as a baseline of comparison against different multiphysics ensemble configurations. The fixed-physics ensemble configuration (FIXED) uses the MYNN PBL scheme and the Thompson MP scheme. The first multiphysics ensemble (MULTI) uses the configuration shown in Table 3. The MULTI configuration was chosen based on three considerations. First, early tests indicated relatively similar performance among the different PBL schemes, so all five PBL schemes that were considered are evenly distributed across the ensemble members. Second, early tests indicated a more pronounced advantage of Thompson relative to the other microphysics schemes so the distribution of microphysics schemes is skewed toward Thompson while also including at least one member with each of the other four microphysics schemes. Third, a 10-member subensemble with Thompson microphysics and two of each PBL scheme is included to facilitate planned future studies on PBL scheme sensitivities. The second multiphysics ensemble

TABLE 5. Analysis time and date of the 20 cases from 2014 that were used for the experiments in section 3. The second column notes whether (subjectively) robust convection was ongoing within the 4-km domain during radar DA.

Analysis time	Convection ongoing during radar DA?
1300 UTC 3 Jun	Yes
1300 UTC 4 Jun	No
1300 UTC 5 Jun	Yes
1300 UTC 6 Jun	Yes
1300 UTC 8 Jun	Yes
1300 UTC 11 Jun	No
1300 UTC 14 Jun	No
1300 UTC 18 Jun	No
1300 UTC 20 Jun	No
1300 UTC 21 Jun	No
1300 UTC 22 Jun	Yes
1300 UTC 24 Jun	Yes
1300 UTC 26 Jun	Yes
1300 UTC 27 Jun	Yes
1300 UTC 28 Jun	Yes
1300 UTC 30 Jun	Yes
1300 UTC 6 Jul	No
1300 UTC 8 Jul	No
1300 UTC 9 Jul	Yes
1300 UTC 12 Jul	No

(MULTI2) provides an alternative configuration based on evenly distributing the three of the better performing PBL and MP schemes across the ensemble members (Table 6).

3. Results

The experiments in this study are performed on 20 cases selected during the period 1 June–15 July 2014 based on the presence of nocturnal convection similar to what PECAN forecasters would be interested in predicting during field operations on 1 June–15 July 2015. The selected cases are listed in Table 5. To calculate the Brier score of the deterministic forecasts, they are first converted to a neighborhood probability forecast with a radius of 48 km, which reduces the oversensitivity to small spatial forecast errors (Schwartz et al. 2010; Johnson and Wang 2012). The ensemble forecasts are also evaluated using the neighborhood ensemble probability. A Brier skill score is then calculated using a deterministic forecast with no-radar DA as a reference forecast. Statistical significance tests on the differences between experiments shown in a given figure are conducted using a one-sided permutation resampling test (Hamill 1999; Johnson and Wang 2012) and plotted at the 90% confidence levels in Fig. 2 (as well as in Fig. 8). In sections 3a and 3b, the available computational resources are used to conduct many experiments with the deterministic forecasts instead of fewer experiments

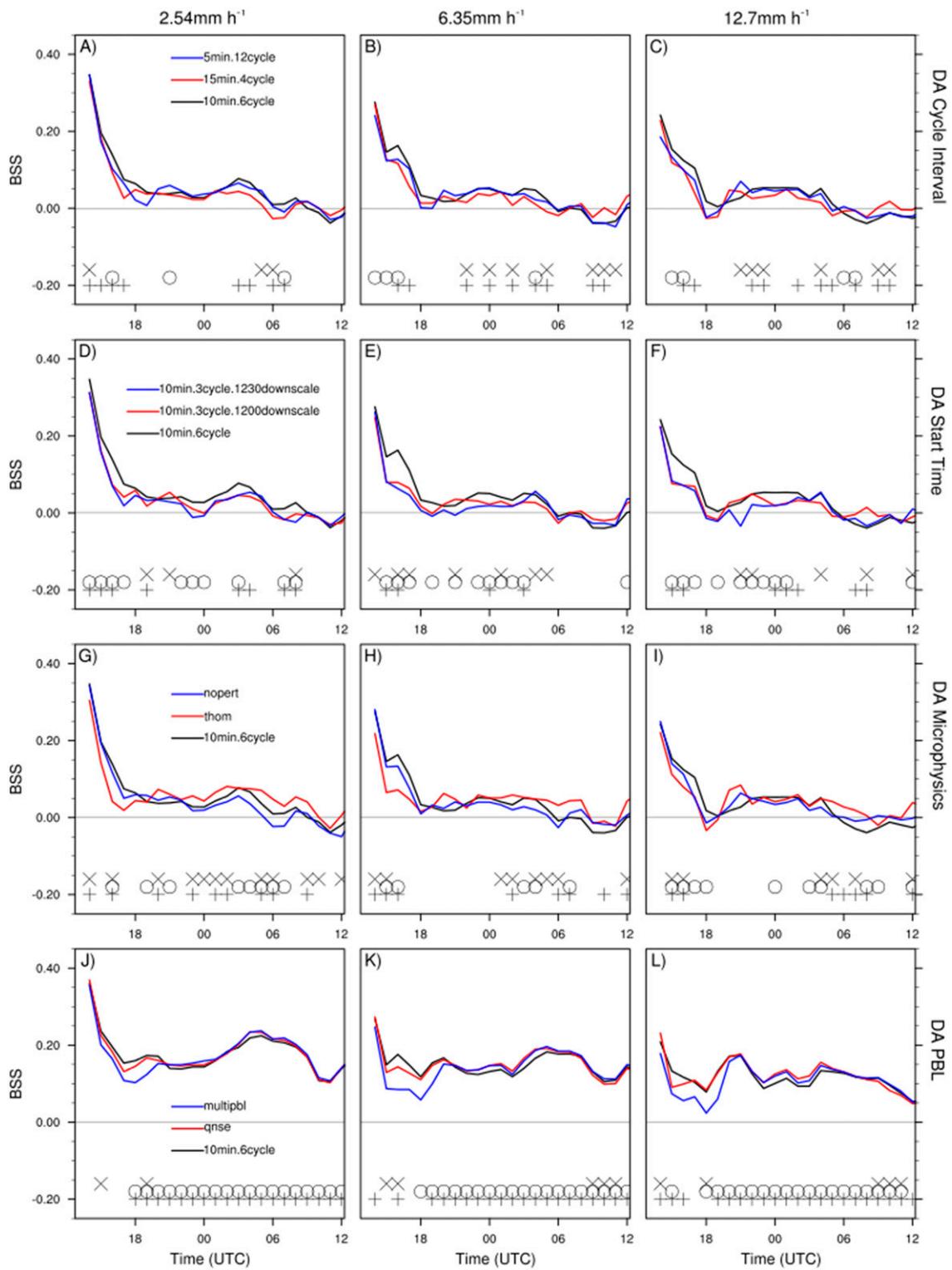


FIG. 2. BSS, using a deterministic forecast without radar DA as the reference forecast, of neighborhood probability (48-km radius) of deterministic forecasts initialized from the ensemble mean analyses from the experiments with different data assimilation configurations (Tables 1 and 2). Statistically significant differences at the 90% confidence level between the black vs red, black vs blue, and blue vs red are indicated by the plus sign, circle, and ex symbols along the bottom axes, respectively.

with a full ensemble of forecasts (Table 4). Observations of precipitation for verification are the gauge-adjusted radar-derived quantitative precipitation estimates in the National Severe Storms Laboratory next-generation QPE products (Q2; Zhang et al. 2011).

a. Radar DA configuration

1) OBJECTIVE EVALUATION USING BRIER SKILL SCORES

The first question to be answered with these experiments is whether it is even necessary to assimilate radar observations at 1200 UTC for forecasts intended to predict nocturnal convection at least 12 h later. Earlier studies have suggested that the advantage of radar DA may be too short lived to justify the computational expense (e.g., Johnson et al. 2015). However, our experiments show that although the advantage of assimilating radar data is most pronounced during the first ~6 h there is still an advantage of assimilating radar data at 1200 UTC during the nocturnal forecast period of 0300–0900 UTC, as indicated by the positive skill relative to the no-radar DA reference forecast through about 18-h lead time (i.e., 0600 UTC; Figs. 2a–c).

Another question that these experiments aim to answer is how frequently radar observations should be assimilated. In addition to the experiment with 1 h of radar DA, cycled every 10 min (i.e., 10min.6cycle), experiments were also run with the hour of radar DA cycled every 5 and 15 min (Figs. 2a–c). The experiment with 10-min cycling has the greatest skill during about the first 6 h (Figs. 2a–c). During the early nocturnal period (i.e., before ~0600 UTC) there is little difference in skill between 5- and 10-min cycling while 15-min cycling still shows slightly degraded performance (Figs. 2a–c). However, after 0600 UTC all three of the experiments have slightly less skill than the no-radar experiment, with the 15-min cycling experiment having the least negative skill at the two higher thresholds (Figs. 2a–c). Unlike the differences after 0600 UTC, the differences before 0600 UTC are consistent across many lead times and thresholds (Figs. 2a–c).

The question of how long the radar DA should continue for is addressed next with two experiments exploring the impact of the length of radar DA (Figs. 2d–f). In addition to the shorter period of DA compared with 10min.6cycle, 10min.3cycle.1230-downscale also has less time for convective scale features to “spin up” after being downscaled from the 12-km domain at 1230 UTC instead of 1200 UTC. The experiment 10min.3cycle.1200downscale controls for this difference, although it introduces a 30-min free forecast before beginning the radar DA, making both

experiments necessary. Both experiments with 30 min of radar DA have less skill than 10min.6cycle during the first ~6 h as well as much of the ~2100–0300 UTC period at lower precipitation thresholds (Figs. 2d–f). However, there is not a consistent advantage of the full hour of DA during the 0300–0900 nocturnal period (Figs. 2d–f).

The importance of the choice of microphysics scheme during radar DA is another question that we address with the next three experiments. The 10min.6cycle configuration uses the WSM6 microphysics scheme with parameter perturbations following Yussouf and Stensrud (2012). Additional experiments are also conducted without the parameter perturbations and with a different (Thompson; experiment referred to as THOM) microphysics scheme to evaluate the impact of this design choice (Figs. 2g–i). The reader will recall that all of the forecasts here used the same forecast physics configuration (Thompson for microphysics), so the DA configuration impacts are evaluated through the differences in the resulting ensemble mean analysis as it is used to initialize a deterministic forecast. The WSM6 experiment without parameter perturbations (i.e., NOPERT; Figs. 2g–i) results in generally lower skill than the experiment with the perturbations at most lead times for the 2.54 and 6.35 mm h⁻¹ thresholds (10min.6cycle; Figs. 2g,h) and at early lead times only for the 12.7 mm h⁻¹ threshold (Fig. 2i). The THOM experiment (Figs. 2g–i) also results in lower skill at early lead times, but greater skill at later lead times for most thresholds and lead times (Figs. 2g–i). The Thompson experiment is most skillful during the nocturnal period of ~0300–0900 UTC (Figs. 2g–i). This result is not changed by using WSM6, instead of Thompson, as the forecast microphysics scheme for all experiments (not shown). All of the above experiments are performed with MYNN as the PBL scheme during DA, based on its superior performance in initial test cases.

The final set of experiments in this section is designed to address the question of how the PBL parameterization ensemble should be configured during DA on both the outer and inner domains (Table 2). The ensemble forecast skill at early (i.e., ~6 h) lead times is lowest for MULTIPBL and greatest for MYNN and QNSE (Figs. 2j–l). To test the hypothesis that the lower performance of the MULTIPBL ensemble is due to including relatively low-skill PBL schemes, an additional experiment with ACM2 as the PBL scheme was also conducted (not shown). ACM2 was selected because it is one of the poorest-performing forecast PBL schemes at ~1200 UTC in our experiments (Figs. 8d–f). The ACM2 experiment deterministic forecasts were also more skillful than the

MULTIPBL deterministic forecasts at these lead times (not shown¹). Therefore, although a multi-PBL scheme configuration may be advantageous during the forecast period, it is actually disadvantageous during the DA period. During the early part of the nocturnal period (i.e., ~0000–0600 UTC) the ensemble forecast skill is generally lowest for MYNN and greatest for QNSE and MULTIPBL (Figs. 2j–l). Thus, QNSE is the only one of the three DA configurations that provides analyses leading to forecasts that perform relatively well for both the shorter- and longer-term forecasts.

Many of the above differences are statistically significant at the 90% level at some lead times but not others. This is likely due to the limited sample size of 20 cases. Although 20 cases provided much more useful information for designing the real-time ensemble configuration than a single or a few case studies, multiple seasons of data would likely be needed to obtain highly significant results. In particular, the degraded performance of MULTIPBL at early lead times in Figs. 2j–l is not significant at the 90% confidence level at most lead times, likely because of a large number of cases without active convection at this time of day. This result becomes more uniformly significant at the 80% level (not shown). At later times, the degraded performance of 10min.6cycle is not as large, but is highly significant because of the larger sample size resulting from greater convective activity at this time of day.

2) SUBJECTIVE EVALUATION USING A REPRESENTATIVE CASE STUDY

The forecasts initialized at 1300 UTC 6 June 2014 are used to qualitatively understand some of the systematic quantitative differences discussed above (Figs. 3 and 4). This case is chosen because the differences in quantitative skill in this case are similar to those in Fig. 2 (not shown). Figures 3 and 4 show neighborhood probability forecasts of hourly accumulated precipitation exceeding 6.35 mm in order to directly relate them to the objective metrics in Fig. 2. Our conclusions are not substantially changed by considering details such as convective mode and structure (not shown). In the experiment without radar DA, the convection ongoing at early lead times is absent in the forecast (Fig. 3a). There is also spurious precipitation that develops in northern Oklahoma by 0000 UTC without radar data assimilation (Figs. 3b,c). The nocturnal MCS that develops the following night in Oklahoma is also largely

absent from the NORADAR forecast at 0600 UTC (Fig. 3d). Figures 3e–p show the impacts of the different cycling intervals for this case. Although the differences are subtle among these experiments, during the first forecast hour 15min.4cycle has lower probabilities than 10min.6cycle and 5min.12cycle in central Oklahoma while 5min.12cycle has a maximum in the southern storm that is displaced slightly west of the other forecasts and observations (Figs. 3e,i,m). By 0000 UTC, 10min.6cycle has a maximum in southwestern Kansas that is closer to the observed precipitation location than do 5min.12cycle and 15min.4cycle (Figs. 3f,j,n). During the nocturnal period at 0300 UTC, all three forecasts have two precipitation maxima in northern Oklahoma: one in or near the observed precipitation and one to the east of the observed precipitation (Figs. 3g,k,o). However, at this time 10min.6cycle shows the western maximum to be more dominant than the spurious eastern maximum (Fig. 3k), unlike 5min.12cycle (Fig. 3g). The 15min.4cycle experiment shows even the western maximum too far east (Fig. 3o). By 0600 UTC, 15min.4cycle has no overlap between the forecast probability and observed precipitation (Fig. 3p), 5min.12cycle has a prominent orientation angle error (Fig. 3h), and 10min.6cycle has a prominent displacement error (Fig. 3i). Therefore, like the systematic results, 10min.6cycle is the best compromise between too frequent and too infrequent radar DA cycling, at least for the forecasts during the morning and early nocturnal periods (Figs. 3i–k). Comparison of Figs. 3i–l with Figs. 4a–h shows a better-developed MCS during the first forecast hour with the full hour of radar DA, compared to only 30 min (Fig. 3i vs Figs. 4a,e). The advantage of the full hour of radar DA in this case is also evident throughout the nocturnal period (Figs. 3k,l vs Figs. 4c,d,g,h). There is very little difference between 10min.6cycle and NOPERT during the first forecast hour, except for slightly higher probabilities in 10min.6cycle where precipitation is also observed in central Oklahoma (Fig. 3i vs Fig. 4i). At this time, THOM is clearly inferior to 10min.6cycle and NOPERT (Figs. 3i and 4i vs Fig. 4m). However, during the early nocturnal period (i.e., 0300 UTC) in this case the probability maximum for THOM is closer to the observed MCS than both 10min.6cycle and NOPERT, and THOM does not have a second spurious maximum farther east (Fig. 4o vs Figs. 3k and 4k). However, in this case there is a slight advantage of 10min.6cycle over NOPERT during the nocturnal period since the 10min.6cycle forecasts have greater overlap with the observation at 0300–0600 UTC, and less spurious precipitation at 0300 UTC, compared with NOPERT (Figs. 3k,l vs Figs. 4k,l).

¹ ACM2 is not added to the ensemble verification in Figs. 2j–l because of limited computational resources.

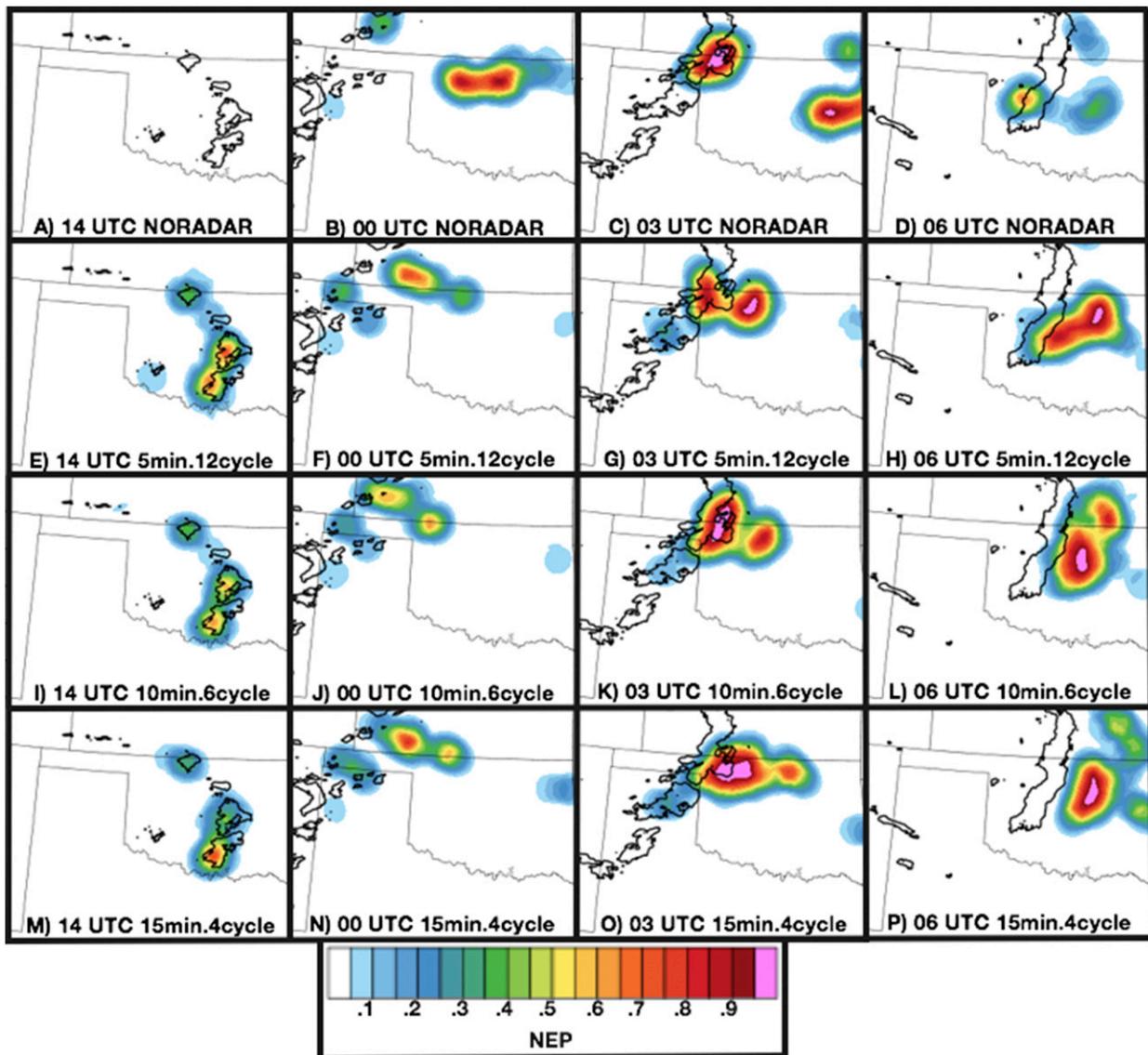


FIG. 3. Neighborhood probability forecasts (color scale) initialized at 1300 UTC 6 Jun 2014 of hourly precipitation at the 6.35 mm h^{-1} threshold from the NORADAR experiment at lead times of (a) 1 h (1400 UTC), (b) 11 h (0000 UTC), (c) 14 h (0300 UTC), and (d) 17 h (0600 UTC). Also shown are the 5min.12cycle forecasts at (e) 1400, (f) 0000, (g) 0300, and (h) 0600 UTC; the 10min.6cycle forecasts at (i) 1400, (j) 0000, (k) 0300, and (l) 0600 UTC; and the 15min.4cycle forecasts at (m) 1400, (n) 0000, (o) 0300, and (p) 0600 UTC. Black lines are the observed 6.35 mm h^{-1} precipitation contour.

It is somewhat counterintuitive that the analyses generated with Thompson microphysics during DA result in less skillful forecasts at early lead times but more skillful forecasts at later lead times, compared with analyses generated with WSM6 microphysics during DA (Figs. 2g–i). The forecasts from each case were further evaluated subjectively to better understand this difference. The forecasts initialized at 1300 UTC 6 June 2014 are representative of the general trend that is seen systematically in Figs. 2g–i. While both forecasts failed to evolve the morning

convection exactly as observed at early lead times the precipitation in central Oklahoma, which was not observed, is less extensive in the 10min.6cycle forecast than the THOM forecast (Figs. 5a,b vs Figs. 5e,f). While both forecasts were also too far east with the MCS that developed the following evening, the nocturnal MCS forecast by THOM at 0300 UTC is closer to the observed MCS than the MCS forecast by 10min.6cycle (Fig. 5h vs Fig. 5d). It is not clear whether the eastward bias that is common to both experiments results from an initial condition error or a physics-related

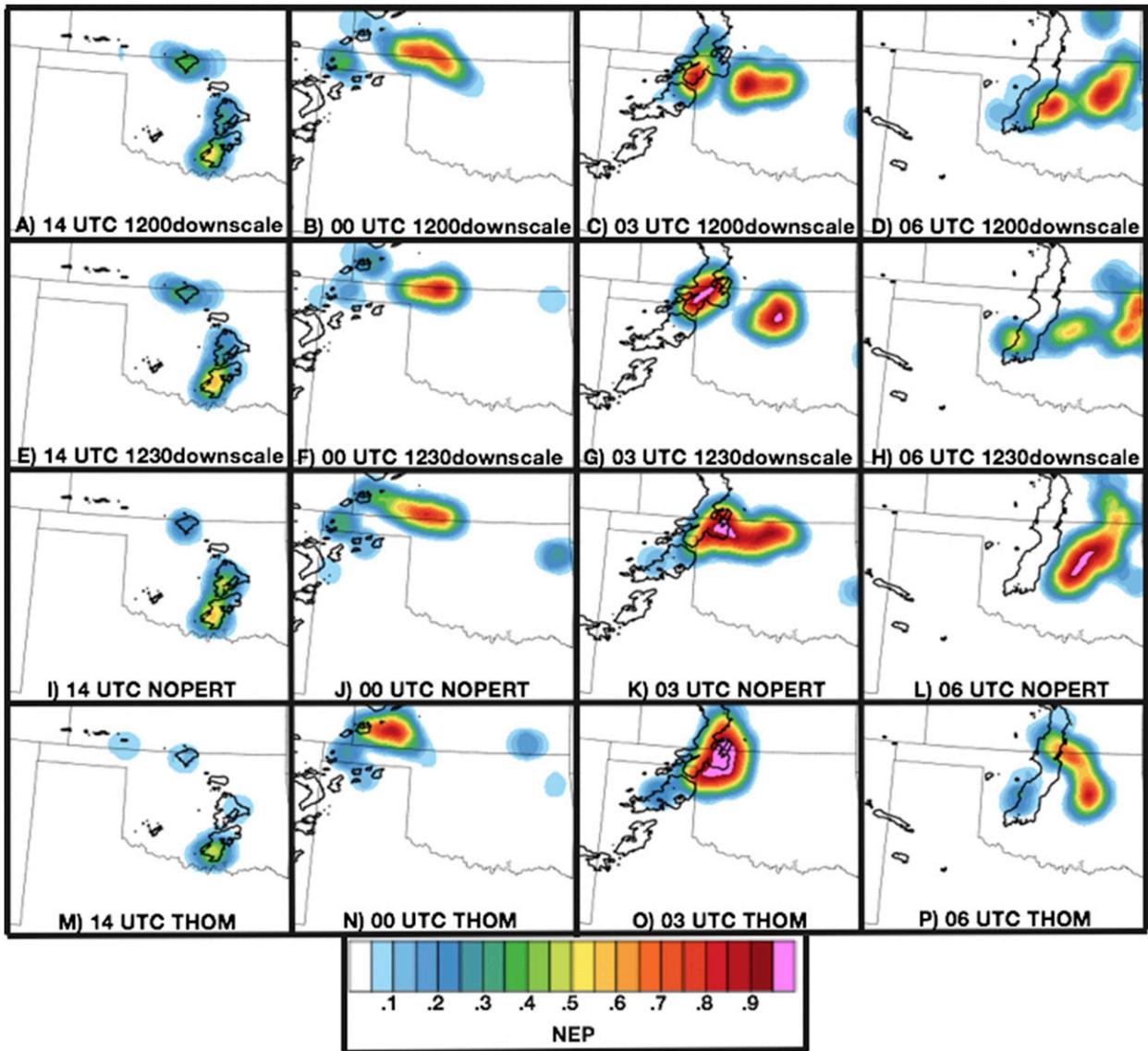


FIG. 4. As in Fig. 3, but for (a)–(d) the 10min.3cycle.1200downscale forecast, (e)–(h) the 10min.3cycle.1230downscale forecast, (i)–(l) the NOPERT forecast, and (m)–(p) the THOM forecast.

error, such as the PBL scheme that is common to both forecasts.

The reason for the better nocturnal MCS forecast when Thompson is the microphysics scheme for DA can be seen by comparing the surface dewpoints in the two forecasts (Fig. 6). The surface air mass in northwestern Oklahoma (Figs. 6d,i; blue circles), upstream of the spurious convection in 10min.6cycle, has dewpoints around 58°–62°F at 2100 UTC in THOM while 10min.6cycle dewpoints are about 62°–66°F. The 10min.6cycle dewpoints are too moist, compared with the observed dewpoints of ~61°F in this area (not shown). It is in this region of excess moisture where the

10min.6cycle forecast develops spurious convection in advance of the observed MCS (Figs. 5c,d). The low-level air that had been advected into northwest Oklahoma and southwest Kansas by 1800–2100 UTC originated from an area of ongoing convection at 1300 UTC (Figs. 5 and 6). The 1300 UTC analysis generated with Thompson microphysics had a shallow moist layer with much drier air aloft than the analysis generated with WSM6. The drier air in THOM was then mixed down to the surface as the boundary layer deepened. In this case (Fig. 6) and several others (not shown), Thompson microphysics during DA provided a more accurate analysis of the impacts on the mesoscale

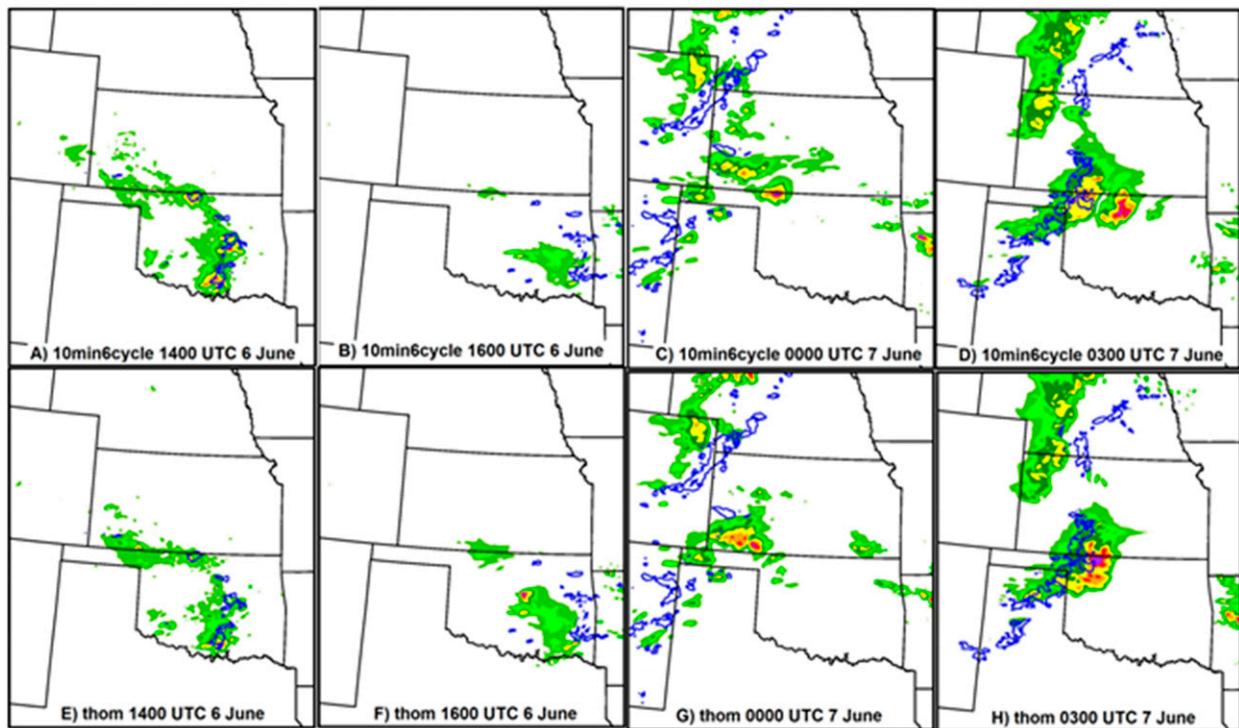


FIG. 5. Hourly accumulated precipitation forecasts for the (a)–(d) 10min.6cycle and (e)–(h) THOM experiments initialized at 1300 UTC 6 Jun 2014, with the 12.7 mm h^{-1} observation contour overlaid in blue.

environment of the convection ongoing during the DA period before 1300 UTC. This resulted in a better forecast of the environment in which convection developed the following evening. The better performance of 10min.6cycle over THOM at earlier lead times, which corresponds to early morning forecasts, was due to the overprediction of weak stratiform precipitation in the Thompson analyses and the underprediction of heavier convective precipitation in the Thompson analyses. However, since this early convection generally dissipates or moves out of the domain before the following evening, it is mainly the impact of that convection on the mesoscale environment that dominates the forecast of new convection the following night.

The difference between THOM and 10min.6cycle for radar DA is further illustrated by the impacts on specific hydrometeor species, especially rain and snow (Fig. 7). Figure 7 shows the ensemble mean background (solid) and analysis (dashed) profiles of rain and snow hydrometeor mixing ratios, averaged over a 40×40 grid point region of central Oklahoma for the 6 June case. Consistent with the findings of previous studies (e.g., Wheatley et al. 2014), Thompson produces more snow hydrometeors than WSM6 (e.g., Fig. 7b). After spinning up the hydrometeors during

the first few cycles (not shown), each experiment reaches a stable amount of snow hydrometeors for its microphysics scheme that is only slightly adjusted by the radar DA (Figs. 7b,d,f). However, for rain hydrometeors in THOM, the radar DA reduces the mixing ratios below about 800 mb during each cycle; then, the mixing ratio increases again in the next background forecast (Figs. 7a,c,e). This is only seen for 10min.6cycle at very low levels below about 975 mb (Figs. 7a,c,e). Therefore, the THOM radar DA cycles are a greater sink of rainwater, and thus environmental moisture because the rain is then regenerated during each forecast cycle, than 10min.6cycle. This is consistent with the weaker convective cores, and drier low-level mesoscale environment, in the THOM experiment.

b. Forecast physics configuration

The general question that is addressed in this section is how to choose an ensemble of microphysics and boundary layer parameterizations during the forecast period for nocturnal convection and related features. The impacts of the different physics configurations outlined in Table 4 are evaluated in this section within the specific contexts of predicting nocturnal MCSs, atmospheric bores, nocturnal LLJs, and nocturnal CI.

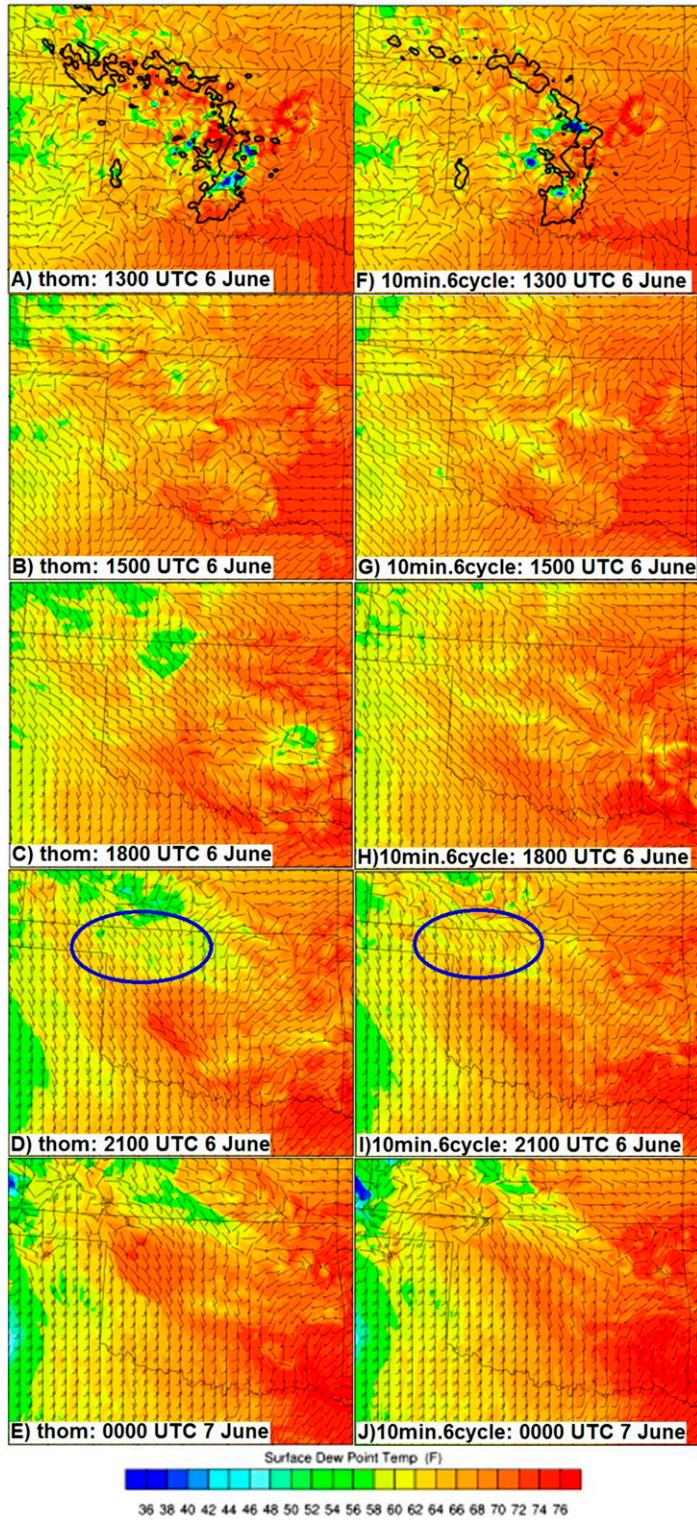


FIG. 6. Surface dewpoint forecast for the (a)–(e) THOM and (f)–(j) 10min.6cycle forecasts initialized at 1300 UTC 6 Jun 2014. The location mentioned in the text is circled in blue in (d) and (i), and analyzed reflectivity is contoured at the 35-dBZ threshold in black in (a) and (f).

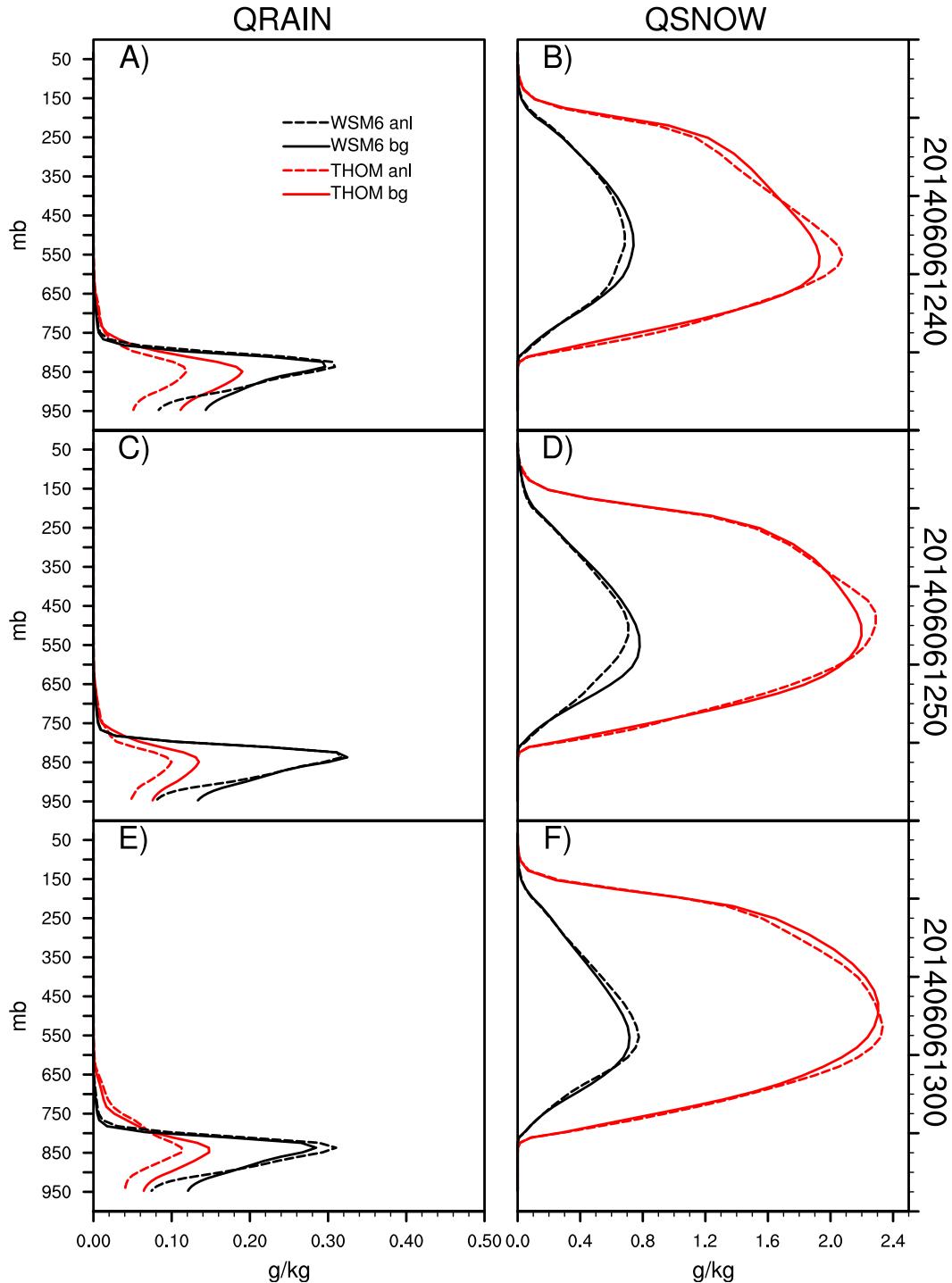


FIG. 7. Ensemble mean background (solid) and analysis (dashed) profiles of (left) QRAIN and (right) QSNOW (rain and snow hydrometeor mixing ratios, respectively) averaged over a 40×40 grid point region of central OK for the 6 Jun case at (a),(b) 1240, (c),(d) 1250, and (e),(f) 1300 UTC.

The deterministic forecast experiments are initialized with the same analyses as 10min.6cycle for consistency with the results in [section 3a](#). In other words, the deterministic forecasts are initialized from the analyses

generated with MYNN as the PBL scheme during DA. However, the ensemble forecasts are generated with the analyses from the QNSE experiment in [Table 2](#) because of that experiment's superior performance

(Figs. 2j,k,l) so that the results will be directly applicable to the configuration of the real-time PECAN ensemble (presented in Part II).

1) PRECIPITATION VERIFICATION AS A PROXY FOR MCS PREDICTION

Quantitative precipitation forecasts are first evaluated as a proxy for nocturnal MCS predictability. The objective precipitation forecast verification during the overnight hours was found to be consistent with subjective comparisons of the nocturnal MCS predictions (not shown). The precipitation forecast skill of the deterministic forecasts with different microphysics schemes is shown in Figs. 8a–c for both the night 1 and night 2 forecasts (i.e., out to 48-h lead time), consistent with the forecast lead times required for the PECAN field experiment. This differs from the more common convection-allowing forecast evaluations of 1-day forecasts, which also tend to focus on afternoon rather than nocturnal convection (e.g., Clark et al. 2012; Johnson and Wang 2012; Duda et al. 2014). In Fig. 8, significance tests are conducted for differences from the CTL experiment (see also Fig. 8 caption). The Thompson scheme in the CTL experiment has a clear advantage over the others during the early nocturnal period from ~0300 to 0600 UTC for both the night 1 and night 2 forecasts (Figs. 8a–c), especially at the higher threshold (Fig. 8c). This result is consistent with the results in Duda et al. (2014), using springtime forecasts from a different year. During the night 2 forecasts, Thompson only has an advantage for the higher thresholds (Figs. 8b,c), where Thompson has no skill relative to the NORADAR forecast, while the other microphysics schemes perform even worse than the NORADAR forecast. The precipitation forecast skill of the deterministic forecasts with different PBL schemes is shown in Figs. 8d–f. Again there is a clear advantage during much of the ~0300–0900 UTC nocturnal period; this time for the MYNN PBL scheme in the CTL forecast for both night 1 and night 2 (Figs. 8d–f). These results support the choice to pair Thompson microphysics with the MYNN PBL scheme in the control configuration for nocturnal MCS prediction. These differences are generally highly significant during the overnight hours (Figs. 8a–f).

Two different ensemble physics configurations are evaluated in addition to a fixed-physics ensemble where all members use the most skillful MP and PBL scheme configuration as determined in the preceding paragraph (FIXED, MULTI, and MULTI2 in Table 4). The differences in precipitation forecast skill among these ensembles are generally small, although often statistically significant (Figs. 8g–i). At certain times during both the

nocturnal (e.g., ~0600) and daytime (e.g., ~1800 UTC) periods the fixed-physics ensemble is noticeably less skillful than both multiphysics ensembles (Figs. 8g–i). The differences between the two different multiphysics ensembles are generally less than the differences between the fixed-physics and multiphysics ensembles (Figs. 8g–i).

A forecaster using the ensemble precipitation forecasts to predict nocturnal MCSs would not only be interested in the overall skill of the forecasts, but also in whether a range of potential outcomes is reflected in the ensemble spread. One way of quantifying the ensemble spread is with the correspondence ratio (CR; Stensrud and Wandishin 2000). In short, the CR is the ratio of the ensemble intersection area to the union area. The union area is the number of grid points where any ensemble member forecasts a threshold to be exceeded and the intersection area is the number of grid points where a prespecified number of members agree in the forecast of the threshold being exceeded. Thus, a CR of 1.0 indicates a minimum of spread since all members agree wherever precipitation is forecast while a CR of 0.0 indicates a maximum of spread since the members do not agree on where to forecast precipitation. The CR (Fig. 9) shows a more pronounced difference between the fixed-physics and multiphysics configurations than the Brier skill score (BSS; Figs. 8g–i). During the nocturnal period, the fixed-physics ensemble has greater ensemble agreement (i.e., less spread) than the multiphysics ensembles for a variety of thresholds and number of members used to determine ensemble agreement (Fig. 9). Thus, the impact of the ensemble physics configuration on forecast spread is particularly important during the nocturnal period of interest for PECAN. The differences between the two multiphysics ensembles are again generally smaller than the differences between the fixed and multiphysics ensembles, with MULTI2 showing slightly more spread than MULTI at the 2.54 mm h⁻¹ threshold and MULTI showing slightly more spread than MULTI2 at the 12.7 mm h⁻¹ threshold (Fig. 9). Although the CR does not reveal how much spread is needed to fully capture the forecast uncertainty, subjective analysis of individual cases can help to answer this question.

Figure 10 provides a subjective example of the differences in spread among MULTI, MULTI2, and FIXED. The observations at 0600 UTC 27 June 2014 show an MCS along the western part of the Kansas–Nebraska border and a cluster of storms in northwestern Missouri (Fig. 10d). FIXED shows strong clustering of the forecasts in west-central Nebraska and in southeastern Nebraska (Fig. 10a). MULTI and MULTI2

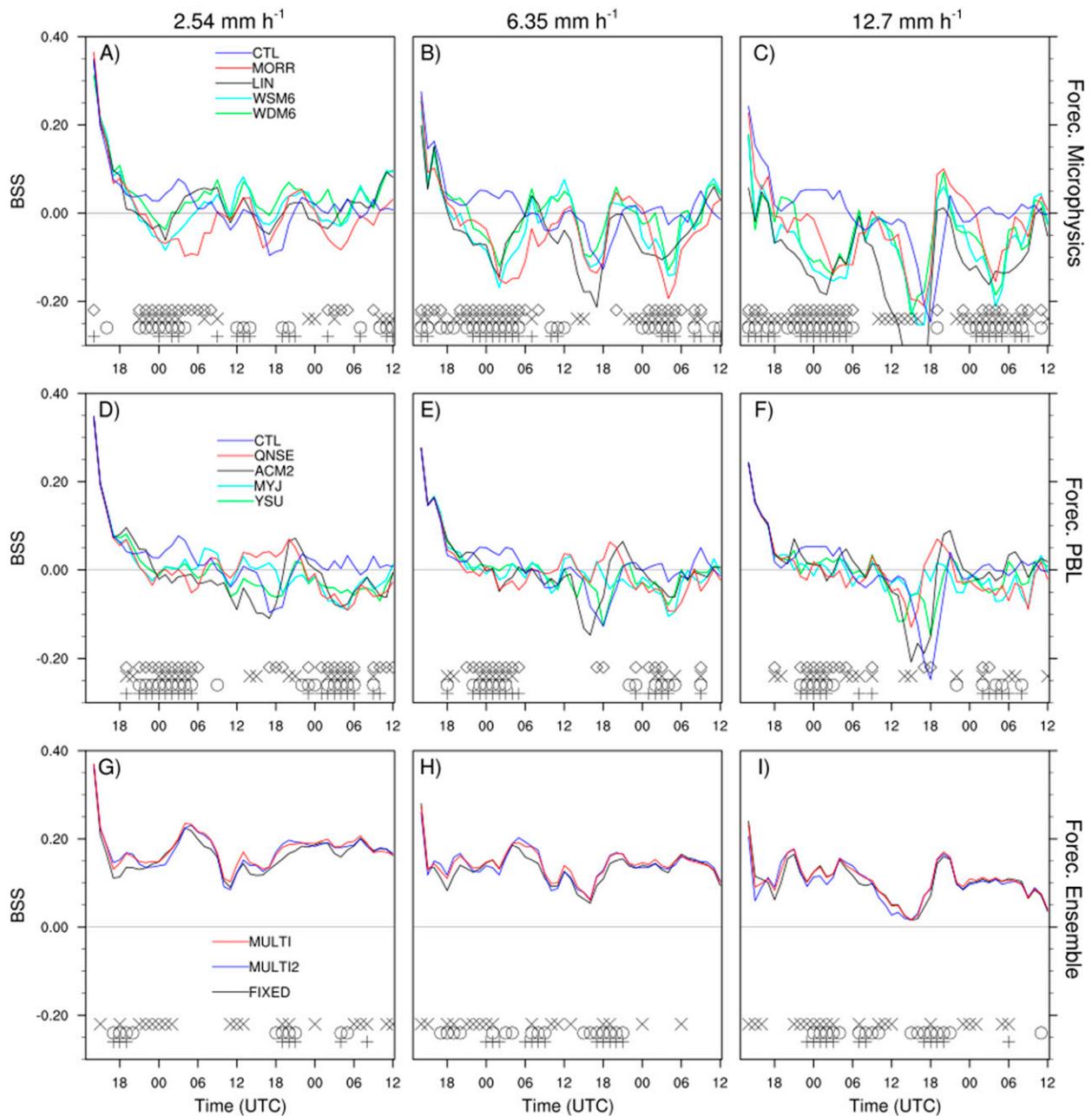


FIG. 8. As in Fig. 2, but for the physics configuration experiments in Table 4. The same reference forecast is used for both figures in the BSS calculation. Significant differences at the 90% confidence level in (a)–(f) for green vs blue, cyan vs blue, black vs blue, and red vs blue are indicated, respectively, by plus sign, circle, ex, and diamond symbols along the bottom axes. Statistical significance in (g)–(i) is indicated for black vs blue, black vs red, and red vs blue the by plus sign, circle, and ex symbols, respectively.

show weaker clustering and greater spread in the location of features in both of these areas (Figs. 10b,c). The greater spread in MULTI and MULTI2 is more consistent with the location errors of the subjective center of the ensemble distributions on this case than the strong clustering in FIXED. In other words, while all three ensembles indicate approximately the same places as the

most likely locations of the two storm clusters, the greater spread in MULTI and MULTI2 better reflects the location error in the ensemble mean forecast. The spread in MULTI is also subjectively larger than the spread in MULTI2 (Figs. 10b,c), likely because of the greater number of different physics schemes in the MULTI ensemble than the MULTI2 ensemble (Table 3 vs Table 6).

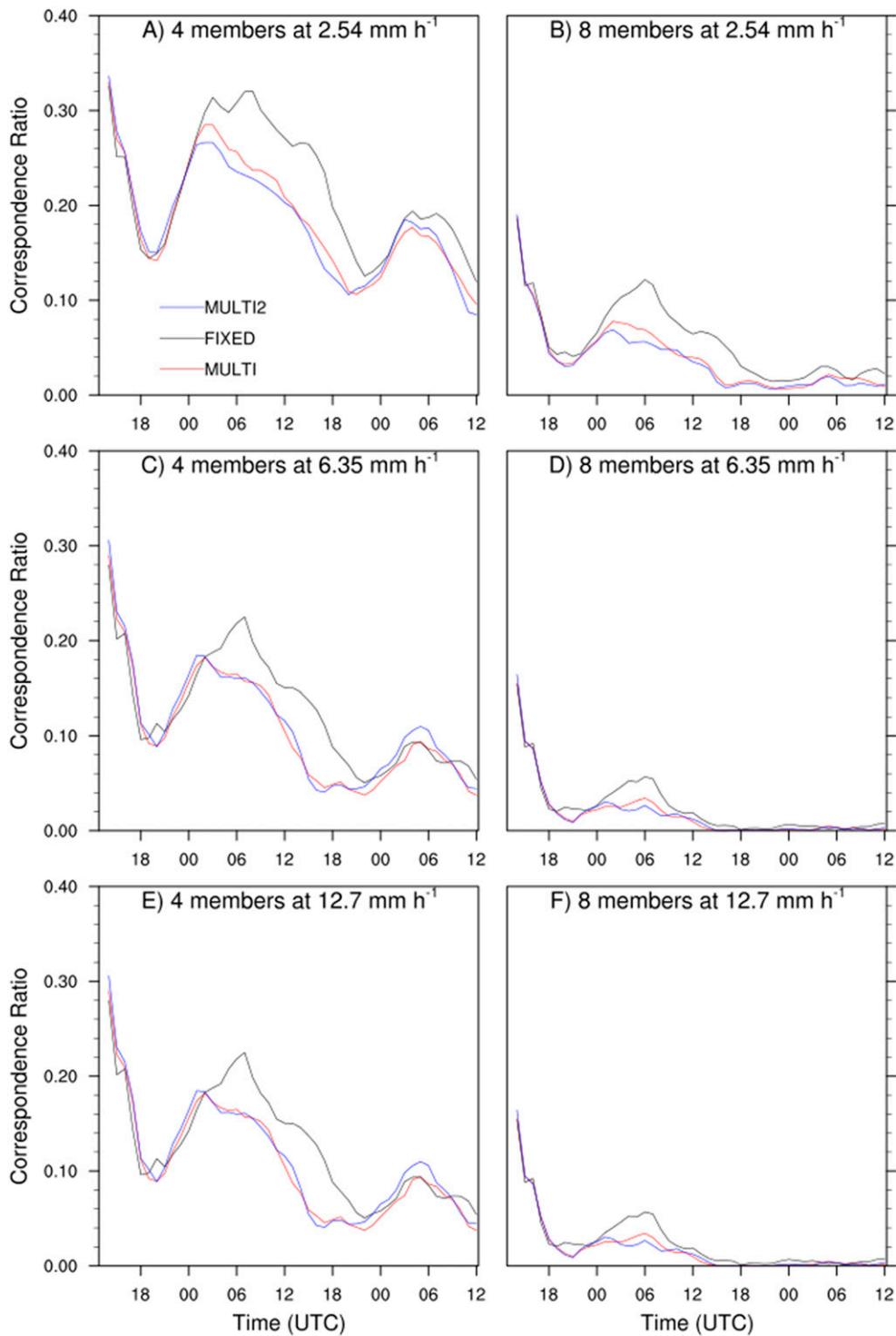


FIG. 9. CR for (a) 2.54 mm h⁻¹ four-member agreement, (b) 2.54 mm h⁻¹ eight-member agreement, (c) 6.35 mm h⁻¹ four-member agreement, (d) 6.35 mm h⁻¹ eight-member agreement, (e) 12.7 mm h⁻¹ four-member agreement, and (f) 12.7 mm h⁻¹ eight-member agreement.

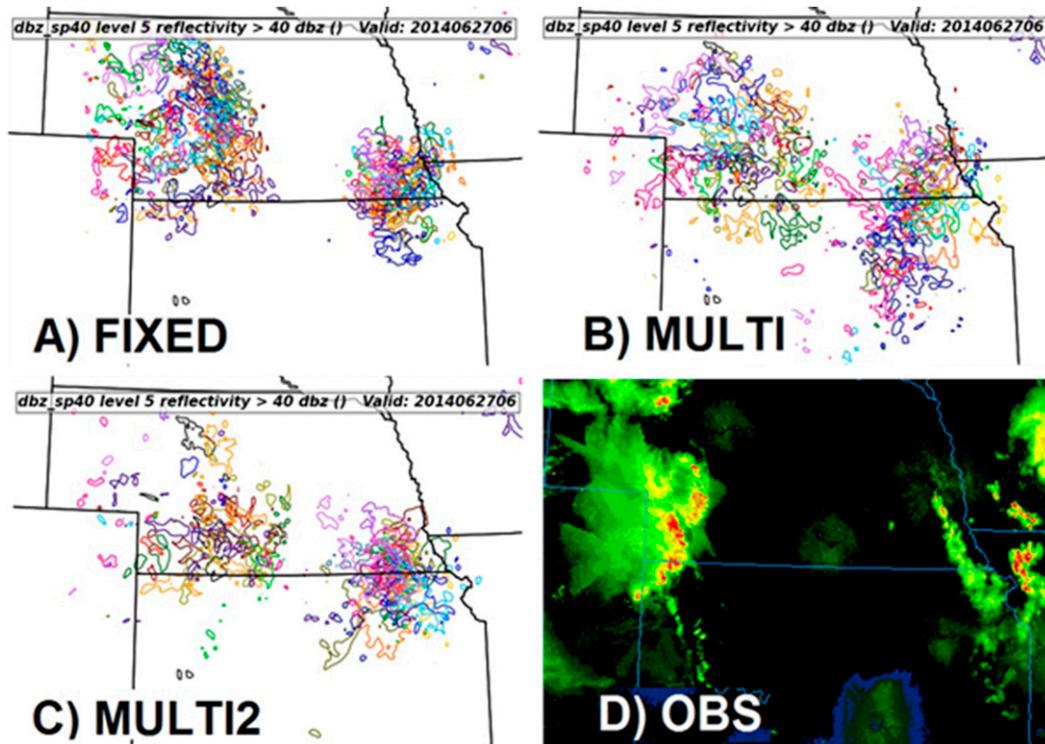


FIG. 10. Ensemble spaghetti plots of the 40-dBZ model level 5 (~ 900 hPa) reflectivity contour from all members overlaid on the same plot for the (a) FIXED, (b) MULTI, and (c) MULTI2 ensemble forecasts initialized at 1300 UTC 26 Jun 2014 and valid at 0600 UTC 27 Jun. (d) Also shown is the observed NEXRAD reflectivity mosaic at 0600 UTC 27 Jun.

Since the observed storms in northwest Missouri are farther east than all three ensemble distributions for this case (Fig. 10d), the MULTI ensemble forecast spread may still be undersampling the forecast uncertainty.

2) SUBJECTIVE VERIFICATION OF NOCTURNAL CI PREDICTION

Given the difficulty of objectively identifying CI in the observations and forecasts (e.g., Kain et al. 2013) and the limited number of pristine CI events in the dataset, the ensemble forecasts are subjectively evaluated for the apparently elevated nocturnal CI within environments undisturbed by ongoing convection. Three events that are representative of this subjective analysis are shown in Fig. 11. These events are chosen to represent both night 1 and night 2 forecasts, since the PECAN forecasters were interested in the forecasts for both of those nights. In the first event, a line of convection was observed to initiate between ~ 0600 and 0900 UTC (Fig. 11d), following the expected pattern of elevated nocturnal CI. The FIXED ensemble is the only configuration that does not have any members forecasting this

nocturnal CI event (Figs. 11a–c). Between the two multiphysics ensembles, MULTI (Fig. 11a) has more members forecasting the CI event. In the second event, a similar but less extensive and shorter-duration line of elevated convection was observed to initiate during the overnight hours (Fig. 11h). This event is the only one of the three events in Fig. 11 to occur during the second night forecast (i.e., lead time ~ 42 h), further reducing the expected predictability of the event. In this event, only the MULTI ensemble shows any indication of nocturnal CI. These first two events are representative of other similar cases in that the possibility of nocturnal elevated CI, a typically low predictability event, is much better forecast by the MULTI ensemble as a result of the greater number of physics configurations being more likely to have some members reflect the conditions leading to the CI event. Studying the driving mechanism for nocturnal CI events is an active research topic in PECAN and is left for future papers. The third event is a more atypical nocturnal CI event in that it does not follow the typical pattern for nocturnal CI of elevated cells oriented parallel to the cold side of a warm or stationary front and perpendicular to a

TABLE 6. Forecast physics configuration used for the experiment labeled MULTI2 in Table 4. Column labels are the same as in Table 3.

Member	MP scheme	PBL scheme	CP scheme (outer domain only)
001	Thompson	MYNN	Grell-3
002	Lin	YSU	Kain-Fritsch
003	WDM6	MYJ	Kain-Fritsch
004	Thompson	YSU	Grell-Freitas
005	Lin	MYJ	Grell-3
006	WDM6	MYNN	Kain-Fritsch
007	Thompson	MYJ	Kain-Fritsch
008	Lin	MYNN	Grell-Freitas
009	WDM6	YSU	Grell-3
010	Thompson	MYNN	Kain-Fritsch
011	Lin	YSU	Kain-Fritsch
012	WDM6	MYJ	Grell-Freitas
013	Thompson	YSU	Grell-3
014	Lin	MYJ	Kain-Fritsch
015	WDM6	MYNN	Kain-Fritsch
016	Thompson	MYJ	Grell-Freitas
017	Lin	MYNN	Grell-3
018	WDM6	YSU	Kain-Fritsch
019	Thompson	MYNN	Kain-Fritsch
020	Lin	YSU	Grell-Freitas

strong LLJ (Fig. 11i). All three ensembles successfully predicted this event to occur (Figs. 11i–k). However, while the observed convection shows some embedded linear and multicellular clusters (Fig. 11i), the FIXED ensemble shows very cellular convection in all members (Fig. 11k). In this case the greater “spread” of the multiphysics ensembles is manifested as greater diversity in the types of convective structures within the region of nocturnal CI in MULTI2 (Fig. 11j) and, especially, in MULTI (Fig. 11i).

Another case of nocturnal elevated CI, is selected to demonstrate typical differences among the deterministic forecasts in Table 4 within the context of nocturnal CI (Fig. 12). In this case, a line of elevated convection was observed to initiate between ~0600 and 0700 UTC (Figs. 12 γ , δ , ϵ). The deterministic forecasts with different MP schemes all show CI in approximately the same location and at the same time (between 0600 and 0700 UTC) as each other (Figs. 12a–c and 12p– β). The main difference among the forecasts with different MP schemes is the intensity of the convection. The deterministic forecasts with different PBL schemes show less variation in the intensity of the convection but more variation in the time of CI (Figs. 12a–o). The time of CI ranges from before 0600 UTC in QNSE (Figs. 12e–f) and MYJ (Figs. 12j–l) to after 0700 UTC in ACM2 (Figs. 12g–i). There is also more variation in the orientation of the convective line, ranging from a more east–west

orientation in CTL (Fig. 12c) to a more north–south orientation in QNSE (Fig. 12f).

3) SOUNDING VERIFICATION AS PROXY FOR LLJ PREDICTION

The ensemble forecasts are also verified against routine 0000 and 1200 UTC National Weather Service (NWS) soundings (e.g., Fig. 13). This verification is useful not only as a complement to the precipitation verification in section 3b(1), but also as a way to understand how well the different physics configurations predict nocturnal LLJs and the nocturnal environments in which they form. The FIXED ensemble shows generally larger RMSEs and generally smaller spread than the multiphysics ensembles at low levels for temperature, moisture, and wind at both the 0000 and 1200 UTC verification times (Fig. 13). All ensembles are generally overdispersive for moisture at low levels (Figs. 13c,d) and underdispersive for temperature at low levels (Figs. 13a,b). The differences between MULTI and MULTI2 are again much smaller than the differences between the fixed and multiphysics ensembles.

The NWS soundings are also used to evaluate the differences among the deterministic forecasts with different PBL schemes (Fig. 14). None of the PBL schemes consistently performs either best or worst across different variables and verification times (Fig. 14). For example, differences in temperature RMSEs are most pronounced around 850–900 hPa, where QNSE has the smallest error and ACM2 has the largest error (Figs. 14a,b). However, QNSE has the largest low-level error for water vapor at 1200 UTC (Fig. 14d) while ACM2 has one of the smallest water vapor errors at low levels at 1200 UTC (Fig. 14d) but one of the largest water vapor errors at low levels at 0000 UTC (Fig. 14c). CTL (i.e., MYNN) generally has the smallest errors at most levels for wind (Figs. 14e,f).

Forecasts initialized on 26 June, and valid on 27 June, are used to demonstrate the sensitivity of a nocturnal LLJ forecast to the PBL scheme (Fig. 15). At 0000 UTC 27 June at Dodge City, Kansas, the QNSE forecast shows the strongest low-level wind speeds (Fig. 15a, red line), although all of the forecasts have weaker wind speeds than observed (Fig. 15a, black line). QNSE maintains the strongest low-level wind speeds at 0300 UTC (Fig. 15b) while ACM2 has the strongest low-level wind speeds at 0600–0900 UTC (Figs. 15c,d). At the next observation time of 1200 UTC 27 June all of the forecasts again show weaker wind speeds than observed (Fig. 15e). Figure 15 demonstrates that the temporal evolution, strength, and direction of the

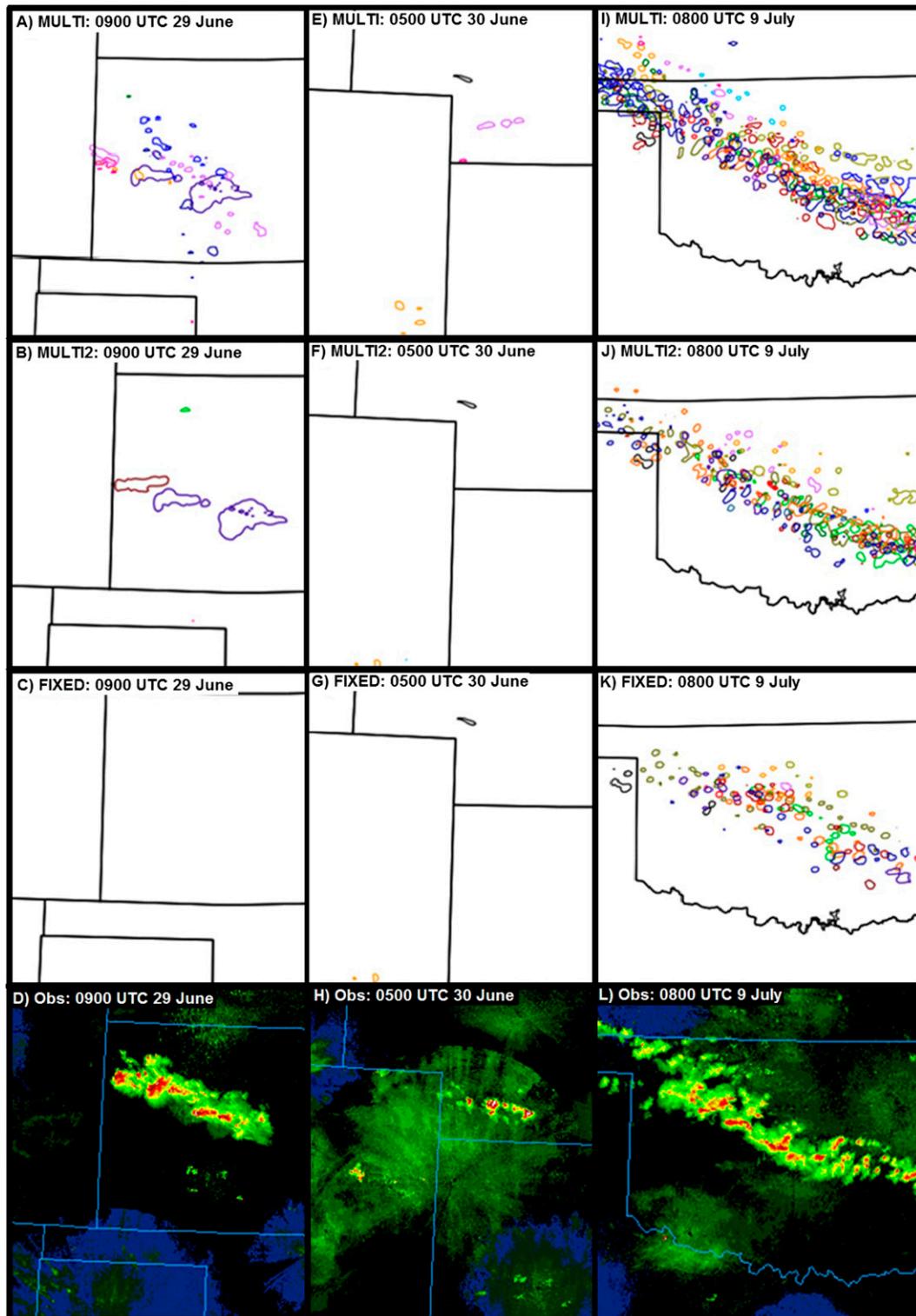


FIG. 11. Ensemble spaghetti plots for forecast reflectivity at the 30-dBZ threshold for the (a),(e),(i) MULTI, (b),(f),(j) MULT12, and (c),(g),(k) FIXED ensembles, as well as (d),(h),(l) the corresponding observed reflectivity, for the three cases of nocturnal CI discussed in the text.

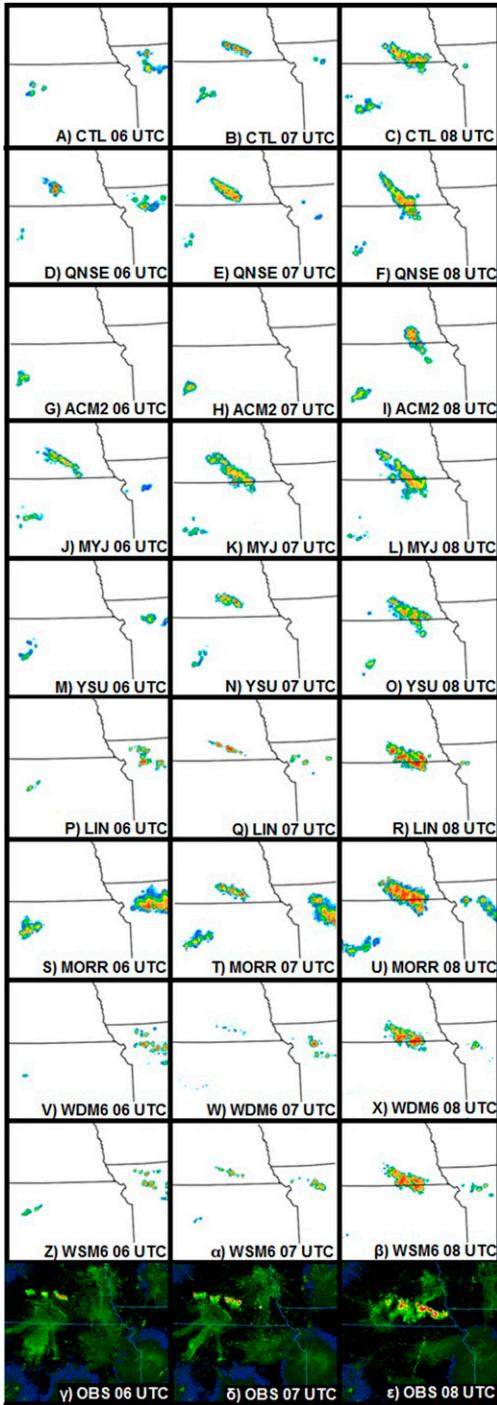


FIG. 12. Reflectivity forecasts for the experiments in Table 4 for observations and forecasts initialized at 1300 UTC 6 Jul 2014 and valid at (left) 0600 UTC 7 Jul, (center) 0700 UTC 7 Jul, and (right) 0800 UTC 7 Jul.

nocturnal LLJ can be quite sensitive to the PBL scheme. Given the lack of upper-air observations between 0000 and 1200 UTC, a more detailed analysis is saved for the verification in Part II, which utilizes the

PECAN field experiment observations, including 3-hourly sounding data.

4. Summary and discussion

One of the goals of PECAN is to better understand and improve the predictability of nocturnal convective features, including nocturnal MCSs, nocturnal CI, the nocturnal LLJ, and atmospheric bores on the nocturnal stable layer. As part of this effort, a GSI-based ensemble DA and forecast system is implemented with a specific focus on nocturnal convection. In this first part of a two-part study, extensive experiments were conducted to understand the optimal DA cycling and forecast physics configurations for such a focus. The results in this paper are used to guide the real-time implementation of the DA and forecast system during the PECAN field experiment, as described in Part II.

Experiments are first conducted to understand the optimal configuration of the GSI-based ensemble DA component of the forecast system. It is shown that the advantage of radar DA between 1200 and 1300 UTC in the morning extends through the following night in terms of precipitation forecast skill. It is also shown that 10-min cycling of the radar DA provides more skillful short-term (0–6 h) forecasts than 5- or 15-min cycling, while 5- and 10-min cycling generally provide similarly skillful longer-term (6–18 h) forecasts during the nocturnal period that are slightly more skillful than with 15-min cycling. We speculate that a possible reason for the lower skill with the more frequent 5-min cycling is that imbalances introduced during DA may not have enough time to adjust to the model before the next assimilation time (Wang et al. 2013). Possible reasons for the lower skill with the less frequent 15-min cycling include the overall assimilation of less data and the greater and more nonlinear perturbation growth in the longer (15 min) first-guess forecasts more severely violating the EnKF assumptions of small perturbations and linear growth. An advantage of 60 min of radar DA over just 30 min of radar DA is also found to extend into the overnight period. The WSM6 microphysics scheme with parameter perturbations during DA provides more skillful short-term forecasts than both the WSM6 and Thompson microphysics schemes without parameter perturbations. However, Thompson microphysics during DA provides more skillful longer-term, overnight forecasts because the impacts of convection ongoing during the morning DA on low- to midlevel moisture are more accurately analyzed with Thompson microphysics. Both the degraded performance of Thompson at the early lead times and the improved analysis of the mesoscale environment above the boundary layer during

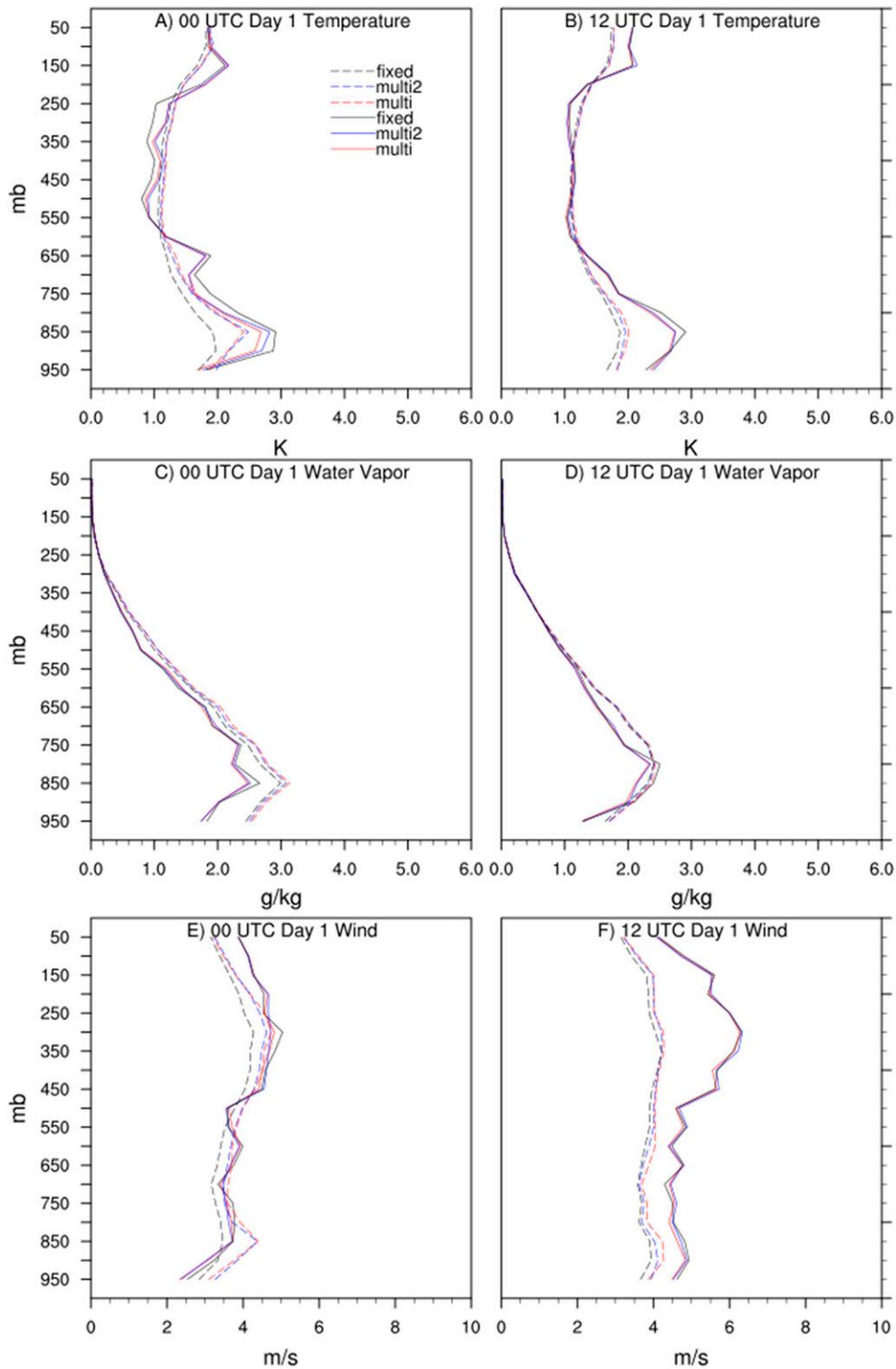


FIG. 13. Ensemble mean RMSE (solid), verified against NWS-observed soundings, and ensemble spread (dashed; includes observation error) from the three experiments on ensemble physics configurations for (a) temperature at 0000 UTC day 1 (i.e., 11-h forecast), (b) temperature at 1200 UTC day 1 (i.e., 23-h forecast), (c) water vapor mixing ratio at 0000 UTC day 1, (d) water vapor mixing ratio at 1200 UTC day 1, (e) U - and V -wind components at 0000 UTC day 1, and (f) U - and V -wind components at 1200 UTC day 1.

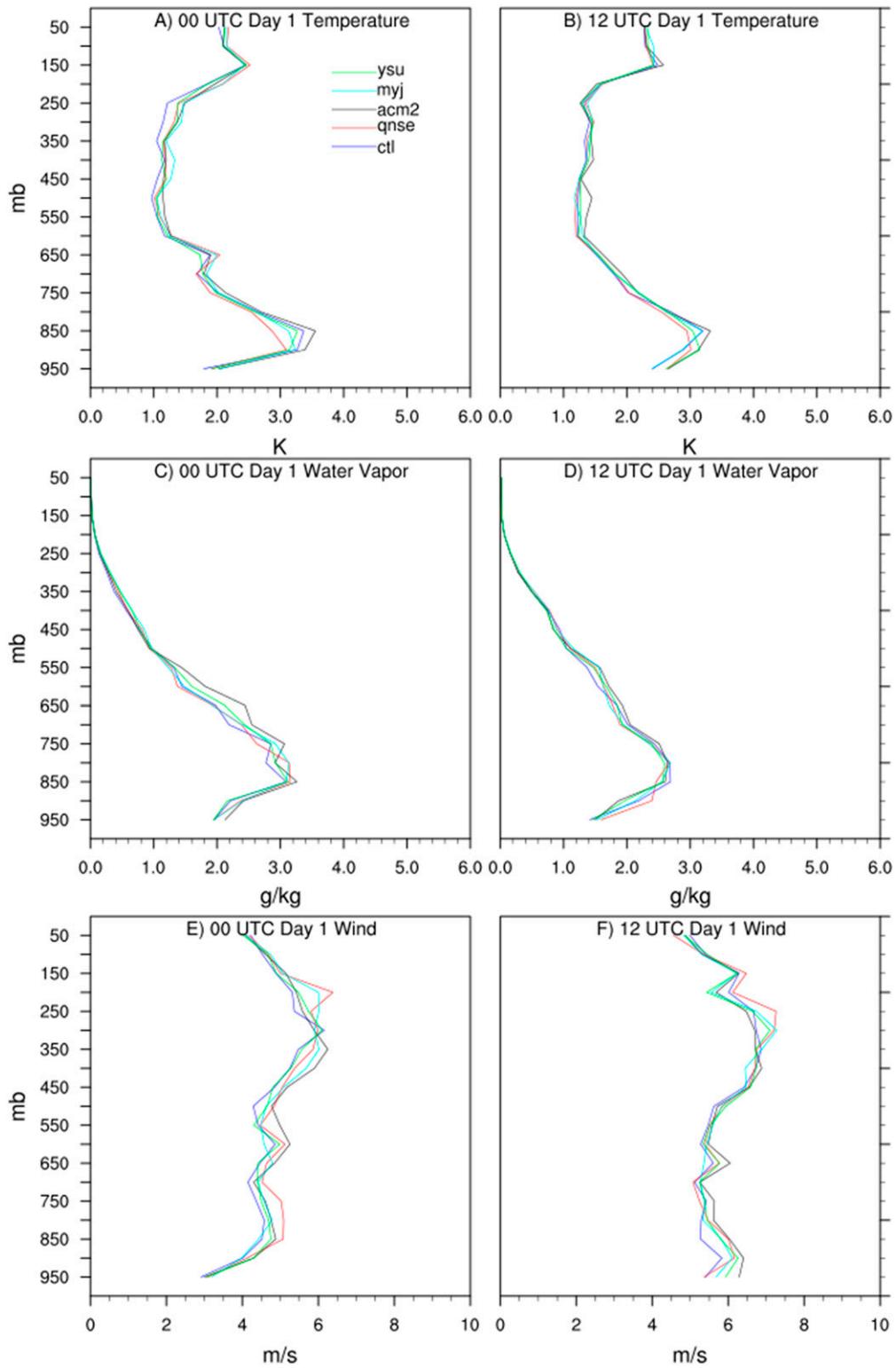


FIG. 14. As in Fig. 13, but for deterministic forecasts with different PBL schemes.

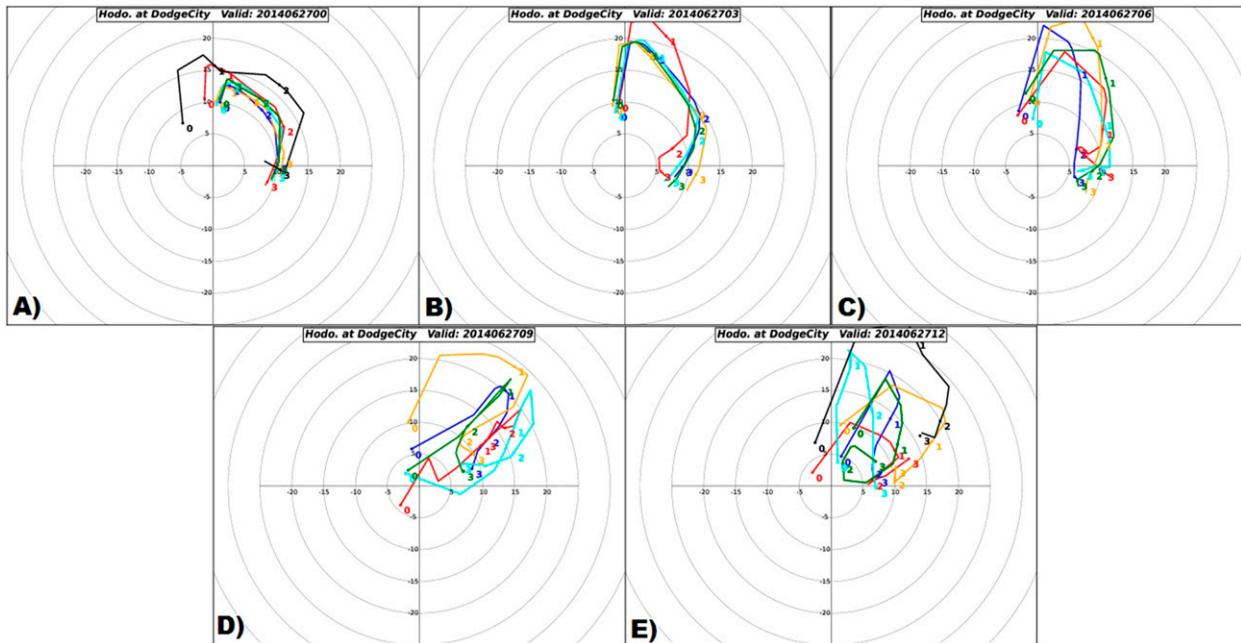


FIG. 15. The 0–3.5-km hodographs at Dodge City from deterministic forecasts with different PBL schemes, initialized at 1300 UTC 26 Jun 2014 and valid at (a) 0000, (b) 0300, (c) 0600, (d) 0900, and (e) 1200 UTC 27 Jun. Overlaid on panels (a) and (e) are the corresponding observed hodographs (black line). Color scheme is as in Fig. 14, but here ACM2 is orange instead of black, and the observations are black.

the 6 June 2014 case appear related to a systematic tendency of the radar DA cycles to remove low-level rainwater that is replaced at the expense of environmental moisture during each background forecast cycle. In situ aircraft observations of hydrometeor concentrations collected during the 2015 PECAN field experiment may allow for better understanding of how the DA system should distribute reflectivity increments among the different hydrometeors. Past studies comparing microphysics schemes for MCS and supercell prediction have typically focused on cold pool temperature and late afternoon storms with a well-developed convective boundary layer (e.g., Wheatley et al. 2014; Dawson et al. 2015). However, in this study the differences in surface temperature are less pronounced, and less important, than the differences in above-surface moisture. Further studies on the interactions between microphysics and radar DA are still needed for different ambient environments (e.g., convective versus stable PBL). Forecasts are also sensitive to the PBL scheme configuration of the DA ensemble. At ~0–6-h lead times, forecasts initialized using a single-PBL scheme during DA are more skillful than forecasts initialized with a multi-PBL configuration during DA. This result holds regardless of the relative skill of the single-PBL scheme. We speculate that the unequal skill and different systematic biases of different PBL schemes too severely violate the EnKF assumption of random, Gaussian errors, and result in

systematic clustering of the ensemble (Johnson et al. 2011). This clustering may be more problematic for DA ensembles than forecast ensembles. During the ~0000–0900 UTC nocturnal period, forecasts initialized with a QNSE PBL configuration during DA are slightly more skillful than forecasts initialized with an MYNN or multi-PBL configuration during DA.

While the above experiments are all conducted using the same forecast physics, varying only the DA configuration, experiments are also conducted using the same DA configuration and varying only the forecast physics configuration. These experiments are evaluated within the context of both deterministic and ensemble forecasts. Deterministic forecasts with different physics schemes show that the greatest precipitation forecast skill during the nocturnal period is obtained with Thompson as the microphysics scheme and MYNN as the PBL scheme. Subjective evaluation of nocturnal elevated CI events also reveals that while the intensity of the initiating convection is more sensitive to the microphysics scheme than the PBL scheme, the time and orientation of the initiating convection are more sensitive to the PBL scheme than the microphysics scheme. Strong sensitivities to PBL schemes are also shown for the low-level winds and stability during the overnight period, which are important for predicting nocturnal LLJ and bore events. Further analysis of these sensitivities is provided in Part II because of the more

comprehensive overnight observations available during the PECAN field experiment in 2015.

A fixed-physics ensemble configuration using the most skillful physics schemes during the period of interest (i.e., MYNN and Thompson) is also compared with two different multiphysics configurations. The fixed-physics ensemble shows greater error and lower spread than both multiphysics ensembles for both precipitation and non-precipitation forecasts. For nocturnal MCS cases, the greater spread in the multiphysics ensembles subjectively corresponds better to the forecast location errors. For nocturnal CI cases, the greater spread in the multiphysics ensembles increases the number of ensemble members forecasting observed CI events and more accurately reflects the types of storm structures observed in the initiating convection. The advantages of the greater spread in the multiphysics ensemble are more pronounced in the MULTI ensemble than the MULTI2 ensemble. Therefore, improvements to the forecast ensemble configuration can be obtained by thoughtfully choosing the physics configurations (i.e., MULTI) rather than evenly distributing the three best PBL and microphysics schemes across the ensemble members (i.e., MULTI2).

In Part II of this two-part study, the implementation of the GSI-based ensemble DA and forecast system for real-time nocturnal convection prediction during the 2015 PECAN experiment will be described and evaluated. A more quantitative analysis of LLJ and CI events is also saved for Part II because of the greater number of events during the 2015 PECAN field experiment.

Acknowledgments. The work is primarily supported by NSF Awards AGS-1359703 and AGS-1046081. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation Grant ACI-1053575, and the Yellowstone machine (ark:/85065/d7wd3xhc) at NCAR's Computational and Information Systems Laboratory, which is sponsored by NSF.

REFERENCES

- Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805–1824, doi:10.1175/2008MWR2691.1.
- Bright, D. R., and S. L. Mullen, 2002: The sensitivity of the numerical simulation of the Southwest monsoon boundary layer to the choice of PBL turbulence parameterization in MM5. *Wea. Forecasting*, **17**, 99–114, doi:10.1175/1520-0434(2002)017<0099:TSOTNS>2.0.CO;2.
- Carbone, R. E., J. W. Conway, N. A. Crook, and M. W. Moncrieff, 1990: The generation and propagation of a nocturnal squall line. Part I: Observations and implications for mesoscale predictability. *Mon. Wea. Rev.*, **118**, 26–49, doi:10.1175/1520-0493(1990)118<0026:TGAPOA>2.0.CO;2.
- Cintineo, R., J. A. Otkin, M. Xue, and F. Kong, 2014: Evaluating the performance of planetary boundary layer and cloud microphysical parameterization schemes in convection-permitting ensemble forecasts using synthetic *GOES-13* satellite observations. *Mon. Wea. Rev.*, **142**, 163–182, doi:10.1175/MWR-D-13-00143.1.
- Clark, A. J., W. A. Gallus Jr., and T.-C. Chen, 2008: Contributions of mixed physics versus perturbed initial/lateral boundary conditions to ensemble-based precipitation forecast skill. *Mon. Wea. Rev.*, **136**, 2140–2156, doi:10.1175/2007MWR2029.1.
- , and Coauthors, 2012: An overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. *Bull. Amer. Meteor. Soc.*, **93**, 55–74, doi:10.1175/BAMS-D-11-00040.1.
- Coffer, B. E., L. C. Maudlin, P. G. Veals, and A. J. Clark, 2013: Dryline position errors in experimental convection-allowing NSSL-WRF model forecasts and the operational NAM. *Wea. Forecasting*, **28**, 746–761, doi:10.1175/WAF-D-12-00092.1.
- Coleman, T. A., and K. R. Knupp, 2011: Radiometer and profiler analysis of the effects of a bore and a solitary wave on the stability of the nocturnal boundary layer. *Mon. Wea. Rev.*, **139**, 211–223, doi:10.1175/2010MWR3376.1.
- Colman, B. R., 1990: Thunderstorms above frontal surfaces in environments without positive CAPE. Part I: A climatology. *Mon. Wea. Rev.*, **118**, 1103–1122, doi:10.1175/1520-0493(1990)118<1103:TAFSIE>2.0.CO;2.
- Coniglio, M. C., J. Y. Hwang, and D. J. Stensrud, 2010: Environmental factors in the upscale growth and longevity of MCSs derived from Rapid Update Cycle analyses. *Mon. Wea. Rev.*, **138**, 3514–3539, doi:10.1175/2010MWR3233.1.
- , J. Correia Jr., P. T. Marsh, and F. Kong, 2013: Verification of convection-allowing WRF model forecasts of the planetary boundary layer using sounding observations. *Wea. Forecasting*, **28**, 842–862, doi:10.1175/WAF-D-12-00103.1.
- Davis, C. A., K. W. Manning, R. E. Carbone, S. B. Trier, and J. D. Tuttle, 2003: Coherence of warm-season continental rainfall in numerical weather prediction models. *Mon. Wea. Rev.*, **131**, 2667–2679, doi:10.1175/1520-0493(2003)131<2667:COWCRI>2.0.CO;2.
- Dawson, D. T., II, M. Xue, J. A. Milbrandt, and A. Shapiro, 2015: Sensitivity of real-data simulations of the 3 May 1999 Oklahoma City tornadic supercell and associated tornadoes to multimoment microphysics. Part I: Storm- and tornado-scale numerical forecasts. *Mon. Wea. Rev.*, **143**, 2241–2265, doi:10.1175/MWR-D-14-00279.1.
- Dowell, D. C., L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139**, 272–294, doi:10.1175/2010MWR3438.1.
- Duda, J. D., X. Wang, F. Kong, and M. Xue, 2014: Using varied microphysics to account for uncertainty in warm-season QPF in a convection-allowing ensemble. *Mon. Wea. Rev.*, **142**, 2198–2219, doi:10.1175/MWR-D-13-00297.1.
- Easterling, D. R., and P. J. Robinson, 1985: The diurnal variation of thunderstorm activity in the United States. *J. Climate Appl. Meteor.*, **24**, 1048–1058, doi:10.1175/1520-0450(1985)024<1048:TDVOTA>2.0.CO;2.
- Ek, M. B., K. E. Mitchell, Y. Lin, P. Grunmann, E. Rogers, G. Gayno, and V. Koren, 2003: Implementation of the upgraded Noah land surface model in the NCEP operational

- mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851, doi:10.1029/2002JD003296.
- Erlingis, J. M., and A. P. Barros, 2014: A study of the role of daytime land-atmosphere interactions on nocturnal convective activity in the southern Great Plains during CLASIC. *J. Hydrometeorol.*, **15**, 1932–1953, doi:10.1175/JHM-D-14-0016.1.
- French, A. J., and M. D. Parker, 2010: The response of simulated nocturnal convective systems to a developing low-level jet. *J. Atmos. Sci.*, **67**, 3384–3408, doi:10.1175/2010JAS3329.1.
- Fritsch, J. M., and R. E. Carbone, 2004: Improving quantitative precipitation forecasts in the warm season: A USWRP research and development strategy. *Bull. Amer. Meteor. Soc.*, **85**, 955–965, doi:10.1175/BAMS-85-7-955.
- Geerts, B. D., and Coauthors, 2017: The 2015 Plains Elevated Convection at Night (PECAN) field project. *Bull. Amer. Meteor. Soc.*, doi:10.1175/BAMS-D-15-00257.1, in press.
- Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004: Precipitation and evolution sensitivity in simulated deep convective storms: Comparisons between liquid-only and simple ice and liquid phase microphysics. *Mon. Wea. Rev.*, **132**, 1897–1916, doi:10.1175/1520-0493(2004)132<1897:PAESIS>2.0.CO;2.
- Hamill, T. M., 1999: Hypothesis tests for evaluating numerical precipitation forecasts. *Wea. Forecasting*, **14**, 155–167, doi:10.1175/1520-0434(1999)014<0155:HTFENP>2.0.CO;2.
- Heideman, K. F., and J. M. Fritsch, 1988: Forcing mechanisms and other characteristics of significant summertime precipitation. *Wea. Forecasting*, **3**, 115–130, doi:10.1175/1520-0434(1988)003<0115:FMAOCO>2.0.CO;2.
- Hill, K. A., and G. M. Lackmann, 2009: Influence of environmental humidity on tropical cyclone size. *Mon. Wea. Rev.*, **137**, 3294–3315, doi:10.1175/2009MWR2679.1.
- Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.
- Hou, D., E. Kalnay, and K. K. Droegemeier, 2001: Objective verification of the SAMEX '98 ensemble forecasts. *Mon. Wea. Rev.*, **129**, 73–91, doi:10.1175/1520-0493(2001)129<0073:OVOTSE>2.0.CO;2.
- Hu, X.-M., J. W. Nielsen-Gammon, and F. Zhang, 2010: Evaluation of three planetary boundary layer schemes in the WRF model. *J. Appl. Meteor. Climatol.*, **49**, 1831–1844, doi:10.1175/2010JAMC2432.1.
- Janjić, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945, doi:10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.
- , 2001: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model. NCEP Office Note 437, 61 pp. [Available online at http://www2.mmm.ucar.edu/wrf/users/phys_refs/SURFACE_LAYER/eta_part4.pdf.]
- Johnson, A., and X. Wang, 2012: Verification and calibration of neighborhood and object-based probabilistic precipitation forecasts from a multimodel convection-allowing ensemble. *Mon. Wea. Rev.*, **140**, 3054–3077, doi:10.1175/MWR-D-11-00356.1.
- , and —, 2013: Object-based evaluation of a storm-scale ensemble during the 2009 NOAA Hazardous Weather Testbed Spring Experiment. *Mon. Wea. Rev.*, **141**, 1079–1098, doi:10.1175/MWR-D-12-00140.1.
- , —, M. Xue, and F. Kong, 2011: Hierarchical cluster analysis of a convection-allowing ensemble during the Hazardous Weather Testbed 2009 Spring Experiment. Part II: Ensemble clustering over the whole experiment period. *Mon. Wea. Rev.*, **139**, 3694–3710, doi:10.1175/MWR-D-11-00016.1.
- , and Coauthors, 2014: Multiscale characteristics and evolution of perturbations for warm season convection-allowing precipitation forecasts: Dependence on background flow and method of perturbation. *Mon. Wea. Rev.*, **142**, 1053–1073, doi:10.1175/MWR-D-13-00204.1.
- , X. Wang, J. R. Carley, L. J. Wicker, and C. Karstens, 2015: A comparison of multiscale GSI-based EnKF and 3DVar data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. *Mon. Wea. Rev.*, **143**, 3087–3108, doi:10.1175/MWR-D-14-00345.1.
- Kain, J. S., and Coauthors, 2013: A feasibility study for probabilistic convection initiation forecasts based on explicit numerical guidance. *Bull. Amer. Meteor. Soc.*, **94**, 1213–1225, doi:10.1175/BAMS-D-11-00264.1.
- Kincer, J. B., 1916: Daytime and nighttime precipitation and their economic significance. *Mon. Wea. Rev.*, **44**, 628–633, doi:10.1175/1520-0493(1916)44<628:DANPAT>2.0.CO;2.
- Lakshmanan, V., T. Smith, G. Stumpf, and K. Hondl, 2007: The Warning Decision Support System–Integrated Information. *Wea. Forecasting*, **22**, 596–612, doi:10.1175/WAF1009.1.
- Lim, K.-S. S., and S.-Y. Hong, 2010: Development of an effective double-moment cloud microphysics scheme with prognostic cloud condensation nuclei (CCN) for weather and climate models. *Mon. Wea. Rev.*, **138**, 1587–1612, doi:10.1175/2009MWR2968.1.
- Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092, doi:10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2.
- Marshall, J. H., S. B. Trier, T. M. Weckwerth, and J. W. Wilson, 2011: Observations of elevated convection initiation leading to a surface-based squall line during 13 June IHOP_2002. *Mon. Wea. Rev.*, **139**, 247–271, doi:10.1175/2010MWR3422.1.
- Mittermaier, M. P., 2007: Improving short-range high-resolution model precipitation forecast skill using time-lagged ensembles. *Quart. J. Roy. Meteor. Soc.*, **133**, 1487–1500, doi:10.1002/qj.135.
- 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, doi:10.1029/97JD00237.
- Morrison, H., G. Thompson, and V. Tatarskii, 2009: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Mon. Wea. Rev.*, **137**, 991–1007, doi:10.1175/2008MWR2556.1.
- Nakanishi, M., and H. Niino, 2009: Development of an improved turbulence closure model for the atmospheric boundary layer. *J. Meteor. Soc. Japan*, **87**, 895–912, doi:10.2151/jmsj.87.895.
- Noh, Y., W. G. Cheon, S. Y. Hong, and S. Raasch, 2003: Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Bound.-Layer Meteorol.*, **107**, 401–427, doi:10.1023/A:1022146015946.
- Parker, M. D., 2008: Response of simulated squall lines to low-level cooling. *J. Atmos. Sci.*, **65**, 1323–1341, doi:10.1175/2007JAS2507.1.
- Pleim, J. E., 2007: A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing. *J. Appl. Meteor. Climatol.*, **46**, 1383–1395, doi:10.1175/JAM2539.1.
- Schwartz, C. S., and Coauthors, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing

- probabilistic guidance with small ensemble membership. *Wea. Forecasting*, **25**, 263–280, doi:10.1175/2009WAF2222267.1.
- Stensrud, D. J., and M. S. Wandishin, 2000: The correspondence ratio in forecast evaluation. *Wea. Forecasting*, **15**, 593–602, doi:10.1175/1520-0434(2000)015<0593:TCRIFE>2.0.CO;2.
- , J.-W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 2077–2107, doi:10.1175/1520-0493(2000)128<2077:UICAMP>2.0.CO;2.
- Sukoriansky, S., B. Galperin, and V. Perov, 2005: Application of a new spectral theory of stably stratified turbulence to atmospheric boundary layers over sea ice. *Bound.-Layer Meteor.*, **117**, 231–257, doi:10.1007/s10546-004-6848-4.
- Tao, W.-K., and Coauthors, 2003: Microphysics, radiation, and surface processes in the Goddard Cumulus Ensemble (GCE) model. *Meteor. Atmos. Phys.*, **82**, 97–137, doi:10.1007/s00703-001-0594-7.
- 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, doi:10.1175/2008MWR2387.1.
- 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, doi:10.1175/1520-0493(1993)121<1078:EOECPT>2.0.CO;2.
- , C. A. Davis, D. A. Ahijevych, M. L. Weisman, and G. H. Bryan, 2006: Mechanisms supporting long-lived episodes of propagating nocturnal convection within a 7-day WRF model simulation. *J. Atmos. Sci.*, **63**, 2437–2461, doi:10.1175/JAS3768.1.
- , —, and —, 2010: Environmental controls on the simulated diurnal cycle of warm-season precipitation in the continental United States. *J. Atmos. Sci.*, **67**, 1066–1090, doi:10.1175/2009JAS3247.1.
- , J. H. Marsham, C. A. Davis, and D. A. Ahijevych, 2011: Numerical simulations of the postsunrise reorganization of a nocturnal mesoscale convective system during 13 June IHOP_2002. *J. Atmos. Sci.*, **68**, 2988–3011, doi:10.1175/JAS-D-11-0112.1.
- Tuttle, J. D., and C. A. Davis, 2006: Corridors of warm season precipitation in the central United States. *Mon. Wea. Rev.*, **134**, 2297–2317, doi:10.1175/MWR3188.1.
- Wallace, J. M., 1975: Diurnal variations in precipitation and thunderstorm frequency over the conterminous United States. *Mon. Wea. Rev.*, **103**, 406–419, doi:10.1175/1520-0493(1975)103<0406:DVIPAT>2.0.CO;2.
- Wang, S., M. Xue, and J. Min, 2013: A four-dimensional asynchronous ensemble square-root filter (4Den SRF) algorithm and tests with simulated radar data. *Quart. J. Roy. Meteor. Soc.*, **139**, 805–819, doi:10.1002/qj.1987.
- Weisman, M. L., C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0–36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting*, **23**, 407–437, doi:10.1175/2007WAF2007005.1.
- Wheatley, D. M., N. Yussouf, and D. J. Stensrud, 2014: Ensemble Kalman filter analyses and forecasts of a severe mesoscale convective system using different choices of microphysics schemes. *Mon. Wea. Rev.*, **142**, 3243–3263, doi:10.1175/MWR-D-13-00260.1.
- , K. H. Knopfmeier, T. A. Jones, and G. J. Creager, 2015: Storm-scale data assimilation and ensemble forecasting with the NSSL Experimental Warn-on-Forecast System. Part I: Radar data experiments. *Wea. Forecasting*, **30**, 1795–1817, doi:10.1175/WAF-D-15-0043.1.
- Yussouf, N., and D. J. Stensrud, 2012: Comparison of single-parameter and multiparameter ensembles for assimilation of radar observations using the ensemble Kalman filter. *Mon. Wea. Rev.*, **140**, 562–586, doi:10.1175/MWR-D-10-05074.1.
- , E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, 2013: The ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell storm using single- and double-moment microphysics schemes. *Mon. Wea. Rev.*, **141**, 3388–3412, doi:10.1175/MWR-D-12-00237.1.
- Zhang, J., and Coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) system: Description, results, and future plans. *Bull. Amer. Meteor. Soc.*, **92**, 1321–1338, doi:10.1175/2011BAMS-D-11-00047.1.